Chapter 3
Robust Optimization for Environmental and Energy Planning

F. Babonneau, J.-P. Vial, and R. Apparigliato

Abstract Uncertainty is often present in environmental and energy economics. Traditional approaches to optimization under uncertainty, e.g., *stochastic programming*, *chance-constrained programming* or *stochastic dynamic programming*, encounter the most severe numerical difficulties because models in this area are large and complex, already in their deterministic formulation. The goal of the present chapter is to introduce a relatively new field, known as *robust optimization*, as an alternative to traditional methods and formulations. Through an illustrative example, we suggest ways of putting robust optimization at work in environmental and energy optimization models.

3.1 Robust Optimization in short

Uncertainty is often present in environmental and energy economics. As models in this area are often large and complex, introducing uncertainty with traditional approaches, e.g., *stochastic programming* (26; 25; 18; 33), *chance-constrained programming* (23; 22; 34; 36) or *stochastic dynamic programming* (13), generally leads to numerical intractable model as soon as a relevant representation of uncertainty is sought.

The goal of the present contribution is to introduce a relatively new field, known as *robust optimization* (37), which is an alternative to traditional methods and for-

F. Babonneau
ORDECSYS, 4 place de l'Etrier, 1224, Chêne-Bougeries, Suisse, e-mail: fbabonneau@ordecsys.com

J.-P. Vial
ORDECSYS, 4 place de l'Etrier, 1224, Chêne-Bougeries, Suisse, e-mail: jpvial@ordecsys.com

R. Apparigliato
GdF-Suez, Paris, France, e-mail: romain.apparigliato@gmail.com

J.A. Filar and A. Haurie (eds.), *Uncertainty and Environmental Decision Making*, International Series in Operations Research & Management Science 138,
DOI 10.1007/978-1-4419-1129-2_3,

mulations. The main feature of this approach is that it is does not resort to the calculus of probability, which makes it immune against the curse of dimensionality and computational intractability. In the meantime, robust optimization turns out to be a safe approximation of chance constrained programming (23). We shall illustrate our presentation through an example, a simplified version in power supply subject to constraints on admissible concentration of pollutants. This will provide a support to the presentation of the main ideas, to the methodology and to the value of the solutions.

Dealing with uncertainty raises formidable theoretical and practical modeling problems. On the other hand, almost all solution methods stumble on intractability issues. In the last decade, a new approach has emerged, which provides new ways to attack the problem. It is named Robust Optimization, and works on a new paradigm. To present it in few words only, we shall contrast it with traditional methods, such as stochastic programming, chance-constrained programming and/or dynamic programming. Roughly speaking, traditional methods posit the prerequisite of a well-defined probability model of the uncertainties involved in the problem; they next expand the mathematical programming model of the deterministic version of the problem to incorporate the uncertainty. This last operation usually goes with an increase of complexity that, most of the time, puts the computation of the solution out of reach of the current optimization methods. This phenomenon is sometimes described as the *curse of dimensionality* and/or *computational intractability* of the model[1]. In contrast, the primary concern of Robust Optimization is to overcome the complexity issue in adopting a non-probabilistic formulation of the uncertainty. The main underlying idea is to start with a much simplified, if not simplistic, description of uncertainty and look for solutions that remain satisfactory for all realizations of the uncertain parameters that are allowed by the uncertainty model. Solutions having this property are named *robust*. In Robust Optimization, no probability model is assigned to the uncertainty, which makes it possible to avoid expensive computations of multi-dimensional integrals associated with probabilities and expectations. Computing robust solution becomes a numerically tractable operation.

At first sight Robust Optimization can be assimilated to a worst case approach on a selected subset of possible realizations of the uncertainty parameters. This may give mixed feelings to people attached to the probabilistic aspect of the problem of interest. Of course, one may object that Robust Optimization has the definite advantage of avoiding the dramatic computational shortcomings of traditional approaches, in particular in multistage problems. Indeed, representing with a minimum of accuracy probabilities as well as computing expectations with multivariate distributions are formidable handicaps in the framework of optimization problems. But, recent results in Robust Optimization (5; 22; 23; 20) offer a more positive view. In few words, these results consist in a lower bound on the probability that the computed solution remains robust when the whole set of possible realizations—those in the uncertainty set and those outside of it— is considered. These strong results are obtained at the cost of rather mild assumptions on the probability model of uncertainty

[1] In very loose terms, a model is numerically intractable if no method can guarantee that a solution can be obtained in polynomial time.

(independence, a bounded range and an average value at the middle of the range). Probabilities, which are discarded in the initial model, reappear unexpectedly and give stronger confidence in Robust Optimization.

The first robust formulation for an optimization problem with uncertainty parameters has been proposed by Soyster (37) at the beginning of the seventies. The concept was taken over in the nineties by El-Gahoui and Lebret (28) and by Ben-Tal and Nemirovski (9). Since then, robust optimization has been intensively studied in the literature. The main theoretical contributions are (28; 10; 8; 4; 12; 15; 14).

Robust optimization is operational and useful for a large set of decision problems with uncertainty. One can mention decision problems in finance and risk measure (17; 19; 21; 30), in supply chain and inventory management (1; 6; 16), in telecommunications (35) and in management of electricity production in hydraulic valley (2; 3). This approach is also adapted in complement of optimization techniques such as constraints in probability (20), dynamic optimization (32), or stochastic optimization (24).

The chapter is organized as follows. In Section 2, we present our illustrative example and show that its deterministic solution performs poorly in an uncertainty context. Section 3 is concerned with an application of the robust optimization concepts developed in the next sections to the illustrative example with static uncertainty. In Section 4, we introduce those basic concepts of robust optimization in a static framework. In Section 5, we examine the dynamic case and propose the concept of Linear Decision Rules to cope with the adaptive nature of decisions. Section 6 is devoted to applications to the illustrative example, when the demand over time is the source of uncertainty. We compare the robust optimization solution with a stochastic programming approach. In Section 7 one explores a hybrid approach mixing stochastic programming and robust optimization. In Section 8 we show that bounds on the probability that the robust solution satisfies the constraint with uncertain coefficients can be obtained at the cost of a mild assumption on the probability distribution of the uncertain coefficients. This result shows an obvious relationship between chance constrained programming and robust optimization. Finally, in Section 9 we present an extension of the robust optimization concept that covers both the cases when the uncertainty lies in the uncertainty set and when it lies out of it. A short conclusion discusses possible issues in dealing with practical models in environmental and energy planning that are necessarily of much larger size than the illustrative example of this note. It also gives hints on new developments in the field of robust optimization.

3.2 An example in power supply under pollution constraint

To help the reader on the concepts of robust optimization that will be presented in this chapter, we illustrate them on a simple environmental and energy planning problem, inspired by the MARKA-Geneva model (27). The energy planning side concerns the simulation of the evolution of the energy system for a region. The model is based on the assumption of an optimal use of the resources and the simulated planning is obtained by minimizing a welfare function. The energy planning model is enriched to account for a cap on pollutant level in the various subregions. The pollutant level in a given subregion results from local emissions by energy producing sources and from transfers from other subregions. The model is dynamic with three periods.

In this model, we consider two types of uncertainties. One is related to the pollutant transfer rates from one subregion to another; the other one concerns the demand. Those two sources of uncertainties are of different natures. We model the demand as an auto-regressive process: a high demand at a time period is an indication that the demand will also be relatively high in the next time period. The decision process must make use of this information and be adaptive. The pollution effect does not enjoy this property. The transfer rates are not stable, but they do not follow any evolution pattern over time. The process is described by an i.i.d. process. In the rest of the chapter, we shall use the terminology region for subregion.

3.2.1 Deterministic formulation

The problem we consider here is the planning of energy production under environmental constraints. The model has four regions, three periods, and three technologies TEM. We denote $\mathscr{R} = \{r_1, r_2, r_3, r_4\}$ the set of regions, $\mathscr{T} = \{t_1, t_2, t_3\}$ the set of time periods and $\mathscr{P} = \{p_1, p_2, p_3\}$ the set of production technologies. The model assumes a single output with production target 10, 12 and 14 for the periods t_1, t_2 and t_3 respectively. The company is committed to serve the demand at all times. Nevertheless, delivery failures may occur if the demand is higher than the production capacity, but the shortage costs, presumably very high, are not known. Thus violations on the demands should be avoided by all means.

Three technologies can be used to produce the output. The admissible production level for each technology depends on the installed capacity. This installed capacity depends on installations performed prior to the planning horizon, thereafter named *residual capacity*, and on investments in the successive periods. These technologies can be installed in four different regions. The residual capacities are given in Table 3.2.

The total system cost is the sum over the three periods and the four regions of the capacity investment, maintenance and operation costs for the three technologies. The costs are reported in Table 3.1.

Table 3.1 Maintenance, investment and operation costs for the three technologies.

Cost	r_1	r_2	r_3	r_4
Maintenance (M_r)	1	1	1	1
Investment (I_r)	5	3	4	5
Operation (O_r)	2	2	3	2

Table 3.2 Residual capacities ($res_{p,t,r}$).

	Technology p_1			Technology p_2			Technology p_3		
	t_1	t_2	t_3	t_1	t_2	t_3	t_1	t_2	t_3
r_1	1	1	-	0.5	-	-	-	-	-
r_2	-	-	-	-	-	-	-	-	-
r_3	-	-	-	-	-	-	-	-	-
r_4	-	-	-	-	-	-	1	1	-

The technologies have different pollutant emission rates, given in Table 3.3, in each region. A pollution transport and dispersion process takes place. A source-receptor matrix specifies, for each source location, the proportion of the emitted pollutant that is deposited in the different receptor locations (see Table 3.4). These data are used to determine the level of concentration (immission) in each region. A standard environmental quality, $Q = 1.5$, is imposed on this concentration for all regions and for all periods.

Table 3.3 Emission rates ($E_{p,r}$) by production technology and by region.

	r_1	r_2	r_3	r_4
p_1	0.7	0.8	0.8	0.6
p_2	0.5	0.4	0.7	0.7
p_3	0.8	0.9	0.6	0.7

Table 3.4 Source-receptor transfer matrix (G_{r_i,r_j}).

	r_1	r_2	r_3	r_4
r_1	0.5	0.1	0.1	0.05
r_2	0.1	0.4	0.04	0.1
r_3	0.09	0.05	0.5	0.1
r_4	0.05	0.1	0.1	0.6

The model gives conditions under which a joint investment and production plan meets the demand needs and the environmental standards at minimal cost.

For the model, we define the following variables:

- $x_{p,t,r}$: production of technology p, at period t and in region r.
- $y_{p,t,r}$: capacity investment of technology p, at period t and in region r.
- $z_{p,t,r}$: installed capacity of technology p, at period t and in region r.
- $e_{t,r}$: emission at period t and in region r.

The optimization problem is given by

$$\min_{x\geq 0, y\geq 0, z, e} \sum_r (O_r \sum_{p,t} x_{p,t,r} + I_r \sum_{p,t} y_{p,t,r} + M_r \sum_{p,t} z_{p,t,r}) \tag{1a}$$

$$z_{p,t,r} = res_{p,t,r} + \sum_{\tau \leq t} y_{p,\tau,r} \quad \forall t \in \mathscr{T}, \forall p \in \mathscr{P}, \forall r \in \mathscr{R} \tag{1b}$$

$$x_{p,t,r} \leq z_{p,t,r} \quad \forall t \in \mathscr{T}, \forall p \in \mathscr{P}, \forall r \in \mathscr{R} \tag{1c}$$

$$\sum_{p,r} x_{p,t,r} = d_t \quad \forall t \in \mathscr{T} \tag{1d}$$

$$\sum_{p} E_{p,r} x_{p,t,r} = e_{t,r} \quad \forall t \in \mathscr{T}, \forall r \in \mathscr{R} \tag{1e}$$

$$\sum_{\rho \in \mathscr{R}} e_{t,\rho} G_{\rho,r} \leq Q \quad \forall t \in \mathscr{T}, \forall r \in \mathscr{R}. \tag{1f}$$

In that formulation, the equality constraints (1b) express the total installed capacities, $z_{p,t,r}$, as the sum of the residual capacities and the capacities that are installed in the previous periods. Inequality constraints (1c) limit production to the installed capacities. Equations (1d) are demand constraints. The last two sets of constraints firstly compute the pollutant emissions resulting from all technologies (1e), and secondly impose bounds on the final emissions resulting from the dispersion process (1f).

Problem (3.1) has 108 variables and 99 constraints. Note that (3.1) can be formulated in a more compact way by removing the z and e variables and the equality constraints (1b) and (1e) (see Subsection 3.6 for more details). Solving problem (3.1) gives the optimal objective value 162.056. The optimal investments, productions and emissions are reported in Table 3.5, 3.6 and 3.7, respectively.

Table 3.5 Optimal investment schedule.

	t_1			t_2			t_3			
	p_1	p_2	p_3	p_1	p_2	p_3	p_1	p_2	p_3	Total
r_1	-	1.41	-	-	0.77	-	-	-	-	2.18
r_2	-	6.09	-	-	1.73	-	-	0.15	-	7.97
r_3	-	-	-	-	-	-	-	-	3.85	3.85
r_4	-	-	-	-	-	-	-	-	-	0
Sum	-	7.50	-	-	2.50	-	-	0.15	3.85	
Total per period	7.50			2.50			4.00			

We observe in Table 3.5 and Table 3.6 that the optimal solution is to invest in technology p_2 only, except in period t_3 where technology p_3 is also selected. As far as production is concerned, the three technologies are used in the first two periods, but the first production technology is abandoned in the last period. Table 3.7 shows that only two emission constraints are active at the optimum (in region r_2 at periods t_2 and t_3), and one is close to be active (in region r_3 at period t_3). Should the optimal investment/production plan be implemented within an uncertain environment, then,

Table 3.6 Optimal production schedule.

	t_1			t_2			t_3			
	p_1	p_2	p_3	p_1	p_2	p_3	p_1	p_2	p_3	Total
r_1	1.00	1.91	-	1.00	2.18	-	-	2.18	-	8.27
r_2	-	6.09	-	-	7.82	-	-	7.97	-	21.88
r_3	-	-	-	-	-	-	-	-	3.85	3.85
r_4	-	-	1.00	-	-	1.00	-	-	-	2.00
Sum	1.00	8.00	1.00	1.00	10.00	1.00	-	10.15	3.85	
Total per period		10.00			12.00			14.00		

Table 3.7 Emission associated with the optimal solution.

	t_1	t_2	t_3
r_1	1.16	1.24	1.07
r_2	1.21	**1.50**	**1.50**
r_3	0.33	0.37	1.39
r_4	0.75	0.82	0.60

those constraints would potentially be critical and one should expect violations for some variations in the entries of the source-receptor transfer matrix.

3.2.2 Uncertainties in the power supply model

In the power supply model, we consider two sources of uncertainties, one on the pollutant diffusion coefficients and the second one on the demand.

First we define the uncertainty on the coefficients of the source-receptor matrix G. Let ξ be a random variable with values in the interval $[-1,1]$, each coefficient of the matrix G is given by

$$G_{i,j}(\xi) = \bar{G}_{i,j} + \hat{G}_{i,j}\xi_{i,j}, \tag{3.2}$$

where $\bar{G}_{i,j}$ is the average coefficient reported in Table 3.4 and $\hat{G}_{i,j}$ corresponds to the coefficient variability. Here we set $\hat{G}_{i,j} = 0.1\bar{G}_{i,j}$.

