

Lecture: Introduction to LP, SDP and SOCP

Zaiwen Wen

*Beijing International Center For Mathematical Research
Peking University*

<http://bicmr.pku.edu.cn/~wenzw/bigdata2019.html>
wenzw@pku.edu.cn

Acknowledgement: this slides is based on Prof. Farid Alizadeh lecture notes

Outline

- 1 Linear Programming (LP)
- 2 Semidefinite Programming (SDP)
- 3 Second Order Cone Programming (SOCP)
- 4 Conic LP

Linear Programming (LP)

Primal

$$\begin{aligned} \min_{x_i} \quad & c_1x_1 + \dots + c_nx_n \\ \text{s.t.} \quad & a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ & \dots \\ & a_{m1}x_1 + \dots + a_{mn}x_n = b_m \\ & x_i \geq 0 \end{aligned}$$

Dual

$$\begin{aligned} \max_{y_i} \quad & b_1y_1 + \dots + b_my_m \\ \text{s.t.} \quad & a_{11}y_1 + \dots + a_{m1}y_m \leq c_1 \\ & \dots \\ & a_{1n}y_1 + \dots + a_{mn}y_m \leq c_n \end{aligned}$$

Linear Programming (LP)

more succinctly

Primal (P)

$$\begin{aligned} \min_x \quad & c^\top x \\ \text{s.t.} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

Dual (D)

$$\begin{aligned} \max_{y,s} \quad & b^\top y \\ \text{s.t.} \quad & A^\top y + s = c \\ & s \geq 0 \end{aligned}$$

Problems with absolute values

$$\begin{aligned} \min \quad & \sum_i c_i |x_i|, & \text{assume } c \geq 0 \\ \text{s.t.} \quad & Ax \geq b \end{aligned}$$

- Reformulation 1:

$$\begin{aligned} \min \quad & \sum_i c_i z_i & \min \quad & \sum_i c_i z_i \\ \text{s.t.} \quad & Ax \geq b & \iff & \text{s.t.} \quad Ax \geq b \\ & |x_i| \leq z_i & & -z_i \leq x_i \leq z_i \end{aligned}$$

- Reformulation 2: $x_i = x_i^+ - x_i^-$, $x_i^+, x_i^- \geq 0$. Then $|x_i| = x_i^+ + x_i^-$

$$\begin{aligned} \min \quad & \sum_i c_i (x_i^+ + x_i^-) \\ \text{s.t.} \quad & Ax^+ - Ax^- \geq b, x^+, x^- \geq 0 \end{aligned}$$

Problems with absolute values

- data fitting:

$$\min_x \|Ax - b\|_\infty$$

$$\min_x \|Ax - b\|_1$$

- Compressive sensing

$$\min \|x\|_1, \text{ s.t. } Ax = b \quad (LP)$$

$$\min \mu \|x\|_1 + \frac{1}{2} \|Ax - b\|^2 \quad (QP, SOCP)$$

$$\min \|Ax - b\|, \text{ s.t. } \|x\|_1 \leq 1$$

An example of linear programming: 菜鸟

2018 杭州·云栖大会 Alibaba Group

智能履行决策

商家

杭州-上海 杭州-广州 杭州-北京 杭州-武汉 ...

YTO ZTO YUNDA

商家

菜鸟智能发货引擎

时效	服务	成本	单量平衡	...
线路容量	网点容量	局部优化	全局优化	...

