PART II

5

Introduction to Least Squares Adjustment

5.1. GENERAL

In Part I the reader was introduced to the basic concepts underlying observations, the mathematical model, and adjustment. It was pointed out that adjustment is in general meaningful only in those cases in which redundant observations are available. In a statistical sense, adjustment is a method of deriving estimates for stochastic variables and their distribution parameters from observed samples. Of the different adjustment methods *least squares* is by far the most common.

Since its first application to an astronomical problem by C. F. Gauss, least squares adjustment has been introduced and applied in a vast number of fields in science and engineering. Its practical importance has recently been enhanced by the introduction of electronic computers, by the formulation of its techniques in matrix notation, and by connecting its concept to statistics.

Before planning observations, a general functional model about the system to be assessed must be specified. Such a functional model, which refers to a finite closed system, is determined by a certain number of var-

iables (parameters or observations or both) and the relationships between them.

There is always a *minimum number* of independent variables that uniquely determines a chosen model. It is important to note that the same physical object or set of events can be associated with different functional models. These models would have different elements depending on the particular choice made for the purpose or extent to which information is needed. Once that choice is made and the problem is specified, the *minimum number* of variables is consequently fixed—although the choice of actual variables still remains optional. In other words, we would determine the minimum number of variables without specifying the particular variables, as illustrated by the following examples.

1. The *shape of a plane triangle* (a particular geometric functional model) is uniquely determined by *two distinct* variables (minimum number). These may be chosen as *any* two angles of the three, or as the ratios of the three sides (optional choice of actual variables).
2. The *size and shape of a plane triangle* (another particular functional model) require a minimum of *three distinct* variables (which therefore cannot be all angles since a linear dimension is necessary for determining size). Several choices remain open, such as two angles and one side, two sides and one angle, or three sides.
3. If, in addition, the location and orientation with respect to a specified cartesian coordinate system of a plane triangle (yet another functional model) are of interest, a *total* of *six* variables becomes necessary. Obviously more choices also exist in this case.

As has been indicated above, once a model is selected it is expressed by a minimum number of distinct variables. This minimum number of variables will always be referred to by n_0. Unless the given observations are sufficient for determining the n_0 variables, the situation will obviously be deficient. We denote the number of observations by n. These observations must be functionally independent, that is, not one of the n observations can be derived from any or all remaining $(n - 1)$ observations. When n is larger than n_0, *redundancy* is said to exist and *adjustment* is needed in order to obtain a unique set of estimates for the model variables. The redundancy, which is denoted by r, is given by

$$r = n - n_0 \tag{5.1}$$

and is equal to the *(statistical) degrees of freedom*.

The redundancy r is meaningful only if the observations and the functional model are mutually consistent. The observations should be sufficient to determine the model, that is, they should refer to at least the minimum

number of distinct random variables. Otherwise there will be a deficiency in the model.

A deficiency can exist in spite of apparent existence of redundancy as, for instance, when multiple observations are present (see the following example). In complex systems such as large photogrammetric blocks, deficiency may be encountered without the presence of multiple observations.

Example 5.1. Denote the three interior angles of a plane triangle by α, β, γ, and assume that determining the shape of the triangle is of interest. Therefore the functional model is defined as that of the shape of a triangle for which $n_0 = 2$, each referring to one distinct angle α, β, or γ. Now, there are several possibilities for actual observations.

1. Each angle is observed once, that is, α_1, β_1, γ_1. Thus $n = 3$ and $r = n - n_0 = 1$.
2. Each angle is observed twice, or α_1, α_2, β_1, β_2, γ_1, γ_2. Hence $n = 6$ and $r = 4$.
3. Angle α observed twice, and β, γ each observed once, or α_1, α_2, β, γ. Here $n = 4$ and $r = 2$.
4. Angle α is observed four times, α_1, α_2, α_3, α_4. In this case although $n = 4$, we actually have a deficient situation, because the observations do not refer to the minimum of $n_0 = 2$ distinct variables of the specified model. Therefore these observations are no longer adequate for the model and the shape of the triangle cannot be determined.

If we are asked to indicate what can be determined from these observations, the model must first be reduced to that of one angle α, for which $n_0 = 1$. In this case $r = 3$ and the adjustment would involve obtaining a unique estimate for the angle α from the four given observations.

Part (4) of the above example demonstrates that before we can evaluate the value of the redundancy r, a check must be made to ascertain that the total model selected is determinable from the given observations.

5.2. THE LEAST SQUARES PRINCIPLE

Because of the inherent stochastic properties of observations, redundant observations are not usually compatible with the functional model. Any sufficient subset of functionally (but not necessarily stochastically) independent observations can be used to assess the model. Due to random variability of the observational sampling, each minimum subset would yield a different result. For example, any two of three measured angles in a plane triangle suffice to determine its shape. However, for every possible pair of angles (for example, α and β; β and γ; or γ and α) a triangle shape results that is in general different from that obtained from another pair; that is, no unique result is possible from redundant observations unless an additional criterion is introduced. In this situation it is the basic principle of adjustment

to derive a unique set of estimates for all the model variables, with certain optimum properties.

The original set of observations, which will be denoted by the vector l and which includes redundant observations and is inconsistent with the model, is replaced in the adjustment by another set of estimates $\hat{l}$ which satisfies the model. This implies that the functional relationships comprising the model are accepted as superior to the given sample values of the observations. Consequently, after the adjustment is performed, model relationships remain unaltered and will be strictly satisfied by the newly estimated set of values $\hat{l}$.

The set of estimates $\hat{l}$ is in general different from the original set l. The difference between the two sets is

$$v = \hat{l} - l \qquad\qquad (5.2)$$

which has been termed in the classical theory as either "corrections" or "residuals." Neither of the two terms is descriptive or precise, and other suggested terms such as "alterations" or "shifts" are equally so. For lack of a better term, the term "residuals" is selected as adequate, and will be used throughout this book.

The vector of residuals v plays an important role after the adjustment process. It is often possible to analyze the elements of v in order to test the adequacy of the model. In certain instances this analysis may even lead to discarding a chosen model completely and to remodeling the problem.

Due to redundancy, there would be an infinite number of estimates for v, or $\hat{l}$, which would satisfy the model. Amongst all the possibilities there exists one set of estimates that, in addition to being consistent with the model, satisfies another criterion commonly referred to as the *least squares principle*. This principle endeavors to ascertain that the new estimates $\hat{l}$ are as close as possible to the sample values of the observations l taking their stochastic properties also into consideration. Such a criterion is obviously plausible since the available sample of observations is the best that we have, although it is superseded by the model functions. Therefore any variations in l necessitated by the existence of inconsistencies with the model due to redundancy must be as small as possible. In the limit when $r = 0$, $\hat{l} = l$ and the residuals will all be zero.

The least squares principle states that

$$\phi = v^t W v \rightarrow \text{minimum} \qquad\qquad (5.3)$$

where W is the weight matrix of the observations. It should be denoted by $W_{ll} = Q_{ll}^{-1}$ with Q_{ll} being the cofactor matrix. However, since they will be used extensively, neither Q nor W will have a subscript for reasons of simplification. In the exceptional cases when ambiguity may arise, the subscripts will be restored. The weight matrix W is square and of order equal to n, the number of observations. Its elements reflect the stochastic properties,

such as variation and correlation, of all the observations as discussed in Part I.

The minimum criterion given above is the most general expression. Several special cases may be derived from it by considering weight matrices of special structure. The first specialization is when the observations are considered as uncorrelated. The weight matrix $\mathbf{W}$ would be a diagonal matrix, and the minimum principle becomes

$$\phi = \sum_{i=1}^{n} (w_i v_i^2) \to \text{minimum} \qquad (5.4)$$

where w_i is the ith diagonal element of $\mathbf{W}$ and v_i is the residual associated with the corresponding ith observation.

A yet simpler case is when the observations are taken as of equal precision in addition to being uncorrelated. Then $\mathbf{W}$ becomes the identity matrix and

$$\phi = \sum_{i=1}^{n} (v_i^2) \to \text{minimum} \qquad (5.5)$$

will be the form of the minimum principle. This last case is the oldest and most classical, and possibly the one that gave rise to the name " least squares," since in this case we seek the " least " of the sum of the squares of the residuals.

It is important to note that the application òf least squares principle given above *does not require a priori knowledge of the distribution associated with the observations*. All that is necessary is to have either $\mathbf{W}$ or $\mathbf{Q}$ defined and known. In the past it has been erroneously stated that least squares adjustment requires normal distribution. Perhaps the reason for this incorrect assertion is that when the observations are normally distributed, the least squares estimates will have some special properties, such as being identical to those from the method of maximum likelihood.

Least squares has been by far the predominant technique of data adjustment in photogrammetry, geodesy, surveying, and many other fields. The reason for such extensive use lies in the fact that it yields a computational algorithm for a unique answer even in very complicated cases. This was essential from the beginning, as early attempts at adjustment of geodetic nets in the nineteenth century failed mainly due to the lack of a suitable algorithm. The advantage of such an algorithm has been enhanced lately by the adoption of matrix notation and the use of electronic computers for performing the computations.

5.3. THE TECHNIQUES OF LEAST SQUARES

The mathematical model represents the point of beginning at which the basic philosophy of the adjustment task is established. Once this has been accomplished, the model remains in the background and consideration is

then given to the practical and computational aspects of selecting a particular technique of least squares. The relationship between these phases is shown schematically in Figure 5.1. The least squares computations yield updated estimates of all model variables as well as their respective cofactor or covariance matrices. After the computational algorithm is applied to the given data, another step is necessary where the results are statistically evaluated. This is again a judgment operation which lies on the same side as the model. In fact, depending on the degree of sophistication, the results of the statistical evaluation could lead to remodeling of the adjustment task if the original model is found to be inadequate.
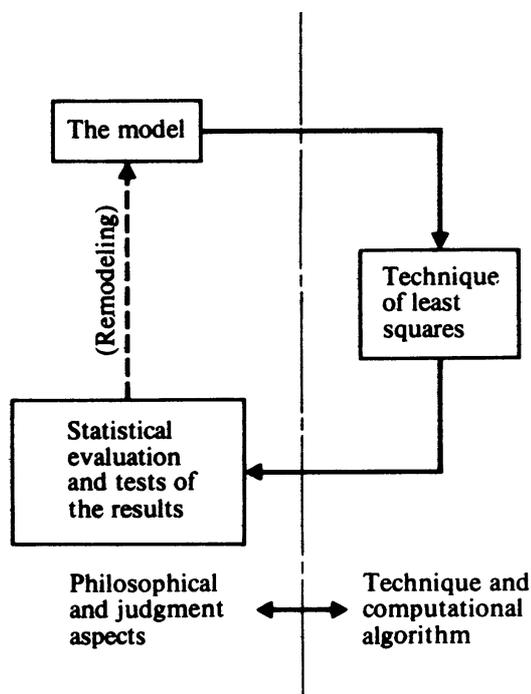

Figure 5.1

Although for a specified model and a given set of data the least squares yields unique results, there are several techniques that can be employed. It should be emphasized that whatever technique is used, the final answers are always the same.

In addition to the observations, the model may also include other variables and numerical constants. The family of those other variables, which are also stochastic, will be termed "parameters" to distinguish them from the observations with given a priori sample values. Invariably these parameters have unknown values at the beginning and estimates are therefore derived for them in the adjustment. They represent functional unknowns, although after the adjustment they are treated as stochastic variables undergoing statistical testing just like the variables representing the observations.

The number of parameters carried in the adjustment will be designated by the letter u (for unknowns). The vector $\Delta_{u,1}$ will be used to represent those parameters. Once the functional and stochastic models are established, the least squares technique or algorithm operates with a set of mathematical functions or equations. These equations describe the functional model of the adjustment problem. For example, the equation for the closure of a plane triangle and the relationship equating the final values of two observations on a distance or angle express the functional models of the respective problems. We shall distinguish between two types of equations: condition equations and constraint equations. Any equation that includes one or more *observations* will be called a *condition equation*. Consequently, every adjustment problem will entail condition equations. The total number of these equations or conditions in the adjustment will always be denoted by c. By contrast, those equations that *do not* include any observations will be called *constraint equations*, *parameter constraints*, or simply *constraints*. By definition, these equations will be functions of only parameters and constants. If the same vector of parameters Δ appears in both the condition equations and constraint equations, the number of the latter (constraints) will be given the letter s. In some cases, in formulating the constraints it may be convenient to include in addition to Δ another vector of *added parameters* Δ'. In those cases s' will represent the number of constraints and q the number of added parameters.

To recapitulate then, Table 5.1 gives the names of the different symbols that have been introduced.

TABLE 5.1

Symbol	Explanation
l	Vector of sample values of the observational variables $\bar{l}$ which will also be denoted by l for simplification
$\hat{l}$	Estimated ("corrected") observations
v	Vector of residuals
Δ	(Unknown) parameters
Δ'	Added parameters (appearing in constraint equations)
n_0	Minimum number of distinct variables expressing a specified model
n	Number of given observations (estimates)
r	The redundancy or (statistical) degrees of freedom
u	Number of (unknown) parameters (in Δ)
q	Number of added parameters (in Δ')
c	Number of condition equations
s	Number of constraints (in terms of only Δ)
s'	Number of constraints with added parameters Δ'
Q	Cofactor matrix of the observations l (replacing Q_{ll})
W	Weight matrix of the observations l (replacing W_{ll})

There are two groups of adjustment techniques that will be treated as one encompassing technique at the end. The two groups are entitled, (a) *least squares adjustment with conditions only*, and (b) *least squares adjustment with conditions and constraints*. These two groups of techniques arise mainly because of making a distinction between observations and parameters (hence conditions and constraints). Recently, it was found that this distinction can be eliminated by considering all model variables as observations, thus leading to only one inclusive method called (c) *a unified least squares adjustment*.

Since many practical adjustment problems can be treated by a technique from the first two groups, they will therefore be developed in detail in the remainder of Part II. Then the unified approach will be introduced in Part III.

The schematic of Figure 5.1 showed that we should, for practical reasons, differentiate clearly between the setting up of the model and the least squares adjustment technique. It must be stressed that the observations, or rather the variables they refer to, including additional parameters associated with them (calibration parameters, orientation parameters, and so on) are part of the total system. The eventual merits of an adjustment are mostly decided by the choice of the model at the beginning.

For a long time the importance of the model concept has not been realized, because in conventional applications the model is given a priori and therefore it is not problematic. But even if no specific reference to the model is made when performing an adjustment, it is nevertheless fixed by implication. Many questions about least squares adjustment, and also practical difficulties, refer in fact to the problem of adequate modeling.

Although the techniques of least squares adjustment are mainly concerned with obtaining the least squares estimates for a given model, the real problem is often to refine or change the model if preliminary adjustments turn out to be unsatisfactory.

5.4. LINEAR AND NONLINEAR FUNCTIONS IN THE MODEL

The conditional as well as the constraint equations involved in an adjustment problem can, in general, be nonlinear. However, least squares treatments are generally performed with linear functions, since it is rather difficult and often impractical, at least at present, to seek a least squares solution of nonlinear equations. Consequently, whenever the equations in the model are originally nonlinear, some means of linearization must be used to get linear equations. Series expansions, and Taylor's series in particular, are often used for the purpose, where only the zero and first-order terms are used and all other higher-order terms are neglected.

When applying a series expansion, a set of approximate values for the unknowns in the equations must be chosen. The choice of those approximations is an important aspect of solving the problem at hand. Unfortunately

there is no concrete and unique way of choosing approximations that can be applied to all adjustment problems. Sometimes experience is relied on, and in other situations some computational shortcuts may be employed. In all cases the attempt should always be made to obtain close approximations that can be obtained by using relatively simple and uninvolved techniques.

To demonstrate how linearization is performed, let any set of m nonlinear functions be denoted by

$$F(\mathbf{x}) = \mathbf{0} \tag{5.6}$$

where $\mathbf{x}_{p,1}$ is a vector of unknown variables (which in the context of adjustment may be observations, parameters, and so on). If $\mathbf{x}^0_{p,1}$ denotes the vector of approximate values of the variables, the zero and first-order terms of a series expansion would be

$$F(\mathbf{x}^0) + \frac{\partial F(\mathbf{x})}{\partial \mathbf{x}}\bigg|_{\mathbf{x}=\mathbf{x}^0} \cdot \Delta\mathbf{x} = \mathbf{0} \tag{5.7}$$

The set of partial derivatives of the functions with respect to the elements of the vector of variables is a rectangular matrix $\mathbf{U}$ of dimensions $m \times p$. The vector $\Delta\mathbf{x}$ is a $p \times 1$ vector of corrections to the approximations which replaces the vector of unknowns $\mathbf{x}$. The result of applying the series expansion is that the nonlinear equation (5.6) becomes a set of linear equations of the general form

$$\underset{m,\,p}{\mathbf{U}} \ \underset{p,\,1}{\Delta\mathbf{x}} = \underset{m,\,1}{\mathbf{u}} \tag{5.8}$$

where $\mathbf{u} = -F(\mathbf{x}^0)$.

After least squares adjustment, we get as a solution the vector $\Delta\mathbf{x}$. If the original approximation vector $\mathbf{x}^0$ was sufficiently close for equation (5.7) to be an adequate substitute for equation (5.6), that is, for second- and higher-order terms of the series to be in fact negligible, then the final least squares estimate is $(\mathbf{x}^0 + \Delta\mathbf{x})$. Often, however, $\mathbf{x}^0$ is not that close an approximation and adding $\Delta\mathbf{x}$ to $\mathbf{x}^0$ yields *only an improved* approximation. The updated vector of approximations must now be used again to formulate equation (5.8) or equation (5.7) and least squares is used to obtain an updated vector $\Delta\mathbf{x}$ the elements of which are, in general, smaller than those of the first one. The process of relinearization at an updated vector of approximation continues until the last value of $\Delta\mathbf{x}$ is insignificantly small and the iterative procedure terminates. The final estimate $\hat{\mathbf{x}}$ will be the sum of the original approximation $\mathbf{x}^0$ and *all* the correction vectors $\Delta\mathbf{x}$. (Or equivalently the last updated approximation vector plus the last correction vector.)

One must note that the method of least squares is applied only within each iteration and on a linear (linearized) set of equations. It does not really have anything to do with the linearization process. The linearization is only a scheme to treat a problem of nonlinear estimation which although possible in principle, is not normally solved directly by least squares.

6

Adjustment with Conditions Only—General Case

Adjustment of Observations and Functionally Independent Parameters

6.1. INTRODUCTION

In this group of techniques the parameters, if they exist, are functionally independent. The existence of the parameters in the adjustment is relatively arbitrary and will depend mostly on the type of problem to be solved. In principle, we can perform the adjustment using only the observations. However, it may be considerably more convenient to include a parameter vector. This is particularly true when the parameters themselves are the variables of interest.

When starting an adjustment, the first thing to be specified is the model, particularly the functional model, which will specify n_0. Then the given n observations are examined with respect to the model in order to ascertain that there is no deficiency. Thus the redundancy $r = n - n_0$ can be computed. This redundancy, or degrees of freedom, may be interpreted to mean that among the n observations there exist r functions (conditions) that must be satisfied. This can be explained by considering that when $r = 0$, then $n = n_0$ and the observations fit the model perfectly. When n exceeds n_0 by one, that is, $r = 1$, one function must be written relating the n observations

together. The sum of the three angles of a plane triangle is such a function, for example, involving one redundancy. This analysis can obviously be extended to include all r degrees of freedom in the adjustment.

Now we consider the existence of parameters in addition to observations. If we have to write r conditions when there are no parameters, then $r + 1$ conditions need to be written if one (unknown) parameter is added. This added condition is necessary to allow for the determination of that extra unknown, and to *retain the same number of degrees of freedom r* in the system. Thus if there are u unknown independent parameters in the adjustment, the number of conditions will, in general, be

$$c = r + u \tag{6.1}$$

This is regarded as the general case within the group of techniques containing independent parameters. Special cases are designated according to the number of parameters u carried in the adjustment. The lower limit occurs obviously when $u = 0$. The upper limit occurs when $u = n_0$, since in this case $c = n$. If u is taken larger than n_0, then the number of conditions would exceed the number of observations, which is not possible unless the parameters are not independent. Hence, the following two inequalities must be satisfied

$$r \leq c \leq n \tag{6.2}$$

$$0 \leq u \leq n_0 \tag{6.3}$$

For the present case the number of parameters u may take any value larger than zero but less than n_0. The chosen parameters may be only those of direct interest, or those of interest plus others. While we are analyzing the adjustment problem, the number of parameters u can be composed not only of parameters of direct interest but also of other parameters that make the writing of the condition equations either easier or less complicated.

As soon as these u parameters are specified, a total of $c = r + u$ *independent* condition equations must be set up between the n observations and u parameters. These equations may be linear or linearized as pointed out in the preceding chapter. We let the condition equations take the general form

$$\mathbf{A}(l + \mathbf{v}) + \mathbf{B}\Delta = \mathbf{d} \tag{6.4}$$

in which $\mathbf{A}$ and $\mathbf{B}$ are coefficient matrices, $\mathbf{d}$ is a column vector of constants, and the remaining symbols have been previously defined (see Table 5.1). Rearranging equation (6.4) leads to the following form which will be used as a standard throughout this book

$$\mathbf{A}\mathbf{v} + \mathbf{B}\Delta = \mathbf{f} \tag{6.5}$$

with

$$\mathbf{f} = \mathbf{d} - \mathbf{A}l \tag{6.6}$$

In view of the inequalities of equations (6.2) and (6.3) and other explanatory remarks, the following are the characteristics of the matrices in equations (6.5) and (6.6):

$\mathbf{A} = c \times n$ matrix   rank $(\mathbf{A}) = c$

$\mathbf{v} = n \times 1$ vector

$\mathbf{B} = c \times u$ matrix   rank $(\mathbf{B}) = u$

$\Delta = u \times 1$ vector

$\mathbf{f} = c \times 1$ vector   (6.7)

There are three other different cases that may be considered as special examples of the present case. These are when $u = 0$, giving the lower limit of $c = r$ and no parameters; when $u = n_0$, giving the upper limit when $c = n$; and a further simplification of the case of $c = n$, with the matrix $\mathbf{A}$ being the identity matrix. Those will be dealt with in later sections.

6.2. DERIVATION

Equation (6.5) is the fundamental form of condition equations for the adjustment of observations and independent parameters combined. It represents c linear equations in $(n + u)$ unknowns, which are the elements of the two vectors $\mathbf{v}$ and Δ. Since from equations (6.1) and (6.2) c is less than $(n + u)$, there exist many solutions for equation (6.5). A unique least squares solution is obtained by adding the basic criterion of equation (5.3), or

$$\phi = \mathbf{v}'\mathbf{W}\mathbf{v} \to \text{minimum}$$

To enforce this criterion and at the same time have a solution for equation (6.5), the method of constrained minima by Lagrange multipliers is used. Thus if $\mathbf{k}_{c,1}$ represents the yet unknown Lagrange multipliers, then we should seek the minimum for the following function, noting that the quantity between parenthesis is zero when equation (6.5) is satisfied.

$$\phi' = \mathbf{v}'\mathbf{W}\mathbf{v} - 2\mathbf{k}'(\mathbf{A}\mathbf{v} + \mathbf{B}\Delta - \mathbf{f}) \tag{6.8}$$

To minimize ϕ', its partial derivatives with respect to $\mathbf{v}$ and to Δ are equated to zero. Consequently, according to the rules in Section A8 (Appendix A), we get

$$\frac{\partial \phi'}{\partial \mathbf{v}} = 2\mathbf{v}'\mathbf{W} - 2\mathbf{k}'\mathbf{A} = \mathbf{0}'$$

and

$$\frac{\partial \phi'}{\partial \Delta} = -2\mathbf{k}'\mathbf{B} = \mathbf{0}'$$

which are in row form. Transposing and rearranging yields

$$- \underset{n,\,n\ \ n,\,1}{\mathbf{W}\ \mathbf{v}} + \underset{n,\,c\ \ c,\,1}{\mathbf{A}^t\ \mathbf{k}} = 0 \tag{6.9}$$

$$\underset{u,\,c\ \ c,\,1}{\mathbf{B}^t\ \mathbf{k}} = 0 \tag{6.10}$$

realizing of course that $\mathbf{W}$ is a symmetric matrix.

It is important to pause here and consider the number of unknowns versus the number of linear equations. There is a total of $(n + u + c)$ unknowns. Correspondingly, there are c linear equations in equation (6.5), n equations in equation (6.9), and u equations in equation (6.10). In matrix form the total system is

$$\begin{array}{c} \ \\ n \\ c \\ u \end{array} \begin{array}{ccc} n & c & u \\ \left[\begin{array}{ccc} -\mathbf{W} & \mathbf{A}^t & 0 \\ \mathbf{A} & 0 & \mathbf{B} \\ 0 & \mathbf{B}^t & 0 \end{array}\right] \end{array} \begin{bmatrix} \mathbf{v} \\ \mathbf{k} \\ \Delta \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{f} \\ 0 \end{bmatrix} \tag{6.11}$$

This system of equations has usually been referred to as *the total system of normal equations*. The matrix of coefficients is a square symmetric matrix of order $(n + c + u)$, which is always *nonsingular* (that is, its rank is equal to its order), unless the model is improperly constructed. In view of this fact the least squares problem will be solved by inverting the system of equation (6.11), or

$$\begin{bmatrix} \mathbf{v} \\ \mathbf{k} \\ \Delta \end{bmatrix} = \begin{bmatrix} -\mathbf{W} & \mathbf{A}^t & 0 \\ \mathbf{A} & 0 & \mathbf{B} \\ 0 & \mathbf{B}^t & 0 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ \mathbf{f} \\ 0 \end{bmatrix} \tag{6.12}$$

The existence of zero matrices along the main diagonal of the coefficient matrix should not be disturbing as long as the matrix is nonsingular and because the zero matrices can be avoided simply by rearranging the equations.

For relatively small problems in which the *total* system of normal equations is limited in number, equation (6.12) may be used and the problem may be directly solved in this manner. However, many practical problems are not that small and the resulting total normal equations constitute a rather large system for which a direct solution may not be convenient, or practical. Furthermore, we may not be interested in both $\mathbf{v}$ and Δ but only in one of them, and we rarely need $\mathbf{k}$ for its own sake. Henceforth, an alternative scheme may be desirable. Fortunately, the system of equation (6.11) contains many zero submatrices and a solution by partitioning is relatively simple. From equation (6.9),

$$\mathbf{v} = \mathbf{W}^{-1}\mathbf{A}^t\mathbf{k} = \mathbf{Q}\mathbf{A}^t\mathbf{k} \tag{6.13}$$

and substituting in equation (6.5) gives

$$\mathbf{AQA}^t\mathbf{k} + \mathbf{B}\Delta = \mathbf{f} \tag{6.14}$$

Let

$$l_e = \mathbf{A}l \tag{6.15}$$

represent an "equivalent" set of observations. Applying the propagation rule, the cofactor matrix $\mathbf{Q}_e$ for l_e is

$$\mathbf{Q}_e = \mathbf{AQA}^t \tag{6.16}$$

which when used in equation (6.14) yields

$$\mathbf{Q}_e\,\mathbf{k} + \mathbf{B}\Delta = \mathbf{f}$$

solving for $\mathbf{k}$

$$\mathbf{k} = \mathbf{Q}_e^{-1}(-\mathbf{B}\Delta + \mathbf{f}) = \mathbf{W}_e(-\mathbf{B}\Delta + \mathbf{f}) \tag{6.17}$$

where

$$\mathbf{W}_e = \mathbf{Q}_e^{-1} = (\mathbf{AQA}^t)^{-1} \tag{6.18}$$

substituting equation (6.17) into (6.10) and reducing

$$(\mathbf{B}^t\mathbf{W}_e\,\mathbf{B})\Delta = (\mathbf{B}^t\mathbf{W}_e\,\mathbf{f}) \tag{6.19}$$

or

$$[\mathbf{B}^t(\mathbf{AQA}^t)^{-1}\mathbf{B}]\Delta = [\mathbf{B}^t(\mathbf{AQA}^t)^{-1}\mathbf{f}] \tag{6.20}$$

Equation (6.19) or (6.20) represents a set of u equations in u unknown parameters (the elements of Δ) which are termed *partially reduced normal equations*. With the auxiliaries

$$\mathbf{N} = \mathbf{B}^t\mathbf{W}_e\,\mathbf{B} = \mathbf{B}^t(\mathbf{AQA}^t)^{-1}\mathbf{B} \tag{6.21}$$

$$\mathbf{t} = \mathbf{B}^t\mathbf{W}_e\,\mathbf{f} = \mathbf{B}^t(\mathbf{AQA}^t)^{-1}\mathbf{f} \tag{6.22}$$

a more compact form of equations (6.19) and (6.20) will be

$$\underset{d,\,u\ u,\,1}{\mathbf{N}\ \ \Delta} = \underset{u,\,1}{\mathbf{t}} \tag{6.23}$$

The vector Δ may be obtained from equation (6.23) by direct inversion such that

$$\Delta = \mathbf{N}^{-1}\mathbf{t} \tag{6.24}$$

In the derivation above, several inverses were taken. First, $\mathbf{W}^{-1}$ in equation (6.13) is all right since $\mathbf{W}$ is nonsingular because the observations are functionally independent. Then $\mathbf{Q}_e^{-1}$ in equation (6.17) is allowable because the rank of $\mathbf{Q}_e$ is equal to the rank of $\mathbf{A}$, or c, which is equal to its order. Finally, $\mathbf{N}^{-1}$ in equation (6.24) is also possible because $\mathbf{N}$ has a rank and order that are equal ($= u$). Having the value of the parameters Δ, we can compute the

vector **k** from equation (6.17) and substitute it into equation (6.13) to evaluate the vector of residuals **v**. With this vector the quadratic form of the residuals may be computed and an unbiased estimate of the reference variance obtained from (a derivation and discussion of this aspect is given in Section 11.3)

$$\hat{\sigma}_0^2 = \frac{\mathbf{v}^t\mathbf{W}\mathbf{v}}{r} \tag{6.25}$$

where r is the number of degrees of freedom (or redundancy) as defined earlier and is equal to the number of condition equations c minus the number of unknown parameters u.

Finally, if the least squares estimate of the observations (which satisfy the model) is also desired, it may be obtained by adding the residuals to the given observations.

6.2.1. A Computational Check The equality

$$\mathbf{a} = \mathbf{B}^t\mathbf{W}_e \, \mathbf{A}\mathbf{v} = \mathbf{B}^t(\mathbf{AQA}^t)^{-1}\mathbf{A}\mathbf{v} = 0 \tag{6.26}$$

may be used for checking the correctness of the computation, particularly for smaller problems. Equation (6.26) may be verified by first substituting for **v** from equation (6.13); then for **k** from equation (6.17), and finally referring to equation (6.12). Thus

$$\begin{aligned}
\mathbf{a} &= \mathbf{B}^t(\mathbf{AQA}^t)^{-1}\mathbf{AQA}^t\mathbf{k} \\
&= \mathbf{B}^t(\mathbf{AQA}^t)^{-1}(-\mathbf{B}\Delta + \mathbf{f}) \\
&= -\mathbf{B}^t(\mathbf{AQA}^t)^{-1}\mathbf{B}\Delta + \mathbf{B}^t(\mathbf{AQA}^t)^{-1}\mathbf{f} \\
&= 0
\end{aligned}$$

6.2.2. Computation of Quadratic Form Equation (6.25) shows that in order to compute the a posteriori estimate of the reference variance, $\hat{\sigma}_0^2$, we must first compute the quadratic form ($\mathbf{v}^t\mathbf{W}\mathbf{v}$). To do this we recall equation (6.13),

$$\mathbf{v} = \mathbf{QA}^t\mathbf{k}$$

and from equation (6.17)

$$\mathbf{v} = \mathbf{QA}^t\mathbf{W}_e(-\mathbf{B}\Delta + \mathbf{f})$$

Thus

$$\mathbf{v}^t\mathbf{W}\mathbf{v} = (-\Delta^t\mathbf{B}^t + \mathbf{f}^t)\mathbf{W}_e \, \mathbf{AQWQA}^t\mathbf{W}_e(-\mathbf{B}\Delta + \mathbf{f})$$

or

$$\begin{aligned}
\mathbf{v}^t\mathbf{W}\mathbf{v} &= \mathbf{f}^t\mathbf{W}_e \, \mathbf{f} - 2\mathbf{f}^t\mathbf{W}_e \, \mathbf{B}\Delta + \Delta^t\mathbf{B}^t\mathbf{W}_e \, \mathbf{B}\Delta \\
&= \mathbf{f}^t\mathbf{W}_e \, \mathbf{f} - 2\mathbf{t}^t\Delta + \Delta^t\mathbf{N}\Delta \\
&= \mathbf{f}^t\mathbf{W}_e \, \mathbf{f} - 2\Delta^t\mathbf{t} + \Delta^t\mathbf{t}
\end{aligned}$$

or

$$v'Wv = f'W_e\, f - \Delta't \qquad\qquad (6.27)$$

In a nonlinear problem (see Section 11.1) if the iterative process has been carried out sufficiently, such that the last correction vector Δ is essentially zero, we can disregard the last term of the right-hand side of equation (6.27) and use the following relation for computing the quadratic form

$$v'Wv = f'W_e\, f = f'(AQA')^{-1}f \qquad\qquad (6.28)$$

6.2.3. Precision Estimation A second, and quite important, part of adjustment is the determination of the precision of the quantities that have been estimated in the adjustment. Such precision estimation is in the form of cofactor matrices, and it may also, of course, be in the form of covariance matrices. These matrices can be derived from the least squares solution by simply applying the rules of propagation of Chapter 4. For completeness, all matrices are given here.

$$l = Il$$

$$f = d - Al$$

$$\Delta = N^{-1}B'W_e\, f = N^{-1}B'W_e(d - Al)$$

$$k = W_e[-B\Delta + f] = W_e(-BN^{-1}B'W_e + I)(d - Al)$$

$$v = QA'k = QA'W_e(-BN^{-1}B'W_e + I)(d - Al)$$

$$\hat{l} = l + v$$

Autocofactor Matrices

$$Q_{ll} = Q$$

$$Q_{ff} = (-A)Q(-A)' = Q_e$$

$$Q_{\Delta\Delta} = (-N^{-1}B'W_e\, A)Q(-A'W_e\, BN^{-1})$$

$$= N^{-1}B'W_e\, BN^{-1} = N^{-1} \qquad\qquad (6.29)$$

$$Q_{kk} = [-W_e(-BN^{-1}B'W_e + I)A]Q[-A'(-W_e\, BN^{-1}B' + I)W_e]$$

$$= (-W_e\, BN^{-1}B'W_e + W_e)Q_e(-W_e\, BN^{-1}B' + I)W_e$$

$$= (-W_eBN^{-1}B' + I)^2W_e = W_e - W_e\, BN^{-1}B'W_e \quad \text{(idempotent)}$$

$$Q_{vv} = QA'(W_e - W_e\, BN^{-1}B'W_e)AQ$$

$$= QA'W_e\, AQ - QA'W_e\, BQ_{\Delta\Delta}\, B'W_e\, AQ \qquad\qquad (6.30)$$

$$Q_{\hat{l}\hat{l}} = Q_{ll} + Q_{lv} + Q_{vl} + Q_{vv} = Q - Q_{vv} \qquad\qquad (6.31)$$

Crosscofactor Matrices

$$\mathbf{Q}_{lf} = -\mathbf{Q}\mathbf{A}^t$$

$$\mathbf{Q}_{l\Delta} = -\mathbf{Q}\mathbf{A}^t\mathbf{W}_e \, \mathbf{BN}^{-1} \tag{6.32}$$

$$\mathbf{Q}_{lk} = -\mathbf{Q}\mathbf{A}^t(\mathbf{W}_e - \mathbf{W}_e \, \mathbf{BN}^{-1}\mathbf{B}^t\mathbf{W}_e)$$

$$\mathbf{Q}_{lv} = -\mathbf{Q}\mathbf{A}^t(\mathbf{W}_e - \mathbf{W}_e \, \mathbf{BN}^{-1}\mathbf{B}^t\mathbf{W}_e)\mathbf{A}\mathbf{Q} = -\mathbf{Q}_{vv}$$

$$\mathbf{Q}_{li} = \mathbf{Q}_{ll} + \mathbf{Q}_{lv} = \mathbf{Q} - \mathbf{Q}_{vv} = \mathbf{Q}_{ii}$$

$$\mathbf{Q}_{f\Delta} = +\mathbf{AQ}\mathbf{A}^t\mathbf{W}_e \, \mathbf{BN}^{-1} = \mathbf{BN}^{-1}$$

$$\mathbf{Q}_{fk} = \mathbf{Q}_e(\mathbf{W}_e - \mathbf{W}_e \, \mathbf{BN}^{-1}\mathbf{B}^t\mathbf{W}_e) = (\mathbf{I} - \mathbf{BN}^{-1}\mathbf{B}^t\mathbf{W}_e)$$

$$\mathbf{Q}_{fv} = \mathbf{Q}_e(\mathbf{I} - \mathbf{W}_e \, \mathbf{BN}^{-1}\mathbf{B}^t)\mathbf{W}_e \, \mathbf{AQ} = (\mathbf{I} - \mathbf{BN}^{-1}\mathbf{B}^t\mathbf{W}_e)\mathbf{AQ}$$

$$\mathbf{Q}_{fi} = \mathbf{Q}_{fl} + \mathbf{Q}_{fv} = -\mathbf{AQ} + \mathbf{AQ} - \mathbf{BN}^{-1}\mathbf{B}^t\mathbf{W}_e \, \mathbf{AQ}$$

$$= -\mathbf{BN}^{-1}\mathbf{B}^t\mathbf{W}_e \, \mathbf{AQ}$$

$$\mathbf{Q}_{\Delta k} = (\mathbf{N}^{-1}\mathbf{B}^t\mathbf{W}_e)\mathbf{Q}_{ff}(-\mathbf{W}_e \, \mathbf{BN}^{-1}\mathbf{B}^t\mathbf{W}_e + \mathbf{W}_e)^t$$

$$= \mathbf{N}^{-1}\mathbf{B}^t\mathbf{W}_e \, \mathbf{Q}_e \, \mathbf{W}_e - \mathbf{N}^{-1}\mathbf{B}^t\mathbf{W}_e \, \mathbf{Q}_e \, \mathbf{W}_e \, \mathbf{BN}^{-1}\mathbf{B}^t\mathbf{W}_e$$

$$= \mathbf{N}^{-1}\mathbf{B}^t\mathbf{W}_e - \mathbf{N}^{-1}\mathbf{B}^t\mathbf{W}_e = \mathbf{0}$$

or

$$\mathbf{Q}_{\Delta k} = \mathbf{Q}_{\Delta\Delta}(-\mathbf{W}_e \, \mathbf{B})^t + \mathbf{Q}_{\Delta f} \, \mathbf{W}_e^t$$

$$= -\mathbf{N}^{-1}\mathbf{B}^t\mathbf{W}_e + \mathbf{N}^{-1}\mathbf{B}^t\mathbf{W}_e = \mathbf{0}$$

$$\mathbf{Q}_{\Delta v} = \mathbf{Q}_{\Delta k}(\mathbf{Q}\mathbf{A}^t)^t = \mathbf{0}$$

$$\mathbf{Q}_{\Delta i} = \mathbf{Q}_{\Delta l} + \mathbf{Q}_{\Delta v} = -\mathbf{N}^{-1}\mathbf{B}^t\mathbf{W}_e \, \mathbf{AQ} \tag{6.33}$$

$$\mathbf{Q}_{kv} = \mathbf{Q}_{kk}(\mathbf{Q}\mathbf{A}^t)^t = (\mathbf{W}_e - \mathbf{W}_e \, \mathbf{BN}^{-1}\mathbf{B}^t\mathbf{W}_e)\mathbf{AQ}$$

$$= \mathbf{W}_e \, \mathbf{AQ} - \mathbf{W}_e \, \mathbf{BN}^{-1}\mathbf{B}^t\mathbf{W}_e \, \mathbf{AQ}$$

$$= \mathbf{Q}_{kl}$$

$$\mathbf{Q}_{ki} = \mathbf{Q}_{kl} + \mathbf{Q}_{kv} = \mathbf{Q}_{kl} - \mathbf{Q}_{kl} = \mathbf{0}$$

$$\mathbf{Q}_{vi} = \mathbf{Q}_{vl} + \mathbf{Q}_{vv} = -\mathbf{Q}_{vv} + \mathbf{Q}_{vv} = \mathbf{0}$$

Of the cofactor matrices, equation (6.29) gives a relation that is of fundamental importance in least squares estimation. It states that the cofactor matrix of the parameters is simply the inverse of the partially reduced normal equation coefficient matrix. Thus as far as the precision of the unknown parameters is concerned, it is obtained as a by-product of the least squares solution, particularly if such a solution was performed by inverting the coefficient matrix in equation (6.24). If $\mathbf{x}^0$ is the vector of approxima-

tions, which is a numerical constant as far as propagation is concerned, then $\hat{x} = x^0 + \Delta$ for nonlinear systems, and

$$Q_{\hat{x}\hat{x}} = Q_{\Delta\Delta} = N^{-1} \qquad (6.34)$$

In certain situations it is possible, and even necessary, to obtain a solution by stopping the elimination procedure after replacing v. This means that the following system is solved directly.

$$\begin{bmatrix} Q_e & B \\ B^t & 0 \end{bmatrix} \begin{bmatrix} k \\ \Delta \end{bmatrix} = \begin{bmatrix} f \\ 0 \end{bmatrix} \qquad (6.35)$$

from which

$$\begin{bmatrix} k \\ \Delta \end{bmatrix} = \begin{bmatrix} Q_e & B \\ B^t & 0 \end{bmatrix}^{-1} \begin{bmatrix} f \\ 0 \end{bmatrix} \qquad (6.36)$$

The desired vector Δ can be taken as the appropriate subvector. However, it is now an important question to find out how $Q_{\Delta\Delta}$ can be evaluated. In order to do this, we seek the inverse in equation (6.36) by partitioning (see Section A3.8, Appendix A.) Thus if

$$\begin{bmatrix} Q_e & B \\ B^t & 0 \end{bmatrix}^{-1} = \begin{bmatrix} \alpha & \beta \\ \beta^t & \gamma \end{bmatrix} \qquad (6.37)$$

and, with $N = B^t W_e B$, from equation (6.21), we write

$$\alpha = W_e - W_e BN^{-1}B^t W_e$$

$$\beta = W_e BN^{-1}$$

$$\gamma = -N^{-1} \qquad (6.38)$$

or

$$\begin{bmatrix} Q_e & B \\ B^t & 0 \end{bmatrix}^{-1} = \begin{bmatrix} (W_e - W_e BN^{-1}B^t W_e) & W_e BN^{-1} \\ N^{-1}B^t W_e & -N^{-1} \end{bmatrix} \qquad (6.39)$$

Although $Q_{kk} = (W_e - W_e BN^{-1}B^t W_e)$, the other two submatrices are not $Q_{k\Delta}$ and $Q_{\Delta\Delta}$, respectively. As derived above $Q_{k\Delta} = 0$ and $Q_{\Delta\Delta} = N^{-1}$ and *not* $-N^{-1}$. Consequently

$$\begin{bmatrix} Q_e & B \\ B^t & 0 \end{bmatrix}^{-1} \neq \begin{bmatrix} Q_{kk} & Q_{k\Delta} \\ Q_{\Delta k} & Q_{\Delta\Delta} \end{bmatrix} \qquad (6.40)$$

But the cofactor matrix of interest, $Q_{\Delta\Delta}$, can in fact be extracted from the inverse in equation (6.36) by taking the negative of the diagonal submatrix corresponding to Δ, or

$$Q_{\Delta\Delta} = -\gamma \qquad (6.41)$$

6.2.4. Functional Dependence and a Posteriori Cofactor Matrices It has been indicated so far that, a priori, both the observations and parameters are functionally independent. Functional independence of the n observations was said to mean that no one observation can be deduced or computed from the remaining $(n-1)$ observations. Also the difference between functional independence and stochastical independence or lack of correlation was stressed. Although the n observations are functionally independent, they may or may not be correlated. The difference between these two concepts can be related to the covariance or cofactor matrix associated with the vector of stochastic variables under consideration. If these variables are *functionally independent*, then the *rank of the covariance* or cofactor matrix is equal to its order, thus implying nonsingularity. On the other hand if the variables are *stochastically independent* or uncorrelated, then all elements off the main diagonal of the covariance, or cofactor, matrix must be zero. This establishes the fundamental difference between the two concepts. (In addition, perfect stochastic correlation is equivalent to linear functional dependence.)

The a priori vectors l and Δ are functionally independent and thus $\mathbf{Q}$ and $\mathbf{Q}_{\Delta\Delta}$ are both nonsingular, the first having a rank of n and the second a rank of u. Other vectors computed from the adjustment include the residuals $\mathbf{v}$ and the estimated observations $\hat{l}$ with corresponding cofactor matrices $\mathbf{Q}_{vv}$ and $\mathbf{Q}_{ii}$. The elements of both of these vectors are functionally dependent with the result that both $\mathbf{Q}_{vv}$ and $\mathbf{Q}_{ii}$ are singular matrices. Perhaps it is easier to show why the elements of $\hat{l}$ are functionally dependent. After the adjustment, $\hat{l}$ satisfies the condition equations. Thus since there are c conditions, then at least c estimated observations (that is, c elements of $\hat{l}$) can be computed from the remaining $(n-c)$ elements. Consequently, the rank of $\mathbf{Q}_{ii}$ is less than its order. In fact,

$$\text{rank } (\mathbf{Q}_{vv}) = \text{redundancy} = r \tag{6.42}$$

$$\text{rank } (\mathbf{Q}_{ii}) = (n - r)$$

$$= \text{the minimum number of variables}$$
$$\text{specifying the model, } n_0 \tag{6.43}$$

Equations (6.42) and (6.43) are derived in Appendix C for this as well as other cases of adjustment.

Example 6.1

Given: Figure 6.1 shows a much simplified problem of an object point which is photographed by three terrestrial cameras. All three camera stations are assumed to lie on the same line which is taken to be the X_1 axis of the object coordinate system. The X_2 axis is taken to coincide with the optical axis of the first camera, S_1. The camera axes are horizontal and parallel. All interior orientations of the cameras are
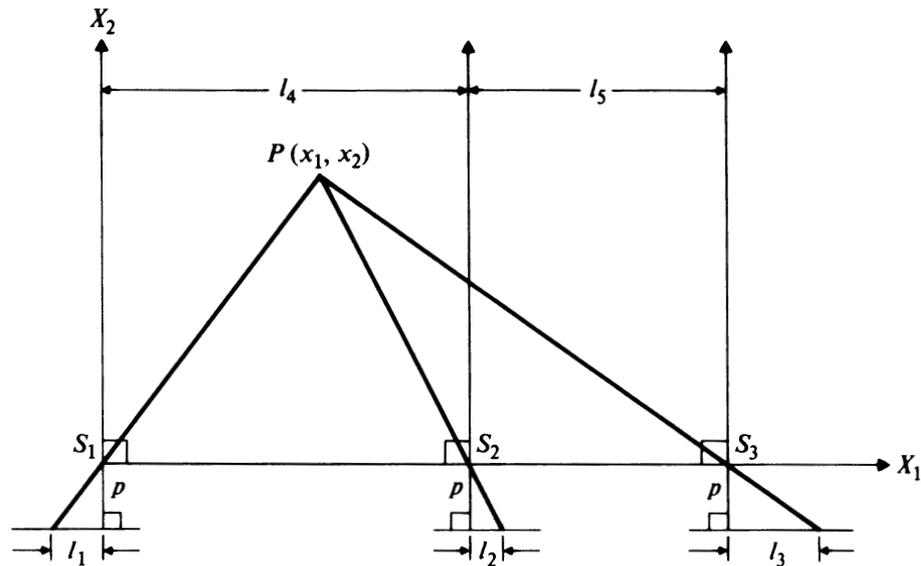
Figure 6.1. (Not to scale).

assumed known and without errors with principal distance $p = 100$ mm (as a constant value). The observations l are those five distances denoted by that symbol in Figure 6.1. Their absolute values (as distances and not coordinates) as well as their standard deviations are given in Table 6.1. All those observations are assumed, for simplicity, to be in the $X_1 X_2$ plane. We shall assume further that no correlation exists between all the observations.

TABLE 6.1

OBSERVATION	VALUE	STANDARD DEVIATION
l_1	16.5 mm	0.10 mm
l_2	3.8 mm	0.10 mm
l_3	20.4 mm	0.10 mm
l_4	10.0 m	0.05 m
l_5	8.0 m	0.05 m

Required: We require the least squares estimate of the coordinates (X_1, X_2) of point P, and the a posteriori estimates of the reference variance, the cofactor matrices, and the covariance matrices.

Solution: The first step is to analyze the mathematical model of the problem. This includes all the information that in this case specifies the geometry of the problem and all the assumptions inherent in that information, whether explicit or implied. To begin with, the object coordinate system is fixed, the camera axes are parallel to each other and normal to the X_1 axis, and the profiles of all three photographs are parallel

to the X_1 axis and at the equal distance of $p = 100$ mm from it. The given observations fix, with respect to the framework given, *three* points: S_2, S_3, and P. Next, we need to determine the minimum number of variables, n_0. Since both S_2 and S_3 lie on the X_1 axis, by fundamental assumption of the given model, only two parameters need to be determined to fix their positions. Added to those are the two that would determine point P, which makes a total of *four* variables as the necessary minimum number for determining the total geometry of the problem. Since n is equal to *five* observations, then there is *1 degree of freedom*, or $r = 1$.

In the statement of the problem it was indicated that two parameters, the pair of coordinates of point P, are desired. Thus $u = 2$ and the number of independent condition equations is $c = r + u = 3$ [see equation (6.1)]. Consulting Figure 22.1 one equation may be written for each ray from P to a camera station. (A sort of simplified version of the collinearity condition equation, as it is known in photogrammetry). Thus

$$\frac{l_1}{p} = \frac{x_1}{x_2}, \qquad \frac{l_2}{p} = \frac{l_4 - x_1}{x_2}, \qquad \text{and} \qquad \frac{l_3}{p} = \frac{l_4 + l_5 - x_1}{x_2}$$

or

$$f_1 = l_1 x_2 - px_1 \qquad\qquad = 0$$

$$f_2 = l_2 x_2 - p(l_4 - x_1) \qquad = 0$$

$$f_3 = l_3 x_2 - p(l_4 + l_5 - x_1) = 0$$

In these equations l_i refer to the *variables* representing the observations and not the numerical values of the observations. Linearization by Taylor series according to Section 11.1.2 leads to the form $\mathbf{Av} + \mathbf{B\Delta} = \mathbf{f}$ where

$$\mathbf{A}_{3,5} = \frac{\partial f_i}{\partial l_j} = \begin{bmatrix} x_2^0 & 0 & 0 & 0 & 0 \\ 0 & x_2^0 & 0 & -p & 0 \\ 0 & 0 & x_2^0 & -p & -p \end{bmatrix} \qquad \mathbf{B}_{3,2} = \frac{\partial f_i}{\partial x_j} = \begin{bmatrix} -p & l_1 \\ p & l_2 \\ p & l_3 \end{bmatrix}$$

$$\mathbf{f}_{3,1} = \begin{bmatrix} px_1^0 - l_1 x_2^0 \\ p(l_4 - x_1^0) - l_2 x_2^0 \\ p(l_4 + l_5 - x_1^0) - l_3 x_2^0 \end{bmatrix}$$

At this early stage, Δl in equations (11.5) and (11.6) reduces to $\mathbf{v}$ because the approximations for the observational variables will be taken equal to the numerical values of the observations, and no iterations on the observations are effected. Approximations for x_1 and x_2, however, are needed and may be evaluated from the following geometric relationships:

$$x_2^0 = \frac{l_4 p}{l_1 + l_2} = \frac{10(\text{m}) \times 100(\text{mm})}{20.3(\text{mm})} \simeq 50 \text{ m}$$

$$x_1^0 = \frac{l_1 x_2^0}{p} = \frac{16.5(\text{mm}) \times 50(\text{m})}{100(\text{mm})} \simeq 8 \text{ m}$$

The stochastic model in the problem is specified by the variances of the observations and the fact that there are no correlations. Assuming that the *reference variance is unity*, the covariance matrix is used as the cofactor matrix, or

$$
\mathbf{Q}_{5,5} = \begin{bmatrix}
0.01 \text{ mm}^2 & 0 & 0 & 0 & 0 \\
 & 0.01 \text{ mm}^2 & 0 & 0 & 0 \\
 & & 0.01 \text{ mm}^2 & 0 & 0 \\
\text{symmetric} & & & 0.0025 \text{ m}^2 & 0 \\
 & & & & 0.0025 \text{ m}^2
\end{bmatrix}
$$

Using the numerical values of the observations and the approximations of the parameters computed above, the elements of **A**, **B**, and **f** may be computed as

$$
\mathbf{A} = \begin{bmatrix}
50(\text{m}) & 0 & 0 & 0 & 0 \\
0 & 50(\text{m}) & 0 & -100(\text{mm}) & 0 \\
0 & 0 & 50(\text{m}) & -100 & -100
\end{bmatrix}
$$

$$
\mathbf{B} = \begin{bmatrix}
-100 & 16.5 \\
100 & 3.8 \\
100 & 20.4
\end{bmatrix} (\text{mm})
$$

$$
f = \begin{bmatrix}
-25 \\
10 \\
-20
\end{bmatrix} (\text{mm m})
$$

Then

$$
\mathbf{Q}_e = (\mathbf{AQA}^t) = 25 \begin{bmatrix}
1 & 0 & 0 \\
0 & 2 & 1 \\
0 & 1 & 3
\end{bmatrix} (\text{mm}^2 \text{ m}^2)
$$

$$
\mathbf{W}_e = \mathbf{Q}_e^{-1} = \tfrac{1}{125} \begin{bmatrix}
5 & 0 & 0 \\
0 & 3 & -1 \\
0 & -1 & 2
\end{bmatrix} (\text{mm}^{-2} \text{ m}^{-2})
$$

and

$$
\mathbf{N} = \mathbf{B}^t \mathbf{W}_e \mathbf{B} = \tfrac{1}{125} \begin{bmatrix}
8 \times 10^4 & -5.45 \times 10^3 \\
-5.45 \times 10^3 & 2.082 \times 10^3
\end{bmatrix} (\text{m}^{-2})
$$

$$
\mathbf{t} = \mathbf{B}^t \mathbf{W}_e f = \tfrac{1}{125} \begin{bmatrix}
1.25 \times 10^4 \\
-2.893 \times 10^3
\end{bmatrix} (\text{m}^{-1})
$$

$$
\mathbf{N}^{-1} = \frac{125}{13684.55 \times 10^4} \begin{bmatrix}
2.082 \times 10^3 & 5.45 \times 10^3 \\
5.45 \times 10^3 & 8 \times 10^4
\end{bmatrix} (\text{m}^2)
$$

and, finally,

$$
\mathbf{\Delta} = \mathbf{N}^{-1} \mathbf{t} = \begin{bmatrix}
0.075(\text{m}) \\
-1.193(\text{m})
\end{bmatrix}
$$

Adding Δ to the approximations yields

$$x_1 = 8.075(\text{m}) \quad \text{and} \quad x_2 = 48.807(\text{m})$$

The two values computed here will be considered the final estimates for the coordinates of point P. It should be mentioned, however, that the problem is nonlinear and more consideration should be given. Since the concern at this stage is mainly with linear (or linearized) conditions, there will be no elaboration on this aspect of nonlinearity now. Suffice it to say that if the solution is properly iterated to account for the neglected higher-order terms, the final answer would have been $x_1 = 8.0749(\text{m})$ and $x_2 = 48.8093(\text{m})$. More discussion of this subject will be given in later chapters and examples.

In order to compute the vector of residuals, we first compute $\mathbf{k}$

$$\mathbf{k} = \mathbf{W}_e(-\mathbf{B}\Delta + \mathbf{f}) = \tfrac{1}{125}[10.95 \quad 24.16 \quad -13.32]^t$$

Then

$$\mathbf{v} = \mathbf{Q}\mathbf{A}^t\mathbf{K} = [0.04 \text{ mm} \quad 0.10 \text{ mm} \quad -0.05 \text{ mm} \quad -0.022 \text{ m} \quad 0.027 \text{ m}]^t$$

The values in $\mathbf{v}$ were truncated to two decimals in millimetres (mm) and three decimals in metres (m), as indicated by the given data and variances.

Using this vector of residuals, the vector of estimated observations may be computed from $\hat{l} = l + \mathbf{v}$, or

$$\hat{l} = [16.54 \text{ mm} \quad 3.90 \text{ mm} \quad 20.35 \text{ mm} \quad 9.978 \text{ m} \quad 8.027 \text{ m}]^t$$

As a *computational check*, the expression of equation (6.26) may be evaluated to

$$\mathbf{a} = \mathbf{B}^t(\mathbf{A}\mathbf{Q}\mathbf{A}^t)^{-1}\mathbf{A}\mathbf{v} = 10^{-9}[6.17 \quad 6.10]^t$$

which is practically a zero vector as expected. The cofactor matrices are

$$\mathbf{Q}_{\Delta\Delta} = \mathbf{N}^{-1} = 10^{-3}\begin{bmatrix} 1.90 & 4.98 \\ 4.98 & 73.08 \end{bmatrix}(\text{m}^2) \quad \text{rank} = 2$$

$$\mathbf{Q}_{\hat{l}\hat{l}} = 10^{-3}\begin{bmatrix} 8.993 & -2.238 & 1.231 & 0.503 & -0.616 \\ & 5.025 & 2.737 & 1.119 & -1.369 \\ & & 8.494 & -0.616 & 0.753 \\ & & & 2.248 & 0.308 \\ \text{symmetric} & & & & 2.124 \end{bmatrix} \quad \text{rank} = 4 = n_0$$

The reference variance is computed, realizing that here there is 1 degree of freedom,

$$\hat{\sigma}_0^2 = 1.9367 \quad \text{and} \quad \hat{\sigma}_0 = 1.39$$

It is important to note that the above values were computed after the first iteration of a nonlinear problem. Although this is, in the case of this example, adequate,

here are the corresponding values after the last (fifth) iteration for the sake of comparison with other methods later on (the subscript f is used to signify final),

$$\mathbf{Q}_{\Delta\Delta_{(f)}} = 10^{-3} \begin{bmatrix} 1.86 & 5.03 \\ 5.03 & 71.33 \end{bmatrix} (m^2)$$

$$\mathbf{Q}_{\hat{u}_{(f)}} = 10^{-3} \begin{bmatrix} 9.024 & -2.189 & 1.213 & 0.500 & -0.621 \\ & 5.090 & 2.721 & 1.121 & -1.394 \\ & & 8.492 & -0.621 & 0.772 \\ & & & 2.244 & 0.318 \\ \text{symmetric} & & & & 2.104 \end{bmatrix}$$

The reference variance is also recomputed after the last iteration to give

$$\hat{\sigma}_{0_{(f)}}^2 = 1.9634 \quad \text{and} \quad \hat{\sigma}_{0_{(f)}} = 1.40$$

(For a posteriori statistical analysis on $\hat{\sigma}_0^2$ and $\mathbf{Q}_{\Delta\Delta}$ refer to Examples 11.4 and 11.6).

Example 6.2

Given: The two-parameter transformation, rotation and scale (see Chapter 8, for more detail), between two coordinate systems $\mathbf{x}$ and $\mathbf{y}$ is represented by

$$y_{1i} = ax_{1i} - bx_{2i}$$

$$y_{2i} = bx_{1i} + ax_{2i}$$

where $(x_1, x_2)_i$ are the $\mathbf{x}$ coordinates of any point i, and $(y_1, y_2)_i$ are its corresponding $\mathbf{y}$ coordinates, and (a, b) are the transformation parameters. In order to estimate a and b, Table 6.2 gives the three points of known coordinates in both systems.

TABLE 6.2

i	x_{1i}	x_{2i}	y_{1i}	y_{2i}
1	0.0	1.0	-2.1	1.1
2	1.0	0.0	1.0	2.0
3	1.0	1.0	-0.9	2.8

For all three given points, the cofactor matrix is equal to the covariance matrix for the pair of coordinates in the $\mathbf{x}$ coordinate system and is equal to

$$\mathbf{Q}_{xx} = \mathbf{\Sigma}_{xx} = \begin{bmatrix} 0.01 & 0 \\ 0 & 0.01 \end{bmatrix} = 0.01\mathbf{I}_2$$

All y coordinates are to be considered as constants insofar as the adjustment is concerned.

Required: We require the least squares estimates of the transformation parameters a and b and the a posteriori estimates of the reference variance and covariance matrices.

Solution: Similar to Example 6.1 the functional model of this example involves plane geometry, which is demonstrated schematically in Figure 6.2. As implied from the given information, we are interested in transforming from the x system to the y system. Therefore the y coordinates of the points are those desired. Having three points of known coordinates in both systems, a total of six observations, that is, $n = 6$, are available. The minimum number of variables necessary for a unique determination of a and b is $n_0 = 2$. Consequently the degrees of freedom or redundancy is $r = 4$. For the purposes of this example there are two unknown parameters, or $u = 2$, and the number of condition equations is $c = 6$. The stochastic model is relatively simple since the cofactor matrices for the x coordinates of all points are the same. Then the total cofactor matrix of the observations is

$$\underset{6,\,6}{\mathbf{Q}} = \begin{bmatrix} \mathbf{Q}_{xx} & 0 & 0 \\ & \mathbf{Q}_{xx} & 0 \\ \text{symmetric} & & \mathbf{Q}_{xx} \end{bmatrix} = 0.01\mathbf{I}_6$$
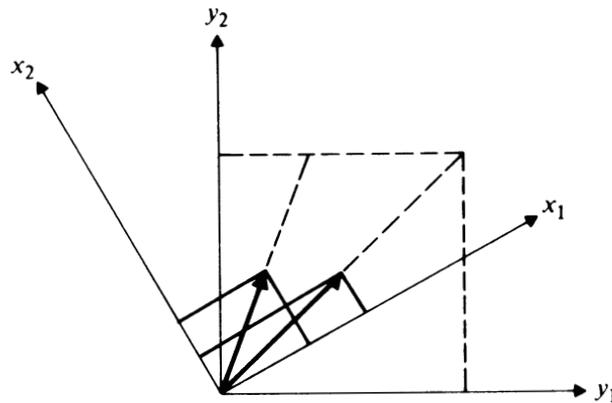

Figure 6.2

The six condition equations may be obtained by writing the pair of transformation equations for each of the three given points (a more general algorithm is given in Chapter 8),

$$x_{1i}a - x_{2i}b - y_{1i} = 0$$

$$x_{2i}a + x_{1i}b - y_{2i} = 0 \qquad i = 1, 2, 3$$

Linearization yields

$$\underset{6,\,6}{\mathbf{A}} \underset{6,\,6}{\mathbf{v}} + \underset{6,\,2}{\mathbf{B}} \underset{2,\,1}{\mathbf{\Delta}} = \underset{6,\,1}{\mathbf{f}}$$

where $\mathbf{A}$ is a block diagonal matrix of 2×2 submatrices along the main diagonal (see Section A9, Appendix A). If the approximations† a^0, b^0 are evaluated by simple computation to be 1 and 2, respectively, each submatrix on the diagonal of $\mathbf{A}$ will be

$$\mathop{\mathbf{A}_i}_{2,2} = \begin{bmatrix} a^0 & -b^0 \\ b^0 & a^0 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}$$

Similarly, the matrices $\mathbf{B}$ and $\mathbf{f}$ are evaluated as

$$\mathop{\mathbf{B}}_{6,2} = \begin{bmatrix} x_{11} & -x_{21} \\ x_{21} & x_{11} \\ x_{12} & -x_{22} \\ x_{22} & x_{12} \\ x_{13} & -x_{23} \\ x_{23} & x_{13} \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}$$

$$\mathop{\mathbf{f}}_{6,1} = \begin{bmatrix} y_{11} - x_{11}a^0 + x_{21}b^0 \\ y_{21} - x_{21}a^0 - x_{11}b^0 \\ y_{12} - x_{12}a^0 + x_{22}b^0 \\ y_{22} - x_{22}a^0 - x_{12}b^0 \\ y_{13} - x_{13}a^0 + x_{23}b^0 \\ y_{23} - x_{23}a^0 - x_{13}b^0 \end{bmatrix} = \begin{bmatrix} -0.1 \\ 0.1 \\ 0 \\ 0 \\ 0.1 \\ -0.2 \end{bmatrix}$$

noting that x_{ij} denotes the x_i coordinate of point j, and the same applies for y_{ij}. A set of two reduced normal equations may be formed from the given condition equations the solution of which is

$$\Delta = [0.00 \quad -0.05]^t$$

Thus the estimates of a and b, obtained by adding Δ to the approximations, are $\hat{a} = 1.00$ and $\hat{b} = 1.95$. Again it should be mentioned that for the purposes of this chapter we shall stop after the first iteration. In order to evaluate $\mathbf{a}$ for the computational check of equation (6.26), first the numerical values are

$$\mathbf{v} = [0.01 \quad 0.08 \quad 0.02 \quad 0.01 \quad -0.05 \quad -0.05]^t$$

$$\mathbf{Q} = 0.01\mathbf{I}_6$$

$$\mathbf{A} = \begin{bmatrix} 1 & -2 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -2 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 2 & 1 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 0 & 1 & 1 & 0 & 1 & 1 \\ -1 & 0 & 0 & 1 & -1 & 1 \end{bmatrix}^t$$

† No iterations on the observations is done here; therefore their approximations are the same as their a priori estimates given in the data table.

Then

$$\mathbf{a} = \begin{bmatrix} 1.2 \times 10^{-10} \\ -5.8 \times 10^{-11} \end{bmatrix} \cong \mathbf{0}$$

Next, the cofactor matrices are

$$\mathbf{Q}_{\Delta\Delta} = \mathbf{N}^{-1} = 10^{-2} \begin{bmatrix} 1.25 & 0 \\ 0 & 1.25 \end{bmatrix} \qquad \text{rank} = 2$$

$$\mathbf{Q}_{ii} = 10^{-3} \begin{bmatrix} 2.5 & 0 & 0 & -2.5 & 2.5 & -2.5 \\ & 2.5 & 2.5 & 0 & 2.5 & 2.5 \\ & & 2.5 & 0 & 2.5 & 2.5 \\ & & & 2.5 & -2.5 & 2.5 \\ & & & & 5.0 & 0 \\ \text{symmetric} & & & & & 5.0 \end{bmatrix}$$

and with 4 degrees of freedom, $\hat{\sigma}_0^2 = 0.3123$ and $\hat{\sigma}_0 = 0.56$. Again, for the sake of later checking, the final cofactor matrices are

$$\mathbf{Q}_{\Delta\Delta_{(f)}} = 10^{-2} \begin{bmatrix} 1.212 & 0 \\ 0 & 1.212 \end{bmatrix}$$

$$\mathbf{Q}_{ii_{(f)}} = 10^{-3} \begin{bmatrix} 2.916 & 0 & 0.052 & -2.750 & 2.579 & -2.538 \\ & 2.916 & 2.750 & 0.052 & 2.538 & 2.579 \\ & & 2.595 & 0 & 2.439 & 2.387 \\ & & & 2.595 & -2.387 & 2.439 \\ & & & & 4.489 & 0 \\ \text{symmetric} & & & & & 4.489 \end{bmatrix}$$

$$\hat{\sigma}_{0_{(f)}}^2 = 0.3114 \qquad \text{and} \qquad \hat{\sigma}_{0_{(f)}} = 0.56$$

The a posteriori reference variance value may be statistically tested against the a priori value according to Section 11.5.1. The a priori value is $\sigma_0^2 = 1.0000$ (with infinite degrees of freedom), whereas the a posteriori estimate is $\hat{\sigma}_0^2 = 0.3114$ with $r = 4$ degrees of freedom. Hence, according to Section 2.8.3

$$\chi_m^2 = \frac{ms^2}{\sigma_0^2} \qquad \text{or} \qquad \chi_4^2 = \frac{(4)(0.3114)}{1.0000} = 1.2446$$

and H_0: $\sigma^2 = \sigma_0^2$ versus H_1: $\sigma^2 > \sigma_0^2$. The null hypothesis H_0 is rejected when $\chi_4^2 > \chi_{\alpha, 4}^2$. At the level of significance $\alpha = 0.05$, Table D3 lists $\chi_{0.05, 4}^2 = 9.49$ which is larger than χ_4^2 and therefore H_0 cannot be rejected. Hence the a posteriori estimate of the reference variance is not statistically significantly different from its a priori value.

It should be noted that the differences between the results after the first and last iteration are not significant from the practical standpoint. However, differences can exist and there may be situations in which they are significant and therefore nonlinear problems should be appropriately iterated.

6.3. ADJUSTMENT WITH MAXIMUM NUMBER OF INDEPENDENT PARAMETERS (ADJUSTMENT WITH n CONDITION EQUATIONS)

We now discuss a special case that represents the upper limit and for which $u = n_0$ and $c = n$. This leads to the same form of conditions but with **A** being an $n \times n$ nonsingular matrix, and **B** an $n \times u$ matrix with a rank of u. All relationships developed for the general case apply directly here while taking the change in dimensions of **A** and **B** into account. For example, $\mathbf{Q}_e$ may be computed from equation (6.16) but will be of order n; no change occurs in **N**, **t**, and Δ as they may be obtained from equations (6.21), (6.22), and (6.24), respectively; and **k** can be computed from equation (6.17) noting that it is an $n \times 1$ vector in this case. Without undue repetition, all relations for precision estimation given in Section 6.2.3 also apply here.

There are two possibilities for using this present case of adjustment. The first is when all $u = n_0$ parameters are of interest and need to be estimated. The second is when those parameters of interest are less than n_0, but we wish to write all possible conditions to avoid errors in constructing the model. The Example 6.3 is a useful illustration of the case.

Example 6.3. Refer to Example 6.1, Figure 6.1, dealing with determination of the coordinates of an object point P, from data obtained from three terrestrial photographs. In example 6.1 it was determined that $n_0 = 4$ and $r = 1$. To demonstrate the case of $c = n$ condition equations, the four variables used to determine n_0 are used as unknown parameters, namely,

x_1 as the x_1 coordinate of point P

x_2 as the x_2 coordinate of point P

x_3 as the x_1 coordinate of point S_2

x_4 as the x_1 coordinate of point S_3

With $u = 4$ the number of condition equations will be equal to $n = 5$,

$$l_1 x_2 - p x_1 \qquad = 0$$
$$l_2 x_2 - p(l_4 - x_1) = 0$$
$$l_3 x_2 - p(x_4 - x_1) = 0$$
$$l_4 - x_3 \qquad\quad = 0$$
$$l_4 + l_5 - x_4 \qquad = 0$$

in which l_i refer to the variables representing the observations. Linearization by Taylor series (without iteration on the observations) gives

$$\mathbf{A}_{5,\,5} = \begin{bmatrix} x_2^0 & 0 & 0 & 0 & 0 \\ 0 & x_2^0 & 0 & -p & 0 \\ 0 & 0 & x_2^0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} \qquad \mathbf{B}_{5,\,4} = \begin{bmatrix} -p & l_1 & 0 & 0 \\ p & l_2 & 0 & 0 \\ p & l_3 & 0 & -p \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

$$\mathbf{f}_{5,\,1} = \begin{bmatrix} px_1^0 - l_1 x_2^0 \\ p(l_4 - x_1^0) - l_2 x_2^0 \\ p(x_4^0 - x_1^0) - l_3 x_2^0 \\ x_3^0 - l_4 \\ x_4^0 - l_4 - l_5 \end{bmatrix}$$

Take the approximations as $x_1^0 = 8$ m, $x_2^0 = 50$ m, $x_3^0 = 10$ m, $x_4^0 = 18$ m, and the same $\mathbf{Q}$ matrix given in Example 6.1. The $\mathbf{A}$ and $\mathbf{B}$ matrices are very similar to those computed previously noting that they are augmented by simple numbers and there is no need to rewrite them. The vector $\mathbf{f}$, however, will be

$$\mathbf{f} = [-25 \quad 10 \quad -20 \quad 0 \quad 0]^t$$

For the sake of variation use the partially reduced system of normal equations given in equation (6.35),

$$\begin{bmatrix} \mathbf{Q}_e & \mathbf{B} \\ \mathbf{B}^t & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{k} \\ \Delta \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ \mathbf{0} \end{bmatrix} \qquad \text{or} \qquad \mathbf{M}_{9,\,9} \begin{bmatrix} \mathbf{k} \\ \Delta \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ \mathbf{0} \end{bmatrix}$$

where

$$\mathbf{Q}_e = \begin{bmatrix} 25.0 & 0 & 0 & 0 & 0 \\ & 25.0 & 0 & 0 & 0 \\ & & 25.0 & 0 & 0 \\ & & & 0.0025 & 0.0025 \\ \text{symmetric} & & & & 0.0050 \end{bmatrix} \qquad \text{and}$$

$$\mathbf{B} = \begin{bmatrix} -100 & 16.5 & 0 & 0 \\ 100 & 3.8 & 0 & 0 \\ 100 & 20.4 & 0 & -100 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

The inverse of the 9×9 matrix $\mathbf{M}$, is computed as

$$\mathbf{M}^{-1} = \begin{bmatrix} 0.0040 & 0.0090 & -0.0049 & -0.8952 & 0.4925 & -0.0043 & 0.0283 & -0.0010 & 0.0002 \\ & 0.0199 & -0.0109 & -1.990 & 1.0948 & 0.0027 & 0.0027 & -0.0022 & -0.0003 \\ & & 0.0060 & 1.0948 & -0.6023 & 0.0030 & 0.0256 & 0.0012 & -0.0003 \\ & & & 198.9996 & -109.4767 & -0.2684 & -0.2704 & -0.7762 & -0.0499 \\ & & & & 60.2270 & -0.2995 & -2.5613 & -0.1231 & -0.9726 \\ & & & & & -0.0019 & -0.0050 & -0.0014 & -0.0022 \\ & & & & & \text{II} & -0.0731 & -0.0071 & -0.0135 \\ & & & & & & & -0.0022 & -0.0026 \\ & \text{symmetric} & & & & \text{I} & & \text{III} & -0.0050 \end{bmatrix}$$

The solution of the normal equations would give a 9×1 vector containing both the Lagrange multipliers $\mathbf{k}$ and $\boldsymbol{\Delta}$. The latter is of more interest to our present problem, and thus

$$\boldsymbol{\Delta} = [0.075 \quad -1.193 \quad -0.022 \quad 0.005]' \text{ (m)}$$

Adding these values to the specified approximations gives the estimates of the four parameters

$$\hat{x}_1 = 8.075 \text{ m} = x_1 \text{ coordinate of point } P$$
$$\hat{x}_2 = 48.807 \text{ m} = x_2 \text{ coordinate of point } P$$
$$\hat{x}_3 = 9.978 \text{ m} = x_1 \text{ coordinate of point } S_2$$
$$\hat{x}_4 = 18.005 \text{ m} = x_1 \text{ coordinate point } S_3$$

It is clear that $\hat{x}_1$ and $\hat{x}_2$ are identical to the parameters obtained in Example 6.1. To check $\hat{x}_3$ and $\hat{x}_4$, we should realize by consulting Figure 6.1 that these two parameters are closely related to two of the observations and their determination is tantamount to computing estimated observations. For example, $\hat{x}_3$ is nothing but $\hat{l}_4$, and the answers from here and from Example 6.1 show that they are indeed equal. The estimate $\hat{x}_4$, however, should be equal to $\hat{l}_4 + \hat{l}_5$ or $(9.978 + 8.027) = 18.005$ m, which is identical to the value computed above.

