Notes 01: Semidefinite Programs and positive semidefiniteness

1. Semidefinite programs

Semidefinite programs generalize linear programs (LP). Recall that a linear program looks like the following:

$$\begin{aligned} \max \quad & 2x_1 + 3x_2 - 4x_3 \\ & 5x_1 - 8x_2 + 4x_3 \leqslant 10 \\ & 4x_1 + 3x_2 - x_3 \leqslant 5 \\ & x_1, x_2, x_3 \geqslant 0 \end{aligned} \tag{1}$$

More generally, a linear program (in canonical form) takes the form

$$\begin{aligned} \max \quad & c^\top x \\ & a_1^\top x \leqslant b_1 \\ & \vdots \\ & a_m^\top x \leqslant b_m \\ & x \geqslant 0 \end{aligned}$$

where $x, c, a_1, \ldots, a_m \in \mathbb{R}^n$ are all n-dimensional real vectors, and $b_1, \ldots, b_m \in \mathbb{R}$ are real scalars. Here x represents our LP variables, c represents our linear objective function, and $a_1, \ldots, a_m$ are the linear constraints. The last inequality constraint $x \geqslant 0$ means that x has to be entry-wise nonnegative.

By contrast, a semidefinite program (SDP) looks like the following:

$$\begin{aligned} \max \quad & \begin{pmatrix} 2 & 3/2 \\ 3/2 & -4 \end{pmatrix} \bullet \begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix} \\ & \begin{pmatrix} 5 & -4 \\ -4 & 4 \end{pmatrix} \bullet \begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix} \leqslant 10 \\ & \begin{pmatrix} 4 & 3/2 \\ 3/2 & -1 \end{pmatrix} \bullet \begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix} \leqslant 5 \\ & \begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix} \succcurlyeq 0 \end{aligned}$$

Here $\bullet$ denotes the Frobenius/Hadamard inner product between two matrices, defined as the entry-wise inner product between two n-by-n matrices (treating them as length-n^2 vectors)

$$A \bullet B \stackrel{\text{def}}{=} \sum_{1 \leqslant i,j \leqslant n} A_{ij} B_{ij} .$$

The above semidefinite program has exactly the same objective function as the linear program (1) above, because

$$\begin{pmatrix} 2 & 3/2 \\ 3/2 & -4 \end{pmatrix} \bullet \begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix} = 2x_1 + 3x_2 - 4x_3 .$$

Compared to (1), the main difference is that the final nonnegative constraint is replaced with a *positive semidefinite* constraint, as defined now.

Definition 1.1. A symmetric n-by-n matrix M is *positive semidefinite* if for every $y \in \mathbb{R}^n$, the quadratic form $y^\top M y \geqslant 0$.

A general semidefinite program takes the form

$$\begin{aligned} \max \quad & C \bullet X \\ & A_1 \bullet X \leqslant b_1 \\ & \vdots \\ & A_m \bullet X \leqslant b_m \\ & X \succcurlyeq 0 \end{aligned}$$

where $X, C, A_1, \ldots, A_m$ are all n-by-n real symmetric matrices, and $b_1, \ldots, b_m \in \mathbb{R}$ are real scalars. The matrix X represents our SDP variables.

2. Quadratic forms

Given a real symmetric matrix M, the expression $y^\top M y$ in Definition 1.1 represents a quadratic form. A quadratic form in $\mathbb{R}^n$ is a homogeneous polynomial of degree 2, without linear or constant terms, such as $f(y_1, y_2) = 2y_1^2 + 3y_1y_2 - 4y_2^2$. This quadratic form corresponds to the real symmetric matrix

$$\begin{pmatrix} 2 & 3/2 \\ 3/2 & -4 \end{pmatrix}, \qquad \text{because} \qquad \begin{pmatrix} y_1 & y_2 \end{pmatrix} \begin{pmatrix} 2 & 3/2 \\ 3/2 & -4 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = f(y_1, y_2).$$

Every real symmetric matrix corresponds to a unique quadratic form, and vice versa. Definition 1.1 says that the a real symmetric matrix is positive semidefinite if its corresponding quadratic form is nonnegative at every input y.

A quadratic form whose input is a scalar (as opposed to a vector) must be of the form $g(y) = \lambda y^2$ for some real number λ. Such a quadratic form is positive semidefinite (that is, the corresponding matrix is positive semidefinite) if and only if $\lambda \geqslant 0$.

