Chapter 2

Uncertain Linear Optimization Problems and their Robust Counterparts

In this Chapter, we introduce the concept of uncertain Linear Optimization problem and its Robust Counterpart.

2.1 Data uncertainty in Linear Optimization

Recall that a *Linear Optimization* (LO) *problem* is of the form

$$\min_x \left\{ c^T x + d : Ax \leq b \right\}, \tag{2.1}$$

where $x \in \mathbb{R}^n$ is the vector of *decision variables*, $c \in \mathbb{R}^n$ and $d \in \mathbb{R}$ form the *objective*, A is an $m \times n$ *constraint matrix*, and $b \in \mathbb{R}^m$ is the *right hand side vector*.

> Clearly, the constant term d in the objective, while affecting the optimal value, does not affect the optimal solution, this is why it traditionally is skipped. As we shall see, when treating the LO problems with *uncertain data* there are good reasons not to neglect this constant term.

The *structure* of problem (2.1) is given by the number m of constraints and the number n of variables, while the *data* of the problem is the collection (c, d, A, b) which we will arrange into $(m+1) \times (n+1)$ *data matrix*

$$D = \left[\begin{array}{c|c} c^T & d \\ \hline A & b \end{array} \right].$$

Usually not all constraints of an LO program, as it arises in applications, are of the form $a^T x \leq$ const; there can be linear "$\geq$"-inequalities and linear equalities as well. Clearly, the constraints of the latter two types can be represented equivalently by linear "$\leq$"-inequalities, and we will assume henceforth that these are the only constraints in the problem.

The data of real world LO's (Linear Optimization problems) is typically not known exactly when the problem is to be solved. The most common reasons for data uncertainty are as follows:

- Some of data entries (future demands, returns, etc.) do not exist when the problem is solved and hence are replaced with their forecasts. These data entries are thus subject to *prediction errors*;
- Some of the data (parameters of technological devices and processes, contents associated with raw materials, etc.) cannot be measured exactly – in reality their values drift around the measured "nominal" values; these data are subject to *measurement errors*;
- Some of the decision variables (intensities with which we intend to use various technological processes, parameters of physical devices we are designing, etc.) cannot be implemented exactly as computed. The resulting *implementation errors* are equivalent to appropriate artificial data uncertainties.

  Indeed, the contribution of a particular decision variable x_j to the left hand side of constraint i is the product $a_{ij}x_j$. Hence the consequences of an additive implementation error $x_j \mapsto x_j + \epsilon$ are as if there were no implementation error at all, but the left hand side of the constraint got an extra additive term $a_{ij}\epsilon$, which, in turn, is equivalent to the perturbation $b_i \mapsto b_j - a_{ij}\epsilon$ in the right hand side of the constraint. The consequences of a more typical *multiplicative* implementation error $x_j \mapsto (1+\epsilon)x_j$ are as if there were no implementation error, but each of the data coefficients a_{ij} was subject to perturbation $a_{ij} \mapsto (1+\epsilon)a_{ij}$. Similarly, the influence of additive and multiplicative implementation error in x_j on the value of the objective can be mimicked by appropriate perturbations in d or c_j.

In the traditional LO methodology, a small data uncertainty (say, 1% or less) is just ignored; the problem is solved *as if* the given ("nominal") data were exact, and the resulting *nominal* optimal solution is what is recommended for use, in hope that small data uncertainties will not affect significantly feasibility and optimality properties of this solution, or that small

Parameter	DrugI	DrugII
Selling price, \$ per 1000 packs	6,200	6,900
Content of agent A, g per 1000 packs	0.500	0.600
Manpower required, hours per 1000 packs	90.0	100.0
Equipment required, hours per 1000 packs	40.0	50.0
Operational costs, \$ per 1000 packs	700	800

(a) Drug production data

Raw material	Purchasing price, \$ per kg	Content of agent A, g per kg
RawI	100.00	0.01
RawII	199.90	0.02

(b) Contents of raw materials

Budget, \$	Manpower, hours	Equipment, hours	Capacity of raw materials storage, kg
100,000	2,000	800	1,000

(c) Resources

Table 2.1: Data for Example 2.1.

adjustments of the nominal solution will be sufficient to make it feasible. We are about to demonstrate that these hopes are not necessarily justified, and sometimes even small data uncertainty deserves significant attention.

Introductory example

Consider a toy linear optimization problem as follows:

Example 2.1 A company produces two kinds of drugs, DrugI and DrugII, containing a specific active agent A, which is extracted from raw materials purchased on the market. There are two kinds of raw materials, RawI and RawII, which can be used as sources of the active agent. The related production, cost and resource data are given in Table 2.1. The goal is to find the production plan which maximizes the profit of the company.

The problem can be immediately posed as the following linear programming program:

(Drug):

$$\begin{array}{rcl}
\text{Opt} & = & \min\Big\{ \overbrace{[100 \cdot RawI + 199.90 \cdot RawII + 700 \cdot DrugI + 800 \cdot DrugII]}^{\text{purchasing and operational costs}} \\
& & - \underbrace{[6200 \cdot DrugI + 6900 \cdot DrugII]}_{\text{income from selling the drugs}} \Big\} \\
& & \text{[minus total profit]}
\end{array}$$

subject to

$$\begin{array}{rcl}
0.01 \cdot RawI + 0.02 \cdot RawII - 0.500 \cdot DrugI - 0.600 \cdot DrugII & \geq & 0 \\
& & \text{[balance of active agent]} \\
RawI + RawII & \leq & 1000 \\
& & \text{[storage restriction]} \\
90.0 \cdot DrugI + 100.0 \cdot DrugII & \leq & 2000 \\
& & \text{[manpower restriction]} \\
40.0 \cdot DrugI + 50.0 \cdot DrugII & \leq & 800 \\
& & \text{[equipment restriction]} \\
100.0 \cdot RawI + 199.90 \cdot RawII + 700 \cdot DrugI + 800 \cdot DrugII & \leq & 100000 \\
& & \text{[budget restriction]} \\
RawI, RawII, DrugI, DrugII & \geq & 0
\end{array}$$

The problem has four variables – the amounts *RawI*, *RawII* (in kg) of raw materials to be purchased and the amounts *DrugI*, *DrugII* (in 1000 of packs) of drugs to be produced.

The optimal solution of our LO problem is

$$\text{Opt} = -8819.658;\, RawI = 0, RawII = 438.789, DrugI = 17.552, DrugII = 0.$$

Note that both the budget restriction and the balance constraint are active (that is, the production process utilizes the entire 100,000 budget and the full amount of active agent contained in the raw materials). The solution promises the company modest, but quite respectful profit 8.8%.

