13 Robust conic programs

The theory of robust conic programming is presented in [1]. We consider a conic program in the form

$$\min_x \langle c, x\rangle : Ax + b \in K,$$

where $K \subset R^N$ is a regular convex cone, and $x \in \mathbb{R}^n$ is the vector of decision variables.

We assume that the *data* A, b of the problem is *uncertain* and varies in an uncertainty set U around a *nominal* data set (A^0, b^0). Let x^* be the nominal optimal solution of the problem for this data set. If the real data is perturbed, $A' = A^0 + \delta A$, $b' = b^0 + \delta b$, then the conic constraint might be violated, i.e., we might have that $A'x^* + b' \notin K$.

Our goal is to safeguard against this situation by choosing a sub-optimal, but *robust* solution. Instead of the nominal problem we shall solve its *robust counterpart* (RC)

$$\min_x \langle c, x\rangle \quad : \quad Ax + b \in K \quad \forall\ (A, b) \in U$$

In other words, we restrict the feasible set of the problem such that its elements satisfy the constraint for *all* realizations of the uncertainty.

The complexity of the resulting robust conic program depends both on K and U. We suppose that the uncertainty set U is parameterized affinely and its elements are given by

$$(A, b) = (A^0, b^0) + \sum_{k=1}^{m-1} u_k \cdot (A^k, b^k), \quad u \in B.$$

Here (A^0, b^0) is the centre of the uncertainty set, (A^k, b^k) the directions of the uncertainty, and $B \subset \mathbb{R}^{m-1}$ is a compact convex set which determines the shape of U.

Example: Finite number of scenarios. In this case B is a polytope with a small number of vertices. Each vertex yields an extreme point of the set U and represents a *scenario*. The robust counterpart of the problem then optimizes the worst case over all scenarios. Let (A_j, b_j), $j = 1, \dots, M$ be the vertices of U, then the RC can be written as

$$\min_x \langle c, x\rangle \quad : \quad A_j x + b_j \in K \quad \forall\ j = 1, \dots, M.$$

Hence the RC is an ordinary conic program with M conic constraints over the same cone K. Equivalently, there is one conic constraint over the direct product K^M.

Note that this reduction of the RC was possible because the expression $Ax + b$ is affine in A, b and K is convex. Hence if $A_j x + b_j \in K$ for all j, then also $Ax + b \in K$ for every convex combination (A, b) of the extreme points (A_j, b_j), for every fixed x.

For a general uncertainty set the RC can also be rewritten as a conic program, but not over the original cone K or its powers. Define the cone

$$K_B = \{(\tau; \tau u) \in \mathbb{R}^m \mid \tau \geq 0,\ u \in B\}.$$

In other words, K_B is the homogenization of the compact convex set B. Then the robust counterpart can be written

$$\min_x \langle c, x\rangle : \left(\sum_{k=0}^{m-1} u_k A^k\right) x + \sum_{k=0}^{m-1} u_k b^k \in K \quad \forall\ u \in K_B$$

by the homogeneity of the cone K. Equivalently we obtain

$$\min_x \langle c, x\rangle : \ \mathcal{A}_x[K_B] \subset K,$$

where the linear map $\mathcal{A}_x : \mathbb{R}^m \to \mathbb{R}^N$ is given by

$$\mathcal{A}_x(u) = \left(\sum_{k=0}^{m-1} u_k A^k\right) x + \sum_{k=0}^{m-1} u_k b^k.$$

Note that the coefficients of the linear map $\mathcal{A}_x$ are affine in x and can be arranged in a real $m \times N$ matrix. We shall then consider the inclusion $\mathcal{A}_x[K_B] \subset K$ in the formulation of the RC as a *conic constraint* in this matrix space $\mathbb{R}^{m\times N}$. This can be formalized by the following definition.

Definition 13.1. Let $K_1 \subset \mathbb{R}^{n_1}$, $K_2 \subset \mathbb{R}^{n_2}$ be regular convex cones. Call a linear map $\mathcal{A} : \mathbb{R}^{n_1} \to \mathbb{R}^{n_2}$ *K_1-to-K_2 positive* if $\mathcal{A}[K_1] \subset K_2$.

