

# Robust linear optimization under general norms

Dimitris Bertsimas<sup>a,\*</sup>, Dessislava Pachamanova<sup>b</sup>, Melvyn Sim<sup>c</sup>

<sup>a</sup>*Sloan School of Management, Massachusetts Institute of Technology, Building E52-363, 50 Memorial Drive, Cambridge, MA 02142-1347, USA*

<sup>b</sup>*Department of Mathematics and Sciences, Babson College, Forest Street, Babson Park, MA 02457, USA*

<sup>c</sup>*NUS Business School, National University of Singapore, 1 Business Link, Singapore 117592, Singapore*

Received 28 August 2003; received in revised form 2 December 2003; accepted 19 December 2003

## Abstract

We explicitly characterize the robust counterpart of a linear programming problem with uncertainty set described by an arbitrary norm. Our approach encompasses several approaches from the literature and provides guarantees for constraint violation under probabilistic models that allow arbitrary dependencies in the distribution of the uncertain coefficients.

© 2004 Elsevier B.V. All rights reserved.

*Keywords:* Robust optimization; Stochastic programming; Linear programming; Norms; Dual norms; Constraint violation

## 1. Introduction

Robust optimization is emerging as a leading methodology to address optimization problems under uncertainty. Applied to the linear programming problem

$$\max \{ \mathbf{c}'\mathbf{x} \mid \tilde{\mathbf{A}}\mathbf{x} \leq \mathbf{b}, \mathbf{x} \in P^x \} \quad (1)$$

with  $\mathbf{x} \in \mathcal{R}^{n \times 1}$ ,  $\tilde{\mathbf{A}} \in \mathcal{R}^{m \times n}$  a matrix of uncertain coefficients belonging to a known uncertainty set  $\mathcal{U}$ ,  $\mathbf{c} \in \mathcal{R}^{n \times 1}$  and  $P^x$  a given set representing the constraints involving certain coefficients, the *robust counterpart* of Problem (1) is

$$\max \{ \mathbf{c}'\mathbf{x} \mid \tilde{\mathbf{A}}\mathbf{x} \leq \mathbf{b}, \mathbf{x} \in P^x, \forall \tilde{\mathbf{A}} \in \mathcal{U} \}. \quad (2)$$

An optimal solution  $\mathbf{x}^*$  is robust with respect to any realization of the data, that is, it satisfies the constraints for any  $\tilde{\mathbf{A}} \in \mathcal{U}$ .

The first approach to modeling coefficient uncertainty in the optimization literature was proposed by Soyster [12]. He considers problems of the kind

$$\max \left\{ \mathbf{c}'\mathbf{x} \mid \sum_{j=1}^n \tilde{\mathbf{A}}_j \mathbf{x}_j \leq \mathbf{b}, \forall \tilde{\mathbf{A}}_j \in K_j, j=1, \dots, n, \mathbf{x} \geq 0 \right\},$$

where  $\mathbf{x} \in \mathcal{R}^{n \times 1}$ ,  $\mathbf{c} \in \mathcal{R}^{n \times 1}$  and  $\mathbf{b} \in \mathcal{R}^{m \times 1}$  are data, and each column  $\tilde{\mathbf{A}}_j$  of the matrix  $\tilde{\mathbf{A}}$  belongs to a convex set  $K_j$ . Soyster shows that the problem is equivalent to a problem with linear constraints

$$\max \left\{ \mathbf{c}'\mathbf{x} \mid \sum_{j=1}^n \bar{\mathbf{A}}_j \mathbf{x}_j \leq \mathbf{b}, \mathbf{x} \geq 0 \right\},$$

\* Corresponding author.

E-mail addresses: [dbertsim@mit.edu](mailto:dbertsim@mit.edu) (D. Bertsimas),

[dpachamanova@babson.edu](mailto:dpachamanova@babson.edu) (D. Pachamanova),

[dscsim@nus.edu.sg](mailto:dscsim@nus.edu.sg) (M. Sim).

<sup>1</sup> Research supported by the Singapore-MIT Alliance.

where the entries of the matrix  $\tilde{\mathbf{A}}$ ,  $\tilde{a}_{ij}$ , satisfy  $\tilde{a}_{ij} = \sup_{\tilde{\mathbf{A}}_j \in \mathcal{K}_j} (\tilde{\mathbf{a}}_{ij})$ . Some work followed Soyster’s note (see, for example, [8]). However, the approaches suggested in this early literature solve a limited range of problems, are not easy to generalize, and are very conservative in the sense that they protect against the worst-case scenario.

