8 Self-concordant barriers

All the difficulty of a conic program is hidden in the conic constraint \( x \in K \). The basic idea of interior-point methods for solving conic programs is to eliminate the conic constraint by the addition of a penalty function \( F : K^o \to \mathbb{R} \) to the linear objective. For convenience we may define \( F(x) = +\infty \) for \( x \notin \text{int} K \). The penalty function should satisfy the following requirements:

- \( F(x) \) is convex (the problem should remain convex),
- \( F(x) \) is sufficiently smooth (to be able to use second-order methods),
- \( \lim_{x \to \partial K} F(x) = +\infty \) (acts as a barrier for \( K \)),
- \( F \) behaves well with the Newton method.

The last requirement needs a formalization, which is achieved by the introduction of self-concordancy.

8.1 Self-concordant functions

In this section we introduce a class of convex functions which is well suited for minimization by the Newton method. The proofs of the statements can be found in [8].

We consider the problem of minimizing a convex \( C^3 \) function \( f : D \to \mathbb{R} \) defined on a convex domain \( D \subset A \), where \( A \) is some affine space. For simplicity we assume that the Hessian \( f'' \) of the function is positive definite everywhere on \( D \). Given an iterate \( x_k \in D \), the Newton algorithm computes the next iterate as

\[
x_{k+1} = x_k - (f''(x_k))^{-1} f'(x_k).
\]

The point \( x_{k+1} \) can be interpreted as the minimizer of the strictly convex second order Taylor polynomial of \( f \) at \( x_k \),

\[
q_k(x) = f(x_k) + \langle f'(x_k), x - x_k \rangle + \frac{1}{2} (x - x_k)^T f''(x_k)(x - x_k).
\]

The above formulae need an explanation. The gradient \( f'(x_k) \) and the Hessian \( f''(x_k) \) are a linear form and a quadratic form, respectively, on the tangent space to \( A \) at \( x_k \), which can be identified with the vector space \( V \) underlying the affine space. The difference \( x - x_k \) is a vector, and these forms can be applied to it. The expression \( (f''(x_k))^{-1} f'(x_k) \) is also a vector and can be subtracted from a point in affine space.

The Newton algorithm is affinely invariant, i.e., its output does not change when computed in another coordinate system on the affine space. Therefore there exists no other natural norm on \( V \) than the Euclidean norm \( \| \cdot \|_{x_k} \), defined by the Hessian \( f''(x_k) \) at the current iterate. In this norm the level subsets \( \{ x \in A \mid q_k(x) \leq c \} \) are norm balls around the minimizer \( x_{k+1} \) of the second order approximation \( q_k(x) \). The current iterate lies at a distance

\[
\rho = \sqrt{(x_{k+1} - x_k)^T f''(x_k)(x_{k+1} - x_k)} = \sqrt{f'(x_k)^T (f''(x_k))^{-1} f'(x_k)} \tag{1}
\]

from the minimizer, which at the same time equals the norm of the gradient at the current iterate. This quantity can also be expressed through the difference between the current function value and the minimum value of \( q_k(x) \),

\[
f(x_k) - q_k(x_{k+1}) = -\langle f'(x_k), x_{k+1} - x_k \rangle - \frac{1}{2} (x_{k+1} - x_k)^T f''(x_k)(x_{k+1} - x_k) = \frac{1}{2} f'(x_k)^T (f''(x_k))^{-1} f'(x_k) = \frac{\rho^2}{2}.
\]

In order to be able to make assertions about the behaviour of the Newton algorithm, we must ensure that the norm defined by the Hessian \( f'' \) does not change too much when passing from the current iterate to the next. On the other hand, in an affinely invariant framework any such changes can be compared only against the step length measured in the norm defined by the current iterate. This leads to the following definition.

**Definition 8.1.** Let \( a > 0 \). A convex \( C^3 \) function \( f : D \to \mathbb{R} \) on a domain \( D \) in some affine space is called \( a \)-self-concordant if it satisfies the inequality

\[
|f''(x)[h, h, h]| \leq 2a^{-1/2}(f''(x)[h, h])^{3/2}
\]

for all \( x \in D \) and all tangent vectors \( h \).
It is called strongly $a$-self-concordant if in addition $\lim_{x \to \partial D} f(x) = +\infty$.