The cone of K_1-to-K_2 positive maps is itself a regular convex cone in $\mathbb{R}^{n_1\times n_2}$, the K_1-to-K_2 positive cone.

Note also that the $(K_1 \times \cdots \times K_m)$-to-$(K'_1 \times \cdots \times K'_{m'})$ positive cone, defined by cones being products of smaller cones, is itself the product $\prod_{k=1}^{m}\prod_{k'=1}^{m'}$ of the K_k-to-$K'_{k'}$ positive cones defined by the smaller factor cones.

Lemma 13.2. *Let $K_1 \subset \mathbb{R}^{n_1}$, $K_2 \subset \mathbb{R}^{n_2}$ be regular convex cones. The K_1-to-K_2 positive cone is isomorphic to the K_2^*-to-K_1^* positive cone. The isomorphism is given by the adjoint operator $\mathcal{A} \mapsto \mathcal{A}^T$.*

Proof. Let $\mathcal{A}$ be K_1-to-K_2 positive, i.e., $\mathcal{A}(x) \in K_2$ for all $x \in K_1$. For all $y \in K_2^*$ we then have $\langle y, \mathcal{A}(x)\rangle = \langle \mathcal{A}^T(y), x\rangle \geq 0$. But since this holds also for all $x \in K_1$, it implies $\mathcal{A}^T(y) \in K_1^*$.

Let on the other hand $\mathcal{A}$ be not K_1-to-K_2 positive. Then there exists $x \in K_1$ such that $\mathcal{A}(x) \notin K_2$, and there exists $y \in K_2^*$ such that $\langle y, \mathcal{A}(x)\rangle = \langle \mathcal{A}^T(y), x\rangle < 0$. But then $\mathcal{A}^T(y) \notin K_1^*$, and $\mathcal{A}^T$ is not K_2^*-to-K_1^* positive. □

Example: For every $K \subset \mathbb{R}^n$, the K-to-$\mathbb{R}_+$ positive cone is the dual cone K^*.

The inclusion $\mathcal{A}_x[K_B] \subset K$ is hence equivalent to $\mathcal{A}_x$ being a K_B-to-K positive map. The solvability of the robust counterpart thus depends on the availability of a nice description of the K_B-to-K positive cone.

Let us consider some common types of uncertainty, classified according to the shape of the compact set B.

- L_1-ball (hyper-octahedron): due to its small number of vertices this is a case of the finite number of scenarios considered above and hence readily solvable, but it may poorly describe the true uncertainty in higher dimensions;
- L_2-ball: more generally ellipsoidal uncertainty, well-balanced uncertainty naturally occurring when data is obtained from parametric estimation, but less tractable than the L_1-ball;
- L_∞-ball (box uncertainty): occurs if we have interval uncertainty, still less tractable.

Robust linear programs: If the original conic program is an LP, then $K = \mathbb{R}_+^N$. By the preceding lemma a map $\mathcal{A} : \mathbb{R}^N \to \mathbb{R}^m$ is K_B-to-$\mathbb{R}_+^N$ positive if and only if $\mathcal{A}^T$ is $\mathbb{R}_+^N$-to-K_B^* positive. Equivalently, $\mathcal{A}^T$ maps every extreme ray of $\mathbb{R}_+^N$ into K_B^*. However, the extreme rays of $\mathbb{R}_+^N$ are generated by the basis vectors e_j, $j = 1, \dots, N$. Hence $\mathcal{A}^T$ is K_B-to-$\mathbb{R}_+^N$ positive if and only if every column of the matrix $\mathcal{A}^T$, or equivalently every row of $\mathcal{A}$, is in K_B^*. Thus the $\mathbb{R}_+^N$-to-K_B positive cone is isomorphic to a direct product of N copies of cones K_B^*.

The RC of a linear program is hence in the class determined by the uncertainty. For polyhedral uncertainty it is an LP, for ellipsoidal uncertainty it is an SOCP.

