8 Symmetric cones

In this section we shall introduce a class of cones which is very important for convex programming, the symmetric cones. These cones generalize the positive orthant $\mathbb{R}_+^n$, and the programs associated to the symmetric cones accordingly generalize the class of linear programs, while retaining most of their favorable properties. The class of symmetric cones is defined as the intersection of two classes of cones characterized by geometric properties, the *homogeneous* cones and the *self-dual* cones.

Definition 8.1. A closed convex cone K is called *symmetric* if it is both homogeneous and self-dual.

8.1 Homogeneous cones

Definition 8.2. Let $K \subset V$ be a closed convex cone in some vector space. An *automorphism* of K is an automorphism A of V such that $A[K] = K$.

The automorphisms of K form a group, the *automorphism group* $\operatorname{Aut} K$ of the cone. The automorphism group always contains the 1-parametric subgroup of *homotheties*, i.e., maps $g_\lambda : v \mapsto \lambda v$ which multiply every vector by a positive constant λ.

Definition 8.3. A regular convex cone K is called *homogeneous* if its automorphism group acts transitively on the interior of K, i.e., if for every $x, y \in K^o$ there exists an automorphism $A \in \operatorname{Aut} K$ such that $Ax = y$.

In a homogeneous cone every interior point is hence equivalent to any other interior point, and every interior point can be considered as the center of the cone. The homogeneous cones can be put into 1-to-1 correspondence with algebraic structures, so-called *T-algebras* [9]. A more explicit classification is also available [7, 4].

An example of a homogeneous cone which is not symmetric is given by the 5-dimensional cone $\{A \in \mathcal{S}_+^3 \mid A_{23} = 0\}$.

8.2 Self-dual cones

Let the vector space V be equipped with a Euclidean scalar product $\langle \cdot, \cdot \rangle$. This allows to identify the dual vector space V^* with V itself. The linear functional $p \in V^*$ is identified with the element $y \in V$ such that $\langle p, x \rangle = \langle y, x \rangle$. Here on the left we have the dual pairing, and on the right the scalar product.

Definition 8.4. Let $K \subset V$ be a closed convex cone, with the ambient vector space V equipped with a scalar product. Then K is called *self-dual* if $K^* = K$ under the identification of V with V^* by the scalar product.

If no scalar product is defined a priori, then K is called *self-dual* if there exists a scalar product on V such that $K^* = K$ under the identification of V with V^* generated by this scalar product.

There exists also a wider definition of self-dual cones, namely a cone is called self-dual if it is linearly isomorphic to its dual.

An example of a self-dual cone which is not symmetric is given by the *power cone* $\{(x, y, z) \in \mathbb{R}^3 \mid x \geq 0,\ y \geq 0,\ |z| \leq x^{1/p} y^{1/q}\}$, where $\frac{1}{p} + \frac{1}{q} = 1$ and $p \in (2, +\infty)$.

8.3 Jordan algebras

Symmetric cones possess a rich algebraic structure and are closely linked to a class of non-associative algebras.

Definition 8.5. A *real algebra* A is a real vector space equipped with a bilinear multiplication $\bullet : A \times A \to A$. It is called *commutative* if $a \bullet b = b \bullet a$ for all $a, b \in A$, and *associative* if $(a \bullet b) \bullet c = a \bullet (b \bullet c)$ for all $a, b, c \in A$.

Definition 8.6. A real algebra A is called a *Jordan algebra* if it is commutative and satisfies the Jordan identity $(x \bullet x) \bullet (x \bullet y) = x \bullet ((x \bullet x) \bullet y)$ for all $x, y \in A$.

This means it is the same if we multiply y first by x and then by x^2 or first by x^2 and then by x.

A Jordan algebra is called *Euclidean* if $\sum_{k=1}^m x_k \bullet x_k = 0$ implies $x_1 = \cdots = x_m = 0$ for all $x_1, \ldots, x_m \in A$.

Then the symmetric cones are exactly the *cones of squares* of the Euclidean Jordan algebras.

Theorem 8.7. *Let $K \subset V$ be a symmetric cone. Then there exists a Euclidean Jordan algebra structure with multiplication $\bullet$ on V such that $K = \{x \bullet x \,|\, x \in V\}$.*

On the other hand, for every Euclidean Jordan algebra A the set $\{x \bullet x \,|\, x \in A\}$ is a symmetric cone, and non-isomorphic algebras produce non-isomorphic cones.

Euclidean Jordan algebras are equipped with a natural scalar product given by

$$\langle x, y\rangle = \mathrm{tr}\ L_{x \bullet y},$$

where $L_x : A \to A$ is the linear operator of multiplication by x, $L_x u := x \bullet u$.