If we are also interested in the a posteriori cofactor matrices, they can be obtained in the usual manner. For example, the cofactor matrix for all four parameters is the negative of the 4×4 submatrix designated (I) in $\mathbf{M}^{-1}$ [see Equation (6.38)]. Within it, the negative of submatrix (II) is the cofactor matrix for the first two parameters that were used in Example 6.1. Considering that all these answers are taken after the first iteration, only the negative of submatrix (II) agrees quite well with $\mathbf{Q}_{\Delta\Delta}$ computed in Example 6.1.

The negative of submatrix (III) in $\mathbf{M}^{-1}$ represents the cofactor matrix of the two parameters $\hat{x}_3$ and $\hat{x}_4$. It can be verified by recognizing that $\hat{x}_3 = \hat{l}_4$ and $\hat{x}_4 = \hat{l}_4 + \hat{l}_5$ or

$$\begin{bmatrix} \hat{x}_3 \\ \hat{x}_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \hat{l}_4 \\ \hat{l}_5 \end{bmatrix}$$

and that from Example 6.1 the cofactor matrix for $\hat{l}_4, \hat{l}_5$ is

$$\begin{bmatrix} 0.00224 & 0.00032 \\ 0.00032 & 0.00211 \end{bmatrix}$$

Thus

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0.00224 & 0.00032 \\ 0.00032 & 0.00211 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0.0022 & 0.0026 \\ 0.0026 & 0.0050 \end{bmatrix}$$

which is identical to the negative of submatrix (III).

To recapitulate, then, the cofactor matrix of all four parameters is

$$\mathbf{Q}_{\Delta\Delta} = 10^{-3} \begin{bmatrix} 1.9 & 5.0 & 1.4 & 2.2 \\ & 73.1 & 7.1 & 13.5 \\ & & 2.2 & 2.6 \\ \text{symmetric} & & & 5.0 \end{bmatrix}$$

and Q_{ii} when computed turn out to be identical to that evaluated in Example 6.1. Finally, the estimate of the reference variance is $\hat{\sigma}_0^2 = 1.9727$.

It should therefore be noted that the choice of four parameters here as compared to two parameters in Example 6.1 has no effect on the evaluation of the remaining variables, v, Q_{vv}, ..., and so on.

For the sake of completeness the final cofactor matrix for the parameters after the last iteration is

$$Q_{\Delta\Delta_{(f)}} = 10^{-3} \begin{bmatrix} 1.862 & 5.029 & 1.420 & 2.170 \\ & 71.332 & 7.106 & 1.347 \\ & & 2.244 & 2.562 \\ \text{symmetric} & & & 4.985 \end{bmatrix}$$

The value of $Q_{ii_{(f)}}$ was also computed and found to be identical to that given in Example 6.1. Finally, the last value of the reference variance is

$$\hat{\sigma}_{0_{(f)}}^2 = 1.9634$$

6.4. GEOMETRIC INTERPRETATION OF THE LEAST SQUARES PRINCIPLE

Although the foundation and derivation of the least squares principle are best done starting from the minimum variance property, a geometrical demonstration of its meaning can give the reader a better appreciation and familiarity with the method. One interpretation is expressed by Brown (see Bibliography):

Consider an n-dimensional coordinate system with orthogonal axes v_1, v_2, ..., v_n (the residuals). Then the quadratic form $\phi = v^t\Sigma^{-1}v$, being positive definite, will represent an n-dimensional ellipsoid. The ellipsoid is centered at the origin. If Σ is diagonal, the axes of the ellipsoid will coincide with the coordinate axes, while for Σ nondiagonal the ellipsoid will be in a tilted orientation. It is clear that by a rotational coordinate transformation, a tilted ellipsoid can be reoriented into a standard position. Such a transformation is specified by $v = R\dot{v}$ where R is an $n \times n$ matrix whose rows (or columns) are composed of the normalized characteristic vectors (eigenvectors) of Σ. Thus a problem involving correlated observations can be reduced to one involving derived observations which are uncorrelated. The dimensions of the hyperellipsoid are, of course, unaffected by a rotation. In fact, the lengths of the axes are directly proportional to the square roots of the characteristic roots (eigenvalues) of Σ. The constant of proportionality which determines their absolute dimensions is simply $\phi^{1/2}$. It thus follows that the volume of the ellipsoid is directly proportional to $\phi^{n/2}$. Therefore, minimizing ϕ is equivalent to minimizing the volume of the ellipsoid, it being understood, naturally, that the condition equations must be satisfied by some point on the ellipsoid. To simplify matters we may assume that any parameters have been eliminated from the linearized condition equations, leaving r relations between the residuals alone. Each condition equation then represents a hyperplane, and the residuals must lie on the intersection of the r hyperplanes. Now consider the family of hyperellipsoids defined by varying ϕ. The orientation and relative dimensions of such ellipsoids will be constant, and all will be centered at the origin. We may think of the family as being formed by the balloonlike expansion of an initial infinitesimal ellipsoid. Let the ellipsoid expand until it becomes tangent to the

intersection of the condition equations hyperplanes. For this point all the condition equations are satisfied and the volume of the ellipsoid (and consequently ϕ) is obviously minimum. Hence the coordinates of the point of tangency give the most probable residuals. . . ."

To demonstrate the statements quoted above we consider the simple case of two observations with the residuals v_1 and v_2.

A two-dimensional coordinate system as shown in Figure 6.3 would represent the space expressed by these two residuals. The condition that may relate v_1 to v_2 can be of the general form $a_0 + a_1 v_1 + a_2 v_2 = 0$ and is depicted by a straight line in the figure. The quadratic form $\phi = \mathbf{v}^t \mathbf{W} \mathbf{v}$ represents in this case a family of ellipses. The dotted ellipse in Figure 6.3, which is tangent to the line of condition, corresponds to the minimum value for ϕ. The point of tangency yields the least squares estimates of v_1 and v_2.
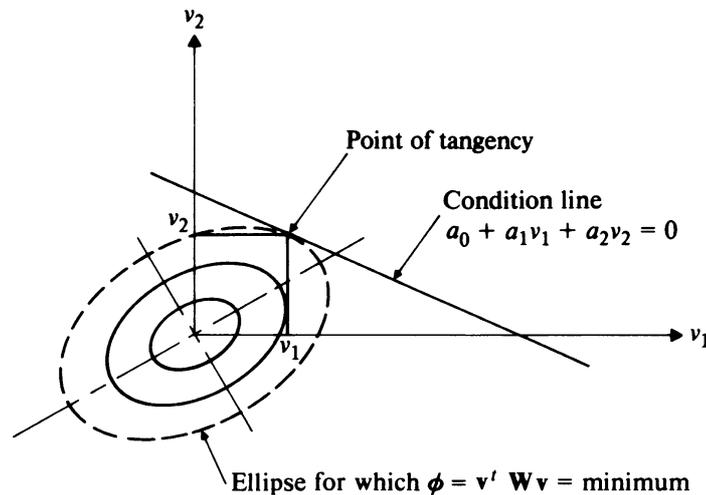

Figure 6.3

The case demonstrated in Figure 6.3 is general in nature and special cases may be derived from it. For example, if we are interested in the equality of the two observations, the straight line of condition would be inclined 45 degrees to the v_1 and v_2 axes. If, further, the two observations are of equal precision (that is, $\sigma_1^2 = \sigma_2^2$), then the ellipse will become a circle. In such a case the point of tangency will be at the foot of the normal from the origin to the line with the obvious result that the estimates of v_1 and v_2 will be equal.

A slightly different interpretation is possible if the observations themselves, and not the residuals, are used as the axes of the coordinate system. As an example consider that the three angles α, β, and γ of a plane triangle are measured with equal precision. With a redundancy of $r = 1$, the well-known condition $\alpha + \beta + \gamma - \pi = 0$ should be satisfied.

The condition represents a plane E_1 in the three-dimensional vector space α, β, γ in Figure 6.4. The a priori values of the observations α, β, γ determine the point P which would not necessarily lie on the plane. Adjustment of the
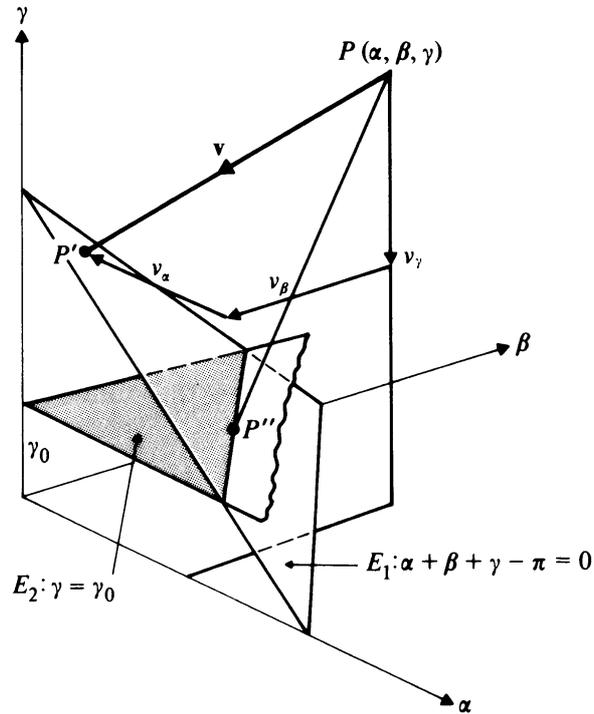
Figure 6.4

given data would replace P by P' such that the least squares principle is satisfied, making the vector v as short as possible, which means making the line $\overline{PP'}$ normal to the plane E_1. This implies directly that the least squares solution would give

$$v_\alpha = v_\beta = v_\gamma = \tfrac{1}{3}(\pi - \alpha - \beta - \gamma)$$

because the plane E_1 is equally inclined to all three axes. All three residuals will have the same sign, being positive if P lies below the plane and negative if it lies above it.

Assume, now, that for one reason or another we want to impose the constraint that the angle γ should have a certain fixed and known value γ_0 in addition to the given data. This additional condition refers to a geometrical model having redundancy $r = 2$, since with γ given, only one observation would suffice to fix the shape of the triangle. The condition $\gamma = \gamma_0$ in the three-dimensional vector space is a plane parallel to the α, β plane. The adjustment in this case has to provide a point P'' that would satisfy both conditions. Thus the point P'' must lie on the intersection of the two planes or on the line ab in Figure 6.4. It will be fixed on this line of intersection by the least squares principle, which would minimize the distance $\overline{PP''}$.

In case the observations α, β, γ are given different a priori weights or are considered correlated, the geometrical explanation of the case remains unaltered if the vector space is spanned by a skew axis system, according to the weight relation.

6.5. SUMMARY OF EQUATIONS

This section is included as an easy reference for those readers who have become familiar with the concepts and wish to extract equations quickly for use in actual adjustment problems. It is a summary of the basic symbols and equations, with reference to the numbers used in the text. Although concise, it is intended to be self-sufficient.

n = number of observations in the vectors l (a priori observations); v (residuals); and $\hat{l}$ (estimated observations)

r = redundancy = $n - n_0$ (where n_0 is the minimum specifying the model)

u = number of parameters in $\Delta (0 < u \leq n_0)$

$c = r + u$ = number of condition equations in

$$\mathbf{Av + B\Delta = d - Al = f} \tag{6.5, 6.6}$$

$\mathbf{A} = c \times n$ matrix of coefficients; rank $(\mathbf{A}) = c$ (upper limit, $\mathbf{A}$ is $n \times n$ with rank n)

$\mathbf{B} = c \times u$ matrix of coefficients; rank $(B) = u$ (upper limit, $\mathbf{B}$ is $n \times u$ with rank $= u$)

$\mathbf{d}$ and $\mathbf{f} = c \times 1$ vectors of constants (upper limit, dimension of $n \times 1$)

$$\mathbf{Q}_e = \mathbf{AQA}^t \tag{6.16}$$

$\mathbf{N} = \mathbf{B}^t\mathbf{W}_e\,\mathbf{B}$ = partially reduced normal

$_{u,\,u}$ equations coefficient matrix   (6.21)

$\mathbf{t} = \mathbf{B}^t\mathbf{W}_e\,\mathbf{f}$ = partially reduced normal
               equations constant vector   (6.22)

$\Delta = \mathbf{N}^{-1}\mathbf{t}$ = estimate of parameters (In the event that the condi-
$_{u,\,1}$ tions are originally nonlinear, then $\mathbf{x} = \mathbf{x}^0 + \Delta$, $\mathbf{x}^0$ is
             a vector of approximations.)   (6.24)

$$\mathbf{v} = \mathbf{QA}^t\mathbf{W}_e(-\mathbf{B}\Delta + \mathbf{f}) \tag{6.13, 6.17}$$

$$\hat{l} = l + \mathbf{v} = \text{estimated observations} \tag{5.2}$$

$$\hat{\sigma}_0^2 = \mathbf{v}^t\mathbf{W}\mathbf{v}/r = (\mathbf{f}^t\mathbf{W}_e\,\mathbf{f} - \Delta^t\mathbf{t})/r \tag{6.25, 6.27}$$

$\mathbf{Q}_{\Delta\Delta} = \mathbf{N}^{-1}$ = cofactor matrix of the parameters
           (rank $\mathbf{Q}_{\Delta\Delta} = u$) For nonlinear con-
           ditions, $\mathbf{Q}_{\hat{x}\hat{x}}$ is also equal to $\mathbf{N}^{-1}$ (6.29)

$$\mathbf{Q}_{vv} = \mathbf{QA}^t(\mathbf{W}_e - \mathbf{W}_e\,\mathbf{BN}^{-1}\mathbf{B}^t\mathbf{W}_e)\mathbf{AQ}$$

$$= \mathbf{QA}^t\mathbf{W}_e\,\mathbf{AQ} - \mathbf{QA}^t\mathbf{W}_e\,\mathbf{BQ}_{\Delta\Delta}\,\mathbf{B}^t\mathbf{W}_e\,\mathbf{AQ} \tag{6.30}$$

= cofactor matrix of the residuals
           (rank $(\mathbf{Q}_{vv}) = r$)   (6.42)

$$\mathbf{Q}_{ll} = \mathbf{Q} - \mathbf{Q}_{vv} \tag{6.31}$$

$$= \text{cofactor matrix of the estimated}$$
$$\text{observations (rank } (\mathbf{Q}_{ll}) = n - r = n_0) \tag{6.43}$$

$$\mathbf{Q}_{\Delta l} = \mathbf{Q}_{\Delta l} = -\mathbf{N}^{-1}\mathbf{B}'\mathbf{W}_e\,\mathbf{A}\mathbf{Q} = \mathbf{Q}_{l\Delta}^t = \mathbf{Q}_{l\Delta}^t \tag{6.32, 6.33}$$

7

Adjustment with Conditions Only—Special Cases

Adjustment of Functionally Independent Parameters and Observations

7.1. INTRODUCTION

The method of least squares has been conventionally used with two cases that are simpler than and special cases of the general case given in the preceding chapter. In this chapter these two cases will be treated and their relationships to the general case shown. The choice from the three techniques will depend mostly on the mathematical model of the problem to be solved, and to a lesser degree on other factors. Such factors may include the type of computational facilities available, the size of the problem, and the preference of the individual involved.

The first of the two special cases involves situations in which only observations appear in the condition equations, thus $u = 0$. The number of condition equations would be equal to the redundancy. Therefore this method is often used for simple geometric problems, such as the adjustment of a plane triangle, a small level net, and the like. The solution of the normal equations, whose number is equal to the number of conditions, yields values for Lagrange multipliers that may be used in turn to compute the vector of residuals, v. By adding v to the given observations l, we compute the

estimated observations, $\hat{l}$. Should we be interested in another set of variables which are related to the elements of $\hat{l}$ by specified functions, we compute such variables *after* the adjustment is completed. The cofactor, or covariance, matrix of these variables is also computed after the adjustment, using $Q_{\hat{l}\hat{l}}$ and the given functions and applying the principle of propagation. Thus the above two operations are considered outside the process of least squares adjustment.

The second special case is more like the general case given in the preceding chapter. In certain simple circumstances it is possible to formulate the condition equations in such a way that each condition equation contains *only one observation*. This implies that parameters are present, that the number of conditions is equal to the number of observations, and that u is at its upper limit of n_0. Because of having only one observation per condition equation, the A matrix reduces to the identity matrix. The number of reduced normal equations is less than the number of condition equations in this case, as it is equal to the number of parameters to be determined. Consequently, their solution yields the values of the parameters and their cofactor matrix directly. If we are also interested in the estimated observations $\hat{l}$, they can still be obtained from the adjustment.

Each of these two cases will now be discussed in detail in separate sections, with numerical examples worked out by both techniques for the sake of comparison.

7.2. ADJUSTMENT OF OBSERVATIONS ONLY

Adjustment of observations only is the technique conventionally referred to as the method of "adjustment by conditions." But since we have utilized the term "condition equations" in a much more general sense, we shall avoid this term. We will simply refer to the technique as "adjustment of observations only."

7.2.1. Formulation Since there are no parameters, the linear condition equations will then simply be

$$A(l + v) = d \tag{7.1}$$

or

$$Av = d - Al = f \tag{7.2}$$

in which d and f are constant vectors.

The application of the least squares in this case follows exactly the same derivation steps as those given in the preceding chapter. Therefore for the sake of brevity, such a derivation is left as an exercise for the reader and only

the summary results will be given here. The total system of normal equations may be written directly from equation (6.11) by enforcing $\mathbf{B} = \mathbf{0}$, thus

$$\begin{bmatrix} -\mathbf{W} & \mathbf{A}^t \\ \mathbf{A} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{v} \\ \mathbf{k} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{f} \end{bmatrix} \qquad (7.3)$$

Solution of equation (7.3) for the two vectors $\mathbf{k}$ and $\mathbf{v}$ yields

$$\mathbf{k} = \mathbf{W}_e\,\mathbf{f} = \mathbf{Q}_e^{-1}\mathbf{f} = (\mathbf{AQA}^t)^{-1}\mathbf{f} \qquad (7.4)$$

$$\mathbf{v} = \mathbf{QA}^t\mathbf{k} = \mathbf{QA}^t\mathbf{W}_e\,\mathbf{f} \qquad (7.5)$$

The matrix $\mathbf{Q}_e$ is the partially reduced normal equations coefficient matrix in this case. The unknown vector in these equations is obviously the vector of Lagrange multipliers, $\mathbf{k}$.

The a priori estimate of the reference variance may be computed from equation (6.25). The quadratic form can be evaluated directly using $\mathbf{v}$ and $\mathbf{W}$, or using equation (6.28) or its equivalent in view of equation (7.4),

$$\mathbf{v}^t\mathbf{W}\mathbf{v} = \mathbf{f}^t\mathbf{k} = \mathbf{k}^t\mathbf{f} \qquad (7.6)$$

7.2.2. Estimation of Precision   In order to apply the principle of propagation to obtain autocofactor and crosscofactor matrices, the pertinant relations are first assembled as follows:

$l = l$ denote $\mathbf{Q}_{ll}$ by $\mathbf{Q}$

$\mathbf{f} = \mathbf{d} - \mathbf{A}l$

$\mathbf{k} = (\mathbf{AQA}^t)^{-1}\mathbf{f} = \mathbf{Q}_e^{-1}\mathbf{f} = \mathbf{W}_e\,\mathbf{d} - \mathbf{W}_e\,\mathbf{A}l = -\mathbf{W}_e\mathbf{A}l + \mathbf{k}_0$

$\mathbf{v} = \mathbf{QA}^t\mathbf{k} = -\mathbf{QA}^t\mathbf{W}_e\,\mathbf{A}l + \mathbf{v}_0$

$\hat{l} = l + \mathbf{v} = (\mathbf{I} - \mathbf{QA}^t\mathbf{W}_e\,\mathbf{A})l + \mathbf{v}_0$

Using the method of substitution we get the following autocofactor and crosscofactor matrices.

Autocofactor Matrices

$$\mathbf{Q}_{ff} = (-\mathbf{A})\mathbf{Q}(-\mathbf{A})^t = \mathbf{AQA}^t = \mathbf{Q}_e$$

$$\mathbf{Q}_{kk} = \mathbf{W}_e\,\mathbf{AQA}^t\mathbf{W}_e = \mathbf{W}_e\,\mathbf{Q}_e\mathbf{W}_e = \mathbf{W}_e$$

$$\mathbf{Q}_{vv} = \mathbf{QA}^t\mathbf{W}_e\,\mathbf{AQA}^t\mathbf{W}_e\,\mathbf{AQ} = \mathbf{QA}^t\mathbf{W}_e\,\mathbf{AQ}$$

$$\mathbf{Q}_{\hat{l}\hat{l}} = (\mathbf{I} - \mathbf{QA}^t\mathbf{W}_e\,\mathbf{A})\mathbf{Q}(\mathbf{I} - \mathbf{A}^t\mathbf{W}_e\,\mathbf{AQ})$$
$$= (\mathbf{I} - \mathbf{QA}^t\mathbf{W}_e\,\mathbf{A})^2\mathbf{Q}$$

(Employing the property of idempotent matrices, see Appendix A.)

$$\mathbf{Q}_{\hat{l}\hat{l}} = \mathbf{Q} - \mathbf{QA}^t\mathbf{W}_e\,\mathbf{AQ}$$
$$= \mathbf{Q} - \mathbf{Q}_{vv}$$

Crosscofactor Matrices

$$Q_{lf} = (I)Q(-A)^t = -QA^t$$

$$Q_{lk} = (I)Q(-W_e A)^t = -QA^tW_e$$

$$Q_{lv} = (I)Q(-QA^tW_e A)^t = -QA^tW_e AQ = -Q_{vv}$$

$$Q_{ll} = (I)Q(I - QA^tW_e A)^t = Q - QA^tW_e AQ = Q_{ll}$$

$$Q_{fk} = (-A)Q(-W_e A)^t = AQA^tW_e = Q_e W_e = I$$

$$Q_{fv} = (-A)Q(-QA^tW_e A)^t = AQA^tW_e AQ = AQ$$

$$Q_{fl} = (-A)Q(I - QA^tW_e A)^t = -AQ + AQ = 0$$

$$Q_{kv} = (-W_e A)Q(-QA^tW_e A)^t = W_e AQA^tW_e AQ = W_e AQ$$

$$Q_{kl} = (-W_e A)Q(I - QA^tW_e A)^t$$
$$= -W_e AQ + W_e AQA^tW_e AQ = 0$$

$$Q_{vl} = (-QA^tW_e A)Q(I - QA^tW_e A)^t$$
$$= -QA^tW_e AQ + QA^tW_e AQA^tW_e AQ = 0$$

Using stepwise propagation we get the same results as follows

Autocofactor Matrices

$$Q_{ff} = AQA^t = Q_e$$

$$Q_{kk} = W_e Q_{ff} W_e = W_e$$

$$Q_{vv} = (QA^t)Q_{kk}(QA^t)^t = QA^tW_e AQ \qquad (7.7)$$

$$Q_{ll} = Q + Q_{vv} + Q_{vl} + Q_{lv}$$

which when using Q_{vv} and $Q_{lv}(= Q_{vl}^t)$ given next, becomes

$$Q_{ll} = Q - QA^tW_e AQ = Q - Q_{vv} \qquad (6.31)$$

Crosscofactor Matrices

$$Q_{lf} = -QA^t$$

$$Q_{lk} = Q_{lf} W_e = -QA^tW_e$$

$$Q_{lv} = Q_{lk}(QA^t)^t = -QA^tW_e AQ = -Q_{vv}$$

$$Q_{ll} = Q + Q_{lv} = Q - Q_{vv} = Q_{ll}$$

$$Q_{fk} = (-A)Q_{lf} W_e = AQA^tW_e = I$$

$$Q_{fv} = (-A)Q_{lk}(QA^t)^t = AQA^tW_e AQ = AQ$$

$$Q_{fl} = (-A)(Q + Q_{lv}) = -AQ + AQA^tW_e AQ = 0$$

or

$$= \mathbf{Q}_{fl} + \mathbf{Q}_{fv} = -\mathbf{AQ} + \mathbf{AQ} = 0$$

$$\mathbf{Q}_{kv} = (\mathbf{W}_e)\mathbf{Q}_{fk}(\mathbf{AQ}) = \mathbf{W}_e \, \mathbf{AQ}$$

$$\mathbf{Q}_{kl} = \mathbf{W}_e(\mathbf{Q}_{fl} + \mathbf{Q}_{fv}) = \mathbf{W}_e(-\mathbf{AQ} + \mathbf{AQ}) = 0$$

or

$$= \mathbf{Q}_{kl} + \mathbf{Q}_{kv} = -\mathbf{W}_e \, \mathbf{AQ} + \mathbf{W}_e \, \mathbf{AQ} = 0$$

$$\mathbf{Q}_{vl} = \mathbf{Q}_{vl} + \mathbf{Q}_{vv} = -\mathbf{Q}_{vv} + \mathbf{Q}_{vv} = 0$$

As shown in Appendix C, the ranks of $\mathbf{Q}_{vv}$ and $\mathbf{Q}_{ll}$ are

$$\text{rank}(\mathbf{Q}_{vv}) = r \tag{6.42}$$

$$\text{rank}(\mathbf{Q}_{ll}) = n - r = n_0 \tag{6.43}$$

Example 7.1. Consider a simple plane triangle with the following observed angles

Angle A: $40°19'02'' = l_1$

Angle B: $70°30'01'' = l_2$

Angle C: $69°11'00'' = l_3$

All angles were measured with equal precision and are assumed to be uncorrelated. It is required to compute the least squares estimates of these three angles, their cofactor matrix and its rank, and an estimate of the reference variance.

Solution: There is one redundant observation and therefore 1 degree of freedom. Thus with $n = 3$ and $c = r = 1$, the condition among the observations is

$$l_1 + l_2 + l_3 - \pi = 0$$

or

$$[1 \quad 1 \quad 1]\mathbf{v} = \pi - l_1 - l_2 - l_3 = -3''$$

From the simple stochastic model given, $\mathbf{Q} = \mathbf{I}$, compute

$$\mathbf{Q}_e = \mathbf{AQA}^t = 3 \qquad \mathbf{k} = \mathbf{W}_e \, \mathbf{f} = -1''$$

$$\mathbf{v} = \mathbf{QA}^t\mathbf{k} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}(-1'') = \begin{bmatrix} -1'' \\ -1'' \\ -1'' \end{bmatrix}$$

Finally, $\hat{l} = l + \mathbf{v}$, or

$\hat{l}_1 = $ angle $A = 40°19'01''$

$\hat{l}_2 = $ angle $B = 70°30'00''$

$\hat{l}_3 = $ angle $C = 69°10'59''$

It can be easily ascertained that $\hat{l}_1 + \hat{l}_2 + \hat{l}_3 - \pi = 0$, which shows that the condition is exactly satisfied after the adjustment. This is equivalent to saying that $\hat{l}$ is consistent with the model.

To compute the cofactor matrix of $\hat{l}$, first compute the cofactor matrix of the residuals, $\mathbf{Q}_{vv}$, using equation (7.7),

$$\mathbf{Q}_{vv} = \tfrac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Although this is a 3×3 matrix, its rank can be easily computed to be 1 (which is equal to r). Next, $\mathbf{Q}_{ll}$ is computed from equation (6.31) as

$$\mathbf{Q}_{ll} = \tfrac{1}{3}\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$$

It can be readily shown that the rank of $\mathbf{Q}_{ll}$ is 2 which is equal to n_0.

Finally, to compute $\hat{\sigma}_0^2$, first $\mathbf{v}^t\mathbf{W}\mathbf{v} = 3 \text{ sec}^2$; and with $r = 1$, then $\hat{\sigma}_0^2 = 3 \text{ sec}^2$. An alternative would be to use equation (7.6), which would give the same value.

Example 7.2. Refer to Example 6.1 and solve the same problem using the technique developed in this section.

Consulting Figure 6.1 and the data given in Example 6.1 for the basic information on the problem, the redundancy is $r = 1$. The stochastic model is expressed by the a priori cofactor matrix,

$$\mathbf{Q} = \text{diag. } \{0.01, 0.01, 0.01, 0.0025, 0.0025\}$$

The first three terms are in square millimeteres and the last two in square metres. In the present case of adjustment, given 1 degree of freedom, one condition equation in terms of *only* the observations is needed. Such a condition equation is not easily obtainable from the geometry of the problem. Therefore first write the three condition equations in Example 6.1 and then eliminate the parameters x_1 and x_2. (It can be argued that if such three condition equations are written first, then adjustment using the technique discussed in the past chapter should be applied. This would be, in fact, the direct way to solve the problem. Elimination of the parameters is done here mainly to provide an exercise for the present technique.)

Using the first two equations solve for x_1 and x_2

$$x_1 = \frac{l_1 l_4}{l_1 + l_2} \quad \text{and} \quad x_2 = \frac{p l_4}{l_1 + l_2}$$

substituting into the third equation leads to

$$-l_1 l_5 - l_2 l_4 - l_2 l_5 + l_3 l_4 = 0$$

where l_i refers to the observational variables.

This is the one condition equation in terms of the five observations. Linearization by Taylor series to the form $\mathbf{A}\mathbf{v} = \mathbf{f}$ gives

$$\begin{aligned} \mathbf{A} &= [-l_5 \quad -(l_4 + l_5) \quad l_4 \quad (l_3 - l_2) \quad -(l_1 + l_2)] \\ &= [-8 \quad -18 \quad 10 \quad 16.6 \quad -20.3] \end{aligned}$$

$$\mathbf{f} = l_1 l_5 + l_2 l_4 + l_2 l_5 - l_3 l_4 = -3.6 \text{ (mm m)}$$

With $\mathbf{Q}$ given above, the normal equation matrix,

$$\mathbf{Q}_e = 6.6 \ (\text{mm}^2 \ \text{m}^2)$$

the Lagrange multiplier is

$$\mathbf{k} = -(6.6)^{-1}3.6 = -0.56 \ (\text{mm}^{-1} \ \text{m}^{-1})$$

and the vector of observational residuals

$$\mathbf{v} = \begin{bmatrix} 0.04 & (\text{mm}) \\ 0.10 & (\text{mm}) \\ -0.05 & (\text{mm}) \\ -0.022 & (\text{m}) \\ 0.027 & (\text{m}) \end{bmatrix}$$

Using these values, the final estimates of the observations will be

$$\hat{l}_1 = 16.54 \ (\text{mm}), \qquad \hat{l}_2 = 3.90 \ (\text{mm}), \qquad \hat{l}_3 = 20.35 \ (\text{mm})$$

$$\hat{l}_4 = 9.978 \ (\text{m}) \qquad \text{and} \qquad \hat{l}_5 = 8.027 \ (\text{m})$$

These values are identical to those computed in Example 6.1. In that example, the coordinates x_1 and x_2 of point P were the quantities of interest. Their estimates are

$$\hat{x}_1 = \frac{\hat{l}_1 \hat{l}_4}{\hat{l}_1 + \hat{l}_2} = \frac{16.54 \ (\text{mm}) \ 9.978 \ (\text{m})}{(16.54 + 3.90) \ (\text{mm})} = 8.07 \ (\text{m})$$

$$\hat{x}_2 = \frac{p \hat{l}_4}{(\hat{l}_1 + \hat{l}_2)} = \frac{100 \ (\text{mm}) \ 9.978 \ (\text{m})}{20.44 \ (\text{mm})} = 48.82$$

These two values are nearly the same as the answers obtained in Example 6.1. Actually, we would obtain exactly the same answers if both cases were iterated properly. The a posteriori cofactor matrix of the estimated observations is

$$\mathbf{Q}_{ll} = 10^{-3} \begin{bmatrix} 9.03 & -2.18 & 1.21 & 0.50 & -0.62 \\ & 5.09 & 2.73 & 1.13 & -1.38 \\ & & 8.48 & -0.63 & 0.77 \\ & & & 2.24 & 0.32 \\ \text{symmetric} & & & & 2.11 \end{bmatrix}$$

It can be shown that the rank of this matrix is $(n - r) = (5 - 1) = 4$.

To compute $\mathbf{Q}_{\hat{x}\hat{x}}$ we apply propagation principle on the formulas

$$\hat{\mathbf{x}} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} = \begin{bmatrix} \dfrac{\hat{l}_1 \hat{l}_4}{\hat{l}_1 + \hat{l}_2} \\[2mm] \dfrac{p \hat{l}_4}{\hat{l}_1 + \hat{l}_2} \end{bmatrix} = F(\hat{l}')$$

$$\underset{2,\,3}{\mathbf{J}} = \frac{\partial \hat{\mathbf{x}}}{\partial \hat{l}'} = \begin{bmatrix} \left(\dfrac{\hat{x}_1}{\hat{l}_1} - \dfrac{\hat{x}_1}{\hat{l}_1 + \hat{l}_2} \right) & -\dfrac{\hat{x}_1}{\hat{l}_1 \hat{l}_2} & \dfrac{\hat{x}_1}{\hat{l}_4} \\[4mm] \dfrac{-\hat{x}_2}{\hat{l}_1 + \hat{l}_2} & -\dfrac{\hat{x}_2}{\hat{l}_1 + \hat{l}_2} & \dfrac{\hat{x}_2}{p} \end{bmatrix}$$

or

$$\mathbf{J} = \begin{bmatrix} 0.0931 & -0.3950 & 0.8093 \\ -2.3877 & -2.3877 & 4.8919 \end{bmatrix}$$

Because the two values $\hat{x}_1$ and $\hat{x}_1$ are computed from only a subvector $\hat{l}'$ of the total vector $\hat{l}$, the pertinent cofactor matrix $\mathbf{Q}_{l'l'}$ is constructed from $\mathbf{Q}_{ll}$ by deleting the latter's third and fifth rows and columns. Thus

$$\mathbf{Q}_{l'l'} = 10^{-3} \begin{bmatrix} 9.03 & -2.18 & 0.50 \\ & 5.09 & 1.13 \\ \text{symmetric} & & 2.24 \end{bmatrix}$$

The cofactor matrix of the derived parameters is therefore

$$\mathbf{Q}_{\hat{x}\hat{x}} = \mathbf{J}\mathbf{Q}_{l'l'}\mathbf{J}' = 10^{-3} \begin{bmatrix} 1.851 & 4.968 \\ 4.968 & 71.006 \end{bmatrix} \quad \text{rank} = 2$$

Finally, the estimate of the reference variance is $\hat{\sigma}_0^2 = 2.016$.

For reference, the values after the final iteration (f) are

$$\mathbf{Q}_{ll_{(f)}} = 10^{-3} \begin{bmatrix} 9.024 & -2.189 & 1.213 & 0.500 & -0.621 \\ & 5.090 & 2.721 & 1.121 & -1.394 \\ & & 8.492 & -0.621 & 0.772 \\ & & & 2.244 & 0.318 \\ \text{symmetric} & & & & 2.104 \end{bmatrix}$$

$$\mathbf{Q}_{\hat{x}\hat{x}_{(f)}} = 10^{-3} \begin{bmatrix} 1.86 & 5.03 \\ 5.03 & 71.33 \end{bmatrix}$$

and

$$\hat{\sigma}_{0_{(f)}}^2 = 1.9633$$

which agree well with those computed in Example 6.1.

Example 7.3. Consider the problem of Example 6.2, the two-parameter transformation, for adjustment with observations only.

In Example 6.2 the redundancy was evaluated as $r = 4$ for the given six observations. Inspection of the problem does not lead, at least not directly, to a simple way of writing four equations in terms of the observations. In fact this type of problem (transformation) is rather awkward to solve using the technique of adjustment of observations only.

Writing a pair of transformation equations for any of the three given points, solve for a and b in terms of observations and constants. For example, using the pair for the first point,

$$y_{11} = ax_{11} - bx_{21}$$

$$y_{21} = bx_{11} + ax_{21}$$

then

$$a = \frac{x_{11}y_{11} + x_{21}y_{21}}{x_{11}^2 + x_{21}^2} = \alpha \qquad b = \frac{x_{11}y_{21} - x_{21}y_{11}}{x_{11}^2 + x_{21}^2} = \beta$$

Substituting these in the transformation equations gives

$$y_{1i} = \alpha x_{1i} - \beta x_{2i}$$

$$y_{2i} = \alpha x_{2i} + \beta x_{1i}$$

or

$$f_{1i} = (x_{11}y_{11} + x_{21}y_{21})x_{1i} - (x_{11}y_{21} - x_{21}y_{11})x_{2i} - y_{1i}(x_{11}^2 + x_{21}^2) = 0$$

$$f_{2i} = (x_{11}y_{11} + x_{21}y_{21})x_{2i} + (x_{11}y_{21} - x_{21}y_{11})x_{1i} - y_{2i}(x_{11}^2 + x_{21}^2) = 0$$

with $i = 2, 3$.

These are four condition equations involving the six observations only. From the given data, the stochastic model is represented by the cofactor matrix of the observations $\mathbf{Q} = 0.01 \, \mathbf{I}_6$.

Linearization: The condition equations may now be linearized to the form,

$$\underset{4, 6}{\mathbf{A}} \; \underset{6, 1}{\mathbf{v}} = \underset{4, 1}{\mathbf{f}}$$

where

$$\mathbf{v} = [v_{x_{11}} v_{x_{21}} v_{x_{12}} v_{x_{22}} v_{x_{13}} v_{x_{23}}]^t$$

and

$$a_{11} = \frac{\partial f_{12}}{\partial x_{11}} = y_{11}x_{12} - y_{21}x_{22} - 2y_{12}x_{11}$$

$$a_{12} = \frac{\partial f_{12}}{\partial x_{21}} = y_{21}x_{12} + y_{11}x_{22} - 2y_{12}x_{21}$$

$$a_{13} = \frac{\partial f_{12}}{\partial x_{12}} = x_{11}y_{11} + x_{21}y_{21}$$

$$a_{14} = \frac{\partial f_{12}}{\partial x_{22}} = -x_{11}y_{21} + x_{21}y_{11}$$

$$a_{15} = \frac{\partial f_{12}}{\partial x_{13}} = 0$$

$$a_{16} = \frac{\partial f_{12}}{\partial x_{23}} = 0$$

$$a_{21} = \frac{\partial f_{22}}{\partial x_{11}} = y_{11}x_{22} + y_{21}x_{12} - 2y_{22}x_{11}$$

$$a_{22} = \frac{\partial f_{22}}{\partial x_{21}} = y_{21}x_{22} - y_{11}x_{12} - 2y_{22}x_{21}$$

$$a_{23} = \frac{\partial f_{22}}{\partial x_{12}} = x_{11}y_{21} - x_{21}y_{11}$$

$$a_{24} = \frac{\partial f_{22}}{\partial x_{22}} = x_{11} y_{11} + x_{21} y_{21}$$

$$a_{25} = \frac{\partial f_{22}}{\partial x_{13}} = 0$$

$$a_{26} = \frac{\partial f_{22}}{\partial x_{23}} = 0$$

$$a_{31} = \frac{\partial f_{13}}{\partial x_{11}} = y_{11} x_{13} - y_{21} x_{23} - 2y_{13} x_{11}$$

$$a_{32} = \frac{\partial f_{13}}{\partial x_{21}} = y_{21} x_{13} + y_{11} x_{23} - 2y_{13} x_{21}$$

$$a_{33} = \frac{\partial f_{13}}{\partial x_{12}} = 0$$

$$a_{34} = \frac{\partial f_{13}}{\partial x_{22}} = 0$$

$$a_{35} = \frac{\partial f_{13}}{\partial x_{13}} = a_{13}$$

$$a_{36} = \frac{\partial f_{13}}{\partial x_{23}} = a_{14}$$

$$a_{41} = \frac{\partial f_{23}}{\partial x_{11}} = y_{11} x_{23} + y_{21} x_{13} - 2y_{23} x_{11}$$

$$a_{42} = \frac{\partial f_{23}}{\partial x_{21}} = y_{21} x_{23} - y_{11} x_{13} - 2y_{23} x_{21}$$

$$a_{43} = \frac{\partial f_{23}}{\partial x_{12}} = 0$$

$$a_{44} = \frac{\partial f_{23}}{\partial x_{22}} = 0$$

$$a_{45} = \frac{\partial f_{23}}{\partial x_{13}} = a_{23}$$

$$a_{46} = \frac{\partial f_{23}}{\partial x_{23}} = a_{24}$$

The elements of the vector $\mathbf{f}_{4,1}$ is the evaluation of the four equations with the given observations. The following are the numerical values:

$$\mathbf{A}_{4,6} = \begin{bmatrix} -2.1 & -0.9 & 1.1 & -2.1 & 0.0 & 0.0 \\ 1.1 & -1.9 & 2.1 & 1.1 & 0.0 & 0.0 \\ -3.2 & 0.8 & 0.0 & 0.0 & 1.1 & -2.1 \\ -1.0 & -2.4 & 0.0 & 0.0 & 2.1 & 1.1 \end{bmatrix}$$

$$\mathop{\mathbf{f}}_{4,\,1} = \begin{bmatrix} 0.10 \\ 0.10 \\ -0.10 \\ 0.40 \end{bmatrix}$$

$$\mathbf{Q}_e = \begin{bmatrix} 10.84 & -0.60 & 6.0 & 4.26 \\ & 10.44 & -5.04 & 3.46 \\ & & 16.50 & 1.28 \\ \text{symmetric} & & & 12.38 \end{bmatrix}$$

$$\mathbf{W}_e = \begin{bmatrix} 13.36 & -0.074 & -4.56 & -4.10 \\ & 13.00 & 4.31 & -4.05 \\ & & 9.08 & -0.57 \\ \text{symmetric} & & & 10.68 \end{bmatrix}$$

$$\mathbf{k} = \mathbf{W}_e\,\mathbf{f} = \begin{bmatrix} -0.143 \\ 0.760 \\ 1.163 \\ -3.514 \end{bmatrix}$$

$$\mathbf{v} = \begin{bmatrix} 0.0093 \\ 0.0805 \\ 0.0144 \\ 0.0114 \\ -0.0610 \\ -0.0631 \end{bmatrix} \qquad \begin{aligned} \hat{x}_{11} &= 0.0093 \\ \hat{x}_{21} &= 1.0805 \\ \hat{x}_{12} &= 1.0144 \\ \hat{x}_{22} &= 0.0114 \\ \hat{x}_{13} &= 0.9390 \\ \hat{x}_{23} &= 0.9369 \end{aligned}$$

Using the estimated observations compute the transformation parameters

$$\hat{a} = \frac{\hat{x}_{11}\,\hat{y}_{11} + \hat{x}_{21}\,\hat{y}_{21}}{\hat{x}_{11}^2 + \hat{x}_{21}^2} = 1.00$$

$$\hat{b} = \frac{\hat{x}_{11}\,\hat{y}_{21} - \hat{x}_{21}\,\hat{y}_{11}}{\hat{x}_{11}^2 + \hat{x}_{21}^2} = 1.95$$

which are identical to those computed in Example 6.2. It should be obvious that for this type of problem, and for transformation problems in general, the general case of adjustment of observations and parameters combined given in the preceding chapter is more appropriate. No special effort is necessary in formulating the condition equations. Furthermore, for this particular example the general technique gave rise to only two reduced normal equations, a number that remains fixed even when the number of observations is increased. In addition, the 2 × 2 matrix of the inverse of the normal equations coefficient matrix is automatically the cofactor matrix of the parameters. By contrast, adjustment of observations only requires special effort to formulate the condition equations, and the number of these equations is dependent on the number of observations. Furthermore, both the parameters and their cofactor matrix require further computation after obtaining the estimated observations and their cofactor matrix from the adjustment.

For the sake of completeness, Q_{ll} and $\hat{\sigma}_0^2$ are

$$\mathbf{Q}_{ll} = 10^{-3} \begin{bmatrix} 2.501 & 0.055 & 0.049 & -2.499 & 2.541 & -2.459 \\ & 3.420 & 3.030 & 0.067 & 2.586 & 2.574 \\ & & 2.685 & 0.059 & 2.292 & 2.281 \\ & & & 2.501 & -2.449 & 2.550 \\ & & & & 4.456 & -0.554 \\ \text{symmetric} & & & & & 4.437 \end{bmatrix}$$

$$\hat{\sigma}_0^2 = 0.2656$$

7.2.3. Further Examples on Adjustment of Observations Only

Example 7.4. *Adjustment of a Level Net*

Given: Figure 7.1 shows a small level net involving eight lines, four points (B, C, D, and E), and a known elevation of the starting point A of 800.000 metres. The observations, which are the differences in elevation along the indicated lines, are given in Table 7.1. These observations (the elevation differences) are assumed to be uncorrelated. However, the weight of each elevation difference in square metres has
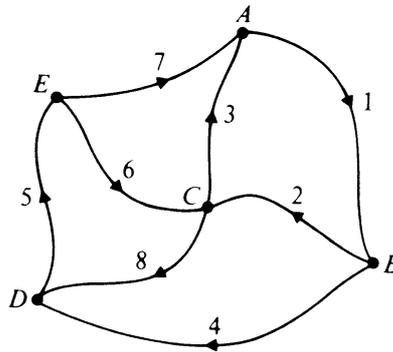

Figure 7-1

TABLE 7-1

LINE NUMBER	DISTANCE (km)	ELEVATION DIFFERENCE (m)
1	18.1	+ 25.42
2	9.4	+ 10.34
3	14.2	− 35.20
4	17.6	− 15.54
5	13.5	+ 21.32
6	9.9	+ 4.82
7	13.8	− 31.02
8	14.0	− 26.11

been found empirically to be equal to the reciprocal of the corresponding distance taken in 10 km units. Thus, for example, the weight for the first elevation difference (line 1 between points A and B) is $w_1 = (1/1.8)$ m^{-2}.

Required: Compute the least squares estimates of the elevations of points B, C, D, and E, and their precision.

Solution: The functional model of the problem is concerned with a simple geometric case of vertical distances between points on the earth's surface within a relatively small area. *Four variables* are needed to determine uniquely the elevations of the four points in the problem. Therefore with eight observations, the redundancy is $r = 4$.

There are two possible adjustment techniques. In the first, four condition equations involving only observations are written and values for all eight estimated observations, $\hat{l}$, are obtained. Using these and simple geometric relations, we would then compute the required elevations of the four points. Also, using $Q_{\hat{l}\hat{l}}$ from the adjustment and applying the rules of propagation, the cofactor matrix of the computed elevations can be evaluated. In the second method the elevations of points B, C, D, and E are introduced as unknown parameters and their values as well as the corresponding cofactor matrix are computed directly from the adjustment. In the present example we shall consider the first case of adjustment with observations only.

Going around a "loop," *in one direction*, the added elevation differences must be zero. Using the four obvious loops (ABC, ACE, BDC, and DEC), and noting that a value is positive along the given arrow and negative when opposite to the given arrow, we may write the following four independent condition equations:

$$l_1 + l_2 + l_3 \qquad\qquad = 0$$
$$\qquad\quad - l_3 \qquad\quad - l_6 + l_7 \quad = 0$$
$$- l_2 \qquad + l_4 \qquad\qquad - l_8 = 0$$
$$\qquad\qquad l_5 + l_6 \qquad + l_8 = 0$$

It should be pointed out that several other equations may be written. It is therefore important not only to have the right number of equations, but also that they be independent. For example, the sum of the first two equations would correspond to the loop $ABCE$, but it is obviously a dependent equation.

Rewriting these equations in the matrix form $Av = f$, thus

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \\ v_8 \end{bmatrix} = \begin{bmatrix} -l_1 - l_2 - l_3 \\ l_3 + l_6 - l_7 \\ l_2 - l_4 + l_8 \\ -l_5 - l_6 - l_8 \end{bmatrix} = \begin{bmatrix} -0.56 \text{ m} \\ +0.64 \text{ m} \\ -0.23 \text{ m} \\ -0.03 \text{ m} \end{bmatrix}$$

From the information given in the problem, the weight matrix of the given observations is the following diagonal matrix:

$$W = \text{diag.} \left\{ \frac{1}{1.81}, \frac{1}{0.94}, \frac{1}{1.42}, \frac{1}{1.76}, \frac{1}{1.35}, \frac{1}{0.99}, \frac{1}{1.38}, \frac{1}{1.40} \right\} (\text{m}^{-2})$$

and the cofactor matrix is

$$\mathbf{Q} = \text{diag. } \{1.81, 0.94, 1.42, 1.76, 1.35, 0.99, 1.38, 1.40\}(\text{m}^2)$$

Using the above values for $\mathbf{A}$ and $\mathbf{Q}$, the reduced normal equations coefficient matrix is

$$\mathbf{Q}_e = \begin{bmatrix} 4.17 & -1.42 & -0.94 & 0 \\ -1.42 & 3.79 & 0 & -0.99 \\ -0.94 & 0 & 4.10 & -1.40 \\ 0 & -0.99 & -1.40 & 3.74 \end{bmatrix}(\text{m}^2)$$

and its inverse

$$\mathbf{W}_e = \begin{bmatrix} 0.3065 & 0.1333 & 0.0944 & 0.0706 \\ 0.1333 & 0.3445 & 0.0707 & 0.1177 \\ 0.0944 & 0.0707 & 0.3118 & 0.1354 \\ 0.7060 & 0.1177 & 0.1354 & 0.3492 \end{bmatrix}(\text{m}^{-2})$$

The vector of Lagrange multipliers is computed as

$$\mathbf{k} = \begin{bmatrix} -0.1102 & 0.1261 & -0.0834 & -0.0059 \end{bmatrix}^t (\text{m}^{-1})$$

from which, the vector of residuals is computed to be

$$\mathbf{v} = \begin{bmatrix} -0.199 & -0.025 & -0.335 & -0.147 & -0.008 \\ -0.131 & 0.174 & 0.109 \end{bmatrix}^t (\text{m})$$

The estimated observations $\hat{l}$ are obtained by adding the vector of residuals $\mathbf{v}$ to the given observations l.

$$\hat{l} = \begin{bmatrix} 25.220 & 10.315 & -35.535 & -15.685 & 21.312 & 4.689 & -30.846 \\ -26.001 \end{bmatrix}^t (\text{m})$$

To compute the elevations of points, B, C, D, and E, denote them by x_1, x_2, x_3, and x_4, respectively, and the elevation of point A by x_0. Thus, from Figure 7.1,

$$x_1 = x_0 + \hat{l}_1 \quad = 825.220 \text{ m} = \text{elev. of point } B$$

$$x_2 = x_0 - \hat{l}_3 \quad = 835.535 \text{ m} = \text{elev. of point } C$$

$$x_3 = x_0 + \hat{l}_1 + \hat{l}_4 = 809.535 \text{ m} = \text{elev. of point } D$$

$$x_4 = x_0 - \hat{l}_7 \quad = 830.846 \text{ m} = \text{elev. of point } E$$

In order to compute the cofactor matrix for these elevations, first compute $\mathbf{Q}_{ll}$

$$\mathbf{Q}_{ll} = \begin{bmatrix} 0.806 & -0.361 & -0.445 & -0.301 & -0.173 & 0.112 & -0.333 & 0.061 \\ & 0.560 & -0.200 & 0.360 & 0.082 & 0.118 & -0.081 & -0.201 \\ & & 0.645 & -0.059 & 0.090 & -0.231 & 0.414 & 0.141 \\ & & & 0.794 & -0.322 & -0.113 & -0.172 & -0.434 \\ & & & & 0.714 & -0.309 & -0.219 & -0.404 \\ & & & & & 0.541 & 0.310 & -0.091 \\ & & & & & & 0.724 & -0.091 \\ \text{symmetric} & & & & & & & 0.635 \end{bmatrix}$$

Referring to the vector of four elevations by $\mathbf{x}$, we can write $\mathbf{x} = \mathbf{J}\hat{\mathbf{l}}$ where

$$\mathbf{J} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \end{bmatrix}$$

and through propagation compute

$$\mathbf{Q}_{xx} = \begin{bmatrix} 0.806 & 0.445 & 0.505 & 0.333 \\ & 0.645 & 0.504 & 0.414 \\ & & 0.999 & 0.505 \\ \text{symmetric} & & & 0.724 \end{bmatrix} (\text{m}^2)$$

If we were interested in the covariance matrix of the elevations, instead of the cofactor matrix which gives only relative values, we would need to scale the latter by the reference variance $\hat{\sigma}_0^2$. Since the a priori precision estimation was only in the form of an empirically derived weight matrix, we can only use the a posteriori estimate of the reference variance (that is, in this case there is no possibility of performing a statistical test on $\hat{\sigma}_0^2$).

The computation of the reference variance leads to

$$\hat{\sigma}_0^2 = 0.0404$$

Thus the covariance matrix of the elevations is

$$\hat{\mathbf{\Sigma}}_{xx} = \hat{\sigma}_0^2 \mathbf{Q}_{xx} = \begin{bmatrix} 0.0326 & 0.0180 & 0.0204 & 0.0135 \\ & 0.0261 & 0.0204 & 0.0167 \\ & & 0.0404 & 0.0204 \\ \text{symmetric} & & & 0.0293 \end{bmatrix} (\text{m}^2)$$

One last remark regarding the computation of $\mathbf{Q}_{xx}$ is worth mentioning here. Examining the matrix $\mathbf{J}$, we see that it contains four zero columns, the second, fifth, sixth, and eighth. This is because the computation of $\mathbf{x}$ requires the use of $\hat{l}_1, \hat{l}_3, \hat{l}_4$, and $\hat{l}_7$ only. Therefore if the computations are performed by simple means (by hand, desk calculator, or a small desk computer), it would be economical to shorten the operation of matrix multiplication in the following way. First compact $\mathbf{J}$ to a size 4×4, thus

$$\begin{array}{cccc} \underline{\quad 1 \quad} & \underline{\quad 3 \quad} & \underline{\quad 4 \quad} & \underline{\quad 7 \quad} \end{array} \text{ columns of } \mathbf{J}$$

$$\underset{4,\,4}{\mathbf{J}'} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

noting that the columns now correspond to $\hat{l}_1, \hat{l}_3, \hat{l}_4$, and $\hat{l}_7$, respectively. Using this fact, the submatrix of $\mathbf{Q}_{l'l'}$ which is needed for computing $\mathbf{Q}_{xx}$ would be a 4×4 matrix and is constructed using elements pertinent to the above four quantities, or

$$\underset{4,\,4}{\mathbf{Q}_{l'l'}} = \begin{bmatrix} 0.806 & -0.445 & -0.301 & -0.333 \\ & 0.645 & -0.059 & 0.414 \\ & & 0.794 & -0.172 \\ \text{symmetric} & & & 0.724 \end{bmatrix}$$

Using $\mathbf{J}'$ and $\mathbf{Q}_{l'l'}$ into the propagation formula, we would get exactly the same $\mathbf{Q}_{xx}$ as computed above. It must be emphasized, however, that if the problem is being solved by a standardized adjustment program, the above modification need not be applied.

Example 7.5. *Quadrilateral Adjustment by Angles and by Directions.* There has been some discussion in the literature regarding the use of angles versus the use of directions when adjusting triangulation nets. Although the method of observing angles may be different and preferable over that for observing directions, the concern here is not with particular surveying techniques. What should be demonstrated is that as far as adjustment computation is concerned, the two techniques of using angles and directions are identical; that is, starting with the same numerical data, we would arrive at exactly the same answer if we construct corresponding mathematical models properly and carry out the computations correctly. In this example we shall ascertain this fact by a numerical example concerning the adjustment of a single quadrilateral.

Figure 7.2 shows a quadrilateral (in a plane) *ABCD* with the two diagonals *AC* and *BD* connected. It shall be assumed at the beginning that the area covered by the quadrilateral is sufficiently small to allow working with plane trigonometry.
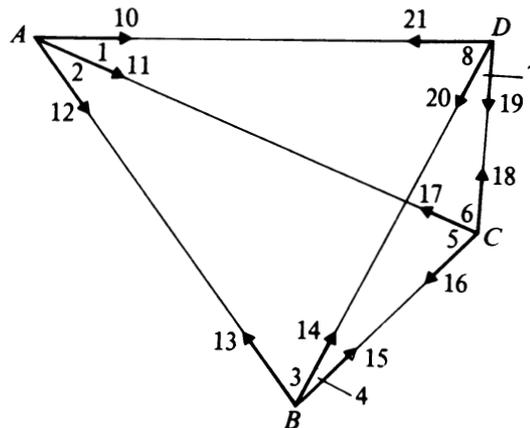

Figure 7-2

At each of the four vertices a theodolite was set up and readings along all three lines were taken. These will be considered here as directions. Any angle at one of the vertices will be simply the difference between the two directions which form that angle. Under these assumptions, then, the given observed data is listed in Table 7.2. For simplicity, all the directions in the table will be assumed as uncorrelated and of equal precision. The corresponding observed angles are fixed as in Table 7.3 once the directions are given.

Adjustment Using Directions: The fundamental observation is assumed to be a single direction which is one reading on a horizontal circle. Such a reading is actually meaningless in itself since no useful information can be derived from a *single* direction. Unless *at least* two directions are designated, we would actually establish nothing by an individual direction. Therefore with a set of directions from one point a

TABLE 7-2			TABLE 7-3	
DIRECTION, d	VALUE		ANGLE l	VALUE
10	$00°00'00''00$		1	$22°01'42''51$
11	$22°01'42''51$		2	$16°44'31''20$
12	$38°46'13''71$		3	$57°08'57''10$
13	$00°00'00''00$		4	$19°33'14''13$
14	$57°08'57''10$		5	$86°33'13''45$
15	$76°42'11''23$		6	$58°46'35''93$
16	$00°00'00''00$		7	$15°06'52''28$
17	$86°33'13''45$		8	$84°04'50''66$
18	$145°19'49''38$			
19	$00°00'00''00$			
20	$15°06'52''28$			
21	$99°11'42''94$			

reference is attached that may be the zero direction and that may or may not be known. To analyze the given information, which is composed of four sets of three directions each, begin by determining the minimum number of observations that are necessary to define the quadrilateral. Starting with point A, use directions 10, 11, and 12. Since the size of the quadrilateral is immaterial as the problem involves only its shape, choose point C anywhere along the direction 11. From C, use directions 16, 17, and 18 to complete the quadrilateral. At this point we would suspect that six observations have been used to determine the required figure leaving six extra observed directions. *This does not mean, however, that there are 6 degrees of freedom* $(r = 6)$. As explained above, a direction taken at a point does not provide a useful and independent piece of information unless it is referred to another from the same point. Consequently, for each of the two sets of directions observed at points B and D, there exists an inherent parameter representing the zero direction. Since these two zero directions were not designated, they are to be considered as unknown variables of the model. Hence, six extra observations minus two parameters lead to *only 4 degrees of freedom*, or a redundancy of $r = 4$.

In this problem the interest is in showing that the use of either directions or angles leads to the same adjustment results, and therefore the technique of adjustment to be chosen is not of primary importance. In the present case we shall be comparing the estimated observations and their cofactor matrix, and thus the case of adjustment involving observations only using $c = r = 4$ independent condition equations is appropriate.

The reader who is familiar with surveying techniques will recognize that there are, for the present case, *only three independent angular conditions* in a quadrilateral and *one* so-called *side condition*. The side condition guarantees that one side computed from two different sets of triangles must be the same. It is actually the one condition which ascertains that we do have a quadrilateral (otherwise, we would have adjacent triangles). Designating directions by d (as used in Table 7.2) the following are four independent condition equations:

$$-d_{10} + d_{12} - d_{13} + d_{14} - d_{20} + d_{21} = 0 \left.\vphantom{\begin{matrix}1\\1\\1\end{matrix}}\right\}$$
$$-d_{14} + d_{15} - d_{16} + d_{18} - d_{19} + d_{20} = 0 \left.\vphantom{\begin{matrix}1\\1\\1\end{matrix}}\right\} \quad \text{angle conditions}$$
$$-d_{10} + d_{11} - d_{17} + d_{18} - d_{19} + d_{21} = 0 \left.\vphantom{\begin{matrix}1\\1\\1\end{matrix}}\right.$$

$$\frac{\sin(d_{11} - d_{10}) \sin(d_{20} - d_{19}) \sin(d_{17} - d_{16}) \sin(d_{14} - d_{13})}{\sin(d_{12} - d_{11}) \sin(d_{21} - d_{20}) \sin(d_{18} - d_{17}) \sin(d_{15} - d_{14})} - 1 = 0$$

side condition

Linearization by Taylor series: Let us denote these four condition equations by $\dot{f}_1$, $\dot{f}_2$, $\dot{f}_3$, and $\dot{f}_4$. Then

$$
\begin{bmatrix}
-1 & 0 & 1 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \\
0 & 0 & 0 & 0 & -1 & 1 & -1 & 0 & 1 & -1 & 1 & 0 \\
-1 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & -1 & 0 & 1 \\
a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} & a_{47} & a_{48} & a_{49} & a_{4,10} & a_{4,11} & a_{4,12}
\end{bmatrix}
$$

$$
\times
\begin{bmatrix}
v_{10} \\
v_{11} \\
v_{12} \\
\vdots \\
v_{21}
\end{bmatrix}
=
\begin{bmatrix}
-\dot{f}_1 \\
-\dot{f}_2 \\
-\dot{f}_3 \\
-\dot{f}_4
\end{bmatrix}
$$

which is of the form

$$\underset{4,12}{\mathbf{A}} \quad \underset{12,1}{\mathbf{v}} = \underset{4,1}{\mathbf{f}}$$

where

		Value (sec)

$$f_1 = +d_{10} - d_{12} + d_{13} - d_{14} + d_{20} - d_{21} \qquad = -1.4700$$

$$f_2 = +d_{14} - d_{15} + d_{16} - d_{18} + d_{19} - d_{20} \qquad = 4.2100$$

$$f_3 = +d_{10} - d_{11} + d_{17} - d_{18} + d_{19} - d_{21} \qquad = -1.3800$$

$$f_4 = 1 - \frac{\sin(d_{11} - d_{10}) \sin(d_{20} - d_{19}) \sin(d_{17} - d_{16}) \sin(d_{14} - d_{13})}{\sin(d_{12} - d_{11}) \sin(d_{21} - d_{20}) \sin(d_{18} - d_{17}) \sin(d_{15} - d_{14})} = -6.2670$$

If the fourth equation is

$$\dot{f}_4 = \frac{U}{V} - 1$$

then the elements of the fourth row of the **A** matrix can be evaluated as

$$a_{4,1} = -(U/V) \cot(d_{11} - d_{10}) \qquad\qquad = -2.47162$$

$$a_{4,2} = (U/V)[\cot(d_{11} - d_{10}) + \cot(d_{12} - d_{11})] = 5.79604$$

$$a_{4,3} = -(U/V) \cot(d_{12} - d_{11}) \qquad\qquad = -3.32242$$

$$a_{4,4} = -(U/V) \cot(d_{14} - d_{13}) \qquad\qquad = -0.64573$$

$$a_{4,5} = (U/V)[\cot(d_{14} - d_{13}) + \cot(d_{15} - d_{14})] = 3.46131$$

$$a_{4, 6} = -(U/V) \cot (d_{15} - d_{14}) \qquad\qquad = -2.81557$$

$$a_{4, 7} = -(U/V) \cot (d_{17} - d_{16}) \qquad\qquad = -0.06022$$

$$a_{4, 8} = (U/V)[\cot (d_{17} - d_{16}) + \cot (d_{18} - d_{17})] = \quad 0.66642$$

$$a_{4, 9} = -(U/V) \cot (d_{18} - d_{17}) \qquad\qquad = -0.60620$$

$$a_{4, 10} = -(U/V) \cot (d_{20} - d_{19}) \qquad\qquad = -3.70255$$

$$a_{4, 11} = (U/V)[\cot (d_{20} - d_{19}) + \cot (d_{21} - d_{20})] = -3.80623$$

$$a_{4, 12} = -(U/V) \cot (d_{21} - d_{20}) \qquad\qquad = -0.10368$$

Since the observed directions were given as uncorrelated and of equal precision, then

$$\mathbf{Q}_e = \begin{bmatrix} 6.0 & -2.0 & 2.0 & -0.6557 \\ & 6.0 & 2.0 & 0.6859 \\ & & 6.01 & 10.5939 \\ \text{symmetric} & & & 100.1021 \end{bmatrix} (\text{sec of arc})^2$$

and

$$\mathbf{Q}_e^{-1} = \mathbf{W}_e = \begin{bmatrix} 0.2773 & 0.1490 & -0.1765 & 0.0195 \\ & 0.2712 & -0.1703 & 0.0171 \\ & & 0.3471 & -0.0367 \\ \text{symmetric} & & & 0.0139 \end{bmatrix} (\text{sec of arc})^{-2}$$

The vector of four Lagrange multipliers is

$$\mathbf{k} = [0.3414 \quad 1.0502 \quad -0.7066 \quad 0.0072]^t (\text{sec of arc})^{-1}$$

and then the vector of residuals $\mathbf{v}$, and estimated directions $\hat{\mathbf{d}}$, are

$$\mathbf{v} = \begin{bmatrix} 0.3474 \\ -0.6648 \\ 0.3174 \\ -0.3460 \\ -0.6838 \\ 1.0299 \\ -1.0506 \\ 0.7114 \\ 0.3392 \\ -0.3703 \\ 0.7362 \\ -0.3660 \end{bmatrix} (\text{sec}) \qquad \mathbf{d} = \begin{bmatrix} \hat{d}_{10} \\ \hat{d}_{11} \\ \hat{d}_{12} \\ \hat{d}_{13} \\ \hat{d}_{14} \\ \hat{d}_{15} \\ \hat{d}_{16} \\ \hat{d}_{17} \\ \hat{d}_{18} \\ \hat{d}_{19} \\ \hat{d}_{20} \\ \hat{d}_{21} \end{bmatrix} = \begin{bmatrix} 0° & 0' & 0''.3474 \\ 22 & 1 & 41.8452 \\ 38 & 46 & 14.0274 \\ 359 & 59 & 59.6540 \\ 57 & 8 & 56.4162 \\ 76 & 42 & 12.2599 \\ 359 & 59 & 58.9494 \\ 86 & 33 & 14.1614 \\ 145 & 19 & 49.7192 \\ 359 & 59 & 59.6297 \\ 15 & 6 & 53.0162 \\ 99 & 11 & 42.5740 \end{bmatrix}$$

From the above directions the following eight angles of the quadrilateral are computed (see Figure 7.2):

$$\hat{l}_1 = \hat{d}_{11} - \hat{d}_{10} = 22°1'41''4978$$

$$\hat{l}_2 = \hat{d}_{12} - \hat{d}_{11} = 16°44'32''1822$$

$$\hat{l}_3 = \hat{d}_{14} - \hat{d}_{13} = 57°8'56''7622$$

$$\hat{l}_4 = \hat{d}_{15} - \hat{d}_{14} = 19°33'15''8437$$

$$\hat{l}_5 = \hat{d}_{17} - \hat{d}_{16} = 86°33'15''2120$$

$$\hat{l}_6 = \hat{d}_{18} - \hat{d}_{17} = 58°46'35''5578$$

$$\hat{l}_7 = \hat{d}_{20} - \hat{d}_{19} = 15°6'53''3865$$

$$\hat{l}_8 = \hat{d}_{21} - \hat{d}_{20} = 84°4'49''5578$$

Values for these eight angles will be computed in the second part of the problem directly from a least squares adjustment and compared to those given above. To complete the comparison, we also check on the a posteriori cofactor matrix $\mathbf{Q}_{\hat{a}\hat{a}}$.

$\mathbf{Q}_{\hat{a}\hat{a}} =$

0.7290	0.1788	0.0922	−0.1600	−0.1870	−0.0270	−0.0221	−0.0685	0.0906	−0.1641	−0.0629	0.2270
	0.6121	0.2091	−0.0354	−0.1588	0.1942	−0.0683	0.1051	−0.0368	0.2253	−0.1592	−0.0661
		0.6987	0.1953	−0.0281	−0.1672	0.0904	−0.0366	−0.1538	−0.0120	0.2221	−0.1609
			0.6917	0.2282	0.0800	−0.1618	0.2192	−0.0573	−0.0652	−0.0215	0.0867
				0.5671	0.2047	−0.0597	−0.1669	0.2266	−0.0095	0.0666	−0.0572
					0.7153	0.2216	−0.0523	−0.1693	0.0747	−0.0451	−0.0295
						0.7267	0.1845	0.0887	−0.1662	0.1904	−0.0242
							0.5978	0.2177	−0.0196	−0.1673	0.1869
								0.6936	0.1858	−0.0231	−0.1627
symmetric									0.6770	0.2408	0.8210
										0.5662	0.1930
											0.7248

The quadratic form may also be computed as $\mathbf{v}^t\mathbf{W}_{dd}\mathbf{v} = 4.8492$ (sec of arc)^2.
In order to compute the cofactor matrix of the eight angles first write

$$
\begin{bmatrix} \hat{l}_1 \\ \hat{l}_2 \\ \hat{l}_3 \\ \hat{l}_4 \\ \hat{l}_5 \\ \hat{l}_6 \\ \hat{l}_7 \\ \hat{l}_8 \end{bmatrix} =
\begin{bmatrix}
-1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1
\end{bmatrix}
\begin{bmatrix} \hat{d}_{10} \\ \hat{d}_{11} \\ \vdots \\ \hat{d}_{21} \end{bmatrix}
$$

or $\hat{l} = \mathbf{F}\hat{d}$ and then apply propagation to compute $\mathbf{Q}_{\hat{l}\hat{l}}$.

$\mathbf{Q}_{\hat{l}\hat{l}} =$

0.9835	−0.3164	−0.4704	0.5669	0.2198	−0.3009	−0.4858	−0.1968
	0.8926	−0.1000	−0.4920	−0.3005	0.1247	0.6679	−0.4762
		0.8024	−0.2143	−0.4881	0.6700	0.0324	−0.2320
			0.8731	−0.1668	−0.5104	−0.1959	0.1394
				0.9554	−0.2843	−0.5044	−0.5688
symmetric					0.8559	−0.0612	−0.4937
						0.7615	−0.2145
							0.9050

Adjustment Using Angles: It is of fundamental importance here to recognize that although the directions are the original uncorrelated observations, the angles that are computationally derived from them need not be uncorrelated. As a matter of fact,

since $l = \mathbf{Fd}$, the a priori cofactor matrix of the angles is $\mathbf{Q}_{ll} = \mathbf{FQ}_{dd}\mathbf{F}' = \mathbf{FF}'$ since $\mathbf{Q}_{dd} = \mathbf{I}$.

This is a basic point that involves choosing the correct stochastic model to fit the variations in the functional model. Thus

$$\mathbf{Q}_{ll} = \begin{bmatrix} 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & 2 & -1 & 0 & 0 & 0 & 0 \\ & & & 2 & 0 & 0 & 0 & 0 \\ & & & & 2 & -1 & 0 & 0 \\ & & & & & 2 & 0 & 0 \\ & & & & & & 2 & -1 \\ \text{symmetric} & & & & & & & 2 \end{bmatrix}$$

This matrix indicates a fact that should be obvious from Figure 7.2; that is, angles 1 and 2 are correlated, angles 3 and 4 are correlated, and so on, but each pair is not correlated with any other. This is because each pair of angles is derived from the directions measured at one point. Consequently, it would be an inaccurate stochastic model if we were to use the identity matrix instead of the given $\mathbf{Q}_{ll}$ above.

To formulate the number of condition equations we have only to realize that it takes a minimum of four angles to construct the shape of a unique quadrilateral thus leaving a redundancy of *four*. Four independent condition equations (three angular and one side) are

$$f_1 = l_1 + l_2 + l_3 + l_8 - \pi = 0$$

$$f_2 = l_4 + l_5 + l_6 + l_7 - \pi = 0$$

$$f_3 = l_1 + l_6 + l_7 + l_8 - \pi = 0$$

$$f_4 = \frac{\sin(l_1)\sin(l_3)\sin(l_5)\sin(l_7)}{\sin(l_2)\sin(l_4)\sin(l_6)\sin(l_8)} - 1 = \frac{Y}{Z} - 1 = 0$$

Linearization

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} & a_{47} & a_{48} \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_8 \end{bmatrix} = \begin{bmatrix} -f_1 \\ -f_2 \\ -f_3 \\ -f_4 \end{bmatrix} = \begin{bmatrix} 0.3414 \\ 1.0502 \\ -0.7066 \\ 0.0072 \end{bmatrix} \text{(sec of arc)}^{-1}$$

where

$$a_{41} = (Y/Z)\cot(l_1) = 2.4716$$

$$a_{42} = -(Y/Z)\cot(l_2) = -3.3244$$

$$a_{43} = (Y/Z)\cot(l_3) = 0.6457$$

$$a_{44} = -(Y/Z)\cot(l_4) = -2.8156$$

$$a_{45} = (Y/Z)\cot(l_5) = 0.0602$$

$$a_{46} = -(Y/Z)\cot(l_6) = -0.6062$$

$$a_{47} = (Y/Z)\cot(l_7) = 3.7025$$

$$a_{48} = -(Y/Z)\cot(l_8) = 0.1037$$

The reduced normal equations coefficient matrix is

$$
\mathbf{Q}_e = \begin{bmatrix} 6.0 & -2.0 & 2.0 & -0.6557 \\ & 6.0 & 2.0 & 0.6859 \\ & & 6.0 & 10.5939 \\ \text{symmetric} & & & 100.1021 \end{bmatrix} (\text{sec of arc})^2
$$

and its inverse

$$
\mathbf{W}_e = \begin{bmatrix} 0.2773 & 0.1490 & -0.1765 & 0.0195 \\ & 0.2712 & -0.1703 & 0.0171 \\ & & 0.3471 & -0.0367 \\ \text{symmetric} & & & 0.0139 \end{bmatrix} (\text{sec of arc})^{-2}
$$

The vector of Lagrange multipliers is

$$
\mathbf{k} = [0.3414 \quad 1.0502 \quad -0.7066 \quad 0.0072]'(\text{sec of arc})^{-1}
$$

The vectors of residuals $\mathbf{v}$ and estimated observations $\hat{l}$ are

$$
\mathbf{v}_{8,1} = \begin{bmatrix} -1.0122 \\ 0.9822 \\ -0.3378 \\ 1.7137 \\ 1.7620 \\ -0.3722 \\ 1.1065 \\ -1.1022 \end{bmatrix} (\text{sec}) \quad \text{and} \quad \begin{matrix} \hat{l}_1 = 22°1'41''4978 \\ \hat{l}_2 = 16°44'32''1822 \\ \hat{l}_3 = 57°8'56''7622 \\ \hat{l}_4 = 19°33'15''8437 \\ \hat{l}_5 = 86°33'15''2120 \\ \hat{l}_6 = 58°46'35''5578 \\ \hat{l}_7 = 15°6'53''3865 \\ \hat{l}_8 = 84°4'49''5578 \end{matrix}
$$

Finally, the cofactor matrix of the estimated angles is

$$
\mathbf{Q}_{ll} = \begin{bmatrix} 0.9835 & -0.3164 & -0.4704 & 0.5669 & 0.2198 & -0.3009 & -0.4858 & -0.1968 \\ & 0.8926 & -0.1000 & -0.4920 & -0.3005 & 0.1247 & 0.6679 & 0.4762 \\ & & 0.8024 & -0.2143 & -0.4881 & 0.6700 & 0.0324 & -0.2320 \\ & & & 0.8731 & -0.1668 & -0.5104 & -0.1959 & 0.1394 \\ & & & & 0.9554 & -0.2843 & -0.5044 & 0.5688 \\ & & & & & 0.8559 & -0.0612 & -0.4937 \\ & & & & & & 0.7615 & -0.2145 \\ \text{symmetric} & & & & & & & 0.9050 \end{bmatrix}
$$

and the computed quadratic form $\mathbf{v}'\mathbf{W}_{ll}\mathbf{v} = 4.8492$.

It should be clear now that the results obtained above match identically with those computed from an adjustment using observed directions.