最优快递

智能决策
ML & Optimization

$C_{i,j} = c1 * \text{成本} + c2 * \text{服务} + c3 * \text{时效}$

决策变量

$\max_x \quad \sum_{i=1}^n \sum_{j=1}^m C_{i,j} x_{i,j}$ 将订单 i 匹配合适快递公司 j 与否

s.t. $\sum_{j=1}^m x_{i,j} \leq 1$

$\sum_{i=1}^n x_{i,j} * a_j \leq u_j$ 商家发货 CP 总单量比例约束

$\sum_{i=1}^n \sum_{j=1}^m x_{i,j} b_{k,i,j} \leq B_k$ 全局约束值, 比如总成本

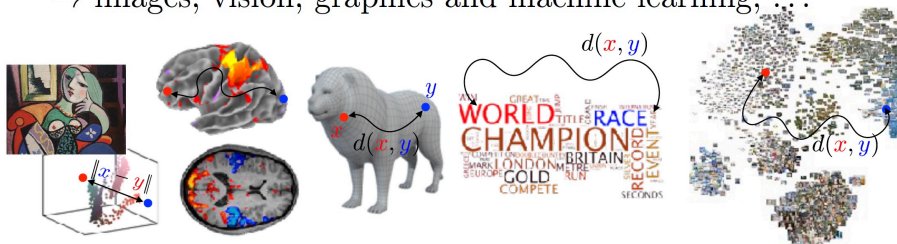
订单的履行是带有全局约束的序列执行决策

- Online assignment problem
- Control based method
- Online linear programming

Ref: Agrawal, Shipra, Zizhuo Wang, and Yinyu Ye. "A dynamic near-optimal algorithm for online linear programming." *Operations Research* 62.4 (2014): 876-890.

Optimal transport

→ images, vision, graphics and machine learning, ...



Monge



Kantorovich



Koopmans



Dantzig



Brenier



Otto



McCann



Villani



Figalli

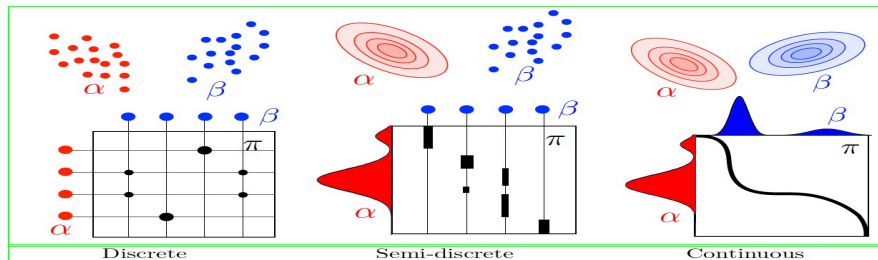
Nobel '75

Fields '10

Fields'18

Optimal transport: LP

$$\begin{aligned} \min_{\pi \in \mathbb{R}^{m \times n}} \quad & \sum_{i=1}^m \sum_{j=1}^n c_{ij} \pi_{ij} \\ \text{s.t.} \quad & \sum_{j=1}^n \pi_{ij} = \mu_i, \quad \forall i = 1, \dots, m, \\ & \sum_{i=1}^m \pi_{ij} = \nu_j, \quad \forall j = 1, \dots, n \\ & \pi \geq 0 \end{aligned}$$



Weak duality

Suppose

- x is feasible to (P)
- (y, s) is feasible to (D)

Then

$$\begin{aligned} 0 &\leq x^\top s \quad \text{because } x_i s_i \geq 0 \\ &= x^\top (c - A^\top y) \\ &= c^\top x - (Ax)^\top y \\ &= c^\top x - b^\top y \\ &= \text{duality gap} \end{aligned}$$

Key Properties of LP

- Strong duality: If both Primal and Dual are feasible then at the optimum

$$c^\top x = b^\top y \iff x^\top s = 0$$

- complementary slackness: This implies

$$\begin{aligned}x^\top s &= x_1 s_1 + \dots + x_n s_n = 0 \quad \text{and therefore} \\x_i s_i &= 0\end{aligned}$$

complementarity

- Putting together primal feasibility, dual feasibility and complementarity together we get a square system of equations

$$\begin{aligned}Ax &= b \\ A^T y + s &= c \\ x_i s_i &= 0 \quad \text{for } i = 1, \dots, n\end{aligned}$$

- At least in principle this system determines the primal and dual optimal values

Outline

- 1 Linear Programming (LP)
- 2 Semidefinite Programming (SDP)**
- 3 Second Order Cone Programming (SOCP)
- 4 Conic LP

Semidefinite Programming (SDP)

- $X \succeq Y$ means that the symmetric matrix $X - Y$ is positive semidefinite
- X is positive semidefinite

$$a^T X a \geq 0 \text{ for all vector } a \iff X = B^T B \iff$$

all eigenvalues of X is nonnegative

Semidefinite Programming (SDP)

- $\langle X, Y \rangle = \sum_{ij} X_{ij}Y_{ij} = \text{Tr}(XY)$

Primal (P)

$$\begin{aligned} \min_{X_i} \quad & \langle C_1, X_1 \rangle + \dots + \langle C_n, X_n \rangle \\ \text{s.t.} \quad & \langle A_{11}, X_1 \rangle + \dots + \langle A_{1n}, X_n \rangle = b_1 \\ & \dots \\ & \langle A_{m1}, X_1 \rangle + \dots + \langle A_{mn}, X_n \rangle = b_m \\ & X_i \succeq 0 \end{aligned}$$

