Solutions to Selected Problems in

NUMERICAL OPTIMIZATION

by J. Nocedal and S.J. Wright

Second Edition

Solution Manual Prepared by:

Frank Curtis
Long Hei
Gabriel López-Calva
Jorge Nocedal
Stephen J. Wright

Contents

1 Introduction

No exercises assigned.

2 Fundamentals of Unconstrained Optimization

Problem 2.1

$$
\begin{aligned}
\frac{\partial f}{\partial x_1} &= 100 \cdot 2(x_2 - x_1^2)(-2x_1) + 2(1-x_1)(-1) \\
&= -400x_1(x_2 - x_1^2) - 2(1-x_1) \\
\frac{\partial f}{\partial x_2} &= 200(x_2 - x_1^2) \\
\Longrightarrow \nabla f(x) &= \begin{bmatrix} -400x_1(x_2 - x_1^2) - 2(1-x_1) \\ 200(x_2 - x_1^2) \end{bmatrix} \\
\frac{\partial^2 f}{\partial x_1^2} &= -400[x_1(-2x_1) + (x_2 - x_1^2)(1)] + 2 = -400(x_2 - 3x_1^2) + 2 \\
\frac{\partial^2 f}{\partial x_2 \partial x_1} &= \frac{\partial^2 f}{\partial x_1 \partial x_2} = -400x_1 \\
\frac{\partial^2 f}{\partial x_2^2} &= 200 \\
\Longrightarrow \nabla^2 f(x) &= \begin{bmatrix} -400(x_2 - 3x_1^2) + 2 & -400x_1 \\ -400x_1 & 200 \end{bmatrix}
\end{aligned}
$$

1. $\nabla f(x^*) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and $x^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is the only solution to $\nabla f(x) = 0$

2. $\nabla^2 f(x^*) = \begin{bmatrix} 802 & -400 \\ -400 & 200 \end{bmatrix}$ is positive definite since $802 > 0$, and $\det(\nabla^2 f(x^*)) = 802(200) - 400(400) > 0$.

3. $\nabla f(x)$ is continuous.

(1), (2), (3) imply that x^* is the only strict local minimizer of $f(x)$.

$\square$

Problem 2.2

$$
\begin{aligned}
\frac{\partial f}{\partial x_1} &= 8 + 2x_1 \\
\frac{\partial f}{\partial x_2} &= 12 - 4x_2 \\
\Longrightarrow \nabla f(x) &= \begin{bmatrix} 8+2x_1 \\ 12-4x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\
& \text{One solution is} \quad x^* = \begin{pmatrix} -4 \\ 3 \end{pmatrix}. \\
& \text{This is the only point satisfying the first order necessary conditions.}
\end{aligned}
$$

$$
\nabla^2 f(x) = \begin{bmatrix} 2 & 0 \\ 0 & -4 \end{bmatrix} \quad \text{is not positive definite, since } \det(\nabla^2 f(x)) = -8 < 0.
$$

Therefore, x^* is **NOT** a minimizer. Consider $\min(-f(x))$. It is seen that $\nabla^2[-f(x)]$ is also not positive definite. Therefore x^* is **NOT** a maximizer. Thus x^* is a saddle point and only a stationary point.

□

The contour lines of $f(x)$ are shown in Figure 1.

Problem 2.3

(1)

$$
\begin{aligned}
f_1(x) &= a^T x \\
&= \sum_{i=1}^{n} a_i x_i
\end{aligned}
$$

$$
\begin{aligned}
\nabla f_1(x) &= \begin{bmatrix} \frac{\partial f_1}{\partial x_1} \\ \dots \\ \frac{\partial f_1}{\partial x_n} \end{bmatrix} = \begin{bmatrix} a_1 \\ \dots \\ a_n \end{bmatrix} = a \\
\nabla^2 f_1(x) &= \begin{bmatrix} \frac{\partial^2 f_1}{\partial x_1^2} & \frac{\partial^2 f_1}{\partial x_2 \partial x_1} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} = \left[\frac{\partial^2 \sum_i a_i x_i}{\partial x_s \partial x_t} \right] \begin{matrix} s = 1 \cdots n \\ t = 1 \cdots n \end{matrix} = 0
\end{aligned}
$$

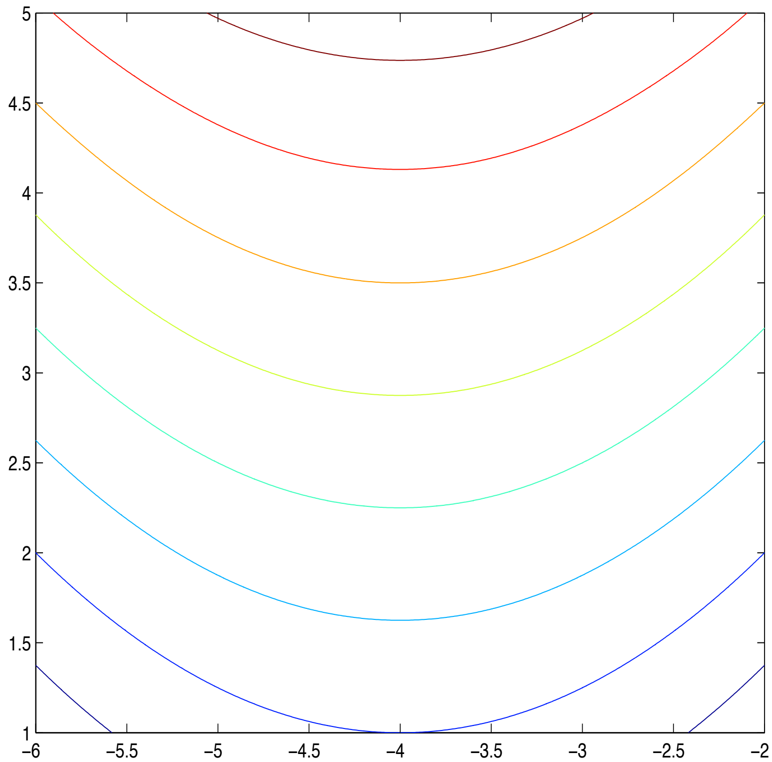

Figure 1: Contour lines of $f(x)$.

(2)

$$\begin{aligned}
f_2(x) &= x^T A x = \sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}x_i x_j \\
\nabla f_2(x) &= \left[\frac{\partial f_2}{\partial x_s}\right]_{s=1\cdots n} = \left[\sum_j A_{sj}x_j + \sum_i A_{is}x_i\right]_{s=1\cdots n} \\
&= \left[2\sum_{j=1}^{n} A_{sj}x_j\right]_{s=1\cdots n} \qquad \text{(since } A \text{ is symmetric)} \\
&= 2Ax
\end{aligned}$$

$$\begin{aligned}
\nabla^2 f_2(x) &= \left[\frac{\partial^2 f_2}{\partial x_s \partial x_t}\right]_{\substack{s=1\cdots n\\ t=1\cdots n}} = \left[\frac{\partial^2 \sum_i\sum_j A_{ij}x_i x_j}{\partial x_s \partial x_t}\right]_{\substack{s=1\cdots n\\ t=1\cdots n}} \\
&= \left[A_{st} + A_{ts}\right]_{\substack{s=1\cdots n\\ t=1\cdots n}} = 2A
\end{aligned}$$

□

Problem 2.4

For any univariate function $f(x)$, we know that the second oder Taylor expansion is

$$f(x+\Delta x) = f(x) + f^{(1)}(x)\Delta x + \frac{1}{2}f^{(2)}(x+t\Delta x)\Delta x^2,$$

and the third order Taylor expansion is

$$f(x+\Delta x) = f(x) + f^{(1)}(x)\Delta x + \frac{1}{2}f^{(2)}(x)\Delta x^2 + \frac{1}{6}f^{(3)}(x+t\Delta x)\Delta x^3,$$

where $t \in (0,1)$.

For function $f_1(x) = \cos(1/x)$ and any nonzero point x, we know that

$$f_1^{(1)}(x) = \frac{1}{x^2}\sin\frac{1}{x}, \quad f_1^{(2)}(x) = -\frac{1}{x^4}\left(\cos\frac{1}{x} + 2x\sin\frac{1}{x}\right).$$

So the second order Taylor expansion for $f_1(x)$ is

$$\begin{aligned}\cos\tfrac{1}{x+\Delta x} = \; & \cos\tfrac{1}{x} + \left(\tfrac{1}{x^2}\sin\tfrac{1}{x}\right)\Delta x \\ & -\tfrac{1}{2(x+t\Delta x)^4}\left[\cos\tfrac{1}{x+t\Delta x} - 2(x+t\Delta x)\sin\tfrac{1}{x+t\Delta x}\right]\Delta x^2,\end{aligned}$$

where $t \in (0,1)$. Similarly, for $f_2(x) = \cos x$, we have

$$f_2^{(1)}(x) = -\sin x, \quad f_2^{(2)}(x) = -\cos x, \quad f_2^{(3)}(x) = \sin x.$$

Thus the third order Taylor expansion for $f_2(x)$ is

$$\cos(x+\Delta x) = \cos x - (\sin x)\Delta x - \frac{1}{2}(\cos x)\Delta x^2 + \frac{1}{6}[\sin(x+t\Delta x)]\Delta x^3,$$

where $t \in (0,1)$. When $x = 1$, we have

$$\cos(1+\Delta x) = \cos 1 - (\sin 1)\Delta x - \frac{1}{2}(\cos 1)\Delta x^2 + \frac{1}{6}[\sin(1+t\Delta x)]\Delta x^3,$$

where $t \in (0,1)$.

Problem 2.5

Using a trig identity we find that

$$f(x_k) = \left(1 + \frac{1}{2^k}\right)^2 (\cos^2 k + \sin^2 k) = \left(1 + \frac{1}{2^k}\right)^2,$$

from which it follows immediately that $f(x_{k+1}) < f(x_k)$.

Let θ be any point in $[0, 2\pi]$. We aim to show that the point $(\cos\theta, \sin\theta)$ on the unit circle is a limit point of $\{x_k\}$.

From the hint, we can identify a subsequence $\xi_{k_1}, \xi_{k_2}, \xi_{k_3}, \ldots$ such that $\lim_{j\to\infty} \xi_{k_j} = \theta$. Consider the subsequence $\{x_{k_j}\}_{j=1}^{\infty}$. We have

$$\begin{aligned}
\lim_{j\to\infty} x_{k_j} &= \lim_{j\to\infty} \left(1 + \frac{1}{2^k}\right) \begin{bmatrix} \cos k_j \\ \sin k_j \end{bmatrix} \\
&= \lim_{j\to\infty} \left(1 + \frac{1}{2^k}\right) \lim_{j\to\infty} \begin{bmatrix} \cos \xi_{k_j} \\ \sin \xi_{k_j} \end{bmatrix} \\
&= \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}.
\end{aligned}$$

Problem 2.6

We need to prove that "isolated local min" $\Rightarrow$ "strict local min." Equivalently, we prove the contrapositive: "not a strict local min" $\Rightarrow$ "not an isolated local min."

If x^* is not even a local min, then it is certainly not an isolated local min. So we suppose that x^* is a local min but that it is not strict. Let $\mathcal{N}$ be any nbd of x^* such that $f(x^*) \le f(x)$ for all $x \in \mathcal{N}$. Because x^* is not a strict local min, there is some other point $x_{\mathcal{N}} \in \mathcal{N}$ such that $f(x^*) = f(x_{\mathcal{N}})$. Hence $x_{\mathcal{N}}$ is also a local min of f in the neighborhood $\mathcal{N}$ that is different from x^*. Since we can do this for *every* neighborhood of x^* within which x^* is a local min, x^* cannot be an isolated local min.

Problem 2.8

Let S be the set of global minimizers of f. If S only has one element, then it is obviously a convex set. Otherwise for all $x, y \in S$ and $\alpha \in [0, 1]$,

$$f(\alpha x + (1-\alpha)y) \le \alpha f(x) + (1-\alpha) f(y)$$

since f is convex. $f(x) = f(y)$ since x, y are both global minimizers. Therefore,

$$f(\alpha x + (1-\alpha)y) \le \alpha f(x) + (1-\alpha) f(x) = f(x).$$

But since $f(x)$ is a global minimizing value, $f(x) \leq f(\alpha x + (1-\alpha)y)$. Therefore, $f(\alpha x + (1-\alpha y) = f(x)$ and hence $\alpha x + (1-\alpha)y \in S$. Thus S is a convex set. $\square$

Problem 2.9

$-\nabla f$ indicates steepest descent. $(p_k)\cdot(-\nabla f) = \|p_k\|\cdot\|\nabla f\|\cos\theta$. p_k is a descent direction if $-90^\circ < \theta < 90^\circ \iff \cos\theta > 0$.

$$\frac{p_k \cdot -\nabla f}{\|p_k\|\|\nabla f\|} = \cos\theta > 0 \qquad \iff \qquad p_k \cdot \nabla f < 0.$$

$$\nabla f = \begin{bmatrix} 2(x_1 + x_2^2) \\ 4x_2(x_1 + x_2^2) \end{bmatrix}$$

$$p_k \cdot \nabla f_k \Big|_{x=\begin{pmatrix}1\\0\end{pmatrix}} = \begin{pmatrix}-1\\1\end{pmatrix} \cdot \begin{pmatrix}2\\0\end{pmatrix} = -2 < 0$$

which implies that p_k is a descent direction.

$$p_k = \begin{pmatrix}-1\\1\end{pmatrix}, \qquad x = \begin{pmatrix}1\\0\end{pmatrix}$$

$$f(x_k + \alpha_k p_k) = f((1-\alpha,\alpha)^T) = ((1-\alpha)+\alpha^2)^2$$

$$\implies \frac{d}{d\alpha} f(x_k + \alpha_k p_k) = 2(1-\alpha+\alpha^2)(-1+2\alpha) = 0 \quad \text{only when } \alpha = \frac{1}{2}.$$

It is seen that $\frac{d^2}{d\alpha^2} f(x_k + \alpha_k p_k)\Big|_{\alpha=\frac{1}{2}} = 6(2\alpha^2 - 2\alpha + 1)\Big|_{\alpha=\frac{1}{2}} = 3 > 0$, so $\alpha = \frac{1}{2}$ is indeed a minimizer.

$\square$

Problem 2.10

Note first that

$$x_j = \sum_{i=1}^{n} S_{ji} z_i + s_j.$$

By the chain rule we have

$$\frac{\partial}{\partial z_i}\tilde{f}(z) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}\frac{\partial x_j}{\partial z_i} = \sum_{j=1}^{n} S_{ji}\frac{\partial f}{\partial x_j} = \left[S^T \nabla f(x)\right]_i .$$

For the second derivatives, we apply the chain rule again:

$$\begin{aligned}\frac{\partial^2}{\partial z_i \partial z_k}\tilde{f}(z) &= \frac{\partial}{\partial z_k}\sum_{j=1}^{n} S_{ji}\frac{\partial f(x)}{\partial x_j} \\ &= \sum_{j=1}^{n}\sum_{l=1}^{n} S_{ji}\frac{\partial^2 f(x)}{\partial x_j \partial x_l}\frac{\partial x_l}{\partial z_k} S_{lk} \\ &= \left[S^T \nabla^2 f(x) S\right]_{ki} .\end{aligned}$$

Problem 2.13

$$x^* = 0$$

$$\frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} = \left|\frac{k}{k+1}\right| < 1 \qquad \text{and} \qquad \frac{k}{k+1} \to 1.$$

$$\text{For any } r \in (0,1), \exists\, k_0 \text{ such that } \ \forall\, k > k_0,\ \frac{k}{k+1} > r.$$

This implies x_k is **not** Q-linearly convergent. □

Problem 2.14

$$\frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^2} = \frac{(0.5)^{2^{k+1}}}{((0.5)^{2^k})^2} = \frac{(0.5)^{2^{k+1}}}{(0.5)^{2^{k+1}}} = 1 < \infty.$$

Hence the sequence is Q-quadratic.

Problem 2.15

$$x_k = \frac{1}{k!} \qquad x^* = \lim_{n\to\infty} x_k = 0$$

$$\frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} = \frac{k!}{(k+1)!} = \frac{1}{k+1} \xrightarrow{k\to\infty} 0.$$

This implies x_k is Q-superlinearly convergent.

$$\frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^2} = \frac{k!k!}{(k+1)!} = \frac{k!}{k+1} \longrightarrow \infty.$$

This implies x_k is **not** Q-quadratic convergent.

□

Problem 2.16

For k even, we have

$$\frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} = \frac{x_k/k}{x_k} = \frac{1}{k} \to 0,$$

while for k odd we have

$$\frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} = \frac{(1/4)^{2^k}}{x_{k-1}/k} = k\frac{(1/4)^{2^k}}{(1/4)^{2^{k-1}}} = k(1/4)^{2^{k-1}} \to 0,$$

Hence we have

$$\frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} = \to 0,$$

so the sequence is Q-superlinear. The sequence is not Q-quadratic because for k even we have

$$\frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^2} = \frac{x_k/k}{x_k^2} = \frac{1}{k}4^{2^k} \to \infty.$$

The sequence is however R-quadratic as it is majorized by the sequence $z_k = (0.5)^{2^k}$, $k = 1, 2, \ldots$. For even k, we obviously have

$$x_k = (0.25)^{2^k} < (0.5)^{2^k} = z_k,$$

while for k odd we have

$$x_k < x_{k-1} = (0.25)^{2^{k-1}} = ((0.25)^{1/2})^{2^k} = (0.5)^{2^k} = z_k.$$

A simple argument shows that z_k is Q-quadratic.

3 Line Search Methods

Problem 3.2

Graphical solution
We show that if c_1 is allowed to be greater than c_2, then we can find a function for which no steplengths $\alpha > 0$ satisfy the Wolfe conditions.

Consider the convex function depicted in Figure 2, and let us choose $c_1 = 0.99$.

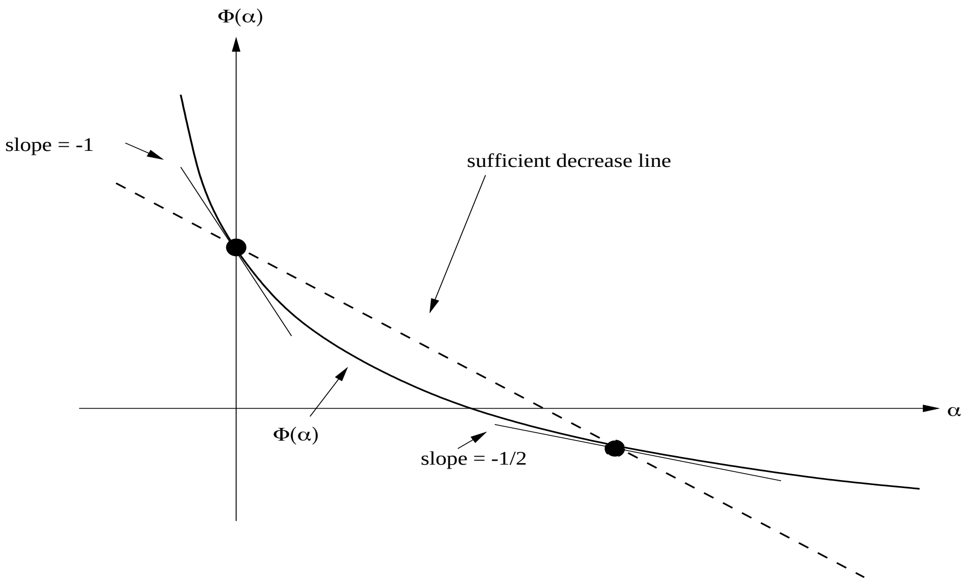

Figure 2: Convex function and sufficient decrease line

We observe that the sufficient decrease line intersects the function only once. Moreover for all points to the left of the intersection, we have

$$\phi'(\alpha) \leq -\frac{1}{2}.$$

Now suppose that we choose $c_2 = 0.1$ so that the curvature condition requires

$$\phi'(\alpha) \geq -0.1. \tag{1}$$

Then there are clearly no steplengths satisfying the inequality (1) for which the sufficient decrease condition holds.