The interest in robust formulations in the optimization community was revived in the 1990s. Ben-Tal and Nemirovski [1–4] introduced a number of important formulations and applications, and provided a detailed analysis of the robust optimization framework in linear programming and general convex programming. Independently, El Ghaoui et al. [9,10] derived similar results, and dealt with robust formulations of optimization problems by adapting robust control techniques. They showed that under the assumption that the coefficient matrix data vary in *ellipsoidal* uncertainty sets, the robust counterparts of some important generic convex optimization problems (linear programming problems (LP), second-order cone problems (SOCP), semidefinite programming problems (SDP)) are either exactly, or approximately tractable problems that are efficiently solvable via interior point methods. However, under ellipsoidal uncertainty sets, the robust counterpart of an LP becomes an SOCP, of an SOCP becomes an SDP, while the robust counterpart of an SDP is NP-hard to solve. In other words, the difficulty of the robust problem increases, as SDPs are numerically harder to solve than SOCPs, which in turn are harder to solve than LPs.

Bertsimas and Sim [7] consider LPs such that each entry  $\tilde{a}_{ij}$  of  $\tilde{\mathbf{A}} \in \mathcal{R}^{m \times n}$  is assumed to take values in the interval  $[\tilde{a}_{ij} - \Delta_{ij}, \tilde{a}_{ij} + \Delta_{ij}]$  and protect for the case that up to  $\Gamma_i$  of the uncertain coefficients in constraint  $i$ ,  $i = 1, \dots, m$ , can take their worst-case values at the same time. The parameter  $\Gamma_i$  controls the tradeoff between robustness and optimality. Bertsimas and Sim show that the robust counterpart is still an LP.

In this paper, we propose a framework for robust modeling of linear programming problems using uncertainty sets described by an arbitrary norm. We explicitly characterize the robust counterpart as a convex optimization problem that involves the dual norm of the given norm. Under a Euclidean norm we recover the second order cone formulation in [1,2,9,10], while under a particular  $D$ -norm we introduce we recover the linear programming formulation proposed

in [7]. In this way, we shed some new light to the nature and structure of the robust counterpart of an LP.

The paper is structured as follows. In Section 2, we review from [5] the robust counterpart of a linear programming problem when the deviations of the uncertain coefficients lie in a convex set and characterize the robust counterpart of an LP when the uncertainty set is described by a general norm, as a convex optimization problem that involves the dual norm of the given norm. In Section 3, we show that by varying the norm used to define the uncertainty set, we recover the second order cone formulation in [1,2,9,10], while under a particular  $D$ -norm we introduce we recover the linear programming formulation proposed in [7]. In Section 4, we provide guarantees for constraint violation under general probabilistic models that allow arbitrary dependencies in the distribution of the uncertain coefficients.

**Notation.** In this paper, lowercase boldface will be used to denote vectors, while uppercase boldface will denote matrices. Tilde ( $\tilde{\mathbf{a}}$ ) will denote uncertain coefficients, while check ( $\check{\mathbf{a}}$ ) will be used for nominal values.  $\tilde{\mathbf{A}} \in \mathcal{R}^{m \times n}$  will usually be the matrix of uncertain coefficients in the linear programming problem, and  $\text{vec}(\tilde{\mathbf{A}}) \in \mathcal{R}^{(m \cdot n) \times 1}$  will denote the vector obtained by stacking its rows on top of one another.

## 2. Uncertainty sets defined by a norm

In this section, we review from [5] the structure of the robust counterpart for uncertainty sets defined by general norms. These characterizations are used to develop the new characterizations in Section 3.

Let  $S$  be a closed, bounded convex set and consider an uncertainty set in which the uncertain coefficients are allowed to vary in such a way that the deviations from their nominal values fall in a convex set  $S$

$$\mathcal{U} = \{\tilde{\mathbf{A}} | (\text{vec}(\tilde{\mathbf{A}}) - \text{vec}(\mathbf{A})) \in S\}.$$

The next theorem characterizes the robust counterpart.

