Deriving Robust Counterparts of Nonlinear Uncertain Inequalities

Ben-Tal, A.; den Hertog, D.; Vial, J.P.

Publication date:
2012

CentER

Discussion Paper

No. 2012-0053

DERIVING ROBUST COUNTERPARTS OF NONLINEAR UNCERTAIN INEQUALITIES

By

Aharon Ben-Tal, Dick den Hertog,
Jean-Philippe Vial

July 2, 2012

ISSN 0924-7815

TILBURG UNIVERSITY

Deriving robust counterparts of nonlinear uncertain inequalities

Aharon Ben-Tal *†

Department of Industrial Engineering and Management,
Technion – Israel Institute of Technology, Haifa 32000, Israel
CentER Extramural Fellow, CentER, Tilburg University, The Netherlands

Dick den Hertog

Ticopt, Department of Econometrics and Operations Research,
Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands

Jean-Philippe Vial

Ordecsys, Place de l'Etrier 4, CH-1224, Chêne-Bougeries, Switzerland

July 2, 2012

Abstract

In this paper we provide a systematic way to construct the robust counterpart of a nonlinear uncertain inequality that is concave in the uncertain parameters. We use convex analysis (support functions, conjugate functions, Fenchel duality) and conic duality in order to convert the robust counterpart into an explicit and computationally tractable set of constraints. It turns out that to do so one has to calculate the support function of the uncertainty set and the concave conjugate of the nonlinear constraint function. Conveniently, these two computations are completely independent. This approach has several advantages. First, it provides an easy structured way to construct the robust counterpart both for linear and nonlinear inequalities. Second, it shows that for new classes of uncertainty regions and for new classes of nonlinear optimization problems tractable counterparts can be derived. We also study some cases where the inequality is nonconcave in the uncertain parameters.

Keywords: Fenchel duality, robust counterpart, nonlinear inequality, robust optimization, support functions
JEL codes: C61

1 Introduction

Robust Optimization (RO) has become an important field in the last decade. For a comprehensive treatment of RO we refer to [5] or the recent survey [10]. The number of applications

*Part of this paper was written when the author was visiting Centrum Wiskunde & Informatica in Amsterdam, The Netherlands, as a CWI Distinguished Scientist.

†Research partly supported by BSF Grant 2008302.

of RO has increased rapidly in recent years. Moreover, the RO methodology has recently been implemented into (commercial) mathematical modelling and optimization systems (e.g. AIMMS [1], ROME [14], and YALMIP [18]). The goal of RO is to immunize an optimization problem against uncertain parameters in the problem. Such uncertain parameters may arise as a result of estimation or rounding errors in the parameter values, or due to implementation errors. Therefore, a so-called *uncertainty region* for the uncertain parameters is defined, and then it is required that the constraints should hold for all parameter values that reside in the uncertain region. For several optimization problems, and for several choices of this uncertainty region, the so-called *Robust Counterpart (RC)* can be formulated as a tractable optimization problem. For example the robust counterpart for a linear programming problem with polyhedral or ellipsoidal uncertainty regions can be reformulated as a linear programming or conic quadratic programming problem, respectively. We refer to [5] for an extensive treatment of these cases.

In this paper we show how Fenchel duality can be used to construct the RC for a nonlinear constraint that is concave in the uncertain parameters. This approach has several advantages. First of all, it shows that for new classes of optimization problems a tractable robust counterpart can be derived. Secondly, for linear programming it provides an easy way to construct the RC, e.g. in case the uncertainty region is the intersection of several regions. In this paper many examples are given, some being of direct applications to known problems, some being candidates for future applications.

We also analyze the tractability of the resulting robust counterparts and show that many RCs can be written as linear, quadratic, or conic quadratic constraints, or else they can be shown to admit a self-concordant barrier function. The latter implies that the optimization problem is solvable in polynomial-time (see [20]).

We now discuss other literature on deriving robust counterparts for nonlinear inequalities following the description in [5]. Exact tractable reformulations for Conic Quadratic Problems are derived when the uncertainty is a simple interval or unstructured norm-bounded. Approximation results are derived for structured norm-bounded uncertainty or intersection of ellipsoids. For Semidefinite Problems exact results are known for unstructured norm-bounded uncertainty, and approximation results for structured norm-bounded uncertainty. In [3] an explicit tractable reformulation is derived for a robust inequality that is linear in the uncertain parameters and the uncertainty set is given by convex inequalities. This is a special case of the type of robust optimization problems that we study in this paper.

This paper is organized as follows. In Section 2 we derive the RC of a general nonlinear robust constraint by using Fenchel duality. The RC involves calculating the support function of the uncertainty region, which is treated in Section 3, and calculating the conjugate function of the constraint function, which is treated in Section 4, using among other methods the theory of Conic Quadratic representable (CQr) functions ([6]). Many concrete examples are treated in Sections 3 and 4. In Section 5 we treat the complexity of the resulting RCs. In the appendices, we recall known theorems on conjugate functions, support functions, Fenchel duality, Conic Quadratic duality and Conic Quadratic representability. Those results will be used in the paper and for the sake of clarity the lemmas and theorems in the appendix will be unambiguously numbered as such.

Notation. Throughout this paper we use the following notation.

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a closed convex function with $\mathrm{dom}(f) = \{x \mid f(x) < \infty\}$. Let $g : \mathbb{R}^n \to \mathbb{R}$ be a closed concave function with $\mathrm{dom}(g) = \{x \mid g(x) > -\infty\}$. The *convex conjugate* of f is defined as

$$f^*(y) = \sup_{x \in \mathrm{dom}(f)} \{y^T x - f(x)\}.$$

The *concave conjugate* of g is defined as

$$g_*(y) = \inf_{x \in \mathrm{dom}(g)} \{y^T x - g(x)\}. \tag{1}$$

For a function $g(.,.)$ of two vector variables, $g_*(.,.)$ will denote the partial concave conjugate with respect to the first variable.

A special conjugate function is the *support function*, which is the conjugate of the *indicator function*. The indicator function on the set S is defined as:

$$\delta(x|S) = \begin{cases} 0 & \text{if} \quad x \in S \\ \infty & \text{otherwise.} \end{cases} \tag{2}$$

Then the conjugate function of $\delta(x|S)$,

$$\delta^*(y|S) = \sup_{x \in S} \; y^T x, \tag{3}$$

is the so-called support function of the set S.

The *adjoint function* of $f : \mathbb{R}_+ \longrightarrow \mathbb{R}$ is defined as:

$$f^\diamondsuit(x) = x f\left(\frac{1}{x}\right), \quad \forall x \in \mathbb{R}_+. \tag{4}$$

Note that $f^\diamondsuit(x)$ is convex if $f(x)$ is convex.

Let A be a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$, then the function gA is defined by

$$(gA)(x) = g(Ax).$$

The relative interior of a set S is denoted by $\mathrm{ri}(S)$.

Throughout the paper a subindex i denotes a scalar (the ith element of the vector) and a superindex denotes a vector.

2 Robust Counterpart (RC) of a general inequality

The Robust Optimization (RO) methodology addresses optimization problems affected by parameter uncertainty in constraints. We focus here on a nonlinear constraint

$$f(a, x) \leq 0, \tag{5}$$

where $x \in \mathbb{R}^n$ is the optimization variable, $f(., x)$ is concave for all $x \in \mathbb{R}^n$, and $a \in \mathbb{R}^m$ is an uncertain vector, which is only known to reside in a set U (called the *uncertainty set*). The robust counterpart of (5) is then

$$(RC) \qquad f(a, x) \leq 0, \quad \forall a \in U, \tag{6}$$

where the uncertainty set U is modeled as follows:

$$U = \{a = a^0 + A\zeta \mid \zeta \in Z \subset \mathbb{R}^L\}.$$

Here $a^0 \in \mathbb{R}^m$ is the so-called "nominal value", the matrix A is given column wise: $A = (a^1\ a^2\ ...\ a^L) \in \mathbb{R}^{m\times L}$, ζ is called the vector of "primitive uncertainties", and Z is a given nonempty, convex and compact set, with $0 \in \mathrm{ri}(Z)$.

With this formulation, it is required to determine the value of x before the actual realization of a is available ("here and now" decisions).

Definition 1 *The nominal vector a^0 is called* regular *if $a^0 \in ri(domf(.,x)), \forall x$.*

Note that when a^0 is regular and since $0 \in \mathrm{ri}(Z)$ then the following holds:

$$ri(U) \cap ri(domf(.,x)) \neq \emptyset,\ \forall x. \tag{7}$$

The robust inequality (RC) can be rewritten as

$$\max_{a\in U} f(a,x) \leq 0. \tag{8}$$

In this paper we frequently use a *general principle* to process (RC). For that purpose we use the dual of the optimization problem in the left-hand side of (8). The dual has the general form

$$\min\{g(b,x) \mid b \in Z(x)\}. \tag{9}$$

Under suitable convexity and regularity conditions on $f(.,x)$ and U (such as (7)) strong duality holds between the maximization problem in (8) and (9), hence x is robust feasible if and only if

$$\min\{g(b,x) \mid b \in Z(x)\} \leq 0. \tag{10}$$

So finally, x is robust feasible for (6) if and only if x and b solve the system

$$\begin{cases} g(b,x) \leq 0 \\ b \in Z(x). \end{cases} \tag{11}$$

In case strong duality does not hold, we still have (by weak duality) that (10) implies (9). Hence, whenever x and some b solve (11), then x satisfies (6), i.e. it is robust feasible.

The next basic result gives an equivalent reformulation for (RC) which is used throughout the paper to derive tractable RCs.

Theorem 2 *Let a^0 be regular. Then the vector $x \in \mathbb{R}^n$ satisfies (RC) if and only if $x \in \mathbb{R}^n$, $v \in \mathbb{R}^m$ satisfy the single inequality*

$$(FRC) \qquad (a^0)^T v + \delta^*(A^T v|Z) - f_*(v,x) \leq 0, \tag{12}$$

in which the support function δ^* and the partial concave conjugate function f_* are defined in (3) and (1), respectively.

Proof: Using the definition of indicator functions (2) and using Fenchel duality (Theorem A.1), we have

$$
\begin{aligned}
F(x) &:= \max_{a \in U} f(a,x) && (13)\\
&= \max_{a \in \mathbb{R}^m} \{f(a,x) - \delta(a|U)\} && (14)\\
&= \min_{v \in \mathbb{R}^m} \{\delta^*(v|U) - f_*(v,x)\}, && (15)
\end{aligned}
$$

where

$$
f_*(v,x) = \inf_{a \in \mathbb{R}^m} \{a^T v - f(a,x)\},
$$

and

$$
\begin{aligned}
\delta^*(v|U) &= \sup_{\zeta \in Z} \{a^T v \mid a = a_0 + A\zeta\} && (16)\\
&= (a^0)^T v + \sup_{\zeta \in Z} v^T A \zeta && (17)\\
&= (a^0)^T v + \delta^*(A^T v|Z). && (18)
\end{aligned}
$$

The passage from (14) to (15) is justified, since the regularity of a^0 implies condition (7), which is the condition needed by Fenchel Duality Theorem to obtain strong duality. Since (RC) in (6) can be written as $F(x) \leq 0$, the theorem follows immediately. □

The following corollary is used frequently in this paper.