We now focus on the demand uncertainty. We adopt the following autoregressive model

$$d_{t+1} = d_t + \alpha_{t+1} + \hat{d}_{t+1}\eta_{t+1},$$

where η_{t+1} is a random variable with values in the interval $[-1,1]$ and α_{t+1} is a deterministic trend factor. This formula can be explicited using backward substitution. Each demand d_t appears then as a function of the past random factor η_τ for $\tau = 1,\dots,t$. In our case with horizon of length 3, we have

$$d_1 = \bar{d}_1 + \hat{d}_1\eta_1 \tag{3a}$$
$$d_2 = \bar{d}_2 + \hat{d}_1\eta_1 + \hat{d}_2\eta_2 \tag{3b}$$
$$d_3 = \bar{d}_3 + \hat{d}_1\eta_1 + \hat{d}_2\eta_2 + \hat{d}_3\eta_3 \tag{3c}$$

where $\bar{d}_i$ and $\hat{d}_i$ are the average demand and the variability of the demand, respectively. We have $\bar{d}_1 = 10$, $\bar{d}_2 = 12$, $\bar{d}_3 = 14$ and $\hat{d}_i = 0.1\bar{d}_i$, $\forall i$.

3.2.3 Validation process

The validation process is an empirical study of the behavior of the solution under a set of simulated values for the uncertain parameters. In our study, an item of the simulation process is a set of independent realizations of the random factors (demand and/or emission transfer parameters) over the three periods. We shall use the terminology "scenario" to name one such set of realizations. To generate a scenario, one needs a probabilistic model of the random factors. In our illustrative example, we choose to have the underlying factor to be i.i.d. with a uniform distribution on the range $[-1,1]$. The validation is performed on a sample of scenarios of size 1000; each scenario in the sample is a multidimensional vector of realizations of the random factors η and ξ. The demands and the pollutant diffusion coefficients are computed via (3.3) and (3.2), respectively.

In the evaluation of performance, we must differentiate among the two types of uncertainties. Let us start with the emission transfer coefficients. Since there is no recourse associated with this type of uncertainty, the investment/production can be implemented as such, without a risk of violating the demand constraint. Of course, the constraint on the air quality may not be satisfied, but this is just recorded without any modification of the solution itself.

The case of an uncertain demand raises a new issue. One could just record the violations of the demand constraint from above or from below as argued in the previous case, but this would not be realistic. Since the production is adapted to the manifested demand, it is difficult to stick to the view that the production should be maintained at its scheduled value if the demand turns out to be smaller than the production. In the validation process we adopt the following strategy. Let $x_{p,t,r}$ and $y_{p,t,r}$ be the production and the investment in a particular solution. If the total planned production $\sum_{p,r} x_{p,t,r}$ at period t is less than the demand d_t, the production is kept as such, and we record a shortage. The ensuing constraint violation is measured in relative value $(\max\{d_t - \sum_{p,r} x_{p,t,r}, 0\})/d_t$. If the demand is less than the planned production, production in each region with each technology is uniformly downsized by the common factor $d_t / \sum_{p,r} x_{p,t,r}$. The treatment of the constraint on the pollutant level in each region is simpler. It measures the amount of relative violation.

3.2.4 Evaluation of the deterministic solution in the uncertain environment

We now subject the deterministic optimal solution computed in the previous section (see Tables 3.5, 3.6 and 3.7) to the simulation process.

Impact of uncertainty on pollutant diffusion

Table 3.8 gives the simulation results with uncertainty on the pollutant diffusion coefficients. We observe that for only 46.5% of the simulations all the air quality constraints are satisfied. For 4.2%, 47.8%, and 1.5% of the cases, one, two and three air quality constraints are violated, respectively. The average and maximum relative violations are about 4.3% and 9.3%, respectively. An analysis of the constraint violations has revealed that the phenomenon occurred in the three pairs (region and periods) that were critical or near critical in the deterministic study. In the other pairs there is enough slack to absorb variations in the transfer rates.

Table 3.8 Simulation results with uncertain source-receptor matrix.

Average relative violation (%)	4.3
Maximum relative violation (%)	9.3
% of satisfaction	46.5
% of one violation	4.2
% of two violations	47.8
% of three violations	1.5

Impact of uncertainty on the demand

In Table 3.9, we report the simulation results conducted on the optimal deterministic solution with a demand uncertainty.

Table 3.9 Simulation results with uncertain demands.

Predicted cost performance	162.056
Observed cost performance	160.871
Scenarios with demand violation(s) in %	62.0
Conditional average relative violation in %	2.5
Average number of violations per scenario	2.0

Contrary to the previous experiment, the uncertainty on the demand has an impact on the cost performance. As a result we distinguish between the predicted performance (the optimal cost in the deterministic model) and the observed performance (the average cost in the simulation). Surprisingly enough, the observed performance is better than the predicted one. The paradox is apparent and the explanation is straightforward. When the demand is lower than the average value used in the deterministic model, the production levels are scaled down to match the actual demand and the production costs less than predicted by the deterministic plan. But when the demand is higher than this average, the production cannot be increased, because the capacity constraints in the deterministic optimal solutions are tight; the production costs remain as computed in the deterministic model, but part of the demand is not satisfied. Thus, the seemingly improved cost performance must be balanced with demand violations. In that respect, we observe a demand violation in 62% of the scenarios and a conditional average of unsatisfied demand of 2.5%. Finally, when a violation occurs in a scenario the average number of periods with violation is approximately 2.0 (over the total of 3). The conditional average demand violation is small with respect to the average demand, but one should recall that the range of variation of the demand around the average is ±5%, ±10% and ±15%, in periods 1, 2 and 3, respectively.

3.3 Case study: robust solution to a problem with uncertain pollutant transfer coefficients

To motivate robust optimization, we propose a simple fix to handle uncertainties in the power supply problem. We limit this illustration to uncertainties on the pollutant diffusion coefficients, and consider that the demands are fixed. We shall use heuristic arguments only and postpone theoretical justifications to later sections.

Let us start with a general formulation of a linear constraint with uncertain coefficients

$$\sum_{j=1}^{n} \tilde{a}_j x_j \leq b, \tag{3.4}$$

where $\tilde{a}_j$ are uncertain. For the sake of a simpler presentation, we assume b to be certain. We further describe the uncertain coefficients as linear functions of an underlying random factor ξ

$$\tilde{a} = \bar{a} + P\xi,$$

where $\xi \in \mathbb{R}^m$ and P is an $n \times m$ matrix. We further assume that the random factor has a symmetric distribution with mean 0. The certain vector $\bar{a}$ is usually named the normal factor. We can thus focus on the uncertain component of the constraint

$$\underbrace{\bar{a}^T x}_{\text{certain}} + \underbrace{(P^T x)^T \xi}_{\text{uncertain}} \leq b.$$

We now evoke a common sense engineering approach that consists in replacing the uncertain term by a safety term $\kappa > 0$. By taking a large enough safety term, we give a sufficient guarantee that the solutions of $\bar{a}^T x + \kappa \leq b$ will almost always remain feasible to (3.4). This says nothing on the critical way to choose the safety factor, but many practitioners will be receptive to a so-called 2σ, 3σ, possibly 6σ approach. This industrial practice is often justified through the following probabilistic and statistical argument. Assuming that the underlying random factor ξ is a vector with probabilistically independent components with mean 0 and standard deviation σ, we can state that the uncertain component $(P^T x)^T \xi$ is a random variable with mean 0 and standard deviation $||P^T x||_2$. In the $k \times \sigma$ approach, the safety term is $\kappa = k||P^T x||_2$, with the implicit argument that for most distributions a large enough k, say $k = 3$, will guarantee that a solution to $\bar{a}^T x + k||P^T x||_2 \leq b$ will satisfy the probabilistic constraint $\sum_{j=1}^n \tilde{a}_j x_j \leq b$ with high probability.

The two salient features of this "engineer-like" approach can be summarized as follows. An uncertain constraint is replaced by a deterministic constraint with a safety term, and the safety term depends on the decision variables x. Applying this idea to our problem of interest, we obtain that the robust equivalent of the air quality constraints (1f) is

$$\sum_{\rho \in \mathscr{R}} e_{t,\rho} \bar{G}_{r,\rho} + k|| \sum_{\rho \in \mathscr{R}} e_{t,\rho} \hat{G}_{r,\rho} ||_2 \leq Q \quad \forall t \in \mathscr{T}, r \in \mathscr{R}. \tag{3.5}$$

This formulation does not increase the problem size, but the linear air quality constraints are replaced by nonlinear ones. Those constraints are convex conic quadratic, and problems with such constraints can be solved very efficiently by modern solvers (e.g., the open source code (38)).

If we wish to remain in the realm of linear programming, we have to replace the safety factor by terms that are amenable to linear inequalities. The following bound will be established in the next section

$$||\alpha||_2 \leq \min_{\beta} \{\sqrt{m}\, ||\alpha - \beta||_\infty + ||\beta||_1\},$$

where $||\alpha||_1 = \sum_{i=1}^m |\alpha_i|$ and $||\alpha||_\infty = \max_i |\alpha_i|$. Hence, the alternative, linear but more restrictive formulation,

$$\sum_{\rho \in \mathscr{R}} e_{p,\rho} \bar{G}_{r,\rho} + k\sqrt{m}\, || \sum_{\rho \in \mathscr{R}} e_{p,\rho} \hat{G}_{r,\rho} - \beta ||_\infty + k||\beta||_1 \leq Q \quad \forall p \in \mathscr{P}, r \in \mathscr{R}. \tag{3.6}$$

In this experiment, we solve the robust equivalent problem with the two formulations, the first one (3.5) with the conic quadratic constraint, and the other one (3.6) with linear constraints. We use the same k in the two formulations, but make this k vary to experiment with different degrees of safety. We shall compare the behavior of the robust solutions with the deterministic solution. The results are reported in Table 3.10.

In the second set of experiments we use the formulation with linear constraints (3.6). We shall prove in the next section that (3.6) is equivalent to a system of linear inequalities. It turns out that the robust formulation of a typical air quality constraint (1f) corresponding to the pair $(t \in \mathscr{T}, r \in \mathscr{R})$ is

Table 3.10 Variable transfer coefficients: behavior on the sample of 1000 scenarios of a robust solution with the conic quadratic formulation (3.5).

	Deterministic	Robust		
		$k = 0.8$	$k = 1$	$k = 1.2$
Cost performance	162.06	163.76	164.21	164.64
Conditional average of relative violation (%)	4.3	1.0	0.5	-
Maxi relative violation (%)	9.3	3.0	1.5	-
proportion of scenarios with violation (%)	53.5	25.4	6	0
proportion of scenarios with one violation (%)	4.2	17.2	5	-
proportion of scenarios with two violations (%)	47.8	6.4	1	-
proportion of scenarios with three violations (%)	1.5	1.8	-	-

$$\sum_{i \in \mathscr{R}} e_{t,i} \bar{G}_{r,i} + k(\sum_{i \in \mathscr{R}} u_{t,r,i} + 2v_{t,r}) \leq Q \tag{7a}$$

$$u_{t,r,\rho'} + v_{t,r} \geq \sum_{\rho \in \mathscr{R}} e_{t,\rho} \hat{G}_{\rho',\rho} \quad \forall \rho' \in \mathscr{R} \tag{7b}$$

$$u \geq 0, v \geq 0 \tag{7c}$$

The above robust counterpart is obtained from the set of inequalities (3.10) in Proposition 1 (Section 3.4). The reader will notice that (3.10) also includes constraints like (7b), but with right-hand side $-\sum_{\rho \in \mathscr{R}} e_{t,\rho} \hat{G}_{\rho',\rho}$. In this particular problem, we know that $\sum_{\rho \in \mathscr{R}} e_{t,\rho} \hat{G}_{\rho',\rho} \geq 0$, which makes one half of the constraints redundant. So, we eliminate them and obtain a robust equivalent model that is still linear but with 60 additional variables u's and v's and 48 additional constraints.

We solve the robust equivalent problem with different values of k and we report the results of the simulations for each robust optimal solution in Table 3.11.

Table 3.11 Variable transfer coefficients: behavior on the sample of 1000 scenarios of a robust solution with respect to the linear formulation (3.7).

	Deterministic	Robust		
		$k = 0.6$	$k = 0.8$	$k = 1$
Cost performance	162.06	163.49	163.99	164.49
Conditional average of relative violation (%)	4.3	1.2	0.7	-
Max relative violation (%)	9.3	3.5	1.7	-
proportion of scenarios with violation (%)	53.5	29.1	12.4	0
proportion of scenarios with one violation (%)	4.2	14.9	9.4	-
proportion of scenarios with two violations (%)	47.8	12.7	2.9	-
proportion of scenarios with three violations	1.5	1.5	0.1	-

The results of Table 3.10 and Table 3.11 are very much alike. Nevertheless, we notice that with the same safety level k, say $k = 0.8$ for a typical example, the robust solution with respect to the linear approximation (3.6) achieves a slightly worst cost objective (163.99 vs. 163.76) than the solution with respect to the conic quadratic formulation (3.5). In the meantime, the first solution achieves a better protection

Table 3.12 Robust investment schedule for $k = 0.8$ and the linear formulation (3.7).

	t_1			t_2			t_3			
	p_1	p_2	p_3	p_1	p_2	p_3	p_1	p_2	p_3	Total
r_1	-	1.39	-	-	1.51	-	-	0.10	-	3.00
r_2	-	5.94	-	-	0.91	-	-	0.16	-	7.01
r_3	-	-	-	-	-	-	-	-	3.74	3.74
r_4	0.17	-	-	0.08	-	-	-	-	-	0.25
Sum	0.17	7.33	-	0.08	2.42	-	-	0.26	3.74	
Total per period		7.50			2.50			4.00		

against constraint violation. These observations conform to the fact that for identical k the linear safety factor in (3.6) in Table 3.11 is an upper bound of the conic quadratic safety factor (3.5) in Table 3.10. The first one is more constrained is thus more constrained. The cost performance in not so good, but it ensures a lesser protection against violations. The chances are that in a simulation we observe a lesser number of constraint violations.

3.4 Robust Optimization for the static problem: theoretical developments

Let us recall the formulation of a linear constraint with uncertain coefficients

$$\sum_{j=1}^{n} \tilde{a}_j x_j \leq b.$$

The first basic assumption on the uncertain parameters is that they depend on some random factor ξ in a linear way.

Assumption 1 *The uncertain vector $\tilde{a}$ is written as*

$$\tilde{a} = \bar{a} + P\xi,$$

where $\xi \in \mathbb{R}^m$ and P is an $n \times m$ matrix.