### Theorem 1. Problem

$$\begin{aligned} \max \quad & \mathbf{c}'\mathbf{x} \\ \text{s.t.} \quad & \tilde{\mathbf{A}}\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \in P^x \\ & \forall \tilde{\mathbf{A}} \in \mathcal{R}^{m \times n} \text{ such that } (\text{vec}(\tilde{\mathbf{A}}) - \text{vec}(\mathbf{A})) \in S \end{aligned} \tag{3}$$

can be formulated as

$$\begin{aligned} \max \quad & \mathbf{c}'\mathbf{x} \\ \text{s.t.} \quad & \check{\mathbf{a}}_i\mathbf{x} + \max_{\mathbf{y} \in S} \{\mathbf{y}'\mathbf{x}\} \leq \mathbf{b}_i, \quad i = 1, \dots, m \\ & \mathbf{x} \in P^x. \end{aligned} \quad (4)$$

**Proof.** Clearly since  $S$  is compact, for each constraint  $i$ ,  $\check{\mathbf{a}}_i\mathbf{x} \leq b_i$  for all  $\text{vec}(\check{\mathbf{A}}) - \text{vec}(\check{\mathbf{A}}) \in S$  if and only if

$$\max_{\{\text{vec}(\check{\mathbf{A}}) - \text{vec}(\check{\mathbf{A}}) \in S\}} \{\check{\mathbf{a}}_i'\mathbf{x}\} \leq b_i.$$

Since

$$\begin{aligned} \max_{\{\text{vec}(\check{\mathbf{A}}) - \text{vec}(\check{\mathbf{A}}) \in S\}} \{\check{\mathbf{a}}_i'\mathbf{x}\} &= \max_{\{\text{vec}(\check{\mathbf{A}}) - \text{vec}(\check{\mathbf{A}}) \in S\}} \{(\text{vec}(\check{\mathbf{A}}))'\mathbf{x}_i\} \\ &= (\text{vec}(\check{\mathbf{A}}))'\mathbf{x}_i + \max_{\{\mathbf{y} \in S\}} \{\mathbf{y}'\mathbf{x}\}, \end{aligned}$$

the theorem follows.  $\square$

We next consider uncertainty sets that arise from the requirement that the distance (as measured by an arbitrary norm) between uncertain coefficients and their nominal values is bounded. Specifically, we consider an uncertainty set  $\mathcal{U}$  given by

$$\mathcal{U} = \{\check{\mathbf{A}} \mid \|\mathbf{M}(\text{vec}(\check{\mathbf{A}}) - \text{vec}(\check{\mathbf{A}}))\| \leq \Delta\}, \quad (5)$$

where  $\mathbf{M}$  is an invertible matrix,  $\Delta \geq 0$  and  $\|\cdot\|$  a general norm.

Given a norm  $\|\mathbf{x}\|$  for a real space of vectors  $\mathbf{x}$ , its dual norm induced over the dual space of linear functionals  $\mathbf{s}$  is defined as follows:

**Definition 1** (Dual Norm).

$$\|\mathbf{s}\|^* \doteq \max_{\{\|\mathbf{x}\| \leq 1\}} \mathbf{s}'\mathbf{x}. \quad (6)$$

The next result is well known (see, for example, [11]).

**Proposition 1.** (a) *The dual norm of the dual norm is the original norm.*

(b) *The dual norm of the  $L_p$  norm*

$$\|\mathbf{x}\|_p \doteq \left( \sum_{j=1}^n |x_j|^p \right)^{1/p} \quad (7)$$

is the  $L_q$  norm  $\|\mathbf{s}\|_q$  with  $q = 1 + 1/(p - 1)$ .

The next theorem derives the form of the robust counterpart, when the uncertainty set is given by Eq. (5).