Data uncertainty and its consequences. Clearly, even in our simple problem some of the data cannot be "absolutely reliable"; e.g., one can hardly believe that the contents of the active agent in the raw materials are exactly 0.01 g/kg for RawI and 0.02 g/kg for RawII. In reality, these contents vary around the indicated values. A natural assumption here is that the actual contents of active agent *aI* in RawI and *aII* in RawII are realizations of random variables somehow distributed around the "nominal contents" $anI = 0.01$ and $anII = 0.02$. To be more specific, assume that *aI* drifts in a 0.5%-margin of *anI*, thus taking values in the segment $[0.00995, 0.01005]$. Similarly, assume that *aII* drifts in the 2% margin of *anII*, thus taking values in the segment $[0.0196, 0.0204]$. Moreover, assume that *aI*, *aII* take, with probabilities 0.5, extreme values in the respective segments. How do these perturbations of the contents of the active agent affect the production process? The optimal solution prescribes to purchase 438.8 kg of RawII and to produce 17552 packs of DrugI. With the above random fluctuations in the content of the active agent in RawII, this production plan, with probability 0.5, will be infeasible, i.e., the actual content of active agent in raw materials will be less than the one required to produce the planned amount of DrugI. For the sake of simplicity, assume that this difficulty is resolved in the simplest way: when the actual content of active agent in raw materials is insufficient, the output of the drug is reduced accordingly. With this policy, the actual production of DrugI becomes random variable which takes with equal probabilities, the nominal value of 17552 packs and the by 2% less value of 17201 packs. These 2% fluctuations in the production affect the profit as well; it becomes a random variable taking, with probabilities 0.5, the nominal value 8,820 and the 21% (!) less value 6,929. The expected profit is 7,843, which is by 11% less than the nominal profit 8,820 promised by the optimal solution of the nominal problem.

We see that in our toy example pretty small (and unavoidable in reality) perturbations of the data may make the nominal optimal solution infeasible. Moreover, a straightforward adjustment of the nominally optimal solution to the actual data may heavily affect solution's quality.

Similar phenomenon can be met in many practical linear programs where at lest part of the data are not known exactly and can vary around their nominal values. The consequences of data uncertainty can be much more severe than in our toy example. The analysis of linear optimization problems from the NETLIB collection[1] reported in [3] reveals that for 13 of 94 NETLIB problems, already 0.01%-perturbations of "clearly uncertain" data can make the nominal optimal solution severely infeasible: with a non-negligible probability, it violates some of the constraints by 50% and more. It should be added that in the general case (in contrast to our toy example) there is no evident way to adjust the optimal solution to the actual values of the data by a small modification, and there are cases when such an adjustment is in fact impossible - in order to become feasible for the perturbed data, the nominal optimal solution should be "completely reshaped".

[1] A collection over 100 LP programs, mainly of real world origin, used a standard benchmark for testing LP solvers.

The conclusion is as follows:

> *In applications of LO, there exists a real need of a technique capable of detecting cases when data uncertainty can heavily affect the quality of the nominal solution, and in these cases to generate a "reliable" solution, one which is immunized against uncertainty.*

We are about to introduce the *Robust Counterpart* approach to uncertain LO aimed at coping with data uncertainty.

2.2 Uncertain Linear Programs and their Robust Counterparts

Definition 2.1 *An uncertain Linear Optimization problem is a collection*

$$\left\{\min_x \left\{c^T x + d : Ax \leq b\right\}\right\}_{(c,d,A,b)\in\mathcal{U}} \tag{$LO_{\mathcal{U}}$}$$

of LO problems (instances) $\min_x \left\{c^T x + d : Ax \leq b\right\}$ *of common structure* (i.e., with common numbers m of constraints and n of variables) *with the data varying in a given* <u>*uncertainty set*</u> $\mathcal{U} \subset \mathbb{R}^{(m+1)\times(n+1)}$.

We always assume that the uncertainty set is parameterized, in an affine fashion, by *perturbation vector* ζ varying in a given *perturbation set* $\mathcal{Z}$:

$$\mathcal{U} = \left\{ \left[\begin{array}{c|c} c^T & d \\ \hline A & b \end{array}\right] = \underbrace{\left[\begin{array}{c|c} c_0^T & d_0 \\ \hline A_0 & b_0 \end{array}\right]}_{\substack{\text{nominal} \\ \text{data } D_0}} + \sum_{\ell=1}^{L} \zeta_\ell \underbrace{\left[\begin{array}{c|c} c_\ell^T & d_\ell \\ \hline A_\ell & b_\ell \end{array}\right]}_{\substack{\text{basic} \\ \text{shifts } D_\ell}} : \zeta \in \mathcal{Z} \subset \mathbb{R}^L \right\}. \tag{2.2}$$

For example, the story told in Section 2.1 makes (Drug) an uncertain LO problem as follows:

- *Decision vector:*

$$x = [\mathit{RawI}; \mathit{RawII}; \mathit{DrugI}; \mathit{DrugII}];$$

- *Nominal data:*

$$D_0 = \left[\begin{array}{cccc|c} 100 & 199.9 & -5500 & -6100 & 0 \\ \hline -0.01 & -0.02 & 0.500 & 0.600 & 0 \\ 1 & 1 & 0 & 0 & 1000 \\ 0 & 0 & 90.0 & 100.0 & 2000 \\ 0 & 0 & 40.0 & 50.0 & 800 \\ 100.0 & 199.9 & 700 & 800 & 100000 \\ -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \end{array}\right]$$

- *Two basic shifts:*

$$D_1 = 5.0\text{e-}5\cdot \left[\begin{array}{cccc|c} 0 & 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right], \quad D_2 = 4.0\text{e-}4\cdot \left[\begin{array}{cccc|c} 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

- *Perturbation set:*

$$\mathcal{Z} = \left\{\zeta \in \mathbb{R}^2 : -1 \leq \zeta_1, \zeta_2 \leq 1\right\}.$$

This description says, in particular, that the only uncertain data in (Drug) are the coefficients *anI*, *anII* of the variables *RawI*, *RawII* in the balance inequality (which is the first constraint in (Drug)), and that these coefficients vary in the respective segments $[0.01\cdot(1-0.005), 0.01\cdot(1+0.005)]$, $[0.02\cdot(1-0.02), 0.02\cdot(1+0.02)]$ around the nominal values 0.01, 0.02 of the coefficients – which is exactly what was said by words in Section 2.1.