The certain vector $\bar{a}$ is usually named the normal factor. We can thus focus on the uncertain component of the constraint

$$\underbrace{\bar{a}^T x}_{\text{certain}} + \underbrace{(P^T x)^T \xi}_{\text{uncertain}} \leq b. \tag{3.8}$$

Our present goal is to use information on ξ to build the safety factor introduced in the previous section (see inequalities (3.5) and (3.6)). The idea is to focus on a subset of all possible events that it is made of all realizations of the underlying uncertain factor ξ that the modeler deems necessary to protect against. This is the so-called *uncertainty set*. The robust version of the initial uncertain constraint $\tilde{a}^T x \le b$ consists in enforcing the uncertain constraint (3.8), not for *all* possible realizations, but only by those in the uncertainty set; that is, the less restrictive constraint

$$\bar{a}^T x + (P^T x)^T \xi \le b, \text{ for all } \xi \in \Xi,$$

where $\Xi \subset \mathbb{R}^m$ is the uncertainty set. A solution to this constraint is called *robust* with respect to Ξ. If Ξ is a continuous set, the robust constraint is a short-hand writing of an infinite number of simple linear constraints. This seems to put the whole approach into the realm of semi-infinite programming, making the computation of solutions a real issue. It turns out that this is not so, for a large variety of uncertainty sets.

We shall consider a few different types of uncertainty sets. Let us start with the ellipsoidal uncertainty set

$$\Xi = \{\xi \mid ||\xi||_2 \le k\}.$$

In the rest of the chapter, we shall describe the ellipsoidal uncertainty set as the ball in the 2-norm, centered at the origin and with radius k. We denote it $B_2(0,k)$. We shall also use more general balls, such as

$$B_p(0,k) = \{\xi \mid ||\xi||_p \le k\}$$

with $1 < p < \infty$ and $||\xi||_p = (\sum_{i=1}^m |\xi_i|^p)^{1/p}$. The following lemma gives the clue to the replacement of the robust constraint by a single finite-dimensional constraint which will be named the *robust equivalent*.

Lemma 1. *Let* $1 < p < \infty$ *and let* q *be such that* $\frac{1}{p} + \frac{1}{q} = 1$. *For any* d

$$\max\{d^T \xi \mid ||\xi||_p \le k\} = k||d||_q.$$

Proof. The proof in the general case is simple but tedious: we omit it. We just illustrate the case of the 2-norm. (See Figure 3.1.) □

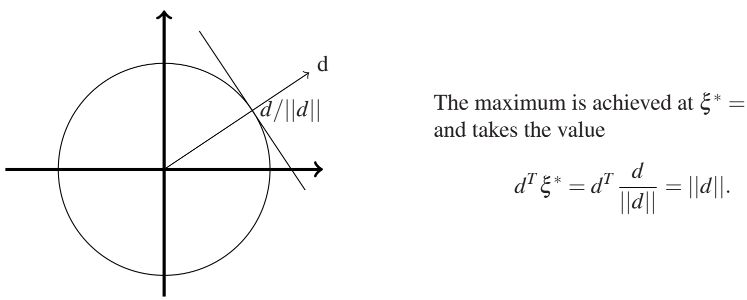

Fig. 3.1 Maximum of a linear form over a ball in the 2-norm.

3.4.1 Robust equivalent: the case of the 2-norm

We can now state the robust equivalent of the robust constraint.

Theorem 1. *The robust equivalent of the constraint*

$$\bar{a}^T x + (P^T x)^T \xi \leq b, \text{ for all } \xi \in \Xi = \{\xi \mid ||\xi||_2 \leq k\},$$

is

$$\bar{a}^T x + k||P^T x||_2 \leq b.$$

Proof. The proof follows directly from the above lemma. It suffices to replace d by $P^T x$. □

The factor k plays a crucial role in Theorem 1. The larger its value, the greater the number of realizations ξ against which a solution of $\bar{a}^T x + k||P^T x||_2 \leq b$ is *immunized* in the constraint $\bar{a}^T x + (P^T x)^T \xi \leq b$ with uncertain coefficients . In the sequel we shall use the terminology *immunization factor k*, or *immunization level k*.

At this stage we should raise some fundamental issues:

1. Is the nonlinear formulation of the robust equivalent a potential source of complexity from a numerical point of view?
2. Should, or could, one consider alternative uncertainty sets?
3. What is the value of starting with the new concept of uncertainty set if one ends up with the same formulation of an engineering safety factor?

We shall answer the first two questions, leaving the answer of the last question to the final section of the chapter. Let us start with the first question. It is easy to show that the new constraint is convex (a property of the 2-norm). Moreover, it can be reformulated as

$$\bar{a}^T x + kz \leq b \tag{9a}$$

$$P^T x = u \tag{9b}$$

$$||u||_2 \leq z. \tag{9c}$$

The last constraint is conic quadratic, a feature that modern convex optimization codes handle about as efficiently as a linear constraint.

Taking the uncertainty set as the primary concept shifts the focus on numerically tractable robust equivalent. However, people concerned with a probabilistic approach may feel ill at ease with the apparent arbitrariness of the uncertainty set. However, there is a powerful theorem in probability theory that assets that under a very mild probabilistic assumption on the random factor ξ, one can provide a surprisingly strong lower bound on the probability that the initial uncertain constraint be satisfied by a robust solution (i.e., a solution to the robust equivalent).

The answer to the second question motivates the next section.

3.4.2 Robust equivalent: the case of the ℓ_1 and ℓ_∞ norms

For some reasons, e.g., a lack of access to a conic quadratic solver, one may want to remain in the realm of linear programming. This can be achieved by resorting to polyhedral uncertainty sets. To link the results with the previous study with ellipsoidal uncertainty sets, we resort to an approximation of the unit ball in the 2-norm by the intersection of balls in the 1-norm and the infinity norm.

The ℓ_1 and ℓ_∞ norms are natural extensions of the p-norm with $p=1$ and $p=\infty$. It generates the balls

$$B_1(0,k) = \{\xi \mid \sum_{i=1}^{m} |\xi_i| \leq k\}$$

and

$$B_\infty(0,k) = \{\xi \mid \max_{i=1,\ldots,m} |\xi_i| \leq k\}.$$

Both are polyhedral sets that can be represented by simple inequalities. The next lemma extends Lemma 1 to the two limit cases $p=\infty$ and $p=1$.

Lemma 2. *We have*

$$\max\{d^T\xi \mid ||\xi||_\infty \leq k\} = k||d||_1$$

and

$$\max\{d^T\xi \mid ||\xi||_1 \leq k\} = k||d||_\infty.$$

The $B_1(0,k\sqrt{m})$ and $B_\infty(0,k)$ balls can be jointly used to approximate the $B_2(0,k)$ ball in the m-dimensional space as it is illustrated in Figure 3.2. We can now give

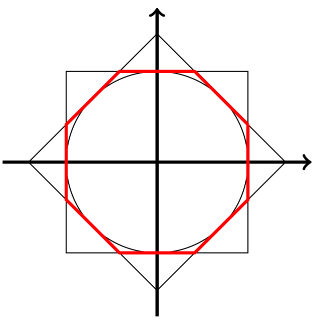

Fig. 3.2 The set $B_1(0,\sqrt{2}) \cap B_\infty(0,1)$

the robust equivalent relative to the uncertainty set $\Xi = B_1(0,k_1) \cap B_\infty(0,k_\infty)$.

Lemma 3. *We have*

$$\max_{\xi}\{d^T\xi \mid \xi \in B_1(0,k_1) \cap B_\infty(0,k_\infty)\} = \min_{w}\{k_1||d-w||_\infty + k_\infty||w||_1\}.$$

Proof. Let

$$z^* = \max_{\xi}\{d^T\xi \mid \xi \in B_1(0,k_1) \cap B_\infty(0,k_\infty)\}.$$

Let us replace $B_\infty(0,k_\infty)$ by linear inequalities in the above maximization problem. We obtain the alternative expression

$$\begin{aligned} \max_{\xi} \quad & d^T\xi \\ & -k_\infty e \leq \xi \leq k_\infty e \\ & \xi \in B_1(0,k_1), \end{aligned}$$

where e is the vector of all ones of appropriate dimension. We form the partial Lagrangian

$$L(\xi,u,v) = d^T\xi + u^T(\xi + k_\infty e) + v^T(-\xi + k_\infty e) = k_\infty e^T(u+v) + (d+u-v)^T\xi.$$

Clearly

$$z^* = \max_{\xi\in B_1(0,k_1)} \min_{u\geq 0, v\geq 0} L(\xi,u,v)$$

and by linear programming duality

$$z^* = \min_{u\geq 0, v\geq 0} \max_{\xi\in B_1(0,k_1)} L(\xi,u,v).$$

The inner minimization problem is

$$\mathscr{L}(u,v) = \max_{\xi}\{k_\infty e^T(u+v) + (d+u-v)^T\xi \mid \xi \in B_1(0,k_1)\}$$

By Lemma 2 we have

$$\mathscr{L}(u,v) = k_\infty e^T(u+v) + k_1||d+u-v||_\infty.$$

One easily checks that an optimal solution of the minimization $\min_{u\geq 0, v\geq 0}\mathscr{L}(u,v)$ satisfies the complementary condition $u_i v_i = 0$ for $i = 1,\ldots,m$. Hence u and v can be viewed as the positive and negative parts of a real vector $w = u - v$. Therefore $e^T(u+v) = ||w||_1$ and one can write

$$z^* = \min_{w}\{k_1||d-w||_\infty + k_\infty||w||_1\}.$$

□

We can now state the robust equivalent for the new polyhedral uncertainty set.

Theorem 2. *The robust equivalent of the robust constraint*

$$\bar{a}^T x + (P^T x)^T\xi \leq b, \textit{ for all } \xi \in \Xi = \{\xi \mid B_1(0,k_1) \cap B_\infty(0,k_\infty)\},$$

is the constraint in x **and** w

$$\bar{a}^T x + k_1 ||P^T x - w||_\infty + k_\infty ||w||_1 \leq b.$$

Proof. Using Lemma 3, we may write the robust equivalent as $\bar{a}^T x + z^* \leq b$. Since for any w one has $z^* \leq k_1 ||P^T x - w||_\infty + k_\infty ||w||_1$, we have that the new inequality implies the robust inequality. By the strong linear duality theorem, the equivalence is achieved because there always exists a w^* such that $z^* = k_1 ||P^T x - w^*||_\infty + k_\infty ||w^*||_1$. □

Remark 1. The ball $B_1(0, k_1)$ in Theorem 2 can be replaced by a ball $B_p(0, k_p)$ in the norm ℓ_p, for any $1 < p < \infty$. One obtains the following robust equivalent

$$\bar{a}^T x + k_p ||P^T x - w||_q + k_\infty ||w||_1 \leq b$$

with q such that $1/p + 1/q = 1$.

For implementation purposes, it is convenient to replace the norm expressions into linear inequalities.

Proposition 1. *The robust counterpart*

$$\bar{a}^T x + k_1 ||P^T x - w||_\infty + k_\infty ||w||_1 \leq b$$

has the same set of solutions as the system of linear inequalities

$$\bar{a}^T x + k_1 t + k_\infty e^T w \leq b \tag{10a}$$
$$w + te \geq P^T x \tag{10b}$$
$$w + te \geq -P^T x \tag{10c}$$
$$w \geq 0, \; t \geq 0.$$

Proof. Let t be a scalar variable and write the minimization problem yielding $z^* = \min_w k_1 ||P^T x - w||_\infty + k_\infty ||w||_1$ as the linear program

$$\begin{aligned} \min \quad & k_1 t + k_\infty e^T (u + v) \leq b \\ & -te \leq P^T x + u - v \leq te \\ & u \geq 0, \; v \geq 0, \; t \geq 0. \end{aligned}$$

This program can be simplified as follows. Consider first the case $(P^T x)_i > 0$. The minimization operation entails $u_i = 0$, and $v_i = \max\{(P^T x)_i - t, 0\}$. A similar reasoning yields $v_i = 0$ and $u_i = \max\{(-P^T x)_i - t, 0\}$, when $(P^T x)_i \leq 0$. Define $w \geq 0$ by $w_i = v_i$ if $(P^T x)_i > 0$ and $w_i = u_i$ if $(P^T x)_i \leq 0$. By construction, w is such that $w + te \geq P^T x$ and $w + te \geq -P^T x$.

Conversely, a solution (w, t) to (3.10) generates a solution (u, v, t) of the initial system. Indeed, set $u_i = 0$ and $v_i = w_i$ if $(P^T x)_i > 0$ and $u_i = w_i$ and $v_i = 0$ if $(P^T x)_i \leq 0$. We have thus that the minimization problem is equivalent to

$$\begin{aligned} \min \quad & k_1 t + k_\infty e^T w \\ & w + te \geq P^T x \\ & w + te \geq -P^T x \\ & w \geq 0, \; t \geq 0. \end{aligned}$$

This concludes the proof. $\square$

3.4.3 Robust equivalent: bounded polyhedral uncertainty set

The balls in ℓ_1 and ℓ_∞ are special cases of polyhedra. We extend the construction of the robust counterpart to the case of uncertainty sets defined by a bounded nonempty polyhedral set defined by $\{\xi \mid Q\xi \leq t\}$. In that case the robustness test for the constraint $(\bar{a}+P\xi)^T x \leq b$, $\forall \xi \in \Xi$ is given by

$$\max_{\xi}\{(P^T x)^T \xi \mid Q\xi \leq t\} \leq -\bar{a}^T x + b.$$

Proposition 2. *The condition*

$$\max_{\xi}\{(P^T x)^T \xi \mid Q\xi \leq t\} \leq -\bar{a}^T x + b$$

is satisfied if and only if there exists a solution (x,u) *for the linear system*

$$\begin{aligned} \bar{a}^T x + t^T u &\leq b \\ P^T x - Q^T u &= 0 \\ u &\geq 0. \end{aligned}$$

Proof. In the proof, x is a fixed parameter. A direct derivation is obtained using the theory of duality in linear programming. To define the worst case in the robust constraint, we consider the linear programming problem

$$\max_{\xi}\{(P^T x)^T \xi \mid Q\xi \leq t\}.$$

This problem is feasible and bounded. It has an optimal solution and its dual

$$\min_{u}\{t^T u \mid Q^T u - P^T x = 0,\ u \geq 0\}$$

also has an optimal value, which is equal to the optimal primal value. For all feasible primal-dual pair (ξ,u) we have the implication

$$\left.\begin{aligned} Q\xi &\leq t \\ Q^T u - P^T x = 0,\ u &\geq 0 \end{aligned}\right\} \Rightarrow (P^T x)^T \xi \leq t^T u,$$

with equality if and only if the pair (ξ,u) is optimal. If we substitute the bounding value $t^T u$ to the maximization in the left-hand side of

$$\bar{a}^T x + \max_{\xi}\{(P^T x)^T \xi \mid Q\xi \leq t\} \leq b,$$

we obtain a stronger inequality, which coincides with the initial one when u is dual optimal. Consequently, the system

$$\begin{aligned}\bar{a}^T x + t^T u &\leq b \\ P^T x - Q^T u &= 0 \\ u &\geq 0,\end{aligned}$$

which contains the feasibility constraints of the dual is equivalent to the initial robust constraint. □

3.4.4 Uncertainties in the objective function

Let us consider the following linear problem with uncertainty data

$$\begin{aligned}\min_x \ & \sum_{j=1}^{n} c_j(\xi) x_j \\ & Ax \leq b.\end{aligned}$$

There is a big difference difference between the objective and the constraints. In the latter, the values achieved by the left-hand side are irrelevant, insofar as they are less than the right-hand side. In particular, the average value is not relevant. In contrast, the average value of the objective function value is by all means an important factor to be considered. Let us assume that the uncertain factor has a symmetric distribution around 0 and that $c(\xi)$ is linear. The average value of the objective is thus $\sum_{j=1}^{n} c_j(0)x_j$ and one can simply minimize it. But in so doing, one completely ignore possible bad outcomes with high values for the objective function.