$\square$

Problem 3.3

Suppose p is a descent direction and define

$$\phi(\alpha) = f(x + \alpha p), \qquad \alpha \geq 0.$$

Then any minimizer α^* of $\phi(\alpha)$ satisfies

$$\phi'(\alpha^*) = \nabla f(x + \alpha^* p)^T p = 0. \tag{2}$$

A strongly convex quadratic function has the form

$$f(x) = \frac{1}{2} x^T Q x + b^T x, \qquad Q > 0,$$

and hence

$$\nabla f(x) = Qx + b. \tag{3}$$

The one-dimensional minimizer is unique, and by Equation (2) satisfies

$$[Q(x + \alpha^* p) + b]^T p = 0.$$

Therefore

$$(Qx + b)^T p + \alpha^* p^T Q p = 0$$

which together with Equation (3) gives

$$\alpha^* = -\frac{(Qx + b)^T p}{p^T Q p} = -\frac{\nabla f(x)^T p}{p^T Q p}.$$

□

Problem 3.4

Let $f(x) = \frac{1}{2} x^T Q x + b^T x + d$, with Q positive definite. Let x_k be the current iterate and p_k a non-zero direction. Let $0 < c < \frac{1}{2}$.

The one-dimensional minimizer along $x_k + \alpha p_k$ is (see the previous exercise)

$$\alpha_k = -\frac{\nabla f_k^T p_k}{p_k^T Q p_k}$$

Direct substitution then yields

$$f(x_k) + (1 - c)\alpha_k \nabla f_k^T p_k = f(x_k) - \frac{(\nabla f_k^T p_k)^2}{p_k^T Q p_k} + c \frac{(\nabla f_k^T p_k)^2}{p_k^T Q p_k}$$

Now, since $\nabla f_k = Qx_k + b$, after some algebra we get

$$f(x_k + \alpha_k p_k) = f(x_k) - \frac{(\nabla f_k^T p_k)^2}{p_k^T Q p_k} + \frac{1}{2}\frac{(\nabla f_k^T p_k)^2}{p_k^T Q p_k},$$

from which the first inequality in the Goldstein conditions is evident. For the second inequality, we reduce similar terms in the previous expression to get

$$f(x_k + \alpha_k p_k) = f(x_k) - \frac{1}{2}\frac{(\nabla f_k^T p_k)^2}{p_k^T Q p_k},$$

which is smaller than

$$f(x_k) + c\alpha_k \nabla f_k^T p_k = f(x_k) - c\frac{(\nabla f_k^T p_k)^2}{p_k^T Q p_k}.$$

Hence the Goldstein conditions are satisfied.

Problem 3.5

First we have from (A.7)

$$\|x\| = \|B^{-1}Bx\| \leq \|B^{-1}\| \cdot \|Bx\|,$$

Therefore

$$\|Bx\| \geq \|x\|/\|B^{-1}\|$$

for any nonsingular matrix B.

For symmetric and positive definite matrix B, we have that the matrices $B^{1/2}$ and $B^{-1/2}$ exist and that $\|B^{1/2}\| = \|B\|^{1/2}$ and $\|B^{-1/2}\| = \|B^{-1}\|^{1/2}$. Thus, we have

$$\begin{aligned}
\cos\theta &= -\frac{\nabla f^T p}{\|\nabla f\| \cdot \|p\|} = \frac{p^T B p}{\|Bp\| \cdot \|p\|} \\
&\geq \frac{p^T B p}{\|B\| \cdot \|p\|^2} = \frac{p^T B^{1/2} B^{1/2} p}{\|B\| \cdot \|p\|^2} \\
&= \frac{\|B^{1/2} p\|^2}{\|B\| \cdot \|p\|^2} \geq \frac{\|p\|^2}{\|B^{-1/2}\|^2 \cdot \|B\| \cdot \|p\|^2} \\
&= \frac{1}{\|B^{-1}\| \cdot \|B\|} \geq \frac{1}{M}.
\end{aligned}$$

We can actually prove the stronger result that $\cos\theta \geq 1/M^{1/2}$. Defining $\tilde{p} = B^{1/2}p = -B^{-1/2}\nabla f$, we have

$$\begin{aligned}\cos\theta &= \frac{p^T B p}{\|\nabla f\| \cdot \|p\|} = \frac{\tilde{p}^T \tilde{p}}{\|B^{1/2}\tilde{p}\| \cdot \|B^{-1/2}\tilde{p}\|} \\ &= \frac{\|\tilde{p}\|^2}{\|B^{1/2}\| \cdot \|\tilde{p}\| \cdot \|B^{-1/2}\| \cdot \|\tilde{p}\|} = \frac{1}{\|B^{1/2}\| \cdot \|B^{-1/2}\|} \geq \frac{1}{M^{1/2}}.\end{aligned}$$

Problem 3.6

If $x_0 - x^*$ is parallel to an eigenvector of Q, then

$$\begin{aligned}\nabla f(x_0) &= Qx_0 - b = Qx_0 - Qx^* + Qx^* - b \\ &= Q(x_0 - x^*) + \nabla f(x^*) \\ &= \lambda(x_0 - x^*)\end{aligned}$$

for the corresponding eigenvalue λ. From here, it is easy to get

$$\begin{aligned}\nabla f_0^T \nabla f_0 &= \lambda^2 (x_0 - x^*)^T (x_0 - x^*), \\ \nabla f_0^T Q \nabla f_0 &= \lambda^3 (x_0 - x^*)^T (x_0 - x^*), \\ \nabla f_0^T Q^{-1} \nabla f_0 &= \lambda (x_0 - x^*)^T (x_0 - x^*).\end{aligned}$$

Direct substitution in equation (3.28) yields

$$\|x_1 - x^*\|_Q^2 = 0 \text{ or } x_1 = x^*.$$

Therefore the steepest descent method will find the solution in one step.

Problem 3.7

We drop subscripts on $\nabla f(x_k)$ for simplicity. We have

$$x_{k+1} = x_k - \alpha\nabla f,$$

so that

$$x_{k+1} - x^* = x_k - x^* - \alpha\nabla f,$$

By the definition of $\|\cdot\|_Q^2$, we have

$$\begin{aligned}\|x_{k+1} - x^*\|_Q^2 &= (x_{k+1} - x^*)^T Q (x_{k+1} - x^*) \\ &= (x_k - x^* - \alpha\nabla f)^T Q (x_k - x^* - \alpha\nabla f) \\ &= (x_k - x^*)^T Q (x_k - x^*) - 2\alpha\nabla f^T Q (x_k - x^*) + \alpha^2 \nabla f^T Q \nabla f \\ &= \|x_k - x^*\|_Q^2 - 2\alpha\nabla f^T Q (x_k - x^*) + \alpha^2 \nabla f^T Q \nabla f\end{aligned}$$

Hence, by substituting $\nabla f = Q(x_k - x^*)$ and $\alpha = \nabla f^T \nabla f/(\nabla f^T Q \nabla f)$, we obtain

$$
\begin{aligned}
\|x_{k+1} - x^*\|_Q^2 &= \|x_k - x^*\|_Q^2 - 2\alpha \nabla f^T \nabla f + \alpha^2 \nabla f^T Q \nabla f \\
&= \|x_k - x^*\|_Q^2 - 2(\nabla f^T \nabla f)^2/(\nabla f^T Q \nabla f) + (\nabla f^T \nabla f)^2/(\nabla f^T Q \nabla f) \\
&= \|x_k - x^*\|_Q^2 - (\nabla f^T \nabla f)^2/(\nabla f^T Q \nabla f) \\
&= \|x_k - x^*\|_Q^2 \left[1 - \frac{(\nabla f^T \nabla f)^2}{(\nabla f^T Q \nabla f)\|x_k - x^*\|_Q^2}\right] \\
&= \|x_k - x^*\|_Q^2 \left[1 - \frac{(\nabla f^T \nabla f)^2}{(\nabla f^T Q \nabla f)(\nabla f^T Q^{-1} \nabla f)}\right],
\end{aligned}
$$

where we used

$$\|x_k - x^*\|_Q^2 = \nabla f^T Q^{-1} \nabla f$$

for the final equality.

Problem 3.8

We know that there exists an orthogonal matrix P such that

$$P^T Q P = \Lambda = \text{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}.$$

So

$$P^T Q^{-1} P = (P^T Q P)^{-1} = \Lambda^{-1}.$$

Let $z = P^{-1}x$, then

$$\frac{(x^T x)^2}{(x^T Q x)(x^T Q^{-1} x)} = \frac{(z^T z)^2}{(z^T \Lambda z)(z^T \Lambda^{-1} z)} = \frac{(\sum_i z_i^2)^2}{(\sum_i \lambda_i z_i^2)(\sum_i \lambda_i^{-1} z_i^2)} = \frac{1}{\frac{\sum_i \lambda_i z_i^2}{\sum_i z_i^2} \cdot \frac{\sum_i \lambda_i^{-1} z_i^2}{\sum_i z_i^2}}.$$

Let $u_i = z_i^2 / \sum_i z_i^2$, then all u_i satisfy $0 \le u_i \le 1$ and $\sum_i u_i = 1$. Therefore

$$\frac{(x^T x)^2}{(x^T Q x)(x^T Q^{-1} x)} = \frac{1}{(\sum_i u_i \lambda_i)(\sum_i u_i \lambda_i^{-1})} = \frac{\phi(u)}{\psi(u)}, \tag{4}$$

where $\phi(u) = \frac{1}{\sum_i u_i \lambda_i}$ and $\psi(u) = \sum_i u_i \lambda_i^{-1}$.

Define function $f(\lambda) = \frac{1}{\lambda}$, and let $\bar{\lambda} = \sum_i u_i \lambda_i$. Note that $\bar{\lambda} \in [\lambda_1, \lambda_n]$. Then

$$\phi(u) = \frac{1}{\sum_i u_i \lambda_i} = f(\bar{\lambda}). \tag{5}$$

Let $h(\lambda)$ be the linear function fitting the data $(\lambda_1, \frac{1}{\lambda_1})$ and $(\lambda_n, \frac{1}{\lambda_n})$. We know that

$$h(\lambda) = \frac{1}{\lambda_n} + \frac{\frac{1}{\lambda_1} - \frac{1}{\lambda_n}}{\lambda_n - \lambda_1}(\lambda_n - \lambda).$$

Because f is convex, we know that $f(\lambda) \leq h(\lambda)$ holds for all $\lambda \in [\lambda_1, \lambda_n]$. Thus

$$\psi(\lambda) = \sum_i u_i f(\lambda_i) \leq \sum_i u_i h(\lambda_i) = h(\sum_i u_i \lambda_i) = h(\bar{\lambda}). \tag{6}$$

Combining (4), (5) and (6), we have

$$\begin{aligned}
\frac{(x^T x)^2}{(x^T Q x)(x^T Q^{-1} x)} &= \frac{\phi(u)}{\psi(u)} \geq \frac{f(\bar{\lambda})}{h(\bar{\lambda})} \geq \min_{\lambda_1 \leq \lambda \leq \lambda_n} \frac{f(\lambda)}{h(\lambda)} \quad (\text{since } \bar{\lambda} \in [\lambda_1, \lambda_n]) \\
&= \min_{\lambda_1 \leq \lambda \leq \lambda_n} \frac{\lambda^{-1}}{\frac{1}{\lambda_n} + \frac{\lambda_n - \lambda}{\lambda_1 \lambda_n}} \\
&= \lambda_1 \lambda_n \cdot \min_{\lambda_1 \leq \lambda \leq \lambda_n} \frac{1}{\lambda(\lambda_1 + \lambda_n - \lambda)} \\
&= \lambda_1 \lambda_n \cdot \frac{1}{\frac{\lambda_1 + \lambda_n}{2}(\lambda_1 + \lambda_n - \frac{\lambda_1 + \lambda_n}{2})} \quad (\text{since the minimum happens at } d = \frac{\lambda_1 + \lambda_n}{2}) \\
&= \frac{4\lambda_1 \lambda_n}{(\lambda_1 + \lambda_n)^2}.
\end{aligned}$$

This completes the proof of the Kantorovich inequality.

Problem 3.13

Let $\phi_q(\alpha) = a\alpha^2 + b\alpha + c$. We get a, b and c from the interpolation conditions

$$\begin{aligned}
\phi_q(0) = \phi(0) &\Rightarrow c = \phi(0), \\
\phi_q'(0) = \phi'(0) &\Rightarrow b = \phi'(0), \\
\phi_q(\alpha_0) = \phi(\alpha_0) &\Rightarrow a = (\phi(\alpha_0) - \phi(0) - \phi'(0)\alpha_0)/\alpha_0^2.
\end{aligned}$$

This gives (3.57). The fact that α_0 does not satisfy the sufficient decrease condition implies

$$\begin{aligned}
0 &< \phi(\alpha_0) - \phi(0) - c_1 \phi'(0)\alpha_0 \\
&< \phi(\alpha_0) - \phi(0) - \phi'(0)\alpha_0,
\end{aligned}$$

where the second inequality holds because $c_1 < 1$ and $\phi'(0) < 0$. From here, clearly, $a > 0$. Hence, ϕ_q is convex, with minimizer at

$$\alpha_1 = -\frac{\phi'(0)\alpha_0^2}{2\left[\phi(\alpha_0) - \phi(0) - \phi'(0)\alpha_0\right]}.$$

Now, note that

$$\begin{aligned} 0 &< (c_1 - 1)\phi'(0)\alpha_0 \\ &= \phi(0) + c_1\phi'(0)\alpha_0 - \phi(0) - \phi'(0)\alpha_0 \\ &< \phi(\alpha_0) - \phi(0) - \phi'(0)\alpha_0, \end{aligned}$$

where the last inequality follows from the violation of sufficient decrease at α_0. Using these relations, we get

$$\alpha_1 < -\frac{\phi'(0)\alpha_0^2}{2(c_1 - 1)\phi'(0)\alpha_0} = \frac{\alpha_0}{2(1 - c_1)}.$$

4 Trust-Region Methods

Problem 4.4

Since $\liminf \|g_k\| = 0$, we have by definition of the lim inf that $v_i \to 0$, where the scalar nondecreasing sequence v_i is defined by $v_i = \inf_{k \geq i} \|g_k\|$. In fact, since $\{v_i\}$ is nonnegative and nondecreasing and $v_i \to 0$, we must have $v_i = 0$ for all i, that is,

$$\inf_{k \geq i} \|g_k\| = 0, \quad \text{for all } i.$$

Hence, for any $i = 1, 2, \ldots,$ we can identify an index $j_i \geq i$ such that $\|g_{j_i}\| \leq 1/i$, so that

$$\lim_{i \to \infty} \|g_{j_i}\| = 0.$$

By eliminating repeated entries from $\{j_i\}_{i=1}^{\infty}$, we obtain an (infinite) subsequence $\mathcal{S}$ of such that $\lim_{i \in \mathcal{S}} \|g_i\| = 0$. Moreover, since the iterates $\{x_i\}_{i \in \mathcal{S}}$ are all confined to the bounded set $\mathcal{B}$, we can choose a further subsequence $\bar{\mathcal{S}}$ such that

$$\lim_{i \in \bar{\mathcal{S}}} x_i = x_\infty,$$

for some limit point x_∞. By continuity of g, we have $\|g(x_\infty)\| = 0$, so $g(x_\infty) = 0$, so we are done.

Problem 4.5

Note first that the scalar function of τ that we are trying to minimize is

$$\phi(\tau) \stackrel{\text{def}}{=} m_k(\tau p_k^{\text{S}}) = m_k(-\tau \Delta_k g_k / \|g_k\|) = f_k - \tau \Delta_k \|g_k\| + \frac{1}{2}\tau^2 \Delta_k^2 g_k^T B_k g_k / \|g_k\|^2,$$

while the condition $\|\tau p_k^{\mathrm{S}}\| \leq \Delta_k$ and the definition $p_k^{\mathrm{S}} = -\Delta_k g_k/\|g_k\|$ together imply that the restriction on the scalar τ is that $\tau \in [-1, 1]$.

In the trivial case $g_k = 0$, the function ϕ is a constant, so any value will serve as the minimizer; the value $\tau = 1$ given by (4.12) will suffice.

Otherwise, if $g_k^T B_k g_k = 0$, ϕ is a linear decreasing function of τ, so its minimizer is achieved at the largest allowable value of τ, which is $\tau = 1$, as given in (4.12).

If $g_k^T B_k g_k \neq 0$, ϕ has a parabolic shape with critical point

$$\tau = \frac{\Delta_k \|g_k\|}{\Delta_k^2 g_k^T B_k g_k / \|g_k\|^2} = \frac{\|g_k\|^3}{\Delta_k g_k^T B_k g_k}.$$

If $g_k^T B_k g_k \leq 0$, this value of τ is negative and is a *maximizer*. Hence, the minimizing value of τ on the interval $[-1, 1]$ is at one of the endpoints of the interval. Clearly $\phi(1) < \phi(-1)$, so the solution in this case is $\tau = 1$, as in (4.12).

When $g_k^T B_k g_k \geq 0$, the value of τ above is positive, and is a *minimizer* of ϕ. If this value exceeds 1, then ϕ must be decreasing across the interval $[-1, 1]$, so achieves its minimizer at $\tau = 1$, as in (4.12). Otherwise, (4.12) correctly identifies the formula above as yielding the minimizer of ϕ.

Problem 4.6

Because $\|g\|^2 = g^T g$, it is sufficient to show that

$$(g^T g)(g^T g) \leq (g^T B g)(g^T B^{-1} g). \tag{7}$$

We know from the positive definiteness of B that $g^T B g > 0$, $g^T B^{-1} g > 0$, and there exists nonsingular square matrix L such that $B = LL^T$, and thus $B^{-1} = L^{-T} L^{-1}$. Define $u = L^T g$ and $v = L^{-1} g$, and we have

$$u^T v = (g^T L)(L^{-1} g) = g^T g.$$

The Cauchy-Schwarz inequality gives

$$(g^T g)(g^T g) = (u^T v)^2 \leq (u^T u)(v^T v) = (g^T L L^T g)(g^T L^{-T} L^{-1} g) = (g^T B g)(g^T B^{-1} g). \tag{8}$$

Therefore (7) is proved, indicating

$$\gamma = \frac{\|g\|^4}{(g^T B g)(g^T B^{-1} g)} \leq 1. \tag{9}$$

The equality in (8) holds only when $L^T g$ and $L^{-1}g$ are parallel. That is, when there exists constant $\alpha \neq 0$ such that $L^T g = \alpha L^{-1} g$. This clearly implies that $\alpha g = LL^T g = Bg$, $\frac{1}{\alpha} g = L^{-T}L^{-1}g = B^{-1}g$, and hence the equality in (9) holds only when g, Bg and $B^{-1}g$ are parallel.