Euclidean Jordan algebras always possess an identity element I.

Definition 8.8. An *idempotent* of an algebra A is an element $x \in A$ obeying $x^2 = x$.

Two idempotents x, y are *orthogonal* if $x \bullet y = 0$.

A *primitive* idempotent is an idempotent which can not be written as a non-trivial sum of orthogonal idempotents.

A *Jordan frame* if a set of orthogonal primitive idempotents which sum to the identity I.

The number of elements in the Jordan frame is always the same and defines the *rank* of the Euclidean Jordan algebra.

Theorem 8.9. *(Spectral decomposition) Every element $x \in A$ of a Euclidean Jordan algebra can in a unique manner be decomposed as a sum $x = \sum_i \lambda_i e_i$, where λ_i are distinct real numbers, the* eigenvalues *of x, and e_i are orthogonal idempotents which sum to the identity element.*

The *multiplicity* of the eigenvalue λ_i is the number of primitive idempotents which have to be summed to obtain the idempotent e_i.

The *rank* of x is the number of non-zero eigenvalues of x, counting their multiplicities.

The *determinant* of x is the product of its eigenvalues, counting their multiplicities.

As a function of x the determinant is a homogeneous polynomial of degree equal to the rank of the algebra.

Definition 8.10. Let $x \in A$ be an element of a Jordan algebra with identity I. An element $y \in A$ is called *inverse* of x, denoted by x^{-1}, if $x \bullet y = I$ and the operators L_x, L_y commute.

If such an element y exists, then x is called *invertible*.

In a Euclidean Jordan algebra an element x is invertible if and only if its determinant is non-zero. If $x = \sum_i \lambda_i e_i$ is the spectral decomposition of x, then the inverse is given by $x^{-1} = \sum_i \lambda_i^{-1} e_i$.

An element x is in the symmetric cone of the algebra A if and only if all its eigenvalues are nonnegative. The symmetric cone can hence also be seen as the cone of "positive semi-definite" elements of the algebra.

Attention: The same symmetric cone can be produced by different Euclidean Jordan algebras, but all these algebras will be isomorphic. The identity elements, scalar products, idempotents, inverses etc. will, however, depend on the chosen algebra. Every element in the interior of the cone may, e.g., take over the role of the identity element, in accordance with the homogeneity of the cone.

8.4 Classification

The symmetric cones have been fully classified. The classification is obtained via the equivalent classification of Euclidean Jordan algebras [6].

Theorem 8.11. *Every symmetric cone K is a direct product of a finite number of symmetric cones $K_1, \dots, K_m$, each of which is either a member of one of the following families, indexed by a natural number $n \geq 1$:*

- *Lorentz cone* $L_n = \{x \in \mathbb{R}^n \,|\, x_0 \geq \sqrt{\sum_{j=1}^{n-1} x_j^2}\}$,
- *real symmetric positive semi-definite matrix cone* $\mathcal{S}_+^n = \{A = A^T \,|\, x^T A x \geq 0 \ \forall\ x \in \mathbb{R}^n\}$,
- *complex Hermitian positive semi-definite matrix cone* $\mathcal{H}_+^n = \{A = A^* \,|\, x^* A x \geq 0 \ \forall\ x \in \mathbb{C}^n\}$,

- *quaternionic Hermitian positive semi-definite matrix cone* $\mathcal{Q}_+^n = \{A = A^* \mid x^*Ax \geq 0 \ \forall \ x \in \mathbb{H}^n\}$*;*

or the exceptional 27-dimensional Albert cone, which is the cone $\mathcal{O}_+^3$ *of octonionic Hermitian positive semi-definite* 3×3 *matrices.*

Every cone except $L_2 = L_1^2$ *which was listed above is symmetric and cannot be further decomposed in a non-trivial manner (i.e., it is* irreducible*).*

For $n = 1$ all families yield the 1-dimensional cone $\mathbb{R}_+$, and the matrix cones for $n = 2$ are all isomorphic to a Lorentz cone. Apart from these exceptions, the irreducible cones listed above are mutually non-isomorphic. Note that the orthant $\mathbb{R}_+^n$ is a direct product of n copies of the 1-dimensional cone and is hence also symmetric.

The Jordan multiplication for the Lorentz cone is given by

$$\begin{pmatrix} a \\ u \end{pmatrix} \bullet \begin{pmatrix} b \\ v \end{pmatrix} = \begin{pmatrix} ab + u^T v \\ bu + av \end{pmatrix}.$$

The Jordan multiplication for the matrix cones is given by

$$A \bullet B = \frac{AB + BA}{2},$$

where the product on the right-hand side is the ordinary matrix multiplication.

