Lecture 10: Semidefinite Programming and Conic Optimization

Fall 2025 *Prof. Udell*

1 Semidefinite programs

1.1 Definition and notation

Let $\mathbb{S}^n$ denote the space of $n \times n$ real symmetric matrices, and $\langle A, B\rangle := \operatorname{tr}(A^T B) = \sum_{i,j} A_{ij}B_{ij}$ the trace inner product. We write $X \succeq 0$ to mean X is *positive semidefinite (psd)*, i.e., $v^T X v \geq 0$ for all $v \in \mathbb{R}^n$.

Definition 1.1 (Semidefinite program (SDP)). An SDP is an optimization problem of the form

$$\begin{array}{ll} \text{minimize} & \langle C, X\rangle \\ \text{subject to} & \langle A_i, X\rangle = b_i, \quad i = 1, \dots, m, \\ & X \succeq 0 \\ \text{variable} & X \in \mathbb{S}^n, \end{array}$$

where $C, A_1, \dots, A_m \in \mathbb{S}^n$ and $b \in \mathbb{R}^m$.

Remark 1.2 (Why SDPs matter). SDPs are convex optimization problems: any local optimum is globally optimal. They strictly generalize linear programs (LPs) and admit efficient algorithms (e.g., interior-point methods; first-order methods for large scale). They arise across control (Lyapunov inequalities), combinatorial optimization (convex relaxations such as MaxCut), and eigenvalue optimization (e.g., minimizing $\lambda_{\max}$).

Recall some facts about psd matrices:

Proposition 1.3 (Equivalent characterizations of $X \succeq 0$). *For $X \in \mathbb{S}^n$, the following are equivalent:*

(a) $X \succeq 0$ (i.e., $v^T X v \geq 0$ for all v).

(b) All eigenvalues of X are nonnegative.

(c) There exists a matrix R such that $X = R^T R$. Any such R is called a square root *of X and may be written as $X^{1/2}$.*

Proof. $(a) \Rightarrow (b)$: for any eigenpair (λ, u) with $\|u\|_2 = 1$, $u^T X u = \lambda \geq 0$. $(b) \Rightarrow (c)$: take $R = \Lambda^{1/2} U^T$ when $X = U \Lambda U^T$ with $\Lambda \succeq 0$. $(c) \Rightarrow (a)$: $v^T X v = \|Rv\|_2^2 \geq 0$. $\square$

Proposition 1.4 (The psd cone is closed and convex). *The set $\mathbb{S}^n_+ := \{X \in \mathbb{S}^n \mid X \succeq 0\}$ is a closed convex cone.*

Proof sketch. Cone and convexity follow from linearity of the quadratic form: if $X \succeq 0$ and $\alpha \geq 0$, then $v^T(\alpha X)v = \alpha v^T X v \geq 0$, and sums preserve psd. Closedness follows from spectral continuity: if $X_k \to X$ and $X_k \succeq 0$, then eigenvalues $\lambda_i(X_k) \geq 0$ converge to $\lambda_i(X)$, so $\lambda_i(X) \geq 0$. $\square$

Example 1.5 (A 2×2 psd matrix)**.** For $X = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \in \mathbb{S}^2$, $X \succeq 0$ iff $a \geq 0$, $c \geq 0$, and $ac - b^2 \geq 0$. Equivalently, defining $t = \frac{a+c}{2}$ and $u = \frac{a-c}{2}$, $X \succeq 0$ iff

$$t \geq \sqrt{u^2 + b^2},$$

so the cone $\mathbb{S}^2_+$ is linearly isomorphic to the second-order cone $\{(u, b, t) \mid \sqrt{u^2 + b^2} \leq t\}$.

1.2 Geometric interpretation

Proposition 1.6 (Affine slice of a cone)**.** *The feasible set of the SDP is the intersection*

$$\mathcal{F} = \{X \in \mathbb{S}^n \mid \langle A_i, X \rangle = b_i, i = 1, \ldots, m\} \cap \mathbb{S}^n_+.$$

Hence $\mathcal{F}$ is convex.

