Operations Research Letters 37 (2009) 1–6

Contents lists available at ScienceDirect

Operations Research Letters

journal homepage: www.elsevier.com/locate/orl

Duality in robust optimization: Primal worst equals dual best

Amir Beck *, Aharon Ben-Tal

Department of Industrial Engineering and Management, Technion—Israel Institute of Technology, Haifa 32000, Israel

ARTICLE INFO

Article history:
Received 1 April 2008
Accepted 26 September 2008
Available online 17 October 2008

Keywords:
Robust optimization
Optimistic counterpart
Convex programming duality

ABSTRACT

We study the dual problems associated with the robust counterparts of uncertain convex programs. We show that while the primal robust problem corresponds to a decision maker operating under the worst possible data, the dual problem corresponds to a decision maker operating under the *best* possible data.

1. Introduction

Consider a general uncertain optimization problem

$$(\mathrm{P}) \quad \begin{array}{ll} \min\limits_{\mathbf{x}} & g(\mathbf{x};\mathbf{u}) \\ \text{s.t.} & f_i(\mathbf{x};\mathbf{v}_i) \le \mathbf{0}, \quad i=1,\dots,m, \\ & \mathbf{x}\in\mathbb{R}^n, \end{array} \tag{1.1}$$

where g and f_i are convex functions with respect to the decision variable $\mathbf{x}$ and $\mathbf{u}\in\mathbb{R}^p$, $\mathbf{v}_i\in\mathbb{R}^{q_i}$ are the uncertain parameters of the problem.

Robust optimization (RO) [1,2] is one of the basic methodologies that deals with the case in which the parameters $\mathbf{u},\mathbf{v}_i$ are not exactly known. The setting in RO is that the information available on the uncertainties is crude: $\mathbf{u}$ and $\mathbf{v}_i$ are only known to reside in certain convex compact uncertainty sets:

$$\mathbf{u}\in\mathcal{U}, \mathbf{v}_i\in\mathcal{V}_i, \quad i=1,\dots,m.$$

A vector $\mathbf{x}$ is a *robust feasible* solution of (P) if it satisfies the constraints for every possible realization of the parameters. That is, it satisfies for every $i=1,\dots,m$:

$$f_i(\mathbf{x};\mathbf{v}_i)\le\mathbf{0} \text{ for every } \mathbf{v}_i\in\mathcal{V}_i.$$

We emphasize the fact that RO deals exclusively with problems modelled as (1.1), i.e., all constraints are *inequalities* and the uncertainty is *constraint-wise*. The constraints in problem (1.1) can be written as

$$F_i(\mathbf{x})\le\mathbf{0}, \quad i=1,\dots,m,$$

where

$$F_i(\mathbf{x})=\max_{\mathbf{v}_i\in\mathcal{V}_i} f_i(\mathbf{x};\mathbf{v}_i). \tag{1.2}$$

If we will also denote

$$G(\mathbf{x})=\max_{\mathbf{u}\in\mathcal{U}} g(\mathbf{x},\mathbf{u}), \tag{1.3}$$

then the *robust counterpart* (RC) of the original problem is given by

$$\begin{array}{ll} \min & G(\mathbf{x}) \\ \text{s.t.} & F_i(\mathbf{x})\le 0, \quad i=1,\dots,m, \\ & \mathbf{x}\in\mathbb{R}^n. \end{array} \tag{1.4}$$

The functions F and G_i are convex as pointwise maxima of convex functions [6]. Therefore, the robust counterpart (1.4) is *always* a convex optimization problem and thus it has a dual convex problem (call it (DR-P)), which under some regularity conditions has the same value as the primal problem. At the same time the primal uncertain problem (P) itself has an uncertain dual problem (call it (D)), with the same uncertain parameters. In this paper we study the relation between (DR-P) and the uncertain dual problem (D). The relation involves the notion of "optimistic counterpart" which is introduced in Section 2. We then show in Section 3 that for linear programming (LP) problems the dual of the robust counterpart is the optimistic counterpart of the uncertain dual problem. For LP problems the dual of the robust counterpart can be computed explicitly and so the above relation is explicitly revealed. In the last section we study the same relation for a general uncertain convex program. Although here the robust counterpart cannot be computed explicitly, we employ a minimax theorem to show that the relation "primal worst equals dual best" is valid; so, while in the primal robust problem we have a decision maker operating under the worst possible data, in the dual problem we have a decision maker operating under the best possible data.

* Corresponding author.
E-mail addresses: becka@ie.technion.ac.il (A. Beck), abental@ie.technion.ac.il (A. Ben-Tal).

0167-6377/$ – see front matter
doi:10.1016/j.orl.2008.09.010

2. The optimistic counterpart

A vector $\mathbf{x}$ is an *optimistic feasible solution* of (P) if it satisfies the constraints for *at least* one realization of the uncertainty set. That is, $\mathbf{x}$ is optimistic feasible solution if and only if for every $i = 1, \ldots, m$

$$f_i(\mathbf{x}; \mathbf{v}_i) \le 0 \quad \text{for some } \mathbf{v}_i \in \mathcal{V}.$$

The optimistic counterpart of problem (P) consists of minimizing the best possible objective function (i.e., minimal with respect to the parameters) over the set of optimistic feasible solutions:

$$\begin{array}{ll} \min\limits_{\mathbf{x}} & \left[\min\limits_{\mathbf{u}\in\mathcal{U}} g(\mathbf{x}; \mathbf{u}_i)\right] \\ \text{s.t.} & f_i(\mathbf{x}; \mathbf{v}_i) \le 0 \text{ for some } \mathbf{v}_i \in \mathcal{V}_i, \quad i = 1, \ldots, m, \\ & \mathbf{x} \in \mathbb{R}^n. \end{array}$$

Denoting $\hat{G}(\mathbf{x}) = \min_{\mathbf{u}\in\mathcal{U}} g(\mathbf{x}; \mathbf{u})$ and $\hat{F}_i(\mathbf{x}) = \min_{\mathbf{v}_i\in\mathcal{V}_i} f_i(\mathbf{x}; \mathbf{v}_i)$, the above problem can be equivalently written as

$$\text{(OC)} \quad \begin{array}{ll} \min\limits_{\mathbf{x}} & \hat{G}(\mathbf{x}) \\ \text{s.t.} & \hat{F}_i(\mathbf{x}) \le 0, \quad i = 1, \ldots, m, \\ & \mathbf{x} \in \mathbb{R}^n. \end{array}$$

As opposed to the robust counterpart, the above problem is in general not convex. The objective function $\hat{G}$ and constraints $\hat{F}_i$ are pointwise *minima* of convex functions and as such do not necessarily posses any convexity/concavity properties.

