ELSEVIER

Operations Research Letters 25 (1999) 1–13

operations research letters

www.elsevier.com/locate/orms

Robust solutions of uncertain linear programs ☆

A. Ben-Tal, A. Nemirovski *

Technion-Israel Institute of Technology, Faculty of Industrial Engineering and Management, Technion City, 32000 Haifa, Israel

Received 1 December 1996; received in revised form 1 November 1998

Abstract

We treat in this paper linear programming (LP) problems with uncertain data. The focus is on uncertainty associated with *hard* constraints: those which *must* be satisfied, whatever is the actual realization of the data (within a prescribed *uncertainty set*). We suggest a modeling methodology whereas an uncertain LP is replaced by its robust counterpart (RC). We then develop the analytical and computational optimization tools to obtain robust solutions of an uncertain LP problem via solving the corresponding explicitly stated convex RC program. In particular, it is shown that the RC of an LP with ellipsoidal uncertainty set is computationally tractable, since it leads to a conic quadratic program, which can be solved in polynomial time. © 1999 Published by Elsevier Science B.V. All rights reserved.

Keywords: Linear programming; Data uncertainty; Robustness; Convex programming; Interior-point methods

1. Introduction

The data A, b associated with a linear program

$$\min\{c^{\mathrm{T}}x \mid Ax \geqslant b\} \tag{1}$$

are "uncertain" to some degree in most real world problems. [1] In many models the uncertainty is ignored altogether, and a representative nominal value of the data is used (e.g. expected values). The classical approach in operations research/management science to deal with uncertainty is *stochastic programming* (SP) (see e.g., [6,13] and references therein). But even in this approach constraints may be violated, with certain penalty (this is the case for *SP with recourse* [6,9], *scenario optimization* [14], *entropic penalty methods* [1]) or with certain probability (chance constraints). In the dominating penalty approach, even when the random variables are degenerate (deterministic), the corresponding SP model does not recover necessarily the original LP constraints, but only a *relaxation* of these constraints. Thus, although this was not stated explicitly in the past, SP treats in fact mainly *soft constraints*. These remarks apply also to the recent scenario-based penalty approach of Mulvey, Vanderbei and Zenios [11].

In this paper we study uncertainty associated with *hard* constraints, i.e., those which *must* be satisfied whatever is the realization of the data (A, b) within, of

☆ The research was partly supported by the German-Israeli Foundation, contract I-0455 214.06/95, and by the Israel Ministry of Science grant #9636-1-96.

* Corresponding author. Fax: +972-4 8235194.

[1] If the vector c is also uncertain, we could look at the equivalent formulation of (1): $\min_{x,t}\{t:\ c^{\mathrm{T}}x \leqslant t,\ Ax \geqslant b\}$ and thus without loss of generality we may restrict the uncertainty to the constraints only.

PII: S0167-6377(99)00016-4

course, a reasonable prescribed "uncertainty set" $\mathscr{U}$. Feasibility of a vector x is thus interpreted as

$$Ax \geqslant b \quad \forall (A,b) \in \mathscr{U}. \tag{2}$$

Consequently, we call the problem

$$\min\{c^{\mathrm{T}}x \mid Ax \geqslant b \quad \forall (A,b) \in \mathscr{U}\} \tag{3}$$

the *robust counterpart* of the uncertain LP problem (1), and we call a vector x^* solving Eq. (3) a *robust solution* of the uncertain problem.

No underlying stochastic model of the data is assumed to be known (or even to exist), although such knowledge may be of use to obtain reasonable uncertainty sets (see Section 4).

Dealing with uncertain hard constraints is perhaps a novelty in Mathematical Programming; to the best of our knowledge, the only previous example related to this question is due to Soyster [16] (see below). The issue of hard uncertain constraints, however, is not a novelty for Control Theory, where it is a well-studied subject forming the area of Robust Control (see, e.g., [17] and references therein).

The approach reflected by Eq. (3) may look at a first glance too "pessimistic": the point is that there are indeed situations in reality when an applicable solution must be feasible for all realizations of the data, and even a small violation of the constraints cannot be tolerated. We have discussed such a situation elsewhere, for problems of designing engineering structures (e.g., bridges), see [2]. In these problems, ignoring even small changes in the forces acting on the structure may cause "violent" displacements and results in a severely unstable structure. This example is not from the world of LP, but we can easily imagine similar situations in LP models. Consider, e.g., a chemical plant which takes raw materials, decomposes them into components and then recombine the components to get the final product (there could be several "decomposition–recombination" stages in the production process). The corresponding LP model includes the inequality constraints expressing the fact that when recombining the intermediate components, you cannot use more than what is given by the preceding decomposition phase, and these constraints indeed are "hard". If, as it is normally the case, the content of the components in raw materials is "uncertain" or/and the yield of the decomposition process depends on uncontrolled factors, we end up with an LP program with uncertain data in *hard* constraints.

As it was already mentioned, uncertain hard constraints in LP models were discussed (in a very specific setting) by Soyster [16] (for further developments, see also [7,15]). The case considered in these papers is the one of "column-wise" uncertainty, i.e., the columns a_i of the constraint matrix in the constraints $Ax \leqslant b,\ x \geqslant 0$ are known to belong to a given convex sets K_i, and the linear objective is minimized over those x for which

$$\sum_{i=1}^{n} x_i a_i \leqslant b \quad \forall (a_i \in K_i,\ i = 1,\ldots,n). \tag{4}$$

As it is shown in [16], the constraints (4) are equivalent to a system of *linear* constraints

$$A^* x \leqslant b, \quad x \geqslant 0, \qquad a^*_{ij} = \sup_{a_i \in K_i} (a_i)_j. \tag{5}$$

It turns out that the phenomenon that (4) is equivalent to a *linear* system (5) is specific for "column-wise" uncertainty. As we shall see later, the general case is the one of "row-wise" uncertainty (the one where the *rows* of the constraint matrix are known to belong to given convex sets). In this latter case, the robust counterpart of the problem is typically *not* an LP program. E.g., when the uncertainty sets for the rows of A are ellipsoids, the robust counterpart turns out to be a conic quadratic problem. Note also that the case of column-wise uncertainty is extremely conservative: the constraints (5) of the robust counterpart correspond to the case when every entry in the constraint matrix is as "bad" (as large) as it could be. At the same time, in the case of row-wise uncertainty the robust counterpart is capable to reflect the fact that generically the coefficients of the constraints cannot be simultaneously at their worst values.

The rest of the paper is organized as follows. Section 2 contains formal definition of the "robust counterpart" of an "uncertain Linear Programming program" (which is the one just defined, but in a slightly more convenient form). Here we also answer some natural questions about the links between this counterpart and the usual LP program corresponding to the "worst" realization of the data from the uncertainty set $\mathscr{U}$. Note that the robust counterpart problem $(P_{\mathscr{U}})$ has a continuum of constraints, i.e., is a *semi-infinite* problem, and as such it looks computationally intractable. In the last part of Section 2 we discuss the crucial issue of which geometries of the uncertainty set $\mathcal{U}$ result in a "computationally tractable" robust counterpart of Eq. (1). Since we have in mind large scale applications of robust LPs, we focus on geometries of the uncertainty set leading to "explicit" robust counterpart of nice analytical structure, and which can be solved by high-performance optimization algorithms (like the interior point methods). Specifically, in Section 3 we consider the case of our preferred structure where $\mathcal{U}$ in an intersection of finitely many ellipsoids and derive for this case the explicit form of the robust counterpart of Eq. (1), which turns out to be a conic quadratic problem. Section 4 contains a simple portfolio selection example illustrating our robust solution and comparing it to the solution obtained by the scenario-based approach of [11].