**Theorem 2.** *Problem*

$$\begin{aligned} \max \quad & \mathbf{c}'\mathbf{x} \\ \text{s.t.} \quad & \check{\mathbf{A}}\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \in P^x \\ & \forall \check{\mathbf{A}} \in \mathcal{U} = \{\check{\mathbf{A}} \mid \|\mathbf{M}(\text{vec}(\check{\mathbf{A}}) - \text{vec}(\check{\mathbf{A}}))\| \leq \Delta\}, \end{aligned} \quad (8)$$

where  $\mathbf{M}$  is an invertible matrix, can be formulated as

$$\begin{aligned} \max \quad & \mathbf{c}'\mathbf{x} \\ \text{s.t.} \quad & \check{\mathbf{a}}_i\mathbf{x} + \Delta \|\mathbf{M}'^{-1}\mathbf{x}_i\|^* \leq \mathbf{b}_i, \quad i = 1, \dots, m \\ & \mathbf{x} \in P^x, \end{aligned} \quad (9)$$

where  $\mathbf{x}_i \in \mathfrak{R}^{(m-n) \times 1}$  is a vector that contains  $\mathbf{x} \in \mathfrak{R}^{n \times 1}$  in entries  $(i - 1) \cdot n + 1$  through  $i \cdot n$ , and 0 everywhere else.

**Proof.** Introducing  $\mathbf{y} = (\mathbf{M}(\text{vec}(\check{\mathbf{A}}) - \text{vec}(\check{\mathbf{A}})))/\Delta$ , we can write  $\mathcal{U}$  as  $\mathcal{U} = \{\mathbf{y} \mid \|\mathbf{y}\| \leq 1\}$  and obtain

$$\begin{aligned} \max_{\{\text{vec}(\check{\mathbf{A}}) \in \mathcal{U}\}} \{\check{\mathbf{a}}_i'\mathbf{x}\} &= \max_{\{\text{vec}(\check{\mathbf{A}}) \in \mathcal{U}\}} \{(\text{vec}(\check{\mathbf{A}}))'\mathbf{x}_i\} \\ &= \max_{\{\mathbf{y} \mid \|\mathbf{y}\| \leq 1\}} \{(\text{vec}(\check{\mathbf{A}}))'\mathbf{x}_i + \Delta(\mathbf{M}^{-1}\mathbf{y})'\mathbf{x}\} \\ &= \check{\mathbf{a}}_i'\mathbf{x} + \Delta \max_{\{\mathbf{y} \mid \|\mathbf{y}\| \leq 1\}} \{\mathbf{y}'(\mathbf{M}'^{-1}\mathbf{x})\}. \end{aligned}$$

By Definition 1, the second term in the last expression is  $\Delta \|\mathbf{M}'^{-1}\mathbf{x}\|^*$ . The theorem follows by applying Theorem 1.  $\square$

In applications, [1,2,9,10] work with uncertainty sets given by the Euclidean norm, i.e.,

$$\mathcal{U} = \{\check{\mathbf{A}} \mid \|\mathbf{M}(\text{vec}(\check{\mathbf{A}}) - \text{vec}(\check{\mathbf{A}}))\|_2 \leq \Delta\},$$

where  $\|\cdot\|_2$  denotes the Euclidean norm. Since the Euclidean norm is self dual, it follows that the robust counterpart is a second order cone problem. If the uncertainty set is described by either  $\|\cdot\|_1$  or  $\|\cdot\|_\infty$

(these norms are dual to each other), then the resulting robust counterpart can be formulated as an LP.

**3. The  $D$ -norm and its dual**

In this section, we show that the approach of [7] follows also from Theorem 2 by simply considering a different norm, called the  $D$ -norm and its dual as opposed to the Euclidean norm considered in [1,2,9,10]. Furthermore, we show worst case bounds on the proximity of the  $D$ -norm to the Euclidean norm.

Bertsimas and Sim [7] consider uncertainty sets given by

$$\| \| \mathbf{M}(\text{vec}(\tilde{\mathbf{A}}) - \text{vec}(\tilde{\mathbf{A}})) \| \|_p \leq \Delta$$

with  $p \in [1, n]$  and for  $\mathbf{y} \in \mathfrak{R}^{n \times 1}$

$$\| \| \mathbf{y} \| \|_p = \max_{\{S \cup \{t\} | S \subseteq N, |S| \leq \lfloor p \rfloor, t \in N \setminus S\}} \left\{ \sum_{j \in S} |y_j| + (p - \lfloor p \rfloor) |y_t| \right\}.$$

The fact that  $\| \| \mathbf{y} \| \|_p$  is indeed a norm, i.e.,  $\| \| \mathbf{y} \| \|_p \geq 0$ ,  $\| \| c\mathbf{y} \| \|_p = |c| \cdot \| \| \mathbf{y} \| \|_p$ ,  $\| \| \mathbf{y} \| \|_p = 0$  if and only if  $\mathbf{y} = \mathbf{0}$ , and  $\| \| \mathbf{x} + \mathbf{y} \| \|_p \leq \| \| \mathbf{x} \| \|_p + \| \| \mathbf{y} \| \|_p$  follows easily. Specifically [7] considers constraint-wise uncertainty sets,  $\mathbf{M}$  a diagonal matrix containing the inverses of the ranges of coefficient variation, and  $\Delta = 1$ . We next derive the dual norm.