Example 2.1 (*Linear Programming*). Consider a linear programming (LP) problem

$$\begin{array}{ll} \min & \mathbf{c}^{\mathrm{T}}\mathbf{x} \\ \text{s.t.} & \mathbf{a}_i^{\mathrm{T}}\mathbf{x} \le b_i, \quad i = 1, \ldots, m, \\ & \mathbf{x} \in \mathbb{R}^n, \end{array} \tag{2.1}$$

where $\mathbf{c}, \mathbf{a}_i \in \mathbb{R}^n$, $i = 1, \ldots, m$ and $b_i \in \mathbb{R}$ for every $i = 1, \ldots, m$. Suppose now that the vectors $\mathbf{a}_i$ are not fixed but rather known to reside in an uncertainty set $\mathcal{V}_i$ which is an l_∞ balls of the form:

$$\mathcal{V}_i = \{\tilde{\mathbf{a}}_i + \mathbf{v}_i : \|\mathbf{v}_i\|_\infty \le \rho\},$$

where $\tilde{\mathbf{a}}_i$ is the nominal value of $\mathbf{a}_i$ and $\rho > 0$. This means that the coefficients a_{ij} (jth component of $\mathbf{a}_i$) have an interval uncertainty: $|a_{ij} - \tilde{a}_{ij}| \le \rho$. The ith constraint function is $F_i(\mathbf{x}) \equiv \max_{\mathbf{v}_i}(\tilde{\mathbf{a}}_i + \mathbf{v}_i)^{\mathrm{T}}\mathbf{x} = \tilde{\mathbf{a}}_i^{\mathrm{T}}\mathbf{x} + \rho\|\mathbf{x}\|_1$, and consequently the robust counterpart becomes

$$\begin{array}{ll} \min & \mathbf{c}^{\mathrm{T}}\mathbf{x} \\ \text{s.t.} & \tilde{\mathbf{a}}_i^{\mathrm{T}}\mathbf{x} + \rho\|\mathbf{x}\|_1 \le b_i, \quad i = 1, \ldots, m, \\ & \mathbf{x} \in \mathbb{R}^n. \end{array}$$

The above problem is of course a convex optimization problem and can be cast as an LP, thus rendering it tractable. On the other hand, the ith constraint in the optimistic counterpart is given by $\hat{F}_i(\mathbf{x}) = \min_{\mathbf{v}_i}(\tilde{\mathbf{a}}_i + \mathbf{v}_i)^{\mathrm{T}}\mathbf{x} = \tilde{\mathbf{a}}_i^{\mathrm{T}}\mathbf{x} - \rho\|\mathbf{x}\|_1$. Consequently, the optimistic counterpart of (2.1) is

$$\begin{array}{ll} \min & \mathbf{c}^{\mathrm{T}}\mathbf{x} \\ \text{s.t.} & \tilde{\mathbf{a}}_i^{\mathrm{T}}\mathbf{x} - \rho\|\mathbf{x}\|_1 \le b_i, \quad i = 1, \ldots, m, \\ & \mathbf{x} \in \mathbb{R}^n, \end{array}$$

which is clearly not a convex problem.

Note however that if instead of (2.1) the nonlinear LP is

$$\begin{array}{ll} \min & \mathbf{c}^{\mathrm{T}}\mathbf{x} \\ \text{s.t.} & \mathbf{a}_i^{\mathrm{T}}\mathbf{x} \le b_i, \quad i = 1, \ldots, m, \\ & \mathbf{x} \ge 0, \end{array} \tag{2.2}$$

then the corresponding optimistic counterpart associated with the same uncertainty sets $\mathcal{V}_i$ is

$$\begin{array}{ll} \min & \mathbf{c}^{\mathrm{T}}\mathbf{x} \\ \text{s.t.} & \tilde{\mathbf{a}}_i^{\mathrm{T}}\mathbf{x} - \rho\sum\limits_{i=1}^{n} x_i \le b_i, \quad i = 1, \ldots, m, \\ & \mathbf{x} \ge 0, \end{array}$$

which is a convex (in fact linear) program. Thus we observe that the convexity status of the optimistic counterpart depends among other things on the representation of the nominal problem. It also depends on the choice of the uncertainty set; for example, for the same nominal problem (2.2) if the uncertainty sets $\mathcal{U}_i$ are

$$\mathcal{U}_i = \{\tilde{\mathbf{a}}_i + \mathbf{v}_i : \|\mathbf{v}_i\|_2 \le \rho\}, \quad i = 1, \ldots, m,$$

then the optimistic counterpart is again a *nonconvex problem*:

$$\begin{array}{ll} \min & \mathbf{c}^{\mathrm{T}}\mathbf{x} \\ \text{s.t.} & \tilde{\mathbf{a}}_i^{\mathrm{T}}\mathbf{x} - \rho\|\mathbf{x}\|_2 \le b_i, \quad i = 1, \ldots, m, \\ & \mathbf{x} \ge 0. \quad \square \end{array}$$

Example 2.2 (*Robust and Optimistic Least Squares*). Given a matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ and a vector $\mathbf{b} \in \mathbb{R}^m$, the celebrated least squares problem [3] consists of minimizing the data error over the entire space:

$$\text{(LS)} \quad \min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2.$$

Now, assume that the matrix $\mathbf{A}$ is not fixed but is rather known to reside in the uncertainty set:

$$\mathcal{U} = \{\tilde{\boldsymbol{\Delta}} + \mathbf{U} : \|\boldsymbol{\Delta}\|_F \le \rho\},$$

where $\|\cdot\|_F$ stands for the Frobenius norm and $\tilde{\mathbf{A}}$ is the fixed nominal matrix. The robust counterpart of (LS) is given by

$$\text{(R-LS)} \quad \min_{\mathbf{x}} \max_{\mathbf{A}\in\mathcal{U}} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2.$$

This problem was introduced and studied in [5] where it was called *robust least squares*. By explicitly solving the inner maximization problem (see [5]), (R-LS) becomes

$$\min_{\mathbf{x}} \|\tilde{\mathbf{A}}\mathbf{x} - \mathbf{b}\|_2 + \rho\|\mathbf{x}\|_2.$$

The above is of course a convex problem, and more specifically a conic quadratic problem, that can be solved efficiently.