The Jordan multiplication for a product cone $K_1 \times K_2$ is given by $(x_1, x_2) \bullet (y_1, y_2) = (x_1 \bullet y_1, x_2 \bullet y_2)$, $x_i, y_i \in K_i$, $i = 1, 2$.

The Jordan multiplication for the orthant $\mathbb{R}_+^n$ is hence the element-wise multiplication of real vectors in $\mathbb{R}^n$.

If A is a real symmetric (complex Hermitian, quaternionic Hermitian) matrix, we shall write $A \succeq 0$ for $A \in \mathcal{S}_+^n$ ($A \in \mathcal{H}_+^n$, $A \in \mathcal{Q}_+^n$) and call A *positive semi-definite*, and we shall write $A \succ 0$ for $A \in \operatorname{int} \mathcal{S}_+^n$ ($A \in \operatorname{int} \mathcal{H}_+^n$, $A \in \operatorname{int} \mathcal{Q}_+^n$) and call A *positive definite*.

8.5 Facial structure

Definition 8.12. A convex subset F of a convex set X is called a *face* of X if for every line segment $l \subset X$ such that $F \cap ri\, l \neq \emptyset$ we have $l \subset F$.

A face F of X is called *proper* if $F \neq \emptyset$ and $F \neq X$.

A face F of a convex set X is called *exposed* if there exists a hyperplane H such that $F = X \cap H$ and $X \not\subset H$.

The faces of $\mathbb{R}_+^n$ are indexed by the 2^n subsets of indices $1, \dots, n$. A subset $I \subset \{1, \dots, n\}$ defines a face by

$$F_I = \{x \in \mathbb{R}_+^n \mid x_j = 0 \ \forall \ j \notin I\}.$$

So, the whole index set $I = \{1, \dots, n\}$ corresponds to the cone itself, and the set $I = \emptyset$ corresponds to the face $\{0\}$. The dimension of the face F_I equals $\#I$ (the cardinality of I), and it is isomorphic to the cone $\mathbb{R}_+^{\#I}$.

The face $F_{\bar{I}}$ with $\bar{I} = \{1, \dots, n\} \setminus I$ is the *complementary* face to F_I, it is the maximal face such that $F_I \bullet F_{\bar{I}} = 0$.

The non-zero faces of $\mathcal{S}_+^n$ are of the form

$$\mathcal{F}_H = \{F \cdot A \cdot F^T \mid A \in \mathcal{S}_+^k\},$$

where $k \in \{1, \dots, n\}$, H is a k-dimensional linear subspace of $\mathbb{R}^n$ (a point of the Grassmanian $Gr(k, \mathbb{R}^n)$), and F is a $n \times k$ matrix containing a basis of H as columns. Multiplication of F from the right by a non-degenerate matrix (an element of $GL_k(\mathbb{R})$) changes the basis, but not its column span, which equals the subspace H, and hence leads to the same face. The face $\mathcal{F}_H$ is isomorphic to the cone $\mathcal{S}_+^k$, and its dimension equals $\frac{k(k+1)}{2}$. The zero face can be seen as the face $\mathcal{F}_{\{0\}}$ belonging to the zero subspace.

The face $\mathcal{F}_{H^\perp}$ with $H^\perp$ the orthogonal complement to H is the complementary face to $\mathcal{F}_H$.

The Lorentz cone L_n has the improper faces $\{0\}$ and L_n. All proper faces are extreme rays of L_n, and every extreme ray is a proper face.

The face complementary to the face $\mathbb{R}_+ \cdot (1, u^T)^T$ is the face $\mathbb{R}_+ \cdot (1, -u^T)^T$, where $||u|| = 1$.