Proof. Direct from definitions and Proposition 1.4: since the equality constraints define an affine subspace and the psd constraint defines a convex cone, their intersection is convex. $\square$

Remark 1.7 (Visual intuition)*.* For $n = 2$, $\mathbb{S}^2$ is 3-dimensional (coordinates (a, b, c) or (u, b, t) above). The set $\mathbb{S}^2_+$ looks like a rotational "ice-cream" (second-order) cone in (u, b, t)-coordinates. Imposing the affine equations $\langle A_i, X \rangle = b_i$ slices this cone with a plane; the feasible set is a convex (possibly empty or unbounded) cross-section.

1.3 Applications

Control (Lyapunov inequalities). A continuous-time linear system $\dot{x} = Ax$ is exponentially stable iff there exists $P \in \mathbb{S}^n$, $P \succ 0$ such that

$$A^T P + PA \prec 0.$$

This is a *linear matrix inequality* (LMI) in the unknown P; feasibility is an SDP (minimize 0 subject to $P \succ 0$ and the LMI). Lyapunov functions and LMIs are a central SDP application area.

Combinatorial optimization. SDPs provide convex relaxations for many NP-hard problems. These relaxations use the psd constraint to encode nonconvex quadratic constraints. Consider a constraint $x_i \in \{\pm 1\}$ for each $i = 1, \ldots, n$. This constraint is equivalent to $x_i^2 = 1$. Define $X = xx^T$; then we can encode the same constraint as $X_{ii} = 1$ for all i together with the nonconvex rank constraint $\text{rank}(X) = 1$ and $X \succeq 0$. Relaxing the rank constraint gives an SDP relaxation.

Example 1.8 (Combinatorial relaxations: MaxCut). Given weights w_{ij}, the (NP-hard) Max-Cut problem admits the standard SDP relaxation

$$\begin{array}{ll} \text{maximize} & \frac{1}{4}\sum_{i,j} w_{ij}\,(1 - X_{ij}) \\ \text{subject to} & X_{ii} = 1, \quad i = 1,\dots,n, \\ & X \succeq 0, \\ \text{variable} & X \in \mathbb{S}^n, \end{array}$$

obtained by lifting $x_i \in \{\pm 1\}$ to unit vectors v_i with $X_{ij} = v_i^T v_j$. The relaxation is tight when $X^\star$ is rank one; in general it gives an upper bound and supports randomized rounding with a 0.878 approximation ratio (Goemans–Williamson).

Eigenvalue optimization. The spectral radius surrogates $\lambda_{\max}$ and $\lambda_{\min}$ are SDP-representable:

$$\lambda_{\max}(X) \le t \iff tI - X \succeq 0 \qquad \text{and} \qquad \lambda_{\min}(X) \ge \ell \iff X - \ell I \succeq 0.$$

Thus problems like $\min\{\lambda_{\max}(X) : \ X \in \mathcal{A}\}$ reduce to an SDP by introducing a scalar t and enforcing $tI - X \succeq 0$.

Exercise. Verify the equivalence $\lambda_{\max}(X) \le t \iff tI - X \succeq 0$. Then, formulate the problem

$$\begin{array}{ll} \text{minimize} & \lambda_{\max}(X) \\ \text{subject to} & \langle A_i, X\rangle = b_i, \quad i = 1,\dots,m, \\ \text{variable} & X \in \mathbb{S}^n \end{array}$$

as an SDP in the scalar t and matrix X.

2 Convex cones and conic form

2.1 Convex cones

Recall the definition of a cone and convex cone.

Definition 2.1 (Cone, convex cone). A set $K \subseteq \mathbb{R}^n$ is a *cone* if $x \in K$ and $\alpha \ge 0$ imply $\alpha x \in K$. It is a *convex cone* if, in addition, $x, y \in K$ implies $x + y \in K$. Equivalently,

$$x, y \in K,\ \alpha, \beta \ge 0 \quad \Longrightarrow \quad \alpha x + \beta y \in K.$$

Example 2.2 (Canonical cones used in optimization).

1. **Zero cone** $\{0\}$.
2. **Nonnegative orthant** $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n \mid x_i \ge 0\}$.
3. **Second-order (Lorentz) cone** $\mathcal{Q}^{n+1} = \{(x, t) \in \mathbb{R}^n \times \mathbb{R} \mid \|x\|_2 \le t\}$.
4. **Positive semidefinite (psd) cone** $\mathbb{S}^n_+ = \{X \in \mathbb{S}^n \mid X \succeq 0\}$.