A possible alternative is to focus on the worst case and consider the problem

$$\min_{x,z} \ z \tag{11a}$$

$$\sum_{j=1}^{n} c_j(\xi) x_j - z \leq 0, \ \forall \xi \in \Xi \tag{11b}$$

$$Ax \leq b. \tag{11c}$$

The z value in (3.11) is an upper bound for the objective function for all possible realizations in the uncertainty set. Minimizing this worst case performance is an acceptable decision criterion, but one may argue that shifting from one problem to the other amounts to pass from one extreme (no concern for the risk) to another (no concern for the average performance).

At minimal computational cost, one may consider a bi-criteria problem involving the two objectives. The Pareto frontier is obtained by solving the parametric problem

$$\begin{aligned}
\min_x \quad & \sum_{j=1}^{n} c_j(0)x_j^* \\
& \sum_{j=1}^{n} c_j(\xi)x_j \le z^* + \gamma, \ \forall \xi \in \Xi \\
& Ax \le b.
\end{aligned}$$

In that formulation $\gamma > 0$ determines an arbitrage between the average performance and the guaranteed worst case performance.

3.4.5 Two-sided inequality constraints

In the robust optimization paradigm, the constraints are immunized separately with respect to their specific uncertainty sets. In the case of the two-sided constraint

$$\underline{b} \le \sum_{j=1}^{n} \tilde{a}_j x_j \le \bar{b}$$

the robust formulation involves the two constraints

$$\underline{b} \le \bar{a}^T x + (P^T x)^T \xi, \ \forall \xi \in \Xi$$

and

$$\bar{a}^T x + (P^T x)^T \xi \le \bar{b}, \ \forall \xi \in \Xi.$$

Assume the polyhedral uncertainty set $\Xi = \{\xi \mid B_1(0,k_1) \cap B_\infty(0,k_\infty)\}$. In view of Theorem 2, the two robust constraints generate their own robust counterpart

$$\underline{b} \le \bar{a}^T x - k_1 ||P^T x - v||_\infty - k_\infty ||v||_1 \tag{3.12}$$

$$\bar{a}^T x + k_1 ||P^T x - w||_\infty + k_\infty ||w||_1 \le \bar{b}. \tag{3.13}$$

In view of Proposition 1, those two inequalities boil down into independent sets of inequalities. The point of this section is that the two sets of extra inequalities don't need to be independent. Actually the same set is valid for both, thus saving significantly on the problem size.

To see this, it suffices to notice that one need not to choose different vectors v and w in the robust counterparts (3.12) and (3.13). Indeed, for a fixed x, the constraints are equivalent to

$$\begin{aligned}
k_1 ||P^T x - v||_\infty + k_\infty ||v||_1 &\le a^T x - \underline{b} \\
k_1 ||P^T x - w||_\infty + k_\infty ||w||_1 &\le \bar{b} - a^T x.
\end{aligned}$$

To check satisfiability of these constraints, it suffices to replace v and w by the value that minimizes the left-hand sides. This can be conveniently formalized into a theorem similar to Proposition 1.

Proposition 3. *The robust counterpart of the robust two-sided constraint*

$$\underline{b} \leq \bar{a}^T x + (P^T x)^T \xi \leq \bar{b}, \text{ for all } \xi \in \Xi = \{\xi \mid B_1(0,k_1) \cap B_\infty(0,k_\infty)\},$$

is

$$k_1 ||P^T x - w||_\infty + k_\infty ||w||_1 \leq \min\{a^T x - \underline{b}, \bar{b} - a^T x\}.$$

It can be represented by the system of linear inequalities

$$k_1 t + k_\infty e^T w \leq a^T x - \underline{b} \tag{14a}$$

$$k_1 t + k_\infty e^T w \leq \bar{b} - a^T x \tag{14b}$$

$$w + te \geq P^T x \tag{14c}$$

$$w + te \geq -P^T x \tag{14d}$$

$$w \geq 0, \quad t \geq 0.$$

The practical importance of this theorem is that the robust counterpart of a two-sided inequality is a set of inequalities with only one more inequality than in the case of a simple one-sided inequality.

A similar treatment applies to a two-sided inequality with ellipsoidal uncertainty. We then would get the robust counterpart

$$k||P^T x||_2 \leq \min\{a^T x - \underline{b}, \bar{b} - a^T x\}.$$

3.4.6 Equality constraints

The meaning of an equality constraint with uncertain coefficients is questionable. Using our standard representation of the uncertain parameters, we can formulate the robust equality constraint $\tilde{a}(\xi)^T x = b$, with $\tilde{a}(\xi) = \bar{a} + P\xi$ and $\xi \in \Xi$. The uncertainty sets that have been considered so far have a non-empty interior and $0 \in \Xi$. Consequently $\tilde{a}^T x = \bar{a}^T x + (P^T x)^T \xi = b$ for all $\xi \in \Xi$ if and only if $(P^T x)^T \xi \equiv 0$ for all $\xi \in \Xi$. This implies $P^T x = 0$. In other words, the vector x should lie in the null space of P^T. This is conveniently summarized in a proposition.

Proposition 4. *Suppose the uncertainty set Ξ has a non-empty interior. The robust equality constraint*

$$\bar{a}^T x + (P\xi)^T x = b, \ \forall \xi \in \Xi$$

is equivalent to the system of equations

$$\bar{a}^T x = b \ \text{ and } \ P^T x = 0.$$

The condition on x looks very restrictive. A possibility is that the presence of uncertain coefficients in an equality constraint signals an error in passing from a deterministic version of the problem to an uncertain one. For instance, some inequality constraints in the deterministic version are known to be necessarily tight at the optimum. This implicit knowledge makes it possible to write them as equalities in the deterministic formulation, but the true inequality formulation should be used in the uncertainty case.

A more difficult situation occurs in multistage problems in which some decisions are recourses that can be adjusted once the uncertainty is revealed. We shall see an illustration of this problem in our energy environmental planning example in a later section dealing with robust optimization in the dynamic case.

3.5 Dynamic problem with recourse

Multistage problems under uncertainty introduce a quantum of difficulties in that the successive decisions can be made on the basis of the information revealed so far. Those decisions will be named *recourses* thereafter, to emphasize the relation with the revealed information. Recourses are functions defined on richer and richer spaces. To cope with this extreme difficulty, it is reasonable to restrict the recourses to functions in a limited class, a class simple enough to allow the formulation of robust constraints, but rich enough to capture a meaningful part of the recourse possibilities. The proposed class is the set of linear functions of the revealed values of the uncertain factor ξ. Recourses in this class will be named Linear Decision Rules, LDR in short. This concept has been introduced for control problem (31) and for stochastic programming (29; 23). It is used in Robust Optimization under the name of Affinely Adjustable Robust Counterpart (AARC) (8).

3.5.1 Linear decision rules

Consider a typical constraint in a two-stage problem:

$$a_1(\xi_1)^T x_1 + a_2(\xi_1,\xi_2)^T x_2 \le b(\xi_1,\xi_2).$$

We assume that x_1 and x_2 are variables with dimension, respectively, n and m. The coefficients a_1 and a_2 have also, respectively, dimension n and m. The dependence of coefficients a_1, a_2 and b with the uncertainty underlying (ξ_1,ξ_2) is given by the functions

$$a_1(\xi_1) = \bar{a}_1 + P_1\xi_1 \tag{15a}$$
$$a_2(\xi_1,\xi_2) = \bar{a}_2 + P_{21}\xi_1 + P_{22}\xi_2 \tag{15b}$$
$$b(\xi_1,\xi_2) = \bar{b} + b_1^T\xi_1 + b_2^T\xi_2. \tag{15c}$$

In the last equality, b_1 and b_2 have same dimension than ξ_1 and ξ_2.

To capture the adaptive property of the recourse in stage 2, the variable x_2 is replaced by a LDR, that is by a linear (more accurately, an affine) function of ξ_1, that is:

$$x_2(\xi_1) = \bar{x}_2 + D\xi_1.$$

In the above expression, the components of $\bar{x}_2$ and D are the new decision variables in the formulation of the problem with uncertain parameters. Note that those variables are to be determined prior to knowing the value of ξ_1, but the actual value of the recourse $x_2(\xi_1)$ will be determined after the value of ξ_1 has become known.

The robust equivalent of the constraint of the two-stage problem is then

$$(\bar{a}_1 + P_1\xi_1)^T x_1 + (\bar{a}_2 + P_{21}\xi_1 + P_{22}\xi_2)^T(\bar{x}_2 + D\xi_1) \\ -(\bar{b} + b_1\xi_1 + b_2\xi_2) \leq 0,\ \forall(\xi_1,\xi_2) \in \Xi.$$

To build the robust counterpart, one must solve the optimization problem

$$\max_{(\xi_1,\xi_2)\in\Xi} (P_1^T x_1 + P_{21}^T\bar{x}_2 + D^T\bar{a}_2 - b_1)\xi_1 + (P_{22}^T\bar{x}_2 - b_2)^T\xi_2 \\ +(P_{21}\xi_1 + P_{22}\xi_2)^T D\xi_1 \tag{3.16}$$

in which the variables x_1, $\bar{x}_2$ and D are parameters. We notice that this problem is quadratic with second order term

$$(P_{21}\xi_1 + P_{22}\xi_2)^T D\xi_1.$$

Remark 2. It is worth noticing that a LDR introduces uncertainty of its own. A constraint with deterministic coefficients, e.g., with $P_1 = P_{21} = 0$, $P_{22} = 0$, $b_1 = 0$ and $b_2 = 0$, becomes uncertain because of the term $(D^T\bar{a}_2 - b_1)\xi_1$.

3.5.2 Problems with uncertain recourse parameters

If the coefficient $a_2(\xi_1,\xi_2)$ of the recourse variable is not fixed, the matrices P_{21} and P_{22} are not identically zero. Since D is not restricted to be in a certain class, the quadratic term in the objective of (3.16) can be indefinite. The maximization problem in (3.16) becomes non-convex and one cannot resort to duality theory to build a simple robust counterpart, as it has be done until now. There is an exception to this bad situation. If the uncertainty set Ξ is an ellipsoid, it can be shown that the problem (3.16) is equivalent to a convex problem on the cone of the positive semi-definite matrices. It becomes possible to derive a robust counterpart. We refer the reader to (11) and (8) for more details. We just mention that there exist efficient codes, e.g., the open source Sedumi (38), that can solve very efficiently problems with constraints on the set of positive semi-definite matrices.

3.5.3 Problems with fixed recourse

To remain in the realm linear programming, we introduce the following assumption.

Assumption 2 *The $a_2(\xi)$ vector associated to the recourse is fixed ($a_2(\xi) \equiv a_2$).*

Proposition 5. *Assume the dynamic problem has fixed recourse. The robust counterpart of a dynamic constraint with LDR and a polyhedral uncertainty set is given by a set of linear constraints.*

Proof. From the fixed-recourse assumption, we have $P_{21}\xi_1 + P_{22}\xi_2 \equiv 0$. Since the quadratic term of the objective disappears, we can apply Proposition 2 and express the robust counterpart as a set of linear constraints. □

The fixed-recourse assumption confines the uncertainty to the first stage coefficients and to the right-hand sides. This is obviously limitative, but still relevant to interesting applications problems, for instance in supply chain management problem (6) or on the management of a hydraulic valley with uncertainty on the water supply (3; 2).

In some dynamic problems, we encounter constraints $a^T x = b$, with a deterministic and b uncertain. We shall get an example of it in our case problem. There, $a^T x$ is a production that is achieved by combining factors in proportion given by x and b is a demand. The only way to match an uncertain demand is to make the x vector uncertain and have it adjusted to the random behavior of b. This is achieved by a LDR incorporating the random factor in b. Two options are open. The first one is to keep working with the equality constraint and achieve global robustness as described in subsection 3.4.6. This approach has been followed in (35). An alternative consists in relaxing the equality into the inequality $a^T x \geq b$, guaranteeing that the demand will be satisfied at the possible cost of disposal if the production exceeds the demand. This approach has been followed in (6).

3.6 Case study: uncertain demand with fixed recourse

In this section, the case study deals with the model with uncertain demand and no uncertainty in the pollutant transfer coefficients. This is a multistage problem with fixed recourse for which the LDR methodology of the previous section is appropriate. To contrast Robust Optimization with more classical approaches, we have implemented a simple Stochastic Programming version. We built a discrete distribution to approximate the stochastic demand process used in the validation process. The number of branches is such that the so-called *deterministic equivalent*, i.e., the extensive formulation of the problem relatively to the event tree, has about the same size as the robust counterpart.

3.6.1 LDR and relaxed demand constraints

We reformulate Problem (3.1) to eliminate equality constraints, either by simple elimination (constraints (1e)) or by relaxation (constraints (1d)). We get the following relaxed version of (3.1)

$$\min_{x\geq 0, y\geq 0} \sum_r (O_r \sum_{p,t} x_{p,t,r} + I_r \sum_{p,t} y_{p,t,r} + M_r \sum_{p,t} (res_{p,t,r} + \sum_{\tau\leq t} y_{p,\tau,r})) \tag{17a}$$

$$x_{p,t,r} \leq res_{p,t,r} + \sum_{\tau\leq t} y_{p,\tau,r} \quad \forall t \in \mathscr{T}, p \in \mathscr{P}, r \in \mathscr{R} \tag{17b}$$

$$\sum_{t,r} x_{p,t,r} \geq d_t \quad \forall t \in \mathscr{T} \tag{17c}$$

$$\sum_{\rho\in\mathscr{R}} \sum_p (E_{p,\rho} x_{p,t,\rho}) G_{\rho,r} \leq Q \quad \forall t \in \mathscr{T}, r \in \mathscr{R} \tag{17d}$$

This more compact formulation has 72 variables and 51 constraints. We now define, for all pairs $(p \in \mathscr{P}, r \in \mathscr{R})$, the following linear decision rules for the production variables $x_{p,t,r}$ and the investment variables $y_{p,t,r}$.

$$x_{p,1,r} = \alpha^0_{p,1,r} + \alpha^1_{p,1,r}\eta_1 \tag{18a}$$

$$x_{p,2,r} = \alpha^0_{p,2,r} + \alpha^1_{p,2,r}\eta_1 + \alpha^2_{p,2,r}\eta_2 \tag{18b}$$

$$x_{p,3,r} = \alpha^0_{p,3,r} + \alpha^1_{p,3,r}\eta_1 + \alpha^2_{p,3,r}\eta_2 + \alpha^3_{p,3,r}\eta_3 \tag{18c}$$

$$y_{p,1,r} = \beta^0_{p,1,r} \tag{18d}$$

$$y_{p,2,r} = \beta^0_{p,2,r} + \beta^1_{p,2,r}\eta_1 \tag{18e}$$

$$y_{p,3,r} = \beta^0_{p,3,r} + \beta^1_{p,3,r}\eta_1 + \beta^2_{p,3,r}\eta_2. \tag{18f}$$

In that definition the random variables η_i are the ones of the demand uncertainty defined in (3.3). Thus the variables x and y are function of the demand uncertainty. Note that the investment in t is decided before the demand at t is known, while the production is set after, hence the difference in the decision rules.

