ELE539A: Optimization of Communication Systems
Lecture 15: Semidefinite Programming, Detection and Estimation Applications

Professor M. Chiang
Electrical Engineering Department, Princeton University

March 14, 2007

Lecture Outline

- Cones and dual cones
- Generalized inequality and convexity
- Conic programming
- Semidefinite programming (SDP)
- Application: ML estimation
- Application: Covariance estimation
- Application: Multi-user detection

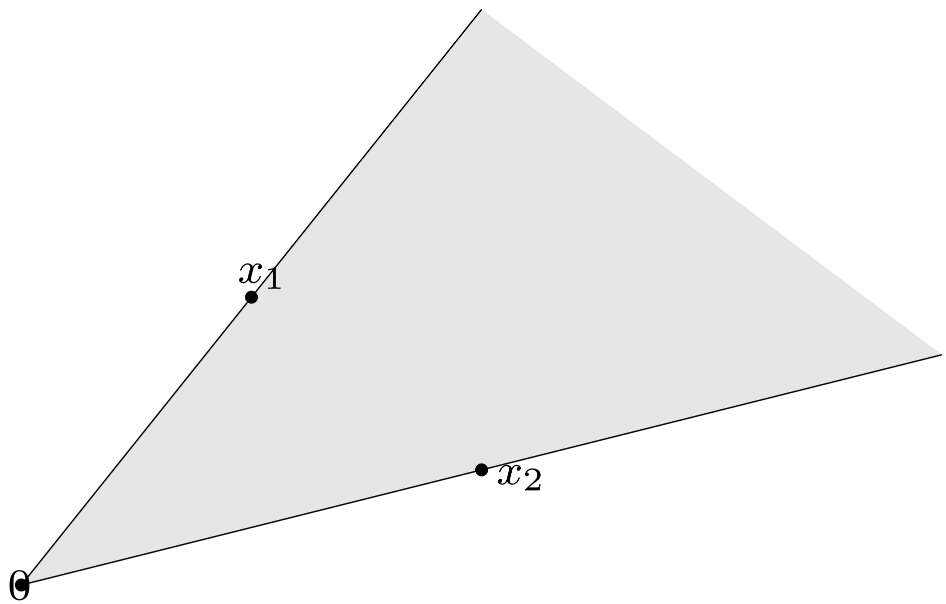
Thanks: Stephen Boyd (Some materials from Boyd and Vandenberghe)

Cones and Convex Cones

C is a cone if for every $x \in C$ and $\theta \geq 0$, we have $\theta x \in C$

C is a convex cone if it is convex and a cone: for any $x_1, x_2 \in C$ and $\theta_1, \theta_2 \geq 0$

$$\theta_1 x_1 + \theta_2 x_2 \in C$$



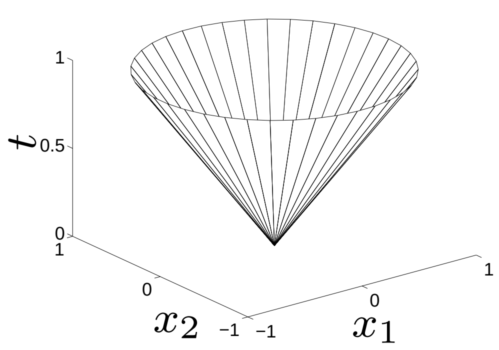
Norm Cones

Given a norm, norm cone is a convex cone:

$$C = \{(x, t) \in \mathbf{R}^{n+1} | \|x\| \leq t\}$$

Example: second order cone:

$$\begin{aligned} C &= \{(x, t) \in \mathbf{R}^{n+1} | \|x\|_2 \leq t\} \\ &= \left\{ \begin{bmatrix} x \\ t \end{bmatrix} | \begin{bmatrix} x \\ t \end{bmatrix}^T \begin{bmatrix} I & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ t \end{bmatrix} \leq 0, t \geq 0 \right\} \end{aligned}$$