Corollary 3 *Let $Z_1, ..., Z_K$ be closed convex sets with $0 \in ri(Z_k), \forall k$, such that $Z = \bigcap_{k=1}^K ri(Z_k) \neq \emptyset$. Moreover, let $f(a,x) = \sum_{j=1}^J f_j(a,x)$. If a^0 is regular then $x \in \mathbb{R}^n$ satisfies (RC) if and only if $x \in \mathbb{R}^n$, $v \in \mathbb{R}^m$, $y^k \in \mathbb{R}^L$, $z^j \in \mathbb{R}^m$ satisfy the following system of constraints:*

$$
(SFRC) \qquad \begin{cases} (a^0)^T v + \sum_{i=1}^K \delta^*(y^i|Z_i) - \sum_{j=1}^J (f_j)_*(z^j, x) \leq 0 \\ \sum_{k=1}^K y^k = A^T v \\ \sum_{j=1}^J z^j = v. \end{cases} \qquad (19)
$$

Proof: Using Lemmas A.2 and A.4, (FRC) can be written as:

$$
(a^0)^T v + \min \left\{ \sum_{k=1}^K \delta^*(y^k|Z_k) \mid \sum_{k=1}^K y^k = A^T v \right\} - \max \left\{ \sum_{j=1}^J (f_j)_*(z^j, x) \mid \sum_{j=1}^J z^j = v \right\} \leq 0,
$$

which is equivalent to $(SFRC)$, by using the general principle mentioned in the beginning of this section. □

The following remarks with respect to Theorem 2 are important:

1. In (FRC) the computation involving f are completely independent from those involving Z.

2. To derive (FRC) we did not assume $f(a,x)$ to be convex in x. However, if $f(a,x)$ is convex in x, $\forall a \in U$, then $f_*(v,x)$ is concave in (v,x), so then (FRC) is a convex inequality in (v,x).

3. It is interesting to observe that robustifying a nonlinear constraint may have a "convexification effect". This is illustrated in the next nominal constraint which is nonconvex, but whose robust counterpart is convex. We consider the following robust counterpart:

$$f(a,x) := \sum_{i=1}^{m} a_i f_i(x) \leq b \quad \forall a \in U,$$

where $U = \{a \in \mathbb{R}^m \mid \|a - a^0\|_\infty \leq \rho\}$, and let $\rho + a_i^0 \geq 0,\ i = 1, ..., m$. We assume that $f_i(x), i = 1, ..., m$, are convex and $f_i(x) \geq 0$, $\forall x$. Suppose (some of) the nominal values a_i^0 are negative, which means that the nominal inequality

$$\sum_{i=1}^{m} a_i^0 f_i(x) \leq b$$

may not be convex. However, in this case the (FRC)

$$\sum_{i=1}^{m} (a_i^0 + \rho) f_i(x) \leq b$$

is indeed a convex inequality.

4. If we consider additive implementation error

$$f(x + a) \leq 0, \quad \forall a \in U, \tag{20}$$

then, except for the linear case, it can not happen that f is convex in x and concave in a. However, in several cases, e.g. when $f(x)$ is quadratic, one can reformulate (20) in such a way that this assumption does hold. See also Section 4.3.

5. If $f(a, x)$ is not concave in a, then it can be easily verified (by weak duality) that

$$F(x) = \max_{a \in U} f(a, x) \leq (a^0)^T v + \delta^*(A^T v|Z) - f_*(v, x),$$

and thus (FRC) is a tighter constraint, i.e., if a pair (x, v) satisfies (FRC) then x is robust feasible.

6. Suppose that there is also uncertainty in the right-hand-side, i.e., the RC is

$$f(a, x) \leq b, \quad \forall (a, b) \in \tilde{U},$$

$$\tilde{U} = \left\{ \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} + \sum_{i=1}^{L} \zeta_i \begin{pmatrix} a_i \\ b_i \end{pmatrix} \mid \zeta \in Z \subset \mathbb{R}^L \right\}, \tag{21}$$

where $a_0, a_1, ..., a_L \in \mathbb{R}^m$ are given (a_0 is the "nominal value"), $b_0, b_1, ..., b_L \in \mathbb{R}$ are given (b_0 is the "nominal value"), and Z is a given nonempty, convex and compact set. Then it can easily be verified that, under the condition that a^0 is regular, this is equivalent to

$$(a^0)^T v + \delta^*(A^T v - b|Z) - f_*(v, x) \leq b_0, \tag{22}$$

where $b = (b_1, ..., b_L)^T$ and $A = (a^1\ a^2\ ...\ a^L)$.

To process (FRC) we have to compute $\delta^*(A^T v|Z)$ and $f_*(v, x)$. This is treated in Sections 3 and 4, respectively.

3 Computing support functions

In this section we illustrate how to compute $\delta^*(A^T v|Z)$ in (FRC) for several important choices of Z. We start with an example for which the support function can explicitly be constructed. It includes interval and ellipsoidal uncertainty.

Example 4 **p-norm.** *Let $Z = \{\zeta \mid \|\zeta\|_p \leq \rho\}$, where $\|.\|_p$ is the p-norm. For the support function of Z we have:*

$$\delta^*(y|Z) = \rho\|y\|_p^* = \rho\|y\|_q, \quad \textit{where } 1/p + 1/q = 1.$$

Here $\|.\|_p^*$ is the dual norm of the p-norm. We conclude that for this choice of Z, (FRC) is equivalent to

$$(a^0)^T v + \rho\|A^T v\|_q - f_*(v,x) \leq 0.$$

This result is also obtained in [12] for linear constraints. □

For most examples given in this section we apply again the general principle, i.e., we write the support function as a minimization problem (e.g. by using duality) and omit the 'min'.

Example 5 **Polyhedron.** *Let $Z = \{\zeta \mid B\zeta \leq b\}$, where $B \in \mathbb{R}^{K \times L}, b \in \mathbb{R}^K$. For the support function of Z we have:*

$$\delta^*(y|Z) = \max_\zeta\{y^T\zeta \mid B\zeta \leq b\} = \min_z\{b^T z \mid B^T z = y,\ z \geq 0\},$$

where the last equality follows from LP duality. We conclude that for this choice of Z, (FRC) is equivalent to

$$\begin{cases} (a^0)^T v + b^T z - f_*(v,x) \leq 0 \\ B^T z = A^T v \\ z \geq 0. \end{cases}$$

□

The next example is a generalization of polyhedral uncertainty.

Example 6 **Cone-based set.** *Let $Z = \{\zeta \mid b - B\zeta \in C\}$, where $B \in \mathbb{R}^{K \times L}, b \in \mathbb{R}^K$, and C is a pointed cone that contains a strictly feasible solution (i.e., there exists a $\bar{\zeta}$ such that $b - B\bar{\zeta} \in int(C)$). By using conic duality, we have for the support function of Z:*

$$\delta^*(y|Z) = \max_\zeta\{y^T\zeta \mid b - B\zeta \in C\} = \min_z\{b^T z \mid B^T z = y,\ z \in C^*\}. \tag{23}$$

We conclude that for this choice of Z, (FRC) is equivalent to

$$\begin{cases} (a^0)^T v + b^T z - f_*(v,x) \leq 0 \\ B^T z = A^T v \\ z \in C^*. \end{cases}$$

□

Example 7 Conic Quadratic representable sets. *In [6] it is shown that many sets are Conic Quadratic representable (CQr). Many simple operations have been proved to preserve Conic Quadratic representability of sets. Using this calculus it is rather easy to examine whether a set is CQr. See Appendix B for a short survey on CQr. If the uncertainty set is CQr, we can use the result of the previous example to derive a tractable formulation of the robust counterpart. See Appendix B and [6] for many examples of such sets.* □

Example 8 Semi Definite representable sets. *In [6] it is shown that many sets are Semi Definitie representable (SDr). As an example, let us consider the set*

$$Z = \{\zeta \mid \lambda_{max}(B(\zeta)) \leq \rho\},$$

where $B(\zeta)$ is a matrix whose elements are linear in ζ, and λ_{max} denotes the maximum eigenvalue. This set can be reformulated as

$$Z = \{\zeta \mid \rho I - B(\zeta) \succeq 0\},$$

which is a Semi Definite Representation, for which the result of Example 6 can be used to derive the support function. □

To compute δ^* for uncertainty regions specified by convex constraints, we first give some helpful lemmas.

Lemma 9 *Let $Z = \{\zeta \mid h_k(\zeta) \leq 0,\ k = 1, ..., K\}$, where $h_k(.)$ is convex. Moreover, we assume that $\cap_{k=1}^K ri(dom\, h_k) \neq \emptyset$. Then*

$$\delta^*(y|Z) = \min_{u \geq 0} \left\{ \sum_{k=1}^K u_k h_k^* \left(\frac{v^k}{u_k} \right) \;\middle|\; \sum_{k=1}^K v^k = y \right\}. \tag{24}$$

Moreover, if $h_k(\zeta) = \sum_{l=1}^{L_k} h_{kl}(\zeta)$, then

$$\delta^*(y|Z) = \min_{u \geq 0, \{w^{kl}\}} \left\{ \sum_{k=1}^K \sum_{l=1}^{L_k} u_k h_{kl}^* \left(\frac{w^{kl}}{u_k} \right) \;\middle|\; \sum_{k=1}^K \sum_{l=1}^{L_k} w^{kl} = y \right\}, \tag{25}$$

and if $h_k(\zeta) = \sum_{l=1}^{L_k} h_{kl}(\zeta_l)$ (separable case), then

$$\delta^*(y|Z) = \min_{u \geq 0, \{w_{kl}\}} \left\{ \sum_{k=1}^K \sum_{l=1}^{L_k} u_k h_{kl}^* \left(\frac{w_{kl}}{u_k} \right) \;\middle|\; \sum_{k=1}^K w_{kl} = y_l,\ l = 1, ..., L \right\}. \tag{26}$$

Proof: It can easily be verified that

$$\begin{aligned}
\delta^*(y|Z) &= \max_{\zeta} \{y^T \zeta \mid h_k(\zeta) \leq 0,\ k = 1, ..., K\} \\
&= \min_{u \geq 0} \left(\sum_{k=1}^K u_k h_k \right)^* (y) \\
&= \min_{u \geq 0} \left\{ \sum_{k=1}^K u_k h_k^* \left(\frac{v^k}{u_k} \right) \;\middle|\; \sum_{k=1}^K v^k = y \right\},
\end{aligned}$$

where the last equality follows by Lemma A.1 and property (51). In the case of separability we have for the support function of Z, starting from (24) and using Corollary A.1,

$$\begin{aligned}
\delta^*(y|Z) &= \max_{\zeta} \left\{ y^T\zeta \;\middle|\; \sum_{l=1}^{L_k} h_{kl}(x_l) \leq 0,\ k=1,...,K \right\} \\
&= \min_{u\geq 0,\{v_k\}} \left\{ \sum_{k=1}^{K} u_k \left(\sum_{l=1}^{L_k} h_{kl} \right)^* \left(\frac{v^k}{u_k} \right) \;\middle|\; \sum_{k=1}^{K} v^k = y \right\} \\
&= \min_{u\geq 0,\{w_{kl}\}} \left\{ \sum_{k=1}^{K}\sum_{l=1}^{L_k} u_k h_{kl}^* \left(\frac{w_{kl}}{u_k} \right) \;\middle|\; \sum_{k=1}^{K} w_{kl} = y_l,\ l=1,...,L \right\}.
\end{aligned}$$

□

Corollary 10 *Let a^0 be regular and let $Z = \{\zeta \mid h_k(\zeta) \leq 0,\ k=1,...,K\}$, where $h_k(.)$ is convex. Moreover, we assume that $\cap_{k=1}^K ri(dom\, h_k) \neq \emptyset$. Then (FRC) is equivalent to*

$$\begin{cases} (a^0)^T v + \sum_{k=1}^K u_k h_k^* \left(\frac{v^k}{u_k} \right) - f_*(v,x) \leq 0 \\ \sum_{k=1}^K v^k = A^T v \\ u \geq 0. \end{cases} \tag{27}$$

□

We now give some concrete examples of uncertainty regions defined by nonlinear inequalities. The first three examples are separable cases.