Remark 2.1 *If the perturbation set $\mathcal{Z}$ in (2.2) is itself represented as the image of another set $\widehat{\mathcal{Z}}$ under affine mapping $\xi \mapsto \zeta = p + P\xi$, then we can pass from perturbations ζ to perturbations ξ:*

$$\begin{array}{rcl}
\mathcal{U} & = & \left\{\left[\begin{array}{c|c} c^T & d \\ \hline A & b \end{array}\right] = D_0 + \sum\limits_{\ell=1}^{L} \zeta_\ell D_\ell : \zeta \in \mathcal{Z}\right\} \\
& = & \left\{\left[\begin{array}{c|c} c^T & d \\ \hline A & b \end{array}\right] = D_0 + \sum\limits_{\ell=1}^{L} [p_\ell + \sum\limits_{k=1}^{K} P_{\ell k}\xi_k] D_\ell : \xi \in \widehat{\mathcal{Z}}\right\} \\
& = & \left\{\left[\begin{array}{c|c} c^T & d \\ \hline A & b \end{array}\right] = \underbrace{\left[D_0 + \sum\limits_{\ell=1}^{L} p_\ell D_\ell\right]}_{\widehat{D}_0} + \sum\limits_{k=1}^{K} \xi_k \underbrace{\left[\sum\limits_{\ell=1}^{L} P_{\ell k} D_\ell\right]}_{\widehat{D}_\ell} : \xi \in \widehat{\mathcal{Z}}\right\}.
\end{array}$$

It follows that when speaking about perturbation sets with simple geometry (parallelotopes, ellipsoids, etc.), we can normalize these sets to be "standard". E.g., a parallelotope is by definition an affine image of a unit box $\{\xi \in \mathbb{R}^k : -1 \leq \xi_j \leq 1, j = 1, ..., k\}$, *which gives us the possibility to work with the unit box instead of general parallelotope. Similarly, an ellipsoid is by definition the image of a unit Euclidean ball* $\{\xi \in \mathbb{R}^k : \|x\|_2^2 \equiv x^T x \leq 1\}$ *under affine mapping, so that we can work with the standard ball instead of ellipsoid, etc. We will use this normalization whenever possible.*

Note that a *family* of optimization problems like (LO$_{\mathcal{U}}$), in contrast to a single optimization problem, is not associated by itself with the concepts of feasible/optimal solution and optimal value. How to define these concepts, it depends of course on the underlying "decision environment". Here we focus on the environment characterized by the following assumptions:

A.1. All decision variables in (LO$_{\mathcal{U}}$) represent "here and now" decisions; they should get specific numerical values as a result of solving the problem *before* the actual data "reveals itself".

A.2. The decision maker is fully responsible for consequences of the decisions to be made when, and only when, the actual data is within the prespecified uncertainty set $\mathcal{U}$ given by (2.2).

A.3. The constraints in (LO$_{\mathcal{U}}$) are "hard" – we cannot tolerate violations of constraints, even small ones, when the data is in $\mathcal{U}$.

The above assumptions determine, in a more or less unique fashion, what are the meaningful feasible solutions to the uncertain problem (LO$_{\mathcal{U}}$). By A.1, these should be fixed vectors; by A.2 and A.3, they should be *robust feasible* – should satisfy all the constraints, whatever be a realization of the data from the uncertainty set. We have arrived at the following definition.

Definition 2.2 *A vector* $x \in \mathbb{R}^n$ *is a* <u>robust feasible</u> *solution to* (LO$_{\mathcal{U}}$), *if it satisfies all realizations of the constraints from the uncertainty set, that is,*

$$Ax \leq b \quad \forall (c, d, A, b) \in \mathcal{U}. \tag{2.3}$$

As about the objective value to be associated with a meaningful (i.e., robust feasible) solution, assumptions A.1 – A.3 do not prescribe it in a unique fashion. However, "the spirit" of these worst-case-oriented assumptions make natural the following definition:

Definition 2.3 *Given a candidate solution x, the* <u>*robust*</u> *value $\widehat{c}(x)$ of the objective in* $(\mathrm{LO}_{\mathcal{U}})$ *at x is the largest value of the "true" objective $c^Tx + d$ over all realizations of the data from the uncertainty set:*

$$\widehat{c}(x) = \sup_{(c,d,A,b)\in\mathcal{U}} [c^Tx + d]. \tag{2.4}$$

After we agree what are meaningful candidate solutions to the uncertain problem $(\mathrm{LO}_{\mathcal{U}})$ and how to quantify their quality, we can seek for the best, in terms of the robust value of the objective, among all robust feasible solutions to the problem. We have arrived at the central for us concept of *Robust Counterpart* of uncertain optimization problem as follows:

Definition 2.4 *The Robust Counterpart of the uncertain LO problem* $(\mathrm{LO}_{\mathcal{U}})$ *is the optimization problem*

$$\min_x \left\{ \widehat{c}(x) = \sup_{(c,d,A,b)\in\overline{c}U} [c^Tx + d] : Ax \le b \ \forall (c,d,A,b) \in \mathcal{U} \right\} \tag{2.5}$$

of minimizing the robust value of the objective over all robust feasible solutions to the uncertain problem.

An optimal solution to the Robust Counterpart is called robust optimal solution to $(\mathrm{LO}_{\mathcal{U}})$*, and the optimal value of the Robust Counterpart is called the robust optimal value of* $(\mathrm{LO}_{\mathcal{U}})$*.*

In a nutshell, the robust optimal solution is simply "the best uncertainty-immunized" solution we can associate with our uncertain problem, and this is the solution to be actually used.

Example 2.1 [continued] Let us find the robust optimal solution to the uncertain problem (Drug). There is exactly one uncertainty-affected "block" in the data, namely, the coefficients of *RawI*, *RawII* in the balance constraint. A candidate solution is thus robust feasible if and only if it satisfies all constraints of (Drug), except for the balance one, as they are, *and* satisfies the "worst" realization of the balance constraint. Since *RawI*, *RawII* are nonnegative, the worst realization of the balance constraint is the one where the uncertain coefficients *anI*, *anII* are set to their minimal values allowed by the uncertainty set (these values are 0.00995 and 0.0196, respectively). Since the objective is not affected by the uncertainty, the robust objective values are the same as the original ones. Thus, the RC (Robust Counterpart) of our uncertain problem is the LO problem

$$\begin{array}{l}
\text{RC(Drug):}\\
\text{RobOpt} = \min\left\{-100 \cdot \mathit{RawI} - 199.9 \cdot \mathit{RawII} + 5500 \cdot \mathit{DrugI} + 6100 \cdot \mathit{DrugII}\right\}\\
\text{subject to}\\
\begin{array}{rcl}
0.00995 \cdot \mathit{RawI} + 0.0196 \cdot \mathit{RawII} - 0.500 \cdot \mathit{DrugI} - 0.600 \cdot \mathit{DrugII} & \geq & 0\\
\mathit{RawI} + \mathit{RawII} & \leq & 1000\\
90.0 \cdot \mathit{DrugI} + 100.0 \cdot \mathit{DrugII} & \leq & 2000\\
40.0 \cdot \mathit{DrugI} + 50.0 \cdot \mathit{DrugII} & \leq & 800\\
100.0 \cdot \mathit{RawI} + 199.90 \cdot \mathit{RawII} + 700 \cdot \mathit{DrugI} + 800 \cdot \mathit{DrugII} & \leq & 100000\\
\mathit{RawI}, \mathit{RawII}, \mathit{DrugI}, \mathit{DrugII} & \geq & 0
\end{array}
\end{array}$$