If we replace the variables in the constraints of (3.17) by their linear decision rules, we obtain the LDR formulation

$$\min_{\alpha,\beta,\xi} \sum_r (O_r \sum_{p,t} \alpha^0_{p,t,r} + I_r \sum_{p,t} \beta^0_{p,t,r} + M_r \sum_{p,t} (res_{p,t,r} + (4-p)\beta^0_{p,t,r}) \tag{19a}$$

$$\alpha^1_{p,1,r}\eta_1 + \alpha^0_{p,1,r} - \beta^0_{p,1,r} \leq res_{p,1,r} \quad \forall p,r \tag{19b}$$

$$\alpha^2_{p,2,r}\eta_2 + (\alpha^1_{p,2,r} - \beta^1_{p,2,r})\eta_1 + \\ + \alpha^0_{p,2,r} - \beta^0_{p,1,r} - \beta^0_{p,2,r} \leq res_{p,2,r} \quad \forall p,r \tag{19c}$$

$$\alpha^3_{p,3,r}\eta_3 + (\alpha^2_{p,3,r} - \beta^2_{p,3,r})\eta_2 + (\alpha^1_{p,3,r} - \beta^1_{p,2,r} - \beta^1_{p,3,r})\eta_1 + \\ + \alpha^0_{p,3,r} - \beta^0_{p,1,r} - \beta^0_{p,2,r} - \beta^0_{p,3,r} \leq res_{p,3,r} \quad \forall p,r \tag{19d}$$

$$(\sum_{p,r} \alpha^1_{p,1,r} - \hat{d}_1)\eta_1 + \sum_{p,r} \alpha^0_{p,1,r} \geq \bar{d}_1 \tag{19e}$$

$$(\sum_{p,r} \alpha^2_{p,2,r} - \hat{d}_2)\eta_2 + (\sum_{p,r} \alpha^1_{p,2,r} - \hat{d}_1)\eta_1 + \sum_{p,r} \alpha^0_{p,2,r} \geq \bar{d}_2 \tag{19f}$$

$$(\sum_{p,r} \alpha^3_{p,3,r} - \hat{d}_3)\eta_3 + (\sum_{p,r} \alpha^2_{p,3,r} - \hat{d}_2)\eta_2 + \\ + (\sum_{p,r} \alpha^1_{p,3,r} - \hat{d}_1)\eta_1 + \sum_{p,r} \alpha^0_{p,3,r} \geq \bar{d}_3 \tag{19g}$$

$$\sum_{i\in\mathscr{R}} \sum_p (E_{p,i}\alpha^1_{p,1,r})G_{i,r}\eta_1 + \sum_{i\in\mathscr{R}} \sum_p (E_{p,i}\alpha^0_{p,1,r})G_{i,r} \leq Q \quad \forall r \tag{19h}$$

$$\sum_{i\in\mathscr{R}} \sum_p (E_{p,i}\alpha^2_{p,2,r})G_{i,r}\eta_2 + \sum_{i\in\mathscr{R}} \sum_p (E_{p,i}\alpha^1_{p,2,r})G_{i,r}\eta_1 + \\ + \sum_{i\in\mathscr{R}} \sum_p (E_{p,i}\alpha^0_{p,2,r})G_{i,r} \leq Q \quad \forall r \tag{19i}$$

$$\sum_{i\in\mathscr{R}} \sum_p (E_{p,i}\alpha^3_{p,3,r})G_{i,r}\eta_3 + \sum_{i\in\mathscr{R}} \sum_p (E_{p,i}\alpha^2_{p,3,r})G_{i,r}\eta_2 + \\ + \sum_{i\in\mathscr{R}} \sum_p (E_{p,i}\alpha^1_{p,3,r})G_{i,r}\eta_1 + \sum_{i\in\mathscr{R}} \sum_p (E_{p,i}\alpha^0_{p,3,r})G_{i,r} \leq Q \quad \forall r \tag{19j}$$

$$\alpha^0_{p,1,r} + \alpha^1_{p,1,r}\eta_1 \geq 0 \quad \forall p,r \tag{19k}$$

$$\alpha^0_{p,2,r} + \alpha^1_{p,2,r}\eta_1 + \alpha^2_{p,2,r}\eta_2 \geq 0 \quad \forall p,r \tag{19l}$$

$$\alpha^0_{p,3,r} + \alpha^1_{p,3,r}\eta_1 + \alpha^2_{p,3,r}\eta_2 + \alpha^3_{p,3,r}\eta_3 \geq 0 \quad \forall p,r \tag{19m}$$

$$\beta^0_{p,1,r} \geq 0 \quad \forall p,r \tag{19n}$$

$$\beta^0_{p,2,r} + \beta^1_{p,2,r}\eta_1 \geq 0 \quad \forall p,r \tag{19o}$$

$$\beta^0_{p,3,r} + \beta^1_{p,3,r}\eta_1 + \beta^2_{p,3,r}\eta_2 \geq 0 \quad \forall p,r. \tag{19p}$$

In that formulation all constraints depend on the random variables ξ and must be given a robust counterpart. For each constraint we choose the uncertainty set to be the intersection of the $B_1(0,k\sqrt{m})$ and $B_\infty(0,k)$ balls, where m is the number of random variables in the constraint. We also select different immunization factors k according to the type of constraints. Namely, we use k_{cap} for the capacity constraints (19b)–(19d), k_{dem} for the demand constraints (19e)–(19g) and k_{emi} for the air quality constraints (19h)–(19j). In the experiments, we set $k_{cap} = k_{emi} = 1$ and we test different values for k_{dem}. The robust equivalent of (3.19) has 419 variables and 502 constraints. It corresponds to an increase with a multiplicative factor 4 of size of the deterministic version (3.17).

We report the results in Table 3.13.

Table 3.13 Variable demand: behavior on the sample of 1000 scenarios of a robust solution with LDR.

	Deterministic	LDR with robust		
		$k_{dem} = 0.4$	$k_{dem} = 0.6$	$k_{dem} = 0.7$
Objective function				
Predicted cost performance	162.06	165.84	167.09	167.40
Observed cost performance	160.87	165.07	166.14	167.47
Constraints on the demand				
Scenarios with demand violation(s) in %	62.0	20.4	9.9	0
Conditional average relative violation in %	2.5	0.8	0.6	-
Average number of violations per scenario	2.0	1	1	-
Constraints on the air quality				
Total number of violated air quality constraints	0	53	39	82

The conclusions to be derived from Table 3.13 are twofold. If we increase the immunization level k_{dem}, the robust solution better and better tracks the (relaxed) demand constraint. On the other hand, a larger immunization factor induces more variability in the LDR and increases risks of violations in the (otherwise deterministic) air quality constraints. Note also the degradation of the cost performance when the immunization increases.

3.6.2 LDR and exact demand constraints

In the previous section, we transformed the equality demand constraint in the deterministic model (1d) into the inequality constraint (17c) to match the uncertainty in the right-hand side. We now follow the approach of Section 3.4.6 and keep the equality in the demand constraint. In view of Proposition 4, the linear decision rule coefficients must satisfy the following conditions. Namely, we replace the demand equations (19e)–(19g) by the following set of equations

$$\sum_{p,r} \alpha^0_{p,1,r} = \bar{d}_1, \quad \sum_{p,r} \alpha^1_{p,1,r} = \hat{d}_1, \tag{20a}$$

$$\sum_{p,r} \alpha^0_{p,2,r} = \bar{d}_2, \quad \sum_{p,r} \alpha^1_{p,2,r} = \hat{d}_1, \quad \sum_{p,r} \alpha^2_{p,2,r} = \hat{d}_2, \tag{20b}$$

$$\sum_{p,r} \alpha^0_{p,3,r} = \bar{d}_3, \quad \sum_{p,r} \alpha^1_{p,3,r} = \hat{d}_1, \quad \sum_{p,r} \alpha^2_{p,3,r} = \hat{d}_2, \quad \sum_{p,r} \alpha^3_{p,3,r} = \hat{d}_3. \tag{20c}$$

These conditions ensure that the demand constraints will be globally satisfied for all possible demands in the 3-dimensional space ($\mathbb{R}^3$). The model has now 412 variables and 487 constraints.

In the numerical example, we keep replacing the other constraints (the capacity constraints (19b)–(19d), the air quality constraints (19h)–(19j) and the non-negativity constraints on the production and capacities) with their equivalent robust counterpart. To make the results comparable, we set $k_{cap} = k_{emi} = 1$ as before. The last set of constraints deals with non-negativity. To enforce those constraints with maximum chances we use an immunization level k equal to 1. The results are displayed in Table 3.14. Surprisingly enough, the results are very much alike the

Table 3.14 Variable demand: behavior on the sample of 1000 scenarios of a globally robust solution with respect to the demand constraint.

Predicted cost performance	167.40
Observed cost performance	167.47
Scenarios with demand violation(s) in %	0
Total number of violated air quality constraints	82

relaxed case with $k_{dem} = 0.7$ in Table 3.13. This reflect the fact that it is relatively easy to track the demand by simple adjustment in the current period.

3.6.3 Stochastic programming

To apply stochastic programming, we model the demand process by a finite event tree. To end up with a problem with a size comparable to (3.19), we use a tree with 3 branches at each node and thus 27 scenarios in total. We assume that the probability factors η_1, η_2 and η_3 are i.i.d. with a uniform distribution. The best approximation of the uniform distribution by a finite distribution with 3 elements consists in partitioning the range space $[-1, 1]$ into 3 elements and choose the mid-point value of each subinterval to be the representative of the subinterval. The probability for each representative is $\frac{1}{3}$. The tree is represented in Figure 3.3.

If we start from the compact problem (3.17), the stochastic formulation has 624 variables and 663 constraints (419 variables and 502 constraints for the LDR formulation).

Table 3.15 Variable demand: behavior on the sample of 1000 scenarios of a stochastic programming solution built on an approximating event tree with 3 branches per node.

	Deterministic	Stochastic
Predicted cost performance	162.06	165.18
Observed cost performance	160.87	164.80
Scenarios with demand violation(s) in %	62.0	50.4
Conditional average relative violation in %	2.5	1.0
Average number of violations per scenario	2.0	1.4

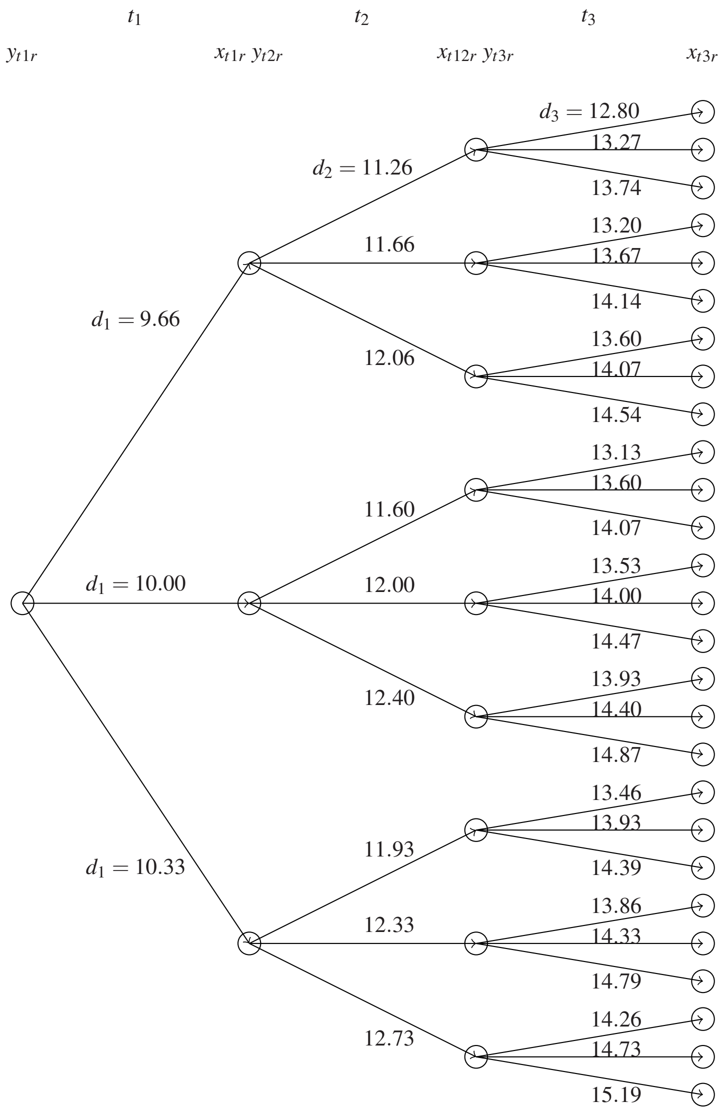

Fig. 3.3 Scenario decision tree

Table 3.15 suggests that the model cannot really cope with the uncertainty in the demand constraints. Indeed, production is determined relatively to the representative demand at the node. In the simulations, the demand is never equal to the value assigned to the node. With the symmetric distribution it lies above this value in 50% of the cases. However, we observe that the stochastic programming solution improves upon the deterministic solution, and contrary to the robust solution with LDR, it never induces air quality constraints. This is so because productions at each node are chosen to meet those constraints. In the simulations, the productions are either kept unchanged or lowered to match lower demand; so, no air quality constraint violation is to be expected.

It would be preposterous to derive definitive conclusions from this short example, but we can stress few facts. Stochastic Programming is at pain to meet constraints, while Robust Optimization controls violations pretty well. Stochastic programming treats constraints at privileged points: the nodes on the event tree, but Robust Optimization deals with the continuum of events in the event tree. To improve the control on the constraint in Robust Optimization, it suffices to increase the immunization level. A similar objective of better constraint satisfaction with Stochastic Programming calls for a larger event tree to better approximate the stochastic process, an approach that quickly leads to an numerically intractable optimization problem.

On the other side, Stochastic Programming delivers a richer output, provided one has reliable and tractable information on the probability distribution. This takes the form of the objective performance on each branch on the event tree, that gives an interesting approximation of the probability distribution of the optimal function value. It is even possible to add constraints, such as CVaR (conditional value at risk), to take into account the amount of risk.

3.7 Case study: uncertainty in demands and pollutant transfers

We consider now the more involved case where both the demands and the pollutant transfer coefficients are uncertain. As it has been pointed out earlier, LDR introduce an uncertainty of their own in the constraints in which they appear. If some of these constraints have themselves uncertain coefficients, the product of an uncertain coefficient with the uncertain decision variable (defined by the LDR) generates a bilinear form in the uncertainty. One cannot anymore use duality as in Section 3.4 to exhibit the robust counterpart.

To cope with this difficulty, we propose two empirical approaches, one that fully remains in the realm of Robust Optimization and a hybrid one that mixes Robust Optimization and Stochastic Programming.