Here the optimistic counterpart of (LS) is given by

$$\text{(O-LS)} \quad \min_{\mathbf{x}} \min_{\mathbf{A}\in\mathcal{U}} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2.$$

Using the variational form of the Euclidean norm, the inner minimization problem becomes

$$\begin{aligned} \min_{\|\Delta\|_F\le\rho} \|\tilde{\mathbf{A}}\mathbf{x} - \mathbf{b} + \boldsymbol{\Delta}\mathbf{x}\|_2 &= \min_{\|\Delta\|_F\le\rho} \max_{\mathbf{y}:\|\mathbf{y}\|_2\le 1} \mathbf{y}^{\mathrm{T}}(\tilde{\mathbf{A}}\mathbf{x} - \mathbf{b} + \boldsymbol{\Delta}\mathbf{x}) \\ &= \max_{\mathbf{y}:\|\mathbf{y}\|_2\le 1} \min_{\|\Delta\|_F\le\rho} \mathbf{y}^{\mathrm{T}}(\tilde{\mathbf{A}}\mathbf{x} - \mathbf{b} + \boldsymbol{\Delta}\mathbf{x}) \\ &= \max_{\mathbf{y}:\|\mathbf{y}\|_2\le 1} \mathbf{y}^{\mathrm{T}}(\tilde{\mathbf{A}}\mathbf{x} - \mathbf{b}) - \rho\|\mathbf{x}\|_2\|\mathbf{y}\|_2, \end{aligned} \tag{2.3}$$

where the second equality follows from the equality between min–max and max–min for convex–concave functions [6, Corollary 37.3.2]. An explicit expression for (2.3) can be found as follows:

$$\begin{aligned} \max_{\mathbf{y}:\|\mathbf{y}\|\le 1} \mathbf{y}^{\mathrm{T}}(\tilde{\mathbf{A}}\mathbf{x} - \mathbf{b}) - \rho\|\mathbf{x}\|\|\mathbf{y}\| &= \max_{0\le\alpha\le 1} \max_{\|\mathbf{y}\|=\alpha} \mathbf{y}^{\mathrm{T}}(\tilde{\mathbf{A}}\mathbf{x} - \mathbf{b}) - \rho\|\mathbf{x}\|\alpha, \\ &= \max_{0\le\alpha\le 1} \alpha(\|\tilde{\mathbf{A}}\mathbf{x} - \mathbf{b}\| - \rho\|\mathbf{x}\|) \\ &= \left[\|\tilde{\mathbf{A}}\mathbf{x} - \mathbf{b}\| - \rho\|\mathbf{x}\|\right]_+ \end{aligned}$$

where for a scalar x, $[x]_+$ denotes the positive part of x:

$$[x]_+ = \begin{cases} x & x > 0, \\ 0 & x \leq 0. \end{cases}$$

Therefore, (O-LS) simplifies to

$$\min_{\mathbf{x}} [\|\tilde{\mathbf{A}}\mathbf{x} - \mathbf{b}\|_2 - \rho\|\mathbf{x}\|_2]_+,$$

which is not a convex optimization problem. It is interesting to note that as was shown in [5], the solution of (R-LS) is of the form $\hat{\mathbf{x}} = (\tilde{\mathbf{A}}^{\mathrm{T}}\tilde{\mathbf{A}} + \alpha\mathbf{I})^{-1}\tilde{\mathbf{A}}^{\mathrm{T}}\mathbf{b}$ where $\alpha \geq 0$ and in [4] it was shown, under the assumption that ρ is "small enough", that the solution of (O-LS) is of the form $\hat{\mathbf{x}} = (\tilde{\mathbf{A}}^{\mathrm{T}}\tilde{\mathbf{A}} - \beta\mathbf{I})^{-1}\tilde{\mathbf{A}}^{\mathrm{T}}\mathbf{b}$ where $\beta \geq 0$. Both α and β are solutions of certain one-dimensional secular equations. Therefore, the optimal solution of (R-LS) is a *regularization* of the least squares solution while the optimal solution of (O-LS) has an opposite *de-regularization* effect. □

We end this section by defining the optimistic counterpart for a more general optimization model in which equality constraints are also present and with uncertainty parameters which are not necessarily constraint-wise. Specifically, consider the general model

$$(\mathrm{G}) \quad \begin{array}{ll} \min\limits_{\mathbf{x}} & g(\mathbf{x}; \mathbf{u}) \\ \text{s.t.} & \mathbf{H}(\mathbf{x}; \mathbf{v}_1) \leq \mathbf{0}, \\ & \mathbf{K}(\mathbf{x}; \mathbf{v}_2) = \mathbf{0}, \\ & \mathbf{x} \in \mathbb{R}^n \end{array}$$

where $\mathbf{H} : \mathbb{R}^n \to \mathbb{R}^{m_1}$ and $\mathbf{K} : \mathbb{R}^n \to \mathbb{R}^{m_2}$ are vector functions and $g : \mathbb{R}^n \to \mathbb{R}$ is a scalar function of n variables. The parameters $\mathbf{u} \in \mathbb{R}^p$, $\mathbf{v}_1 \in \mathbb{R}^{q_1}$, $\mathbf{v}_2 \in \mathbb{R}^{q_2}$ are uncertain and only known to reside in compact sets $\mathcal{U}_1$, $\mathcal{V}_1$ and $\mathcal{V}_2$ respectively. The optimistic counter part of problem (G) is defined as

$$(\text{O-G}) \quad \begin{array}{ll} \min\limits_{\mathbf{x}} & \min\limits_{\mathbf{u} \in \mathcal{U}} g(\mathbf{x}; \mathbf{u}) \\ \text{s.t.} & \mathbf{H}(\mathbf{x}; \mathbf{v}_1) \leq \mathbf{0} \text{ for some } \mathbf{v}_1 \in \mathcal{V}_1, \\ & \mathbf{K}(\mathbf{x}; \mathbf{v}_2) = \mathbf{0} \text{ for some } \mathbf{v}_2 \in \mathcal{V}_2, \\ & \mathbf{x} \in \mathbb{R}^n. \end{array}$$

Note that as opposed to the robust counterpart, in the context of the optimistic counterpart, it *does* make sense to deal with equality constraints.