**Proposition 2.** *The dual norm of the norm  $\| \| \cdot \| \|_p$  is given by*

$$\| \| \mathbf{s} \| \|_p^* = \max(\| \mathbf{s} \| \|_\infty, \| \mathbf{s} \| \|_1 / p).$$

**Proof.** The norm  $\| \| \mathbf{y} \| \|_p$  can be written as

$$\begin{aligned} \| \| \mathbf{y} \| \|_p &= \max \sum_{j=1}^n u_j y_j \\ \text{s.t. } \sum_{j=1}^n u_j &\leq p, 0 \leq u_j \leq 1 \\ &= \min \quad pr + \sum_{j=1}^n t_j \end{aligned}$$

$$\begin{aligned} \text{s.t. } r + t_j &\geq |y_j|, t_j \geq 0, \\ j &= 1, \dots, n, r \geq 0, \end{aligned}$$

where the second equality follows by linear programming strong duality. Thus,  $\| \| \mathbf{y} \| \|_p \leq 1$  if and only if

$$\begin{aligned} pr + \sum_{j=1}^n t_j &\leq 1, r + t_j \geq |y_j|, t_j \geq 0, \\ j &= 1, \dots, n, r \geq 0 \end{aligned} \tag{10}$$

is feasible. The dual norm  $\| \| \mathbf{s} \| \|_p^*$  is given by

$$\| \| \mathbf{s} \| \|_p^* = \max_{\| \| \mathbf{y} \| \|_p \leq 1} \mathbf{s}'\mathbf{y}.$$

From Eq. (10) we have that

$$\begin{aligned} \| \| \mathbf{s} \| \|_p^* &= \max \quad \mathbf{s}'\mathbf{y} \\ \text{s.t. } pr + \sum_{j=1}^n t_j &\leq 1, \\ y_j - t_j - r &\leq 0, \quad j = 1, \dots, n, \\ -y_j - t_j - r &\leq 0, \quad j = 1, \dots, n, \\ r &\geq 0, t_j \geq 0, \quad j = 1, \dots, n. \end{aligned}$$

Applying LP duality again we obtain

$$\begin{aligned} \| \| \mathbf{s} \| \|_p^* &= \min \quad \theta \\ \text{s.t. } p\theta - \sum_{j=1}^n u_j - \sum_{j=1}^n v_j &\geq 0, \\ \theta - u_j - v_j &\geq 0, \quad j = 1, \dots, n, \\ u_j - v_j &= s_j, \quad j = 1, \dots, n, \\ \theta &\geq 0, u_j, v_j \geq 0, \quad j = 1, \dots, n. \end{aligned}$$

Thus,

$$\begin{aligned} \| \| \mathbf{s} \| \|_p^* &= \min \quad \theta \\ \text{s.t. } \theta &\geq |s_j|, \quad j = 1, \dots, n, \\ \theta &\geq \sum_{j=1}^n |s_j| / p, \end{aligned}$$

and hence,

$$\| \| \mathbf{s} \| \|_p^* = \max(\| \mathbf{s} \| \|_\infty, \| \mathbf{s} \| \|_1 / p). \quad \square$$

Combining Theorem 2 and Proposition 2 leads to an LP formulation of the robust counterpart of the

uncertainty set in [7]. We thus observe that Theorem 2 provides a unified treatment of the approaches in [1,2,7,9,10].

3.1. Comparison with the Euclidean norm

Since uncertainty sets in the literature have been described using the Euclidean and the  $D$ -norm it is of interest to understand the proximity between these two norms.

