# Lecture 17: Primal-Dual Interior-Point Methods for Conic Optimization

Lecture 16 proved a feasible-start path-following method for minimizing a linear objective over a convex set equipped with a self-concordant barrier. Lecture 17 explains why cone formulations are the standard language of modern interior-point solvers. A cone constraint exposes a dual-cone slack, the duality gap becomes the pairing $\langle x, s\rangle$, and primal-dual interior-point methods become Newton's method applied to perturbed KKT residuals.

## 17.1 From Applications to Standard Conic Form

The point of conic form is not only uniform notation. It changes the objects that the algorithm sees. Second-order cones keep robust Euclidean constraints as explicit vector inequalities; semidefinite cones turn nonlinear quadratic relations into linear objectives over matrix variables.

**Example 17.1** (Robust linear constraints lead to SOCP)**.** Suppose the nominal inequality

$$a_i^\top x \le b_i$$

must hold even if the coefficient vector is perturbed to

$$a_i + P_i u, \qquad \|u\|_2 \le 1.$$

Robust feasibility means

$$\sup_{\|u\|_2\le 1} (a_i + P_i u)^\top x \le b_i.$$

Since

$$\sup_{\|u\|_2\le 1} u^\top P_i^\top x = \left\|P_i^\top x\right\|_2,$$

the robust constraint is

$$a_i^\top x + \left\|P_i^\top x\right\|_2 \le b_i.$$

Equivalently,

$$(b_i - a_i^\top x,\ P_i^\top x) \in \mathcal{Q}^{k_i+1}, \qquad \mathcal{Q}^{k+1} := \left\{(\tau, z) \in \mathbb{R}\times\mathbb{R}^k : \tau \ge \|z\|_2\right\}.$$

Thus ellipsoidal uncertainty naturally produces second-order cone constraints.

**Example 17.2** (Quadratic binary optimization leads to SDP)**.** Let $G = (V, E)$ be an undirected graph with weights $w_{ij} \ge 0$. A cut can be encoded by signs $\varepsilon_i \in \{\pm 1\}$. The Max-Cut objective is

$$\max_{\varepsilon_i \in \{\pm 1\}} \sum_{\{i,j\}\in E} w_{ij} \frac{1-\varepsilon_i\varepsilon_j}{2}.$$

This is a nonconvex quadratic optimization problem in the signs. If we set

$$X = \varepsilon\varepsilon^\top,$$

then $X \succeq 0$, $X_{ii} = 1$, and $\varepsilon_i\varepsilon_j = X_{ij}$. Dropping the rank-one constraint gives the SDP relaxation

$$\begin{array}{ll} \text{maximize} & \displaystyle\sum_{\{i,j\}\in E} w_{ij} \frac{1-X_{ij}}{2} \\ \text{subject to} & X_{ii} = 1 \quad i = 1, \dots, n, \\ & X \succeq 0. \end{array}$$

The objective is now linear in the matrix variable $X$, and the nonlinear "legal Gram matrix" condition is carried by the cone $X \succeq 0$. This is the modeling move behind many SDP relaxations: pairwise products become matrix entries, and the semidefinite cone records which matrices can be Gram matrices.

The relaxation *Example* 17.2 is convex but its optimal $X^\star$ is not in general rank-one, so we cannot just read off a cut. Goemans and Williamson [GW95] closed this gap with a random rounding scheme whose expected cut value is within a universal constant of the SDP value.

**Definition 17.1** (Goemans–Williamson randomized rounding)**.** Let $X^\star \succeq 0$ with $X^\star_{ii} = 1$ be the SDP optimum of Example 17.2. Factor

$$X^\star = V^\top V, \qquad V = [v_1, \dots, v_n],$$

so that $v_i \in \mathbb{R}^n$, $\|v_i\|_2 = 1$, and $\langle v_i, v_j\rangle = X^\star_{ij}$. Sample $g \sim \mathcal{N}(0, I_n)$ and set

$$\hat{\varepsilon}_i := \operatorname{sign}(\langle g, v_i\rangle).$$

**Theorem 17.1** (Goemans–Williamson approximation ratio)**.** *Let* $\mathrm{SDP}^\star$ *be the optimal value of the Max-Cut SDP relaxation in Example 17.2, and let* $\mathrm{MaxCut}^\star$ *be the optimal value of the original Max-Cut problem. The random cut produced by Definition 17.1 satisfies*

$$\mathbb{E}\left[\sum_{\{i,j\}\in E} w_{ij} \frac{1-\hat{\varepsilon}_i\hat{\varepsilon}_j}{2}\right] \;\geq\; \alpha_{\mathrm{GW}} \cdot \mathrm{SDP}^\star \;\geq\; \alpha_{\mathrm{GW}} \cdot \mathrm{MaxCut}^\star,$$

*where*

$$\alpha_{\mathrm{GW}} := \min_{0\leq\theta\leq\pi} \frac{2\theta}{\pi(1-\cos\theta)} \approx 0.87856.$$

*The proof is given in Section 17.7 at the end of this lecture.*

The examples above are named subclasses of one common template. The cone encodes the nonlinear-looking part of the feasible set, while all remaining constraints and objectives are linear.