2. Robust counterpart of an uncertain Linear Programming program

2.1. The construction

For convenience and notational simplicity we choose to cast the LP problem in the form

$$(\mathrm{P}) \quad \min\{c^{\mathrm{T}}x \mid Ax \geqslant 0, f^{\mathrm{T}}x = 1\}. \tag{6}$$

The canonical LP program (1) in the Introduction can be obtained from Eq. (6) by the correspondence

$$c := \begin{pmatrix} c \\ 0 \end{pmatrix}, \quad A : [A; -b], \;\; f = (0, \ldots, 0, 1)^{\mathrm{T}}.$$

In the formulation (6), the vectors $c, f \in \mathbb{R}^n$ are *fixed data*, and the uncertainty is associated only with the $m \times n$ matrix A. The *uncertainty set*, of which A is a member, is a set $\mathcal{U}$ of $m \times n$ real matrices. The *robust counterpart* of (6) is defined to be the optimization problem

$$(P_{\mathcal{U}}) \quad \min\{c^{\mathrm{T}}x \mid Ax \geqslant 0 \quad \forall A \in \mathcal{U};\; f^{\mathrm{T}}x = 1\}. \tag{7}$$

Thus, a *robust feasible* (r-feasible for short) solution to the "robust counterpart" of (P) should, by definition, satisfy all realizations of the constraints from the uncertainty set $\mathcal{U}$, and a *robust optimal* (r-optimal for short) solution to $(P_{\mathcal{U}})$ is an r-feasible solution with the best possible value of the objective. In the rest of this section we investigate the basic mathematical properties of program $(P_{\mathcal{U}})$, with emphasis on the crucial issue of its "computational tractability".

From now on we fix certain LP data $\mathcal{U}$, c, f and denote by $\mathcal{P}$ the corresponding "uncertain LP program", i.e., the family of all LP programs (P) with given c, f and some $A \in \mathcal{U}$. Each program of this type will be called *an instance* (or a realization) of the uncertain LP program.

Evidently, the robust counterpart $(P_{\mathcal{U}})$ of an uncertain LP program remains unchanged when we replace the uncertainty set $\mathcal{U}$ by its closed convex hull. Consequently, from now on we always assume that $\mathcal{U}$ is convex and closed.

2.2. Robust counterpart and the "worst" LP program from $\mathcal{P}$

The first issue concerning the robust counterpart of an uncertain LP is how "conservative" it is. From the very beginning our approach is "worst case oriented"; but could $(P_{\mathcal{U}})$ be even worse than the "worst" instance from $\mathcal{P}$? E.g.,

(A) Assume $(P_{\mathcal{U}})$ is infeasible. Does it mean that there is an infeasible instance $(P) \in \mathcal{P}$?

Further, assume that $(P_{\mathcal{U}})$ is feasible, and its optimal value c^* is finite. By construction, c^* is greater than or equal to the optimal value $c^*(P)$ of every problem instance $(P) \in \mathcal{P}$. The question of how conservative is the robust counterpart now may be posed as

(B) Assume that $(P_{\mathcal{U}})$ is feasible with finite optimal value c^*. Does it mean that there exists a problem instance $(P) \in \mathcal{P}$ with $c^*(P) = c^*$?

A "gap" between the solvability properties of the instances of an uncertain LP and those of its robust counterpart may indeed exist, as is demonstrated by the following simple example:

$$\begin{array}{ll} \min & x_1 + x_2 \\ \text{s.t.} & a_{11}x_1 + x_2 \geqslant 1, \\ & x_1 + a_{22}x_2 \geqslant 1, \\ & x_1 + x_2 = 1, \\ & x_1,\; x_2 \geqslant 0. \end{array}$$

The uncertain elements are a_{11} and a_{22}, and the uncertainty set is $\mathcal{U} = \{a_{11} + a_{22} = 2, \frac{1}{2} \leqslant a_{11} \leqslant \frac{3}{2}\}$. Here every problem instance clearly is solvable with optimal value 1 (if $a_{11} \geqslant 1$, then the optimal solution is $(1, 0)$, otherwise it is $(0, 1)$), while the robust counterpart, which clearly is

$$\begin{array}{ll} \min & x_1+x_2 \\ \text{s.t.} & \frac{1}{2}x_1+x_2\geqslant 1, \\ & x_1+\frac{1}{2}x_2\geqslant 1, \\ & x_1+x_2=1, \\ & x_1,\ x_2\geqslant 0 \end{array}$$

is infeasible.

We are about to demonstrate that there are quite natural assumptions under which the indicated gap disappears, and the answers to the above two questions (A) and (B) become positive, so that $(P_{\mathscr{U}})$ under these assumptions is not worse than the "worst" instance from $\mathscr{P}$.

The main part of the aforementioned assumptions is *that the uncertainty should be "constraint wise"*. To explain this concept, let $\mathscr{U}_i$ be the set of all possible realizations of i-th row in the constraint matrix, i.e., be the projection of $\mathscr{U}\subset\mathbb{R}^{m\times n}=\mathbb{R}^n\times\cdots\times\mathbb{R}^n$ onto i-th direct factor of the right hand side of the latter relation.[2] We say that the uncertainty is *constraint-wise*, if the uncertainty set $\mathscr{U}$ is the direct product of the "partial" uncertainty sets $\mathscr{U}_i$:

$$\mathscr{U}=\mathscr{U}_1\times\mathscr{U}_2\times\cdots\times\mathscr{U}_m.$$

By construction, x is r-feasilbe if and only if

$$a^{\mathrm{T}}x\geqslant 0\quad \forall a\in\mathscr{U}_i\ \forall i;\ \ f^{\mathrm{T}}x=1. \tag{8}$$

In other words, $(P_{\mathscr{U}})$ remains unchanged when we extend the initial uncertainty set $\mathscr{U}$ to the direct product $\hat{\mathscr{U}}=\mathscr{U}_1\times\cdots\times\mathscr{U}_m$ of its projections on the sub-spaces of data of different constraints; the robust counterpart "feels" only the possible realizations of ith constraint, $i=1,\dots,m$ and does not feel the dependencies (if any) between these constraints in the instances. Thus, given an arbitrary uncertainty set $\mathscr{U}$, we can always extend it to a constraint-wise uncertainty set resulting in the same robust counterpart.

In view of the above remarks, when speaking about questions (A) and (B) it is quite natural to assume constraint-wise uncertainty. An additional technical assumption needed to answer affirmatively to (A) and (B) is the following:

[2] In order not to introduce both column and row vectors, we represent a row in an $m\times n$ matrix by a *column* vector.

Boundedness assumption. There exists a convex compact set $Q\subset\mathbb{R}^n$ which for sure contains feasible sets of all problem instances $(P)\in\mathscr{P}$.

To ensure this assumption, it suffices to assume that the constraints of all instances have a common "certain" part (e.g., box constraints on the design variables) which defines a bounded set in $\mathbb{R}^n$.

Now we can formulate the main results of the subsection.

Proposition 2.1. *If the uncertainty $\mathscr{U}$ is constraint-wise and the boundedness assumption holds, then*

(i) *$(P_{\mathscr{U}})$ is infeasible if and only if there exists an infeasible instance $(P)\in\mathscr{P}$.*

(ii) *If $(P_{\mathscr{U}})$ is feasible and c^* is the optimal value of the problem, then*

$$c^*=\sup\{c^*(P)\,|\,(P)\in\mathscr{P}\}. \tag{9}$$

Proof. (i): since the feasible set of $(P_{\mathscr{U}})$ is contained in the feasible set of every problem instance, the "if" part of the statement is evident. To prove the "only if" part, assume that $(P_{\mathscr{U}})$ is infeasible, and let us prove that there exists an infeasible instance $(P)\in\mathscr{P}$. Consider the system of linear inequalities (8) with unknown x and additional restriction $x\in Q$, Q being given by the Boundedness assumption. Since $(P_{\mathscr{U}})$ is infeasible, this system has no solution. By the standard compactness arguments it follows that already certain finite subsystem $a_p^{\mathrm{T}}x\geqslant 0$, $p=1,\dots,N$, $f^{\mathrm{T}}x=1$ of (8) has no solution in Q. Now let $A_1,\dots,A_N\in\mathscr{U}$ be the instances from which the a_p's come. We claim that the system of inequalities

$$A_1x\geqslant 0,\dots,A_Nx\geqslant 0,\quad f^{\mathrm{T}}x=1 \tag{10}$$

has no solutions. Indeed, by its origin it has no solutions in Q, and due to the Boundedness assumption it has no solutions outside Q as well. By Farkas' lemma, inconsistency of Eq. (10) implies existence of nonnegative λ_{ip}, $i=1,\dots,m$, $p=1,\dots,N$, and a real μ such that

$$\sum_{i=1}^{m}\sum_{p=1}^{N}\lambda_{ip}a_i^p+\mu f=0,\quad \mu>0, \tag{11}$$

where a_i^p is ith row of A_p. Let $\lambda_i=\sum_{p=1}^N\lambda_{ip}$. For all i with nonzero λ_i, let a_i be defined as $a_i=\lambda_i^{-1}\sum_{p=1}^N\lambda_{ip}a_i^p$; for i with zero λ_i let $a_i=a_i^1$.

