3 The positive semidefinite cone - Conic programming

In this course we will focus a lot of our attention on the *positive semidefinite cone*. Let $\mathbf{S}^n$ denote the vector space of $n \times n$ real symmetric matrices. Recall that by the spectral theorem any matrix $A \in \mathbf{S}^n$ is diagonalisable in an orthonormal basis and has real eigenvalues. Let $\mathbf{S}^n_+$ (resp. $\mathbf{S}^n_{++}$) denote the set of positive semidefinite matrices, i.e., the set of real symmetric matrices having nonnegative (resp. strictly positive) eigenvalues. For a matrix $A \in \mathbf{S}^n_+$ we will use the following convenient notations:

$$A \succeq 0 \iff A \text{ positive semidefinite}$$

and

$$A \succ 0 \iff A \text{ positive definite.}$$

Proposition 3.1. *Let $A \in \mathbf{S}^n$. The following conditions are equivalent:*

(i) $A \in \mathbf{S}^n_+$

(ii) The eigenvalues of A are nonnegative

(iii) $x^T A x \geq 0$ *for all* $x \in \mathbb{R}^n$

(iv) There exists $L \in \mathbb{R}^{n\times n}$ lower triangular such that $A = LL^T$ (Cholesky factorization)

(v) All the principal minors of A are nonnegative, i.e., $\det A[S,S] \geq 0$ *for any nonempty $S \subseteq \{1,\ldots,n\}$ where $A[S,S]$ is the submatrix of A consisting of the rows and columns indexed by S (Sylvester criterion)*

Also the following are all equivalent:

(i) $A \in \mathbf{S}^n_{++}$

(ii) The eigenvalues of A are strictly positive

(iii) $x^T A x > 0$ *for all* $x \in \mathbb{R}^n \setminus \{0\}$

(iv) There exists $L \in \mathbb{R}^{n\times n}$ lower triangular with $L_{ii} > 0$ for all $i = 1,\ldots,n$ such that $A = LL^T$ (Cholesky factorization)

(v) All the leading principal minors of A are positive, i.e., $\det A[\{1,\ldots,k\},\{1,\ldots,k\}] > 0$ *for all $k = 1,\ldots,n$ (Sylvester criterion)*

Proof. We leave it as an exercise to the reader. □

Theorem 3.1. $\mathbf{S}^n_+$ *is a closed pointed convex cone in* $\mathbf{S}^n$ *with interior* $\mathbf{S}^n_{++}$.

Proof. That $\mathbf{S}^n_+$ is closed and convex follows from

$$\mathbf{S}^n_+ = \{A \in \mathbf{S}^n : x^T A x \geq 0 \;\forall x \in \mathbb{R}^n\} = \bigcap_{x\in\mathbb{R}^n} \underbrace{\{A \in \mathbf{S}^n : x^T A x \geq 0\}}_{H_x}.$$

H_x is a closed halfspace in $\mathbf{S}^n$ for any fixed x thus $\mathbf{S}^n_+$ is closed and convex as an intersection of closed halfspaces. To show that $\mathbf{S}^n_+$ is pointed we need to show that $(\mathbf{S}^n_+) \cap (-\mathbf{S}^n_+) = \{0\}$. This is easy to see because if $A \in \mathbf{S}^n_+ \cap (-\mathbf{S}^n_+)$ then all the eigenvalues of A must be equal to zero which means that $A = 0$. It remains to show that $\operatorname{int}(\mathbf{S}^n_+) = \mathbf{S}^n_{++}$. To do so we first define the *spectral norm* of a matrix $A \in \mathbf{S}^n$ as:

$$\|A\| = \max_{x \in \mathbb{R}^n : \|x\|_2 = 1} \|Ax\|_2 = \max\{-\lambda_{\min}(A), \lambda_{\max}(A)\}.$$

Note that this is the $\ell_2 \to \ell_2$ induced norm. We now show that $\operatorname{int}(\mathbf{S}^n_+) = \mathbf{S}^n_{++}$.

- We first show the inclusion $\operatorname{int}(\mathbf{S}^n_+) \subseteq \mathbf{S}^n_{++}$. If $A \in \operatorname{int}(\mathbf{S}^n_+)$ then there exists small enough $\epsilon > 0$ such that $\|A - X\| \le \epsilon \Rightarrow X \in \mathbf{S}^n_+$. Let $X = A - \epsilon I$ where I is the $n \times n$ identity matrix, and note that $\|A - X\| = \|\epsilon I\| \le \epsilon$. It thus follows that $X = A - \epsilon I \in \mathbf{S}^n_+$. Since the eigenvalues of $A - \epsilon I$ are the $(\lambda_i - \epsilon)$ (where (λ_i) are the eigenvalues of A) it follows that $\lambda_i \ge \epsilon > 0$ and thus A is positive definite, i.e., $A \in \mathbf{S}^n_{++}$.

