Chapter 3: Canonical Problem Forms

Chao Wang

SUSTech

Outline

1. Linear Program LP
2. Convex Quadratic Program QP
3. Semidefinite Program SDP
4. Cone Program
(最大类)

Linear Program

Linear program

A linear program or LP is an optimization problem of the form

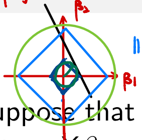
$$\begin{array}{ll} \max_x & c^T x \\ \text{subject to} & Dx \leq d \\ & Ax = b \end{array}$$

Observe that this is always a convex optimization problem

- First introduced by Kantorovich in the late 1930s and Dantzig in the 1940 s
- Dantzig's simplex algorithm gives a direct (noniterative) solver for LPs
- Fundamental problem in convex optimization. Many diverse applications, rich history

Example: basis pursuit

$\min \|\beta\|_1$ s.t. $X\beta=y$ $\beta \in \mathbb{R}^2$. $|\beta_1|+|\beta_2|$ $x_1\beta_1+x_2\beta_2=y$.



$\|\ \|_1$ 等高线 → 缩小

$\|\ \|_2$ 等高线 → 缩小 相切.

Given $y \in \mathbb{R}^n$ and $X \in \mathbb{R}^{n\times p}$, where $p > n$. Suppose that we seek the sparsest solution to underdetermined linear system $X\beta = y$

Nonconvex formulation:

近可能多的使 $\beta=0$.

$$\begin{array}{ll} \min_\beta & \|\beta\|_0 \\ \text{subject to} & X\beta = y \end{array}$$

where recall $\|\beta\|_0 = \sum_{j=1}^p 1\{\beta_j \neq 0\}$, the ℓ_0 "norm"

The ℓ_1 approximation, often called basis pursuit:

ℓ_1 逼近.

$$\min_\beta \|\beta\|_1 \qquad \sum_{i=1}^n |\beta|$$

转化为 Linear Program

subject to $X\beta = y$

Example: basis pursuit

$|\beta_1| + |\beta_2| + \cdots + |\beta_n|$

$\downarrow$

$z_1 = \max(\beta_1, -\beta_1) \Rightarrow$ 形成线性问题.

Basis pursuit is a linear program. Reformulation:

同步效果.

max 则失效.

$$\min_\beta \|\beta\|_1 \quad \text{subject to } X\beta = y \qquad \Longleftrightarrow \qquad \begin{aligned} \min_{\beta,z} \quad & 1^T z \\ \text{subject to} \quad & z \geq \beta \\ & z \geq -\beta \\ & X\beta = y \end{aligned}$$

Standard form

$$
\begin{array}{rll}
\text{Maximize} & c_1x_1 + c_2x_2 + \ldots + c_nx_n & \\
\text{subject to} & a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = b_1 \\
 & a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n & = b_2 \\
 & \vdots & \vdots \\
 & a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = b_m \\
\text{and} & x_1 \geq 0, x_2 \geq 0, \ldots, x_n \geq 0 &
\end{array} \tag{1}
$$

where $b_i \geq 0$.

Any linear program can be rewritten in standard form (check this!) We will give more explanations on Chapter 4.

Convex Quadratic Program

A convex quadratic program or QP is an optimization problem of the form

$$\begin{array}{ll} \min_x & c^T x + \frac{1}{2} x^T Q x \\ \text{subject to} & Dx \le d \\ & Ax = b \end{array}$$

$Q \equiv 0 \Rightarrow LP$

Hession $= Q \succeq 0$: second order ... $\Rightarrow$ convex

where $Q \succeq 0$, i.e., positive semidefinite Note that this problem is not convex when $Q \not\succeq 0$.

From now on, when we say quadratic program or QP, we implicitly assume that $Q \succeq 0$ (so the problem is convex)

Example: lasso

Given $y \in \mathbb{R}^n, X \in \mathbb{R}^{n\times p}$, recall the lasso problem:

$(y - X\beta)^T (y - X\beta)$

$= (y^T y) - \underline{2y^T X\beta} + \beta^T X^T X\beta$

$c^T x \quad x^T Q x$

$$\begin{array}{ll} \min_\beta & \|y - X\beta\|_2^2 \\ \text{subject to} & \|\beta\|_1 \le s \end{array}$$

Here $s \ge 0$ is a tuning parameter. Indeed, this can be reformulated as a quadratic program (check this!)