Example 11 **ϕ-divergence uncertainty.** *In [8] it is proposed, in case the uncertain parameter a in (6) can be considered as a probability vector, to use uncertainty regions defined by so-called ϕ-divergence functions. Let us concentrate in this example on the well-known Kullback-Leibler function:*

$$h_l(\zeta_l) = \zeta_l \log\left(\frac{\zeta_l}{\zeta_l^0} \right) - \frac{\rho}{L},$$

where ζ_l^0 $(l=1,...,L)$ are given nominal values. In this case the uncertainty region is defined by

$$Z = \left\{ \zeta \in \mathbb{R}^L \mid \sum_{l=1}^{L} \zeta_l \log\left(\frac{\zeta_l}{\zeta_l^0} \right) \leq \rho \right\}.$$

Thus $h_l^*(s) = \zeta_l^0 e^{s-1} + \frac{\rho}{L}$ and (FRC) is

$$(a^0)^T v + \sum_{l=1}^{L} \zeta_l^0 u e^{\frac{(a^l)^T v}{u} - 1} + \rho u - f_*(v,x) \leq 0.$$

In Table 3 many more examples for h_l and their conjugates are given. □

Example 12 **Using the adjoint.** *Let $Z = \{\zeta \mid \sum_l h_l(\zeta_l) \leq 0\}$, where h_l is convex. Using Lemma 9, the (FRC) can be written as*

$$(a^0)^T v + \sum_{l=1}^{L} u h_l^* \left(\frac{(a^l)^T v}{u} \right) - f_*(v,x) \leq 0.$$

It may be the case that there is no closed form for h_l^*. Then, one can check whether there is a closed form for the conjugate of the adjoint (see definition (4)), and then Lemma A.5 can be used. We obtain for (FRC):

$$\begin{cases} (a^0)^T v + \sum_{l=1}^L y_l - f_*(v,x) \le 0 \\ u\left(h_l^{\diamondsuit}\right)^* \left(-\frac{y_l}{u}\right) \le -(a^l)^T v \quad \forall l. \end{cases}$$

As a concrete example, take $h_l(t) = te^{1/t}$, for which there is no closed form conjugate. However, $h_l^{\diamondsuit}(t) = e^t$, which has a closed form conjugate (see Table 3). □

We now give some examples in which the constraint functions that define the uncertainty region are not separable. For ease of notation we assume that there is only one such constraint $h(\zeta) \le 0$ that defines the uncertainty region. The extension to multiple constraints is straightforward. The first class of problems is where $h(\zeta) = g(D^T\zeta) = \sum_i g_i(d_i^T\zeta)$, and $h_i(\zeta) = g_i(d_i^T\zeta)$, where $D \in \mathbb{R}^{L\times r}$, d_i is the i-th column of D, $g : \mathbb{R}^r \longrightarrow \mathbb{R}$, and $g_i : \mathbb{R} \longrightarrow \mathbb{R}$, $\forall i$. Using Corollary 10 and Lemma A.7 we obtain for the support function

$$\begin{aligned} \delta^*(y|Z) &= \min_{u\ge 0,\ \{w^i\}} \left\{ \sum_{i=1}^r u h_i^*\left(\frac{w^i}{u}\right) \mid \sum_{i=1}^r w^i = y \right\} \\ &= \min_{u\ge 0,\ \{w^i\}} \left\{ u \sum_{i=1}^r \inf_{z_i} \left(g_i^*\left(\frac{z_i}{u}\right) \mid z_i d^i = w^i \right) \mid \sum_{i=1}^r w^i = y \right\}. \end{aligned}$$

Hence, for this choice of the uncertainty region, (FRC) becomes:

$$\begin{cases} (a^0)^T v + u\sum_{i=1}^r g_i^*\left(\frac{z_i}{u}\right) - f_*(v,x) \le 0 \\ Dz = A^T v \\ u \ge 0. \end{cases} \tag{28}$$

We apply this result to uncertainty regions defined by geometric and l_p-programming constraints.

Example 13 Uncertainty region defined by geometric programming constraints. *Let $h(\zeta) = \sum_{i=1}^r \alpha_i e^{(d^i)^T\zeta}$, where $\alpha_i > 0$, for all i. Hence $g_i(t) = \alpha_i e^t$, and $g_i^*(s)$ is given in Table 3. Then (28) becomes*

$$\begin{cases} (a^0)^T v + \sum_{i=1}^r \left\{ z_i \log\left(\frac{z_i}{\alpha_i u}\right) - z_i \right\} - f_*(v,x) \le 0 \\ Dz = A^T v \\ u, z \ge 0. \end{cases}$$

□

Example 14 Uncertainty region defined by l_p-programming constraints. *Let $h(\zeta) = \sum_{i=1}^r \frac{\alpha_i}{p_i} |(d^i)^T\zeta - \beta_i|^{p_i}$, where $\alpha_i > 0$ and $p_i > 1$, for all i. Hence $g_i(t) = \frac{\alpha_i}{p_i}|t-\beta_i|^{p_i}$, and $g_i^*(s)$ can be easily derived from Table 3. Then (28) becomes*

$$\begin{cases} (a^0)^T v + u\sum_{i=1}^r \frac{\alpha_i}{q_i} \left|\frac{z_i}{\alpha_i u}\right|^{q_i} + \beta^T z - f_*(v,x) \le 0 \\ Dz = A^T v \\ u \ge 0, \end{cases}$$

where $1/p_i + 1/q_i = 1$. □

Example 15 Uncertainty based on Anderson-Darling test. *Suppose the uncertain parameter ζ in (6) can be considered as a probability vector as in Example 15. Instead of using ϕ-divergence test statistics, one could also use the well-known Anderson-Darling test statistic [2] for the definition of the uncertainty region. This leads to*

$$Z = \left\{ \zeta \in \mathbb{R}^L \;\middle|\; -\sum_{k=1}^{N} \left(\ln(\mathbf{1}^k)^T\zeta + \ln(\mathbf{1}^{-k})^T\zeta \right) \le \rho \right\},$$

where N is the number of observations and where the first k elements of the vector $\mathbf{1}^k \in \mathbb{R}^L$ are 1, and the others 0, and the last k elements of the vector $\mathbf{1}^{-k}$ are 1, and the others 0, i.e., $\mathbf{1}^{-k} = \mathbf{1} - \mathbf{1}^k$, where $\mathbf{1}$ is the all one vector. Applying (28) and Corollary 3 we get for (FRC):

$$\begin{cases} (a^0)^T v - \sum_{k=1}^N u\left[1 + \ln \frac{-z_k^+}{u}\right] - \sum_{k=1}^N u\left[1 + \ln \frac{-z_k^-}{u}\right] + \rho u - f_*(v,x) \le 0 \\ z_k^+ e^k = w^k, \quad k = 1, ..., N \\ z_k^- e^{-k} = w^{-k}, \quad k = 1, ..., N \\ \sum_{k=1}^N (w^k + w^{-k}) = A^T v \\ z_k^+, z_k^- \le 0 \quad k = 1, ..., N \\ u \ge 0. \end{cases}$$

□

Example 16 Conic Quadratic representable functions. *Let $z = \{\zeta \mid h_k(\zeta) \le 0,\ k = 1, ..., K\}$. If the functions $h_k(\zeta)$ that define the uncertainty set Z are CQr, then there are two ways to obtain a tractable robust counterpart. The first way is to construct a Conic Quadratic representation of Z, and then apply conic duality (see Example 8). The second way is to use the result of Lemma 9 and construct a Conic Quadratic representation for the convex functions $u_k h_k^*(\frac{v^k}{u_k})$. In Appendix B it is shown that these functions are CQr if h_k, $k = 1, ..., K$, are CQr.*
□

Computing the support function for the intersection or Minkowski sum of sets is relatively easy by using Lemmas A.4 and A.3, respectively. This is illustrated in the next examples.

Example 17 Intersection of 1, 2, and ∞ norm ball. *Let $Z_k = \{\zeta \mid \|\zeta\|_k \le \rho_k\}$, $\rho_k > 0$, for $k = 1, 2, \infty$, and let $Z = Z_1 \cap Z_2 \cap Z_\infty$. Then*

$$\delta^*(y|Z) = \min_{w^1,w^2,w^\infty} \{\delta^*(w^1|Z_1) + \delta^*(w^2|Z_2) + \delta^*(w^\infty|Z_\infty) \mid w^1 + w^2 + w^\infty = y\} \tag{29}$$

$$= \min_{w^1,w^2,w^\infty} \{\rho_1\|w^1\|_\infty + \rho_2\|w^2\|_2 + \rho_\infty\|w^\infty\|_1 \mid w^1 + w^2 + w^\infty = y\}, \tag{30}$$

where the first equality follows from Lemma A.4 and the second equality from Example 4. We conclude that for this choice of Z, (FRC) is equivalent to

$$\begin{cases} (a^0)^T v + \rho_1\|w^1\|_\infty + \rho_2\|w^2\|_2 + \rho_\infty\|w^\infty\|_1 - f_*(v,x) \le 0 \\ w^1 + w^2 + w^\infty = A^T v. \end{cases}$$

□

Example 18 Entropy uncertainty region. *In [19] a so-called entropy uncertainty region is derived to obtain a safe approximation of a chance constraint. This region is defined as*

$$Z = \left\{ \zeta \;\middle|\; \|\zeta\|_\infty \leq 1,\ \sum_{l=1}^{L} \{(1+\zeta_l)\ln(1+\zeta_l) + (1-\zeta_l)\ln(1-\zeta_l)\} \leq \beta \right\}.$$

To derive a tractable (FRC), let us define

$$Z_1 = \{\zeta \mid \|\zeta\|_\infty \leq 1\} \quad \text{and} \quad Z_2 = \left\{ \zeta \;\middle|\; \sum_{l=1}^{L} \{(1+\zeta_l)\ln(1+\zeta_l) + (1-\zeta_l)\ln(1-\zeta_l)\} \leq \beta \right\}.$$

Then, using Lemma A.4 we have

$$\delta^*(y|Z) = \delta^*(y|Z_1 \cap Z_2) = \min_{w,z}\{\delta^*(w|Z_1) + \delta^*(z|Z_2) \mid w + z = y\}.$$

Moreover, if we define $h = h_1 + h_2$, where $h_1(\zeta_l) = (1+\zeta_l)\ln(1+\zeta_l)$ and $h_2(\zeta_l) = (1-\zeta_l))\ln(1-\zeta_l)$, then it can be verified (by using Lemma A.1) that

$$h^*(z) = \min_{s,t}\{h_1^*(s) + h_2^*(t) \mid s + t = z\} = \min_{s,t}\{e^{s-1} - s + e^{-t-1} + t \mid s + t = z\}.$$

Hence, using Lemma 9, the (FRC) can be written as

$$\begin{cases} (a^0)^T v + \|w\|_\infty + \sum_{l=1}^{L} \left\{ t_l - s_l + u \;\left[e^{s_l/u-1} + e^{-t_l/u-1} \right] \right\} - f_*(v,x) \leq 0 \\ w + s + t = A^T v \\ u \geq 0. \end{cases}$$

□

Example 19 Minkowski sum of sets. *Let $Z_1 = \{\zeta \mid \|\zeta\|_\infty \leq \rho_\infty\}$ and $Z_2 = \{\zeta \mid \| \zeta\|_2 \leq \rho_2\}$. Then $Z = Z_1 + Z_2$ is a box with "rounded corners". The support function for Z is:*

$$\delta^*(y|Z) = \delta^*(y|Z_1) + \delta^*(y|Z_2) = \rho_\infty \|y\|_1 + \rho_2 \|y\|_2,$$

where the first equality follows from Lemma A.3 and the second equality from Example 4. We conclude that for this choice of Z, (FRC) is equivalent to

$$(a^0)^T v + \rho_\infty \|A^T v\|_1 + \rho_2 \|A^T v\|_2 - f_*(v,x) \leq 0.$$

□

The following example shows the (FRC) for cases where the uncertainty set is the convex hull of sets.