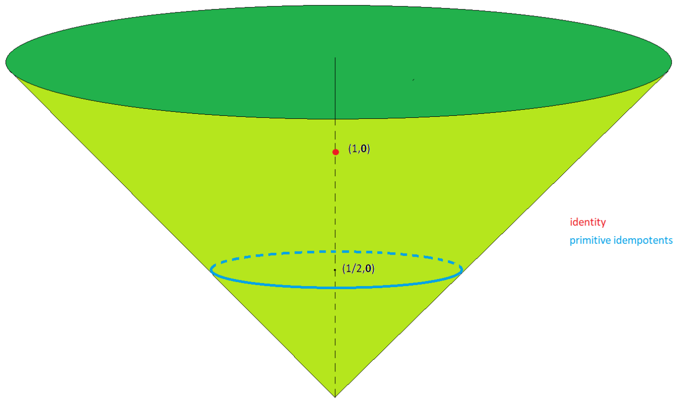

Figure 1: Identity and primitive idempotents in the algebra of the Lorentz cone

8.6 Summary

symmetric cone	$\mathcal{S}_+^n$	$\mathbb{R}_+^n$	L_n
dimension	$\frac{n(n+1)}{2}$	n	n
multiplication	$A \bullet B = \frac{AB+BA}{2}$	$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \bullet \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 y_1 \\ \vdots \\ x_n y_n \end{pmatrix}$	$\begin{pmatrix} a \\ u \end{pmatrix} \bullet \begin{pmatrix} b \\ v \end{pmatrix} = \begin{pmatrix} ab + u^T v \\ bu + av \end{pmatrix}$
identity	I_n	$\mathbf{1} = (1, \dots, 1)^T$	$(1, \mathbf{0})^T$
primitive idempotents	xx^T, $\|\|x\|\| = 1$	basis vectors e_k	$\frac{1}{2}(1, u^T)^T$, $\|\|u\|\| = 1$
rank	n	n	2
determinant	det A	$\prod_{j=1}^n x_j$	$\det \begin{pmatrix} a \\ u \end{pmatrix} = a^2 - \|\|u\|\|^2$
eigenvalues	$\lambda_j(A)$	x_j	$a \pm \|\|u\|\|$
inverse	A^{-1}	$(x_1^{-1}, \dots, x_n^{-1})^T$	$\begin{pmatrix} a \\ u \end{pmatrix}^{-1} = \frac{1}{a^2 - \|\|u\|\|^2} \begin{pmatrix} a \\ -u \end{pmatrix}$
proper rank k faces	$F \cdot \mathcal{S}_+^k \cdot F^T$, $F \in \mathbb{R}^{n \times k}/GL_k$	$\prod_j \mathbb{R}_+^{\sigma_j}$, $\sigma \in \{0,1\}^n$, $\mathbf{1}^T \sigma = k$	$k = 1$, boundary rays

An extensive treatment of symmetric cones can be found in [5].

8.7 Universality of the real symmetric semi-definite cone

All symmetric cones can be represented as intersections of a cone of positive semi-definite real symmetric matrices with a linear subspace.

We have $(x_0, x_1, \dots, x_{n-1})^T \in L_n$ if and only if

$$\begin{pmatrix} x_0 + x_1 & x_2 & \cdots & x_{n-1} \\ x_2 & x_0 - x_1 & \mathbf{0} & 0 \\ \vdots & \mathbf{0} & \ddots & \mathbf{0} \\ x_{n-1} & 0 & \mathbf{0} & x_0 - x_1 \end{pmatrix} \succeq 0.$$

We have $x \in \mathbb{R}_+^n$ if and only if $\operatorname{diag}(x) \succeq 0$.

For a complex Hermitian matrix $A = S + iK$, where S is real symmetric and K is real skew-symmetric, we have $A \succeq 0$ if and only if

$$\begin{pmatrix} S & K \\ -K & S \end{pmatrix} \succeq 0.$$

For a quaternionic Hermitian matrix $A = S + iK + jL + kM$, where S is real symmetric and K, L, M are real skew-symmetric, we have $A \succeq 0$ if and only if

$$\begin{pmatrix} S & K & L & M \\ -K & S & -M & L \\ -L & M & S & -K \\ -M & -L & K & S \end{pmatrix} \succeq 0.$$

Let $K = K_1 \times K_2$ be a product cone, and suppose $A_1 \in K_1$ if and only if $\mathcal{L}_1(A_1) \succeq 0$, and $A_2 \in K_2$ if and only if $\mathcal{L}_2(A_2) \succeq 0$, where $\mathcal{L}_1$, $\mathcal{L}_2$ are linear matrix-valued maps encoding the representations of K_1, K_2 as intersections of linear subspaces with real symmetric matrix cones. Construct the linear matrix valued map $\mathcal{L}$ on the product $A_1 \times A_2$ by $\mathcal{L}(A_1, A_2) = \mathrm{diag}(\mathcal{L}_1(A_1), \mathcal{L}_2(A_2))$. Then we have $(A_1, A_2) \in K$ if and only if $\mathcal{L}(A_1, A_2) \succeq 0$, and $\mathcal{L}$ represents the product cone K as intersection of a real symmetric matrix cone with a linear subspace.

Definition 8.13. Intersections of linear subspaces with the cone of positive semi-definite real symmetric matrices are called *spectrahedral* cones.

Definition 8.14. A cone which can be represented as a linear image of a spectrahedral cone is called *semi-definite representable.*

The above results can be rephrased as: every symmetric cone is a spectrahedral cone. Actually, even every homogeneous cone is a spectrahedral cone [2].