Solving this problem, we get

$$\text{RobOpt} = -8294.567; \mathit{RawI} = 877.732, \mathit{RawII} = 0, \mathit{DrugI} = 17.467, \mathit{DrugII} = 0.$$

The "price" of robustness is the reduction in the promised profit from its nominal optimal value 8819.658 to its robust optimal value 8294.567, that is, by 5.954%. This is much less than the reduction of the actual profit to 7,843 (by 11%) which we may suffer when sticking to the nominal optimal solution when the "true" data are "against" it. Note also that the structure of the robust optimal solution is quite different from the one of the nominal optimal solution: with the robust solution, we shall buy only raw materials RawI, while with the nominal one – only raw materials RawII. The explanation is clear: with the nominal data, RawII as compared to RawI results in a bit smaller per unit price of the active agent (9,995 \$/g vs. 10,000 \$/g), this is why with the nominal data, it does not make sense to use RawI. The robust optimal solution takes into account that "uncertainty" in *anI* (i.e., variability of contents of active agent in RawI) is 4 times smaller than that of *anII* (0.5% vs. 2%), which ultimately makes it better to use RawI.

More o Robust Counterparts

We start with several useful observations.

A. The Robust Counterpart (2.5) of (LO$\mathcal{U}$) can be rewritten equivalently as the problem

$$\min_{x,t}\left\{t : \left.\begin{array}{rcl} c^T x - t & \leq & -d \\ Ax & \leq & b \end{array}\right\} \forall (c,d,A,b) \in \mathcal{U}\right\}. \tag{2.6}$$

Note that we can arrive at this problem in another fashion: we first introduce extra variable t and rewrite instances of our uncertain problem (LO$_{\mathcal{U}}$)

equivalently as

$$\min_{x,t}\left\{t: \begin{array}{rcl} c^T x - t & \leq & -d \\ Ax & \leq & b \end{array}\right\},$$

thus arriving at an equivalent to $(\mathrm{LO}_{\mathcal{U}})$ uncertain problem in variables x, t with the objective t which is not affected by uncertainty at all. We now can build the Robust Counterpart of the resulting uncertain problem, and the RC of the reformulated problem is exactly (2.6). We see that

> *An uncertain LO problem always can be reformulated as an uncertain LO problem with certain objective. The Robust Counterpart of the reformulated problem has the same objective as this problem and is equivalent to the RC of the original uncertain problem.*

As a consequence, *we lose nothing when restricting ourselves with uncertain LO programs with certain objective*, and we shall frequently use this option in the sequel.

B. Assuming that $(\mathrm{LO}_{\mathcal{U}})$ is with certain objective, the Robust Counterpart of the problem is

$$\min_x \left\{c^T x + d : Ax \leq b,\ \forall (A, b) \in \mathcal{U}\right\} \tag{2.7}$$

(note that the uncertainty set is now a set in the space of the constraint data $[A, b]$). We see that

> *The Robust Counterpart of uncertain LO problem with certain objective is purely "constraint-wise" construction: to get RC, we*
>
> - *preserve the original certain objective as it is, and*
> - *replace every one of the original constraints*
>
> $$(Ax_i) \leq b_i \Leftrightarrow a_i^T x \leq b_i \tag{C$_i$}$$
>
> *(a_i^T is i-th row in A) with its Robust Counterpart*
>
> $$a_i^T x \leq b_i \ \forall [a_i; b_i] \in \mathcal{U}_i \tag{RC(C$_i$)}$$
>
> *where $\mathcal{U}_i$ is the projection of $\mathcal{U}$ on the space of data of i-th constraint:*
>
> $$\mathcal{U}_i = \{[a_i; b_i] : [A, b] \in \mathcal{U}\}.$$

In particular,

> The RC of uncertain LO problem with certain objective remains intact when the original uncertainty set $\mathcal{U}$ is extended to the direct product
> $$\widehat{\mathcal{U}} = \mathcal{U}_1 \times ... \times \mathcal{U}_m$$
> of its projections onto the spaces of data of respective constraints.

Example 2.2 The RC of the system of uncertain constraints

$$\{x_1 \geq \zeta_1,\, x_2 \geq \zeta_2\} \tag{2.8}$$

with $\zeta \in \mathcal{U} := \{\zeta_1 + \zeta_2 \leq 1, \zeta_1, \zeta_2 \geq 0\}$ is the infinite system of constraints

$$x_1 \geq \zeta_1,\, x_1 \geq \zeta_2 \;\; \forall \zeta \in \mathcal{U};$$

on variables x_1, x_2. The latter system is clearly equivalent to the pair of constraints

$$x_1 \geq \max_{\zeta \in \mathcal{U}} \zeta_1 = 1,\;\; x_2 \geq \max_{\zeta \in \mathcal{U}} \zeta_2 = 1. \tag{2.9}$$

The projections of $\mathcal{U}$ to the spaces of data of the two uncertain constraints (2.8) are the segments $\mathcal{U}_1 = \{\zeta_1 : 0 \leq \zeta_1 \leq 1\}$, $\mathcal{U}_2 = \{\zeta_2 : 0 \leq \zeta_2 \leq 1\}$, and the RC of (2.8) w.r.t. the uncertainty set $\widehat{\mathcal{U}} = \mathcal{U}_1 \times \mathcal{U}_2 = \{\zeta \in \mathbb{R}^2 : 0 \leq \zeta_1, \zeta_2 \leq 1\}$ clearly is (2.9).

> The conclusion we have arrived at seems to be counter-intuitive: it says that it is immaterial whether the perturbations of data in different constraints are or are not linked to each other, while intuition says that such a link should be important. We shall see later that this intuition makes sense when a more advanced concept of *Adjustable* Robust Counterpart is considered.

C. If x is a robust feasible solution of (C_i), then x remains robust feasible when we extend the uncertainty set $\mathcal{U}_i$ to its convex hull $\mathrm{Conv}(\mathcal{U}_i)$. Indeed, if $[\bar{a}_i; \bar{b}_i] \in \mathrm{Conv}(\mathcal{U}_i)$, then

$$[\bar{a}_i; \bar{b}_i] = \sum_{j=1}^{J} \lambda_j [a_i^j; b_i^j]$$

with appropriately chosen $[a_i^j; b_i^j] \in \mathcal{U}_i$, $\lambda_j \geq 0$ such that $\sum\limits_j \lambda_j = 1$. We now have

$$\bar{a}_i^T x = \sum_{j=1}^{J} \lambda_j [a_i^j]^T x \underbrace{\leq}_{(a)} \sum_j \lambda_j b_i^j = b_i$$

where the concluding inequality if given by the fact that x is feasible for $\mathrm{RC}(\mathrm{C}_i)$ and $[a_i^j; b_i^j] \in \mathcal{U}_i$. We see that $\bar{a}_i^T x \leq \bar{b}_i$ for all $[\bar{a}_i; \bar{b}_i] \in \mathrm{Conv}(\mathcal{U}_i)$, Q.E.D.

