
Lecture 19

Semidefinite programming (SDP) is a far-reaching generalization of linear programming (LP), in which vectors are replaced by matrices and the non-negativity condition is replaced with the condition that the matrices be positive semidefinite.

19.1 Semidefinite programming

Recall the linear programming formulation

$$\begin{aligned} & \text{minimize} && \langle \mathbf{c}, \mathbf{x} \rangle \\ & \text{subject to} && \mathbf{A}\mathbf{x} = \mathbf{b} \\ & && \mathbf{x} \geq \mathbf{0}, \end{aligned} \tag{LP}$$

with $\mathbf{c}, \mathbf{x} \in \mathbb{R}^n$, $\mathbf{A} \in \mathbb{R}^{m \times n}$, and $\mathbf{b} \in \mathbb{R}^m$. Semidefinite programming is a far-reaching generalization of linear programming, in which the vector space \mathbb{R}^n is replaced by the vector space of real symmetric matrices,

$$\text{SYM}_n = \{\mathbf{X} \in \mathbb{R}^{n \times n} : x_{ij} = x_{ji}, 1 \leq i < j \leq n\},$$

with the trace inner product,

$$\langle \mathbf{X}, \mathbf{Y} \rangle = \mathbf{X} \bullet \mathbf{Y} = \text{tr}(\mathbf{XY}) = \sum_{i=1}^n \sum_{j=1}^n x_{ij} y_{ij},$$

where $\text{tr}(\mathbf{X}) = \sum_{i=1}^n x_{ii}$ is the *trace* of a matrix \mathbf{X} . The matrix \mathbf{A} in (LP) is replaced by a linear map $\mathcal{A}: \text{SYM}_n \rightarrow \mathbb{R}^m$. So far, this new setting does not fall out of the general framework of linear programming (note that (LP) does not make any reference on the nature of the inner product and the linear map \mathbf{A}). The point of departure of semidefinite programming is that the inequality constraints $\mathbf{x} \geq \mathbf{0}$ are replaced with the constraint $\mathbf{X} \succeq \mathbf{0}$, meaning that \mathbf{X} is positive semidefinite. A *semidefinite programming* (SDP) problem thus has the form

$$\begin{aligned} & \text{minimize} && \mathbf{C} \bullet \mathbf{X} \\ & \text{subject to} && \mathcal{A}(\mathbf{X}) = \mathbf{b} \\ & && \mathbf{X} \succeq \mathbf{0}. \end{aligned} \tag{SDP-P}$$

Remark 19.1. Recall that $\mathbf{X} \succeq \mathbf{0}$ means that for all $\mathbf{v} \in \mathbb{R}^n$, $\mathbf{v}^\top \mathbf{X} \mathbf{v} \geq 0$. Therefore, SDP can be viewed as linear programming with an uncountable number of inequality constraints. Recall also that a symmetric matrix is positive semidefinite if and only if the (necessarily real) eigenvalues are non-negative. While the feasible sets of linear programming are polyhedra, the feasible sets of SDP are called *spectrahedra*, since they are “polyhedra in the spectrum (set of eigenvalues)”.

The vector space SYM_n has dimension $d := n(n+1)/2$, the number of entries on or above the main diagonal. Therefore, if we identify $\text{SYM}_n \cong \mathbb{R}^d$, the linear map \mathcal{A} can be identified with a “big” $m \times d$ matrix. Alternatively, the condition $\mathcal{A}(\mathbf{X}) = \mathbf{b}$ can be seen as collection of m linear constraints, each of the form $\mathbf{A}_i \bullet \mathbf{X} = b_i$ with a symmetric matrix \mathbf{A}_i , so that we can express (SDP-P) as

$$\begin{aligned} & \text{minimize} && \mathbf{C} \bullet \mathbf{X} \\ & \text{subject to} && \mathbf{A}_i \bullet \mathbf{X} = b_i, \quad 1 \leq i \leq m, \\ & && \mathbf{X} \succeq \mathbf{0}. \end{aligned}$$

Example 19.2. Consider an SDP with $n = 3$ and $m = 2$ (3×3 symmetric matrices and 2 constraints), with the following data

$$\mathbf{A}_1 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 3 & 7 \\ 1 & 7 & 5 \end{pmatrix}, \quad \mathbf{A}_2 = \begin{pmatrix} 0 & 2 & 8 \\ 2 & 6 & 0 \\ 8 & 0 & 4 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 9 & 0 \\ 3 & 0 & 7 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 11 \\ 19 \end{pmatrix}.$$

The decision variable is the symmetric matrix

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{12} & x_{22} & x_{23} \\ x_{13} & x_{23} & x_{33} \end{pmatrix}.$$