**Definition 17.2** (Standard conic programs: LP, SOCP, and SDP)**.** A conic linear program is an optimization problem that can be written as

$$\inf \{\langle c, x\rangle : Ax = b,\ x \in K\},$$

where $A : E \to Y$ is linear, $b \in Y$, $c \in E^*$, and $K \subseteq E$ is a closed convex cone. The standard named subclasses are obtained by choosing $K$ from the standard cone families:

$$K = \mathbb{R}^m_+ \qquad \text{gives a linear program (LP),}$$

$$K = \mathbb{R}^p_+ \times \prod_{i=1}^{q} \mathcal{Q}^{k_i+1} \qquad \text{gives a second-order cone program (SOCP),}$$

and

$$K = \mathbb{R}^p_+ \times \prod_{j=1}^{r} \mathbb{S}^{n_j}_+ \qquad \text{gives a semidefinite program (SDP).}$$

Here

$$\mathcal{Q}^{k+1} := \left\{(\tau, z) \in \mathbb{R} \times \mathbb{R}^k : \tau \geq \|z\|_2\right\}, \qquad \mathbb{S}^n_+ := \{X \in \mathbb{S}^n : X \succeq 0\}.$$

Products of these cones allow several cone blocks in one optimization problem.

Both examples have the same pattern: a nonlinear convex constraint is represented as membership in a cone after a linear map. Once the feasible set is written this way, the right notion of a linear certificate is not a scalar multiplier but an element of the dual cone.

**Definition 17.3** (Dual cone)**.** Let $K \subseteq E$ be a cone in a finite-dimensional real vector space. Its dual cone is

$$K^* := \{s \in E^* : \langle s, x\rangle \geq 0\ \forall x \in K\}.$$

**Definition 17.4** (Conic primal-dual pair)**.** Let $A : E \to Y$ be linear, let $b \in Y$, $c \in E^*$, and let $K \subseteq E$ be a closed convex cone. The standard conic primal problem is

$$(P) \qquad p^\star = \inf \{\langle c, x\rangle : Ax = b,\ x \in K\}.$$

The standard conic dual problem is

$$(D) \qquad d^\star = \sup \{\langle b, y\rangle : A^* y + s = c,\ s \in K^*\}.$$

Here $y \in Y^*$ is the equality multiplier and $s \in E^*$ is the dual-cone slack.

## 17.2 Conic Duality, KKT, and Perturbations

Everything in this subsection is still the same lower-bound logic as Lecture 5. The new feature is that a cone constraint gives a vector slack $s \in K^*$, rather than a list of scalar nonnegative multipliers.

**Theorem 17.2** (Conic lower-bound certificates)**.** *Consider the conic primal problem in Defini-*

*tion 17.4. If* $y \in Y^*$ *and* $s \in K^*$ *satisfy*

$$A^*y + s = c,$$

*then* $\langle b, y\rangle$ *is a lower bound on every primal feasible objective value. Equivalently, the Lagrangian*

$$L(x; y, s) := \langle c, x\rangle + \langle y, b - Ax\rangle - \langle s, x\rangle, \qquad s \in K^*,$$

*in which the dual-cone slack* $s$ *is kept as an explicit variable rather than eliminated, satisfies*

$$\inf_{x \in E} L(x; y, s) = \begin{cases} \langle b, y\rangle, & A^*y + s = c, \\ -\infty, & A^*y + s \neq c. \end{cases}$$

*Consequently,* $(D)$ *is the conic Lagrange dual. Moreover, if* $x$ *is primal feasible and* $(y, s)$ *is dual feasible, then*

$$\langle c, x\rangle - \langle b, y\rangle = \langle x, s\rangle \geq 0.$$

*Proof of Theorem 17.2.* If $x$ is primal feasible and $A^*y + s = c$, then

$$\langle c, x\rangle = \langle A^*y + s, x\rangle = \langle y, Ax\rangle + \langle s, x\rangle = \langle y, b\rangle + \langle s, x\rangle \geq \langle b, y\rangle,$$

because $x \in K$ and $s \in K^*$. For the Lagrangian formula,

$$L(x; y, s) = \langle b, y\rangle + \langle c - A^*y - s, x\rangle.$$

Taking the infimum over unrestricted $x \in E$ gives $\langle b, y\rangle$ if the linear coefficient vanishes and $-\infty$ otherwise. The gap identity is the same calculation with both $x$ and $(y, s)$ feasible. □

*Remark* 17.1 (Why the slack is introduced explicitly)*.* The Lagrangian is not a different source of duality from the perturbation value function. As in Lecture 5, it is the computational device that turns value-function slopes into lower-bound certificates. We introduce $s \in K^*$ explicitly because this is the object primal-dual IPM maintains: it is the dual slack, and the gap is exactly $\langle x, s\rangle$.

