15. Conic optimization

- conic linear program
- examples
- modeling
- duality

Generalized (conic) inequalities

Conic inequality: a constraint $x \in K$ where K is a convex cone in $\mathbf{R}^m$

we will require that the cone K is **proper:**

- closed
- pointed: $K \cap (-K) = \{0\}$
- with nonempty interior: $\operatorname{int} K \neq \emptyset$; equivalently, $K + (-K) = \mathbf{R}^m$

Notation (for proper K)

$$x \succeq_K y \quad \Longleftrightarrow \quad x - y \in K$$
$$x \succ_K y \quad \Longleftrightarrow \quad x - y \in \operatorname{int} K$$

Inequality notation

we will use a different convention than in EE236B

Vector inequalities: for $x, y \in \mathbf{R}^m$

- $x > y,\ x \geq y$ denote componentwise inequality
- $x \succ y,\ x \succeq y$ denote conic inequality for general (unspecified) proper cone K
- $x \succ_K y,\ x \succeq_K y$ denote conic inequality for specific K

Matrix inequality: for $X, Y \in \mathbf{S}^p$

$X \succ Y,\ X \succeq Y$ mean $X - Y$ is positive (semi-)definite

Properties of conic inequalities

preserved by nonnegative linear combinations: if $x \preceq y$ and $u \preceq v$, then

$$\alpha x + \beta u \preceq \alpha y + \beta v \qquad \forall \alpha, \beta \geq 0$$

define a **partial order** of vectors

- $x \preceq x$
- $x \preceq y \preceq z$ implies $x \preceq z$
- $x \preceq y$ and $y \preceq x$ imply $y = x$

in general, not a *total* order (requires that $x \preceq y$ or $y \preceq x$ for all x, y)

Conic linear program

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \\ & Fx = g \end{array}$$

- $A \in \mathbf{R}^{m \times n}$, $F \in \mathbf{R}^{p \times n}$; without loss of generality, we can assume

$$\mathbf{rank}(F) = p, \qquad \mathbf{rank}(\left[\begin{array}{c} A \\ F \end{array}\right]) = n$$

- K is a proper cone in $\mathbf{R}^m$
- for $K = \mathbf{R}^m_+$, problem reduces to regular linear program (LP)
- by defining $K = K_1 \times \cdots \times K_r$, this can represent multiple conic inequalities

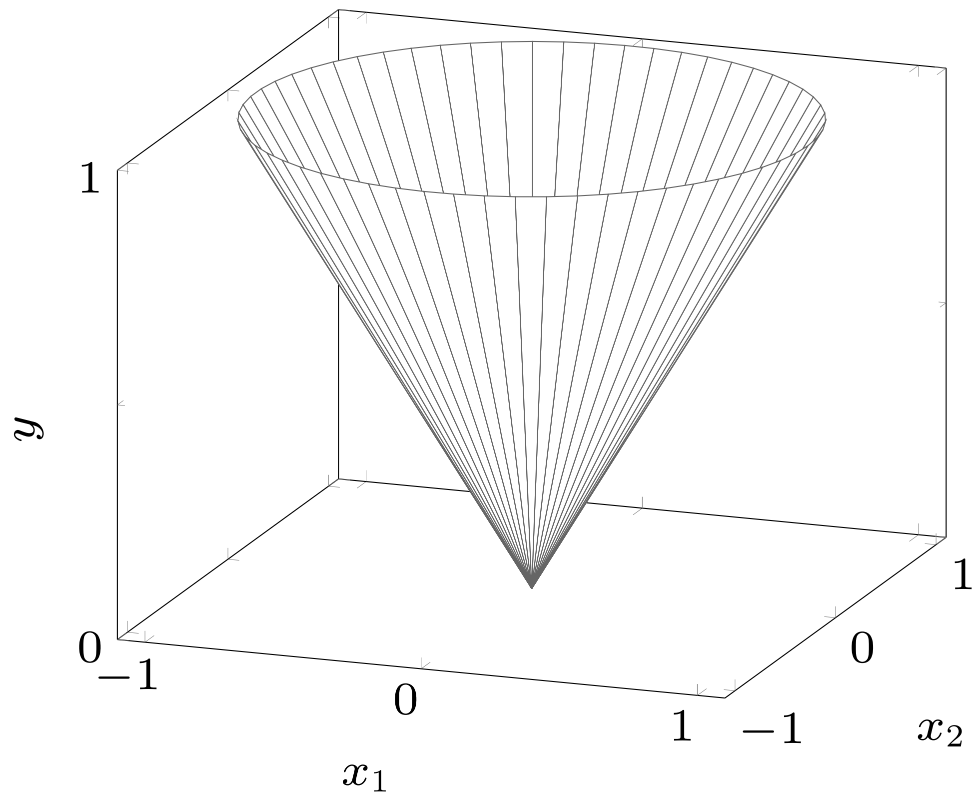
$$A_1 x \preceq_{K_1} b_1, \qquad \ldots, \qquad A_r x \preceq_{K_r} b_r$$