- We now prove the reverse inclusion $\mathbf{S}^n_{++} \subseteq \operatorname{int}(\mathbf{S}^n_+)$. Let $A \in \mathbf{S}^n_{++}$. Let $\lambda_{\min} > 0$ be the smallest eigenvalue of A and define the spectral norm ball $B = \{M \in \mathbf{S}^n : \|M - A\| \le \lambda_{\min}/2\}$. We will show that the ball B is included in $\mathbf{S}^n_+$ which will establish our claim. Let M such that $\|M - A\| \le \lambda_{\min}/2$. Then this means that for any x with $\|x\| = 1$, $x^T(A - M)x \le \lambda_{\min}/2$ and so $x^T M x \ge x^T A x - \lambda_{\min}/2 \ge \lambda_{\min}/2 > 0$. We have shown that $x^T M x \ge 0$ for any unit normed x thus M is positive semidefinite. This completes the proof.

□

The real vector space $\mathbf{S}^n$ has dimension $\binom{n+1}{2}$. We equip this vector space with the (trace) inner product

$$\langle A, B \rangle := \operatorname{Tr}[AB] = \sum_{1 \le i,j \le n} A_{ij} B_{ij}.$$

With this inner product the cone $\mathbf{S}^n_+$ is self-dual, meaning that $(\mathbf{S}^n_+)^* = \mathbf{S}^n_+$.

Theorem 3.2. *With the trace inner product on $\mathbf{S}^n$ we have $(\mathbf{S}^n_+)^* = \mathbf{S}^n_+$.*

Proof. By definition $(\mathbf{S}^n_+)^* = \{B \in \mathbf{S}^n : \operatorname{Tr}(AB) \ge 0 \ \forall A \in \mathbf{S}^n_+\}$. We first show that $\mathbf{S}^n_+ \subseteq (\mathbf{S}^n_+)^*$. Assume B is positive semidefinite. The eigenvalue decomposition of B takes the form $B = \sum_{i=1}^n \lambda_i v_i v_i^T$ where $\lambda_i \ge 0$ for $i = 1, \dots, n$ and the v_i are the unit-normed eigenvectors of B. Now for any $A \in \mathbf{S}^n_+$ we have $\operatorname{Tr}(AB) = \sum_{i=1}^n \lambda_i \operatorname{Tr}(A v_i v_i^T) = \sum_{i=1}^n \lambda_i v_i^T A v_i$. Since $A \in \mathbf{S}^n_+$ we have $v_i^T A v_i \ge 0$ for all $i = 1, \dots, n$ and thus, since $\lambda_i \ge 0$ we get $\operatorname{Tr}(AB) \ge 0$. This shows $\mathbf{S}^n_+ \subseteq (\mathbf{S}^n_+)^*$.

To show the reverse inclusion, assume $B \in \mathbf{S}^n$ is such that $\operatorname{Tr}(AB) \ge 0$ for all $A \in \mathbf{S}^n_+$. We want to show that B is positive semidefinite. By taking $A = xx^T$ for any $x \in \mathbb{R}^n$ we get that $\operatorname{Tr}(xx^T B) = x^T B x \ge 0$. This is true for all $x \in \mathbb{R}^n$ and thus shows that B is positive semidefinite. □

Theorem 3.3. *The extreme rays of $\mathbf{S}^n_+$ are the rays spanned by rank-one matrices, i.e., of the form $S_x = \{\lambda xx^T, \lambda \ge 0\}$ where $x \in \mathbb{R}^n$.*

Proof. We first show that any ray spanned by a matrix of the form xx^T is extreme for $\mathbf{S}^n_+$. Then we will show that these are the only ones.