Alternative parametrization (called Lagrange, or penalized form):

$$\min_\beta \frac{1}{2}\|y - X\beta\|_2^2 + \underline{\lambda\|\beta\|_1},$$

$y^T y$

Now $\lambda \ge 0$ is a tuning parameter. And again, this can be rewritten as a quadratic program (check this!)

$(y - X\beta)^T (y - X\beta)$

$= y^T y \; \underline{-2y^T X\beta} + \beta^T X^T X\beta + \underline{\lambda z}$

$z > \beta$

$z > -\beta.$

Standard form

确定后

可解性

A quadratic program is in standard form if it is written as

$$\begin{array}{ll} \min_x & c^T x + \frac{1}{2} x^T Q x \qquad Q \in S_+^n \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array}$$

Any quadratic program can be rewritten in standard form

Semidefinite Program

Motivation for semidefinite programs

Consider linear programming again:

$$\begin{aligned} \min_x \quad & c^T x \\ \text{subject to} \quad & Dx \leq d \\ & Ax = b \end{aligned}$$

Can generalize by changing $\leq$ to different (partial) order.

- $\mathbb{S}^n$ is space of $n \times n$ symmetric matrices
- $\mathbb{S}^n_+$ is the space of positive semidefinite matrices, i.e.,
  $\mathbb{S}^n_+ = \{X \in \mathbb{S}^n : u^T X u \geq 0 \text{ for all } u \in \mathbb{R}^n\}$
- $\mathbb{S}^n_{++}$ is the space of positive definite matrices, i.e.,
  $\mathbb{S}^n_{++} = \{X \in \mathbb{S}^n : u^T X u > 0 \text{ for all } u \in \mathbb{R}^n \backslash \{0\}\}$

Motivation for semidefinite programs

- Basic linear algebra facts, here $\lambda(X) = (\lambda_1(X), \ldots, \lambda_n(X))$:

将λ(x) 转为矩阵 正定 or 半正定 关系.

$$X \in \mathbb{S}^n \Longrightarrow \lambda(X) \in \mathbb{R}^n$$

对称矩阵：特征值为实.

$$X \in \mathbb{S}^n_+ \overset{\text{半正定}}{\Longleftrightarrow} \lambda(X) \in \mathbb{R}^n_+ \ \geq 0$$

$$X \in \mathbb{S}^n_{++} \overset{\text{正定}}{\Longleftrightarrow} \lambda(X) \in \mathbb{R}^n_{++} \ > 0$$

- We can define a partial ordering over $\mathbb{S}^n$: given $X, Y \in \mathbb{S}^n$,

$$X \succeq Y \Longleftrightarrow X - Y \in \mathbb{S}^n_+$$

列向量.

Note: for $x, y \in \mathbb{R}^n$, $\text{diag}(x) \succeq \text{diag}(y) \Longleftrightarrow x \geq y$ (recall, the latter is interpreted elementwise)

线性 → 半正定.

$F_0 - (x_1 F_1 + \dots + x_n F_n) \in S_+^n$

A semidefinite program or SDP is an optimization problem of the form

$c^T x$

$$
\begin{array}{ll}
\min_x & c^T x \\
\text{subject to} & x_1 F_1 + \ldots + x_n F_n \preceq F_0 \\
& Ax = b
\end{array}
$$

$Ax = b$

$Dx \le d$

D

$m \times 1$

$F_1 =$

Here $F_j \in \mathbb{S}^d$, for $j = 0, 1, \ldots n$, and $A \in \mathbb{R}^{m \times n}$, $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$. Observe that this is always a convex optimization problem Also, any linear program is a semidefinite program (check this!)

$F_0 - \sum_{i=1}^n (\theta x_i + (1-\theta) y_i) F_i$

$x, y \in \Omega_1$

$F_0 - \sum_{i=1}^n x_i F_i \in S_+^n$

$= \theta (F_0 - \sum_{i=1}^n x_i F_i) + (1-\theta)(F_0 - \sum_{i=1}^n y_i F_i) \succeq 0.$

$\theta \in [0,1]$

$F_0 - \sum_{i=1}^n y_i F_i \in S_+^n$

$\theta x + (1-\theta) y \overset{?}{\in} \Omega_1$

Cone Program

A conic program is an optimization problem of the form:

$$\begin{aligned} \min_x \quad & c^T x \\ \text{subject to} \quad & Ax = b \\ & \underbrace{D(x)}_{\text{linear Map.}} + d \in K \end{aligned}$$