9 Conic programs

In this section we consider a large class of convex optimization problems, the *conic programs*. Actually, every convex optimization problem can be formulated as a conic program. However, this form is more convenient for certain problems and less convenient for others.

9.1 Formulation

Definition 9.1. A *conic program* over a regular convex cone K is an optimization problem of the form

$$\inf_{x \in K} c^T x \ : \quad Ax = b. \tag{1}$$

Here $x \in K$ is the *conic constraint*, $Ax = b$ is a linear constraint, and the objective function is linear. The feasible set of a conic program is hence an intersection of the cone with an affine subspace. As the intersection of convex sets it is convex. The objective function is convex too, and hence a conic program formally belongs to the class of convex problems. The complexity of a conic program is encoded in the cone K, and whether the conic program is efficiently solvable depends on which descriptions of the cone K are available.

Since every closed convex set can be written as an intersection of an affine subspace with a regular convex cone, and minimizing a convex function is equivalent to minimizing a linear function over the epigraph of the original function, every convex problem can actually be rewritten as a conic program.

We may write a conic program also in the form

$$\min_x c^T x : \quad Ax + b \in K.$$

Here the decision variable x directly parameterizes the affine hull of the feasible set.

9.2 Equivalence of conic programs over different cones

Conic solvers are able to handle problems over a restricted collection of standard cones and their direct products. Usually these are the nonnegative orthants $\mathbb{R}^n_+$ (for LP solvers), plus the Lorentz cones L_n (for SOCP solvers), plus the matrix cones $\mathcal{S}^n_+$ (for SDP solvers). Some solvers are able to handle the matrix cones $\mathcal{H}^n_+$ too. In recent years solvers appear that are able to solve conic programs over a very limited number of nonsymmetric cones, namely the exponential cone $K_{\exp}$ and the power cones.

It is therefore necessary to convert the optimization problem to solve into a standard form which can be handled by a solver. Sometimes this is a non-trivial task, and the corresponding representability results employ advanced mathematics. Here we consider a more common and simple case to convert a conic problem over a cone K into a conic problem over another cone K', where K is a linear section or a linear projection of K'.

K is a section of K': Let $V \subset V'$ be a linear subspace of a real vector space, $K' \subset V'$ be a regular convex cone, such that $K = K' \cap V$. Let a conic program of the form (1) over K be given. Then this problem is equivalent to the conic program

$$\min_{y \in K'} \langle c', y \rangle : \qquad A'y = b, \quad y \in V$$

over the cone K'. Here c' is an arbitrary extension of the linear functional c, and A' is an arbitrary extension of the linear operator A from V to V'. This means that c' and A' coincide with c and A on $V \subset V'$, respectively. Note that the condition $y \in V$ is a linear equality constraint, so the problem is indeed a conic program over K'.

Usually this type of equivalence is already implicitly assumed in the formulation by considering x as an element of V' instead an element of V.

K is a projection of K': This type of equivalence is called *lifting* and is sometimes not so trivial to find. Let $K \subset V$ be a regular convex cone. Let further $\Pi : V' \to V$ be a linear map and $K' \subset V'$ be a regular convex cone such that $\Pi[K'] = K$. Then (1) is equivalent to the program

$$\min_{y \in K'} \langle \Pi^\dagger c, y \rangle : \qquad A\Pi(y) = b,$$

where $c' = \Pi^\dagger c$ is the linear functional on V' defined by $\langle c', y \rangle := \langle c, \Pi(y) \rangle$.

Often a lifting is accomplished by adding variables $z \in V''$ in the problem formulation such that $V' = V \otimes V''$, where V'' is yet another real vector space. Then $y := (x, z)$, $\Pi(y) := x$, and the problem takes the form

$$\min_{(x,z) \in K'} \langle c, x \rangle : \qquad Ax = b.$$

In Section (8.7) it was demonstrated that all symmetric cone programs can be reduced to SDPs.

9.3 Examples

Linear programs: If the conic constraints involve orthants $\mathbb{R}^n_+$ only, then the resulting conic program is a Linear Program (LP).

Second order cone programs: If the conic constraints involve only products $L_{n_1} \times \cdots \times L_{n_m}$ of Lorentz cones, then the resulting conic program is a Second Order Cone Program (SOCP).

An LP is also an SOCP, because $\mathbb{R}^n_+ = L_1^n$, but not every SOCP can be cast as an LP. An SOCP can, however, be approximated by LPs with a polynomial increase in complexity [1].