Problem 4.8

On one hand, $\phi_2(\lambda) = \frac{1}{\Delta} - \frac{1}{\|p(\lambda)\|}$ and (4.39) gives

$$\begin{aligned} \phi_2'(\lambda) &= -\frac{d}{d\lambda}\frac{1}{\|p(\lambda)\|} = -\frac{d}{d\lambda}(\|p(\lambda)\|^2)^{-1/2} = \frac{1}{2}(\|p(\lambda)\|^2)^{-3/2}\frac{d}{d\lambda}(\|p(\lambda)\|^2) \\ &= \frac{1}{2}\|p(\lambda)\|^{-3}\frac{d}{d\lambda}\sum_{j=1}^{n}\frac{(q_j^T g)^2}{(\lambda_j+\lambda)^2} = -\|p(\lambda)\|^{-3}\sum_{j=1}^{n}\frac{(q_j^T g)^2}{(\lambda_j+\lambda)^3} \end{aligned}$$

where q_j is the j-th column of Q. This further implies

$$\frac{\phi_2(\lambda)}{\phi_2'(\lambda)} = \frac{\|p(\lambda)\|^{-1}\frac{\|p(\lambda)\|-\Delta}{\Delta}}{-\|p(\lambda)\|^{-3}\sum_{j=1}^{n}\frac{(q_j^T g)^2}{(\lambda_j+\lambda)^3}} = -\frac{\|p(\lambda)\|^{2}\frac{\|p(\lambda)\|-\Delta}{\Delta}}{\sum_{j=1}^{n}\frac{(q_j^T g)^2}{(\lambda_j+\lambda)^3}}. \tag{10}$$

On the other hand, we have from Algorithm 4.3 that $q = R^{-T}p$ and $R^{-1}R^{-T} = (B+\lambda I)^{-1}$. Hence (4.38) and the orthonormality of $q_1, q_2, \ldots, q_n$ give

$$\begin{aligned} \|q\|^2 &= p^T(R^{-1}R^{-T})p = p^T(B+\lambda I)^{-1}p = p^T\sum_{j=1}^{n}\frac{q_j^T q_j}{\lambda_j+\lambda}p \\ &= \left(\sum_{j=1}^{n}\frac{q_j^T g}{\lambda_j+\lambda}q_j^T\right)\left(\sum_{j=1}^{n}\frac{q_j^T q_j}{\lambda_j+\lambda}\right)\left(\sum_{j=1}^{n}\frac{q_j^T g}{\lambda_j+\lambda}q_j\right) \\ &= \sum_{j=1}^{n}\frac{(q_j^T g)^2}{(\lambda_j+\lambda)^3}. \end{aligned} \tag{11}$$

Substitute (11) into (10), then we have that

$$\frac{\phi_2(\lambda)}{\phi_2'(\lambda)} = -\frac{\|p\|^2}{\|q\|^2}\cdot\frac{\|p\|-\Delta}{\Delta}. \tag{12}$$

Therefore (4.43) and (12) give (in the l-th iteration of Algorithm 4.3)

$$\lambda^{(l+1)} = \lambda^{(l)} + \frac{\|p_l\|^2}{\|q_l\|^2}\cdot\frac{\|p_l\|-\Delta}{\Delta} = \lambda^{(l)} + \left(\frac{\|p_l\|}{\|q_l\|}\right)^2\left(\frac{\|p_l\|-\Delta}{\Delta}\right).$$

This is exactly (4.44).

Problem 4.10

Since B is symmetric, there exist an orthogonal matrix Q and a diagonal matrix Λ such that $B = Q\Lambda Q^T$, where $\Lambda = \text{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ and $\lambda_1 \leq \lambda_2 \leq \ldots \lambda_n$ are the eigenvalues of B. Now we consider two cases:

(a) If $\lambda_1 > 0$, then all the eigenvalues of B are positive and thus B is positive definite. In this case $B + \lambda I$ is positive definite for $\lambda = 0$.

(b) If $\lambda_1 \leq 0$, we choose $\lambda = -\lambda_1 + \epsilon > 0$ where $\epsilon > 0$ is any fixed real number. Since λ_1 is the most negative eigenvalue of B, we know that $\lambda_i + \lambda \geq \epsilon > 0$ holds for all $i = 1, 2, \ldots, n$. Note that $B + \lambda I = Q(\Lambda + \epsilon I)Q^T$, and therefore $0 < \lambda_1 + \epsilon \leq \lambda_2 + \epsilon \leq \ldots \leq \lambda_n + \epsilon$ are the eigenvalues of $B + \lambda I$. Thus $B + \lambda I$ is positive definite for this choice of λ.

5 Conjugate Gradient Methods

Problem 5.2

Suppose that $p_0, \ldots, p_l$ are conjugate. Let us express one of them, say p_i, as a linear combination of the others:

$$p_i = \sigma_0 p_0 + \cdots + \sigma_l p_l \tag{13}$$

for some coefficients $\sigma_k (k = 0, 1, \ldots, l)$. Note that the sum does not include p_i. Then from conjugacy, we have

$$\begin{aligned} 0 = p_0^T A p_i &= \sigma_0 p_0^T A p_0 + \cdots + \sigma_l p_0^T A p_l \\ &= \sigma_0 p_0^T A p_0. \end{aligned}$$

This implies that $\sigma_0 = 0$ since the vectors $p_0, \ldots, p_l$ are assumed to be conjugate and A is positive definite. The same argument is used to show that all the scaler coefficients $\sigma_k (k = 0, 1, \ldots, l)$ in (13) are zero. Equation (13) indicates that $p_i = 0$, which contradicts the fact that p_i is a nonzero vector. The contradiction then shows that vectors $p_0, \ldots, p_l$ are linearly independent.

□

Problem 5.3

Let

$$\begin{aligned} g(\alpha) &= \phi(x_k + \alpha p_k) \\ &= \frac{1}{2}\alpha^2 p_k^T A p_k + \alpha (A x_k - b)^T p_k + \phi(x_k). \end{aligned}$$

Matrix A is positive definite, so α_k is the minimizer of $g(\alpha)$ if $g'(\alpha_k) = 0$. Hence, we get

$$g'(\alpha_k) = \alpha_k p_k^T A p_k + (Ax_k - b)^T p_k = 0,$$

or

$$\alpha_k = -\frac{(Ax_k - b)^T p_k}{p_k^T A p_k} = -\frac{r_k^T p_k}{p_k^T A p_k}.$$

Problem 5.4

To see that $h(\sigma) = f(x_0 + \sigma_0 p_0 + \cdots + \sigma_{k-1} p_{k-1})$ is a quadratic, note that

$$\sigma_0 p_0 + \cdots + \sigma_{k-1} p_{k-1} = P\sigma$$

where P is the $n \times k$ matrix whose columns are the $n \times 1$ vectors p_i, i.e.

$$P = \begin{bmatrix} | & \cdots & | \\ p_0 & \cdots & p_{k-1} \\ | & \cdots & | \end{bmatrix}$$

and σ is the $k \times 1$ matrix

$$\sigma = \begin{bmatrix} \sigma_0 & \cdots & \sigma_{k-1} \end{bmatrix}^T.$$

Therefore

$$\begin{aligned} h(\sigma) &= \frac{1}{2}(x_0 + P\sigma)^T A (x_0 + P\sigma) + b^T(x_0 + P\sigma) \\ &= \frac{1}{2} x_0^T A x_0 + x_0^T A P \sigma + \frac{1}{2}\sigma^T P^T A P \sigma + b^T x_0 + (b^T P)\sigma \\ &= \frac{1}{2} x_0^T A x_0 + b^T x_0 + [P^T A^T x_0 + P^T b]^T \sigma + \frac{1}{2}\sigma^T (P^T A P)\sigma \\ &= C + \hat{b}^T \sigma + \frac{1}{2}\sigma^T \hat{A} \sigma \end{aligned}$$

where

$$C = \frac{1}{2} x_0^T A x_0 + b^T x_0, \qquad \hat{b} = P^T A^T x_0 + P^T b \qquad \text{and} \qquad \hat{A} = P^T A P.$$

If the vectors $p_0 \cdots p_{k-1}$ are linearly independent, then P has full column rank, which implies that

$$\hat{A} = P^T A P$$

is positive definite. This shows that $h(\sigma)$ is a strictly convex quadratic.

$\square$

Problem 5.5

We want to show

$$\text{span}\,\{r_0, r_1\} = \text{span}\,\{r_0, Ar_0\} = \text{span}\,\{p_0, p_1\}\,. \tag{14}$$

From the CG iteration (5.14) and $p_0 = -r_0$ we know

$$r_1 = Ax_1 - b = A(x_0 + \alpha_0 p_0) - b = (Ax_0 - b) - \alpha_0 Ar_0 = r_0 - \alpha_0 Ar_0. \tag{15}$$

This indicates $r_1 \in \text{span}\,\{r_0, Ar_0\}$ and furthermore

$$\text{span}\,\{r_0, r_1\} \subseteq \text{span}\,\{r_0, Ar_0\}\,. \tag{16}$$

Equation (15) also gives

$$Ar_0 = \frac{1}{\alpha_0}(r_0 - r_1) = \frac{1}{\alpha_0}r_0 - \frac{1}{\alpha_0}r_1.$$

This shows $Ar_0 \in \text{span}\,\{r_0, r_1\}$ and furthermore

$$\text{span}\,\{r_0, r_1\} \supseteq \text{span}\,\{r_0, Ar_0\}\,. \tag{17}$$

We conclude from (16) and (17) that $\text{span}\,\{r_0, r_1\} = \text{span}\,\{r_0, Ar_0\}$.

Similarly, from (5.14) and $p_0 = -r_0$, we have

$$p_1 = -r_1 + \beta_1 p_0 = -\beta_1 r_0 - r_1 \quad \text{or} \quad r_1 = \beta_1 p_0 - p_1.$$

Then $\text{span}\,\{r_0, r_1\} \subseteq \text{span}\,\{p_0, p_1\}$, and $\text{span}\,\{r_0, r_1\} \supseteq \text{span}\,\{p_0, p_1\}$. So $\text{span}\,\{r_0, r_1\} = \text{span}\,\{p_0, p_1\}$. This completes the proof.

Problem 5.6

By the definition of r, we have that

$$\begin{aligned} r_{k+1} &= Ax_{k+1} - b = A(x_k + \alpha_k p_k) - b \\ &= A_k x_k + \alpha_k Ap_k - b = r_k + \alpha_k Ap_k. \end{aligned}$$

Therefore

$$Ap_k = \frac{1}{\alpha_k}(r_{k+1} - r_k). \tag{18}$$

Then we have

$$p_k^T Ap_k = p_k^T(\frac{1}{\alpha_k}(r_{k+1} - r_k)) = \frac{1}{\alpha_k}p_k^T r_{k+1} - \frac{1}{\alpha_k}p_k^T r_k.$$

The expanding subspace minimization property of CG indicates that $p_k^T r_{k+1} = p_{k-1}^T r_k = 0$, and we know $p_k = -r_k + \beta_k p_{k-1}$, so

$$p_k^T A p_k = -\frac{1}{\alpha_k}(-r_k^T + \beta_k p_{k-1}^T) r_k = \frac{1}{\alpha_k} r_k^T r_k - \frac{\beta_k}{\alpha_k} p_{k-1}^T r_k = \frac{1}{\alpha_k} r_k^T r_k. \quad (19)$$

Equation (18) also gives

$$\begin{aligned}
r_{k+1}^T A p_k &= r_{k+1}^T\left(\frac{1}{\alpha_k}(r_{k+1} - r_k)\right) \\
&= \frac{1}{\alpha_k} r_{k+1}^T r_{k+1} - \frac{1}{\alpha_k} r_{k+1}^T r_k \\
&= \frac{1}{\alpha_k} r_{k+1}^T r_{k+1} - \frac{1}{\alpha_k} r_{k+1}^T(-p_k + \beta_k p_{k-1}) \\
&= \frac{1}{\alpha} r_{k+1}^T r_{k+1} + \frac{1}{\alpha} r_{k+1}^T p_k - \frac{\beta_k}{\alpha_k} r_{k+1}^T p_{k-1} \\
&= \frac{1}{\alpha} r_{k+1}^T r_{k+1}.
\end{aligned}$$

This equation, together with (19) and (5.14d), gives that

$$\beta_{k+1} = \frac{r_{k+1}^T A p_k}{p_k^T A p_k} = \frac{\frac{1}{\alpha} r_{k+1}^T r_{k+1}}{\frac{1}{\alpha_k} r_k^T r_k} = \frac{r_{k+1}^T r_{k+1}}{r_k^T r_k}.$$

Thus (5.24d) is equivalent to (5.14d). □

Problem 5.9

Minimize $\hat{\Phi}(\hat{x}) = \frac{1}{2}\hat{x}^T(C^{-T}AC^{-1})\hat{x} - (C^{-T}b)^T\hat{x} \iff$ solve $(C^{-T}AC^{-1})\hat{x} = C^{-T}b$. Apply CG to the transformed problem:

$$\hat{r}_0 = \hat{A}\hat{x}_0 - \hat{b} = (C^{-T}AC^{-1})Cx_0 - C^{-T}b = C^{-T}(Ax_0 - b) = C^{-T}r_0.$$

$$\left\{\begin{array}{c} \hat{p}_0 = -\hat{r}_0 = -C^{-T}r_0 \\ My_0 = r_0 \end{array}\right\} \Longrightarrow \quad \hat{p}_0 = -C^{-T}(My_0) = -C^{-T}C^TCy_0 = -Cy_0.$$

$$\Longrightarrow \quad \hat{\alpha}_0 = \frac{\hat{r}_0^T r_0}{\hat{p}_0^T \hat{A} \hat{p}_0} = \frac{r_0^T C^{-1} C^{-T} r_0}{y_0^T C^T C^{-T} A C^{-1} C y_0} = \frac{r_0^T M^{-1} r_0}{y_0^T A y_0} = \frac{r_0^T y_0}{p_0^T A y_0} = \alpha_0.$$

$$\hat{x}_1 = \hat{x}_0 + \hat{\alpha}_0 p_0 \Rightarrow C x_1 = C x_0 + \frac{r_0^T y_0}{p_0^T A y_0}(-C y_0)$$

$$\Longrightarrow \quad x_1 = x_0 - \frac{r_0^T y_0}{p_0^T A y_0} y_0 = x_0 + \alpha_0 p_0$$

$$\hat{r}_1 = \hat{r}_0 + \hat{\alpha}_0 \hat{A} \hat{p}_0 \Rightarrow C^{-T} r_1 = C^{-T} r_0 + \frac{r_0^T y_0}{p_0^T A y_0} C^{-T} A C^{-1}(-C y_0)$$

$$\Longrightarrow \quad r_1 = r_0 + \frac{r_0^T y_0}{p_0^T A y_0} A(-y_0) = r_0 + \alpha_0 A p_0$$

$$\hat{\beta}_1 = \frac{\hat{r}_1^T \hat{r}_1}{\hat{r}_0^T \hat{r}_0} = \frac{r_1^T C^{-1} C^{-T} r_1}{r_0^T C^{-1} C^{-T} r_0} = \frac{r_1^T M^{-1} r_1}{r_0^T M^{-1} r_0} = \frac{r_1^T y_1}{r_0^T y_0} = \beta_1$$

$$\hat{p}_1 = -\hat{r}_1 + \hat{\beta}_1 \hat{p}_0 \Rightarrow -C y_1 = -C^{-T} r_1 + \beta_1(-C y_0)$$

$$\Longrightarrow \quad y_1 = M^{-1} r_1 + \beta_1 y_0 \Rightarrow p_1 = -y_1 + \beta_1 p_0 \qquad (\text{ because } \hat{p}_1 = C p_1).$$

By comparing the formulas above with Algorithm 5.3, we can see that by applying CG to the problem with the new variables, then transforming back into original variables, the derived algorithm is the same as Algorithm 5.3 for $k = 0$. Clearly, the same argument can be used for any k; the key is to notice the relationships:

$$\left\{ \begin{array}{l} \hat{x}_k = C x_k \\ \hat{p}_k = C p_k \\ \hat{r}_k = C^{-T} r_k \end{array} \right\}.$$

□

Problem 5.10

From the solution of Problem 5.9 it is seen that $\hat{r}_i = C^{-T} r_i$ and $\hat{r}_j = C^{-T} r_j$. Since the unpreconditioned CG algorithm is applied to the transformed problem, by the orthogonality of the residuals we know that $\hat{r}_i^T \hat{r}_j = 0$ for all $i \neq j$. Therefore

$$0 = \hat{r}_i^T \hat{r}_j = r_i^T C^{-1} \cdot C^{-T} r_j = r_i^T M^{-1} r_j.$$

Here the last equality holds because $M^{-1} = (C^T C)^{-1} = C^{-1} C^{-T}$.

6 Quasi-Newton Methods

Problem 6.1

(a) A function $f(x)$ is strongly convex if all eigenvalues of $\nabla^2 f(x)$ are positive and bounded away from zero. This implies that there exists $\sigma > 0$ such that

$$p^T \nabla^2 f(x) p \geq \sigma \|p\|^2 \qquad \text{for any } p. \tag{20}$$

By Taylor's theorem, if $x_{k+1} = x_k + \alpha_k p_k$, then

$$\nabla f(x_{k+1}) = \nabla f(x_k) + \int_0^1 [\nabla^2 f(x_k + z\alpha_k p_k) \alpha_k p_k] dz.$$

By (20) we have

$$\begin{aligned} \alpha_k p_k^T y_k &= \alpha_k p_k^T [\nabla f(x_{k+1} - \nabla f(x_k)] \\ &= \alpha_k^2 \int_0^1 \left[p_k^T \nabla^2 f(x_k + z\alpha_k p_k) p_k \right] dz \\ &\geq \sigma \|p_k\|^2 \alpha_k^2 > 0. \end{aligned}$$

The result follows by noting that $s_k = \alpha_k p_k$.

□

(b) For example, when $f(x) = \dfrac{1}{x+1}$, we have $g(x) = -\dfrac{1}{(x+1)^2}$. Obviously

$$f(0) = 1, f(1) = \frac{1}{2}, g(0) = -1, g(1) = -\frac{1}{4}.$$

So

$$s^T y = (f(1) - f(0)) (g(1) - g(0)) = -\frac{3}{8} < 0$$

and (6.7) does not hold in this case.

Problem 6.2

The second strong Wolfe condition is

$$\left|\nabla f(x_k + \alpha_k p_k)^T p_k\right| \leq c_2 \left|\nabla f(x_k)^T p_k\right|$$

which implies

$$\begin{aligned}\nabla f(x_k + \alpha_k p_k)^T p_k &\geq -c_2 \left|\nabla f(x_k)^T p_k\right| \\ &= c_2 \nabla f(x_k)^T p_k\end{aligned}$$

since p_k is a descent direction. Thus

$$\begin{aligned}\nabla f(x_k + \alpha_k p_k)^T p_k - \nabla f(x_k)^T p_k &= (c_2 - 1)\nabla f(x_k)^T p_k \\ &> 0\end{aligned}$$

since we have assumed that $c_2 < 1$. The result follows by multiplying both sides by α_k and noting $s_k = \alpha_k p_k, y_k = \nabla f(x_k + \alpha_k p_k) - \nabla f(x_k)$.