Outline

- conic linear program
- **examples**
- modeling
- duality

Norm cones

$$K = \left\{(x, y) \in \mathbf{R}^{m-1} \times \mathbf{R} \mid \|x\| \leq y\right\}$$



for the Euclidean norm this is the second-order cone (notation: $\mathcal{Q}^m$)

Second-order cone program

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \|B_{k0}x + d_{k0}\|_2 \leq B_{k1}x + d_{k1}, \quad k = 1, \ldots, r \end{array}$$

Conic LP formulation: express constraints as $Ax \preceq_K b$

$$K = \mathcal{Q}^{m_1} \times \cdots \times \mathcal{Q}^{m_r}, \qquad A = \begin{bmatrix} -B_{10} \\ -B_{11} \\ \vdots \\ -B_{r0} \\ -B_{r1} \end{bmatrix}, \qquad b = \begin{bmatrix} d_{10} \\ d_{11} \\ \vdots \\ d_{r0} \\ d_{r1} \end{bmatrix}$$

(assuming B_{k0}, d_{k0} have $m_k - 1$ rows)

Vector notation for symmetric matrices

- vectorized symmetric matrix: for $U \in \mathbf{S}^p$

$$\mathrm{vec}(U) = \sqrt{2}\left(\frac{U_{11}}{\sqrt{2}}, U_{21}, \ldots, U_{p1}, \frac{U_{22}}{\sqrt{2}}, U_{32}, \ldots, U_{p2}, \ldots, \frac{U_{pp}}{\sqrt{2}}\right)$$

- inverse operation: for $u = (u_1, u_2, \ldots, u_n) \in \mathbf{R}^n$ with $n = p(p+1)/2$

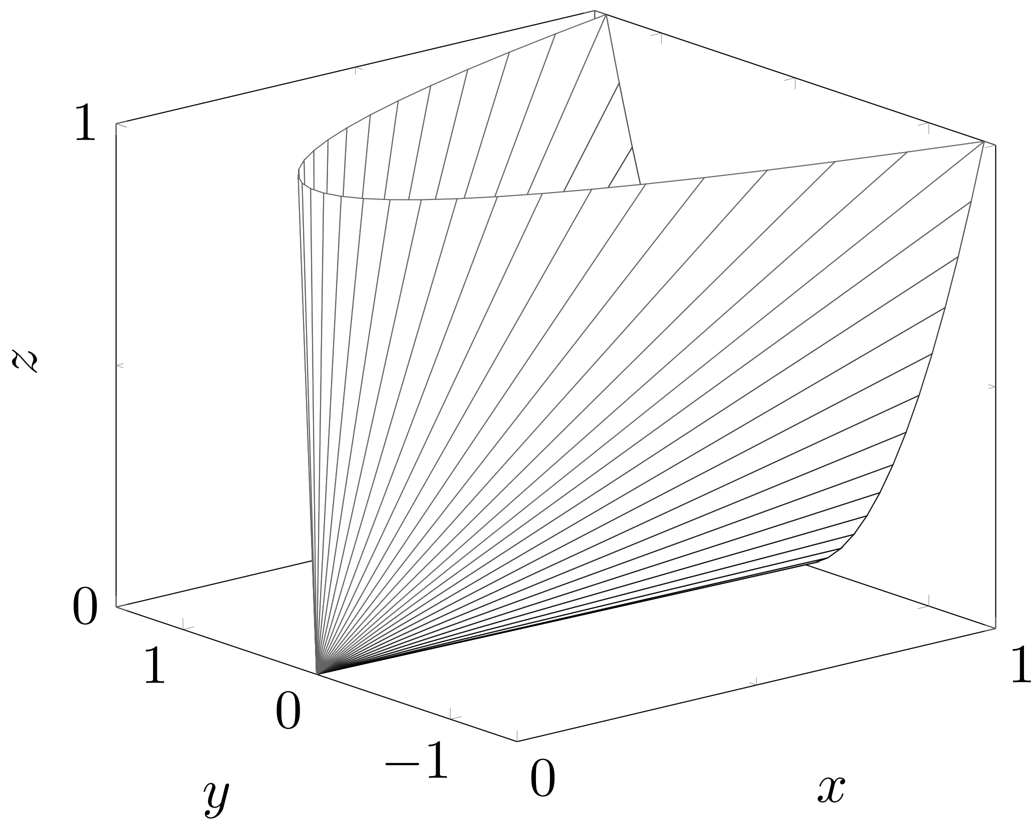
$$\mathrm{mat}(u) = \frac{1}{\sqrt{2}}\begin{bmatrix} \sqrt{2}u_1 & u_2 & \cdots & u_p \\ u_2 & \sqrt{2}u_{p+1} & \cdots & u_{2p-1} \\ \vdots & \vdots & & \vdots \\ u_p & u_{2p-1} & \cdots & \sqrt{2}u_{p(p+1)/2} \end{bmatrix}$$