Semi-definite programming: If the conic constraints involve matrix cones only, then the resulting conic program is a Semi-Definite Program (SDP). Any LP or SOCP, and more generally every symmetric cone program, can be written as an SDP by the universality property of the semi-definite matrix cone. However, it is not advisable to treat LPs and SOCPs as SDPs due to their better structure and the availability of dedicated solution methods. SDPs cannot in general be approximated by LPs with a polynomial increase in complexity only.

Copositive programs: Let $\mathcal{C}^n$ be the cone of real symmetric $n \times n$ matrices A such that $x^T Ax \geq 0$ for every $x \in \mathbb{R}^n_+$. This cone is called the *copositive cone.* Its difference with the positive semi-definite cone is that we require A to be nonnegative only on *nonnegative* vectors. The cone $\mathcal{C}^n$ is regular, but to decide whether a given point is not in $\mathcal{C}^n$ is NP-complete [8]. Efficient methods for solving generic programs over the copositive cone are hence unlikely to exist.

Geometric programs: A geometric program (GP) aims to minimize an objective function $f_0(x)$ over the positive orthant $\mathbb{R}^n_{++}$ with respect to inequality constraints $f_i(x) \leq 1$ and equality constraints $h_j(x) = 1$. Here the f_i are *posynomials*, i.e., of the form $\sum_k c_k \prod_l x_l^{\alpha_{kl}}$ with $c_k > 0$, and the h_j are monomials, i.e., of the form $c\prod_l x_l^{\alpha_l}$ with $c > 0$. The problem can be formalized as

$$\min_{x\in\mathbb{R}^n_{++}} \sum_k c_{0k} \prod_l x_l^{\alpha_{0kl}} \;:\quad \sum_k c_{ik} \prod_l x_l^{\alpha_{ikl}} \leq 1,\ i = 1,\dots,m; \quad c_j \prod_l x_l^{\alpha_{jl}} = 1,\ j = 1,\dots,m'.$$

Introducing the variable change $y = \log x$ and $b = \log c$, the problem can be rewritten as

$$\min_{y\in\mathbb{R}^n} \sum_k e^{b_{0k}+\sum_l \alpha_{0kl}y_l} \;:\quad \sum_k e^{b_{ik}+\sum_l \alpha_{ikl}y_l} \leq 1,\ i = 1,\dots,m; \quad b_j + \sum_l \alpha_{jl}y_l = 0,\ j = 1,\dots,m'.$$

Introducing additional variables and defining the 3-dimensional convex cone

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

the problem can be rewritten as

$$\min_{y\in\mathbb{R}^n} \sum_k d_{0k} \;:\quad (b_{ik} + \sum_l \alpha_{ikl}y_l, d_{ik}, 1) \in K_{\exp}\ \forall\, k,\ i = 0,\dots,m; \quad \sum_k d_{ik} = 1,\ i = 1,\dots,m;$$

$$b_j + \sum_l \alpha_{jl}y_l = 0,\ j = 1,\dots,m'.$$

The problem now has a linear objective function and involves only linear equality constraints and conic constraints described by the cone $K_{\exp}$.

The cone $K_{\exp}$ is the closure of the homogenization of the epigraph of the exponential function. It is self-dual in the wider sense that it is isomorphic to $K^*_{\exp}$, but it is not homogeneous.

9.4 Duality

To any conic program over a cone K one can define a *dual* program over the dual cone K^*. To contrast it with the dual problem, the original problem is called the *primal*. In order to obtain the dual program we dualize the conic constraint by means of a variable $\lambda \in K^*$, as follows:

$$\left(\min_{x\in K} c^Tx :\ Ax = b\right) = \left(\min_x \max_{\lambda\in K^*} (c^Tx - \lambda^Tx) :\ Ax = b\right) \geq$$

$$\geq \left(\max_{\lambda\in K^*} \min_x (c^Tx - \lambda^Tx) :\ Ax = b\right) = \left(\max_{\lambda\in K^*} s^Tb :\ c - \lambda = A^Ts\right) = \left(\max_s s^Tb :\ c - A^Ts \in K^*\right).$$

In the third step we used that the minimum $\min_x(c^Tx - \lambda^Tx)$: $Ax = b$ is finite if and only if the linear functional $c - \lambda$ is constant on the affine subspace defined by the equations $Ax = b$. This happens if and only if $c - \lambda$ can be expressed as a product A^Ts for some vector s.

The dual program is hence a *maximization* problem, and every feasible point for the dual problem yields a *lower bound* on the objective value of the primal problem. Vice versa, every feasible point for the primal problem yields an *upper bound* on the objective value of the dual problem.