By construction, a_i is a convex combination of ith rows of certain instances, so that $a_i \in \mathscr{U}_i$. Since the uncertainty is constraint-wise, the matrix A with the rows a_i^{T}, $i=1,\dots,m$, belongs to $\mathscr{U}$. On the other hand, Eq. (11) implies that $\sum_{i=1}^{m} \lambda_i a_i + \mu f = 0$ & $\mu > 0$; but this means exactly that the problem instance given by A is infeasible □

Claim (ii) is an immediate consequence of (i). Indeed, let us denote the right hand side of (9) by d. Evidently $d \leqslant c^*$, and all we need to prove is that the strict inequality here is impossible. Assume, to the contrary, that $d < c^*$, and let us add to all our instances the common "certain" inequality $d f^{\mathrm{T}}x - c^{\mathrm{T}}x \geqslant 0$ (i.e., let us add an upper bound d on the objective value). The new uncertain LP problem clearly has constraint-wise uncertainty and satisfies the boundedness assumption, and due to the origin of d all its instances are feasible. By (i), the robust counterpart $(P'_{\mathscr{U}})$ of this problem also is feasible. But this counterpart is nothing but $(P_{\mathscr{U}})$ with added inequality $c^{\mathrm{T}}x \leqslant d$, and since $d < c^*$, $(P'_{\mathscr{U}})$ is infeasible, which is the desired contradiction.

2.3. Computational tractability of $(P_{\mathscr{U}})$

Problem $(P_{\mathscr{U}})$ can be equivalently rewritten as

$$\min\{c^{\mathrm{T}}x \mid x \in G_{\mathscr{U}}\},$$

$$G_{\mathscr{U}} = \{x \mid Ax \geqslant 0 \quad \forall A \in \mathscr{U};\, f^{\mathrm{T}}x = 1\}. \tag{12}$$

It is clearly seen that $G_{\mathscr{U}}$ is a closed convex set, so that $(P_{\mathscr{U}})$ is a convex program. It is known [8] that in order to minimize in a theoretically efficient manner a linear objective over a closed convex set $G \subset \mathbb{R}^n$ it suffices to equip the set G with an efficient *separation oracle*. The latter is a routine which, given as input a point $x \in \mathbb{R}^n$, reports whether $x \in G$, and if it is not the case, returns a *separator* of x and G, i.e., a vector $e_x \in \mathbb{R}^n$ such that $e^{\mathrm{T}}x > \sup_{x' \in G} e^{\mathrm{T}}x'$.

Thus, the question "what are the geometries of the uncertainty set $\mathscr{U}$ which lead to a computationally tractable robust counterpart $(P_{\mathscr{U}})$ of the original LP program (P)" can be reformulated as follows: "when can $G_{\mathscr{U}}$ be equipped with an efficient separation oracle?" Recall that when answering this question, without loss of generality we may restrict ourselves to the case of *closed convex* uncertainty sets $\mathscr{U}$. An intermediate answer to our question is as follows: in order to equip $G_{\mathscr{U}}$ with an efficient separation oracle, it suffices to build an efficient *inclusion oracle* – an efficient routine which, given on input an $x \in \mathbb{R}^n$, verifies whether the convex set $\mathscr{U}(x) = \{Ax \mid A \in \mathscr{U}\}$ is contained in the nonnegative orthant $\mathbb{R}^m_+$, and if it is not the case, returns a matrix $A_x \in \mathscr{U}$ such that the point $A_x x$ does not belong to $\mathbb{R}^m_+$. Indeed, given an inclusion oracle $\mathscr{R}$, we can imitate a separation oracle for $G_{\mathscr{U}}$ as follows. In order to verify whether $x \in G_{\mathscr{U}}$, we first check whether $f^{\mathrm{T}}x = 1$; if it is not the case, then x clearly is not in $G_{\mathscr{U}}$, and we can take as a separator of x and $G_{\mathscr{U}}$ either f or $-f$, depending on the sign of the difference $f^{\mathrm{T}}x - 1$. Now assume that $f^{\mathrm{T}}x = 1$. Let us call $\mathscr{R}$ to check whether $\mathscr{U}(x) \subset \mathbb{R}^m_+$. If it is the case, then $x \in G_{\mathscr{U}}$; otherwise $\mathscr{R}$ returns a matrix $A_x \in \mathscr{U}$ such that $y = A_x x \notin \mathbb{R}^m_+$, so that at least one of the coordinates of the vector y, say, ith, is negative. Now consider the linear form $e_x^{\mathrm{T}}x' \equiv -e_i^{\mathrm{T}}A_x x'$ of $x' \in \mathbb{R}^n$, where e_i is the ith standard orth of $\mathbb{R}^m$. By construction, at the point $x' = x$ this form is positive, while at the set $G_{\mathscr{U}}$ it clearly is nonpositive (indeed, if $x' \in G_{\mathscr{U}}$, then $Ax' \in \mathbb{R}^m_+$ for all $A \in \mathscr{U}$ and, in particular, for $A = A_x$; consequently, the vector $y' = A_x x'$ is nonnegative and therefore $e_x^{\mathrm{T}}\, x' = -e_j^{\mathrm{T}}y' \leqslant 0$). Thus, e_x separates x from $G_{\mathscr{U}}$, as required.

We see that in order to equip $G_{\mathscr{U}}$ with an efficient separation oracle, it suffices to have in our disposal an efficient inclusion oracle $\mathscr{R}$. When does the latter oracle exist? An immediate example is the one when $\mathscr{U}$ is given as a convex hull of a finite set of "scenarios" $A_1,\dots,A_M$. Indeed, in this case to get $\mathscr{R}$, it suffices, given an input x to $\mathscr{R}$, to verify whether all the vectors $A_i x$, $i = 1,\dots,M$, are nonnegative. If it is the case, then $\mathscr{U}(x) = \mathrm{Conv}\{A_1x,\dots,A_Mx\}$ is contained in $\mathbb{R}^m_+$, and if for some i the vector $A_i x$ is not nonnegative, we can take A_i as A_x. The case in question is of no specific interest, since here $(P_{\mathscr{U}})$ simply is an LP program $\min\{c^{\mathrm{T}}x \mid A_1x \geqslant 0,\dots,A_Mx \geqslant 0,\ f^{\mathrm{T}}x = 1\}$. In fact, basically all we need to get an efficient inclusion oracle is "computational tractability of $\mathscr{U}$" – $\mathscr{U}$ itself should admit an efficient separation oracle. Namely, we can formulate the following

Tractability principle. An efficient Separation oracle for $\mathscr{U}$ implies an efficient Inclusion oracle and, consequently, implies an efficient separation oracle for $G_{\mathscr{U}}$ (and thus implies computational tractability of $(P_{\mathscr{U}})$).

Note that the formulated statement is a principle, not a *theorem*; to make it a theorem, we should add a number of unrestrictive technical assumptions (for details, see [8]). The principle is almost evident: in order to detect whether $\mathcal{U}(x) \subset \mathbb{R}^m_+$ for a given x, it suffices to solve m auxiliary convex programs

$$\min\{e_i^{\mathrm{T}} Ax \mid A \in \mathcal{U}\}, \quad i = 1, \dots, m,$$

A being the "design vector". (Indeed, if the optimal values in all these programs are nonnegative, $\mathcal{U}(x) \subseteq \mathbb{R}^m_+$; if the optimal value in ith of the problems is negative, then any feasible solution A_i to the ith problem with negative value of the objective can be taken as "separator" A_x). Now, our auxiliary problems are problems of minimizing a linear objective over a closed convex set $\mathcal{U}$, and we already have mentioned that basically all we need in order to solve efficiently a problem of minimizing a linear objective over a convex set is an efficient separation oracle for the set).