coefficients $\sqrt{2}$ are added so that standard inner products are preserved:

$$\mathrm{tr}(UV) = \mathrm{vec}(U)^T \mathrm{vec}(V), \qquad u^T v = \mathrm{tr}(\mathrm{mat}(u)\,\mathrm{mat}(v))$$

Positive semidefinite cone

$$\mathcal{S}^p = \{\mathrm{vec}(X) \mid X \in \mathbf{S}_+^p\} = \{x \in \mathbf{R}^{p(p+1)/2} \mid \mathrm{mat}(x) \succeq 0\}$$



$$\mathcal{S}^2 = \left\{(x, y, z) \;\middle|\; \begin{bmatrix} x & y/\sqrt{2} \\ y/\sqrt{2} & z \end{bmatrix} \succeq 0\right\}$$

Semidefinite program

$$
\begin{array}{ll}
\text{minimize} & c^T x \\
\text{subject to} & x_1 A_{11} + x_2 A_{12} + \cdots + x_n A_{1n} \preceq B_1 \\
& \ldots \\
& x_1 A_{r1} + x_2 A_{r2} + \cdots + x_n A_{rn} \preceq B_r
\end{array}
$$

with $A_{ij}, B_i \in \mathbf{S}^{p_i}$

Conic LP formulation

$$
K = \mathcal{S}^{p_1} \times \mathcal{S}^{p_2} \times \cdots \times \mathcal{S}^{p_r}
$$

