Lecture: Introduction to LP, SDP and SOCP

Zaiwen Wen

Beijing International Center For Mathematical Research
Peking University

http://bicmr.pku.edu.cn/~wenzw/bigdata2015.html
wenzw@pku.edu.cn

Linear Programming (LP)

Primal

$\min \quad c_1 x_1 + \ldots + c_n x_n$

$\text{s.t.} \quad a_{11} x_1 + \ldots + a_{1n} x_n = b_1$

$\qquad \ldots$

$\qquad a_{m1} x_1 + \ldots + a_{mn} x_n = b_m$

$\qquad x_i \geq 0$

Dual

$\max \quad b_1 y_1 + \ldots + b_m y_m$

$\text{s.t.} \quad a_{11} y_1 + \ldots + a_{m1} y_m \leq c_1$

$\qquad \ldots$

$\qquad a_{1n} y_1 + \ldots + a_{mn} y_m \leq c_n$

Linear Programming (LP)

more succinctly

Primal (P)	**Dual (D)**
$\min \quad c^\top x$	$\max \quad b^\top y$
$\text{s.t.} \quad Ax = b$	$\text{s.t.} \quad A^\top y + s = c$
$\qquad \quad x \geq 0$	$\qquad \quad s \geq 0$

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

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

- complementary slackness: This implies

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

complementarity

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

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

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

Algebraic characterization

- We can define $x \circ s = (x_1 s_1, \ldots, x_n s_n)^\top$ and

$$\mathsf{L}_x : y \to (x_1 y_1, \ldots, x_n y_n)^\top \text{ i.e. } \mathsf{L}_x = \mathrm{Diag}(x)$$

- We can write complementary slackness conditions as

$$x \circ s = \mathsf{L}_x s = \mathsf{L}_x \mathsf{L}_s 1 = 0$$

- 1, the vector of all ones, is the identity element:

$$x \circ 1 = x$$

Semidefinite Programming (SDP)

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

- X is positive semidefinite

$$a^\top X a \geq 0 \text{ for all vector } a \Longleftrightarrow X = B^\top B \Longleftrightarrow$$

all eigenvalues of X is nonnegative

Semidefinite Programming (SDP)

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

Primal (P)	**Dual (D)**

min $\quad \langle C_1, X_1 \rangle + \ldots + \langle C_n, X_n \rangle$

s.t. $\quad \langle A_{11}, X_1 \rangle + \ldots + \langle A_{1n}, X_n \rangle = b_1$

$\quad \ldots$

$\quad \langle A_{m1}, X_1 \rangle + \ldots + \langle A_{mn}, X_n \rangle = b_m$

$\quad X_i \succeq 0$

max $\quad b_1 y_1 + \ldots + b_m y_m$

s.t. $\quad A_{11} y_1 + \ldots + A_{m1} y_m + S_1 = c_1$

$\quad \ldots$

$\quad A_{1n} y_1 + \ldots + A_{mn} y_m + S_n = c_n$

$\quad S_i \succeq 0$

Simplified SDP

- For simplicity we deal with single variable SDP:

| **Primal (P)** | **Dual (D)** |

$$\min \quad \langle C, X \rangle$$
$$\text{s.t.} \quad \langle A_1, X \rangle = b_1$$
$$\cdots$$
$$\langle A_m, X \rangle = b_m$$
$$X \succeq 0$$

$$\max \quad b^\top y$$
$$\text{s.t.} \quad \sum_i y_i A_i + S = C$$
$$S \succeq 0$$

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

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

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

$$S^{1/2}XS^{1/2} = 0 \implies S^{1/2}X^{1/2}X^{1/2}S^{1/2} = 0$$
$$X^{1/2}S^{1/2} = 0 \implies XS = 0$$

Algebraic properties of SDP

- 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

$$XS + SX = 0$$

Algebraic properties of SDP

- Definition: $X \circ S = \frac{XS + SX}{2}$

- The binary operation $\circ$ is commutative $X \circ S = S \circ X$

- $\circ$ is not associative: $X \circ (Y \circ Z) \neq (X \circ Y) \circ Z$ in general

- But $X \circ (X \circ X) = (X \circ X) \circ X$. Thus $X^{\circ p} = X^p$ is well defined

- In general $X \circ (X^2 \circ Y) = X^2 \circ (X \circ Y)$

- The identity matrix I is identity w.r.t $\circ$

- Define the operator

$$\mathsf{L}_X : Y \to X \circ Y, \text{ thus } X \circ S = \mathsf{L}_X(S) = \mathsf{L}_X(\mathsf{L}_S(I))$$

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

Second Order Cone Programming (SOCP)

- For simplicity we deal with single variable SOCP:

Primal (P)	**Dual (D)**

$$\begin{aligned} \min \quad & c^\top x \\ \text{s.t.} \quad & Ax = b \\ & x_{\mathcal{Q}} \succeq 0 \end{aligned} \qquad\qquad \begin{aligned} \max \quad & b^\top y \\ \text{s.t.} \quad & A^\top y + s = c \\ & s_{\mathcal{Q}} \succeq 0 \end{aligned}$$
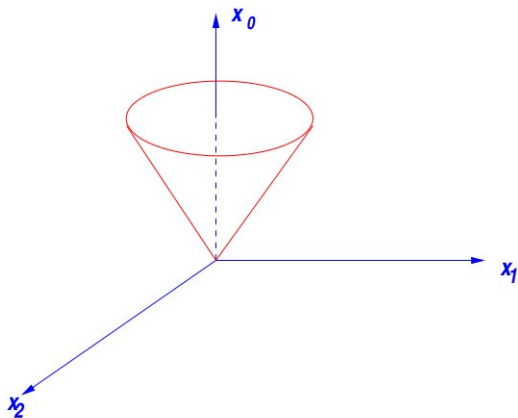
- the vectors x, s, c are indexed from zero

- If $z = (z_0, z_1, \ldots, z_n)^\top$ and $\bar{z} = (z_1, \ldots, z_n)^\top$

$$z_{\mathcal{Q}} \geq 0 \iff z_0 \geq \|\bar{z}\|$$

Illustration of SOC

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

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

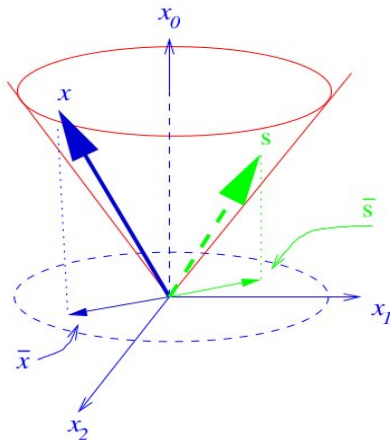$$(*) \quad x_0^2 \geq \sum_{i=1}^{n} x_i^2$$

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

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

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

- This implies

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

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 $x \succ_{\mathcal{Q}} 0$

- If there is dual-feasible $s \succ_{\mathcal{Q}} 0$

Complementary Slackness for SOCP

- Thus again we have a square system

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

Algebraic properties of SOCP

- Let us define a binary operation for vectors x and s both indexed from zero

$$\begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} \circ \begin{pmatrix} s_0 \\ s_1 \\ \vdots \\ s_n \end{pmatrix} = \begin{pmatrix} x^\top s \\ x_0 s_1 + s_0 x_1 \\ \vdots \\ x_0 s_n + s_0 x_n \end{pmatrix}$$

Algebraic properties of SOCP

- The binary operation $\circ$ is commutative $x \circ s = s \circ x$

- $\circ$ is not associative: $x \circ (y \circ z) \neq (x \circ y) \circ z$ in general

- But $x \circ (x \circ x) = (x \circ x) \circ x$. Thus $x^{\circ p} = x^p$ is well defined

- In general $x \circ (x^2 \circ y) = x^2 \circ (x \circ y)$

- The identity matrix I is identity w.r.t $\circ$

- $e = (1, 0, \ldots, 0)^\top$ is the identity: $x \circ e = x$

- Define the operator

$$L_x : y \to x \circ y$$

$$L_x = \mathrm{Arw}(x) = \begin{pmatrix} x_0 & \bar{x}^\top \\ \bar{x} & x_0 I \end{pmatrix}$$

$$x \circ s = \mathrm{Arw}(x)s = \mathrm{Arw}(x)\mathrm{Arw}(s)e$$

Summary

Properties

	LP	SDP	SOCP
binary operator	$x \circ s = (x_i s_i)$	$X \circ S = \frac{XS + SX}{2}$	$x \circ s = \begin{pmatrix} x^{\top} s \\ x_0 \bar{s} + s_0 \bar{x} \end{pmatrix}$
identity	1	I	$e = (1, 0, \ldots, 0)^{\top}$
associative	yes	no	no
L_X	$y \to \mathrm{Diag}(x) y$	$Y \to \frac{XY + YX}{2}$	$y \to \mathrm{Arw}(x) y$
Primal feasibility	$Ax = b$	$\langle A_i, X \rangle = b_i$	$Ax = b$
dual feasility	$A^{\top} y + s = c$	$\sum_i y_i A_i + S = C$	$A^{\top} y + s = c$
complementarity	$L_x L_s 1 = 0$	$L_X(L_S(I)) = 0$	$L_x L_s e = 0$

Conic LP

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

- It is a cone: If $x \in K \implies ax \in K$ 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)	**Dual (D)**

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

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

- 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 $\mathsf{K} = \mathsf{K}^*$