Dual (D)

$$\begin{aligned} \max_{y_i} \quad & b_1 y_1 + \dots + b_m y_m \\ \text{s.t.} \quad & A_{11}^\top y_1 + \dots + A_{m1}^\top y_m + S_1 = c_1 \\ & \dots \\ & A_{1n}^\top y_1 + \dots + A_{mn}^\top y_m + S_n = c_n \\ & S_i \succeq 0 \end{aligned}$$

Simplified SDP

- For simplicity we deal with single variable SDP:

Primal (P)

$$\begin{aligned} \min_X \quad & \langle C, X \rangle \\ \text{s.t.} \quad & \langle A_1, X \rangle = b_1 \\ & \dots \\ & \langle A_m, X \rangle = b_m \\ & X \succeq 0 \end{aligned}$$

Dual (D)

$$\begin{aligned} \max_{y, S} \quad & b^\top y \\ \text{s.t.} \quad & \sum_i y_i A_i + S = C \\ & S \succeq 0 \end{aligned}$$

- A single variable LP is trivial
- But a single matrix SDP is as general as a multiple matrix

Facts on matrix calculation

- If $A, B \in \mathbb{R}^{m \times n}$, then $\text{Tr}(AB^\top) = \text{Tr}(B^\top A)$
- If $U, V \in \mathcal{S}^n$ and Q is orthogonal, then $\langle U, V \rangle = \langle Q^\top U Q, Q^\top V Q \rangle$
- If $X \in \mathcal{S}^n$, then $U = Q^\top \Lambda Q$, where $Q^\top Q = I$ and Λ is diagonal.
- Matrix norms: $\|X\|_F = \|\lambda(X)\|_2$, $\|X\|_2 = \|\lambda(X)\|_\infty$, $\lambda(X) = \text{diag}(\Lambda)$
- $X \succeq 0 \iff v^\top X v \geq 0$ for all $v \in \mathbb{R}^n \iff \lambda(X) \geq 0 \iff X = B^\top B$
- The dual cone of \mathcal{S}_+^n is \mathcal{S}_+^n
- If $X \succeq 0$, then $X_{ii} \geq 0$. If $X_{ii} = 0$, then $X_{ik} = X_{ki} = 0$ for all k .
- If $X \succeq 0$, then $PXP^\top \succeq 0$ for any P of appropriate dimensions
- If $X = \begin{pmatrix} X_{11} & X_{12} \\ X_{12}^\top & X_{22} \end{pmatrix} \succeq 0$, then $X_{11} \succeq 0$.
- $X \succeq 0$ iff every principal submatrix is positive semidefinite (psd).

Facts on matrix calculation

- Let $U = \begin{pmatrix} A & B \\ B^\top & C \end{pmatrix}$ with A and C symmetric and $A \succ 0$. Then

$$U \succeq 0 \text{ (or } \succ 0) \iff C - B^\top A^{-1} B \succeq 0 \text{ (or } \succ 0).$$

The matrix $C - B^\top A^{-1} B$ is the **Schur complement** of A in U :

$$\begin{pmatrix} A & B \\ B^\top & C \end{pmatrix} = \begin{pmatrix} I & 0 \\ B^\top A^{-1} & I \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & C - B^\top A^{-1} B \end{pmatrix} \begin{pmatrix} I & A^{-1} B \\ 0 & I \end{pmatrix}$$

- If $A \in \mathcal{S}^n$, then $x^\top A x = \langle A, x x^\top \rangle$
- If $A \succ 0$, then $\langle A, B \rangle > 0$ for every nonzero $B \succeq 0$ and $\{B \succeq 0 \mid \langle A, B \rangle \leq \beta\}$ is bounded for $\beta > 0$
- If $A, B \succeq 0$, then $\langle A, B \rangle = 0$ iff $AB = 0$
- $A, B \in \mathcal{S}^n$, then A and B commute iff AB is symmetric, iff A and B can be simultaneously diagonalized