With $r = 4$, the a posteriori estimate of the reference variance is

$$
\hat{\sigma}_0^2 = \frac{4.8492}{4} = 1.2123
$$

Using this value the "standard deviation-correlation" matrix† for $\hat{l}$ in seconds is

$$
\mathbf{M}_{\sigma\rho} =
\begin{bmatrix}
1.092 & -0.338 & -0.529 & 0.612 & 0.227 & -0.328 & -0.561 & -0.209 \\
 & 1.040 & -0.118 & -0.557 & -0.325 & 0.143 & 0.810 & -0.530 \\
 & & 0.986 & -0.256 & -0.557 & 0.808 & 0.041 & -0.272 \\
 & & & 1.029 & -0.183 & -0.590 & -0.240 & 0.157 \\
 & & & & 1.076 & -0.314 & -0.591 & 0.612 \\
 & & & & & 1.019 & -0.076 & -0.561 \\
 & & & & & & 0.961 & -0.258 \\
 \text{symmetric} & & & & & & & 1.047
\end{bmatrix}
$$

Difficulties may arise when the a priori correlation between angles is neglected, if such angles are computed from measured directions. To demonstrate this point, the same problem was solved one more time except that instead of using the correct cofactor matrix used above, $\mathbf{Q}_{ll} = 2\mathbf{I}$ was used. The reason for the multiplier of 2 should be obvious since each angle is compiled from two directions each having a cofactor of unity, thus it will have a cofactor of two. The results of this nonrigorous adjustment are

$$
\mathbf{v} =
\begin{bmatrix}
-1.02 \\
0.26 \\
0.38 \\
1.69 \\
1.78 \\
0.30 \\
0.43 \\
-1.10
\end{bmatrix}
\text{(sec)} \quad \text{and} \quad
\begin{aligned}
\hat{l}_1 &= 22°1'41''49 \\
\hat{l}_2 &= 16°44'31''46 \\
\hat{l}_3 &= 57°8'57''48 \\
\hat{l}_4 &= 19°33'15''82 \\
\hat{l}_5 &= 86°33'15''23 \\
\hat{l}_6 &= 58°46'36''23 \\
\hat{l}_7 &= 15°6'52''71 \\
\hat{l}_8 &= 84°4'49''56
\end{aligned}
$$

Comparing the set of residuals computed here and those computed using the correct adjustment we find some differences. The variation is both in the magnitude and sign of the corrections. The maximum difference in the values of the estimated angles is 0.72 seconds. However, this value is within the reference standard deviation and may be tolerated.

This example represents a case of "approximate adjustment," inasmuch as a simplified stochastic model was used. Such simplified computational methods have been, and still are, used in practice. They lose importance, however, with the development of computer techniques which increasingly make the use of simplified procedures unwarranted. The example further demonstrates that the estimated observations and the estimates of the parameters can be close to the proper least squares estimates. By contrast to estimated observations and parameters, however, the a posteriori cofactor and covariance matrices computed on the basis of a priori simplified matrices are usually not close to those obtained from proper least squares.

† The "standard deviation-correlation" matrix is used frequently in statistics and is given here for the sake of introducing the reader to another useful concept. Its elements on the main diagonal are the standard deviations, whereas those off the diagonal are the correlation coefficients.

Because this is quite frequent in practice, a general warning is appropriate against relying on such a posteriori matrices. Instead, new matrices should be computed on the basis of the nonsimplified cofactors of the original observations.

Let $\mathbf{Q}$ be the original rigorous a priori cofactor matrix of the observations and $\bar{\mathbf{Q}}$ the simplified matrix. Let $\mathbf{s}$ represent the set of variables estimated using $\mathbf{Q}$. Instead of using the a posteriori cofactor matrix for the estimates $\hat{\mathbf{s}}$ as computed with $\bar{\mathbf{Q}}$, we should apply the principle of propagation to get a $\mathbf{Q}_{\hat{s}\hat{s}}$ using the original $\mathbf{Q}$.

7.3. ADJUSTMENT OF INDIRECT OBSERVATIONS

Adjustment of indirect observations is the second of the two classical cases of least squares adjustment. Here the adjustment is performed with both observations and parameters but with the restriction that each condition equation contains only *one observation*. Therefore the number of condition equations is the same as the number of observations, or $c = n$.

7.3.1. Formulation Because of the restriction on this case, the general linear functional equations would be of the form

$$l + \mathbf{v} + \mathbf{B}\Delta = \mathbf{d} \tag{7.8}$$

or

$$\mathbf{v} + \mathbf{B}\Delta = (-l + \mathbf{d}) = \mathbf{f} \tag{7.9}$$

where $\mathbf{d}$ and $\mathbf{f}$ are constant vectors.

Since the residual vector $\mathbf{v}$ is explicitly separable in equation (7.9), the application of the least squares principle can be done more directly than in the cases discussed previously. For this reason, and as a further demonstration of the minimum criterion, we give the relatively short development here, although the results can be obtained directly from the equations obtained in Chapter 6. Thus the scalar to be minimized is

$$\phi = \mathbf{v}'\mathbf{W}\mathbf{v}$$

$$= (\mathbf{f} - \mathbf{B}\Delta)'\mathbf{W}(\mathbf{f} - \mathbf{B}\Delta)$$

$$= (\mathbf{f}' - \Delta'\mathbf{B}')(\mathbf{W}\mathbf{f} - \mathbf{W}\mathbf{B}\Delta)$$

$$= \Delta'\mathbf{B}'\mathbf{W}\mathbf{B}\Delta - \Delta'\mathbf{B}'\mathbf{W}\mathbf{f} + \mathbf{f}'\mathbf{W}\mathbf{f} - \mathbf{f}'\mathbf{W}\mathbf{B}\Delta$$

Realizing that the second and last terms on the right-hand side are scalars (as well as all other terms) then

$$\phi = \Delta'\mathbf{B}'\mathbf{W}\mathbf{B}\Delta - 2\mathbf{f}'\mathbf{W}\mathbf{B}\Delta + \mathbf{f}'\mathbf{W}\mathbf{f} \tag{7.10}$$

The free variable in equation (7.10) is the vector of parameters Δ, thus for ϕ to be a minimum $\partial\phi/\partial\Delta$ must be zero, or

$$\frac{\partial\phi}{\partial\Delta} = 2\Delta'\mathbf{B}'\mathbf{W}\mathbf{B} - 2\mathbf{f}'\mathbf{W}\mathbf{B} = \mathbf{0}' \tag{7.11}$$

recognizing that the last term in equation (7.10) is a constant insofar as the partial differentiation is concerned. Transposing equation (7.11) and rearranging

$$(\mathbf{B}'\mathbf{W}\mathbf{B})\Delta = \mathbf{B}'\mathbf{W}\mathbf{f} \tag{7.12}$$

Equation (7.12) is the u reduced normal equations in the u unknown parameters Δ. With the auxiliaries used previously [see equations (6.21) and (6.22)], then

$$\mathbf{N}\Delta = \mathbf{t} \tag{7.13}$$

or

$$\Delta = \mathbf{N}^{-1}\mathbf{t} \tag{7.14}$$

It is important to note here that all the relations developed in this section are essentially the same as those derived in Section 6.2. Having obtained Δ, the estimated observations $\hat{l}$ can be directly obtained from

$$\hat{l} = l + \mathbf{f} - \mathbf{B}\Delta \tag{7.15}$$

7.3.2. Computational Check and Reference Variance   The computational check in this case is

$$\mathbf{a} = \mathbf{B}'\mathbf{W}\mathbf{A}\mathbf{v} = \mathbf{0} \tag{7.16}$$

The proof of equation (7.16) is exactly the same as that given previously (see Section 6.2.1).

The a posteriori estimate of the reference variance $\hat{\sigma}_0^2$ is given by equation (6.25). The quadratic form $\mathbf{v}'\mathbf{W}\mathbf{v}$ may be evaluated directly using a $\mathbf{v}$ computed from equation (7.9), or from

$$\mathbf{v}'\mathbf{W}\mathbf{v} = \mathbf{f}'\mathbf{W}\mathbf{f} - \Delta'\mathbf{t} \tag{7.17}$$

7.3.3. Estimation of Precision   The autocofactor and crosscofactor matrices are

$$l = \mathbf{I}l$$

$$\mathbf{f} = -l + \mathbf{d}$$

$$\Delta = \mathbf{N}^{-1}\mathbf{B}'\mathbf{W}\mathbf{f} = -\mathbf{N}^{-1}\mathbf{B}'\mathbf{W}l + \Delta_0$$

$$\mathbf{v} = \mathbf{f} - \mathbf{B}\Delta = \mathbf{f} - \mathbf{B}\mathbf{N}^{-1}\mathbf{B}'\mathbf{W}\mathbf{f}$$

$$\hat{l} = l + \mathbf{v}$$

Autocofactor Matrices

$$\mathbf{Q}_{ll} = \mathbf{Q}$$

$$\mathbf{Q}_{ff} = \mathbf{Q}$$

$$\mathbf{Q}_{\Delta\Delta} = \mathbf{N}^{-1}\mathbf{B}'\mathbf{W}\mathbf{Q}_{ff}\,\mathbf{W}\mathbf{B}\mathbf{N}^{-1} = \mathbf{N}^{-1} \tag{7.18}$$

$$\mathbf{Q}_{vv} = \mathbf{Q} - \mathbf{B}\mathbf{N}^{-1}\mathbf{B}^{t} \tag{7.18a}$$

$$\mathbf{Q}_{\hat{l}\hat{l}} = \mathbf{Q} - \mathbf{Q}_{vv} \tag{7.19}$$

Crosscofactor Matrices

$$\mathbf{Q}_{lf} = -\mathbf{Q}$$

$$\mathbf{Q}_{l\Delta} = -\mathbf{Q}(\mathbf{W}\mathbf{B}\mathbf{N}^{-1}) = -\mathbf{B}\mathbf{N}^{-1}$$

$$\mathbf{Q}_{lv} = \mathbf{Q}_{lf} - \mathbf{Q}_{lf}(\mathbf{W}\mathbf{B}\mathbf{N}^{-1}\mathbf{B}^{t}) = -\mathbf{Q} + \mathbf{B}\mathbf{N}^{-1}\mathbf{B}^{t} = -\mathbf{Q}_{vv}$$

$$\mathbf{Q}_{l\hat{l}} = \mathbf{Q} + \mathbf{Q}_{lv} = \mathbf{Q} - \mathbf{Q}_{vv} = \mathbf{Q}_{\hat{l}\hat{l}}$$

$$\mathbf{Q}_{f\Delta} = \mathbf{Q}(\mathbf{W}\mathbf{B}\mathbf{N}^{-1}) = \mathbf{B}\mathbf{N}^{-1}$$

$$\mathbf{Q}_{fv} = \mathbf{Q}_{ff} - \mathbf{Q}_{f\Delta}\mathbf{B}^{t} = \mathbf{Q} - \mathbf{B}\mathbf{N}^{-1}\mathbf{B}^{t}$$

$$\mathbf{Q}_{f\hat{l}} = \mathbf{Q}_{fl} + \mathbf{Q}_{fv} = -\mathbf{Q} + \mathbf{Q} - \mathbf{B}\mathbf{N}^{-1}\mathbf{B}^{t} = -\mathbf{B}\mathbf{N}^{-1}\mathbf{B}^{t}$$

$$\mathbf{Q}_{\Delta v} = \mathbf{Q}_{\Delta f} - \mathbf{Q}_{\Delta\Delta}\mathbf{B}^{t} = \mathbf{N}^{-1}\mathbf{B}^{t} - \mathbf{N}^{-1}\mathbf{B}^{t} = \mathbf{0}$$

$$\mathbf{Q}_{\Delta\hat{l}} = \mathbf{Q}_{\Delta l} + \mathbf{Q}_{\Delta v} = -\mathbf{N}^{-1}\mathbf{B}^{t}$$

$$\mathbf{Q}_{v\hat{l}} = \mathbf{Q}_{vl} + \mathbf{Q}_{vv} = -\mathbf{Q}_{vv} + \mathbf{Q}_{vv} = \mathbf{0}$$

Example 7.6. Reconsider the case of the plane triangle given in Example 7.1. Although it was worked out in the most straightforward manner there, we shall solve it again here using the technique of adjustment of indirect observations.

Consider the two angles l_1 and l_2 also as unknown parameters x_1 and x_2, respectively. The three condition equations will therefore be

$$l_1 - x_1 = 0$$

$$l_2 - x_2 = 0$$

$$l_3 + x_1 + x_2 - \pi = 0$$

or

$$\mathbf{v} + \begin{bmatrix} -1 & 0 \\ 0 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -l_1 \\ -l_2 \\ \pi - l_3 \end{bmatrix} = \begin{bmatrix} -40°19'02'' \\ -70°30'01'' \\ +110°49'00'' \end{bmatrix}$$

With $\mathbf{W} = \mathbf{Q}^{-1} = \mathbf{I}$, the reduced normal equations may be computed as

$$\mathbf{N} = \mathbf{B}'\mathbf{W}\mathbf{B} = \mathbf{B}'\mathbf{B} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

$$\mathbf{t} = \mathbf{B}'\mathbf{W}\mathbf{f} = \mathbf{B}'\mathbf{f} = \begin{bmatrix} 151°08'02'' \\ 181°19'01'' \end{bmatrix}$$

$$\mathbf{N}^{-1} = \tfrac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

$$\Delta = \hat{\mathbf{x}} = \mathbf{N}^{-1}\mathbf{t} = \begin{bmatrix} 40°19'01'' \\ 70°30'00'' \end{bmatrix}$$

These two values are the same as the values of $\hat{l}_1$ and $\hat{l}_2$ obtained in Example 7.1. The cofactor matrix of the parameters is

$$\mathbf{Q}_{\Delta\Delta} = \mathbf{N}^{-1} = \tfrac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

which is also the submatrix of $\mathbf{Q}_{ll}$ corresponding to $\hat{l}_1$ and $\hat{l}_2$ in that example. To complete the problem we compute

$$\mathbf{v} = \mathbf{f} - \mathbf{B}\Delta = \begin{bmatrix} -40 & 19 & 02 \\ -70 & 30 & 01 \\ 110 & 49 & 00 \end{bmatrix} - \begin{bmatrix} -40 & 19 & 01 \\ -70 & 30 & 00 \\ 110 & 49 & 01 \end{bmatrix} = \begin{bmatrix} -1'' \\ -1'' \\ -1'' \end{bmatrix}$$

and

$$\hat{l}_1 = 40°19'01'' \qquad \hat{l}_2 = 70°30'00'' \qquad \hat{l}_3 = 69°10'59''$$

$$\mathbf{Q}_{ll} = \mathbf{B}'\mathbf{N}^{-1}\mathbf{B} = \tfrac{1}{3}\begin{bmatrix} 2 & -1 & -1 \\ & 2 & -1 \\ \text{symmetric} & & 2 \end{bmatrix}$$

which is the same as the matrix $\mathbf{Q}_{ll}$ computed in Example 7.1.

Another verification for the computations is possible when applying equation (7.16).

$$\mathbf{a} = \mathbf{B}'\mathbf{W}\mathbf{v} = \mathbf{B}'\mathbf{v} = \begin{bmatrix} -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \mathbf{0}$$

For the quadratic form we use the computed $\mathbf{v}$ directly, or

$$\mathbf{v}'\mathbf{W}\mathbf{v} = \begin{bmatrix} -1 & -1 & -1 \end{bmatrix}\begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix} = 3 \ \text{sec}^2$$

then

$$\hat{\sigma}_0^2 = 3 \ \text{sec}^2$$

Example 7.7. Consider the problem of Example 6.1 according to the technique of indirect observations. As indicated above, as many condition equations as there are observations (or five equations) are needed for this problem. From Figure 6.1 we can readily write the following five equations with the parameters being (a) x_1 is the x_1 coordinate of point P; (b) x_2 is the x_2 coordinate of point P; (c) x_3 is the distance $S_1 S_2$; and (d) x_4 is the distance $S_2 S_3$.

$$l_1 - \frac{px_1}{x_2} = 0 \qquad l_2 - \frac{p(x_3 - x_1)}{x_2} = 0 \qquad l_3 - \frac{p(x_3 + x_4 - x_1)}{x_2} = 0$$

$$l_4 - x_3 = 0 \qquad l_5 - x_4 = 0$$

l_i represent observational variables. Linearization leads to

$$\mathbf{B}_{5,4} = \begin{bmatrix} -p/x_2^0 & px_1^0/x_2^{0^2} & 0 & 0 \\ p/x_2^0 & p(x_3^0 - x_1^0)/x_2^{0^2} & -p/x_2^0 & 0 \\ p/x_2^0 & p(x_4^0 - x_1^0)/x_2^{0^2} & -p/x_2^0 & -p/x_2^0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

Approximations may be computed as follows:

$$x_1^0 = \frac{l_1 l_4}{(l_1 + l_2)} = 8.128 \text{ m}$$

$$x_2^0 = \frac{pl_4}{(l_1 + l_2)} = 49.261 \text{ m}$$

$$x_3^0 = l_4 = 10.0 \text{ m}$$

$$x_4^0 = l_5 = 8.0 \text{ m}$$

with $p = 100$ the numerical value of $\mathbf{B}$ becomes

$$\mathbf{B} = \begin{bmatrix} -2.03 & 0.335 & 0 & 0 \\ 2.03 & 0.077 & -2.03 & 0 \\ 2.03 & 0.407 & -2.03 & -2.03 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} (\text{mm m}^{-1})$$

With the covariance matrix used before, the weight matrix may be obtained from

$$\frac{1}{\sigma_0^2} \mathbf{W} = \Sigma^{-1} = \begin{bmatrix} 100 & & & & \\ & 100 & & & \\ & & 100 & & \\ & & & 400 & \\ & & & & 400 \end{bmatrix} = 100 \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & 4 & \\ & & & & 4 \end{bmatrix}$$

From which the a priori value of σ_0^2 is 0.01 and the weight matrix is

$$\mathbf{W} = \text{diag.} \{1, 1, 1, 4, 4\}$$

Then

$$\mathbf{N} = \begin{bmatrix} 12.363 & 0.303 & -8.242 & -4.121 \\ & 0.284 & -0.982 & -0.826 \\ & & 12.242 & 4.121 \\ \text{symmetric} & & & 8.121 \end{bmatrix}$$

In order to compute $\mathbf{t}$, first evaluate $\mathbf{f}$ as follows:

$$f_1 = \frac{p x_1^0}{x_2^0} - l_1 = 0$$

$$f_2 = \frac{p(x_3^0 - x_1^0)}{x_2^0} - l_2 = -0.175 \text{ mm}$$

$$f_3 = \frac{p(x_3^0 + x_4^0 - x_1^0)}{x_2^0} - l_3 = -0.360 \text{ mm}$$

$$f_4 = x_3^0 - l_4 = 0$$

$$f_5 = x_4^0 - l_5 = 0$$

Then

$$\mathbf{t} = \begin{bmatrix} -1.065 \\ -1.160 \\ +1.065 \\ +0.831 \end{bmatrix}$$

With

$$\mathbf{N}^{-1} = \begin{bmatrix} 0.189 & 0.513 & 0.143 & 0.075 \\ & 7.334 & 0.718 & 0.642 \\ & & 0.224 & 0.031 \\ \text{symmetric} & & & 0.211 \end{bmatrix}$$

then the solution is

$$\Delta = \begin{bmatrix} -0.054 & -0.455 & -0.022 & 0.028 \end{bmatrix}^t \text{ (mm)}$$

leading to the parameter values

$$\hat{x}_1 = 8.074 \text{ m} \qquad \hat{x}_2 = 48.806 \text{ m}$$

$$\hat{x}_3 = 9.978 \text{ m} \qquad \hat{x}_4 = 8.028 \text{ m}$$

which agree well with the values computed previously.

This example brings up one last remark. In order that we may use this case of adjustment of indirect observations, it is necessary that the *maximum number of independent parameters* involved in the model be carried in the adjustment. Although this may be desirable when all such parameters are of interest, it is not convenient or useful to have to do this when only a subset is needed. This example illustrates the point since we were only interested in two parameters, x_1 and x_2, but had to solve for four.

Example 7.8. In this example, solve the two-parameter transformation problem of Example 6.2 using the technique of adjustment of indirect observations. The transformation equations were given as

$$y_{1i} = ax_{1i} - bx_{2i}$$

$$y_{2i} = bx_{1i} + ax_{2i}$$

or in matrix form

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_i = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}_i$$

Since the x coordinates are the observations, first write the inverse transformation, or

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}_i = \frac{1}{(a^2 + b^2)} \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_i$$

or

$$f_{1i} = x_{1i} - \frac{a}{(a^2 + b^2)} y_{1i} - \frac{b}{(a^2 + b^2)} y_{2i} = 0 \tag{a}$$

$$f_{2i} = x_{2i} + \frac{b}{(a^2 + b^2)} y_{1i} - \frac{a}{(a^2 + b^2)} y_{2i} = 0$$

with $i = 1, 2, 3$, which are the six condition equations ($n = 6$). Linearization by Taylor series to the form $\mathbf{v} + \mathbf{B\Delta} = \mathbf{f}$ yields

$$\underset{2,1}{\mathbf{v}_i} + \frac{1}{(a^2 + b^2)^2} \begin{bmatrix} [(a^2 - b^2)y_{1i} + 2aby_{2i}] & [(-a^2 + b^2)y_{2i} + 2aby_{1i}] \\ [(a^2 - b^2)y_{2i} - 2aby_{1i}] & [(a^2 - b^2)y_{1i} + 2aby_{2i}] \end{bmatrix} \mathbf{\Delta}$$

$$= \begin{bmatrix} -f_{1i} \\ -f_{2i} \end{bmatrix}$$

with the given data

$$\mathbf{x}_1 = (0.0, 1.0), \ \mathbf{x}_2 = (1.0, 0.0), \ \mathbf{x}_3 = (1.0, 1.0)$$

$$\mathbf{y}_1 = (-2.1, 1.1), \ \mathbf{y}_2 = (1.0, 2.0), \ \mathbf{y}_3 = (-0.9, 2.8)$$

and the approximations $a^0 = 1$ and $b^0 = 2$, the numerical values of the matrices become

$$\mathbf{B} = \begin{bmatrix} 0.428 & -0.204 \\ 0.204 & 0.428 \\ 0.200 & 0.400 \\ -0.400 & 0.200 \\ 0.556 & 0.192 \\ -0.192 & 0.556 \end{bmatrix} \quad \text{and} \quad \mathbf{f} = \begin{bmatrix} 0.02 \\ 0.06 \\ 0.00 \\ 0.00 \\ -0.06 \\ -0.08 \end{bmatrix}$$

Remembering that the x coordinates are uncorrelated and of equal precision, for the sake of variation take an a priori $\sigma_0^2 = 0.01$ and thus a $\mathbf{W} = \mathbf{I}$. Then

$$\mathbf{N} = \begin{bmatrix} 0.7708 & 0.0 \\ 0.0 & 0.7708 \end{bmatrix} \qquad \mathbf{N}^{-1} = \begin{bmatrix} 1.2974 & 0.0 \\ 0.0 & 1.2974 \end{bmatrix}$$

$$\Delta = \begin{bmatrix} 0.0036 \\ -0.0446 \end{bmatrix}$$

and

$$\hat{a} = 1.0036 \qquad \hat{b} = 1.9554$$

These answers agree with those obtained twice before. They, however, represent values obtained after the first iteration of a nonlinear problem. If we were to continue the iterations until the last Δ is insignificantly small, we would get as a final answer $\hat{a} = 1.0031$ and $\hat{b} = 1.9561$.

We have solved the problem above in a straightforward manner without any attempt at simplification. It is possible, however, to eliminate the nonlinearity in the problem as well as simplify the algebraic operations by using two new parameters c and d, mainly because the y coordinates were given in the problem as constants. Referring to equation (a), make the substitution

$$c = \frac{a}{a^2 + b^2} \qquad d = \frac{b}{a^2 + b^2} \tag{b}$$

It is important to emphasize that reparameterization is possible provided we always make certain that the *number of the new parameters is the same as the old parameters, and that the new parameters are also functionally independent just as the old ones were.* In our present case both restrictions are met by the choice of c and d. Thus

$$f_{1i} = x_{1i} - cy_{1i} - dy_{2i} = 0 \tag{c}$$

$$f_{2i} = x_{2i} + dy_{1i} - cy_{2i} = 0$$

with $i = 1, 2, 3$; and then in the usual form $\mathbf{v} + \mathbf{B}\Delta = f$, or

$$\mathbf{v}_i + \begin{bmatrix} -y_{1i} & -y_{2i} \\ -y_{2i} & y_{1i} \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} -x_{1i} \\ -x_{2i} \end{bmatrix} \tag{d}$$

again with $i = 1, 2, 3$. With $\mathbf{Q} = \mathbf{I}_6$, then

$$\mathbf{N} = \begin{bmatrix} 19.27 & 0 \\ 0 & 19.27 \end{bmatrix} \qquad \mathbf{t} = \begin{bmatrix} 4.0 \\ 7.8 \end{bmatrix}$$

and

$$\Delta = \begin{bmatrix} \hat{c} \\ \hat{d} \end{bmatrix} = \begin{bmatrix} 0.207576 \\ 0.404774 \end{bmatrix}$$

To get the estimates of the original parameters a and b we solve for them from (b), or

$$\hat{a} = \frac{\hat{c}}{\hat{c}^2 + \hat{d}^2} = 1.0031$$

$$b = \frac{\hat{d}}{\hat{c}^2 + \hat{d}^2} = 1.9561$$

These values are identical to those obtained from the first solution, *but after* the problem has been iterated sufficiently. This should be clear to the reader since in the linear case we would obtain the exact solution directly, which in the nonlinear case is attainable only after the effects of linearization have been reduced to a negligible level.

Example 7.9. *Simple Mean, Weighted Mean, and Correlated Mean*

a. *Simple Mean.* Given n repeated measurements on a variate l, derive the relation for the best estimate if all the measurements are uncorrelated and of equal precision (weight).

Solution: Denoting the best estimate by x (a parameter), the n condition equations would be

$$l_1 + v_1 - x = 0$$

$$l_2 + v_2 - x = 0$$

$$\dots\dots\dots\dots$$

$$l_n + v_n - x = 0$$

or

$$\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} + \begin{bmatrix} -1 \\ -1 \\ \vdots \\ -1 \end{bmatrix} \underset{1,\,1}{\mathbf{x}} = \begin{bmatrix} -l_1 \\ -l_2 \\ \vdots \\ -l_n \end{bmatrix}$$

With $\mathbf{W} = \mathbf{Q}^{-1} = \mathbf{I}$, then

$$\mathbf{N} = n \qquad \mathbf{t} = \sum_i l_i$$

$$\mathbf{x} = \frac{1}{n} \sum_i l_i = \bar{l} = \text{the arithmetic mean} \tag{7.20}$$

$$\mathbf{Q}_{xx} = \frac{1}{n} = \text{cofactor matrix of arithmetic mean} \tag{7.21}$$

b. *Weighted Mean.* Given n uncorrelated observations on the same stochastic variable l, each of which having a weight w_i, derive the formula for the best estimate of that variable.

Solution: The condition equations are the same as those in (a) above, but the weight matrix $\mathbf{W}$ in this case is

$$\mathbf{W} = \text{diag. } [w_1, w_2, \ldots, w_n]$$

and

$$\mathbf{N} = \sum_{i=1}^{n} w_i$$

$$\mathbf{t} = \sum_{i=1}^{n} w_i l_i$$

$$\mathbf{x} = \frac{\sum_{i=1}^{n} w_i l_i}{\sum_{i=1}^{n} w_i} = \text{the weighted arithmetic mean} \qquad (7.22)$$

$$\mathbf{Q}_{xx} = \left[\sum_{i=1}^{n} w_i \right]^{-1} \qquad (7.23)$$

$$\mathbf{W}_{xx} = \sum_{i=1}^{n} w_i$$

which says that the weight of the weighted mean is equal to the sum of all weights of constituent observations.

c. *Correlated Mean.* Given n observations on the same random variable l that are correlated and of unequal precision, if the weight matrix of these observations is $\mathbf{W}$, compute the best estimate of l.

Solution: The weight matrix may be written in expanded form as

$$\mathbf{W} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{bmatrix}$$

The condition equations are as in (a), but the normal equations become

$$\mathbf{N} = \mathbf{B}'\mathbf{W}\mathbf{B} = \begin{bmatrix} -1 & -1 & \cdots & -1 \end{bmatrix} \mathbf{W} \begin{bmatrix} -1 \\ -1 \\ \vdots \\ -1 \end{bmatrix}$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}$$

and, similarly

$$t = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} l_i$$

The best estimate is therefore

$$\mathbf{x} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} l_i}{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}} \qquad (7.24)$$

and

$$\mathbf{Q}_{xx} = \left[\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \right]^{-1}$$

$$\mathbf{W}_{xx} = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \tag{7.25}$$

which equals the sum of all the elements of the weight matrix of the given observational vector l.

Example 7.10. *The Vector Mean.* The plane coordinates of a point have been independently determined twice by two methods with the following results. By the first method

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \qquad \text{with a covariance matrix } \underset{2,\,2}{\Sigma_{yy}}$$

By the second method

$$\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \qquad \text{with a covariance matrix } \underset{2,\,2}{\Sigma_{zz}}$$

If there is no correlation between $\mathbf{y}$ and $\mathbf{z}$, find by the method of least squares the best estimate of the coordinates.

Solution: Let the final estimate coordinate vector be denoted by the parameter vector

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Realizing that the $\mathbf{y}$ and $\mathbf{z}$ are observational vectors, we can write the following pair of vector condition equations

$$\mathbf{y} + \mathbf{v}_y - \mathbf{x} = \mathbf{0}$$

$$\mathbf{z} + \mathbf{v}_z - \mathbf{x} = \mathbf{0}$$

or

$$\begin{bmatrix} \mathbf{v}_y \\ \mathbf{v}_z \end{bmatrix} + \begin{bmatrix} -\mathbf{I} \\ -\mathbf{I} \end{bmatrix} \mathbf{x} = \begin{bmatrix} -\mathbf{y} \\ -\mathbf{z} \end{bmatrix}$$

which is of the form $\mathbf{v} + \mathbf{B}\Delta = \mathbf{f}$.

Since there is no correlation between the $\mathbf{y}$ and $\mathbf{z}$ observational vectors, the total covariance matrix of the observations can be constructed as

$$\underset{4,\,4}{\Sigma} = \begin{bmatrix} \Sigma_{yy} & \mathbf{0} \\ \mathbf{0} & \Sigma_{zz} \end{bmatrix}$$

and hence

$$\mathbf{W} = \Sigma^{-1} = \begin{bmatrix} \Sigma_{yy}^{-1} & \mathbf{0} \\ \mathbf{0} & \Sigma_{zz}^{-1} \end{bmatrix}$$

The normal equations are

$$\mathbf{N} = [-\mathbf{I} \quad -\mathbf{I}] \begin{bmatrix} \Sigma_{yy}^{-1} & \mathbf{0} \\ \mathbf{0} & \Sigma_{zz}^{-1} \end{bmatrix} \begin{bmatrix} -\mathbf{I} \\ -\mathbf{I} \end{bmatrix}$$

$$= (\Sigma_{yy}^{-1} + \Sigma_{zz}^{-1})$$

$$\mathbf{t} = [-\mathbf{I} \quad -\mathbf{I}] \begin{bmatrix} \Sigma_{yy}^{-1} & \mathbf{0} \\ \mathbf{0} & \Sigma_{zz}^{-1} \end{bmatrix} \begin{bmatrix} -\mathbf{y} \\ -\mathbf{z} \end{bmatrix} = \Sigma_{yy}^{-1}\mathbf{y} + \Sigma_{zz}^{-1}\mathbf{z}$$

Hence

$$\mathbf{x} = (\Sigma_{yy}^{-1} + \Sigma_{zz}^{-1})^{-1}(\Sigma_{yy}^{-1}\mathbf{y} + \Sigma_{zz}^{-1}\mathbf{z}) \tag{7.26}$$

and

$$\Sigma_{xx} = (\Sigma_{yy}^{-1} + \Sigma_{zz}^{-1})^{-1} \tag{7.27}$$

This relationship gives the vector mean of two independent vectors in two-dimensional space, each given with its covariance matrix as well. Obviously the same relationship holds for two vectors of more than two components (that is, in multidimensional space). Also, either cofactor or weight matrices can be used in place of the covariance matrices.

The case of two determinations only was used as an example. Actually, both equations (7.26) and (7.27) hold for multiple determinations as long as there are no correlations among them. For example, if m determinations $\mathbf{x}_i$, $i = 1, 2, \ldots, m$ are given, each having a covariance matrix Σ_i, the following relations hold provided $\Sigma_{jk} = \mathbf{0}$ for all $j \neq k$:

$$\mathbf{x} = (\Sigma_1^{-1} + \Sigma_2^{-1} + \cdots + \Sigma_m^{-1})^{-1}(\Sigma_1^{-1}\mathbf{x}_1 + \Sigma_2^{-1}\mathbf{x}_2 + \cdots + \Sigma_m^{-1}\mathbf{x}_m) \tag{7.28}$$

and

$$\Sigma_{xx} = (\Sigma_1^{-1} + \Sigma_2^{-1} + \cdots + \Sigma_m^{-1})^{-1} \tag{7.29}$$

If $\Sigma_{jk} \neq \mathbf{0}$ for some or all $j \neq k$ neither equation (7.28) nor (7.29) will apply, and a more general treatment must be done.

Example 7.11. This example is to rework the problem of Example 7.4, of the five-point level net, using the method of adjustment of indirect observations. Recall that there are $n = 8$ observations, and that the redundancy was $r = 4$. If we introduce the four elevations of points B, C, D, and E as parameters, then $u = 4$. Thus the number of independent condition equations will in this case be $c = r + u = 8$, which is also equal to n. In this case it is likely that we can readily write the condition equations such that each equation includes only one observation. Referring to Figure 7.1, the following *eight* equations can be directly written. (For simplicity in the derivation and computation the elevation of point A will be considered temporarily equal to zero. At the end of the adjustment, the elevation of A, which is

800.00 m will be added to all computed elevations.) In these equations the elevations of points B, C, D, and E are designated x_1, x_2, x_3, and x_4, respectively.

$$l_1 - x_1 = 0$$

$$l_2 + x_1 - x_2 = 0$$

$$l_3 + x_2 = 0$$

$$l_4 + x_1 - x_3 = 0$$

$$l_5 + x_3 - x_4 = 0$$

$$l_6 - x_2 + x_4 = 0$$

$$l_7 + x_4 = 0$$

$$l_8 + x_2 - x_3 = 0$$

These equations are linear and can therefore be written directly in the form

$$\mathbf{v} + \mathbf{B}\Delta = \mathbf{f}$$

$$
\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \\ v_8 \end{bmatrix}
+
\begin{bmatrix}
-1 & 0 & 0 & 0 \\
1 & -1 & 0 & 0 \\
0 & 1 & 0 & 0 \\
1 & 0 & -1 & 0 \\
0 & 0 & 1 & -1 \\
0 & -1 & 0 & 1 \\
0 & 0 & 0 & 1 \\
0 & 1 & -1 & 0
\end{bmatrix}
\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}
=
\begin{bmatrix} -l_1 \\ -l_2 \\ -l_3 \\ -l_4 \\ -l_5 \\ -l_6 \\ -l_7 \\ -l_8 \end{bmatrix}
=
\begin{bmatrix} -25.42 \text{ m} \\ -10.34 \text{ m} \\ 35.20 \text{ m} \\ 15.54 \text{ m} \\ -21.32 \text{ m} \\ -4.82 \text{ m} \\ 31.02 \text{ m} \\ 26.11 \text{ m} \end{bmatrix}
$$

The weight matrix of the observations is

$$\mathbf{W} = \text{diag.} \ [1/1.81, \ 1/0.94, \ 1/1.42, \ 1/1.76, \ 1/1.35, \ 1/0.99, \ 1/1.38, \ 1/1.40]$$

The normal equations coefficient matrix $\mathbf{N}$ is

$$
\mathbf{N} =
\begin{bmatrix}
2.1845 & -1.0638 & -0.5682 & 0 \\
 & 3.4924 & -0.7143 & -1.0101 \\
 & & 2.0232 & -0.7407 \\
\text{symmetric} & & & 2.4755
\end{bmatrix}
$$

and its inverse

$$
\mathbf{N}^{-1} =
\begin{bmatrix}
0.806 & 0.445 & 0.505 & 0.333 \\
 & 0.645 & 0.504 & 0.414 \\
 & & 0.999 & 0.505 \\
\text{symmetric} & & & 0.724
\end{bmatrix}
$$

In a similar fashion the constant term vector may be computed and used to compute the vector Δ,

$$\Delta = [25.220 \quad 35.535 \quad 9.534 \quad 30.846]^t (\text{m})$$

Thus the elevations of points B, C, D, and E may be obtained by adding 800.00 m to each element of the vector Δ,

elevation of point $B = 825.220$ m

elevation of point $C = 835.535$ m

elevation of point $D = 809.534$ m

elevation of point $E = 830.846$ m

which agree exactly with the values computed in Example 7.4 using the method of adjustment of observations only.

The cofactor matrix of these four elevations is equal to $\mathbf{N}^{-1}$, which is the same as $\mathbf{Q}_{xx}$ computed in Example 7.4. If *only* the elevations and their *cofactor* matrix are required, this would conclude the present problem and we would agree that this method of adjustment is shorter than that used in the other example, at least for the current problem. If, however, it is the *covariance* matrix of the elevations that is required, further computation is necessary to determine the a posteriori estimate of the reference variance $\hat{\sigma}_0^2$. First, compute the vector of residuals

$$\mathbf{v} = [-0.199 \quad -0.025 \quad -0.335 \quad -0.147 \quad -0.008 \quad -0.131 \quad 0.174 \quad 0.109]^t \text{ (m)}$$

and

$$\hat{\sigma}_0^2 = 0.0404$$

Multiplying $\mathbf{Q}_{xx}$ by $\hat{\sigma}_0^2$ gives the covariance matrix of the elevations

$$\Sigma_{xx} = \begin{bmatrix} 0.0326 & 0.0180 & 0.0204 & 0.0135 \\ & 0.0261 & 0.0204 & 0.0167 \\ & & 0.0404 & 0.0204 \\ \text{symmetric} & & & 0.0293 \end{bmatrix} \text{ (m}^2\text{)}$$

which as expected is identical to that computed by the other method. The reader is left the task of determining which technique of adjustment is preferable. He should compare such factors as setting up the condition equations, the size of the normal equations to be inverted, the amount of effort required for precision estimation, and above all the way the computations will be performed (by limited means or by electronic computer). Before leaving this example we perform one last computation to demonstrate the computational check of equation (7.16).

$$\mathbf{B}^t\mathbf{W}\mathbf{v} = \begin{bmatrix} -0.555 & 1.064 & 0 & 0.568 & 0 & 0 & 0 & 0 \\ 0 & -1.064 & 0.704 & 0 & 0 & -1.010 & 0 & 0.714 \\ 0 & 0 & 0 & -0.568 & 0.741 & 0 & 0 & -0.714 \\ 0 & 0 & 0 & 0 & -0.741 & 1.010 & 0.725 & 0 \end{bmatrix} \begin{bmatrix} -0.199 \\ -0.025 \\ -0.335 \\ -0.147 \\ -0.008 \\ -0.131 \\ 0.174 \\ 0.109 \end{bmatrix}$$

$$= \begin{bmatrix} 0.0 \\ 0.0 \\ 0.0 \\ 0.0 \end{bmatrix} = \mathbf{0}$$

7.4. SUMMARY OF EQUATIONS AND CONCLUDING REMARKS

7.4.1. Adjustment of Observations Only

The given information would be a vector of n observations l and its cofactor (or covariance) matrix Q. After determining the redundancy r, a set of r linear condition equations are written

$$Av = d - Al = f$$

v is the $n \times 1$ vector of residuals, and d and f are $r \times 1$ vectors of constants

A is an $r \times n$ coefficient matrix of rank $= r$

$Q_e = AQA^t = r \times r$ reduced normal equations coefficient matrix

$k = Q_e^{-1}f = W_e\,f =$ vector of r Lagrange multipliers   (7.4)

$v = QA^t k = QA^t W_e\,f$ (7.5)

$\hat{l} = l + v =$ estimated observations

$\hat{\sigma}_0^2 = v^t W v / r = f^t k / r$ (7.6)

$Q_{vv} = QA^t W_e\,AQ$ rank $(Q_{vv}) = r$ (7.7), (6.42)

$Q_{\hat{l}\hat{l}} = Q - Q_{vv}$ rank $(Q_{\hat{l}\hat{l}}) = n - r = n_0$ (6.31), (6.43)

This case of adjustment of observations only is suitable for use when the problem is relatively simple and uninvolved. It is also suitable when either the estimated observations themselves, or rather simple functions of them, are the variables of interest from the adjustment. For good size problems, and for situations in which the values of clearly specified parameters are required, this technique of adjustment is usually not utilized. Either the case covered in Chapter 6 or the simpler technique of indirect observations is often more appropriate.

7.4.2. Adjustment of Indirect Observations

As in the preceding section we begin with the observational vector l and its cofactor matrix Q. Since the number of condition equations is specified as equal to the observations n, an $n_0 = n - r$ parameters in Δ, where r is the redundancy, must be carried in the adjustment. The condition equations are of the form

$$v + B\Delta = -l + d = f$$ (7.8), (7.9)

d and f are $n \times 1$ vectors of constants

B is an $n \times u$ $(= n \times n_0)$ matrix of rank $= u$

$N = B^t Q^{-1} B = B^t W B =$ reduced normal coefficient matrix

$t = B^t W f =$ reduced normal equations constant vector

$\Delta = N^{-1} t$ (7.14)

(If conditions were originally nonlinear, then $\hat{x} = x^0 + \Delta$, where x^0 is a vector of approximations.)

$$v = f - B\Delta \tag{7.9}$$

$\hat{l} = l + v = $ estimated observations

$$\hat{\sigma}_0^2 = v'Wv/r = (f'Wf - \Delta't)/r \tag{7.17}$$

$$Q_{\Delta\Delta} = N^{-1}$$

$$Q_{vv} = Q - BN^{-1}B'$$

$$Q_{\hat{l}\hat{l}} = BN^{-1}B' = Q - Q_{vv} \tag{7.19}$$

The technique of adjustment of indirect observations is more closely related to that presented in Chapter 6. The only exception is that the coefficient matrix A of the residual vector is in the present case the identity matrix.

8

Examples and General Discussion on Adjustment with Conditions Only

8.1. GENERAL

The purpose of this chapter is to present the reader with a number of worked examples. For each example, we will first present a problem, discuss it in detail, determine the model, analyze the merits of different adjustment techniques, and then choose and apply the one that is considered most appropriate. It may be necessary at times to solve the same problem by more than one technique to stress a particular point.

Example 8.1. Consider mass production of metal cubes of which a particular unit is selected as a sample for check. Let us assume that all face angles in the cube *are* in fact 90-degree angles, and that the volume of the cube is to be checked. Under the assumption above, all that is necessary for computing the volume of a cube is to measure one of its sides. Consider now two possibilities: (a) that the same side of the cube is measured three times; or (b) that three different sides of the cube are measured, each once, as shown in Figure 8.1. In both cases there are two redundant observations and adjustment is necessary. The question is whether there is any difference between the two cases insofar as the adjustment problem is concerned.

For the first case, denote by l_1, l_2, and l_3 the three repetitive measurements of the
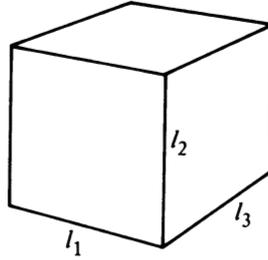
Figure 8-1

one side of the cube. If all three measurements are uncorrelated and of equal precision, the least squares estimate is simply the arithmetic mean of the three observations (see Example 7.9).

$$\hat{l} = \tfrac{1}{3}(l_1 + l_2 + l_3)$$

The volume of the cube is then

$$y = \hat{l}^3$$

For the sake of comparison later on, take the numerical values

$$l_1 = 50.0 \text{ mm} \qquad l_2 = 50.2 \text{ mm} \qquad l_3 = 49.8 \text{ mm}$$

thus

$$\hat{l} = 50.0 \text{ mm} \qquad \text{and} \qquad y = 125 \times 10^3 \text{ mm}^3$$

In the second case, the three observations are of three separate sides of the cube as shown in Figure 8.1. Since it takes only one side to fix a cube, there are 2 degrees of freedom, that is $r = 2$. There are, of course, three possibilities of adjustment techniques based on the three cases presented in the preceding two chapters.

Consider first the case of adjustment with observations only. Thus two condition equations are

$$l_1 - l_2 = 0 \qquad l_2 - l_3 = 0$$

or

$$\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} l_2 - l_1 \\ l_3 - l_2 \end{bmatrix} = \begin{bmatrix} 0.2 \text{ mm} \\ -0.4 \text{ mm} \end{bmatrix}$$

With $\mathbf{Q} = \mathbf{I}$ as assumed above, the normal equations are

$$\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \mathbf{k} = \begin{bmatrix} 0.2 \text{ mm} \\ -0.4 \text{ mm} \end{bmatrix} \qquad \text{or} \qquad \mathbf{k} = \begin{bmatrix} 0 \\ -0.2 \text{ mm} \end{bmatrix}$$

and

$$\mathbf{v} = \begin{bmatrix} 0 & -0.2 \text{ mm} & 0.2 \text{ mm} \end{bmatrix}^t$$

Thus the least squares estimate of each of the three sides is in fact 50.0 mm with the volume being $y = 125 \times 10^3$.

So it can be seen that the two solutions are identical both as regards to the estimate of the observations and the estimate of the volume. The reason for this

possibility is the fact that only one observation is required and also that the type of condition between the observations in both cases is the same—being the equality of all observations—and that the observations were uncorrelated and of equal precision.

For completeness, we work out the above adjustment using the case of explicit observations and parameters. Let x be the parameter representing the best estimate of the cube's side. Then

$$l_1 - x = 0$$

$$l_2 - x = 0$$

$$l_3 - x = 0$$

are three condition equations $(c = r + u = 2 + 1 = 3)$, or

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} + \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix} x = \begin{bmatrix} -l_1 \\ -l_2 \\ -l_3 \end{bmatrix}$$

With $\mathbf{Q} = \mathbf{I}$, the normal equations would simply be

$$3x = (l_1 + l_2 + l_3)$$

Realizing that x is the same as $\hat{l}$, we get

$$\hat{l} = \tfrac{1}{3}(l_1 + l_2 + l_3)$$

which agrees with earlier results.

Example 8.2. In the cases of the preceding example the mathematical equations have all been originally linear and the adjustment proceeded directly. Furthermore the computation of the volume of the cube was a postadjustment operation. In this example we shall deviate from both points and introduce the volume as the parameter required and estimate it directly from the adjustment. If y denotes the volume, the condition equations will be

$$l_1 - y^{1/3} = 0$$

$$l_2 - y^{1/3} = 0$$

$$l_3 - y^{1/3} = 0$$

Linearization yields

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} + \begin{bmatrix} -\tfrac{1}{3}y^{0-2/3} \\ -\tfrac{1}{3}y^{0-2/3} \\ -\tfrac{1}{3}y^{0-2/3} \end{bmatrix} \Delta = \begin{bmatrix} y^{01/3} - l_1 \\ y^{01/3} - l_2 \\ y^{01/3} - l_3 \end{bmatrix}$$

The normal equations matrix (in this case a scalar) is

$$\mathbf{N} = \tfrac{1}{3}y^{0-4/3}$$

If the approximation is taken as $y^0 = (49.5)^3$, then

$$\mathop{\mathbf{f}}_{3,1} = \begin{bmatrix} -0.5 \text{ mm} \\ -0.7 \text{ mm} \\ -0.3 \text{ mm} \end{bmatrix} \qquad \text{and} \qquad \mathbf{t} = 0.5y^{0-2/3}$$

Thus

$$\Delta = 3y^{0+4/3}(0.5y^{0-2/3}) = 3675.37 \text{ mm}^3$$

and

$$\hat{y} = 121287.375 + 3675.37 = 124962.745 \text{ mm}^3$$

If this new value of the volume is used as a fresh approximation, another correction Δ may be computed as approximately 37.25 mm^3. This would give the expected answer of 125000.00 mm^3.

The choice of $y^0 = (49.5)^3$ was intentionally done so that both linearization and iteration may be demonstrated. An obvious choice should have been $(50)^3$, but that would have led directly to a $\Delta = 0$, and the adjustment may have appeared trivial.

Example 8.3. For the adjustment case of observations only the linearized condition equations were of the form

$$\underset{r,n\ n,1}{\mathbf{A}\ \mathbf{v}} = \underset{r,1}{\mathbf{f_c}} \tag{a}$$

For the case of adjustment of indirect observations, the condition equation were of the form

$$\underset{n,1}{\mathbf{v}} + \underset{n,u\ u,1}{\mathbf{B}\ \Delta} = \underset{n,1}{\mathbf{f_0}} \tag{b}$$

Show that for the *same adjustment problem* the following is true: $\mathbf{AB} = \mathbf{0}$.

Solution: Partition the second set of equations (b) into two systems, the first containing u condition equations,

$$\underset{u,1}{\mathbf{v_1}} + \underset{u,u\ u,1}{\mathbf{B_1}\ \Delta} = \underset{u,1}{\mathbf{f_{01}}}$$

$$\underset{(n-u),1}{\mathbf{v_2}} + \underset{(n-u),u\ u,1}{\mathbf{B_2}\ \Delta} = \underset{u,1}{\mathbf{f_{02}}}$$

Under the assumption that the condition equations are independent, solve the upper set for Δ (here $\mathbf{B_1}$ will be nonsingular),

$$\Delta = \mathbf{B_1}^{-1}(-\mathbf{v_1} + \mathbf{f_{01}})$$

and substitute into the lower set

$$\mathbf{v_2} + \mathbf{B_2}\ \mathbf{B_1}^{-1}(-\mathbf{v_1} + \mathbf{f_{01}}) = \mathbf{f_{02}}$$

$$[-\mathbf{B_2}\ \mathbf{B_1}^{-1}\quad \mathbf{I}]\begin{bmatrix}\mathbf{v_1}\\\mathbf{v_2}\end{bmatrix} = [-\mathbf{B_2}\ \mathbf{B_1}^{-1}\mathbf{f_{01}} + \mathbf{f_{02}}]$$

This equation is now of the same form as the first set of condition equations (a), for the adjustment of observations only,

$$\underset{(n-u),n\ n,1}{\mathbf{A}\ \mathbf{v}} = \underset{(n-u),1}{\mathbf{f_c}}$$

Thus

$$AB = [-B_2 \ B_1^{-1} \quad I] \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$

$$= -B_2 + B_2 = 0$$

An Alternative Solution: Using information from precision estimation recall equation (7.7) from adjustment of observations only

$$\underset{n,n}{Q_{vv}} = \underset{n,n}{Q} \underset{n,r}{A^t} \underset{r,r}{W_e} \underset{r,n}{A} \underset{n,n}{Q} \tag{7.7}$$

and equation (7.18a) for adjustment of indirect observations

$$\underset{n,n}{Q_{vv}} = \underset{n,n}{Q} - \underset{n,u}{B} \underset{u,u}{N^{-1}} \underset{u,n}{B^t} \tag{7.18a}$$

Since it is an adjustment of the same problem but with two different techniques, first, W_e is different from N^{-1} and, second, Q_{vv} must be the same from both solutions. Thus

$$QA^t W_e \ AQ = Q - BN^{-1}B^t$$

Premultiplying both sides by A, and realizing that $W_e = (AQA^t)^{-1}$,

$$(AQA^t)W_e \ AQ = AQ - ABN^{-1}B^t$$

$$AQ = AQ - ABN^{-1}B^t$$

$$ABN^{-1}B^t = 0$$

Postmultiplying by (WB) and recalling that $N = B^t WB$,

$$ABN^{-1}(B^t WB) = 0$$

or

$$AB = 0$$

As a numerical demonstration refer to Examples 7.4 and 7.11 from the adjustment of a level net given in the preceding chapter:

$$\underset{4,8}{A} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}$$

$$\underset{8,4}{B} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 \end{bmatrix} \qquad \underset{4,4}{AB} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$
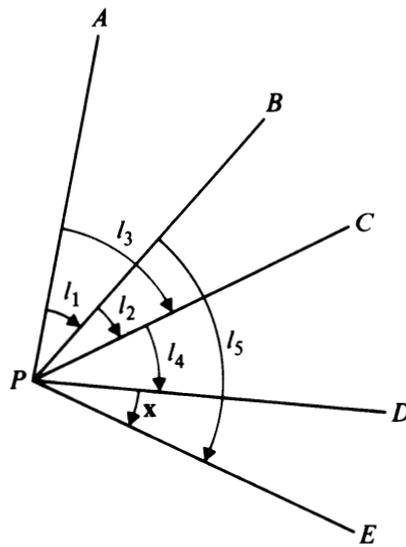
Figure 8-2

Example 8.4. Figure 8.2 shows five plane angles measured at point P with the following values:

l_1 = angle APB = 30°15′15″
l_2 = angle BPC = 20°00′00″
l_3 = angle APC = 50°15′18″
l_4 = angle CPD = 30°00′00″
l_5 = angle BPE = 70°00′01″

All five angles are uncorrelated and of equal precision. Compute the best estimate of the angle DPE, its cofactor, and the estimate of the reference variance from the adjustment.

Solution: The mathematical model of this problem is comprised of five lines in a plane all emanating from one point P. Since there are four angles constructed by those five lines, it would take a minimum of four observed angles to have a unique determination of the problem. Given five observations there is one redundant observation leading to 1 degree of freedom, $r = 1$. If the value of the required angle DPE is considered an unknown parameter x, then with $n = 5$, $r = 1$, and $u = 1$, the number of independent condition equations will be $c = r + u = 2$. Two equations can be readily written from Figure 8.2 as

$$l_1 + l_2 - l_3 = 0$$

$$l_2 + l_4 - l_5 + x = 0$$

or

$$\begin{bmatrix} 1 & 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 1 & -1 \end{bmatrix}_{5,1} \mathbf{v} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}_{1,1} \mathbf{x} = \begin{bmatrix} -l_1 - l_2 + l_3 \\ -l_2 - l_4 + l_5 \end{bmatrix}$$

$$= \begin{bmatrix} 3″ \\ 20°00′01″ \end{bmatrix}$$

Given $\mathbf{Q} = \mathbf{I}$, then

$$\mathbf{Q}_e = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \qquad \mathbf{W}_e = \frac{1}{8} \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}$$

and

$$\mathbf{N} = \frac{3}{8} \qquad \mathbf{N}^{-1} = \frac{8}{3} \qquad t = \frac{60°}{8}$$

Therefore

$$x = \mathbf{N}^{-1}t = 20°00'00'' \qquad \text{and} \qquad \mathbf{Q}_{xx} = \mathbf{N}^{-1} = \frac{8}{3}$$

Thus the estimate of the angle DPE is 20 degrees and its cofactor is $\frac{8}{3}$. To compute the a posteriori estimate of the reference variance, first

$$\mathbf{v} = \mathbf{Q}\mathbf{A}'\mathbf{W}_e(-\mathbf{B}\Delta + \mathbf{f}) = [1'' \quad 1'' \quad -1'' \quad 0 \quad 0]^t$$

and with

$$r = 1 \qquad \hat{\sigma}_0^2 = 3 \ \text{sec}^2 \qquad \text{and} \qquad \sigma_x^2 = \hat{\sigma}_0^2\mathbf{Q}_{xx} = 8 \ \text{sec}^2$$

Example 8.5. Rework the problem of Example 8.4 using a different technique of least squares.

Solution: There are two other techniques that can be used, one using $\mathbf{A}\mathbf{v} = \mathbf{f}$ and the other using $\mathbf{v} + \mathbf{B}\Delta = \mathbf{f}$. In the first possibility, we would adjust the given observations to obtain $\hat{l}$ and $\mathbf{Q}_{ll}$ and then seek a function relating x and $\hat{l}$ to compute both x and $\mathbf{Q}_{xx}$. In the second possibility, we would attempt to reduce the situation covered in Example 8.4 such that the $\mathbf{A}$ matrix becomes the identity matrix. Let us consider the first case. Here, with $r = 1$ we need to write only one condition equation thus leading to only one normal equation in one unknown. This means that the entire problem would probably be done by hand without need of any computational aid. In the second case, however, we must write *five* condition equations. Furthermore, since $r = 1 = n - u$ and with $n = 5$, it will be necessary to have *four* parameters in the condition equations in order to construct them properly. This means that with only one parameter being of interest we have to carry three more just for the sake of computation but without any real need of them. More serious, however, is the fact that there will be four normal equations, the solution of which will require inverting a square matrix of order 4. This obviously requires much more computation. It is clear then that we should choose the adjustment technique involving only observations as the more economical. From the figure one condition equation is

$$l_1 + l_2 - l_3 = 0$$

This condition equation includes a subset of three observations only from the total of five observations given. The problem can be continued from here and the adjustment answer will be correct *if and only if* there is no correlation between the observations appearing in the condition and those not used. This is true in the present case as

$\mathbf{Q} = \mathbf{I}$. However, we shall change the data of the problem later to demonstrate the point raised here. Continuing now, first,

$$[1 \quad 1 \quad -1]\mathbf{v} = -l_1 - l_2 + l_3 = 3''$$

$$\mathbf{N} = 3, \quad \mathbf{k} = 1'', \quad \text{and} \quad \mathbf{v} = [1'' \quad 1'' \quad -1'']^t$$

Thus

$$\hat{l}_1 = 30°15'16' \qquad \hat{l}_2 = 20°00'01'' \qquad \hat{l}_3 = 50°15'17''$$

$$\hat{l}_4 = l_4 = 30°00'00'' \qquad \hat{l}_5 = l_5 = 70°00'01''$$

$$\mathbf{Q}_{vv} = \tfrac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ & 1 & 1 \\ \text{symmetric} & & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{Q}_{\hat{l}\hat{l}} = \tfrac{1}{3}\begin{bmatrix} 2 & -1 & -1 \\ & 2 & -1 \\ \text{symmetric} & & 2 \end{bmatrix}$$

or, for all five values

$$\mathbf{Q}_{\hat{l}\hat{l}} = \tfrac{1}{3}\begin{bmatrix} 2 & -1 & -1 & 0 & 0 \\ & 2 & -1 & 0 & 0 \\ & & 2 & 0 & 0 \\ & & & 3 & 0 \\ \text{symmetric} & & & & 3 \end{bmatrix}$$

The angle of interest is then

$$\text{angle } DPE = \mathbf{x} = -\hat{l}_2 - \hat{l}_4 + \hat{l}_5 = 20°00°00''$$

$$\mathbf{Q}_{xx} = \tfrac{1}{3}[-1 \quad -1 \quad 1]\begin{bmatrix} 2 & 0 & 0 \\ & 3 & 0 \\ \text{symmetric} & & 3 \end{bmatrix}\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} = \tfrac{8}{3}$$

Note that the 3×3 matrix in the middle of the above expression is formed from the $\mathbf{Q}_{\hat{l}\hat{l}}$ matrix above by deleting the first and third columns and rows and taking the elements delineated by dotted lines. Finally, from the above computed vector of residuals, it is straightforward to compute $\hat{\sigma}_0^2 = 3 \text{ sec}^2$.

Actually the solution according to this case may be considered as simpler than that used in Example 8.4. Both techniques, however, required nothing more than simple computation.

Now we return to the consideration of taking all the observations into the condition equation even though some of them will have zero coefficients. Thus

$$[1 \quad 1 \quad -1 \quad 0 \quad 0]\underset{5,\,1}{\mathbf{v}} = -l_1 - l_2 + l_3 + 0l_4 + 0l_5 = 3''$$

Since $\mathbf{Q} = \mathbf{I}$, then

$$\mathbf{N} = 3 \quad \mathbf{k} = 1'' \quad \text{and} \quad \mathbf{v} = [1'' \quad 1'' \quad -1'' \quad 0 \quad 0]^t$$

which shows that the residuals for l_4 and l_5 are zero. The rest of the solution follows the same steps carried out above.

A very interesting and important case arises when the given observations are correlated a priori. This is treated in the following example.

Example 8.6. Given the same observed angles of Example 8.4 (see Figure 8.2) but with the a priori cofactor matrix

$$\mathbf{Q} = \begin{bmatrix} 2 & 1 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix}$$

instead of the identity matrix used before. Interest remains in computing the angle x using specifically the technique of adjustment of observations only.

Solution: Suppose that we begin by considering the one condition in terms of the three observations l_1, l_2, l_3 as in the preceding example. Thus

$$[1 \quad 1 \quad -1]\mathbf{v} = -l_1 - l_2 + l_3 = 3''$$

Taking the appropriate $\mathbf{Q}$ for the first three observations, then

$$\mathbf{N} = 6 \qquad \mathbf{k} = 0\overset{''}{.}5 \qquad \text{and} \qquad \mathbf{v} = [1\overset{''}{.}5 \quad 1'' \quad -0\overset{''}{.}5]^t$$

Thus

$$\hat{l}_1 = 30°15'16\overset{''}{.}5 \qquad \hat{l}_2 = 20°00'01'' \qquad \hat{l}_3 = 50°15'17\overset{''}{.}5$$

Examining the figure, the required angle would be

$$x = l_5 - \hat{l}_2 - l_4 = 20°00'00''$$

Although this answer agrees with the value computed in the preceding two examples, it is in fact the wrong answer. This is because the other two observations, l_4 and l_5, are correlated with the first three and should be carried in the adjustment as newer estimates for them can be obtained. To prove this, we repeat the above adjustment such that all observations are used.

$$[1 \quad 1 \quad -1 \quad 0 \quad 0]\underset{5,\,1}{\mathbf{v}} = 3''$$

$$\mathbf{N} = [1 \quad 1 \quad -1 \quad 0 \quad 0] \begin{bmatrix} 2 & 1 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} = 6$$

$$\mathbf{k} = 0\overset{''}{.}5 \qquad \text{and} \qquad \mathbf{v} = [1\overset{''}{.}5 \quad 1'' \quad -0\overset{''}{.}5 \quad -0.5'' \quad 0]^t$$

$$\hat{l}_1 = 30°15'16\overset{''}{.}5 \qquad \hat{l}_2 = 20°00'01'' \qquad \hat{l}_3 = 50°15'17\overset{''}{.}5$$

$$\hat{l}_4 = 29°59'59'' \qquad \hat{l}_5 = 70°00'01''$$

$$x = \hat{l}_5 - \hat{l}_2 - \hat{l}_4 = 20°00'00\overset{''}{.}5$$

The value of x is, as expected, different from that computed above, but is the *correct answer*. As a matter of fact it should be made a rule that we must consider first the most general case and then reduce it to a specific one if we can safely do so. In the case when $\mathbf{Q}$ was $\mathbf{I}$ this was possible, but in the presence of correlation this becomes

dangerous. The emphasis on using the most general formulation should apply to all techniques of adjustment and not only to the case discussed.

To complete the example, compute the a posteriori cofactor matrices

$$\mathbf{Q}_{vv} = \tfrac{1}{6} \begin{bmatrix} 9 & 6 & -3 & -3 & 0 \\ & 4 & -2 & -2 & 0 \\ & & 1 & 1 & 0 \\ & & & 1 & 0 \\ \text{symmetric} & & & & 0 \end{bmatrix} \qquad \mathbf{Q}_{ll} = \tfrac{1}{6} \begin{bmatrix} 3 & 0 & 3 & 3 & 0 \\ & 8 & 8 & 2 & 0 \\ & & 11 & 5 & 0 \\ & & & 11 & 6 \\ \text{symmetric} & & & & 12 \end{bmatrix}$$

$$\mathbf{Q}_{xx} = \begin{matrix} (\hat{l}_2 & \hat{l}_4 & \hat{l}_5) \\ [-1 & -1 & 1] \end{matrix} \tfrac{1}{6} \begin{bmatrix} 8 & 2 & 0 \\ 2 & 11 & 6 \\ 0 & 6 & 12 \end{bmatrix} \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} = \frac{23}{6}$$

8.2. COORDINATE TRANSFORMATIONS

Transformation from one system of coordinates to another is a very useful operation that is used frequently in photogrammetry, geodesy, and surveying. There is a whole family of transformations, some of which are applicable to two-dimensional space and others for use in three-dimensional problems. Furthermore the effect of a transformation varies from simple changes of location and direction (without any change in shape and size), to a uniform change in scale (no change in shape), and, finally, to change in shape and size of different degree of nonlinearity.

8.2.1. Transformations in Two-Dimensional Space

1. From polar coordinate system to a cartesian coordinate system

$$x_1 = r \cos \theta$$
$$x_2 = r \sin \theta \tag{8.1}$$

2. From one cartesian coordinate system to another
   a. *Rotation*

$$y_1 = x_1 \cos \alpha - x_2 \sin \alpha$$
$$y_2 = x_1 \sin \alpha + x_2 \cos \alpha \tag{8.2}$$

or

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{8.3}$$

   b. *Rotation and a Scale Change*

$$y_1 = x_1(s \cos \alpha) - x_2(s \sin \alpha)$$
$$y_2 = x_1(s \sin \alpha) + x_2(s \cos \alpha) \tag{8.4}$$

or

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = s \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{8.5}$$

c. *Rotation, Shift, and Scale Change*

$$y_1 = ax_1 - bx_2 + c$$

$$y_2 = bx_1 + ax_2 + d \tag{8.6}$$

This transformation is referred to by several names, such as four-parameter transformation, two-dimensional linear conformal transformation, and similarity transformation.

All three types of transformation given above do not allow for any deformation of the space transformed. The maximum change in the shape is a uniform scaling. The following transformations, however, do not possess that property.

a. Six-parameter transformation

$$y_1 = ax_1 + bx_2 + c$$

$$y_2 = dx_1 + ex_2 + f \tag{8.7}$$

b. Eight-parameter projective transformation (fractional linear)

$$y_1 = \frac{a_1 x_1 + a_2 x_2 + a_3}{c_1 x_1 + c_2 x_2 + 1}$$

$$y_2 = \frac{b_1 x_1 + b_2 x_2 + b_3}{c_1 x_1 + c_2 x_2 + 1} \tag{8.8}$$

c. General two-dimensional polynomial transformation

$$y_1 = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_1 x_2 + a_4 x_1^2 + a_5 x_2^2 + \cdots$$

$$y_2 = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_1 x_2 + b_4 x_1^2 + b_5 x_2^2 + \cdots \tag{8.9}$$

d. Conformal two-dimensional polynomial transformation

$$y_1 = A_0 + A_1 x_1 + A_2 x_2 + A_3(x_1^2 - x_2^2) + A_4(2x_1 x_2) + \cdots$$

$$y_2 = B_0 - A_2 x_1 + A_1 x_2 - A_4(x_1^2 - x_2^2) + A_3(2x_1 x_2) + \cdots \tag{8.10}$$

These are only a few examples of such type of transformation and readers who are familiar with computational photogrammetry recognize that there are many more possibilities.

8.2.2. Transformations in Three-Dimensional Space

1. One of the most commonly used transformations is the *seven-parameter transformation* which allows for three rotations, three translations, and one scale change.

$$\underset{3,1}{\mathbf{y}} = s\underset{3,3}{\mathbf{M}}\underset{3,1}{\mathbf{x}} + \underset{3,1}{\mathbf{k}} \tag{8.11}$$

where

$\mathbf{y} = [y_1\, y_2\, y_3]^t$ coordinates after transformation

$\mathbf{x} = [x_1\, x_2\, x_3]^t$ coordinates before transformation

$\mathbf{k} = [k_1\, k_2\, k_3]^t$ vector of three shifts

$\mathbf{M}$ = an orthogonal matrix in terms of *only* three independent variables (often taken as three sequential rotation angles)

s = scale change

This transformation is also referred to as a linear three-dimensional conformal transformation.

2. General polynomial in three-dimensions

$$y_1 = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_1^2 + a_5 x_2^2 + a_6 x_3^2 + a_7 x_1 x_2$$
$$+ a_8 x_2 x_3 + a_9 x_3 x_1 + \cdots$$

$$y_2 = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_4 x_1^2 + b_5 x_2^2 + b_6 x_3^2 + b_7 x_1 x_2$$
$$+ b_8 x_2 x_3 + b_9 x_3 x_1 + \cdots \tag{8.12}$$

$$y_3 = c_0 + c_1 x_1 + c_2 x_2 + c_3 x_3 + c_4 x_1^2 + c_5 x_2^2 + c_6 x_3^2 + c_7 x_1 x_2$$
$$+ c_8 x_2 x_3 + c_9 x_3 x_1 + \cdots$$

3. Polynomials in three-dimensions which are conformal in the three planes[†]

$$y_1 = A_0 + A_1 x_1 + A_2 x_2 - A_3 x_3 + A_5(x_1^2 - x_2^2 - x_3^2) + 0$$
$$+ 2A_7 x_3 x_1 + 2A_6 x_1 x_2$$

$$y_2 = B_0 - A_2 x_1 + A_1 x_2 + A_4 x_3 + A_6(-x_1^2 + x_2^2 - x_3^2) + 2A_7 x_2 x_3$$
$$+ 0 + 2A_5 x_1 x_2$$

$$y_3 = C_0 + A_3 x_1 - A_4 x_2 + A_1 x_3 + A_7(-x_1^2 - x_2^2 + x_3^2) + 2A_6 x_2 x_3$$
$$+ 2A_5 x_3 x_1 + 0 \tag{8.13}$$

We have given a few different transformations that perform different functions. Now we discuss aspects of adjustment with regard to these transformations.

Example 8.7. (This is a more extensive example than those given on the two-parameter transformation in previous chapters). A transformation between two coordinate systems $\mathbf{x}$ and $\mathbf{y}$ is given by

$$y_1 = ax_1 - bx_2$$

$$y_2 = bx_1 + ax_2$$

† See Mikhail, 1964, in the Bibliography.

where a and b are the parameters in the transformation, (x_1, x_2) the coordinates before transformation, and (y_1, y_2) the coordinates after transformation. In order to estimate the two parameters, three points of known coordinates in both systems are given. A fourth point is known in the x-coordinate system only and its corresponding coordinates in the y-coordinate system is required. The data are given in Table 8.1.

TABLE 8-1

POINT NUMBER	x_1	x_2	y_1	y_2
1	0.0	1.0	-2.1	1.1
2	1.0	0.0	1.0	2.0
3	1.0	1.0	-0.9	2.8
4	1.0	2.0	?	?

Every point in the first coordinate system has the covariance matrix

$$\Sigma_{xx} = \begin{bmatrix} 0.01 & 0.0 \\ 0.0 & 0.01 \end{bmatrix}$$

and in the second system the matrix

$$\Sigma_{yy} = \begin{bmatrix} 0.04 & 0.0 \\ 0.0 & 0.04 \end{bmatrix}.$$

The units are all of the same type, for example, metres for coordinates and (metre)^2 for covariance matrices' elements.

Required: First compute the estimates of a and b and their covariance matrix, then using these compute the y coordinates of point 4 and its covariance matrix. Use point 1 to analyze the question of final coordinates and covariance matrices of the given points in the y-coordinate system.

Solution: This problem of two-parameter transformation was given in a little more simplified form in Example 6.2. In that example a detailed analysis of the mathematical model was given and the reader may wish to review that portion before proceeding here. The position of a point in the y-coordinate system is obviously the end product since the transformation is given as *from* the x *to* the y system. The points that are given by both x and y coordinates present the interesting question of how they should be treated. In Example 6.2 only the x coordinates were considered as observations, whereas here *both* sets of coordinates are observations with given a priori covariance matrices. Therefore for the three points there are $n = 12$ observations and the determination of the minimum number of elements, n_0, is an important question. Each one of these points contributes four observations to the amount of known information, and adds two parameters as well. These two parameters are the final y coordinates of that point which must eventually be determined. Consequently, for all three points a total of *eight* variables or elements

must be considered: two transformation coefficients and six parameters for the y coordinates of the three points. This means that $n_0 = 8$ and then $r = 12 - 8 = 4$ degrees of freedom.

The above analysis for determining n_0 is rather important since it points out important facts. The first point to be observed is that n_0 varies with the amount of given information; that is, it is not independent of n. For example, consider the cases when one, two, or three points are given as in Table 8.2. This fact of different n_0 for different n is often encountered when adjusting transformation problems.

TABLE 8-2

NUMBER OF POINTS	n	n_0	r (redundancy)
1	4	4	0
2	8	6	2
3	12	8	4

Having determined $r = 4$, the question of which technique of least squares to use follows. An attempt could be made to write four conditions among the given 12 observations. However, from the given transformation equations and the treatment of an even simpler case of two-parameter transformation problem in previous examples (for instance, Example 7.2), it is clear that writing such equations is not obvious. Equally unduly complex is to write condition equations for adjustment of indirect observations since this would require the carrying of eight parameters. Since interest is in only two parameters, a and b, and the form of condition equations is already given, then the logical choice would be to have $u = 2$, $c = 6$ and write the linearized conditions for the ith point as

$$\underset{2,4}{\mathbf{A}_i} \; \underset{4,1}{\mathbf{v}_i} + \underset{2,2}{\mathbf{B}_i} \; \underset{2,1}{\Delta} = \underset{2,1}{\mathbf{f}_i}$$

If this equation is written three times for the given three points, then

$$\mathbf{A}_1 \mathbf{v}_1 + \mathbf{B}_1 \Delta = \mathbf{f}_1$$

$$\mathbf{A}_2 \mathbf{v}_2 + \mathbf{B}_2 \Delta = \mathbf{f}_2$$

$$\mathbf{A}_3 \mathbf{v}_3 + \mathbf{B}_3 \Delta = \mathbf{f}_3$$

or

$$\begin{bmatrix} \mathbf{A}_1 & 0 & 0 \\ 0 & \mathbf{A}_2 & 0 \\ 0 & 0 & \mathbf{A}_3 \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \mathbf{v}_3 \end{bmatrix} + \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \\ \mathbf{B}_3 \end{bmatrix} \Delta = \begin{bmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \\ \mathbf{f}_3 \end{bmatrix}$$

or

$$\underset{6,12}{\mathbf{A}} \; \underset{12,1}{\mathbf{v}} + \underset{6,2}{\mathbf{B}} \; \underset{2,1}{\Delta} = \underset{6,1}{\mathbf{f}}$$

The cofactor matrix $\mathbf{Q}_i$ for each point i is a 4×4 matrix. If the reasonable assumption is made that no correlation exists between the coordinates of different points (either the x or the y coordinates), then the total cofactor matrix of the three points is

$$\mathbf{Q}_{12,12} = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Q}_3 \end{bmatrix}$$

Then

$$\mathbf{Q}_e = \mathbf{AQA}^t = \begin{bmatrix} (\mathbf{AQA}^t)_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & (\mathbf{AQA}^t)_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & (\mathbf{AQA}^t)_3 \end{bmatrix}$$

and

$$\mathbf{W}_e = \begin{bmatrix} \mathbf{W}_{e_1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_{e_2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{W}_{e_3} \end{bmatrix}$$

The normal equations coefficient matrix is

$$\mathbf{N}_{2,2} = \mathbf{B}^t(\mathbf{AQA}^t)^{-1}\mathbf{B} = \mathbf{B}^t\mathbf{W}_e\mathbf{B}$$

$$= [\mathbf{B}_1^t \ \mathbf{B}_2^t \ \mathbf{B}_3^t]\mathbf{W}_e \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \\ \mathbf{B}_3 \end{bmatrix} = (\mathbf{B}^t\mathbf{W}_e\mathbf{B})_1 + (\mathbf{B}^t\mathbf{W}_e\mathbf{B})_2 + (\mathbf{B}^t\mathbf{W}_e\mathbf{B})_3$$

or

$$\mathbf{N}_{2,2} = \sum_{i=1}^{3} (\mathbf{B}^t\mathbf{W}_e\mathbf{B})_i = \sum_{i=1}^{3} \mathbf{N}_i \tag{8.14}$$

Similarly the constant term vector $\mathbf{t}$ may be developed as

$$\mathbf{t}_{2,1} = \mathbf{B}^t\mathbf{W}_e\mathbf{f} = [\mathbf{B}_1^t \ \mathbf{B}_2^t \ \mathbf{B}_3^t]\mathbf{W}_e \begin{bmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \\ \mathbf{f}_3 \end{bmatrix}$$

$$= (\mathbf{B}^t\mathbf{W}_e\mathbf{f})_1 + (\mathbf{B}^t\mathbf{W}_e\mathbf{f})_2 + (\mathbf{B}^t\mathbf{W}_e\mathbf{f})_3$$

or

$$\mathbf{t}_{2,1} = \sum_{i=1}^{3} (\mathbf{B}^t\mathbf{W}_e\mathbf{f})_i \tag{8.15}$$

The two results above are interesting because they show that the normal equations may be formed by a summation process. This has the distinct advantage that for programming on a computer, data for each point may be handled at one time, and more important considerable savings in computer space can be effected. Note, for example, that for the $\mathbf{A}$, $\mathbf{Q}$, and $\mathbf{B}$ matrices space required is 2×4, 4×4, and 2×2 instead of 6×12, 12×12, and 6×2, respectively. The space for these component matrices can be used over and over for all points instead of having one set for

each point. This characteristic of summation is, under the above assumption of block diagonality of $\mathbf{A}$ and $\mathbf{Q}$ matrices, common to most transformations. Computer space savings increase as the number of parameters in the transformation increases.