Example 20 Convex hull of sets. *Let $Z = conv(Z_1, ..., Z_K)$. If $f(a,x)$ is linear in a, then the corresponding robust counterpart can be split over the individual sets Z_i, i.e.,*

$$f(a,x) \leq 0 \quad \forall a = a^0 + A\zeta, \quad \zeta \in Z_i \ \forall i,$$

and there is no need to explicitly construct the convex hull. If $f(a,x)$ is nonlinear then this does not hold anymore. However, since

$$\delta^*(A^T v|Z) = \max_i \delta^*(A^T v|Z_i),$$

we obtain for (FRC):

$$\left(a^0\right)^T v + \max_i \delta^*(A^T v|Z_i) - f_*(v,x) \leq 0.$$

Hence, even in the nonlinear case there is no need for an explicit construction of the convex hull of the sets Z_i. □

An overview of many classes of uncertainty regions and their support functions treated in this section, is given in Table 1.

Uncertainty region	Z	Robust Counterpart
Box	$\|\zeta\|_\infty \leq \rho$	$(a^0)^T v + \rho \|A^T v\|_1 - f_*(v,x) \leq 0$
Ball	$\|\zeta\|_2 \leq \rho$	$(a^0)^T v + \rho \|A^T v\|_2 - f_*(v,x) \leq 0$
Polyhedral	$b - B\zeta \geq 0$	$\begin{cases} (a^0)^T v + b^T z - f_*(v,x) \leq 0 \\ B^T z = A^T v \\ z \geq 0 \end{cases}$
Cone	$b - B\zeta \in C$	$\begin{cases} (a^0)^T v + b^T z - f_*(v,x) \leq 0 \\ B^T z = A^T v \\ z \in C^* \end{cases}$
Convex functions	$h_k(\zeta) \leq 0$ $k = 1, ..., K$	$\begin{cases} (a^0)^T v + \sum_k u_k h_k^* \left(\frac{w^k}{u_k}\right) - f_*(v,x) \leq 0 \\ \sum_k w^k = A^T v \\ u \geq 0 \end{cases}$
Separable functions	$\sum_{l=1}^{L_k} h_{kl}(\zeta_l) \leq 0$ $k = 1, ..., K$	$\begin{cases} (a^0)^T v + \sum_{k,l} u_k h_{kl}^* \left(\frac{w_{kl}}{u_k}\right) - f_*(v,x) \leq 0 \\ \sum_k w_{kl} = (a^l)^T v, \ \ l = 1, ..., L \\ u \geq 0 \end{cases}$
example	$\sum_l \zeta_l \log\left(\frac{\zeta_l}{\zeta_l^0}\right) \leq \rho$	$\begin{cases} (a^0)^T v + \sum_l \zeta_l^0 u e^{(a^l)^T v/u - 1} + \rho u - f_*(v,x) \leq 0 \\ u \geq 0 \end{cases}$
Composite separable functions	$g(D^T\zeta) \leq 0$	$\begin{cases} (a^0)^T v + u g^* \left(\frac{z}{u}\right) - f_*(v,x) \leq 0 \\ Dz = A^T v \\ u \geq 0 \end{cases}$
example	$\sum_i \alpha_i e^{(d_i)^T \zeta} \leq \rho$	$\begin{cases} (a^0)^T v + \sum_i \{z_i \log\left(\frac{z_i}{\alpha_i u}\right) - z_i\} - f_*(v,x) \leq 0 \\ Dz = A^T v \\ u \geq 0 \\ z \geq 0 \end{cases}$
Intersection	$Z = \cap_i Z_i$	$\begin{cases} (a^0)^T v + \sum_i \delta(w^i \mid Z_i) - f_*(v,x) \leq 0 \\ \sum_i w^i = A^T v \end{cases}$
example	$Z_k = \{\zeta \mid \|\zeta\|_k \leq \rho_k\}$ $k = 1, 2, \infty$	$\begin{cases} (a^0)^T v + \rho_1 \|w^1\|_\infty + \rho_2 \|w^2\|_2 + \rho_\infty \|w^\infty\|_1 \\ \qquad -f_*(v,x) \leq 0 \\ w^1 + w^2 + w^\infty = A^T v \end{cases}$
Minkowski sum	$Z = Z_1 + ... + Z_K$	$(a^0)^T v + \sum_i \delta(A^T v \mid Z_i) - f_*(v,x) \leq 0$
example	$Z_1 = \{\zeta \mid \|\zeta\|_\infty \leq \rho_\infty\}$ $Z_2 = \{\zeta \mid \|\zeta\|_2 \leq \rho_2\}$	$\begin{cases} (a^0)^T v + \rho_\infty \|A^T v\|_1 + \rho_2 \|A^T v\|_2 \\ \qquad -f_*(v,x) \leq 0 \end{cases}$
Convex hull	$Z = \mathrm{conv}(Z_1, ..., Z_K)$	$(a^0)^T v + \max_i \delta(A^T v \mid Z_i) - f_*(v,x) \leq 0$
example	$Z_1 = \{\zeta \mid \|\zeta\|_\infty \leq \rho_\infty\}$ $Z_2 = \{\zeta \mid \|\zeta - \zeta^0\|_2 \leq \rho_2\}$	$\begin{cases} (a^0)^T v + \max\{\rho_\infty \|A^T v\|_1, \left(A\zeta^0\right)^T v + \rho_2 \|A^T v\|_2\} \\ \qquad -f_*(v,x) \leq 0 \end{cases}$

Table 1: Robust optimization for different choices of the uncertainty region. The functions h_k, h_{kl}, g are assumed to be convex, with conjugates h_k^*, h_{kl}^*, g^* respectively. C^* denotes the dual cone of C.

4 Computing conjugate functions

In this section we show how to construct f_*. Besides several examples for which the conjugate function can be explicitly constructed, we also give examples for which we apply again the general principle, i.e., we write the conjugate function as a maximization problem and omit the 'max'.

We distinguish three categories: $f(a,x)$ is linear in a, concave in a, and nonconcave in a.

4.1 Linear uncertainty

Let $f(a,x)=\sum_{i=1}^m a_i f_i(x) = a^T f(x)$, $a \geq 0$, and $f_i(x)$ $(i=1,...,m)$ be convex. Since

$$f_*(v,x) = \inf_{a\in\mathbb{R}^m}\{a^T v - a^T f(x)\} = \begin{cases} 0 & v = f(x) \\ -\infty & v \neq f(x), \end{cases} \tag{31}$$

(FRC) becomes

$$(a^0)^T f(x) + \delta^*(A^T f(x), Z) \leq 0. \tag{32}$$

This is basically the type of inequalities studied in [3], where a similar result is obtained as (32). We now give some concrete examples.

Example 21 Linear in the optimization variables and ellipsoidal uncertainty. *Suppose* $f(a,x)=a^T x - \beta$, *and* $Z = \{\zeta \mid \|\zeta\|_2 \leq \rho\}$. *Note that we have*

$$f_*(v,x) = \begin{cases} \beta & \text{if} \quad v = x \\ -\infty & \text{if} \quad v \neq x. \end{cases} \tag{33}$$

Moreover, $\delta^*(A^T v|Z) = \rho\|A^T v\|_2$. *Hence,* (FRC) *in this case becomes:*

$$(a^0)^T x + \rho\|A^T x\|_2 \leq \beta.$$

$\square$

Example 22 Linear in the optimization variables and cone uncertainty. *Suppose* $f(a,x) = a^T x - \beta$, *and* $Z = \{\zeta \mid b - B\zeta \in C\}$, *where* C *is a pointed cone that contains a strictly feasible solution. By (23) we have* $\delta^*(A^T v|Z) = \min_z\{b^T z \mid B^T z = A^T v,\ z \in C^*\}$. *Hence,* (FRC) *in this case becomes:*

$$\begin{cases} (a^0)^T x + b^T z \leq \beta \\ B^T z = A^T x \\ z \in C^*. \end{cases}$$

$\square$

Example 23 Linear in the optimization variables and intersections of uncertainty regions. *Suppose* $f(a,x) = a^T x - \beta$, *and* $Z = Z_1 \cap Z_2 \cap Z_\infty$, *where* $Z_p = \{\zeta \mid \|\zeta\|_p \leq \rho_p\}$, *for* $p = 1, 2, \infty$. *Note that* $\delta^*(A^T v|Z_p) = \rho_p\|A^T v\|_q$, *where* $1/p + 1/q = 1$. *Using Lemma A.4 we obtain that* (FRC) *can be rewritten as as set of conic quadratic constraints:*

$$\begin{cases} (a^0)^T x + \rho_1\|A^T v^1\|_\infty + \rho_2\|A^T v^2\|_2 + \rho_\infty\|A^T v^\infty\|_1 \leq \beta \\ v^1 + v^2 + v^\infty = x. \end{cases} \tag{34}$$

$\square$

4.2 Concave uncertainty

We first study the case that $f(a,x)$ can be written as $f(a,x) = \sum_{i=1}^n f_i(a)x_i$, where $f_i(a)$ is concave, and $x_i \geq 0$. Note that

$$f_*(v,x) = \sup_{\{s^i\}} \left\{ \sum_{i=1}^n (x_i f_i)_* (s^i) \mid \sum_{i=1}^n s^i = v \right\} = \sup_{\{s^i\}} \left\{ \sum_{i=1}^n x_i (f_i)_* \left(\frac{s^i}{x_i} \right) \mid \sum_{i=1}^n s^i = v \right\}.$$

Hence (FRC) becomes

$$\begin{cases} (a^0)^T v + \delta^*(A^T v|Z) - \sum_{i=1}^n x_i (f_i)_*(s^i/x_i) \leq 0 \\ \sum_{i=1}^n s^i = v. \end{cases} \tag{35}$$

This is a convex system in the variables $v, \{s^i\}$, and x. When f is separable, i.e., $f_i(a) = f_i(a_i)$, it can easily be verified that (FRC) becomes

$$(a^0)^T v + \delta^*(A^T v|Z) - \sum_{i=1}^n x_i (f_i)_*(v_i/x_i) \leq 0.$$

We now consider examples in which $f_i(a)$ is not separable.

Example 24 Quadratic uncertainty. *Suppose that*

$$f_i(a) = -\frac{1}{2} a^T Q_i a,$$

where $Q_i \in \mathbb{R}^{m \times m}$ is a symmetric positive semi-definite matrix. If Q_i is nonsingular we obtain for (FRC):

$$\begin{cases} (a^0)^T v + \delta^*(A^T v|Z) + \frac{1}{2} \sum_{i=1}^n \frac{(s^i)^T Q_i^{-1} s^i}{x_i} \leq 0 \\ \sum_{i=1}^n s^i = v, \end{cases}$$

and if Q is singular then (FRC) becomes:

$$\begin{cases} (a^0)^T v + \delta^*(A^T v|Z) + \frac{1}{2} \sum_{i=1}^n \frac{(s^i)^T \tilde{Q}_i s^i}{x_i} \leq 0 \\ \sum_{i=1}^n s^i = v \\ s^i \in L_i \quad \forall i, \end{cases}$$

where $\tilde{Q}_i$ is the unique symmetric positive semi-definite matrix such that $Q_i \tilde{Q}_i = \tilde{Q}_i Q_i = P_i$, where P_i is the matrix of the linear transformation which projects $\mathbb{R}^m$ orthogonally onto the orthogonal complement L_i of the subspace $\{a \mid Q_i a = 0\}$. See [21] for more details on the conjugate of $f_i(a)$.