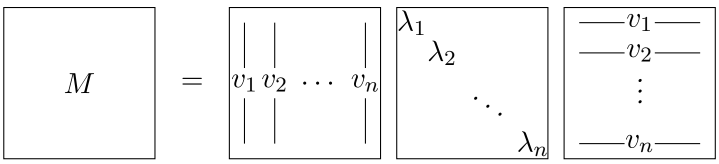
We can add two quadratic forms (coefficient-wise) to get another quadratic form, just like we can add two real symmetric matrices to get another real symmetric matrix. By adding together "simple" quadratic forms, we get complicated ones. Here a quadratic form is "simple" if, roughly speaking, it depends only on one dimension. Formally, a quadratic form $f(y)$ is simple if $f(y) = g(y^\top v)$ for some vector v and quadratic form g that takes a scalar input. In other words, f is constructed by first projecting y along direction v and then evaluating scalar quadratic form g. The real symmetric matrices corresponding to "simple" quadratic forms are precisely those of rank 1, that is, of the form $\lambda v v^\top$ for some $\lambda \in \mathbb{R}$ and $v \in \mathbb{R}^n$.

The following theorem (that we state without proof) tells us the structure of every quadratic form and their corresponding real symmetric matrices.

Theorem 2.1 (Spectral theorem for real symmetric matrices)**.** *Any n-by-n real symmetric matrix M has n real eigenvalues $\lambda_1, \ldots, \lambda_n$ and n orthonormal eigenvectors $v_1, \ldots, v_n$. Equivalently, we can express any such an M as*

$$M = V \Lambda V^\top \,, \tag{2}$$

where V is an n-by-n matrix whose columns are precisely the eigenvectors $v_1, \ldots, v_n$, and Λ is a diagonal *matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$ on its diagonal ($\Lambda_{ii} = \lambda_i$). Since the eigenvectors are orthonormal, we also have $V^\top V = V V^\top = I$. Decomposition (2) can be represented in picture as*



,

or as a sum of outer products

$$M = \sum_{1 \leqslant i \leqslant n} \lambda_i v_i v_i^\top \,.$$

This theorem says that every quadratic form in $\mathbb{R}^n$ is a sum of n "simple" quadratic forms, and these simple quadratic forms depend on orthogonal directions.

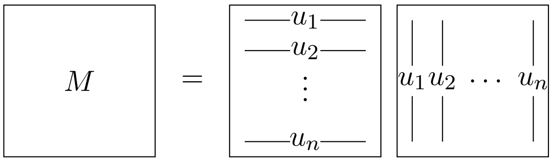
3. Positive semidefiniteness

The positive semidefinite (PSD) condition has a number of equivalent definitions.

Proposition 3.1. *Given a real symmetric n-by-n matrix M, the following are equivalent:*

(a) For every $y \in \mathbb{R}^n$, we have $y^\top M y \geqslant 0$
(b) All eigenvalues of M are nonnegative
(c) $M = U^\top U$ for some m-by-n matrix U (U is not necessarily symmetric or square)

Condition (c) is equivalent to saying that there are n vectors $u_1, \ldots, u_n \in \mathbb{R}^m$ such that M encodes the inner products between them. More precisely, $M_{ij} = u_i^\top u_j$, namely the ij-entry of X equals the inner product between the i- and the j-vectors. To see this, simply define $u_1, \ldots, u_n$ as the column vectors of U, and condition (c) becomes



.

Proof of the proposition. (a) $\Rightarrow$ (b): Consider each eigenvalue λ_i and its eigenvector v_i of M. Take y to be v_i, and positive semidefiniteness implies

$$0 \leqslant y^\top M y = v_i^\top M v_i = \lambda_i \ .$$

This inequality is true for every eigenvalue λ_i, so all eigenvalues are nonnegative.

(b) $\Rightarrow$ (c): Let $\sqrt{\Lambda}$ be the diagonal matrix with $\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n}$ on its diagonal ($\sqrt{\Lambda}_{ii} = \sqrt{\lambda_i}$), and let $U = \sqrt{\Lambda} V^\top$. Since M has only nonnegative eigenvalues (and they lie on the diagonal), $\sqrt{\Lambda}$ has only real entries. Also $(\sqrt{\Lambda})^\top = \sqrt{\Lambda}$ because it is a diagonal matrix. Then the spectral decomposition (2) becomes

$$M = V \sqrt{\Lambda}^\top \sqrt{\Lambda} V^\top = U^\top U \ ,$$

where $U = \sqrt{\Lambda} V^\top$.

(c) $\Rightarrow$ (a): For any $y \in \mathbb{R}^n$,

$$y^\top M y = y^\top U^\top U y = \|Uy\|_2^2 \geqslant 0 \ .$$

This proof says that when $M = U^\top U$, the quadratic form $y^\top M y$ amounts to measuring the norm squared of the vector Uy (the image of y under the linear map U from $\mathbb{R}^n$ to $\mathbb{R}^m$). And the norm squared of any vector is nonnegative. ∎