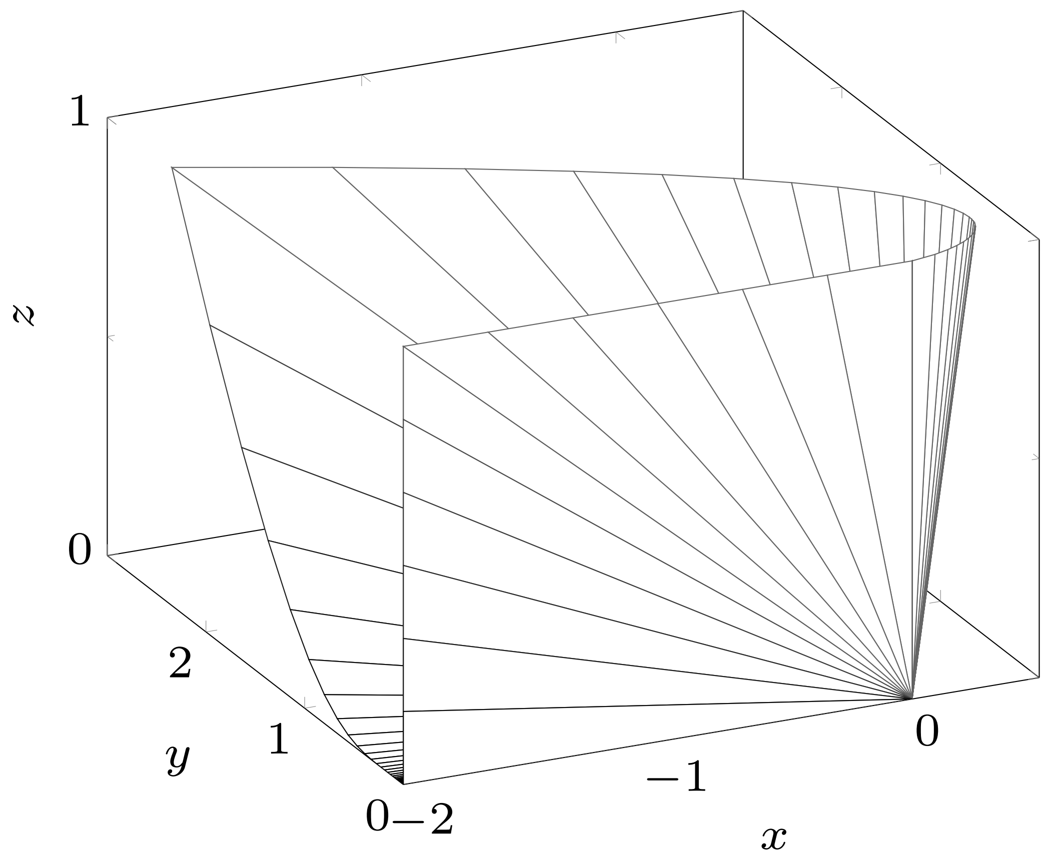
$$
A = \left[\begin{array}{cccc}
\mathrm{vec}(A_{11}) & \mathrm{vec}(A_{12}) & \cdots & \mathrm{vec}(A_{1n}) \\
\mathrm{vec}(A_{21}) & \mathrm{vec}(A_{22}) & \cdots & \mathrm{vec}(A_{2n}) \\
\vdots & \vdots & & \vdots \\
\mathrm{vec}(A_{r1}) & \mathrm{vec}(A_{r2}) & \cdots & \mathrm{vec}(A_{rn})
\end{array} \right], \qquad
b = \left[\begin{array}{c}
\mathrm{vec}(B_1) \\
\mathrm{vec}(B_2) \\
\vdots \\
\mathrm{vec}(B_r)
\end{array} \right]
$$

Exponential cone

the epigraph of the perspective of $\exp x$ is a non-proper cone

$$K = \left\{ (x, y, z) \in \mathbf{R}^3 \mid ye^{x/y} \le z,\ y > 0 \right\}$$

the exponential cone is $K_{\exp} = \mathbf{cl}\, K = K \cup \{(x, 0, z) \mid x \le 0, z \ge 0\}$



Geometric program

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \log \sum_{k=1}^{n_i} \exp(a_{ik}^T x + b_{ik}) \leq 0, \quad i = 1, \ldots, r \end{array}$$