Note that both robust counterpart problems can be reformulated as a conic quadratic optimization problem. It is well-known that if Q_i is not positive semi-definite, and the uncertainty region is ellipsoidal, then the RC can be reformulated as a system of LMIs. See also Section 4.3. □

Example 25 Variance uncertainty. *Suppose that R is a discrete random variable, and that a_j, $j = 1, ..., m$, is the probability of outcome r_j, i.e., $a_j = P(R = r_j)$, and that $f(a,x) = var[h(x,R)]$, where $h(x,R)$ is a given function that is convex in x for each R. Note that*

$$f(a,x) = var[h(x,R)] = \sum_{j=1}^m h^2(x,r_j) a_j - \left(\sum_{j=1}^m h(x,r_j) a_j \right)^2 = \alpha(x)^T a - (\beta(x)^T a)^2,$$

where $\alpha(x), \beta(x) \in \mathbb{R}^m$, $\beta_j(x) = h(x, r_j)$, and $\alpha_j(x) = h^2(x, r_j) = \beta_j^2(x)$. Note that $f(a, x)$ is concave quadratic in a for all x. Define $g(t_1, t_2) = t_1 - t_2^2$, then $f(a, x) = g(\alpha(x)^T a, \beta(x)^T a)$. It can easily be verified that[1]

$$g_*(w, z) = \begin{cases} -\frac{1}{4}z^2 & \text{for } w = 1 \\ \infty & \text{else.} \end{cases}$$

Hence, (FRC) becomes

$$\begin{cases} (a^0)^T v + \delta^*(A^T v|Z) + \frac{z^2}{4} \leq 0 \\ \alpha(x) + z\beta(x) = v. \end{cases}$$

It can be shown that the equality in this robust counterpart problem can be replaced by an $\leq$ inequality:

$$\begin{cases} (a^0)^T v + \delta^*(A^T v|Z) + \frac{z^2}{4} \leq 0 \\ h(x, r_j)^2 + z h(x, r_j) \leq v, \quad j = 1, ..., m. \end{cases}$$

This robust counterpart is convex for fixed values of z. Hence, to solve the robust counterpart we can optimize for different values of z.

This example is very useful since the variance is often used as a risk measure. A famous example is the mean-variance portfolio optimization problem. $\square$

We observe that the results above can also be extended to the more general case where $f(a, x)$ can be written as $\sum_i f_i(a) g_i(x)$, where $g_i(x) \geq 0$.

Example 26 Transformed uncertainty region. *Suppose $f(\tilde{a}, x) = \tilde{a}^T x - \beta$ and the uncertainty region $\tilde{U}$ is defined as follows:*

$$\tilde{U} = \{\tilde{a} \mid h(\tilde{a}) = (h_1(\tilde{a}_1), ..., h_m(\tilde{a}_m))^T \in U\},$$

where $h_i(.)$ is convex for each i, and moreover we assume that h_i^{-1} exists for all i. This case is in fact the linear case, but the difference is that the uncertainty region is now stated in a transformed space. By substituting $\tilde{a} = h^{-1}(a)$, we obtain for (RC)

$$h^{-1}(a)^T x \leq \beta \quad \forall a \in U.$$

The corresponding (FRC) becomes:

$$(a^0)^T v + \delta^*(A^T v|Z) + \sum_{i=1}^n x_i((h_i)^{-1})_*(v_i/x_i) \leq \beta,$$

which is by Theorem A.6 equivalent to

$$(a^0)^T v + \delta^*(A^T v|Z) + \sum_{i=1}^n v_i(h_i)^*(x_i/v_i) \leq \beta. \tag{36}$$

The result shows that even if we cannot compute a closed form for h_i^{-1}, we can still construct the robust counterpart. As an example, take $h_i(t) = -(t + \log t)$. There is no closed form for h_i^{-1}, but we can still construct the robust counterpart (36) since $h_i^*(s) = -1 - \log(-(s+1))$. $\square$

[1]Here the conjugate is with respect to both arguments.

Finally we give some examples in which $f(a,x)$ cannot be written as $f(a)^T g(x)$, but still $f_*(v,x)$ can be computed.

Example 27 *Suppose $f(a,x) = -\sum_{i=1}^m x_i^{a_i}$, $x_i > 1$, and $0 \le a \le 1$. It can be verified that $f(a,x)$ is concave in a and convex in x. The robust counterpart (FRC) for this case becomes:*

$$\begin{cases} (a^0)^T v + \delta^*(A^T v|Z) - \sum_{i=1}^m \left(\frac{v_i}{\ln x_i} \ln \frac{-v_i}{\ln x_i} - \frac{v_i}{\ln x_i} \right) \le 0 \\ v \le 0. \end{cases}$$

□

Example 28 *Suppose $f(a,x) = g(a,x) + \theta(a,x)$, in which $g(a,x)$ is concave in a for each x, and*

$$\theta(a,x) = (\varphi)^{-1}(a_1 + \sum_{i=2}^m a_i \varphi(\alpha_i(x))),$$

where $\varphi(.)$ is convex and such that φ^{-1} exists, $a_i > 0, i = 1,...,m$, and $\alpha_i(x)$ is a linear function in x. Using the notation $\beta(x)^T = (1, \varphi(\alpha_1(x)),...,\varphi(\alpha_m(x)))^T$, we have

$$\theta(a,x) = \varphi^{-1}(a^T \beta(x)).$$

It can be verified that $\theta(a,x)$ is concave in a. For f_* we have

$$\begin{aligned} f_*(v,x) &= sup_w\{g_*(w,x) + \theta_*(v-w,x)\} \\ &= sup_w\{g_*(w,x) + sup_z\{(\varphi^{-1})_*(z) \mid z\beta(x) = v-w\}\} \\ &= sup_w\{g_*(w,x) - inf_z\{(\varphi^*)^\Diamond(z) \mid z\beta(x) = v-w\}\}, \end{aligned}$$

where the second equality follows from Lemma A.7 and the last equality from Lemma A.6. Hence, the robust counterpart (FRC) for this case becomes:

$$\begin{cases} (a^0)^T v + \delta^*(A^T v|Z) - g_*(w,x) + (\varphi^*)^\Diamond(z) \le 0 \\ z\beta(x) = v - w. \end{cases}$$

Since $g_*(w,x)$ is increasing in w, we may replace the "$=$" in the second constraint by a "$\le$". Hence, we obtain

$$\begin{cases} (a^0)^T v + \delta^*(A^T v|Z) - g_*(w,x) + (\varphi^*)^\Diamond(z) \le 0 \\ z \begin{pmatrix} 1 \\ \varphi(\alpha_2(x)) \\ \vdots \\ \varphi(\alpha_m(x)) \end{pmatrix} \le v - w. \end{cases} \tag{37}$$

For fixed z this problem is convex, hence we can solve this problem for different values of z, and search for the best value of z. However, for special choices of φ, (37) can be solved directly. As an example, take $\theta(a,x) = \sqrt{a_1 + \sum_{i=2}^m a_i(\alpha_i(x))^2}$, hence $\varphi(t) = t^2$. From Table 3 we have $\varphi^*(s) = s^2/4$, and hence $(\varphi^*)^\Diamond(s) = 1/(4s)$. The robust counterpart (37) for this case then becomes (also substitute $y = 1/z$):

$$\begin{cases} (a^0)^T v + \delta^*(A^T v|Z) - g_*(w,x) + y/4 \le 0 \\ \frac{1}{y} \begin{pmatrix} 1 \\ (\alpha_2(x))^2 \\ \vdots \\ (\alpha_m(x))^2 \end{pmatrix} \le v - w, \end{cases} \tag{38}$$

where the last inequalities can be rewritten as conic quadratic constraints.

Also note that (37) can be given explicitly, even if there is no closed form for φ^{-1}. *As an example take* $\varphi(t) = -t - \log t$.

Finally, we observe that the so-called certainty equivalent used as a risk measure (see [4]) is of the same form as $\theta(a, x)$, *in which a is interpreted as the probability vector. This means that we can derive tractable robust counterparts for constraints in which the uncertain parameters are probabilities.* $\square$

In several examples (see e.g. Section 4.2, formula (35) and Example 26) there is a need to compute $-f_*(.)$. A powerful technique to accomplish this task is based on a Conic Quadratic representation (CQr) of the original concave function f (see also Appendix B.2). We now describe this technique. Let $g(x) = -f(x)$ and suppose g (a convex function) has the following CQr of its epigraph:

$$epi(g) = \{(t, x) \mid t \geq g(x)\} = \{(t, x) \mid \exists u : tq + Qx + Ru + r \in K\},$$

for some vectors q and r, and matrices Q and R, where K is the Lorentz cone of appropriate dimension. Then (see Proposition 3.3.4, p. 98, in [6])

$$g^*(s) = \min_y \left\{ r^T y \mid q^T y = 1,\ Q^T y + s = 0,\ R^T y = 0,\ y \in K \right\}. \tag{39}$$

So, finally $-f_*(s) = g^*(-s)$. Recall that in the final (FRC), in which $-f_*(.)$ appears, the min operator in (39) is omitted and we shall get a "clean" explicit conic quadratic inequality.

4.3 Nonconcave uncertainty

In Theorem 2 we assume that $f(a, x)$ is concave in a for each x. If this assumption does not hold, one may try to reformulate inequality (6) in such a way that the assumption does hold. In this subsection we describe different ways for such a reformulation.

1. **Reparametrizing and computing convex hull.** Suppose $f(a, x)$ can be written as $f(a)^T g(x)$, where $f(a)$ is not necessarily concave and/or $g(x)$ may attain positive and negative values. Hence, we cannot apply Theorem 2. Let us for ease of notation also assume that $m = L$, $A = I$, and $a^0 = 0$. By the parametrization $b = f(a)$, (RC) becomes

   $$b^T g(x) \leq 0 \quad \forall (a, b) \in \bar{U},$$

   where $\bar{U} = \{(a, b) \mid a \in U,\ b = f(a)\}$. Since the left-hand side of this constraint is linear in the uncertain parameter b, we may replace $\bar{U}$ by $\mathrm{conv}(\bar{U})$. Hence, if we can compute $\mathrm{conv}(\bar{U})$ we can apply Theorem 2.

   We do not give an extensive treatment of how to construct $\mathrm{conv}(\bar{U})$; this is the subject of a future paper. We only give two examples. A well-known example is quadratic uncertainty and an ellipsoidal uncertainty region. For this case it has been proved that $\mathrm{conv}(\bar{U})$ can be formulated as an LMI. The resulting robust counterpart is therefore a system of LMIs. See [5].

Constraint function	$f(a,x)$	**Robust Counterpart**
linear in a	$a^T f(x)$	$(a^0)^T v + \delta^*(A^T f(x) \vert Z) \le 0$
concave in a **separable in** a **and** x	$f(a)^T x$	$\begin{cases} (a^0)^T v + \delta^*(A^T v \vert Z) - \sum_{i=1}^n x_i (f_i)_*(s^i/x_i) \le 0 \\ \sum_{i=1}^n s^i = v. \end{cases}$
example	$-\sum_i \left(a^T Q_i a\right) x_i$	$\begin{cases} (a^0)^T v + \delta^*(A^T v \vert Z) + \frac{1}{2}\sum_{i=1}^n \frac{(s^i)^T Q_i^{-1} s^i}{x_i} \le 0 \\ \sum_{i=1}^n s^i = v, \end{cases}$
example	$\sum_i h_i^2(x) a_i - \left(\sum_i h_i(x) a_i\right)^2$	$\begin{cases} (a^0)^T v + \delta^*(A^T v \vert Z) + \frac{z^2}{4} \le 0 \\ h^2(x) + z h(x) = v \end{cases}$
example	$h^{-1}(a)^T x$	$(a^0)^T v + \delta^*(A^T v \vert Z) + \sum_{i=1}^n v_i (h_i)^*(x_i/v_i) \le \beta$
concave in a **not separable**	$f(a,x)$	$(a^0)^T v + \delta^*(A^T v \vert Z) - f_*(v,x) \le 0$
example	$-\sum_{i=1}^m x_i^{a_i}, x_i > 1, 0 \le a \le 1$	$\begin{cases} (a^0)^T v + \delta^*(A^T v \vert Z) - \sum_{i=1}^m \left(\frac{v_i}{\ln x_i} \ln \frac{-v_i}{\ln x_i} - \frac{v_i}{\ln x_i}\right) \le 0 \\ v \le 0 \end{cases}$
example	$g(a,x) + \theta(a,x)$	$\begin{cases} (a^0)^T v + \delta^*(A^T v \vert Z) - g_*(w,x) + (\varphi^*)^\Diamond(z) \le 0 \\ z \begin{pmatrix} 1 \\ \varphi(\alpha_2(x)) \\ \vdots \\ \varphi(\alpha_m(x)) \end{pmatrix} \le v - w \end{cases}$
sum of functions	$\sum_i f_i(a,x)$	$\begin{cases} (a^0)^T v + \delta^*(A^T v \vert Z) - \sum_{i=1}^n (f_i)_*(s^i, x) \le 0 \\ \sum_{i=1}^n s^i = v. \end{cases}$
sum of separable functions	$\sum_i f_i(a_i,x)$	$(a^0)^T v + \delta^*(A^T v \vert Z) - \sum_{i=1}^n (f_i)_*(v_i, x) \le 0$

Table 2: Robust counterparts for different choices for the constraint function $f(a,x)$.