By similar reasons, the set of robust feasible solutions to (C_i) remains intact when we extend $\mathcal{U}_i$ to the closure of this set. Combining these observations with **B.**, we arrive at the following conclusion:

The Robust Counterpart of uncertain LO problem with certain objective remains intact when we extend the sets $\mathcal{U}_i$ of uncertain data of respective constraints to their closed convex hulls, and extend $\mathcal{U}$ to the direct product of the resulting sets.

In other words, we lose nothing when assuming from the very beginning that the sets $\mathcal{U}_i$ of uncertain data of constraints are closed and convex, and $\mathcal{U}$ is the direct product of these sets.

In terms of the parameterization (2.2) of the uncertainty sets, the latter conclusion means that

When speaking about Robust Counterpart of uncertain LO problem with certain objective, we lose nothing when assuming that the set $\mathcal{U}_i$ of uncertain data of i-th constraint is given as

$$\mathcal{U}_i = \left\{ [a_i; b_i] = [a_i^0; b_i^0] + \sum_{\ell=1}^{L_i} \zeta_\ell [a_i^\ell; b_i^\ell] : \zeta \in \mathcal{Z}_i \right\} \tag{2.10}$$

with closed and convex perturbation set $\mathcal{Z}_i$.

What is ahead. After introducing the concept of Robust Counterpart of an uncertain LO problem, we arrive at two major questions as follows:

1. What is the "computational status" of the RC? When it is possible to process RC efficiently?
2. How to define a meaningful uncertainty set?

The first of these questions, to be addressed in-depth in Section 2.3, is a "structural" one: what should be the structure of the uncertainty set in order to make the RC computationally tractable? Note that the RC as given by (2.6) or (2.7) is a *semi-infinite* LO program, that is, an optimization program with simple linear objective and *infinitely many* linear constraints.

In principle, such a problem can be "computationally intractable" – NP-hard.

Example 2.3 Consider an uncertain "nearly linear" constraint

$$\{\|Px - p\|_1 \le 1\}_{[P;p]\in\mathcal{U}}\,, \tag{2.11}$$

where $\|z\|_1 = \sum\limits_j |z_j|$, and assume that the matrix P is certain, while the vector p is uncertain and is parameterized by perturbations from the unit box:

$$p \in \{p = B\zeta : \|\zeta\|_\infty \le 1\}\,,$$

where $\|\zeta\|_\infty = \max\limits_\ell |\zeta_\ell|$ and B is a given positive semidefinite matrix. To check whether $x = 0$ is robust feasible is exactly the same as to verify whether $\|B\zeta\|_1 \le 1$ whenever $\|\zeta\|_\infty \le 1$, or, due to the evident relation $\|u\|_1 = \max\limits_{\|\eta\|_\infty\le 1} \eta^T u$, the same as to check, given B, whether $\max\limits_{\eta,\zeta}\left\{\eta^T B\zeta : \|\eta\|_\infty \le 1, \|\zeta\|_\infty \le 1\right\} \le 1$. The maximum of the bilinear form $\eta^T B\zeta$ with positive semidefinite B over η, ζ varying in a convex symmetric neighborhood of the origin is always achieved when $\eta = \zeta$ (you may check it by using the polarization identity $\eta^T B\zeta = \frac{1}{4}(\eta+\zeta)^T B(\eta+\zeta) + \frac{1}{4}(\eta-\zeta)^T B(\eta-\zeta)$). Thus, to check whether $x = 0$ is robust feasible for (2.11) is the same as to check whether the maximum of a given nonnegative quadratic form $\zeta^T B\zeta$ over the unit box is ≤ 1. The latter problem is known to be NP-hard[2], and therefore so is the problem of checking robust feasibility for (2.11).

The second of the above questions is a modelling one, and as such, goes beyond the scope of purely theoretical considerations. However, theory, as we shall see in Section 3.1, allows to contribute significantly to this modelling issue.

2.3 Tractability of Robust Counterpart

In this Section, we investigate the "computational status" of the RC of uncertain LO problem. The situation here turns out to be as nice as it could be: we shall see, essentially, that *the RC of uncertain LO problem with uncertainty set $\mathcal{U}$ is computationally tractable whenever the convex uncertainty set $\mathcal{U}$ is computationally tractable.* This can be reformulated as a precise mathematical statement; we, however, will prove a slightly restricted version of this statement which does not require long excursions into complexity theory.

[2] In fact, it is NP-hard to compute the maximum of a nonnegative quadratic form over the unit box with accuracy like 4% [22].

2.3.1 The strategy

Our strategy will be as follows. First, we can restrict ourselves with uncertain LO problems with certain objective – we remember from item **A** in Section 2.2 that we lose nothing when assuming the objective certain. Second, all we need is a "computationally tractable" representation of the RC of a *single* uncertain linear constraint, that is, an equivalent representation of such a RC by an explicit (and "short") system of explicit convex inequalities. Given such representations for the RC's of every one of the constraints of our uncertain problem and putting them together (cf. item **B** in Section 2.2), we reformulate the RC of the problem as the problem of minimizing the original linear objective under a finite (and short) system of explicit convex constraints, and thus – as a computationally tractable problem.

To proceed, we should explain first what does it mean "an equivalent representation of a constraint by a system of convex inequalities". Every one understands that the system of 4 constraints on 2 variables

$$x_1 + x_2 \leq 1, x_1 - x_2 \leq 1, -x_1 + x_2 \leq 1, -x_1 - x_2 \leq 1 \tag{2.12}$$

"represents equivalently" the nonlinear inequality

$$|x_1| + |x_2| \leq 1 \tag{2.13}$$

– both (2.13) and (2.12) define the same feasible set. Well, what about the claim that the system of 5 linear inequalities

$$-u_1 \leq x_1 \leq u_1, -u_2 \leq x_2 \leq u_2, u_1 + u_2 \leq 1 \tag{2.14}$$

represents the same set as (2.13)? Here again every one will agree with the claim, although we cannot justify the claim in the former fashion: the feasible sets of (2.13) and (2.14) live in different spaces and therefore cannot be equal to each other!