According to the tractability principle, all "reasonable" closed convex uncertainty sets $\mathcal{U}$ lead to "computationally tractable" $(P_{\mathcal{U}})$, e.g., all sets given by finitely many "efficiently computable" convex inequalities.[3] Note, however, that "computational tractability" is a theoretical property which is not exactly the same as efficient solvability in practice, especially given the huge sizes of some real world LP programs. In order to end up with "practically solvable" robust counterpart $(P_{\mathcal{U}})$, it is highly desirable to ensure a "simple" analytical structure of the latter problem, which in turn requires $\mathcal{U}$ to be "relatively simple". On the other hand, when restricting ourselves with "too simple" geometries of $\mathcal{U}$, we loose in flexibility of the above approach – in its ability to model diverse actual uncertainties. In our opinion, a reasonable solution to the conflicting goals is the one where $\mathcal{U}$ is restricted to be an "ellipsoidal" uncertainty, i.e., an intersection of finitely many "ellipsoids" – sets given by convex quadratic inequalities. Some arguments in favour of this choice are as follows:

- ellipsoidal uncertainty sets form relatively wide family including, as we shall see, polytopes (bounded sets given by finitely many linear inequalities) and can be used to approximate well many cases of complicated convex sets.
- an ellipsoid is given parametrically by data of moderate size, hence it is convenient to represent "ellipsoidal uncertainty" as input;
- in some important cases there are "statistical" reasons which give rise to ellipsoidal uncertainty (see Section 4);
- last (and most important), problem $(P_{\mathcal{U}})$ associated with an ellipsoidal $\mathcal{U}$ possesses a very nice analytical structure – as we shall see in a while, it turns out to be a *conic quadratic program* (CQP), i.e., a program with linear objective and the constraints of the type

$$a_i^{\mathrm{T}} x + \alpha_i \geqslant \|B_i x + b_i\|, \quad i = 1, \dots, M, \tag{13}$$

where α_i are fixed reals, a_i and b_i are fixed vectors, and B_i are fixed matrices of proper dimensions; $\|\cdot\|$ stands for the usual Euclidean norm. Recent progress in interior point methods (see, e.g., [3,4]) makes it possible to solve truly large-scale CQPs, so that "ellipsoidal uncertainty" leads to "practically solvable" robust counterparts $(P_{\mathcal{U}})$.

[3] Indeed, in order to get an efficient Separation oracle for a set $\mathcal{U}$ defined by finitely many convex constraints $g_i(A) \leqslant 0$ with "efficiently computable" $g_i(\cdot)$, it suffices, given A, to verify whether $g_i(A) \leqslant 0$ for all i: if it is the case, $A \in \mathcal{U}$, otherwise the (taken at A) subgradient of the violated constraint is a separator of A and $\mathcal{U}$.

3. Problem $(P_{\mathcal{U}})$ in the case of ellipsoidal uncertainty

3.1. Ellipsoids and ellipsoidal uncertainties

In geometry, a K-dimensional ellipsoid in $\mathbb{R}^K$ can be defined as an image of K-dimensional Euclidean ball under a one-to-one affine mapping from $\mathbb{R}^K$ to $\mathbb{R}^K$. For our purposes this definition is not that convenient. On one hand, we would like to consider "flat" ellipsoids in the space $E = \mathbb{R}^{m \times n}$ of data matrices A – usual ellipsoids in proper affine subspaces of E (such an ellipsoid corresponds to the case when we deal with "partial uncertainty", e.g., some of the entries in A are "certain"). On the other hand, we want to incorporate also "ellipsoidal cylinders" – sets of the type "sum of a flat ellipsoid and a linear subspace". The latter sets occur when we impose on A several ellipsoidal restrictions, each of them dealing with part of the entries; e.g., an "interval" $m \times n$ matrix A ($\mathcal{U}$ is given then by upper and lower bounds on the entries of the matrix). In order to cover all these cases, we define an ellipsoid in $\mathbb{R}^K$ as a set of the form

$$U = \{\Pi(u) \mid ||Qu|| \leqslant 1\}, \tag{14}$$

where $u \to \Pi(u)$ is an affine embedding of certain $\mathbb{R}^L$ into $\mathbb{R}^k$ and Q is an $M \times L$ matrix. This definition covers all previously discussed cases: when $L = M = K$ and Q is nonsingular. $\mathscr{U}$ is the standard K-dimensional ellipsoid in $\mathbb{R}^K$; "flat" ellipsoids correspond to the case when $L = M < K$ and Q is nonsingular; and the case when Q is singular corresponds to ellipsoidal cylinders.

From now on we shall say that $\mathscr{U} \in \mathbb{R}^{m\times n}$ is an *ellipsoidal uncertainty*, if

- A. $\mathscr{U}$ is given as an intersection of finitely many ellipsoids:

$$\mathscr{U} = \bigcap_{\ell=0}^{k} U(\Pi_\ell, Q_\ell) \tag{15}$$

  with explicitly given data Q_ℓ and $\Pi_\ell(\cdot)$;
- B. $\mathscr{U}$ is bounded;
- C. ["Slater condition"] there is at least one matrix $A \in \mathscr{U}$ which belongs to the "relative interior" of every ellipsoid $U(\Pi_\ell, Q_\ell)$, $\ell = 1,\dots,k$;

$$\forall \ell \leqslant k\ \exists u_\ell\colon A = \Pi_\ell(u^\ell)\ \&\ ||Q_\ell u_\ell|| < 1.$$

3.2. Problem $(P_{\mathscr{U}})$ for the case of ellipsoidal uncertainty U

Our current goal is to derive explicit representation of problem $(P_{\mathscr{U}})$ associated with the uncertainty set (15); as we shall see, in this case $(P_{\mathscr{U}})$ is a CQP.

3.2.1. Simplest cases

Let us start with the simplest cases where $\mathscr{U}$ is a usual ellipsoid or is a constraint-wise uncertainty with every constraint uncertainty set $\mathscr{U}_i$ being an ellipsoid.

$\mathscr{U}$ is a usual ellipsoid: $\mathscr{U} = \{A = P^0 + \sum_{j=1}^k u_j P^j \mid u^{\mathrm T} u \leqslant 1\}$, where P^j, $j = 0,\dots,k$, are $m \times n$ matrices.

Let us denote by $r_i^{(j)}$ ith row of P^j (recall that we always represent rows of a matrix by column vectors), and let R_i be the $n \times k$ matrix with the columns $r_i^{(1)},\dots,r_i^{(k)}$ so that ith row of $\Pi(u)$ is exactly $r_i^{(0)} + R_i u$. A point $x \in \mathbb{R}^n$ is r-feasible if and only if $f^{\mathrm T}x = 1$ and, for all $i = 1,\dots,m$, the inner product of ith row in $\Pi(u)$ and x is nonnegative whenever $||u|| \leqslant 1$, i.e.,

$$[r_i^{(0)}]^{\mathrm T}x + (R_i u)^{\mathrm T}x \geqslant 0, \forall(u, ||u|| \leqslant 1) \quad [\forall i = 1,\dots,m].$$

In other words, x is r-feasible if and only if $f^{\mathrm T}x = 1$ and

$$\min_{u:|u|\leqslant 1} [[r_i^{(0)}]^{\mathrm T}x + u^{\mathrm T}R_i^{\mathrm T}x] = [r_i^{(0)}]^{\mathrm T}x - ||R_i^{\mathrm T}x|| \geqslant 0, \quad i = 1,\dots,m.$$

We see that $(P_{\mathscr{U}})$ is nothing but the CQP

$$\min\{c^{\mathrm T}x \mid [r_i^{(0)}]^{\mathrm T}x \geqslant ||R_i^{\mathrm T}x||,\ i = 1,\dots,m;\ f^{\mathrm T}x = 1\} \tag{16}$$

$\mathscr{U} = \Pi_{i=1}^m \mathscr{U}_i$ is constraint-wise uncertainty with ellipsoids $\mathscr{U}_i$:

$$\mathscr{U}_i = \{A \mid A^{\mathrm T}e_i = r_i + R_i u^i \text{ for some } u^i \in \mathbb{R}^{L_i} \text{ with } ||u^i|| \leqslant 1,\ i = 1,\dots,m\};$$

here e_i are the standard orths in $\mathbb{R}^m, r_i \in \mathbb{R}^n$ and R_i are $n \times L_i$ matrices.