Here the derivatives of $f$ are as above treated as multi-linear maps on the space of tangent vectors, i.e.,
\[
\frac{\partial^2 f(x)}{\partial x_i \partial x_j} h_i h_j, \quad \frac{\partial^3 f(x)}{\partial x_i \partial x_j \partial x_k} h_i h_j h_k,
\]
where $n$ is the dimension of the space. The exponent $\frac{3}{2}$ on the right-hand side has been introduced to obtain the same degree of homogeneity in $h$ on both sides of the inequality.

The quantity $2a^{-1/2}$ can be interpreted as Lipschitz constant of the Hessian $f''$. The larger $a$, the stronger the Lipschitz condition. Note, however, that an $a$-self-concordant function can always be normalized to a $1$-self-concordant function (or just self-concordant function for short) by multiplication with the constant $a^{-1}$.

The limit condition means that $f$ tends to $+\infty$ if evaluated at any sequence of points in $D$ which tend to a boundary point of $D$. This condition ensures that the level sets $\{x \in D \mid f(x) \leq c\}$ are closed for all constants $c \in \mathbb{R}$.

The following result guarantees that the Newton algorithm can safely make steps of finite length, where "safely" means that the next iterate stays in the domain $D$ and features a lower objective value than the previous one.

**Lemma 8.2.** Let $f : D \to \mathbb{R}$ be a strongly $a$-self-concordant function. Then for every $x^* \in D$ the Dikin ellipsoid
\[
E_{2^*, a} = \{x \mid (x - x^*)^T f''(x^*)(x - x^*) < a\}
\]
around $x^*$ is contained in the domain $D$.

It follows that a Newton step of full length is safe if the gradient $f'(x_k)$ at the current iterate has norm smaller than $\sqrt{a}$ as measured in the norm defined by $f''(x_k)$. If the norm of the gradient is larger, then we must make shorter steps, i.e., employ a damped Newton algorithm. We define the following normalized version of the gradient norm.

**Definition 8.3.** Let $f : D \to \mathbb{R}$ be a strongly $a$-self-concordant function with non-degenerate Hessian. The Newton decrement of $f$ at $x \in D$ is defined by
\[
\lambda(x; f) = a^{-1/2} \rho = a^{-1/2} \sqrt{f'(x_k)^T f''(x_k)^{-1} f'(x_k)}.
\]

We may then propose the following damped Newton algorithm.

We are given a strongly $a$-self-concordant function $f : D \to \mathbb{R}$ with non-degenerate Hessian to minimize.

choose a starting point $x_0 \in D$, fix an accuracy $\epsilon > 0$, and set $k = 0$

while $\lambda(x_k; f) \geq \epsilon$,
\[
\begin{align*}
\gamma_k &= \frac{1}{1 + a \lambda(x_k; f)} \quad & \lambda(x_k; f) \geq \frac{1}{3}, \\
\lambda(x_k; f) &= (2 - \sqrt{3}, \frac{1}{3}, \lambda(x_k; f)) \quad & \lambda(x_k; f) \in (2 - \sqrt{3}, \frac{1}{3}), \\
1, & \quad & \lambda(x_k; f) \leq 2 - \sqrt{3}.
\end{align*}
\]
\[
x_{k+1} = x_k - \gamma_k f''(x_k)^{-1} f'(x_k)
\]
\[
k \leftarrow k + 1
\]
end

Hence while the Newton decrement $\lambda(x; f)$ is large, the algorithm makes small steps. Below the threshold of $2 - \sqrt{3}$ the algorithm makes the full Newton step. The algorithm has the following properties:

- for $\lambda(x_k; f) \geq \frac{1}{3}$ we have $f(x_{k+1}) \leq f(x_k) - a(\lambda(x_k; f) - \log(1 + \lambda(x_k; f)))$,
- a value in the intermediate range $(2 - \sqrt{3}, \frac{1}{3})$ occurs at most once,
- for $\lambda(x_k; f) \leq 2 - \sqrt{3}$ we have $\lambda(x_{k+1}; f) \leq \left(\frac{\lambda(x_k; f)}{1 - \lambda(x_k; f)}\right)^2$. 

Thus below the threshold $2 - \sqrt{3}$ the algorithm converges \textit{quadratically} (at each step the number of known digits approximately doubles).

It may happen that the algorithm does not converge at all, namely in the case when the function $f$ is unbounded from below. However, in this case the Newton decrement stays above the threshold value 1 and the sequence $f(x_k)$ tends to $-\infty$. If at some step the Newton decrement falls below 1, then the function is guaranteed to be bounded from below and the iterates will converge to the minimizer.