**Definition 17.5** (Conic KKT point)**.** A triple $(x, y, s) \in E \times Y^* \times E^*$ is a KKT point of the conic pair in Definition 17.4 if

$$Ax = b, \qquad x \in K,$$

$$A^*y + s = c, \qquad s \in K^*,$$

and

$$\langle x, s\rangle = 0.$$

**Proposition 17.3** (Conic KKT points are optimality certificates)**.** *Every KKT point* $(x, y, s)$ *of the conic pair in Definition 17.4 certifies primal and dual optimality:*

$$\langle c, x\rangle = p^\star = d^\star = \langle b, y\rangle.$$

*Proof of Proposition 17.3.* By Theorem 17.2, every dual feasible pair gives a lower bound, and for this particular triple

$$\langle c, x\rangle - \langle b, y\rangle = \langle x, s\rangle = 0.$$

Thus the lower bound is tight. □

**Theorem 17.4** (Conic Slater)**.** *Let $K \subseteq E$ be a closed convex cone and consider*

$$p^\star = \inf\left\{\langle c, x\rangle : Ax = b,\ x \in K\right\}.$$

*Assume $p^\star \in \mathbb{R}$ and that there exists $\bar{x} \in \operatorname{int}(K)$ with $A\bar{x} = b$. Then the conic dual attains its optimum and*

$$d^\star = p^\star.$$

*Consequently, if the primal optimum is attained at $x^\star$, then there exist $y^\star$ and $s^\star$ such that $(x^\star, y^\star, s^\star)$ is a conic KKT point.*

*Proof of Theorem 17.4.* Define the equality-perturbation value function on $\operatorname{range}(A)$ by

$$p(u) := \inf\left\{\langle c, x\rangle : Ax = b + u,\ x \in K\right\}.$$

Since $\bar{x} \in \operatorname{int}(K)$, small perturbations of $\bar{x}$ remain in $K$. Choose a linear right inverse

$$B : \operatorname{range}(A) \to E, \qquad ABu = u.$$

For all sufficiently small $u \in \operatorname{range}(A)$,

$$\bar{x} + Bu \in K, \qquad A(\bar{x} + Bu) = b + u.$$

Hence $0 \in \operatorname{ri}(\operatorname{dom} p)$. By the marginal-duality theorem Theorem 4.4, applied on the perturbation space $\operatorname{range}(A)$, there exists a subgradient $\tilde{y} \in \operatorname{range}(A)^*$ attaining the conjugate-side dual formula at $u = 0$. Extend $\tilde{y}$ to some $y \in Y^*$. Then

$$p(0) = -p^*(\tilde{y}).$$

Computing the conjugate gives

$$\begin{aligned} p^*(\tilde{y}) &= \sup_{u \in \operatorname{range}(A)} \left(\langle \tilde{y}, u\rangle - p(u)\right) \\ &= \sup_{x \in K} \left(\langle y, Ax - b\rangle - \langle c, x\rangle\right) \\ &= -\langle b, y\rangle + \sup_{x \in K} \langle A^* y - c, x\rangle. \end{aligned}$$

The last supremum is finite exactly when

$$A^* y - c \in K^\circ = -K^*.$$

Equivalently, $s := c - A^* y \in K^*$, so $(y, s)$ is dual feasible. In this case $p^*(\tilde{y}) = -\langle b, y\rangle$, and therefore

$$p^\star = p(0) = \langle b, y\rangle.$$

Thus the dual optimum is attained and $d^\star = p^\star$. If $x^\star$ is primal optimal, then the gap identity in Theorem 17.2 gives

$$\langle x^\star, s\rangle = p^\star - d^\star = 0,$$

so $(x^\star, y, s)$ is a KKT point. □

*Remark* 17.2 (General set constraints versus cone constraints). The conic dual is the cone-specialized version of the Lecture 4 support-function dual. For a closed convex set $C$, the equality-constrained problem

$$\inf\{\langle c, x\rangle : Ax = b,\ x \in C\}$$

has the perturbation dual

$$\sup_y \{\langle b, y\rangle - \sigma_C(A^* y - c)\}.$$

If $C = K$ is a closed convex cone, then

$$\sigma_K(z) = \begin{cases} 0, & z \in K^\circ = -K^*, \\ +\infty, & z \notin K^\circ. \end{cases}$$

Thus the support-function penalty turns into the dual feasibility condition $A^* y + s = c,\ s \in K^*$. This is why cones are algorithmically cleaner than arbitrary convex-set constraints.

## 17.3 Dual Feasibility and Certificates in the Examples

We now return to the opening SOCP and SDP examples. The goal is to identify what the abstract dual-feasibility object $s \in K^*$ becomes in concrete coordinates, and how it certifies bounds.