What actually is meant when speaking about "equivalent representations of problems/constraints" in Optimization can be formalized as follows:

Definition 2.5 *A set $X^+ \subset \mathbb{R}^n_x \times \mathbb{R}^k_u$ is said to represent a set $X \subset \mathbb{R}^n_x$, if the projection of X^+ onto the space of x-variables is exactly X. In other words, "X^+ represents X" means that whenever $x \in \mathbb{R}^n_x$ we have $x \in X$ if and only if there exists $u \in \mathbb{R}^k_u$ such that $(x,u) \in X^+$:*

$$X = \left\{x : \exists u : (x,u) \in X^+\right\}.$$

A system of constraints $\mathcal{S}^+$ in variables $x \in \mathbb{R}^n_x$, $u \in \mathbb{R}^k_u$ is said to represent a system of constraints $\mathcal{S}$ in variables $x \in \mathbb{R}^n_x$, if the feasible set of the former system represents the feasible set of the latter one. In other words, $\mathcal{S}^+$ represents $\mathcal{S}$, if, whenever $x \in \mathbb{R}^n_x$, x is feasible for $\mathcal{S}$ if and only if x can be extended, by a $u \in \mathbb{R}^k_u$, to a feasible solution (x,u) of $\mathcal{S}^+$.

With this definition, it is clear that the system (2.14) indeed represents the constraint (2.13), and, more generally, that the system of $2n+1$ linear inequalities

$$-u_j \leq x_j \leq u_j,\ j = 1,...,n, \sum_j u_j \leq 1$$

in variables x, u represents the constraint

$$\sum_j |x_j| \leq 1.$$

To understand how "powerful" this representation is, note that to represent the same constraint in the style of (2.12), that is, without extra variables, it would take as much as 2^n linear inequalities.

Coming back to the general case, assume that we are given an optimization problem

$$\min_x \{f(x) \text{ s.t. } x \text{ satisfies } \mathcal{S}_i,\ i = 1,...,m\} \tag{P}$$

where $\mathcal{S}_i$ are systems of constraints in variables x, and that we have in our disposal systems $\mathcal{S}_i^+$ of constraints in variables x, v^i which represent the systems $\mathcal{S}_i$. Clearly, the problem

$$\min_{x,v^1,...,v^m} \left\{f(x) \text{ s.t. } (x,v^i) \text{ satisfies } \mathcal{S}_i^+,\ i = 1,...,m\right\} \tag{P$^+$}$$

is equivalent to (P): the x-component of every feasible solution to (P$^+$) is feasible for (P) with the same value of the objective, and the optimal values in the problems are equal to each other, so that the x-component of an ϵ-optimal, in terms of the objective, feasible solution to (P$^+$) is an ϵ-optimal feasible solution to (P). We shall say that (P^+) represents equivalently the original problem (P). What is important here, is that a representation can possess desired properties which are absent in the original problem. For example, an appropriate representation can convert the problem of the form $\min\limits_x\{\|Px-p\|_1 : Ax \le b\}$ with n variables, m linear constraints and k-dimensional vector p, into a LO problem with $n+k$ variables and $m+2k+1$ linear inequality constraints, etc. Our goal now is to build a representation capable to express equivalently a semi-infinite linear constraint (specifically, the robust counterpart of an uncertain linear inequality) as a finite system of explicit convex constraints, with the ultimate goal to use these representations in order to convert the RC of an uncertain LO problem into an explicit (and as such, computationally tractable) convex program.

The outlined strategy allows us to focus on a *single* uncertainty-affected linear inequality – a family

$$\left\{a^T x \le b\right\}_{[a;b]\in\mathcal{U}} \tag{2.15}$$

of linear inequalities with the data varying in the uncertainty set

$$\mathcal{U} = \left\{[a;b] = [a^0;b^0] + \sum_{\ell=1}^{L} \zeta_\ell [a^\ell;b^\ell] : \zeta \in \mathcal{Z}\right\} \tag{2.16}$$

– and on "tractable representation" of the RC

$$a^T x \le b \quad \forall \left([a;b] = [a^0;b^0] + \sum_{\ell=1}^{L} \zeta_\ell [a^\ell;b^\ell] : \zeta \in \mathcal{Z}\right). \tag{2.17}$$

of this uncertain inequality.

By reasons indicated in item **C** of Section 2.2, we assume from now on that the associated perturbation set $\mathcal{Z}$ is convex.

2.3.2 Tractable representation of (2.17): simple cases

We start with the cases where the desired representation can be found by "bare hands", specifically, the cases of *interval* and *simple ellipsoidal* uncertainty.

Example 2.4 Consider the case of *interval uncertainty*, that is, the case where $\mathcal{Z}$ in (2.17) is a box. W.l.o.g. we can normalize the situation by assuming that

$$\mathcal{Z} = \mathrm{Box}_1 \equiv \{\zeta \in \mathbb{R}^L : \|\zeta\|_\infty \le 1\}.$$

In this case, (2.17) reads

$$\begin{array}{rl} & [a^0]^T x + \sum\limits_{\ell=1}^{L} \zeta_\ell [a^\ell]^T x \le b^0 + \sum\limits_{\ell=1}^{L} \zeta_\ell b^\ell \ \ \forall (\zeta : \|\zeta\|_\infty \le 1) \\ \Leftrightarrow & \sum\limits_{\ell=1}^{L} \zeta_\ell [[a^\ell]^T x - b^\ell] \le b^0 - [a^0]^T x \ \ \forall (\zeta : |\zeta_\ell| \le 1,\ \ell = 1,...,L) \\ \Leftrightarrow & \max\limits_{-1\le\zeta_\ell\le 1} \left[\sum\limits_{\ell=1}^{L} \zeta_\ell [[a^\ell]^T x - b^\ell]\right] \le b^0 - [a^0]^T x \end{array}$$

The concluding maximum in the chain is clearly $\sum\limits_{\ell=1}^{L} |[a^l]^T x - b^\ell|$, and we arrive at a representation of (2.17) by the explicit convex constraint

$$[a^0]^T x + \sum_{\ell=1}^{L} |[a^\ell]^T x - b^\ell| \le b^0, \tag{2.18}$$

which in turn admits a representation by a system of linear inequalities:

$$\left\{ \begin{array}{l} -u_\ell \le [a^\ell]^T x - b^\ell \le u_\ell, \ \ell = 1, ..., L, \\ [a^0]^T x + \sum\limits_{\ell=1}^{L} u_\ell \le b^0. \end{array} \right. \tag{2.19}$$

Example 2.5 Consider the case of *ellipsoidal uncertainty*, that is, the case where $\mathcal{Z}$ in (2.17) is an ellipsoid. W.l.o.g. we can normalize the situation by assuming that $\mathcal{Z}$ is merely the ball of radius Ω centered at the origin:

$$\mathcal{Z} = \text{Ball}_\Omega = \{\zeta \in \mathbb{R}^L : \|\zeta\|_2 \le \Omega\}.$$

In this case, (2.17) reads

$$\begin{array}{rc} & [a^0]^T x + \sum\limits_{\ell=1}^{L} \zeta_\ell [a^\ell]^T x \le b^0 + \sum\limits_{\ell=1}^{L} \zeta_\ell b^\ell \ \forall (\zeta : \|\zeta\|_2 \le \Omega) \\ \Leftrightarrow & \max\limits_{\|\zeta\|_2 \le \Omega} \left[\sum\limits_{\ell=1}^{L} \zeta_\ell [[a^\ell]^T x - b^\ell] \right] \le b^0 - [a^0]^T x \\ \Leftrightarrow & \Omega \sqrt{\sum\limits_{\ell=1}^{L} ([a^\ell]^T x - b^\ell)^2} \le b^0 - [a^0]^T x, \end{array}$$

and the concluding line provides a representation of (2.17) by the explicit convex constraint

$$[a^0]^T x + \Omega \sqrt{\sum_{\ell=1}^{L} ([a^\ell]^T x - b^\ell)^2} \le b^0. \tag{2.20}$$