\textbf{Geometric interpretation:} We shall now give a geometric interpretation of the self-concordance condition. The minimizer $x^*$ of the function $f$ is characterized by the condition $f'(x^*) = 0$. We may then ignore the information on the values of $f$ and consider only its gradient.

If $V$ is the vector space underlying the affine space $A$ containing the domain of interest, then the gradient is an element of the dual vector space $V^*$. The graph of the map $\nabla f : x \mapsto f'(x)$ is then an $n$-dimensional sub-manifold $M$ of the $2n$-dimensional product space $A \times V^*$. The minimizer $x^*$ is then given by the point $(x^*,0)$ of intersection of $M$ with the horizontal subspace $H_0 = \{(x,0) \mid x \in A\}$ of $A \times V^*$.

Let us interpret the Newton algorithm in these terms. At the point $x_k$, we construct a quadratic approximation $q_k(x)$ of $f$. In the same way as the gradient map $\nabla f$ the gradient map $\nabla q_k$ has a graph $M_k \subset A \times V^*$. However, since $q_k(x)$ is a quadratic function, its gradient will be \textit{linear}, and hence $M_k$ is actually even a \textit{subspace} of $A \times V^*$. Since the gradient and the Hessian of $q_k(x)$ coincide with those of $f(x)$ at $x_k$, this subspace is the \textit{tangent space} to $M$ at the point $(x_k,f'(x_k))$.

The quality of approximation of a sub-manifold in the neighbourhood of some point by its tangent subspace at that point is given by the \textit{curvature} of the sub-manifold. The self-concordance condition then states that the
curvature of $M$ is uniformly bounded. Here curvature can be interpreted as the curvature of the sub-manifold $M$ in the ambient space $A \times V^*$, but also as the curvature of the domain $D$, equipped with the Riemannian metric defined by the Hessian $f''$. The latter is a quadratic function of the former.

### 8.2 Self-concordant barriers

In conic optimization the domain is a cone, which has an additional structure, namely invariance with respect to homotheties. It is natural to demand this invariance from the penalty function $F$ too.

**Definition 8.4.** Let $K \subset \mathbb{R}^n$ be a regular convex cone. A convex $C^3$ function $F : K^\circ \to \mathbb{R}$ with positive definite Hessian is called a logarithmically homogeneous self-concordant barrier with barrier parameter $\nu$ if $F$ is a strongly 1-self-concordant function on $K^\circ$ and $F$ is logarithmically homogeneous of degree $-\nu$, i.e.,

$$F(\alpha x) = -\nu \log \alpha + F(x)$$

(2)

for all $x \in K^\circ$ and $\alpha > 0$.

Homotheties thus act by adding constants to the barrier. It turns out that this condition also uniformly bounds the Newton decrement.

The gradient norm of a barrier with parameter $\nu$ is equal to the constant $\sqrt{\nu}$. Indeed, differentiating (2) with respect to $x$ we obtain $\alpha F''(\alpha x) = F'(x)$. Differentiating this and (2) with respect to $\alpha$ at $\alpha = 1$, we obtain

$$F'(x) + F''(x) \cdot x = 0, \quad \langle F'(x), x \rangle = -\nu.$$

It follows that $(F''(x))^{-1} F'(x) = -x$ and hence $(F'(x)) (F''(x))^{-1} F'(x) = -(F'(x), x) = \nu$.

The Newton decrement $\lambda(x; F)$ is hence everywhere equal to $\sqrt{\nu}$. It is desirable to have barriers with parameter as low as possible at our disposal, because a lower value of the parameter corresponds to larger step lengths and faster convergence. Actually, the tractability of a conic program is essentially determined by the availability of a computable self-concordant barrier on the underlying cone. We now give concrete examples of such barriers.

**Symmetric cones:** For symmetric cones a barrier is available via the Jordan algebra structure, namely the logarithm of the inverse of the determinant.

<table>
<thead>
<tr>
<th>symmetric cone</th>
<th>$S_+^n$</th>
<th>$\mathbb{R}^n_+$</th>
<th>$L_n$</th>
<th>$K = \bigcap_j K_j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>barrier</td>
<td>$-\log \det A$</td>
<td>$-\sum_{j=1}^n \log x_j$</td>
<td>$-\log(x_0^2 - x_1^2 - \cdots - x_n^2)$</td>
<td>$\sum_j F_j$</td>
</tr>
<tr>
<td>parameter $\nu$</td>
<td>$n$</td>
<td>$n$</td>
<td>$\frac{2}{n}$</td>
<td>$\sum_j \nu_j$</td>
</tr>
</tbody>
</table>

The barrier parameter of these barriers is optimal. These barriers possess an additional advantageous property, they are self-scaled. This property allows for especially efficient interior-point methods for solving symmetric cone programs, the long-step methods [5],[6].