**Example 17.3** (SOCP robust dual certificate)**.** Write the robust LP from Example 17.1, with linear objective $c^\top x$ and cone block

$$(t_i, z_i) := (b_i - a_i^\top x,\ P_i^\top x) \in \mathcal{Q}^{k_i+1},$$

as the conic primal

$$\min\ c^\top x \qquad \text{subject to} \qquad (t_i, z_i) \in \mathcal{Q}^{k_i+1}, \quad i = 1, \dots, m.$$

Its conic dual, with dual variables $s_i = (\alpha_i, v_i) \in \mathcal{Q}^{k_i+1}$ (the second-order cone is self-dual), is

$$\max\ -\sum_{i=1}^m \alpha_i b_i \qquad \text{subject to} \qquad \sum_{i=1}^m (P_i v_i - \alpha_i a_i) = c, \quad (\alpha_i, v_i) \in \mathcal{Q}^{k_i+1}.$$

The SOC slack is

$$s_i = (\alpha_i, v_i) \in \mathcal{Q}^{k_i+1}.$$

For every primal-feasible $x$ and dual-feasible $(\alpha, v)$,

$$c^\top x + \sum_{i=1}^m \alpha_i b_i = \sum_{i=1}^m [\alpha_i t_i + v_i^\top z_i] = \sum_{i=1}^m \langle s_i, (t_i, z_i)\rangle \geq 0,$$

since each block-pair $s_i, (t_i, z_i)$ lies in the self-dual cone $\mathcal{Q}^{k_i+1}$. So a feasible SOC slack is a vector lower-bound certificate for the robust LP: each block contributes a scalar margin term $\alpha_i t_i$ and a vector direction term $v_i^\top z_i$, independently.

**Example 17.4** (Max-Cut SDP dual certificate)**.** Write the Max-Cut SDP relaxation from Example 17.2 as

$$\max \{\langle C, X\rangle : X_{ii} = 1,\ X \succeq 0\}.$$

Its dual can be written as

$$\min_{y\in\mathbb{R}^n} \sum_{i=1}^n y_i \qquad \text{subject to} \qquad \mathrm{Diag}(y) - C \succeq 0.$$

The PSD slack is

$$S := \mathrm{Diag}(y) - C \succeq 0.$$

For every feasible $X$ and $(y, S)$,

$$\sum_{i=1}^n y_i - \langle C, X\rangle = \langle \mathrm{Diag}(y) - C, X\rangle = \langle S, X\rangle \geq 0.$$

So $S \succeq 0$ is a matrix upper-bound certificate for all legal Gram matrices $X \succeq 0$ with unit diagonal.

*Remark* 17.3 (A scalar spectral formulation hides the matrix slack)*.* The feasible set $X \succeq 0$ can also be written as the scalar nonsmooth inequality

$$\lambda_{\max}(-X) \leq 0.$$

This describes the same primal feasible matrices, but it uses a different perturbation space. If we perturb the cone constraint, the perturbation is matrix-valued and the dual object is a matrix slack $S \succeq 0$. If we perturb the scalar inequality $\lambda_{\max}(-X) \leq 0$, the Lagrange multiplier is only a scalar $\lambda \geq 0$, and stationarity contains the nonsmooth inclusion

$$0 \in C - \mathrm{Diag}(y) - \lambda\, \partial\lambda_{\max}(-X).$$

Unpacking the subgradient reconstructs a PSD matrix certificate of the form

$$S = \lambda R, \qquad R \succeq 0, \quad \mathrm{tr}(R) = 1,$$

supported on the active eigenspace. Thus the final reduced dual bound can be the same, but the KKT system is not presented in the same variables. The conic formulation keeps the matrix slack explicit, which is exactly what the primal-dual Newton system needs.

## 17.4 Logarithmically Homogeneous Barriers on Cones

Conic duality and KKT do not require barriers. They give the target equations: primal feasibility, dual feasibility, and complementarity. An interior-point method cannot impose $\langle x, s\rangle = 0$ while staying strictly inside the cones. The role of a barrier is to replace exact complementarity by a controlled interior perturbation. On cones, the useful barriers are compatible with positive scaling.

A cone $K \subseteq E$ is called *proper* if it is closed, convex, pointed, and has nonempty interior. Properness is exactly what we need to support a self-concordant barrier in the next definition: closedness and convexity make the cone a well-defined feasibility object, pointedness rules out lines through the origin (so a barrier on the interior cannot be invariant under a one-parameter group), and the interior is the natural domain of the barrier.