The second example is when $f_i(a) = f_i(a_i)$ and the uncertainty region is a box, i.e. $\{a \mid \|a\|_\infty \leq 1\}$. Since $\bar{U}$ is completely separable with respect to i, we basically have to compute

$$\tilde{U}_i = \mathrm{conv}((a_i, b_i) \mid -1 \leq a_i \leq 1,\ b_i = f_i(a_i)).$$

It is easy to construct an explicit description of this bivariate convex hull. Suppose, for example, that $f_i(a_i)$ is convex. Then, it can easily be verified that

$$\tilde{U}_i = \{(a_i, b_i) \mid -1 \leq a_i \leq 1,\ b_i \geq f_i(a_i),\ 2b_i \leq f_i(1)(a_i+1) - f_i(-1)(a_i-1)\}, \quad \forall i.$$

Hence, (RC) becomes

$$b^T g(x) \leq 0 \quad \forall (a_i, b_i) \in \mathrm{conv}(\bar{U}_i)\ \forall i,$$

for which now (32) can be used. Since we have $\delta^*(y|\tilde{U}_1 \times ... \times \tilde{U}_m) = \sum_{i=1}^m \delta^*(y^i|\tilde{U}_i)$, see Corollary A.2 we have to derive $\delta^*(y^i|\tilde{U}_i))$. By defining

$$\tilde{U}_{i1} = \{(a_i, b_i) \mid |a_i| \leq 1\}$$

$$\tilde{U}_{i2} = \{(a_i, b_i) \mid f_i(a_i) - b_i \leq 0\},$$

and

$$\tilde{U}_{i3} = \{(a_i, b_i) \mid a_i(f_i(-1) - f_i(1)) + 2b_i - f_i(1) - f_i(-1) \leq 0\},$$

we have $\tilde{U}_i = \tilde{U}_{i1} \cap \tilde{U}_{i2} \cap \tilde{U}_{i3}$. Using Lemma A.4 we have

$$\delta^*(y^i|\tilde{U}_i)) = \min_{y^{i1}, y^{i2}, y^{i3}} \{\delta^*(y^{i1}|\tilde{U}_{i1}) + \delta^*(y^{i2}|\tilde{U}_{i2}) + \delta^*(y^{i3}|\tilde{U}_{i3}) \mid y^{i1} + y^{i2} + y^{i3} = y^i\}.$$

Using the result of Example 4 we obtain

$$\begin{aligned}
\delta^*(y^{i1}|\tilde{U}_{i1}) &= \max_{a_i, b_i} \{y_1^{i1} a_i + y_2^{i1} b_i \mid |a_i| \leq 1\} \\
&= \begin{cases} |y_1^{i1}| & \text{if } y_2^{i1} = 0 \\ \infty & \text{else.} \end{cases}
\end{aligned}$$

Using Lemma 9 we obtain

$$\begin{aligned}
\delta^*(y^{i2}|\tilde{U}_{i2}) &= \max_{a_i, b_i} \{y_1^{i2} a_i + y_2^{i2} b_i \mid f_i(a_i) - b_i \leq 0\} \\
&= \min_{u_i \geq 0} \{u_i f_i^* \left(\frac{y_1^{i2}}{u_i}\right) \mid y_2^{i2} = -u_i\}.
\end{aligned}$$

Using the result of Example 5 we obtain

$$\begin{aligned}
\delta^*(y^{i3}|\tilde{U}_{i3}) &= \max_{a_i, b_i} \{y_1^{i3} a_i + y_2^{i3} b_i \mid \alpha_i^T \begin{pmatrix} a_i \\ b_i \end{pmatrix} \leq \gamma_i\} \\
&= \min_{z_i \geq 0} \{\gamma_i z_i \mid z_i \alpha_i = y^{i3}\},
\end{aligned}$$

where

$$\alpha_i = \begin{pmatrix} f_i(-1) - f_i(1) \\ 2 \end{pmatrix},$$

and $\gamma_i = f_i(1)+f_i(-1)$. Substituting all these expressions into (32) we obtain the following equivalent system of inequalities for (FRC):

$$\begin{cases} (a_0)^T x + \sum_i |y_1^{i1}| + \sum_i u_i f_i^* \left(\frac{y_1^{i2}}{u_i}\right) + \frac{1}{2}\sum_i \gamma_i u_i \leq 0 \\ y_1^{i1} + y_1^{i2} + \frac{1}{2}u_i(f_i(-1) - f_i(1)) = v_i \\ u \geq 0. \end{cases}$$

Explicit examples for such f_i and their conjugates can be found in Table 3.

Above we assumed $f_i(a_i)$ to be convex. However, it can easily be verified that if $f_i(a_i)$ is concave then

$$\tilde{U}_i = \{(a_i, b_i) \mid -1 \leq a_i \leq 1,\ b_i \leq f_i(a_i),\ 2b_i \geq f_i(1)(a_i+1) - f_i(-1)(a_i-1)\}, \quad \forall i,$$

and now the same approach as above can be followed to derive the robust counterpart.

2. **Reparametrizing and using "hidden concavity" results.** For special types of constraints we can use the "hidden concavity" result of [9]. Suppose (RC) is of the following form

$$a^T g_1(x) + f(a)^T g_2(x) \leq 0 \quad \forall a \ : \ \alpha_k^T a + \beta_k^T f(a) \leq \gamma_k,\ k = 1, ..., K. \tag{40}$$

After reparametrizing $b = f(a)$ this robust counterpart problem becomes:

$$a^T g_1(x) + b^T g_2(x) \leq 0 \quad \forall (a,b) \ : \ \alpha_k^T a + \beta_k^T b \leq \gamma_k,\ k = 1, ..., K,\ b = f(a).$$

Note that the difficulty is now in the nonlinear equality constraint in the definition of the uncertainty region. We therefore replace $b = f(a)$ in the definition of the uncertainty region by $b \geq f(a)$:

$$a^T g_1(x) + b^T g_2(x) \leq 0 \quad \forall (a,b) \ : \ \alpha_k^T a + \beta_k^T b \leq \gamma_k,\ k = 1, ..., K,\ b \geq f(a), \tag{41}$$

and under certain assumptions it can be proved that (41) is equivalent to (40). More precisely, we assume that f_i is convex $\forall i$, and the constraint functions that define the uncertainty are *termwise parallel*, which is defined as follows.

Definition 29 *The inequalities:* $\alpha_k^T a + \beta_k^T f(a) \leq \gamma_k$ $(k = 1, ..., K)$ *are* termwise parallel *if*

$$\text{rank}\{(\alpha_{ki}, \beta_{ki}) : k = 1, \ldots, K\} \leq 1 \ \text{ for all } i = 1, \ldots, m, \tag{42}$$

where $\alpha_k = (\alpha_{ki})_{i=1}^m,\ \beta_k = (\beta_{ki})_{i=1}^m$. □

Now, define

$$\bar{a} = \text{argmax}\{a^T g_1(x) + b^T g_2(x) \mid \alpha_k^T a + \beta_k^T b \leq \gamma_k,\ k = 1, ..., K,\ b \geq f(a)\}.$$

It has been proved in [9] that if $b > f(\bar{a})$, then there is another optimal solution for a that satisfies $b = f(a)$. Hence, the RC (40) is equivalent to (41), for which Theorem 2 can be applied. The final robust counterpart is:

$$\begin{cases} \sum_{i=1}^m (g_{2i}(x) + \sum_{k=1}^K \beta_{ki} v_k) f_i^* \left(-\frac{g_{1i}(x) + \sum_{k=1}^K \alpha_{ki} v_k}{g_{2i}(x) + \sum_{k=1}^K \beta_{ki} v_k}\right) + \sum_{k=1}^K \gamma_k v_k \leq 0 \\ g_{2i}(x) + \sum_{k=1}^K \beta_{ki} v_k \geq 0 \ (i = 1, \ldots, m) \\ v \geq 0. \end{cases} \tag{43}$$

3. **Miscellaneous.** It might be the case that $f(a,x)$ is convex in a for each x. It is well-known that a convex function on a polyhedral region attains its maximum in one of the vertices of the feasible region. Hence, if the uncertainty region is e.g. a box (i.e. interval uncertainty) then we might obtain a tractable reformulation by enumerating all possible vertices. Let us illustrate this by the following examples.

   Example 30 Implementation error and separable constraint. *Consider the following robust constraint:*

$$\alpha^T(x+a) + \sum_{j=1}^{n} f_j(x_j + a_j) \leq 0, \quad \forall a \in U, \tag{44}$$

   where f_j is convex, and $U = \{a \mid \|a\|_\infty \leq \rho\}$. Note that the uncertain parameter a can be considered as additive implementation error. Since $b^T g(x)$ is linear in the uncertain parameter b, we have

$$\max_{a\in U} \sum_{j=1}^{n} \{\alpha_j(x_j+a_j) + f_j(x_j+a_j)\} = \sum_{j=1}^{n} \max\left[\alpha_j(x_j-\rho) + f_j(x_j-\rho), \alpha_j(x_j+\rho) + f_j(x_j+\rho)\right].$$

   Hence, (44) is equivalent to

$$\begin{cases} \alpha^T x + \sum_{j=1}^{n} y_j \leq 0 \\ y_j \geq -\alpha_j\rho + f_j(x_j-\rho) \quad \forall j \\ y_j \geq \alpha_j\rho + f_j(x_j+\rho) \quad \forall j. \end{cases} \tag{45}$$

   An other example that can be reformulated such that the constraint is concave in a is the following robust objective function:

$$\min_x \max_{a\in U} \frac{f(a,x)}{\alpha(a)},$$

   where $f(a,x)$ is concave in a for each x and $\alpha(a) > 0$ is a linear function in a. This objective can be reformulated as

$$\min_{x,y} \{y \mid f(a,x) - \alpha(a)y \leq 0, \quad \forall a \in U\},$$

   in which the left-hand side is concave in a for each x and y. A similar reformulation can be carried out when $\alpha(a)$ is convex and $y \geq 0$.

5 Complexity results

In this section we analyze the complexity of the resulting robust counterparts (12) studied in the previous sections. A key notion in complexity theory for convex optimization problems is *self-concordance*. In [20] it is shown that if the optimization problem admits a self-concordant barrier, then there are interior point methods that yields an ϵ-optimal solution in $O\left(\sqrt{n}|\ln\epsilon|\right)$ Newton steps, where n is the number of variables. See also [17]. In this section we show that the logarithmic barrier function for most of the optimization problems studied in the previous sections are self-concordant, thereby proving that there are interior point methods that solve such problems in polynomial time.

We first start with the formal definition of the logarithmic barrier function and self-concordance.