Exactly as above, we immediately conclude that in the case in question $(P_{\mathscr{U}})$ is the CQP

$$\min\{c^{\mathrm T}x \mid r_i^{\mathrm T}x \geqslant ||R_i^{\mathrm T}x||,\ i = 1,\dots,m;\ f^{\mathrm T}x = 1\}. \tag{17}$$

The case of general ellipsoidal uncertainty. In the case of general ellipsoidal uncertainty, the robust counterpart also turns out to be a *conic quadratic program*, i.e., a program with finitely many constraints of the form (13):

Theorem 3.1. *The robust counterpart $(P_{\mathscr{U}})$ of an uncertain LP problem with general ellipsoidal uncertainty can be converted to a conic quadratic program.*

Proof. See Appendix.

Note that, as it was already mentioned, conic quadratic problems can be solved by polynomial time interior point methods at basically the same computational complexity as LP problems of similar size.

Remark 3.1. *It was mentioned earlier that the case when the uncertainty set $\mathscr{U}$ is a polytope*, i.e.,

$$\mathscr{U} = \{u \in \mathbb{R}^k \mid d_i^{\mathrm T}u \leqslant r_i,\ i = 1,\dots,M\} \quad [d_i \neq 0] \tag{18}$$

is in fact a case of ellipsoidal uncertainty. This is not absolutely evident in advance: a polytope by definition is an intersection of finitely many half-spaces, and a half-space is not an ellipsoid – an ellipsoid should have a symmetry center. Nevertheless, a polytope is an intersection of finitely many "ellipsoidal cylinders". Indeed, since $\mathscr{U}$ *given by Eq.* (18) *is bounded, we can represent this set as an intersection of "stripes":* $\mathscr{U} = \{u \mid s_i \leqslant d_i^{\mathrm{T}} u \leqslant r_i,\ i = 1,\dots,M\}$, *ensuring the differences* $r_i - S_i$ *to be large enough. And a stripe is a very simple "ellipsoidal cylinder":*

$$\{u \in \mathbb{R}^k \mid s_i \leqslant d_i^{\mathrm{T}} u \leqslant r_i\} = U(p_i, I, Q_i),$$

where p_i *is an arbitrary point with* $d_i^{\mathrm{T}} p_i = (s_i + r_i)/2$, I *is the unit* $k \times k$ *matrix and* Q_i *is the* $1 \times k$ *matrix (i.e., a row vector) given by* $Q_i u = \frac{r_i - s_i}{2} d_i^{\mathrm{T}} u$. *Thus, a polytopic uncertainty is indeed ellipsoidal.*

4. Example

It is time now to discuss the following important issue: where could an ellipsoidal uncertainty come from? What could be the ways to define the ellipsoids constituting the uncertainty set?

One natural source of ellipsoidal uncertainty was already mentioned: approximation of more complicated uncertainty sets, which by themselves would lead to difficult robust counterparts of the uncertain problem. We can hardly say anything more on this issue – here everything depends on the particular situation we meet.

In case we are given several *primary scenarios* of the data, we could construct the uncertainty ellipsoid as the minimal volume ellipsoid containing these scenarios (or a small neighbourhood of this finite set).

There is, however, another source of ellipsoidal uncertainties which comes from statistical considerations. To explain it, let us look at the following example:

A simple portfolio problem. \$ 1 *is to be invested at the beginning of the year in a portfolio comprised of n shares. The end-of-the-year return per* \$ 1 *invested in share i is* $p_i > 0$. *At the end of the year you sell the portfolio. The goal is to determine the amount* x_i *to be invested in share i,* $i = 1,\dots,n$, *so as to maximize the end-of-the-year portfolio value* $\sum_{i=1}^n p_i x_i$.

When the quantities p_i are known in advance, the situation is modeled by the following simple LP program:

$$\max \left\{ \sum_{i=1}^n p_i x_i \,\middle|\, \sum_{i=1}^n x_i = 1,\ x_i \geqslant 0 \right\}, \tag{19}$$

and the optimal solution is evident: we should invest all we have in the "most promising" (with the largest p_i) share. Assuming that the coefficients p_i are distinct and arranged in ascending order: $p_1 < p_2 < \cdots < p_n$, the solution is

$$x_n = 1, \quad x_i = 0, \quad i = 1,\dots,n-1.$$

Now, what happens if the coefficients p_i are uncertain, as it is in reality? Assume that what we know are "nominal" values of these coefficients $p_i^*, p_1^* < p_2^* < \cdots < p_n^*$ and bounds $\sigma_i < p_i^*$ such that the actual values of p_i are within the "uncertainty intervals" $\Delta_i = [p_i^* - \sigma_i, p_i^* + \sigma_i]$. Assume, moreover, that p_i are of statistical nature and that they are mutually independent and distributed in Δ_i, symmetrically with respect to the nominal values p_i^*.

Under the indicated assumptions the simplest Stochastic Programming reformulation of the problem – the one where we are interested to maximize the expectation of the final portfolio value – is given by the same LP program (19) with p_i replaced by their "nominal" values p_i^*; the "nominal" solution is simply to invest all the money in the "most promising" share n ($x_n^* = 1, x_j^* = 0, j = 1,\dots,n-1$). This policy will result in random yield x^{nom} with the expected value

$$\mathscr{E}\{x^{\mathrm{nom}}\} = p_n^*. \tag{20}$$

Now let us look what could be done with the Robust Counterpart approach. The only question here is how to specify the uncertainty set, and the most straightforward answer is as follows: "in the situation in question, the vector p of uncertain coefficients runs through the box $B = \{(p_1,\dots,p_n) \mid |p_i - p_i^*| \leqslant \sigma_i, \forall i\}$, every point of the box being a possible value of the vector. Thus, the uncertainty set is B (here this is exactly the approach proposed by Soyster). The corresponding robust optimal policy would be to invest everything in the shares with the largest worst-case return $p_i^* - \sigma_i$. Of course, this policy is too conservative to be of actual interest. Note, however, that when applying the Robust Counterpart approach, we are not obliged to include into the uncertainty set *all* which may happen.

For the LP problem (19) and the underlying assumption on the uncertain coefficients p_i, we propose the uncertainty ellipsoidal set:

$$\mathscr{U}^\theta = \left\{ p \in \mathbb{R}^n \,\middle|\, \sum_{i=1}^n \sigma_i^{-2}(p_i - p_i^*)^2 \leqslant \theta^2 \right\}.$$

The parameter θ is a subjective value chosen by the decision maker to reflect his attitude towards risk; the larger is θ, the more risk averse he is. Note that for $\theta = 0, \mathscr{U}^\theta$ shrinks to the singleton $\mathscr{U}^0 = \{p^*\}$ – the nominal data; for $\theta = 1, \mathscr{U}^1$ is the largest volume ellipsoid contained in the box

$$B = \{p \mid |p_i - p_i^*| \leqslant \sigma_i, \ i = 1, \ldots, n\},$$

and for $\theta = \sqrt{n}, \mathscr{U}^\theta$ is the smallest volume ellipsoid containing the box.

Writing the LP problem (19) in the equivalent form

$$\max \left\{ y \,\middle|\, y \leqslant \sum_{i=1}^n p_i x_i, \sum_{i=1}^n x_1 = 1, \ x \geqslant 0 \right\}, \qquad (21)$$

we can use the result of Section 3.2.1 to derive the following robust counterpart of Eq. (21) with respect to the uncertainty set $\mathscr{U}^\theta$:

$$\max \left\{ \sum_{i=1}^n p_i^* x_i - \theta V^{1/2}(x) \,\middle|\, \sum_{i=1}^n x_i = 1, x \geqslant 0 \right\},$$

$$V(x) = \sum_{i=1}^n \sigma_i^2 x_i^2. \qquad (22)$$

Problem (22) resembles much the Markovitz approach to portfolio selection, although in this classical approach the role of $V^{1/2}(x)$ is played by $V(x)$.