**Definition 17.6** (Logarithmically homogeneous self-concordant barrier)**.** Let $K \subseteq E$ be a proper cone. A $\nu$-logarithmically homogeneous self-concordant barrier, abbreviated $\nu$-LHSCB, is a $\nu$-self-concordant barrier $F : \operatorname{int}(K) \to \mathbb{R}$ satisfying

$$F(\tau x) = F(x) - \nu \log \tau \qquad \forall x \in \operatorname{int}(K),\ \forall \tau > 0.$$

**Proposition 17.5** (Basic LHSCB identities)**.** *Let $F$ be a $\nu$-LHSCB for a proper cone $K$. Then for every $x \in \operatorname{int}(K)$,*

$$\langle \nabla F(x), x \rangle = -\nu, \qquad \nabla^2 F(x) x = -\nabla F(x), \qquad \|\nabla F(x)\|_{x,*} = \sqrt{\nu}.$$

*Moreover,*

$$-\nabla F(x) \in \operatorname{int}(K^*).$$

*Proof of Proposition 17.5.* Differentiate $F(\tau x) = F(x) - \nu \log \tau$ in $\tau$ at $\tau = 1$. This gives

$$\langle \nabla F(x), x \rangle = -\nu.$$

Differentiating this identity with respect to $x$ gives

$$\nabla^2 F(x) x = -\nabla F(x).$$

Therefore

$$\|\nabla F(x)\|_{x,*}^2 = \left\langle \nabla F(x), (\nabla^2 F(x))^{-1} \nabla F(x) \right\rangle = \langle \nabla F(x), -x \rangle = \nu.$$

The inclusion $-\nabla F(x) \in \operatorname{int}(K^*)$ is a standard duality property of logarithmically homogeneous barriers: the barrier gradient map sends $\operatorname{int}(K)$ into $-\operatorname{int}(K^*)$. We use this standard fact to construct central-path slacks below. □

**Example 17.5** (Standard conic barriers)**.** The basic barriers are

$$K = \mathbb{R}^n_+, \qquad F(x) = -\sum_{i=1}^n \log x_i, \qquad \nu = n;$$

$$K = \mathcal{Q}^{d+1}, \qquad F(\tau, u) = -\log(\tau^2 - \|u\|_2^2), \qquad \nu = 2;$$

$$K = \mathbb{S}^n_+, \qquad F(X) = -\log\det X, \qquad \nu = n.$$

For the PSD cone,

$$\nabla F(X) = -X^{-1}, \qquad \nabla^2 F(X)[H] = X^{-1} H X^{-1}.$$

Products of cones use sums of barriers, and the parameters add.

Before specializing further to cones, it is useful to record the general existence theorem behind the word "universal." We state the clean bounded version, where the geometry is easiest to write down. It explains why self-concordant barriers are not a special accident of polytopes or PSD cones; the price is that the construction is usually not a computational recipe.

**Theorem 17.6** (Universal barrier for a convex body)**.** *Let $D \subseteq E$ be compact, convex, and have nonempty interior, and let $n := \dim E$. For $x \in \operatorname{int}(D)$, define the polar of $D$ with respect to $x$ by*

$$D_x^\circ := \{s \in E^* : \langle s, y - x\rangle \le 1 \ \forall y \in D\}.$$

*Then*

$$F_D(x) := \log \operatorname{vol}(D_x^\circ)$$

*is an $n$-self-concordant barrier for $D$. Here* vol *denotes any Lebesgue measure on $E^*$; changing the measure only adds a constant to $F_D$.*

Theorem 17.6 is a general existence theorem: every compact full-dimensional convex set has a barrier with parameter controlled by dimension. For conic optimization we need a version compatible with positive scaling of the cone. This leads to the characteristic barrier.

**Theorem 17.7** (Characteristic barrier for cones)**.** *Let $K \subseteq E$ be a proper cone, and let $n := \dim E$. Define the characteristic function of $K$ by*

$$\phi_K(x) := \int_{K^*} e^{-\langle s,x\rangle}\, ds, \qquad x \in \operatorname{int}(K),$$

*where $ds$ is any Lebesgue measure on $E^*$. The integral is finite on* $\operatorname{int}(K)$*, and changing $ds$ only adds a constant to* $\log \phi_K$*. Then*

$$F_K(x) := \log \phi_K(x)$$

*is an $n$-logarithmically homogeneous self-concordant barrier for $K$.*

For cones, this construction is usually called the characteristic barrier. It is the barrier-existence result compatible with conic scaling. We do not prove the self-concordance parts of Theorems 17.6 and 17.7 in this course. The universal-barrier construction goes back to Nesterov and Nemirovskii [NN94], and the sharp $n$-self-concordance statement is due to Lee and Yue [LY21]. For cones, the logarithm of the characteristic function was identified by Güler [Gül96].

The logarithmic homogeneity is easy to check. For $\tau > 0$, change variables $r = \tau s$ in the integral:

$$\phi_K(\tau x) = \int_{K^*} e^{-\langle s,\tau x\rangle}\, ds = \tau^{-n} \int_{K^*} e^{-\langle r,x\rangle}\, dr = \tau^{-n}\phi_K(x).$$

Therefore

$$F_K(\tau x) = F_K(x) - n \log \tau.$$

The role of these universal constructions here is orientation. They give barriers in principle, but not necessarily barriers that a solver can use efficiently: practical algorithms need fast formulas for the value, gradient, Hessian, and Newton systems.