13.1 Robust programs with ellipsoidal uncertainty

If the uncertainty set B is an L_2-ball, then its homogenization is a Lorentz cone L_m. We shall now consider the cone of L_m-to-K positive maps for different cones K appearing in symmetric cone programming.

Robust LP: As seen above, a robust LP can be written as an SOCP.

Robust SOCP: The cone underlying an SOCP is a direct product of Lorentz cones. Hence the L_m-to-K positive cone is a product of L_m-to-L_n positive cones for different n. We shall now describe the L_m-to-L_n positive cone by a *linear matrix inequality* (LMI).

We start with the standard description of a single Lorentz cone L_r by an LMI. Define a linear map $\mathcal{W}_r : \mathbb{R}^r \to \mathcal{S}^{r-1}$ into the space of real symmetric $(r-1)\times(r-1)$ matrices by

$$\mathcal{W}_r(x) = \begin{pmatrix} x_0 + x_1 & x_2 & \cdots & \cdots & x_{r-1} \\ x_2 & x_0 - x_1 & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & 0 & 0 \\ \vdots & 0 & 0 & x_0 - x_1 & 0 \\ x_{r-1} & 0 & \cdots & 0 & x_0 - x_1 \end{pmatrix}.$$

Denote also by $\mathcal{A}(r)$ the space of skew-symmetric $r \times r$ matrices. Then the L_m-to-L_n positive cone can be described as follows.

Theorem 13.3. *Consider a map $A : \mathbb{R}^m \to \mathbb{R}^n$ given by a real $n \times m$ matrix. Then A is L_m-to-L_n positive if and only if there exists $X \in \mathcal{A}(n-1) \otimes \mathcal{A}(m-1)$ such that*

$$(\mathcal{W}_n \otimes \mathcal{W}_m)(A) + X \succeq 0.$$

Here the matrix $A \in \mathbb{R}^{m\times n}$ is considered as the tensor product space $\mathbb{R}^n \otimes \mathbb{R}^m$. The map $\mathcal{W}_n \otimes \mathcal{W}_m$ acts on rank 1 matrices as $xy^T \mapsto \mathcal{W}_n(y) \otimes \mathcal{W}_m(x)$ and is extended to arbitrary matrices by linearity. On the right-hand side the symbol $\otimes$ denotes the Kronecker product of matrices.

The theorem yields a (lifted) LMI representation of the L_m-to-L_n positive cone.

Example: The map $\mathcal{W}_4 \otimes \mathcal{W}_4$ takes 4×4 matrices A to symmetric 9×9 matrices, given by

$$(\mathcal{W}_4 \otimes \mathcal{W}_4)(A) = \begin{pmatrix} A_{++} & A_{+2} & A_{+3} & A_{2+} & A_{22} & A_{23} & A_{3+} & A_{32} & A_{33} \\ A_{+2} & A_{+-} & & A_{22} & A_{2-} & & A_{32} & A_{3-} & \\ A_{+3} & & A_{+-} & A_{23} & & A_{2-} & A_{33} & & A_{3-} \\ A_{2+} & A_{22} & A_{23} & A_{-+} & A_{-2} & A_{-3} & & & \\ A_{22} & A_{2-} & & A_{-2} & A_{--} & & & & \\ A_{23} & & A_{2-} & A_{-3} & & A_{--} & & & \\ A_{3+} & A_{32} & A_{33} & & & & A_{-+} & A_{-2} & A_{-3} \\ A_{32} & A_{3-} & & & & & A_{-2} & A_{--} & \\ A_{33} & & A_{3-} & & & & A_{-3} & & A_{--} \end{pmatrix},$$

where $A_{+\pm} = A_{00} \pm A_{01} + A_{10} \pm A_{11}$, $A_{-\pm} = A_{00} \pm A_{01} - A_{10} \mp A_{11}$, $A_{\pm k} = A_{0k} \pm A_{1k}$, $A_{k\pm} = A_{k0} \pm A_{k1}$. A generic matrix from $\mathcal{A}(3) \times \mathcal{A}(3)$ has the form