□

7 Large-Scale Unconstrained Optimization

Problem 7.2

Since $s_k \neq 0$, the product

$$\begin{aligned}\hat{H}_{k+1} s_k &= \left(I - \frac{s_k y_k^T}{y_k^T s_k}\right) s_k \\ &= s_k - \frac{y_k^T s_k}{y_k^T s_k} s_k \\ &= 0\end{aligned}$$

illustrates that $\hat{H}_{k+1}$ is singular.

Problem 7.3

We assume line searches are exact, so $\nabla f_{k+1}^T p_k = 0$. Also, recall $s_k = \alpha_k p_k$. Therefore,

$$
\begin{aligned}
p_{k+1} &= -H_{k+1}\nabla f_{k+1} \\
&= -\left(\left(I - \frac{s_k y_k^T}{y_k^T s_k}\right)\left(I - \frac{y_k s_k^T}{y_k^T s_k}\right) + \frac{s_k s_k^T}{y_k^T s_k}\right)\nabla f_{k+1} \\
&= -\left(\left(I - \frac{p_k y_k^T}{y_k^T p_k}\right)\left(I - \frac{y_k p_k^T}{y_k^T p_k}\right) + \alpha_k \frac{p_k p_k^T}{y_k^T p_k}\right)\nabla f_{k+1} \\
&= -\left(I - \frac{p_k y_k^T}{y_k^T p_k}\right)\nabla f_{k+1} \\
&= -\nabla f_{k+1} + \frac{\nabla f_{k+1}^T y_k}{y_k^T p_k} p_k,
\end{aligned}
$$

as given.

Problem 7.5

For simplicity, we consider $(x_3 - x_4)$ as an element function despite the fact that it is easily separable. The function can be written as

$$
f(x) = \sum_{i=1}^{3} \phi_i(U_i x)
$$

where

$$
\begin{aligned}
\phi_i(u_1, u_2, u_3, u_4) &= u_2 u_3 e^{u_1 + u_3 - u_4}, \\
\phi(v_1, v_2) &= (v_1 v_2)^2, \\
\phi(w_1, w_2) &= w_1 - w_2,
\end{aligned}
$$

and

$$
\begin{aligned}
U_1 &= I, \\
U_2 &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \\
U_3 &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.
\end{aligned}
$$

Problem 7.6

We find

$$\begin{aligned} Bs &= \left(\sum_{i=1}^{ne} U_i^T B_{[i]} U_i \right) s \\ &= \sum_{i=1}^{ne} U_i^T B_{[i]} s_{[i]} \\ &= \sum_{i=1}^{ne} U_i^T y_{[i]} \\ &= y, \end{aligned}$$

so the secant equation is indeed satisfied.

8 Calculating Derivatives

Problem 8.1

Supposing that L_c is the constant in the central difference formula, that is,

$$\left| \frac{\partial f}{\partial x_i} - \left[\frac{f(x+\epsilon e_i) - f(x-\epsilon e_i)}{2\epsilon} \right] \right| \le L_c \epsilon^2,$$

and assuming as in the analysis of the forward difference formula that

$$\begin{aligned} |\text{comp}(f(x+\epsilon e_i)) - f(x+\epsilon e_i))| &\le L_f \mathbf{u}, \\ |\text{comp}(f(x-\epsilon e_i)) - f(x-\epsilon e_i))| &\le L_f \mathbf{u}, \end{aligned}$$

the total error in the central difference formula is bounded by

$$L_c \epsilon^2 + \frac{2\mathbf{u} L_f}{2\epsilon}.$$

By differentiating with respect to ϵ, we find that the minimizer is at

$$\epsilon = \left(\frac{L_f \mathbf{u}}{2L_c} \right)^{1/3},$$

so when the ratio L_f/L_c is reasonable, the choice $\epsilon = \mathbf{u}^{1/3}$ is a good one. By substituting this value into the error expression above, we find that both terms are multiples of $\mathbf{u}^{2/3}$, as claimed.

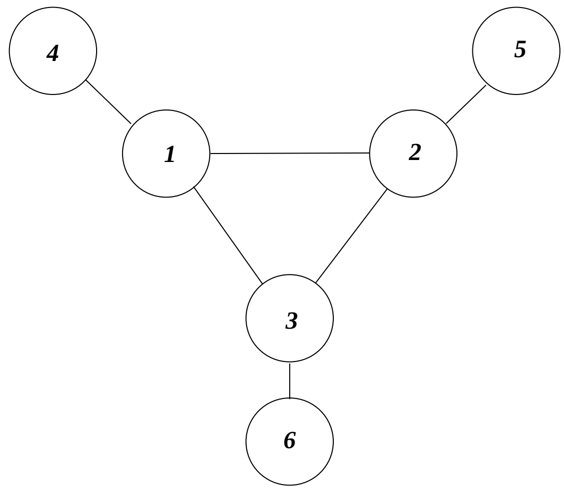

Figure 3: Adjacency Graph for Problem 8.6

Problem 8.6

See the adjacency graph in Figure 3.

Four colors are required; the nodes corresponding to these colors are $\{1\}$, $\{2\}$, $\{3\}$, $\{4, 5, 6\}$.

Problem 8.7

We start with

$$\nabla x_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \qquad \nabla x_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \qquad \nabla x_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

By applying the chain rule, we obtain

$$\nabla x_4 = x_1 \nabla x_2 + x_2 \nabla x_1 = \begin{bmatrix} x_2 \\ x_1 \\ 0 \end{bmatrix},$$

$$\nabla x_5 = (\cos x_3) \nabla x_3 = \begin{bmatrix} 0 \\ 0 \\ \cos x_3 \end{bmatrix},$$

$$\nabla x_6 = e^{x_4} \nabla x_4 = e^{x_1 x_2} \begin{bmatrix} x_2 \\ x_1 \\ 0 \end{bmatrix},$$

$$\nabla x_7 = x_4 \nabla x_5 + x_5 \nabla x_4 = \begin{bmatrix} x_2 \sin x_3 \\ x_1 \sin x_3 \\ x_1 x_2 \cos x_3 \end{bmatrix},$$

$$\nabla x_8 = \nabla x_6 + \nabla x_7 = e^{x_1 x_2} \begin{bmatrix} x_2 \\ x_1 \\ 0 \end{bmatrix} + \begin{bmatrix} x_2 \sin x_3 \\ x_1 \sin x_3 \\ x_1 x_2 \cos x_3 \end{bmatrix},$$

$$\nabla x_9 = \frac{1}{x_3} \nabla x_8 - \frac{x_8}{x_3^2} \nabla x_3.$$

9 Derivative-Free Optimization

Problem 9.3

The interpolation conditions take the form

$$(\hat{s}^l)^T \hat{g} = f(y^l) - f(x_k) \qquad l = 1, \ldots, q-1, \tag{21}$$

where

$$\hat{s}^l \equiv \left((s^l)^T, \{s_i^l s_j^l\}_{i<j}, \left\{ \tfrac{1}{\sqrt{2}} (s_i^l)^2 \right\} \right)^T \qquad l = 1, \ldots, m-1,$$

and s^l is defined by (9.13). The model (9.14) is uniquely determined if and only if the system (21) has a unique solution, or equivalently, if and only if the set $\{\hat{s}^l : l = 1, \ldots, q-1\}$ is linearly independent.

Problem 9.10

It suffices to show that for any v, we have $\max_{j=1,2,\ldots,n+1} v^T d_j \geq (1/4n)\|v\|_1$. Consider first the case of $v \geq 0$, that is, all components of v are nonnegative.

We then have

$$\max_{j=1,2,\ldots,n+1} v^T d_j \geq v^T d_{n+1} \geq \frac{1}{2n} e^T v = \frac{1}{2n} \|v\|_1.$$

Otherwise, let i be the index of the most negative component of v. We have that

$$\|v\|_1 = -\sum_{v_j<0} v_j + \sum_{v_j\geq 0} v_j \leq n|v_i| + \sum_{v_j \geq 0} v_j.$$

We consider two cases. In the first case, suppose that

$$|v_i| \geq \frac{1}{2n} \sum_{v_j \geq 0} v_j.$$

In this case, we have from the inequality above that

$$\|v\|_1 \leq n|v_i| + (2n)|v_i| = (3n)|v_i|,$$

so that

$$\begin{aligned}
\max_{d\in\mathcal{D}_k} d^T v &\geq d_i^T v \\
&= (1-1/2n)|v_i| + (1/2n)\sum_{j\neq i} v_j \\
&\geq (1-1/2n)|v_i| - (1/2n) \sum_{j\neq i, v_j<0} v_j \\
&\geq (1-1/2n)|v_i| - (1/2n)n|v_i| \\
&\geq (1/2 - 1/2n)|v_i| \\
&\geq (1/4)|v_i| \\
&\geq (1/12n)\|v\|_1,
\end{aligned}$$

which is sufficient to prove the desired result. We now consider the second case, for which

$$|v_i| < \frac{1}{2n} \sum_{v_j \geq 0} v_j.$$

We have here that

$$\|v\|_1 \leq n \frac{1}{2n} \sum_{v_j\geq 0} v_j + \sum_{v_j \geq 0} v_j \leq \frac{3}{2} \sum_{v_j \geq 0} v_j,$$

so that

$$\begin{aligned}
\max_{d\in\mathcal{D}_k} d^T v &\geq d_{n+1}^T v \\
&= \frac{1}{2n}\sum_{v_j\leq 0} v_j + \frac{1}{2n}\sum_{v_j\geq 0} v_j \\
&\geq -\frac{1}{2n} n|v_i| + \frac{1}{2n}\sum_{v_j\geq 0} v_j \\
&= -\frac{1}{2}|v_i| + \frac{1}{2n}\sum_{v_j\geq 0} v_j \\
&\geq -\frac{1}{4n}\sum_{v_j\geq 0} v_j + \frac{1}{2n}\sum_{v_j\geq 0} v_j \\
&= \frac{1}{4n}\sum_{v_j\geq 0} v_j \\
&\geq \frac{1}{6n}\|v\|_1.
\end{aligned}$$

which again suffices.

10 Least-Squares Problems

Problem 10.1

Recall:
(i) "J has full column rank" is equivalent to "$Jx = 0 \Rightarrow x = 0$";
(ii) "$J^T J$ is nonsingular" is equivalent to "$J^T Jx = 0 \Rightarrow x = 0$";
(iii) "$J^T J$ is positive definite" is equivalent to "$x^T J^T Jx \geq 0 (\forall x)$" and "$x^T J^T Jx = 0 \Rightarrow x = 0$".

(a) We want to show (i) $\Leftrightarrow$ (ii).

- (i) $\Rightarrow$ (ii). $J^T Jx = 0 \Rightarrow x^T J^T Jx = 0 \Rightarrow \|Jx\|_2^2 = 0 \Rightarrow Jx = 0 \Rightarrow$ (by (i)) $x = 0$.
- (ii) $\Rightarrow$ (i). $Jx = 0 \Rightarrow J^T Jx = 0 \Rightarrow$ (by (ii)) $x = 0$.

(b) We want to show (i) $\Leftrightarrow$ (iii).

- (i) $\Rightarrow$ (iii). $x^T J^T Jx = \|Jx\|_2^2 \geq 0 (\forall x)$ is obvious. $x^T J^T Jx = 0 \Rightarrow \|Jx\|_2^2 = 0 \Rightarrow Jx = 0 \Rightarrow$ (by (i)) $x = 0$.

- (iii) $\Rightarrow$ (i). $Jx = 0 \Rightarrow x^T J^T J x = \|Jx\|_2^2 = 0 \Rightarrow$ (by (iii)) $x = 0$.

Problem 10.3

(a) Let Q be a $n \times n$ orthogonal matrix and x be any given n-vector. Define $q_i (i = 1, 2, \cdots, n)$ to be the i-th column of Q. We know that

$$q_i^T q_j = \begin{cases} \|q_i\|^2 = 1 & (\text{if } i = j) \\ 0 & (\text{if } i \neq j). \end{cases} \tag{22}$$

Then

$$\begin{aligned} \|Qx\|^2 &= (Qx)^T (Qx) \\ &= (x_1 q_1 + x_2 q_2 + \cdots + x_n q_n)^T (x_1 q_1 + x_2 q_2 + \cdots + x_n q_n) \\ &= \sum_{i=1}^n \sum_{j=1}^n x_i x_j q_i^T q_j \qquad (\text{by } (22)) \\ &= \sum_{i=1}^n x_i^2 = \|x\|^2. \end{aligned}$$

(b) If $\Pi = I$, then $J^T J = (Q_1 R)^T (Q_1 R) = R^T R$. We know that the Cholesky decomposition is unique if the diagonal elements of the upper triangular matrix are positive, so $\bar{R} = R$.

Problem 10.4

(a) It is easy to see from (10.19) that

$$J = \sum_{i=1}^n \sigma_i u_i v_i^T = \sum_{i:\sigma_i \neq 0} \sigma_i u_i v_i^T.$$

Since the objective function $f(x)$ defined by (10.13) is convex, it suffices to show that $\nabla f(x^*) = 0$, where x^* is given by (10.22). Recall $v_i^T v_j = 1$ if

$i = j$ and 0 otherwise, $u_i^T u_j = 1$ if $i = j$ and 0 otherwise. Then

$$\begin{aligned}
\nabla f(x^*) &= J^T(Jx^* - y) \\
&= J^T\left[\sum_{i:\sigma_i \neq 0} \sigma_i u_i v_i^T \left(\sum_{i:\sigma_i \neq 0} \frac{u_i^T y}{\sigma_i} v_i + \sum_{i:\sigma_i = 0} \tau_i v_i\right) - y\right] \\
&= J^T\left[\sum_{i:\sigma_i \neq 0} \frac{\sigma_i}{\sigma_i}(u_i^T y) u_i (v_i^T v_i) - y\right] \\
&= \left(\sum_{i:\sigma_i \neq 0} \sigma_i v_i u_i^T\right)\left[\sum_{i:\sigma_i \neq 0} (u_i^T y) u_i - y\right] \\
&= \sum_{i:\sigma_i \neq 0} \sigma_i (u_i^T y) v_i (u_i^T u_i) - \sum_{i:\sigma_i \neq 0} \sigma_i v_i (u_i^T y) \\
&= \sum_{i:\sigma_i \neq 0} \sigma_i (u_i^T y) v_i - \sum_{i:\sigma_i \neq 0} \sigma_i v_i (u_i^T y) = 0.
\end{aligned}$$

(b) If J is rank-deficient, we have

$$x^* = \sum_{i:\sigma_i \neq 0} \frac{u_i^T y}{\sigma_i} v_i + \sum_{i:\sigma_i = 0} \tau_i v_i.$$

Then

$$\|x^*\|_2^2 = \sum_{i:\sigma_i \neq 0} \left(\frac{u_i^T y}{\sigma_i}\right)^2 + \sum_{i:\sigma_i = 0} \tau_i^2,$$

which is minimized when $\tau_i = 0$ for all i with $\sigma_i = 0$.

□

Problem 10.5

For the Jacobian, we get the same Lipschitz constant:

$$\begin{aligned}
& \|J(x_1) - J(x_2)\| \\
= \; & \max_{\|u\|=1} \|(J(x_1) - J(x_2))u\| \\
= \; & \max_{\|u\|=1} \left\| \begin{pmatrix} (\nabla r_1(x_1) - \nabla r_1(x_2))^T u \\ \vdots \\ (\nabla r_m(x_1) - \nabla r_m(x_2))^T u \end{pmatrix} \right\| \\
\leq \; & \max_{\|u\|=1} \max_{j=1,\dots,m} |(\nabla r_j(x_1) - \nabla r_j(x_2))^T u| \\
\leq \; & \max_{\|u\|=1} \max_{j=1,\dots,m} \|\nabla r_j(x_1) - \nabla r_j(x_2)\| \|u\| |\cos(\nabla r_j(x_1) - \nabla r_j(x_2), u)| \\
\leq \; & L\|x_1 - x_2\|.
\end{aligned}$$

For the gradient, we get $\tilde{L} = L(L_1 + L_2)$, with $L_1 = \max_{x \in \mathcal{D}} \|r(x)\|_1$ and $L_2 = \max_{x \in \mathcal{D}} \sum_{j=1}^m \|\nabla r_j(x)\|$:

$$\begin{aligned}
& \|\nabla f(x_1) - \nabla f(x_2)\| \\
= \; & \left\| \sum_{j=1}^m \nabla r_j(x_1) r_j(x_1) - \sum_{j=1}^m \nabla r_j(x_2) r_j(x_2) \right\| \\
= \; & \left\| \sum_{j=1}^m (\nabla r_j(x_1) - \nabla r_j(x_2)) r_j(x_1) + \sum_{j=1}^m \nabla r_j(x_2)(r_j(x_1) - r_j(x_2)) \right\| \\
\leq \; & \sum_{j=1}^m \|\nabla r_j(x_1) - \nabla r_j(x_2)\| \, |r_j(x_1)| + \sum_{j=1}^m \|\nabla r_j(x_2)\| \, |r_j(x_1) - r_j(x_2)| \\
\leq \; & L\|x_1 - x_2\| \sum_{j=1}^m |r_j(x_1)| + L\|x_1 - x_2\| \sum_{j=1}^m \|\nabla r_j(x_2)\| \\
\leq \; & \tilde{L}\|x_1 - x_2\|.
\end{aligned}$$

Problem 10.6

If $J = U_1 S V^T$, then $(J^T J + \lambda I) = V(S^2 + \lambda I)V^T$. From here,

$$\begin{aligned} p^{LM} &= -V(S^2+\lambda I)^{-1} S U_1^T r \\ &= -\sum_{i=1}^{n} \frac{\sigma_i}{\sigma_i^2+\lambda}(u_i^T r) v_i \\ &= -\sum_{i:\sigma_i\neq 0} \frac{\sigma_i}{\sigma_i^2+\lambda}(u_i^T r) v_i. \end{aligned}$$

Thus,

$$\|p^{LM}\|^2 = \sum_{i:\sigma_i\neq 0} \left(\frac{\sigma_i}{\sigma_i^2+\lambda}(u_i^T r) v_i \right)^2,$$

and

$$\lim_{\lambda\to 0} p^{LM} = -\sum_{i:\sigma_i\neq 0} \frac{u_i^T r}{\sigma_i} v_i.$$

□

11 Nonlinear Equations

Problem 11.1

Note $s^T s = \|s\|_2^2$ is a scalar, so it sufficies to show that $\|ss^T\| = \|s\|_2^2$. By definition,

$$\|ss^T\| = \max_{\|x\|_2=1} \|(ss^T)x\|_2.$$

Matrix multiplication is associative, so $(ss^T)x = s(s^T x)$, and $s^T x$ is a scalar. Hence,

$$\max_{\|x\|_2=1} \|s(s^T x)\|_2 = \max_{\|x\|_2=1} |s^T x| \|s\|_2.$$

Last,

$$|s^T x| = |\|s\|_2 \|x\|_2 \cos\theta_{s,x}| = \|s\|_2 |\cos\theta_{s,x}|,$$

which is maximized when $|\cos\theta_{s,x}| = 1$. Therefore,

$$\max_{\|x\|_2=1} |s^T x| = \|s\|_2,$$

which yields the result.

