第 3 章 凸优化模型

SMaLL

[1] 中国石油大学（华东）
SMaLL 课题组
small.sem.upc.edu.cn
liangxijunsd@163.com

2023

本章要点

- 掌握凸优化问题的形式，能够识别凸优化问题
- 了解拟凸优化的形式
- 掌握线性规划和二次规划的形式, 能够识别这两类优化问题, 并能够调用软件包解决问题
- 掌握二阶锥规划和半定规划 (semidefinite prog.) 的的形式, 能够识别这两类优化问题，并能够调用软件包解决问题

标准形式的优化问题

$$
\begin{aligned}
\text{minimize} \quad & f_0(x) \\
\text{subject to} \quad & f_i(x) \le 0, \quad i = 1, \ldots, m \\
& h_i(x) = 0, \quad i = 1, \ldots, p
\end{aligned}
\tag{1}
$$

- $x \in \mathbb{R}^n$ 是优化变量
- $f_0 : \mathbb{R}^n \to \mathbb{R}$ 是目标函数或代价函数
- $f_i : \mathbb{R}^n \to \mathbb{R}, i = 1, \ldots, m,$ 是不等式约束函数
- $h_i : \mathbb{R}^n \to \mathbb{R}$ 是等式约束函数

隐式约束

$$x \in \mathcal{D} = \bigcap_{i=0}^{m} \textbf{dom}\ f_i \ \cap \ \bigcap_{i=1}^{p} \textbf{dom}\ h_i, \tag{2}$$

- 我们称 $\mathcal{D}$ 为问题的**定义域**
- 约束 $f_i(x) \leq 0, h_i(x) = 0$ 是显式约束
- 如果一个问题没有明确的约束 $(m = p = 0)$，它就是**不受约束的**

例.

$$\text{minimize} \quad f_0(x) = -\sum_{i=1}^{k} \log\left(b_i - a_i^T x\right) \tag{3}$$

是一个具有隐式约束 $a_i^T x < b_i$ 的无约束问题

最优点和局部最优点

定义. 如果 $x \in \textbf{dom}\, f_0$ 并且满足约束条件，则称 x 是 **可行的**

最优值 p*:

$$p^\star = \inf\{f_0(x) \mid f_i(x) \le 0, i = 1, \dots, m, h_i(x) = 0, i = 1, \dots, p\} \qquad (4)$$

如果 $f_0(x) = p^\star$ 称这个可行的 x 为 **最优解**; X_{opt} 是所有最优解的集合

x 是 **局部最优的**如果存在 $R > 0$ 使得 x 是下面关于 z 优化问题的解

$$\begin{array}{ll} \text{minimize (over } z) & f_0(z) \\ \text{subject to} & f_i(z) \le 0, \quad i = 1, \dots, m, \quad h_i(z) = 0, \quad i = 1, \dots, p \\ & \|z - x\|_2 \le R \end{array} \qquad (5)$$

例. ($n=1, m=p=0$)

- $f_0(x)=1/x$, $\textbf{dom } f_0=\mathbb{R}_{++}: p^\star=0$, 无最优点
- $f_0(x)=-\log x$, $\textbf{dom } f_0=\mathbb{R}_{++}: p^\star=-\infty$
- $f_0(x)=x\log x$, $\textbf{dom } f_0=\mathbb{R}_{++}: p^\star=-1/e, x=1/e$ 是最优点
- $f_0(x)=x^3-3x$, $p^\star=-\infty$, 在 $x=1$ 处局部最优

可行性问题

$$
\begin{array}{lll}
\text{find} & x & \\
\text{subject to} & f_i(x) \leq 0, & i = 1, \ldots, m \\
& h_i(x) = 0, & i = 1, \ldots, p
\end{array}
\tag{6}
$$

当 $f_0(x) = 0$ 时，可以认为是一般问题的特例:

$$
\begin{array}{lll}
\text{minimize} & 0 & \\
\text{subject to} & f_i(x) \leq 0, & i = 1, \ldots, m \\
& h_i(x) = 0, & i = 1, \ldots, p
\end{array}
\tag{7}
$$