Conic LP formulation

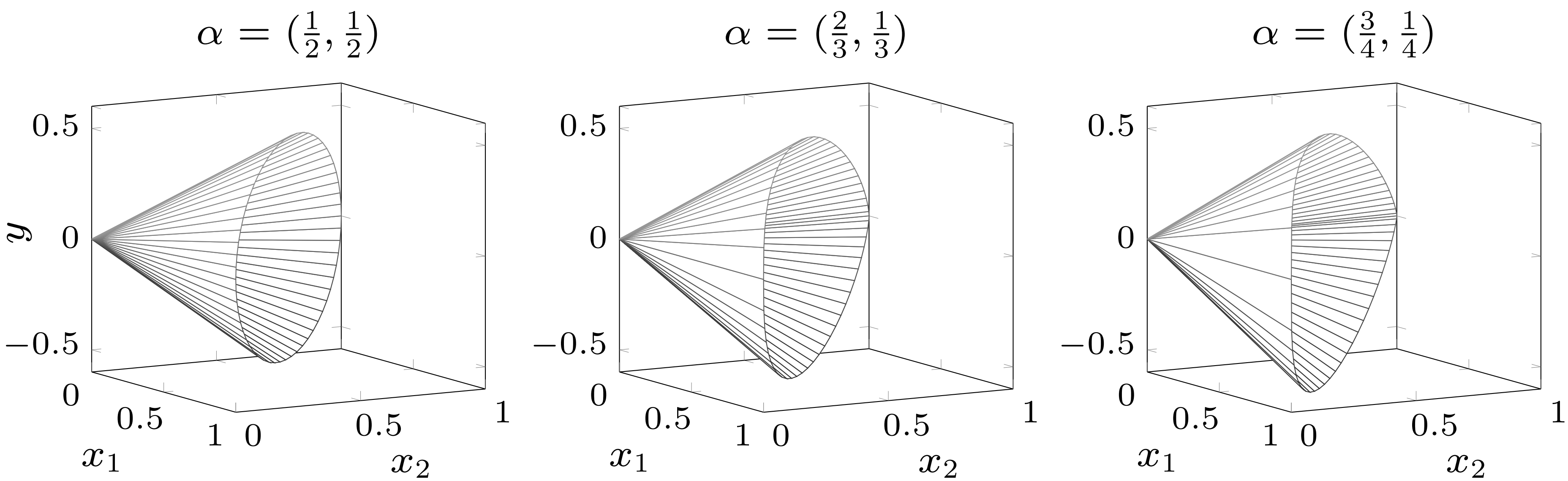
$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \begin{bmatrix} a_{ik}^T x + b_{ik} \\ 1 \\ z_{ik} \end{bmatrix} \in K_{\exp}, \quad k = 1, \ldots, n_i, \quad i = 1, \ldots, r \\ & \sum_{k=1}^{n_i} z_{ik} \leq 1, \quad i = 1, \ldots, m \end{array}$$

Power cone

Definition: for $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_m) > 0$ and $\sum_{i=1}^{m} \alpha_i = 1$

$$K_\alpha = \left\{(x, y) \in \mathbf{R}_+^m \times \mathbf{R} \mid |y| \le x_1^{\alpha_1} \cdots x_m^{\alpha_m}\right\}$$

Examples for $m = 2$



Cones constructed from convex sets

Inverse image of convex set under perspective

$$K = \{(x, y) \in \mathbf{R}^n \times \mathbf{R} \mid y > 0,\ x/y \in C\}$$

- $K \cup \{(0, 0)\}$ is a convex cone if C is a convex set
- $\mathrm{cl}\, K$ is proper if C has nonempty interior, does not contain straight lines

Consequence

any convex constraint $x \in C$ can be represented as a conic inequality

$$x \in C \qquad \Longleftrightarrow \qquad (x, 1) \in K$$

(with minor modifications to make K proper)

Cones constructed from functions

Epigraph of perspective of convex function

$$K = \{(x, y, z) \in \mathbf{R}^n \times \mathbf{R} \times \mathbf{R} \mid y > 0,\ yf(x/y) \le z\}$$

- $K \cup \{(0, 0, 0)\}$ is a convex cone if f is convex
- $\mathrm{cl}\, K$ is proper if $\mathrm{int}\,\mathrm{dom}\, f \neq \emptyset$, $\mathbf{epi}\, f$ does not contain straight lines

Consequence

can represent any convex constraint $f(x) \le t$ as a conic linear inequality

$$f(x) \le t \qquad \Longleftrightarrow \qquad (x, 1, t) \in K$$

(with minor modifications to make K proper)

Outline

- conic linear program
- examples
- **modeling**
- duality

Modeling software

Modeling packages for convex optimization

- CVX, YALMIP (MATLAB)
- CVXPY, PICOS (Python)
- MOSEK Fusion (different platforms)

assist in formulating convex problems by automating two tasks:

- verifying convexity from convex calculus rules
- transforming problem in input format required by standard solvers