Problem 11.2

Starting at $x_0 \neq 0$, we have $r'(x_0) = q x_0^{q-1}$. Hence,

$$x_1 = x_0 - \frac{x_0^q}{q x_0^{q-1}} = \left(1 - \frac{1}{q}\right) x_0.$$

A straghtforward induction yields

$$x_k = \left(1 - \frac{1}{q}\right)^k x_0,$$

which certainly converges to 0 as $k \to \infty$. Moreover,

$$\frac{x_{k+1}}{x_k} = 1 - \frac{1}{q},$$

so the sequence converges Q-linearly to 0, with convergence ratio $1 - 1/q$.

Problem 11.3

For this function, Newton's method has the form:

$$x_{k+1} = x_k - \frac{r(x)}{r'(x)} = x_k - \frac{-x^5 + x^3 + 4x}{-5x^4 + 3x^2 + 4}.$$

Starting at $x_0 = 1$, we find

$$\begin{array}{rcl} x_1 & = & x_0 - \dfrac{-x_0^5 + x_0^3 + 4x_0}{-5x_0^4 + 3x_0^2 + 4} = 1 - \dfrac{4}{2} = -1, \\ x_2 & = & x_1 - \dfrac{-x_1^5 + x_1^3 + 4x_1}{-5x_1^4 + 3x_1^2 + 4} = -1 + \dfrac{4}{2} = 1, \\ x_3 & = & -1, \\ \vdots & \vdots & \vdots \end{array}$$

as described.

A trivial root of $r(x)$ is $x = 0$, i.e.,

$$r(x) = (-x - 0)(x^4 - x^2 - 4).$$

The remaining roots can be found by noticing that $f(x) = x^4 - x^2 - 4$ is quadratic in $y = x^2$. According to the quadratic equation, we have the roots

$$y = \frac{1 \pm \sqrt{17}}{2} = x^2 \quad \Rightarrow \quad x = \pm\sqrt{\frac{1 \pm \sqrt{17}}{2}}.$$

As a result,

$$r(x) = (-x)\left(x-\sqrt{\frac{1-\sqrt{17}}{2}}\right)\left(x-\sqrt{\frac{1+\sqrt{17}}{2}}\right)\left(x+\sqrt{\frac{1-\sqrt{17}}{2}}\right)\left(x+\sqrt{\frac{1+\sqrt{17}}{2}}\right).$$

Problem 11.4

The sum-of-squares merit function is in this case

$$f(x) = \frac{1}{2}\left(\sin(5x) - x\right)^2.$$

Moreover, we find

$$\begin{aligned} f'(x) &= (\sin(5x) - x)\,(5\cos(5x) - 1)\,, \\ f''(x) &= -25\sin(5x)\,(\sin(5x) - x) + (5\cos(5x) - 1)^2\,. \end{aligned}$$

The merit function has local minima at the roots of r, which as previously mentioned are found at approximately $x \in S = \{-0.519148, 0, 0.519148\}$. Furthermore, there may be local minima at points where the Jacobian is singular, i.e., x such that $J(x) = 5\cos(5x) - 1 = 0$. All together, there are an infinite number of local minima described by

$$x^* \in S \cup \left\{x \mid 5\cos(5x) = 1 \ \wedge \ f''(x) \geq 0\right\}.$$

Problem 11.5

First, if $J^T r = 0$, then $\phi(\lambda) = 0$ for all λ.

Suppose $J^T r \neq 0$. Let the singular value decomposition of $J \in \Re^{m\times n}$ be

$$J = USV$$

where $U \in \Re^{m\times n}$ and $V \in \Re^{n\times n}$ are orthogonal. We find (let $z = S^T U^T r$):

$$\begin{aligned} \phi(\lambda) = &\ \|(J^TJ + \lambda I)^{-1}J^Tr\| \\ = &\ \|(V^TS^TU^TUSV + \lambda V^TV)^{-1}V^Tz\| \\ = &\ \|(V^T(S^TS + \lambda I)V)^{-1}V^Tz\| \\ = &\ \|V^T(S^TS + \lambda I)^{-1}VV^Tz\| && (\text{since} V^{-1} = V^T) \\ = &\ \|V^T(S^TS + \lambda I)^{-1}z\| \\ = &\ \|(S^TS + \lambda I)^{-1}z\| && (\text{since} V^T \text{ is orthogonal}) \\ = &\ \|(D(\lambda))^{-1}z\| \end{aligned}$$

where $D(\lambda)$ is a diagonal matrix having

$$[D(\lambda)]_{ii} = \begin{cases} \sigma_i^2 + \lambda, & i = 1, \ldots, \min(m, n) \\ \lambda, & i = \min(m, n) + 1, \ldots, \max(m, n). \end{cases}$$

Each entry of $y(\lambda) = (D(\lambda))^{-1} z$ is of the form

$$y_i(\lambda) = \frac{z_i}{[D(\lambda)]_{ii}}.$$

Therefore, $|y_i(\lambda_1)| < |y_i(\lambda_2)|$ for $\lambda_1 > \lambda_2 > 0$ and $i = 1, \ldots, n$, which implies $\phi(\lambda_1) < \phi(\lambda_2)$ for $\lambda_1 > \lambda_2 > 0$.

Problem 11.8

Notice that

$$JJ^T r = 0 \Rightarrow r^T JJ^T r = 0.$$

If $v = J^T r$, then the above implies

$$r^T JJ^T r = v^T v = \|v\|^2 = 0$$

which must mean $v = J^T r = 0$.

Problem 11.10

The homotopy map expands to

$$\begin{aligned} H(x, \lambda) &= \lambda\left(x^2 - 1\right) + (1 - \lambda)(x - a) \\ &= \lambda x^2 + (1 - \lambda)x - \frac{1}{2}(1 + \lambda). \end{aligned}$$

For a given λ, the quadratic formula yields the following roots for the above:

$$\begin{aligned} x &= \frac{\lambda - 1 \pm \sqrt{(1 - \lambda)^2 + 2\lambda(1 + \lambda)}}{2\lambda} \\ &= \frac{\lambda - 1 \pm \sqrt{1 + 3\lambda^2}}{2\lambda}. \end{aligned}$$

By choosing the positive root, we find that the zero path defined by

$$\begin{cases} \lambda = 0 & \Rightarrow \ x = 1/2, \\ \lambda \in (0, 1] & \Rightarrow \ x = \frac{\lambda - 1 + \sqrt{1 + 3\lambda^2}}{2\lambda}, \end{cases}$$

connects $(\frac{1}{2}, 0)$ to $(1, 1)$, so continuation methods should work for this choice of starting point.

12 Theory of Constrained Optimization

Problem 12.4

First, we show that local solutions to problem 12.3 are also global solutions. Take any local solution to problem 12.3, denoted by x_0. This means that there exists a neighborhood $N(x_0)$ such that $f(x_0) \leq f(x)$ holds for any $x \in N(x_0) \cap \Omega$. The following proof is based on contradiction.

Suppose x_0 is not a global solution, then we take a global solution $\tilde{x} \in \Omega$, which satisfies $f(x_0) > f(\tilde{x})$. Because Ω is a convex set, there exists $\alpha \in [0, 1]$ such that $\alpha x_0 + (1-\alpha)\tilde{x} \in N(x_0) \cap \Omega$. Then the convexity of $f(x)$ gives

$$\begin{aligned} f(\alpha x_0 + (1-\alpha)\tilde{x}) &\leq \alpha f(x_0) + (1-\alpha) f(\tilde{x}) \\ &< \alpha f(x_0) + (1-\alpha) f(x_0) \\ &= f(x_0), \end{aligned}$$

which contradicts the fact that x_0 is the minimum point in $N(x_0) \cap \Omega$. It follows that x_0 must be a global solution, and that any local solution to problem 12.3 must also be a global solution.

Now, let us prove that the set of global solutions is convex. Let

$$S = \{x \mid x \text{ is a global solution to problem 12.3}\},$$

and consider any $x_1, x_2 \in S$ such that $x_1 \neq x_2$ and $x = \alpha x_1 + (1-\alpha)x_2$, $\alpha \in (0, 1)$. By the convexity of $f(x)$, we have

$$\begin{aligned} f(\alpha x_1 + (1-\alpha)x_2) &\leq \alpha f(x_1) + (1-\alpha) f(x_2) \\ &= \alpha f(x_1) + (1-\alpha) f(x_1) \\ &= f(x_1). \end{aligned}$$

Since $x \in \Omega$, the above must hold as an equality, or else x_1 would not be a global solution. Therefore, $x \in S$ and S is a convex set.

Problem 12.5

Recall

$$\begin{aligned} f(x) &= \|v(x)\|_\infty \\ &= \max |v_i(x)|, \ i = 1, \ldots, m. \end{aligned}$$

Minimizing f is equivalent to minimizing t where $|v_i(x)| \leq t,\ i = 1, \ldots, m$; i.e., the problem can be reformulated as

$$\begin{aligned} &\min_x \ t \\ &\text{s.t. } t - v_i(x) \geq 0,\ i = 1, \ldots, m, \\ &\qquad t + v_i(x) \geq 0,\ i = 1, \ldots, m. \end{aligned}$$

Similarly, for $f(x) = \max\ v_i(x),\ i = 1, \ldots, m$, the minimization problem can be reformulated as

$$\begin{aligned} &\min_x \ t \\ &\text{s.t. } t - v_i(x) \geq 0,\ i = 1, \ldots, m. \end{aligned}$$

Problem 12.7

Given

$$d = -\left(I - \frac{\nabla c_1(x) \nabla c_1^T(x)}{\|\nabla c_1(x)\|^2}\right) \nabla f(x),$$

we find

$$\begin{aligned} \nabla c_1^T(x) d = &\ -\nabla c_1^T(x) \left(I - \frac{\nabla c_1(x) \nabla c_1^T(x)}{\|\nabla c_1(x)\|^2}\right) \nabla f(x) \\ = &\ -\nabla c_1^T(x) \nabla f(x) + \frac{(\nabla c_1^T(x) \nabla c_1(x))(\nabla c_1^T(x) \nabla f(x))}{\|\nabla c_1(x)\|^2} \\ = &\ 0. \end{aligned}$$

Furthermore,

$$\begin{aligned} \nabla f^T(x) d = &\ -\nabla f^T(x) \left(I - \frac{\nabla c_1(x) \nabla c_1^T(x)}{\|\nabla c_1(x)\|^2}\right) \nabla f(x) \\ = &\ -\nabla f^T(x) \nabla f(x) + \frac{(\nabla f^T(x) \nabla c_1(x))(\nabla c_1^T(x) \nabla f(x))}{\|\nabla c_1(x)\|^2} \\ = &\ -\|\nabla f(x)\|^2 + \frac{(\nabla f^T(x) \nabla c_1(x))^2}{\|\nabla c_1(x)\|^2} \end{aligned}$$

The Hölder Inequality yields

$$\begin{aligned} |\nabla f^T(x) \nabla c_1(x)| \leq &\ \|\nabla f^T(x)\| \|\nabla c_1(x)\| \\ \Rightarrow (\nabla f^T(x) \nabla c_1(x))^2 \leq &\ \|\nabla f^T(x)\|^2 \|\nabla c_1(x)\|^2, \end{aligned}$$

and our assumption that (12.10) does not hold implies that the above is satisfied as a strict inequality. Thus,

$$\begin{aligned}\nabla f^T(x)d &= -\|\nabla f(x)\|^2 + \frac{(\nabla f^T(x)\nabla c_1(x))^2}{\|\nabla c_1(x)\|^2} \\ &< -\|\nabla f(x)\|^2 + \frac{\|\nabla f(x)\|^2\|\nabla c_1(x)\|^2}{\|\nabla c_1(x)\|^2} \\ &= 0.\end{aligned}$$

Problem 12.13

The constraints can be written as

$$\begin{aligned} c_1(x) = 2 - (x_1-1)^2 - (x_2-1)^2 \geq 0, \\ c_2(x) = 2 - (x_1-1)^2 - (x_2+1)^2 \geq 0, \\ c_3(x) = x_1 \geq 0, \end{aligned}$$

so

$$\nabla c_1(x) = \begin{bmatrix} -2(x_1-1) \\ -2(x_2-1) \end{bmatrix}, \quad \nabla c_2(x) = \begin{bmatrix} -2(x_1-1) \\ -2(x_2+1) \end{bmatrix}, \quad \nabla c_3(x) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

All constraints are active at $x^* = (0,0)$. The number of active constraints is three, but the dimension of the problem is only two, so $\{\nabla c_i \mid i \in A(x^*)\}$ is not a linearly independent set and LICQ does not hold. However, for $w = (1,0)$, $\nabla c_i(x^*)^T w > 0$ for all $i \in A(x^*)$, so MFCQ does hold.

Problem 12.14

The optimization problem can be formulated as

$$\begin{aligned} &\min_x \; f(x) = \|x\|^2 \\ &\text{s.t. } c(x) = a^T x + \alpha \geq 0. \end{aligned}$$

The Lagrangian function is

$$\begin{aligned} L(x,\lambda) &= f(x) - \lambda c(x) \\ &= \|x\|^2 - \lambda(a^T x + \alpha) \end{aligned}$$

and its derivatives are

$$\begin{aligned}\nabla_x L(x,\lambda) &= 2x - \lambda a \\ \nabla_{xx} L(x,\lambda) &= 2I.\end{aligned}$$

Notice that the second order sufficient condition $\nabla_{xx}L(x,\lambda) = 2I > 0$ is satisfied at all points.

The KKT conditions $\nabla_x L(x^*,\lambda^*) = 0$, $\lambda^* c(x^*) = 0$, $\lambda^* \geq 0$ imply

$$x^* = \frac{\lambda^*}{2}a$$

and

$$\lambda^* = 0 \quad \text{or} \quad a^T x^* + \alpha = \frac{\lambda^* \|a\|^2}{2} + \alpha = 0.$$

There are two cases. First, if $\alpha \geq 0$, then the latter condition implies $\lambda^* = 0$, so the solution is $(x^*,\lambda^*) = (0,0)$. Second, if $\alpha < 0$, then

$$(x^*,\lambda^*) = -\left(\frac{\alpha}{\|a\|^2}a, \frac{2}{\|a\|^2}\right)$$

Problem 12.16

Eliminating the x_2 variable yields

$$x_2 = \pm\sqrt{1 - x_1^2}$$

There are two cases:

Case 1: Let $x_2 = \sqrt{1-x_1^2}$. The optimization problem becomes

$$\min_{x_1} \; f(x_1) = x_1 + \sqrt{1 - x_1^2}.$$

The first order condition is

$$\nabla f = 1 - \frac{x_1}{\sqrt{1-x_1^2}} = 0,$$

which is satisfied by $x_1 = \pm 1/\sqrt{2}$. Plugging each into f and choosing the value for x_1 that yields a smaller objective value, we find the solution to be $(x_1, x_2) = (-1/\sqrt{2}, 1/\sqrt{2})$.

Case 2: Let $x_2 = -\sqrt{1-x_1^2}$. The optimization problem becomes

$$\min_{x_1} \; f(x_1) = x_1 - \sqrt{1-x_1^2}.$$

The first order condition is

$$\nabla f = 1 + \frac{x_1}{\sqrt{1-x_1^2}} = 0,$$

which is satisfied by $x_1 = \pm 1/\sqrt{2}$. Plugging each into f and choosing the value for x_1 that yields a smaller objective value, we find the solution to be $(x_1, x_2) = (-1/\sqrt{2}, -1/\sqrt{2})$.

Each choice of sign leads to a distinct solution. However, only case 2 yields the optimal solution

$$x^* = \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right).$$

Problem 12.18

The problem is

$$\begin{aligned} &\min_{x,y} \; (x-1)^2 + (y-2)^2 \\ &\text{s.t. } (x-1)^2 - 5y = 0. \end{aligned}$$

The Lagrangian is

$$\begin{aligned} L(x,y,\lambda) = &\; (x-1)^2 + (y-2)^2 - \lambda((x-1)^2 - 5y) \\ = &\; (1-\lambda)(x-1)^2 + (y-2)^2 + 5\lambda y, \end{aligned}$$

which implies

$$\begin{aligned} \frac{\partial}{\partial x} L(x,y,\lambda) &= 2(1-\lambda)(x-1) \\ \frac{\partial}{\partial y} L(x,y,\lambda) &= 2(y-2) + 5\lambda. \end{aligned}$$

The KKT conditions are

$$\begin{aligned} 2(1-\lambda^*)(x^*-1) = &\; 0 \\ 2(y^*-2) + 5\lambda^* = &\; 0 \\ (x^*-1)^2 - 5y^* = &\; 0. \end{aligned}$$

Solving for x^*, y^*, and λ^*, we find $x^* = 1$, $y^* = 0$, and $\lambda^* = \frac{4}{5}$ as the only real solution. At $(x^*, y^*) = (1, 0)$, we have

$$\nabla c(x,y)|_{(x^*,y^*)} = \begin{bmatrix} 2(x-1) \\ -5 \end{bmatrix}\Bigg|_{(x^*,y^*)} = \begin{bmatrix} 0 \\ -5 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

so LICQ is satisfied.

Now we show that $(x^*, y^*) = (1, 0)$ is the optimal solution, with $f^* = 4$. We find

$$\begin{aligned} w \in F_2(\lambda^*) &\Leftrightarrow w = (w_1, w_2) \text{ satisfies } [\nabla c(x^*, y^*)]^T w = 0 \\ &\Leftrightarrow \begin{bmatrix} 0 & -5 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = 0 \\ &\Rightarrow w_2 = 0, \end{aligned}$$

then for all $w = (w_1, 0)$ where $w_1 \neq 0$,

$$\begin{aligned} w^T \nabla^2 L(x^*, y^*, \lambda^*) w &= \begin{bmatrix} w_1 & 0 \end{bmatrix} \begin{bmatrix} 2(1-\frac{4}{5}) & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} w_1 \\ 0 \end{bmatrix} \\ &= \frac{2}{5} w_1^2 > 0 \text{ (for } w_1 \neq 0). \end{aligned}$$

Thus from the second-order sufficient condition, we find $(1, 0)$ is the optimal solution.

Finally, we substitute $(x-1)^2 = 5y$ into the objective function and get the following unconstrained optimization problem:

$$\min_y \; 5y + (y-2)^2 = y^2 + y + 4.$$

Notice that $y^2 + y + 4 = (y + \frac{1}{2})^2 + \frac{15}{4} \geq \frac{15}{4}$, so $\tilde{y} = -1/2$ yields an objective value of $15/4 < 4$. Therefore, optimal solutions to this problem cannot yield solutions to the original problem.

Problem 12.21

We write the problem in the form:

$$\begin{aligned} &\min_{x_1, x_2} \; -x_1 x_2 \\ &\text{ s.t. } 1 - x_1^2 - x_2^2 \geq 0. \end{aligned}$$

The Lagrangian function is

$$L(x_1, x_2, \lambda) = -x_1 x_2 - \lambda(1 - x_1^2 - x_2^2).$$

The KKT conditions are

$$\begin{aligned} -x_2 - \lambda(-2x_1) &= 0 \\ -x_1 - \lambda(-2x_2) &= 0 \\ \lambda &\geq 0 \\ \lambda(1 - x_1^2 - x_2^2) &= 0. \end{aligned}$$

We solve this system to get three KKT points:

$$(x_1, x_2, \lambda) \in \left\{ (0,0,0), \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, \frac{1}{2} \right), \left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, \frac{1}{2} \right) \right\}$$

Checking the second order condition at each KKT point, we find

$$(x_1, x_2) \in \left\{ \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right), \left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2} \right) \right\}$$

are the optimal points.