Eigenvalue optimization

- minimizing the largest eigenvalue $\lambda_{\max}(A_0 + \sum_i x_i A_i)$:

$$\min \lambda_{\max}(A_0 + \sum_i x_i A_i)$$

can be expressed as an SDP

and its dual is

$$\min z$$

$$\text{s.t. } zI - \sum_i x_i A_i \succeq A_0$$

$$\max \langle A_0, Y \rangle$$

$$\text{s.t. } \langle A_i, Y \rangle = k$$

$$\langle I, Y \rangle = 1$$

$$Y \succeq 0$$

- follows from

$$\lambda_{\max}(A) \leq t \iff A \preceq tI$$

Eigenvalue optimization

- Let $A_i \in \mathbb{R}^{m \times n}$. Minimizing the 2-norm of $A(x) = A_0 + \sum_i x_i A_i$:

$$\min_x \|A(x)\|_2$$

can be expressed as an SDP

$$\begin{aligned} \min_{x,t} \quad & t \\ \text{s.t.} \quad & \begin{pmatrix} tI & A(x) \\ A(x)^\top & tI \end{pmatrix} \succeq 0 \end{aligned}$$

- Constraint follows from

$$\begin{aligned} \|A\|_2 \leq t & \iff A^\top A \preceq t^2 I, \quad t \geq 0 \\ & \iff \begin{pmatrix} tI & A(x) \\ A(x)^\top & tI \end{pmatrix} \succeq 0 \end{aligned}$$

Eigenvalue optimization

- Let $\Lambda_k(A)$ indicate sum of the k largest eigenvalues of A . Then minimizing $\Lambda_k(A_0 + \sum_i x_i A_i)$:

$$\min \quad \Lambda_k(A_0 + \sum_i x_i A_i)$$

can be expressed as an SDP

$$\begin{aligned} \min \quad & kz + \text{Tr}(X) \\ \text{s.t.} \quad & zI + X - \sum_i x_i A_i \succeq A_0 \\ & X \succeq 0 \end{aligned}$$

since

$$\Lambda_k(A) \leq t \iff t - kz - \text{Tr}(X) \geq 0, zI + X \succeq A, X \succeq 0$$

The following problems can be expressed as SDP

- maximizing sum of the k smallest eigenvalues of $A_0 + \sum_i x_i A_i$
- minimizing sum of the k absolute-value-wise largest eigenvalues
- minimizing sum of the k largest singular values of $A_0 + \sum_i x_i A_i$

Quadratically Constrained Quadratic Programming

Consider QCQP

$$\begin{aligned} \min \quad & x^\top A_0 x + 2b_0^\top x + c_0 \quad \text{assume } A_i \in \mathcal{S}^n \\ \text{s.t.} \quad & x^\top A_i x + 2b_i^\top x + c_i \leq 0, \quad i = 1, \dots, m \end{aligned}$$

- If $A_0 \succ 0$ and $A_i = B_i^\top B_i$, $i = 1, \dots, m$, then it is a SOCP
- If $A_i \in \mathcal{S}^n$ but may be indefinite

$$x^\top A_i x + 2b_i^\top x + c_i = \langle A_i, xx^\top \rangle + 2b_i^\top x + c_i$$

- The original problem is equivalent to

$$\begin{aligned} \min \quad & \text{Tr}A_0 X + 2b_0^\top x + c_0 \\ \text{s.t.} \quad & \text{Tr}A_i X + 2b_i^\top x + c_i \leq 0, \quad i = 1, \dots, m \\ & X = xx^\top \end{aligned}$$

- If $A_i \in S^n$ but may be indefinite

$$x^\top A_i x + 2b_i^\top x + c_i = \left\langle \begin{pmatrix} A_i & b_i \\ b_i^\top & c_i \end{pmatrix}, \begin{pmatrix} X & x \\ x^\top & 1 \end{pmatrix} \right\rangle := \langle \bar{A}_i, \bar{X} \rangle$$

$\bar{X} \succeq 0$ is equivalent to $X \succeq xx^\top$

- The SDP relaxation is

$$\begin{aligned} \min \quad & \text{Tr} \bar{A}_0 \bar{X} \\ \text{s.t.} \quad & \text{Tr} \bar{A}_i \bar{X} \leq 0, \quad i = 1, \dots, m \\ & \bar{X} \succeq 0 \end{aligned}$$

- Maxcut: $\max x^\top Wx, \quad \text{s.t.} \quad x_i^2 = 1$
- Phase retrieval: $|a_i^\top x| = b_i$, the value of $a_i^\top x$ is complex

Max cut

- For graph (V, E) and weights $w_{ij} = w_{ji} \geq 0$, the maxcut problem is

$$(Q) \quad \max_x \frac{1}{2} \sum_{i < j} w_{ij} (1 - x_i x_j), \quad \text{s.t. } x_i \in \{-1, 1\}$$