Here:

- $c, x \in \mathbb{R}^n$, and $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$
- $D : \mathbb{R}^n \to Y$ is a linear map, $d \in Y$, for Euclidean space Y
- $K \subseteq Y$ is a closed convex cone Both LPs and SDPs are special cases of conic programming. For LPs, $K = \mathbb{R}^n_+$; for SDPs, $K = \mathbb{S}^n_+$

$\forall\ x_1, x_2 \in C \quad \theta_1, \theta_2 \geq 0$

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

Example: second-order cone programs

A second-order cone program or SOCP is an optimization problem of the form:

$$
\begin{array}{ll}
\min_x & c^T x \\
\text{subject to} & \|D_i x + d_i\|_2 \leq e_i^T x + f_i, i = 1, \ldots p \\
 & Ax = b
\end{array}
$$

This is indeed a cone program. Why?
The second-order cone

$$Q = \{(x, t) : \|x\|_2 \leq t\}$$

So we have

$$\|D_i x + d_i\|_2 \leq e_i^T x + f_i \Longleftrightarrow \left(D_i x + d_i, e_i^T x + f_i\right) \in Q_i$$

for second-order cone Q_i of appropriate dimensions. Now take $K = Q_1 \times \ldots \times Q_p$

Observe that every LP is an SOCP. Further, every SOCP is an SDP Why?
Turns out that

列向量.　　也行

$$\|\underline{x}\|_2 \le t \iff \begin{bmatrix} tI & x \\ x^T & t \end{bmatrix} \succeq 0$$

$$\begin{bmatrix} tI & x \\ x^T & t \end{bmatrix}$$

Hence we can write any SOCP constraint as an SDP constraint.
The above is a special case of the Schur complement theorem:

$$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \succeq 0 \iff A - BC^{-1}B^T \succeq 0$$

$$tI - x\frac{1}{t}x^T = \underbrace{tI - \frac{1}{t}xx^T}_{E} \succeq 0$$

$$\frac{1}{t}x^T(xx^T)x$$

for A, C symmetric and $C \succ 0$

$\forall y \quad y^T E y \ge 0$

也行证明.

取 $y = x \Rightarrow t x^T x - \frac{1}{t}(x^T x)^2$

$\Rightarrow t\|x\|_2^2 \ge \frac{1}{t}\|x\|_2^4$

$\Rightarrow \|x\|_2 \le t.$

Connections

$\Sigma^{\frac{1}{2}} = \begin{pmatrix} \sqrt{\sigma_1} & & 0 \\ & \sqrt{\sigma_2} & \\ 0 & & \ddots & \sqrt{\sigma_n} \end{pmatrix}$

Finally, our old friend QPs "sneak" into the hierarchy.
Turns out QPs are SOCPs, which we can see by rewriting a QP as

找到最 min 时 x 和 t.

$A \in S^n$ $A = U \Sigma U^T$ 半正定矩阵可分解为2个矩阵.

$= \underline{\underline{U\Sigma^{\frac{1}{2}}}}\,\underline{\underline{\Sigma^{\frac{1}{2}}U^T}} = LL^T = A^{\frac{1}{2}}(A^{\frac{1}{2}})^T$

$$\min_{x,t} c^T x + t$$

subject to $Dx \leq d, \underline{\frac{1}{2}x^T Qx \leq t}$ $\min t = \frac{1}{2}x^T Q x$

$(\frac{1}{\sqrt{2}}Q^{1/2}x, \frac{1}{2}(1-t))^T(\frac{1}{\sqrt{2}}Q^{1/2}x, \frac{1}{2}(1-t))$
$= \frac{1}{2}x^T Q x + \frac{1}{2}(1-t)^T?$

$$Ax = b$$

Now write $\frac{1}{2}x^T Qx \leq t \iff \left\| \left(\frac{1}{\sqrt{2}} Q^{1/2} x, \frac{1}{2}(1-t) \right) \right\|_2 \leq \frac{1}{2}(1+t)$
Take a breath (phew!). Thus we have established the hierachy
$\mathrm{LPs} \subseteq \mathrm{QPs} \subseteq \mathrm{SOCPs} \subseteq \mathrm{SDPs} \subseteq$ Conic programs