- 如果约束条件可行，那么 $p^\star = 0$ ；任何可行的 x 都是最佳的
- 如果约束不可行，那么 $p^\star = \infty$

凸优化问题

标准形式的凸优化问题

$$\begin{array}{lll} \text{minimize} & f_0(x) & \\ \text{subject to} & f_i(x) \leq 0, & i = 1, \ldots, m \\ & a_i^T x = b_i, & i = 1, \ldots, p \end{array} \tag{8}$$

- $f_0, f_1, \ldots, f_m$ 均为凸函数; 等式约束是仿射函数

重要性质：凸优化问题的可行集是凸的

示例

凸优化问题

$$
\begin{array}{ll}
\text{minimize} & f_0(x) = x_1^2 + x_2^2 \\
\text{subject to} & f_1(x) = x_1/(1+x_2^2) \leq 0 \\
& h_1(x) = (x_1+x_2)^2 = 0
\end{array}
\tag{9}
$$

- f_0 是凸函数; 可行集 $\{(x_1, x_2) \mid x_1 = -x_2 \leq 0\}$ 是凸集
- (根据我们的定义) 并非一个凸优化问题: f_1 不是凸函数, h_1 不是仿射函数
- 与凸问题等价（但不相同）

$$
\begin{array}{ll}
\text{minimize} & x_1^2 + x_2^2 \\
\text{subject to} & x_1 \leq 0 \\
& x_1 + x_2 = 0
\end{array}
\tag{10}
$$

局部和全局最优

定理 1

凸问题的任何局部最优点都是（全局）最优的

证明. 假设 x 是局部最优的，但存在一个可行的 y ，$f_0(y) < f_0(x)$

x 局部最优意味着存在 $R > 0$ 使得

$$z \text{ 可行}, \quad \|z - x\|_2 \leq R \quad \Longrightarrow \quad f_0(z) \geq f_0(x) \tag{11}$$

令 $z = \theta y + (1-\theta)x$ ，其中 $\theta = R/\left(2\|y-x\|_2\right)$

- $\|y - x\|_2 > R$, 所以 $0 < \theta < 1/2$
- z 是两个可行点的凸组合，因此也是可行的
- $\|z - x\|_2 = R/2$ 且

$$f_0(z) \leq \theta f_0(y) + (1-\theta) f_0(x) < f_0(x) \tag{12}$$

这与我们的假设相悖，即 x 是局部最优的

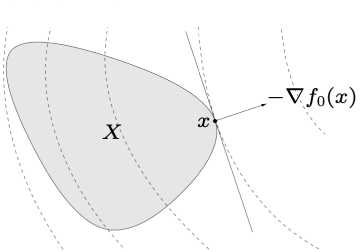
□

可微函数 f_0 的最优性准则

定理 2

对于任何凸优化问题，x 是最优的，当且仅当其可行且

$$\nabla f_0(x)^T(y-x) \geq 0 \quad 对所有可行\ y \tag{13}$$



其中可行集 x 由阴影表示，如果非零，$\nabla f_0(x)$ 在 x 处定义了可行集 X 的一个支撑超平面

证明.

证明. $\Longleftarrow$: f_0 为凸函数: $f_0(y) \geq f_0(x) + \langle \nabla f_0(x), y - x \rangle$, $\forall x, y \in \operatorname{dom} f_0$

$\Longrightarrow$: 假设 x 是最优的, $\exists y \in X$, 满足 $\langle \nabla f_0(x), y - x \rangle < 0$.

考虑点 $z(t) = ty + (1-t)x$, 其中 $t \in [0,1]$. X 是凸集 $\Rightarrow$ $z(t)$ 是可行的.

对于较小的正数 t: $f_0(z(t)) < f_0(x)$, $\Rightarrow$ x 不是最优的. 为了说明这点:

$$\left.\frac{d}{dt} f_0(z(t))\right|_{t=0} = \nabla f_0(x)^T (y - x) < 0$$

对于较小的正数 t: $f_0(z(t)) < f_0(x)$.