3. The dual of robust linear programming problems

In this section we consider an LP of the following general form:

$$(\mathrm{LP}) \quad \begin{array}{lll} \max & \mathbf{c}^{\mathrm{T}}\mathbf{x} & \\ \text{s.t.} & \mathbf{a}_i^{\mathrm{T}}\mathbf{x} \leq b_i, & i = 1, \ldots, m, \\ & \mathbf{x} \geq \mathbf{0}, & \end{array} \qquad (3.1)$$

where $\mathbf{c}, \mathbf{a}_i \in \mathbb{R}^n$. Assume now that for every $i = 1, \ldots, m$, the vector $\mathbf{a}_i$ is not fixed but rather known to reside in some uncertainty set $\mathcal{U}_i$:

$$\mathbf{a}_i \in \mathcal{U}_i,$$

where $\mathcal{U}_i$ is a nonempty compact set. We assume that $\mathbf{b}$ and $\mathbf{c}$ are certain data vectors for the sake of exposition. Indeed a problem with uncertainty in the objective function and righthand side of the constraints can be reduced to an equivalent one with uncertainty only in the lefthand side of the constraints. Throughout, the letters "D", "R", "O" stand for "dual", "robust" and "optimistic" respectively. In particular, (R-LP) is the robust counterpart of (LP), (DR-LP) is the dual of (R-LP), (D-LP) is the dual of (LP) and (OD-LP) is the optimistic counterpart of (D-LP).

Now, consider (DR-LP) – the dual of the robust counterpart of (LP). A very natural question is whether or not the dual of the robust counterpart is the same as the robust counterpart of the dual problem. The answer is *certainly not*. On the contrary, we will show in this section (Theorem 3.1) that regardless of the choice of the uncertainty set, problem (DR-LP) – the dual of the robust counterpart of (LP) – is the same as the *optimistic* counterpart of the dual problem of (LP), that is, problem (OD-LP). The following sketch illustrates this relation.

$$\begin{array}{ccccc} & & (\mathrm{LP}) & & \\ \text{robust} & \swarrow & & \searrow & \text{dual} \\ & (\text{R-LP}) & & (\text{D-LP}) & \\ \text{dual} & \downarrow & & \downarrow & \text{optimistic} \\ & (\text{DR-LP}) & = & (\text{OD-LP}) & \end{array}$$

We use the following notation. For a compact nonempty set S, the support function of S at a point $\mathbf{y}$ is denoted by

$$\sigma_S(\mathbf{y}) \equiv \max\{\mathbf{y}^{\mathrm{T}}\mathbf{z} : \mathbf{z} \in S\}.$$

Theorem 3.1. *The dual of the robust counterpart of (LP) and the optimistic counterpart of the dual problem of (LP) are both given by the convex program*

$$(\textit{DR-LP}),(\textit{OD-LP}) \quad \begin{array}{ll} \min & \mathbf{b}^{\mathrm{T}}\mathbf{y} \\ \textit{s.t.} & g(\mathbf{y}) \geq 0, \\ & \mathbf{y} \geq \mathbf{0}, \end{array} \qquad (3.2)$$

where

$$g(\mathbf{y}) = \min_{\mathbf{x} \geq \mathbf{0}} \left\{ \sum_{i=1}^m y_i \sigma_{\mathcal{U}_i}(\mathbf{x}) - \mathbf{c}^{\mathrm{T}}\mathbf{x} \right\}. \qquad (3.3)$$

Proof. The ith constraint of the robust counterpart of (3.1) is

$$\max\{\mathbf{a}_i^{\mathrm{T}}\mathbf{x} : \mathbf{a}_i \in \mathcal{U}_i\} \leq b_i,$$

which can also be written as

$$(\text{R-LP}) \quad \begin{array}{lll} \max\limits_{\mathbf{x}} & \mathbf{c}^{\mathrm{T}}\mathbf{x} & \\ \text{s.t.} & \sigma_{\mathcal{U}_i}(\mathbf{x}) \leq b_i, & i = 1, \ldots, m, \\ & \mathbf{x} \geq \mathbf{0}. & \end{array}$$

To construct the dual, let us write the Lagrangian:

$$\mathcal{L}(\mathbf{x}, \mathbf{y}) = \mathbf{c}^{\mathrm{T}}\mathbf{x} - \sum_{i=1}^m y_i(\sigma_{\mathcal{U}_i}(\mathbf{x}) - b_i).$$

By the homogenicity of the support function we have

$$\max_{\mathbf{x} \geq 0} \mathcal{L}(\mathbf{x}, \mathbf{y}) = \begin{cases} \mathbf{b}^{\mathrm{T}}\mathbf{y} & g(\mathbf{y}) \geq 0, \\ \infty & \text{else} \end{cases}$$

where g is given in (3.3). Therefore, the dual problem of (R-LP) is given by

$$(\text{DR-LP}) \quad \begin{array}{ll} \min & \mathbf{b}^{\mathrm{T}}\mathbf{y} \\ \text{s.t.} & g(\mathbf{y}) \geq 0, \\ & \mathbf{y} \geq \mathbf{0}, \end{array}$$

where g is given in (3.3). Let us now consider the uncertain dual problem of (LP):

$$(\text{D-LP}) \quad \begin{array}{ll} \min\limits_{\mathbf{y}} & \mathbf{b}^{\mathrm{T}}\mathbf{y} \\ \text{s.t.} & \sum\limits_{i=1}^m y_i \mathbf{a}_i \geq \mathbf{c}, \\ & \mathbf{y} \geq \mathbf{0}. \end{array} \qquad (3.4)$$