Now we return to our numerical problem and evaluate the matrices for any point i, after writing the equations in the functional form

$$f_{1i} = ax_{1i} - bx_{2i} - y_{1i} = 0$$

$$f_{2i} = bx_{1i} + ax_{2i} - y_{2i} = 0$$

$$\mathbf{A}_i = \begin{bmatrix} \partial f_1/\partial x_1 & \partial f_1/\partial x_2 & \partial f_1/\partial y_1 & \partial f_1/\partial y_2 \\ \partial f_2/\partial x_1 & \partial f_2/\partial x_2 & \partial f_2/\partial y_1 & \partial f_2/\partial y_2 \end{bmatrix}_i = \begin{bmatrix} a^0 & -b^0 & -1 & 0 \\ b^0 & a^0 & 0 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{C} & -\mathbf{I} \\ 2,2 & 2,2 \end{bmatrix} \qquad \text{(independent of "}i\text{")}$$

in which a^0 and b^0 are approximations for the two parameters

$$\mathbf{B}_i = \begin{bmatrix} \partial f_1/\partial a & \partial f_1/\partial b \\ \partial f_2/\partial a & \partial f_2/\partial b \end{bmatrix}_i = \begin{bmatrix} x_1 & -x_2 \\ x_2 & x_1 \end{bmatrix}_i$$

and

$$\Sigma_i \atop 4,4 = \begin{bmatrix} \Sigma_{xx} & \mathbf{0} \\ \mathbf{0} & \Sigma_{yy} \end{bmatrix}_i = 10^{-2} \begin{bmatrix} 1 & 0 & 0 & 0 \\ & 1 & 0 & 0 \\ & & 4 & 0 \\ \text{symmetric} & & & 4 \end{bmatrix} = \sigma_0^2 \, \mathbf{Q}_i$$

With the a priori $\sigma_0^2 = 0.01$, the cofactor matrix of the observations will be taken as

$$\mathbf{Q}_i = \begin{bmatrix} 1 & 0 & 0 & 0 \\ & 1 & 0 & 0 \\ & & 4 & 0 \\ \text{symmetric} & & & 4 \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & 4\mathbf{I} \end{bmatrix}$$

then

$$\mathbf{Q}_{e_i} = \mathbf{A}_i \, \mathbf{Q}_i \, \mathbf{A}_i^t = \mathbf{C}\mathbf{C}^t + 4\mathbf{I} = (a^{02} + b^{02})\mathbf{I} + 4\mathbf{I}$$

$$= (a^{02} + b^{02} + 4)\mathbf{I}$$

$$\mathbf{W}_{e_i} = (a^{02} + b^{02} + 4)^{-1}\mathbf{I} = \alpha\mathbf{I}$$

where $\alpha = 1/(a^{02} + b^{02} + 4)$, and

$$\mathbf{N}_i = \mathbf{B}_i^t \, \mathbf{W}_{e_i} \, \mathbf{B}_i = \alpha(x_1^2 + x_2^2)_i \, \mathbf{I}$$

Thus the total normal equation matrix will then be

$$\mathbf{N} = \alpha \sum_{i=1}^{3} (x_1^2 + x_2^2)_i \, \mathbf{I} = \alpha(1 + 1 + 2)\mathbf{I} = 4\alpha\mathbf{I}$$

noting that α is the only value that varies from one iteration to the next because of updating the approximations a^0 and b^0.

In a similar manner the constant term vector is

$$
\mathbf{t} = \sum_{i=1}^{3} \mathbf{t}_i = \sum_{i=1}^{3} \alpha \mathbf{B}_i^t \mathbf{f}_i = \alpha \sum_{i=1}^{3} \mathbf{B}_i^t \mathbf{f}_i = \alpha \begin{bmatrix} \sum (x_1 \, f_1 + x_2 \, f_2)_i \\ \sum (-x_2 \, f_1 + x_1 \, f_2)_i \end{bmatrix}
$$

It is important to note that the vector $\mathbf{t}$ is more dependent on the approximations a^0 and b^0 because it is directly related to the vector of so-called discrepancies $\mathbf{f}$. The latter vector will approach a minimum when the solution is iterated until practically no improvement takes place.

Using the matrix $\mathbf{N}$ and vector $\mathbf{t}$ the correction vector $\mathbf{\Delta}$ may be computed as

$$
\underset{2,\,2}{\mathbf{\Delta}} = \mathbf{N}^{-1}\mathbf{t} = \frac{1}{4} \begin{bmatrix} \sum_{i=1}^{3} (x_1 \, f_1 + x_2 \, f_2)_i \\ \sum_{i=1}^{3} (-x_2 \, f_1 + x_1 \, f_2)_i \end{bmatrix}
$$

Now, consider the question of choosing approximations for the parameters. Taking the rather easy choice first begin by $a^0 = b^0 = 0$. Thus

Case 1. $a^0 = 0$ and $b^0 = 0$

$$
\mathbf{f}_1 = \begin{bmatrix} -2.1 \\ 1.1 \end{bmatrix} \qquad \mathbf{f}_2 = \begin{bmatrix} 1.0 \\ 2.0 \end{bmatrix} \qquad \mathbf{f}_3 = \begin{bmatrix} -0.9 \\ 2.8 \end{bmatrix}
$$

$$
\sum (x_1 \, f_1 + x_2 \, f_2)_i = 1.1 + 1.0 + 1.9 = 4.0
$$

$$
\sum (-x_2 \, f_1 + x_1 \, f_2)_i = 2.1 + 2.0 + 3.7 = 7.8
$$

Then

$$
\mathbf{\Delta} = \frac{1}{4} \begin{bmatrix} 4.0 \\ 7.8 \end{bmatrix} = \begin{bmatrix} 1.0 \\ 1.95 \end{bmatrix}
$$

and

$$
\mathbf{Q}_{\Delta\Delta} = \mathbf{N}^{-1} = \frac{1}{4\alpha}\mathbf{I} = \mathbf{I}
$$

since $\alpha = \frac{1}{4}$ when $a^0 = b^0 = 0$ which with a priori $\sigma_0^2 = 0.01$ leads to the covariance matrix of the parameters

$$
\mathbf{\Sigma}_{\Delta\Delta} = \begin{bmatrix} 0.01 & 0 \\ 0 & 0.01 \end{bmatrix}
$$

This, in fact, ends only the first iteration, and since the corrections are obviously quite large, at least another iteration is necessary. To start the second iteration, therefore,

$$
a^0 = 1.0 \qquad b^0 = 1.95
$$

and

$$
\mathbf{f}_1 = \begin{bmatrix} -0.15 \\ 0.10 \end{bmatrix} \qquad \mathbf{f}_2 = \begin{bmatrix} 0.0 \\ 0.05 \end{bmatrix} \qquad \mathbf{f}_3 = \begin{bmatrix} +0.05 \\ -0.15 \end{bmatrix}
$$

$$
\sum (x_1 \, f_1 + x_2 \, f_2)_i = 0.10 + 0.0 - 0.10 = 0
$$

$$
\sum (-x_2 \, f_1 + x_1 \, f_2)_i = 0.15 + 0.05 - 0.20 = 0
$$

With a zero t vector then $\Delta = [0 \quad 0]^t$ and the final parameters are $\hat{a} = 1.0$ and $\hat{b} = 1.95$

It is very important to realize that in order to get the correct cofactor matrix of the parameters, the last N^{-1} must be computed, otherwise the one already available may be grossly in error. In this case the new N^{-1} is

$$N^{-1} = \frac{1}{4\alpha} I = \frac{a^{02} + b^{02} + 4}{4} I = \frac{8.8025}{4} I = 2.2006I$$

Thus the covariance matrix of the final estimates of the parameters is

$$\Sigma_{xx} = \begin{bmatrix} 0.022 & 0.0 \\ 0.0 & 0.022 \end{bmatrix}$$

Comparing this matrix with the one computed at the end of the first iteration reveals a variation of over 200%, indicating how serious an error may be committed if proper attention is not paid when solving a problem by iteration.

Case 2: In this case consider an approach of computing approximations for the parameters with reasonable effort. For example, we might apply a minimum solution by choosing, say, the first point to solve for a and b. Thus

$$-2.1 = -b \quad \text{or} \quad b^0 = 2.0 \quad \text{(rounded off to whole number)}$$
$$1.1 = \quad a \quad \text{or} \quad a^0 = 1.0 \quad \text{(rounded off to whole number)}$$

(Note that if the second point was used, the above approximations would have been obtained without rounding off. If the third point is used, a little more computation will be involved but with about the same resulting approximations.)

Another possibility, which in principle would yield slightly better approximations, is to compute the means of all coordinates and use them to solve a minimum case. Although this is not a demanding effort in the present example, it could be a rather unwarranted computation in more involved problems.

With $a^0 = 1.0$ and $b^0 = 2.0$, then

$$f_1 = \begin{bmatrix} -0.10 \\ 0.10 \end{bmatrix} \quad f_2 = \begin{bmatrix} 0.0 \\ 0.0 \end{bmatrix} \quad f_3 = \begin{bmatrix} 0.10 \\ -0.20 \end{bmatrix}$$

$$\sum (x_1 \, f_1 + x_2 \, f_2)_i = 0.10 + 0.0 - 0.10 = 0$$

$$\sum (-x_2 \, f_1 + x_1 \, f_2)_i = 0.10 + 0.0 - 0.3 = -0.2$$

$$\Delta = \frac{1}{4} \begin{bmatrix} 0 \\ -0.2 \end{bmatrix} = \begin{bmatrix} 0 \\ -0.05 \end{bmatrix}$$

leading directly to the final result of $\hat{a} = 1.0$ and $\hat{b} = 1.95$. Now, to compute $Q_{\Delta\Delta}$ we should advisably use $\hat{a}$ and $\hat{b}$ directly to compute the final N^{-1}. If we were to use the original approximations, we would get

$$Q_{\Delta\Delta} = N^{-1} = \frac{a^{02} + b^{02} + 4}{4} I = 2.25I$$

which would lead to

$$\Sigma_{xx} = \begin{bmatrix} 0.0225 & 0 \\ 0 & 0.0225 \end{bmatrix}$$

The variation in this case is only slightly over 2% and this is because the computed Δ is about the same percentage different from the final answer. Although such difference may be sufficiently small to be negligible, it is advisable always to update the cofactor matrix at the last iteration.

An important final remark is worth mentioning here, which is that we should always obtain the best possible approximation vector that requires a reasonable effort of computation. As has been shown above, with good approximations it was not even necessary to perform another iteration. The reader must be cautioned, however, that this is true here only because the condition equations are almost linear (of course, if they were in fact linear, no iteration would be necessary). Nevertheless, we would expect that, in general, choosing better approximations would lead to fewer iterations.

Coordinates of Point 4 in the **y** *System*

$$\hat{y}_{14} = \hat{a}x_{14} - \hat{b}x_{24} = -2.90$$

$$\hat{y}_{24} = \hat{b}x_{14} + \hat{a}x_{24} = 3.95$$

To compute the covariance matrix of the vector $\hat{\mathbf{y}}_4$ we should realize that all variables $(x_1, x_2)_4$ and $(\hat{a}, \hat{b})$ are random variables with known covariance matrices. Since point 4 was not one of the points used in estimating $\hat{a}, \hat{b}$, there would be no correlation between the covariance matrix of the coordinates and that of the parameters. To facilitate the computation introduce the auxiliary vector

$$\mathbf{r} = [\hat{a} \quad \hat{b} \quad x_{14} \quad x_{24}]^t$$

and thus

$$\underset{4,\,4}{\Sigma_{rr}} = \begin{bmatrix} 0.022\mathbf{I}_2 & \mathbf{0} \\ \mathbf{0} & 0.01\mathbf{I}_2 \end{bmatrix}$$

Hence

$$\Sigma_{\hat{y}_4\hat{y}_4} = \mathbf{J}\Sigma_{rr}\mathbf{J}^t$$

where

$$\mathbf{J} = \frac{\partial \mathbf{y}_4}{\partial \mathbf{r}} = \begin{bmatrix} x_{14} & -x_{24} & \hat{a} & -\hat{b} \\ x_{24} & x_{14} & \hat{b} & \hat{a} \end{bmatrix} = \begin{bmatrix} 1 & -2 & 1 & -1.95 \\ 2 & 1 & 1.95 & 1 \end{bmatrix}$$

Thus

$$\Sigma_{\hat{y}_4\hat{y}_4} = \begin{bmatrix} 0.158 & 0.0 \\ 0.0 & 0.158 \end{bmatrix}$$

The Coordinates of Point 1: Point 1 was used in the transformation to estimate the two parameters a and b. It was therefore given with coordinates in both systems of coordinates **x** and **y**. However, using the same procedure that was applied to point 4 above from the known·**x** coordinates and estimated parameters $\hat{a}$ and $\hat{b}$, another pair of "transformed" **y** coordinates may be computed:

$$y'_{11} = \hat{a}x_{11} - \hat{b}x_{21} = -1.95$$

$$y'_{12} = \hat{b}x_{11} + \hat{a}x_{21} = 1.00$$

These coordinates are obviously different from those given a priori, and it is not the final position of the point. The final coordinates in the y system will be computed shortly, but first we evaluate the covariance matrix of these coordinates. Again, let an auxiliary vector be defined as

$$\mathbf{t} = [a \quad b \quad x_{11} \quad x_{21}]^t$$

The covariance matrix Σ_{tt} may be constructed in the form
$$\underset{4,\,4}{}$$

$$\Sigma_{tt} = \begin{bmatrix} 0.022\mathbf{I} & \Sigma_{\Delta x1} \\ \Sigma_{x1\Delta} & 0.01\mathbf{I}_2 \end{bmatrix}$$

Unlike the case of point 4, the matrix $\Sigma_{\Delta x_1}$ is not necessarily zero because point 1 was used in the transformation and therefore $(\hat{a}, \hat{b})$ and $(x_1, x_2)_1$ are expected to be correlated. To compute this covariance matrix it should be realized that $(x_1, x_2)_1$ is a part of the total observation vector $l_{12,\,1}$. Referring to equation (6.32) the total cofactor matrix would be

$$\underset{2,\,12}{\mathbf{Q}_{\Delta l}} = -\underset{2,\,2}{\mathbf{N}^{-1}} \quad \underset{2,\,6}{\mathbf{B}^t} \quad \underset{\leftarrow 6,\,6\rightarrow}{\left(\mathbf{A}\mathbf{Q}\mathbf{A}^t\right)^{-1}} \quad \underset{6,\,12}{\mathbf{A}} \quad \underset{12,\,12}{\mathbf{Q}}$$

$$= -\mathbf{N}^{-1}[\mathbf{B}_1^t \ \mathbf{B}_2^t \ \mathbf{B}_3^t]\begin{bmatrix} \mathbf{W}_{e_1} & & \\ & \mathbf{W}_{e_2} & \\ & & \mathbf{W}_{e_3} \end{bmatrix}\begin{bmatrix} \mathbf{A}_1 & & \\ & \mathbf{A}_2 & \\ & & \mathbf{A}_3 \end{bmatrix}\begin{bmatrix} \mathbf{Q}_1 & & \\ & \mathbf{Q}_2 & \\ & & \mathbf{Q}_3 \end{bmatrix}$$

$$= -\underset{2,\,2}{\mathbf{N}^{-1}} \left[\underset{\leftarrow 2,\,4\rightarrow}{\left(\mathbf{B}^t\mathbf{W}_e\mathbf{A}\mathbf{Q}\right)_1} \underset{\leftarrow 2,\,4\rightarrow}{\left(\mathbf{B}^t\mathbf{W}_e\mathbf{A}\mathbf{Q}\right)_2} \underset{\leftarrow 2,\,4\rightarrow}{\left(\mathbf{B}^t\mathbf{W}_e\mathbf{A}\mathbf{Q}\right)_3}\right]$$

Since we are interested in point 1 only, then

$$\underset{2,\,4}{\mathbf{Q}_{\Delta l_1}} = -\mathbf{N}^{-1}\mathbf{B}_1^t \ \mathbf{W}_{e_1}\mathbf{A}_1 \ \mathbf{Q}_1 = -\mathbf{N}^{-1}\mathbf{B}_1^t \ \mathbf{W}_{e_1}[\mathbf{C}\mathbf{Q}_{x_1} - \mathbf{Q}_{y_1}] \qquad (8.16)$$

where

$$\mathbf{C} = \begin{bmatrix} \hat{a} & -\hat{b} \\ \hat{b} & \hat{a} \end{bmatrix}, \qquad \mathbf{Q}_{x_1} = \mathbf{I}_2, \qquad \text{and} \qquad \mathbf{Q}_{y_1} = 4\mathbf{I}_2$$

The cofactor matrix required is that which relates the estimated parameters and the x coordinates only. Thus

$$\mathbf{Q}_{\Delta x_1} = -\mathbf{N}^{-1}\mathbf{B}_1^t \ \mathbf{W}_{e1}\mathbf{C}\mathbf{Q}_{x1}$$

or

$$\mathbf{Q}_{\Delta x_1} = \frac{-1}{4\alpha}\mathbf{I}_2\begin{bmatrix} 0 & +1 \\ -1 & 0 \end{bmatrix}\alpha\mathbf{I}_2\begin{bmatrix} 1 & -1.95 \\ 1.95 & 1 \end{bmatrix}\mathbf{I}_2$$

$$= \frac{1}{4}\begin{bmatrix} -1.95 & -1 \\ 1 & -1.95 \end{bmatrix} = \begin{bmatrix} -0.4875 & -0.25 \\ 0.25 & -0.4875 \end{bmatrix}$$

or, the covariance matrix (remembering that $\sigma_0^2 = 0.01$)

$$\Sigma_{\Delta x_1} = 0.01\begin{vmatrix} -0.4875 & -0.25 \\ 0.25 & -0.4875 \end{vmatrix}$$

The covariance matrix of the vector **t** is therefore

$$\underset{4,\,4}{\Sigma_{tt}} = 10^{-2} \begin{bmatrix} 2.2 & 0 & -0.4875 & -0.25 \\ & 2.2 & 0.25 & -0.4875 \\ & & 1.0 & 0 \\ \text{symmetric} & & & 1.0 \end{bmatrix}$$

Finally, the covariance matrix of the transformed coordinates $\Sigma_{y_1'y_1'}$ is computed as

$$\Sigma_{y_1'y_1'} = \mathbf{J}'\Sigma_{tt}\,\mathbf{J}'^t$$

where

$$\mathbf{J}' = \begin{bmatrix} x_{11} & -x_{21} & \hat{a} & -\hat{b} \\ x_{21} & x_{11} & \hat{b} & \hat{a} \end{bmatrix} = \begin{bmatrix} 0 & -1 & 1 & -1.95 \\ 1 & 0 & 1.95 & 1 \end{bmatrix}$$

or

$$\Sigma_{y_1'y_1'} = 10^{-2}\begin{bmatrix} 4.6019 & 0 \\ 0 & 4.6019 \end{bmatrix}$$

Thus the "transformed" coordinates of point 1 are

$$y_1' = [-1.95 \quad 1.00]^t$$

and the covariance matrix is

$$\Sigma_{y_1'y_1'} = 10^{-2}\begin{bmatrix} 4.6019 & 0 \\ 0 & 4.6019 \end{bmatrix}$$

A more interesting question still remains, however; and that is what is the best estimate $\hat{y}_1$ for the coordinates of this point in the y system and its covariance matrix? To answer this question, recall the available information.

First, the given observational vector $y = [-2.1 \quad 1.1]^t$ and its covariance matrix $\Sigma_{yy} = 0.04\mathbf{I}_2$; and another estimate $y' = [-1.95 \quad 1.0]^t$ and its covariance matrix $\Sigma_{y'y'} = 0.0460191\mathbf{I}_2$. (We have dropped the subscript 1 for simplicity). These two vectors together may be constructed as an observational vector

$$\underset{4,\,1}{l} = \begin{bmatrix} y' \\ y \end{bmatrix} = [-1.95 \quad 1 \quad -2.1 \quad 1.1]^t$$

To construct the covariance matrix Σ_{ll} of these observations, care must be exercised in establishing whether the two vectors y and y' are correlated or not. The vector y' was computed in the adjustment and we must find out if y was involved in its computation. Recalling those equations from which y' was computed

$$y' = \begin{bmatrix} x_1 & -x_2 \\ x_2 & x_1 \end{bmatrix}\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = \mathbf{B}\Delta$$

In the process of estimating $\Delta = [\hat{a} \quad \hat{b}]^t$, the vector y was in fact used in the least squares solution. Therefore, working with cofactors for the moment, $\mathbf{Q}_{y'y}$ may be computed as follows: Given

$$y' = \mathbf{B}\Delta \qquad \text{and} \qquad y = y$$

and

$$Q_{\Delta y_1} = -N^{-1}B_1^t \ W_{e_1}(-Q_{y_1}) \qquad \text{[from equation (8.16)]}$$

Again, dropping the subscript 1, then

$$Q_{\Delta y} = N^{-1}B^t W_e \ Q_y$$

Then

$$Q_{y'y} = \frac{\partial y'}{\partial \Delta} Q_{\Delta y} \frac{\partial y}{\partial y}$$

or

$$Q_{y'y} = BN^{-1}B^t W_e \ Q_y$$

$$= \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \frac{1}{4\alpha} I_2 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \alpha I_2 \cdot 4I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$$

Therefore

$$Q_{ll} = \begin{bmatrix} 4.6019I_2 & I_2 \\ I_2 & 4I_2 \end{bmatrix} \qquad W_{ll} \atop 4,4 = \frac{1}{17.4076} \begin{bmatrix} 4I & -I \\ -I & 4.6019I \end{bmatrix}$$

To obtain the best estimate $\hat{y}$ we apply the least squares method of adjustment of indirect observations. Also, to further the reader's familiarity with shorter notations, we perform the operations directly in vectors. Thus

$$v_1 - \hat{y} = -y'$$

$$v_2 - \hat{y} = -y$$

or

$$\underset{4,1}{v} + \begin{bmatrix} -I \\ -I \end{bmatrix} \underset{2,1}{\hat{y}} = \begin{bmatrix} -y' \\ -y \end{bmatrix}$$

which is of the form $v + B\Delta = f$. Next,

$$N = \frac{6.6019}{17.4076} I, \qquad N^{-1} = 2.6368I, \qquad \text{and} \qquad t = \frac{1}{17.4076} (3y' + 3.6019y)$$

$$y = N^{-1}t = \frac{1}{6.6019} (3y' + 3.6019y) = \begin{bmatrix} -2.03184 \\ 1.05456 \end{bmatrix}$$

$$Q_{\hat{y}\hat{y}} = N^{-1} = 2.6368I \qquad \Sigma_{\hat{y}\hat{y}} = (0.01)(2.6368)I = 0.026368I$$

Thus the final estimate of the y coordinates of point 1 are $(-2.0384, 1.05456)$ and its covariance matrix is $0.026368I$.

In the preceding analysis and computation many rather important steps were involved. Such steps demonstrate several features of the concepts of the functional and, particularly, the stochastic models. We shall give the discussion of these features at the end of the example. We shall now continue the analysis of obtaining the final y-coordinate estimates for the reference points and the covariance matrices by another process.

It should be remembered that both the **x** and **y** coordinates were considered as observations. This means that both sets of coordinates, for each of the reference points, receive residuals after the adjustment in order to obtain the estimated $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ coordinates. These estimated coordinates must satisfy the transformation equations since this is one of the basic premises of the least squares adjustment. Therefore when $\hat{\mathbf{x}}$ coordinates are transformed by $\hat{a}$ and $\hat{b}$, the resulting transformed coordinates must be equal to the $\hat{\mathbf{y}}$ coordinates, and that these coordinates are the final best estimate of the position of the point in the **y** system. The total vector of residuals is (recalling that the final Δ is zero)

$$\begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \mathbf{v}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_1 & & \\ & \mathbf{Q}_2 & \\ & & \mathbf{Q}_3 \end{bmatrix} \begin{bmatrix} \mathbf{A}_1^t & & \\ & \mathbf{A}_2^t & \\ & & \mathbf{A}_3^t \end{bmatrix} \begin{bmatrix} \mathbf{W}_{e_1} & & \\ & \mathbf{W}_{e_2} & \\ & & \mathbf{W}_{e_3} \end{bmatrix} \begin{bmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \\ \mathbf{f}_3 \end{bmatrix}$$

from which we can write for the residual vector of point 1

$$\mathbf{v}_1 = \mathbf{Q}_1 \, \mathbf{A}_1^t \, \mathbf{W}_{e_1} \, \mathbf{f}_1$$

Continuing further with the partitioning

$$\begin{bmatrix} \mathbf{v}_x \\ \mathbf{v}_y \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_{xx} & 0 \\ 0 & \mathbf{Q}_{yy} \end{bmatrix}_1 \begin{bmatrix} \mathbf{C}^t \\ -\mathbf{I} \end{bmatrix} \alpha \mathbf{I} \mathbf{f}_1$$

$$\mathbf{v}_{x_1} = \alpha \mathbf{Q}_{xx} \, \mathbf{C}^t \mathbf{f}_1 = \alpha \mathbf{C}^t \mathbf{f}_1$$

$$\mathbf{v}_{y_1} = -\alpha \mathbf{Q}_{yy} \mathbf{f}_1 = -4\alpha \mathbf{f}_1$$

with

$$\alpha = (\hat{a}^2 + \hat{b}^2 + 4)^{-1} = \frac{1}{8.8025} = 0.1136$$

$$\mathbf{C} = \begin{bmatrix} \hat{a} & -\hat{b} \\ \hat{b} & \hat{a} \end{bmatrix} = \begin{bmatrix} 1.0 & -1.95 \\ 1.95 & 1.0 \end{bmatrix}, \quad \mathbf{f} = \begin{bmatrix} -0.15 \\ 0.10 \end{bmatrix}$$

then

$$\mathbf{v}_x = \begin{bmatrix} 0.00511 \\ 0.04459 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_y = \begin{bmatrix} 0.0682 \\ -0.0454 \end{bmatrix}$$

The estimated "observations" are therefore

$$\hat{\mathbf{x}} = \mathbf{x} + \mathbf{v}_x = \begin{bmatrix} 0.00511 \\ 1.04459 \end{bmatrix} \quad \hat{\mathbf{y}} = \mathbf{y} + \mathbf{v}_y = \begin{bmatrix} -2.0318 \\ 1.0546 \end{bmatrix}$$

Using $\hat{a}$, $\hat{b}$, and $\hat{\mathbf{x}}$, an estimate for the **y** coordinates may be computed again

$$\hat{\mathbf{y}} = \begin{bmatrix} \hat{a} & -\hat{b} \\ \hat{b} & \hat{a} \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} = \begin{bmatrix} 1.0 & -1.95 \\ 1.95 & 1.0 \end{bmatrix} \begin{bmatrix} 0.00511 \\ 1.04459 \end{bmatrix} = \begin{bmatrix} -2.03183 \\ 1.05455 \end{bmatrix}$$

Thus all three procedures of computing $\hat{\mathbf{y}}$ for reference point number 1 resulted in exactly the same values. To complete the analysis, evaluate the corresponding cofactor matrix $\mathbf{Q}_{\hat{y}\hat{y}}$. First

$$\mathbf{Q}_{vv} = \mathbf{Q}\mathbf{A}^t\mathbf{W}_e(\mathbf{I} - \mathbf{B}\mathbf{N}^{-1}\mathbf{B}^t\mathbf{W}_e)\mathbf{A}\mathbf{Q}$$

Remembering that $\mathbf{Q}$, $\mathbf{A}$, and $\mathbf{W}_e$ are block diagonal matrices, examine the expression

$$\underset{6,6}{\mathbf{K}} = \underset{6,2}{\mathbf{B}} \ \underset{2,2}{\mathbf{N}^{-1}} \ \underset{2,6}{\mathbf{B}^t} \ \underset{6,6}{\mathbf{W}_e} = \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \\ \mathbf{B}_3 \end{bmatrix} \mathbf{N}^{-1} [\mathbf{B}_1^t \ \mathbf{B}_2^t \ \mathbf{B}_3^t] \begin{bmatrix} \mathbf{W}_{e_1} & & \\ & \mathbf{W}_{e_2} & \\ & & \mathbf{W}_{e_3} \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{B}_1 \mathbf{N}^{-1}\mathbf{B}_1^t \mathbf{W}_{e_1} & \mathbf{B}_1 \mathbf{N}^{-1}\mathbf{B}_2^t \mathbf{W}_{e_2} & \mathbf{B}_1 \mathbf{N}^{-1}\mathbf{B}_3^t \mathbf{W}_{e_3} \\ \mathbf{B}_2 \mathbf{N}^{-1}\mathbf{B}_1^t \mathbf{W}_{e_1} & \mathbf{B}_2 \mathbf{N}^{-1}\mathbf{B}_2^t \mathbf{W}_{e_2} & \mathbf{B}_2 \mathbf{N}^{-1}\mathbf{B}_3^t \mathbf{W}_{e_3} \\ \mathbf{B}_3 \mathbf{N}^{-1}\mathbf{B}_1^t \mathbf{W}_{e_1} & \mathbf{B}_3 \mathbf{N}^{-1}\mathbf{B}_2^t \mathbf{W}_{e_2} & \mathbf{B}_3 \mathbf{N}^{-1}\mathbf{B}_3^t \mathbf{W}_{e_3} \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{K}_{11} & \mathbf{K}_{12} & \mathbf{K}_{13} \\ \mathbf{K}_{21} & \mathbf{K}_{22} & \mathbf{K}_{23} \\ \mathbf{K}_{31} & \mathbf{K}_{32} & \mathbf{K}_{33} \end{bmatrix}$$

and then

$$\underset{12,12}{\mathbf{Q}_{vv}} =$$

$$\begin{bmatrix} \mathbf{Q}_1 \mathbf{A}_1^t \mathbf{W}_{e1} & & \\ & \mathbf{Q}_2 \mathbf{A}_2^t \mathbf{W}_{e2} & \\ & & \mathbf{Q}_3 \mathbf{A}_3^t \mathbf{W}_{e3} \end{bmatrix} (\mathbf{I} - \mathbf{K}) \begin{bmatrix} \mathbf{A}_1 \mathbf{Q}_1 & & \\ & \mathbf{A}_2 \mathbf{Q}_2 & \\ & & \mathbf{A}_3 \mathbf{Q}_3 \end{bmatrix}$$

From the above structure of $\mathbf{Q}_{vv}$, any submatrix of interest can be directly extracted. For example, the 4×4 submatrix of cofactors for the first point (with respect to itself) is

$$\mathbf{Q}_{v_1 v_1} = \mathbf{Q}_1 \mathbf{A}_1^t \mathbf{W}_{e_1}(\mathbf{I} - \mathbf{B}_1 \mathbf{N}^{-1}\mathbf{B}_1^t \mathbf{W}_{e_1})\mathbf{A}_1 \mathbf{Q}_1$$

for which

$$(\mathbf{I} - \mathbf{B}_1 \mathbf{N}^{-1}\mathbf{B}_1^t \mathbf{W}_{e_1}) = \mathbf{I} - \mathbf{B}_1 \frac{1}{4\alpha}\mathbf{I}\mathbf{B}_1^t \alpha\mathbf{I} = \mathbf{I} - \frac{1}{4}\mathbf{B}_1 \mathbf{B}_1^t$$

$$= \mathbf{I} - \frac{(x_1^2 + x_2^2)}{4}\mathbf{I} = 0.75\mathbf{I}$$

and

$$\mathbf{Q}_{v_1 v_1} = \begin{bmatrix} \mathbf{I} & 0 \\ 0 & 4\mathbf{I} \end{bmatrix}\begin{bmatrix} \mathbf{C}^t \\ -\mathbf{I} \end{bmatrix} 0.75\alpha\mathbf{I}[\mathbf{C} \quad -\mathbf{I}]\begin{bmatrix} \mathbf{I} & 0 \\ 0 & 4\mathbf{I} \end{bmatrix}$$

$$= 0.0852\begin{bmatrix} \mathbf{C}^t\mathbf{C} & 4\mathbf{C}^t \\ 4\mathbf{C} & 16\mathbf{I} \end{bmatrix}$$

$$\mathbf{Q}_{v_1 v_1} = \begin{bmatrix} 0.4092 & 0 & -0.6632 & -0.3408 \\ & 0.4092 & 0.3408 & -0.6632 \\ & & 1.3632 & 0 \\ \text{symmetric} & & & 1.3632 \end{bmatrix}$$

$$\mathbf{Q}_{l_1 l_1} = \mathbf{Q}_1 - \mathbf{Q}_{v_1 v_1} = \begin{bmatrix} 0.5908 & 0 & \vdots & 0.6632 & 0.3408 \\ & 0.5908 & \vdots & -0.3403 & 0.6632 \\ \hdashline & & \vdots & 2.6368 & 0 \\ \text{symmetric} & & \vdots & & 2.6368 \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{Q}_{\hat{x}\hat{x}} & \mathbf{Q}_{\hat{x}\hat{y}} \\ \mathbf{Q}_{\hat{y}\hat{x}} & \mathbf{Q}_{\hat{y}\hat{y}} \end{bmatrix}$$

with the numerical values of the submatrices easily obtained from the partitioning lines.

At this point, it is clear that $Q_{\hat{y}\hat{y}} = 2.6368I$ is exactly the same as was computed before. To complete the computation, derive $Q_{\hat{y}\hat{y}}$ by propagation applied to $\hat{y} = C\hat{x}$. As before, both $\hat{a}$, $\hat{b}$ (elements of C) and $\hat{x}_1$, $\hat{x}_2$ (elements of $\hat{x}$) are random variables. Placing these in a vector

$$t' = [\hat{a} \quad \hat{b} \quad \hat{x}_1 \quad \hat{x}_2]'$$

the corresponding cofactor matrix would be

$$Q_{t't'} = \begin{bmatrix} Q_{\Delta\Delta} & Q_{\Delta\hat{x}} \\ Q_{\hat{x}\Delta} & Q_{\hat{x}\hat{x}} \end{bmatrix}$$

From paragraph 6.2.3 the cross cofactor matrix between parameters and original observations is the same as between parameters and estimated observations, thus

$$Q_{\Delta\hat{x}} = Q_{\Delta x} = \begin{bmatrix} -0.4875 & -0.25 \\ 0.25 & -0.4875 \end{bmatrix}$$

Having

$$Q_{\Delta\Delta} = 2.2006I \quad \text{and} \quad Q_{\hat{x}\hat{x}} = 0.5908I$$

then

$$Q_{t't'} = \begin{bmatrix} 2.2006 & 0 & -0.4875 & -0.25 \\ & 2.2006 & 0.25 & -0.4875 \\ & & 0.5908 & 0 \\ \text{symmetric} & & & 0.5908 \end{bmatrix}$$

Cofactor propagation yields

$$Q_{\hat{y}\hat{y}} = JQ_{\hat{x}} J'$$

where

$$J = \begin{bmatrix} \hat{x}_1 & -\hat{x}_2 & \hat{a} & -\hat{b} \\ \hat{x}_2 & \hat{x}_1 & \hat{b} & \hat{a} \end{bmatrix} = \begin{bmatrix} 0 & -1.0 & 1.0 & -1.95 \\ 1.0 & 0 & 1.95 & 1.0 \end{bmatrix}$$

or

$$Q_{\hat{y}\hat{y}} = \begin{bmatrix} 2.6368 & 0 \\ 0 & 2.6368 \end{bmatrix} = 2.6368I$$

which is identical to the two answers already obtained.

GENERAL DISCUSSION

This example has been treated extensively because several important aspects were demonstrated in the process of solving what appeared to be a relatively simple problem. In the course of this example a number of points should be recognized.

1. Either or both sets of coordinates may be taken as observations. In Example 6.2, only the x coordinates were considered as observations, but in

this example both sets were considered. A comparison of both examples would be useful for recognizing differences in procedure.

2. The technique of adjusting observations and parameters is more suitable for transformation problems than other techniques, particularly when all coordinates are taken as observations with cofactor matrices different from unit matrices.

3. If the a priori cofactor matrix of the observations is block diagonal (by point), then the normal equations can be formed in a summation process in which only the contribution of one point is computed at one time. This technique may lead to considerable savings in computer space and possibly computing time.

4. We should try to obtain the best approximations for the parameters that are practicable, without excessive computational effort. One practical way with transformation problems, and for that matter most other problems, is to perform a unique solution using a minimum data set. Having good approximations may result in having to iterate the solution only once or twice.

5. It is possible that the first set of corrections Δ will be all that is needed to yield the required answer. In such cases it might happen that the magnitude of Δ is large relative to the vector of approximations x^0. Here we must be sure to update the inverse of the normal equations, N^{-1}, using $x = x^0 + \Delta$ in order to obtain the correct estimate of $Q_{\hat{x}\hat{x}}$. If this is not done, the error could be considerable (200% in this example).

6. The computation of transformed coordinates brought to light several quite important points. For example, it is very important when propagating cofactors (or covariances) to determine *all* the random variables in the equations *and* to construct the correct cofactor (or covariance) matrix of the independent variables (that is, the variables on the right side of the equal sign). Of utmost importance here is the determination of correlation whenever it exists.

7. A fact, which is probably obvious but does merit mentioning, is that the estimated observations $\hat{l}$ satisfy the condition equations. In the case of transformation, realizing this fact facilitates considerably the computation of the final estimates of the coordinates of the reference points used in the transformation.

8. This last remark is perhaps the most important as it relates to the fundamental concepts of observations and the stochastic model. At the beginning of the example, the given coordinates were considered the observations with their given cofactor matrices. In the computations subsequent to estimating the transformation parameters, these parameter estimates $(\hat{a}, \hat{b})$ were then treated as observations, or stochastic variables. The same applies to the estimated observations as well. In these computations several quite different procedures were used to compute the same quantity of interest. The reason all the answers turned out to be identical is that in each case the

proper stochastic model was constructed and the resulting correct cofactor matrix used. Any deviation from that would have resulted in inconsistent answers. In brief, then, observations may be the actual numeric outputs of the operations we call "measurements," but may also be derived quantities. Whichever quantities we choose to term observations, we must set up the proper stochastic model that fits our chosen values.

Example 8.8. *Four-Parameter Transformation.* As before, the four-parameter transformation is given by

$$y_1 = ax_1 - bx_2 + c$$

$$y_2 = bx_1 + ax_2 + d \tag{8.6}$$

This transformation is one of the most commonly used types in photogrammetry and allied areas. Under certain assumptions, we can apply a simplified approach to least squares adjustment by writing the conditions as

$$v_1 + ax_1 - bx_2 + c - y_1 = 0$$

$$v_2 + bx_1 + ax_2 + d - y_2 = 0$$

or

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} + \begin{bmatrix} x_1 & -x_2 & 1 & 0 \\ x_2 & x_1 & 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} - \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = 0$$

This is a pair of condition equations that arises for each point of known coordinates in both the x and y systems. Obviously two such points are necessary and sufficient to compute the four parameters of the transformation (a, b, c, and d) without adjustment. Whenever more than two points are given, a least squares adjustment becomes necessary.

Suppose that we have m points that are known in both coordinate systems. It can easily be shown that (see the pre Example, 8.7), in case $\mathbf{W} = \mathbf{I}$,

$$\begin{bmatrix} \sum_1^m (x_1^2 + x_2^2) & 0 & \sum_1^m x_1 & \sum_1^m x_2 \\ & \sum_1^m (x_1^2 + x_2^2) & -\sum_1^m x_2 & \sum_1^m x_1 \\ & & m & 0 \\ \text{symmetric} & & & m \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$$

$$= \begin{bmatrix} \sum_1^m (x_1 y_1 + x_2 y_2) \\ \sum_1^m (x_1 y_2 - x_2 y_1) \\ \sum_1^m y_1 \\ \sum_1^m y_2 \end{bmatrix} \tag{8.17}$$

From these four normal equations the four parameters may be obtained. In the preceding treatment no attempt was made to specify the observations. As a matter of fact, we simply attached a residual to each equation in a manner similar to that used

in statistical regression problems. On analysis, however, we can see that each residual, v, is composed of two parts, one associated with a y coordinate and the other with the correspondingly "transformed" coordinate from the x system. We shall see subsequently how this type of adjustment compares to what we would normally do. But first let us note a computational simplification.

If in the x system, we would use the m points of the transformation and compute

$$\bar{x}_1 = \frac{\sum x_1}{m} \quad \bar{x}_2 = \frac{\sum x_2}{m}$$

and subtract $\bar{x}_1$ from x_1 coordinates, and $\bar{x}_2$ from x_2 coordinates, for *all* points in the x system, then the normal equations matrix above would reduce to a diagonal form. This is because in such a case

$$\sum_1^m x_1 = 0 \quad \text{and} \quad \sum_1^m x_2 = 0$$

Thus

$$\hat{a} = \frac{\sum_1^m (x_1 y_1 + x_2 y_2)}{\sum_1^m (x_1^2 + x_2^2)}$$

$$\hat{b} = \frac{\sum_1^m (x_1 y_2 - x_2 y_1)}{\sum_1^m (x_1^2 + x_2^2)} \tag{8.18}$$

$$\hat{c} = \frac{\sum_1^m y_1}{m}$$

$$\hat{d} = \frac{\sum_1^m y_2}{m}$$

noting that x_1 and x_2 are referred to the *centroid* as an origin. In this case also, the parameters would be uncorrelated and have the diagonal cofactor matrix as follows:

$$\mathop{Q_{\Delta\Delta}}_{4,4} = N^{-1} = \begin{bmatrix} [\sum_1^m (x_1^2 + x_2^2)]^{-1} & & & \\ & [\sum_1^m (x_1^2 + x_2^2)]^{-1} & & \\ & & 1/m & \\ & & & 1/m \end{bmatrix} \tag{8.19}$$

The covariance matrix of the parameters can be computed if the a posteriori estimate of the reference variance is computed. To do this first compute the residuals

$$v_1 = y_1 - \hat{a}x_1 + \hat{b}x_2 - \hat{c}$$

$$v_2 = y_2 - \hat{b}x_1 - \hat{a}x_2 - \hat{d}$$

Then

$$v'Wv = v'v = \sum_1^m [(y_1 - \hat{a}x_1 + \hat{b}x_2 - \hat{c})^2 + (y_2 - \hat{b}x_1 - \hat{a}x_2 - \hat{d})^2]$$

and the reference variance $\hat{\sigma}_0^2 = v'v/r$ where r = number of degrees of freedom $= 2m - 4$.

This gives the following covariance matrix for the parameters:

$$\Sigma_{\Delta\Delta}_{4,4} = \frac{1}{r} \sum_1^m \left[(y_1 - ax_1 + bx_2 - c)^2 + (y_2 - bx_1 - ax_2 - d)^2 \right]$$

$$\times \begin{bmatrix} [\sum_1^m (x_1^2 + x_2^2)]^{-1} & & & \\ & [\sum_1^m (x_1^2 + x_2^2)]^{-1} & & \\ & & 1/m & \\ & & & 1/m \end{bmatrix} \qquad (8.20)$$

The **y** coordinates of any point of given **x** coordinates may be computed from the transformation equations using the estimated parameters ($\hat{a}$, $\hat{b}$, $\hat{c}$, and $\hat{d}$). The corresponding covariance matrix of these coordinates may be computed by propagation techniques.

Returning now to the original transformation equations, rewrite them in functional form

$$f_1 = ax_1 - bx_2 + c - y_1 = 0$$
$$f_2 = bx_1 + ax_2 + d - y_2 = 0$$

As would normally be done we consider the general situation in which both the **x** and **y** coordinates are observations. Thus

$$\begin{bmatrix} a^0 & -b^0 & -1 & 0 \\ b^0 & a^0 & 0 & -1 \end{bmatrix} \begin{bmatrix} v_{x_1} \\ v_{x_2} \\ v_{y_1} \\ v_{y_2} \end{bmatrix} + \begin{bmatrix} x_1 & -x_2 & 1 & 0 \\ x_2 & x_1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \delta_a \\ \delta_b \\ \delta_c \\ \delta_d \end{bmatrix}$$

$$= \begin{bmatrix} y_1 - a^0 x_1 + b^0 x_2 - c^0 \\ y_2 - b^0 x_1 - a^0 x_2 - d^0 \end{bmatrix}$$

is the linearized form of the condition equations for one point. It should be noted that in the simplified case above the equations were directly linear, but when the regular adjustment treatment is pursued, the equations become nonlinear.

The first and simplest case to consider is when all coordinates (both **x** and **y**) are uncorrelated and of equal precision, or when $\mathbf{W} = \mathbf{Q} = \mathbf{I}$. If we choose the first approximations to be $a^0 = b^0 = c^0 = d^0 = 0$, the matrix $(\mathbf{AQA'})^{-1}$ reduces to the identity matrix and $\mathbf{f} = [y_1 \, y_2]^t$. In this situation the normal equations become identical to those given by equation (8.17). Taking, for the sake of simplicity of derivation, the case of referring the **x** coordinates to their centroid, the first solution becomes that given by equation (8.18); that is,

$$\delta a_1 = \frac{\sum (x_1 y_1 + x_2 y_2)}{\sum (x_1^2 + x_2^2)} = a_1^0$$

$$\delta b_1 = \frac{\sum (x_1 y_2 - x_2 y_1)}{\sum (x_1^2 + x_2^2)} = b_1^0$$

$$\delta c_1 = \frac{\sum y_1}{m} = c_1^0 \qquad\qquad (8.21)$$

$$\delta d_1 = \frac{\sum y_2}{m} = d_1^0$$

because the first approximations were all zero. The second iteration can be performed in the following manner:

$$(\mathbf{A}\mathbf{Q}\mathbf{A}')^{-1} = (1 + a_1^{02} + b_1^{02})^{-1}\mathbf{I} = \alpha\mathbf{I}$$

The scalar factor α is the *only change that will occur to the normal equations*. Next

$$\mathbf{t} = \mathbf{B}'(\mathbf{A}\mathbf{Q}\mathbf{A}')^{-1}\mathbf{f} = \alpha\mathbf{B}'\mathbf{f}$$

Each element of $\mathbf{t}$ may be derived as follows:

$$t_1 = \alpha \sum [x_1(y_1 - a_1^0 x_1 + b_1^0 x_2 - c_1^0)]$$
$$+ \alpha \sum [x_2(y_2 - b_1^0 x_1 - a_1^0 x_2 - d_1^0)]$$
$$= \alpha[\sum x_1 y_1 - a_1^0 \sum x_1^2 + b_1^0 \sum x_1 x_2 - c_1^0 \sum x_1]$$
$$+ \alpha[\sum x_2 y_2 - b_1^0 \sum x_1 x_2 - a_1^0 \sum x_2^2 - d_1^0 \sum x_2]$$

Since $\sum x_1 = 0$ and $\sum x_2 = 0$, and referring to equation (8.21), then

$$t_1 = \alpha[\sum (x_1 y_1 + x_2 y_2) - a_1^0 \sum (x_1^2 + x_2^2)] = 0$$
$$t_2 = \alpha \sum [-x_2(y_1 - a_1^0 x_1 + b_1^0 x_2 - c_1^0)]$$
$$+ \alpha \sum [x_1(y_2 - b_1^0 x_1 - a_1^0 x_2 - d_1^0)]$$
$$= \alpha[-\sum x_2 y_1 + a_1^0 \sum x_1 x_2 - b_1^0 \sum x_2^2 + c_1^0 \sum x_2]$$
$$+ \alpha[\sum x_1 y_2 - b_1^0 \sum x_1^2 - a_1^0 \sum x_1 x_2 - d_1^0 \sum x_1]$$
$$= \alpha[\sum (x_1 y_2 - x_2 y_1) - b_1^0 \sum (x_1^2 + x_2^2)] = 0$$
$$t_3 = \alpha \sum (y_1 - a_1^0 x_1 + b_1^0 x_2 - c_1^0)$$
$$= \alpha(\sum y_1 - a_1^0 \sum x_1 + b_1^0 \sum x_2 - mc_1^0) = \alpha(\sum y_1 - \sum y_1) = 0$$
$$t_4 = \alpha \sum [y_2 - b_1^0 x_1 - a_1^0 x_2 - d_1^0] = \alpha(\sum y_2 - md_1^0) = 0$$

Thus $\mathbf{t} = \mathbf{0}$ and $\Delta_2 = \mathbf{0}$, which means that the final estimates of the parameters are those given by equation (8.21). To complete this section, we must evaluate the updated normal equations in order to have the proper cofactor matrix, or

$$\mathbf{Q}_{\Delta\Delta} = \frac{1}{\alpha}\mathbf{N}^{-1} \tag{8.22}$$

where $\mathbf{N}^{-1}$ is that given by equation (8.19).

At first glance this may lead us to believe that the simplified approach, although giving the exact parameter estimates as the regular approach, does not yield the *same cofactor matrix* for the parameters as exhibited by the factor $1/\alpha$ in equation (8.22). However, since cofactor matrices are related to the covariance matrices by scale factors and since the only difference between equations (8.22) and (8.19) is a scale factor $(1/\alpha)$, it would be advisable to continue the derivation and obtain the covariance matrix of the parameters. Therefore the vector of residuals is

$$\mathbf{v} = \mathbf{Q}\mathbf{A}'(\mathbf{A}\mathbf{Q}\mathbf{A}')^{-1}(-\mathbf{B}\Delta + \mathbf{f})$$

but $\mathbf{Q} = \mathbf{I}$, $(\mathbf{A}\mathbf{Q}\mathbf{A}')^{-1} = \alpha\mathbf{I}$, and $\Delta = \mathbf{0}$; then

$$\mathbf{v} = \alpha\mathbf{A}'\mathbf{f}$$

The quadratic form is therefore, $\mathbf{v}^t\mathbf{W}\mathbf{v} = \mathbf{v}^t\mathbf{v} = \alpha^2\mathbf{f}^t\mathbf{A}\mathbf{A}^t\mathbf{f}$ but, again $\mathbf{A}\mathbf{A}^t = 1/\alpha\mathbf{I}$ so

$$\mathbf{v}^t\mathbf{v} = \alpha\mathbf{f}^t\mathbf{f}$$

or

$$\mathbf{v}^t\mathbf{W}\mathbf{v} = \alpha\sum_1^m [(y_1 - \hat{a}x_1 + \hat{b}x_2 - \hat{c})^2 + (y_2 - \hat{b}x_1 - \hat{a}x_2 - \hat{d})^2]$$

With r as the number of degrees of freedom, the covariance matrix of the parameters becomes identical to that in equation (8.20) indicating that the *estimate of the covariance matrix of the parameters* obtained from the simplified treatment given at the beginning of this example is also the correct one. Consequently, such a simplified use of adjustment is totally practical and rigorous as long as the *observations, all of them, have a unit cofactor matrix.* The question then arises as to whether it also applies in other situations of cofactor matrices that are different from the identity matrix. The first of these would be a diagonal matrix, with the x's having the same precision but which is different from that for the y's, or

$$\mathop{\mathbf{Q}}_{4,\,4} = \begin{bmatrix} q_x & & & \\ & q_x & & \\ & & q_y & \\ & & & q_y \end{bmatrix} \quad \text{for each point}$$

Thus again if $a^0 = b^0 = c^0 = d^0 = 0$, then

$$(\mathbf{A}\mathbf{Q}\mathbf{A}^t)^{-1} = \begin{bmatrix} q_y & 0 \\ 0 & q_y \end{bmatrix} = \mathop{\mathbf{W}}_{2,\,2}{}_{yy} = q_y^{-1}\mathbf{I}$$

The normal equations matrix

$$\mathbf{N} = \mathbf{B}^t\mathbf{W}_{yy}\,\mathbf{B} = q_y^{-1}\begin{bmatrix} \sum(x_1^2 + x_2^2) & & & \\ & \sum(x_1^2 + x_2^2) & & \\ & & m & \\ & & & m \end{bmatrix} \tag{8.23}$$

and

$$\mathbf{t} = \mathbf{B}^t\mathbf{W}_{yy}\,\mathbf{f} = q_y^{-1}\begin{bmatrix} \sum(x_1 y_1 + x_2 y_2) \\ \sum(x_1 y_2 - x_2 y_1) \\ \sum y_1 \\ \sum y_2 \end{bmatrix}$$

The solution of this system is identical to that given by equation (8.21). The second iteration must then be performed.

$$(\mathbf{A}\mathbf{Q}\mathbf{A}^t)^{-1} = \left\{ \begin{bmatrix} a^0 & -b^0 & -1 & 0 \\ b^0 & a^0 & 0 & -1 \end{bmatrix} \begin{bmatrix} q_x & & & \\ & q_x & & \\ & & q_y & \\ & & & q_y \end{bmatrix} \begin{bmatrix} a^0 & b^0 \\ -b^0 & a^0 \\ -1 & 0 \\ 0 & -1 \end{bmatrix} \right\}^{-1}$$

$$= \begin{bmatrix} (a^{0^2} + b^{0^2})q_x + q_y & 0 \\ 0 & (a^{0^2} + b^{0^2})q_x + q_y \end{bmatrix}^{-1} = \beta\mathbf{I}$$

where

$$\beta = [(a^{0^2} + b^{0^2})q_x + q_y]^{-1}$$

The normal equation matrix will change only by being multiplied by the scalar β. The constant term vector will however be

$$\mathbf{t} = \mathbf{B}^t \mathbf{W} \mathbf{f} = \beta \mathbf{B}^t \mathbf{f}$$

Since this is, to a scalar variation, the same as that computed previously, it follows that $\mathbf{t} = \mathbf{0}$ and the second correction vector is also zero. Thus the estimates of the parameters are identical to those given by equation (8.21). The cofactor matrix of the parameters can be shown to be

$$\mathbf{Q}_{\Delta\Delta} = q_y \beta^{-1} \mathbf{N}^{-1}$$

where $\mathbf{N}$ is computed from equation (8.23). Finally, in order to compute the covariance matrix $\boldsymbol{\Sigma}_{\Delta\Delta}$ then

$$\mathbf{v} = \mathbf{Q}\mathbf{A}^t (\mathbf{A}\mathbf{Q}\mathbf{A}^t)^{-1}(-\mathbf{B}\Delta + \mathbf{f}) = \beta \mathbf{Q}\mathbf{A}^t \mathbf{f}$$

and

$$\mathbf{v}^t \mathbf{W} \mathbf{v} = \beta^2 \mathbf{f}^t \mathbf{A}\mathbf{Q}\mathbf{W}\mathbf{Q}\mathbf{A}^t \mathbf{f} = \beta^2 \mathbf{f}^t (\mathbf{A}\mathbf{Q}\mathbf{A}^t)\mathbf{f} = \beta \mathbf{f}^t \mathbf{f}$$

$$= \beta \sum [(y_1 - \hat{a}x_1 + \hat{b}x_2 - \hat{c})^2 + (y_2 - \hat{b}x_1 - \hat{a}x_2 - \hat{d})^2]$$

The covariance matrix of the parameters $\boldsymbol{\Sigma}_{\Delta\Delta}$ is thus identical to that of equation (8.20).

Clearly the simplified adjustment is still rigorous and practical in the case in which $\mathbf{Q}_{xx}$ is different from $\mathbf{Q}_{yy}$ and each is a scalar matrix.

The next case would be no correlation, but each coordinate in each coordinate system has a different cofactor, and that each point has the same cofactor matrix, or,

$$\mathbf{Q} = \text{diag.} \{q_{x_1}, q_{x_2}, q_{y_1}, q_{y_2}\} \qquad \textit{for each point}$$

If the first set of approximations are again taken to be $a^0 = b^0 = c^0 = d^0 = 0$,

$$(\mathbf{A}\mathbf{Q}\mathbf{A}^t)^{-1} = \begin{bmatrix} q_{y_1} & 0 \\ 0 & q_{y_2} \end{bmatrix}^{-1} = \begin{bmatrix} w_{y_1} & 0 \\ 0 & w_{y_2} \end{bmatrix} = \mathbf{W}_{yy}$$

$$\mathbf{N} = \mathbf{B}^t \mathbf{W}_{yy} \mathbf{B} = \begin{bmatrix} \sum (w_{y_1} x_1^2 + w_{y_2} x_2^2) & & & \\ & \sum (w_{y_1} x_1^2 + w_{y_2} x_2^2) & & \\ & & mw_{y_1} & \\ & & & mw_{y_2} \end{bmatrix}$$

and

$$\mathbf{t} = \mathbf{B}^t \mathbf{W}_y \mathbf{f} = \begin{bmatrix} \sum (w_{y_1} x_1 y_1 + w_{y_2} x_2 y_2) \\ \sum (-w_{y_1} x_2 y_1 + w_{y_2} x_1 y_2) \\ \sum w_{y_1} y_1 \\ \sum w_{y_2} y_2 \end{bmatrix}$$

Then

$$a_1^0 = \frac{\sum (w_{y1} x_1 y_1 + w_{y2} x_2 y_2)}{\sum (w_{y1} x_1^2 + w_{y2} x_2^2)}$$

$$b_1^0 = \frac{\sum (-w_{y1} x_2 y_1 + w_{y2} x_1 y_2)}{\sum (w_{y1} x_1^2 + w_{y2} x_2^2)}$$

$$c_1^0 = \frac{\sum w_{y1} y_1}{m w_{y1}} = \frac{\sum y_1}{m}$$

$$d_1^0 = \frac{\sum w_{y2} y_2}{m w_{y2}} = \frac{\sum y_2}{m}$$

Second iteration

$$(AQA')^{-1} = \left\{ \begin{bmatrix} a_1^0 & -b_1^0 & -1 & 0 \\ b_1^0 & a_1^0 & 0 & -1 \end{bmatrix} \begin{bmatrix} q_{x1} & & & \\ & q_{x2} & & \\ & & q_{y1} & \\ & & & q_{y2} \end{bmatrix} \begin{bmatrix} a_1^0 & b_1^0 \\ -b_1^0 & a_1^0 \\ -1 & 0 \\ 0 & -1 \end{bmatrix} \right\}^{-1}$$

$$= \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{12} & \gamma_{22} \end{bmatrix} \tag{8.24}$$

From equation (8.24) it becomes obvious that the situation is no longer simple and that iterating the solution is necessary. Consequently, we can no longer use the simplified treatment and the solution must be carried out rigorously and iterated sufficiently. Of course, if we do not already have a program to handle the general case $(Av + B\Delta = f)$, or if the transformation is a part of a more elaborate program, a simpler computational approach may be effected using the relations above. After the iterations are terminated, the cofactor matrix of the parameters will be the inverse of the last normal equations' coefficient matrix.

Other cases that may be considered for four-parameter transformation would be those involving correlation. First, there would be cases in which Q_{xx} and Q_{yy} are both full matrices, but $Q_{xy} = 0$, meaning no correlation between the x coordinates and y coordinates. Finally, there would be the case in which Q is, in principle, a full matrix. All such cases become rather too involved to analyze by algebraic derivations as was done above. Instead, it would be by far better to perform the adjustment by rigorous and direct means using conditions of the form $Av + B\Delta = f$.

Before leaving the subject of four-parameter transformation, we should point out one of the common cases that often arises in practice. It is the case when the y coordinates are of such a superior quality that they may be considered as constants in the adjustment. Although we can derive the appropriate relations from the start by applying the above assumption (something that is left to the reader as an exercise), the case can be easily

accommodated with what we have presented so far. Instead of strictly enforcing the y coordinates to be constants, we introduce them as observations, but with *very small cofactors* compared to the x coordinates. This has practically the same effect, and no individual treatment is therefore necessary. More elaboration on this aspect will be given under the unified approach to adjustment in Chapter 12, Part III.

Example 8.9. *Six-Parameter Transformation.* The basic equations are

$$y_1 = ax_1 + bx_2 + c$$

$$y_2 = dx_1 + ex_2 + f$$

One of the simplest cases is that in which *only* the y coordinates are considered as the observations, with unit cofactor matrices. The adjustment problem is then linear and the condition equations for one point are

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} + \begin{bmatrix} -x_1 & -x_2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -x_1 & -x_2 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix} = \begin{bmatrix} -y_1 \\ -y_2 \end{bmatrix} \quad (8.25)$$

Assuming that the x coordinates are shifted to their centroid as origin the normal equations take the form

$$\begin{bmatrix} \sum x_1^2 & \sum x_1 x_2 & 0 & 0 & 0 & 0 \\ & \sum x_2^2 & 0 & 0 & 0 & 0 \\ & & m & 0 & 0 & 0 \\ & & & \sum x_1^2 & \sum x_1 x_2 & 0 \\ & & & & \sum x_2^2 & 0 \\ \text{symmetric} & & & & & m \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix} = \begin{bmatrix} \sum x_1 y_1 \\ \sum x_2 y_1 \\ \sum y_1 \\ \sum x_1 y_2 \\ \sum x_2 y_2 \\ \sum y_2 \end{bmatrix} \quad (8.26)$$

An interesting observation can be made regarding equation (8.26), which is that the normal equations' coefficient matrix can be divided into two *distinct and identical* 3 × 3 matrices. The reason is that each of the two condition equations contains three parameters that are completely different from the three appearing in the other equation. This leads to an economical way of solving the normal equations, as equation (8.26) may be rewritten as

$$\begin{bmatrix} \sum x_1^2 & \sum x_1 x_2 & 0 \\ & \sum x_2^2 & 0 \\ \text{symmetric} & & m \end{bmatrix} \begin{bmatrix} a & d \\ b & e \\ c & f \end{bmatrix} = \begin{bmatrix} \sum x_1 y_1 & \sum x_1 y_2 \\ \sum x_2 y_1 & \sum x_2 y_2 \\ \sum y_1 & \sum y_2 \end{bmatrix} \quad (8.27)$$

The form of equation (8.27) requires the inversion of only a 3 × 3 matrix, and the inverse when multiplied by the right-hand side yields all six parameters in the form of

two columns of three elements each. Such a scheme can be used under the same assumptions of observations and cofactor matrices to treat transformation equations of many parameters as long as such parameters are distinctly divided between the two condition equations. A good example of this is given by equation (8.9) where 12 or more parameters may be involved. Another example, which requires the extension of the technique to three equal groups of parameters, is found in the adjustment of transformation equations given by equation (8.12).

Returning to a more general case in which the x coordinates are also considered as observations, the problem becomes nonlinear. The linearized condition equations become

$$
\begin{bmatrix} a^0 & b^0 & -1 & 0 \\ d^0 & e^0 & 0 & -1 \end{bmatrix}
\begin{bmatrix} v_{x_1} \\ v_{x_2} \\ v_{y_1} \\ v_{y_2} \end{bmatrix}
+
\begin{bmatrix} x_1 & x_2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_1 & x_2 & 1 \end{bmatrix}
\begin{bmatrix} \delta_a \\ \delta_b \\ \delta_c \\ \delta_d \\ \delta_e \\ \delta_f \end{bmatrix}
$$

$$
= \begin{bmatrix} y_1 - a^0 x_1 - b^0 x_2 - c^0 \\ y_2 - d^0 x_1 - e^0 x_2 - f^0 \end{bmatrix} \qquad (8.28)
$$

Even under the assumption of unit cofactor matrices, the matrix

$$
(\mathbf{AQA'})^{-1} = \begin{bmatrix} (1 + a^{02} + b^{02}) & (a^0 d^0 + b^0 e^0) \\ (a^0 d^0 + b^0 e^0) & (1 + d^{02} + e^{02}) \end{bmatrix}^{-1}
$$

does not reduce to a simple enough form. Consequently, even though the first iteration solution is identical to that obtained above when $a^0 = b^0 = c^0 = d^0 = e^0 = f^0 = 0$, the second and subsequent iterations must be performed until the solution converges. A less than rigorous solution would be obtained if equation (8.25) were used instead of equation (8.28). Of course this deviation from rigor increases as the cofactor matrix of the observations increases in complexity from an identity matrix to a full matrix.

Example 8.10. *Seven-Parameter Transformation.* A seven-parameter transformation is a conformal transformation between two 3-dimensional coordinate systems allowing for a uniform scale change, an orthogonal rotation matrix that rotates the space solidly, and a vector of three mutually orthogonal shifts. Therefore it preserves shape (particularly angles) and size to the extent of the uniform scale change. It is normally written as

$$
\underset{3, 1}{\mathbf{y}} = s \underset{3, 3}{\mathbf{M}} \underset{3, 1}{\mathbf{x}} + \underset{3, 1}{\mathbf{k}}
$$

which was given by equation (8.11) and explained thereafter. It is important to recall that $\mathbf{M}$ is an orthogonal matrix or that $\mathbf{MM'} = \mathbf{I}$. Rewriting that equation for one point in a functional form

$$
s\mathbf{Mx} + \mathbf{k} - \mathbf{y} = 0 \qquad (8.29)
$$

and considering, for the sake of generality, both x and y as observations, we linearize equation (8.29) as $\mathbf{Av} + \mathbf{B\Delta} = \mathbf{f}$ where

$$\underset{3,6}{\mathbf{A}} = \begin{bmatrix} \underset{3,3}{s^0\mathbf{M}^0} & \underset{3,3}{-\mathbf{I}} \end{bmatrix}$$

$$\underset{6,1}{\mathbf{v}} = \begin{bmatrix} \underset{3,1}{\mathbf{v}_x} \\ \underset{3,1}{\mathbf{v}_y} \end{bmatrix} \qquad \underset{3,7}{\mathbf{B}} = \begin{bmatrix} \underset{3,1}{\mathbf{M}^0\mathbf{x}} & \underset{3,1}{s^0\mathbf{M}_1^0\mathbf{x}} & \underset{3,1}{s^0\mathbf{M}_2^0\mathbf{x}} & \underset{3,1}{s^0\mathbf{M}_3^0\mathbf{x}} & \underset{3,3}{\mathbf{I}} \end{bmatrix}$$

in which $\mathbf{M}_1$, $\mathbf{M}_2$, and $\mathbf{M}_3$ are the partial derivatives of $\mathbf{M}$ with respect to each of the three independent parameters, respectively.

$$\underset{7,1}{\mathbf{\Delta}} = [\delta s \quad \delta_1 \quad \delta_2 \quad \delta_3 \quad \delta k_1 \quad \delta k_2 \quad \delta k_3]^t$$

in which δ_1, δ_2, and δ_3 designate corrections to approximations of the three parameters of $\mathbf{M}$, respectively. Consider first the simplest case when $\mathbf{Q} = \mathbf{W} = \mathbf{I}$.

Examining the term $\mathbf{Q}_e = \mathbf{AQA}^t$ gives

$$\underset{3,3}{\mathbf{Q}_e} = \mathbf{AQA}^t = \mathbf{AA}^t = [s^0\mathbf{M}^{0t} - \mathbf{I}] \begin{bmatrix} s^0\mathbf{M}^{0t} \\ -\mathbf{I} \end{bmatrix} = s^{02}\mathbf{I} + \mathbf{I}$$

or

$$\mathbf{Q}_e = (1 + s^{02})\mathbf{I} \qquad \text{and} \qquad \mathbf{W}_e = (1 + s^{02})^{-1}\mathbf{I} = \alpha\mathbf{I}$$

The contribution from one point to the normal equations' matrix and constant vector becomes

$$\mathbf{N}_i = \mathbf{B}_i^t \mathbf{W}_{e_i} \mathbf{B}_i = \alpha\mathbf{B}_i^t \mathbf{B}_i \qquad \text{and} \qquad \mathbf{t}_i = \alpha\mathbf{B}_i^t \mathbf{f}_i$$

from which

$$\mathbf{\Delta} = \left[\sum_i (\mathbf{B}^t\mathbf{B})_i \right]^{-1} \left[\sum_i (\mathbf{B}^t\mathbf{f}_i) \right]$$

which is the solution that would be obtained if we simply wrote the equations in the form $\mathbf{v} + \mathbf{B\Delta} = \mathbf{f}$. In both cases, however, iteration is a necessary process because the equations are nonlinear in the parameters. The scalar α, although having no effect on the computation of $\mathbf{\Delta}$, should be evaluated in order to compute the cofactor matrix of the parameters. It depends only on the scale factor s and can therefore be computed only once after the iterative procedure is ended. Thus

$$\mathbf{Q}_{\Delta\Delta} = \alpha^{-1} \left[\sum_i (\mathbf{B}_i^t \mathbf{B}_i) \right]^{-1} = (1 + s^{02}) \left[\sum_i (\mathbf{B}_i^t \mathbf{B}_i) \right]^{-1}$$

In order to compute the corresponding covariance matrix, we need to evaluate the a posteriori value of the reference variance.

$$\mathbf{v} = \mathbf{QA}^t(\mathbf{AQA}^t)^{-1}(-\mathbf{B\Delta} + \mathbf{f}) = \alpha\mathbf{A}^t(-\mathbf{B\Delta} + \mathbf{f})$$

If $\mathbf{\Delta}$ is practically zero at the end of iterations, then $\mathbf{v} = \alpha\mathbf{A}^t\mathbf{f}$ and

$$\mathbf{v}^t\mathbf{W}\mathbf{v} = \alpha^2\mathbf{f}^t\mathbf{AA}^t\mathbf{f} = \alpha\mathbf{f}^t\mathbf{f}, \qquad \hat{\sigma}_0^2 = \alpha\mathbf{f}^t\mathbf{f}/r$$

and

$$\Sigma_{\Delta\Delta} = \hat{\sigma}_0^2 \mathbf{Q}_{\Delta\Delta} = \alpha\alpha^{-1}\frac{\mathbf{f'f}}{r}\left|\sum_i (\mathbf{B'B})_i\right|^{-1}$$

$$\Sigma_{\Delta\Delta} = \frac{1}{r}\mathbf{f'f}\left|\sum_i (\mathbf{B'B})_i\right|^{-1}$$

which is an interesting result because it is independent of α and would be directly obtainable from the case of $\mathbf{v} + \mathbf{B}\Delta = \mathbf{f}$.

The second case to discuss is when $\mathbf{Q}_{xx} = a\mathbf{I}$ and $\mathbf{Q}_{yy} = b\mathbf{I}$ for each point used in the transformation. Again, first evaluate the term $\mathbf{Q}_e$ for each point

$$\mathbf{Q}_e = \mathbf{AQA'} = [s^0\mathbf{M}^0 - \mathbf{I}]\begin{bmatrix} a\mathbf{I} & 0 \\ 0 & b\mathbf{I} \end{bmatrix}\begin{bmatrix} s^0\mathbf{M}^{0\prime} \\ -\mathbf{I} \end{bmatrix}$$

$$= as^{0^2}\mathbf{I} + b\mathbf{I} = (as^{0^2} + b)\mathbf{I}$$

$$\mathbf{W}_e = (as^{0^2} + b)^{-1}\mathbf{I} = \beta\mathbf{I}$$

The contribution of one point to the normal equation is

$$\mathbf{N}_i = \beta(\mathbf{B'B})_i \qquad \text{and} \qquad \mathbf{t}_i = \beta(\mathbf{B'f})_i$$

and thus

$$\Delta = \left|\sum_i (\mathbf{B'B})_i\right|^{-1}\left|\sum_i (\mathbf{B'f})_i\right|$$

which is the same as the simplified solution from $\mathbf{v} + \mathbf{B}\Delta = \mathbf{f}$. Next

$$\mathbf{Q}_{\Delta\Delta} = \mathbf{N}^{-1} = \beta^{-1}\left|\sum_i (\mathbf{B'B})_i\right|^{-1}$$

and

$$\mathbf{v} = \mathbf{QA'W}_e(-\mathbf{B}\Delta + \mathbf{f}) = \beta\mathbf{QA'f}$$

$$\mathbf{v'Wv} = \beta^2\mathbf{f'AQWQA'f} = \beta\mathbf{f'f}$$

$$\hat{\sigma}_0^2 = \frac{1}{r}\beta\mathbf{f'f}$$

$$\Sigma_{\Delta\Delta} = \hat{\sigma}_0^2 \mathbf{Q}_{\Delta\Delta} = \frac{1}{r}\mathbf{f'f}\left|\sum_i (\mathbf{B'B})_i\right|^{-1}$$

Another case would be when

$$\mathbf{Q} = \text{diag.}\{q_{x_1}, q_{x_2}, q_{x_3}, q_{y_1}, q_{y_2}, q_{y_3}\} = \begin{bmatrix} \mathbf{Q}_{xx} & 0 \\ 0 & \mathbf{Q}_{yy} \end{bmatrix}$$

$$\mathbf{Q}_e = \mathbf{AQA'} = [s^0\mathbf{M}^0 \quad \mathbf{I}]\begin{bmatrix} \mathbf{Q}_{xx} & 0 \\ 0 & \mathbf{Q}_{yy} \end{bmatrix}\begin{bmatrix} s^0\mathbf{M}^{0\prime} \\ -\mathbf{I} \end{bmatrix}$$

$$= s^{0^2}\mathbf{M}^0\mathbf{Q}_{xx}\mathbf{M}^{0\prime} + \mathbf{Q}_{yy}$$

which is a matrix that is not necessarily diagonal.

Consequently, from this case on (that is, with more general cofactor matrices), we should not bother with seeking simpler approaches and it would be advisable to apply directly the general case of $\mathbf{A}\mathbf{v} + \mathbf{B}\Delta = \mathbf{f}$.

For those readers who might want to program this transformation, a little more detail is given about the orthogonal matrix $\mathbf{M}$. One way of constructing it is in the form of three sequential rotation matrices, or $\mathbf{M} = \mathbf{M}_\kappa \mathbf{M}_\phi \mathbf{M}_\omega$ where

$$\mathbf{M}_\omega = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\omega & \sin\omega \\ 0 & -\sin\omega & \cos\omega \end{bmatrix} \qquad \text{rotation about } x_1 \text{ axis}$$

$$\mathbf{M}_\phi = \begin{bmatrix} \cos\phi & 0 & -\sin\phi \\ 0 & 1 & 0 \\ +\sin\phi & 0 & \cos\phi \end{bmatrix} \qquad \begin{array}{l}\text{rotation about once-} \\ \text{rotated } x_2 \text{ axis}\end{array}$$

$$\mathbf{M}_\kappa = \begin{bmatrix} \cos\kappa & +\sin\kappa & 0 \\ -\sin\kappa & \cos\kappa & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{array}{l}\text{rotation about twice-rotated } x_3 \\ \text{axis}\end{array}$$

The partial derivatives of $\mathbf{M}$ with respect to ω, ϕ, and κ, respectively, can be shown to be

$$\mathbf{M}_1 = \frac{\partial \mathbf{M}}{\partial \omega} = \mathbf{M}_\kappa \mathbf{M}_\phi \frac{\partial \mathbf{M}_\omega}{\partial \omega} = \mathbf{M} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & +1 \\ 0 & -1 & 0 \end{bmatrix}$$

$$\mathbf{M}_2 = \frac{\partial \mathbf{M}}{\partial \phi} = \mathbf{M}_\kappa \frac{\partial \mathbf{M}_\phi}{\partial \phi} \mathbf{M}_\omega = \mathbf{M} \begin{bmatrix} 0 & \sin\omega & \cos\omega \\ -\sin\omega & 0 & 0 \\ \cos\omega & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & -\cos\kappa \\ 0 & 0 & \sin\kappa \\ +\cos\kappa & -\sin\kappa & 0 \end{bmatrix} \mathbf{M}$$

$$\mathbf{M}_3 = \frac{\partial \mathbf{M}}{\partial \kappa} = \frac{\partial \mathbf{M}_\kappa}{\partial \kappa} \mathbf{M}_\phi \mathbf{M}_\omega = \begin{bmatrix} 0 & +1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{M}$$

9

Least Squares Adjustment with Conditions and Constraints

(Adjustment with Functionally Dependent Parameters)

9.1. INTRODUCTION

A constraint equation has been defined as that which relates only parameters to each other. Its presence in the functional model implies that the parameters are functionally dependent.