3.7.1 A fully robust optimization approach

In the fully robust optimization approach, we eliminate the LDR in the air quality constraints by the following gimmick. We introduce an artificial variable to bound the value produced by the LDR. In our problem of interest the artificial variable represents a deterministic upper bound on the productions prescribed by the LDR. In view of the nature of air quality constraints with nonnegative coefficients only, higher productions entail higher emissions and, after pollutant transmission, deteriorate the local air quality. Therefore, it makes sense to replace the LDR value by an upper bound to comply with the worst case requirement.

This idea is implemented as follows. The air quality constraints (19h)–(19j) in the LDR formulation (3.19) are replaced by

$$\sum_{i\in\mathscr{R}}\sum_{p}(E_{p,i}x_{p,t,r})G_{i,r} \leq Q \quad \forall r,t \tag{21a}$$

$$\alpha^0_{p,1,r}+\alpha^1_{p,1,r}\eta_1 \leq x_{p,1,r} \quad \forall p,r \tag{21b}$$

$$\alpha^0_{p,2,r}+\alpha^1_{p,2,r}\eta_1+\alpha^2_{p,2,r}\eta_2 \leq x_{p,2,r} \quad \forall p,r \tag{21c}$$

$$\alpha^0_{p,3,r}+\alpha^1_{p,3,r}\eta_1+\alpha^2_{p,3,r}\eta_2+\alpha^3_{p,3,r}\eta_3 \leq x_{p,1,r} \quad \forall p,r. \tag{21d}$$

where the matrix G is uncertain. Because the upper bound x is deterministic, we now can deal with uncertain coefficients G in (21a) and write their robust counterpart. Note that constraints (21b)–(21d) together with the non-negativity constraints (19k)–(19m) form a group of two-sided constraints. We use Proposition 3 to cut down the number of constraints and variables in the robust counterpart of those two-sided constraints. Note that the production LDR variables are replaced by an upper bound in the air quality constraints (21a), but not in the capacity constraints (19b)–(19d). The formulation has 487 variables and 542 constraints

The numerical experiments aim to compare the LDR solution computed in subsection 3.6.1 and the solution of the fully robust optimization approach. In the latter, we used different immunization levels k_{emi} for the air quality constraints. In all cases we used the same immunization level $k_{dem} = 0.7$ for the demand constraints. The results are displayed in Table 3.16.

Table 3.16 Variable demand and variable transfer coefficients: behavior on the sample of 1000 scenarios of a robust solution with LDR.

	Standard LDR	LDR with robust const.		
		$k_{emi}=0$	$k_{emi}=0.5$	$k_{emi}=1$
Objective function				
Predicted cost performance	167.40	167.66	169.21	170.72
Observed cost performance	167.47	167.51	169.28	170.78
Constraints on the demand				
Scenarios with demand violation(s) in %	0	0	0	0
Conditional average relative violation in %	-	-	-	-
Average number of violations per scenario	-	-	-	-
Constraints on the air quality				
Total number of violated air quality constraints	829	802	150	11

The reader will notice that the new approach with an the immunization level in the air quality constraint set to $k_{emi} = 0$, yields a solution such that the number of violated air quality constraints is slightly less (802 instead of 829) than with the standard LDR approach, while the cost performance is slightly worse (167.66 instead of 167.40). In both cases, the model treats the air quality constraint in a deterministic way, but in the new approach the LDR production variable is replaced by an upper bound in the air quality constraint. This more conservative approach is a safer for the constraint but also more costly. However the difference almost negligible. If we increase the immunization level k_{emi}, the solution with the new approach ensures better and better protection against violations of the air quality constraints.

3.7.2 A hybrid approach: stochastic programming with robust constraints

We now assume both the coefficients of the source-receptor matrix G and the demands are uncertain. As in the previous section, the demands are defined by (3.3). Because we cannot use LDR in a constraint with other uncertain coefficients, we choose to represent the demand uncertainty via a finite event tree. We shall use stochastic programming to define the decision to be taken at each node of the tree. The new feature is that those decisions will be requested to be robust with respect to the air quality constraints. Therefore, we shall have to introduce different robust counterparts of those constraints at each node of the tree. As before, we use the event tree displayed in Figure 3.3. The hybrid formulation with a demand event tree and robust air quality constraints at the tree nodes has 819 variables and 819 constraints. The results are displayed in Table 3.17.

Table 3.17 Variable demand and variable transfer coefficients: behavior on the sample of 1000 scenarios of a solution obtained by the mixed approach.

	Deterministic	Hybrid		
		$k = 0.5$	$k = 1.0$	$k = 1.2$
Objective function				
Predicted cost performance	165.18	166.35	167.46	167.88
Observed cost performance	164.80	165.98	167.09	167.51
Constraints on the demand				
Scenarios with demand violation(s) in %	50.4	50.4	50.4	50.4
Conditional average relative violation in %	1.0	1.0	1.0	1.0
Average number of violations per scenario	1.40	1.36	1.36	1.36
Constraints on the air quality				
% of simulations with air quality violations	54.6	27.1	1.5	0
Average number of air quality violations	1.7	1.5	1.2	0

If we compare the results of the LDR solution in Table 3.14 with those of the hybrid approach in Table 3.17, we observe that the latter can achieve much better control of the air quality constraints (even though there was no exogenous uncertainty in the first case and some in the other), while the former enables full control of the demand constraint, something that stochastic programming cannot achieve.

3.8 Probability of constraint satisfaction

Prior to applying the robust optimization to the example, it is worth relating robust optimization to chance-constrained programming. This alternative approach, that was introduced in the late fifties by Charnes and Cooper (22), replaces the deterministic constraint by a constraint in probability. Namely, it considers that a solution is admissible if the probability that the constraint be satisfied with this solution is higher than a certain threshold. Since robust optimization is not built on probability grounds, comparing the two approaches seems to be out of order. Surprisingly enough, it turns out that one can assess bounds on the probability of satisfaction of a constraint by a robust solution, at the cost of mild assumptions on the probability distributions of the uncertain parameters. The result stems from a theorem to be found in Chapter 3 of (5). We give here a slightly stronger version than the one to be found in the literature.

3.8.1 Bounds on the probability of constraint satisfaction

Theorem 3. *Let ξ_i, $i = 1,\dots,m$ be independent random variables with values in interval $[-1,1]$ and with average zero: $E(\xi_i) = 0$. If z_i, $i = 1,\dots,m$ are deterministic coefficients, we have for all $k \geq 0$*

$$Prob\left\{\xi \mid \sum_{i=1}^{m} z_i\xi_i > k\sqrt{\sum_{i=1}^{m} z_i^2}\right\} \leq \exp(-\frac{k^2}{1.5}).$$

Remark 3. The assumption in Theorem 3 deals with the support of the random variables, their expectation and their independence. No other probabilistic assumption is made, in particular, nothing concerning distributions.

The following two lemmas are used in the proof of Theorem 3.

Lemma 4 (Chebytchev inequality).

$$Prob(X \geq a) \leq e^{-a}E(e^X).$$

Proof.

$$\begin{aligned} E(e^X) &= E(e^X|X < a)\text{Prob}(X < a) + E(e^X \mid X \geq a)\text{Prob}(X \geq a) \\ &\geq E(e^X|X \geq a)\text{Prob}(X \geq a) \qquad (\text{because } e^X > 0 \Rightarrow E(e^X \mid X < a) \geq 0) \\ &\geq e^a\text{Prob}(X \geq a). \end{aligned}$$

The last inequality is from

$$E(e^X|X \geq a) \geq E(e^a|X \geq a) = e^a.$$

□

Lemma 5. *The inequality* $e^{\frac{t^2}{\alpha}} + t \geq e^t$ *with* $1 \leq \alpha \leq \sqrt{\frac{8}{3}}$ *is valid for all* $t \geq 0$.

Proof. The function $e^{t^2/\alpha} + t - e^t$ is decreasing in α for all $t > 0$. To prove the property, we consider the extreme value of α. We assume in the following $\alpha = \sqrt{\frac{8}{3}}$. The derivatives of function $g(t) = e^{t^2/\alpha} + t - e^t$ are

$$g'(t) = \frac{2}{\alpha} t e^{t^2/\alpha} + 1 - e^t$$

and

$$g''(t) = \frac{2}{\alpha}(\frac{2}{\alpha}t^2 + 1)e^{t^2/\alpha} - e^t = e^t\left(\frac{2}{\alpha}(\frac{2}{\alpha}t^2 + 1)e^{t^2/\alpha - t} - 1\right). \tag{3.22}$$

We note that $g'(0) = 0$ and $g''(0) = 2/\alpha - 1 \geq \sqrt{3/2} - 1 > 0$. We show first that $g''(t) \geq 0$ for all $t \geq 0$. Thus we show that the term in parenthesis in the right hand side of (3.22) is positive. From the inequality $e^x \geq 1 + x$, we have

$$\frac{2}{\alpha}(\frac{2}{\alpha}t^2 + 1)e^{t^2/\alpha - t} - 1 \geq h(t) = \frac{2}{\alpha}\left(\frac{2}{\alpha}t^2 + 1\right)\left(1 + \frac{t^2}{\alpha} - t\right) - 1.$$

The right component is a fourth degree polynomial that is convex because $\alpha < 4$ and

$$h''(t) = \frac{12}{\alpha^2}\left(\frac{4}{\alpha}t^2 - 2t + 1\right) = \frac{12}{\alpha^2}\left((1-t)^2 + (\frac{4}{\alpha} - 1)t^2\right) > 0.$$

The polynomial

$$h'(t) = \frac{2}{\alpha}\left(\frac{8}{\alpha^2}t^3 - \frac{6}{\alpha}t^2 + \frac{6}{\alpha}t - 1\right) = \sqrt{\frac{3}{2}}\left(3t^3 - 3\sqrt{\frac{3}{2}}t^2 + 3\sqrt{\frac{3}{2}}t - 1\right)$$

is thus increasing; its value at 0 is $-\sqrt{3/2}$ and it tends to $+\infty$ when t tends to $+\infty$. It has a unique real root in $\bar{t} = 0.366$. The function has its minimum in $\bar{t}$ and thus $g''(t) \geq h(t) \geq h(\bar{t}) = 0.0208$.

We proved that $g''(t) \geq 0$ on the interval $[0,1]$ and thus that the function $g'(t)$ is increasing. Since $g'(0) = 0$, we conclude that $g'(t) \geq 0$ and thus that $g(t)$ is increasing. Finally since $g(0) = 0$, we have $g(t) \geq 0$ on $0 \leq t \leq 1$. □

Proof. Now we can prove Theorem 3. This proof is derived from (5). From Lemma 4 we write

$$\text{Prob}(\sum_i z_i\zeta_i > k||z||) \leq e^{-k||z||}E(e^{\sum_i z_i\zeta_i}).$$

Since the variables ζ_i are independent, then $E(e^{\sum_i z_i\zeta_i}) = E(\prod_i e^{z_i\zeta_i}) = \prod_i E(e^{z_i\zeta_i})$. We also have

$$\begin{aligned} E(e^{z_i\zeta_i}) &= 1+\sum_{k\geq 2} E(\frac{(z_i\zeta_i)^k}{k!}) \quad \text{(because } E(\zeta_i)=0) \\ &\leq 1+\sum_{k\geq 2}\frac{|z_i|^k}{k!} \quad \text{(because } |z_i\zeta_i|\leq |z_i|) \\ &\leq e^{|z_i|}-|z_i| \leq e^{\frac{z_i^2}{\alpha}}. \end{aligned}$$

The last inequality comes from Lemma 5. Thus we can write

$$\text{Prob}(\sum_i z_i\zeta_i > k||z||) \leq e^{-k||z||}\prod_i e^{\frac{z_i^2}{\alpha}} = e^{-k||z||+\frac{||z||^2}{\alpha}}.$$

As the relation

$$\text{Prob}(\sum_i z_i\zeta_i > k||z||) = \text{Prob}(\gamma\sum_i z_i\zeta_i > \gamma k||z||)$$

is true for all $\gamma > 0$, then we have

$$\begin{aligned} \text{Prob}(\sum_i z_i\zeta_i > k||z||) &\leq \min_{\gamma>0} e^{-\gamma k||z||+\gamma^2\frac{||z||^2}{\alpha}} \\ &\leq e^{-\frac{\alpha^2 k^2}{4}} \leq e^{-\frac{k^2}{1.5}}. \end{aligned}$$

Remark 4. The standard result gives the bound $\exp(-k^2/2)$. An experimental study of the function $\exp(t^2/\alpha)+t-\exp(t)$ shows that the maximum value of α, for which Lemma 5 remains true, is 1.79. The bound in the proposition is tighten to $\exp(-t^2/1.25)$.

Let us apply the result when the uncertainty set is the ball $B_2(0,k)$ and the variable support for ξ_j is the interval $[-1,1]$. We consider the robust constraint

$$\bar{a}^T x + (P\xi)^T x \leq 0,\ \forall \xi \in B_2(0,k).$$

From Lemma 1

$$\max_{\xi\in\Xi}(P^T x)^T\xi \leq \max_{\xi}\{(P^T x)^T\xi \mid ||\xi||_2 \leq k\} = k||P^T x||_2.$$

Letting $z = P^T x$, Theorem 3 yields

$$\text{Prob}\left\{(\bar{a}+\ P\xi)^T x \leq 0 \mid \xi\in\Xi\right\} \geq \text{Prob}\left\{(P^T x)^T\xi \leq k||P^T x||_2\right\} \geq 1-\exp(-\frac{k^2}{1.5}).$$

Note that for $k = 2.63$ we get the bound 0.99 on the probability.

Using Theorem 3, we can also bound the probability associated with an uncertainty set defined as the intersection of two balls in the ℓ_1 and ℓ_∞ norms respectively. This done in the next corollary.

Corollary 1. *Let ξ_i, $i = 1,\dots,m$ be independent random variables with values in the interval $[-1,1]$ and average zero: $E(\xi_i) = 0$. If z_i, $i = 1,\dots,m$ are deterministic coefficients, we have for all $k \geq 0$*

$$Prob\left\{\xi \mid \sum_{i=1}^{m} z_i\xi_i > k\sum_{i=1}^{m}|z_i|\right\} \leq \exp(-\frac{k^2}{1.5\,m}).$$

Proof. The proof results from the relation between the norms ℓ_1 and ℓ_2 in $\mathbb{R}^m$ that implies

$$B_2(0,k/\sqrt{m}) \subseteq B_1(0,k).$$

Replacing k by $k/\sqrt{m}$ in Theorem 3 we get the result. □

3.8.2 Uncertainty set versus constraint satisfaction

To get a flavor of the implication of Theorem 3 and its corollary, it is worth comparing the volume of the certainty set $\Xi = B_\infty(0,1)$ with the volumes of the balls $B_1(0,k\sqrt{m})$ and $B_2(0,k)$, for some k guaranteeing a large probability of satisfying the uncertain constraint. Since the ball $B_1(0,k\sqrt{m})$ is a rotation of the ball $B_\infty(0,k)$, it has the same volume as the ball $B_\infty(0,k)$, and hence a volume larger than the ball $B_\infty(0,1)$ by a factor k^m (for $k \geq 1$). Notwithstanding, the ball $B_1(0,k\sqrt{m})$ intersects the ball $B_\infty(0,1)$ for $k \leq \sqrt{m}$. Hence the uncertainty set $\Xi = B_\infty(0,1) \cap B_1(0,k\sqrt{m})$ is smaller than the certainty set $B_\infty(0,1)$, but not in large proportion.