A vector $\mathbf{y} \geq \mathbf{0}$ is optimistic-feasible solution of (D) if and only if

$$\exists \mathbf{a}_i \in \mathcal{U}_i : \sum_{i=1}^m y_i \mathbf{a}_i \geq \mathbf{c}.$$

The latter is satisfied if and only if

$$\max_{\mathbf{a}_i \in \mathcal{U}_i} \min_{\mathbf{z} \geq \mathbf{0}} \left\{ \mathbf{z}^T \left(\sum_{i=1}^m y_i \mathbf{a}_i - \mathbf{c} \right) \right\} \geq 0.$$

Using the equality of min–max and max–min for convex–concave functions [6, Corollary 37.3.2], we can replace the order of the min and the max, thus resulting with

$$\min_{\mathbf{z} \geq \mathbf{0}} \max_{\mathbf{a}_i \in \mathcal{U}_i} \left\{ \sum_{i=1}^m y_i \mathbf{a}_i^T \mathbf{z} - \mathbf{c}^T \mathbf{z} \right\} \geq 0,$$

which is the same as

$$g(\mathbf{y}) = \min_{\mathbf{z} \geq \mathbf{0}} \left\{ \sum_{i=1}^m y_i \sigma_{\mathcal{U}_i}(\mathbf{z}) - \mathbf{c}^T \mathbf{z} \right\} \geq 0.$$

Therefore, the optimistic counterpart of the dual problem (D-LP) coincides with (DR-LP) — the dual of the robust counterpart of the primal problem. □

Example 3.1. Suppose that the uncertainty sets $\mathcal{U}_i$ are ellipsoidal sets given by

$$\mathcal{U}_i = \left\{ \mathbf{a}_i = \tilde{\mathbf{a}}_i + \mathbf{v}_i : \|\mathbf{v}_i\|_2 \leq \rho \right\}, \quad i = 1, \ldots, m.$$

In this case

$$\sigma_{U_i}(\mathbf{x}) = \max_{\mathbf{z} \in \mathcal{U}_i} \mathbf{x}^T \mathbf{z} = \max_{\|\mathbf{v}_i\|_2 \leq \rho} (\mathbf{x}^T \tilde{\mathbf{a}}_i + \mathbf{x}^T \mathbf{v}_i) = \tilde{\mathbf{a}}_i^T \mathbf{x} + \rho \|\mathbf{x}\|_2.$$

Therefore,

$$\begin{aligned} g(\mathbf{y}) &= \min_{\mathbf{x}} \left\{ \sum_{i=1}^m y_i \sigma_{\mathcal{U}_i}(\mathbf{x}) - \mathbf{c}^T \mathbf{x} \right\} \\ &= \min_{\mathbf{x}} \left(\sum_{i=1}^m y_i \tilde{\mathbf{a}}_i - \mathbf{c} \right)^T \mathbf{x} + \rho \left(\sum_{i=1}^m y_i \right) \|\mathbf{x}\|_2 \\ &= \begin{cases} 0 & \left\| \sum_{i=1}^m y_i \tilde{\mathbf{a}}_i - \mathbf{c} \right\|_2 \leq \rho \left(\sum_{i=1}^m y_i \right) \\ -\infty & \text{else.} \end{cases} \end{aligned}$$

Plugging this in (3.2) we conclude that the dual of the robust counterpart of LP, which is the same as the optimistic counterpart of the dual of LP is given by the conic quadratic problem

$$\begin{aligned} \max \quad & \mathbf{b}^T \mathbf{y} \\ \text{s.t.} \quad & \left\| \sum_{i=1}^m y_i \tilde{\mathbf{a}}_i - \mathbf{c} \right\|_2 \leq \rho \left(\sum_{i=1}^m y_i \right), \\ & \mathbf{y} \geq \mathbf{0}. \quad \square \end{aligned}$$

Theorem 3.1 shows that val(DR-LP) = val(OD-LP). In addition, if (R-LP), the robust counterpart of (LP), is bounded below and satisfies a regularity condition such as Slater condition, then strong duality holds [6] and val(R-LP) = val(DR-LP). From this we conclude that

$$\text{val(R-LP)} = \text{val(OD-LP)}. \tag{3.5}$$

Namely, the value of the robust counterpart of (LP) ("primal worst") is equal to the value of the optimistic counterpart of the dual problem ("dual best"). To interpret this result, assume that problem (LP) models a standard production problem. In this context x_j, the decision variable, stands for the amount to be produced of item j; c_j is the profit from the sale of one unit of item j; a_{ij} (the jth component of the vector $\mathbf{a}_i$) is the quantity of resource i needed to produce one unit of item j, and b_i is the the available supply of resource i. The dual problem of (3.1) is

$$\min \left\{ \mathbf{b}^T \mathbf{y} : \sum_{i=1}^m y_i \mathbf{a}_i \geq \mathbf{c}, \mathbf{y} \geq \mathbf{0} \right\}. \tag{3.6}$$

A common interpretation of the dual is as follows: suppose that a merchant wants to buy the resources of the manufacturer. For every i, his decision variable y_i stands for the price he offers to pay for purchasing one unit of resource i. Of course he wants to pay as little as possible and thus his objective is to minimize $\mathbf{b}^T\mathbf{y}$. The set of constraints $\sum_{i=1}^m a_{ij} y_i \geq c_j$ indicates that the merchant has to suggest competitive prices (so that the profit c_j the manufacturer can get from selling a unit of item j is no more than the value he can get from what the merchant is paying for the resources needed to produce a unit of item j).

Now suppose that the vectors are uncertain. The robust counterpart of the production problem (3.1) reflects in a sense a pessimistic manufacturer that wishes to maximize his profit under a worst case scenario. The equality (3.5) states that what is worst for the factory is best for the merchant. Specifically, since the manufacturer underestimates the value of the resources, the merchant needs to be less competitive and can offer lower prices for the resources, thus reducing his total payment.