2.3.3 Tractable representation of (2.17): general case

Now consider a rather general case when the perturbation set $\mathcal{Z}$ in (2.17) is given by a *conic representation* (cf. Appendix **??**:

$$\mathcal{Z} = \left\{ \zeta \in \mathbb{R}^L : \exists u \in \mathbb{R}^K : P\zeta + Qu + p \in \mathbf{K} \right\}, \tag{2.21}$$

where $\mathbf{K}$ is a closed convex pointed cone in $\mathbb{R}^N$ with a nonempty interior. In the case when $\mathbf{K}$ is *not* a polyhedral cone, assume that this representation is strictly feasible:

$$\exists (\bar{\zeta}, \bar{u}) : P\bar{\zeta} + Q\bar{u} + p \in \text{int} K. \tag{2.22}$$

Theorem 2.1 *Let the perturbation set $\mathcal{Z}$ be given by (2.21), and in the case of non-polyhedral $\mathbf{K}$, let also (2.22) take place. Then the semi-infinite constraint (2.17) can be represented by the following system of conic inequalities in variables $x \in \mathbb{R}^n, y \in \mathbb{R}^L$:*

$$\begin{array}{l} p^T y + [a^0]^T x \leq b^0, \\ Q^T y = 0, \\ (P^T y)_\ell + [a^\ell]^T x = b^\ell, \ell = 1, ..., L, \\ y \in \mathbf{K}_*, \end{array} \tag{2.23}$$

where $\mathbf{K}_* = \{y : y^T z \geq 0 \,\forall z \in \mathbf{K}\}$ is the cone dual to $\mathbf{K}$.

Proof.
We have

$$\begin{array}{rc} & x \text{ is feasible for (2.17)} \\ \Leftrightarrow & \sup\limits_{\zeta \in \mathcal{Z}} \Big\{ \underbrace{[a^0]^T x - b^0}_{d[x]} + \sum\limits_{\ell=1}^{L} \zeta_\ell \underbrace{\left[[a^\ell]^T x - b^\ell\right]}_{c_\ell[x]} \Big\} \leq 0 \\ \Leftrightarrow & \sup\limits_{\zeta \in \mathcal{Z}} \left\{ c^T[x]\zeta + d[f] \right\} \leq 0 \\ \Leftrightarrow & \sup\limits_{\zeta \in \mathcal{Z}} c^T[x]\zeta \leq -d[x] \\ \Leftrightarrow & \max\limits_{\zeta, v} \left\{ c^T[x]\zeta : P\zeta + Qv + p \in \mathbf{K} \right\} \leq -d[x]. \end{array}$$

The concluding relation says that x is feasible for (2.17) if and only if the optimal value in the conic program

$$\max_{\zeta, v} \left\{ c^T[x]\zeta : P\zeta + Qv + p \in \mathbf{K} \right\} \tag{CP}$$

is $\leq -d[x]$. Assume, first, that (2.22) takes place. Then (CP) is strictly feasible, and therefore, applying the Conic Duality Theorem (Theorem **??**), the optimal value in (CP) is $\leq -d[x]$ if and only if the optimal value in the conic dual to (CP) problem

$$\min_{y} \left\{ p^T y : Q^T y = 0, P^T y = -c[x], y \in \mathbf{K}_* \right\}, \tag{CD}$$

is achieved and is $\leq -d[x]$. Now assume that (2.22) does not take place. Under assumptions of Theorem, the latter is possible only when $\mathbf{K}$ is a polyhedral cone, in which case the usual LO Duality Theorem yields exactly the same conclusion: the optimal value in (CP) is $\leq -d[x]$ if and only if the optimal value in (CD) is achieved and is $\leq -d[x]$. In other words, under the premise of Theorem, x is feasible for (2.17) if and only if (CD) has a feasible solution y with $p^T y \leq -d[x]$. □

Observing that nonnegative orthants, Lorentz and Semidefinite cones are self-dual, we derive from Theorem 2.1 the following corollary:

Corollary 2.1 *Let the nonempty perturbation set in* (2.17) *be*

(i) *polyhedral, i.e., given by* (2.21) *with a nonnegative orthant* $\mathbb{R}^N_+$ *in the role of* $\mathbf{K}$*, or*

(ii) *conic quadratic representable, i.e., given by* (2.21) *with a direct product* $\mathbf{L}^{k_1} \times ... \times \mathbf{L}^{k_m}$ *of Lorentz cones* $\mathbf{L}^k = \{x \in \mathbb{R}^k : x_k \geq \sqrt{x_1^2 + ... + x_{k-1}{}^2}\}$ *in the role of* $\mathbf{K}$*, or*

(iii) *semidefinite representable, i.e., given by* (2.21) *with the positive semidefinite cone* $\mathbf{S}^k_+$ *in the role of* $\mathbf{K}$.

In the cases of (ii), (iii)*, assume that* (2.22) *holds true. Then the Robust Counterpart* (2.17) *of uncertain linear inequality with the perturbation set* $\mathcal{Z}$ *admits equivalent reformulation as an explicit system of*

— *linear inequalities, in the case of* (i),

— *conic quadratic inequalities, in the case of* (ii),

— *linear matrix inequalities, in the case of* (iii).

In all cases, the size of the reformulation is polynomial in the number of variables in (2.17) *and the size of the conic description of* $\mathcal{Z}$*, while the data of the reformulation is readily given by the data describing, via* (2.21)*, the perturbation set* $\mathcal{Z}$.

Remark 2.2 Usually, the cone $\mathbf{K}$ participating in (2.21) is the direct product of simpler cones $\mathbf{K}^1, ..., \mathbf{K}^S$, so that representation (2.21) takes the form

$$\mathcal{Z} = \{\zeta : \exists u^1, ..., u^S : P_s\zeta + Q_s u^s + p_s \in \mathbf{K}^s,\ s = 1, ..., S\}. \tag{2.24}$$

In this case, (2.23) becomes the system of conic constraints in variables $x, y^1, ..., y^S$ as follows:

$$\begin{array}{l} \sum\limits_{s=1}^{S} p_s^T y^s + [a^0]^T x \leq b^0, \\ Q_s^T y^s = 0,\ s = 1, ..., S, \\ \sum\limits_{s=1}^{S} (P_s^T y^s)_\ell + [a^\ell]^T x = b^\ell, \ell = 1, ..., L, \\ y^s \in \mathbf{K}^s_*, s = 1, ..., S, \end{array} \tag{2.25}$$

where K^s_* is the cone dual to K^s.