$$X = \begin{pmatrix} 0 & 0 & 0 & 0 & X_{15} & X_{16} & 0 & X_{18} & X_{19} \\ 0 & 0 & 0 & -X_{15} & 0 & X_{26} & -X_{18} & 0 & X_{29} \\ 0 & 0 & 0 & -X_{16} & -X_{26} & 0 & -X_{19} & -X_{29} & 0 \\ 0 & -X_{15} & -X_{16} & 0 & 0 & 0 & 0 & X_{48} & X_{49} \\ X_{15} & 0 & -X_{26} & 0 & 0 & 0 & -X_{48} & 0 & X_{59} \\ X_{16} & X_{26} & 0 & 0 & 0 & 0 & -X_{49} & -X_{59} & 0 \\ 0 & -X_{18} & -X_{19} & 0 & -X_{48} & -X_{49} & 0 & 0 & 0 \\ X_{18} & 0 & -X_{29} & X_{48} & 0 & -X_{59} & 0 & 0 & 0 \\ X_{19} & X_{29} & 0 & X_{49} & X_{59} & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The matrix A is L_4-to-L_4 positive if and only if there exists X of the above form such that $(\mathcal{W}_4 \otimes \mathcal{W}_4)(A) + X \succeq 0$.

Thus the robust counterpart of an SOCP with ellipsoidal uncertainty can be written as an SDP.

Robust SDP: To determine whether a linear map $A : \mathbb{R}^m \to \mathcal{S}^n$ is in the L_m-to-$\mathcal{S}^n_+$ positive cone is equivalent to the NP-hard matrix ellipsoid problem [3]. Therefore we are not able to solve the RC of an SDP with ellipsoidal uncertainty exactly. However, we may approximate it by an SDP.

Lemma 13.4. *Consider a map $A : \mathbb{R}^m \to \mathcal{S}^n$ given by*

$$x \mapsto \sum_{k=0}^{m-1} x_k A_k, \quad A_k \in \mathcal{S}^n$$

Define an associated matrix

$$\mathcal{M}_A = \begin{pmatrix} A_0 + A_1 & A_2 & \cdots & \cdots & A_{m-1} \\ A_2 & A_0 - A_1 & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & 0 & 0 \\ \vdots & 0 & 0 & A_0 - A_1 & 0 \\ A_{m-1} & 0 & \cdots & 0 & A_0 - A_1 \end{pmatrix}$$

and suppose there exists a matrix $X \in \mathcal{A}(m-1) \otimes \mathcal{A}(n)$ such that $\mathcal{M}_A + X \succeq 0$. Then A is L_m-to-$\mathcal{S}^n_+$ positive.

Proof. First we prove the desired condition $A(x) \succeq 0$ for special elements $x \in L_m$, namely such that $x_0 + x_1 = 1$ and $x_0^2 = x_1^2 + \cdots + x_{n-1}^2$. Then we have $x_0 - x_1 = ||\tilde{x}||_2^2$ with $\tilde{x} = (x_2, \ldots, x_{n-1})^T$.

Let $z \in \mathbb{R}^n$ be arbitrary and set $v = (1, \tilde{x}^T)^T \otimes z \in \mathbb{R}^{(m-1)n}$. We then get $v^T X v = 0$, because every block of X is skew-symmetric, and therefore

$$v^T \mathcal{M}_A v = z^T [A_0 + A_1 + 2 \sum_{k=2}^{m-1} x_k A_k + ||\tilde{x}||_2^2 (A_0 - A_1)] z = 2 z^T A(x) z \geq 0$$

by the assumption on A. Hence $A(x) \succeq 0$ for all such special x.

However, the whole Lorentz cone is the closure of the conic convex hull of such special elements x, and hence A is L_m-to-$\mathcal{S}^n_+$ positive. □

Theorem 13.5. *The above inner approximation of the L_m-to-$\mathcal{S}^n_+$ positive cone is exact for $n \leq 3$.*

13.2 Matrix cube

The case of an uncertain SDP with L_∞ uncertainty (box uncertainty) has been considered in [2]. Although the robust counterpart is hard to solve exactly, in the case of low rank perturbations an approximation of the RC can be formulated which has a quantitative bound on the error.