Positive Semidefinite Cone

Matrix $A \in \mathbf{R}^{n \times n}$ is positive semidefinite $A \succeq 0$ if for all $x \in \mathbf{R}^n$,

$$x^T A x \geq 0$$

Matrix $A \in \mathbf{R}^{n \times n}$ is positive definite $A \succ 0$ if for all $x \in \mathbf{R}^n$,

$$x^T A x > 0$$

Set of symmetric positive semidefinite matrices:

$$\mathbf{S}_+^n = \{X \in \mathbf{R}^{n \times n} | X = X^T, X \succeq 0\}$$

$\mathbf{S}_+^n$ is a convex cone: if $\theta_1, \theta_2 \geq 0$ and $A, B \in \mathbf{S}_+^n$, then $\theta_1 A + \theta_2 B \in \mathbf{S}_+^n$, since for all $x \in \mathbf{R}^n$:

$$x^T(\theta_1 A + \theta_2 B)x = \theta_1 x^T A x + \theta_2 x^T B x \geq 0$$

Proper Cones and Generalized Inequalities

A cone K is a proper cone if

- K is convex
- K is closed
- K has nonempty interior
- K has no lines $(x \in K, -x \in K \Rightarrow x = 0)$

Proper cone K induces a generalized inequality (partial ordering on $\mathbf{R}^n$):

$$x \preceq_K y \quad \Leftrightarrow \quad y - x \in K$$

$$x \prec_K y \quad \Leftrightarrow \quad y - x \in \textbf{int}\ K$$

Examples

- Nonnegative orthant and componentwise inequality:

$K = \mathbf{R}^n_+$ is a proper cone

$x \preceq_K y$ means $x_i \leq y_i, \;\; i = 1, \ldots, n$

$x \prec_K y$ means $x_i < y_i, \;\; i = 1, \ldots, n$

- Positive semidefinite cone and matrix inequality:

$K = \mathbf{S}^n_+$ is a proper cone in the set of symmetric matrices $\mathbf{S}^n$

$X \preceq_K Y$ means $Y - X$ is positive semidefinite

$X \prec_K Y$ means $Y - X$ is positive definite.

Properties of Generalized Inequalities

- If $x \preceq_K y$ and $u \preceq_K v$, then $x + u \preceq_K y + v$
- If $x \preceq_K y$ and $y \preceq_K z$, then $x \preceq_K z$
- If $x \preceq_K y$ and $\alpha \geq 0$, then $\alpha x \preceq_K \alpha y$
- If $x \preceq_K y$ and $y \preceq_K x$, then $x = y$
- If $x_i \preceq_K y_i$ for $i = 1, \ldots,$ and $x_i \to x$ and $y_i \to y$ as $i \to \infty$, then $x \preceq_K y$

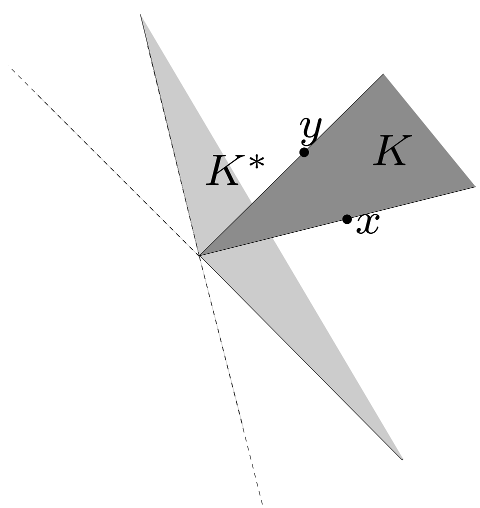
Dual Cones

Given a cone K. Dual cone of K:

$$K^* = \{y | x^T y \geq 0 \ \forall x \in K\}$$

K^* is always a convex cone

K is proper cone $\Rightarrow K^*$ is a proper cone and $K^{**} = K$



- Nonnegative orthant cone is self-dual

Dual PSD Cone

Consider inner product $X^TY = \mathbf{tr}(XY) = \sum_{i,j} X_{ij}Y_{ij}$ on $\mathbf{S}^n$. PSD cone $\mathbf{S}^n_+$ is self-dual:

$$\mathbf{tr}(XY) \geq 0 \ \forall X \succeq 0 \Leftrightarrow Y \succeq 0$$

Proof: Forward direction: suppose $Y \not\succeq 0$. Then $\exists q \in \mathbf{R}^n$ such that $q^TYq = \mathbf{tr}(qq^TY) < 0$. Therefore, $X = qq^T$ satisfies $\mathbf{tr}(XY) < 0$.