## 17.5 Conic Central Path as Perturbed KKT

From now on, fix a computable $\nu$-LHSCB $F$ for the cone under consideration. For the standard cones this will usually be one of the explicit barriers in Example 17.5, not necessarily the characteristic barrier from Theorem 17.7. The primal barrier subproblem is

$$x(t) = \arg\min \{t\,\langle c, x\rangle + F(x) : Ax = b,\ x \in \operatorname{int}(K)\}.$$

The equality-constrained optimality condition gives some $\eta(t) \in Y^*$ such that

$$tc + \nabla F(x(t)) + A^*\eta(t) = 0.$$

Define

$$\mu := \frac{1}{t}, \qquad y(t) := -\mu\,\eta(t), \qquad s(t) := -\mu\,\nabla F(x(t)).$$

By Proposition 17.5,

$$s(t) \in \operatorname{int}(K^*).$$

The central-path equations become

$$Ax(t) = b, \qquad A^*y(t) + s(t) = c, \qquad s(t) = -\mu\nabla F(x(t)).$$

Moreover,

$$\langle x(t), s(t)\rangle = -\mu\,\langle x(t), \nabla F(x(t))\rangle = \nu\mu.$$

Thus the conic central path is a strictly feasible perturbation of KKT: exact complementarity $\langle x, s\rangle = 0$ is replaced by

$$\langle x(t), s(t)\rangle = \nu\mu.$$

For LP, $F(x) = -\sum_i \log x_i$, so

$$s_i(t) = \frac{\mu}{x_i(t)} \qquad \Longleftrightarrow \qquad x_i(t)s_i(t) = \mu.$$

Writing

$$(x \odot s)_i := x_i s_i$$

for the coordinatewise product, this is the familiar perturbed complementarity equation

$$x(t) \odot s(t) = \mu\mathbf{1}.$$

For the other standard cones, the same rule $s(t) = -\mu\nabla F(x(t))$ gives different explicit formulas; we return to these in Section 17.7.

## 17.6 Primal-Dual Newton Systems

Lecture 14 introduced Newton's method through optimization. Primal-dual IPM uses Newton in its equation-solving form: for a residual map $R(z)$, solve

$$DR(z)\Delta z = -R(z), \qquad z^+ = z + \Delta z.$$

Here $R$ is the residual of primal feasibility, dual feasibility, and perturbed complementarity.

For LP,

$$\min\left\{c^\top x : Ax = b,\ x \geq 0\right\},$$

the dual is

$$\max\left\{b^\top y : A^\top y + s = c,\ s \geq 0\right\}.$$

The perturbed KKT system at complementarity level $\mu > 0$ is

$$Ax = b, \qquad A^\top y + s = c, \qquad x \odot s = \mu\mathbf{1}, \qquad x > 0,\ s > 0,$$

where $\odot$ denotes coordinatewise product. With residuals

$$r_p := b - Ax, \qquad r_d := c - A^\top y - s, \qquad r_c := \mu\mathbf{1} - x \odot s,$$

linearization gives the primal-dual Newton system

$$A\Delta x = r_p, \qquad A^\top \Delta y + \Delta s = r_d, \qquad S\Delta x + X\Delta s = r_c,$$

where $X = \operatorname{Diag}(x)$ and $S = \operatorname{Diag}(s)$. A step length is then chosen so that

$$x + \alpha\Delta x > 0, \qquad s + \alpha\Delta s > 0.$$

**Algorithm 1** One primal-dual path-following update for LP

**Require:** Strictly positive triple $(x_k, y_k, s_k)$ with $x_k > 0$, $s_k > 0$, current $\mu_k$, and target $\mu_{k+1} < \mu_k$.

1: Form the residuals at the new target:
$$r_p := b - Ax_k, \quad r_d := c - A^\top y_k - s_k, \quad r_c := \mu_{k+1}\mathbf{1} - x_k \odot s_k.$$
2: Solve the Newton system
$$A\Delta x = r_p, \quad A^\top \Delta y + \Delta s = r_d, \quad S_k\Delta x + X_k\Delta s = r_c.$$
3: Choose $\alpha \in (0, 1]$ so that $x_k + \alpha\Delta x > 0$ and $s_k + \alpha\Delta s > 0$.
4: Set $(x_{k+1}, y_{k+1}, s_{k+1}) := (x_k, y_k, s_k) + \alpha(\Delta x, \Delta y, \Delta s)$.

This box is only a one-step update template. A rigorous short-step method must define a central neighborhood and prove, by induction, that the update maps a point in the $\mu_k$-neighborhood to a point in the $\mu_{k+1}$-neighborhood while preserving positivity. Such analyses are standard for LP, and extend to the standard conic families using their additional scaling structure; see Nesterov–Nemirovskii [NN94], Wright [Wri97], and Renegar [Ren01]. Modern primal-dual solvers often use predictor-corrector and infeasible-start variants of the same perturbed-KKT Newton idea. We do not analyze these variants in this course.