The situation with the ball $B_2(0,k)$ is dramatically different. This fact is discussed in (5) and is summarized by the inequality

$$\frac{\mathrm{Vol}B_2(0,k)}{\mathrm{Vol}B_\infty(0,1)} = \frac{(k\sqrt{\pi})^m}{2^m\Gamma(m/2+1)} \leq \left(k\sqrt{\frac{e\pi}{2m}}\right)^m.$$

For $k = 2.63$ and $m = 30$ the ratio on the right hand-side is strictly less than 1 and goes super-exponentially fast to zero as the dimension of the space of uncertainty factor goes to infinity. Hence in large dimension the uncertainty set may be incommensurably smaller than the certainty set, yet a solution which is robust with respect to this uncertainty set may achieve a very high probability of satisfying the constraint.

At this point, it is worth wondering whether it is appropriate to rely on an intuitive justification of the robust optimization based on the idea that it is necessary to select a large uncertainty set to achieve a high probability of constraint satisfaction. For sure, the probability of satisfying the constraint is at least as large as the probability associated with the uncertainty set, but the reverse is not true. We had a first confirmation of this fact by comparing the volume of the $B_2(0,k)$ uncertainty set with the volume of the certainty set $B_\infty(0,1)$. We now provide an example that shows that on some circumstances, it is possible to achieve very high probability while the uncertainty set has very small probability, even a zero probability!

Consider the constraint

$$\sum_{i=1}^{m} a_i x_i \leq b \tag{3.23}$$

with uncertain coefficients $a_i = \bar{a}_i + \hat{a}_i \xi_i$, where $\xi_i \in \Omega$. (Note that it is a particular case of the model $a = \bar{a} + P\xi$ where P is the diagonal matrix with main diagonal $\hat{a}$.) We define an uncertainty set $\Xi \subset \Omega$ and work with the robust constraint

$$\sum_{i=1}^{m} (\bar{a}_i + \hat{a}_i \xi_i) x_i \leq b, \; \forall \xi \in \Xi.$$

Suppose we can assess a probability distribution for ξ on Ω. We have the implication

$$\begin{gathered}\textstyle\sum_{i=1}^{m} (\bar{a}_i + \hat{a}_i \xi_i) x_i \leq b, \; \forall \xi \in \Xi \text{ and } \mathrm{Prob}(\xi \in \Xi) \geq 1 - \alpha \\ \Downarrow \\ \mathrm{Prob}(\textstyle\sum_{i=1}^{m} (\bar{a}_i + \hat{a}_i \xi_i) x_i \leq b) \geq 1 - \alpha.\end{gathered}$$

Truly enough, the reverse implication does not hold. Yet, the direct implication suggests that choosing an uncertainty set with large probability might be a good strategy. The following example reveals that the strategy can be overly conservative.

In this example we assume the a_i are independent random variables

$$a_i = \begin{cases} 1 \text{ with probability } 1/2 \\ 0 \text{ with probability } 1/2. \end{cases} \tag{3.24}$$

This uncertain constraint is cast into a robust optimization framework as follows. For the time being, forget about the probability distribution and write the coefficients in the form $a_i = \frac{1}{2} + \frac{1}{2}\xi_i$ with $-1 \leq \xi_i \leq 1$. Let the uncertainty set be the ellipsoid $\Xi = \{\xi \mid ||\xi||_2 \leq k\}$. The robust counterpart of the constraint

$$\frac{1}{2}\sum_i x_i + \frac{1}{2}\sum_i x_i \xi_i, \; \forall \xi \in \Xi$$

is

$$\frac{1}{2}\sum_i x_i + \frac{k}{2}\sqrt{\sum_i x_i^2} \leq b.$$

Since the ξ_i are independent random variables with zero mean with range $[-1,1]$, we can apply Theorem 3 and get

$$\mathrm{Prob}\left\{\tilde{\xi} \mid \sum_{i=1}^{m} z_i \tilde{\xi}_i > k\sqrt{\sum_i^{m} z_i^2}\right\} \leq \exp(-\frac{k^2}{1.5})$$

for arbitrary z_i. For $k = 2$ we get the bound 0.07 on the probability of constraint violation. If we apply the theorem to the robust counterpart, we have that a solution to

$$\frac{1}{2}\sum_i x_i + \sqrt{\sum_i x_i^2} \leq b$$

has a probability $1 - 0.07 = 0.93$ to satisfy the probabilistic constraint

$$\sum_{i=1}^{m} a_i(\xi_i)x_i \leq b$$

for any probability distribution satisfying the (weak) hypothesis of the theorem. In particular, this is true for (3.24).

On the other hand, the set $\Xi = \{\xi \mid ||\xi||_2 \leq 2\}$ *does not contain any single realization* of the random variable ξ as soon as ξ has dimension larger than 4. In other words $\mathrm{Prob}(\xi \in \Xi) = 0$. So, we imposed x to be robust with respect to an uncertainty set[2] having probability 0, but we still guarantee that the robust solution satisfies the uncertain constraint with a probability at least 0.93. Clearly, Robust Optimization does much more than the intuition suggests.

Let us pursue the discussion with this example. Consider now the same version of the problem but with binary variables x_i and $b = 60$. The robust counterpart (still with $k = 2$)

$$\frac{1}{2}\sum_i x_i + \sqrt{\sum_i x_i^2} \leq 60$$

is equivalent to a bound on the number variables that can set to the value 1. We easily find that this number is $N = 100$. We now ask the question: how good is the lower bound 0.93 on the true probability of constraint satisfaction? Without loss of generality, we assume that the robust solution is $x_i = 1$, $i = 1, \ldots 100$ and zero otherwise. The left-hand side in the constraint $\sum_{i=1}^{m} a_i x_i \leq 60$ is a binomial random variable, with parameter 1/2. From the tables, we get that the probability of satisfaction is 0.98. It is certainly an improvement upon 0.93, but still it is quite an achievement to get 0.93 with a robust approach, in view of the weak assumption on the true distribution in Theorem 3.

3.9 Extension: Globalized robust optimization

As we have seen it again and again, robust optimization deals with worst cases with respect to uncertainty sets. Robust optimization concentrates on solutions that remain feasible for all realizations within the uncertainty sets, but is silent about realizations that lie outside. In particular it does not take into account the magnitude of

[2] Note that in this example, the ellipsoidal set has either probability 0, if k is small or probability one.

the violation. Globalized robust optimization (4) proposes an extension that admits possible constraint violations, but control their magnitude.

3.9.1 The concept of globalized robust optimization

A globalized robust solution must satisfy the following two criteria:

1. The solution is robust for all realizations in the uncertainty set.
2. If a realization falls outside the uncertainty set, a violation of the constraint is tolerated, but this violation must "remain under control".

The fuzzy concept of remaining under control must be clarified, but we already see that a globalized robust solution is robust. The concept is then more restrictive, but it covers all cases, inside and outside the uncertainty set. To make the definition operational, one needs to be more specific about the meaning of "being under control". In (4) the authors suggest to consider the distance from a current realization to the uncertainty set. If this distance is positive, i.e., if the realization is strictly exterior to the uncertainty set, a violation is acceptable, but *its magnitude should be less than a fixed multiple of this distance*. This definition applies to the realizations within the uncertainty set, because, the distance is null, and no violation is permitted.

To formalize the idea, we first define an arbitrary distance to the uncertainty set. As in the previous examples, we consider the constraint

$$(\bar{a}+P\xi)^T x = \bar{a}^T x + (P^T x)^T \xi \leq b$$

to be satisfied for any ξ in the uncertainty set. We assume that $\Omega = \mathbb{R}^m$ and that Ξ is convex. We introduce the convex distance function $d(\xi, \Xi)$ between ξ and the set Ξ. A solution is called *globally robust* if it satisfies

$$\bar{a}^T x + (P^T x)^T \xi \leq b + \alpha d(\xi, \Xi), \forall \xi \in \Omega$$

where $\alpha > 0$ is a user parameter. This parameter is chosen in function of the tolerated violation. In this formulation, ξ can be given any value[3] in $\mathbb{R}^m$.

To check whether a solution x meets the globalized robustness requirements, it suffices to replace the uncertain terms by its maximal value

$$\bar{a}^T x + \max_{\xi \in \Omega} \{(P^T x)^T \xi - \alpha d(\xi, \Xi)\} \leq b.$$

To make things more precise, let us specify that the distance function is generated by a norm. Let $\delta : \mathbb{R}^m \to \mathbb{R}_+$ be this norm, and define the distance as

$$d(\xi, \Xi) = \min_{\xi' \in \Xi} \delta(\xi - \xi').$$

[3] This requirement may be excessive and unrealistic. One could think that the set is bounded, possibly with large bounds. This possibility is shortly discussed at the end of this section.

The maximization operation in the left-hand side of the globalized robust equivalent constraint is a convex problem (maximizing a concave function). It can be written as

$$\begin{aligned}\max_{\xi\in\mathbb{R}^m}\{(P^T x)^T\xi - \alpha \min_{\xi'\in\Xi}\delta(\xi-\xi')\} \\ &= \max_{\xi\in\mathbb{R}^m}\{(P^T x)^T\xi + \alpha \max_{\xi'\in\Xi}\{-\delta(\xi-\xi')\}\} \\ &= \max\{(P^T x)^T\xi - \alpha\delta(\xi-\xi') \mid \xi\in\mathbb{R}^m,\, \xi'\in\Xi\}.\end{aligned}$$

We have thus the globalized robust equivalent written as

$$\bar{a}^T x + \max\{(P^T x)^T\xi - \alpha\delta(\xi-\xi') \mid \xi\in\mathbb{R}^m,\, \xi'\in\Xi\} \le b. \tag{3.25}$$

3.9.2 Globalized robustness with linear programming

We shall use this expression with a choice of the norm that keeps the whole formulation in the realm of linear programming. To this end, we select a polyhedral uncertainty set and an appropriate norm that can be described by a finite number of linear inequalities, such as those generated by the balls in the ℓ_∞ and/or ℓ_1 norms. The distance to this uncertainty set will also be defined with respect to ℓ_∞ and/or ℓ_1, so that everything can be translated into linear inequalities. In principle, the combination of the ℓ_∞ and ℓ_1 norms used in defining the distance may be totally independent of the structure of the uncertainty set, but we choose to analyze a case where the uncertainty set and the distance function derive both from the same norm.

Consider the uncertainty set $\Xi = \{\xi \in \Omega \mid ||\xi||_1 \le k_1,\, ||\xi||_\infty \le k_\infty\}$ defined by the intersection of the ball with radius k_∞ in the norm ℓ_∞ and the ball with radius k_1 in the norm ℓ_1. (In practice we often choose $k_1 = k_\infty\sqrt{m}$.) This polyhedron can be used to defined a norm as follows. Consider the homothetic set

$$\Xi(t) = \{\xi \in \Omega \mid ||\xi||_1 \le k_1 t,\, ||\xi||_\infty \le k_\infty t\}$$

with $t > 0$. We can use this polyhedron to define the norm

$$\delta(\mu) = \min_{t>0}\{t \mid \mu \in \Xi(t)\} = \max\{\frac{1}{k_1}||\mu||_1, \frac{1}{k_\infty}||\mu||_\infty\}.$$

The distance from a point ξ to the set Ξ, is now defined by the optimization problem

$$d(\xi,\Xi) = \min_{\nu}\{\delta(\xi-\nu) \mid \nu \in \Xi\}.$$

Figure 3.4 illustrates the distance from ξ to the set Ξ with the selected norm in the 2-dimensional space with $k_1 = k_\infty\sqrt{2}$. The blue dotted curves represent the sets of points that are at the same distance of Ξ. The point $\nu \in \Xi$ is one of the closest point from ξ. (The vector ν is not unique because the distance function is not strictly convex.) The computed distance is $d(\xi,\Xi) = \delta(\nu,\xi) = 2.53$.

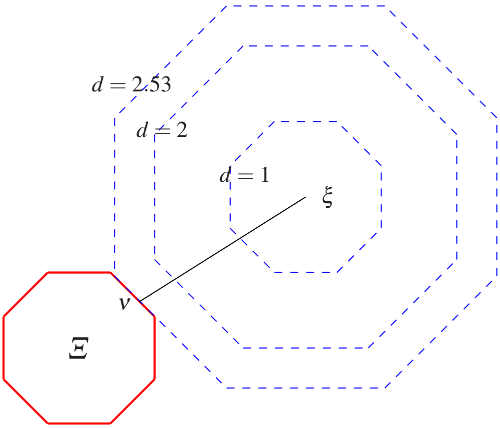

Fig. 3.4 Distance $d(\xi, \Xi)$ with the composite norm

Proposition 6. *Let $\tilde{a}^T x \leq b$ be a constraint with uncertain elements $\tilde{a} = \bar{a} + P\xi$. Let $\Xi = \{\xi \in \Omega \mid ||\xi||_1 \leq k_1, ||\xi||_\infty \leq k_\infty\}$ be the uncertainty set and let $d(\xi, \Xi) = \min_{\xi' \in \Xi} \delta(\xi - \xi')$ be the distance of a point ξ to Ξ, where $\delta(\mu) = \min_{t>0}\{t \mid \mu \in \Xi(t)\} = \max\{\frac{1}{k_1}||\mu||_1, \frac{1}{k_\infty}||\mu||_\infty\}$ is the norm induced by the polyhedron Ξ. Finally, let α be the coefficient of globalized robustness.*

The globalized robust equivalent of the globalized robust constraint

$$\bar{a}^T x + (P^T x)^T \xi \leq b + \alpha d(\xi, \Xi), \forall \xi \in \Omega$$

is given by the two constraints

$$\bar{a}^T x + k_1 ||P^T x - u||_\infty + k_\infty ||u||_1 \leq b \tag{26a}$$

$$k_1 ||P^T x - v||_\infty + k_\infty ||v||_1 \leq \alpha. \tag{26b}$$

Proof. Let us study the maximization problem yielding the worst case in the definition of the robust equivalent and show it to be a linear programming problem. First, let us introduce the auxiliary variable $\zeta \in \mathbb{R}$ and write the maximization problem yielding the globalized robust equivalent as

$$z^* = \max_{\xi, \xi', \zeta} \{(P^T x)^T \xi - \alpha \zeta \mid \delta(\xi - \xi') \leq \zeta,\ \zeta \geq 0,\ \xi' \in \Xi,\ \xi \in \Omega\}.$$

Note that we have added the redundant constraint $\zeta \geq 0$. The objective maximization ensures ζ to be as small as possible, enforcing the equality $\delta(\xi' - \xi) = \zeta$. Introducing the new variable $\eta = \xi - \xi'$ and recalling that the norm δ is the maximum of two norms, we obtain the alternative formulation

$$z^* = \max_{\xi',\eta,\zeta} \ (P^T x)^T (\xi' + \eta) - \alpha\zeta$$
$$||\eta||_\infty \leq k_\infty \zeta, \ ||\eta||_1 \leq k_1 \zeta, \ \zeta \geq 0$$
$$||\xi'||_\infty \leq k_\infty, \ ||\xi'||_1 \leq k_1.$$

For fixed ζ, the problem is separable in ξ' and η. By Lemma 2, we have

$$\begin{aligned} z^* &= \max_{\zeta \geq 0} \min_{u,v} \left\{ k_1 ||P^T x - u||_\infty + k_\infty ||u||_1 + \zeta (k_1 ||P^T x - v||_\infty + k_\infty ||v||_1 - \alpha) \right\} \\ &= \min_{u,v} \max_{\zeta \geq 0} \left\{ k_1 ||P^T x - u||_\infty + k_\infty ||u||_1 + \zeta (k_1 ||P^T x - v||_\infty + k_\infty ||v||_1 - \alpha) \right\}. \end{aligned}$$

The interchange between the max and min operators is allowed by the strong duality theorem in linear programming. The last expression can be simplified upon noticing that the inner maximization is either unbounded if $k_1||P^T x - v||_\infty + k_\infty||v||_1 > \alpha$, or takes the value $k_1||P^T x - u||_\infty + k_\infty||u||_1$ if $k_1||P^T x - v||_\infty + k_\infty||v||_1 \leq \alpha$ (because $\zeta = 0$ in that case). Thus

$$z^* = \min_{u,v} \{ k_1 ||P^T x - u||_\infty + k_\infty ||u||_1 \mid k_1 ||P^T x - v||_\infty + k_\infty ||v||_1 \leq \alpha \}.$$

As shown in Proposition 1, the above linear programming problem can be explicitly recast in terms of a linear objective and linear inequality constraints. Moreover, if feasible, it achieves its optimal value. This makes it possible to write the deterministic equivalent as

$$\bar{a}^T x + k_1 ||P^T x - u||_\infty + k_\infty ||u||_1 \leq b$$
$$k_1 ||P^T x - v||_\infty + k_\infty ||v||_1 \leq \alpha.$$

□

Remark 5. A theorem similar to Proposition 1 can be stated for the globalized robust equivalent (3.26) with norms l_1 and l_∞. We leave it to the reader to perform the simple transformations that lead to equivalent linear inequalities.