Parameter constraints occur in practice when some or all of the parameters in the adjustment must conform to some relationships arising from either geometric or physical characteristics of the model. For example, in photogrammetry, points on a lake shore can be constrained to have the same elevation, which may or may not be known; points chosen on a straight portion of a railroad are constrained to lie on a straight line; two points in the photogrammetric model may be constrained to known distance; camera stations from an orbital vehicle may be constrained to an orbit; and so on. The functional dependence of the parameters leads to having as many dependent parameters as there are constraint equations. This can be explained in the following manner.

First, from n observations and n_0 minimum number of model variables, the redundancy r is determined for which r conditions may be formulated

among the observations. If we now wish to include u' *functionally indepen-dent* parameters, the number of conditions increases accordingly to $(r + u')$ with a maximum number of n. Progressing further we would like to accept the possibility of *functionally dependent* parameters in the model. Thus if $(u' + 1)$ parameters are considered, then one of them will be dependent on the u' independent parameters. For the parameter, one *constraint equation* must be written to reflect that dependency. In general, if as usual there are u parameters $(u > u')$, some of which are functionally dependent, then s con-straints must be written to account for (and be equal to) the number of dependent parameters, or $s = u - u'$. Evidently s must always be less than u, or the constraints become an inconsistent set of equations,

$$s < u \qquad (9.1)$$

Since the s constraints are equal to the number of dependent parameters, it follows from equation (6.1) that

$$c + s = r + u \qquad (9.2)$$

or the number of conditions plus constraints must be equal to the redun-dancy plus the total number of (dependent) parameters.

The number of conditions still must not be larger than the number of observations as given by equation (6.2). However, the limitations placed on the number of parameters in equation (6.3) when the parameters are func-tionally independent do not hold here. In other words, there is essentially no upper limit on the number of parameters in this case, as u can readily exceed n_0, *provided* the proper number of constraints is formulated according to equation (9.2).

9.2. GENERAL CASE FOR ADJUSTMENT WITH CONDITIONS AND CONSTRAINTS

The two sets of equations, conditions and constraints, are

$$\underset{c,\,n}{\mathbf{A}} \underset{n,\,1}{\mathbf{v}} + \underset{c,\,u}{\mathbf{B}} \underset{u,\,1}{\mathbf{\Delta}} = \underset{c,\,1}{\mathbf{f}} \qquad (9.3a)$$

$$\underset{s,\,u}{\mathbf{C}} \underset{u,\,1}{\mathbf{\Delta}} = \underset{s,\,1}{\mathbf{g}} \qquad (9.3b)$$

with

$$\text{rank } (\mathbf{A}) = c \qquad \text{rank } (\mathbf{B}) = u \qquad \text{rank } (\mathbf{C}) = s \qquad (9.3c)$$

In the presence of constraints it is, in general, possible that $u > c$ leading to rank of $\mathbf{B}$ being c. In such a situation, part of the development to follow (solution by partitioning) would not work. Here $u < c$ is selected and the other possibility (for $u > c$) is addressed later as a special case. Two different procedures are possible and each is presented in a separate section.

9.2.1. Direct Treatment Since there are two separate sets of equations, then two sets or vectors of Lagrange multipliers

$$\underset{n,\,1}{\mathbf{k}} \quad \text{and} \quad \underset{s,\,1}{\mathbf{k}_c}$$

are needed in order to use the "constrained minima" technique. The symbol $\mathbf{k}$ without subscript is still retained to mean that for the conditions. Consequently, the quadratic form to be minimized is

$$\phi = \mathbf{v}^t\mathbf{W}\mathbf{v} - 2\mathbf{k}^t(\mathbf{A}\mathbf{v} + \mathbf{B}\Delta - \mathbf{f}) - 2\mathbf{k}_c^t(\mathbf{C}\Delta - \mathbf{g})$$

$$= \text{minimum} \tag{9.4}$$

Equation (9.4) contains two free variable vectors, $\mathbf{v}$ and Δ, thus

$$\frac{\partial\phi}{\partial\mathbf{v}} = 2\mathbf{v}^t\mathbf{W} - 2\mathbf{k}^t\mathbf{A} = \mathbf{0}^t$$

$$\frac{\partial\phi}{\partial\Delta} = 2\mathbf{k}^t\mathbf{B} - 2\mathbf{k}_c^t\,\mathbf{C} = \mathbf{0}^t$$

or

$$-\mathbf{W}\mathbf{v} + \mathbf{A}^t\mathbf{k} = \mathbf{0} \tag{9.5}$$

$$\mathbf{B}^t\mathbf{k} + \mathbf{C}^t\mathbf{k}_c = \mathbf{0} \tag{9.6}$$

The relations (9.5) and (9.6) represent a set of $(n + u)$ linear equations which when combined with equations (9.3a) and (9.3b) give a total of $(n + c + u + s)$ equations in as many unknowns. The unknown vectors are

$$\underset{n,\,1}{\mathbf{v}}, \quad \underset{c,\,1}{\mathbf{k}}, \quad \underset{u,\,1}{\Delta}, \quad \text{and} \quad \underset{s,\,1}{\mathbf{k}_c}$$

The four sets of equations combined constitute the total system of normal equations. The least squares problem is resolved by solving that system, or

$$\begin{bmatrix} \mathbf{v} \\ \mathbf{k} \\ \Delta \\ \mathbf{k}_c \end{bmatrix} = \begin{bmatrix} -\mathbf{W} & \mathbf{A}^t & 0 & 0 \\ \mathbf{A} & 0 & \mathbf{B} & 0 \\ 0 & \mathbf{B}^t & 0 & \mathbf{C}^t \\ 0 & 0 & \mathbf{C} & 0 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ \mathbf{f} \\ 0 \\ \mathbf{g} \end{bmatrix} \tag{9.7}$$

This matrix has a rank of $(n + c + u + s)$ and is therefore nonsingular (assuming of course that the model is properly constructed). Difficulties in inverting the coefficient matrix would be experienced only if the mathematical model is improperly constructed. Therefore the functional model must be correctly established and the proper conditional and constraint equations written. (Of course unusual cases may exist in which the coefficient matrix is ill conditioned.)

In general, the inversion in equation (9.7) could involve quite a large matrix. Of the four vectors computed from that equation only Δ, and perhaps v, may be of direct interest. Therefore, for many problems a solution by partitioning is more appropriate.

Thus from equations (9.5), (9.3a), and (9.6), we get, respectively,

$$v = W^{-1}A^t k = QA^t k \tag{9.8}$$

$$k = (AQA^t)^{-1}(-B\Delta + f)$$

$$= W_e(-B\Delta + f) \tag{9.9}$$

$$-(B^t W_e B)\Delta + B^t W_e f + C^t k_c = 0$$

or

$$-N\Delta + C^t k_c = -t \tag{9.10}$$

Solving for Δ and k_c from equations (9.10) and (9.3b) gives

$$\Delta = N^{-1}(t + C^t k_c) = N^{-1}t + N^{-1}C^t k_c = \Delta^0 + \delta\Delta \tag{9.11}$$

$$k_c = (CN^{-1}C^t)^{-1}(g - CN^{-1}t) = M^{-1}(g - C\Delta^0) \tag{9.12}$$

where

$$M = (CN^{-1}C^t), \qquad \Delta^0 = N^{-1}t, \qquad \delta\Delta = N^{-1}C^t k_c \tag{9.13}$$

The inverse of M exists because its rank is equal to its order s.

The vector Δ^0 represents the solution for the parameter vector in the absence of constraints. The final parameter vector is therefore

$$\Delta = \Delta^0 + N^{-1}C^t(CN^{-1}C^t)^{-1}(g - C\Delta^0)$$

$$= \Delta^0 + N^{-1}C^t M^{-1}(g - C\Delta^0) \tag{9.14}$$

which is the u partially *reduced normal equations*, assuming that Δ is the unknown vector of primary importance.

Several observations can be made regarding the treatment of constraint equations. First, equation (9.14) shows that the parameter vector Δ can be computed in two parts: Δ^0 as the least squares solution of the problem as if the constraints did not exist; then, with matrices C and g from the constraint equations $\delta\Delta$ can be computed. Once Δ is computed, the rest of the problem proceeds in a manner identical to that covered previously. For example, using Δ, k may be computed from equation (9.9) and substituted into equation (9.8) to compute v. Adding v to l gives $\hat{l}$. The a posteriori estimate for the reference variance, $\hat{\sigma}_0^2$ may be computed from equation (6.25).

9.2.2. Precision Estimation for the Direct Treatment
Applying the propagation rule, it can be shown that

$$Q_{\Delta\Delta} = N^{-1}(I - C^t M^{-1}CN^{-1}) \tag{9.15a}$$

or

$$Q_{\Delta\Delta} = N^{-1}[I - C'(CN^{-1}C')^{-1}CN^{-1}] \qquad (9.15b)$$

An interesting equality exists on the basis of equation (9.15b), if we post-multiply both its sides by C', thus

$$Q_{\Delta\Delta} C' = N^{-1}C' - N^{-1}C'(CN^{-1}C')^{-1}CN^{-1}C'$$

$$= N^{-1}C' - N^{-1}C'$$

or

$$Q_{\Delta\Delta} C' = 0 \qquad (9.16)$$

which may be used for checking purposes. Similarly,

$$Q_{vv} = QA'W_e AQ - QA'W_e B(N^{-1} - N^{-1}C'M^{-1}CN^{-1})B'W_e AQ \qquad (9.17a)$$

which, from equation (9.15a) becomes

$$Q_{vv} = QA'W_e AQ - QA'W_e BQ_{\Delta\Delta} B'W_e AQ \qquad (9.17b)$$

The reader can ascertain that equation (9.17b) is a general relationship since it is identical to equation (6.30). From equation (9.17b) an interesting equality is possible if both sides are premultiplied by A and postmultiplied by A', thus

$$AQ_{vv} A' = (AQA')W_e(AQA') - (AQA')W_e BQ_{\Delta\Delta} B'W_e(AQA')$$

or

$$AQ_{vv} A' = Q_e - BQ_{\Delta\Delta} B' \qquad (9.18)$$

Finally, the cofactor matrix for the estimated observations is given by the same standard relation or

$$Q_{\hat{l}\hat{l}} = Q - Q_{vv} \qquad (6.31)$$

9.2.3. Solution by Elimination of Constraints   The constraint equations (9.3b) may be used to solve, functionally, for as many parameters as there are constraints. These parameters are then substituted in turn into the first set of condition equations, thus reducing the number of parameters. The resulting condition equations can be utilized in a direct least squares solution for the remaining parameters. Such an approach has two advantages: First, it avoids the direct use of the constraints into the least squares estimation, and, second, it leads to a reduction of the number of parameters to be simultaneously estimated. However, the process of eliminating part of the parameters using the constraints needs care since the constraint equations may not, and often do not, include all the parameters. This may make it difficult to standardize the elimination procedure in a computer program. In rather com-

plex cases, prior evaluation of the equations may be necessary to effect a rearrangement of their order so that they may be used in a prewritten program.

Partitioning the parameter vector $\boldsymbol{\Delta}$ into two subvectors

$$\underset{s,\,1}{\boldsymbol{\Delta}_1} \quad \text{and} \quad \underset{a,\,1}{\boldsymbol{\Delta}_2}$$

where $a = u - s$, gives

$$\underset{c,\,n}{\mathbf{A}} \; \underset{n,\,1}{\mathbf{v}} + \underset{c,\,s}{\mathbf{B}_1} \; \underset{s,\,1}{\boldsymbol{\Delta}_1} + \underset{c,\,a}{\mathbf{B}_2} \; \underset{a,\,1}{\boldsymbol{\Delta}_2} = \underset{c,\,1}{\mathbf{f}} \tag{9.19a}$$

$$\underset{s,\,s}{\mathbf{C}_{11}} \; \underset{s,\,1}{\boldsymbol{\Delta}_1} + \underset{s,\,a}{\mathbf{C}_{12}} \; \underset{a,\,1'}{\boldsymbol{\Delta}_2} = \underset{s,\,1}{\mathbf{g}} \tag{9.19b}$$

Using equation (9.19b), we can solve for $\boldsymbol{\Delta}_1$, in terms of $\boldsymbol{\Delta}_2$,

$$\boldsymbol{\Delta}_1 = \mathbf{C}_{11}^{-1}(\mathbf{g} - \mathbf{C}_{12}\,\boldsymbol{\Delta}_2) \tag{9.20}$$

Care must be exercised in the partitioning operation to render $\mathbf{C}_{11}$ nonsingular. If the constraint equations as they are generated lead to a singular matrix, they must be rearranged to avoid that difficulty. Thus

$$\mathbf{A}\mathbf{v} + (-\mathbf{B}_1\mathbf{C}_{11}^{-1}\mathbf{C}_{12} + \mathbf{B}_2)\boldsymbol{\Delta}_2 = (\mathbf{f} - \mathbf{B}_1\mathbf{C}_{11}^{-1}\mathbf{g})$$

and with the auxiliaries

$$\bar{\mathbf{B}} = (-\mathbf{B}_1\mathbf{C}_{11}^{-1}\mathbf{C}_{12} + \mathbf{B}_2) \tag{9.21a}$$

$$\bar{\mathbf{f}} = (\mathbf{f} - \mathbf{B}_1\mathbf{C}_{11}^{-1}\mathbf{g}) \tag{9.21b}$$

becomes

$$\mathbf{A}\mathbf{v} + \bar{\mathbf{B}}\boldsymbol{\Delta}_2 = \bar{\mathbf{f}} \tag{9.22}$$

Equation (9.22) represents a system of c condition equations including only $a = u - s$ parameters. In a straightforward manner, the solution for $\boldsymbol{\Delta}_2$ by least squares would be

$$\boldsymbol{\Delta}_2 = \bar{\mathbf{N}}^{-1}\bar{\mathbf{t}} \tag{9.23}$$

where

$$\underset{a,\,a}{\bar{\mathbf{N}}} = \bar{\mathbf{B}}^t \underset{a,\,c}{\left(\underset{c,\,c}{\mathbf{A}\mathbf{Q}\mathbf{A}^t}\right)^{-1}} \bar{\mathbf{B}} = \bar{\mathbf{B}}^t\mathbf{W}_e\,\bar{\mathbf{B}} \tag{9.24a}$$

$$\underset{a,\,1}{\bar{\mathbf{t}}} = \bar{\mathbf{B}}^t(\mathbf{A}\mathbf{Q}\mathbf{A}^t)^{-1}\bar{\mathbf{f}} = \bar{\mathbf{B}}^t\mathbf{W}_e\,\bar{\mathbf{f}} \tag{9.24b}$$

Once $\boldsymbol{\Delta}_2$ is computed, $\boldsymbol{\Delta}_1$ may be evaluated from equation (9.20) and the total vector of parameters would simply be

$$\boldsymbol{\Delta} = \begin{bmatrix} \boldsymbol{\Delta}_1 \\ \boldsymbol{\Delta}_2 \end{bmatrix} \tag{9.25}$$

With the value of Δ thus computed, the rest of the variables can be evaluated exactly as before.

9.2.4. Precision Estimation for the Solution by Constraint Elimination

From the propagation rule,

$$\mathbf{Q}_{\Delta_2\Delta_2} = \bar{\mathbf{N}}^{-1} \tag{9.26}$$

$$\mathbf{Q}_{\Delta_1\Delta_1} = \mathbf{C}_{11}^{-1}\mathbf{C}_{12}\bar{\mathbf{N}}^{-1}\mathbf{C}_{12}^{t}(\mathbf{C}_{11}^{-1})^{t} \tag{9.27}$$

$$\mathbf{Q}_{\Delta_1\Delta_2} = -\mathbf{C}_{11}^{-1}\mathbf{C}_{12}\,\bar{\mathbf{N}}_{1}^{-1} \tag{9.28}$$

Thus the total cofactor matrix for Δ is

$$\mathbf{Q}_{\Delta\Delta} = \begin{bmatrix} \mathbf{C}_{11}^{-1}\mathbf{C}_{12}\,\bar{\mathbf{N}}^{-1}\mathbf{C}_{12}^{t}(\mathbf{C}_{11}^{-1})^{t} & -\mathbf{C}_{11}^{-1}\mathbf{C}_{12}\,\bar{\mathbf{N}}^{-1} \\ -\bar{\mathbf{N}}^{-1}\mathbf{C}_{12}^{t}(\mathbf{C}_{11}^{-1})^{t} & \bar{\mathbf{N}}^{-1} \end{bmatrix} \tag{9.29}$$

Then

$$\mathbf{Q}_{vv} = \mathbf{Q}\mathbf{A}^{t}[\mathbf{W}_e - \mathbf{W}_e\,\mathbf{B}\mathbf{Q}_{\Delta\Delta}\,\mathbf{B}^{t}\mathbf{W}_e]\mathbf{A}\mathbf{Q} \tag{9.30}$$

in which $\mathbf{Q}_{\Delta\Delta}$ would be evaluated from equation (9.29). As before, the cofactor matrix of estimated observations $\mathbf{Q}_{ll}$ can be computed from equation (6.31).

Example 9.1. Figure 9.1 shows a simple problem of trilateration to determine the two coordinates of point R. Three ranges to three perfectly known points, P, O, and Q, are measured with the following results:

$$l_1 = \overline{RP} = 5.10 \text{ m}$$

$$l_2 = \overline{RO} = 7.07 \text{ m}$$

$$l_3 = \overline{RQ} = 4.90 \text{ m}$$

All three observations are of equal precision and uncorrelated. The coordinates of the fixed points are $P = (0, 5)$, $O = (0, 0)$, and $Q = (5, 0)$.
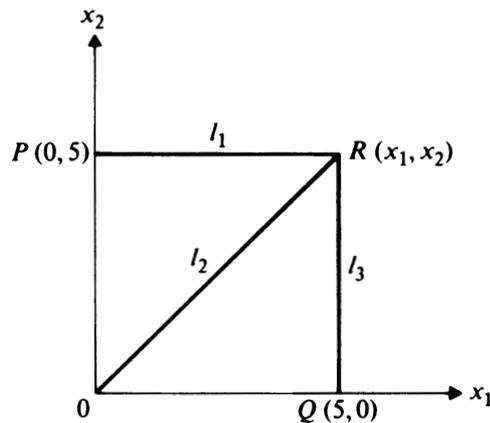

Figure 9.1

a. Assuming the approximate values for the unknown coordinates to be $x_1^0 = 5.0$ m and $x_2^0 = 5.0$ m, compute the correction vector from the first iteration.

b. If x_2 is to be constrained to the value 5.0 m, compute the corresponding value of x_1.

Solution: The model here is relatively simple, being composed of a plane with a two-dimensional cartesian coordinate system. There are four points in that plane involved in the problem, three of which are given as perfectly known. In order to determine the coordinates of the fourth point, two ranges will be the minimum required. (The point of intersection in the first quadrant will be selected as implied from Figure 9.1.) Since three ranges are given, there exists 1 degree of freedom, that is, $r = 1$. Given $u = 2$, the number of condition equations will be $c = 3$. Obviously a condition equation can be written for each given range, thus

$$l_1^2 - (x_1 - 0)^2 - (x_2 - 5)^2 = 0$$

$$l_2^2 - (x_1 - 0)^2 - (x_2 - 0)^2 = 0$$

$$l_3^2 - (x_1 - 5)^2 - (x_2 - 0)^2 = 0$$

If these equations are left as they are, an **A** matrix would have to be evaluated since they are nonlinear in the observations. However, with simple manipulation we can reduce **A** to the identity matrix and work with the simpler case of adjustment of indirect observations. (Of course, as a price for this, the condition equations may become slightly more complicated, thus requiring a little more work in linearization.)

$$l_1 - [x_1^2 + (x_2 - 5)^2]^{1/2} = 0$$

$$l_2 - (x_1^2 + x_2^2)^{1/2} = 0$$

$$l_3 - [(x_1 - 5)^2 + x_2^2]^{1/2} = 0$$

Linearization to the form $v + B\Delta = f$ leads to

$$\underset{3,2}{\mathbf{B}} = \begin{bmatrix} \dfrac{-x_1^0}{[x_1^{02} + (x_2^0 - 5)^2]^{1/2}} & \dfrac{-(x_2^0 - 5)}{[x_1^{02} + (x_2^0 - 5)^2]^{1/2}} \\[2ex] \dfrac{-x_1^0}{(x_1^{02} + x_2^{02})^{1/2}} & \dfrac{-x_2^0}{(x_1^{02} + x_2^{02})^{1/2}} \\[2ex] \dfrac{-(x_1^0 - 5)}{[(x_1^0 - 5)^2 + x_2^{02})]^{1/2}} & \dfrac{-x_2^0}{[(x_1^0 - 5)^2 + x_2^{02})]^{1/2}} \end{bmatrix} = \begin{bmatrix} -1 & 0 \\[2ex] \dfrac{-1}{\sqrt{2}} & \dfrac{-1}{\sqrt{2}} \\[2ex] 0 & -1 \end{bmatrix}$$

$$\underset{3,1}{\mathbf{f}} = \begin{bmatrix} [x_1^{02} + (x_2^0 - 5)^2]^{1/2} - l_1 \\[1.5ex] (x_1^{02} + x_2^{02})^{1/2} - l_2 \\[1.5ex] [(x_1^0 - 5)^2 + x_2^{02}]^{1/2} - l_3 \end{bmatrix} = \begin{bmatrix} -0.1 \\ 0 \\ 0.1 \end{bmatrix}$$

Given $\mathbf{W} = \mathbf{I}$ and the above values, then

$$\underset{2,2}{\mathbf{N}} = \begin{bmatrix} \frac{3}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{3}{2} \end{bmatrix}$$

$$\mathbf{t} = \begin{bmatrix} 0.1 \\ -0.1 \end{bmatrix} \qquad \Delta = \begin{bmatrix} 0.1 \\ -0.1 \end{bmatrix}$$

which is the first answer required.

Given the constraint, n_0 will reduce to 1 and the redundancy will increase to $r = 2$; thus $c = 3$, $u = 2$, and $s = 1$. First, using the *direct treatment* the constraint equation is simply $x_2 - 5 = 0$. This equation, although linear, must, however, be linearized. This is because the condition equations are nonlinear and have therefore to be linearized. Such linearization leads to estimating a correction vector Δ instead of the coordinates themselves. Consequently, x_2 in the constraint equation is replaced by $(x_2^0 + \delta x_2)$, where x_2^0 *must be the same approximate* value used in the condition equations. Thus with the given $x_2^0 = 5$ m, the constraint equation becomes

$$x_2^0 + \delta x_2 - 5 = 0 \qquad \text{or} \qquad \delta x_2 = 0$$

which, in the usual matrix from of $\mathbf{C\Delta} = \mathbf{g}$, becomes

$$[0 \quad 1] \begin{bmatrix} \delta x_1 \\ \delta x_2 \end{bmatrix} = 0$$

Using the $\mathbf{C}$ matrix and $\mathbf{N}^{-1}$ above leads to

$$\mathbf{M} = \tfrac{3}{4}$$

and

$$\delta \Delta = \begin{bmatrix} -0.033 \\ 0.1 \end{bmatrix}$$

Thus

$$\Delta = \Delta^0 + \delta\Delta = \begin{bmatrix} 0.1 \\ -0.1 \end{bmatrix} + \begin{bmatrix} -0.033 \\ 0.1 \end{bmatrix} = \begin{bmatrix} 0.067 \\ 0.0 \end{bmatrix}$$

which implies that there will be no correction to $x_2^0 = 5.00$ m, thus satisfying the constraint equation of fixing the value of x_2 at 5.00 m. In the second process, we eliminate x_2 from the original condition equations using the constraint equations. Thus with $x_2 = 5$, δx_2 must be zero (since $x_2^0 = 5$) and the original condition equations become

$$\mathbf{v} + \begin{bmatrix} -1 \\ \dfrac{-1}{\sqrt{2}} \\ 0 \end{bmatrix} (\delta x_1) = \begin{bmatrix} -0.1 \\ 0 \\ 0.1 \end{bmatrix} - \begin{bmatrix} 0 \\ \dfrac{-1}{\sqrt{2}} \\ -1 \end{bmatrix} (0) = \begin{bmatrix} -0.1 \\ 0 \\ 0.1 \end{bmatrix}$$

and

$$\mathbf{\bar{N}}_{1,1} = \tfrac{3}{2}, \qquad \mathbf{\bar{t}} = 0.1, \qquad \mathbf{x}_1 = 0.067$$

which is exactly the same value we obtained above.

Example 9.2

Given: The problem of Example 6.1 is expanded to allow for the use of constraints. Three terrestrial camera stations S_1, S_2, and S_3 photograph two points A and B (see Figure 9.2). All three camera axes are parallel and normal to the base which is taken as the X_1 axis. The X_2 axis coincides with the optical axis of camera
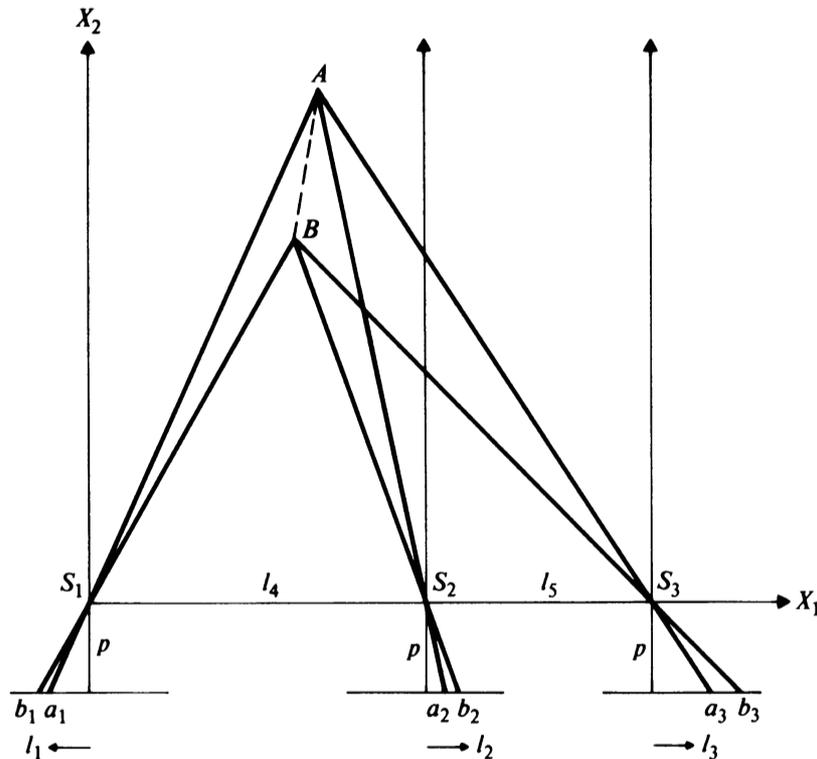
Figure 9.2. (Not to Scale)

S_1. The measured values of the two bases are $l_4 = 10$ m and $l_5 = 8$ m, each with a standard deviation, $\sigma = 0.05$ m. Denoting by l_1, l_2, and l_3 the observations on the three photographs S_1, S_2, and S_3, respectively, the observation are

POINT	l_1 mm	l_2 mm	l_3 mm
a	14.1	6.1	22.1
b	16.6	7.1	26.3

The standard deviation for all photo observations is 0.10 mm, and the principal distance is $p = 100$ mm (assumed without error, that is, a constant).

Required: Using the data given above compute the coordinates of both points A and B in an x_1 and x_2 coordinate system. Then constraining the distance between the two points to 7.8 m, compute the new set of coordinates, using both the direct treatment and the method of elimination.

Solution: To demonstrate first the use of the direct treatment, solve the first part by the method of combined observations and parameters. The number of condition equations is computed as follows:

+ Minimum number of variables to determine the geometry of the problem is six (x_1 for S_2 and S_3, and x_1 and x_2 for both points A and B), or $n_0 = 6$.
+ The total number of observed quantities is $n = 8$.
+ The redundancy is $r = 2$.
+ The number of parameters to be solved for is $u = 4$ (coordinates of A and B).
+ The total number of independent condition equations is $c = r + u = 6$.

Referring to Figure 9.2, write one condition equation for each photogrammetric ray:

$$l_{1a}x_{2a} - px_{1a} = 0$$

$$l_{1b}x_{2b} - px_{1b} = 0$$

$$l_{2a}x_{2a} - pl_4 + px_{1a} = 0$$

$$l_{2b}x_{2b} - pl_4 + px_{1b} = 0$$

$$l_{3a}x_{2a} - pl_4 - pl_5 + px_{1a} = 0$$

$$l_{3b}x_{2b} - pl_4 - pl_5 + px_{1b} = 0$$

Linearization:

$$\mathbf{Av} = \begin{bmatrix} x_{2a}^0 & 0 & 0 & & & & & 0 \\ 0 & x_{2b}^0 & 0 & & & & & 0 \\ & 0 & x_{2a}^0 & 0 & 0 & 0 & -p & 0 \\ \vdots & \vdots & 0 & x_{2b}^0 & 0 & 0 & -p & 0 \\ & & \vdots & 0 & x_{2a}^0 & 0 & -p & -p \\ 0 & 0 & 0 & 0 & 0 & x_{2b}^0 & -p & -p \end{bmatrix} \begin{bmatrix} v_{1a} \\ v_{1b} \\ v_{2a} \\ v_{2b} \\ v_{3a} \\ v_{3b} \\ v_4 \\ v_5 \end{bmatrix}$$

$$\mathbf{B\Delta} = \begin{bmatrix} -p & l_{1a} & 0 & 0 \\ 0 & 0 & -p & l_{1b} \\ p & l_{2a} & 0 & 0 \\ 0 & 0 & p & l_{2b} \\ p & l_{3a} & 0 & 0 \\ 0 & 0 & p & l_{3b} \end{bmatrix} \begin{bmatrix} \delta x_{1a} \\ \delta x_{2a} \\ \delta x_{1b} \\ \delta x_{2b} \end{bmatrix}$$

$$\mathbf{f}_{6,1} = \begin{bmatrix} px_{1a}^0 - l_{1a}x_{2a}^0 \\ px_{1b}^0 - l_{1b}x_{2b}^0 \\ pl_4 - px_{1a}^0 - l_{2a}x_{2a}^0 \\ pl_4 - px_{1b}^0 - l_{2b}x_{2b}^0 \\ pl_4 + pl_5 - px_{1a}^0 - l_{3a}x_{2a}^0 \\ pl_4 + pl_5 - px_{1b}^0 - l_{3b}x_{2b}^0 \end{bmatrix}$$

Using simple geometric relationships compute the following approximations:

$$x_{1a}^0 = 8.00 \text{ m}, \quad x_{2a}^0 = 51.00 \text{ m}, \quad x_{1b}^0 = 7.00 \text{ m}, \quad \text{and}$$

$$x_{2b}^0 = 41.00 \text{ m}$$

from which, with the observations, evaluate the numerical values of the matrices as

$$
\underset{6,\,8}{\mathbf{A}} =
\begin{bmatrix}
51.0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 41.0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 51.0 & 0 & 0 & 0 & -100.0 & 0 \\
0 & 0 & 0 & 41.0 & 0 & 0 & -100.0 & 0 \\
0 & 0 & 0 & 0 & 51.0 & 0 & -100.0 & -100.0 \\
0 & 0 & 0 & 0 & 0 & 41.0 & -100.0 & -100.0
\end{bmatrix}
$$

This matrix includes parameter approximations and therefore would have to be recomputed each iteration.

$$
\underset{6,\,4}{\mathbf{B}} =
\begin{bmatrix}
-100.0 & 14.1 & 0 & 0 \\
0 & 0 & -100.0 & 16.6 \\
100.0 & 6.1 & 0 & 0 \\
0 & 0 & 100.0 & 7.1 \\
100.0 & 22.1 & 0 & 0 \\
0 & 0 & 100.0 & 26.3
\end{bmatrix}
$$

This matrix contains no parameter approximations. It remains unaltered through the iterative process.

$$
\underset{6,\,1}{\mathbf{f}} = [80.90 \quad 19.40 \quad -111.10 \quad 8.90 \quad -127.10 \quad 21.70]^t
$$

This vector needs to be recomputed every iteration.

From the given standard deviations the cofactor matrix of the given observation (assuming a priori $\sigma_0^2 = 1$) is

$$
\underset{8,\,8}{\mathbf{Q}} = \text{diag.} \left\{ \underbrace{0.01, 0.01, 0.01, 0.01, 0.01, 0.01,}_{\text{mm}^2} \underbrace{0.0025, 0.0025}_{\text{m}^2} \right\}
$$

With these numerical values, the first correction vector to the coordinate approximations may be computed as (by a computer program)

$$
\Delta^0_{(1)} =
\begin{bmatrix}
\delta x_{1a} \\
\delta x_{2a} \\
\delta x_{1b} \\
\delta x_{2b}
\end{bmatrix}
=
\begin{bmatrix}
-1.005 \text{ m} \\
-1.284 \text{ m} \\
-0.019 \text{ m} \\
+0.969 \text{ m}
\end{bmatrix}
$$

The updated values of the coordinates of the two points are therefore

$$x_{1a} = 6.995 \text{ m} \qquad x_{1b} = 6.981 \text{ m}$$

$$x_{2a} = 49.716 \text{ m} \qquad x_{2b} = 41.969 \text{ m}$$

With these values we recompute the matrices $\mathbf{A}$ and $\mathbf{f}$, form normal equations, and solve for a new correction vector.

$$
\Delta^0_{(2)} = [0.00026 \quad -0.00042 \quad 0.00026 \quad -0.00038]^t (\text{m})
$$

which is zero to three decimal places, thus the iteration procedure may be terminated. The final coordinates required in the problem are then

$$\hat{x}_{1a} = 6.995 \text{ m} \qquad \hat{x}_{2a} = 49.716 \text{ m} \qquad \hat{x}_{1b} = 6.982 \text{ m} \qquad \hat{x}_{2b} = 41.968 \text{ m}$$

Introduction of Distance Constraint: The second part of the required information for this problem is to compute the coordinates of points A and B when the distance between them is constrained to be exactly 7.8 m. From the coordinates computed above, the said distance is less than 7.8 m and therefore when the constraint is introduced, these coordinates will change.

Without the constraint, n_0 is equal to six; that is, six independent variables are necessary for a unique determination of the geometry of the problem. This is true because all the parameters are functionally independent. With *one* geometric constraint, the functional model must be re-examined and a new n_0 computed. It can be easily shown that in this case $n_0 = 5$. Thus $n = 8$, $u = 4$, $c = 6$, $s = 1$, and all pertinent relationships are satisfied. Therefore

$$0 < u = 4 < n_0 = 5$$

$$c + s = 6 + 1 = (n - n_0) + u = 3 + 4$$

$$(n - n_0) = 3 < c = 6 < n = 8$$

$$s = 1 < u = 4$$

The constraint equation is

$$(x_{1b} - x_{1a})^2 + (x_{2b} - x_{2a})^2 - (7.8)^2 = 0$$

which when linearized yields

$$C\Delta = [-2(x_{1b}^0 - x_{1a}^0) \quad -2(x_{2b}^0 - x_{2a}^0) \quad 2(x_{1b}^0 - x_{1a}^0) \quad 2(x_{2b}^0 - x_{2a}^0)]$$

$$\times \begin{bmatrix} \delta x_{1a} \\ \delta x_{2a} \\ \delta x_{1b} \\ \delta x_{2b} \end{bmatrix}$$

$$g = [(7.8)^2 - (x_{1b}^0 - x_{1a}^0)^2 - (x_{2b}^0 - x_{2a}^0)^2]$$

Using the *same starting approximations*, the C and g matrices become

$$C = [2.0 \quad 20.0 \quad -2.0 \quad -20.0] \qquad \text{and} \qquad g = -40.16$$

With these matrices, the addition to $\Delta_{(1)}^0$ due to the constraint for the first iteration may be computed as the vector

$$\delta\Delta_{(1)} = [-0.006 \quad 0.267 \quad 0.011 \quad -0.078]^t (m)$$

Thus the total correction vector for the first iteration is

$$\Delta_{(1)} = [-1.011 \quad -1.017 \quad -0.008 \quad 0.891]^t (m)$$

The new approximation vector is therefore

$$x_{(1)}^0 = [6.989 \quad 49.983 \quad 6.992 \quad 41.891]^t (m)$$

Since this vector is different from that computed after the first iteration of the solution without constraint, it is very important to note that the value $\Delta_{(2)}^0$, computed there, *cannot be used here*. This is because the second iteration is now starting with a different approximation vector than that used for the linearization of condition equations used to compute $\Delta_{(2)}^0$ for the case of no constraint.

With $x^0_{(1)}$ computed above, a new set of matrices A, f, C, and g must be evaluated. From these a new total correction vector $\Delta_{(2)}$ can be obtained as

$$\Delta_{(2)} = [0.005 \quad -0.224 \quad -0.009 \quad 0.063]^t(m)$$

Some of the elements of that vector are still significant and a third iteration should be performed. The new approximation vector is

$$x^0_{(2)} = [6.994 \quad 49.761 \quad 6.983 \quad 41.954]^t(m)$$

and the total correction vector from the third iteration is

$$\Delta_{(3)} = [0.00013 \quad -0.00417 \quad -0.00013 \quad 0.00112]^t(m)$$

which is sufficiently small to terminate the relinearization process. The final answer for the coordinates of points A and B with distance between them constrained to 7.8 m is

$$\hat{x}_{1a} = 6.994 \text{ m} \qquad \hat{x}_{2a} = 49.757 \text{ m} \qquad \hat{x}_{1b} = 6.983 \text{ m} \qquad \hat{x}_{2b} = 41.955 \text{ m}$$

A check on the distance between the points computed from the above coordinates would show that it is equal to 7.8 m to within the fifth decimal place.

The a posteriori cofactor matrices are as follows.

a. *Without Constraint.* After terminating the iterations, the final results are

$$Q_{\Delta\Delta(f)} = 10^{-3} \begin{bmatrix} 1.693 & 0.784 & 3.470 & 4.419 \\ & 1.430 & 5.225 & 3.342 \\ & & 75.348 & 31.902 \\ \text{symmetric} & & & 46.003 \end{bmatrix} (m^2) \qquad \text{rank} = 4 = u$$

$$Q_{ll(f)} = 10^{-3}$$

$$\times \begin{bmatrix} 8.932 & 0.291 & -2.401 & 0.654 & 1.330 & -0.363 & 0.363 & -0.454 \\ & 9.031 & 0.654 & -2.178 & -0.363 & 1.209 & 0.431 & -0.538 \\ & & 4.602 & 1.471 & 2.997 & -0.817 & 0.817 & -1.020 \\ & & & 5.103 & 0.817 & 2.719 & 0.968 & -1.209 \\ & & & & 8.336 & 0.454 & -0.454 & 0.560 \\ & & & & & 8.491 & -0.538 & 0.671 \\ & & & & & & 2.061 & 0.548 \\ \text{symmetric} & & & & & & & 1.815 \end{bmatrix}$$

$$\text{rank} \, (Q_{ll})_{(f)} = 6$$

$$\hat{\sigma}^2_{0(f)} = 0.8228$$

$$\hat{\sigma}_{0(f)} = 0.9071$$

b. *With Distance Constraint.* After the last (sixth iteration) the results are

$$Q_{\Delta\Delta(f)} = 10^{-3} \begin{bmatrix} 1.678 & 0.815 & 4.184 & 4.185 \\ & 1.368 & 3.803 & 3.802 \\ & & 42.532 & 42.533 \\ \text{symmetric} & & & 42.533 \end{bmatrix} \qquad \text{rank} = 3 = u - s$$

$\mathbf{Q}_{ll(f)} = 10^{-3}$

$$\times \begin{bmatrix} 5.422 & 2.774 & -2.665 & 0.815 & -1.914 & 1.961 & 0.647 & -0.207 \\ & 7.274 & 0.841 & -2.291 & 1.935 & -0.436 & 0.230 & -0.712 \\ & & 4.581 & 1.483 & 2.754 & -0.642 & 0.838 & -1.001 \\ & & & 5.097 & -0.668 & 2.612 & 0.056 & -1.220 \\ & & & & 5.330 & 2.604 & -0.191 & 0.794 \\ & & & & & 6.951 & -0.726 & 0.508 \\ & & & & & & 2.038 & 0.528 \\ \text{symmetric} & & & & & & & 1.798 \end{bmatrix}$$

rank $\mathbf{Q}_{ll(f)} = 5$

$\hat{\sigma}_{0(f)}^2 = 0.5644$

$\hat{\sigma}_{0(f)} = 0.7513$

Example 9.3. In this example the problem of the three photographs with distance constraint in the preceding example is reworked using the technique of eliminating the constraint. Recall that

$$n = 8, \quad n_0 = 5, \quad r = 3, \quad u = 4, \quad c = 6, \quad \text{and} \quad s = 1$$

With the approximations

$$\mathbf{x}^0_{4,1} = [x_{1a}^0 \quad x_{2a}^0 \quad x_{1b}^0 \quad x_{2b}^0]^t = [8.0 \quad 51.0 \quad 7.0 \quad 41.0]^t$$

the constraint equation was

$$[2 \quad 20 \quad -2 \quad -20][\delta x_{1a} \quad \delta x_{2a} \quad \delta x_{1b} \quad \delta x_{2b}]^t = -40.16$$

Since $s = 1$, the partitioning becomes (note that $\mathbf{\Delta}_1$ is a 1×1 only and corresponds to the correction δx_{1a}).

$$[2]\mathbf{\Delta}_1 + [20 \quad -2 \quad -20]\mathbf{\Delta}_2 = -40.16$$

or

$$\mathbf{C}_{11} = [2] \quad \mathbf{C}_{12} = [20 \quad -2 \quad -20] \quad \mathbf{g} = [-40.16]$$

Further, partition the **B** matrix (evaluated at the same approximation vector $\mathbf{x}^0$) into

$$\mathbf{B}_1 \quad \text{and} \quad \mathbf{B}_2$$
$$\scriptstyle c, s \qquad\qquad \scriptstyle c, a$$

or

$$\mathbf{B}_1 = \begin{bmatrix} -100 \\ 0 \\ 100 \\ 0 \\ 100 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{B}_2 = \begin{bmatrix} 14.1 & 0 & 0 \\ 0 & -100 & 16.6 \\ 6.1 & 0 & 0 \\ 0 & 100 & 7.1 \\ 22.1 & 0 & 0 \\ 0 & 100 & 26.3 \end{bmatrix}$$
$$\scriptstyle 6,1 \qquad\qquad\qquad \scriptstyle 6,3$$

With B_1, B_2, f, C_{11}, C_{12}, and g the matrices $\bar{B}$ and $\bar{f}$ are

$$
\underset{6,\,3}{\bar{B}} = \begin{bmatrix} 1014.1 & -100.0 & -1000.0 \\ 0 & -100.0 & 16.6 \\ -993.9 & 100.0 & 1000.0 \\ 0 & 100.0 & 7.1 \\ -977.9 & 100.0 & 1000.0 \end{bmatrix} \qquad \bar{f} = \begin{bmatrix} -1927.1 \\ 19.4 \\ 1896.9 \\ 8.9 \\ 1880.9 \end{bmatrix}
$$

In addition to $\bar{B}$ and $\bar{f}$, the matrices A and Q from the preceding example are used in equations (9.23) and (9.24) to compute Δ_2 as

$$
\Delta_2 = \begin{bmatrix} \delta x_{2a} \\ \delta x_{1b} \\ \delta x_{2b} \end{bmatrix} = \begin{bmatrix} -1.017 \\ -0.008 \\ 0.891 \end{bmatrix} \text{(m)}
$$

and using it in equation (9.20) gives

$$
\Delta_1 = \delta x_{1a} = 1.011 \text{(m)}
$$

These four values are the same as those computed after the first iteration of the second part of Example 9.2. After the proper iteration, the final estimates for the coordinates are

$$
\hat{x}_{1a} = 6.994 \text{ m} \qquad \hat{x}_{2a} = 49.756 \text{ m} \qquad \hat{x}_{1b} = 6.983 \text{ m} \qquad \hat{x}_{2b} = 41.957 \text{ m}
$$

As regards a posteriori estimates, at the end of the iterations the results obtained are precisely the same as those given in Example 9.2.

9.3. SPECIAL CASES

From the general case covered in Section 9.2, several special cases can be obtained and solved using the relations derived so far. A more unusual case which requires a slightly different treatment is given in the following section.

9.3.1. Adjustment with Constraints When the Parameters Exceed the Conditions
It may be recalled that at the early discussion on adjustment with constraints it was indicated that the number of parameters u is rather arbitrary, particularly in regard to its upper limit. It was also said that although u has no upper limit, the number of condition equations *must* not exceed the number of observations n. Consequently, a rather interesting situation arises when there are more parameters than condition equations. This special case is therefore characterized by the inequality

$$
c < u \tag{9.31}
$$

while all other previously given relationships hold. In light of equation (9.31), the ranks of the matrices are

$$
\text{rank } (A) = c \qquad \text{rank } (B) = c \qquad \text{rank } (C) = s \tag{9.32}
$$

Because of equation (9.32) both treatments given in Section 9.2 must be modified to fit the present case.

DIRECT TREATMENT   Following the development in Section 9.2.1, the partitioning must stop with N, because in the present case it is of order u, but of rank $c < u$. Combining equations (9.10) and (9.3b) gives

$$\begin{bmatrix} -N & C^t \\ C & 0 \end{bmatrix} \begin{bmatrix} \Delta \\ k_c \end{bmatrix} = \begin{bmatrix} -t \\ g \end{bmatrix} \tag{9.33a}$$

which may be solved directly

$$\begin{bmatrix} \Delta \\ k_c \end{bmatrix} = \begin{bmatrix} -N & C^t \\ {}_{u,u} & {}_{u,s} \\ C & 0 \\ {}_{s,u} & {}_{s,s} \end{bmatrix}^{-1} \begin{bmatrix} -t \\ g \end{bmatrix} \tag{9.33b}$$

and from which Δ can be obtained. Note that although N is singular, the coefficient matrix in equation (9.33b) is nonsingular. With respect to precision estimation, special consideration is needed. Denote the inverse in equation (9.33b) by

$$\Omega = \begin{bmatrix} -N & C^t \\ C & 0 \end{bmatrix}^{-1} = \begin{bmatrix} \alpha & \beta^t \\ \beta & \gamma \end{bmatrix} \tag{9.34}$$

and expand the constant term vector to

$$a = \begin{bmatrix} -t \\ g \end{bmatrix} = \begin{bmatrix} -B^t W_e f \\ g \end{bmatrix} = \begin{bmatrix} B^t W_e A l \\ 0 \end{bmatrix} + \begin{bmatrix} -B^t W_e d \\ g \end{bmatrix} \tag{9.35}$$

The last vector in equation (9.35) is a constant, and thus from cofactor propagation

$$Q_{aa} = \begin{bmatrix} B^t W_e A \\ 0 \end{bmatrix} Q[A^t W_e B \quad 0^t] = \begin{bmatrix} B^t W_e B & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} N & 0 \\ 0 & 0 \end{bmatrix} \tag{9.36}$$

Applying propagation principle to equation (9.33b) also (referring to $[\Delta \quad k_c]^t$ by u) we get

$$Q_{uu} = \begin{bmatrix} Q_{\Delta\Delta} & Q_{\Delta k_c} \\ Q_{k_c\Delta} & Q_{k_c k_c} \end{bmatrix} = \Omega Q_{aa} \Omega \qquad (\Omega \text{ is symmetric})$$

$$= \begin{bmatrix} \alpha & \beta^t \\ \beta & \gamma \end{bmatrix} \begin{bmatrix} N & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha & \beta^t \\ \beta & \gamma \end{bmatrix} = \begin{bmatrix} \alpha N\alpha & \alpha N\beta^t \\ \beta N\alpha & \beta N\beta^t \end{bmatrix}. \tag{9.37}$$

from which

$$Q_{\Delta\Delta} = \alpha N\alpha \tag{9.38}$$

but from equation (9.34) we may write

$$\begin{bmatrix} -\mathbf{N} & \mathbf{C}^t \\ \mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \alpha & \beta^t \\ \beta & \gamma \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \tag{9.39}$$

then

$$-\mathbf{N}\alpha + \mathbf{C}^t\beta = \mathbf{I}$$

$$\mathbf{N}\alpha = -\mathbf{I} + \mathbf{C}^t\beta \tag{9.40}$$

$$\mathbf{C}\alpha = \mathbf{0}$$

$$\alpha\mathbf{C}^t = \mathbf{0} \qquad (\alpha \text{ is symmetric}) \tag{9.41}$$

Substituting equation (9.40) into (9.38)

$$\mathbf{Q}_{\Delta\Delta} = \alpha(-\mathbf{I} + \mathbf{C}^t\beta)$$

$$= -\alpha + \alpha\mathbf{C}^t\beta$$

or since $\alpha\mathbf{C}^t = \mathbf{0}$ from equation (9.41)

$$\mathbf{Q}_{\Delta\Delta} = -\alpha \tag{9.42}$$

Next

$$\mathbf{Q}_{\Delta k_c} = \alpha\mathbf{N}\beta^t = (-\mathbf{I} + \beta^t\mathbf{C})\beta^t$$

$$= -\beta^t + \beta^t\mathbf{C}\beta^t$$

$$\mathbf{C}\beta^t = \mathbf{I}$$

$$\mathbf{Q}_{\Delta k_c} = \beta^t - \beta^t = \mathbf{0} \tag{9.43}$$

$$\mathbf{Q}_{k_c k_c} = \beta\mathbf{N}\beta^t$$

$$-\mathbf{N}\beta^t + \mathbf{C}^t\gamma = \mathbf{0}$$

$$\mathbf{N}\beta^t = \mathbf{C}^t\gamma$$

$$\mathbf{Q}_{k_c k_c} = \beta\mathbf{C}^t\gamma$$

$$\mathbf{Q}_{k_c k_c} = \gamma \tag{9.44}$$

Collecting all the cofactor matrices then

$$\begin{bmatrix} \mathbf{Q}_{\Delta\Delta} & \mathbf{Q}_{\Delta k_c} \\ \mathbf{Q}_{k_c\Delta} & \mathbf{Q}_{k_c k_c} \end{bmatrix} = \begin{bmatrix} -\alpha & \mathbf{0} \\ \mathbf{0} & \gamma \end{bmatrix} \neq \begin{bmatrix} -\mathbf{N} & \mathbf{C}^t \\ \mathbf{C} & \mathbf{0} \end{bmatrix}^{-1} \tag{9.45}$$

These results when checked with those developed in equations (6.35) through (6.41) show that the two problems are essentially the same. The only difference is that here the upper left submatrix is singular, whereas in the

other case it was not and there the solution was possible using inversion by partitioning.

CONSTRAINT ELIMINATION   From the general relationship (9.2)

$$c + s = r + u$$

the number of elements in the subvector $\underset{a,\,1}{\Delta_2}$ remaining after eliminating $\underset{s,\,1}{\Delta_1}$ from $\underset{u,\,1}{\Delta}$ (see Section 9.2.3) is given by

$$a = u - s = c - r \tag{9.46}$$

Equation (9.46) indicates that the number of parameters a remaining after eliminating the constraints is less than the number of conditions c. Henceforth the matrix $\bar{\mathbf{B}}$ formed by equation (9.21a) will have a rank equal to a, leading to a nonsingular coefficient matrix $\bar{\mathbf{N}}$ in equation (9.23). Consequently, the solutions for Δ_2 in equation (9.23) and Δ_1 in equation (9.20) are both possible in the present case. This means that all equations and steps given in Section 9.2.3 apply directly to this case without any modification.

Example 9.4.† Suppose that there are $n = 2$ observations, $c = 2$ condition equations (with an $\mathbf{A} = \mathbf{I}$ for simplicity), $u = 3$ parameters, and $s = 2$ constraint equations.

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} + \begin{bmatrix} -3 & 1 & -1 \\ 1 & -4 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 18 \end{bmatrix} \quad \text{or} \quad \underset{2,\,1}{\mathbf{v}} + \underset{2,\,3}{\mathbf{B}}\,\underset{3,\,1}{\Delta} = \underset{2,\,1}{\mathbf{f}}$$

$$\begin{bmatrix} -1 & 1 & 1 \\ 1 & -3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 26 \end{bmatrix} \quad \text{or} \quad \underset{2,\,3}{\mathbf{C}}\,\underset{3,\,1}{\Delta} = \underset{2,\,1}{\mathbf{g}}$$

noting that $c + s = 2 + 2 = 4$ and $u + r = 3 + 1 = 4$ (one redundancy). It is required to compute, by two methods, the estimates of x_1, x_2, x_3, and their cofactor matrix. Take the a priori weight matrix

$$\mathbf{W} = \begin{bmatrix} 2 & 0 \\ 0 & 0.5 \end{bmatrix}$$

Direct Solution:

$$\mathbf{N} = \begin{bmatrix} 18.5 & -8 & 5 \\ -8 & 10 & 2 \\ 5 & 2 & 4 \end{bmatrix} \qquad \mathbf{t} = \begin{bmatrix} -3 \\ -32 \\ -22 \end{bmatrix}$$

† The numbers for this example are adopted from an example in reference Linkwitz, 1971, in the Bibliography.

According to equation (9.33b)

$$
\begin{bmatrix} \Delta \\ {}_{3,1} \\ \mathbf{k}_c \\ {}_{2,1} \end{bmatrix} = \begin{bmatrix} -\mathbf{N} & \mathbf{C}^t \\ \mathbf{C} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} -\mathbf{t} \\ \mathbf{g} \end{bmatrix}
$$

$$
= \begin{bmatrix}
-18.5 & 8 & -5 & -1 & 1 \\
 & -10 & -2 & 1 & -3 \\
 & & -4 & 1 & 1 \\
 & & & 0 & 0 \\
\text{symmetric} & & & & 0
\end{bmatrix}^{-1}
\begin{bmatrix} 3 \\ 32 \\ 22 \\ 2 \\ 26 \end{bmatrix}
$$

$$
= \begin{bmatrix}
-0.05 & -0.025 & -0.025 & -0.275 & -0.125 \\
 & -0.0125 & -0.0125 & 0.1125 & -0.3125 \\
1 & & -0.0125 & 0.6125 & 0.1875 \\
 & & & 2.1125 & -0.8125 \\
\text{symmetric} & & & & 0.3125
\end{bmatrix}
\begin{bmatrix} 3 \\ 32 \\ 22 \\ 2 \\ 26 \end{bmatrix}
$$

From which

$$
\Delta = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -5.30 \\ -8.65 \\ +5.35 \end{bmatrix}
$$

The cofactor matrix of the parameters $\mathbf{Q}_{\Delta\Delta}$ is the negative of submatrix 1 according to equation (9.42).

Solution by Elimination: First partition Δ to

$$
\begin{bmatrix} \Delta_1 \\ \Delta_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad\text{and}\quad \mathbf{v} + \mathbf{B}_1 \Delta_1 + \mathbf{B}_2 \Delta_2 = \mathbf{f}
$$

and correspondingly the constraints into $\mathbf{C}_{11}\Delta_1 + \mathbf{C}_{12}\Delta_2 = \mathbf{g}$

$$
\mathbf{C}_{11} = \begin{bmatrix} -1 & 1 \\ 1 & -3 \end{bmatrix} \quad
\mathbf{C}_{12} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad
\mathbf{B}_1 = \begin{bmatrix} -3 & 1 \\ 1 & -4 \end{bmatrix} \quad
\mathbf{B}_2 = \begin{bmatrix} -1 \\ -2 \end{bmatrix}
$$

and compute

$$
\bar{\mathbf{N}} = 80 \qquad \bar{\mathbf{t}} = 428 \qquad \Delta_2 = x_3 = 5.35
$$

$$
\Delta_1 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -5.30 \\ -8.65 \end{bmatrix}
$$

Thus the total answer vector is identical to that computed above. To compute $\mathbf{Q}_{\Delta\Delta}$, equation (9.29) is applied and the result is identical to that obtained in the direct method.

9.4. CONSTRAINTS WITH ADDED PARAMETERS

In all the cases covered in the previous sections of this chapter a tacit assumption was made that the constraint equations involve only the parameters used in the condition equations. In practice, however, there are situations in which the writing of constraints may introduce a new set of parameters, which may not be of interest themselves. Take the simple example of constraining a set of points in a plane to lie on a circle. Writing an equation of a circle may require, in addition to the coordinates of the points in question (which would belong to the parameters of the problem), a set of three added parameters. These may be, for example, the two coordinates of its center and its radius. It is, of course, possible to combine the added parameters with the original parameters and treat the problem with one of the cases already discussed. Although this is in principle possible (although singular matrices could be encountered), there may be situations in which keeping the two sets of parameters separate is easier to handle. Consequently, for the sake of completeness, this will be briefly treated hereafter.

As introduced in Section 5.3 (see Table 5.1) the symbol s' denotes the number of constraint equations when a set of q added parameters exists. In this case the total number of equations will be $(c + s')$ and the total number of parameters $(u + q)$. Therefore equations (9.1) and (9.2) become

$$s' < u + q \tag{9.47}$$

$$(c + s') = r + (u + q) \tag{9.48}$$

and furthermore

$$s' > q \tag{9.49}$$

or the number of constraints must exceed the number of added parameters, otherwise the constraints would not contribute to the adjustment problem.

The linear, or linearized, equations for the case of added parameters take the form

$$\underset{c,n}{\mathbf{A}}\ \underset{n,1}{\mathbf{v}} + \underset{c,u}{\mathbf{B}}\ \underset{u,1}{\mathbf{\Delta}} = \underset{c,1}{\mathbf{f}} \tag{9.50a}$$

$$\underset{s',u}{\mathbf{D}_1}\ \underset{u,1}{\mathbf{\Delta}} + \underset{s',q}{\mathbf{D}_2}\ \underset{q,1}{\mathbf{\Delta}'} = \underset{s',1}{\mathbf{h}} \tag{9.50b}$$

with rank $(\mathbf{A}) = c$, rank $(\mathbf{B})$ = the smaller of c or u, rank $(\mathbf{D}_1) = s'$, and rank $(\mathbf{D}_2) = q$. The vector $\mathbf{\Delta}$ is as usual the original parameter vector of interest, whereas $\mathbf{\Delta}'$ is the vector of added parameters that appears only in the constraints.

9.4.1. The General Case In this case the selection of $u \leq c$, which makes rank $(\mathbf{B}) = u$, is specified.

DIRECT TREATMENT The minimum principle is in this case

$$\phi = \mathbf{v}^t \mathbf{W} \mathbf{v} - 2\mathbf{k}^t(\mathbf{Av} + \mathbf{B}\Delta - \mathbf{f}) - 2\mathbf{k}_c^t(\mathbf{D}_1\Delta + \mathbf{D}_2\Delta' - \mathbf{h}) \qquad (9.51)$$

The partial derivatives of ϕ with respect to the three free variables $\mathbf{v}$, Δ, and Δ' may be evaluated and equated to zero, to obtain a minimum for ϕ. This leads to three equations that may be combined with equations (9.50a) and (9.50b) to give the following *total system of normal equations*:

$$
\begin{array}{c}
\\
n\\
c\\
u\\
s'\\
q
\end{array}
\begin{array}{ccccc}
n & c & u & s' & q
\end{array}
\begin{bmatrix}
-\mathbf{W} & \mathbf{A}^t & 0 & 0 & 0 \\
\mathbf{A} & 0 & \mathbf{B} & 0 & 0 \\
0 & \mathbf{B}^t & 0 & \mathbf{D}_1^t & 0 \\
0 & 0 & \mathbf{D}_1 & 0 & \mathbf{D}_2 \\
0 & 0 & 0 & \mathbf{D}_2^t & 0
\end{bmatrix}
\begin{bmatrix}
\mathbf{v} \\
\mathbf{k} \\
\Delta \\
\mathbf{k}_c \\
\Delta'
\end{bmatrix}
=
\begin{bmatrix}
0 \\
\mathbf{f} \\
0 \\
\mathbf{h} \\
0
\end{bmatrix}
\qquad (9.52)
$$

Obviously this equation is an extension of equation (9.7) by the addition of the Δ' vector. The direct solution of equation (9.52) yields the total answer by supplying all five vectors. For practical reasons, however, it is advisable to solve the system by partitioning. Thus we recall that

$$\mathbf{v} = \mathbf{Q}\mathbf{A}^t\mathbf{k} \qquad (9.8)$$

$$\mathbf{k} = \mathbf{W}_e(-\mathbf{B}\Delta + \mathbf{f}) \qquad (9.9)$$

$$\mathbf{N} = \mathbf{B}^t\mathbf{W}_e\mathbf{B} \qquad \text{rank } (\mathbf{N}) = u \text{ because } u < c$$

$$\mathbf{t} = \mathbf{B}^t\mathbf{W}_e\mathbf{f}$$

$$\Delta^0 = \mathbf{N}^{-1}\mathbf{t}$$

With these relationships, the third vector equation in (9.52) becomes

$$\Delta = \mathbf{N}^{-1}(\mathbf{t} + \mathbf{D}_1^t\mathbf{k}_c) = \Delta^0 + \mathbf{N}^{-1}\mathbf{D}_1^t\mathbf{k}_c \qquad (9.53)$$

which reduces equation (9.52) to the form

$$
\begin{bmatrix}
(\mathbf{D}_1\mathbf{N}^{-1}\mathbf{D}_1^t) & \mathbf{D}_2 \\
\mathbf{D}_2^t & 0
\end{bmatrix}
\begin{bmatrix}
\mathbf{k}_c \\
\Delta'
\end{bmatrix}
=
\begin{bmatrix}
\mathbf{h} - \mathbf{D}_1\Delta^0 \\
0
\end{bmatrix}
\qquad (9.54)
$$

Equation (9.54) may be solved directly for both $\mathbf{k}_c$ and Δ' or it may be reduced further

$$\mathbf{k}_c = \mathbf{P}^{-1}(\mathbf{h} - \mathbf{D}_1\Delta^0 - \mathbf{D}_2\Delta') \qquad (9.55)$$
$$_{s,1}$$

$$\Delta' = \mathbf{R}^{-1}[\mathbf{D}_2^t\mathbf{P}^{-1}(\mathbf{h} - \mathbf{D}_1\Delta^0)] \qquad (9.56)$$
$$_{q,1}$$

$$\Delta = \Delta^0 + \mathbf{N}^{-1}\mathbf{D}_1^t\mathbf{P}^{-1}(\mathbf{h} - \mathbf{D}_1\Delta^0 - \mathbf{D}_2\Delta') \qquad (9.57)$$
$$_{u,1}$$

where

$$\mathbf{P} = \mathbf{D}_1 \mathbf{N}^{-1} \mathbf{D}_1^t \tag{9.58}$$

$$\mathbf{R} = \mathbf{D}_2^t \, \mathbf{P}^{-1} \mathbf{D}_2 \tag{9.59}$$

The computation would be carried out in this sequence, first $\mathbf{N}$ and $\mathbf{t}$ are computed and used to compute $\boldsymbol{\Delta}^0$. Next, with $\mathbf{D}_1$, $\mathbf{D}_2$, and $\mathbf{h}$ from the constraint equations, the vector of added parameters $\boldsymbol{\Delta}'$ is evaluated. Using $\boldsymbol{\Delta}'$ thus computed, the original parameters $\boldsymbol{\Delta}$ can then be computed.

The a posteriori cofactor matrices are

$$\mathbf{Q}_{\Delta'\Delta'} = \mathbf{R}^{-1} \tag{9.60}$$

$$\mathbf{Q}_{\Delta\Delta} = \mathbf{N}^{-1}[\mathbf{I} - \mathbf{D}_1^t \mathbf{P}^{-1} \mathbf{D}_1 \mathbf{N}^{-1} + \mathbf{D}_1^t \mathbf{P}^{-1} \mathbf{D}_2 \, \mathbf{R}^{-1} \mathbf{D}_2^t \, \mathbf{P}^{-1} \mathbf{D}_1 \mathbf{N}^{-1}] \tag{9.61}$$

It is important to note that in case $\boldsymbol{\Delta}'$ is not present, equation (9.61) reduces to equation (9.15b) by enforcing $\mathbf{D}_2 = 0$.

An interesting case arises when the matrix $\mathbf{D}_1$ happens to be a square nonsingular matrix. In this case the first two terms of equation (9.61) cancel out and only the last term remains, or

$$\mathbf{Q}_{\Delta\Delta} = \mathbf{D}_1^{-1} \mathbf{D}_2 \, \mathbf{R}^{-1} \mathbf{D}_2^t (\mathbf{D}_1^t)^{-1} \tag{9.62}$$

$$\mathbf{Q}_{\Delta\Delta'} = -\mathbf{N}^{-1} \mathbf{D}_1^t \mathbf{P}^{-1} \mathbf{D}_2 \, \mathbf{R}^{-1} \tag{9.63}$$

Two useful computational checks are possible,

$$\mathbf{F} = \mathbf{Q}_{\Delta\Delta} \mathbf{D}_1^t + \mathbf{Q}_{\Delta\Delta'} \mathbf{D}_2^t = 0 \tag{9.64}$$

$$\mathbf{G} = \mathbf{Q}_{\Delta\Delta'} \mathbf{D}_1^t + \mathbf{Q}_{\Delta'\Delta'} \mathbf{D}_2^t = 0 \tag{9.65}$$

The cofactor matrices of $\mathbf{v}$ and $\tilde{l}$ are the same as those given in equations (9.17b) and (6.31).

SOLUTION BY ELIMINATING CONSTRAINTS   From q of the s' constraint equations $(q < s')$ eliminate the q added parameters $\boldsymbol{\Delta}'$. The remaining $(s' - q)$ constraint equations are then used to eliminate as many parameters from $\boldsymbol{\Delta}$. Partition equation (9.50b),

$$\underset{\substack{s,u}}{\mathbf{D}_{11}} \underset{\substack{u,1}}{\boldsymbol{\Delta}} + \underset{\substack{s,q}}{\mathbf{D}_{12}} \underset{\substack{q,1}}{\boldsymbol{\Delta}'} = \underset{\substack{s,1}}{\mathbf{h}_1} \tag{9.66a}$$

$$\underset{\substack{q,u}}{\mathbf{D}_{21}} \underset{\substack{u,1}}{\boldsymbol{\Delta}} + \underset{\substack{q,q}}{\mathbf{D}_{22}} \underset{\substack{q,1}}{\boldsymbol{\Delta}'} = \underset{\substack{q,1}}{\mathbf{h}_2} \tag{9.66b}$$

where

$$s = s' - q$$

Solve equation (9.66b) for Δ' in terms of Δ,

$$\underset{q,\,1}{\Delta'} = D_{22}^{-1}(h_2 - D_{21}\Delta) \tag{9.66c}$$

making sure that D_{22} is nonsingular during the partitioning operation, and substitute into equation (9.66a)

$$(D_{11} - D_{12}\,D_{22}^{-1}\,D_{21})\Delta = h_1 - D_{12}\,D_{22}^{-1}\,h_2 \tag{9.66d}$$

Let

$$\underset{s,\,u}{C} = D_{11} - D_{12}\,D_{22}^{-1}D_{21} \tag{9.67a}$$

$$\underset{s,\,1}{g} = h_1 - D_{12}\,D_{22}^{-1}h_2 \tag{9.67b}$$

then equation (9.66d) combined with equation (9.50a) becomes the same as equations (9.3a) and (9.3b) and can be treated in a manner identical to that explained previously.

With respect to precision estimation, the cofactor matrix of the original parameters $Q_{\Delta\Delta}$ may be computed from equation (9.29).

On the other hand $Q_{\Delta'\Delta'}$ may be computed directly from equation (9.66c) as

$$Q_{\Delta'\Delta'} = D_{22}^{-1}D_{21}Q_{\Delta\Delta}\,D_{21}^{t}(D_{22}^{-1})^{t} \tag{9.68}$$

Example 9.5. The following case of adjustment is to estimate Δ, Δ', and their combined cofactor matrix:

$$Av + B\Delta = f \qquad \text{with } Q = I$$

$$D_1\Delta + D_2\,\Delta' = h$$

with

$$\underset{3,\,3}{A} = I \qquad \underset{3,\,2}{B} = \begin{bmatrix} 2 & -3 \\ -1 & 2 \\ 0 & 1 \end{bmatrix} \qquad \underset{3,\,1}{f} = \begin{bmatrix} -1.1 \\ 1.2 \\ 1.0 \end{bmatrix} \qquad \underset{2,\,1}{h} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$$

$$\underset{2,\,2}{D_1} = \begin{bmatrix} 1 & -1 \\ 2 & -1 \end{bmatrix} \qquad \underset{2,\,1}{D_2} = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \qquad \underset{2,\,1}{\Delta} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad \Delta' = x_3$$

Solution:

$$N = \begin{bmatrix} 5 & -8 \\ -8 & 14 \end{bmatrix} \qquad t = \begin{bmatrix} -3.4 \\ 6.7 \end{bmatrix} \qquad \Delta^0 = \begin{bmatrix} 1.00 \\ 1.05 \end{bmatrix}$$

Using equation (9.56)

$$\Delta' = -0.97922$$

Next from equation (9.57)

$$\Delta = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1.00 \\ 1.05 \end{bmatrix} + \begin{bmatrix} 0.06233 \\ 0.03311 \end{bmatrix} = \begin{bmatrix} 1.06233 \\ 1.08311 \end{bmatrix}$$

hence the final vector is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1.06233 \\ 1.08311 \\ -0.97922 \end{bmatrix}$$

To compute $\mathbf{Q}_{\Delta\Delta}$ for x_1 and x_2, use equation (9.62) because $\mathbf{D}_1$ is square and nonsingular. Thus

$$\mathbf{Q}_{\Delta\Delta} = \begin{bmatrix} 0.116883 & 0.155844 \\ \text{symmetric} & 0.207792 \end{bmatrix}$$

Next

$$\mathbf{Q}_{\Delta'\Delta'} = \mathbf{R}^{-1} = 1/77 = 0.012987$$

Finally, the crosscofactor matrix is computed from equation (9.63) as

$$\mathbf{Q}_{\Delta\Delta'} = [0.038961 \quad 0.051948]^t$$

leading to the final cofactor matrix for all three variables

$$\mathbf{Q}_{xx} = \begin{bmatrix} 0.116883 & 0.155844 & 0.038961 \\ & 0.207792 & 0.051948 \\ \text{symmetric} & & 0.012987 \end{bmatrix}$$

Solution by elimination is as follows:

$$\mathbf{D}_{11} = [1 \quad -1] \qquad \mathbf{D}_{12} = 1 \qquad \mathbf{h}_1 = -1$$

$$\mathbf{D}_{21} = [2 \quad -1] \qquad \mathbf{D}_{22} = -2 \qquad \mathbf{h}_2 = -3$$

$$\mathbf{C} = \mathbf{D}_{11} - \mathbf{D}_{12}\mathbf{D}_{22}^{-1}\mathbf{D}_{21} = [4 \quad -3]$$

$$\mathbf{g} = \mathbf{h}_1 - \mathbf{D}_{12}\mathbf{D}_{22}^{-1}\mathbf{h}_2 = 1$$

$$\mathbf{B}_1 = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} \qquad \mathbf{B}_2 = \begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix} \qquad \mathbf{C}_{11} = 4 \qquad \mathbf{C}_{12} = -3$$

Thus

$$\bar{\mathbf{B}} = -\mathbf{B}_1\mathbf{C}_{11}^{-1}\mathbf{C}_{12} + \mathbf{B}_2 = [-1.50 \quad 1.25 \quad 1.00]^t$$

$$\bar{\mathbf{f}} = -\mathbf{B}_1\mathbf{C}_{11}^{-1}\mathbf{g} + \mathbf{f} = [-1.60 \quad 1.45 \quad 1.00]^t$$

from which, because both **A** and **Q** are identity matrices, we get,

$$\Delta_2 = x_2 = (\bar{\mathbf{B}}^t\bar{\mathbf{B}})^{-1}(\bar{\mathbf{B}}^t\mathbf{f}) = 1.08311$$

Then

$$\Delta_1 = x_1 = \mathbf{C}_{11}^{-1}(\mathbf{g} - \mathbf{C}_{12}\,\Delta_2) = 1.06233$$

$$\Delta' = x_3 = \mathbf{D}_{22}^{-1}(\mathbf{h}_2 - \mathbf{D}_{21}\Delta) = -0.97933$$

thus the total vector is identical to that computed before. As for a posteriori cofactor matrices, suffice it to compute $\mathbf{Q}_{\Delta\Delta}$ for x_1 and x_2. Using equation (9.29),

$$\mathbf{Q}_{x_1x_1} = \mathbf{C}_{11}^{-1}\mathbf{C}_{12}\,\bar{\mathbf{N}}^{-1}\mathbf{C}_{12}^t\,\mathbf{C}_{11}^{-1} = 0.116883$$

$$\mathbf{Q}_{x_1x_2} = -\mathbf{C}_{11}^{-1}\mathbf{C}_{12}\,\bar{\mathbf{N}}^{-1} = -0.155844$$

$$\mathbf{Q}_{x_2x_2} = \bar{\mathbf{N}}^{-1} = 0.207792$$

with the final result the same as that evaluated previously.

Example 9.6

Given: Figure 9.3 shows essentially the same geometric layout of the problem in example 9.2, but expanded further to include four object points *A*, *B*, *C*, and *D*. The measurements for these four points on the three photographs are given in Table 9-1. The standard deviation for all the above measurements is 0.10 mm, the principal distance is $p = 100$ mm; $l_4 = 10$ m and $l_5 = 8$ m, and their standard deviation is 0.05 m (no correlation).
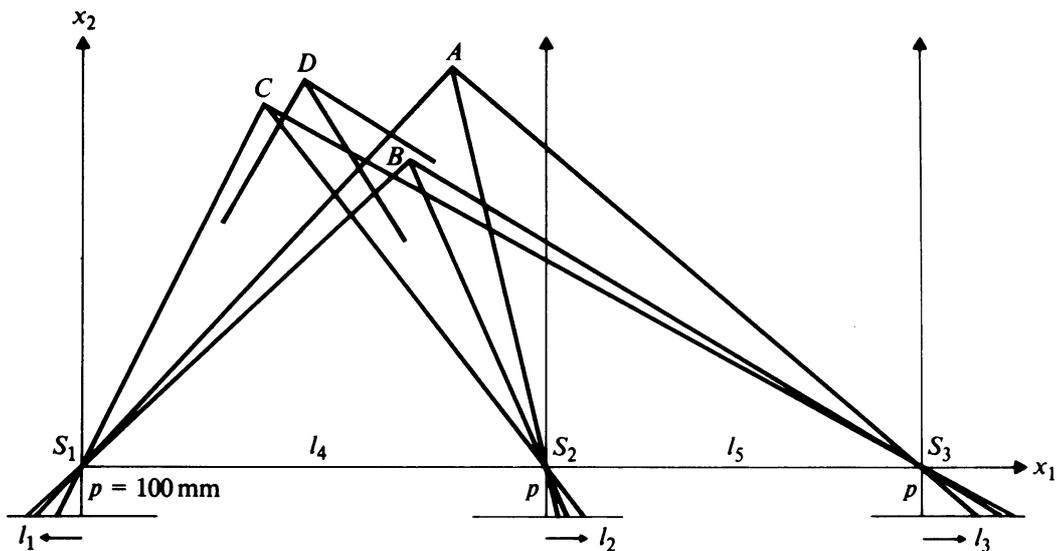
Figure 9.3. (Not to Scale)

TABLE 9-1

IMAGE POINT	$l_1 = \lvert x_1 \rvert$ (mm) ON $S1$	$l_2 = \lvert x_2 \rvert$ (mm) ON $S2$	$l_3 = \lvert x_3 \rvert$ (mm) on $S3$
a	14.1	6.1	22.1
b	16.6	7.1	26.3
c	6.5	15.2	32.5
d	8.3	12.3	28.7

Required: (a) Compute the coordinates of the four points A, B, C, and D. (b) Compute a new set of coordinates for the four points when they are constrained to lie on a circle whose equation is

$$(x_{1i} - x_{10})^2 + (x_{2i} - x_{20})^2 - R^2 = 0 \qquad i = 1, \ldots, 4 \text{ (for } A, B, C, D)$$

where x_{10}, x_{20}, and R (representing the coordinates of the center of the circle and its radius) are three added parameters. Use the following approximations for the original and added parameters (the numbers between parenthesis represent the x_1 and x_2 coordinates, respectively).

$$A = (8, 51), \qquad B = (7, 41), \qquad C = (4, 47), \qquad D = (4, 48)$$

$$x_{10} = 7, \qquad x_{20} = 46, \qquad R = 4$$

all values are in metres. Use both the direct procedure and the one by elimination of constraints.