Reverse direction: suppose $X, Y \succeq 0$. Express X by eigenvalue decomposition: $X = \sum_{i=1}^n \lambda_i q_i q_i^T$ where $\lambda_i \geq 0$, $i = 1, \ldots, n$. Then

$$\mathbf{tr}XY = \mathbf{tr}\left(Y \sum_{i=1}^n \lambda_i q_i q_i^T\right) = \sum_{i=1}^n \lambda_i q_i^T Y q_i \geq 0$$

Dual Generalized Inequality

Given proper cone K and generalized inequality $\succeq_K$. Dual cone K^* is also proper and induces dual generalized inequality $\succeq_{K^*}$

Relationship between $\succeq_K$ and $\succeq_{K^*}$:

1. $x \preceq_K y$ iff $\lambda^T x \leq \lambda^T y$ for all $\lambda \succeq_{K^*} 0$
2. $x \prec_K y$ iff $\lambda^T x < \lambda^T y$ for all $\lambda \succeq_{K^*} 0, \lambda \neq 0$

Since $K^{**} = K$, these properties hold if $\succeq_K$ and $\succeq_{K^*}$ are interchanged

Generalized Inequality Induced Monotinicity

$f : \mathbf{R}^n \to \mathbf{R}$ is K-nondecreasing (increasing) if

$$x \preceq_K y \Rightarrow f(x) \leq (<) f(y)$$

First order condition for differentiable f with convex domain: f is K-nondecreasing iff:

$$\nabla f(x) \succeq_{K^*} 0$$

f is K-increasing if:

$$\nabla f(x) \succ_{K^*} 0$$

Example: For PSD cone, $\mathbf{tr}(X^{-1})$ is matrix decreasing on $\mathbf{S}^n_{++}$ and $\det X$ is matrix increasing on $\mathbf{S}^n_+$

Generalized Inequality Induced Convexity

$f : \mathbf{R}^n \to \mathbf{R}^m$ is K-convex if for all x, y and $\theta \in [0, 1]$,

$$f(\theta x + (1-\theta)y) \preceq_K \theta f(x) + (1-\theta)f(y)$$

$f : \mathbf{R}^n \to \mathbf{R}^m$ is strictly K-convex if for all $x \neq y$ and $\theta \in (0, 1)$,

$$f(\theta x + (1-\theta)y) \prec_K \theta f(x) + (1-\theta)f(y)$$

First order condition: For differentiable f with convex domain, f is K-convex iff for all $x, y \in \mathbf{dom}\, f$,

$$f(y) \succeq_K f(x) + Df(x)(y - x)$$

Matrix Convexity

f is a symmetric-matrix-valued function. $f : \mathbf{R}^{n\times m} \to \mathbf{S}^m$. f is convex with respect to matrix inequality if for any $x, y \in \mathbf{dom}\, f$ and $\theta \in [0,1]$,

$$f(\theta x + (1-\theta)y) \preceq \theta f(x) + (1-\theta) f(y)$$

Equivalently, f is matrix convex iff scaler-valued function $z^T f(x) z$ is convex for all z

- $f(X) = XX^T$ is matrix convex
- $f(X) = X^p$ is matrix convex for $p \in [1,2]$ or $p \in [-1,0]$, and matrix concave for $p \in [0,1]$
- $f(X) = e^X$ is not matrix convex (unless X is a scalar)

Generalized Inequality Constraints

Convex optimization with generalized inequality constraints on vector-valued functions:

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \preceq_{K_i} 0, \;\; i = 1, 2, \ldots, m \end{array}$$

$f_0 : \mathbf{R}^n \to \mathbf{R}$, $f_i : \mathbf{R}^n \to \mathbf{R}^{k_i}$ are K_i-convex for some proper cones K_i

- Feasible set is convex
- Local optimality $\Rightarrow$ global optimality

Conic Programming

Linear programming with linear generalized inequality constraint:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Fx + G \preceq_K 0 \\ & Ax = b \end{array}$$

• When K is nonnegative orthant, conic program reduces to LP

• When K is PSD cone, write inequality constraints as Linear Matrix Inequalities (LMI):

$$x_1 F_1 + \ldots + x_n F_n + G \preceq 0$$

where $F_i, G \in \mathbf{S}^k$. When they are diagonal, LMI reduces to linear inequalities

SDP

SDP: Minimize linear objective over linear equalities and LMI on variables $x \in \mathbf{R}^n$

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & x_1 F_1 + \ldots + x_n F_n + G \preceq 0 \\ & Ax = b \end{array}$$

SDP in standard form: Minimize a matrix inner product over equality constraints on inner products on variables $X \in \mathbf{S}^n$

$$\begin{array}{ll} \text{minimize} & \mathbf{tr}(CX) \\ \text{subject to} & \mathbf{tr}(A_i X) = b_i, \;\; i = 1, 2, \ldots, p \\ & X \succeq 0 \end{array}$$