13 Linear Programming: The Simplex Method

Problem 13.1

We first add slack variables z to the constraint $A_2 x + B_2 y \leq b_2$ and change it into

$$A_2 x + B_2 y + z = b_2, \quad z \geq 0.$$

Then we introduce surplus variables s_1 and slack variables s_2 into the two-sided bound constraint $l \leq y \leq u$:

$$y - s_1 = l, \quad y + s_2 = u, \quad s_1 \geq 0, \quad s_2 \geq 0.$$

Splitting x and y into their nonnegative and nonpositive parts, we have

$$\begin{aligned} &x = x^+ - x^-, \quad x^+ = \max(x, 0) \geq 0, \quad x^- = \max(-x, 0) \geq 0, \\ &y = y^+ - y^-, \quad y^+ = \max(y, 0) \geq 0, \quad y^- = \max(-y, 0) \geq 0. \end{aligned}$$

Therefore the objective function and the constraints can be restated as:

$$\begin{array}{rcl}
\max c^T x + d^T y & \Leftrightarrow & \min -c^T(x^+ - x^-) - d^T(y^+ - y^-) \\
A_1 x = b_1 & \Leftrightarrow & A_1(x^+ - x^-) = b_1 \\
A_2 x + B_2 y \le b_2 & \Leftrightarrow & A_2(x^+ - x^-) + B_2(y^+ - y^-) + z = b_2 \\
l \le y \le u & \Leftrightarrow & y^+ - y^- - s_1 = l, y^+ - y^- + s_2 = u,
\end{array}$$

with all the variables $(x^+, x^-, y^+, y^-, z, s_1, s_2)$ nonnegative. Hence the standard form of the given linear program is:

$$\text{minimize}_{x^+,x^-,y^+,y^-,z,s_1,s_2} \quad \begin{bmatrix} -c \\ c \\ -d \\ d \\ 0 \\ 0 \\ 0 \end{bmatrix}^T \begin{bmatrix} x^+ \\ x^- \\ y^+ \\ y^- \\ z \\ s_1 \\ s_2 \end{bmatrix}$$

$$\text{subject to} \quad \begin{bmatrix} A_1 & -A_1 & 0 & 0 & 0 & 0 & 0 \\ A_2 & -A_2 & B_2 & -B_2 & I & 0 & 0 \\ 0 & 0 & I & -I & 0 & -I & 0 \\ 0 & 0 & I & -I & 0 & 0 & I \end{bmatrix} \begin{bmatrix} x^+ \\ x^- \\ y^+ \\ y^- \\ z \\ s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ l \\ u \end{bmatrix}$$

$$x^+, x^-, y^+, y^-, z, s_1, s_2 \ge 0.$$

Problem 13.5

It is sufficient to show that the two linear programs have identical KKT systems. For the first linear program, let π be the vector of Lagrangian multipliers associated with $Ax \ge b$ and s be the vector of multipliers associated with $x \ge 0$. The Lagrangian function is then

$$L_1(x, \pi, s) = c^T x - \pi^T(Ax - b) - s^T x.$$

The KKT system of this problem is given by

$$\begin{array}{rcl}
A^T \pi + s & = & c \\
Ax & \ge & b \\
x & \ge & 0 \\
\pi & \ge & 0 \\
s & \ge & 0 \\
\pi^T(Ax - b) & = & 0 \\
s^T x & = & 0.
\end{array}$$

For the second linear program, we know that $\max b^T\pi \Leftrightarrow \min -b^T\pi$. Similarly, let x be the vector of Lagrangian multipliers associated with $A^T\pi \leq c$ and y be the vector of multipliers associated with $\pi \geq 0$. By introducing the Lagrangian function

$$L_2(\pi, x, y) = -b^T\pi - x^T(c - A^T\pi) - y^T\pi,$$

we have the KKT system of this linear program:

$$\begin{array}{rcl} Ax - b & = & y \\ A^T\pi & \leq & c \\ \pi & \geq & 0 \\ x & \geq & 0 \\ y & \geq & 0 \\ x^T(c - A^T\pi) & = & 0 \\ y^T\pi & = & 0. \end{array}$$

Defining $s = c - A^T\pi$ and noting that $y = Ax - b$, we can easily verify that the two KKT systems are identical, which is the desired argument.

Problem 13.6

Assume that there does exist a basic feasible point $\hat{x}$ for linear program (13.1), where $m \leq n$ and the rows of A are linearly dependent. Also assume without loss of generality that $\mathcal{B}(\hat{x}) = \{1, 2, \ldots, m\}$. The matrix $B = [A_i]_{i=1,2,\ldots,m}$ is nonsingular, where A_i is the i-th column of A.

On the other hand, since $m \leq n$ and the rows of A are linearly dependent, there must exist $1 \leq k \leq m$ such that the k-th row of A can be expressed as a linear combination of other rows of A. Hence, with the same coefficients, the k-th row of B can also expressed as a linear combination of other rows of B. This implies that B is singular, which obviously contradicts the argument that B is nonsingular. Then our assumption that there is a basic feasible point for (13.1) must be incorrect. This completes the proof.

Problem 13.10

By equating the last row of L_1U_1 to the last row of $P_1L^{-1}B^+P_1^T$, we have the following linear system of 4 equations and 4 unknowns:

$$\begin{array}{lllllllll} l_{52}u_{33} & & & & & & & = & u_{23} \\ l_{52}u_{34} & + & l_{53}u_{44} & & & & & = & u_{24} \\ l_{52}u_{35} & + & l_{53}u_{45} & + & l_{54}u_{55} & & & = & u_{25} \\ l_{52}w_3 & + & l_{53}w_4 & + & l_{54}w_5 & + & \hat{w}_2 & = & w_2. \end{array}$$

We can either successively retrieve the values of l_{52}, l_{53}, l_{54} and $\hat{w}_2$ from

$$\begin{aligned}
l_{52} &= u_{23}/u_{33} \\
l_{53} &= (u_{24} - l_{52}u_{34})/u_{44} \\
l_{54} &= (u_{25} - l_{52}u_{35} - l_{53}u_{45})/u_{55} \\
\hat{w}_2 &= w_2 - l_{52}w_3 - l_{53}w_4 - l_{54}w_5,
\end{aligned}$$

or calculate these values from the unknown quantities using

$$\begin{aligned}
l_{52} &= u_{23}/u_{33} \\
l_{53} &= (u_{24}u_{33} - u_{23}u_{34})/(u_{33}u_{44}) \\
l_{54} &= (u_{25}u_{33}u_{44} - u_{23}u_{35}u_{44} - u_{24}u_{33}u_{45} + u_{23}u_{34}u_{45})/(u_{33}u_{44}u_{55}) \\
\hat{w}_2 &= w_2 - w_3\frac{u_{23}}{u_{33}} - w_4\frac{u_{24}u_{33} - u_{23}u_{34}}{u_{33}u_{44}} \\
&\quad - w_5\frac{u_{25}u_{33}u_{44} - u_{23}u_{35}u_{44} - u_{24}u_{33}u_{45} + u_{23}u_{34}u_{45}}{u_{33}u_{44}u_{55}}.
\end{aligned}$$

14 Linear Programming: Interior-Point Methods

Problem 14.1

The primal problem is

$$\begin{aligned}
\min_{x_1,x_2} \; & x_1 \\
\text{s.t. } & x_1 + x_2 = 1 \\
& (x_1, x_2) \geq 0,
\end{aligned}$$

so the KKT conditions are

$$F(x, \lambda, s) = \begin{pmatrix} x_1 + x_2 - 1 \\ \lambda + s_1 - 1 \\ \lambda + s_2 \\ x_1 s_1 \\ x_2 s_2 \end{pmatrix} = 0,$$

with $(x_1, x_2, s_1, s_2) \geq 0$. The solution to the KKT conditions is

$$(x_1, x_2, s_1, s_2, \lambda) = (0, 1, 1, 0, 0),$$

but $F(x, \lambda, s)$ also has the spurious solution

$$(x_1, x_2, s_1, s_2, \lambda) = (1, 0, 0, -1, 1).$$

Problem 14.2

(i) For any $(x, \lambda, s) \in \mathcal{N}_2(\theta_1)$, we have

$$Ax = b \tag{23a}$$
$$A^{\mathrm{T}}\lambda + s = c \tag{23b}$$
$$x > 0 \tag{23c}$$
$$s > 0 \tag{23d}$$
$$\|XSe - \mu e\|_2 \leq \theta_1 \mu. \tag{23e}$$

Given $0 \leq \theta_1 < \theta_2 < 1$, equation (23e) implies

$$\|XSe - \mu e\|_2 \leq \theta_1 \mu < \theta_2 \mu. \tag{24}$$

From equations (23a)–(23d),(24), we have $(x, \lambda, s) \in \mathcal{N}_2(\theta_2)$. Thus $\mathcal{N}_2(\theta_1) \subset \mathcal{N}_2(\theta_2)$ when $0 \leq \theta_1 < \theta_2 < 1$.

For any $(x, \lambda, s) \in \mathcal{N}_{-\infty}(\gamma_1)$, we have

$$Ax = b \tag{25a}$$
$$A^{\mathrm{T}}\lambda + s = c \tag{25b}$$
$$x > 0 \tag{25c}$$
$$s > 0 \tag{25d}$$
$$x_i s_i \geq \gamma_1 \mu, \quad i = 1, 2, \ldots, n. \tag{25e}$$

Given $0 < \gamma_2 \leq \gamma_1 \leq 1$, equation (25d) implies

$$x_i s_i \geq \gamma_1 \mu \geq \gamma_2 \mu. \tag{26}$$

We have from equations (25a)–(25d),(26) that $(x, \lambda, s) \in \mathcal{N}_{-\infty}(\gamma_2)$. This shows that $\mathcal{N}_{-\infty}(\gamma_1) \subset \mathcal{N}_{-\infty}(\gamma_2)$ when $0 < \gamma_2 \leq \gamma_1 \leq 1$.

(ii) For any $(x, \lambda, s) \in \mathcal{N}_2(\theta)$, we have

$$Ax = b \tag{27a}$$
$$A^{\mathrm{T}}\lambda + s = c \tag{27b}$$
$$x > 0 \tag{27c}$$
$$s > 0 \tag{27d}$$
$$\|XSe - \mu e\|_2 \leq \theta \mu. \tag{27e}$$

Equation (27e) implies

$$\sum_{i=1}^{n}(x_i s_i - \mu)^2 \le \theta^2 \mu^2. \tag{28}$$

Suppose that there exists some $k \in 1, 2, \ldots, n$ satisfying

$$x_k s_k < \gamma\mu \quad \text{where} \quad \gamma \le 1 - \theta. \tag{29}$$

We have

$$\begin{aligned} & x_k s_k < \gamma\mu \le (1-\theta)\mu \\ \Longrightarrow\ & x_k s_k - \mu < -\theta\mu < 0 \\ \Longrightarrow\ & (x_k s_k - \mu)^2 > \theta^2\mu^2. \end{aligned}$$

Obviously, this contradicts equation (28), so we must have $x_k s_k \ge \gamma\mu$ for all $k = 1, 2, \ldots, n$. This conclusion, together with equations (27a)–(27d), gives $(x, \lambda, s) \in N_{-\infty}(\gamma)$. Therefore $N_2(\theta) \subset N_{-\infty}(\gamma)$ when $\gamma \le 1 - \theta$.

Problem 14.3

For $(\bar{x}, \bar{\lambda}, \bar{s}) \in \mathcal{N}_{-\infty}(\gamma)$ the following conditions hold:

$$(\bar{x}, \bar{\lambda}, \bar{s}) \in \mathcal{F}^0, \tag{30}$$

$$\bar{x}_i \bar{s}_i \ge \gamma\mu, \quad i = 1, \ldots, n. \tag{31}$$

Therefore, for an arbitrary point $(x, \lambda, s) \in \mathcal{F}^0$ we have $(x, \lambda, s) \in \mathcal{N}_{-\infty}(\gamma)$ if and only if condition (31) holds. Notice that

$$x_i s_i \ge \gamma\mu \;\Leftrightarrow\; \frac{x_i s_i}{\mu} \ge \gamma \;\Leftrightarrow\; \frac{n x_i s_i}{x^T s} \ge \gamma.$$

Therefore, the range of γ such that $(x, \lambda, s) \in \mathcal{N}_{-\infty}(\gamma)$ is equal to the set

$$\Gamma = \left\{ \gamma \;:\; \gamma \le \min_{1 \le i \le n} \frac{n x_i s_i}{x^T s} \right\}.$$

Problem 14.4

First, notice that if $\|XSe - \mu e\|_2 > \theta\mu$ holds for $\theta = 1$, then it must hold for every $\theta \in [0, 1)$. For $n = 2$,

$$\begin{aligned}
\|XSe - \mu e\|_2 > \mu \;&\Leftrightarrow\; (x_1s_1 - \mu)^2 + (x_2s_2 - \mu)^2 > \mu^2 \\
&\Leftrightarrow\; \left(\frac{x_1s_1 - x_2s_2}{2}\right)^2 + \left(\frac{x_2s_2 - x_1s_1}{2}\right)^2 > \left(\frac{x_1s_1 + x_2s_2}{2}\right)^2 \\
&\Leftrightarrow\; 2(x_1s_1 - x_2s_2)^2 > (x_1s_1 + x_2s_2)^2 \\
&\Leftarrow\; \sqrt{2}(x_1s_1 - x_2s_2) > x_1s_1 + x_2s_2 \\
&\Leftrightarrow\; \frac{x_1s_1}{x_2s_2} > \frac{\sqrt{2}+1}{\sqrt{2}-1} \approx 5.8284,
\end{aligned}$$

which holds, for example, when

$$x = \begin{pmatrix} 6 \\ 1 \end{pmatrix} \quad \text{and} \quad s = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Problem 14.5

For $(x, \lambda, s) \in \mathcal{N}_{-\infty}(1)$ the following conditions hold:

$$(x, \lambda, s) \in \mathcal{F}^0 \tag{32}$$

$$x_is_i \geq \mu,\; i = 1, \dots, n. \tag{33}$$

Assume that for some $i = 1, \dots, n$ we have $x_is_i > \mu$. Then,

$$\sum_{i=1}^{n} x_is_i > n\mu \;\Leftrightarrow\; \frac{x^Ts}{n} > \mu \;\Leftrightarrow\; \mu > \mu,$$

which is a contradiction. Therefore, $x_is_i = \mu$ for $i = 1, \dots, n$. Along with condition (32), this coincides with the central path $\mathcal{C}$.

For $(x, \lambda, s) \in \mathcal{N}_2(0)$ the following conditions hold:

$$(x, \lambda, s) \in \mathcal{F}^0 \tag{34}$$

$$\sum_{i=1}^{n} (x_is_i - \mu)^2 \leq 0. \tag{35}$$

If $x_i s_i \neq \mu$ for some $i = 1, \ldots, n$, then

$$\sum_{i=1}^{n} (x_i s_i - \mu)^2 > 0,$$

which contradicts condition (35). Therefore, $x_i s_i = \mu$ for $i = 1, \ldots, n$ which, along with condition (34), coincides with $\mathcal{C}$.

Problem 14.7

Assuming

$$\lim_{x_i s_i \to 0} \mu = \lim_{x_i s_i \to 0} x^T s/n \neq 0,$$

i.e., $x_k s_k > 0$ for some $k \neq i$, we also have

$$\lim_{x_i s_i \to 0} x^T s \neq 0 \quad \text{and} \quad \lim_{x_i s_i \to 0} \log x^T s \neq -\infty.$$

Consequently,

$$\begin{aligned} \lim_{x_i s_i \to 0} \Phi_\rho &= \lim_{x_i s_i \to 0} \left(\rho \log x^T s - \sum_{i=1}^{n} \log x_i s_i \right) \\ &= \rho \lim_{x_i s_i \to 0} \log x^T s - \lim_{x_i s_i \to 0} \log x_1 s_1 - \cdots - \lim_{x_i s_i \to 0} \log x_n s_n \\ &= c - \lim_{x_i s_i \to 0} \log x_i s_i \\ &= \infty, \end{aligned}$$

as desired, where c is a finite constant.

Problem 14.8

First, assume the coefficient matrix

$$M = \begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix}$$

is nonsingular. Let

$$M_1 = \begin{bmatrix} 0 & A^T & I \end{bmatrix}, \quad M_2 = \begin{bmatrix} A & 0 & 0 \end{bmatrix}, \quad M_3 = \begin{bmatrix} S & 0 & X \end{bmatrix},$$

then the nonsingularity of M implies that the rows of M_2 are linearly independent. Thus, A has full row rank.

Second, assume A has full row rank. If M is singular, then certain rows of M can be expressed as a linear combination of its other rows. We denote one of these such rows as row m. Since I, S, X are all diagonal matrices with positive diagonal elements, we observe that m is neither a row of M_1 nor a row of M_3. Thus m must be a row of M_2. Due to the structure of I, S, and X, m must be expressed as a linear combination of rows of M_2 itself. However, this contradicts our assumption that A has full row rank, so M must be nonsingular.

Problem 14.9

According to the assumptions, the following equalities hold

$$\begin{aligned} A\Delta x &= 0 \quad (36) \\ A^T\Delta\lambda + \Delta s &= 0. \quad (37) \end{aligned}$$

Multiplying equation (36) on the left by $\Delta\lambda^T$ and equation (37) on the left by Δx^T yields

$$\begin{aligned} \Delta\lambda^T A\Delta x &= 0 \quad (38) \\ \Delta x^T A^T\Delta\lambda + \Delta x^T\Delta s &= 0. \quad (39) \end{aligned}$$

Subtracting equation (38) from (39) yields

$$\Delta x^T\Delta s = 0,$$

as desired.

Problem 14.12

That AD^2A^T is symmetric follows easily from the fact that

$$\left(AD^2A^T\right)^T = \left(A^T\right)^T\left(D^2\right)^T(A)^T = AD^2A^T$$

since D^2 is a diagonal matrix.

Assume that A has full row rank, i.e.,

$$A^Ty = 0 \quad \Rightarrow \quad y = 0.$$

Let $x \neq 0$ be any vector in $\mathcal{R}^m$ and notice:

$$\begin{aligned} x^TAD^2A^Tx &= x^TADDA^Tx \\ &= \left(DA^Tx\right)^T\left(DA^Tx\right) \\ &= v^Tv \\ &= ||v||_2^2, \end{aligned}$$

where $v = DA^T x$ is a vector in $\mathcal{R}^m$. Due to the assumption that A has full row rank it follows that $A^T x \neq 0$, which implies $v \neq 0$ (since D is diagonal with all positive diagonal elements). Therefore,

$$x^T AD^2 A^T x = ||v||_2^2 > 0,$$

so the coefficient matrix AD^2A^T is positive definite whenever A has full row rank.