- Relaxation:

$$(P) \quad \max_{v_i \in \mathbb{R}^n} \frac{1}{2} \sum_{i < j} w_{ij} (1 - v_i^\top v_j), \quad \text{s.t. } \|v_i\|_2 = 1$$

- Equivalent SDP of (P):

$$(SDP) \quad \max_{X \in \mathcal{S}^n} \frac{1}{2} \sum_{i < j} w_{ij} (1 - X_{ij}), \quad \text{s.t. } X_{ii} = 1, X \succeq 0$$

Max cut: rounding procedure

Goemans and Williamson's randomized approach

- Solve (SDP) to obtain an optimal solution X . Compute the decomposition $X = V^T V$, where

$$V = [v_1, v_2, \dots, v_n]$$

- Generate a vector r uniformly distributed on the unit sphere, i.e., $\|r\|_2 = 1$
- Set

$$x_i = \begin{cases} 1 & v_i^T r \geq 0 \\ -1 & \text{otherwise} \end{cases}$$

Max cut: theoretical results

- Let W be the objective function value of x and $E(W)$ be the expected value. Then

$$E(W) = \frac{1}{\pi} \sum_{i < j} w_{ij} \arccos(v_i^\top v_j)$$

- Goemans and Williamson showed:

$$E(W) \geq \alpha \frac{1}{2} \sum_{i < j} w_{ij} (1 - v_i^\top v_j)$$

where

$$\alpha = \min_{0 \leq \theta < \pi} \frac{2}{\pi} \frac{\theta}{1 - \cos \theta} > 0.878$$

- Let $Z_{(SDP)}^*$ and $Z_{(Q)}^*$ be the optimal values of (SDP) and (Q)

$$E(W) \geq \alpha Z_{(SDP)}^* \geq \alpha Z_{(Q)}^*$$

Weak duality in SDP

- Just as in LP

$$\langle X, S \rangle = \langle C, X \rangle - b^\top y$$

- Also if both $X \succeq 0$ and $S \succeq 0$ then

$$\langle X, S \rangle = \text{Tr}(XS^{1/2}S^{1/2}) = \text{Tr}(S^{1/2}XS^{1/2}) \geq 0$$

because $S^{1/2}XS^{1/2} \succeq 0$

- Thus

$$\langle X, S \rangle = \langle C, X \rangle - b^\top y \geq 0$$

Complementarity Slackness Theorem

- $X \succeq 0$ and $S \succeq 0$ and $\langle X, S \rangle = 0$ implies

$$XS = 0$$

- Proof:

$$\langle X, S \rangle = \text{Tr}(XS^{1/2}S^{1/2}) = \text{Tr}(S^{1/2}XS^{1/2})$$

Thus $\text{Tr}(S^{1/2}XS^{1/2}) = 0$. Since $S^{1/2}XS^{1/2} \succeq 0$, then

$$\begin{aligned} S^{1/2}XS^{1/2} = 0 &\implies S^{1/2}X^{1/2}X^{1/2}S^{1/2} = 0 \\ X^{1/2}S^{1/2} = 0 &\implies XS = 0 \end{aligned}$$

Equivalent complementarity slackness

- For reasons to become clear later it is better to write complementary slackness conditions as

$$\frac{XS + SX}{2} = 0$$

- It can be shown that if $X \succeq 0$ and $S \succeq 0$, then $XS = 0$ iff

$$XS + SX = 0$$

Constraint Qualifications

- Unlike LP we need some conditions for the optimal values of Primal and Dual SDP to coincide
- Here are two:
 - If there is primal-feasible $X \succ 0$ (i.e. X is positive definite)
 - If there is dual-feasible $S \succ 0$
- When strong duality holds $\langle X, S \rangle = 0$

KKT Condition

- Thus just like LP The system of equations

$$\begin{aligned}\langle A_i, X \rangle &= b_i, \quad \text{for } i = 1, \dots, m \\ \sum_i y_i A_i + S &= C \\ X \circ S &= 0\end{aligned}$$

Gives us a square system

Outline

- 1 Linear Programming (LP)
- 2 Semidefinite Programming (SDP)
- 3 Second Order Cone Programming (SOCP)**
- 4 Conic LP

Second Order Cone Programming (SOCP)

- For simplicity we deal with single variable SOCP:

Primal (P)

$$\min \quad c^\top x$$

$$\text{s.t.} \quad Ax = b$$

$$x_Q \succeq 0$$

Dual (D)

$$\max \quad b^\top y$$

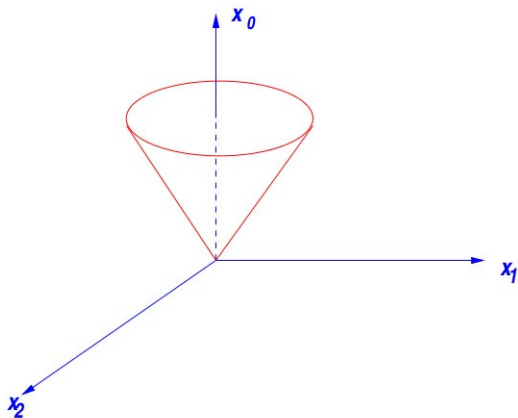
$$\text{s.t.} \quad A^\top y + s = c$$

$$s_Q \succeq 0$$

- the vectors x, s, c are indexed from zero
- If $z = (z_0, z_1, \dots, z_n)^\top$ and $\bar{z} = (z_1, \dots, z_n)^\top$

$$z_Q \geq 0 \iff z_0 \geq \|\bar{z}\|$$

Illustration of SOC



$$\mathcal{Q} = \{z \mid z_0 \geq \|\bar{z}\|\}$$

Quadratic Programming (QP)

$$\begin{array}{ll} \min & q(x) = x^\top Qx + a^\top x + \beta \quad \text{assume } Q \succ 0, Q = Q^\top \\ \text{s.t.} & Ax = b \\ & x \geq 0 \end{array}$$

- $q(x) = \|\bar{u}\|^2 + \beta - \frac{1}{4}a^\top Q^{-1}a$, where $\bar{u} = Q^{1/2}x + \frac{1}{2}Q^{-1/2}a$.
- equivalent SOCP

$$\begin{array}{ll} \min & u_0 \\ \text{s.t.} & \bar{u} = Q^{1/2}x + \frac{1}{2}Q^{-1/2}a \\ & Ax = b \\ & x \geq 0, \quad (u_0, \bar{u}) \succeq_Q 0 \end{array}$$

Quadratic constraints

$$q(x) = x^\top B^\top Bx + a^\top x + \beta \leq 0$$

is equivalent to

$$(u_0, \bar{u}) \succeq_{\mathcal{Q}} 0,$$

where

$$\bar{u} = \begin{pmatrix} Bx \\ \frac{a^\top x + \beta + 1}{2} \end{pmatrix} \quad \text{and} \quad u_0 = \frac{1 - a^\top x - \beta}{2}$$

Norm minimization problems

Let $\bar{v}_i = A_i x + b_i \in \mathbb{R}^{n_i}$.

- $\min_x \sum_i \|\bar{v}_i\|$ is equivalent to

$$\begin{aligned} \min \quad & \sum_i v_{i0} \\ \text{s.t.} \quad & \bar{v}_i = A_i x + b_i \\ & (v_{i0}, \bar{v}_i) \succeq_{\mathcal{Q}} 0 \end{aligned}$$

- $\min_x \max_{1 \leq i \leq r} \|\bar{v}_i\|$ is equivalent to

$$\begin{aligned} \min \quad & t \\ \text{s.t.} \quad & \bar{v}_i = A_i x + b_i \\ & (t, \bar{v}_i) \succeq_{\mathcal{Q}} 0 \end{aligned}$$

Norm minimization problems

Let $\bar{v}_i = A_i x + b_i \in \mathbb{R}^{n_i}$.

- $\|\bar{v}_{[1]}\|, \dots, \|\bar{v}_{[r]}\|$ are the norms $\|\bar{v}_1\|, \dots, \|\bar{v}_r\|$ sorted in nonincreasing order
- $\min_x \sum_i \|\bar{v}_{[i]}\|$ is equivalent to

$$\min \sum_i u_i + kt$$

$$\text{s.t. } \bar{v}_i = A_i x + b_i$$

$$\|\bar{v}_i\| \leq u_i + t$$

$$u_i \geq 0$$

Rotated Quadratic Cone

- rotated cone $w^\top w \leq xy$, where $x, y \geq 0$, is equivalent to

$$\left\| \begin{pmatrix} 2w \\ x - y \end{pmatrix} \right\| \leq x + y$$

- Minimize the harmonic mean of positive affine functions

$$\min \sum_i 1/(a_i^\top x + \beta_i), \text{ s.t. } a_i^\top x + \beta_i > 0$$

is equivalent to

$$\begin{aligned} \min \quad & \sum_i u_i \\ \text{s.t.} \quad & \bar{v}_i = a_i^\top x + \beta_i \\ & 1 \leq u_i \bar{v}_i \\ & u_i \geq 0 \end{aligned}$$