Solution:

Part a. The number of observations is $n = 14$, the minimum number of variables necessary to define the geometry of the problem is $n_0 = 10$, redundancy $r = 4$; unknown parameters to be solved for are $u = 8$; and number of condition equations are $c = r + u = 12$. The condition equations are as follows

$$
\begin{aligned}
l_{1a}x_{2a} - px_{1a} &= 0 & \qquad l_{1c}x_{2c} - px_{1c} &= 0 \\
l_{2a}x_{2a} - pl_4 + px_{1a} &= 0 & l_{2c}x_{2c} - p_4 + px_{1c} &= 0 \\
l_{3a}x_{2a} - pl_4 - pl_5 + px_{1a} &= 0 & l_{3c}x_{2c} - pl_4 - pl_5 + px_{1c} &= 0 \\
l_{1b}x_{2b} - px_{1b} &= 0 & l_{1d}x_{2d} - px_{1d} &= 0 \\
l_{2b}x_{2b} - pl_4 + px_{1b} &= 0 & l_{2d}x_{2d} - pl_4 + px_{1d} &= 0 \\
l_{3b}x_{2b} - pl_4 - pl_5 + px_{1b} &= 0 & l_{3d}x_{2d} - pl_4 - pl_5 + px_{1d} &= 0
\end{aligned}
$$

where l_i represent observational variables.

Linearization: For linearization purposes choose the sequence of the elements in the observation and parameter vectors as follows:

$$\underset{14,\,1}{l} = [l_{1a} \quad l_{2a} \quad l_{3a} \quad l_{1b} \quad l_{2b} \quad l_{3b} \quad l_{1c} \quad l_{2c} \quad l_{3c} \quad l_{1d} \quad l_{2d} \quad l_{3d} \quad l_4 \quad l_5]^t$$

$$\underset{8,\,1}{x} = [x_{1a} \quad x_{2a} \quad x_{1b} \quad x_{2b} \quad x_{1c} \quad x_{2c} \quad x_{1d} \quad x_{2d}]^t$$

$$\mathbf{A}_{12,14} =$$

$$
\begin{bmatrix}
x^0_{2a} & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & 0 & 0 \\
0 & x^0_{2a} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & -p & 0 \\
0 & 0 & x^0_{2a} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & -p & -p \\
0 & & 0 & x^0_{2b} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & 0 & 0 \\
0 & & & 0 & x^0_{2b} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & -p & 0 \\
0 & & & & 0 & x^0_{2b} & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & -p & -p \\
0 & & & & & 0 & x^0_{2c} & \cdot & \cdot & \cdot & \cdot & 0 & 0 & 0 \\
0 & & & & & & 0 & x^0_{2c} & \cdot & \cdot & \cdot & 0 & -p & 0 \\
0 & & & & & & & 0 & x^0_{2c} & \cdot & \cdot & 0 & -p & -p \\
0 & & & & & & & & 0 & x^0_{2d} & \cdot & 0 & 0 & 0 \\
0 & & & & & & & & & 0 & x^0_{2d} & 0 & -p & 0 \\
0 & & \cdot & & \cdot & & \cdot & & & & 0 & x^0_{2d} & -p & -p
\end{bmatrix}
$$

$$\mathbf{B}_{12,8} =
\begin{bmatrix}
-p & l_{1a} & 0 & 0 & 0 & 0 & 0 & 0 \\
p & l_{2a} & 0 & 0 & \cdot & \cdot & & \\
p & l_{3a} & 0 & 0 & \cdot & \cdot & & \\
0 & 0 & -p & l_{1b} & \cdot & \cdot & \cdot & \cdot \\
& & p & l_{2b} & & & \cdot & \cdot \\
\cdot & \cdot & p & l_{3b} & 0 & 0 & \cdot & \cdot \\
\cdot & \cdot & 0 & 0 & -p & l_{1c} & & \\
\cdot & \cdot & & & p & l_{2c} & & \\
\cdot & \cdot & & & p & l_{3c} & 0 & 0 \\
\cdot & \cdot & & & 0 & 0 & -p & l_{1d} \\
& & & & 0 & 0 & p & l_{2d} \\
0 & 0 & 0 & 0 & 0 & 0 & p & l_{3d}
\end{bmatrix}
$$

$$\mathbf{f}_{12,1} =
\begin{bmatrix}
px^0_{1a} - l_{1a}x^0_{2a} \\
pl_4 - l_{2a}x^0_{2a} - px^0_{1a} \\
pl_4 + pl_5 - l_{3a}x^0_{2a} - px^0_{1a} \\
px^0_{1b} - l_{1b}x^0_{2b} \\
pl_4 - l_{2b}x^0_{2b} - px^0_{1b} \\
pl_4 + pl_5 - l_{3b}x^0_{2b} - px^0_{1b} \\
px^0_{1c} - l_{1c}x^0_{2c} \\
pl_4 - l_{2c}x^0_{2c} - px^0_{1c} \\
pl_4 + pl_5 - l_{3c}x^0_{2c} - px^0_{1c} \\
px^0_{1d} - l_{1d}x^0_{2d} \\
pl_4 - l_{2d}x^0_{2d} - px^0_{1d} \\
pl_4 + pl_5 - l_{3d}x^0_{2d} - px^0_{1d}
\end{bmatrix}
$$

The a priori cofactor matrix is diagonal (assuming $\sigma_0^2 = 1$)

$$\mathop{\mathbf{Q}}_{14,\,14} = \text{diag.}\{0.01,\ 0.01,\ 0.01,\ 0.01,\ 0.01,\ 0.01,\ 0.01,\ 0.01,\ 0.01,$$

$$0.01,\ 0.01,\ 0.01,\ 0.0025,\ 0.0025\}$$

Using the given approximations and observations, the above matrices may be evaluated and used to construct and solve the reduced normal equations. For purposes of checking, the values from the first and last iterations are given as follows:

$$\Delta_{(1)} = [-1.00344 \quad -1.28713 \quad -0.01703 \quad -0.96588 \quad -1.00480$$

$$-0.85048 \quad 0.03100 \quad 0.64280]^t (m)$$

$$\Delta_{(2)} = [0.00019 \quad -0.00029 \quad 0.00019 \quad -0.00026 \quad 0.00022$$

$$-0.00027 \quad 0.00021 \quad -0.00028]^t (m)$$

The elements of this vector are obviously sufficiently small that the process of iterations can be terminated giving the final required answer.

$$\hat{x}_{1a} = 6.997 \text{ m} \qquad \hat{x}_{1c} = 2.995 \text{ m}$$

$$\hat{x}_{2a} = 49.713 \text{ m} \qquad \hat{x}_{2c} = 49.149 \text{ m}$$

$$\hat{x}_{1b} = 6.983 \text{ m} \qquad \hat{x}_{1d} = 4.031 \text{ m}$$

$$\hat{x}_{2b} = 41.966 \text{ m} \qquad \hat{x}_{2d} = 48.643 \text{ m}$$

Part (b). First, reconsider the minimum number of variables that established the geometry of the problem (Table 9-2). Consequently, the minimum number of variables in this case is $n_0 = 9$, thus $r = 14 - 9 = 5$. Therefore

$$c + s = (r) + (u + q) = 5 + (8 + 3) = 16 \text{ equations}$$

TABLE 9-2

	NUMBER OF VARIABLES
Point S_1 is fixed by virtue of being the origin of the coordinate system	
Points S_2, S_3 being on the x_1 axis require one coordinate each to locate	2
Point A, two coordinates	2
Point B, two coordinates	2
Point C, two coordinates	2
(These three points would now determine a unique circle on which point D must lie. With this information about point D, only one more parameter is needed in order to fix its position)	1
Total	9

A logical choice would be to have $c = 12$ as in part (a) above and therefore write $s = 4$ constraints, one for each point. This simplifies the writing of both condition and constraint equations considerably, and satisfies all pertinent relations, particularly $s' > q$.

The constraint equations therefore take the following form:

$$(x_{1a} - x_{10})^2 + (x_{2a} - x_{20})^2 - R^2 = 0$$

$$(x_{1b} - x_{10})^2 + (x_{2b} - x_{20})^2 - R^2 = 0$$

$$(x_{1c} - x_{10})^2 + (x_{2c} - x_{20})^2 - R^2 = 0$$

$$(x_{1d} - x_{10})^2 + (x_{2d} - x_{20})^2 - R^2 = 0$$

Linearization:

$$\mathbf{D}_1 = \begin{bmatrix} 2(x_{1a}^0 - x_{10}^0) & 2(x_{2a}^0 - x_{20}^0) & 0 & 0 \\ 0 & 0 & 2(x_{1b}^0 - x_{10}^0) & 2(x_{2b}^0 - x_{20}^0) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 2(x_{1c}^0 - x_{10}^0) & 2(x_{2c}^0 - x_{20}^0) & 0 & 0 \\ 0 & 0 & 2(x_{1d}^0 - x_{10}^0) & 2(x_{2d}^0 - x_{20}^0) \end{bmatrix}$$

$$\mathbf{D}_2 = \begin{bmatrix} -2(x_{1a}^0 - x_{10}^0) & -2(x_{2a}^0 - x_{20}^0) & -2R^0 \\ -2(x_{1b}^0 - x_{10}^0) & -2(x_{2b}^0 - x_{20}^0) & -2R^0 \\ -2(x_{1c}^0 - x_{10}^0) & -2(x_{2c}^0 - x_{20}^0) & -2R^0 \\ -2(x_{1d}^0 - x_{10}^0) & -2(x_{2d}^0 - x_{20}^0) & -2R^0 \end{bmatrix}$$

$$\mathbf{h} = \begin{bmatrix} R^{02} - (x_{1a}^0 - x_{10}^0)^2 - (x_{2a}^0 - x_{20}^0)^2 \\ R^{02} - (x_{1b}^0 - x_{10}^0)^2 - (x_{2b}^0 - x_{20}^0)^2 \\ R^{02} - (x_{1c}^0 - x_{10}^0)^2 - (x_{2c}^0 - x_{20}^0)^2 \\ R^{02} - (x_{1d}^0 - x_{10}^0)^2 - (x_{2d}^0 - x_{20}^0)^2 \end{bmatrix}$$

Again, for checking purposes we have

First Iteration

$$\underset{8,1}{\Delta} = \begin{bmatrix} -1.00598 & -1.24835 & -0.01922 & -0.98883 & -1.02657 \\ -0.74275 & 0.05950 & 0.42954 \end{bmatrix}^t$$

$$\underset{3,\,1}{\Delta'} = [-0.23980 \quad -0.15638 \quad -0.30651]^t$$

Third and Last Iteration

$$\Delta_{(3)} = [-0.00011 \quad 0.00019 \quad -0.00012 \quad 0.00124 \quad 0.00032$$
$$-0.00274 \quad 0.00008 \quad -0.00091]^t$$
$$\Delta'_{(3)} = [-0.00119 \quad 0.00063 \quad -0.00415]^t$$

The final answer is then

$$\hat{x}_{1a} = 6.995 \text{ m} \qquad \hat{x}_{1d} = 4.043 \text{ m}$$

$$\hat{x}_{2a} = 49.774 \text{ m} \qquad \hat{x}_{2d} = 48.552 \text{ m}$$

$$\hat{x}_{1b} = 6.983 \text{ m} \qquad \hat{x}_{10} = 6.879 \text{ m}$$

$$\hat{x}_{2b} = 41.980 \text{ m} \qquad \hat{x}_{20} = 45.876 \text{ m}$$

$$\hat{x}_{1c} = 2.992 \text{ m} \qquad \hat{R} = 3.898 \text{ m}$$

$$\hat{x}_{2c} = 46.166 \text{ m}$$

The a posteriori cofactor matrices are (the differences between first iteration results, given here, and last iteration were insignificant) as follows.

a. *Without Circle Constraint*

$$\underset{8,\,8}{Q_{\Delta\Delta}} = 10^{-3}
\begin{bmatrix}
1.73 & 3.40 & 0.77 & 4.44 & 0.35 & 4.88 & 0.46 & 5.14 \\
 & 77.27 & 5.25 & 31.87 & 2.22 & 35.05 & 3.00 & 36.94 \\
 & & 1.39 & 3.41 & 0.35 & 4.87 & 0.46 & 5.13 \\
 & & & 45.11 & 1.87 & 29.59 & 2.53 & 31.19 \\
 & & & & 1.46 & -1.91 & 0.22 & 2.17 \\
 & & & & & 61.46 & 2.79 & 34.29 \\
 & & & & & & 1.45 & -0.71 \\
\text{symmetric} & & & & & & & 69.66
\end{bmatrix}$$

rank $(Q_{\Delta\Delta} = 8 = u)$

$$Q_{ll} = 10^{-3}
\begin{bmatrix}
8.87 & -2.56 & 1.43 & 0.21 & 0.46 & -0.25 & 0.18 & 0.40 & -0.22 & 0.17 & 0.39 & -0.22 & 0.25 & -0.31 \\
 & 4.20 & 3.24 & 0.47 & 1.04 & -0.57 & 0.40 & 0.91 & -0.50 & 0.39 & 0.89 & -0.49 & 0.55 & -0.71 \\
 & & 8.19 & -0.26 & -0.58 & 0.32 & -0.22 & -0.51 & 0.28 & -0.22 & -0.50 & 0.28 & -0.31 & 0.40 \\
 & & & 8.93 & 2.39 & 1.32 & 0.22 & 0.50 & 0.28 & 0.22 & 0.49 & 0.27 & 0.32 & -0.39 \\
 & & & & 4.56 & 2.96 & 0.50 & 1.13 & -0.63 & 0.49 & 1.10 & -0.61 & 0.81 & -0.86 \\
 & & & & & 8.37 & -0.28 & -0.62 & 0.35 & 1.10 & -0.61 & 0.34 & -0.34 & 0.48 \\
 & & & & & & 8.89 & -2.51 & 1.40 & -0.19 & 0.43 & -0.24 & 0.27 & -0.34 \\
 & & & & & & & 4.35 & 3.15 & 0.43 & 0.96 & -0.54 & 0.61 & -0.77 \\
 & & & & & & & & 3.25 & -0.24 & -0.54 & 0.29 & -0.34 & 0.43 \\
 & & \text{symmetric} & & & & & & & 8.88 & -2.52 & 1.41 & 0.26 & -0.33 \\
 & & & & & & & & & & 4.30 & 3.17 & 0.59 & -0.75 \\
 & & & & & & & & & & & 8.28 & -0.33 & 0.42 \\
 & & & & & & & & & & & & 1.90 & 0.75 \\
 & & & & & & & & & & & & & 1.57
\end{bmatrix}$$

$$\operatorname{rank}(Q_{ll}) = 10 = n - r$$

244

$$Q_{pp} = 10^{-3}$$

1.73										
3.45	76.49									
0.77	5.20	1.39								
4.47	31.41	3.43	44.83							
0.32	2.66	0.33	2.13	1.21						
5.02	32.87	4.99	28.30	−0.69	55.42					
0.50	2.43	0.49	2.19	0.54	1.19	1.02				
4.86	41.24	4.89	33.73	−0.25	46.25	2.45	45.99			
−0.44	39.95	1.51	−4.08	2.28	−3.46	1.01	3.38	38.35		
4.17	50.30	4.29	38.98	2.20	31.43	2.26	37.64	14.06	43.65	
−0.36	23.61	1.07	−7.32	0.08	3.92	0.08	4.88	22.67	5.84	16.46

$(Q_{\Delta\Delta})$ $(Q_{\Delta'\Delta'})$

symmetric

245

The reference variance and reference standard deviation are

$$\hat{\sigma}_0^2 = 0.4852$$

$$\hat{\sigma}_0 = 0.70$$

b. *With Circle Constraint.* Because there are three additional parameters, $\mathbf{Q}_{\Delta\Delta}$, $\mathbf{Q}_{\Delta'\Delta'}$, and $\mathbf{Q}_{\Delta\Delta'}$ were computed separately and collected in a matrix of cofactors $\mathbf{Q}_{pp}$ for all 11 parameters. The value of $\mathbf{Q}_{pp}$ after the first iteration is on page 245.
The computed rank of the 8×8 submatrix $\mathbf{Q}_{\Delta\Delta}$ is 7. This can be explained by recognizing that the number of independent primary parameters is $u - (s' - q) = 8 - (4 - 3) = 7$. The rank of $\mathbf{Q}_{\Delta'\Delta'}$ was also computed and found to be equal to 3 as expected.
The computational checks of equations (9.64) and (9.65) yield

$$\underset{4,\,8}{\mathbf{F}} = 10^{-11} \begin{bmatrix} 0.05 & 0.07 & 0.05 & -0.09 & -1.5 & 2.9 & -0.1 & -0.4 \\ -0.07 & 0.27 & 0 & -0.02 & 2.2 & 0 & 0.4 & 0.4 \\ -0.05 & 0.02 & -0.02 & 0 & -2.9 & 3.6 & -0.4 & -0.4 \\ 0.3 & -0.1 & 0.03 & 0.08 & 1.5 & 1.1 & 0.7 & 1.1 \end{bmatrix}$$

$$\underset{3,\,4}{\mathbf{G}} = 10^{-11} \begin{bmatrix} -1.8 & 0.68 & 0.82 & -0.43 \\ -1.1 & 1.8 & 0.41 & -0.73 \\ -1.5 & 0 & -0.05 & -0.50 \end{bmatrix}$$

both of which are effectively zero. Finally,

$$\hat{\sigma}_0^2 = 0.7724$$
$$\hat{\sigma}_0 = 0.8789$$

Solution by Elimination of Constraints: First, $s = s' - q = 1$. The partitioned matrices have the following submatrices:

$$\underset{1,\,8}{\mathbf{D}}_{11} = [2(x_{1a}^0 - x_{10}^0) \quad 2(x_{2a}^0 - x_{10}^0) \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0]$$

$$\underset{1,\,3}{\mathbf{D}}_{12} = [2(x_{1a}^0 - x_{10}^0) \quad -2(x_{2a}^0 - x_{20}^0) \quad -2R^0]$$

$$\underset{1,\,1}{\mathbf{h}}_1 = (R^{02} - (x_{1a}^0 - x_{10}^0)^2 - (x_{2a}^0 - x_{20}^0)^2)$$

$$\mathbf{D}_{21} \atop 3,8 = \begin{bmatrix} 0 & 0 & 2(x_{1b}^0 - x_{10}^0) & 2(x_{2b}^0 - x_{20}^0) & 0 \\ 0 & 0 & 0 & 0 & 2(x_{1c}^0 - x_{10}^0) \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 & 0 \\ 2(x_{2c}^0 - x_{20}^0) & 0 & 0 \\ 0 & 2(x_{1d}^0 - x_{10}^0) & 2(x_{2d}^0 - x_{20}^0) \end{bmatrix}$$

$$\mathbf{D}_{22} \atop 3,3 = \begin{bmatrix} -2(x_{1b}^0 - x_{10}^0) & -2(x_{2b}^0 - x_{20}^0) & -2R^0 \\ -2(x_{1c}^0 - x_{10}^0) & -2(x_{2c}^0 - x_{20}^0) & -2R^0 \\ -2(x_{1d}^0 - x_{10}^0) & -2(x_{2d}^0 - x_{20}^0) & -2R^0 \end{bmatrix}$$

$$\mathbf{h}_2 \atop 3,1 = \begin{bmatrix} (R^{02} - (x_{1b}^0 - x_{10}^0)^2 - (x_{2b}^0 - x_{20}^0)^2) \\ (R^{02} - (x_{1c}^0 - x_{10}^0)^2 - (x_{2c}^0 - x_{20}^0)^2) \\ (R^{02} - (x_{1d}^0 - x_{10}^0)^2 - (x_{2d}^0 - x_{20}^0)^2) \end{bmatrix}$$

Using the given approximations, the numerical values of the matrices are

$$\mathbf{D}_{11} = [2 \quad 10 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0]$$
$$\mathbf{D}_{12} = [-2 \quad -10 \quad -9]$$
$$\mathbf{h}_1 = -10$$

$$\mathbf{D}_{21} = \begin{bmatrix} 0 & 0 & 0 & -10 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -6 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -6 & 4 \end{bmatrix}$$

$$\mathbf{D}_{22} = \begin{bmatrix} 0 & 10 & -8 \\ 6 & -2 & -8 \\ 6 & -4 & -8 \end{bmatrix} \qquad \mathbf{h}_2 = \begin{bmatrix} -9 \\ 6 \\ 3 \end{bmatrix} \qquad c_{11} = 2$$

$$\mathbf{C}_{12} = [10 \quad 0 \quad 13.33 \quad -74.0 \quad 24.67 \quad 72.0 \quad -48.0] \qquad g = 40$$

The **B** matrix, accordingly partitioned, becomes

$$
\mathbf{B_1} =
\begin{bmatrix}
-100 \\
100 \\
100 \\
0 \\
0 \\
0 \\
0 \\
0 \\
0 \\
0 \\
0 \\
0
\end{bmatrix}
\qquad
\mathbf{B_2} =
\begin{bmatrix}
14.1 & 0 & 0 & 0 & 0 & 0 & 0 \\
6.1 & 0 & 0 & 0 & 0 & 0 & 0 \\
22.1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -100 & 16.6 & 0 & 0 & 0 & 0 \\
0 & 100 & 7.1 & 0 & 0 & 0 & 0 \\
0 & 100 & 26.3 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -100 & 6.5 & 0 & 0 \\
0 & 0 & 0 & 100 & 15.2 & 0 & 0 \\
0 & 0 & 0 & 100 & 32.5 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -100 & 8.3 \\
0 & 0 & 0 & 0 & 0 & 100 & 12.3 \\
0 & 0 & 0 & 0 & 0 & 100 & 28.7
\end{bmatrix}
$$

Then

$$
\mathbf{\Delta_2} = [-1.24835 \quad -0.01922 \quad 0.98883 \quad -1.02657 \quad -0.74275
$$
$$
{}_{7,\,1}
$$

$$
0.05950 \quad 0.42954]^t
$$

$$
\mathbf{\Delta_1} = -1.00598
$$

$$
\mathbf{\Delta'} =
\begin{bmatrix}
-0.23980 \\
-0.15638 \\
-0.30651
\end{bmatrix}
$$

Adding these corrections to the given approximations, the solution is iterated until we get the final answer as

$$\hat{x}_{1a} = 6.994 \text{ m} \qquad \hat{x}_{1b} = 6.983 \text{ m} \qquad \hat{x}_{1c} = 2.992 \text{ m}$$

$$\hat{x}_{2a} = 49.774 \text{ m} \qquad \hat{x}_{2b} = 41.980 \text{ m} \qquad \hat{x}_{2c} = 46.166 \text{ m}$$

$$\hat{x}_{1d} = 4.043 \text{ m} \qquad \hat{x}_{10} = 6.879 \text{ m} \qquad \hat{R} = 3.898 \text{ m}$$

$$\hat{x}_{2d} = 48.552 \text{ m} \qquad \hat{x}_{20} = 45.877 \text{ m}$$

As far as precision estimation is concerned, $\mathbf{Q_{AA}}$ was computed according to equation (9.29), and $\mathbf{Q_{A'A'}}$ according to equation (9.68), and both matrices were identical to those computed previously.

Example 9.7 Piecewise Polynomial Surface Fitting. In a practical application using side-looking airborne radar (SLAR) imagery to determine horizontal positions of terrain points, an interpolation method was used applying piecewise polynomials. The imagery covered a rectangular area with 21 control points. Two "pieces" of polynomial surfaces, the first covering 10 control points and the second 11 control points, were chosen. At the junction between pieces, the x coordinate is constant.

The polynomials in X have coefficients that are independent from those in Y and are also slightly different.

$$X = a_0 + a_1 x + a_2 x^2 + a_3 y + a_4 xy + a_5 x^2 y$$
$$Y = b_0 + b_1 x + b_2 x^2 + b_3 y + b_4 xy + b_5 y^2$$

In these equations X and Y represent the control coordinates; x and y are the coordinates in the imagery either directly from measurements or after it has been subjected to preliminary transformation (for example, four parameter); and $a_0, \ldots, a_5$; $b_0, \ldots, b_5$ are the transformation parameters. The two surfaces represented by the two pieces must have the same value for the functions at the joining line. Since that line is of constant x, then for the X function (the subscript outside the parenthesis refers to the number of the "piece"),

$$(a_0 + a_1 x + a_2 x^2)_1 = (a_0 + a_1 x + a_2 x^2)_2 \qquad \text{for the constant term}$$
$$(a_3 + a_4 x + a_5 x^2)_1 = (a_3 + a_4 x + a_5 x^2)_2 \qquad \text{for the term in } y$$

and for the Y function

$$(b_0 + b_1 x + b_2 x^2)_1 = (b_0 + b_1 x + b_2 x^2)_2 \qquad \text{for the constant term}$$
$$(b_3 + b_4 x)_1 \qquad = (b_3 + b_4 x)_2 \qquad \text{for the term in } y$$
$$(b_5)_1 \qquad = (b_5)_2 \qquad \text{for the term in } y^2$$

with x in these equations being equal to the constant coordinate at the junction between the two pieces (see Tables 9-3 and 9-4).

The stochastic model for this problem will be rather simple and will be expressed by attaching a residual to each of the two polynomial equations for each point. This implies that the coordinates in the ground control system will be considered as observations. All observations will be assumed uncorrelated and of equal precision.

TABLE 9-3   Piece 1

		IMAGE COORDINATES		GROUND COORDINATES	
i	POINT NUMBER	x	y	X	Y
1	1	1042.9	1080.7	1556.2	4146.0
2	5	1011.9	1171.3	1016.6	7900.1
3	9	972.5	1261.7	21.9	11485.2
4	10	1080.7	1076.8	3269.6	3787.9
5	28	1069.1	1155.5	3443.0	6953.0
6	38	1052.4	1273.5	1273.5	11499.0
7	39	1137.8	1137.8	5826.8	3463.0
8	51	1192.2	1079.5	8293.9	3447.4
9	54	1132.0	1188.1	6504.1	7962.7
10	61	1139.0	1271.4	7430.7	11042.6

TABLE 9-4   Piece 2

i	POINT NUMBER	IMAGE COORDINATES		GROUND COORDINATES	
		x	y	X	Y
11	69	1176.9	1170.4	8302.0	7075.7
12	78	1195.7	1267.5	9840.7	10755.3
13	81	1298.0	1082.9	13130.4	3257.7
14	91	1279.5	1200.9	13165.4	7935.2
15	95	1282.9	1243.5	13677.5	9562.8
16	100	1371.3	1231.8	17771.5	8763.4
17	113	1405.7	1257.7	19662.7	9514.6
18	115	1380.8	1096.8	17113.6	3432.4
19	131	1458.5	1071.4	20383.2	2028.9
20	146	1444.0	1139.8	20366.9	4803.9
21	157	1404.4	1236.9	19393.5	8715.8

Solution: Because of the simplified stochastic assumptions, the pair of condition equations for each given point is directly linear and may be written in the following form:

For the x coordinate

$$v_{x_i} + \begin{bmatrix} 1 & x_i & x_i^2 & y_i & x_i y_i & x_i^2 y_i \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix}_j = X_i$$

$j = 1$ (first piece)

$i = 1, 2, \ldots, 10$ (10 points in first piece)

For the y coordinate

$$v_{y_i} + \begin{bmatrix} 1 & x_i & x_i^2 & y_i & x_i y_i & y_i^2 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{bmatrix}_j = Y_i$$

$j = 2$ (second piece)

$i = 11, 12, \ldots, 21$ (11 points in second piece)

both of which are of the form $\mathbf{v} + \mathbf{B\Delta} = \mathbf{f}$.

To enforce the conditions at the joining line, two constraint equations are written for the X coordinate and three for the Y coordinate. Denoting by x_c ($= 1176$) the constant coordinate at the junction, then

For the X coordinate

$$
\begin{bmatrix}
1 & x_c & x_c^2 & 0 & 0 & 0 & -1 & -x_c & -x_c^2 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & x_c & x_c^2 & 0 & 0 & 0 & -1 & -x_c & -x_c^2
\end{bmatrix}
\begin{bmatrix}
a_{01} \\ a_{11} \\ a_{21} \\ a_{31} \\ a_{41} \\ a_{51} \\ a_{02} \\ a_{12} \\ a_{22} \\ a_{32} \\ a_{42} \\ a_{52}
\end{bmatrix}
=
\begin{bmatrix} 0 \\ 0 \end{bmatrix}
$$

For the Y coordinate

$$
\begin{bmatrix}
1 & x_c & x_c^2 & 0 & 0 & 0 & -1 & -x_c & -x_c^2 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & x_c & 0 & 0 & 0 & 0 & -1 & -x_c & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1
\end{bmatrix}
\begin{bmatrix}
b_{01} \\ b_{11} \\ b_{21} \\ b_{31} \\ b_{41} \\ b_{51} \\ b_{02} \\ b_{12} \\ b_{22} \\ b_{32} \\ b_{42} \\ b_{52}
\end{bmatrix}
=
\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}
$$

Both sets of constraint equations are of the general form $\mathbf{C\Delta} = \mathbf{g}$. Thus there are 20 condition equations for piece 1, 22 condition equations for piece 2, and 5 constraint equations. The total number of parameters is obviously 24, thus leading to the degrees of freedom,

$$r = (20 + 22) + 5 - 24 = 23$$

There are, of course, several ways in which this problem may be solved. Using any one of these [for example, the one carrying the Lagrange multipliers $\mathbf{k}_c$ as we used in programming this problem, see equation (9.33b)] leads to the same answer. Because

of the relatively large numbers involved, it was found prudent to premultiply all coordinates x and y by 10^{-3}. The answers are

Piece 1 $(j = 1)$

$$X\ Coordinate \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix}_1 = \begin{bmatrix} -5.04 \\ -57.5 \\ 52.2 \\ -40.9 \\ 101.0 \\ -51.1 \end{bmatrix} \qquad Y\ Coordinate \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{bmatrix}_1 = \begin{bmatrix} -30.5 \\ -27.4 \\ 8.09 \\ 56.0 \\ 4.37 \\ -9.46 \end{bmatrix}$$

Piece 2 $(j = 2)$

$$X\ Coordinate \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix}_2 = \begin{bmatrix} -150.0 \\ 204.0 \\ -65.2 \\ 99.3 \\ -151.0 \\ 62.0 \end{bmatrix} \qquad Y\ Coordinate \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{bmatrix}_2 = \begin{bmatrix} -59.7 \\ 14.4 \\ -6.35 \\ 63.0 \\ -1.57 \\ -9.46 \end{bmatrix}$$

The values of these parameters may now be used in the original polynomial equations to compute new sets of coordinates X and Y for each point. These will in general be different from those originally given, and the differences can be used as an indication of how well the fitting process has been accomplished.

9.5. SUMMARY OF EQUATIONS

When parameters involved in the mathematical model of an adjustment are functionally dependent, constraint equations must be written to account for that dependency. Thus with n observations in l and u parameters in Δ, we write c conditions and s constraints.

$$\mathbf{Av + B\Delta = f} \tag{9.3a}$$

$$\mathbf{C\Delta = g} \tag{9.3b}$$

$\mathbf{f}$ is $c \times 1$ and $\mathbf{g}$ is $s \times 1$, both are vectors of constants.
$\mathbf{A}$ is $c \times n$ with rank $= c$.
$\mathbf{v}$ is $n \times 1$ vector of residuals.
$\mathbf{B}$ is $c \times u$ with rank = smaller of c and u.
Δ is $u \times 1$ vector of parameters having $(u - s)$ independent elements.
$\mathbf{C}$ is $s \times u$ with rank $= s$.

9.5.1. When $u < c$

Direct Solution:

$$\mathbf{Q}_e = \mathbf{AQA}^t \qquad \mathbf{W}_e = \mathbf{Q}_e^{-1} \tag{9.9}$$

$$N = B^t W_e B \qquad t = B^t W_e f$$

$$\Delta^0 = N^{-1} t \qquad M = (CN^{-1}C^t)$$

$$\Delta = \Delta^0 + \delta\Delta \qquad \delta\Delta = N^{-1}C^t M^{-1}(g - C\Delta^0)$$

$$\hat{\sigma}_0^2 = (v^t W v)/r$$

$$Q_{\Delta\Delta} = N^{-1}[I - C^t(CN^{-1}C^t)^{-1}CN^{-1}] \tag{9.15b}$$

$$Q_{\Delta\Delta} C^t = 0 \tag{9.16}$$

$$v = QA^t W_e(-B\Delta + f) \tag{9.8), (9.9}$$

$$\hat{l} = l + v$$

$$Q_{vv} = QA^t W_e AQ - QA^t W_e BQ_{\Delta\Delta} B^t W_e AQ \tag{9.17b}$$

$$Q_{ll} = Q - Q_{vv}$$

9.5.2. Solution by Constraint Elimination

$$a = u - s$$

Partition

$$B = \begin{bmatrix} B_1 & B_2 \\ {\scriptstyle c,s} & {\scriptstyle c,a} \end{bmatrix} \quad C = \begin{bmatrix} C_{11} & C_{12} \\ {\scriptstyle s,s} & {\scriptstyle s,a} \end{bmatrix} \quad \text{with } C_{11} \text{ nonsingular}$$
$$\scriptstyle c,u \qquad\qquad\qquad s,u$$

$$\tag{9.19b}$$

$$\bar{B} = -B_1 C_{11}^{-1} C_{12} + B_2 \qquad \bar{f} = f - B_1 C_{11}^{-1} g \tag{9.21}$$
$$\scriptstyle c,a \qquad\qquad\qquad\qquad c,1$$

$$\bar{N} = \bar{B}^t W_e \bar{B} \qquad\qquad \bar{t} = \bar{B}^t W_e \bar{f} \tag{9.24}$$

$$\Delta_2 = \bar{N}^{-1}\bar{t} \tag{9.23}$$

$$\Delta_1 = C_{11}^{-1}(g - C_{12}\Delta_2) \tag{9.20}$$

$$\Delta = \begin{bmatrix} \Delta_1 \\ \Delta_2 \end{bmatrix}$$

$$Q_{\Delta\Delta} = \begin{bmatrix} C_{11}^{-1}C_{12}\bar{N}^{-1}C_{12}^t(C_{11}^{-1})^t & -C_{11}^{-1}C_{12}\bar{N}^{-1} \\ -\bar{N}^{-1}C_{12}^t(C_{11}^{-1})^t & \bar{N}^{-1} \end{bmatrix} \tag{9.29}$$

9.5.3. When $u > c$

$N = B^t W_e B$ is here a singular matrix,
$t = B^t W_e f.$

$$\begin{bmatrix} \Delta \\ k_c \end{bmatrix} = \begin{bmatrix} -N & C^t \\ C & 0 \end{bmatrix}^{-1} \begin{bmatrix} -t \\ g \end{bmatrix} = \begin{bmatrix} \alpha & \beta^t \\ \beta & \gamma \end{bmatrix} \begin{bmatrix} -t \\ g \end{bmatrix} \tag{9.33b), (9.34}$$

(k_c is $s \times 1$ vector of Lagrange multipliers for the constraints.)

$$Q_{\Delta\Delta} = -\alpha \tag{9.42}$$

All other relations for the remainder of the direct solution as well as the solution by constraint elimination for the case of $u < c$ apply equally here.

9.5.4. Constraints with Added Parameters

If in addition to the u parameters in Δ, a set of q added parameters in Δ' appear in the constraints, the number of constraints becomes s' such that

$$c + s' = r + (u + q) \qquad (r = \text{redundancy}) \tag{9.48}$$

$$q < s' < (u + q) \tag{9.47}, \ (9.49)$$

The conditions and constraints are of the form

$$\mathbf{Av} + \mathbf{B\Delta} = \mathbf{f}$$

$$\mathbf{D_1\Delta} + \mathbf{D_2\Delta'} = \mathbf{h}$$

$\mathbf{h}$ is $s' \times 1$ vector of constants
$\mathbf{D_1}$ is $s' \times u$ with rank $= s$
$\mathbf{D_2}$ is $s' \times q$ with rank $= q$

When $u < c$ all relations up to computing Δ^0 apply here. Then

$$\mathbf{P} = \mathbf{D_1 N^{-1} D_1^t} \tag{9.58}$$

$$\mathbf{R} = \mathbf{D_2^t P^{-1} D_2} \tag{9.59}$$

$$\Delta' = \mathbf{R}^{-1}[\mathbf{D_2^t P^{-1}(h - D_1\Delta^0)}] \tag{9.56}$$

$$\Delta = \Delta^0 + \mathbf{N^{-1} D_1^t P^{-1}(h - D_1\Delta^0 - D_2\Delta')} \tag{9.57}$$

$$\mathbf{Q_{\Delta'\Delta'}} = \mathbf{R}^{-1} \tag{9.60}$$

$$\mathbf{Q_{\Delta\Delta}} = \mathbf{N^{-1}[I - D_1^t P^{-1} D_1 N^{-1} + D_1^t P^{-1} D_2 R^{-1} D_2^t P^{-1} D_1 N^{-1}]} \tag{9.61}$$

$$\mathbf{Q_{\Delta\Delta'}} = -\mathbf{N^{-1} D_1^t P^{-1} D_2 R^{-1}} \tag{9.64}$$

$\mathbf{Q_{vv}}$ and $\mathbf{Q_{ll}}$ are the same as given in Section 9.5.1.

9.5.5. Elimination of Constraints with Added Parameters

With $s = s' - q$ partition the constraints into

$$\mathbf{D_1}_{s',u} = \begin{bmatrix} \mathbf{D_{11}} \\ {}_{s,u} \\ \mathbf{D_{21}} \\ {}_{q,u} \end{bmatrix} \qquad \mathbf{D_2}_{s',q} = \begin{bmatrix} \mathbf{D_{12}} \\ {}_{s,q} \\ \mathbf{D_{22}} \\ {}_{q,q} \end{bmatrix} \qquad \mathbf{h}_{s',u} = \begin{bmatrix} \mathbf{h_1} \\ {}_{s,1} \\ \mathbf{h_2} \\ {}_{q,1} \end{bmatrix} \tag{9.66}$$

such that $\mathbf{D_{22}}$ is nonsingular.

$$\mathbf{C} = (\mathbf{D_{11} - D_{12} D_{22}^{-1} D_{21}}) \tag{9.67a}$$

$$\mathbf{g} = (\mathbf{h_1 - D_{12} D_{22}^{-1} h_2}) \tag{9.67b}$$

With C and g, partition and solve as in Section 9.5.2 to get Δ_1 and Δ_2, thus the total parameter vector Δ. Then

$$\Delta' = D_{22}^{-1}(h_2 - D_{21}\Delta) \tag{9.66c}$$

$Q_{\Delta\Delta}$ as in equation (9.29)

$$Q_{\Delta'\Delta'} = D_{22}^{-1}D_{21}Q_{\Delta\Delta}\,D_{21}^{t}D_{22}^{-1} \tag{9.68}$$

10

Adjustment with Derived Observations and Adjustment in Steps (of Observations Only)

10.1. INTRODUCTION

In earlier discussions it was shown that observational variables need not be the direct numerical outcomes of the physical operations of measurement. In fact, as was shown by Example 7.5, we could use in the adjustment other stochastic vaiables that are derived from the originally available variables. The outcome is exactly the same *provided the stochastic model is modified in such a way that the stochastic properties of the derived observations are also properly obtained*. This rather important concept will be discussed further and demonstrated by examples in the first part of this chapter.

The second part of the chapter will deal with another related concept. It involves what is termed " adjustment in steps " when the technique applied is that of adjusting observations only. In many practical situations a set of observations is obtained that belongs to a certain (geometrical) model. This first set is then adjusted and estimates for the observations are obtained from the adjustment. At a later stage, a second set of observations may be acquired that belongs to either the same model or to an expanded form of it. Thus the adjustment of the second step would involve in it not only the

second set of observations but also some or all of the *new estimates* of the observations in the first set.

By necessity, then, the adjustment is performed in two steps and the final result must be the same as would have been obtained had we performed the total adjustment of all sets of observations together. This will be demonstrated both algebraically and by numerical example.

10.2. ADJUSTMENT WITH DERIVED OBSERVATIONS

Consideration is limited to linear functions (which may be originally linear or may have been linearized). Also only the case of adjustment of observations is considered here, whereas other possibilities, which include parameters, will be treated in Part III of this book.

With n observations and r redundancy the conditions may be written as

$$\underset{r,\,n}{\mathbf{A}_l} \underset{n,\,1}{\mathbf{v}_l} = (-\mathbf{A}_l l + \mathbf{d}_l) = \underset{r,\,1}{\mathbf{f}_l} \tag{10.1}$$

with $\mathbf{d}_l$ and $\mathbf{f}_l$ being constant term vectors. Associated with l is the cofactor matrix $\mathbf{Q}_{ll}$.† If an adjustment is carried out using these conditions, the following relations may be written (refer to Chapter 7):

$$\mathbf{k}_l = (\mathbf{A}_l \mathbf{Q}_{ll} \mathbf{A}_l^t)^{-1}\mathbf{f}_l = \mathbf{W}_{e_l} \mathbf{f}_l \tag{7.4}$$

$$\mathbf{v}_l = \mathbf{Q}_{ll} \mathbf{A}_l^t \mathbf{k}_l = \mathbf{Q}_{ll} \mathbf{A}_l^t \mathbf{W}_{e_l} \mathbf{f}_l \tag{7.5}$$

$$\mathbf{Q}_{v_l v_l} = \mathbf{Q}_{ll} \mathbf{A}_l^t \mathbf{W}_{e_l} \mathbf{A}_l \mathbf{Q}_{ll} \tag{7.7}$$

$$\mathbf{Q}_{\hat{l}\hat{l}} = \mathbf{Q}_{ll} - \mathbf{Q}_{v_l v_l} \tag{6.31}$$

Suppose now that we are to *derive* a new set of observations $\mathbf{m}_{p,\,1}$ from the original set l, such that $p \le n$. This new set can be used in the adjustment in place of the original set with results consistent with those that would have been obtained from l, provided the following assertions are considered:

1. The number of derived observations is less than or equal to the number of original observations, that is, $p \le n$.
2. The cofactor matrix $\underset{p,\,p}{\mathbf{Q}_{mm}}$ for the new set $\mathbf{m}$ must be properly evaluated.

Thus if

$$\underset{p,\,1}{\mathbf{m}} = \underset{p,\,n}{\mathbf{D}} \underset{n,\,1}{l} \quad \text{and} \quad \mathbf{v}_m = \mathbf{D}\mathbf{v}_l \tag{10.2}$$

then

$$\mathbf{Q}_{mm} = \mathbf{D}\mathbf{Q}_{ll}\mathbf{D}^t \tag{10.3}$$

must be computed.

---

† The subscript l is reimposed here to facilitate the comparison with the derived set applying the subscript m.

3. Redundancy, or the number of degrees of freedom r, *must remain the same* after obtaining the derived observations. In other words the model must not change while switching over from one set of observations to another if we are to get consistent adjustment results.

Because of point (1) above, we may write the conditions in terms of the derived observations m as follows:

$$\mathop{\mathbf{A}_m}_{r,\,p} \mathop{\mathbf{v}_m}_{p,\,1} = (-\mathbf{A}_m\,\mathbf{m} + \mathbf{d}_m) = \mathop{\mathbf{f}_m}_{r,\,1} \qquad (10.4)$$

Using a vector $\mathbf{f}_m$ that is different from $\mathbf{f}_l$ allows for the possibility of using a different set of conditions, even though the number of conditions is the same. For the obvious reason that any r conditions chosen are independent, then

$$\mathop{\mathbf{f}_m}_{r,\,1} = \mathop{\mathbf{C}}_{r,\,r} \mathop{\mathbf{f}_l}_{r,\,1} \qquad (10.5)$$

with $\mathbf{C}$ being nonsingular.

Applying least squares to equation (10.4) leads directly to the following relations:

$$\mathop{\mathbf{k}_m}_{r,\,1} = \mathbf{W}_{e_m}\,\mathbf{f}_m = (\mathbf{A}_m\,\mathbf{Q}_{mm}\,\mathbf{A}_m^t)^{-1}\mathbf{f}_m \qquad (10.6a)$$

$$\mathbf{v}_m = \mathbf{Q}_{mm}\,\mathbf{A}_m^t\,\mathbf{k}_m = \mathbf{Q}_{mm}\,\mathbf{A}_m^t\,\mathbf{W}_{e_m}\,\mathbf{f}_m \qquad (10.6b)$$

$$\mathbf{Q}_{v_m v_m} = \mathbf{Q}_{mm}\,\mathbf{A}_m^t\,\mathbf{W}_{e_m}\,\mathbf{A}_m\,\mathbf{Q}_{mm} \qquad (10.6c)$$

$$\mathbf{Q}_{\hat{m}\hat{m}} = \mathbf{Q}_{mm} - \mathbf{Q}_{v_m v_m} \qquad (10.6d)$$

Equations (10.6) give the least squares solution in terms of the new set of observations m and their a priori cofactor matrix $\mathbf{Q}_{mm}$. We must now compute the residuals $\mathbf{v}_l$ and the a posteriori matrix $\mathbf{Q}_{v_l v_l}$ from these, and also prove that they are the same values that would have been obtained if l and $\mathbf{Q}_{ll}$ were used directly.

From equations (10.1), (10.2), (10.4), and (10.5) it follows that

$$\mathbf{A}_m\,\mathbf{D}\mathbf{v}_l = \mathbf{C}\mathbf{f}_l$$

$$(\mathbf{C}^{-1}\mathbf{A}_m\,\mathbf{D})\mathbf{v}_l = \mathbf{f}_l$$

which leads to

$$\mathbf{A}_l = \mathbf{C}^{-1}\mathbf{A}_m\,\mathbf{D} \qquad (10.7)$$

From equations (7.4, 10.3, 10.5, and 10.6a) we write

$$\begin{aligned}
\mathbf{k}_l &= (\mathbf{A}_l\,\mathbf{Q}_{ll}\,\mathbf{A}_l^t)^{-1}\mathbf{f}_l \\
&= [(\mathbf{C}^{-1}\mathbf{A}_m\,\mathbf{D})\mathbf{Q}_{ll}(\mathbf{D}^t\mathbf{A}_m(\mathbf{C}^{-1})^t)]^{-1}\mathbf{f}_l \\
&= [\mathbf{C}^{-1}\mathbf{A}_m\,\mathbf{Q}_{mm}\,\mathbf{A}_m^t(\mathbf{C}^t)^{-1}]^{-1}\mathbf{f}_l \\
&= \mathbf{C}^t(\mathbf{A}_m\,\mathbf{Q}_{mm}\,\mathbf{A}_m^t)^{-1}\mathbf{C}\mathbf{f}_l \\
&= \mathbf{C}^t\mathbf{W}_{e_m}\,\mathbf{f}_m
\end{aligned}$$

or

$$\mathbf{k}_l = \mathbf{C}^t \mathbf{k}_m \tag{10.8}$$

Now, to derive a relation for $\mathbf{v}_l$ in terms of $\mathbf{v}_m$,

$$\mathbf{v}_l = \mathbf{Q}_{ll} \mathbf{A}_l^t \mathbf{k}_l$$

$$= \mathbf{Q}_{ll} \mathbf{D}^t \mathbf{A}_m^t (\mathbf{C}^{-1})^t \mathbf{C}^t \mathbf{k}_m$$

or

$$\mathbf{v}_l = \mathbf{Q}_{ll} \mathbf{D}^t \mathbf{A}_m^t \mathbf{k}_m \tag{10.9a}$$

or, with reference to equation (10.6b)

$$\mathbf{v}_l = \mathbf{Q}_{ll} \mathbf{D}^t \mathbf{Q}_{mm}^{-1} \mathbf{v}_m \tag{10.9b}$$

It should be straightforward to show that premultiplying both sides of equation (10.9b) by $\mathbf{D}$ yields equation (10.2), taking equation (10.3) into account.

Using equation (10.9b) we can derive a relationship for the a posteriori cofactor matrix $\mathbf{Q}_{v_l v_l}$.

$$\mathbf{Q}_{v_l v_l} = (\mathbf{Q}_{ll} \mathbf{D}^t \mathbf{Q}_{mm}^{-1}) \mathbf{Q}_{v_m v_m} (\mathbf{Q}_{ll} \mathbf{D}^t \mathbf{Q}_{mm}^{-1})^t$$

and from equation (10.6c),

$$\mathbf{Q}_{v_l v_l} = \mathbf{Q}_{ll} \mathbf{D}^t \mathbf{Q}_{mm}^{-1} \mathbf{Q}_{mm} \mathbf{A}_m^t \mathbf{W}_{e_m} \mathbf{A}_m \mathbf{Q}_{mm} \mathbf{Q}_{mm}^{-1} \mathbf{D} \mathbf{Q}_{ll}$$

or

$$\mathbf{Q}_{v_l v_l} = \mathbf{Q}_{ll} \mathbf{D}^t \mathbf{A}_m^t \mathbf{W}_{e_m} \mathbf{A}_m \mathbf{D} \mathbf{Q}_{ll} \tag{10.10}$$

Finally, it must be shown that equation (10.10) is equivalent to equation (7.7).

$$\mathbf{Q}_{v_l v_l} = \mathbf{Q}_{ll} \mathbf{D}^t \mathbf{A}_m^t (\mathbf{A}_m \mathbf{D} \mathbf{Q}_{ll} \mathbf{D}^t \mathbf{A}_m^t)^{-1} \mathbf{A}_m \mathbf{D} \mathbf{Q}_{ll}$$

and from equation (10.7),

$$\mathbf{Q}_{v_l v_l} = \mathbf{Q}_{ll} \mathbf{D}^t \mathbf{A}_m^t (\mathbf{C} \mathbf{A}_l \mathbf{Q}_{ll} \mathbf{A}_l^t \mathbf{C}^t)^{-1} \mathbf{A}_m \mathbf{D} \mathbf{Q}_{ll}$$

$$= \mathbf{Q}_{ll} [\mathbf{D}^t \mathbf{A}_m^t (\mathbf{C}^{-1})^t] (\mathbf{A}_l \mathbf{Q}_{ll} \mathbf{A}_l^t)^{-1} (\mathbf{C}^{-1} \mathbf{A}_m \mathbf{D}) \mathbf{Q}_{ll}$$

or

$$\mathbf{Q}_{v_l v_l} = \mathbf{Q}_{ll} \mathbf{A}_l^t \mathbf{W}_{e_l} \mathbf{A}_l^t \mathbf{Q}_{ll}$$

which is identical to equation (7.7). To show that

$$\mathbf{v}_l^t \mathbf{W}_{ll} \mathbf{v}_l = \mathbf{v}_m^t \mathbf{W}_{mm} \mathbf{v}_m$$

we refer to equations (10.9b) and (10.3).

$$\mathbf{v}_l^t \mathbf{W}_{ll} \mathbf{v}_l = (\mathbf{v}_m^t \mathbf{Q}_{mm}^{-1} \mathbf{D} \mathbf{Q}_{ll}) \mathbf{W}_{ll} (\mathbf{Q}_{ll} \mathbf{D}^t \mathbf{Q}_{mm}^{-1} \mathbf{v}_m)$$

$$= \mathbf{v}_m^t \mathbf{Q}_{mm}^{-1} (\mathbf{D} \mathbf{Q}_{ll} \mathbf{D}^t) \mathbf{Q}_{mm}^{-1} \mathbf{v}_m$$

$$= \mathbf{v}_m^t \mathbf{Q}_{mm}^{-1} \mathbf{v}_m = \mathbf{v}_m^t \mathbf{W}_{mm} \mathbf{v}$$

To demonstrate the relations derived above some numerical examples follow.

Example 10.1. As a first example consider the simple case of a plane triangle ABC in Figure 10.1. Six directions are measured, two at each corner. From these six directions, which may be considered as the original observations, three angles may be computed as derived observations. It is required to show that from the adjustment
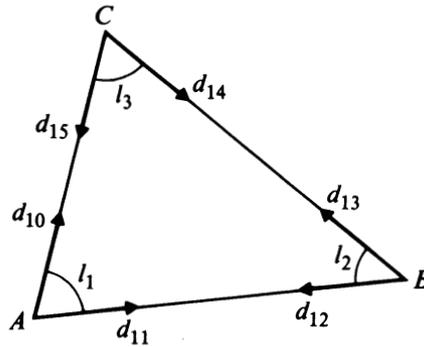

Figure 10.1

with the derived observations (angles), residuals for the original directions may be computed as well as their cofactor matrix which are the same as those obtained from adjustment with the directions themselves. Assume that the cofactor matrix for the directions is the identity matrix and that the triangle's misclosure (contradiction or discrepancy) is w seconds of arc.

Solution: For the sake of later checks, the triangle is first adjusted using directions. Since there are three zero directions, one at each corner, and since two variables are necessary for determining the shape of the triangle, then $r = 1$. This is the same as adjusting by angles, which satisfies requirement (3) in Section 10.2. From this it is obvious that there is only one f, or $f_l = f_m = w$ which leads to $C = I$, in equation (10.5).

The condition for adjustment by directions is

$$[-1 \quad 1 \quad -1 \quad 1 \quad -1 \quad 1]v_d = w \qquad \text{or} \qquad \underset{1,6}{A_d} \underset{6,1}{v_d} = \underset{1,1}{f}$$

with

$$Q_{dd} = I \qquad Q_{e_d} = 6 \qquad k_d = k = \frac{w}{6} \qquad v_d^t = [-1 \quad 1 \quad -1 \quad 1 \quad -1 \quad 1]\frac{w}{6}$$

$$Q_{v_d v_d} = \tfrac{1}{6} \begin{bmatrix} 1 & -1 & 1 & -1 & 1 & -1 \\ & 1 & -1 & 1 & -1 & 1 \\ & & 1 & -1 & 1 & -1 \\ & & & 1 & -1 & 1 \\ & & & & 1 & -1 \\ \text{symmetric} & & & & & 1 \end{bmatrix}$$

To get the angles

$$\begin{bmatrix} l_1 \\ l_2 \\ l_3 \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} d_{10} \\ d_{11} \\ d_{12} \\ d_{13} \\ d_{14} \\ d_{15} \end{bmatrix} \quad \text{or} \quad l = Dd$$

and

$$Q_{ll} = DQ_{dd}\,D^t = 2I_3$$

The condition equation is

$$\begin{bmatrix} 1 & 1 & 1 \end{bmatrix} v_l = w$$

$$Q_{e_m} = 6 \qquad k_l = k = \frac{w}{6}$$

$$V_l^t = \frac{w}{3}\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$$

$$Q_{v_l v_l} = \tfrac{2}{3}\begin{bmatrix} 1 & 1 & 1 \\ & 1 & 1 \\ \text{symmetric} & & 1 \end{bmatrix}$$

Now compute v_d and $Q_{v_d v_d}$ from v_l, $Q_{v_l v_l}$, and other matrices belonging to the adjustment. From equation (10.9a),

$$v_d = Q_{dd}D^tA_l^t k_l = \begin{bmatrix} -1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\frac{w}{6} = w/6\begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \\ -1 \\ 1 \end{bmatrix}$$

Then from equation (10.10),

$$Q_{v_d v_d} = Q_{dd}\,D^tA_l^t\,W_{e_l}\,A_l\,DQ_{dd} = \tfrac{1}{6}\begin{bmatrix} 1 & -1 & 1 & -1 & 1 & -1 \\ & 1 & -1 & 1 & -1 & 1 \\ & & 1 & -1 & 1 & -1 \\ & & & 1 & -1 & 1 \\ & & & & 1 & -1 \\ \text{symmetric} & & & & & 1 \end{bmatrix}$$

Both v_d and $Q_{v_d v_d}$ are the same as was obtained previously. This verifies the relations derived in the present section.

Example 10.2. In this example the results are used from the adjustment in Example 7.5 of a quadrilateral by angles and directions in order to verify further the applicability of equations (10.9) and (10.10).

Solution: Using equation (10.9b)

$$\underset{12,12}{\mathbf{v}_d} = \underset{12,12}{\mathbf{Q}_{dd}} \quad \underset{12,8}{\mathbf{D}^t} \quad \underset{8,8}{\mathbf{Q}_{ll}^{-1}} \quad \underset{8,1}{\mathbf{v}_l}$$

where

$$\mathbf{Q}_{dd} = \mathbf{I}_{12} \quad \text{and} \quad \mathbf{D} = \text{matrix} \underset{8,12}{\mathbf{F}} \text{ in Example 7.5}$$

$$\underset{8,8}{\mathbf{Q}_{ll}^{-1}} = \tfrac{1}{3} \begin{bmatrix}
2 & 1 & 0 & & & \text{(zeros)} & & \\
1 & 2 & 0 & 0 & & & & \\
0 & 0 & 2 & 1 & 0 & & & \\
& 0 & 1 & 2 & 0 & & & \\
& & 0 & 0 & 2 & 1 & 0 & \\
& & & 0 & 1 & 2 & 0 & 0 \\
& & & & 0 & 0 & 2 & 1 \\
& \text{(zeros)} & & & & & 0 & 1 & 2 \\
\end{bmatrix}$$

$$\mathbf{v}_l = [-1.0122 \quad 0.9822 \quad -0.3378 \quad 1.7137$$
$$1.7620 \quad -0.3722 \quad 1.1065 \quad -1.1022]^t$$

$$\mathbf{v}_d = \tfrac{1}{3} \begin{bmatrix}
-2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -2 & -1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 2 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 2 & -1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 2 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -2 & -1 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 \\
\end{bmatrix} \begin{bmatrix}
-1.0122 \\
0.9822 \\
-0.3378 \\
1.7137 \\
1.7620 \\
-0.3722 \\
1.1065 \\
-1.1022 \\
\end{bmatrix} = \begin{bmatrix}
0.3474 \\
-0.6648 \\
0.3174 \\
-0.3460 \\
-0.6838 \\
1.0299 \\
-1.0506 \\
0.7114 \\
0.3392 \\
-0.3703 \\
0.7362 \\
-0.3660 \\
\end{bmatrix}$$

which is exactly the same vector obtained from direct adjustment of directions (see Example 7.5). With regard to equation (10.10), compute $\mathbf{Q}_{v_d v_d}$ as follows:

$$\mathbf{Q}_{v_d v_d} = \mathbf{Q}_{ll} \mathbf{D}^t \mathbf{A}_l^t \mathbf{W}_{el} \mathbf{A}_l \mathbf{D} \mathbf{Q}_{ll} = \underset{12,8}{\mathbf{D}} \; \underset{8,4}{\mathbf{A}_l^t} \; \underset{4,4}{\mathbf{W}_{el}} \; \underset{4,8}{\mathbf{A}_l} \; \underset{8,12}{\mathbf{D}}$$

where

$$\underset{4,8}{\mathbf{A}_l} = \begin{bmatrix}
1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\
1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
2.4716 & -3.3244 & 0.6457 & -2.8156 & 0.0602 & -0.6062 & 3.7025 & -0.1037 \\
\end{bmatrix}$$

$$
\mathbf{W}_{e_l}_{4,4} = \begin{bmatrix} 0.2773 & 0.1490 & -0.1765 & 0.0195 \\ & 0.2712 & -0.1703 & 0.0171 \\ & & 0.3471 & -0.0367 \\ \text{symmetric} & & & 0.0139 \end{bmatrix}
$$

The value of $\mathbf{D}^t \mathbf{A}_l^i$ is first computed

$$
\mathbf{D}^t \mathbf{A}_l^i = \begin{bmatrix} -1 & 0 & -1 & -2.4716 \\ 0 & 0 & 1 & 5.7960 \\ 1 & 0 & 0 & -3.3244 \\ 1 & 0 & 0 & -0.6457 \\ 1 & -1 & 0 & 3.4613 \\ 0 & 1 & 0 & -2.8156 \\ 0 & 1 & 0 & -0.0602 \\ 0 & 0 & -1 & 0.6644 \\ 0 & 1 & 1 & -0.6062 \\ 0 & -1 & -1 & -3.7026 \\ -1 & 1 & 0 & 3.8062 \\ 1 & 0 & 1 & -0.1037 \end{bmatrix}
$$

This is equal to $\mathbf{A}_d^t$, and since $\mathbf{W}_{e_d} = \mathbf{W}_{e_l}$, the computed $\mathbf{Q}_{v_d v_d}$ will be exactly the same as that computed from the adjustment of directions and need not be repeated.

Example 10.3. Refer to Figure 10.1 of the plane triangle and consider that the angle A is measured twice (l_1 and l_2), whereas angles B and C are each measured once l_3 and l_4, respectively. The four measurements, which are uncorrelated and of equal precision, are

$$l_1 = 60°00'$$

$$l_2 = 60°02'$$

$$l_3 = 60°00'$$

$$l_4 = 60°01'$$

It is required to adjust the triangle, first using all four observations and then using derived observations, and to check the consistency of both results.

Solution: The model for this problem requires a minimum of two variables; thus $r = 2$, and two conditions are needed as follows:

$$l_1 - l_2 = 0$$

$$l_1 + l_3 + l_4 - \pi = 0$$

or

$$\begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & 1 & 1 \end{bmatrix} \mathbf{v}_l = \begin{bmatrix} 2' \\ -1' \end{bmatrix} \quad \text{with } \mathbf{Q}_{ll} = \mathbf{I}_4$$

Thus

$$\mathbf{k}_l = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}^{-1} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \tfrac{1}{5} \begin{bmatrix} 7 \\ -4 \end{bmatrix}$$

$$\mathbf{v}_l = \tfrac{1}{5} \begin{bmatrix} 1 & 1 \\ -1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 7 \\ -4 \end{bmatrix} = \tfrac{1}{5} \begin{bmatrix} 3' \\ -7' \\ -4' \\ -4' \end{bmatrix} \qquad \mathbf{Q}_{v_l v_l} = \tfrac{1}{5} \begin{bmatrix} 3 & -2 & 1 & 1 \\ -2 & 3 & 1 & 1 \\ 1 & 1 & 2 & 2 \\ 1 & 1 & 2 & 2 \end{bmatrix}$$

It is possible to first derive one "observation" from l_1 and l_2. Since both are of equal precision, then the derived observations are

$$\begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix} = \begin{bmatrix} \tfrac{1}{2} & \tfrac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} l_1 \\ l_2 \\ l_3 \\ l_4 \end{bmatrix} \qquad \text{or} \qquad \mathbf{m} = \mathbf{D}l$$

where m_1 is the arithmetic mean of l_1 and l_2. With these three observations, we may now adjust the triangle. However, with only three observations, we would have only one redundancy. This means that in the process of obtaining the derived observations we also have reduced the redundancy from two to one. This makes it no longer a technique using derived observations, and thus a treatment using adjustment in steps is warranted instead. First adjust l_1 and l_2, and in the second part use the adjusted values as derived observations. This is an adjustment in steps, as will be explained in the next section.

Adjusting l_1, l_2 first we get

$$v_1 = +1' \qquad v_2 = -1'$$

$$\hat{l}_1 = 60°01' \qquad \hat{l}_2 = 60°01'$$

and

$$\hat{\mathbf{Q}} = \begin{bmatrix} \tfrac{1}{2} & \tfrac{1}{2} \\ \tfrac{1}{2} & \tfrac{1}{2} \end{bmatrix}$$

Now we work with the "derived" observations $\mathbf{m}$ noting that we should use $\hat{l}_1$ and $\hat{l}_2$ instead of l_1 and l_2.

Hence

$$\mathbf{Q}_{mm}_{3,3} = \begin{bmatrix} \tfrac{1}{2} & \tfrac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \tfrac{1}{2} & \tfrac{1}{2} & 0 & 0 \\ \tfrac{1}{2} & \tfrac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \tfrac{1}{2} & 0 & 0 \\ \tfrac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \tfrac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and the condition $\mathbf{A}_m \mathbf{V}_m = \mathbf{f}_m$ is

$$[1 \quad 1 \quad 1]\mathbf{V}_m = 180 - 60°01' - 60°00' - 60°01' = -2'$$

$$k_m = \frac{2}{5}(-2) = -\frac{4'}{5}$$

$$\mathbf{V}_m = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}(-\tfrac{4}{5}) = \tfrac{1}{5}\begin{bmatrix} -2 \\ -4 \\ -4 \end{bmatrix}$$

Now, in order to compute $\mathbf{v}_{l'}$ apply equation (10.9a)

$$\mathbf{v}_{l'} = \mathbf{Q}_{l'l'}\,\mathbf{D}^t\mathbf{A}_m^t k_m = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}(-\tfrac{4}{5}) = \tfrac{1}{5}\begin{bmatrix} -2 \\ -2 \\ -4 \\ -4 \end{bmatrix}$$

Note the first two corrections $-\frac{2}{5}$ and $-\frac{2}{5}$ are to be applied to $\hat{l}_1$ and $\hat{l}_2$, respectively, after the first step of adjustment, and *not* to the original observations l_1 and l_2. In order to get the total residuals, add the contribution from the first step to that from the adjustment with reduced observations, or

$$\mathbf{v}_l = \begin{bmatrix} 1' \\ -1' \\ 0 \\ 0 \end{bmatrix} + \tfrac{1}{5}\begin{bmatrix} -2' \\ -2' \\ -4' \\ -4' \end{bmatrix} = \tfrac{1}{5}\begin{bmatrix} 3' \\ -7' \\ -4' \\ -4' \end{bmatrix}$$

which is identical to that obtained directly from the adjustment of all four observations. Next, we should evaluate and check the cofactor matrices of the estimated observations. (Check the derivations in the next section to ascertain this point.) Thus first compute

$$\mathbf{Q}_{ll} \text{ from } \mathbf{Q}_{v_l v_l}$$
$$\scriptstyle 4, 4$$

evaluated in the first part

$$\mathbf{Q}_{ll} = \mathbf{I} - \tfrac{1}{5}\begin{bmatrix} 3 & -2 & 1 & 1 \\ & 3 & 1 & 1 \\ & & 2 & 2 \\ \text{symmetric} & & & 2 \end{bmatrix} = \tfrac{1}{5}\begin{bmatrix} 2 & 2 & -1 & -1 \\ & 2 & -1 & -1 \\ & & 3 & -2 \\ \text{symmetric} & & & 3 \end{bmatrix}$$

Then compute $\mathbf{Q}_{v_{l'}v_{l'}}$ from the second part (with derived observations) using equation (10.10), or

$$\mathbf{Q}_{v_{l'}v_{l'}} = \mathbf{Q}_{l'l'}\,\mathbf{D}^t\mathbf{A}_m^t\,\mathbf{W}_{e_m}\,\mathbf{A}_m\,\mathbf{D}\mathbf{Q}_{l'l'}$$

$$\mathbf{Q}_{v_{l'}v_{l'}} = \tfrac{1}{10}\begin{bmatrix} 1 & 1 & 2 & 2 \\ & 1 & 2 & 2 \\ & & 4 & 4 \\ \text{symmetric} & & & 4 \end{bmatrix}$$

Finally,

$$\mathbf{Q}_{ll} = \tfrac{1}{3} \begin{bmatrix} 2 & 2 & -1 & -1 \\ & 2 & -1 & -1 \\ & & 3 & -2 \\ \text{symmetric} & & & 3 \end{bmatrix}$$

This value is identical to that computed by the other procedure. A last and *important* remark: It is worth noting that although it may not have been readily evident, it was possible to propagate backward from a smaller set of variables (for example, angles) to a larger set of variables (directions).

10.3. ADJUSTMENT IN STEPS

We begin adjustment in steps by having a total set of $(r + s)$ condition equations

$$\underset{(r+s),\, n}{\mathbf{A}} \left(l + \underset{n,\, 1}{\mathbf{v}} \right) = \underset{(r+s),\, 1}{\mathbf{d}}$$

amongst the observations $l_{n,\,1}$, the cofactor matrix of which is $\mathbf{Q}$. The solution for this total system is given by equations (7.4), (7.5), and (7.7) in Section 10.2. Due to the manner in which the conditions arise, the total system of conditions is partitioned into two systems, the first being

$$\underset{r,\, n}{\mathbf{A}_1} \left(l + \underset{n,\, 1}{\mathbf{v}_1} \right) = \underset{r,\, 1}{\mathbf{d}_1} \tag{10.11}$$

Using the equations just referred to, we get $\mathbf{v}_1$, $\mathbf{Q}_{v_1 v_1}$, and $\mathbf{Q}_2 = \mathbf{Q} - \mathbf{Q}_{v_1 v_1}$ as the solution for the first system in the first step. Next, the second system of conditions is

$$\underset{s,\, n}{\mathbf{A}_2} \left(l + \underset{n,\, 1}{\mathbf{v}_1 + \mathbf{v}_2} \right) = \underset{s,\, 1}{\mathbf{d}_2} \tag{10.12}$$

for which the cofactor matrix is now $\mathbf{Q}_2$. Again using the same equations, solve for $\mathbf{v}_2$ and $\mathbf{Q}_{v_2 v_2}$ and compute

$$\bar{\mathbf{Q}}_{ll} = \mathbf{Q}_2 - \mathbf{Q}_{v_2 v_2} \tag{10.13}$$

The answer obtained from the total system of conditions may be obtained from the adjustment in steps as follows:

$$\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2 \tag{10.14a}$$

$$\mathbf{Q}_{ll} = \bar{\mathbf{Q}}_{ll} \tag{10.14b}$$

$$\phi = \phi_1 + \phi_2 \quad \text{(the quadratic forms)} \tag{10.14c}$$

10.3.1. Proof that $v = v_1 + v_2$

The total system may be partitioned to

$$\begin{bmatrix} A_1 \\ A_2 \end{bmatrix}(l + v) = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$$

or

$$A_1 v = d_1 - A_1 l = f_1 \tag{10.15a}$$

$$A_2 v = d_2 - A_2 l = f_2 \tag{10.15b}$$

Now with regard to the adjustment in steps, from the first step:

$$A_1 v_1 = d_1 - A_1 l = f_1 \tag{10.16a}$$

and

$$Q_{v_1 v_1} = QA_1^t(A_1 QA_1^t)^{-1}A_1 Q \tag{10.16b}$$

$$Q_2 = Q - QA_1^t(A_1 QA_1^t)^{-1}A_1 Q \tag{10.16c}$$

For the second step

$$A_2(l + v_1 + v_2) = d_2$$

$$A_2(v_1 + v_2) = d_2 - A_2 l \tag{10.17}$$

From equations (10.17) and (10.15) it follows that

$$v = v_1 + v_2$$

provided that the consistency between equations (10.15a) and (10.16a) is ascertained. For such a consistency to stand we must prove that

$$A_1 v_2 = 0$$

Solving equation (10.17) for v_2

$$v_2 = Q_2 A_2^t(A_2 Q_2 A_2^t)^{-1}(d_2 - A_2 l - A_2 v_1)$$
$$= Q_2 A_2^t(A_2 Q_2 A_2^t)^{-1}f_2'$$

and using equation (10.16c)

$$v_2 = [Q - QA_1^t(A_1 QA_1^t)^{-1}A_1 Q]A_2^t(A_2 Q_2 A_2^t)^{-1}f_2'$$

Now premultiplying by A_1

$$A_1 v_2 = [A_1 Q - (A_1 QA_1^t)(A_1 QA_1^t)^{-1}A_1 Q]A_2^t(A_2 Q_2 A_2^t)^{-1}f_2'$$
$$= 0$$

thus proving the validity of equation (10.14a).