## 17.7 Symmetric Cones and Self-Scaled Barriers

The LP derivation is explicit because the centrality equation is coordinatewise: $x \odot s = \mu\mathbf{1}$. For a general cone, there is no canonical coordinatewise product. The standard cones are special because they have enough algebraic structure to make primal-dual scaling and centrality computable.

Theorems 17.6 and 17.7 show that barriers exist in broad generality. They do not say that all cones are equally good computationally. LP, SOCP, and SDP have additional symmetry.

A cone in an inner-product space is *self-dual* if $K = K^*$ under the chosen inner product. It is *homogeneous* if its automorphism group acts transitively on $\operatorname{int}(K)$. A proper cone is *symmetric* if it is both self-dual and homogeneous. The standard cones

$$\mathbb{R}^n_+, \qquad \mathcal{Q}^{d+1}, \qquad \mathbb{S}^n_+$$

are symmetric cones.

For symmetric cones, the barrier geometry has algebraic operations analogous to products, inverses, square roots, and determinants. We will not introduce the Euclidean Jordan algebra product in this course. Instead, it is safest to remember centrality directly from the barrier-gradient formula

$$s = -\mu\nabla F(x).$$

For the three standard cones this gives

| Cone | Barrier | Centrality from $s = -\mu \nabla F(x)$ |
|---|---|---|
| $\mathbb{R}^n_+$ | $-\sum_i \log x_i$ | $s_i = \mu / x_i$, equivalently $x_i s_i = \mu$ for all $i$. |
| $\mathcal{Q}^{d+1}$ | $-\log(\tau^2 - \lVert u \rVert_2^2)$ | For $x = (\tau, u)$, if $d := \tau^2 - \lVert u \rVert_2^2$, then $$s = \left( \frac{2\mu\tau}{d}, -\frac{2\mu u}{d} \right).$$ |
| $\mathbb{S}^n_+$ | $-\log \det X$ | $S = \mu X^{-1}$, equivalently $X^{1/2} S X^{1/2} = \mu I$. |

Self-scaled barriers and Nesterov–Todd scaling formalize this structure [NT97]. We treat them only as orientation here: they explain why the standard cones support especially clean primal-dual scaling, but the course will not develop the Jordan-algebra machinery.

## Dependency and Proof Sketch

1. Examples 17.1 and 17.2 motivate conic form from applications: robust linear constraints produce SOCP blocks, while quadratic binary problems can be relaxed to SDPs by lifting products into a Gram matrix.
2. Theorem 17.2 and Proposition 17.3 are the conic versions of the Lecture 5 certificate story: every dual feasible slack gives a lower bound, and zero gap makes the certificate exact.
3. Theorem 17.4 is the Lecture 4 perturbation argument specialized to cone constraints. Slater puts the origin in the relative interior of the equality-perturbation domain.
4. Remark 17.3 explains why formulation matters for algorithms: equivalent primal feasible sets can have different perturbation spaces and therefore different visible KKT variables.
5. Proposition 17.5 turns barrier gradients into dual slacks. This is the bridge from the primal barrier method of Lecture 16 to the primal-dual Newton systems in this lecture.

## Exercises

1. Derive the SOCP reformulation in Example 17.1 from Cauchy–Schwarz, including the equality case in the support-function calculation.
2. Derive the conic dual of the Max-Cut SDP in Example 17.4. Show directly that the slack $S = \operatorname{Diag}(y) - C \succeq 0$ gives an upper bound on every feasible Gram matrix.
3. Numerically minimize $\theta \mapsto 2\theta/(\pi(1 - \cos\theta))$ on $[0, \pi]$ and verify that the optimum $\alpha_{\mathrm{GW}} \approx 0.87856$ is attained near $\theta^\star \approx 2.331$.
4. For LP, starting from
$$Ax = b, \qquad A^\top y + s = c, \qquad x \odot s = \mu \mathbf{1},$$
derive the primal-dual Newton system displayed above.
5. For $X, S \succeq 0$, prove that
$$\operatorname{tr}(XS) = 0 \quad \Longrightarrow \quad XS = SX = 0.$$

6. Verify the PSD barrier identities

$$\nabla(-\log\det X) = -X^{-1}, \qquad \nabla^2(-\log\det X)[H] = X^{-1}HX^{-1},$$

and check logarithmic homogeneity.

## Proof of the Goemans–Williamson Theorem

We prove Theorem 17.1 in two steps: a geometric edge-probability identity, and a worst-case edge ratio.