Related packages

general-purpose optimization modeling: AMPL, GAMS

Modeling and conic optimization

Convex modeling systems

- convert problems stated in standard mathematical notation to conic LPs
- in principle, any convex problem can be represented as a conic LP
- in practice, a small set of primitive cones is used ($\mathbf{R}^n_+$, $\mathcal{Q}^p$, $\mathcal{S}^p$)
- choice of cones is limited by available algorithms and solvers (see later)

modeling systems implement set of rules for expressing constraints

$$f(x) \leq t$$

as conic inequalities for the implemented cones

Examples of second-order cone representable functions

- convex quadratic

$$f(x) = x^T P x + q^T x + r \qquad (P \succeq 0)$$

- quadratic-over-linear function

$$f(x,y) = \frac{x^T x}{y} \quad \text{with } \operatorname{dom} f = \mathbf{R}^n \times \mathbf{R}_+ \qquad \text{(assume } 0/0 = 0\text{)}$$

- convex powers with rational exponent

$$f(x) = |x|^\alpha, \qquad f(x) = \begin{cases} x^\beta & x > 0 \\ +\infty & x \leq 0 \end{cases}$$

  for rational $\alpha \geq 1$ and $\beta \leq 0$

- p-norm $f(x) = \|x\|_p$ for rational $p \geq 1$

Examples of SD cone representable functions

- matrix-fractional function

$$f(X, y) = y^T X^{-1} y \quad \text{with } \mathbf{dom}\, f = \{(X, y) \in \mathbf{S}^n_+ \times \mathbf{R}^n \mid y \in \mathcal{R}(X)\}$$

- maximum eigenvalue of symmetric matrix

- maximum singular value $f(X) = \|X\|_2 = \sigma_1(X)$

$$\|X\|_2 \leq t \quad \Longleftrightarrow \quad \begin{bmatrix} tI & X \\ X^T & tI \end{bmatrix} \succeq 0$$

- nuclear norm $f(X) = \|X\|_* = \sum_i \sigma_i(X)$

$$\|X\|_* \leq t \quad \Longleftrightarrow \quad \exists U, V : \begin{bmatrix} U & X \\ X^T & V \end{bmatrix} \succeq 0, \quad \frac{1}{2}(\mathbf{tr}\, U + \mathbf{tr}\, V) \leq t$$

Functions representable with exponential and power cone

Exponential cone

- exponential and logarithm
- entropy $f(x) = x \log x$

Power cone

- increasing power of absolute value: $f(x) = |x|^p$ with $p \geq 1$
- decreasing power: $f(x) = x^q$ with $q \leq 0$ and domain $\mathbf{R}_{++}$
- p-norm: $f(x) = \|x\|_p$ with $p \geq 1$

Outline

- conic linear program
- examples
- modeling
- **duality**

Dual cone

$$K^* = \{y \mid x^T y \geq 0 \text{ for all } x \in K\}$$

Properties (if K is a proper cone)

- K^* is a proper cone
- $(K^*)^* = K$
- $\operatorname{int} K^* = \{y \mid x^T y > 0 \text{ for all } x \in K, x \neq 0\}$

Dual inequality: $x \succeq_* y$ means $x \succeq_{K^*} y$ for generic proper cone K

Note: dual cone depends on choice of inner product

Examples

- $\mathbf{R}_+^p$, $\mathcal{Q}^p$, $\mathcal{S}^p$ are self-dual: $K = K^*$

- dual of norm cone is norm cone for dual norm

- dual of exponential cone

$$K_{\exp}^* = \left\{(u, v, w) \in \mathbf{R}_- \times \mathbf{R} \times \mathbf{R}^+ \mid -u \log(-u/w) + u - v \leq 0\right\}$$

(with $0 \log(0/w) = 0$ if $w \geq 0$)

- dual of power cone is

$$K_\alpha^* = \left\{(u, v) \in \mathbf{R}_+^m \times \mathbf{R} \mid |v| \leq (u_1/\alpha_1)^{\alpha_1} \cdots (u_m/\alpha_m)^{\alpha_m}\right\}$$

Primal and dual conic LP

Primal (optimal value $p^\star$)

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \end{array}$$

Dual (optimal value $d^\star$)

$$\begin{array}{ll} \text{maximize} & -b^T z \\ \text{subject to} & A^T z + c = 0 \\ & z \succeq_* 0 \end{array}$$