5. **Exponential cone** $\mathcal{K}_{\exp} = \{(x, y, z) \in \mathbb{R}^3 \mid y > 0,\ ye^{x/y} \leq z\}$.
6. **Sums and products:** $K_1 + K_2 = \{x_1 + x_2 \mid x_i \in K_i\}$ and $K_1 \times K_2 = \{(x_1, x_2) \mid x_i \in K_i\}$ are convex cones when K_1, K_2 are.

Proposition 2.3 (Basic properties)**.** *Let $K \subseteq \mathbb{R}^n$ be a convex cone.*

(a) $0 \in K$.

(b) If $A \in \mathbb{R}^{m \times n}$, then the image $AK = \{Ax \mid x \in K\}$ is a convex cone.

(c) If $L : \mathbb{R}^n \to \mathbb{R}^m$ is linear, then the preimage $L^{-1}(K) = \{x \mid Lx \in K\}$ is a convex cone.

Proof. (a) With $x \in K$ and $\alpha = 0$, $\alpha x = 0 \in K$. (b)–(c) follow from linearity and the definition. $\square$

Remark 2.4 (Proper cones). A cone K is called *proper* if it is closed, convex, pointed ($K \cap (-K) = \{0\}$), and solid (has nonempty interior). Many duality results and algorithms assume K is proper; the canonical cones in Example 2.2 are proper.

3 Conic duality

Conic duality generalizes LP duality to optimization problems over convex cones. In contrast to general nonlinear duality, conic duality retains a clean and useful structure. For example, this allows for the development of efficient algorithms with predictable behavior for conic problems, such as interior-point methods for problems with quadratic objectives and inequality constraints (via the second-order cone) and for semidefinite programming.

3.1 Dual cones

We will need the concept of a dual cone to construct conic dual optimization problems.

Definition 3.1 (Dual cone)**.** The *dual cone* of a cone $K \subseteq \mathbb{R}^n$ is

$$K^* = \{y \in \mathbb{R}^n \mid \langle y, x\rangle \geq 0\ \forall x \in K\}.$$

Proposition 3.2 (Basic properties of dual cones)**.** *Let $K \subseteq \mathbb{R}^n$ be a convex cone.*

(a) K^* is a closed convex cone.

(b) If $K_1 \subseteq K_2$, then $K_2^* \subseteq K_1^*$.

(c) $(K^*)^* = \overline{\mathrm{conv}}(K)$, the closed convex hull of K.

Proof. (a) Cone and convexity follow from linearity of the inner product. Closedness follows from continuity of the inner product. (b) If $y \in K_2^*$, then $\langle y, x\rangle \geq 0$ for all $x \in K_2$; since $K_1 \subseteq K_2$, this holds for all $x \in K_1$ as well, so $y \in K_1^*$. (c) If $y \in K^*$, then $\langle y, x\rangle \geq 0$ for all $x \in K$; since $K \subseteq \overline{\mathrm{conv}}(K)$, this holds for all $x \in \overline{\mathrm{conv}}(K)$ as well, so $y \in \overline{\mathrm{conv}}(K)^*$. Conversely, if $y \in \overline{\mathrm{conv}}(K)^*$, then $\langle y, x\rangle \geq 0$ for all $x \in \overline{\mathrm{conv}}(K)$; since $\overline{\mathrm{conv}}(K)$ is the smallest closed convex set containing K, this implies $\langle y, x\rangle \geq 0$ for all $x \in K$, so $y \in K^*$. □

Definition 3.3 (Self-dual cone)**.** A cone K is *self-dual* if $K = K^*$.

Many of the most important cones in optimization are self-dual: $K = K^*$. Examples include the nonnegative orthant, the second-order cone, and the psd cone. We now prove self-duality of the psd cone.

Proposition 3.4. *The psd cone is* self-dual*. Moreover, with the trace inner product,*

$$(\mathbb{S}^n_+)^* = \{Y \in \mathbb{S}^n \mid \langle X, Y\rangle \geq 0 \forall X \in \mathbb{S}^n_+\} = \mathbb{S}^n_+.$$

Proof. If $Y \not\succeq 0$, there exists u with $u^T Y u < 0$; then for $X = uu^T \succeq 0$, $\langle X, Y\rangle = \mathrm{tr}(uu^T Y) = u^T Y u < 0$. Conversely, if $Y \succeq 0$ then $\langle X, Y\rangle = \mathrm{tr}(R^T R Y) = \mathrm{tr}(R Y R^T) \geq 0$ for $X = R^T R \succeq 0$. We can see $(RYR^T) \geq 0$ since for any v, $v^T(RYR^T)v = (R^T v)^T Y (R^T v) \geq 0$. □