Similarly to the case of LPs we may write the pair of conic programs in the following more symmetric form. Let $L, L^\perp$ be complementary linear subspaces of the primal and the dual vector space, respectively. Then the primal-dual pair can be written as

$$\min(c^Tx + \mathit{const}) :\quad x \in K \cap (L + b),$$

$$\max(-b^T s + const): \quad s \in K^* \cap (L^\perp + c).$$

The vectors b, c are not necessarily equal to those in the original problem formulation.

Primal-dual optimization methods treat both problems simultaneously and symmetrically, producing a sequence of primal-dual pairs of iterates whose objective values approach each other.

The optimal values of the primal and dual programs, even if they both exist, do not need to coincide. In this case one speaks of a *duality gap*.

9.5 Second-order cone programs

Linear programs can be used to solve problems with linear constraints and a linear cost function. The descriptive power of second order cone programs (SOCP) is much higher. Below are examples of constraints which can be reformulated as second order cone constraints.

- $||x||_2^2 \leq t$, $x \in \mathbb{R}^n$ is equivalent to $(\frac{t+1}{2}, \frac{t-1}{2}, x) \in L_{n+2}$;
- $\frac{||x||_2^2}{s} \leq t$, $s \geq 0$, $x \in \mathbb{R}^n$ can be expressed as $(\frac{t+s}{2}, \frac{t-s}{2}, x) \in L_{n+2}$;
- $ts \geq 1$, $t, s > 0$ is equivalent to $(\frac{t+s}{2}, \frac{t-s}{2}, 1) \in L_3$;
- $x^T Ax + b^T x + c \leq t$ for $A \succeq 0$ is equivalent to $(\frac{t-b^T x-c+1}{2}, \frac{t-b^T x-c-1}{2}, A^{1/2}x) \in L_{n+2}$;
- $|t| \leq \sqrt{x_1 x_2}$ and $x_1, x_2 \geq 0$ is equivalent to $(\frac{x_1+x_2}{2}, \frac{x_1-x_2}{2}, t) \in L_3$;
- $t \leq \sqrt{x_1 x_2}$ and $x_1, x_2 \geq 0$ is equivalent to $t \leq s$, $s \geq 0$, $(\frac{x_1+x_2}{2}, \frac{x_1-x_2}{2}, s) \in L_3$.

Example: Minimization of a convex quadratic function over an intersection of ellipsoids. Consider the problem

$$\min_{x \in \mathbb{R}^n} (x^T A_0 x + b_0^T x + c_0) \; : \qquad x^T A_i x + b_i^T x + c_i \leq 0, \; i = 1, \dots, m.$$

Here $A_0, \dots, A_m$ are supposed to be positive semi-definite. This problem can be reformulated as the SOCP

$$\min_{x,t} t \; : \qquad \left(\frac{t - b_0^T x - c_0 + 1}{2}, \frac{t - b_0^T x - c_0 - 1}{2}, A_0^{1/2} x \right) \in L_{n+2},$$

$$\left(\frac{1 - b_i^T x - c_i}{2}, \frac{1 + b_i^T x + c_i}{2}, A_i^{1/2} x \right) \in L_{n+2}, \; i = 1, \dots, m.$$

9.6 Semi-definite programs

The descriptive power of SDPs is even higher than that of SOCPs. Below are examples of constraints which can be reformulated as semi-definite constraints.

- $\lambda_{\max}(X) \leq t$ is equivalent to $tI - X \succeq 0$;
- $||X||_\infty \leq t$ is equivalent to $-tI \preceq X \preceq tI$ for symmetric matrices;
- $\sum_{j=1}^k \lambda_j \leq t$, where $\lambda_1, \dots, \lambda_n$ are the ordered eigenvalues of X, is equivalent to $t \geq ks + \text{tr } Z$, $Z \succeq 0$, $Z + sI \succeq X$;
- $A \succeq BC^\dagger B^T$, $C \succeq 0$, $\ker C \subset \ker B$ is equivalent to $\begin{pmatrix} A & B \\ B^T & C \end{pmatrix} \succeq 0$;
- $||A||_\infty \leq t$ is equivalent to $\begin{pmatrix} tI & A \\ A^T & tI \end{pmatrix} \succeq 0$ for rectangular matrices;