**Proposition 3.** For every  $\mathbf{y} \in \mathfrak{R}^n$

$$\min \left\{ 1, \frac{p}{\sqrt{n}} \right\} \leq \frac{\|\mathbf{y}\|_p}{\|\mathbf{y}\|_2} \leq \sqrt{\lfloor p \rfloor + (p - \lfloor p \rfloor)^2}$$

$$\min \left\{ \frac{1}{p}, \frac{1}{\sqrt{n}} \right\} \leq \frac{\|\mathbf{y}\|_p^*}{\|\mathbf{y}\|_2} \leq \max \left\{ \frac{\sqrt{n}}{p}, 1 \right\}.$$

**Proof.** We find a lower bound on  $\|\mathbf{y}\|_p / \|\mathbf{y}\|_2$  by solving the following nonlinear optimization problem:

$$\max \sum_{j \in N} x_j^2 \tag{11}$$

s.t.  $\|\mathbf{x}\|_p = 1.$

Let  $S = \{1, \dots, \lfloor p \rfloor\}$ ,  $t = \lfloor p \rfloor + 1$ . We can impose nonnegativity constraints on  $\mathbf{x}$  and the constraints that  $x_j \geq x_t, \forall j \in S$  and  $x_j \leq x_t, \forall j \in N \setminus S$ , without affecting the objective value of the Problem (11). Observing that the objective can never decrease if we let  $x_j = x_t, \forall j \in N \setminus S$ , we reduce (11) to the following problem:

$$\max \sum_{j \in S} x_j^2 + (n - \lfloor p \rfloor)x_t^2$$

s.t.  $\sum_{j \in S} x_j + (p - \lfloor p \rfloor)x_t = 1,$  (12)

$x_j \geq x_t \quad \forall j \in S,$

$x_t \geq 0.$

Since we are maximizing a convex function over a polytope, there exist an extreme point optimal solution to Problem (12). There are  $|S| + 1$  inequality constraints. Out of those,  $|S|$  need to be tight to establish an extreme point solution. The  $|S| + 1$  extreme points

can be found to be

$$\mathbf{x}^k = \mathbf{e}_k \quad \forall k \in S, \tag{13}$$

$$\mathbf{x}^{|S|+1} = \frac{1}{p} \mathbf{e}, \tag{14}$$

where  $\mathbf{e}$  is the vector of ones and  $\mathbf{e}_k$  is the unit vector with the  $k$ th element equals one, and the rest equal zero. Clearly, substituting all possible solutions, Problem (12) yields the optimum value of  $\max\{1, n/p^2\}$ . Taking the square root, the inequality  $\|\mathbf{y}\|_2 \leq \max\{1, \sqrt{n/p}\} \|\mathbf{y}\|_p$  follows.

Similarly, in order to derive an upper bound of  $\|\mathbf{y}\|_p / \|\mathbf{y}\|_2$ , we solve the following problem:

$$\min \sum_{j \in N} x_j^2 \tag{15}$$

s.t.  $\|\mathbf{x}\|_p = 1.$

Using the same partition of the solution as before, and observing that the objective can never increase with  $x_j = 0, \forall j \in N \setminus S \setminus \{t\}$ , we reduce Problem (15) to the following problem:

$$\min \sum_{j \in S} x_j^2 + x_t^2$$

s.t.  $\sum_{j \in S} x_j + (p - \lfloor p \rfloor)x_t = 1,$  (16)

$x_j \geq x_t \quad \forall j \in S,$

$x_t \geq 0.$

An optimal solution of Problem (16) can be found using the KKT conditions:

$$x_j = \begin{cases} \frac{1}{\lfloor p \rfloor + (p - \lfloor p \rfloor)^2} & \text{if } j \in S, \\ \frac{p - \lfloor p \rfloor}{\lfloor p \rfloor + (p - \lfloor p \rfloor)^2} & \text{if } j = t, \\ 0 & \text{otherwise.} \end{cases}$$

Substituting, the optimal objective value of Problem (16) is  $(\lfloor p \rfloor + (p - \lfloor p \rfloor)^2)^{-1}$ . Hence, taking the square root, we establish the inequality  $\|\mathbf{y}\|_p \leq \sqrt{\lfloor p \rfloor + (p - \lfloor p \rfloor)^2} \|\mathbf{y}\|_2$ .