**Lemma 17.8** (Hyperplane separation probability)**.** *Let $v_i, v_j \in \mathbb{R}^n$ be unit vectors with $\theta_{ij} := \arccos(\langle v_i, v_j\rangle) \in [0, \pi]$, and let $g \sim \mathcal{N}(0, I_n)$. Then*

$$\Pr\big[\mathrm{sign}(\langle g, v_i\rangle) \neq \mathrm{sign}(\langle g, v_j\rangle)\big] = \frac{\theta_{ij}}{\pi}.$$

*Proof of Lemma 17.8.* The pair $(\langle g, v_i\rangle, \langle g, v_j\rangle)$ depends on $g$ only through its orthogonal projection onto the two-dimensional plane $\Pi := \mathrm{span}(v_i, v_j)$. Since $g$ is rotation-invariant, that projection is a Gaussian $g_\Pi \in \mathbb{R}^2$ with covariance $I_2$, and the events $\{\mathrm{sign}\,\langle g, v_k\rangle\}$ only depend on the angular direction of $g_\Pi$, which is uniformly distributed on the unit circle. The signs $\langle g_\Pi, v_i\rangle$ and $\langle g_\Pi, v_j\rangle$ disagree exactly when $g_\Pi$ lies between the hyperplanes orthogonal to $v_i$ and to $v_j$, i.e. inside a pair of antipodal arcs whose total angular length is $2\theta_{ij}$ out of $2\pi$. Hence the probability is $\theta_{ij}/\pi$. □

*Proof of Theorem 17.1.* With $X^\star = V^\top V$ and $\theta_{ij} = \arccos(X^\star_{ij})$, Lemma 17.8 gives

$$\mathbb{E}[\hat\varepsilon_i\hat\varepsilon_j] = 1 - 2\Pr[\hat\varepsilon_i \neq \hat\varepsilon_j] = 1 - \frac{2\theta_{ij}}{\pi},$$

so the expected cut weight is

$$\mathbb{E}\left[\sum_{\{i,j\}\in E} w_{ij}\frac{1-\hat\varepsilon_i\hat\varepsilon_j}{2}\right] = \sum_{\{i,j\}\in E} w_{ij}\frac{\theta_{ij}}{\pi}.$$

On the other hand, the SDP value equals

$$\mathrm{SDP}^\star = \sum_{\{i,j\}\in E} w_{ij}\frac{1-X^\star_{ij}}{2} = \sum_{\{i,j\}\in E} w_{ij}\frac{1-\cos\theta_{ij}}{2}.$$

For each edge $\{i,j\}$ with $w_{ij} \geq 0$, the ratio of the rounding contribution to the SDP contribution is

$$\frac{\theta_{ij}/\pi}{(1-\cos\theta_{ij})/2} = \frac{2\theta_{ij}}{\pi(1-\cos\theta_{ij})} \geq \min_{0\leq\theta\leq\pi}\frac{2\theta}{\pi(1-\cos\theta)} = \alpha_{\mathrm{GW}}.$$

Summing over edges, $\mathbb{E}[\text{cut weight}] \geq \alpha_{\mathrm{GW}}\cdot\mathrm{SDP}^\star$. Since $\mathrm{SDP}^\star \geq \mathrm{MaxCut}^\star$, the second inequality follows. A one-dimensional numerical minimization gives $\alpha_{\mathrm{GW}} \approx 0.87856$, attained near $\theta^\star \approx 2.331$. □

### Extended Reading

The standard reference for self-concordant barriers and polynomial-time interior-point methods is Nesterov–Nemirovskii [NN94]. For primal-dual LP/IPM algorithms, see Wright [Wri97] and Renegar [Ren01]. For self-scaled barriers and symmetric-cone primal-dual scaling, see Nesterov–Todd [NT97]; for the structure of symmetric cones, see Faraut–Koranyi [FK94]. The Max-Cut SDP and rounding theorem are due to Goemans–Williamson [GW95].

## References


[FK94] Jacques Faraut and Adam Korányi. *Analysis on Symmetric Cones.* Oxford University Press, Oxford, 1994.

[Gül96] Osman Güler. Barrier functions in interior point methods. *Mathematics of Operations Research*, 21(4):860–885, 1996.

[GW95] Michel X. Goemans and David P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM*, 42(6):1115–1145, 1995.

[LY21] Yin Tat Lee and Man-Chung Yue. Universal barrier is $n$-self-concordant. *Mathematics of Operations Research*, 46(3):1129–1148, 2021.

[NN94] Yurii Nesterov and Arkadii Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*, volume 13 of *SIAM Studies in Applied Mathematics.* Society for Industrial and Applied Mathematics, Philadelphia, PA, 1994.

[NT97] Yurii E. Nesterov and Michael J. Todd. Self-scaled barriers and interior-point methods for convex programming. *Mathematics of Operations Research*, 22(1):1–42, 1997.

[Ren01] James Renegar. *A Mathematical View of Interior-Point Methods in Convex Optimization.* Society for Industrial and Applied Mathematics, Philadelphia, PA, 2001.

[Wri97] Stephen J. Wright. *Primal-Dual Interior-Point Methods.* Society for Industrial and Applied Mathematics, Philadelphia, PA, 1997.