The robust counterpart (22) can be motivated by reasoning as follows. Assume that we distribute the unit amount of money between the shares as $x = (x_1, \ldots, x_n)$. The corresponding yield $y = \sum_{i=1}^n p_i x_i$ can be expressed as

$$y = \sum_{i=1}^n p_i^* x_i + \zeta, \quad \zeta = \sum_{i=1}^n x_i[p_i - p_i^*], \qquad (23)$$

where the random part ζ has zero mean and variance

$$\mathrm{Var}(\zeta) = \sum_{i=1}^n (x_i)^2 \mathscr{E}\{(p_i - p_i^*)^2\}$$

$$\leqslant V(x) = \sum_{i=1}^n (x_i)^2 \sigma_i^2.$$

Consequently, the "typical" value of y will differ from the "nominal" value $\sum_{i=1}^m p_i^* x_i$ by a quantity of order of $\mathrm{Var}^{1/2}(\zeta) \leqslant V^{1/2}(x)$, variations of both signs being equally probable. A natural idea to handle the uncertainty is as follows. *Let us choose somehow a "reliability level" θ and ignore the events where the "noise" ζ is less than $-\theta V^{1/2}(x)$.*[4] *Among the remaining events we take the worst one – $\zeta = -\theta V^{1/2}(x)$, and act as if it was the only possible event*. With this approach, the "stable" yield of a decision x is $\sum_{i=1}^n p_i^* x_i - \theta V^{1/2}(x)$. These considerations lead precisely to the robust counterpart (22) obtained above.

It should be stressed that the uncertainty ellipsoid U^θ is in no sense an approximation of the support of the distribution of p. Assume, e.g., that p_i takes values $p_i^* \pm \sigma_i$ with probability 1/2 and that $\theta = 6$ (the probability to get $\sum_i p_i x_i < \sum_i p_i^* x_i - \theta V^{1/2}(x)$ is $< 10^{-7}$ for every x (see footnote 4)). Under these assumptions, for not too small n – namely, for $n > \theta^2 = 36$ – the ellipsoid U^6 does not contain a single realization of the random vector p!

Note also that the resulting uncertainty ellipsoid depends on the safety parameter θ. In applications, a decision maker could solve the robust counterparts of the problem for several values of θ and then choose the one which, in his opinion, results in the best tradeoff between "safety" and "greed".

The above considerations illustrate how one can specify ellipsoidal uncertainty sets, starting from stochastic data.

Now let us look at problem (22). In order to demonstrate what could be the effect of passing from the "nominal" program (19) with p_i set to their nominal values p_i^* to the problem (22), consider the following numerical example with $n = 150$. The nominal coefficients p_i^*, $i = 1, \ldots, n$, form an arithmetic progression:

---

[4] It can easily seen that $\Pr\{\zeta < -\theta V^{1/2}(x)\} < \exp\{-\theta^2/2\}$; for $\theta = 6$ the latter quantity is $< 10^{-7}$.

$$p_i^* = \lambda^* + i\delta, \quad \lambda^* = 1.15,$$

$$\delta = \frac{0.05}{150} \quad [p_1^* \approx 1.15, p_{150}^* = 1.2]$$

and the parameters σ_i are chosen to be

$$\sigma_i = \frac{1}{3}\delta\sqrt{2in(n+1)} \approx 0.0236\sqrt{i}$$

$$[\sigma_i = 0.0236, \sigma_{150} = 0.2896].$$

Note that σ_i are increasing, so that more promising investments are also more risky.

With the above data, we consider three candidate investment policies:

- the "nominal" one, $\mathcal{N}$, which is the optimal solution of the "nominal" program (20), and calls for investing all we have in the most promising share. For this policy the expected yield is $y^{\text{nom}} = 1.2$, and the standard deviation of the yield is 0.290. Also, with probability 0.5 we loose 9% of the initial captial, and with the same probability this capital is increased by 45%; one hardly could consider this unstable policy as appropriate; [5]
- the "robust counterpart" one, $\mathcal{RC}$, which is the optimal solution of problem (22) with $\theta = 1.5$. Here, due to our particular choice of σ_i, this policy is to invest equally in all the shares, i.e., $x_i = 1/n$, $i = 1, \ldots, n = 150$, the robust optimal value being 1.15; [6]
- the "robust" policy of Mulvey et al. [11]; this policy, $\mathcal{MVL}$, comes from the "robust optimal solution" given by methodology of [11], i.e., the optimal solution to the problem

$$\max\left\{ y - \frac{\mu}{N}\sum_{t=1}^{n} g(y - x^{\mathrm{T}} p^t) \,\middle|\, \sum_{i=1}^{n} x_i = 1, x_i \geqslant 0 \right\}, \tag{24}$$

  where $p^1, \ldots, p^N$ are given "scenarios" – realizations of the yield coefficients $p = (p_1, \ldots, p_n)$, and the function g is a penalty function for violating the "uncertain constraint" $y \leqslant x^{\mathrm{T}} p$ along the set of scenarios. The parameter $\mu > 0$ reflects the decision maker tradeoff between optimality and feasibility. A typical choice of g in [11], which is also used here, is $g(z) = z_+ \equiv \max[z, 0]$. In our experiments $\mu = 100$ and the scenario #t, $p^t = (p_1^t, \ldots, p_n^t)$, $t = 1, \ldots, N$, is chosen at random from the distribution corresponding to independent entries p_i^t taking with probability 1/2 the values $p_i^* \pm \sigma_i$.

In order to compare the stability properties of the second and the third policies (the characteristics of $\mathcal{N}$ policy are clear in advance, see above), we set the number of scenarios N in $\mathcal{MVL}$ to $1, 2, 4, 8, \ldots, 256$, then generated randomly the corresponding set of scenarios, and solved the resulting problem (24). Given the solution to the problem, we test it against the $\mathcal{RC}$-solution, running 400 simulations of the random yields $p = (p_1, \ldots, p_n)$ and comparing the random yields given by the policies in question.

The results of the comparison are given in Table 1. The headers in the table mean the following:

- N is the number of scenarios used in the $\mathcal{MVL}$-solution;
- prefixes `R`, `M` correspond to the $\mathcal{RC}$ and $\mathcal{MVL}$ policies, respectively;
- the meaning of the symbols `ov`, `mn`, `mx`, `av`, `sd`: `ov` – optimal value, `mn` – minimal observed yield, `mx` – maximal observed yield, `av` – expected yield, `sd` – standard deviation of yield, `ls` – empirical probability (in %) of yield < 1.

We see from the results in Table 1 that, as far as the standard deviation is concerned, the $\mathcal{RC}$-policy is about 15 times more stable than the nominal one; besides this, in all our $400 \times 9 = 3600$ simulations the $\mathcal{RC}$-policy never resulted in losses (and in fact always yielded at least 11% profit), while the nominal policy results in 9% losses with probability 0.5. As about the $\mathcal{MVL}$-policy, it is indeed more stable than the nominal one, although less stable than the $\mathcal{RC}$-policy. The stability properties of $\mathcal{MVL}$ heavily depend on the number of scenarios; with 16 scenarios, the policy in 7% of cases results in losses and has standard deviation of the yield 7.5 times larger than for the $\mathcal{RC}$-policy. As the number of scenarios grow, the stability properties of $\mathcal{MVL}$ approach those of

[5] Of course, there is nothing new in the phenomenon in question; OR financial models take care not only of the expected yields, but also of the variances and other characteristics of risk since the seminal work of Markovitz [10]. Note that in the particular example we treat here the robust counterpart solution resembles a lot the one given by the Markovitz approach.

[6] For different data, the $\mathcal{RC}$ policy gives different portions to the various shares, but still implements the principle "not to put all eggs in the same basket".