**$p$-norm cone:** Consider the cone

$$K_p = \{(x_0, x) \in \mathbb{R} \times \mathbb{R}^n \mid x_0 \geq ||x||_p\}.$$

This cone allows to formulate constraints on the $p$-norm of linear combinations of the decision variables of the conic program. For non-rational $p$ this is a transcendental cone. We shall represent $K_p$ as a linear projection of another cone

$$\hat{K}_p = \{(x_0, x, y) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \mid y \geq 0, \sum_i y_i = x_0, y_i^{1/p} x_0^{1-1/p} \geq |x_i|\}.$$

Indeed, $(x_0, x) \in K_p$ if and only if there exists $y$ such that $(x_0, x, y) \in \hat{K}_p$. The function

$$\Phi(x_0, x, y) = -\sum_i \left( \log(y_i^{2/p} x_0^{2-2/p} - x_i^2) + \log y_i + \log x_0 \right)$$

is then a self-concordant barrier on $\hat{K}_p$ with parameter $\nu = 4n$. Details can be looked up in [7].
**Exponential cone:** Consider the cone

\[ K_{\exp} = \{(x, y, 0) \mid x \leq 0, \ y \geq 0\} \cup \{(x, y, z) \mid z > 0, \ y \geq ze^{x/z}\}, \]

encountered in geometric programming. On this cone we have the barrier

\[ F(x, y, z) = -\log(z \log \frac{y}{z} - x) - \log y - \log z \]

with parameter \( \nu = 3 \). This parameter value is optimal.

### 8.3 Legendre duality

Every barrier \( F : \text{int} \ K \rightarrow \mathbb{R} \) on a regular convex cone \( K \) with parameter \( \nu \) generates a barrier \( F_* : \text{int} \ K^* \rightarrow \mathbb{R} \) on the dual cone with the same parameter. This dual barrier is given by the Legendre transform

\[ F_*(s) = \sup_{x \in K} (-\langle s, x \rangle - F(x)). \]

The supremum is achieved at the point \( x \in K^* \) satisfying \( F'(x) = -s \), which always exists if \( s \in \text{int} \ K^* \). The map \( \mathcal{L} : x \mapsto -F'(x) \) is a bijection between the interiors of \( K \) and \( K^* \).

If we consider these interiors as Riemannian manifolds equipped with the Hessians \( F''(x) \) and \( F''(s) \) as metrics, respectively, then \( \mathcal{L} \) is an isometry.

If \( K = K^* \) is a symmetric cone and \( F(x) = -\log \det x \), then \( F_* = F \) and the map \( \mathcal{L} \) is the inversion \( x \mapsto x^{-1} \) in the Jordan algebra underlying the cone.

### 8.4 Universal barriers

The availability of a logarithmically homogeneous self-concordant barrier on a convex cone enables us to solve conic programs over this cone. Naturally the question arises whether such barriers exist on arbitrary cones. The answer to this question is positive, and we shall learn to know two constructions of such barriers.

**Universal barrier.** For a regular convex cone \( K \subset \mathbb{R}^n \), consider its characteristic function

\[ \varphi(x) = \int_{K^*} e^{-\langle x, y \rangle} \, dy \]

for \( x \) in the interior of \( K \). Then the function \( F(x) = \log \varphi(x) \) is a logarithmically homogeneous self-concordant barrier on \( K \) with barrier parameter \( \nu = n \), the universal barrier. It has been introduced in [8] with a bound \( O(n) \) on the parameter, the actual value of the parameter has been established in [2],[3].

The universal barrier is given by a multi-dimensional integral over the dual cone and is difficult to compute even for relatively simple non-homogeneous cones. It also suffers from the draw-back that its Legendre dual is in general not the universal barrier for \( K^* \).

**Entropic barrier.** This is the dual barrier to the universal barrier. Its parameter also equals the dimension of the cone. For details see [2].

**Canonical barrier.** The construction of this barrier relies on a deep result in the theory of partial differential equations.