Remark 3.1. In Example 2.1 we showed that the optimistic counterpart of the LP (2.1) is *nonconvex*, and thus it cannot be equivalent to the dual of a robust LP (which is always a convex problem). How does this then agree with the duality result in Theorem 3.1? the answer is that the LP in (2.1) is the dual of the following LP:

$$\begin{aligned} \max \quad & \mathbf{b}^T \mathbf{y} \\ \text{s.t.} \quad & \sum_{i=1}^m y_i \mathbf{a}_i = \mathbf{c}, \\ & \mathbf{y} \leq \mathbf{0}, \end{aligned}$$

which is *not* of the general form (3.1), where the constraints are assumed to be inequalities and the uncertainties must be constraint-wise.

4. Primal worst equals dual best

Consider now the general model (P) given in (1.1). The robust counterpart of (P) is the problem (R-P) given in (1.1). The dual of (R-P) is given by

$$\max_{\lambda \geq 0} \min_{\mathbf{x}} \left\{ G(\mathbf{x}) + \sum_{i=1}^m \lambda_i F_i(\mathbf{x}) \right\}.$$

Recalling the definition of G and F_i (see (1.2) and (1.3)), the problem becomes

$$\text{(DR-P)} \quad \max_{\lambda \geq 0} \min_{\mathbf{x}} \max_{\mathbf{u} \in \mathcal{U}, \mathbf{v}_i \in \mathcal{V}_i} \left\{ g(\mathbf{x}; \mathbf{u}) + \sum_{i=1}^m \lambda_i f_i(\mathbf{x}; \mathbf{v}_i) \right\}.$$

We will assume that Slater constraint qualification is satisfied for problem (R-P) and that (R-P) is bounded below. Under these conditions it is well known that val(R-P) = val(DR-P) [6].

On the other hand, the dual of (P) is given by

$$\text{(D-P)} \quad \max_{\lambda \geq 0} q(\boldsymbol{\lambda}; \mathbf{u}, \mathbf{v}_i),$$

where

$$q(\boldsymbol{\lambda}; \mathbf{u}, \mathbf{v}_i) \equiv \min_{\mathbf{x}} \left\{ g_i(\mathbf{x}; \mathbf{u}) + \sum_{i=1}^m \lambda_i f_i(\mathbf{x}; \mathbf{v}_i) \right\}. \tag{4.1}$$

The optimistic counterpart of (D-P) is

$$\max_{\lambda\geq 0}\max_{\mathbf{u}\in\mathcal{U},\mathbf{v}_i\in\mathcal{V}_i} q(\boldsymbol{\lambda};\mathbf{u},\mathbf{v}_i).$$

Plugging the expression (4.1) for q in the above problem, we arrive at the following formulation of the optimistic counterpart of the dual:

$$\text{(OD-P)}\quad \max_{\lambda\geq 0}\max_{\mathbf{u}\in\mathcal{U},\mathbf{v}_i\in\mathcal{V}_i}\min_{\mathbf{x}}\left\{g_i(\mathbf{x};\mathbf{u})+\sum_{i=1}^{m}\lambda_i f_i(\mathbf{x};\mathbf{v}_i)\right\}.$$

The LP case considered in Section 3 suggests that under some conditions the optimal values of (OD-P) and (DR-P) are equal. The following result shows that val(OD-P) is always greater than or equal to val(DR-P) and that under suitable convexity assumptions, equality holds.

Theorem 4.1. *Consider the general convex problem (P) (problem (1.1)). Then*

$$\mathrm{val(OD-P)}\leq \mathrm{val(DR-P)}. \tag{4.2}$$

If in addition the functions g, f_i are concave with respect to the unknown parameters $\mathbf{u}, \mathbf{v}_i$, then the following equality holds:

$$\mathrm{val(OD-P)}= \mathrm{val(DR-P)}. \tag{4.3}$$

Proof. Since min–max is always greater than or equal to max–min we have:

$$\begin{aligned}\mathrm{val(OD\text{-}P)} &= \max_{\lambda\geq 0}\max_{\mathbf{u}\in\mathcal{U},\mathbf{v}_i\in\mathcal{V}_i}\min_{\mathbf{x}}\left\{g_i(\mathbf{x};\mathbf{u})+\sum_{i=1}^{m}\lambda_i f_i(\mathbf{x};\mathbf{v}_i)\right\}\\ &\leq \max_{\lambda\geq 0}\min_{\mathbf{x}}\max_{\mathbf{u}\in\mathcal{U},\mathbf{v}_i\in\mathcal{V}_i}\left\{g_i(\mathbf{x};\mathbf{u})+\sum_{i=1}^{m}\lambda_i f_i(\mathbf{x};\mathbf{v}_i)\right\}\\ &= \mathrm{val(DR\text{-}P)}.\end{aligned}$$

If g, f_i are concave with respect to $\mathbf{u}, \mathbf{v}_i$ (in addition to the convexity with respect to $\mathbf{x}$), and since $U, \mathcal{V}_i$ are convex compact sets, then by [6, Corollary 37.3.2] equality (4.3) holds. □

Clearly, Theorem 4.1 generalizes the result obtained in Section 3.

As was noted in the LP case, by the strong duality result for convex programming we know that val(R-P) = val(DR-P). We therefore conclude that if the functions are all convex with respect to $\mathbf{x}$ and concave with respect to $\mathbf{u}, \mathbf{v}_i$ — the unknown parameters, then

$$\mathrm{val(R\text{-}P)}= \mathrm{val(OD\text{-}P)}. \tag{4.4}$$

Loosely speaking, relation (4.4) states that optimizing under the worst case scenario in the primal is the same as optimizing under the best case scenario in the dual ("primal worst equals dual best").

Theorem 4.1 is not restricted only to LP problems. Another interesting example is that of a convex quadratically constrained quadratic programming (QCQP).