- Assume $A, B \in \mathbf{S}_+^n$ are such that $A + B = \lambda x x^T$ for some $\lambda \geq 0$. We need to show that A and B are both a multiple of xx^T. Let u be any vector orthogonal to x, i.e., $u^T x = 0$. Then $0 \leq u^T A u \leq u^T (A+B) u = u^T (\lambda x x^T) u = 0$. Thus for any $u \in \{x\}^\perp$ we have $u^T A u = 0$ which implies, since $A \succeq 0$, $u \in \ker(A)$. Since $\operatorname{im}(A) = \ker(A)^\perp$ for any symmetric matrix A we get $\operatorname{im}(A) = \ker(A)^\perp \subseteq \operatorname{span}(x)$. This means that A is of the form $A = \lambda x x^T$. One can show in a similar way that B is a nonnegative multiple of xx^T.

- We now show that these are the only extreme rays. Consider a ray $S = \{\lambda A : \lambda \geq 0\}$ spanned by some matrix $A \in \mathbf{S}_+^n$. If $\operatorname{rank}(A) \geq 2$, an eigenvalue decomposition allows us to express A as a sum of elements that are not in S which shows that S cannot be an extreme ray.

$\square$

Conic programming

Definition 3.1. A cone $K \subseteq \mathbb{R}^n$ is called *proper* if it is closed, convex, pointed and has nonempty interior.

From Theorem 2.3 and Question 9 in the first exercise sheet, we know that the dual of a proper cone is also proper. The positive semidefinite cone is an example of a proper cone.

Let $K \subseteq \mathbb{R}^n$ be a proper cone. A *conic program* over K is an optimisation problem of the form:

$$\begin{array}{ll} \text{minimise} & \langle c, x \rangle \\ \text{subject to} & Ax = b \\ & x \in K \end{array} \tag{1}$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$. The optimisation variable here is $x \in \mathbb{R}^n$. The *feasible set* is the set of $x \in \mathbb{R}^n$ that satisfy the constraints $x \in K$ and $Ax = b$. The feasible set is the intersection of the cone K with an affine space $\{x \in \mathbb{R}^n : Ax = b\}$ and thus is a closed convex set as an intersection of closed convex sets.

Linear programming

A *linear program* is a conic program over the cone $K = \mathbb{R}_+^n$ (nonnegative orthant). The constraint $x \in \mathbb{R}_+^n$ means that $x_i \geq 0$ for $i = 1, \ldots, n$. We will often use the abbreviation $x \geq 0$ to denote that $x_i \geq 0$ for $i = 1, \ldots, n$ (here $x \in \mathbb{R}^n$). So a linear program is a problem of the form:

$$\begin{array}{ll} \text{minimise} & \langle c, x \rangle \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array} \tag{2}$$

For example the following optimisation problem is a linear program:

$$\text{minimise} \quad 2x_1 + x_2 \quad \text{s.t.} \quad 3x_1 - x_2 = 1, \;\; x_1 \geq 0, \;\; x_2 \geq 0.$$

This optimisation problem is an instance of (1) where $K = \mathbb{R}_+^2$, the cost vector is $c = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, the matrix A is 1×2 given by $A = \begin{bmatrix} 3 & -1 \end{bmatrix}$ and $b = 1$.

Despite their apparent simplicity, linear programs have applications in many areas of applied sciences, engineering and economics. What makes linear programming appealing is that there are efficient algorithms to solve such optimisation problems. Problems with thousands (even millions) of constraints can be easily solved on a personal computer using current algorithms.

Note: Historically, linear programming appeared in 1940s, much earlier than conic programs. Conic programs were introduced in 1990s as a generalisation of linear programming and were shown to enjoy some of the nice theoretical (and sometimes computational) properties of linear programming. For more historical information, see the bibliography section of Chapter 4 in Boyd & Vandenberghe.

Any linear program can be put in the following form, known as "*inequality form*" (the form (2) is known as "*standard form*"):

$$\begin{array}{ll} \underset{z\in\mathbb{R}^k}{\text{minimise}} & \langle e, z\rangle \\ \text{subject to} & Fz + g \geq 0 \end{array} \tag{3}$$

where $e \in \mathbb{R}^k$, $F \in \mathbb{R}^{n\times k}$ and $g \in \mathbb{R}^n$. To go from (2) to (3), let g be a point in the affine subspace $\{x \in \mathbb{R}^n : Ax = b\}$ and let F be a matrix whose columns form a basis of $\ker(A)$. Then $\{x \in \mathbb{R}^n : Ax = b\} = \{Fz + g : z \in \mathbb{R}^k\}$ (where $k = \dim\ker(A)$). Problem (2) is thus equivalent to:

$$\underset{z\in\mathbb{R}^k}{\text{minimise}} \quad \langle c, Fz + g\rangle \quad \text{s.t.} \quad Fz + g \geq 0. \tag{4}$$

If we let $e = F^T c$ and note that $\langle c, g\rangle$ is a constant, we see that (4) is equivalent to (3).