Weak duality: $p^\star \geq d^\star$ (without exception)

Strong duality

Main theorem: $p^\star = d^\star$ if primal *or* dual problem is strictly feasible

Other implications of strict feasibility

- if primal is strictly feasible, then dual optimum is attained (if $d^\star$ is finite)
- if dual is strictly feasible then primal optimum is attained (if $p^\star$ is finite)

Compare with linear programming duality ($K = \mathbf{R}^m_+$)

- for an LP, only exception to strong duality is $p^\star = +\infty$, $d^\star = -\infty$
- strong duality holds if primal or dual is feasible
- if optimal value is finite then it is attained (in primal and dual)

Example with finite nonzero duality gap

Primal problem

$$\begin{array}{ll} \text{minimize} & x_1 \\ \text{subject to} & \begin{bmatrix} 0 & x_1 \\ x_1 & x_2 \end{bmatrix} \succeq 0 \\ & x_1 \geq -1 \end{array}$$

optimal value $p^\star = 0$

Dual problem

$$\begin{array}{ll} \text{maximize} & -z \\ \text{subject to} & \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12} & Z_{22} \end{bmatrix} \succeq 0, \quad z \geq 0 \\ & 2Z_{12} + z = 1, \quad Z_{22} = 0 \end{array}$$

optimal value $d^\star = -1$

Optimality conditions

if strong duality holds, then x and z are optimal if and only if

$$\begin{bmatrix} 0 \\ s \end{bmatrix} = \begin{bmatrix} 0 & A^T \\ -A & 0 \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} + \begin{bmatrix} c \\ b \end{bmatrix} \qquad (1)$$

$$s \succeq 0, \qquad z \succeq_* 0, \qquad z^T s = 0$$

Primal feasibility: block 2 of (1) and $s \succeq 0$

Dual feasibility: block 1 of (1) and $z \succeq_* 0$

Zero duality gap: inner product of (x, z) and $(0, s)$ gives

$$z^T s = c^T x + b^T z$$

Strong theorems of alternative

Strict primal feasibility

exactly one of the following two systems is solvable

1. $Ax \prec b$
2. $A^T z = 0, \quad z \neq 0, \quad z \succeq_* 0, \quad b^T z \leq 0$

Strict dual feasibility

if $c \in \mathcal{R}(A^T)$, exactly one of the following two systems is solvable

1. $Ax \preceq_K 0, \quad Ax \neq 0, \quad c^T x \leq 0$
2. $A^T z + c = 0, \quad z \succ_{K^*} 0$

solution of one system is a certificate of infeasibility of the other system

Weak theorems of alternative

Primal feasibility

at most one of the following two systems is solvable

1. $Ax \preceq b$
2. $A^T z = 0, \quad z \succeq_* 0, \quad b^T z < 0$

Dual feasibility

at most one of the following two systems is solvable

1. $Ax \preceq 0, \quad c^T x < 0$
2. $A^T z + c = 0, \quad z \succeq_* 0$

these are strong alternatives if a constraint qualification holds

Self-dual embeddings

Idea

embed primal, dual conic LPs into a single (self-dual) conic LP, so that:

- embedded problem is feasible, with known feasible points
- from the solution of the embedded problem, primal and dual solutions of original problem can be constructed, or certificates of primal or dual infeasibility

Purpose: a feasible algorithm applied to the embedded problem

- can detect infeasibility in original problem
- does not require a phase I to find initial feasible points

used by some interior-point solvers

Basic self-dual embedding

$$\begin{array}{ll} \text{minimize} & 0 \\ \text{subject to} & \begin{bmatrix} 0 \\ s \\ \kappa \end{bmatrix} = \begin{bmatrix} 0 & A^T & c \\ -A & 0 & b \\ -c^T & -b^T & 0 \end{bmatrix} \begin{bmatrix} x \\ z \\ \tau \end{bmatrix} \\ & s \succeq 0, \quad \kappa \geq 0, \quad z \succeq_* 0, \quad \tau \geq 0 \end{array}$$

- a self-dual conic LP
- has a trivial solution (all variables zero)
- $z^T s + \tau\kappa = 0$ for all feasible points (follows from equality constraint)
- hence, problem is not strictly feasible