LP and SOCP as SDP

LP as SDP:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \mathbf{diag}(Gx - h) \preceq 0 \\ & Ax = b \end{array}$$

SOCP:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \|A_i x + b_i\|_2 \leq c_i^T x + d_i, \quad i = 1, \ldots, N \\ & Fx = g \end{array}$$

SOCP as SDP:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \begin{bmatrix} (c_i x + d_i) I & A_i x + b_i \\ (A_i x + b_i)^T & (c_i x + d_i) I \end{bmatrix} \succeq 0, \quad i = 1, \ldots, N \\ & Fx = g \end{array}$$

Matrix Norm Minimization

$A(x) = A_0 + x_1 A_1 + \dots x_n A_n$ where $A_i \in \mathbf{R}^{p \times q}$. Consider unconstrained spectral norm (max. singular value) minimization over x:

$$\text{minimize} \|A(x)\|_2$$

which is equivalent to convex optimization with LMI on (x, t)

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & A(x)^T A(x) \preceq t^2 I \end{array}$$

which is equivalent to SDP

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \begin{bmatrix} tI & A(x) \\ A(x)^T & tI \end{bmatrix} \succeq 0 \end{array}$$

Related Problems

SDP problems:

- Minimize largest eigenvalue
- Minimize sum of r largest eigenvalues
- Maximize log determinant

SDP approximations:

- Combinatorial optimization
- Rank minimization

ML Estimation

Estimate $x \in \mathbf{R}^n$ from a set of measurements $y \in \mathbf{R}^m$:

$$y_i = a_i^T x + v_i, \;\; i = 1, \ldots, m$$

where v_i are i.i.d. noise with distribution p on $\mathbf{R}$

By independence, likelihood function is

$$\prod_{i=1}^{m} p(y_i - a_i^T x)$$

Maximum Likelihood estimate is any optimal point for the problem in x:

$$\text{maximize} \;\; \sum_{i=1}^{m} \log p(y_i - a_i^T x)$$

Since many distributions are log-concave, the above problem are often convex optimization

Examples

- Gaussian noise: v_i are Gaussian with zero mean and variance σ^2:
$p(z) = (2\pi\sigma^2)^{-1/2}\exp(-z^2/2\sigma^2)$

Log-likelihood function: $-(1/2)\log(2\pi\sigma) - (1/2\sigma^2)\|Ax - y\|_2^2$

$$\hat{x} = \underset{x}{\operatorname{argmin}} \|Ax - y\|_2$$

where rows of A are a_i^T. Solution of least-squares

- Laplacian noise: v_i are Laplacian: $p(z) = (1/2a)\exp(\frac{-|z|}{a}),\ a > 0$

$$\hat{x} = \underset{x}{\operatorname{argmin}} \|Ax - y\|_1$$

Solution of $l - 1$ norm minimization

Covariance Estimation for Gaussian Distribution

$y \in \mathbf{R}^n$ is multivariate Gaussian with zero mean and covariance matrix $R = \mathbf{E}\, yy^T \in \mathbf{S}^n_{++}$ with distribution:

$$p_R(y) = (2\pi)^{-n/2} \det(R)^{-1/2} \exp(-y^T R^{-1} y/2)$$

Goal: estimate covariance R based on N i.i.d. samples $y_1, \ldots, y_N \in \mathbf{R}^n$

Sample covariance matrix

$$Y = \frac{1}{N} \sum_{k=1}^{N} y_k y_k^T$$

Log likelihood function $\log p_R(y_1, \ldots, y_N)$ not concave in R:

$$l(R) = -(Nn/2) \log(2\pi) - (N/2) \log\det R - (N/2)\mathbf{tr}(R^{-1}Y)$$

Information matrix $S = R^{-1}$, log likelihood function is concave in S:

$$l(S) = -(Nn/2) \log(2\pi) + (N/2) \log\det S - (N/2)\mathbf{tr}(SY)$$

Covariance Estimation for Gaussian Distribution

ML estimate of S found by solving convex optimization:

$$\begin{array}{ll} \text{maximize} & \log\det S - \textbf{tr}(SY) \\ \text{subject to} & S \in \mathcal{S}, \;\; S \succeq 0 \end{array}$$

where $\mathcal{S}$ is a convex constraint set of S based on prior information about R

e.g., bounds on R: $L \preceq R \preceq U$ $(U^{-1} \preceq S \preceq L^{-1})$

e.g., condition number constraint on R: $\lambda_{max}(R) \leq \kappa_{max}\lambda_{min}(R)$
$(uI \preceq S \preceq \kappa_{max}uI$, variables: $(S, u))$