3.2 Primal–dual conic optimization problems

We begin from the conic-form primal introduced earlier:

$$\mathcal{P}: \quad \begin{array}{ll} \text{minimize} & \langle c, x\rangle \\ \text{subject to} & b - Ax \in K \\ \text{variable} & x \in \mathbb{R}^n, \end{array} \tag{1}$$

where $K \subseteq \mathbb{R}^m$ is a convex cone. Define the slack $s = b - Ax \in K$. To construct the dual, we introduce a Lagrange multiplier λ that acts *on the cone*, i.e., $\lambda \in K^* := \{y \mid \langle y, s\rangle \geq 0 \ \forall s \in K\}$ (the *dual cone*).

Definition 3.5 (Lagrangian and dual function in conic form)**.** The Lagrangian of the conic standard form problem (1) is

$$\mathcal{L}(x, \lambda) = \langle c, x\rangle - \langle \lambda, b - Ax\rangle = \langle c + A^*\lambda, x\rangle - \langle \lambda, b\rangle,$$

where A^* is the adjoint of A, defined by $\langle A^* w, x\rangle = \langle w, Ax\rangle$. The dual function is

$$g(\lambda) = \inf_x \mathcal{L}(x, \lambda) = \begin{cases} \langle -b, \lambda\rangle & c + A^*\lambda = 0 \\ -\infty & \text{otherwise} \end{cases}, \qquad \lambda \in K^*.$$

Recall that we construct the Lagrangian to ensure that it provides a lower bound on the primal objective for any feasible x and dual-feasible λ. The adjoint identity defining A^* is the standard Hilbert-space relation and will be used repeatedly below. For real-valued matrices and vectors, A^* is the transpose A^T.

Worked map for A^* (used later for SDPs). If $A : \mathbb{S}^n \to \mathbb{R}^m$ is given by $(AX)_i = \langle A_i, X\rangle$, then $\langle A^*\lambda, X\rangle = \langle \lambda, AX\rangle = \sum_{i=1}^m \lambda_i \langle A_i, X\rangle = \langle \sum_{i=1}^m \lambda_i A_i, X\rangle$,

Definition 3.6 (Conic dual problem)**.** Maximizing the dual function yields the dual problem

$$\mathcal{D}: \quad \begin{array}{ll} \text{maximize} & \langle -b, \lambda\rangle \\ \text{subject to} & c + A^*\lambda = 0, \\ \text{variable} & \lambda \in K^*. \end{array} \tag{2}$$

Remark 3.7 (Sign conventions and an equivalent dual). We have written our standard-form conic optimization problem in inequality form. Some texts (and our LP unit) write the standard-form problem with an equality constraint and $x \in K$. In this case, we arrive at a dual with objective $\langle b, \tilde{\lambda}\rangle$ and constraint $c - A^*\tilde{\lambda} = 0$, where $\tilde{\lambda} := -\lambda \in K^*$; this gives the familiar weak-duality inequality $\langle c, x\rangle \geq \langle b, \tilde{\lambda}\rangle$. We will keep λ and $\mathcal{D}$ as stated above to remain consistent with the slides on conic optimization. When we discuss an explicit dual for the standard-form SDP with equality constraints, we will see the $\langle b, \cdot\rangle$ objective in the SDP dual.

3.3 Weak and strong duality

Proposition 3.8 (Weak duality)**.** *For any primal-feasible x and dual-feasible λ (i.e., $b - Ax \in K$, $\lambda \in K^*$, and $c + A^*\lambda = 0$),*

$$\langle c, x\rangle + \langle b, \lambda\rangle \geq 0.$$

Proof. By feasibility of x and $\lambda \in K^*$, $\langle \lambda, b - Ax\rangle \geq 0$. Hence

$$\langle c, x\rangle \geq \langle c, x\rangle - \langle \lambda, b - Ax\rangle = \langle c + A^*\lambda, x\rangle - \langle \lambda, b\rangle = -\langle \lambda, b\rangle.$$

□

The value of the dual function $g(\lambda) = \langle -b, \lambda\rangle$ at dual-feasible λ, so $\langle c, x\rangle \geq g(\lambda)$.