Table 1
Results of $\mathscr{RC}$ and $\mathscr{MVL}$ solutions

N	R_ov	M_ov	R_mn	M_mn	R_mx	M_mx	R_av	M_av	R_sd	M_sd	R_ls	M_ls
1	1.15	1.49	1.13	0.91	1.23	1.49	1.18	1.20	0.017	0.277	0.00	50.25
2	1.15	1.48	1.12	0.92	1.22	1.48	1.18	1.20	0.017	0.277	0.00	49.75
4	1.15	1.45	1.11	0.93	1.22	1.45	1.18	1.19	0.017	0.263	0.00	49.00
8	1.15	1.45	1.13	0.93	1.23	1.45	1.18	1.19	0.017	0.263	0.00	48.25
16	1.15	1.32	1.12	0.93	1.23	1.45	1.18	1.19	0.017	0.128	0.00	7.00
32	1.15	1.24	1.13	0.98	1.23	1.38	1.18	1.19	0.017	0.069	0.00	0.50
64	1.15	1.20	1.13	1.03	1.22	1.32	1.18	1.18	0.017	0.044	0.00	0.00
128	1.15	1.18	1.13	1.08	1.23	1.28	1.18	1.18	0.017	0.037	0.00	0.00
256	1.15	1.17	1.12	1.08	1.23	1.26	1.18	1.18	0.017	0.030	0.00	0.00

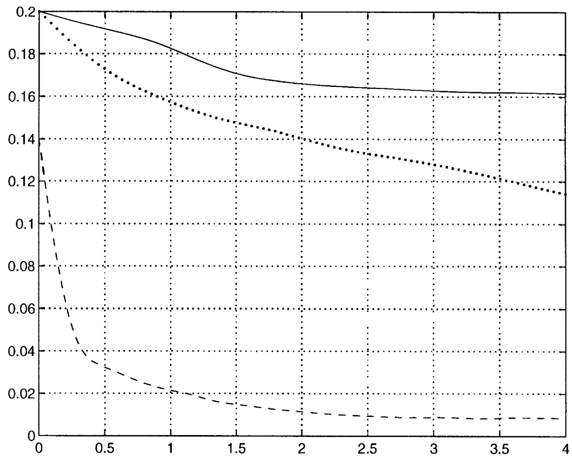

Fig. 1. Characteristics of the $\mathscr{RC}$ policy as functions of θ. dotted: $\rho(\theta)$; solid: $\pi(\theta)$; dashed: $v(\theta)$.

$\mathscr{RC}$, although the standard deviation for $\mathscr{MVL}$ remains 1.8 times worse than the one for $\mathscr{RC}$ even for 256 scenarios. Note also that for small numbers of scenarios $\mathscr{MVL}$ is "less conservative" than $\mathscr{RC}$ – results in a better average yield; this phenomenon, however, disappears already for 64 scenarios, and with this (and larger) # of scenarios $\mathscr{MVL}$ is simply dominated by $\mathscr{RC}$ – both policies result in the same average yield, but the first one has worse standard deviation.

We next study the behaviour of the optimal solution to the $\mathscr{RC}$ problem (22) (denoted by $x^*(\theta)$) as a function of the "safety parameter" θ. Recall that the ellipsoidal uncertainty is the larger the larger is θ. In particular, $x^*(0)$ is the nominal solution.

Three characteristics of the optimal solution are particularly important: – the *expected profit* $\pi(\theta) = \sum_{i=1}^n p_i^* x_i^*(\theta) - 1$, its *standard deviation* $v(\theta) = \sqrt{\sum_{i=1}^n \sigma_i^2 (x_i^*(\theta))^2}$ and the *net robust optimal value* $\rho(\theta)$ – the optimal value in Eq. (22) minus 1. As we can see from Fig. 1, both $\pi(\theta)$ and $\rho(\theta)$, as well as $v(\theta)$, vary slowly with θ, given $\theta \geqslant 1$. This implies that our nominal choice $\theta = 1.5$ is not crucial to the type of results obtained for $\mathscr{RC}$-policy and reported in Table 1.

The very simple single-stage model with independent returns discussed here can be extended to derive robust investment policies for multi-stage Portfolio problem with dependent returns, see [5].

Appendix Proof of Theorem 3.1

Let $\mathscr{U}$ be a general ellipsoidal uncertainty (15):

$$\mathscr{U} = \bigcap_{\ell=0}^{k} U(\Pi_\ell, Q_\ell),$$

$$U(\Pi_\ell, Q_\ell) = \{A = \Pi_\ell(u^\ell) \mid \|Q_\ell u^\ell\| \leqslant 1\}$$
$$[A \in \mathbb{R}^{m \times n}]$$

Let $a_i^{\mathrm{T}}[A]$ be the ith row of a matrix A. We first observe that

(I) *A point $x \in \mathbb{R}^n$ satisfying the normalization constraint $f^{\mathrm{T}} x = 1$ is robust feasible if and only if for every $i = 1, \ldots, m$ the optimal value in the optimization* *problem*

$$a_i[\Pi_0(u^0)]^{\mathrm T}x \to \min$$
$$\text{s.t.}\quad \Pi_\ell(u^\ell) = \Pi_0(u^0),\ \ \ell = 1,\dots,k,$$
$$||Q_\ell(u^\ell)|| \leqslant 1,\ \ \ell = 0,1,\dots,k \qquad (P_i[x])$$

with design variables $u^0,\dots,u^k$ *is non-negative.*

Indeed, it is readily seen that $\mathscr{U}$ is exactly the set of values attained by the matrix-valued function $\Pi_0(u^0)$ at the feasible set of $(P_1[x])$.

Now, $(P_i[x])$ is a conic quadratic optimization problem – one of the generic form

$$e^{\mathrm T}z + \phi \to \min$$
$$\text{s.t.}\quad Rz = r,$$
$$||A_\ell z - b_\ell|| \leqslant c_\ell^{\mathrm T} z - d_\ell,\ \ell = 0,\dots,k \qquad (\mathrm{CQP}_p)$$

(the design vector in the problem is z, while $A_\ell, b_\ell, c_\ell, d_\ell,\ \ell = 0,\dots,k, e, \phi$ are given matrices/vectors/scalars).

It is known [12, Chapter 4] that the Fenchel dual of (CQP_p) is again a conic quadratic problem, namely, the problem

$$r^{\mathrm T}\lambda + \sum_{\ell=0}^{k} [d_\ell v_\ell + b_\ell^{\mathrm T}\mu_\ell] + \phi \to \max$$
$$\text{s.t.}\quad R^{\mathrm T}\lambda + \sum_{\ell=1}^{k} [v_\ell c_\ell + A_\ell^{\mathrm T}\mu_\ell] = e,$$
$$||\mu_\ell|| \leqslant v_\ell,\ \ \ell = 0,\dots,k \qquad (\mathrm{CQP}_d)$$

with design variables $\lambda,\ \{v_\ell,\mu_\ell\}_{\ell=0}^k$ which are scalars (v_ℓ) or vectors (λ,μ_ℓ) of proper dimensions. Moreover, the Fenchel duality theorem in the case in question becomes the following statement (see [12], Theorem 4.2.1):

(II) *Let the primal problem* (CQP_p) *be strictly feasible* (*i.e., admitting a feasible solution* $\hat z$ *with* $||A_\ell\hat z - b_\ell|| < c_\ell^{\mathrm T}\hat z - d_\ell,\ \ell = 0,\dots,k$) *and let the objective of* (CQP_p) *be bounded below on the feasible set of the problem. Then the dual problem* (CQP_d) *is solvable, and the optimal values in the primal and the dual problems are equal to each other.*

Note that when (CQP_p) is the problem $(P_i[x])$ (so that only the objective in the problem depends on the "parameter" x, and this dependence is affine), then (CQP_d) becomes the problem of the form

$$[r^{(i)}]^{\mathrm T}\lambda^{(i)} + \sum_{\ell=0}^{k} [d_\ell^{(i)} v_\ell^{(i)} + [b_\ell^{(i)}]^{\mathrm T}\mu_\ell^{(i)}] + \phi^{(i)}[x] \to \max$$
$$\text{s.t.}\quad [R^{(i)}]^{\mathrm T}\lambda^{(i)} + \sum_{\ell=0}^{k} [v_\ell c_\ell^{(i)} + [A_\ell^{(i)}]^{\mathrm T}\mu_\ell^{(i)}] = e^{(i)}[x],$$
$$||\mu_\ell^{(i)}|| \leqslant v_\ell^{(i)},\ \ \ell = 0,\dots,k, \qquad (D_i[x])$$

where

- $\lambda^{(i)},\ \{\mu_\ell^{(i)}, v_\ell^{(i)}\}_{\ell=0}^k$ are the design variables of the problem,
- $e^{(i)}[x], \phi^{(i)}[x]$ are *affine* vector-valued, respectively, scalar functions of x,
- $r^{(i)}, R^{(i)}, \{A_\ell^{(i)}, b_\ell^{(i)}, c_\ell^{(i)}, d_\ell^{(i)}\}_{\ell=0}^k$ are *independent* of x matrices/vectors/scalars readily given by the coefficients of the affine mappings $\Pi_\ell(\cdot), Q_\ell(\cdot)$.