Now, assume that AD^2A^T is positive definite, i.e.,

$$x^T AD^2 A^T x > 0$$

for all nonzero $x \in \mathcal{R}^m$. If some row of A could be expressed as a linear combination of other rows in A, then $A^T y = 0$ for some nonzero $y \in \mathcal{R}^m$. However, this would imply

$$y^T AD^2 A^T y = \left(y^T AD^2\right)\left(A^T y\right) = 0,$$

which contradicts the assumption that AD^2A^T is positive definite. Therefore, A must have full row rank.

Finally, consider replacing D by a diagonal matrix in which exactly m of the diagonal elements are positive and the remainder are zero. Without loss of generality, assume that the first m diagonal elements of m are positive. A real symmetric matrix M is positive definite if and only if there exists a real nonsingular matrix Z such that

$$M = ZZ^T. \tag{40}$$

Notice that

$$C = AD^2A^T = (AD)(AD)^T = \left(BD'\right)\left(BD'\right)^T,$$

where B is the submatrix corresponding to the first m columns of A and D' is the $m \times m$ diagonal submatrix of D with all positive diagonal elements. Therefore, according to (40), the desired results can be extended in this case if and only if BD' is nonsingular, which is guaranteed if the resulting matrix B has linearly independent columns.

Problem 14.13

A Taylor series approximation to $\mathcal{H}$ near the point (x, λ, s) is of the form:

$$\begin{aligned}\left(\hat{x}(\tau), \hat{\lambda}(\tau), \hat{s}(\tau)\right) = & \left(\hat{x}(0), \hat{\lambda}(0), \hat{s}(0)\right) \\ & + \tau\left(\hat{x}'(0), \hat{\lambda}'(0), \hat{s}'(0)\right) \\ & + \frac{1}{2}\tau^2\left(\hat{x}''(0), \hat{\lambda}''(0), \hat{s}''(0)\right) + \cdots,\end{aligned}$$

where $\left(\hat{x}^{(j)}(0), \hat{\lambda}^{(j)}(0), \hat{s}^{(j)}(0)\right)$ is the jth derivative of $\left(\hat{x}(\tau), \hat{\lambda}(\tau), \hat{s}(\tau)\right)$ with respect to τ, evaluated at $\tau = 0$. These derivatives can be determined by implicitly differentiating both sides of the equality given as the definition of $\mathcal{H}$. First, notice that $\left(\hat{x}'(\tau), \hat{\lambda}'(\tau), \hat{s}'(\tau)\right)$ solves

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ \hat{S}(\tau) & 0 & \hat{X}(\tau) \end{bmatrix} \begin{bmatrix} \hat{x}'(\tau) \\ \hat{\lambda}'(\tau) \\ \hat{s}'(\tau) \end{bmatrix} = \begin{bmatrix} -r_c \\ -r_b \\ -XSe \end{bmatrix}. \tag{41}$$

After setting $\tau = 0$ and noticing that $\hat{X}(0) = X$ and $\hat{S}(0) = S$, the linear system in (41) reduces to

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \hat{x}'(0) \\ \hat{\lambda}'(0) \\ \hat{s}'(0) \end{bmatrix} = \begin{bmatrix} -r_c \\ -r_b \\ -XSe \end{bmatrix}, \tag{42}$$

which is exactly the system in (14.8). Therefore,

$$\left(\hat{x}'(0), \hat{\lambda}'(0), \hat{s}'(0)\right) = \left(\Delta x^{\text{aff}}, \Delta\lambda^{\text{aff}}, \Delta s^{\text{aff}}\right). \tag{43}$$

Differentiating (41) with respect to τ yields

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ \hat{S}(\tau) & 0 & \hat{X}(\tau) \end{bmatrix} \begin{bmatrix} \hat{x}''(\tau) \\ \hat{\lambda}''(\tau) \\ \hat{s}''(\tau) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -2\hat{X}'(\tau)\hat{S}'(\tau)e \end{bmatrix}. \tag{44}$$

If we let $(\Delta x^{\text{corr}}, \Delta\lambda^{\text{corr}}, \Delta s^{\text{corr}})$ be the solution to the corrector step, i.e., when the right-hand-side of (14.8) is replaced by $\left(0, 0, -\Delta X^{\text{aff}}\Delta S^{\text{aff}}e\right)$, then after setting $\tau = 0$ and noting (43) we can see that

$$\left(\hat{x}''(0), \hat{\lambda}''(0), \hat{s}''(0)\right) = \frac{1}{2}\left(\Delta x^{\text{corr}}, \Delta\lambda^{\text{corr}}, \Delta s^{\text{corr}}\right). \tag{45}$$

Finally, differentiating (44) with respect to τ yields

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ \hat{S}(\tau) & 0 & \hat{X}(\tau) \end{bmatrix} \begin{bmatrix} \hat{x}'''(\tau) \\ \hat{\lambda}'''(\tau) \\ \hat{s}'''(\tau) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -3\left(\hat{X}''(\tau)\hat{S}'(\tau) + \hat{S}''(\tau)\hat{X}'(\tau)\right)e \end{bmatrix} \tag{46}$$

Setting $\tau = 0$ and noting (43) and (45), we find

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \hat{x}'''(0) \\ \hat{\lambda}'''(0) \\ \hat{s}'''(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -\frac{3}{2}\left(\Delta X^{\text{corr}}\Delta S^{\text{aff}} + \Delta S^{\text{corr}}\Delta X^{\text{aff}}\right)e \end{bmatrix} \tag{47}$$

In total, a Taylor series approximation to $\mathcal{H}$ is given by

$$\begin{aligned} \left(\hat{x}(\tau), \hat{\lambda}(\tau), \hat{s}(\tau)\right) &= (x, \lambda, s) \\ &\quad + \tau\left(\Delta x^{\text{aff}}, \Delta\lambda^{\text{aff}}, \Delta s^{\text{aff}}\right) \\ &\quad + \tau^2\left(\Delta x^{\text{corr}}, \Delta\lambda^{\text{corr}}, \Delta s^{\text{corr}}\right) \\ &\quad + \frac{\tau^3}{3!}\left(\hat{x}'''(0), \hat{\lambda}'''(0), \hat{s}'''(0)\right), \end{aligned}$$

where $\left(\hat{x}'''(0), \hat{\lambda}'''(0), \hat{s}'''(0)\right)$ solves (47).

Problem 14.14

By introducing Lagrange multipliers for the equality constraints and the nonnegativity constraints, the Lagrangian function for this problem is given by

$$L(x, y, \lambda, s) = c^T x + d^T y - \lambda^T(A_1 x + A_2 y - b) - s^T x.$$

Applying Theorem 12.1, the first-order necessary conditions state that for (x^*, y^*) to be optimal there must exist vectors λ and s such that

$$\begin{aligned} A_1^T\lambda + s &= c, & (48) \\ A_2^T\lambda &= d, & (49) \\ A_1 x + A_2 y &= b, & (50) \\ x_i s_i &= 0, \quad i = 1, \ldots, n, & (51) \\ (x, s) &\geq 0. & (52) \end{aligned}$$

Equivalently, these conditions can be expressed as

$$F(x,y,\lambda,s) = \begin{bmatrix} A_1^T\lambda + s - c \\ A_2^T\lambda - d \\ A_1x + A_2y - b \\ XSe \end{bmatrix} = 0, \tag{53}$$

$$(x,s) \geq 0. \tag{54}$$

Similar to the standard linear programming case, the central path is described by the system including (48)-(52) where (51) is replaced by

$$x_i s_i = \tau, \quad i = 1,\dots,n.$$

The Newton step equations for $\tau = \sigma\mu$ are

$$\begin{bmatrix} 0 & 0 & A_1^T & I \\ 0 & 0 & A_2^T & 0 \\ A_1 & A_2 & 0 & 0 \\ S & 0 & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta\lambda \\ \Delta s \end{bmatrix} = \begin{bmatrix} -r_c \\ -r_d \\ -r_b \\ -XSe + \sigma\mu e \end{bmatrix} \tag{55}$$

where

$$r_b = A_1x + A_2y = b, \quad r_c = A_1^T\lambda + s - c, \quad \text{and} \quad r_d = A_2^T\lambda - d.$$

By eliminating Δs from (55), the augmented system is given by

$$\begin{bmatrix} 0 & 0 & A_2^T \\ A_1 & A_2 & 0 \\ -D^{-2} & 0 & A_1^T \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta\lambda \end{bmatrix} = \begin{bmatrix} -r_d \\ -r_b \\ -r_c + s - \sigma\mu X^{-1}e \end{bmatrix}, \tag{56}$$

$$\Delta s = -s + \sigma\mu X^{-1}e - D^{-2}\Delta x, \tag{57}$$

where $D = S^{-1/2}X^{1/2}$.

We can eliminate Δx from (56) by noting

$$\begin{aligned} -D^{-2}\Delta x + A_1^T\Delta\lambda &= -r_c + s - \sigma\mu X^{-1}e \\ \Longleftrightarrow \qquad \Delta x &= -D^2\left(r_c + s - \sigma\mu X^{-1}e - A_1^T\Delta\lambda\right), \end{aligned}$$

which yields the system

$$\begin{bmatrix} 0 & A_2^T \\ A_2 & A_1D^2A_1^T \end{bmatrix} \begin{bmatrix} \Delta y \\ \Delta\lambda \end{bmatrix} = \begin{bmatrix} -r_d \\ -r_b + A_1D^2\left(-r_c + s - \sigma\mu X^{-1}e\right) \end{bmatrix} \tag{58}$$

$$\Delta x = -D^2\left(-r_c + s - \sigma\mu X^{-1}e - A_1^T\Delta\lambda\right) \tag{59}$$

$$\Delta s = -s + \sigma\mu X^{-1}e - D^{-2}\Delta x. \tag{60}$$

Unfortunately, there is no way to reduce this system any further in general. That is, there is no way to create a system similar to the normal-equations in (14.44).

15 Fundamentals of Algorithms for Nonlinear Constrained Optimization

Problem 15.3

(a) The formulation is

$$\begin{aligned} \min \quad & x_1 + x_2 \\ \text{s.t.} \quad & x_1^2 + x_2^2 = 2 \\ & 0 \le x_1 \le 1 \\ & 0 \le x_2 \le 1. \end{aligned}$$

This problem has only one feasible point, namely $x_1 = x_2 = 1$. Thus it has a solution at $x^* = x_2^* = 1$, and the optimal objective is 2.

(b) The formulation is

$$\begin{aligned} \min \quad & x_1 + x_2 && \text{(61a)} \\ \text{s.t.} \quad & x_1^2 + x_2^2 \le 1 && \text{(61b)} \\ & x_1 + x_2 = 3 && \text{(61c)} \end{aligned}$$

Substituting equation (61c) into (61b), we get

$$x_1^2 + (3 - x_1)^2 \le 1 \quad \text{which implies} \quad x_1^2 - 3x_1 + 4 \le 0.$$

This inequality has no solution; thus the feasible region of the original problem is empty. This shows that the problem has no solution.

(c) The formulation is

$$\begin{aligned} \min \quad & x_1 x_2 \\ \text{s.t.} \quad & x_1 + x_2 = 2 \end{aligned}$$

Since the constraint of this problem is linear, we eliminate x_2 from the objective and get an unconstrained problem, namely

$$\min \quad x_1(2 - x_1) = -(x_1 - 1)^2 + 1.$$

Obviously, when $|x_1 - 1| \to +\infty$, we see that $-(x_1 - 1)^2 + 1 \to -\infty$. This shows that the original problem is unbounded below, hence it has no solution.

Problem 15.4

The optimization problem is

$$\min_{x,y} \; x^2 + y^2$$
$$\text{s.t. } (x-1)^3 = y^2.$$

If we eliminate x by writing it in terms of y, i.e. $x = \sqrt[3]{y^2} + 1$, then the above becomes the unconstrained problem

$$\min \; f(y) \equiv (y^{2/3} + 1)^2 + y^2.$$

Notice $f \geq 0$, so the optimal solution to the unconstrained problem is $y^* = 0$, which corresponds to the optimal solution $(x^*, y^*) = (1, 0)$ to the original problem.

Problem 15.5

We denote the i^{th} column of B^{-1} by y_i, $(i = 1, 2, \ldots, m)$, and the j^{th} column of $-B^{-1}N$ by z_j, $(j = 1, 2, \ldots n-m)$. The existence of B^{-1} shows that $y_1, y_2, \ldots, y_m$ are linearly independent. Let us consider

$$\begin{bmatrix} Y & Z \end{bmatrix} = \begin{bmatrix} B^{-1} & -B^{-1}N \\ 0 & I \end{bmatrix} = \begin{bmatrix} y_1 & y_2 & \ldots & y_m & z_1 & z_2 & \ldots & z_{n-m} \\ 0 & 0 & \ldots & 0 & e_1 & e_2 & \ldots & e_{n-m} \end{bmatrix}.$$

In order to see the linear dependence of $\begin{bmatrix} Y \; Z \end{bmatrix}$, we consider

$$k_1 \begin{bmatrix} y_1 \\ 0 \end{bmatrix} + k_2 \begin{bmatrix} y_2 \\ 0 \end{bmatrix} + \cdots + k_m \begin{bmatrix} y_m \\ 0 \end{bmatrix} + t_1 \begin{bmatrix} z_1 \\ e_1 \end{bmatrix} + t_2 \begin{bmatrix} z_2 \\ e_2 \end{bmatrix} + \cdots + t_{n-m} \begin{bmatrix} z_{n-m} \\ e_{n-m} \end{bmatrix} = 0. \tag{62}$$

The last $(n-m)$ equations of (62) are in fact

$$t_1 e_1 + t_2 e_2 + \cdots + t_{n-m} e_{n-m} = 0,$$

where $e_j = \begin{bmatrix} 0 \; 0 \; \cdots \; 0 \; 1 \; 0 \; 0 \; \cdots \; 0 \end{bmatrix}^{\mathrm{T}}$. Thus $t_1 = t_2 = \cdots = t_{n-m} = 0$. This shows that the first m equations of (62) are

$$k_1 y_1 + k_2 y_2 + \cdots + k_m y_m = 0.$$

It follows immediately that $k_1 = k_2 = \cdots = k_m = t_1 = t_2 = \cdots = t_{n-m} = 0$, which indicates that the collection of columns of $\begin{bmatrix} Y \; Z \end{bmatrix}$ form a linearly independent basis of $\mathbb{R}^n$.

Problem 15.6

Recall $A^T\Pi = YR$. Since Π is a permutation matrix, we know $\Pi^T = \pi^{-1}$. Thus $A = \Pi R^T Y^T$. This gives

$$AA^T = \Pi R^T Y^T Y R \Pi^T. \tag{63}$$

The matrix $\begin{bmatrix} Y & Z \end{bmatrix}$ is orthogonal, so $Y^T Y = I$. Then (63) gives

$$\begin{aligned}
AA^T &= \Pi R^T R \Pi^T \\
(AA^T)^{-1} &= \Pi R^{-1} R^{-T} \Pi^T \\
A^T(AA^T)^{-1} &= (YR\Pi^T)\Pi R^{-1} R^{-T} \Pi^T \\
A^T(AA^T)^{-1} &= YR^{-T}\Pi^T \\
A^T(AA^T)^{-1}b &= YR^{-T}\Pi^T b.
\end{aligned}$$

Problem 15.7

(a) We denote the i^{th} column of matrix

$$\begin{bmatrix} I \\ (B^{-1}N)^T \end{bmatrix} = \begin{bmatrix} | & \dots & | & \dots & | \\ y_1 & \dots & y_i & \dots & y_n \\ | & \dots & | & \dots & | \end{bmatrix} \qquad \text{by } y_i.$$

Then

$$\|y_i\|^2 = 1 + \|(B^{-1}N)_i\|^2 \geq 1.$$

Thus Y is no longer of norm 1. The same argument holds for the matrix

$$Z = \begin{bmatrix} -B^{-1}N \\ I \end{bmatrix}.$$

Furthermore,

$$\begin{aligned}
Y^T Z &= \begin{bmatrix} I & B^{-1}N \end{bmatrix} \begin{bmatrix} -B^{-1}N \\ I \end{bmatrix} = -B^{-1}N + B^{-1}N = 0, \\
AZ &= \begin{bmatrix} B & N \end{bmatrix} \begin{bmatrix} -B^{-1}N \\ I \end{bmatrix} = -BB^{-1}N + N = 0.
\end{aligned}$$

These show that the columns of Y and Z form an independent set and Y, Z are valid basis matrices.

(b) We have from $A = \begin{bmatrix} B & N \end{bmatrix}$ that

$$AA^T = \begin{bmatrix} B & N \end{bmatrix} \begin{bmatrix} B^T \\ N^T \end{bmatrix} = BB^T + NN^T.$$

Therefore,

$$\begin{aligned} AY &= \begin{bmatrix} B & N \end{bmatrix} \begin{bmatrix} I \\ (B^{-1}N)^T \end{bmatrix} = B + N(B^{-1}N)^T = B + NN^TB^{-T} \\ &= (BB^T + NN^T)B^{-T} = (AA^T)B^{-T}. \end{aligned}$$

And then,

$$\begin{aligned} & (AY)^{-1} = B^T(AA^T)^{-1} \\ & \Longrightarrow Y(AY)^{-1} = YB^T(AA^T)^{-1} \\ & \Longrightarrow Y(AY)^{-1}(AA^T) = YB^T(AA^T)^{-1}(AA^T) = YB^T \\ & \qquad = \begin{bmatrix} I \\ (B^{-1}N)^T \end{bmatrix} B^T = \begin{bmatrix} N^TB^{-T}B^T \end{bmatrix} = \begin{bmatrix} B^T \\ N^T \end{bmatrix} = A^T. \end{aligned}$$

This implies $Y(AY)^{-1} = A^T(AA^T)^{-1}$. Thus $Y(AY)^{-1}b = A^T(AA^T)^{-1}b$, which is the minimum norm solution of $Ax = b$.

Problem 15.8

The new problem is:

$$\begin{array}{ll} \min & \sin(x_1 + x_2) + x_3^2 + \frac{1}{3}\left(x_4 + x_5^4 + \frac{1}{2}x_6\right) \\ \text{s.t.} & 8x_1 - 6x_2 + x_3 + 9x_4 + 4x_5 = 6 \\ & 3x_1 + 2x_2 - x_4 + 6x_5 + 4x_6 = -4 \\ & 3x_1 + 2x_3 \geq 1. \end{array}$$

If we eliminate variables with (15.11):

$$\begin{pmatrix} x_3 \\ x_6 \end{pmatrix} = -\begin{pmatrix} 8 & -6 & 9 & 4 \\ \frac{3}{4} & \frac{1}{2} & -\frac{1}{4} & \frac{3}{2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} + \begin{pmatrix} 6 \\ -1 \end{pmatrix},$$

the objective function will turn out to be (15.12). We substitute (15.11) into the inequality constraint:

$$\begin{aligned} 1 &\leq 3x_1 + 2(-8x_1 + 6x_2 - 9x_4 - 4x_5 + 6) \\ &= -13x_1 + 12x_2 - 18x_4 - 8x_5 + 12 \end{aligned}$$

$$\Longrightarrow \qquad -13x_1 + 12x_2 - 18x_4 - 8x_5 \geq -11,$$

which is exactly (15.23). Thus the problem turns out to be minimizing function (15.12) subject to (15.23).

