Math 591 – Real Algebraic Geometry and Convex Optimization
Lecture 3: The cone of positive semidefinite matrices
Cynthia Vinzant, Spring 2019

An important example of a convex cone is the cone of positive semidefinite (PSD) matrices. This cone lives in the vector space of real symmetric matrices, $V = \mathbb{R}^{n\times n}_{\text{sym}}$, which has dimension $\binom{n+1}{2} = \frac{n(n+1)}{2}$. To see this, note that we can identify entries on the diagonal and above with 2-element subsets of the set $\{\text{diag}, 1, 2, \ldots, n\}$. Here $\{i, j\}$ corresponds to the (i, j) entry and (diag, i) corresponds to the (i, i) entry.

Recall that all the eigenvalues of a real symmetric matrix A are real.

Definition. The matrix $A \in \mathbb{R}^{n\times n}_{\text{sym}}$ is **positive semidefinite**, denoted $A \succeq 0$, if all its eigenvalues are nonnegative and **positive definite**, denoted $A \succ 0$, if all its eigenvalues are positive.

To characterize positive semi-definiteness, we use the following fact from linear algebra:

Linear Algebra Fact. For any $A \in \mathbb{R}^{n\times n}_{\text{sym}}$, there exists an orthogonal matrix U (i.e. with $U^T U = UU^T = I_n$) such that

$$A \;=\; U \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix} U^T.$$

Then $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A and

$$A \;=\; \sum_{i=1}^n \lambda_i u_i u_i^T$$

where $u_1, \ldots, u_n$ are the columns of U. For any vector $v \in \mathbb{R}^n$, this gives that

$$v^T A v \;=\; \sum_{i=1}^n \lambda_i v^T u_i u_i^T v \;=\; \sum_{i=1}^n \lambda_i (u_i^T v)^2.$$

Observation 1. *$v^T A v \geq 0$ for all $v \in \mathbb{R}^n$ $\Leftrightarrow$ $\lambda_1, \ldots, \lambda_n \geq 0$*

Proof. ($\Leftarrow$) Since $u_i^T v \in \mathbb{R}$, $(u_i^T v)^2 \geq 0$ and $v^T A v = \sum_{i=1}^n \lambda_i (u_i^T v)^2 \geq 0$.
($\Rightarrow$) If $\lambda_j < 0$ for some j, then consider $v = u_j$. Then

$$u_j^T A u_j \;=\; \sum_{i=1}^n \lambda_i (u_i^T u_j)^2 \;= \lambda_j < 0.$$

The last equality follows from the fact that $\{u_1, \ldots, u_n\}$ are orthonormal. □

Observation 2. *If $A \succeq 0$, then $v^T A v = 0$ if and only if $v \in \ker(A)$.*

Proof. ($\Leftarrow$) Clear. ($\Rightarrow$) Suppose $0 = v^T A v = \sum_{i=1}^n \lambda_{i=1}^n (u_i^T v)^2$. For each $i = 1, \ldots, n$, $\lambda_i (u_i^T v)^2 \geq 0$ and their sum is zero. It follows that each summand, $\lambda_i (u_i^T v)^2$ must be zero. Therefore for each i, either $\lambda_i = 0$ or $u_i^T v = 0$. It follows that

$$v \in \text{span}_{\mathbb{R}}\{u_j : \lambda_j = 0\} = (\text{span}_{\mathbb{R}}\{u_j : \lambda_j > 0\})^{\perp}.$$

□

Here we used the following useful fact:

Useful Fact. If $a_1, \ldots, a_n \in \mathbb{R}_{\geq 0}$ and $\sum_{i=1}^n a_i = 0$, then $a_1 = \ldots = a_n = 0$.

Observation 3. $A \succeq 0 \Leftrightarrow A \in \text{conv}\{xx^T : x \in \mathbb{R}^n\}$.

Proof. ($\Leftarrow$) If $A \in \text{conv}\{xx^T : x \in \mathbb{R}^n\}$, then we can write $A = \sum_i \lambda_i x_i x_i^T$ where $\lambda_i \geq 0$ and $\sum_i \lambda_i = 1$. Then for any $v \in \mathbb{R}^n$,

$$v^T A v \;=\; \sum_i \lambda_i v^T x_i x_i^T v \;=\; \sum_i \lambda_i (x_i^T v)^2 \;\geq\; 0$$

($\Rightarrow$) If $A \succeq 0$, then $A = \sum_{i=1}^n \lambda_i u_i u_i^T$ where $\lambda_i \geq 0$. Then taking $s = \sum_{i=1}^n \lambda_i$ gives

$$A = \sum_{i=1}^n \frac{\lambda_i}{s} (\sqrt{s} u_i)(\sqrt{s} u_i)^T.$$

Here the coefficients λ_i / s are nonnegative and sum to one, writing A as a convex combination of matrices of the form xx^T. □

Let PSD_n denote the set of positive semidefinite matrices in $\mathbb{R}^{n\times n}_{\text{sym}}$. It follows from both Observations 1 and 3 that PSD_n is a convex cone.