Now, it is readily seen that the assumptions B, C (see the definition of ellipsoidal uncertainty) ensure that the problem $(P_i[x])$ is strictly feasible and bounded below, whence by (II) the optimal value in $(P_i[x])$ is equal to one in $(D_i[x])$. Combining this observation with (I), we see that

(III) *A vector* $x \in \mathbb{R}^n$, $f^{\mathrm T}x = 1$, *is robust feasible, the uncertainty set being* $\mathscr{U}$, *if and only if for every* $i = 1,\dots,m$ *there exist* $\lambda^{(i)}, \{\mu_\ell^{(i)}, v_\ell^{(i)}\}_{\ell=0}^k$ *satisfying, along with* x, *the system of constraints*

$$[r^{(i)}]^{\mathrm T}\lambda^{(i)} + \sum_{\ell=0}^{k} [d_\ell^{(i)} v_\ell^{(i)} + [b_\ell^{(i)}]^{\mathrm T}\mu_\ell^{(i)}] + f^{(i)}[x] \geqslant 0,$$
$$[R^{(i)}]^{\mathrm T}\lambda^{(i)} + \sum_{\ell=0}^{k} [v_\ell c_\ell^{(i)} + [A_\ell^{(i)}]^{\mathrm T}\mu_\ell^{(i)}] = e^{(i)}[x],$$
$$||\mu_\ell^{(i)}|| \leqslant v_\ell^{(i)},\ \ell = 0,\dots,k. \qquad (\mathscr{C}_i)$$

We conclude that $(P_{\mathscr U})$ is equivalent to the problem

$$c^{\mathrm T}x \to \min$$
$$\text{s.t.}\quad (x, \lambda^{(i)}, \{\mu_\ell^{(i)}, v_\ell^{(i)}\}_{\ell=0}^k) \text{ satisfy } \mathscr{C}_i,\ \ i = 1,\dots,m,$$
$$f^{\mathrm T}x = 1, \qquad (\mathrm{CQP})$$

with design variables $x, \{\lambda^{(i)}, \{\mu_\ell^{(i)}, v_\ell^{(i)}\}_{\ell=0}^k\}_{i=1}^m$. Specifically, x is robust feasible if and only if it can be extended to a feasible solution of (CQP). To complete the proof of Theorem 3.1, it remains to note that (CQP) is a conic quadratic problem, and that the data specifying this problem are readily given by the coefficients of the affine mappings $\Pi_\ell(\cdot), Q_\ell(\cdot)$.

Remark 4.1. *The outlined proof demonstrates that Theorem* 3.1 *admits several useful modifications*, *in particular the following two*:

(i) *Assume that the uncertainty set* $\mathscr{U}$ *is polyhedral*:

$$\mathscr{U} = \{A \in \mathbb{R}^{m\times n} \mid \exists u \in \mathbb{R}^k\colon A = \Pi(u), Q(u) \geqslant 0\},$$

where $\Pi(\cdot) : \mathbb{R}^m \to \mathbb{R}^{m\times(n+1)}, Q(\cdot) : \mathbb{R}^k \to \mathbb{R}^l$ *are affine mappings. Then* $(P_{\mathscr{U}})$ *is equivalent to an LP program with sizes which are polynomial in* m, n, k, l *and the data readily given by the coefficients of the affine mappings* $\Pi(\cdot), Q(\cdot)$.

(ii) *Assume that the uncertainty set* $\mathscr{U}$ *is semidefinite-representable*:

$$\mathscr{U} = \{A \in \mathbb{R}^{m\times n} \mid \exists u \in \mathbb{R}^k\colon A = \Pi(u), Q(u) \succcurlyeq 0\},$$

where $\Pi(u)$ *is affine*, $Q(u)$ *is an affine mapping taking values in the space of symmetric matrices of a given row size* L *and* $B \succcurlyeq 0$ *means that* B *is a symmetric positive semidefinite matrix. Assume also that* $\mathscr{U}$ *is bounded*, *and that there exists* $\hat{u}$ *with positive definite* $Q(\hat{u})$. *Then* $(P_{\mathscr{U}})$ *is equivalent to a semidefinite program*, *i.e.*, *program of the form*

$$e^{\mathrm{T}}z \to \min$$

$$\text{s.t.} \quad A_0 + \sum_{i=1}^{\dim z} z_i A_i \succcurlyeq 0. \qquad \text{(SDP)}$$

The sizes of (SDP) (*i.e.*, $\dim z$ *and the row sizes of* A_i) *are polynomial in* m, n, k, l, *and the data* $e, A_0, \ldots, A_{\dim z}$ *of the problem are readily given by the coefficients of the affine mappings* $\Pi(\cdot), Q(\cdot)$.

References

[1] A. Ben-Tal, The entropic penalty approach to stochastic programming, Math. Oper. Res. 10 (1985) 263–279.

[2] A. Ben-Tal, A. Nemirovski, Robust truss topology design via semidefinite programming, SIAM J. Optim. 7 (1997) 991–1016.

[3] A. Ben-Tal, A. Nemirovski, Robust convex optimization, Math. Oper. Res. 23 (1998) 769–805.

[4] A. Ben-Tal, M. Zibulevsky, Penalty/barrier multiplier methods for convex programming problems, SIAM J. Optim. 7 (1997) 347–366.

[5] A. Ben-Tal, T. Margelit, A. Nemirovski, Robust modeling of multi-stage portfolio problems, in: Proc. Workshop on High-Performance Optimization, S. Zhang (Ed.), Rotterdam, August 1997, Kluwer Academic Press, Dordrecht, to appear, 1999.

[6] J.R. Birge, F. Louveaux, Introduction to Stochasatic Programming, Springer, Berlin, 1997.

[7] J.E. Falk, Exact solutions to inexact linear programs, Oper. Res. 1976, 783–787.

[8] M. Grötschel, L. Lovasz, A. Schrijver, The Ellipsoid Method and Combinatorial Optimization, Springer, Heidelberg, 1988.

[9] P. Kall, S.W. Wallace, Stochastic Programming, Wiley-Interscience, New York, 1994.

[10] H.M. Markovitz, Portfolio Selection: Efficiency Diversification of Investment, Wiley, New York, 1959.

[11] J.M. Mulvey, R.J. Vanderbei, S.A. Zenios, Robust optimization of large-scale systems, Oper. Res. 43 (1995) 264–281.

[12] Yu. Nesterov, A. Nemirovski, Interior Point Polynomial Algorithms in Convex Programming, SIAM Studies in Applied Mathematics, SIAM, Philadelphia, 1994.

[13] A. Prékopa, Stochastic Programming, Kluwer Academic Publishers, Dordrecht, 1995.

[14] R.T. Rockafellar, R.J.-B. Wets, Scenarios and policy aggregation in optimization under uncertainty, Math. Oper. Res. 16 (1991) 119–147.

[15] C. Singh, Convex programming with set-inclusive constraints and its applications to generalized linear and fractional programming, J. Optim. Theory Appl. 38 (1) (1982) 33–42.

[16] A.L. Soyster, Convex programming with set-inclusive constraints and applications to inexact linear programming, Oper. Res. 1973, 1154–1157.

[17] K. Zhou, J.C. Doyle, K. Glover, Robust and Optimal Control, Prentice-Hall, Englewood Cliffs, NJ, 1996.