10.3.2. Proof that $Q_{ll} = \bar{Q}_{ll}$ We first evaluate Q_{ll} from the total system after it is partitioned.

$$Q_{ll} = Q - QA^t(AQA^t)^{-1}AQ$$

for which we expand

$$(AQA^t)^{-1} = \left\{ \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} Q[A_1^t A_2^t] \right\}^{-1} = \begin{bmatrix} A_1 QA_1^t & A_1 QA_2^t \\ A_2 QA_1^t & A_2 QA_2^t \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} \dot{M} & \bar{M} \\ \bar{M}^t & \ddot{M} \end{bmatrix}^{-1} = \begin{bmatrix} E & G \\ G^t & H \end{bmatrix}$$

where

$$H = (\ddot{M} - \bar{M}^t \dot{M}^{-1} \bar{M})^{-1} \tag{10.18a}$$

$$G = -\dot{M}^{-1} \bar{M} H \tag{10.18b}$$

$$E = \dot{M}^{-1} - \dot{M}^{-1} \bar{M} G^t \tag{10.18c}$$

Thus

$$Q_{ll} = Q - Q[A_1^t A_2^t] \begin{bmatrix} E & G \\ G^t & H \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} Q$$

or

$$Q_{ll} = Q - QA_1^t EA_1 Q - QA_2^t G^t A_1 Q - QA_1^t GA_2 Q - QA_2^t HA_2 Q$$

$$= Q - QA_1^t(\dot{M}^{-1} - \dot{M}^{-1}\bar{M}G^t)A_1 Q - QA_2^t G^t A_1 Q$$

$$- QA_1^t GA_2 Q - QA_2^t HA_2 Q$$

or

$$Q_{ll} = Q - \underset{(2)}{(QA_1^t \dot{M}^{-1} A_1 Q)} + \underset{(3)}{(QA_1^t \dot{M}^{-1}\bar{M}G^t A_1 Q)}$$

$$- \underset{(4)}{(QA_2^t G^t A_1 Q)} - \underset{(5)}{(QA_1^t GA_2 Q)}$$

$$- \underset{(6)}{(QA_2^t HA_2 Q)} \tag{10.19}$$

Then we also evaluate $\bar{Q}_{ll}$ as follows:

$$\bar{Q}_{ll} = Q_2 - Q_2 A_2^t(A_2 Q_2 A_2^t)^{-1}A_2 Q_2$$

$$= \underset{(1)}{Q} - \underset{(2)}{QA_1^t \dot{M}^{-1} A_1 Q}$$

$$- (Q - QA_1^t \dot{M}^{-1} A_1 Q)A_2^t(A_2 Q_2 A_2^t)^{-1}A_2(Q - QA_1^t \dot{M}^{-1} A_1 Q)$$

$$\tag{10.20}$$

We evaluate the following term first:

$$(\mathbf{A}_2 \, \mathbf{Q}_2 \, \mathbf{A}_2^t)^{-1} = [\mathbf{A}_2(\mathbf{Q} - \mathbf{Q}\mathbf{A}_1^t \, \dot{\mathbf{M}}^{-1}\mathbf{A}_1\mathbf{Q})\mathbf{A}_2^t]^{-1}$$

$$= [(\mathbf{A}_2 \, \mathbf{Q}\mathbf{A}_2^t) - (\mathbf{A}_2 \, \mathbf{Q}\mathbf{A}_1^t)\dot{\mathbf{M}}^{-1}(\mathbf{A}_1\mathbf{Q}\mathbf{A}_2^t)]^{-1}$$

$$= (\ddot{\mathbf{M}} - \bar{\mathbf{M}}^t\dot{\mathbf{M}}^{-1}\bar{\mathbf{M}})^{-1}$$

$$= \mathbf{H} \qquad\qquad (10.21)$$

Since the first two terms are the same in both equation (10.19) and equation (10.20), we consider the last term in equation (10.20), which is denoted by t, using the just derived result.

$$\mathbf{t} = \underset{(6)}{-\mathbf{Q}\mathbf{A}_2^t \, \mathbf{H} \mathbf{A}_2 \, \mathbf{Q}} + \mathbf{Q}\mathbf{A}_2^t \, \mathbf{H}\mathbf{A}_2 \, \mathbf{Q}\mathbf{A}_1^t \, \dot{\mathbf{M}}^{-1}\mathbf{A}_1\mathbf{Q} + \mathbf{Q}\mathbf{A}_1^t \, \dot{\mathbf{M}}^{-1}\mathbf{A}_1\mathbf{Q}\mathbf{A}_2^t$$

$$\times \mathbf{H}\mathbf{A}_2 \, \mathbf{Q} - \mathbf{Q}\mathbf{A}_1^t \, \dot{\mathbf{M}}^{-1}\mathbf{A}_1\mathbf{Q}\mathbf{A}_2^t \, \mathbf{H}\mathbf{A}_2 \, \mathbf{Q}\mathbf{A}_1^t \, \dot{\mathbf{M}}^{-1}\mathbf{A}_1\mathbf{Q}$$

Dropping the first term since it is the same as the sixth term of equation (10.19), we continue with

$$\mathbf{t}' = \mathbf{Q}\mathbf{A}_2^t \, \mathbf{H}\bar{\mathbf{M}}^t\dot{\mathbf{M}}^{-1}\mathbf{A}_1\mathbf{Q} + \mathbf{Q}\mathbf{A}_1^t\dot{\mathbf{M}}^{-1}\bar{\mathbf{M}}\mathbf{H}\mathbf{A}_2 \, \mathbf{Q}$$

$$- \mathbf{Q}\mathbf{A}_1^t \, \dot{\mathbf{M}}^{-1}\bar{\mathbf{M}}\mathbf{H}\bar{\mathbf{M}}^t\dot{\mathbf{M}}^{-1}\mathbf{A}_1\mathbf{Q}$$

$$= \underset{(4)}{-\mathbf{Q}\mathbf{A}_2^t \, \mathbf{G}^t\mathbf{A}_1\mathbf{Q}} - \underset{(5)}{\mathbf{Q}\mathbf{A}_1^t\mathbf{G}\mathbf{A}_2 \, \mathbf{Q}} + \underset{(3)}{\mathbf{Q}\mathbf{A}_1\dot{\mathbf{M}}^{-1}\bar{\mathbf{M}}\mathbf{G}^t\mathbf{A}_1\mathbf{Q}} \qquad (10.22)$$

The first term of equation (10.21) is the same as the fourth term in equation (10.19), the second is the same as the fifth, and the last term is the same as the third, respectively. This proves that both $\mathbf{Q}_{ll}$ and $\bar{\mathbf{Q}}_{ll}$ can be expanded to six terms that are identical. Hence

$$\mathbf{Q}_{ll} = \bar{\mathbf{Q}}_{ll}$$

10.3.3. Proof that $\phi = \phi_1 + \phi_2$ In order to compute an a posteriori estimate of the reference variance, the quadratic form $\phi = \mathbf{v}^t\mathbf{W}\mathbf{v}$ must be computed. For the computation of ϕ we use equation (7.6) or

$$\phi = \mathbf{k}^t\mathbf{f} = \mathbf{f}^t(\mathbf{A}\mathbf{Q}\mathbf{A}^t)^{-1}\mathbf{f}$$

From equations (10.15b) and (10.17) we can readily write

$$\mathbf{f}_2' = (\mathbf{d}_2 - \mathbf{A}_2 l) - \mathbf{A}_2 \, \mathbf{v}_1 = \mathbf{f}_2 - \mathbf{A}_2 \, \mathbf{v}_1 \qquad\qquad (10.23)$$

and we also recall that

$$\mathbf{v}_1 = \mathbf{Q}\mathbf{A}_1^t(\mathbf{A}_1\mathbf{Q}\mathbf{A}_1^t)^{-1}\mathbf{f}_1 = \mathbf{Q}\mathbf{A}_1^t \, \dot{\mathbf{M}}^{-1}\mathbf{f}_1 \qquad\qquad (10.24)$$

$$\phi = \mathbf{f}^t \begin{bmatrix} \dot{\mathbf{M}} & \bar{\mathbf{M}} \\ \bar{\mathbf{M}}^t & \ddot{\mathbf{M}} \end{bmatrix}^{-1} \mathbf{f} = [\mathbf{f}_1^t \quad \mathbf{f}_2^t] \begin{bmatrix} \mathbf{E} & \mathbf{G} \\ \mathbf{G}^t & \mathbf{H} \end{bmatrix} \begin{bmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \end{bmatrix}$$

$$= \underset{(1)}{(\mathbf{f}_1^t\dot{\mathbf{M}}^{-1}\mathbf{f}_1} - \underset{(2)}{\mathbf{f}_1^t\dot{\mathbf{M}}^{-1}\bar{\mathbf{M}}\mathbf{G}^t\mathbf{f}_1)} + \underset{(3)}{2\mathbf{f}_1^t\mathbf{G}\mathbf{f}_2} + \underset{(4)}{\mathbf{f}_2^t\mathbf{H}\mathbf{f}_2} \qquad (10.25)$$

and

$$\phi_1 = \mathbf{f}_1^t(\mathbf{A}_1\mathbf{Q}\mathbf{A}_1^t)^{-1}\mathbf{f}_1 = \mathbf{f}_1^t\dot{\mathbf{M}}^{-1}\mathbf{f}_1 \tag{10.26}$$

finally,

$$\phi_2 = \mathbf{f}_2''(\mathbf{A}_2\,\mathbf{Q}_2\,\mathbf{A}_2^t)^{-1}\mathbf{f}_2' \tag{10.27}$$

which according to equations (10.21), (10.23), and (10.24) becomes

$$\phi_2 = \mathbf{f}_2''\mathbf{H}\mathbf{f}_2'$$

$$= (\mathbf{f}_2^t - \mathbf{v}_1^t\mathbf{A}_2^t)\mathbf{H}(\mathbf{f}_2 - \mathbf{A}_2\,\mathbf{v}_1)$$

$$= \mathbf{f}_2^t\,\mathbf{H}\mathbf{f}_2 - 2\mathbf{v}_1^t\mathbf{A}_2^t\,\mathbf{H}\mathbf{f}_2 + \mathbf{v}_1^t\mathbf{A}_2^t\,\mathbf{H}\mathbf{A}_2\,\mathbf{v}_1$$

$$= \mathbf{f}_2^t\,\mathbf{H}\mathbf{f}_2 - 2\mathbf{f}_1^t\dot{\mathbf{M}}^{-1}(\mathbf{A}_1\mathbf{Q}\mathbf{A}_2^t)\mathbf{H}\mathbf{f}_2$$

$$\qquad + \mathbf{f}_1^t\dot{\mathbf{M}}^{-1}(\mathbf{A}_1\mathbf{Q}\mathbf{A}_2^t)\mathbf{H}(\mathbf{A}_2\,\mathbf{Q}\mathbf{A}_1^t)\dot{\mathbf{M}}^{-1}\mathbf{f}_1$$

$$= \mathbf{f}_2^t\,\mathbf{H}\mathbf{f}_2 - 2\mathbf{f}_1^t\dot{\mathbf{M}}^{-1}\bar{\mathbf{M}}\mathbf{H}\mathbf{f}_2 + \mathbf{f}_1^t\dot{\mathbf{M}}^{-1}\bar{\mathbf{M}}(\mathbf{H}\bar{\mathbf{M}}^t\dot{\mathbf{M}}^{-1})\mathbf{f}_1$$

$$\phi_2 = -\mathbf{f}_1^t\dot{\mathbf{M}}^{-1}\bar{\mathbf{M}}\mathbf{G}^t\mathbf{f}_1 + 2\mathbf{f}_1^t\mathbf{G}\mathbf{f}_2 + \mathbf{f}_2^t\,\mathbf{H}\mathbf{f}_2 \tag{10.28}$$

It is obvious from equations (10.25), (10.26), and (10.28) that

$$\phi = \phi_1 + \phi_2$$

Example 10.4. Figure 10.2 shows two adjacent plane triangles where six angles are measured. These are uncorrelated and of equal precision and have the following values:

$l_1 = 60°00'03''$
$l_2 = 60°00'02''$
$l_3 = 60°00'01''$
$l_4 = 120°00'05''$
$l_5 = 120°00'05''$
$l_6 = 60°00'01''$
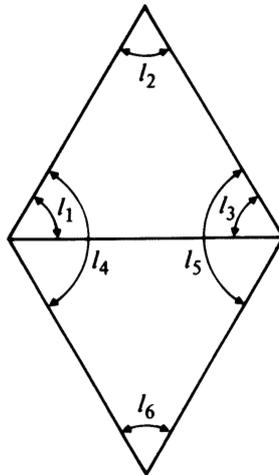
Figure 10.2

It is required to adjust the whole figure with all six observations together. We adjust the figure in two steps: First, using angles l_1, l_2, and l_3, and in the second step, using the remaining data. Then we verify the consistency of the answers from both procedures.

Solution

The Direct Adjustment. There are 2 degrees of freedom (the reader should verify this) for which we write the following two independent conditions:

$$l_1 + l_2 + l_3 - 180 = 0$$

$$l_2 + l_4 + l_5 + l_6 - 360 = 0$$

or

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix} \mathbf{v} = \begin{bmatrix} 180 - l_1 - l_2 - l_3 \\ 360 - l_2 - l_4 - l_5 - l_6 \end{bmatrix} = \begin{bmatrix} -6'' \\ -12 \end{bmatrix}$$

$$\mathbf{k} = \begin{bmatrix} 3 & 1 \\ 1 & 4 \end{bmatrix}^{-1} \begin{bmatrix} -6 \\ -12 \end{bmatrix} = \begin{bmatrix} -1.09'' \\ -2.73'' \end{bmatrix}$$

$$\mathbf{v} = \mathbf{Q}\mathbf{A}^t\mathbf{k} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1.09'' \\ -2.73'' \end{bmatrix} = \begin{bmatrix} -1.09'' \\ -3.82'' \\ -1.09'' \\ -2.73'' \\ -2.73'' \\ -2.73'' \end{bmatrix}$$

$$\mathbf{Q}_{vv} = \tfrac{1}{11} \begin{bmatrix} 4 & 3 & 4 & -1 & -1 & -1 \\ & 5 & 3 & 2 & 2 & 2 \\ & & 4 & -1 & -1 & -1 \\ & & & 3 & 3 & 3 \\ & & & & 3 & 3 \\ \text{symmetric} & & & & & 3 \end{bmatrix}$$

$$\mathbf{Q}_{ll} = \mathbf{I} - \mathbf{Q}_{vv} = \tfrac{1}{11} \begin{bmatrix} 7 & -3 & -4 & 1 & 1 & 1 \\ & 6 & -3 & -2 & -2 & -2 \\ & & 7 & 1 & 1 & 1 \\ & & & 8 & -3 & -3 \\ & & & & 8 & -3 \\ \text{symmetric} & & & & & 8 \end{bmatrix}$$

Adjustment in Steps

First Step: Here there is only one condition,

$$l_1 + l_2 + l_3 - 180 = 0$$

$$[1 \quad 1 \quad 1 \quad 0 \quad 0 \quad 0]\mathbf{v}_1 = -6''$$

$$k_1 = 2'' \qquad \mathbf{v}_1 = [-2'' \quad -2'' \quad -2'' \quad 0 \quad 0 \quad 0]^t$$

$$\mathbf{Q}_{v_1v_1} = \tfrac{1}{3} \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ & 1 & 1 & 0 & 0 & 0 \\ & & 1 & 0 & 0 & 0 \\ & & & 0 & 0 & 0 \\ & & & & 0 & 0 \\ \text{symmetric} & & & & & 0 \end{bmatrix}$$

$$\mathbf{Q}_2 = \mathbf{Q} - \mathbf{Q}_{v_1v_1} = \mathbf{I} - \mathbf{Q}_{v_1v_1} = \tfrac{1}{3} \begin{bmatrix} 2 & -1 & -1 & 0 & 0 & 0 \\ & 2 & -1 & 0 & 0 & 0 \\ & & 2 & 0 & 0 & 0 \\ & & & 3 & 0 & 0 \\ & & & & 3 & 0 \\ \text{symmetric} & & & & & 3 \end{bmatrix}$$

Second Step: The vector of observations in this step is not the one originally given at the beginning of the problem. It is updated by the residuals from the first step. Thus

$$l'_1 = 60°00''01' \qquad l'_2 = 60°00''00' \qquad l'_3 = 59°59''39'$$

$$l'_4 = 120°00''05' \qquad l'_5 = 120°00''04' \qquad l'_6 = 60°00''01'$$

The one condition for this step is

$$l'_2 + l'_4 + l'_5 + l'_6 - 360 = 0$$

$$[0 \quad 1 \quad 0 \quad 1 \quad 1 \quad 1]\mathbf{v}_2 = 360 - l'_2 - l'_4 - l'_5 - l'_6 = -10''$$

Note that $f'_2 = -10''$ is not the same as $f_2 = -12''$. The difference as given by equation (10.23) is $(-\mathbf{A}_2\,\mathbf{v}_1) = +2''$, which is correct. Thus

$$\mathbf{k}_2 = (\mathbf{A}_2\,\mathbf{Q}_2\,\mathbf{A}_2)^{-1}(-10) = -\left(\frac{30}{11}\right)''$$

$$\mathbf{v}_2 = \mathbf{Q}_2\,\mathbf{A}_2\,\mathbf{k}_2$$

$$= [0.91 \quad -1.82 \quad 0.91 \quad -2.73 \quad -2.73 \quad -2.73]^t \qquad \text{(sec of arc)}$$

and

$$\mathbf{v}_1 + \mathbf{v}_2 = [-1.09 \quad -3.82 \quad -1.09$$

$$-2.73 \quad -2.73 \quad -2.73]^t \qquad \text{(sec of arc)}$$

which is identical to v computed in the direct adjustment.

$$\mathbf{Q}_{v2v2} = \tfrac{1}{33} \begin{bmatrix} 1 & -2 & 1 & -3 & -3 & -3 \\ & 4 & -2 & 6 & 6 & 6 \\ & & 1 & -3 & -3 & -3 \\ & & & 9 & 9 & 9 \\ & & & & 9 & 9 \\ \text{symmetric} & & & & & 9 \end{bmatrix}$$

$$\bar{\mathbf{Q}}_{ll} = \mathbf{Q}_2 - \mathbf{Q}_{v2v2} = \tfrac{1}{11} \begin{bmatrix} 7 & -3 & -4 & 1 & 1 & 1 \\ & 6 & -3 & -2 & -2 & -2 \\ & & 7 & 1 & 1 & 1 \\ & & & 8 & -3 & -3 \\ \text{symmetric} & & & & 8 \end{bmatrix}.$$

which is again identical to $\mathbf{Q}_{ll}$ computed in the direct adjustment.

It may be concluded that both adjustment with derived observations and adjustment in steps are feasible provided we make sure of satisfying the requirements of each concept. Most important of these are not changing the model when adjusting with derived observations and making the necessary modification of the stochastic model for both concepts of adjustment.

11

Numerical and Statistical Considerations in Adjustment

This chapter is devoted to several topics that were mentioned but not discussed in detail in the foregoing chapters. Such topics pertain mostly to practical matters of adjustment and include selection of approximations for model variables, criteria for termination of iterations for nonlinear cases, data editing for gross errors, and computational aspects of adjustment, and so on. The problem of linearization of nonlinear equations is discussed first.

11.1. NONLINEARITY OF THE EQUATIONS

Nonlinear equations arise more often in practice than those that are directly linear. Therefore the adjustment problem quite often involves nonlinear estimation. The least squares criterion can, in principle, be applied directly to nonlinear equations but with the consequence that the resulting normal equations would also be nonlinear. Such nonlinear equations may pose considerable difficulties in seeking a unique solution. Furthermore, in cases of large adjustment problems, computational efforts may become so extensive as to render the solution impractical. To demonstrate this point and the possible complexity of this approach, a simple example is given.

274

Example 11.1. In this example reconsider the problem of Example 7.2 and refer to Figure 6.1 for the layout of the three terrestrial photographs. In that example there was 1 degree of freedom, leading to one condition equation in terms of the observations only, of the form

$$\mathbf{F}(l) = -l_1 l_5 - l_2 l_4 - l_2 l_5 + l_3 l_4 = 0$$

where l_i represents the five measurements as shown in Figure 6.1. In Example 7.2 (as in all other treatments) this nonlinear equation was first linearized by Taylor series expansion before the least squares criterion was applied. Here we go directly into the criterion with the nonlinear condition, thus in effect performing nonlinear least squares adjustment. In this case the quadratic form to be minimized is

$$\phi = \mathbf{v}'\mathbf{W}\mathbf{v} - 2\mathbf{k}_n^t(-l_1 l_5 - l_2 l_4 - l_2 l_5 + l_3 l_4)$$

where $\mathbf{k}_n$ is one Lagrange multiplier for the nonlinear condition. Replacing l_i by $(l_i + v_i)$ and remembering (from Example 7.2) that $\mathbf{W}$ is a 5×5 diagonal matrix with elements w_i, the quadratic form ϕ may be expanded to

$$\begin{aligned}
\phi = {}& (w_1 v_1^2 + w_2 v_2^2 + w_3 v_3^2 + w_4 v_4^2 + w_5 v_5^2) \\
& + 2k_n[(l_1 + v_1)(l_5 + v_5) \\
& + (l_2 + v_2)(l_4 + v_4) + (l_2 + v_2)(l_5 + v_5) \\
& - (l_3 + v_3)(l_4 + v_4)]
\end{aligned}$$

To fulfill the minimum criterion, differentiate ϕ with respect to each v_i and equate each partial derivative to zero. Thus

$$\begin{aligned}
w_1 v_1 + k_n(l_5 + v_5) &= 0 \\
w_2 v_2 + k_n[(l_4 + v_4) + (l_5 + v_5)] &= 0 \\
w_3 v_3 - k_n(l_4 + v_4) &= 0 \\
w_4 v_4 + k_n[(l_2 + v_2) - (l_3 + v_3)] &= 0 \\
w_5 v_5 + k_n[(l_1 + v_1) + (l_2 + v_2)] &= 0
\end{aligned}$$

These are five equations that, when combined with the one condition, result in *six nonlinear* normal equations in six unknowns; the unknowns being the five residuals and k_n. It must be recognized that the solution of this set of nonlinear equations is not as easily obtainable as for linear equations, as their solution requires more involved numerical schemes. Another point that is worth mentioning is the fact that here there are *six* equations to solve simultaneously, whereas in the linearized case the v's were easily eliminated from the six equations and *one* normal equation remained in terms of k. Finally, if the weight matrix were a full matrix, the expression for ϕ would have included many more terms thus resulting in more complicated nonlinear equations. This change in weight matrix, on the other hand, would have less effect on the normal equations in the linearized case.

These remarks point out the relative complexity of nonlinear least squares when compared to the linear case. This is the reason that in practice we invariably linearize first, then apply the least squares method.

11.1.1. The Linearization Concept Section 5.4 included a brief description of linearization by Taylor series expansion using zero and first-order terms and dropping second- and higher-order terms. In Appendix B the complete series for one-dimensional and multidimensional functions are given. These are introduced in terms of functional variables without regard to their classification in different groups when used in adjustment. Here, on the other hand, linearization is discussed as it is applied in adjustment techniques.

It has been customary in practice to linearize the conditions at the given observations and at approximations for the parameters. This means that for cases of adjustment of observations only the solution is not iterated, and for all other cases the solution is iterated on the parameters only. Although this procedure may be totally practical in most cases, it is not theoretically sound. Of course, we must admit that in many situations the given a priori values of the observations are sufficiently close to the final estimates that no iterations are needed. Notwithstanding such cases we should, in principle, treat all variables in the model equally and perform the linearization and iteration on all. There are situations in which failure to linearize and iterate on the observable variables leads to incorrect results. As an example, in a unified approach to least squares (see Chapter 12 in Part III) when parameters are treated strictly as observations, lack of proper iteration causes the method to fail. Although the given observations will, as usual, be denoted by l, an approximate vector that varies from one iteration to another will be denoted by l^0. Furthermore, the vector of residuals will be retained as v, but a vector of corrections Δl will also be necessary. For the parameters x, the customary x^0 will be used for approximations and Δ for corrections. On the basis of these symbols the linearization schemes for each of the different techniques of adjustment are given.

11.1.2. Linearization Form for General Case of Adjustment with Conditions Only $F(l, x) = 0$ The original nonlinear conditions are expressed functionally by

$$f(l, x) = 0 \tag{11.1}$$

or

$$F[(l^0 + \Delta l), (x^0 + \Delta)] = 0 \tag{11.2}$$

Applying Taylor's series to equation (11.2), dropping second- and higher-order terms, and using the auxiliaries

$$A = \frac{\partial F}{\partial l}\bigg|_{l^0, x^0} \quad \text{and} \quad B = \frac{\partial F}{\partial x}\bigg|_{l^0, x^0} \tag{11.3}$$

the linearized form of equation (11.2) is

$$F(l^0, x^0) + A\,\Delta l + B\Delta = 0 \tag{11.4}$$

The final estimate of the observations $\hat{l}$ must be equal to the observed vaues plus the least squares residuals, as well as the approximate values plus the corrections, or

$$\hat{l} = l + v = l^0 + \Delta l \tag{11.5}$$

from which

$$\Delta l = v + (l - l^0) \tag{11.6}$$

Using this relation in equation (11.4) and letting

$$f^0 = -[F(l^0, x^0) + A(l - l^0)] \tag{11.7}$$

we get

$$Av + B\Delta = f^0 \tag{11.8}$$

Equation (11.8) is of the same form as the basic equation (6.5). The only apparent difference is in the term f^0 on the right-hand side. When the conditions are originally linear, f as given by equation (6.6) is used. If the conditions are nonlinear and we do not wish to iterate on the observations, $l = l^0$ and f^0 would reduce to

$$f = -F(l, x^0) \tag{11.9}$$

In the general case of iterating on all variables equation (11.8) is used. The details of iterating by updating approximations will be treated in a subsequent section after giving the linearization forms for the remaining techniques.

11.1.3. Linearization Form for Adjustment of Observations Only

$F(l) = 0$ In a manner similar to the general case above, we begin with the nonlinear conditions

$$F(l) = 0 \tag{11.10}$$

which in this case are in terms of observations only (no parameters). Therefore equations (11.7) and (11.8) become

$$f^0 = -[F(l^0) + A(l - l^0)] \tag{11.11}$$

$$Av = f^0 \tag{11.12}$$

If we choose to iterate on the observations, the constant term would be as in equation (11.11). Otherwise the term $A(l - l^0)$ drops out and $f^0 = -F(l^0)$ for nonlinear conditions, but the solution is performed only once.

11.1.4. Linearization Form for Adjustment of Indirect Observations

$l + F(x) = 0$ The nonlinear conditions for this technique of adjustment take the form

$$l + F(x) = 0 \tag{11.13}$$

which is obviously linear in the observations. Therefore we may write the linearized form directly as

$$\mathbf{v} + \mathbf{B\Delta} = \mathbf{f}^0 \tag{11.14}$$

$$\mathbf{f}^0 = -[l + F(\mathbf{x}^0)] \tag{11.15}$$

Note that no iteration on the observations is necessary here since the conditions in them are originally linear.

11.1.5. Linearization Form for Adjustment with Conditions and Constraints

Only the general case is considered here, as other specialized cases can be readily treated in a straightforward manner. The original nonlinear conditions and constraints are functionally expressed as

$$F(l, \mathbf{x}) = 0 \tag{11.16a}$$

$$G(\mathbf{x}) = 0 \tag{11.16b}$$

If

$$\mathbf{C} = \left.\frac{\partial G}{\partial \mathbf{x}}\right|_{\mathbf{x}^0} \tag{11.17}$$

$$\mathbf{g}^0 = -G(\mathbf{x}^0) \tag{11.18}$$

then with equations (11.3) and (11.7) the linearized form of the pair of equations in equation (11.16) becomes

$$\mathbf{Av} + \mathbf{B\Delta} = \mathbf{f}^0 \tag{11.19a}$$

$$\mathbf{C\Delta} = \mathbf{g}^0 \tag{11.19b}$$

The different cases involving $\mathbf{f}^0$ and $\mathbf{f}$ have already been discussed in Section 11.1.2. As regards $\mathbf{g}^0$, it would simply become a vector of numerical constants $\mathbf{g}$, given a priori, when the constraints are originally linear.

Before leaving this section it should be emphasized that linearization, if it is necessary, must be applied to *both* the conditions and constraints, and at the same approximation vector $\mathbf{x}^0$, even if either set is originally linear.

11.1.6. Scheme of Iteration

With the approximate vectors f_0^0 and $\mathbf{x}_0^0$ we would compute the coefficient matrices $\mathbf{A}_0$ and $\mathbf{B}_0$ (either or both depending on the case of adjustment involved). Also the constant term vector $\mathbf{f}_0^0$ (or $\mathbf{f}_0$ depending again on the adjustment case) is evaluated. With these numerical values, normal equations are formed and solved after which we may directly obtain first values for $\mathbf{\Delta}$ and $\mathbf{v}$, which may be denoted by $\mathbf{\Delta}_1$ and $\mathbf{v}_1$. This essentially concludes the computations for one iteration. In order to begin

the second iteration, updated approximations are required. These are computed from

$$\mathbf{x}_1^0 = \mathbf{x}_1^0 + \boldsymbol{\Delta}_1 \tag{11.20a}$$

$$l_1^0 = l + \mathbf{v}_1 \tag{11.20b}$$

Given l_1^0 and $\mathbf{x}_1^0$, the matrices $\mathbf{A}_1$, $\mathbf{B}_1$, and $\mathbf{f}_1^0$ (or $\mathbf{f}_1$) are evaluated anew. Another least squares solution is effected yielding $\boldsymbol{\Delta}_2$ and $\mathbf{v}_2$ as answer vectors. Then

$$\mathbf{x}_2^0 = \mathbf{x}_1^0 + \boldsymbol{\Delta}_2 \tag{11.21a}$$

$$l_2^0 = l + \mathbf{v}_2 \tag{11.21b}$$

Note that although $\boldsymbol{\Delta}_2$ is added to the previous approximation vector $\mathbf{x}_1^0$, the vector $\mathbf{v}_2$ is added to the original observational vector. This is a direct consequence of equation (11.5). In general, then, the approximations l_i^0 and $\mathbf{x}_i^0$ at the end of the ith iteration may be computed from

$$\mathbf{x}_i^0 = \mathbf{x}_{i-1}^0 + \boldsymbol{\Delta}_i \tag{11.22a}$$

$$l_i^0 = l + \mathbf{v}_i \tag{11.22b}$$

At the end of each iteration a check is performed to see if it is necessary to continue with another iteration or to terminate the process. This check is usually performed according to a present criterion of tolerance. This aspect is treated in Section 11.4.

Example 11.2. Recall the problem of the three terrestrial photographs in Example 7.2, where it was solved by the technique of adjustment of observations only. In that example the one nonlinear condition was linearized at the given observations and the solution performed only once. If we now extend the solution such that iterations are performed on observations, equations (11.11) and (11.12) should be applied. In doing so, four iterations were carried out. The vectors of residuals after the first and final iterations are

$$\mathbf{v}_1 = \begin{bmatrix} 0.043 & 642 & 1495 & (\text{mm}) \\ 0.098 & 194 & 8365 & (\text{mm}) \\ -0.054 & 552 & 6869 & (\text{mm}) \\ -0.022 & 639 & 3650 & (\text{m}) \\ 0.027 & 685 & 4886 & (\text{m}) \end{bmatrix} \quad \mathbf{v}_f = \begin{bmatrix} 0.0.43 & 775 & 4674 & (\text{mm}) \\ 0.098 & 182 & 5901 & (\text{mm}) \\ -0.054 & 407 & 1226 & (\text{mm}) \\ -0.022 & 421 & 6799 & (\text{m}) \\ 0.027 & 867 & 1859 & (\text{m}) \end{bmatrix}$$

The vector $\mathbf{v}_1$ would be identical to that obtained in Example 7.2, if that many digits were carried in the latter. In fact, 10 digits to the right of the decimal were carried there for the purpose of illustrating the exactness of the procedures derived above. To ascertain the consistency of these procedures, the same problem was solved by the method of adjustment of indirect observations. Remember that in such a case we compute only one value for the vector $\mathbf{v}$ after terminating the iterations on the parameters. A criterion for iteration termination was selected such that the last corrections $\boldsymbol{\Delta}$ must all be less than 10^{-10}. With this tolerance criterion met, *all* five residuals agreed with the elements of $\mathbf{v}_f$ to *all* 10 decimal places.

As a further illustration the scheme of direct nonlinear least squares discussed in Example 11.1 was also used. Instead of directly solving the six nonlinear equations, a much simpler procedure was applied in order to verify the correctness of v_f. The values of the residuals (in v_f) are substituted in the five equations (obtained by partial differentiation in Example 11.1) and five values for k_n are computed as follows:

$$k_n = -0.545 \quad 293 \quad 867 \quad 41$$
$$= -0.545 \quad 293 \quad 867 \quad 44$$
$$= -0.545 \quad 293 \quad 867 \quad 41$$
$$= -0.545 \quad 293 \quad 867 \quad 43$$
$$= -0.545 \quad 293 \quad 867 \quad 44$$

These values agree with each other to the tenth decimal place. When the elements of v_1 instead of v_f were used (thus representing the noniterative procedure of Example 7.2) that agreement was not attained.

To conclude this example it should be mentioned that the intention here was to carry enough digits in order to verify the theoretical considerations given earlier. In this particular example, we would not really need to iterate, since the differences between v_1 and v_f are of no significance in view of the given σ's. However, there may be situations in which failure to iterate on the observations results in serious problems.

Example 11.3. As a further illustration consider the problem of the two-parameter transformation of Example 7.3. Without undue repetition, the solution of this problem was done using iterations on the observations with the following two vectors being the residuals after the first and the final iterations:

$$v_1 = \begin{bmatrix} 0.009 & 296 & 1919 \\ 0.080 & 495 & 5370 \\ 0.014 & 389 & 2837 \\ 0.011 & 362 & 0128 \\ -0.061 & 013 & 1615 \\ -0.063 & 078 & 9819 \end{bmatrix} \quad v_f = \begin{bmatrix} 0.009 & 340 & 9445 \\ 0.078 & 360 & 1453 \\ 0.017 & 125 & 0649 \\ 0.010 & 378 & 8272 \\ -0.053 & 450 & 9601 \\ -0.054 & 488 & 8428 \end{bmatrix}$$

The problem was also programmed in the form $v + B\Delta = f^0$, the solution iterated on the parameters, and the final residual vector computed and found to agree with v_f above to even more digits than those given.

A possibility may be considered in which the conventional one-run solution is used to obtain a vector of estimates $\tilde{l}$ and a corresponding cofactor matrix Q_{ll}. We may feel that with these new values a relinearization can be performed and another adjustment solution carried out. Such a procedure is incorrect and would fail because Q_{ll} is always a singular matrix (see Appendix C). This point can be demonstrated using the data of the present example. For the first iteration

A_1 is a 4×6 matrix with a rank 4.

Q is a 6×6 matrix with a rank 6. (Original observations are given as functionally independent, thus a priori cofactor matrix is nonsingular.)

$Q_{e_1} = A_1 Q A_1'$ is a 4×4 nonsingular coefficient matrix of partially reduced normal equations (of rank 4).

After the first iteration, the computed l will be functionally dependent, which leads to a singular Q_{ll} matrix of a rank 2. Consequently, Q_{e_2} for the second iteration will have rank 2, and we cannot iterate the solution with this scheme.

11.2. APPROXIMATE VALUES FOR MODEL VARIABLES

With a nonlinear least squares problem, the procedure is to linearize the functional relationships and to apply the linear least squares algorithm repeatedly on the different levels of linearization. In order to linearize or expand a given function into power series, approximate values for all the variables in the function are required. The given a priori estimates of the observations are the best approximations possible when linearizing with respect to the observation variables. In the past, providing parameter approximations was not considered a problem. This is because in conventional applications in geodesy and astronomy it was always possible to arrive at such approximations by performing preliminary computations with a selected small subset of the measurements. At the present time, however, the problem of approximations is not at all trivial, particularly when we consider present least squares applications. They are usually fully automated, carried out on a computer, and involve extensive amounts of data. Obtaining approximations is therefore a problem of a general nature for which there is, unfortunately, no general algorithm available.

11.2.1. General Criteria for Parameter Approximations There are two main reasons for seeking close approximations: First, nonlinear problems may have several real solutions. With coarse initial approximate values, iterative relinearizations may converge, but to a solution that does not correspond to the minimum criterion of least squares. Second, it is, in general, true that the closer the approximations, the lesser is the number of iterations needed to reach the final solution. This obviously means a more economical solution, particularly for large problems.

Both of the above criteria figure more prominantly in cases in which gross data errors confound the solutions. But, in general, we usually strive to satisfy them both. However, with modern computer solutions the tendency is toward automated methods for obtaining approximate values. This added criterion of automation may in certain cases lead to the sacrifice of the second item above, at least to some extent. Of course, the ultimate result of automation is overall economy which obviously offsets any increase in the number of iterations.

11.2.2. Methods for Deriving Approximations Since there is no general algorithm for obtaining approximate values, it is usually the analysis of each particular adjustment problem that leads to an appropriate procedure for deriving approximations. As illustration, the following are several possibilities.

1. The procedure very frequently used in geodesy in the past was to perform computation with a minimum number of the given observations. This often allowed the calculation in a sequential manner of one (or more) approximate value after another. A good example is the case of a trigonometric net where we proceed to solve one triangle at a time until approximations for all the coordinates in the net are computed. In such procedures, partial adjustments may even be possible, as, for instance, in adjusting triangle closures.

This classically successful procedure can be formulated in a general principle. Redundancy is reduced to such a degree that it becomes possible to evaluate sequentially groups of unknowns as explicit functions of observations or other unknowns (or both) for which approximations have already been calculated. Such a scheme can often work quite well. In fact, for some problems it may yield approximations so close to the final estimates that only one linearized adjustment will be needed. The principal weakness of this method is that it relies on subjective judgment rather than an automated approach. In some cases, however, the subjective part can be automated.

For example, the hand method for geodetic nets can be simulated by computer techniques. In the simplest case it means operating on a few subprograms, such as intersection, resection, or traverse computation, where the sequence of operation is fed into the computer by a human operator, perhaps by making use of a graph technique. Another possibility is working out a special search program by which the computer searches by itself as to how to proceed from one point to the next, trying several possibilities in case difficulty is encountered. Here we can distinguish between two cases: whether the computer is at least told where to start, or whether it is to start at random or according to certain features that it would systematically search for.

Although such procedures work satisfactorily in most practical cases, they are by no means general. Depending on the geometrical structure of a system and on redundancy, such systems may break down in unusual cases.

A second weakness is the sensitivity of such systems to blunders in the data. Because the philosophy of such strategies is to find a straightforward way through the system, without actual adjustment, the success depends on the assumption that the observations can be relied on. If a blunder exists in the system, it can make many of the computed approximations grossly erroneous. Even if the program can detect blunders, it will usually not be able to circumvent them. All this leads to the conclusion that in well-determined (well-conditioned) systems the described strategy of automatic procedure for determination of approximate values works very well. As a general algorithm it can break down, however, or lead to unacceptable results.

With the recent progress in computer applications, the requirements for generalization and automation of a program are continuously increasing.

The ultimate goal is to look for a completely automatic system of obtaining approximate values.

2. In many cases, notably in photogrammetry, it is often sufficient simply to assume certain initial values as approximations for some of the unknowns. Such values are easily recognized because of the specific physical and geometric circumstances of the problem. As an example consider block triangulation: (a) for regular aerial photography, zero tilts are quite frequently assumed and the azimuth taken from the flight plan; (b) particularly with high-altitude photography, the flying height may be approximated by its nominally known value; (c) approximate scale may readily be evaluated from the nominal flying height and principal distance; (d) terrain may be taken as being flat, particularly for small-scale photography.

3. It is often sufficient, in certain cases, to have approximations for only one group of unknowns and compute the other group automatically in the adjustment program. A good example of this is block triangulation of bundles using the collinearity equations in analytical photogrammetry. In some systems we need only introduce input approximations pertaining to the exposure station parameters. Built into the triangulation program would be routines for computing approximations for object point coordinates. It is also possible to accept approximations for even a part of the exposure station elements and compute the rest in the program. For example, knowing the nominal direction of flight and overlap, approximations for the position of one or a few camera stations would be sufficient to compute the others in a strip, given that the angular elements are approximately zero (or known).

4. Sometimes enough information may be available to solve part of the problem first. The results from that step may then serve as approximate values for the rest of the problem. As an example, consider the situation in which auxiliary data are available for the orientation parameters of aerial photographs. With such data, preliminary block triangulation may be carried out in order to obtain approximations for the joint adjustment (of photo data and auxiliary data).

As another example, some block triangulation programs contain special distance conditions. It is a good practice to perform a preliminary block adjustment first to obtain approximate coordinates for linearizing the conditions arising from distance measurement.

Yet another example concerns orbital photography. Given orbit tracking data and time, we can perform preliminary computations in order to derive approximations for the camera parameters. Using these, the rigorous adjustment of the photographic block, together with orbital constraints, may be performed.

5. It is sometimes possible to change to a simpler mathematical model in order to obtain approximations. A good example of this is performing planimetric block triangulation using four-parameter transformation at the plate

scale. This reduces a relatively involved nonlinear case to one that is linear and does not require approximations. In fact, in some systems a sequential set of such approximate procedures is applied so that the final rigorous adjustment is only performed once or twice.

Another example concerns the eight-parameter transformation equations that represent the projectivity between two planes. This pair of fractional linear equations may be reduced to two linear equations by disregarding or otherwise reducing the denominator to unity. A unique one-iteration solution can then be effected to compute approximations for six of the eight parameters. Using these and some or all of the observations, approximations for the remaining two can be obtained. This procedure has often been termed "quasi least squares."

This method is rather important when the adjustment problem is extensive. Here, determining approximations is not necessarily a simple computational effort. It can, in fact, become a sizable effort in its own right. But it needs to be said that such an effort for acquiring approximations is justified only if it is more than offset by thee savings that would accrue from having good approximations. This is one reason why, in many instances, the operation of detecting gross data errors (data editing) is combined with deriving approximations. It has a twofold advantage: to use efficiently the apparently extensive computation effort and to detect outliers and gross errors at an early stage. A very good example of this is performing sequential triangulation prior to a large block reduction by bundles. Forming models, then strips, and adjusting a block by strips using polynomials often will form a sequential set of steps in order both to detect and eliminate gross errors and blunders and to obtain the best approximations for use in a subsequent block adjustment by bundles. Experiences with such a system indicate that the extensive block triangulations by bundles, which require long computer time, are usually complete after only one iteration.

11.2.3. Recapitulation All the examples given above illustrate several points that may be summarized as follows.

First, each problem or family of problems usually has its own suitable possibilities for either obtaining approximations or reducing the model such that approximations are not needed.

Second, the degree of generality of the method selected for computing approximations determines the extent of its suitability for rendering adequate approximations; that is, a relatively general procedure may succeed in the majority of cases to yield adequate approximations but may fail badly in certain specific cases. Consequently, whenever generality is sought, we must be aware that problems may arise in some special cases.

Third, whenever consideration is for large and extensive adjustment problems, automated techniques should be sought. Whenever possible such methods should also include the possibility for blunder detection.

It has been said on occasion that poor initial approximations are acceptable since adjustment is performed on a computer. The reasoning goes that in such a case all that is necessary is to let the solution continue for a large number of iterations. Although this approach may succeed for some problems, we must view this reasoning with much caution. If, for instance, the topological structure of the problem is changed due to poor initial approximations, the procedure might not converge to the proper solution.

Finally, we cannot push the system to extremes. For instance, it is not possible to select zero approximations as valid for all coordinates of a geodetic system of points or for, say, the sines and cosines of the same angle in a transformation.

11.3. A POSTERIORI ESTIMATE OF THE REFERENCE VARIANCE $(\hat{\sigma}_0^2)$

Whether the problem is originally linear, or is linearized, an estimate of the reference variance $\hat{\sigma}_0^2$ can be computed from the observational residuals and the a priori weight matrix of the observations, or by other formulas, as was given in all techniques of least squares presented. The reference variance is an essential element of each adjustment problem. It is useful for a number of purposes, including the assessment of the adjustment results (see Section 11.5) and termination of the iterations (see Section 11.4). Although equations for its computations have been given, it is useful to give the corresponding derivations, which is the purpose of this section. Before proceeding, let us enumerate a number of useful relationships.

1. A useful matrix relationship from Appendix A may be recalled first.

$$\mathbf{x}^t\mathbf{A}\mathbf{x} = tr(\mathbf{x}\mathbf{x}^t\mathbf{A}) \qquad \text{(A symmetric)} \tag{A.67}$$

2. Next, it is recalled that the covariance matrix of a random vector $\tilde{\mathbf{x}}$ is Σ_{xx} and its mean or expectation is $\mu_x = E(\tilde{\mathbf{x}})$. Then by definition,

$$\begin{aligned}
\Sigma_{xx} &= E[(\tilde{\mathbf{x}} - \mu_x)(\tilde{\mathbf{x}} - \mu_x)^t] \\
&= E[\tilde{\mathbf{x}}\tilde{\mathbf{x}}^t] - E[\tilde{\mathbf{x}}\mu_x^t] - E[\mu_x \tilde{\mathbf{x}}^t] + E[\mu_x \mu_x^t] \\
&= E[\tilde{\mathbf{x}}\tilde{\mathbf{x}}^t] - 2\mu_x \mu_x^t + \mu_x \mu_x^t
\end{aligned}$$

or

$$\Sigma_{xx} = E[\tilde{\mathbf{x}}\tilde{\mathbf{x}}^t] - \mu_x \mu_x^t \tag{11.23}$$

and alternatively

$$E[\tilde{\mathbf{x}}\tilde{\mathbf{x}}^t] = \Sigma_{xx} + \mu_x\mu_x^t \tag{11.24}$$

3. The elements of the observational vector l are such that their residuals have an expectation or mean vector of zero, or

$$\mu_v = 0 \tag{11.25}$$

This means that there are no biases in the observations.

4. By definition, the relations between covariance, cofactor, and weight matrices are

$$\mathbf{W} = \mathbf{Q}^{-1} = \sigma_0^2 \mathbf{\Sigma}^{-1} \tag{11.26}$$

where $\mathbf{W}$ is the weight matrix, $\mathbf{Q}$ is the cofactor matrix, and σ_0^2 is the reference variance.

With these general remarks, we may now proceed to derive the relations for an unbiased a posteriori estimate for the reference variance. Each technique of adjustment is treated separately.

11.3.1. General Case of Adjustment with Conditions (Observations and Parameters) For completeness, the pertinent relations are recalled

$$\underset{c,\,n\ \ n,\,1}{\mathbf{A}\ \ \mathbf{v}} + \underset{c,\,u\ \ u,\,1}{\mathbf{B}\ \ \mathbf{\Delta}} = \underset{c,\,1}{\mathbf{f}}$$

$$\underset{c,\,c}{\mathbf{Q}_e = \mathbf{AQA}^t = \mathbf{W}_e^{-1}} \qquad \underset{u,\,u}{\mathbf{N} = \mathbf{B}^t\mathbf{W}_e\,\mathbf{B}} \qquad \mathbf{t} = \mathbf{B}^t\mathbf{W}_e\,\mathbf{f}$$

and from equation (11.26), the following relations may be introduced

$$\mathbf{\Sigma}_e^{-1} = \frac{1}{\sigma_0^2}\,\mathbf{W}_e = (\mathbf{A}\mathbf{\Sigma}\mathbf{A}^t)^{-1} \qquad \mathbf{M} = \mathbf{B}^t\mathbf{\Sigma}_e^{-1}\mathbf{B}$$

Further, recalling the relations

$$\mathbf{Q}_{ff} = \mathbf{Q}_e \qquad \mathbf{Q}_{tt} = \mathbf{N} \qquad \mathbf{Q}_{\Delta\Delta} = \mathbf{N}^{-1}$$

and factoring the reference variance leads to

$$\mathbf{\Sigma}_{ff} = \mathbf{\Sigma}_e \qquad \mathbf{\Sigma}_{tt} = \sigma_0^4\,\mathbf{M} \qquad \mathbf{\Sigma}_{\Delta\Delta} = \mathbf{M}^{-1}$$

Finally, we compute the expectation or mean of $\mathbf{f}$ as

$$\mathbf{\mu}_f = E[\mathbf{f}] = \mathbf{A}E[\mathbf{v}] + \mathbf{B}E[\mathbf{\Delta}]$$

or

$$\mathbf{\mu}_f = \mathbf{B}\mathbf{\mu}_\Delta$$

because

$$E[\mathbf{v}] = \mathbf{\mu}_v = 0$$

The quadratic form of the residuals is

$$\mathbf{v}^t\mathbf{W}\mathbf{v} = \sigma_0^2(\mathbf{v}^t\mathbf{\Sigma}^{-1}\mathbf{v}) \tag{11.27}$$

and from equation (6.27)

$$\mathbf{v}^t\mathbf{W}\mathbf{v} = \mathbf{f}^t\mathbf{W}_e\,\mathbf{f} - \mathbf{\Delta}^t\mathbf{t} = \mathbf{f}^t\mathbf{W}_e\,\mathbf{f} - \mathbf{\Delta}^t\mathbf{N}\mathbf{\Delta}$$

and

$$\mathbf{v}^t\boldsymbol{\Sigma}^{-1}\mathbf{v} = \mathbf{f}^t\boldsymbol{\Sigma}_e^{-1}\mathbf{f} - \boldsymbol{\Delta}^t\mathbf{M}\boldsymbol{\Delta}$$

$$E[\mathbf{v}^t\boldsymbol{\Sigma}^{-1}\mathbf{v}] = E[\mathbf{f}^t\boldsymbol{\Sigma}_e^{-1}\mathbf{f}] - E[\boldsymbol{\Delta}^t\mathbf{M}\boldsymbol{\Delta}]$$

or from equation (A67)

$$\begin{aligned}
E[\mathbf{v}^t\boldsymbol{\Sigma}^{-1}\mathbf{v}] &= E[tr(\mathbf{ff}^t\boldsymbol{\Sigma}_e^{-1})] - E[tr(\boldsymbol{\Delta}\boldsymbol{\Delta}^t\mathbf{M})] \\
&= tr(E[\mathbf{ff}^t]\boldsymbol{\Sigma}_e^{-1}) - tr(E[\boldsymbol{\Delta}\boldsymbol{\Delta}^t]\mathbf{M}) \\
&= tr([\boldsymbol{\Sigma}_{ff} + \boldsymbol{\mu}_f\boldsymbol{\mu}_f^t]\boldsymbol{\Sigma}_e^{-1}) - tr([\boldsymbol{\Sigma}_{\Delta\Delta} + \boldsymbol{\mu}_\Delta\boldsymbol{\mu}_\Delta^t]\mathbf{M}) \\
&= tr(\mathbf{I}_c + \boldsymbol{\mu}_f\boldsymbol{\mu}_f^t\boldsymbol{\Sigma}_e^{-1}) - tr(\mathbf{I}_u + \boldsymbol{\mu}_\Delta\boldsymbol{\mu}_\Delta^t\mathbf{M}) \\
&= tr(\mathbf{I}_c - \mathbf{I}_u) + tr(\boldsymbol{\mu}_f\boldsymbol{\mu}_f^t\boldsymbol{\Sigma}_e^{-1} - \boldsymbol{\mu}_\Delta\boldsymbol{\mu}_\Delta^t\mathbf{M}) \\
&= (c - u) + \boldsymbol{\mu}_f^t\boldsymbol{\Sigma}_e^{-1}\boldsymbol{\mu}_f - \boldsymbol{\mu}_\Delta^t\mathbf{M}\boldsymbol{\mu}_\Delta
\end{aligned}$$

which, in view of the fact that $\boldsymbol{\mu}_f = \mathbf{B}\boldsymbol{\mu}_\Delta$, becomes

$$\begin{aligned}
E[\mathbf{v}^t\boldsymbol{\Sigma}^{-1}\mathbf{v}] &= (c - u) + \boldsymbol{\mu}_\Delta^t\mathbf{B}^t\boldsymbol{\Sigma}_e^{-1}\mathbf{B}\boldsymbol{\mu}_\Delta - \boldsymbol{\mu}_\Delta^t\mathbf{M}\boldsymbol{\mu}_\Delta \\
&= (c - u) + \boldsymbol{\mu}_\Delta^t\mathbf{M}\boldsymbol{\mu}_\Delta - \boldsymbol{\mu}_\Delta^t\mathbf{M}\boldsymbol{\mu}_\Delta \\
&= (c - u) \\
&= r + u - u = r = \text{the degree of freedom (or redundancy)}
\end{aligned}$$

Thus according to equation (11.27),

$$E[\mathbf{v}^t\mathbf{W}\mathbf{v}] = \sigma_0^2 E[\mathbf{v}^t\boldsymbol{\Sigma}^{-1}\mathbf{v}] = \sigma_0^2 r$$

$$\sigma_0^2 = \frac{E[\mathbf{v}^t\mathbf{W}\mathbf{v}]}{r} \tag{11.28}$$

Consequently, an unbiased estimate of the reference variance may be computed a posteriori of the least squares adjustment from

$$\hat{\sigma}_0^2 = \frac{\mathbf{v}^t\mathbf{W}\mathbf{v}}{c - u} = \frac{\mathbf{v}^t\mathbf{W}\mathbf{v}}{r} \tag{11.29}$$

11.3.2. Case of Adjustment of Observations Only

In this case the relations are

$$\mathbf{A}\mathbf{v} = \mathbf{f}, \qquad \mathbf{Q}_e = \mathbf{A}\mathbf{Q}\mathbf{A}^t, \qquad \boldsymbol{\Sigma}_e = \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^t, \qquad \mathbf{Q}_{ff} = \mathbf{Q}_e,$$

and

$$\boldsymbol{\Sigma}_{ff} = \boldsymbol{\Sigma}_e$$

Since $\boldsymbol{\mu}_v = \mathbf{0}$, then $\boldsymbol{\mu}_f = \mathbf{0}$ and from equation (11.23)

$$E[\mathbf{ff}^t] = \boldsymbol{\Sigma}_{ff} = \boldsymbol{\Sigma}_e$$

In a manner similar to the derivation above, it can be shown that

$$E(\mathbf{v}^t\mathbf{\Sigma}^{-1}\mathbf{v}) = r$$

and an unbiased a posteriori estimate would be

$$\hat{\sigma}_0^2 = \frac{\mathbf{v}^t\mathbf{W}\mathbf{v}}{r} \tag{11.30}$$

11.3.3. Case of Adjustment of Indirect Observations   Earlier relations are recalled,

$$\mathbf{v} + \mathbf{B}\mathbf{\Delta} = \mathbf{f}, \qquad \mathbf{N} = \mathbf{B}^t\mathbf{W}\mathbf{B}, \qquad \mathbf{t} = \mathbf{B}^t\mathbf{W}\mathbf{f} \qquad \mathbf{M} = \mathbf{B}^t\mathbf{\Sigma}^{-1}\mathbf{B}$$

and

$$\mathbf{Q}_{ff} = \mathbf{Q}$$

thus

$$\mathbf{\Sigma}_{ff} = \mathbf{\Sigma}$$

$$\mathbf{Q}_{\Delta\Delta} = \mathbf{N}$$

hence

$$\mathbf{\Sigma}_{\Delta\Delta} = \mathbf{M}$$

Again because $\mathbf{\mu}_v = \mathbf{0}$, then $\mathbf{\mu}_f = \mathbf{B}\mathbf{\mu}_\Delta$. Next

$$\mathbf{v}^t\mathbf{W}\mathbf{v} = \mathbf{f}^t\mathbf{W}\mathbf{f} - \mathbf{\Delta}^t\mathbf{t} = \mathbf{f}^t\mathbf{W}\mathbf{f} - \mathbf{\Delta}^t\mathbf{N}\mathbf{\Delta}$$

from which it can be shown that

$$E(\mathbf{v}^t\mathbf{\Sigma}^{-1}\mathbf{v}) = (n - u) = r = \text{degrees of freedom}$$

and an unbiased estimate is therefore

$$\hat{\sigma}_0^2 = \frac{\mathbf{v}^t\mathbf{W}\mathbf{v}}{r} \tag{11.31}$$

The above derivations show the generality of the relation for the a posteriori estimate of the reference variance. It is always equal to the value of the quadratic form of the residuals divided by the number of degrees of freedom. Because of this rather general relation, no attempt is made to prove the cases of adjustment with functionally dependent parameters, as the same formula will always hold.

The reader must be aware of the assumptions made during the derivations. Of these, two important ones need to be mentioned: the assumption that $\mathbf{\mu}_v = \mathbf{0}$ or no bias in the observations, and the assumption that one has a priori the population (or distribution) covariance (or cofactor) matrix. The inclusion of these two assumptions provides for obtaining an unbiased estimate of the reference variance.

11.4. ITERATION TERMINATION WITH LINEARIZED CONDITIONS

When setting up an iterative adjustment program, criteria for terminating the iterations must be incorporated in the program. In order to select the appropriate criterion, or criteria, we must recognize that iteration termination is the mechanism by which we ascertain that the solution has converged sufficiently "close" to the theoretical nonlinear solution. The "closeness" of the solution to be accepted is ultimately limited by the word "length" in the computer—that is, the number of digits carried in the computation unless special techniques are applied. Consequently, practical considerations are important when we design an iterative adjustment system.

Before addressing the actual iteration termination subject, it is useful to mention a point of general nature regarding convergence. In principle, not all iterative processes necessarily converge. Some may diverge, others may oscillate or repeat themselves in certain cycles, and yet others may converge but not to the desired solution. It is not the intention here to dwell on all types of iterative procedures, which is a topic for numerical analysis to be found in the corresponding literature. Instead, consideration is given only to one type: repeated relinearization of the nonlinear condition equations and the performance of a complete linear least squares solution within each linearization cycle. This procedure belongs to a class of iterative methods, called "Newton processes," the convergence properties of which are known by general theory. Such processes have a positive definite coefficient matrix and will converge to the correct (minimum) solution. This is always the case for least squares normal equations' coefficient matrix; hence the desired answer will be attained, given certain conditions concerning the approximate values. Convergence in such methods is known to be of second order and is relatively fast. All this implies that unless the system is degenerate or the adjustment model is incorrectly established, convergence would be quickly obtained provided that problems of numerical acuity are not encountered.

The problem of limited numerical acuity is therefore another practical matter that merits consideration. It would make very little sense to continue with meaningless iterations beyond the level dictated by the number of significant digits and accumulating rounding-off errors.

In addition to the two elements—striving to reach as close an answer to the theoretical solution as possible, and the practical limitations due to digital considerations—there is a third element, that of economy. Although we may wish to get the answer with the highest precision, it is unwise to carry on numerous iterations for a relatively small gain in accuracy. Therefore whenever a criterion for terminating iterations is selected, this rather important economic factor must be considered. It actually refers to proper experimental design of balancing accuracy levels versus cost, which in turn affects iteration termination.

11.4.1. Criteria for Iteration Termination   The selected criteria should be representative indicators of convergence. They should have high discriminative power to signalize the attainment of convergence. Several criteria will therefore be given, together with examples, and a discussion as to their power.

FIRST GROUP OF CRITERIA   The first group of criteria involves the quadratic form, $v^t W v$, the reference variance, $\hat{\sigma}_0^2 = (1/r)v^t W v$, or the reference standard deviation (the so-called standard error of unit weight) $\hat{\sigma}_0$, which are obviously all related. We may test the absolute value against a preselected tolerance value, or we may test the numerical change from one iteration to the next against a threshold. It is also possible to test the rate of change, rather than just the change. For instance, if we consider $\hat{\sigma}_0$, then we may test

$$\hat{\sigma}_0 < \delta_1 \tag{11.32a}$$

$$\left| \hat{\sigma}_{0_j} - \hat{\sigma}_{0_{j+1}} \right| < \delta_2 \tag{11.32b}$$

$$\left| \frac{\hat{\sigma}_{0_j} - \hat{\sigma}_{0_{j+1}}}{\hat{\sigma}_{0_j}} \right| = \left| 1 - \frac{\hat{\sigma}_{0_{j+1}}}{\hat{\sigma}_{0_j}} \right| < \delta_3 \tag{11.32c}$$

where δ_i is the appropriate tolerance value and j the iterative index. Equations (11.32) apply equally to $\hat{\sigma}_0^2$ and to $(v^t W v)$. Through the iterative procedure, the quadratic form or the reference variance (or both) can be, and often are, computed and listed; but these criteria are not very discriminative. In fact, one of the general difficulties is to select the numerical value of the threshold, such that "sufficient" convergence is reached. This bears directly on the discriminative power of the criterion and no general scheme is available for the choice of δ. Correspondingly, each case or group of cases, as in a particular field of application, needs to be studied separately.

The use of equation (11.32a) has been found to be unadvisable. The value of $\hat{\sigma}_0$ either does not reach δ_1 at all, or when the iterations are terminated at it, the convergence is not actually reached. It is usually better to use equation (11.32c) as it is more indicative of convergence, although still not with much power. At any rate it is quite difficult to determine a priori the proper ratio that would signify convergence. The value of δ_3 in equation (27.32c) can be as small as 10^{-5} and in some cases even smaller. In other cases there may be constant convergence ratios established experimentally through a number of simulated studies. In these latter cases these constants may be directly used for problems of the same type.

The use of relations depending on the reference variance (or reference standard deviation) can be nearly useless in cases of adjustment problems with low numerical stability. Such cases are encountered when nearly indeterminate (or poorly determinate) situations are involved in which the determinant of the normal equations coefficient matrix approaches zero. An example of this case involves a photogrammetric strip adjustment by

independent models with ground control only at the beginning and end of fairly long strips (30–50 models). This represents a case of low determinacy.

Using σ_0 as a criterion, the Newton process converged after three iterations to the fifth decimal place. By contrast, the method of conjugate gradiant used to reach the least squares minimum required 230 iterations (which is here not directly relevant) to reach σ_0 within 10^{-4} μm. *Nevertheless* the coordinates in the center of the strip were still different from the check values by some 30 μm in the negative scale. This means that in cases of such poor control, the use of σ_0 with the conjugate gradiant method is nearly useless for iteration termination. Even with the very large number of iterations, correct convergence was not reached. It is because of situations like this that the criterion of terminating iterations after a certain prefixed number of iterations is used, as is discussed in the section to follow on the third criterion.

SECOND GROUP OF CRITERIA   The second group of criteria is related to parameter increments or functions of such increments. There are a number of possibilities.

1. The value of *each* increment (or parameter correction) Δ, should approach zero. In this case a threshold ε which is insignificantly small may be used for the test,

$$|\Delta| < \varepsilon \tag{11.33a}$$

2. Instead of having each parameter increment approach zero, it is often sufficient to check on the maximum increment, or,

$$|\Delta_{\max}| < \varepsilon \tag{11.33b}$$

3. In many problems the parameters involved may be in the form of several distinct classes or subsets, each of which contains parameters of the same type (angular, linear, and so on). In such a case the criterion would check the maximum value within each subset i of parameters of a kind, or

$$|\Delta_{i\,\max}| < \varepsilon \tag{11.33c}$$

As a simple example, consider a geodetic net: We may stop iterations when the maximum alteration in coordinates is less than, say, 1 mm, and the maximum change in angular parameters is less than 10^{-6} radians. The same would apply to aerial photography with a threshold for the camera position and another for the camera orientation parameters.

4. Another possibility for a criterion is to use the rate of change of the parameter increments, rather than the increments themselves, or

$$\left| \frac{\Delta_{i-1} - \Delta_i}{\Delta_i} \right| < \varepsilon \tag{11.33d}$$

In some cases, a different form is used

$$\left| \frac{\Delta_{i-1} - \Delta_i}{\sigma_\Delta} \right| < \varepsilon \tag{11.33e}$$

The use of equation (11.33d) or (11.33e) as a test criterion for terminating iterations is somewhat more general than the use of equation (11.33a), (11.33b), or (11.33c) because of being not too dependent on the type of parameters. For example, the rate of change of angular parameters of an aerial photograph may be

$$\left| \frac{\alpha_{i-1} - \alpha_i}{\sigma} \right| < \frac{0.1^{cc}}{10^{cc}} < 10^{-2}$$

5. Yet another group of criteria utilizes a function of the parameters or functions of subsets of them instead of the parameters themselves. A very simple example would be to test on

$$\omega^2 + \phi^2 + \kappa^2 < \varepsilon^2 \tag{11.33f}$$

instead of each of the rotation elements (ω, ϕ, and κ) of an aerial photograph. A more sophisticated example for criteria of this kind concerns photogrammetric block adjustment. Instead of testing on the parameters (exterior orientation elements of the photographs, both positional and rotational), we may test on the maximum changes in point coordinates at the negative scale, or

$$\Delta x_{max}, \Delta y_{max}, \Delta z_{max} < \varepsilon \tag{11.33g}$$

in which ε may take a small value of say 0.1 μm.

THIRD CRITERION   A criterion that is useful to use in addition to other criteria is to specify that iterations must not exceed a certain number which is input to the adjustment program. Such a criterion is used usually as a safeguard against excessive iterations in case other criteria fail. The failure of the built-in criteria could be due to either the presence of large mistakes or the poor numerical condition of the system. Consequently, the choice of the number of iterations after which cutoff is effected should not be taken lightly. Although it obviously should be small when blunders contaminate the data, it must be large enough to allow for the suitable number of iterations in case of quasistable systems. In other words the threshold number of iterations in the latter case should not lead to stopping the iterations too soon.

It is evident from the preceding paragraph that gross data errors affect the iterations and consequently the criterion for their termination. In the following section, this point will be briefly discussed.

11.4.2. Iteration Termination and Gross Errors   The question of terminating iterations is, in practice, always intertwined with the problem of gross data errors—their detection and elimination. In large systems these error

types usually occur and therefore the editing of the data is a necessary operation in order for the solution to converge to the appropriate answer with a reasonable number of iterations. Gross errors and blunders can disturb the iterative process considerably, to the extent that the system may diverge. This is one reason for using a criterion based on σ_0. Here, a check that $\sigma_{0(i+1)} < \sigma_{0i}$ will assure that at least the solution is not diverging. Of course, it has little power in pointing out the rate of convergence. Thus with some criterion on σ_0, very large blunders in the data may be intercepted and eliminated, hence avoiding the possibility of divergence. Other criteria may also be incorporated to detect blunders of moderate size which may slow down, but not impede, convergence, and which may also lead to an unacceptable minimum solution.

In concluding this section, it is worth stressing that iteration termination is an important operation both technically and economically. Well-conditioned and designed adjustment problems, with no gross errors, are, in principle, quite reasonable to handle using the criteria given above. As an example, a quotation† may be given regarding two block adjustments of large-scale photography, summarized in Table 11-1. The method of triangulation is called PAT-M-43 and uses the method of independent models. The first block, Block A, contained 50 models of photo scale 1 : 3400, and the second, Block B, had 129 models of photo scale 1 : 14000. Counting a plan and height adjustment sequence as one iteration, the maximum alteration against the previous iteration were reduced to 6 mm for A, and 0 mm for B, after three steps, and after the first iteration the coordinates changed 1758 m and 10229 m, respectively. It can be seen that only three iterations were more than sufficient for convergence under the apparently good conditions of the problem.

TABLE 11-1 Rates of Convergence of Plane and Height Iterations of Three-Dimensional Block Adjustments with PAT-M-43

BLOCK	COOR-DINATE DIFFER-ENCES	ITERATION STEP 1		ITERATION STEP 2		ITERATION STEP 3	
		PLAN (m)	HEIGHT (m)	PLAN (m)	HEIGHT (m)	PLAN (m)	HEIGHT (m)
A	x_{max}	1501	8	0.78	0.18	0.005	0.001
	y_{max}	1758	23	0.93	0.19	0.006	0.001
	z_{max}	318	147	0.47	0.21	0.005	0.001
B	x_{max}	10,229	30	0.92	0.52	0.000	0.000
	y_{max}	7,303	85	2.68	0.35	0.000	0.000
	z_{max}	389	527	0.78	0.65	0.000	0.000

† See Ackermann (1973) in the Bibliography.

By contrast to well-conditioned systems, the termination tolerances may have to be quite fine for poorly conditioned problems. In such cases there may be large numbers of iterations and we may have to compare the nth iteration not only with $(n - 1)$ but perhaps also with the $(n - 10)$ iteration. Another point is always to ascertain that iterations do not proceed beyond the point at which the numerical acuity of the system is reached. Finally, then, when designing a large adjustment system for nonlinear problems, it is advisable to incorporate several different criteria for iteration termination— as well as, in some cases, automatic data editing. The extensive programs for photogrammetric block triangulation by DBA systems are good examples for both automatic data editing and iteration termination schemes.†

11.5. A POSTERIORI STATISTICAL ANALYSIS

The results of least squares adjustment are composed of (point) estimates for one or more sets of random variables (parameters or observations, or both) as well as estimates of the corresponding cofactor or covariance matrices. Once the adjustment is performed, statistical techniques may be used to analyze and assess the results obtained. Although least squares estimation does not require a priori knowledge of the distributions of the random variables involved, the a posteriori statistical analysis does require such distributions. Therefore it is usually assumed that the observations are normally distributed and that the least squares estimation is performed on a linear functional model. Consequently, for nonlinear cases it is assumed that after iteration termination, the linearized form is sufficiently close to make the statistical testing valid.

11.5.1. Tests on the Reference Variance One of the statistical tests commonly applied to least squares data concerns the a posteriori estimate of the reference variance $\hat{\sigma}_0^2$. This value may be compared to the a priori value σ_0^2 using a χ^2 test (see Section 2.8.3), since

$$\chi_r^2 = \frac{r\hat{\sigma}_0^2}{\sigma_0^2}$$

has a χ^2 distribution with r degrees of freedom (r = redundancy in the adjustment). There are two possible hypothesis tests: First,

$$H_0: \sigma^2 = \sigma_0^2 \qquad \text{vs.} \qquad H_1: \sigma^2 \neq \sigma_0^2$$

and reject H_0 when

$$\chi_r^2 < \chi_{1-\alpha/2, r}^2 \qquad \text{or} \qquad \chi_r^2 > \chi_{\alpha/2, r}^2$$

† See Davis (1967) in the Bibliography.

The value χ_r^2 is computed from the adjustment results as

$$\chi_r^2 = \frac{\phi}{\sigma_0^2} = \frac{\mathbf{v}^t \mathbf{W} \mathbf{v}}{\sigma_0^2}$$

Second,

$$H_0: \sigma^2 = \sigma_0^2 \qquad \text{vs.} \qquad H_1: \sigma^2 > \sigma_0^2$$

and reject H_0 when

$$\chi_r^2 > \chi_{\alpha, r}^2$$

The second hypothesis test, which is a one-tail upper-bound test, is more frequently employed in practice. However, the first two-tail test may be useful, particularly in the cases of the design of new adjustment programs.

The use of the χ^2 test above required the a priori knowledge of the reference variance σ_0^2. In some situations the best available value for the unit variance may itself be an estimate with its own statistical fluctuations. In such instances we may compare the estimate from the adjustment to that a priori estimate and apply the F test for the equality of two variances (see Section 2.8.3).

Let $\hat{\sigma}_0^2$ with r degrees of freedom be the result from the least squares adjustment to be tested, and $\hat{\sigma}_{00}^2$ while r_0 degrees of freedom be the a priori known values. Denote the larger value of the two variances ($\hat{\sigma}_0^2$ or $\hat{\sigma}_{00}^2$) by $\hat{\sigma}_1^2$ and the smaller by $\hat{\sigma}_2^2$, with corresponding degrees of freedom r_1, r_2, and compute $F_{r_1, r_2} = \hat{\sigma}_1^2 / \hat{\sigma}_2^2$. The hypothesis to be tested is $H_0: \sigma_1^2 = \sigma_2^2$ versus the alternative $H_1: \sigma_1^2 > \sigma_2^2$. Thus we would reject H_0 if $F_{r_1, r_2} > F_{\alpha, r_1, r_2}$. In practice, the F ratio is often evaluated as $\hat{\sigma}_0^2$ (from the adjustment) divided by $\hat{\sigma}_{00}^2$ (a priori). If this ratio is less than one, then we usually do not need to test a hypothesis and $\hat{\sigma}_0^2$ is taken to be insignificantly different from $\hat{\sigma}_{00}^2$. Obviously, if the ratio is larger than one, the test proceeds as specified.

If the test on $\hat{\sigma}_0^2$ fails, a poor adjustment is indicated and we must investigate the possible cause. There are several possibilities.

First, we must ascertain that the computations were performed correctly before proceeding to determine other possible sources of difficulty.

Second, we need to investigate how adequately the systematic effects were accounted for. This also relates to the adequacy of the selected model since systematic effects are best interpreted as model deficiencies.

Third, it is very important that the *proper number* of condition (and constraint) equations be determined. Furthermore, the actual functional relationships forming such equations must be correct in the sense of reflecting the physical and geometric situation of the adjustment problem. If both of these points are met, then the linearized form (if nonlinear problem) must also be checked, particularly with the selected approximations, to make sure that numerical instabilities did not occur. All these factors pertain to the functional model.

Finally, the proper stochastic model is equally important. In some instances the a priori cofactor matrix may be accurately known but σ_0^2 is estimated. Here we should apply the F test in order to account for the fact that the a priori value of the reference variance is an estimate and not a numerical constant (representing the population variance). If the a priori variances and covariances are estimated through repeated observations, the number of repetitions should be sufficiently large, that is, 20 to 30.

Example 11.4. Refer to the three-photo problem of Example 6.1 and test if the a posteriori reference variance $\hat{\sigma}_0^2$ is significantly different from its a priori value. The a posteriori value is $\hat{\sigma}_0^2 = 1.9634$ with 1 degree of freedom and the a priori value is $\sigma_0^2 = 1.0000$ (with infinite degrees of freedom). Thus according to Section 2.8.3,

$$\chi_m^2 = \frac{ms^2}{\sigma^2} \qquad \text{or} \qquad \chi_1^2 = \frac{(1)\hat{\sigma}_0^2}{\sigma_0^2} = \frac{(1)(1.9634)}{(1)} = 1.9634$$

and

$$H_0: \sigma^2 = \sigma_0^2 \qquad \text{vs.} \qquad H_1: \sigma^2 > \sigma_0^2$$

The null hypothesis (H_0) would be rejected if $\chi_1^2 > \chi_{\alpha,\,1}^2$. At the level of significance

$$\alpha = 0.05, \qquad \chi_{0.05,\,1}^2 = 3.84 > \chi_1^2$$

and H_0 cannot be rejected. In fact, H_0 would still be accepted at $\alpha = 0.15$ level of significance.

11.5.2. Confidence Regions

It has been shown in Section 2.6 that the covariance matrix may be used to establish regions around the estimated values, each of which having a specified probability. In order to review this concept we begin by a numerical example before discussing its application to least squares results.

Example 11.5. The position of point A, in Figure 11.1 is determined by measuring r, α whose values are

$$r = 206\ 264\ \text{cm} \qquad \sigma_r = 2\ \text{cm}$$

$$\alpha = 60°00'00'' \qquad \sigma_\alpha = 1''$$

with $\sigma_{r\alpha} = 0$ (that is, r and α are uncorrelated). If $\mathbf{v} = [r \quad \alpha]^t$, then

$$\boldsymbol{\Sigma}_{vv} = \begin{bmatrix} 4\ \text{cm}^2 & 0 \\ 0 & 1/\rho^2 \end{bmatrix}$$

where ρ is for conversion from seconds of arc to radian and is equal to 206,264 or r. The position of point A may be expressed by the cartesian coordinates x_1 and x_2 and the covariance matrix $\boldsymbol{\Sigma}_{xx}$. This matrix is obtained from $\boldsymbol{\Sigma}_{vv}$ by propagation rule, or

$$\boldsymbol{\Sigma}_{xx} = \mathbf{J}_{xv}\boldsymbol{\Sigma}_{vv}\mathbf{J}_{xv}^t$$
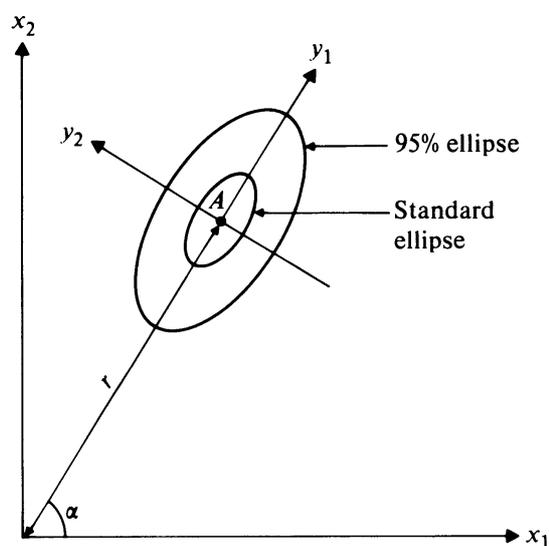
Figure 11.1

To evaluate the Jacobian $\mathbf{J}_{xv}$, the cartesian coordinates

$$\mathbf{x} = [r \cos \alpha \quad r \sin \alpha]^t$$

are differentiated with respect to r and α. Thus

$$\mathbf{J} = \begin{bmatrix} \cos \alpha & -r \sin \alpha \\ \sin \alpha & r \cos \alpha \end{bmatrix} \quad \text{and} \quad \mathbf{\Sigma}_{xx} = \tfrac{1}{4} \begin{bmatrix} 7 & 3\sqrt{3} \\ 3\sqrt{3} & 13 \end{bmatrix}$$

Suppose we are interested in obtaining the standard ellipse associated with $\mathbf{\Sigma}_{xx}$. First the eigenvalues are computed from the characteristic polynomial

$$\lambda^2 - 5\lambda + 4 = 0$$

or

$$\lambda_1 = 4 \quad \text{and} \quad \lambda_2 = 1$$

Second, the eigenvectors are evaluated as

$$\mathbf{y}_1 = [1 \quad \sqrt{3}]^t \quad \text{for } \lambda_1$$

and

$$\mathbf{y}_2 = [1 \quad -\sqrt{3}/2]^t \quad \text{for } \lambda_2$$

Consequently, the semimajor axis of the standard ellipse is $a = \sqrt{\lambda_1} = 2$ cm and is in the direction of $\mathbf{y}_1$, and the semiminor axis is $b = \sqrt{\lambda_2} = 1$ cm in the direction of $\mathbf{y}_2$ (see Figure 11.1). These results are consistent with the original data of the problem since

$$a = \sigma_{y_1} = \sigma_r \quad \text{and} \quad b = \sigma_{y_2} = \rho \sigma_\alpha$$

Furthermore, the diagonalized form of $\mathbf{\Sigma}_{xx}$ reverts back to $\mathbf{\Sigma}_{vv}$.