The objective function is

$$\mathbf{C} \bullet \mathbf{X} = x_{11} + 4x_{12} + 6x_{13} + 9x_{22} + 7x_{33},$$

where we used the symmetry $x_{ij} = x_{ji}$ to simplify the expression. Working out the constraints $\mathbf{A}_i \bullet \mathbf{X} = b_i$ in the same way, we arrive at the following form of the SDP:

$$\begin{aligned} & \text{minimize} && x_{11} + 4x_{12} + 6x_{13} + 9x_{22} + 7x_{33} \\ & \text{subject to} && x_{11} + 2x_{13} + 3x_{22} + 14x_{23} + 5x_{33} = 11 \\ & && 4x_{12} + 16x_{13} + 6x_{22} + 4x_{33} = 19 \\ & && \mathbf{X} \succeq \mathbf{0}. \end{aligned}$$

As mentioned in Remark 19.1, were it not for the positive semidefinite constraint, the above would be an old-fashioned LP in the variables $\mathbf{x} = (x_{11}, x_{12}, x_{13}, x_{22}, x_{23}, x_{33})^\top$.

Linear programming is a special case of semidefinite programming. To see this, start with the problem (LP) and create the new matrices

$$\mathbf{A}_i = \text{diag}(a_{i1}, \dots, a_{in}), \quad \mathbf{C} = \text{diag}(c_1, \dots, c_n).$$

Then the problem (LP) is equivalent to the SDP

$$\begin{aligned} & \text{minimize} && \mathbf{C} \bullet \mathbf{X} \\ & \text{subject to} && \mathbf{A}_i \bullet \mathbf{X} = b_i, \quad 1 \leq i \leq m, \\ & && x_{ij} = 0, \quad i < j, \\ & && \mathbf{X} \succeq \mathbf{0}. \end{aligned}$$

In fact, any solution of the above has the form $\mathbf{X} = \text{diag}(x_{11}, \dots, x_{nn})$, which we identify with a vector $\mathbf{x} \in \mathbb{R}^n$, with the semidefinite constraint, applied to the diagonal matrix, translating into $x_i \geq 0$ for $1 \leq i \leq n$.

19.2 Semidefinite programming duality

Just as with linear programming, we can formulate a dual problem

$$\begin{aligned} & \text{maximize} && \langle \mathbf{y}, \mathbf{b} \rangle \\ & \text{subject to} && \sum_{i=1}^m y_i \mathbf{A}_i + \mathbf{S} = \mathbf{C} \\ & && \mathbf{S} \succeq \mathbf{0}. \end{aligned} \tag{SDP-D}$$

Example 19.3. The dual to the problem in Example 19.2 is the problem

$$\begin{aligned} & \text{maximize} && 11y_1 + 19y_2 \\ & \text{subject to} && y_1 \begin{pmatrix} 1 & 0 & 1 \\ 0 & 3 & 7 \\ 1 & 7 & 5 \end{pmatrix} + y_2 \begin{pmatrix} 0 & 2 & 8 \\ 2 & 6 & 0 \\ 8 & 0 & 4 \end{pmatrix} + \mathbf{S} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 9 & 0 \\ 3 & 0 & 7 \end{pmatrix}, \\ & && \mathbf{S} \succeq \mathbf{0}. \end{aligned}$$

One can also write the constraints more compactly by eliminating \mathbf{S} , in terms of a matrix whose entries are linear functions of y_1 and y_2 and which is required to be positive semidefinite.

We call \mathbf{X} a feasible matrix for (SDP-P) if it satisfies the constraints, and (\mathbf{y}, \mathbf{S}) a feasible pair for (SDP-D) if \mathbf{y} and \mathbf{S} satisfy the constraints. As with the nonlinear optimization setting from Lecture 16, we call the *duality gap* the difference $\mathbf{C} \bullet \mathbf{X} - \langle \mathbf{b}, \mathbf{y} \rangle$ between the primal and dual objective functions at feasible points.

Theorem 19.4. Let \mathbf{X} be a feasible matrix for (SDP-P) and let (\mathbf{y}, \mathbf{S}) be a feasible pair for (SDP-D). Then the duality gap satisfies

$$\mathbf{C} \bullet \mathbf{X} - \langle \mathbf{b}, \mathbf{y} \rangle = \mathbf{S} \bullet \mathbf{X} \geq 0.$$

Moreover, if the duality gap is zero, then \mathbf{X} and (\mathbf{y}, \mathbf{S}) are optimal points for (SDP-P) and (SDP-D), respectively.

Recall that $\mathbf{X} \in \text{SYM}_n$ if and only if $\mathbf{X} = \mathbf{Q}\mathbf{D}\mathbf{Q}^\top$, with \mathbf{D} diagonal (containing the eigenvalues) and \mathbf{Q} orthogonal (that is, $\mathbf{Q}^\top \mathbf{Q} = \mathbf{I}$). If in addition $\mathbf{X} \succeq \mathbf{0}$, then the diagonal entries are nonnegative. We need some more facts about positive semidefinite matrices.