- Logarithmic Tchebychev approximation

$$\min_x \max_{1 \leq i \leq r} |\ln(a_i^\top x) - \ln b_i|$$

Since $|\ln(a_i^\top x) - \ln b_i| = \ln \max(a_i^\top x/b_i, b_i/a_i^\top x)$, the problem is equivalent to

$$\begin{aligned} \min \quad & t \\ \text{s.t.} \quad & 1 \leq (a_i^\top x/b_i)t \\ & a_i^\top x/b_i \leq t \\ & t \geq 0 \end{aligned}$$

- Inequalities involving geometric means

$$\left(\prod_{i=1}^n (a_i^\top x + b_i) \right)^{1/n} \geq t$$

- $n=4$

$$\max \prod_{i=1}^4 (a_i^\top x - b_i) \iff$$

$$\max \quad w_3$$

$$\text{s.t.} \quad a_i^\top x - b_i \geq 0$$

$$(a_1^\top x - b_1)(a_2^\top x - b_2) \geq w_1^2$$

$$(a_3^\top x - b_3)(a_4^\top x - b_4) \geq w_2^2$$

$$w_1 w_2 \geq w_3^2$$

$$w_i \geq 0$$

- This can be extended to products of rational powers of affine functions

Robust linear programming

the parameters in LP are often uncertain

$$\begin{array}{ll} \min & c^\top x \\ \text{s.t.} & a_i^\top x \leq b_i \end{array}$$

There can be uncertainty in c, a_i, b .

two common approaches to handling uncertainty (in a_i , for simplicity)

- deterministic model: constraints must hold for all $a_i \in \mathcal{E}_i$

$$\begin{array}{ll} \min & c^\top x \\ \text{s.t.} & a_i^\top x \leq b_i, \text{ for all } a_i \in \mathcal{E}_i \end{array}$$

- stochastic model: a_i is random variable; constraints must hold with probability η

$$\begin{array}{ll} \min & c^\top x \\ \text{s.t.} & \text{prob}(a_i^\top x \leq b_i) \geq \eta \end{array}$$

deterministic approach via SOCP

- Choose an ellipsoid as \mathcal{E}_i :

$$\mathcal{E}_i = \{\bar{a}_i + P_i u \mid \|u\|_2 \leq 1\}, \quad \bar{a}_i \in \mathbb{R}^n, P_i \in \mathbb{R}^{n \times n}$$

- Robust LP

$$\begin{aligned} \min \quad & c^\top x \\ \text{s.t.} \quad & a_i^\top x \leq b_i, \text{ for all } a_i \in \mathcal{E}_i \end{aligned}$$

is equivalent to the SOCP

$$\begin{aligned} \min \quad & c^\top x \\ \text{s.t.} \quad & \bar{a}_i^\top x + \|P_i^\top x\|_2 \leq b_i \end{aligned}$$

since

$$\sup_{\|u\|_2 \leq 1} (\bar{a}_i + P_i u)^\top x = \bar{a}_i^\top x + \|P_i^\top x\|_2$$

stochastic approach via SOCP

- a_i is Gaussian with mean \bar{a}_i , covariance Σ_i ($a_i \sim \mathcal{N}(\bar{a}_i, \Sigma_i)$)
- $a_i^\top x$ is Gaussian r.v. with mean $\bar{a}_i^\top x$, variance $x^\top \Sigma_i x$; hence

$$\text{prob}(a_i^\top x \leq b_i) = \Phi \left(\frac{b_i - \bar{a}_i^\top x}{\|\Sigma_i^{1/2} x\|_2} \right)$$

where $\Phi(x) = (1/\sqrt{2\pi}) \int_{-\infty}^x e^{-t^2/2} dt$ is CDF of $\mathcal{N}(0, 1)$

- robust LP

$$\begin{aligned} \min \quad & c^\top x \\ \text{s.t.} \quad & \text{prob}(a_i^\top x \leq b_i) \geq \eta \end{aligned}$$

is equivalent to the SOCP

$$\begin{aligned} \min \quad & c^\top x \\ \text{s.t.} \quad & \bar{a}_i^\top x + \Phi^{-1}(\eta) \|\Sigma_i^{1/2} x\|_2 \leq b_i \end{aligned}$$