可微函数 f_0 的最优性准则

- **无约束问题:** x 是最优解，当且仅当

$$x \in \mathbf{dom}\ f_0, \quad \nabla f_0(x) = 0 \tag{14}$$

- **等式约束问题**

$$\text{minimize} \quad f_0(x) \quad \text{subject to} \quad Ax = b \tag{15}$$

  x 是最优解，当且仅当存在一个 ν 使得

$$x \in \mathbf{dom}\ f_0, \quad Ax = b, \quad \nabla f_0(x) + A^T\nu = 0 \tag{16}$$

- **非负整数的最小化**

$$\text{minimize} \quad f_0(x) \quad \text{subject to} \quad x \succeq 0 \tag{17}$$

  x 是最优解，当且仅当

$$x \in \mathbf{dom}\ f_0, \quad x \succeq 0, \quad \begin{cases} \nabla f_0(x)_i \geq 0 & x_i = 0 \\ \nabla f_0(x)_i = 0 & x_i > 0 \end{cases} \tag{18}$$

等价凸问题

如果一个问题的解很容易从另一个问题中得到，则两个问题 (非正式地) 是 **等价的**, 反之亦然

保持凸性的一些常见变换:

- **消除等式约束**

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & Ax = b \end{array} \tag{19}$$

等价于

$$\begin{array}{ll} \text{minimize (over} z) & f_0(Fz + x_0) \\ \text{subject to} & f_i(Fz + x_0) \leq 0, \quad i = 1, \ldots, m \end{array} \tag{20}$$

其中 F 和 x_0 满足

$$Ax = b \quad \Longleftrightarrow \quad x = Fz + x_0 (\text{对一切} z) \tag{21}$$

等价凸问题

- **引入等式约束**

$$\begin{array}{ll} \text{minimize} & f_0\left(A_0 x+b_0\right) \\ \text{subject to} & f_i\left(A_i x+b_i\right) \leq 0, \quad i=1, \ldots, m \end{array} \tag{22}$$

等价于

$$\begin{array}{lll} \text{minimize (over } x, y_i) & f_0\left(y_0\right) & \\ \text{subject to} & f_i\left(y_i\right) \leq 0, & i=1, \ldots, m \\ & y_i=A_i x+b_i, & i=0,1, \ldots, m \end{array} \tag{23}$$

- **引入线性不等式的松弛变量**

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & a_i^T x \leq b_i, \quad i=1, \ldots, m \end{array} \tag{24}$$

等价于

$$\begin{array}{lll} \text{minimize (over } x, s) & f_0(x) & \\ \text{subject to} & a_i^T x+s_i=b_i, & i=1, \ldots, m \\ & s_i \geq 0, & i=1, \ldots m \end{array} \tag{25}$$

等价凸问题

- **上图（epigraph）形式:** 标准形式凸问题等价于

$$
\begin{array}{ll}
\text{minimize (over } x, t) & t \\
\text{subject to} & f_0(x) - t \leq 0 \\
& f_i(x) \leq 0, \quad i = 1, \ldots, m \\
& Ax = b
\end{array}
\tag{26}
$$

- **最小化某些变量**

$$
\begin{array}{ll}
\text{minimize} & f_0(x_1, x_2) \\
\text{subject to} & f_i(x_1) \leq 0, \quad i = 1, \ldots, m
\end{array}
\tag{27}
$$

等价于

$$
\begin{array}{ll}
\text{minimize} & \tilde{f}_0(x_1) \\
\text{subject to} & f_i(x_1) \leq 0, \quad i = 1, \ldots, m
\end{array}
\tag{28}
$$

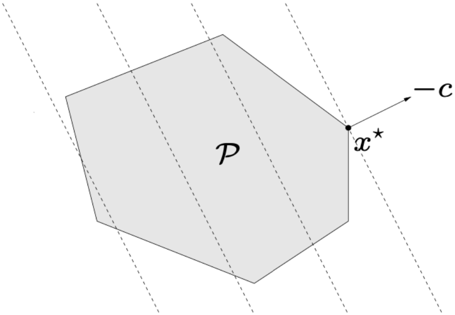
其中 $\tilde{f}_0(x_1) = \inf_{x_2} f_0(x_1, x_2)$