Multi-user Detection

Received signal of K-user basic synchronous CDMA channel:

$$y(t) = \sum_{k=1}^{K} A_k b_k s_k(t) + n(t), \quad t \in [0, T]$$

Amplitude A_k, signature waveform $s_k(t)$, information bit $b_k \in \{-1, +1\}$, noise $n(t)$, period T

ML detection: find b that minimizes

$$\begin{aligned} & \int_0^T \left[y(t) - \sum_{k=1}^{K} A_k b_k s_k(t) \right]^2 dt \\ = \quad & \int_0^T \left[\sum_{k=1}^{K} A_k b_k s_k(t) \right] dt - 2 \int_0^T \left[\sum_{k=1}^{K} A_k b_k s_k(t) \right] y(t) dt \\ = \quad & b^T H b - 2 b^T A y \end{aligned}$$

Boolean Constrained QP Formulation

$A = \mathbf{diag}(A)$: amplitude matrix

$R_{ij} = \int_0^T s_i(t)s_j(t)dt$: cross-correlation matrix

$y = RAb + n$: sampled matched filter output

Notation: $H = ARA$ and $p = -2Ay$

$$\begin{array}{ll} \text{minimize} & x^T H x + p^T x \\ \text{subject to} & x_i \in \{-1, +1\}, \;\; i = 1, \ldots, K \end{array}$$

Equivalent form: let $X = xx^T$

$$\begin{array}{ll} \text{subject to} & X_{ii} = 1, \;\; i = 1, \ldots, K, \\ & X \succeq 0, \;\; \text{rank}(X) = 1 \end{array}$$

SDP Relaxation

Neglect rank constraint:

$$\begin{array}{ll} \text{subject to} & X_{ii} = 1, \;\; i = 1, \ldots, K, \\ & X \succeq 0 \end{array}$$

Let $\hat{X} = \begin{bmatrix} xx^T & x \\ x^T & 1 \end{bmatrix}$ and $C = \begin{bmatrix} H & p/2 \\ p^T/2 & 1 \end{bmatrix}$, reformulate as SDP:

$$\begin{array}{ll} \text{minimize} & \mathrm{tr}(C\hat{X}) \\ \text{subject to} & \hat{X}_{ii} = 1, \;\; i = 1, \ldots, K+1, \\ & \hat{X} \succeq 0 \end{array}$$

Factorization and randomization methods to recover solution to original boolean constrained QP from SDP relaxation solutions

Dual Relaxation

Relax $x^T x = K$ to $x^T x \leq K$:

$$\begin{array}{ll} \text{minimize} & x^T H x + p^T x \\ \text{subject to} & x^T x \leq K \end{array}$$

Lagrange dual problem:

$$\begin{array}{ll} \text{minimize} & -\frac{1}{4} p^T (H + \lambda I)^{-1} p - \lambda K \\ \text{subject to} & \lambda \geq 0 \end{array}$$

Gradient-descent method solves this convex optimization in one variable λ

Recover primal optimal solution: $x^* = (H + \lambda^* I)^{-1} A y$

Complexity Performance Tradeoff

Exhaustive search: NP-hard $\mathcal{O}(2^K)$

Successive interference cancellation

Relaxation methods:

- SDP relaxation: polynomial time $\mathcal{O}(K^3)$ (SDP solution dominates randomization in terms of computational load)
- Unconstrained relaxation: analytic solution (related to decorrelator and linear MMSE)
- Bound relaxation
- Dual relaxation

Better complexity-tradeoff possible through duality, randomization, and other relaxation methods?

Lecture Summary

- Proper cones induce generalized inequalities in $\mathbf{R}^n$ and $\mathbf{S}^n$, which induces generalized convex inequality constraints
- Convex optimization with generalized inequalities: conic programming
- SDP is a conic programming over PSD cone with LMI, includes LP, QP, QCQP, SOCP as special cases
- Relaxation for multi-user detection problems

Reading: Sections 2.1, 2.6, 3.6, 4.6, 5.9 in Boyd and Vandenberghe.

L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Review,* March 1996.

W. K. Ma, T. N. Davidson, K. M. Wong, Z. Q. Luo, and P. C. Ching, "Quasi-maximum-likelihood multiuser detection using semidefinite relaxation with applications to synchronous CDMA," *IEEE Trans. Signal Proc.*, vol. 50, no. 4, pp. 912-922, April 2002.