Although the probability is 0.394 for the point A to lie inside the standard ellipse, a 95 % confidence ellipse can be established. Referring to Table 2.1, for $P = 0.95$ the value of $k = 2.447$. Therefore the 95% confidence ellipse has a semimajor axis

$$a = 2.447 \times 2 = 4.894 \text{ cm}$$

and a semiminor axis

$$b = 2.447 \times 1 = 2.447 \text{ cm}$$

Other confidence ellipses can be constructed through the use of the appropriate values of k from Table 2.1. An equivalent circular confidence region may also be established using the techniques presented in Section 2.6. For example, the circular standard error is

$$\sigma_c = \simeq \tfrac{1}{2}(\sigma_{y_1} + \sigma_{y_2}) \simeq 1.5 \text{ cm}$$

and the circular probable error is

$$\text{CPE} \simeq 0.589(\sigma_{y_1} + \sigma_{y_2}) \simeq 1.767 \text{ cm}$$

The probability for the point to lie within the circle with radius σ_c is approximately equal to 0.394 (same as the standard ellipse), and that inside the circle with radius of circular probable error is approximately equal to 0.500.

The preceding example dealt with a direct point determination and associated error propagation. The same techniques used for getting the confidence regions apply equally to the results of least squares adjustment. Suppose $\hat{\mathbf{x}}_{u,\,1}$ is the vector of least squares estimates of a set of parameters $\mathbf{x}$, where $\mathbf{x}$ is conventionally known as the "true" values, or in more modern terminology it is the mean vector of the multidimensional distribution of these random variables. Again, from the information in Section 2.6 the quadratic form

$$\phi = (\hat{\mathbf{x}} - \mathbf{x})' \Sigma_{xx}^{-1} (\hat{\mathbf{x}} - \mathbf{x})$$

represents an equation of a hyperellipsoid in u-dimensional space, centered at $\hat{\mathbf{x}}$. It also has a χ^2 distribution with u degrees of freedom. Thus for $\phi = 1$ the region of constant probability will be that of the standard hyperellipsoid. For other numerical values a corresponding region may be established and its associated probability determined. For example, consider the spatial determination in three-dimensional space. If we are interested in the 95% confidence ellipsoid, then with $\chi^2_{0.05,\,3} = 7.81$ (see Table D3), its equation would be

$$(\hat{\mathbf{x}} - \mathbf{x})' \Sigma_{xx}^{-1} (\hat{\mathbf{x}} - \mathbf{x}) = 7.81$$
$$\text{3.3}$$

Let the eigenvalues of Σ be denoted, in decreasing numerical value, by $\lambda_1, \lambda_2, \lambda_3$, and the corresponding eigenvectors by $\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3$. The 95% confidence ellipsoid would have the semimajor axes: $a = \sqrt{7.81\lambda_1}$ oriented along $\mathbf{y}_1$,

$b = \sqrt{7.81\lambda_2}$ oriented along $\mathbf{y}_2$, and $c = \sqrt{7.81\lambda_3}$ oriented along $\mathbf{y}_3$. The interpretation of this confidence ellipsoid is that if the determination of $\hat{\mathbf{x}}$ is repeated a large number of times, then we would expect on the average that 95% of the confidence ellipsoids thus determined would include the "true" point.

In the preceding discussion the covariance matrix Σ_{xx} is assumed to be known. This implies that σ_0^2 is known a priori and used to obtain Σ_{xx} from the inverse of the reduced normal equations' coefficient matrix $\mathbf{N}^{-1}$. If instead, the reference variance is estimated by $\hat{\sigma}_0^2$ from the least squares adjustment, then

$$\frac{1}{u\hat{\sigma}_0^2}(\hat{\mathbf{x}} - \mathbf{x})^t \mathbf{N}(\hat{\mathbf{x}} - \mathbf{x}) = F_{u, r} \tag{11.34}$$

and hypothesis tests can be performed using the tabulated F values. For example, for

$$H_0 : \mathbf{x} = \mathbf{x}_0 \qquad \text{vs.} \qquad H_1 : \mathbf{x} \neq \mathbf{x}_0$$

we would reject H_0 if $F_{u, r}$ computed from equation (11.34) exceeds $F_{\alpha, u, r}$ at the α level of significance (r = redundancy in the adjustment). In a similar manner confidence regions can be established using the appropriate F values instead of χ^2 when the estimate $\hat{\sigma}_0^2$ is used.

Example 11.6. Let us continue the analysis of the three-photo problem (Examples 6.1 and 11.4) by considering the a posteriori covariance matrix of the parameters x_1 and x_2, that is, the position of point P (see Figure 6.1). Since $\hat{\sigma}_0^2$ was found to be consistent with the a priori value (Example 11.5), then the covariance matrix Σ_{xx} is ($\mathbf{Q}_{\Delta\Delta(f)}$ in Example 6.1)

$$\Sigma_{xx} = 10^{-3} \begin{bmatrix} 1.86 & 5.03 \\ 5.03 & 71.33 \end{bmatrix} (\text{m}^2)$$

The eigenvalues of this matrix may be computed to be $\lambda_1 = 0.0729$ and $\lambda_2 = 0.0016$, thus $a = 0.27$ m and $b = 0.04$ m for the standard ellipse. The major axis deviates only slightly from the x_2 axis (by about $4°07'$). The rather flattened shape of the ellipse is due to the geometry of the intersection of the three rays. Due to the small angles of intersection, the point P is relatively poorly determined in the direction of the ellipse major axis which is here very nearly the x_2 axis.

The shape of the ellipse is independent of the actual values of the observations. Rather, it depends on the geometry and the a priori covariance matrix. It can therefore be studied without the actual measurements. Such studies are performed in the planning phase so that we may decide beforehand whether the planned observations will yield the required results, or whether additional observations or more precise observations are needed.

11.5.3. Correlation Analysis

The concept of correlation has been introduced and explained in Chapter 2 of Part I, and a relationship for the correlation coefficient ρ between two random variables has been presented.

The numerical value of ρ_{xy}, which by definition varies between $+1$ and -1, reflects the degree of correlation between the two random variables $\tilde{x}$ and $\tilde{y}$. When $|\rho| = 1$, there exists perfect linear correlation, and plotted observations (x as abscissa and y as ordinate) will lie on a straight line. If the plotted points fall in a narrow band about a straight line, linear correlation of moderate value is said to exist. The value of the correlation coefficient will tend toward zero when the plotted points either fall on a curved line (nonlinear correlation) or are widely scattered, thus implying lack of correlation.

An estimate of the correlation coefficient $\hat{\rho}$ may be computed from a sample in the same way as would the sample variance s_x^2 and sample covariance s_{xy}, or

$$\hat{\rho} = \frac{s_{xy}}{s_x s_y} \tag{11.35}$$

where

$$s_{xy} = \sum \frac{(x - \bar{x})(y - \bar{y})}{n - 1} \tag{11.36a}$$

$$s_x^2 = \sum \frac{(x - \bar{x})^2}{n - 1} \tag{11.36b}$$

$$s_y^2 = \sum \frac{(y - \bar{y})^2}{n - 1} \tag{11.36c}$$

with n being the sample size. There are two possible tests on $\hat{\rho}$ under the null hypothesis, H_0: $\rho = 0$ against the alternative H_1: $\rho \neq 0$. The first applies the t distribution by computing

$$t_{n-2} = \left[\frac{(n - 2)\hat{\rho}^2}{1 - \hat{\rho}^2} \right]^{1/2} \tag{11.37}$$

and accepting H_0 when

$$-t_{\alpha/2, n-2} < t_{n-2} < t_{\alpha/2, n-2}$$

The second (for large n) applies the normal distribution by computing

$$z = \frac{\sqrt{n - 3}}{2} \ln \left[\frac{1 + \hat{\rho}}{1 - \hat{\rho}} \right] \tag{11.38}$$

and accepting H_0 when

$$-z_{\alpha/2} < z < z_{\alpha/2}$$

When H_0 is rejected, the two random variables in question are said to be significantly correlated at the selected α level. These tests assume that the random variables involved are normally distributed.

If the problem to be analyzed concerns *only* two random variables

(jointly normally distributed), then $\hat{\rho}$ as computed by equation (11.35) and either of the two tests above will be appropriate. On the other hand if a multinormal set of random variables are involved, it is better to use the *partial correlation coefficient*.

For example, if $\mathbf{\Sigma}_{xx}$ is the covariance matrix of the random vector $\tilde{\mathbf{x}}$ (which is $u \times 1$) and $\mathbf{W}_{xx} = \mathbf{\Sigma}_{xx}^{-1}$, then the correlation coefficient between x_1 and x_2 when the remaining $u - 2$ are fixed is

$$\rho_{12|3\cdots u} = \frac{-w_{12}}{(w_{11}w_{22})^{1/2}} \tag{11.39}$$

where w_{ij} are elements of $\mathbf{\Sigma}_{xx}^{-1}$. The estimate $\hat{\rho}_{12|3\cdots u}$ is defined the same way as in equation (11.39) using the covariance matrix evaluated from the sample values. Note that when $u = 2$, the two correlation coefficients from both equation (11.35) and (11.39) become identical since the inverse of the covariance matrix is

$$\begin{bmatrix} w_{11} & w_{12} \\ w_{12} & w_{22} \end{bmatrix} = \frac{1}{s_1^2 s_2^2 - s_{12}^2} \begin{bmatrix} s_2^2 & -s_{12} \\ -s_{12} & s_1^2 \end{bmatrix}$$

Similar to equation (11.37) the test of hypothesis $H_0: \rho_{12|3\cdots u} = 0$ versus the alternative $H_1: \rho_{12|3\cdots u} \neq 0$ is performed by computing

$$t_{n-u} = \left[\frac{(n-u)\hat{\rho}_{12|3\cdots u}^2}{1 - \hat{\rho}_{12|3\cdots u}^2} \right]^{1/2} \tag{11.40}$$

and accepting H_0 when

$$-t_{\alpha/2,\, n-u} < t_{n-u} < t_{\alpha/2,\, n-u}$$

In the test, n is the number of repeated measurements on the $u \times 1$ vector of observations, and u is the number of parameters to be determined (the size of the random vector). In case of linear least squares, n becomes the number of observations and u the number of parameters estimated in the vector Δ. In this case, $n - u$ is the redundancy r, and the partial correlation coefficient may be evaluated directly from the coefficient matrix $\mathbf{N}$ of the partially reduced normal equations (being the inverse of the scaled covariance matrix). Hence $\hat{\rho}_{12|3\cdots u}$ for the correlation between δ_1 and δ_2 of Δ is evaluated from

$$\hat{\rho}_{12|3\cdots u} = \frac{-n_{12}}{(n_{11}n_{22})^{1/2}} \tag{11.41}$$

where n_{ij} are elements of $\mathbf{N}$. The computed t value is

$$t_r = \left[\frac{r\hat{\rho}_{12|3\cdots u}^2}{1 - \hat{\rho}_{12|3\cdots u}^2} \right]^{1/2} \tag{11.42}$$

and the test is carried out as before.

Example 11.7. Consider the x_1 and x_2 coordinates of point P determined from three photographs (see Example 6.1). The covariance matrix is given in Example 11.6, and we may test if the correlation between x_1 and x_2 is significant. First,

$$\hat{\rho}_{12} = \frac{5.03}{(1.86 \times 71.33)^{1/2}} = 0.437$$

and since $r = 1$, then

$$t_1 = \frac{0.437}{[1 - (.437)^2]^{1/2}} = 0.49$$

At $\alpha = 0.05$, $t_{0.025, 1} = 12.706$, and thus x_1 and x_2 can safely be assumed to be uncorrelated.

Example 11.8. For this example recall the problem of the five-point level net solved by the method of indirect observations in Example 7.11. Recalling $\mathbf{Q}_{xx}$ as $\mathbf{N}^{-1}$, first compute the estimates of the (marginal) correlation coefficients $\hat{\rho}_{ij}$ according to equation (11.35),

$$\mathbf{Q}_{xx} = \mathbf{N}^{-1} = \begin{bmatrix} 0.806 & 0.445 & 0.505 & 0.333 \\ & 0.645 & 0.504 & 0.414 \\ & & 0.999 & 0.505 \\ \text{symmetric} & & & 0.742 \end{bmatrix} \quad \text{and}$$

$$\{\hat{\rho}_m\} = \begin{bmatrix} . & 0.617 & 0.563 & 0.436 \\ & . & 0.628 & 0.606 \\ & & . & 0.594 \\ & & & . \end{bmatrix}$$

The inverse of $\mathbf{Q}_{xx}$ is $\mathbf{N}$ from which the partial correlation coefficients $\hat{\rho}_p$ may be computed according to equation (11.41),

$$N = \begin{bmatrix} 2.185 & -1.064 & -0.568 & 0 \\ & 3.492 & -0.714 & -1.010 \\ & & 2.023 & -0.741 \\ \text{symmetric} & & & 2.476 \end{bmatrix} \quad \text{and}$$

$$\{\hat{\rho}_p\} = \begin{bmatrix} . & 0.385 & 0.270 & 0 \\ & . & 0.269 & 0.344 \\ & & . & 0.331 \\ & & & . \end{bmatrix}$$

The redundancy for the problem is $r = 4$. For the largest element (0.628) in $\{\hat{\rho}_m\}$, the computed t is 1.614; and for the largest element (0.385) in $\{\hat{\rho}_p\}$, the computed t is 0.834. Since $t_{4, 0.025} = 2.776$, no significant correlation exists between the elevations of the four points determined in the adjustment of the level net. It is worth noting that the correlation coefficients in $\{\hat{\rho}_m\}$ are generally higher than those in $\{\hat{\rho}_p\}$. As an example consider $\hat{\rho}_{m12} = 0.617$; it reflects not only the direct correlation between x_1 and x_2, but also the indirect effect of the correlation between x_1 and x_2, and x_3 and x_4. On the other hand $\hat{\rho}_{p12} = 0.385$ is correlation between x_1 and x_2 when x_3 and x_4

remain fixed. Therefore in this particular case the inclusion of the indirect effect (with x_3, x_4) increases the correlation between x_1 and x_2, though still not to a significant level at $\alpha = 0.05$.

In addition to the statistical analyses presented in this section, several others may be employed. These may include: (a) plotting the residuals versus the estimated observations to determine trends, (b) performing regression analysis on the residuals, (c) performing analysis of variance techniques on discrepancies at withheld control to compare different methods of analysis; and so forth. These and other advanced statistical techniques are beyond the scope of this book.

11.6. COMPUTATIONAL AND NUMERICAL CONSIDERATIONS

Once the mathematical model of an adjustment problem is designated and the appropriate equations formulated and linearized, the technique of least squares becomes a sequence of matrix operations. The algorithms written to perform these operations are usually automated in the sense of being carried out on a computer. Consequently, the application of least squares reduces at the end to computational matters. The larger the adjustment problem, the more important is the efficient performance of the computational algorithms. In fact, the full potential of the least squares method lies in applications involving very large adjustment systems. Although the treatment of computational methods is in the field of numerical analysis, a concise discussion of pertinent aspects is given in this section. More detailed information can be obtained from textbooks on numerical methods.

11.6.1. Selection of Adjustment Technique and Parameters The technique of adjustment of observations only ($Av = f$) was frequently chosen in the past for the adjustment of survey problems. The main reason lay in the small number of condition and normal equations (r) and the limited capacity of the computational aids available. With present computers the size of the normal equations is not such a totally overriding factor as it used to be. Instead, the adjustment with conditions containing parameters, particularly the technique of indirect observations wherever possible, is preferred. These techniques make the systematic handling of the adjustment data much easier, even though the resulting normal equations may frequently be larger in number than with the procedure of adjustment of observations only.

Another point that bears on the computational effort is the number of parameters carried in the adjustment. This aspect has been discussed in several places in previous chapters, mainly in regard to whether the parameters are required or not. Here the general remark is made that the ability to set up the computational algorithm systematically and optimally is more

important than attempting to carry the fewest number of parameters in the adjustment. The prime reason is the fact that in most cases the normal equations exhibit a special structure which when exploited leads to using considerably more economical computational techniques. Another consideration is to develop and apply computer adjustment programs that cover a wide range of problems. The effort involved in setting up good programs is such that one program should be applicable to a whole family of cases, even at the expense of not being fully optimized for smaller or simpler special cases. Separate programming of limited algorithms for special cases is not economical.

11.6.2. Computational Organization and Grouping of Variables Because the mechanics of matrix operations for least squares is known beforehand, it is usually not necessary to set up the matrices of the condition and constraint equations. Instead, we proceed to compute directly the elements of the normal equations. In fact, the partially reduced normal equations after eliminating the vectors $\mathbf{v}$, $\mathbf{k}$, and $\mathbf{k}_c$ (if it is present in the adjustment) are usually set up and solved for the unknown parameters. Although this approach requires, in general, special programming effort for certain types of adjustment problems, it leads to increased efficiency and savings in computer storage. This is particularly true when the observations, or certain subsets of the observations, are uncorrelated. The careful arrangement of the variables results in having zero submatrices and other submatrices that may be of diagonal or block diagonal form. The computational algorithm takes advantage of these particular characteristics (see the following section) particularly by not operating on the zero submatrices.

11.6.3. Structure of Normal Equations' Matrices When speaking of matrix structure of normal equations, we refer to the arrangement of zero and nonzero elements. It has been shown repeatedly before that the total system of normal equations contains a large number of zero submatrices. On the basis of such structure it has been possible to solve separately for the vectors $\mathbf{v}$, $\mathbf{k}$, Δ, and so on. In a similar manner the partially reduced normal equations, that is, those in terms of Δ, either have or can be made to have a useful pattern of zero and nonzero submatrices. An excellent example of this is the photogrammetric block triangulation by units (either bundles or models). For instance, Figure 11.2 depicts the structure of a partially reduced normal equations' coefficient matrix for bundle triangulation of a 4×5 block of aerial photographs. It can be readily seen that the matrix is sparse indeed, since the nonzero elements are much smaller in numbers than the zero elements. Furthermore, both submatrices $\dot{\mathbf{N}}$ and $\ddot{\mathbf{N}}$ are block diagonal, whereas $\bar{\mathbf{N}}$ has a banded structure. This particular pattern, which is obtained when cross-strip ordering of both photographs and points is done, is brought about by grouping the photoparameters together and separate from
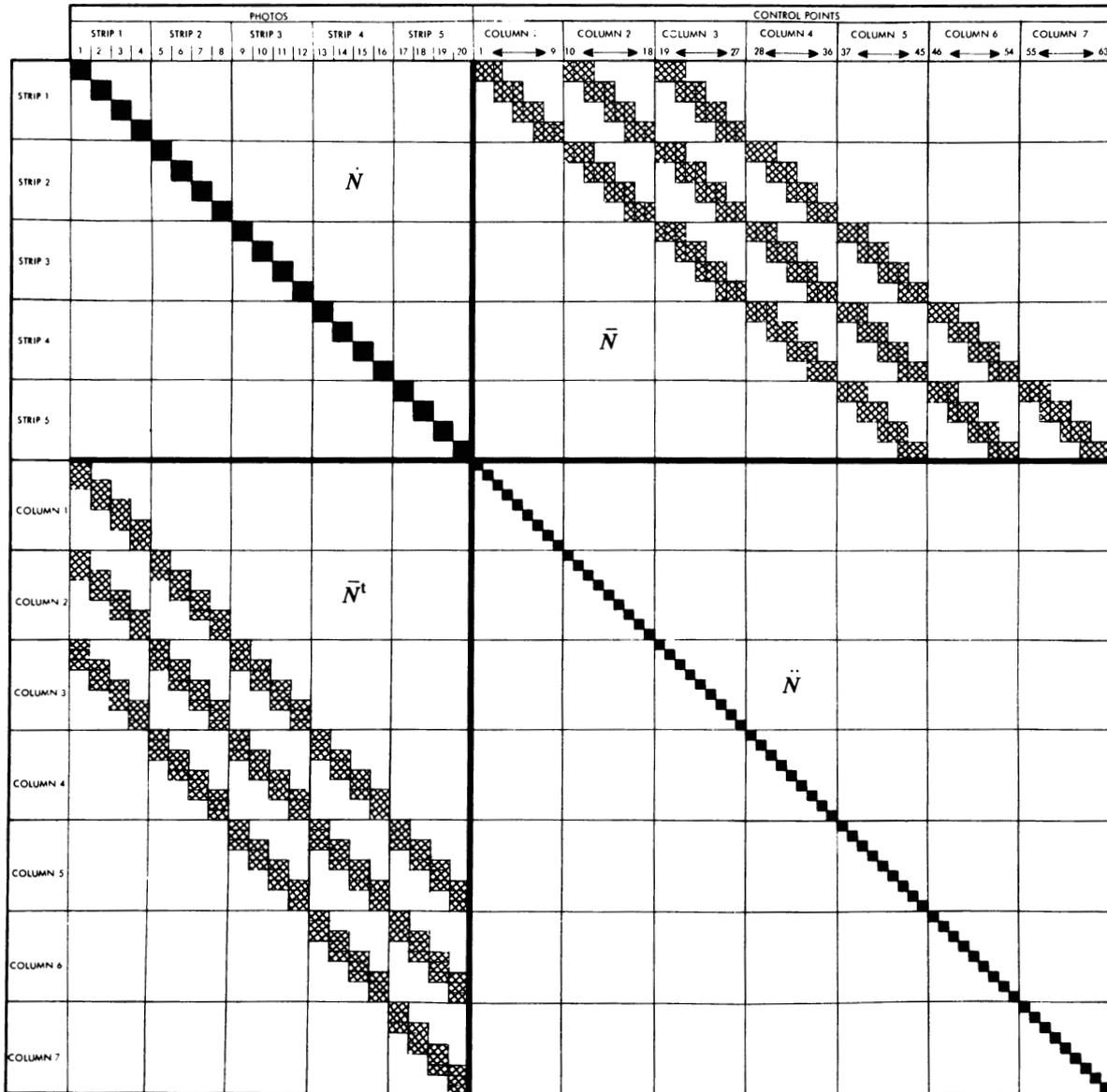
Figure 11.2. Structure of Normal Equations from Bundle Adjustment of 4 × 5 Photo Block of Photographs (From Brown, 1974. See Bibliography.)

the object point parameters. It therefore allows for the elimination of either of the two groups of unknowns by further reducing the normal equations. For example, the reduced normals in terms of photo parameters are $(\dot{N} - \bar{N}\ddot{N}^{-1}\bar{N}^t)$, and in terms of object point parameters are $(\ddot{N} - \bar{N}^t\dot{N}^{-1}\bar{N})$. Not only does either reduction reduce the size of system to be solved, but it also exploits the fact that both $\dot{N}$ and $\ddot{N}$ are block diagonal matrices with submatrices of small order. (For $\dot{N}$, the submatrix is usually of order 6, whereas $\ddot{N}$ has 3×3 submatrices along the main diagonal).

Although the reduction of the normal equations to only one group of parameters yields significant savings in the computational effort, further

study of the structure of reduced normals indicates that there are still a
significant number of zero submatrices. In fact the reduced normal equa-
tions' coefficient matrix is of the banded type where the band width is
dependent only on the number of strips in the block and is independent of
the length of the strips. Figure 11.3 shows four examples illustrating the
structure of the reduced normal equations' systèm for cross-strip ordering.
As the length of the strips is increased, keeping the number of strips
constant, the ratio of nonzero to zero elements decreases. Consequently,
devising an algorithm for solving a system that takes advantage of the
banded structure would make the triangulation of very large blocks both
feasible and economical. Such an algorithm will be discussed in the follow-
ing section.

The banded structure arises when small subsets of the parameters are not
related to each other either by actual condition or constraint equations, or
through stochastical correlation of the observations. The size of such subsets
determines the band width although the length of the band may be, in
principle, indefinitely large. There are situations, however, in which the ele-
ments of one subset of the parameters are related to all other parameters
through condition equations or otherwise. In this case the reduced normal
equations take the form

$$
\begin{bmatrix}
N_{11} & N_{12} & N_{13} & \cdots & N_{1n} \\
N_{12}^{t} & N_{22} & 0 & \cdots & 0 \\
N_{13}^{t} & 0 & N_{33} & \cdots & 0 \\
\vdots & \vdots & \vdots & & \vdots \\
N_{1n}^{t} & 0 & 0 & \cdots & N_{nn}
\end{bmatrix}
\begin{bmatrix}
\Delta_1 \\
\Delta_2 \\
\Delta_3 \\
\vdots \\
\Delta_n
\end{bmatrix}
=
\begin{bmatrix}
t_1 \\
t_2 \\
t_3 \\
\vdots \\
t_n
\end{bmatrix}
\qquad (11.43)
$$

where the coefficient matrix may be expressed schematically as†
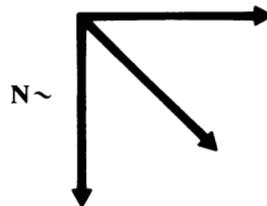
The sense of this scheme is that the horizontal and vertical bars generally
represent nonzero elements, the diagonal bar represents block diagonal non-
zero submatrices, and remaining region (blank) is formed of totally zero
elements. The arrows indicate that the total matrix N can extend indefinitely,
keeping the character the same as depicted. This system is referred to as
banded bordered. Brown† also names it *first-order partitioned* system because

† See Brown and Trotter in the Bibliography.

BW = 96

3 × 7 block

BW = 72

2 × 7 block

BW = 36

1 × 7 strip

BW = 120

4 × 7 block
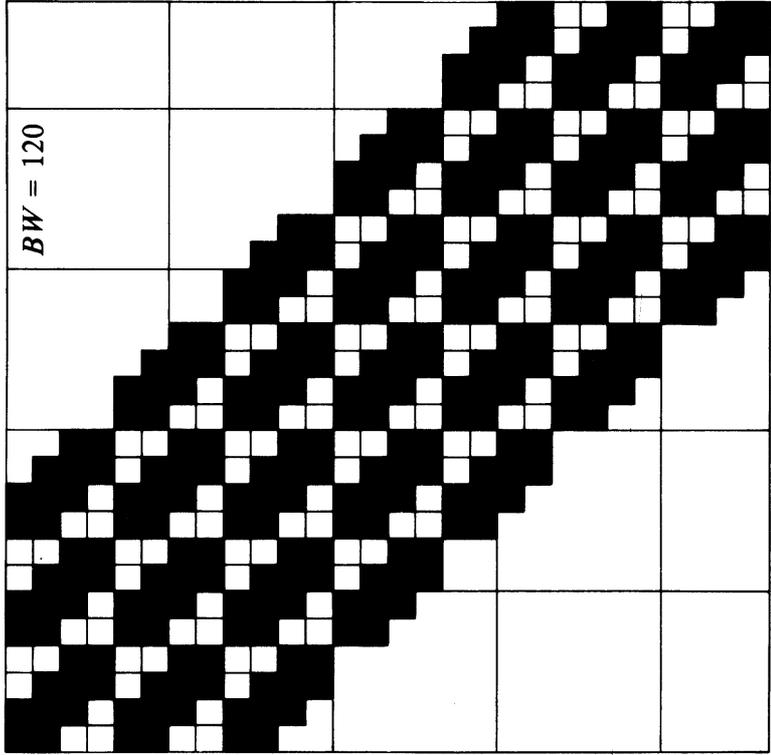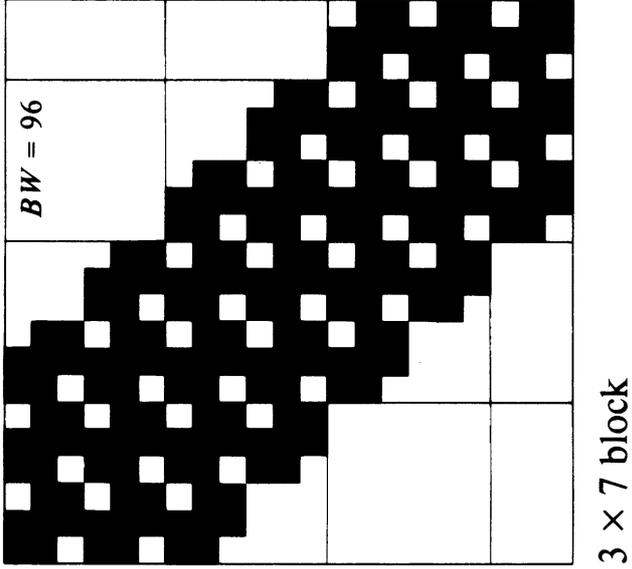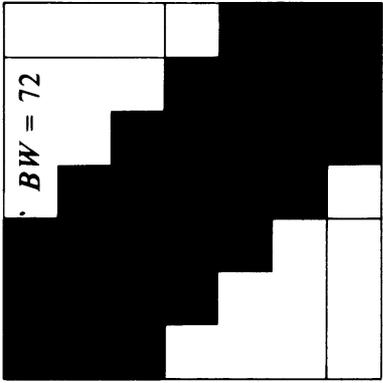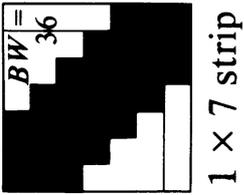
Figure 11.3. Structure of Reduced System of Normal Equations for a Succession of Uniform Blocks with Cross-Strip ordering. (From Brown, 1974. See Bibliography.)
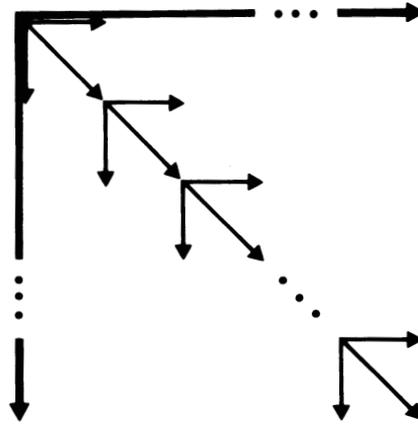
Figure 11.4. Schematic of Coefficient Matrix Characteristic of a Second-Order Partitioned System of Normal Equations (From Brown and Trotter, See Bibliography.)

of the way he extends it to higher-order systems. For example, the second- and third-order partitioned systems are depicted schematically in Figures 11.4 and 11.5.

When setting up the equations and arranging the parameters, the objective is to obtain as narrow a band as possible. As will be seen in the following section the computational time will depend mainly on the band width. As for the border, it may exist at the outset because of the conditions of the adjust-
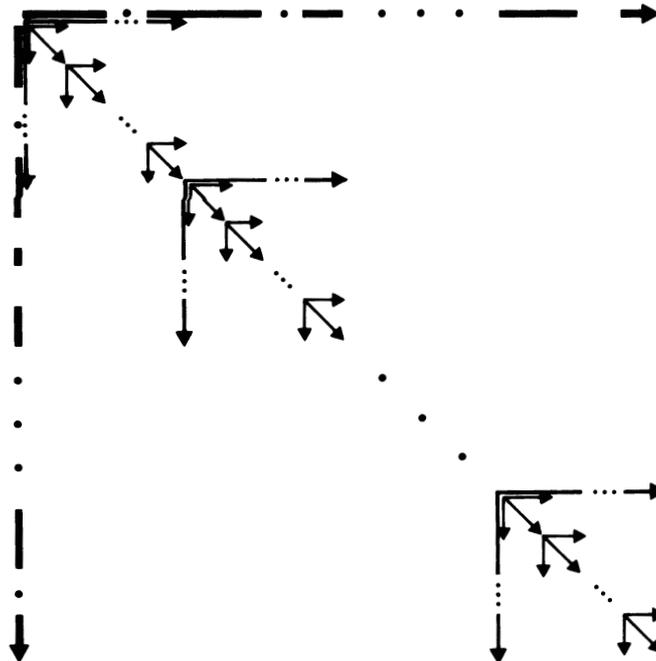
Figure 11.5. Schematic of Coefficient Matrix Characteristic of a Third-Order Partitioned System of Normal Equations (From Brown and Trotter. See Bibliography.)

ment problem or it may actually be introduced in order to minimize the band width. The latter case is common when constraints exist between parameters that are widely separated in the numbering scheme.

11.6.4. Algorithms for Solution of Normal Equations   All adjustment problems eventually come down to the solution of a system of linear equations— the reduced normal equations. In general, the larger the adjustment problem, the larger the system of equations to be solved. In addition to arranging the problem such that special matrix structure results, the selected algorithm for the solution must both exploit the structure and be as efficient as possible. In fact, the capacity and efficiency of the algorithm selected are often the limiting factors on the size of the adjustment system to be solved by a given computer configuration. Therefore it is quite important that the most suitable algorithm be selected for a given adjustment problem with a specific structure of normal equations.

Algorithms for the solution of systems of linear equations constitute a large segment of the activities of the numerical analysis field. We can only briefly comment therefore on the topic here and leave the details to the appropriate sources. There are, in general, two basic classes of algorithms: one group applying iterative processes and the other direct solution procedures. Of the iterative methods the following may be mentioned: block Gauss-Seidel, block Gauss-Seidel with acceleration, block successive over relaxation (BSOR), and conjugate gradient method. Many of these methods have been applied successfully in large adjustment programs, particularly BSOR and the conjugate gradient methods. Their stated advantages refer to required limited core storage, numerical acuity, and relative simplicity in programming. However, as with most iterative procedures, the rate of convergence is an important factor. It often depends on the particular circumstances of a given problem. For example, using BSOR for adjustment of photogrammetric blocks, it was found that the rate of convergence is dependent on the amount and disposition of available ground control. Although its convergence is satisfactory with suitable distribution of control, it is quite slow in cases of sparse control. For this reason and others, the tendency has been lately to apply direct closed solution of the system of reduced normal equations.

There are also many direct solution algorithms. In Appendix A an explanation is given of the classical procedure of Gauss elimination and the Gauss-Jordon method. Other direct techniques include the Cholesky method and the partitioning scheme which is often called the "bordering" technique. Some of these methods take advantage of the symmetry of the reduced normal equations' coefficient matrix. A relatively recent procedure, termed *recursive partitioning*, exploits the special structure of the matrices and is efficiently applied to both banded and banded-bordered matrices. It is therefore selected for further discussion.
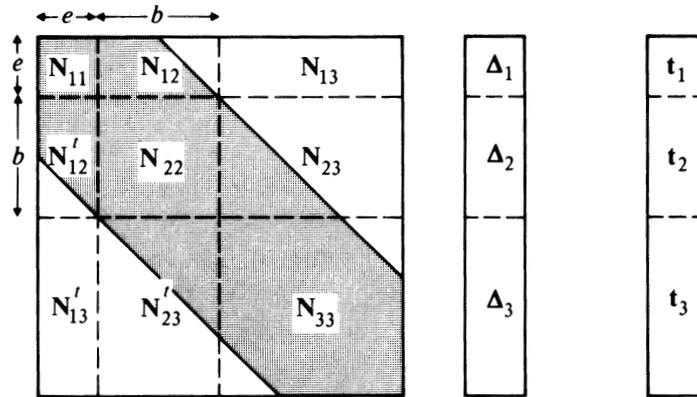
Figure 11.6

11.6.5. Recursive Partitioning†   Consider the banded system of Figure 11.6 with the triple partitioning shown, for which $e \leq b$. The number of elements in the second partition is b or the bandwidth. As shown in the figure this partitioning yields N_{13} (and N_{13}^t) as a null matrix. The vector Δ_1 may be eliminated using the method of partitioning (see Appendix A) thus

$$\left[\begin{bmatrix} N_{22} & N_{23} \\ N_{23}^t & N_{33} \end{bmatrix} - \begin{bmatrix} N_{12}^t \\ N_{13}^t \end{bmatrix} N_{11}^{-1} [N_{12}\ N_{13}]\right] \begin{bmatrix} \Delta_2 \\ \Delta_3 \end{bmatrix} = \begin{bmatrix} t_2 \\ t_3 \end{bmatrix} - \begin{bmatrix} N_{12}^t \\ N_{13}^t \end{bmatrix} N_{11}^{-1} t_1 \tag{11.44}$$

But since $N_{13} = 0$, then

$$\begin{bmatrix} N_{22} - N_{12}^t N_{11}^{-1} N_{12} & N_{23} \\ N_{23}^t & N_{33} \end{bmatrix} \begin{bmatrix} \Delta_2 \\ \Delta_3 \end{bmatrix} = \begin{bmatrix} t_2 - N_{12}^t N_{11}^{-1} t_1 \\ t_3 \end{bmatrix} \tag{11.45}$$

Comparing equation (11.45) to the original system (Figure 11.6) reveals that both the banded form and the bandwidth are preserved. The elimination of Δ_1 affects only the N_{22} portion of the coefficient matrix which lies entirely within the band. The system may be triply partitioned further and the procedure repeated recurrently, hence the name "recursive partitioning." At the end there will remain a system that is sufficiently small to be solved directly for the remaining unknowns. Once the elements of the last unknown vector are computed, a backward procedure can be carried out to obtain each of the unknown subvectors (the Δ_i) eliminated at the different steps of the forward process. As in the forward reduction, the bandwidth is preserved throughout the backward procedure. The computational efficiency of recursive partitioning stems from confining all the numerical operations to a submatrix (N_{22}) which is totally inside the band. The number of arithmetic operations required for the reduction of a $u \times u$ banded system of bandwidth b is proportional to $b^2 u$. By comparison, conventional Guass elimination requires a number of operations which is proportional to u^3. Thus

† See Brown, 1974, in the Bibliography.

recursive partitioning leads to considerable savings in computational effort particularly when the bandwidth b is much smaller than the matrix order u.

The method of recursive partitioning applies equally to banded-bordered systems such that the numerical operations are restricted to the nonzero elements within the band and border. In fact, the number of arithmetic operations required in the reduction of a system of bandwidth b and border-width b' is proportional to $(b + b')^2 u$, where u is the order of the system. Therefore, whenever possible, attempts should be made to keep both the bandwidth and border width as narrow as possible through prudent arrangement of parameters. This leads to the most economical computational algorithm.

It is frequently the practice that the selected algorithm is used to solve the system of reduced normal equations and not invert the coefficient matrix. This is particularly true when the original equations are nonlinear and relinearization is applied. During the last relinearization, however, we often seek the inverse of the coefficient matrix for purposes of a posteriori precision estimation. The reason for not seeking the inverse repeatedly is that it normally requires more computational time. For the particular algorithm of recursive partitioning it is possible to *generate only that portion of the inverse that corresponds to the original band* (which is primarily what is needed for a posteriori precision estimation). Nevertheless, the computation of only a portion of the inverse was found to require approximately twice the time needed for the solution of the system alone.

Problems for Part II

1. The angle α was measured with three different instruments as indicated below. Determine the "best value" of α and its standard deviation.

INSTRUMENT	α	σ
WILD T1 AE	45°27′20″	6″
WILD T16 ED	45°27′24″	9″
WILD T2	45°27′23″	3″

2. The three angles of a triangle were measured with three different instruments as indicated below. Determine the estimates of the angles that satisfy the condition that $\alpha + \beta + \gamma = 180°$.

INSTRUMENT	ANGLE	σ ANGLE
WILD T1 AE	$\alpha = 30°00′05″$	6″
WILD T16 ED	$\beta = 60°00′07″$	9″
WILD T2	$\gamma = 90°00′10″$	3″

3. A baseline was measured with four different measurement systems, yielding the following results:

MEASUREMENT VALUE (m)	STANDARD DEVIATION (cm)
1000.19	12
1000.40	13
1000.50	14
1000.28	15

a. Compute the weighted mean.
b. Compute the standard deviation of the weighted mean.

4. Given the simple level net where A is a reference bench mark whose elevation is, for simplicity, zero, the observed differences in elevation are given in *metres* as

$$l_1 = 4 \qquad l_2 = 2 \qquad l_3 = -6.3 \qquad l_4 = 5$$

Assuming that all the observations are uncorrelated and of equal weight, compute the elevation of point D and its cofactor.
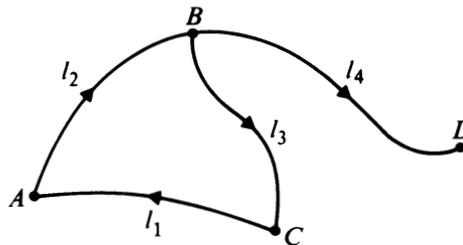
5. Assume that a gravimeter has been used to determine gravity differences along the lines indicated in the diagram below. Assume further, for simplicity, that all systematic corrections have been accounted for and that the terrain is level. The absolute value of gravity at point A has been established from pendulum observations as $g_A = 980000.0$ mgal. The observations (Δg) are uncorrelated and it has been determined empirically that the weight of Δg is given by

$$W_{\Delta g_i}\left(\frac{1}{\text{mgal}^2}\right) = \frac{1}{d} \qquad (d \text{ in } 10 \text{ km units})$$

Thus the following observations are available:

LINE	DISTANCE (km)	Δg (mgal)
1	18.1	+25.42
2	9.4	+10.34
3	14.2	−35.20
4	17.6	−15.54
5	13.5	+21.32
6	9.9	+ 4.82
7	13.8	−31.02
8	14.0	−26.11
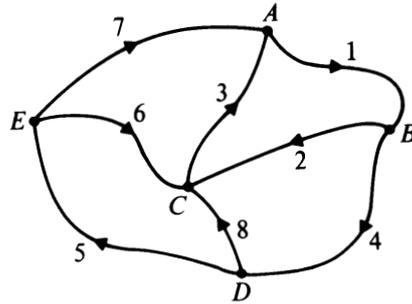
a. Using least squares adjustment of indirect observations compute the gravity values at points B, C, D, and E and their standard deviations.

b. Check the values obtained in (a) above using the method of least squares adjustment of observations only.

6. The figure shows the diagram for a system of measured angles about position 0. The azimuth of line $0A$ is considered known and fixed. It is desired to determine the adjusted azimuth of lines $0B$, $0C$, and $0D$.
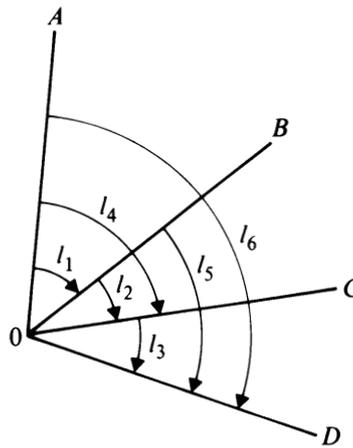
a. Considering application of the method of adjustment of indirect observations (i) How many condition equations should be written? (ii) Write all the condition equations, specifying the parameters.

b. Considering application of the method of adjustment of observations only: (i) How many condition equations should be written? (ii) Write all the condition equations and explain how the required azimuths and their covariance matrix will be determined.

7. Given the following observations with respect to the figure

$$l_1 = A\hat{O}B = 30°15'15'' \qquad \sigma_1 = 1''$$
$$l_2 = A\hat{O}C = 50°37'51'' \qquad \sigma_2 = 2''$$
$$l_3 = C\hat{O}D = 15°31'17'' \qquad \sigma_3 = 1''$$
$$l_4 = B\hat{O}D = 35°54'00'' \qquad \sigma_4 = 1''$$
$$l_5 = C\hat{O}E = 45°31'18'' \qquad \sigma_5 = 2''$$

Compute the angle θ $(D\hat{O}E)$ and its standard deviation σ_θ.
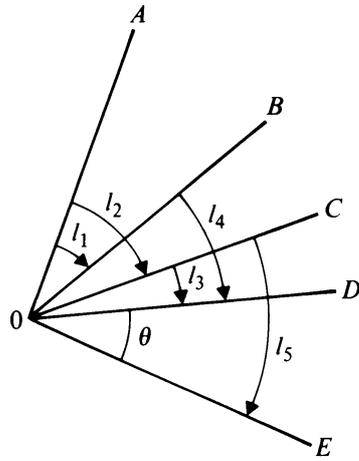
8. Given the following data with respect to the figure:

$$l_1 = A\hat{O}B = 30°15'15''$$

$$l_2 = B\hat{O}C = 20°00'00''$$

$$l_3 = A\hat{O}C = 50°15'18''$$

$$l_4 = C\hat{O}D = 30°00'00''$$

$$l_5 = B\hat{O}E = 70°00'01''$$

Assuming that all observations are *uncorrelated* and of equal precision, compute the least squares estimate of the angle *DOE and its cofactor*.
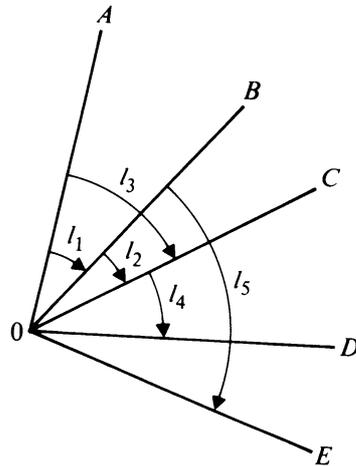
9. Given the following information with respect to the figure:

Direction $l_1 = 10°$	$\sigma_1 = 1''$
Direction $l_2 = 20°$	$\sigma_2 = 1''$
Direction $l_3 = 30°$	$\sigma_3 = 1''$
Direction $l_4 = 60°$	$\sigma_4 = 1''$
Angle $l_5 = 20°$	$\sigma_5 = 2''$

No correlation exists between the observations. Compute the angle θ and its standard deviation.
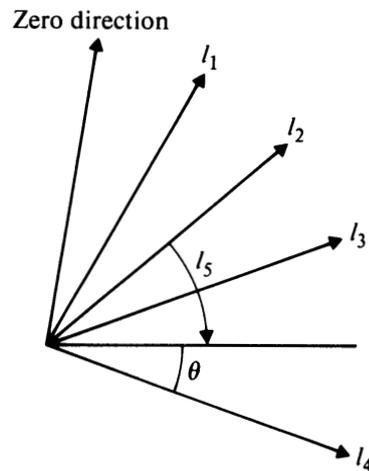
10. In the level circuit shown, the following elevation differences were observed:

Station 1   42.107 m above Station 5

Station 1   12.424 m above Station 2

Station 2   42.251 m above Station 3

Station 4   8.464 m above Station 3

Station 5   4.138 m above Station 4

Station 1   46.269 m above Station 4

Station 2   33.802 m above Station 4

Station 5 is a benchmark with known (assumed fixed) elevation of 500.000 m AMSL. Using the adjustment of indirect observations method compute the elevations of stations 1 through 4 and their a posteriori cofactor matrix. Assume that the variances
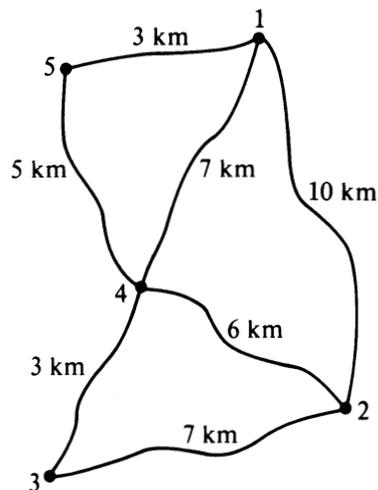
in the observations are directly proportional to the distance leveled (as indicated on the sketch).

11. Solve Problem 10 using the method of adjustment of observations only.

12. Set up each of the following problems such that a least squares solution can be determined. Show the steps used to get to the final equations, specifying the dimensions on all matrices used.

a. The following 11 angles have been observed in a pentagon as shown.
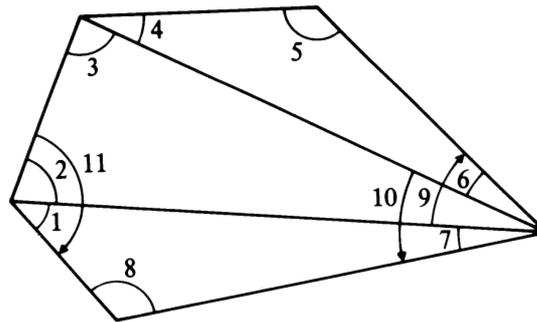
b. The following 16 elevation differences have been observed in a level net as shown. Stations A, B, and C are considered to be of known elevation.
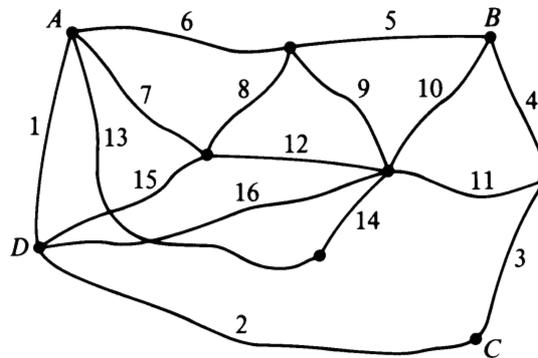
13. The figure depicts a simplified problem of intersection in a plane. The three points a, b, and c are perfectly known control with x y coordinates as follows:

POINT	x	y
a	2.00	19.32
b	2.00	2.00
c	19.32	2.00

An *unknown* point P is to be determined such that its x coordinate is the same as that
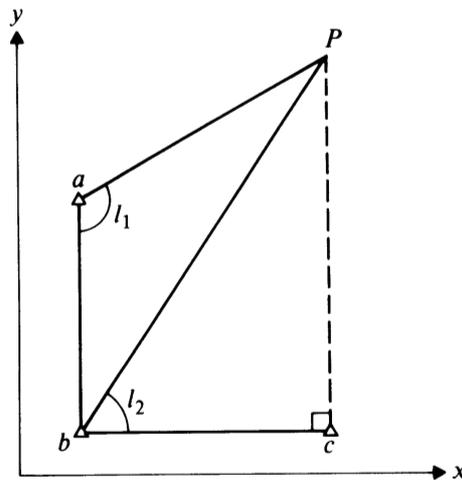
for point c (that is, angle bcP is perfectly known to be 90 degrees). The observations are the two angles shown in the figure,

$$l_1 = 120° \quad \text{and} \quad l_2 = 57°40'$$

where l_1 and l_2 are uncorrelated and of equal weight. Compute the position of point P. (If an approximate value is needed use $y_p^\circ = 29.00$.)

14. In the figure, A and B are *known* horizontal control points. It is desired to determine the position of point P using the three observations shown. Give the *linearized* mathematical model for the adjustment of the problem to obtain the coordinates of point P.
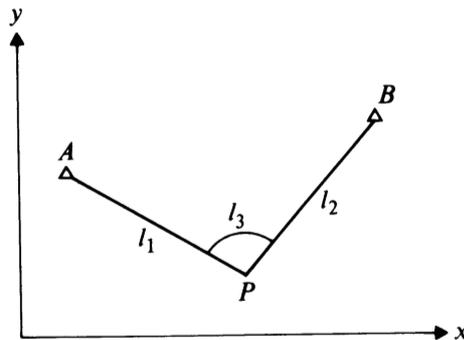
15. The figure shows two *known* horizontal control points A and B. To determine the coordinates of point P, the two lengths l_1 and l_3 and the two angles l_2 and l_4 are
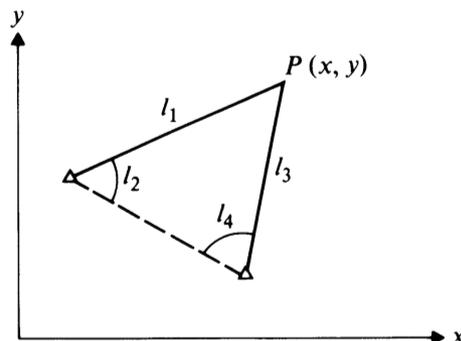
measured. Set this problem up for an adjustment, showing in detail the makeup of the appropriate matrices.

16. For the plane triangle shown, the following quantities are measured:

Angle $A = l_1 = 59°59'58''$ $\sigma_1 = 5''$

Angle $B = l_2 = 59°59'51''$ $\sigma_2 = 5''$

Angle $C = l_3 = 59°59'54''$ $\sigma_3 = 5''$

side $AC = l_4 = 1000$ m $\sigma_4 = 0.05$ m.

Compute the coordinates of point B with respect to the coordinate system shown, and determine their covariance matrix.
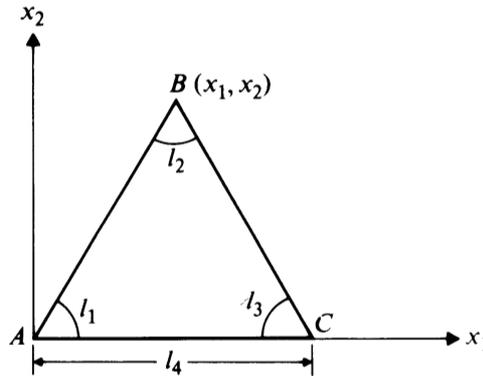
17. The three sides and three angles of a plane triangle were observed as follows:

Sides $l_1 = 1400$ ft $\sigma_1 = 0.01$ ft

Side $l_2 = 1100$ ft $\sigma_2 = 0.02$ ft

Side $l_3 = 1700$ ft $\sigma_3 = 0.01$ ft

Angle $l_4 = 39°23'40''$ $\sigma_4 = 1''$

Angle $l_5 = 88°33'05''$ $\sigma_5 = 1''$

Angle $l_6 = 52°03'17''$ $\sigma_6 = 1''$
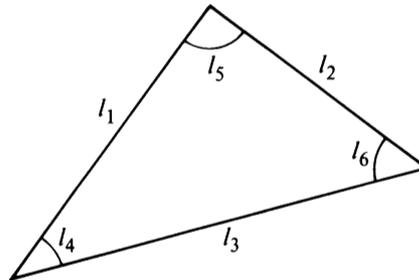
Compute the least squares estimates of the six observations, their a posteriori cofactor matrix, and the a posteriori estimate of the reference variance.

18. The figure shows a simple intersection problem to determine the horizontal position $(X, Y)_p$ of the unknown point P. The coordinates $(X, Y)_a$, $(X, Y)_b$, and $(X, Y)_c$ of the three horizontal control points A, B, and C are *known*. If the four angles l_1, l_2, l_3, and l_4 are observed, give the steps for determining the coordinates of point P by least squares.
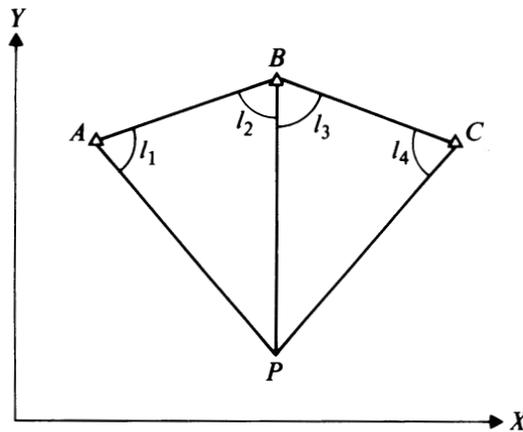
19. The figure shows four (4) positions separated by six (6) measured distances. Assuming station 1 fixed and the Y coordinate of 2 equal to the Y coordinate of 1, *develop* the appropriate equations for a least squares solution to determine the coordinates of points 2, 3, 4, and their covariance matrix. State clearly any assumptions you make.
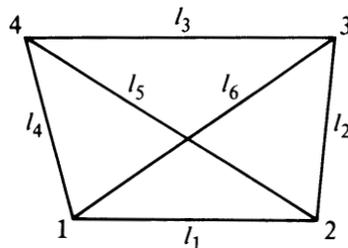
20. Given the triangle shown in the figure with four observed values as follows:

Side AB $= l_1 = 1.1$ m

Side BC $= l_2 = 0.9$ m

Side CA $= l_3 = 1.0$ m

Angle $BAC = l_4 = 60°$

Assuming that the approximate value of the angle ABC is 60 degrees, compute the least squares estimate of the correction δx. Assume that all observations are uncorrelated and of equal weight.
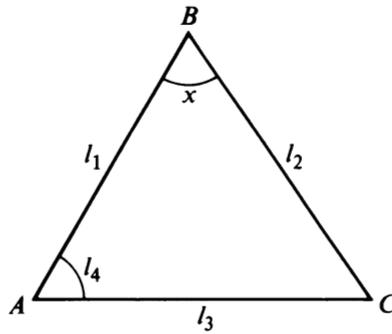
21. Instead of the information given in Problem 20, you are given the information shown in the figure where

Angle $ABC = l_1 = 60°02'$

Angle $BCA = l_2 = 60°01'$

Side $CA \quad = l_3 = 1.0$ m

Angle $BAC = l_4 = 60°00'$
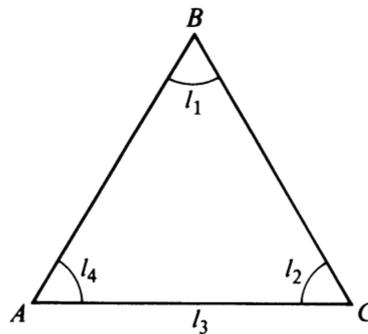
All observations are uncorrelated and of equal weight. How would you compute the least squares estimate of angle ABC? Compare the two adjustment procedures for the two problems.

22. The *positions of all* points in the figure *with respect to the coordinate system shown* are required. The angle α is perfectly known (a constant). All three sides l_1, l_2,
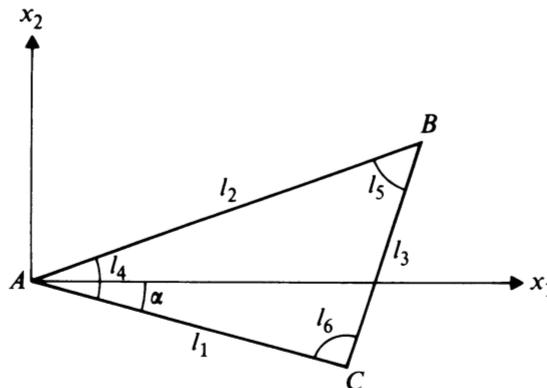
and l_3 and all three interior angles l_4, l_5, and l_6 are measured with an a priori cofactor matrix $\mathbf{Q}$.
$$\underset{6,6}{}$$
a. Set up the mathematical model giving all pertinent information (for example, n, n_0, r, u, c, ..., etc.). Write the equations in detail. If they are nonlinear, explain how you would linearize them. Show how the desired coordinates will be obtained.
b. What would change in the problem if α is also an observation?
c. What would change in the problem if α is not given, that is, if you only have $l_1, \ldots, l_6$?

23. The figure presents a simple triangle in which the vertices Q and S are horizontal control stations with known X, Y coordinates. As shown, the measurements are l_1, l_2, l_3, l_4, l_5, and l_6. Set up the least squares adjustment solution for the determination of the horizontal position of station R.

Azimuth mark
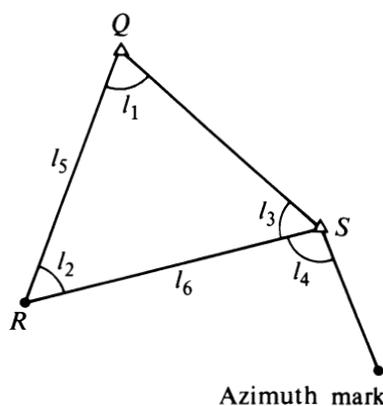
24. a. In the figure the four observations shown are given. Explain the adjustment procedure you would use to determine the *shape* of the figure.
b. If the interior angle at C, which is l_5, is also a given observation, show in detail any change to your answer to (a) above.
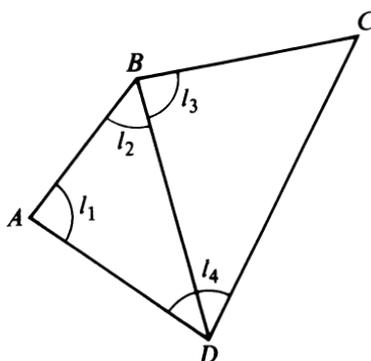

25. The figure shows a *cartesian* coordinate system xy with origin at point B. Points A and C are *known control* points on the y and x axes, respectively. Point D is of unknown coordinates. The observations are the three angles l_1, l_2, and l_3 shown in the figure. Determine n, n_0, r, and then
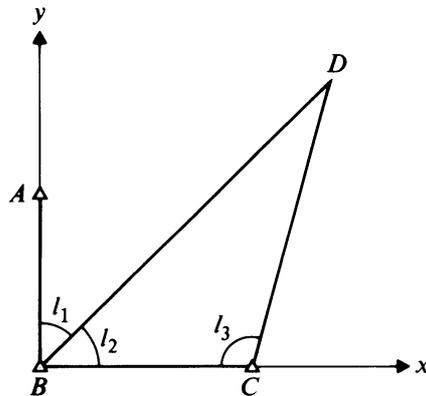
a. For $Av = f$ write the condition equations (no linearization).

b. For $Av + B\Delta = f$ write the condition equations necessary for obtaining the estimates of the coordinates of point D. (No linearization.) State explicitly any assumptions you make.

26. The figure shows a circle for which the line AD is a *diameter*, the estimate of which is desired. Because point D could not be occupied, two other points (B and C) are selected on the circumference. The angles l_1 and l_2 and the distances l_3 and l_4 are measured. *Show in complete detail*, giving the appropriate condition equations, how the method of least squares may be used to estimate $\overline{AD}$ and its variance. Specify any assumptions you may make.
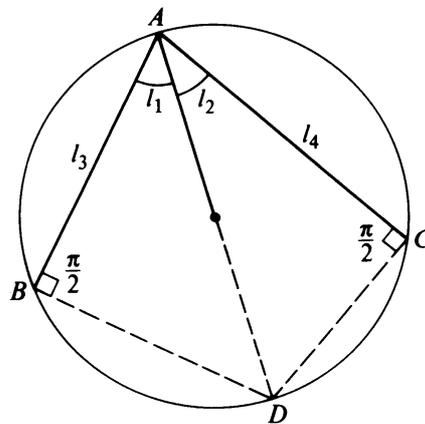
27. The figure shows a photogrammetric model that has been leveled and approximately scaled. All that remains is final scaling. Two distances are *measured* in the model which are $l_1 = 300$ mm and $l_2 = 400$ mm. The corresponding *known* control distances are $D_1 = 305$ mm and $D_2 = 410$ mm, at the nominal model scale. The current value of the base, that is, that corresponding to l_1 and l_2, is $b = 200$ mm.

a. Compute the final value of the base B that, when introduced, would complete the scaling of the model.

b. Compute the observational residuals in l_1 and l_2.

c. Check to verify that the mathematical model is satisfied after the adjustment. Assume all observations to be uncorrelated and of equal weight.
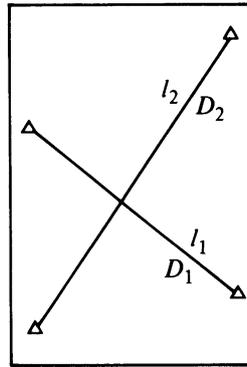
28. The figure shows a photogrammetric model that is scaled and approximately leveled. The four points 3, 4, 5, and 6 are given vertical control points. If Z denotes ground elevation in feet, and z denotes model elevation in feet, let the following values represent the *observations*:

$$l_3 = Z_3 - z_3 = +4 \text{ ft}$$

$$l_4 = Z_4 - z_4 = +2 \text{ ft}$$

$$l_5 = Z_5 - z_5 = +1 \text{ ft}$$

$$l_6 = Z_6 - z_6 = -5 \text{ ft}$$
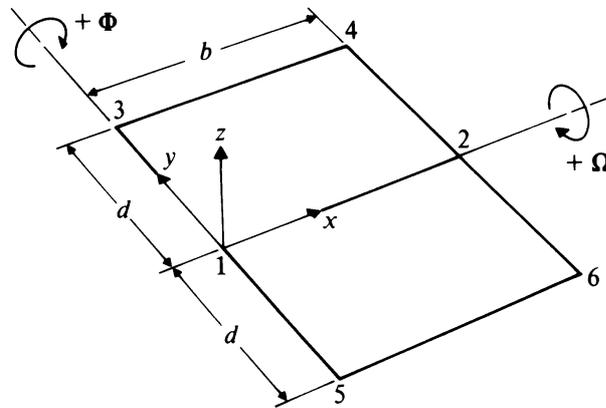
If these observations are uncorrelated and of equal weight, compute the estimated observations using the method of least squares adjustment of observations only.

29. In a chemistry lab a student desired to find the radius of the inside of a glass sphere. He filled the glass sphere with mercury and carefully poured the mercury into a graduated vial from which he read the volume in cubic centimetres. He observed readings of 31.9, 32.2, and 32.0. He considered that the second measurement was about one half as good as the first and the third measurement about one third as good as the first. Considering these as weights, determine what the best estimate of the radius is *and* how good that estimate is.

(Hint: Adjustment of observations only would probably be easiest.)

30. a. Find the best-fitting straight line to the following data:

X (CONSTANT)	Y (OBSERVATIONS)
-4	-1.9
-2	-0.2
0	2.0
2	4.2
4	6.1

The observations are assumed to be independent and have variances of 0.05.
b. Does the a posteriori estimate of reference variance seem reasonable?

31. The equation of a plane is usually given as

$$Ax + By + Cz + D = 0$$

We know that three points determine a plane. Therefore only three of these parameters are necessary. Thus we can write the equation as

$$ax + by + cz + 1 = 0$$

Given the x, y, z coordinates of the four corners of a building from a stereo pair of photographs, and assuming that these four points must lie in a plane,
a. Show that the normal equations for the determination of the best-fitting plane can be formed by a summation process. Be sure to dimension all matrices. Assume

$$\mathop{\mathbf{Q}}_{12,\,12} = \mathbf{I}$$

b. Using as approximations $a^0 = 1$, $b^0 = 1$, and $c^0 = -1$, find the best equation for that plane. Do *one* iteration only.

POINT	MODEL COORDINATES		
	X	Y	Z
1	1.1	-1.0	0.9
2	-2.0	2.0	1.0
3	2.0	-2.0	1.0
4	-1.1	1.0	0.9

32. Two rectangular coordinate systems, A and B, are related through translation and rotation. For the two-dimensional case in transforming from system A to system B, the following expressions apply:

$$x_b = a_0 + a_1 x_a - a_2 y_a$$

$$y_b = a_3 + a_2 x_a + a_1 y_a$$

Measurements on five points in each coordinate system were recorded as follows:

POINT	x_a	y_a	x_b	y_b
1	2.020	4.107	8.457	16.740
2	5.132	1.098	12.472	15.292
3	0.080	6.204	5.863	17.865
4	7.483	0.109	15.155	15.367
5	4.206	8.128	8.818	21.333

Solve for the transformation coefficients a_0, a_1, a_2, and a_3 assuming that all measurements are uncorrelated and of equal precision.

33. The outcome of an experiment is to be plotted with respect to an x-y system of coordinates. The values of the ordinate y are *observations* taken at equal intervals of x. The following table lists the data:

| x	y	σ_y
2	2.1	0.05
4	3.9	0.05
6	6.2	0.05
8	8.1	0.05

If a straight line with the equation:

$$y = ax + b$$

is to be fitted to the data, compute the least squares estimates for a and b and their covariance matrix.

34. A parabola with the equation $y^2 = ax$ is to be fitted to the two points $(1, 2)$ and $(2, 3)$. Considering that the y values are the observations, which are uncorrelated and of equal precision, compute the *first* correction to the approximate value of a, $a^0 = 4.0$.

35. The figure presents a traverse situation in which the endpoints A and D are horizontal control stations with known X, Y coordinates. B and E are the given
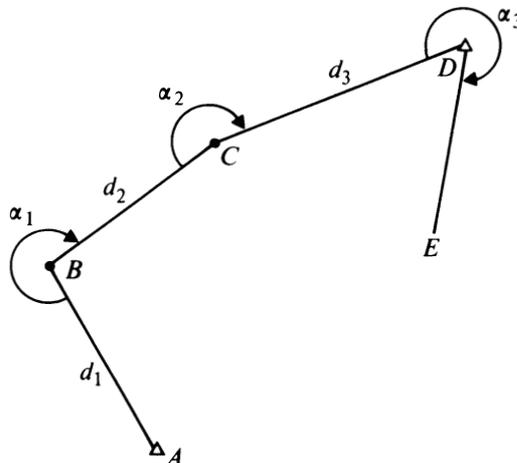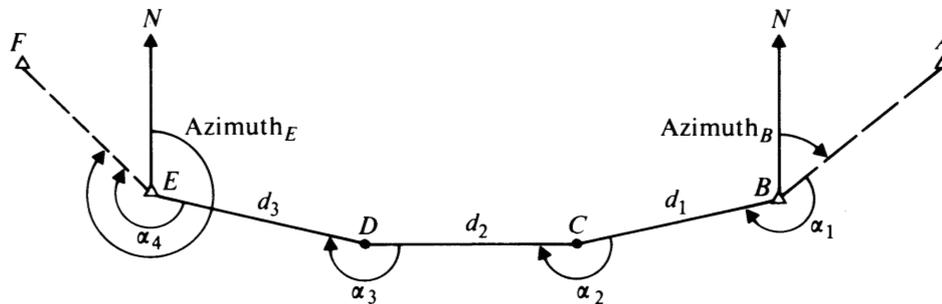
azimuth marks for stations A and D, respectively. Set up the least squares adjustment solution for the determination of the horizontal position of stations B and C, given that stations B, C, and D are occupied to measure α_1, α_2, α_3, d_1, d_2, and d_3.

36. For the traverse shown, the following angle and distance measurements were recorded with their associated standard deviations.

ANGLE	σ	DISTANCE	σ
$\alpha_1 = 172°53'34''$	$2''$	$d_1 = 281.832$ m	0.016 m
$\alpha_2 = 185°22'14''$	$2''$	$d_2 = 271.300$ m	0.016 m
$\alpha_3 = 208°26'19''$	$2''$	$d_3 = 274.100$ m	0.016 m
$\alpha_4 = 205°13'51''$	$2''$		

| POINT	COORDINATES		
	X (m)	Y (m)	AZIMUTH
B	8478.139	2483.826	$68°15'20''7$
E	7709.336	2263.411	$300°11'30''5$

a. Compute the coordinates of points C and D and their respective covariance matrices.

b. Compute the adjusted estimates of the observed angles and distances.

c. Compute the estimate $\hat{\sigma}_0^2$.

d. Compute the covariance matrix for the estimated observations.

37. In the figure $b\hat{0}c = \theta_2$ and $c\hat{0}d = \theta_1$ are unknown parameters. The measurements l_1, l_2, l_3, and l_4 are given by

$$l = \begin{bmatrix} d\hat{0}e \\ a\hat{0}b \\ c\hat{0}e \\ a\hat{0}c \end{bmatrix} = \begin{bmatrix} 51° \\ 38° \\ 80° \\ 98° \end{bmatrix}$$

Let $\theta_1^0 = 29$ and $\theta_2^0 = 60$. Take the covariance matrix of the observations to be the identity matrix

a. Find θ_1 and θ_2.

b. It is known that angle $b0d$ is a right angle. Find θ_1 and θ_2 under this *constraint*.

$$a \hat{o} e = \pi$$
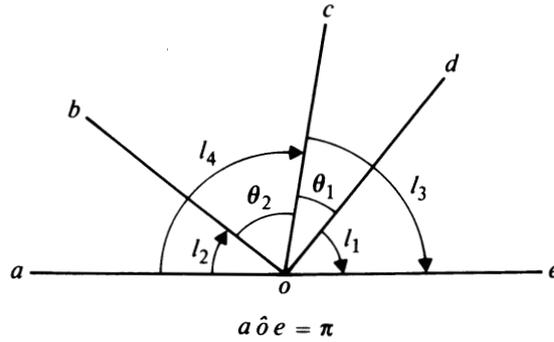
38. A parabola with the equation $y^2 = ax$ is to be fitted to the following data:

| POINT	(1)	(2)	(3)
x	1	2	3
y	2	3	3.5

Consider the y *values as observations* that are uncorrelated and of equal weight.

a. With an approximate value $a^0 = 4$ compute the correction δa from *first iteration only*.

b. What would the solution be if the parabola is *constrained* to pass through point (1) precisely?

39. The figure shows two *perfectly* known points: $A(0, 2)$ and $B(0, 5)$. Two other points 1 and 2 are to be located in the x, y plane through measuring the three distances: $l_1 = \overline{B1} = 10.1$; $l_2 = \overline{B2} = 5.1$; $l_3 = \overline{A2} = 4.0$ and $Q = I$. Also, point 1 has a fixed y coordinate $= 5.0$.

a. Compute the three coordinates x_1, x_2, y_2.

b. If points 1 and 2 are constrained to lie on the line $y = mx$, write the *conditions* and *constraints* giving the numerical values of all matrices. Then eliminate the constraints and give the numerical values of the matrices $\bar{B}$ and $\bar{f}$ of the modified conditions. Use reasonable approximations such as $x_1^0 = 10$, $x_2^0 = 4$, and $y_2^0 = 2$.
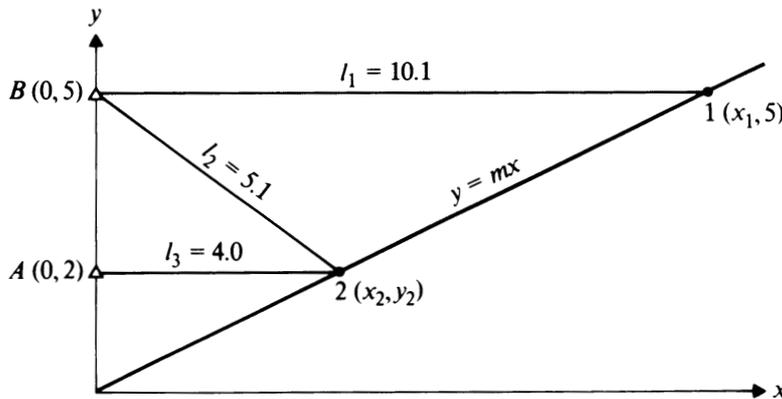
40. Given two coordinate systems $\mathbf{x} = [x_1\, x_2]^t$, $\mathbf{y} = [y_1\, y_2]^t$, and the following three points with measured coordinates in both systems (take a priori $\mathbf{Q} = \mathbf{I}$)

POINT	x_1	x_2	y_1	y_2
1	576.889	286.144	738.284	372.541
2	598.213	449.326	745.184	561.693
3	608.554	597.625	741.096	732.570

The transformation between the two systems is

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = s \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}$$

where s is a scale factor and a, b, c, d, t_1, t_2 are six parameters. Compute the least squares estimates for these seven parameters applying appropriate constraints such that the transformation is in effect the four-parameter linear conformal in two dimensions [equation (8.6)]. Perform the adjustment with the four-parameter transformation and ascertain that the results from both procedures are the same.

41. In the figure, the four observations shown are given. If the four points A, B, C, and D fall on the circumference of a circle, write the elements of the functional model for the least squares adjustment for determining the shape of the figure.
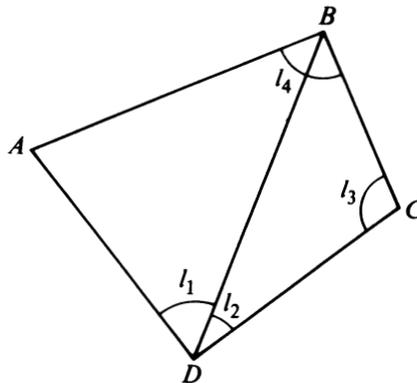
42. The figure shows a large *rectangular* area divided into two *rectangular* areas (that is, all angles are perfectly known). The observations, which are uncorrelated and of equal precision, are

$$l_1 = 2.00 \qquad l_2 = 3.00 \qquad l_3 = 3.10 \qquad l_4 = 1.00 \qquad l_5 = 1.94$$

a. Compute the least squares estimates of the areas X_1 and X_2 and their cofactors.
b. Under the added information of constraining X_1 to be $3X_2$: (i) Write all the elements of the functional model giving the condition and constraint equations; if necessary, linearize these equations and give the numerical values of all matrices involved. (ii) Eliminate the constraints and give the numerical values of the matrices
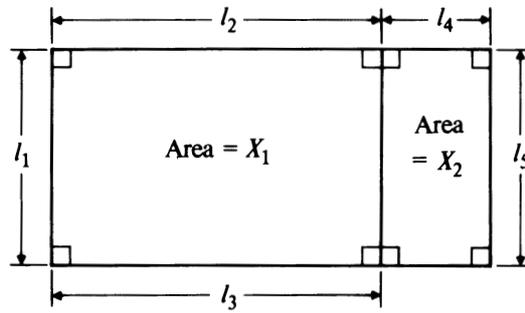
involved in remaining equations. (iii) Eliminate the parameters left after part (ii) above and give the numerical values of the matrices involved in the remaining equations.