Optimality condition for embedded problem

$$\begin{bmatrix} 0 \\ s \\ \kappa \end{bmatrix} = \begin{bmatrix} 0 & A^T & c \\ -A & 0 & b \\ -c^T & -b^T & 0 \end{bmatrix} \begin{bmatrix} x \\ z \\ \tau \end{bmatrix}$$

$$s \succeq 0, \quad \kappa \geq 0, \quad z \succeq_* 0, \quad \tau \geq 0$$

$$z^T s + \tau\kappa = 0$$

- follows from self-dual property
- a (mixed) linear complementarity problem

Classification of nonzero solution

let s, κ, x, z, τ be a nonzero solution of the self-dual embedding

Case 1: $\tau > 0$, $\kappa = 0$

$$\hat{s} = s/\tau, \qquad \hat{x} = x/\tau, \qquad \hat{z} = z/\tau$$

are primal and dual solutions of the conic LPs, *i.e.*, satisfy

$$\begin{bmatrix} 0 \\ \hat{s} \end{bmatrix} = \begin{bmatrix} 0 & A^T \\ -A & 0 \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{z} \end{bmatrix} + \begin{bmatrix} c \\ b \end{bmatrix}$$

$$\hat{s} \succeq 0, \quad \hat{z} \succeq_* 0, \quad \hat{s}^T \hat{z} = 0$$

Classification of nonzero solution

Case 2: $\tau = 0$, $\kappa > 0$; this implies $c^T x + b^T z < 0$

- if $b^T z < 0$, then $\hat{z} = z/(-b^T z)$ is a proof of primal infeasibility:

$$A^T \hat{z} = 0, \qquad b^T \hat{z} = -1, \qquad \hat{z} \succeq_* 0$$

- if $c^T x < 0$, then $\hat{x} = x/(-c^T x)$ is a proof of dual infeasibility:

$$A\hat{x} \preceq 0, \qquad c^T \hat{x} = -1$$

Case 3: $\tau = \kappa = 0$; no conclusion can be made about the original problem

Extended self-dual embedding

$$
\begin{array}{ll}
\text{minimize} & \theta \\
\text{subject to} & \begin{bmatrix} 0 \\ s \\ \kappa \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & A^T & c & q_x \\ -A & 0 & b & q_z \\ -c^T & -b^T & 0 & q_\tau \\ -q_x^T & -q_z^T & -q_\tau & 0 \end{bmatrix} \begin{bmatrix} x \\ z \\ \tau \\ \theta \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \\
& s \succeq 0, \quad \kappa \geq 0, \quad z \succeq_* 0, \quad \tau \geq 0
\end{array}
$$

- q_x, q_z, q_τ chosen so that

$$(s, \kappa, x, z, \tau, \theta) = (s_0, 1, x_0, z_0, 1, z_0^T s_0 + 1)$$

  is feasible, for some given $s_0 \succ 0$, x_0, $z_0 \succ_* 0$

- a self-dual conic LP

Optimality condition

$$\begin{bmatrix} 0 \\ s \\ \kappa \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & A^T & c & q_x \\ -A & 0 & b & q_z \\ -c^T & -b^T & 0 & q_\tau \\ -q_x^T & -q_z^T & -q_\tau & 0 \end{bmatrix} \begin{bmatrix} x \\ z \\ \tau \\ \theta \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$s \succeq 0, \quad \kappa \geq 0, \quad z \succeq_* 0, \quad \tau \geq 0$$

$$z^T s + \tau\kappa = 0$$

- follows from self-dual property
- a (mixed) linear complementarity problem

Properties of extended self-dual embedding

- problem is strictly feasible by construction
- if $s, \ \kappa, \ x, \ \ z, \ \tau, \ \theta$ satisfy equality constraint, then

$$\theta = s^T z + \kappa\tau$$

  (take inner product with (x, z, τ, θ) of each side of the equality)

- at optimum, $\theta = 0$ and problem reduces to the embedding on p.15-30
- classification of p.15-32 also applies to solutions of extended embedding

Reference

A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization. Analysis, Algorithms, and Engineering Applications*, (2001).