Weak Duality in SOCP

- The single block SOCP is not as trivial as LP but it still can be solved analytically
- weak duality: Again as in LP and SDP

$$x^\top s = c^\top x - b^\top y = \text{duality gap}$$

If $x, s \succeq_{\mathcal{Q}} 0$, then

$$\begin{aligned}x^\top s &= x_0 s_0 + \bar{x}^\top \bar{s} \geq \\ &\geq \|\bar{x}\| \cdot \|\bar{s}\| + \bar{x}^\top \bar{s} \quad \text{since } x, s \succeq_{\mathcal{Q}} 0 \\ &\geq |\bar{x}^\top \bar{s}| + \bar{x}^\top \bar{s} \quad \text{Cauchy-Schwartz inequality} \\ &\geq 0\end{aligned}$$

Complementary Slackness for SOCP

- Given $x \succeq_{\mathcal{Q}} 0$, $s \succeq_{\mathcal{Q}} 0$ and $x^\top s = 0$. Assume $x_0 > 0$ and $s_0 > 0$
- We have

$$(*) \quad x_0^2 \geq \sum_{i=1}^n x_i^2$$

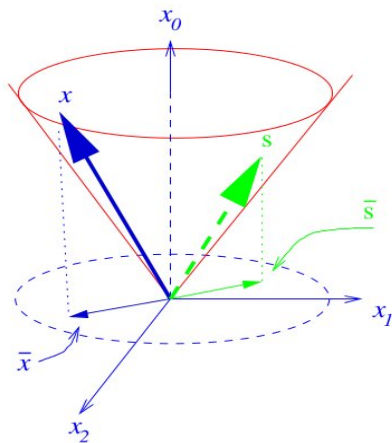
$$(**) \quad s_0^2 \geq \sum_{i=1}^n s_i^2 \iff x_0^2 \geq \sum_{i=1}^n \frac{s_i^2 x_0^2}{s_0^2}$$

$$(***) \quad x^\top s = 0 \iff -x_0 s_0 = \sum_i x_i s_i \iff -2x_0^2 = \sum_{i=1}^n \frac{2x_i s_i x_0}{s_0}$$

- Adding (*), (**), (***), we get $0 \geq \sum_{i=1}^n \left(x_i + \frac{s_i x_0}{s_0} \right)^2$
- This implies

$$x_i s_0 + x_0 s_i = 0, \text{ for } i = 1, \dots, n$$

Illustration of SOC



When $x \succeq_{\mathcal{Q}} 0$, $s \succeq_{\mathcal{Q}} 0$ are orthogonal both must be on the boundary in such a way that their projection on the x_1, \dots, x_n plane is collinear

Strong Duality

- at the optimum

$$c^\top x = b^\top y \iff x^\top s = 0$$

- Like SDP constraint qualifications are required
- If there is primal-feasible $x \succ_Q 0$
- If there is dual-feasible $s \succ_Q 0$

Complementary Slackness for SOCP

- Thus again we have a square system

$$\begin{aligned}Ax &= b, \\A^\top y + s &= c \\x^\top s &= 0 \\x_0 s_i + s_0 x_i &= 0\end{aligned}$$

Outline

- 1 Linear Programming (LP)
- 2 Semidefinite Programming (SDP)
- 3 Second Order Cone Programming (SOCP)
- 4 Conic LP**

A set $K \subseteq \mathbb{R}^n$ is a proper cone if

- It is a cone: If $x \in K \implies \alpha x \in K$ for all $\alpha \geq 0$
- It is convex: $x, y \in K \implies \alpha x + (1 - \alpha)y \in K$ for $\alpha \in [0, 1]$
- It is pointed: $K \cap (-K) = \{0\}$
- It is closed
- It has non-empty interior in \mathbb{R}^n
- dual cone:

$$K^* = \{x \mid \text{for all } z \in K, \langle x, z \rangle \geq 0\}$$

Conic LP

Conic-LP is defined as the following optimization problem:

Primal (P)

$$\begin{aligned} \min \quad & c^\top x \\ \text{s.t.} \quad & Ax = b \\ & x \in K \end{aligned}$$

Dual (D)

$$\begin{aligned} \max \quad & b^\top y \\ \text{s.t.} \quad & A^\top y + s = c \\ & s \in K^* \end{aligned}$$

- For LP K is the nonnegative orthant
- For SDP K is the cone of positive semidefinite matrices
- For SOCP K is the circular or Lorentz cone
- In all three cases the cones are self-dual $K = K^*$