**Theorem 8.5.** Let \( D \subset \mathbb{R}^n \) be a convex domain containing no line. Then there exists a unique smooth solution \( F : D \rightarrow \mathbb{R} \) with positive definite Hessian of the PDE \( \log \det F^* = 2F \) with boundary condition \( \lim_{x \to \partial D} F(x) = +\infty \).

If the domain \( D \) is a regular convex cone, then this solution can be proven to be a logarithmically homogeneous self-concordant barrier with barrier parameter \( \nu = n \), the canonical barrier. Its Legendre dual can be proven to coincide with the canonical barrier on \( K^* \). It is also difficult to compute, but for a few non-homogeneous cones explicit expressions are available.
Example: For the 3-dimensional cone $K_{\text{exp}}$ the canonical barrier is given by

$$F_{\text{can}}(x, y, z) = -\log y - 2 \log z + \phi(\log \frac{y}{z} - \frac{x}{z}),$$

where the scalar function $\phi: \mathbb{R}_{++} \rightarrow \mathbb{R}$ is given implicitly by the curve

$$\left\{ \left( \begin{array}{c} t \\ \phi \end{array} \right) = \frac{1}{2} \left( \log(1 + \kappa) + 2\kappa \\ \log(1 + \kappa) - 3\log \kappa \right) \bigg| \kappa \in \mathbb{R}_{++} \right\}.$$  

For homogeneous cones the three barriers coincide, and for symmetric cones they are proportional to the self-scaled barrier seen above.

8.5 Interior-point methods

We consider only the most basic version of a primal short-step interior-point algorithm. Let $K$ be a regular convex cone and $F$ a logarithmically homogeneous self-concordant barrier on $K$. We approximate the conic program

$$\inf_{x \in K} c^T x : Ax = b,$$

whose solution $x^*$ we suppose to exist, by the family of problems

$$\min_x (\tau \cdot c^T x + F(x)) : Ax = b$$

parameterized by a real parameter $\tau \geq 0$. The objectives of these problems are convex self-concordant functions. The minimizers $x^*(\tau)$ form a curve, the central path, which links the analytic center $x^*(0)$ to the solution $x^* = x^*(+\infty)$ of the original conic program. If the original problem has more than one solution, then the central path converges to a point in the relative interior of the solution set.

The method starts with an iterate $(x_0, \tau_0)$, where $x_0$ is a point close to the analytic center, and $\tau_0$ is a small enough positive number such that $x_0$ is located in the fast convergence region of the Newton method for the function $\tau_0 \cdot c^T x + F(x)$.

The $k$-th iteration consists of one Newton step $x_k \mapsto x_{k+1}$ towards the minimizer of the function $\tau_k \cdot c^T x + F(x)$ and an increase of the parameter $\tau$ by a constant factor $\theta > 1$, $\tau_{k+1} = \theta \cdot \tau_k$. The parameter $\theta$ is chosen such that $x_{k+1}$ still lies in the fast convergence region for the new objective $\tau_{k+1} \cdot c^T x + F(x)$. The larger the barrier parameter $\nu$, the smaller $\theta$ has to be chosen and the shorter are the steps the method can make.

The iterates thus stay in a neighbourhood of the central path and follow this path towards the solution. This class of methods is conditioned on the availability of a barrier only and is applicable to conic programs over arbitrary regular convex cones.

If the cone $K$ is symmetric, then the additional algebraic structure allows to design long-step methods, which use global properties of the barrier and whose iterates can be far away from the central path.
9 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 $K \subseteq \mathbb{R}^N$ is 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)

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

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 $B \subseteq \mathbb{R}^{m-1}$ 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 $(A_j, b_j)$, $j = 1, \ldots, M$ be the vertices of $U$, then the RC can be written as

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

Hence the RC is an ordinary conic program with $M$ conic constraints over the same cone $K$.

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 $K$. 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

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

by the homogeneity of the cone $K$. Equivalently we obtain

$$\min_x \langle c, x \rangle : \quad A_x [K_B] \subseteq K,$$

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

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

Note that the coefficients of the linear map $A_x$ affine in $x$ and can be arranged in a real $m \times N$ matrix. We shall then consider the inclusion $A_x [K_B] \subseteq K$ 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 9.1. Let $K_1 \subset \mathbb{R}^{n_1}$, $K_2 \subset \mathbb{R}^{n_2}$ be regular convex cones. Call a linear map $A : \mathbb{R}^{n_1} \to \mathbb{R}^{n_2}$ $K_1$-to-$K_2$ positive if $A[K_1] \subset K_2$.

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.