Let K_U be the cone obtained by homogenization of the L_∞ unit norm ball. Consider a linear map $L : \mathbb{R}^{m+1} \to \mathcal{S}^n$ determined by matrices $B_0, \ldots, B_m \in \mathcal{S}^n$ according to $x = (x_0, \ldots, x_m)^T \mapsto \sum_{l=0}^m x_l B_l$. We have $L \in \mathrm{Pos}(K_U, \mathcal{S}^n_+)$ if and only if $B_0 + \sum_{l=1}^m \epsilon_l B_l \succeq 0$ for all combinations $\epsilon_l \in \{-1, +1\}$. In order to check this inclusion we have to verify exponentially many, namely 2^m, linear matrix inequalities. This condition can, however, be strengthened with a decrease in complexity down to an order linear in m. Consider the semi-definite relaxation

$$\mathcal{SR} = \left\{ L : x \mapsto \sum_{l=0}^m x_l B_l \mid \exists\, X_l \in \mathcal{S}^n : \; X_l \succeq \pm B_l, \; l = 1, \ldots, m; B_0 \succeq \sum_{l=1}^m X_l \right\}.$$

Clearly $\mathcal{SR} \subset \mathrm{Pos}(K_U, \mathcal{S}^n_+)$ and $\mathcal{SR}$ is an inner approximation of the cone of positive maps. Checking the inclusion $L \in \mathcal{SR}$ is equivalent to checking the feasibility of a system of $2m+1$ LMIs.

In order to bound the error of this approximation we shall need the following quantity. For $k \in \mathbb{N}_+$ define

$$\begin{aligned} \eta(k) &= \min_{V \in \mathcal{S}^k} \mathbb{E}_\xi |\xi^T V \xi| \;:\; \quad ||V||_1 = 1 \\ &= \min_{\lambda \in \mathbb{R}^k} \mathbb{E}_\kappa | \sum_{j=1}^k \lambda_j \kappa_j | \;:\; \quad ||\lambda||_1 = 1, \end{aligned}$$

where $\xi \in \mathbb{R}^k$ is a standard normally distributed random vector, and $\kappa_1, \ldots, \kappa_k$ are independent random scalars with probability density $\mu(t) = (2\pi t e^t)^{-1/2}$, i.e., according to the $\chi^2(1)$ law. The values of $\eta(k)$ for low k are given in the following table.

k	$\eta(k)$, exact	$\eta(k)$, numerical
1	1	1.0000
2	$\frac{2}{\pi}$	0.6366
3	Root$(t^3 + 9t^2 + 135t - 81)$	0.5764
4	$\frac{1}{2}$	0.5000

Note that the expectation is convex in λ and invariant with respect to permutations of the elements of λ. Hence the minimum over $\lambda \in S^{k-1}$ is achieved at a point of the form $\lambda = (\lambda_+, \dots, \lambda_+, \lambda_-, \dots, \lambda_-)$, where $\lambda_+ > 0$, $\lambda_- < 0$ are reals appearing k_+, k_- times, respectively. The value of the minimum is given by

$$\frac{2\Gamma(\frac{k}{2}+1)}{\Gamma(\frac{k_+}{2})\Gamma(\frac{k_-}{2})} \int_0^1 |\lambda_+\tau + \lambda_-(1-\tau)|\tau^{k_+/2-1}(1-\tau)^{k_-/2-1}\, d\tau.$$

Inserting $k_+ = k_- = \frac{k}{2}$, $\lambda_+ = -\lambda_- = \frac{1}{k}$ we obtain the value $\frac{\Gamma(\frac{k}{2}+1)}{2^{k/2}\Gamma(\frac{k}{4}+1)^2} \approx \frac{2}{\sqrt{\pi k}}$. For even k this value is achieved at some vector λ. Further we have the lower bound $\eta(k) \geq \frac{2}{\pi\sqrt{k}}$ [2], and $\eta(k)$ is decreasing by construction. Thus we get the asymptotics $\eta(k) \sim k^{-1/2}$.