Example 4.1. Consider the QCQP problem

$$\text{(QCQP)}\quad \begin{array}{ll}\min & \mathbf{x}^T\mathbf{A}_0\mathbf{x}+2\mathbf{b}_0^T\mathbf{x}+c_0\\ \text{s.t.} & \mathbf{x}^T\mathbf{A}_i\mathbf{x}+2\mathbf{b}_i^T\mathbf{x}+c_i\leq 0,\\ & \mathbf{x}\in\mathbb{R}^n.\end{array}$$

Then for each $i = 0, 1, \ldots, m$ the matrix $\mathbf{A}_i$ is uncertain and resides in the uncertainty set

$$\mathcal{U}_i=\left\{\tilde{\mathbf{A}}_i+\boldsymbol{\Delta}_i:\boldsymbol{\Delta}_i=\boldsymbol{\Delta}_i^T,\ \|\boldsymbol{\Delta}_i\|_2\leq\rho_i\right\}$$

here $\|\mathbf{S}\|_2$ is the spectral norm of a matrix $\mathbf{S}$, i.e., $\|\mathbf{S}\|_2=\sqrt{\lambda_{\max}(\mathbf{S}^T\mathbf{S})}$. We assume that the nominal matrices $\tilde{\mathbf{A}}_i$ are positive definite and that $\rho_i<\lambda_{\min}(\tilde{\mathbf{A}}_i)$ for every $i = 0, 1, \ldots, m$. Under this condition the objective functions and constraints are strictly convex for every possible realization of the uncertain parameters. Therefore, the objective function and all the constraint functions are convex with respect to $\mathbf{x}$ and concave (in fact linear) with respect to $\boldsymbol{\Delta}_i$. By Theorem 4.1 this implies that val(DR-P) = val(OD-P). Let us find explicit expressions for these problems and verify the result. Since for every $i = 0, 1, \ldots, m$

$$\max_{\mathbf{A}_i\in\mathcal{U}_i}\mathbf{x}^T\mathbf{A}_i\mathbf{x}+2\mathbf{b}_i^T\mathbf{x}+c_i=\mathbf{x}^T(\tilde{\mathbf{A}}_i+\rho_i\mathbf{I})\mathbf{x}+2\mathbf{b}_i^T\mathbf{x}+c_i,$$

the robust counterpart of (QCQP) is given by

$$\text{(R-QCQP)}\quad \begin{array}{ll}\min & \mathbf{x}^T\left(\tilde{\mathbf{A}}_0+\rho_0\mathbf{I}\right)\mathbf{x}+2\mathbf{b}_0^T\mathbf{x}+c_0\\ \text{s.t.} & \mathbf{x}^T\left(\tilde{\mathbf{A}}_i+\rho_i\mathbf{I}\right)\mathbf{x}+2\mathbf{b}_i^T\mathbf{x}+c_i\leq 0,\\ & \mathbf{x}\in\mathbb{R}^n.\end{array}$$

The dual problem of (R-QCQP) is

$$\begin{aligned}\text{(DR-QCQP)}\quad &\max_{\lambda\geq 0}-\left(\mathbf{b}_0+\sum\lambda_i\mathbf{b}_i\right)^T\\ &\times\left(\tilde{\mathbf{A}}_0+\rho_0\mathbf{I}+\sum\lambda_i(\tilde{\mathbf{A}}_i+\rho_i\mathbf{I})\right)^{-1}\left(\mathbf{b}_0+\sum\lambda_i\mathbf{b}_i\right)\\ &+c_0+\sum c_i\lambda_i,\end{aligned}$$

where the summation is over $i = 1, \ldots, m$. Now consider the dual problem of (QCQP):

$$\begin{aligned}\text{(D-QCQP)}\quad &\max_{\lambda\geq 0}-\left(\mathbf{b}_0+\sum\lambda_i\mathbf{b}_i\right)^T\\ &\times\left(\mathbf{A}_0+\sum\lambda_i\mathbf{A}_i\right)^{-1}\left(\mathbf{b}_0+\sum\lambda_i\mathbf{b}_i\right)+c_0+\sum c_i\lambda_i.\end{aligned}$$

The optimistic counterpart of (D-QCQP) is given by

$$\begin{aligned}&\max_{\lambda\geq 0}\max_{\|\Delta_i\|_2\leq\rho_i}-\left(\mathbf{b}_0+\sum\lambda_i\mathbf{b}_i\right)^T\\ &\times\left(\tilde{\mathbf{A}}_0+\boldsymbol{\Delta}_0+\sum\lambda_i\left(\tilde{\mathbf{A}}_i+\boldsymbol{\Delta}_i\right)\right)^{-1}\left(\mathbf{b}_0+\sum\lambda_i\mathbf{b}_i\right)\\ &+c_0+\sum c_i\lambda_i.\end{aligned}\tag{4.5}$$

To solve the inner maximization, we will use the following property: if two positive definite matrices satisfy $\mathbf{A}\succeq\mathbf{B}$, then $\mathbf{A}^{-1}\preceq\mathbf{B}^{-1}$. Therefore, since $\boldsymbol{\Delta}_i\preceq\rho_i\mathbf{I}$ for every $i = 0, 1, \ldots, m$, we conclude that

$$\begin{aligned}&-\left(\mathbf{b}_0+\sum\lambda_i\mathbf{b}_i\right)^T\left(\tilde{\mathbf{A}}_0+\boldsymbol{\Delta}_0+\sum\lambda_i(\tilde{\mathbf{A}}_i+\boldsymbol{\Delta}_i)\right)^{-1}\\ &\quad\times\left(\mathbf{b}_0+\sum\lambda_i\mathbf{b}_i\right)\\ &\leq-\left(\mathbf{b}_0+\sum\lambda_i\mathbf{b}_i\right)^T\left(\tilde{\mathbf{A}}_0+\rho_0\mathbf{I}+\sum\lambda_i(\tilde{\mathbf{A}}_i+\rho_i\mathbf{I})\right)^{-1}\\ &\quad\times\left(\mathbf{b}_0+\sum\lambda_i\mathbf{b}_i\right).\end{aligned}$$