Lemma 19.5. *Let $\mathbf{X} \in \mathcal{S}_+^n$. Then:*

1. *The diagonal entries satisfy $x_{ii} \geq 0$ for $1 \leq i \leq n$;*
2. *If $x_{ii} = 0$ for some i , then $x_{ij} = x_{ji} = 0$ for all $1 \leq j \leq n$ (if a diagonal is zero, then the whole corresponding row and column is zero).*

Proof of Theorem 19.4. (Optional.) Note that

$$\mathbf{C} \bullet \mathbf{X} - \langle \mathbf{b}, \mathbf{y} \rangle = \sum_{i=1}^m y_i \mathbf{A}_i \bullet \mathbf{X} + \mathbf{S} \bullet \mathbf{X} - \langle \mathbf{b}, \mathbf{y} \rangle = \mathbf{S} \bullet \mathbf{X},$$

since $\sum_{i=1}^m y_i \mathbf{A}_i \bullet \mathbf{X} = \sum_{i=1}^m y_i b_i = \langle \mathbf{y}, \mathbf{b} \rangle$. We next show that if $\mathbf{X} \succeq \mathbf{0}$ and $\mathbf{S} \succeq \mathbf{0}$, then $\mathbf{X} \bullet \mathbf{S} \geq 0$. Assume $\mathbf{X} = \mathbf{Q}\mathbf{D}\mathbf{Q}^\top$ and $\mathbf{S} = \mathbf{P}\mathbf{E}\mathbf{P}^\top$, with \mathbf{Q}, \mathbf{P} orthogonal and \mathbf{D}, \mathbf{E} diagonal. Then

$$\mathbf{X} \bullet \mathbf{S} = \text{tr}(\mathbf{S}\mathbf{X}) = \text{tr}(\mathbf{Q}\mathbf{D}\mathbf{Q}^\top \mathbf{P}\mathbf{E}\mathbf{P}^\top) = \text{tr}(\mathbf{D}\mathbf{Q}^\top \mathbf{P}\mathbf{E}\mathbf{P}^\top \mathbf{Q}),$$

where we used that $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$. The last expression equals

$$\sum_{i=1}^n D_{ii} (\mathbf{Q}^\top \mathbf{P}\mathbf{E}\mathbf{P}^\top \mathbf{Q})_{ii} \geq 0, \quad (19.1)$$

since $D_{ii} \geq 0$ and $\mathbf{Q}^\top \mathbf{P}\mathbf{E}\mathbf{P}^\top \mathbf{Q} \succeq \mathbf{0}$, which implies that the diagonal entries are also nonnegative by Lemma 19.5.

Now if $\mathbf{X} \bullet \mathbf{S} = 0$, then by (19.1), $\sum_{i=1}^n D_{ii} (\mathbf{Q}^\top \mathbf{P}\mathbf{E}\mathbf{P}^\top \mathbf{Q})_{ii} = 0$. This means that for each i , either $D_{ii} = 0$, or $(\mathbf{Q}^\top \mathbf{P}\mathbf{E}\mathbf{P}^\top \mathbf{Q})_{ii} = 0$. In the latter case, by Lemma 19.5, the whole i -th row of this matrix is zero. It follows that

$$\mathbf{D}(\mathbf{Q}^\top \mathbf{P}\mathbf{E}\mathbf{P}^\top \mathbf{Q}) = (\mathbf{D}\mathbf{Q}^\top)(\mathbf{P}\mathbf{E}\mathbf{P}^\top \mathbf{Q}) = \mathbf{0}.$$

Since $\mathbf{AB} = \mathbf{0}$ implies $\mathbf{BA} = \mathbf{0}$, we get that $\mathbf{XS} = (\mathbf{P}\mathbf{E}\mathbf{P}^\top \mathbf{Q})(\mathbf{D}\mathbf{Q}^\top) = \mathbf{0}$. \square

As with nonlinear convex optimization, some mild conditions (also known as Slater's condition) ensure that the primal and dual optimal values coincide.

Theorem 19.6. *Let p^* be the optimal value of (SDP-P) and d^* the optimal value of (SDP-D). If there exists a feasible matrix $\mathbf{X} \succ \mathbf{0}$ for (SDP-P), and a feasible pair (\mathbf{y}, \mathbf{S}) for (SDP-D) with $\mathbf{S} \succ \mathbf{0}$, then $p^* = d^*$.*

The theory of interior point methods carries over to semidefinite programming, giving us efficient (in theory and practice) algorithms for solving such problem. In the next lecture we will address some concrete problems that can be solved efficiently using semidefinite programming, and for which linear or quadratic programming are not enough.