Definition 31 *The logarithmic barrier function for the set of constraints*

$$\{g_i(z) \leq 0, \ i \in I\}$$

in the variables $z \in \mathbb{R}^n$, is defined as

$$\varphi(z) = -\sum_{i \in I} \ln(-g_i(z)).$$

$\Box$

Definition 32 *Let $F \subset \mathbb{R}^n$ be an open and convex set and κ a nonnegative number. A function $\varphi : F \to \mathbb{R}$ is called κ-*self-concordant on F *if φ is convex, belongs to $C^3(F)$, and satisfies:*

$$|\nabla^3 \varphi(y)[h,h,h]| \leq 2\kappa(h^T \nabla^2 \varphi(y) h)^{\frac{3}{2}}, \quad \forall y \in F \ \forall h \in \mathbb{R}^n,$$

where $\nabla^3 \varphi(y)[h,h,h]$ denotes the third order differential of φ at y and h. We call φ self-concordant if it is κ-self-concordant for some κ. $\Box$

We now study the complexity of (FRC) in (12). We rewrite (FRC) as follows:

$$\begin{cases} (a^0)^T v + t_1 + t_2 \leq 0 \\ \delta^*(y|Z) \leq t_1 \\ -f_*(v,x) \leq t_2 \\ y = A^T v. \end{cases} \tag{46}$$

The *first* constraint of (46) is a linear inequality for which we know that the corresponding logarithmic barrier is self-concordant. We now consider the *second* constraint of (46). When the uncertainty region is polyhedral, ellipsoidal, or norm-based, then it can easily be checked that the logarithmic barrier function for this constraint is self-concordant. If the uncertainty region is e.g. defined by separable functions (see (26))

$$h_k(\zeta) = \sum_{l=1}^{L_k} h_{kl}(\zeta_l) \leq 0, \ k = 1, ..., K,$$

then we have to check whether the logarithmic barrier function for the constraint

$$\sum_{k=1}^{K} \sum_{l=1}^{L_k} u_k h_{kl}^* \left(\frac{w_{kl}}{u_k} \right) \leq t_1 \tag{47}$$

is self-concordant. This inequality can be reformulated as

$$\begin{cases} \sum_{k=1}^{K} \sum_{l=1}^{L_k} q_{kl} = t_1 \\ u_k h_{kl}^* \left(\frac{w_{kl}}{u_k} \right) \leq q_{kl}, \quad k = 1, ..., K, \ l = 1, ..., L_k, \end{cases} \tag{48}$$

which is easier to analyze with respect to self-concordance by using the following lemma, which is taken from [8]:

Lemma 33 *[8] Assume that a convex function $f \in C^3(\mathbb{R}^+)$ satisfies*

$$|f'''(s)| \leq \kappa f''(s)/s \quad \text{for some} \quad \kappa > 0. \tag{49}$$

Then the logarithmic barrier function for the constraints

$$\{rf(s/r) \leq t, \ r \geq 0, \ s \geq 0\},$$

i.e., $-\ln(t - rf(s/r)) - \ln r - \ln s$, is $(2 + \frac{\sqrt{2}}{3})\kappa$-self-concordant. $\Box$

$f(t)$	$f^*(s)$ (domain)
t	$0 \ (s=1)$
t^2	$s^2/4 \ (s \in \mathbb{R})$
$\lvert t\rvert^p/p \ (p>1)$	$\lvert s\rvert^q/q \ (s \in \mathbb{R})$
$-t^p/p \ (t \geq 0, 0<p<1)$	$-(-s)^q/q \ (s \leq 0)$
$-\log t \ (t>0)$	$-1-\log(-s) \ (s<0)$
$t \log t \ (t>0)$	$e^{s-1} \ (s \in \mathbb{R})$
e^t	$\begin{cases} s\log s - s \ (s>0) \\ 0 \ (s=0) \end{cases}$
$\log(1+e^t)$	$\begin{cases} s\log s + (1-s)\log(1-s) \ (0<s<1) \\ 0 \ (s=0,1) \end{cases}$
$\sqrt{1+t^2}$	$-\sqrt{1-s^2} \ (-1 \leq s \leq 1)$

Table 3: Some examples for f, with conjugate f^*. The parameters p and q are related as follows: $1/p + 1/q = 1$.

Moreover, the following lemma is an easy consequence of Lemma 33.

Lemma 34 *[8] Assume that convex functions $f_i \in C^3(\mathbb{R}^+)$ satisfies*

$$|f_i'''(s)| \leq \kappa_i f_i''(s)/s \quad \textit{for some} \quad \kappa_i > 0. \tag{50}$$

Let $f(s) = \sum_i f_i(s)$. Then the logarithmic barrier function for the constraints

$$\{rf(s/r) \leq t, \ r \geq 0, \ s \geq 0\},$$

i.e., $-\ln(t - rf(s/r)) - \ln r - \ln s$, is $(2 + \frac{\sqrt{2}}{3}) \max_i \kappa_i$-self-concordant. □

Hence, if all h^*_{kl} in (48) satisfy condition (49), then the logarithmic barrier function for (48) is self-concordant. In Table 3 several examples are given that satisfy condition (49). Two examples in this table do not directly satisfy this assumption, but we show that using some reformulations leads to self-concordant logarithmic barriers. The first one is: $h(t) = t\log t$. Then $uh^*(w/u) = ue^{-1+w/u} \leq q$ can be rewritten as $-u \log q/u \leq u - w$, and since $-\log t$ satisfies condition (49), we conclude from Lemma 33 that the corresponding logarithmic barrier function for this constraint is self-concordant. The second one is: $h(t) = \sqrt{1+t^2}$. It can easilby be verified that $uh^*(w/u) \leq q$ can be rewritten as $\sqrt{w^2+q^2} \leq u$, which is a conic quadratic constraint.

The conclusion is that if the functions h_{kl} are such that h^*_{kl} satisfies (49) (see the examples in Table 3), then the *second* constraint in (46) allows a self-concordant logarithmic barrier.

Concerning the *third* constraint in (46), it can easily be checked that for f as given in Examples 22–25 the logarithmic barrier function for this constraint is self-concordant. This is also the case for the separable case (26) where f_i are chosen such that (49) holds. Again, see Table 3 for such examples.

Finally, the *fourth* constraint in (46) is a linear equality constraint, that only restricts the domain of the logarithmic barrier function.

Acknowledgements. We would like to thank Bram Gorissen (Tilburg University) for his critical reading of the paper.

A Conjugate functions, support functions and Fenchel duality

In this section we give some basic results on conjugate functions, support functions and Fenchel duality. For a detailed treatment we refer to [21].

We start with some well-known results on conjugate functions. First, note that f^* is closed convex, and g_* is closed concave, and moreover $f^{**} = f$ and $g_{**} = g$. It is well-known that for $a > 0$

$$(af)^*(y) = af^*\left(\frac{y}{a}\right) \quad \text{and} \quad (ag)_*(y) = ag_*\left(\frac{y}{a}\right), \tag{51}$$

and for $\tilde{f}(x) = f(ax)$ and $\tilde{g}(x) = g(ax)$, $a > 0$, we have

$$\tilde{f}^*(y) = f^*\left(\frac{y}{a}\right) \quad \text{and} \quad \tilde{g}_*(y) = g_*\left(\frac{y}{a}\right),$$

and for $\tilde{f}(x) = f(x - a)$ and $\tilde{g}(x) = g(x - a)$ we have

$$\tilde{f}^*(y) = f^*(y) + ay \quad \text{and} \quad \tilde{g}_*(y) = g_*(y) + ay.$$

In this paper wel also frequently use the following sum-rules for conjugate functions.

Lemma A.1 *Assume that f_i, $i = 1, ..., m$, are convex, and the intersection of the relative interiors of the domains of f_i, $i = 1, ..., m$, is nonempty, i.e., $\cap_{i=1}^m ri(dom\, f_i) \neq \emptyset$. Then*

$$\left(\sum_{i=1}^m f_i\right)^*(s) = \inf_{\{v^i\}_{i=1}^m} \left\{ \sum_{i=1}^m f_i^*(v^i) \mid \sum_{i=1}^m v^i = s \right\},$$

and the inf *is attained for some v_i, $i = 1, ..., m$.* □

Corollary A.1 *Assume that f_i, $i = 1, ..., m$, are convex, and separable, i.e. $f_i(x) = f_i(x_i)$. Then*

$$\left(\sum_{i=1}^m f_i\right)^*(s) = \sum_{i=1}^m f_i^*(s_i).$$

□

Lemma A.2 *Assume that g_i, $i = 1, ..., m$, are concave, and the intersection of the relative interiors of the domains of g_i, $i = 1, ..., m$, is nonempty, i.e., $\cap_{i=1}^m ri(dom\, g_i) \neq \emptyset$. Then*

$$\left(\sum_{i=1}^m g_i\right)_*(s) = \sup_{\{v^i\}_{i=1}^m} \left\{ \sum_{i=1}^m (g_i)_*(v^i) \mid \sum_{i=1}^m v^i = s \right\},$$

and the sup *is attained for some v^i, $i = 1, ..., m$.* □

The following useful lemma states that the support function of the Minkowski sum of sets is the sum of the corresponding support functions.

Lemma A.3

$$\delta^*(y|S_1 + S_2 + ...S_k) = \sum_{i=1}^k \delta^*(y|S_i).$$

Proof: The proof easily follows by using the definition of the support function:

$$\begin{aligned}\delta^*(y|S_1+S_2+...+S_k) &= \sup_{x^1\in S_1,...,x^k\in S_k} y^T(x^1+...+x^k)\\ &= \sup_{x^1\in S_1} y^T x^1 + ... + \sup_{x^k\in S_k} y^T x^k\\ &= \sum_{i=1}^{k}\delta^*(y|S_i).\end{aligned}$$

□

The following lemma is a result on the support function for the intersection of several sets.

Lemma A.4 *Let $S_1, ..., S_k$ be closed convex sets, such that $\bigcap_i ri(S_i) \neq \emptyset$, and let $S = \cap_{i=1}^k S_i$. Then*

$$\delta^*(y|S) = \min\left\{\sum_{i=1}^{k}\delta^*(v^i|S_i) \mid \sum_{i=1}^{k} v^i = y\right\}.$$

□

Corollary A.2 *Let $y^{[1]}, y^{[2]}, ..., y^{[k]}$ be a partition of the variables $(y^1, y^2, ..., y^n)$ into k mutually exclusive subvectors. Let $S_1, ..., S_k$ be closed convex sets, and let $S = S_1 \times ... \times S_k$. Then*

$$\delta^*(y|S) = \sum_{i=1}^{k}\delta^*(y^{[i]}|S_i).$$

□

We now state three results which are used in this paper to derive tractable robust counterparts. The first lemma relates the conjugate of the adjoint function (see (4)) to the conjugate of the original function. Note that $f^\diamond(x)$ is convex if $f(x)$ is convex. The next proposition can be used in cases where f^* is not available in closed form, but $(f^\diamond)^*$ is available as such.

Lemma A.5 *[16] For the conjugate of a function $f : \mathbb{R}_+ \longrightarrow \mathbb{R}$ and the conjugate of its adjoint $f^\diamond$, we have*

$$f^*(s) = \inf\{y \in \mathbb{R} : (f^\diamond)^*(-y) \leq -s\}.$$

□

The next proposition can be used in cases where f^{-1} is not available in closed form, but $(f^{-1})^*$ is available as such.

Lemma A.6 *[4] Let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be strictly increasing and concave. Then, for all $y > 0$*

$$\left(f^{-1}\right)^*(y) = -yf_*\left(\frac{1}{y}\right) = -(f_*)^\diamond(y).$$

□

The next proposition gives a usefull result related to the conjugate of a function after linear transformations.

Lemma A.7 *Let A be a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$. Assume there exists an x such that $Ax \in ri(dom\ g)$. Then, for each convex function g on $\mathbb{R}^m$, one has*

$$(gA)^*(z) = \inf_y \{g^*(y) | A^T y = z\},$$

where for each z the infimum is attained, and where the function gA is defined by

$$(gA)(x) = g(Ax).$$

□

We define the primal problem:

$$(P) \qquad \inf\{f(x) - g(x) \mid x \in \text{dom}(f) \cap \text{dom}(g)\}.$$

The Fenchel dual of (P) is given by:

$$(D) \qquad \sup\{g_*(y) - f^*(y) \mid y \in \text{dom}(g_*) \cap \text{dom}(f^*)\}.$$

Now we can give the well-known Fenchel duality theorem.