Examples

We are about to apply Theorem 2.1 to build tractable reformulations of the semi-infinite inequality (2.17) in two particular cases. While at a first glance seemingly no natural "uncertainty models" lead to "strange" perturbation sets we are about to consider, it will become clear in the mean time that these sets are of significant importance – they allow to model *random* uncertainty.

Example 2.6 *$\mathcal{Z}$ is the intersection of concentric co-axial box and ellipsoid, specifically,*

$$\mathcal{Z} = \{\zeta \in \mathbb{R}^L : -1 \leq \zeta_\ell \leq 1, \ell \leq L, \sqrt{\sum_{\ell=1}^{L} \zeta_\ell^2/\sigma_\ell^2} \leq \Omega\}, \tag{2.26}$$

where $\sigma_\ell > 0$ and $\Omega > 0$ are given parameters.

Here representation (2.24) becomes

$$\mathcal{Z} = \{\zeta \in \mathbb{R}^L : P_1\zeta + p_1 \in \mathbf{K}^1, P_2\zeta + p_2 \in \mathbf{K}^2\},$$

where

• $P_1\zeta \equiv [\zeta; 0]$, $p_1 = [0_{L\times 1}; 1]$ and $\mathbf{K}^1 = \{(z,t) \in \mathbb{R}^L \times \mathbb{R} : t \geq \|z\|_\infty\}$, whence $\mathbf{K}^1_* = \{(z,t) \in \mathbb{R}^L \times \mathbb{R} : t \geq \|z\|_1\}$;

• $P_2\zeta = [\Sigma^{-1}\zeta; 0]$ with $\Sigma = \mathrm{Diag}\{\sigma_1, ..., \sigma_L\}$, $p_2 = [0_{L\times 1}; \Omega]$ and $\mathbf{K}^2$ is the Lorentz cone of the dimension $L+1$ (whence $\mathbf{K}^2_* = \mathbf{K}^2$)

Setting $y^1 = [\eta_1; \tau_1]$, $y^2 = [\eta_2; \tau_2]$ with one-dimensional τ_1, τ_2 and L-dimensional η_1, η_2, (2.25) becomes the following system of constraints in variables τ, η, x:

$$\begin{array}{lrcl}
(a) & \tau_1 + \Omega\tau_2 + [a^0]^T x & \leq & b^0, \\
(b) & (\eta_1 + \Sigma^{-1}\eta_2)_\ell & = & b^\ell - [a^\ell]^T x, \ \ell = 1, ..., L, \\
(c) & \|\eta_1\|_1 & \leq & \tau_1 \quad [\Leftrightarrow [\eta_1; \tau_1] \in \mathbf{K}^1_*], \\
(d) & \|\eta_2\|_2 & \leq & \tau_2 \quad [\Leftrightarrow [\eta_2; \tau_2] \in \mathbf{K}^2_*].
\end{array}$$

We can eliminate from this system the variables τ_1, τ_2 – for every feasible solution to the system, we have $\tau_1 \geq \bar{\tau}_1 \equiv \|\eta_1\|_\infty$, $\tau_2 \geq \bar{\tau}_2 \equiv \|\eta_2\|_2$, and the solution obtained when replacing τ_1, τ_2 with $\bar{\tau}_1$, $\bar{\tau}_2$ still is feasible. The reduced system in variables x, $z = \eta_1$, $w = \Sigma^{-1}\eta_2$, namely, the system

$$\boxed{\begin{array}{rcl}
\sum_{\ell=1}^{L} |z_\ell| + \Omega\sqrt{\sum_\ell \sigma_\ell^2 w_\ell^2} + [a^0]^T x & \leq & b^0, \\
z_\ell + w_\ell & = & b^\ell - [a^\ell]^T x, \ \ell = 1, ..., L
\end{array} \qquad (2.27)}$$

also is a representation of (2.17), (2.26).

Example 2.7 *"Budgeted uncertainty"*. Consider the case where

$$\mathcal{Z} = \{\zeta \in \mathbb{R}^L : \|\zeta\|_\infty \equiv \max_\ell |\zeta_\ell| \leq 1, \ \|\zeta\|_1 \equiv \sum_\ell |\zeta_\ell| \leq \gamma\}, \tag{2.28}$$

where γ, $1 \leq \gamma \leq L$, is a given "uncertainty budget".

In the case in question, representation (2.24) becomes

$$\mathcal{Z} = \{\zeta \in \mathbb{R}^L : P_1\zeta + p_1 \in \mathbf{K}^1, P_2\zeta + p_2 \in \mathbf{K}^2\},$$

where

• $P_1\zeta \equiv [\zeta;0]$, $p_1 = [0_{L\times 1};1]$ and $\mathbf{K}^1 = \{[z;t] \in \mathbb{R}^L \times \mathbb{R} : t \geq \|z\|_\infty\}$, whence $\mathbf{K}^1_* = \{[z;t] \in \mathbb{R}^L \times \mathbb{R} : t \geq \|z\|_1\}$;

• $P_2\zeta = [\zeta;0]$, $p_2 = [0_{L\times 1};\gamma]$ and $\mathbf{K}^2 = \mathbf{K}^1_* = \{[z;t] \in \mathbb{R}^L \times \mathbb{R} : t \geq \|z\|_1\}$, whence $\mathbf{K}^2_* = \mathbf{K}^1$.

Setting $y^1 = [z;\tau_1]$, $y^2 = [w;\tau_2]$ with one-dimensional τ and L-dimensional z, w, system (2.25) becomes the following system of constraints in variables τ_1, τ_2, z, w, x:

$$\begin{array}{lrcl} (a) & \tau_1 + \gamma\tau_2 + [a^0]^T x & \leq & b^0, \\ (b) & (z+w)_\ell & = & b^\ell - [a^\ell]^T x,\ \ell = 1,...,L, \\ (c) & \|z\|_1 & \leq & \tau_1 \quad [\Leftrightarrow [\eta_1;\tau_1] \in \mathbf{K}^1_*], \\ (d) & \|w\|_\infty & \leq & \tau_2 \quad [\Leftrightarrow [\eta_2;\tau_2] \in \mathbf{K}^2_*]. \end{array}$$

Same as in Example 2.6, we can eliminate the τ-variables, arriving at the following representation of (2.17), (2.28) by the following system of constraints in variables x, z, w:

$$\begin{array}{rcl} \sum_{\ell=1}^{L} |z_\ell| + \gamma \max_\ell |w_\ell| + [a^0]^T x & \leq & b^0, \\ z_\ell + w_\ell & = & b^\ell - [a^\ell]^T x,\ \ell = 1,...,L. \end{array} \tag{2.29}$$

which can be further converted into the system of linear inequalities in z, w and additional variables.