Since the upper bound is attained at $\boldsymbol{\Delta}_i=\rho_i\mathbf{I}$, we conclude that the solution of the inner maximization is $\boldsymbol{\Delta}_i=\rho_i\mathbf{I}$, thus simplifying (4.5) to

$$\begin{aligned}\text{(OD-QCQP)}\quad &\max_{\lambda\geq 0}-\left(\mathbf{b}_0+\sum\lambda_i\mathbf{b}_i\right)^T\left(\tilde{\mathbf{A}}_0+\rho_0\mathbf{I}\right.\\ &\left.+\sum\lambda_i\left(\tilde{\mathbf{A}}_i+\rho_i\mathbf{I}\right)\right)^{-1}\left(\mathbf{b}_0+\sum\lambda_i\mathbf{b}_i\right)+c_0+\sum c_i\lambda_i,\end{aligned}$$

which is the same as (DR-QCQP). Also note that by using Schur complement, problem (DR-QCQP) can be cast as the linear semidefinite program

$$\begin{aligned} \max_{\lambda\geq 0,t} \quad & -t + c_0 + \sum c_i\lambda_i \\ \text{s.t.} \quad & \begin{pmatrix} \tilde{\mathbf{A}}_0 + \rho_0\mathbf{I} + \sum\lambda_i(\tilde{\mathbf{A}}_i + \rho_i\mathbf{I}) & \mathbf{b}_0 + \sum\lambda_i\mathbf{b}_i \\ \left(\mathbf{b}_0 + \sum\lambda_i\mathbf{b}_i\right)^{\mathrm{T}} & t \end{pmatrix} \succeq \mathbf{0}. \quad \square \end{aligned}$$

The next and last example illustrates that when the concavity assumption of the functions with respect to the unknown parameters fails, strict inequality may occur in (4.2).

Example 4.2. Consider the least squares problem

$$\text{(LS):} \quad \min \|\mathbf{Ax} - \mathbf{b}\|_2.$$

Let us assume that **A** is uncertain and is known to reside in a ball:

$$\mathbf{A} \in \mathcal{U} = \{\mathbf{A}_0 + \boldsymbol{\Delta} : \|\boldsymbol{\Delta}\|_F \leq \rho\}.$$

The robust counterpart of (LS) is:

$$\min_{\mathbf{x}} \max_{\|\Delta\|_F\leq\rho} \|(\mathbf{A}_0 + \boldsymbol{\Delta})\mathbf{x} - \mathbf{b}\|_2.$$

Solving the inner maximization problem, the above problem reduces to

$$\text{(R-LS)} \quad \min_{\mathbf{x}} \|\mathbf{A}_0\mathbf{x} - \mathbf{b}\|_2 + \rho\|\mathbf{x}\|_2.$$

The dual problem to (R-LS) is

$$\text{(DR-LS)} \quad \begin{aligned} \max \quad & \mathbf{b}^{\mathrm{T}}\boldsymbol{\lambda} \\ \text{s.t.} \quad & \|\boldsymbol{\lambda}\|_2 \leq 1, \\ & \|\mathbf{A}_0^{\mathrm{T}}\boldsymbol{\lambda}\|_2 \leq \rho. \end{aligned}$$

In order to write a dual problem of (LS), let us rewrite it first as

$$\text{(LS}'\text{)} \quad \min_{\mathbf{x},\mathbf{y}} \{\|\mathbf{y}\|_2 : \mathbf{y} = \mathbf{Ax} - \mathbf{b}\}.$$

The dual problem to (LS′) is

$$\text{(D-LS)} \quad \begin{aligned} \max \quad & \mathbf{b}^{\mathrm{T}}\boldsymbol{\lambda} \\ \text{s.t.} \quad & \|\boldsymbol{\lambda}\|_2 \leq 1 \\ & \mathbf{A}^{\mathrm{T}}\boldsymbol{\lambda} = \mathbf{0}. \end{aligned}$$

To find the optimistic counterpart of the dual problem we need to write explicitly (i.e., in terms of $\boldsymbol{\lambda}$) the constraint:

$$\mathbf{A}^{\mathrm{T}}\boldsymbol{\lambda} = \mathbf{0} \text{ for some } \mathbf{A} \in \mathcal{U}.$$

The above constraint is the same as

$$\|\mathbf{A}_0^{\mathrm{T}}\boldsymbol{\lambda}\|_2 \leq \rho\|\boldsymbol{\lambda}\|_2$$

and thus the optimistic counterpart of (D-LS) is

$$\text{(OD-LS)} \quad \begin{aligned} \max \quad & \mathbf{b}^{\mathrm{T}}\boldsymbol{\lambda} \\ \text{s.t.} \quad & \|\boldsymbol{\lambda}\|_2 \leq 1 \\ & \|\mathbf{A}_0^{\mathrm{T}}\boldsymbol{\lambda}\|_2 \leq \rho\|\boldsymbol{\lambda}\|_2. \end{aligned}$$

The feasible set of (DR-LS) is contained in the feasible set of (OD-LS), thus verifying the inequality

$$\text{val(OD-LS)} \leq \text{val(DR-LS)}.$$

Strict inequality may occur. For example, let us take $m = n = 1$ and $A_0 = 2$, $b = -2$, $\rho = 1$. (DR-LS) is just the problem

$$\max\{-2\lambda : |\lambda| \leq 1, |2\lambda| \leq 1\},$$

whose optimal solution $\lambda = -0.5$ with an optimal function value of 1. On the other hand, (OD-LS) is

$$\text{(OD-LS)} \max\{-2\lambda : |\lambda| \leq 1, |2\lambda| \leq |\lambda|\}.$$

The only feasible point in (OD-LS) is $\lambda = 0$. Thus, in this case the optimal value of (OD-LS) is zero and is strictly smaller than the optimal value of (DR-LS). □

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 optimization—methodology and applications, Math. Program. Ser. B 92 (3) (2002) 453–480. ISMP 2000, Part 2 (Atlanta, GA).

[3] A. Björck, Numerical Methods for Least-Squares Problems, SIAM, Philadelphia, PA, 1996.

[4] S. Chandrasekaran, G.H. Golub, M. Gu, A.H. Sayed, Efficient algorithms for least squares type problems with bounded uncertainties, in: Recent Advances in Total Least Squares Techniques and Errors-in-variables Modeling (Leuven, 1996), SIAM, Philadelphia, PA, 1997, pp. 171–180.

[5] L. El Ghaoui, H. Lebret, Robust solution to least-squares problems with uncertain data, SIAM J. Matrix Anal. Appl. 18 (4) (1997) 1035–1064.

[6] R.T. Rockafellar, Convex Analysis, Princeton Univ. Press, Princeton, NJ, 1970.