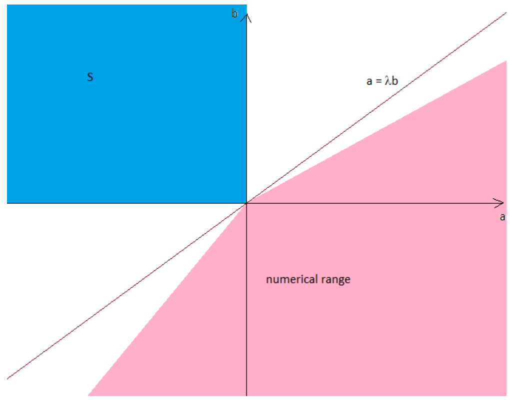

Figure 2: Separation of S from the numerical range

- $(AXB)(AXB)^T + CXD + (CXD)^T + E \preceq Y$ is equivalent to $\begin{pmatrix} I & (AXB)^T \\ AXB & Y - E - CXD - (CXD)^T \end{pmatrix} \succeq 0$ (here X, Y are the design variables and $A, \ldots, E$ are the data of the problem);
- $x^T Ax \geq 0$ for all x such that $x^T Bx \geq 0$ and there exists x_0 such that $x_0^T Bx_0 > 0$ is equivalent to $A - \lambda B \succeq 0$ and $\lambda \geq 0$ (S-lemma).

In the last case A is the original design variable.

The proof of the S-lemma relies on the following theorem by Dines [3].

Theorem 9.2. *Let $A, B \in \mathcal{S}^n$ be real symmetric matrices. Then the set $\{(x^T Ax, x^T Bx) \in \mathbb{R}^2 \mid x \in \mathbb{R}^n\}$ is convex.*

The set $\{(x^T Ax, x^T Bx) \in \mathbb{R}^2 \mid x \in \mathbb{R}^n\}$ is called the *numerical range* of the pair (A, B). The theorem then says that the numerical range is a convex cone.

Lemma 9.3. *Let $A, B \in \mathcal{S}^n$ such that there exists a vector $\hat{x} \in \mathbb{R}^n$ such that $\hat{x}^T B\hat{x} > 0$. Then the following assertions are equivalent:*

- *$x^T Ax \geq 0$ for all x such that $x^T Bx \geq 0$,*
- *there exists $\lambda \geq 0$ such that $A - \lambda B \succeq 0$.*

Proof. If $A - \lambda B \succeq 0$ for some $\lambda \geq 0$, then $x^T Ax \geq \lambda x^T Bx$ for all $x \in \mathbb{R}^n$, and the first assertion is evident.

Let us assume that $x^T Ax \geq 0$ for all x such that $x^T Bx \geq 0$. Then the set $S = \{(a, b) \mid a < 0,\ b \geq 0\}$ has an empty intersection with the numerical range of (A, B). By convexity of the numerical range there exists a 1-dimensional linear subspace of $\mathbb{R}^2$ which separates it from S. Since there exists a point $(\hat{a}, \hat{b})$ in the numerical range with $\hat{a} \geq 0$, $\hat{b} > 0$, this subspace cannot be the abscissa. It is therefore given by the equation $a = \lambda b$ for some $\lambda \geq 0$. But then $a \geq \lambda b$ for all points (a, b) in the numerical range, which is equivalent to the condition $A - \lambda B \succeq 0$. □

More semi-definite representable constraints can be found in the lectures of A. Ben-Tal and A. Nemirovski (`https://www2.isye.gatech.edu/~nemirovs/Lect_ModConvOpt.pdf`)

References

[1] Aharon Ben-Tal and Arkadi Nemirovski. On polyhedral approximations of the second-order cone. *Math. Oper. Res.*, 26(2):193–205, 1999.

[2] Chek Beng Chua. Relating homogeneous cones and positive definite cones via T-algebras. *SIAM J. Optim.*, 14(2):500–506, 2003.

[3] Lloyd L. Dines. On the mapping of quadratic forms. *Bull. Amer. Math. Soc.*, 47(6):494–498, 1941.

[4] Josef Dorfmeister. Inductive construction of homogeneous cones. *Trans. Amer. Math. Soc.*, 252:321–349, 1979.

[5] Jacques Faraut and Adam Koranyi. *Analysis on symmetric cones.* Oxford University Press, London, New York, 1994.

[6] Pascual Jordan, John von Neumann, and Eugene Wigner. On an algebraic generalization of the quantum mechanical formalism. *Ann. Math.*, 35(1):29–64, 1934.

[7] Soji Kaneyuki and Tadashi Tsuji. Classification of homogeneous bounded domains of lower dimension. *Nagoya Math. J.*, 53:1–46, 1974.

[8] Katta G. Murty and Santosh N. Kabadi. Some NP-complete problems in quadratic and nonlinear programming. *Math. Program.*, 39:117–129, 1987.

[9] Ernest Borisovich Vinberg. The theory of convex homogeneous cones. *Trans. Moscow Math. Soc.*, 12:340–403, 1963.