Corollary 3.9 (Weak duality of optimal values)**.** *Let $p^\star$ and $d^\star$ be the optimal values of $\mathcal{P}$ and $\mathcal{D}$. Then $p^\star \geq d^\star$.*

Theorem 3.10 (Strong duality under Slater)**.** *Suppose the primal is feasible and satisfies* Slater's condition*: there exists $\bar{x}$ with $\bar{s} = b - A\bar{x} \in \operatorname{int} K$. Then strong duality holds: $p^\star = d^\star$. Moreover the dual optimum is attained (and likewise by symmetry if a strictly feasible dual exists).*

Remark 3.11 (KKT conditions for conic programs). Under Slater, optimality is characterized by the KKT system

Primal feasibility:	$s = b - Ax \in K,$
Dual feasibility:	$\lambda \in K^*,$
Stationarity:	$c + A^*\lambda = 0,$
Complementary slackness:	$\langle \lambda, s\rangle = 0.$

The last condition is *complementary slackness*: the optimal slack and dual variable are orthogonal. (This is the conic analogue of $y \geq 0,\ s \geq 0,\ y_i s_i = 0$ in LP.)

Geometric picture. If $s = b - Ax \in \partial K$ at optimum, the dual vector $\lambda \in K^*$ defines a supporting hyperplane $\{u \mid \langle \lambda, u \rangle = 0\}$ to K at s, and complementary slackness enforces that s lies on this face.

3.4 Self-dual cones and SDPs

When K is *self-dual* ($K = K^*$), the primal and dual involve the *same cone type*. The three main examples are LP ($\mathbb{R}^m_+$), SOCP ($\mathcal{Q}^{n+1}$), and SDP ($\mathbb{S}^n_+$); see the dual-cones table.

Explicit SDP dual. Consider the standard-form SDP

$$\begin{array}{ll} \text{minimize} & \langle C, X \rangle \\ \text{subject to} & \langle A_i, X \rangle = b_i, \quad i = 1, \dots, m, \\ & X \succeq 0, \\ \text{variable} & X \in \mathbb{S}^n, \end{array}$$

with inner product $\langle U, V \rangle = \mathrm{tr}(U^T V)$. The Lagrangian with multipliers $\lambda \in \mathbb{R}^m$ for the equalities and $S \succeq 0$ for the cone constraint is

$$\mathcal{L}(X, \lambda, S) = \ \langle C, X \rangle \ - \sum_{i=1}^{m} \lambda_i \big(\langle A_i, X \rangle - b_i \big) \ - \ \langle S, X \rangle = \ \big\langle C - \textstyle\sum_i \lambda_i A_i - S,\ X \big\rangle \ + \ b^T \lambda.$$

Minimizing over X forces $C - \sum_i \lambda_i A_i - S = 0$ (otherwise the infimum is $-\infty$). Eliminating $S \succeq 0$ yields the dual:

$$\boxed{\begin{array}{ll} \text{maximize} & b^T \lambda \\ \text{subject to} & C - \sum_{i=1}^m A_i \lambda_i \succeq 0, \end{array}}$$

which is an SDP again (self-duality of $\mathbb{S}^n_+$). Using the adjoint relation $A^* \lambda = \sum_i \lambda_i A_i$ justifies the middle step.

Remark 3.12 (KKT for SDP). Under Slater (e.g., there exists $X \succ 0$ with $\langle A_i, X \rangle = b_i$), optimality is equivalent to

$$X \succeq 0, \quad C - \sum_i A_i \lambda_i \succeq 0, \quad \langle A_i, X \rangle = b_i (i = 1, \dots, m), \quad \big\langle C - \textstyle\sum_i A_i \lambda_i,\ X \big\rangle = 0.$$

The last line is matrix complementary slackness: $\langle S, X \rangle = 0$ with $S = C - \sum_i A_i \lambda_i \succeq 0$.

Summary. Conic duality for $\mathcal{P} : \min\{\langle c, x \rangle \mid b - Ax \in K\}$ yields a clean companion problem in the dual cone K^*; weak duality is immediate from the Lagrangian, and strong duality follows under Slater's condition. For self-dual cones (LP/SOCP/SDP), the dual has the same cone type; for SDPs this produces the familiar dual LMI $C - \sum_i A_i \lambda_i \succeq 0$.