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

Lemma 9.2. 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 $A \mapsto A^T$.

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

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

The inclusion $A_x[K_B] \subset K$ is hence equivalent to $A_x$ 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;
- $L_\infty$-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 $A : \mathbb{R}^N \to \mathbb{R}^m$ is $K_B$-to-$\mathbb{R}_+^N$ positive if and only if $A^T$ is $\mathbb{R}_+^N$-to-$K_B^*$ positive. Equivalently, $A^T$ 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 $e_j$, $j = 1, \ldots, N$. Hence $A^T$ is $K_B$-to-$\mathbb{R}_+^N$ positive if and only if every column of the matrix $A^T$, or equivalently every row of $A$, 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.

9.1 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 $W_r : \mathbb{R}^r \to \mathcal{S}^{r-1}$ into the space of real symmetric $(r - 1) \times (r - 1)$ matrices by

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

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 9.3.** 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

$$(W_n \otimes W_m)(A) + X \succeq 0.$$  

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 $W_n \otimes W_m$ acts on rank 1 matrices as $x y^T \mapsto W_n(y) \otimes W_m(x)$ and is extended to arbitrary matrices by linearity.

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

**Example:** The map $W_4 \otimes W_4$ takes $4 \times 4$ matrices $A$ to symmetric $9 \times 9$ matrices, given by

$$(W_4 \otimes W_4)(A) = \begin{pmatrix}
 A_{++} & A_{+2} & A_{+3} & A_{2+} & A_{22} & A_{23} & A_{4+} & A_{32} & A_{33} \\
 A_{2+} & A_{++} & A_{22} & A_{23} & A_{32} & A_{33} & A_{32} & A_{33} & A_{33} \\
 A_{+3} & A_{+2} & A_{22} & A_{23} & A_{32} & A_{33} & A_{32} & A_{33} & A_{33} \\
 A_{23} & A_{22} & A_{23} & A_{++} & A_{A_2} & A_{A_2} & A_{A_3} & A_{A_3} & A_{A_3} \\
 A_{33} & A_{32} & A_{33} & A_{32} & A_{++} & A_{A_2} & A_{A_3} & A_{A_3} & A_{A_3} \\
 A_{32} & A_{32} & A_{33} & A_{33} & A_{33} & A_{++} & A_{A_2} & A_{A_3} & A_{A_3} \\
 A_{33} & A_{33} & A_{33} & A_{33} & A_{33} & A_{33} & A_{++} & A_{A_2} & A_{A_3} \\
 A_{32} & A_{32} & A_{33} & A_{33} & A_{33} & A_{33} & A_{33} & A_{++} & A_{A_2} \\
 A_{33} & A_{33} & A_{33} & A_{33} & A_{33} & A_{33} & A_{33} & A_{33} & A_{++}
\end{pmatrix},$$

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

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

The matrix $A$ is $L_4$-to-$L_4$ positive if and only if there exists $X$ of the above form such that $(W_4 \otimes W_4)(A) + X \succeq 0$.

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 \cite{4}. 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 9.4. Consider a map \( A : \mathbb{R}^m \to S^n \) given by
\[
x \mapsto \sum_{k=0}^{m-1} x_k A_k, \quad A_k \in S^n
\]
Define an associated matrix
\[
M_A = \begin{pmatrix}
A_0 + A_1 & A_2 & \cdots & \cdots & A_{m-1} \\
A_2 & A_0 - A_1 & 0 & \cdots & 0 \\
\vdots & 0 & \ddots & 0 & 0 \\
A_{m-1} & 0 & \cdots & 0 & A_0 - A_1
\end{pmatrix}
\]
and suppose there exists a matrix \( X \in A(m-1) \otimes A(n) \) such that \( M_A + X \succeq 0 \). Then \( A \) is \( L_m \)-to-\( S^n_+ \) 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 \( v = (1, \tilde{x}^T)^T \times z \in \mathbb{R}(m-1)n \). We then get \( v^T X v = 0 \), because every block of \( X \) is skew-symmetric, and therefore
\[
v^T M_A v = z^T[A_0 + A_1 + 2 \sum_{k=2}^{m-1} x_k A_k + \|\tilde{x}\|_2^2 (A_0 - A_1)] z = 2z^T A(x) z \geq 0
\]
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-\( S^n_+ \) positive.

Theorem 9.5. The above inner approximation of the \( L_m \)-to-\( S^n_+ \) positive cone is exact for \( n \leq 3 \).

References