Since

$$1 \leq \frac{\|\mathbf{y}\|_1}{\|\mathbf{y}\|_2} \leq \sqrt{n}, \frac{1}{\sqrt{n}} \leq \frac{\|\mathbf{y}\|_\infty}{\|\mathbf{y}\|_2} \leq 1,$$

and

$$\| \| y \| \|_p^* = \max \left\{ \frac{\| y \|_1}{p}, \| y \|_\infty \right\}$$

the bounds follow.  $\square$

#### 4. Probabilistic guarantees

In this section, we derive probabilistic guarantees on the feasibility of an optimal robust solution when the uncertainty set  $\mathcal{U}$  is described by a general norm  $\| \cdot \|$  with a dual norm  $\| \cdot \|_*$ .

We assume that the matrix  $\tilde{\mathbf{A}}$  has an arbitrary (and unknown) distribution, but with known expected value  $\mathbf{A} \in \mathfrak{R}^{m \times n}$  and known covariance matrix  $\Sigma \in \mathfrak{R}^{(m \cdot n) \times (m \cdot n)}$ . Note that we allow arbitrary dependencies on  $\tilde{\mathbf{A}}$ . We define  $\mathbf{M} = \Sigma^{-1/2}$ , which is a symmetric matrix.

Let  $\mathbf{x}^* \in \mathfrak{R}^{n \times 1}$  be an optimal solution to Problem (9). Recall that  $\mathbf{x}_i^* \in \mathfrak{R}^{(m \cdot n) \times 1}$  denotes the vector containing  $\mathbf{x}^*$  in entries  $(i - 1) \cdot n$  through  $i \cdot n$ , and 0 everywhere else.

**Theorem 3.** (a) *The probability that  $\mathbf{x}^*$  satisfies constraint  $i$  for any realization of the uncertain matrix  $\tilde{\mathbf{A}}$  is*

$$\begin{aligned} P(\tilde{\mathbf{a}}_i' \mathbf{x}^* \leq b_i) &= P((\text{vec}(\tilde{\mathbf{A}}))' \mathbf{x}_i^* \leq b_i) \\ &\geq 1 - \frac{1}{1 + \Delta^2 (\| \Sigma^{1/2} \mathbf{x}_i^* \|_* / \| \Sigma^{1/2} \mathbf{x}_i^* \|_2)^2}. \end{aligned} \quad (17)$$

(b) *If the norm used in  $\mathcal{U}$  is the  $D$ -norm  $\| \cdot \|_p$ , then*

$$P(\tilde{\mathbf{a}}_i' \mathbf{x}^* \leq b_i) \geq 1 - \frac{1}{1 + \Delta^2 \min \left\{ \frac{1}{p^2}, \frac{1}{n} \right\}}. \quad (18)$$

(c) *If the norm used in  $\mathcal{U}$  is the dual  $D$ -norm  $\| \cdot \|_p^*$ , then*

$$P(\tilde{\mathbf{a}}_i' \mathbf{x}^* \leq b_i) \geq 1 - \frac{1}{1 + \Delta^2 \min \left\{ 1, \frac{p^2}{n} \right\}}. \quad (19)$$

(d) *If the norm used in  $\mathcal{U}$  is the Euclidean norm, then*

$$P(\tilde{\mathbf{a}}_i' \mathbf{x}^* \leq b_i) \geq 1 - \frac{1}{1 + \Delta^2}. \quad (20)$$

**Proof.** Since an optimal robust solution satisfies

$$(\text{vec}(\tilde{\mathbf{A}}))' \mathbf{x}_i^* + \Delta \| \Sigma^{1/2} \mathbf{x}_i^* \|_* \leq b_i,$$

we obtain that

$$\begin{aligned} P((\text{vec}(\tilde{\mathbf{A}}))' \mathbf{x}_i^* > b_i) &\leq P((\text{vec}(\tilde{\mathbf{A}}))' \mathbf{x}_i^* \geq (\text{vec}(\tilde{\mathbf{A}}))' \mathbf{x}_i^* \\ &\quad + \| \Sigma^{1/2} \mathbf{x}_i^* \|_*). \end{aligned} \quad (21)$$

Bertsimas and Popescu [6] show that if  $S$  is a convex set, and  $\tilde{\mathbf{X}}$  is a vector of random variables with mean  $\tilde{\mathbf{X}}$  and covariance matrix  $\Sigma$ , then

$$P(\tilde{\mathbf{X}} \in S) \leq \frac{1}{1 + d^2}, \quad (22)$$

where

$$d^2 = \inf_{\tilde{\mathbf{X}} \in S} (\tilde{\mathbf{X}} - \tilde{\mathbf{X}})' \Sigma^{-1} (\tilde{\mathbf{X}} - \tilde{\mathbf{X}}).$$