16 Quadratic Programming

Problem 16.1

(b) The optimization problem can be written as

$$\begin{aligned} &\min_x \frac{1}{2} x^T G x + d^T x \\ &\text{s.t. } c(x) \geq 0, \end{aligned}$$

where

$$G = \begin{bmatrix} -8 & -2 \\ -2 & -2 \end{bmatrix}, \quad d = \begin{bmatrix} -2 \\ -3 \end{bmatrix}, \quad \text{and} \quad c(x) = \begin{bmatrix} x_1 - x_2 \\ 4 - x_1 - x_2 \\ 3 - x_1 \end{bmatrix}.$$

Defining

$$A = \nabla c(x) = \begin{bmatrix} 1 & -1 \\ -1 & -1 \\ -1 & 0 \end{bmatrix},$$

we have the Lagrangian

$$L(x, \lambda) = \frac{1}{2} x^T G x + d^T x - \lambda^T c(x)$$

and its corresponding derivatives in terms of the x variables

$$\nabla_x L(x, \lambda) = Gx + d - A^T \lambda \quad \text{and} \quad \nabla_{xx} L(x, \lambda) = G.$$

Consider $x = (a, a) \in \Re^2$. It is easily seen that such an x is feasible for $a \leq 2$ and that

$$q(x) = -7a^2 - 5a \to -\infty \quad \text{as} \quad a \to -\infty.$$

Therefore, the problem is unbounded. Moreover, $\nabla_{xx} L = G < 0$, so no solution satisfies the second order necessary conditions are there are no local minimizers.

Problem 16.2

The problem is:

$$\begin{aligned} \min_x \ & \frac{1}{2}(x-x_0)^T(x-x_0) \\ \text{s.t. } & Ax = b. \end{aligned}$$

The KKT conditions are:

$$\begin{aligned} x^* - x_0 - A^T\lambda^* &= 0, & (64) \\ Ax^* &= b. & (65) \end{aligned}$$

Multiplying (64) on the left by A yields

$$Ax^* - Ax_0 - AA^T\lambda = 0. \qquad (66)$$

Substituting (65) into (66), we find

$$b - Ax_0 = AA^T\lambda,$$

which implies

$$\lambda^* = \left(AA^T\right)^{-1}(b - Ax_0). \qquad (67)$$

Finally, substituting (67) into (64) yields

$$x^* = x_0 + A^T(AA^T)^{-1}(b - Ax_0). \qquad (68)$$

Consider the case where $A \in \Re^{1\times n}$. Equation (68) gives

$$x^* - x_0 = A^T(AA^T)^{-1}(b - Ax_0) = \frac{1}{\|A\|_2}A^T(b - Ax_0),$$

so the optimal objective value is given by

$$\begin{aligned} f^* &= \frac{1}{2}(x^*-x_0)^T(x^*-x_0) \\ &= \frac{1}{2}\left(\frac{1}{\|A\|_2^2}\right)^2(b - ax_0)^TAA^T(b - Ax_0) \\ &= \frac{1}{2}\frac{1}{\|A\|_2^4}\left(\|A\|_2^2\right)(b - Ax_0)^T(b - Ax_0) \\ &= \frac{1}{2}\frac{1}{\|A\|_2^2}(b - Ax_0)^2. \end{aligned}$$

and the shortest distance from x_0 to the solution set of $Ax = b$ is

$$\sqrt{2f^*} = \sqrt{\frac{1}{\|A\|_2^2}(b - Ax_0)^2} = \frac{|b - Ax_0|}{\|A\|_2}.$$

Problem 16.6

First, we will show that the KKT conditions for problem (16.3) are satisfied by the point satisfying (16.4). The Lagrangian function for problem (16.3) is

$$L(x, \lambda) = \frac{1}{2}x^T Gx + d^T x - \lambda^T (Ax - b),$$

so the KKT conditions are

$$\begin{aligned} Gx + d - A^T\lambda &= 0 \\ Ax &= b. \end{aligned}$$

The point (x^*, λ^*) satisfies the KKT conditions if and only if

$$\begin{bmatrix} G & -A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x^* \\ \lambda^* \end{bmatrix} = \begin{bmatrix} -d \\ b \end{bmatrix},$$

which is exactly the system given by (16.4).

Now assume that the reduced Hessian $Z^T GZ$ is positive definite. The second order conditions for (16.3) are satisfied if $w^T \nabla_{xx} L(x^*, \lambda^*)w = w^T Gw > 0$ for all $w \in \mathcal{C}(x^*, \lambda^*), w \neq 0$. By definition, $w \in \mathcal{C}(x^*, \lambda^*)$ if $w = Zu$ for any real u, so

$$w^T Gw = u^T Z^T GZu > 0$$

and the second order conditions are satisfied.

Problem 16.7

Let $x = x^* + \alpha Zu, \alpha \neq 0$. We find

$$\begin{aligned} q(x) &= q(x^* + \alpha Zu) \\ &= \frac{1}{2}(x^* + \alpha Zu)^T G(x^* + \alpha Zu) + d^T (x^* + \alpha Zu) \\ &= \frac{1}{2}x^{*T} Gx^* + \alpha x^{*T} GZu + \frac{1}{2}\alpha^2 u^T Z^T GZu + d^T x^* + \alpha d^T Zu \\ &= q(x^*) + \frac{1}{2}\alpha^2 u^T A^T GZu + \alpha(x^{*T} GZu + d^T Zu). \end{aligned}$$

A point (x^*, λ^*) satisfying the KKT conditions yields

$$0 = Gx^* + d - A^T \lambda^*.$$

Taking the transpose and multiplying on the right by Zu, we find

$$0 = x^{*T} GZu + d^T Zu - \lambda^{*T} AZu = x^{*T} GZu + d^T Zu,$$

so in fact

$$q(x) = q(x^*) + \frac{1}{2}\alpha^2 u^T A^T GZu.$$

If there exists a u such that $u^T Z^T GZu < 0$, then $q(x) < q(x^*)$. Hence (x^*, λ^*) is a stationary point.

Problem 16.15

Suppose that there is a vector pair (x^*, λ^*) that satisfies the KKT conditions. Let u be some vector such that $u^T Z^T GZu \le 0$, and set $p = Zu$. Then for any $\alpha \neq 0$, we have

$$A(x^* + \alpha p) = b,$$

so that $x^* + \alpha p$ is feasible, while

$$\begin{aligned} q(x^* + \alpha p) = &\ q(x^*) + \alpha p^T (Gx^* + c) + \frac{1}{2}\alpha^2 p^T Gp \\ = &\ q(x^*) + \frac{1}{2}\alpha^2 p^T Gp \\ \le &\ q(x^*), \end{aligned}$$

where we have used the KKT condition $Gx^* + c = A^T \lambda^*$ and the fact that $p^T A^T \lambda^* = u^T Z^T A^T \lambda^* = 0$. Therefore, from any x^* satisfying the KKT conditions, we can find a feasible direction p along which q does not increase. In fact, we can always find a direction of strict decrease when $Z^T GZ$ has negative eigenvalues.

Problem 16.21

The KKT conditions of the quadratic program are

$$\begin{aligned} Gx + d - A^T \lambda - \bar{A}^T \mu = &\ 0, \\ Ax - b \ge &\ 0, \\ \bar{A}x - \bar{b} = &\ 0, \\ [Ax - b]_i \lambda_i = &\ 0, \ i = 1, \ldots, n \\ \lambda \ge &\ 0. \end{aligned}$$

Introducing slack variables y yields

$$\begin{aligned} Gx + d - A^T\lambda - \bar{A}^T\mu &= 0, \\ Ax - y - b &= 0, \\ \bar{A}x - \bar{b} &= 0, \\ y_i\lambda_i &= 0, \ i = 1, \ldots, n \\ (y, \lambda) &\geq 0, \end{aligned}$$

which can be expressed as

$$F(x, y, \lambda, \mu) = \begin{bmatrix} Gx + d - A^T\lambda - \bar{A}^T\mu \\ Ax - y - b \\ \bar{A}x - \bar{b} \\ Y\Lambda e \end{bmatrix} = 0.$$

The analog of (16.58) is

$$\begin{bmatrix} G & -A^T & -\bar{A}^T & 0 \\ A & 0 & 0 & -I \\ \bar{A} & 0 & 0 & 0 \\ 0 & 0 & Y & \Lambda \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta\lambda \\ \Delta\mu \\ \Delta y \end{bmatrix} = \begin{bmatrix} -r_d \\ -r_b \\ -r_{\bar{b}} \\ -\Lambda Y e + \sigma\mu e \end{bmatrix}$$

where

$$r_d = Gx + d - A^T\lambda, \quad r_b = Ax - y - b, \quad \text{and} \quad r_{\bar{b}} = \bar{A}x - \bar{b}.$$

17 Penalty and Augmented Lagrangian Methods

Problem 17.1

The following equality constrained problem

$$\begin{aligned} \min_x \ & -x^4 \\ \text{s.t. } & x = 0 \end{aligned}$$

has a local solution at $x^* = 0$. The corresponding quadratic penalty function is

$$Q(x; \mu) = -x^4 + \frac{1}{2}\mu x^2,$$

which is unbounded for any value of μ.

The inequality constrained problem

$$\min_x \; x^3 \quad \text{s.t. } x \geq 0$$

has a local solution at $x^* = 0$. The corresponding quadratic penalty function is

$$\begin{aligned} Q(x;\mu) &= x^3 + \frac{1}{2}\mu([x]^-)^2 \\ &= x^3 + \frac{1}{2}\mu(\max(-x,0))^2 \\ &= \begin{cases} x^3 & \text{if } x \geq 0 \\ x^3 + \frac{1}{2}\mu x^2 & \text{if } x < 0, \end{cases} \end{aligned}$$

which is unbounded for any value of μ.

Problem 17.5

The penalty function and its gradient are

$$Q(x;\mu) = -5x_1^2 + x_2^2 + \frac{\mu}{2}(x_1 - 1)^2 \quad \text{and} \quad \nabla Q(x;\mu) = \begin{bmatrix} (\mu - 10)x_1 - \mu \\ 2x_2 \end{bmatrix},$$

respectively. For $\mu = 1$, the stationary point is $(-1/9, 0)$ and the contours are shown in figure 4.

Problem 17.9

For Example 17.1, we know that $x^* = (-1,-1)$ and $\lambda^* = -\frac{1}{2}$. The goal is to show that $\phi_1(x;\mu)$ does not have a local minimizer at $(-1,-1)$ unless $\mu \geq \|\lambda^*\|_\infty = \frac{1}{2}$.

We have from the definition of the directional derivative that for any $p = (p_1, p_2)$,

$$\begin{aligned} D(\phi_1(x^*;\mu), p) &= \nabla f(x^*)^T p + \mu \sum_{i \in \mathcal{E}} |\nabla c_i(x^*)^T p| \\ &= (p_1 + p_2) + \mu|-2(p_1 + p_2)| \\ &= \begin{cases} (1 - 2\mu)(p_1 + p_2) & \text{if } p_1 + p_2 < 0 \\ (1 + 2\mu)(p_1 + p_2) & \text{if } p_1 + p_2 \geq 0. \end{cases} \end{aligned}$$

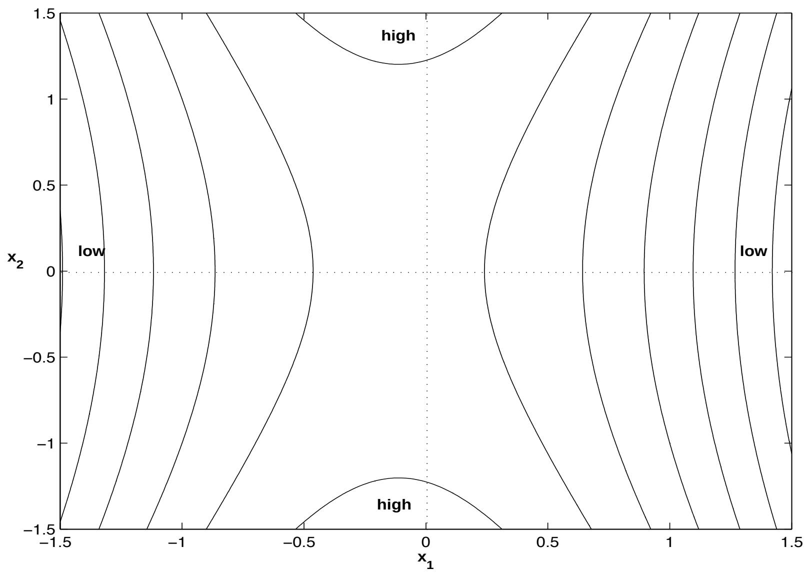

Figure 4: Contours for the quadratic penalty function $Q(x;\mu), \mu = 1$.

It is easily seen that when $\mu < \frac{1}{2}$, we can always choose $p_1 + p_2 < 0$ such that

$$(1 - 2\mu)(p_1 + p_2) < 0,$$

in which case p is a descent direction for $\phi_1(x^*;\mu)$. On the other hand, when $\mu \geq \frac{1}{2}$, there can be no descent directions for $\phi_1(x^*;\mu)$ since $D(\phi_1(x^*;\mu), p) \geq 0$ always holds. This shows that $\phi_1(x;\mu)$ does not have a local minimizer at $x^* = (-1,-1)$ unless $\mu \geq \|\lambda^*\|_\infty = \frac{1}{2}$.

18 Sequential Quadratic Programming

Problem 18.4

When $\theta_k \neq 1$, we have

$$\theta_k = \frac{0.8 s_k^T B_k s_k}{s_k^T B_k s_k - s_k^T y_k}$$

where $s_k^T y_k < 0.2 s_k^T B_k s_k$. Therefore

$$\begin{aligned}
s_k^T r_k &= s_k^T \left(\theta_k y_k + (1-\theta_k) B_k s_k\right) \\
&= \theta_k (s_k^T y_k) + (1-\theta_k) s_k^T B_k s_k \\
&= \frac{0.8 s_k^T B_k s_k}{s_k^T B_k s_k - s_k^T y_k} s_k^T y_k + \frac{0.2 s_k^T B_k s_k - s_k^T y_k}{s_k^T B_k s_k - s_k^T y_k} s_k^T B_k s_k \\
&= \frac{s_k^T B_k s_k}{s_k^T B_k s_k - s_k^T y_k} \left(0.8 s_k^T y_k + 0.2 s_k^T B_k s_k - s_k^T y_k\right) \\
&= \frac{s_k^T B_k s_k}{s_k^T B_k s_k - s_k^T y_k} \left(0.2 s_k^T B_k s_k - 0.2 s_k^T y_k\right) \\
&= 0.2 s_k^T B_k s_k \\
&> 0.
\end{aligned}$$

This shows that the damped BFGS updating satisfies (18.17).

Problem 18.5

We have

$$c(x) = x_1^2 + x_2^2 - 1 \quad \text{and} \quad \nabla c(x) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix},$$

so the linearized constraint at x_k is

$$\begin{aligned}
0 &= c(x_k) + \nabla c(x_k)^T p \\
&= x_1^2 + x_2^2 - 1 + 2x_1 p_1 + 2x_2 p_2.
\end{aligned}$$

(a) At $x_k = (0,0)$, the constraint becomes

$$0 = -1,$$

which is incompatible.

(b) At $x_k = (0,1)$, the constraint becomes

$$0 = 2p_2,$$

which has a solution of the form $p = (q, 0), q \in \Re$.

(c) At $x_k = (0.1, 0.02)$, the constraint becomes

$$0 = -0.9896 + 0.2p_1 + 0.04p_2,$$

which has a solution of the form $p = (4.948, 0) + q(-0.2, 1), q \in \Re$.

(d) At $x_k = -(0.1, 0.02)$, the constraint becomes

$$0 = -0.9896 - 0.2p_1 - 0.04p_2,$$

which has a solution of the form $p = -(4.948, 0) + q(-0.2, 1), q \in \Re$.

19 Interior-Point Methods for Nonlinear Programming

Problem 19.3

Define the vector function

$$c(x) = Dr(x),$$

where D is a diagonal scaling matrix with nonzero diagonal entries. The Jacobian corresponding to $c(x)$ is

$$A(x) = \begin{bmatrix} \nabla c_1(x)^T \\ \vdots \\ \nabla c_n(x)^T \end{bmatrix} = \begin{bmatrix} D_{11}\nabla r_1(x)^T \\ \vdots \\ D_{nn}\nabla r_n(x)^T \end{bmatrix} = DJ(x).$$

Therefore, the Newton step p is obtained via the solution of the linear system

$$DJ(x)p = -Dr(x),$$

which is equivalent to

$$J(x)p = -r(x)$$

since D is nonsingular.

Problem 19.4

Eliminating the linear equation yields $x_1 = 2 - x_2$. Plugging this expression into the second equation implies that the solutions satisfy

$$-3x_2^2 + 2x_2 + 1 = 0. \tag{69}$$

Thus, the solutions are

$$(x_1, x_2) \in \left\{ (1, 1), \left(\frac{7}{3}, \frac{1}{3} \right) \right\}.$$

Similarly, multiplying the first equation by x_2 yields the system

$$\begin{bmatrix} x_1x_2 + x_2^2 - 2x_2 \\ x_1x_2 - 2x_2^2 + 1 \end{bmatrix} = 0.$$

Subtracting the first equation from the second again yields (69), and the solutions remain unchanged.

Newton's method applied to the two systems yields the linear systems

$$\begin{bmatrix} 1 & 1 \\ x_2 & x_1 - 4x_2 \end{bmatrix} d = -\begin{bmatrix} x_1 + x_2 - 2 \\ x_1x_2 - 2x_2^2 + 1 \end{bmatrix}$$

and

$$\begin{bmatrix} x_2 & x_1 + 2x_2 - 2 \\ x_2 & x_1 - 4x_2 \end{bmatrix} d = -\begin{bmatrix} x_1x_2 + x_2^2 - 2x_2 \\ x_1x_2 - 2x_2^2 + 1 \end{bmatrix}.$$

From the point $x = (1, -1)$, the steps are found to be $d = (4/3, 2/3)$ and $d = (1/2, 1/2)$, respectively.

Problem 19.14

For clarity, define

$$U = \begin{bmatrix} W \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad V = \begin{bmatrix} WM^T \\ 0 \\ 0 \\ 0 \end{bmatrix}^T,$$

and

$$C = \begin{bmatrix} D & A^T \\ A & 0 \end{bmatrix},$$

where

$$D = \begin{bmatrix} \xi I & 0 \\ 0 & \Sigma \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} A_E & 0 \\ A_I & I \end{bmatrix}.$$

It can easily be shown that

$$C^{-1} = \begin{bmatrix} D^{-1} - D^{-1}A^T(AD^{-1}A^T)^{-1}AD^{-1} & D^{-1}A^T(AD^{-1}A^T)^{-1} \\ (AD^{-1}A^T)AD^{-1} & -(AD^{-1}A^T)^{-1} \end{bmatrix},$$

so the solution r of the primal-dual system $(C+UV^T)r = -s$ can be obtained via the Sherman–Morrison–Woodbury formula as

$$r = -(C + UV^T)^{-1}s = -\left(C^{-1} - C^{-1}U(I + V^TC^{-1}U)^{-1}V^TC^{-1}\right)s,$$

which requires only solutions of the system $Cv = b$ for various b.