The feasible set of (3) is a *polyhedron*, i.e., an intersection of a finite number of halfspaces; the halfspaces here are defined by $f_i^T z + g_i \geq 0$ $(i = 1, \dots, n)$ where $f_1^T, \dots, f_n^T$ are the rows of F. Thus we see that geometrically, linear programming is the problem of optimizing a linear function over a polyhedron.

An example from signal processing We now give a simple example of a linear program that has attracted a lot of attention in the signal processing community. Let $M \in \mathbb{R}^{m\times n}$ and $d \in \mathbb{R}^m$ with $m < n$. We are interested in finding a solution to $Mx = d$ that has the smallest ℓ_1 norm. Recall that the ℓ_1 norm of a vector is given by $\|x\|_1 = \sum_{i=1}^n |x_i|$. In other words, we want to solve the optimisation problem

$$\underset{x\in\mathbb{R}^n}{\text{minimise}} \quad \|x\|_1 \quad \text{s.t.} \quad Mx = d. \tag{5}$$

Problem (5), as written, is *not* a linear program since the cost function is not linear. We will see however that by adequately introducing new variables we can express it as a linear program. We first claim that (5) is "equivalent" to the following problem:

$$\underset{x,y\in\mathbb{R}^n}{\text{minimise}} \quad \sum_{i=1}^n y_i \quad \text{s.t. } Mx = d, y + x \geq 0, y - x \geq 0. \tag{6}$$

What we will show is that any solution to (5) can be converted to a solution of (6) and vice-versa.

Claim 1. *If $x \in \mathbb{R}^n$ satisfies $Mx = d$ then there is $y \in \mathbb{R}^n$ that satisfies the constraints of (6) and such that $\sum_{i=1}^n y_i \leq \|x\|_1$. Conversely if $x, y \in \mathbb{R}^n$ satisfy the constraints of (6) then $\|x\|_1 \leq \sum_{i=1}^n y_i$. As a consequence the optimal values of (5) and (6) are equal.*

Proof. For the first direction take $y_i = |x_i|$ and note that $y_i + x_i \geq 0$ and $y_i - x_i \geq 0$ and $\sum_{i=1}^n y_i = \|x\|_1$. For the other direction note that if $x, y \in \mathbb{R}^n$ satisfy the constraints of (6) then $|x_i| = \max(x_i, -x_i) \leq y_i$ and so $\|x\|_1 \leq \sum_{i=1}^n y_i$. $\square$

Problem (6) is now much closer to being a linear program in the form (2), however it is not yet exactly in the form (2). We now show how to put it exactly in the form (2). If we define $u = y + x$ and $v = y - x$ then problem (6) can be rewritten as

$$\underset{u,v,x,y \in \mathbb{R}^n}{\text{minimise}} \quad \sum_{i=1}^{n} y_i \quad \text{s.t.} \begin{cases} Mx = d \\ u = y + x \\ v = y - x \\ u \geq 0, \ v \geq 0 \end{cases} \tag{7}$$

Problem (7) is almost of the form (2) except for a small difference: in (2) the variables are all constrained to be nonnegative, whereas in (7) only the variables u, v are nonnegative (and x and y can take arbitrary signs). However one can actually eliminate the variables x, y from the linear program (7) since $x = (u - v)/2$ and $y = (u + v)/2$. Problem (7) is thus equivalent to:

$$\underset{u,v \in \mathbb{R}^n}{\text{minimise}} \quad \frac{1}{2} \sum_{i=1}^{n} (u_i + v_i) \quad \text{s.t.} \quad M(u - v) = 2d, \ u \geq 0, \ v \geq 0. \tag{8}$$

Now this is a linear program in the form (2) with the following choice of matrix A and vectors b and c:

$$A = \begin{bmatrix} M & -M \end{bmatrix} \in \mathbb{R}^{m \times 2n}, \quad b = 2d \in \mathbb{R}^m, \quad c = (1, \ldots, 1) \in \mathbb{R}^{2n}.$$

The conversion from a problem (6) to its LP standard form can be a bit tedious. From now on, we will not do this conversion anymore and it will be taken for granted that problem (6) for example is a linear program (the step of going from (5) to (6) however is less trivial and so cannot be taken for granted in general). It will be assumed that the reader can do the "mechanic" conversion from (6) to a standard linear programming form (2) if needed.