We consider the convex set  $S_i = \{ \text{vec}(\tilde{\mathbf{A}}) | (\text{vec}(\tilde{\mathbf{A}}))' \mathbf{x}_i^* > (\text{vec}(\tilde{\mathbf{A}}))' \mathbf{x}_i^* + \Delta \| \Sigma^{1/2} \mathbf{x}_i^* \|_* \}$ . In our case,

$$\begin{aligned} d_i^2 &= \inf_{\text{vec}(\tilde{\mathbf{A}}) \in S_i} (\text{vec}(\tilde{\mathbf{A}}) - \text{vec}(\tilde{\mathbf{A}}))' \Sigma^{-1} (\text{vec}(\tilde{\mathbf{A}}) \\ &\quad - \text{vec}(\tilde{\mathbf{A}})). \end{aligned}$$

Applying the KKT conditions for this optimization problem we obtain an optimal solution

$$\text{vec}(\tilde{\mathbf{A}}) + \Delta \left( \frac{\| \Sigma^{1/2} \mathbf{x}_i^* \|_*}{\| \Sigma^{1/2} \mathbf{x}_i^* \|_2} \right)^2 \Sigma \mathbf{x}_i^*$$

with

$$d^2 = \Delta^2 \left( \frac{\| \Sigma^{1/2} \mathbf{x}_i^* \|_*}{\| \Sigma^{1/2} \mathbf{x}_i^* \|_2} \right)^2.$$

Applying (22), Eq. (17) follows.

(b) If the norm used in the uncertainty set  $\mathcal{U}$  is the  $D$ -norm, then by applying Proposition 3, we obtain Eq. (18).

(c) If the norm used in the uncertainty set  $\mathcal{U}$  is the dual  $D$ -norm, then by applying Proposition 3, we obtain Eq. (19).

(d) If the norm used in the uncertainty set  $\mathcal{U}$  is the Euclidean norm, then Eq. (20) follows immediately from Eq. (17).  $\square$

Under the stronger assumption that the data in each constraint are independently distributed according to a symmetric distribution [7] proves stronger bounds. In contrast the bounds in Theorem 3 are weaker, but

have wider applicability as they include arbitrary dependencies.

## References

- [1] A. Ben-Tal, A. Nemirovski, Robust convex optimization, *Math. Oper. Res.* 23 (4) (1998) 769–805.
- [2] A. Ben-Tal, A. Nemirovski, Robust solutions of uncertain linear programs, *Oper. Res. Lett.* 25 (1) (1999) 1–13.
- [3] A. Ben-Tal, A. Nemirovski, Robust solutions of linear programming problems contaminated with uncertain data, *Math. Program. A* 88 (2000) 411–424.
- [4] A. Ben-Tal, A. Nemirovski, On polyhedral approximations of the second-order Cone, *Math. Oper. Res.* 26 (2) (2001) 193–228.
- [5] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*, MPR-SIAM Series on Optimization, SIAM, Philadelphia, 2001.
- [6] D. Bertsimas, I. Popescu, Optimal inequalities in probability theory: a convex optimization approach, in *SIAM J. Optim.* 2004, to appear.
- [7] D. Bertsimas, M. Sim, The price of robustness, *Oper. Res.* 52 (1) (2004) 35–53.
- [8] J.E. Falk, Exact solutions of inexact linear programs, *Oper. Res.* 24 (1976) 783–787.
- [9] L. El Ghaoui, H. Lebret, Robust solutions to least-squares problems with uncertain data, *SIAM J Matrix Anal. Appl.* 18 (4) (1997) 1035–1064.
- [10] L. El Ghaoui, F. Oustry, H. Lebret, Robust solutions to uncertain semidefinite programs, *SIAM J. Optim.* 9 (1) (1998) 33–52.
- [11] P.D. Lax, *Linear Algebra*, Wiley, New York, 1997.
- [12] A.L. Soyster, Convex programming with set-inclusive constraints and applications to inexact linear programming, *Oper. Res.* 21 (1973) 1154–1157.