Theorem A.1 *If $ri(dom(f)) \cap ri(dom(g)) \neq \emptyset$ then the optimal values of (P) and (D) are equal and the maximal value of (D) is attained.*
If $ri(dom(g_*)) \cap ri(dom(f^*)) \neq \emptyset$ then the optimal values of (P) and (D) are equal and the minimal value of (P) is attained. □

Note that since $f^{**} = f$ and $g_{**} = g$, we have that the dual of (D) is (P).

B Conic quadratic optimization

B.1 Conic quadratic duality

Consider the following primal conic quadratic optimization problem:

$$(P) \qquad \min_x \{c^T x \mid Rx = r, \|D_i x - d_i\|_2 \leq p_i^T x - q_i, i = 1, ..., K\}.$$

This can be rewritten as

$$(P1) \qquad \min_x \{c^T x \mid Rx = r, B_i x - b_i \in L^{m_i}, i = 1, ..., K\},$$

where

$$B_i = \begin{bmatrix} D_i \\ p_i^T \end{bmatrix}, \quad b_i = \begin{bmatrix} d_i \\ q_i \end{bmatrix},$$

and L^{m_i} is the Lorentz cone of order m_i.

The dual problem of (P) is given by

$$(D) \qquad \begin{cases} \max_{v,y,z} \left\{ r^T v + \sum_{i=1}^K \left(d_i^T z_i + q_i y_i \right) \right\} \\ R^T v + \sum_{i=1}^K (D_i^T z_i + y_i p_i) = c \\ \|z_i\|_2 \leq y_i, \quad i = 1, ..., K. \end{cases}$$

The dual problem of $(P1)$ is given by

$$(D1) \qquad \begin{cases} \max_{v,\eta} \left\{ r^T v + \sum_{i=1}^K b_i^T \eta_i \right\} \\ R^T v + \sum_{i=1}^K B_i^T \eta_i = c \\ \eta_i \in L^{m_i}, \quad i = 1, ..., K. \end{cases}$$

The following theorem states the well-known duality for Conic Quadratic Programming, but first we need the following definition.

Definition B.1 *(P) is regular if $\exists \hat{x} : R\hat{x} = r, \|D_i\hat{x} - d_i\|_2 < p_i^T \hat{x} - q_i, \forall i = 1, ..., K$.* □

Theorem B.1 *(Strong duality) If one of the problems (P) or (D) is regular and bounded, then the other problem is solvable and the optimal values of (P) and (D) are equal. If both (P) and (D) are regular then both problems are solvable and the optimal values of (P) and (D) are equal.* □

B.2 Conic Quadratic representation

We start with the definition of Conic Quadratic representable.

Definition B.2 *A set $X \subset \mathbb{R}^n$ is Conic Quadratic representable (CQr) if there exist:*

- *a vector $u \in \mathbb{R}^l$ of additional variables*
- *an affine mapping:*

$$H(x, u) = \begin{bmatrix} H_1(x, u) \\ H_2(x, u) \\ \vdots \\ H_K(x, u) \end{bmatrix} \; : \; \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}^{m_1} \times ... \times \mathbb{R}^{m_K},$$

such that

$$X = \{x \in \mathbb{R}^n \mid \exists u \in \mathbb{R}^l \; : \; H_j(x, u) \in K^{m_j}, \quad j = 1, ..., K\},$$

where K^{m_j} is the second-order cone of order m_j.

The collection $(l, K, H(.,.), m_1, ..., m_K)$ is called a Conic Quadratic Representation (CQR) of X. □

Given a CQR $(l, K, H(.,.), m_1, ..., m_K)$ of X, the problem

$$\min\{c^T x \mid x \in X\}$$

can be posed as a Conic Quadratic Problem (CQP):

$$\min_{x,u}\{c^T x \mid H_j(x, u) \in K^{m_j}, \; j = 1, ..., K\}.$$

The following definition extends CQr to functions.

Definition B.3 *A function $f : \mathbb{R}^n \to \mathbb{R}$ is CQr if its epigraph*

$$epi(f) = \{(t, x) \in \mathbb{R} \times \mathbb{R}^n \mid f(x) \leq t\}$$

is a CQr set.

B.3 Operations preserving CQr

We now state some operations that preserve CQr of sets (for proofs and a full list we refer to [6]):

- If X_i is CQr $\forall i = 1, ..., N$, then $\cap X_i$ and $X_1 \times ... \times X_N$ and $X_1 + X_2 + ... + X_N$ are CQr;
- If X is CQr, then the set $\{Bx + b \mid x \in X\}$ is CQr;
- If X is CQr, then the set $\{y \mid By + b \in X\}$ is CQr.

We now state some operations that preserve CQr of functions (for proofs and a full list we refer to [6]):

- If $f_i(x)$ is CQr $\forall i = 1, ..., N$, then $\max_i f_i(x)$ and $\sum_i \alpha_i f_i(x)$, $\alpha_i \geq 0$, are CQr.
- If $f_i(x)$ is CQr, then $f(Bx + b)$ is CQr.
- If $f(x)$ is CQr, then the conjugate function $f^*(s)$ is CQr.
- If $f(x)$ is CQr, then the perspective function $f^{per}(x, v) = vf(x/v)$, $v > 0$, is CQr.

Since the last result is new, we give a proof. Note that

$$\text{epi}(f^{per}(x, v)) = \{(t, x, v) \mid vf(x/v) \leq t\} = \{(t, x, v) \mid (x/v, t/v) \in \text{epi}(f)\}.$$

Let $(l, K, H(x, u), m_1, ..., m_K)$ be the CQR of epi(f). Since $H(x, u)$ is an affine mapping, also $H^{per}(x, v, u) := vH(x/v, u/v)$ is an affine mapping. Moreover, since $H(x, u) \in K$, we have $H_j^{per}(x, v, u) \in K^{m_j}$. Hence, we have epi$(f^{per})$ is CQr, and $(l, K, H^{per}(x, v, u), m_1, ..., m_K)$ is its CQR. $\square$

Note that once we have a CQR of f, we immediately have the CQR of f^* and f^{per}.

It can easily be verified that the following functions/sets are CQr:

- $f(x) = x^T Q^T Q x + q^T x + r$;
- $X_m = \{(t, x_1, ..., x_M) \mid t^M \leq x_1 ... x_M\} = \text{epi}(x_1 x_2 ... x_m)^{1/M}$, where $m > 0$ is an integer, and $M = 2^m$;
- $f(x) = \max(x, 0)^\pi$, where $\pi \geq 1$ is rational;
- $f(x) = |x|^\pi$, where $\pi \geq 1$ is rational;
- $f(x) = x_1^{-\pi_1} x_2^{-\pi_2} ... x_m^{-\pi_m}$, $x_i > 0$, where $\pi_i > 0$;
- $f(x) = \|x\|_p$, where $1 \leq p \leq \infty$ is rational.

B.4 Example

Consider the set

$$Z = \{y \in \mathbb{R}^n \mid By + b \in \cap_{i=1}^K Z_i\},$$

where $B \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^n$, and

$$Z_i = \{z \mid (d_i^T z + \beta_i)^2 + 2(d_i^T z + \beta_i)^8 \leq 1\},$$

where $d_i \in \mathbb{R}^n$, $\beta_i \in \mathbb{R}$. Note that

$$Z = \cap_{i=1}^K \{By + b \in Z_i\},$$

with

$$Z_i = \{z \mid d_i^T z + \beta_i \in X\},$$

and

$$X = \{x \in \mathbb{R} \mid x^2 + 2x^8 \leq 1\}.$$

To compute the support function of Z it is enough to compute the CQR of X. Observe that

$$X = \left\{ x \in \mathbb{R} \mid \exists t_1, t_2, t_3 \ : \ \begin{pmatrix} 2x \\ t_1 - 1 \\ t_2 + 1 \end{pmatrix} \in L^3, \ \begin{pmatrix} 2t_1 \\ t_2 - 1 \\ t_2 + 1 \end{pmatrix} \in L^3, \ \begin{pmatrix} 2t_2 \\ t_3 - 1 \\ t_3 + 1 \end{pmatrix} \in L^3, \ t_1 + 2t_3 \leq 1 \right\}.$$

We now can use the rules of *intersection* of CQr sets and *linear transformations* of CQr sets to write down the CQR of Z explicitly. Let $\delta^*(v|Z) = \sup_{y\in Z} v^T y$. Using the CQR of Z this optimization problem is a conic quadratic problem. Its dual can be written down explicitly using conic quadratic duality.

References

[1] AIMMS. Paragon Decision Technology, http://www.aimms.com/operations-research/mathematical-programming/robust-optimization, 2012.

[2] T.W. Anderson, D.A. Darling. Asymptotic theory of certain goodness-of-fit criteria based on stochastic processes. *Annals of Mathematical Statistics* 23(2), 193–212, 1952.

[3] I. Averbakh, Y.-B. Zhao. Explicit reformulations for robust optimization problems with general uncertainty sets. *SIAM Journal on Optimization* 18(4), 1436–1466, 2007.

[4] A. Ben-Tal, A. Ben-Israel, and M. Teboulle. Certainty equivalents and information measures: duality and extremal principles. *Journal of Mathematical Analysis and Applications* 157, 211–236, 1991.

[5] A. Ben-Tal, L. El-Ghaoui, A. Nemirovski. *Robust Optimization.* Princeton Series in Applied Mathematics, 2009.

[6] A. Ben-Tal, A. Nemirovski. *Lectures on Modern Convex Optimization - Analysis, Algorithms, and Engineering Applications.* MPS-SIAM Series on Optimization, 2001.

[7] A. Ben-Tal, D. den Hertog. Immunizing conic quadratic optimization problems against implementation errors. CentER Discussion Paper 2011-060, Tilburg University, 2011. To appear in Mathematical Programming.

[8] A. Ben-Tal, D. den Hertog, A.M.B. De Waegenaere, B. Melenberg, G. Rennen. Robust solutions of optimization problems affected by uncertain probabilities. CentER Discussion Paper 2011-061, Tilburg University, 2011. To appear in *Management Science*.

[9] A. Ben-Tal, D. den Hertog, M. Laurent. Hidden convexity in partially separable optimization. CentER Discussion Paper 2011-070, Tilburg University, 2011.

[10] D. Bertsimas, D. Brown, and C. Caramanis. Theory and applications of robust optimization. *SIAM Review* 53(3), 464–501, 2011.

[11] D. Bertsimas, M. Sim. The price of robustness. *Operations Research* 52, 35–53, 2004.

[12] D. Bertsimas, D. Pachamanova, M. Sim. Robust linear optimization under general norms. *Operations Research Letters* 32, 510–516, 2004.

[13] S. Boyd, L. Vandenberghe. *Convex Optimization.* Cambridge University Press, Cambridge, 2004.

[14] J. Goh, M. Sim. Robust optimization made easy with ROME. *Operations Research* 59(4), 973–985, 2011.

[15] G.H. Golub, C.F. van Loan. *Matrix Computations* (2nd edition). Johns Hopkins University Press, London, UK. 1989.

[16] A.A. Gushchin. On an extension of the notion of f-divergence. *Theory of Probability and its Applications* 52(3), 439–455, 2008.

[17] D. den Hertog. *Interior Point Approach to Linear, Quadratic and Convex Programming.* Kluwer Academic Publishers, Dordrecht, 1994.

[18] J. Löfberg. Automatic robust convex programming. *Optimization Methods and Software*, 27(1), 115–129, 2012.

[19] A. Nemirovski. Lectures on robust convex optimization. *Lecture 2 of Lecture Notes. http://www2.isye.gatech.edu/ nemirovs/Lect2.pdf*, 2009.

[20] Y.E. Nesterov and A. Nemirovski. *Interior Point Polynomial Algorithms in Convex Programming.* SIAM Studies in Applied Mathematics, Philadelphia, 1993.

[21] R.T. Rockafellar. *Convex Analysis.* Princeton Press, Princeton, 1970.