Remark 6. In the definition of the globalized robustness, we assumed that the set Ω of all possible realizations is $\mathbb{R}^m$. This makes the globalized robustness condition rather demanding. It may be worth considering that $\Omega \subset \mathbb{R}^m$ is constrained by box constraints, such as $\Omega = \{\xi \mid \underline{\xi} \leq \xi \leq \bar{\xi}\}$. In that case, the globalized robust equivalent calls for two additional nonnegative variables and the constraints

$$\bar{a}^T x + k_1 ||P^T x - u - w + t||_\infty + k_\infty ||u||_1 + \bar{\xi}^T w - \underline{\xi}^T t \leq b$$
$$k_1 ||P^T x - v - w + t||_\infty + k_\infty ||v||_1 \leq \alpha$$
$$w \geq 0, \ t \geq 0.$$

If we set $w = 0$ and $t = 0$, we retrieve (3.26), confirming that reducing the span of Ω enlarges the set of solutions.

Let us add few words of comments. Because of the additional constraint (26b), the set of globalized robust solutions is more restricted than the set of standard robust solutions. The globalized robustness uses the violation tolerance factor α. The smaller this parameter, the smaller must be the constraint violation. On the contrary, if α tends to infinity, constraint (26b) becomes inactive and we retrieve the plain robust solution.

In this presentation we used polyhedral uncertainty sets. Similar results can be obtained with ellipsoidal uncertainty sets. An interesting application to multi-echelon, multi-period inventory control is reported in (7).

3.9.3 Case study: Globalized robustness with uncertain demands

We now apply the concept of globalized robust optimization to the LDR formulation described in Subsection 3.6.1. We remind the reader that in Subsection 3.6.1 only the demands were uncertain and that the experiments were performed with different values of k_{dem}. Here we propose to test different values of α on the particular case $k_{dem} = 0.4$. We report the simulation results for each globalized robust solution in Table 3.18.

Table 3.18 Variable demand: behavior on the sample of 1000 scenarios of a globalized robust solution with LDR.

	LDR solutions with $k_{dem} = 0.4$		
	$\alpha = 0.3$	$\alpha = 0.2$	$\alpha = 0.1$
Predicted cost performance	165.84	166.26	166.82
Observed cost performance	165.07	165.71	166.62
Scenarios with demand violation(s) in %	20.4	16.0	1.1
Conditional average relative violation in %	0.8	0.6	0.4
Average number of violations per scenario	1	1	1
Total number of violated air quality constraints	69	69	69

We notice that for $\alpha = 0.3$ we retrieve the standard robust solution. Lower values of α permit a better control of the demand violations.

The model formulation has 433 constraints and 509 variables (419 variables and 502 constraints for the standard LDR formulation).

3.10 Conclusion

The main goal of this chapter was to present an alternative approach to the dealing with uncertainties in environmental and energy planning. To this end, we introduced basic concepts in robust optimization and we applied them to an illustrative example.

For the sake of the exposition, we chose an overly simplified example of modest size with three periods only. In practice, the models that are commonly used in the area of environmental and energy planning are much larger and much more complex, with thousands of variables and constraints. One should expect that the robust optimization methodology generates robust counterparts of much larger size, possibly to the point that they become impractical. This view must be qualified, by differentiating the type of uncertainties that are taken into consideration. The main difficulty stems from the multi-period feature of the models with uncertainties described as stochastic process. At this point, there is a sharp distinction to be made, depending on whether the realizations of the stochastic process are progressively revealed to the decision-maker, or only at once at the end of the terminal stage. If the exact value of the process remains unknown till the horizon of the model, the decision process is essentially static: all present and future decisions are to be taken at once, but the eventual effect of uncertainty is revealed in a second phase with no possibility of recourse. We treated an example of this sort in Section 3.3. The fact that the size of robust counterpart increases linearly with the dimension of the uncertainty factors suggests that, in this situation, even large models could be handled by robust optimization.

The situation is much more dramatic if the uncertainty is progressively revealed in time. Then, the decision process must be adapted to the revealed information, a fact that introduces a major, fundamental difficulty for all known methods, except in some special cases[4]. Linear decision rules for problems with fixed recourses may provide an acceptable approximation of an adaptive behavior and yield interesting hindsight, as those obtained in Section 3.6. We even suggested that robust optimization could be used in connection with traditional approaches to handle more complex cases. Other possibilities are to be considered. LDR is a crude approximation of the anticipation mechanism in the decision process. But is is already a highly complex one in regard of the actual decision process in the environmental and energy planning problems (and in many other problems!). Moreover, in an implementation phase of the model solution, the optimal LDR will almost surely never be implemented as such. Rather, at each time stage, a fresh version of the problem will be considered to account for the new situation. This leads to view the LDR not as a practical "open loop control", but like a tool yielding a plausible anticipation of future evolution and providing valuable information in the design of the first stage decisions.

An interesting alternative, though probably a computationally costly one, would be to work in a folding horizon framework and scenario simulations. Suppose a scenario is selected, which prescribes up to the horizon the realization of the various uncertainties. The decision-maker only knows the probabilistic way the scenario is built, but not the scenario itself. The planning model may have a rather rich description of uncertainty in its first stage, in a robust optimization framework, and a cruder description for the later stages based on a LDR. The first stage optimal solution of this model is implemented in connection with the realization of the stochastic pro-

[4] For instance, Markovian decision processes with reduced state space cardinality can be treated very efficiently by dynamic programming

cess in the first stage. This determines the state of the system at the beginning of the second stage, where a problem of the same nature as before, and a shorter horizon, is considered anew. This approach has been implemented in (2; 3; 6) and could prove interesting in more general models.

The secondary goal of this chapter was to give a simple enough introduction to the burgeoning field of robust optimization. We deliberately confined our presentation to the linear programming context, except for a few mention of ellipsoidal uncertainty sets. In the latter case, the robust counterparts fall in the realm of second order cone programming (SCOP) for which highly efficient solution methods exist. The striking fact is that in the case of ellipsoidal uncertainty sets the robust counterpart is obtained via similar duality-based arguments. Further generalizations can be made that take advantage of the power of conic programming with linear constraints and self-dual cones. Most recent contributions in robust optimization are concerned with these extensions and/or use them intensively. They are beyond the scope of this introduction, but they cannot be ignored. The hope is that this introduction will make it easier for the reader to get around with the growing literature.

Finally, we would like to mention that the probabilistic results in Section 3.8 have been considerably strengthened in some recent contributions. The crude hypotheses of Theorem 3 of random factors with zero mean and a symmetric range around the mean can be replaced by a range on the mean and a bound of the variance. Many other results of comparable nature are reported in (5). They give more evidences of the strong link between chance constrained programming and robust optimization.

Acknowledgements This work was partially supported by the European research project FP7-PLANETS

References

[1] E. Adida and G. Perakis. A robust optimization approach to dynamic pricing and inventory control with no backorders. *Mathematical Programming*, 107(1-2):97 – 129, 2006.

[2] R. Apparigliato. *Règles de décision pour la gestion du risque : Application à la gestion hebdomadaire de la production électrique*. PhD thesis, Ecole Polytechnique, Paris, 2008.

[3] R. Apparigliato, J.-P. Vial, and R. Zorgati. Optimisation robuste linéaire : application à la gestion court terme d'une vallée hydraulique. Working paper H-R32-2007-00658-FR, EDF R&D, Département Osiris, 1, avenue du Général de Gaulle, F-92141 Clamart, 2007.

[4] A. Ben-Tal, S. Boyd, and A. Nemirovski. Extending the scope of robust optimization: Comprehensive robust counterparts of uncertain problems. *Mathematical Programming Series B*, 107(1-2):63–89, 2006.

[5] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust Optimization*. Princeton University Press (Book under review), 2009.

[6] A. Ben-Tal, B. Golany, A. Nemirovski, and J.-Ph. Vial. Retailer-supplier flexible commitments contracts: a robust optimization approach. *Manufacturing and Service Operations Management*, 7:248–271, 2005.
[7] A. Ben-Tal, B. Golany, and S. Shtern. Robust multi-echelon, multi-period inventory control. Technical report, 2008. To appear in *European Journal of Operational Research*.
[8] A. Ben-Tal, A. Goryashko, E. Guslitzer, and A. Nemirovski. Adjustable robust solutions of uncertain linear programs. *Mathematical Programming*, 99(2):351–376, 2004.
[9] A. Ben-Tal and A. Nemirovski. Robust convex optimization. *Mathematics of Operations Research*, 23:769 – 805, 1998.
[10] A. Ben-Tal and A. Nemirovski. Robust solutions of linear programming problems contaminated with uncertain data. *Mathematical Programming Series A*, 88:411 – 424, 2000.
[11] A. Ben-Tal and A. Nemirovski. *Lectures on Modern Convex Optimization: Analysis, Algorithms and Engineering Applications*. MPS-SIAM series on Optimization. Society for Industrial and Applied Mathematics and Mathematical Programming Society, 2001.
[12] A. Ben-Tal and A. Nemirovski. Selected topics in robust convex optimization. Technical report, Faculty of Industrial Engineering and Management, Technion, Technion city, Haifa 32000, Israel, 2006. (To appear in *Mathematical Programming*).
[13] D.P. Bertsekas. *Dynamic Programming and Optimal Control*. Athena Scientific, Belmont, Mass., 1995.
[14] D. Bertsimas, D. Pachamanova, and M. Sim. Robust linear optimization under general norms. *Operations Research Letters*, 32:510–516, 2004.
[15] D. Bertsimas and M. Sim. Price of robustness. *Operations Research*, 52:35–53., 2004.
[16] D. Bertsimas and A. Thiele. A robust optimization approach to inventory theory. *Operations Research*, 54(1):150–168, 2006.
[17] D. Bienstock. Experiments in robust portfolio optimization. Technical report, Center for Financial Engineering, Columbia University, January 2007.
[18] J. R. Birge and F. Louveaux. *Introduction do stochastic programming*. Springer-Verlag New York, Inc., 1997.
[19] G. C. Calafiore. Ambiguous risk measures and optimal robust portfolios. Technical report (to appear in *SIAM Journal on Optimization* 2007), Dipartmento di Automatica e Informatica, Politecnico di Torino, Italy, 2007.
[20] G. C. Calafiore and L. El-Gahoui. On distributionally robust chance-constrained linear programs. *Journal of Optimization Theory and Applications*, 130:1–22, 2006.
[21] S. Ceria and R. Stubbs. Incorporating estimation errors into portfolio selection: Robust portfolio construction. *Journal of Asset Management*, 7(2):109 – 127, 2006.
[22] A. Charnes and W.W. Cooper. Chance constrained programming. *Management Science*, 6:73–89, 1959.

[23] A. Charnes, W.W. Cooper, and G.H. Symonds. Cost horizons and certainty equivalents: an approach to stochastic programming of heating oil. *Management Science*, 4:235–263, 1958.

[24] X. Chen, M. Sim, and P. Sun. A robust optimization perspective of stochastic programming. Working paper, NUS Business school, 2005.

[25] G. B. Dantzig and A. Madansky. On the solution of two-stage linear programs under uncertainty. In Jerzy Neyman, editor, *Fourth Berkeley Symposium on Mathematical Statistics and Probability (June 20-July 30, 1960)*, volume 1, pages 165–176. University of California Press, 1961.

[26] G.B. Dantzig. Linear programming under uncertainty. *Management Science*, 1(3-4):197–206, 1956.

[27] Fragnière E. and Haurie A. Markal-geneva: a model to assess energy-environment choices for a swiss canton. In *Carraro C, Haurie A (eds) Operations research and environmental management*. Kluwer Academic Books, Dordrecht, 1996.

[28] L. El Ghaoui and H. Lebret. Robust solutions to least-square problems to uncertain data matrices. *SIAM Journal of Matrix Analysis and Applications*, 18:1035–1064, 1997.

[29] S. J. Garstka and R. J.-B. Wets. On decision rules in stochastic programming. *Mathematical Programming*, 7:117–143, 1974.

[30] D. Goldfarb and G. Iyengar. Robust portfolio selection problems. *Mathematics of Operations Research*, 28:1 – 38, 2003.

[31] C.C. Holt, F. Modigliani, and H.A. Simon. A linear decision rule for production and employment scheduling. *Management Science*, 2(1):1–30, 1955.

[32] G. Iyengar. Robust dynamic programming. *Mathematics of Operations Research*, 30:257 – 280, 2005.

[33] P. Kall and S. W. Wallace. *Stochastic Programming*, volume 2nd edition. John Wiley & Sons, Chichester, 1994.

[34] L.B. Miller and H. Wagner. Chance-constrained programming with joint constraints. *Operations Research*, 13, 1965.

[35] A. Ouorou and J.-P. Vial. A model for robust capacity planning for telecommunications networks under demand uncertainty. In *Proceedings of the 6th International Workshop on Design and Reliable Communication Networks, DRCN 2007*, 2007.

[36] A. Prékopa. On probabilistic constrained programming. In *Proceedings of the Princeton Symposium on Mathematical Programming*, pages 113–138. Princeton University Press, Princeton, 1970.

[37] A. L. Soyster. Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations Research*, 21:1154–1157, 1973.

[38] J.F. Sturm. Using SeDuMi 1.02, a Matlab Toolbox for Optimization over Symmetric Cones (updated for version 1.05). Technical report, 2001.