For any matrices $A, B \in \mathbb{R}^{n\times n}_{\text{sym}}$ define

$$\langle A, B\rangle \;=\; \text{trace}(AB) \;=\; \sum_{i=1}^n \sum_{j=1}^n A_{ij} B_{ij}.$$

Any linear functional on $\mathbb{R}^{n\times n}_{\text{sym}}$ has the form $A \mapsto \langle A, B\rangle$ for some $B \in \mathbb{R}^{n\times n}_{\text{sym}}$. Under this identification of $(\mathbb{R}^{n\times n}_{\text{sym}})^*$ with $\mathbb{R}^{n\times n}_{\text{sym}}$, the PSD cone is *self-dual.*

Theorem. $(\text{PSD}_n)^* = \text{PSD}_n$. *That is, a matrix $B \in \mathbb{R}^{n\times n}_{\text{sym}}$ is positive semidefinite if and only if $\langle A, B\rangle \geq 0$ for all $A \in \text{PSD}_n$.*

Moreover, for $A, B \in \text{PSD}_n$,

$$\langle A, B\rangle = 0 \quad \Leftrightarrow \quad \text{rowspan}(B) \subseteq \ker(A).$$

Proof. ($\Leftarrow$) Suppose that $\langle A, B\rangle \geq 0$ for all $A \in \text{PSD}_n$. In particular, for any $v \in \mathbb{R}^n$, $A = vv^T$ is positive semidefinite so

$$\langle A, B\rangle \;=\; \langle vv^T, B\rangle \;=\; \text{trace}(vv^T B) \;=\; \text{trace}(v^T B v) \;=\; v^T B v \geq 0.$$

Since $v^T B v \geq 0$ for all $v \in \mathbb{R}^n$, by Observation 1, $B \in \text{PSD}_n$.

($\Rightarrow$) If $B \in \text{PSD}_n$, then $B = \sum_{j=1}^n \beta_j w_j w_j^T$ where $\beta_j \geq 0$ and $\{w_1, \ldots, w_n\}$ are orthonormal. Then for real symmetric matrix A,

$$\langle A, B\rangle \;=\; \sum_{j=1}^n \beta_j \langle A, w_j w_j^T\rangle \;=\; \sum_{j=1}^n \beta_j w_j^T A w_j.$$

If A is positive semidefinite, each term $w_j^T A w_j$ is nonnegative, giving that $\langle A, B\rangle \geq 0$.

To prove the "moreover" statement, suppose that A, B are positive semidefinite and consider the decomposition $B = \sum_{j=1}^{n} \beta_j w_j w_j^T$ as above. Then

$$\begin{aligned} \langle A, B \rangle = 0 &\Leftrightarrow \beta_j w_j^T A w_j = 0 \text{ for all } j = 1, \dots, n \\ &\Leftrightarrow w_j^T A w_j = 0 \text{ for all } j \text{ with } \beta_j \neq 0 \\ &\Leftrightarrow w_j \in \ker(A) \text{ for all } j \text{ with } \beta_j \neq 0 \\ &\Leftrightarrow \operatorname{rowspan}(B) = \operatorname{span}_{\mathbb{R}}\{w_j : \beta_j \neq 0\} \subseteq \ker(A). \end{aligned}$$

□

Question. Consider the 3×3 diagonal matrices

$$B_1 = \begin{pmatrix} -1 & & \\ & 0 & \\ & & 0 \end{pmatrix}, \quad B_2 = \begin{pmatrix} -1 & & \\ & -1 & \\ & & 0 \end{pmatrix}, \quad \text{and} \quad B_3 = \begin{pmatrix} -1 & & \\ & -1 & \\ & & -1 \end{pmatrix}.$$

What are the faces of PSD_3 exposed by the linear functionals $A \mapsto \langle A, B_i \rangle$? What are the dimensions of these faces?

Answer. Since $-B_1$ is positive semidefinite, $\langle A, B_1 \rangle = -\langle A, -B_1 \rangle \leq 0$ for all positive semidefinite matrices A. Therefore the face exposed by B_1 is the set of PSD matrices A for which $\langle A, B_1 \rangle = -\langle A, -B_1 \rangle = 0$. Using the theorem above, this is the set of PSD matrices with $\operatorname{rowspan}(B_1) = \operatorname{span}_{\mathbb{R}}\{(1,0,0)\}$ in the kernel. This is

$$\{A \in \mathrm{PSD}_3 : (1,0,0) \in \ker(A)\} \quad = \quad \left\{ \begin{pmatrix} 0 & 0 \\ 0 & \tilde{A} \end{pmatrix} : \tilde{A} \in \mathrm{PSD}_2 \right\}.$$

Since PSD_2 has dimension $\binom{2+1}{2} = 3$, this face is 3-dimensional.

Similarly, since $-B_2$ is positive semidefinite, the face exposed by B_2 is the set of PSD matrices A for which $\langle A, B_2 \rangle = -\langle A, -B_2 \rangle = 0$, which is the set of matrices with $\operatorname{rowspan}(B_2) = \operatorname{span}_{\mathbb{R}}\{(1,0,0),(0,1,0)\}$ in the kernel. This is

$$\{A \in \mathrm{PSD}_3 : (1,0,0),(0,1,0) \in \ker(A)\} \quad = \quad \left\{ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & a \end{pmatrix} : a \in \mathbb{R}_{\geq 0} \right\}.$$

This is just a one-dimensional ray.

Finally, the face exposed by $A \mapsto \langle A, B_3 \rangle$ is the set of positive semidefinite matrices A with $\operatorname{rowspan}(B_3) = \operatorname{span}_{\mathbb{R}}\{(1,0,0),(0,1,0),(0,0,1)\} = \mathbb{R}^3$ in the kernel. This is just the zero matrix:

$$\{A \in \mathrm{PSD}_3 : (1,0,0),(0,1,0),(0,0,1) \in \ker(A)\} \quad = \quad \left\{ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right\},$$

which is zero-dimensional.