Let now $L \notin \mathcal{SR}$, where L is given by a tuple $(B_0, \dots, B_m)$ of matrices. Then there exists a dual $Z = (Z_0, \dots, Z_m) \in \mathcal{SR}^*$ such that $\langle Z, L\rangle = \sum_{j=0}^m \langle Z_j, B_j\rangle < 0$. By construction $\mathcal{SR}$ is a linear projection of the self-dual product cone $(\mathcal{S}_+^n)^{2m+1}$. Hence the dual cone $\mathcal{SR}^*$ is a linear section of this product cone, namely

$$\mathcal{SR}^* = \{(Z_0, Z_1^+ - Z_1^-, \dots, Z_m^+ - Z_m^-) \mid Z_j^+, Z_j^- \succeq 0,\ Z_j^+ + Z_j^- = Z_0\ \forall\, j = 1, \dots, m\}.$$

It is not hard to prove that for matrices $B \in \mathcal{S}^n$, $Z \in \mathcal{S}_+^n$ the optimal value of the semi-definite program

$$\min_{Z^+, Z^- \succeq 0} \langle B, Z^+ - Z^-\rangle \ :\quad Z^+ + Z^- = Z$$

equals $-||Z^{1/2}BZ^{1/2}||_1$.

For a matrix $V \in \mathcal{S}^n$ of rank $\leq k$ we have the estimate

$$\mathbb{E}_\xi|\xi^T V\xi| \geq \eta(k) \cdot ||V||_1,$$

where $\xi \sim \mathcal{N}(\mathbf{0}, I)$ is a random normal vector.

Hence

$$0 > \langle Z, L\rangle \geq \operatorname{tr}(Z_0^{1/2}B_0Z_0^{1/2}) - \sum_{j=1}^m ||Z_0^{1/2}B_jZ_0^{1/2}||_1 \geq \mathbb{E}_\xi \left(\xi^T Z_0^{1/2}B_0Z_0^{1/2}\xi - \sum_{j=1}^m \frac{1}{\eta(\text{rk } B_j)}|\xi^T Z_0^{1/2}B_jZ_0^{1/2}\xi|\right).$$

Therefore there exists $\zeta \in \mathbb{R}^n$ such that

$$\zeta^T B_0\zeta - \sum_{j=1}^m \frac{1}{\eta(\text{rk } B_j)}|\zeta^T B_j\zeta| < 0.$$

Set $\epsilon_j = -\operatorname{sgn}(\zeta^T B_j\zeta)$, then we get $\zeta^T(B_0 + \sum_{j=1}^m \frac{\epsilon_j}{\eta(\text{rk } B_j)}B_j)\zeta < 0$. It follows that the map $\tilde{L}$ determined by $(B_0, \frac{1}{\eta(\text{rk } B_1)}B_1, \dots, \frac{1}{\eta(\text{rk } B_m)}B_m)$, does not belong to the cone of positive maps $\text{Pos}(K_U, \mathcal{S}_+^n)$.

Thus we obtain the following result.

Theorem 13.6. *Let the matrix cube* $\mathbf{U} = \{B_0 + \sum_{j=1}^m u_jB_j \mid u_j \in [-1, 1]\}$ *be given by the tuple of matrices* $(B_0, \dots, B_m) \notin \mathcal{SR}$. *Then the dilated matrix cube* $\tilde{\mathbf{U}} = \{B_0 + \sum_{j=1}^m u_jB_j \mid \eta(\text{rk } B_j) \cdot |u_j| \leq 1\}$ *is not contained in the positive semi-definite cone* $\mathcal{S}_+^n$.

References

[1] Aharon Ben-Tal, Laurent El Ghaoui, and Arkadi Nemirovski. *Robust Optimization.* Princeton University Press, 2009.

[2] Aharon Ben-Tal and Arkadi Nemirovski. On tractable approximations of uncertain linear matrix inequalities affected by interval uncertainty. *SIAM J. Optim.*, 12(3):811–833, 2002.

[3] Yuri Nesterov. Random walk in a simplex and quadratic optimization over convex polytopes. Discussion paper 2003/71, CORE, Louvain-la-Neuve, 2003.