1. 凸优化问题

线性规划 (LP)

$$
\begin{array}{ll}
\text{minimize} & c^T x + d \\
\text{subject to} & Gx \preceq h \\
& Ax = b
\end{array}
\tag{29}
$$

- 目标函数和约束条件都是仿射函数的凸问题
- 可行集是多面体



示例

1. 饮食问题: 选择 n 种食物的数量 $x_1, \ldots, x_n$

- 一个单位的食物 j 花费 c_j, 包含的营养 i 的数量为 a_{ij}
- 健康饮食需要的营养 i 在数量上至少为 b_i

为了找到最便宜的健康饮食方案,

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \succeq b, \quad x \succeq 0 \end{array} \tag{30}$$

2. 分段线性最小化问题

$$\text{minimize} \quad \max_{i=1,\ldots,m} \left(a_i^T x + b_i\right) \tag{31}$$

等价于一个 LP 问题

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & a_i^T x + b_i \leq t, \quad i = 1, \ldots, m \end{array} \tag{32}$$

SMaLL
Statistical Machine Learning Lyceum

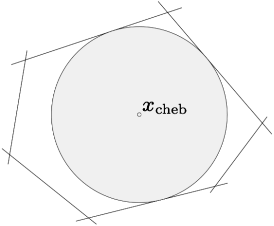
示例

3. 多面体的切比雪夫中心

$$\mathcal{P} = \left\{x \mid a_i^T x \leq b_i, i = 1, \ldots, m\right\} \tag{33}$$

的切比雪夫中心是最大内切球的中心

$$\mathcal{B} = \{x_c + u \mid \|u\|_2 \leq r\} \tag{34}$$



- $a_i^T x \leq b_i \ \forall x \in \mathcal{B}$ 当且仅当

$$\sup\left\{a_i^T (x_c + u) \mid \|u\|_2 \leq r\right\} = a_i^T x_c + r\left\|a_i\right\|_2 \leq b_i \tag{35}$$

- 因此, x_c, r 可以通过求解 LP 问题来确定

$$\begin{array}{ll} \text{maximize} & \text{r} \\ \text{subject to} & a_i^T x_c + r\left\|a_i\right\|_2 \leq b_i, \quad i = 1, \ldots, m \end{array} \tag{36}$$

例. 基于线性规划求解该多面体的切比雪夫中心

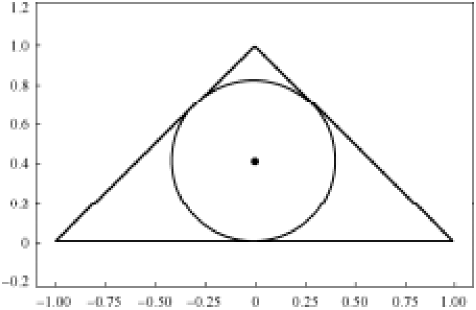
考虑下面的线性不等式组定义的多面体 $\mathcal{P}$:

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ -1 & 0 \end{bmatrix} \cdot \begin{bmatrix} y \\ x \end{bmatrix} \preccurlyeq \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

运行代码, 基于线性规划求解该多面体的切比雪夫中心

基于线性规划求解该多面体的切比雪夫中心 (运行代码)

```
from scipy import optimize
import numpy as np
from matplotlib import pyplot as plt
c = np.array([1,0,0])
A = np.array([[np.sqrt(2),1,1],[np.sqrt(2),1,1],[1,-1,0]])
b = np.array([1,1,0])
res = optimize.linprog(-c,A,b)
res.x
plt.plot([-1,0],[0,1],’b’)
plt.plot([0,1],[1,0],’b’)
plt.plot([-1,1],[0,0],’b’)
r,x,y = res.x[0],res.x[2],res.x[1]
theta = np.arange(0,2*np.pi,0.01)
plt.plot(x+r*np.cos(theta),y+r*np.sin(theta),’r’)
plt.plot(x,y,’k.’)
plt.axis(’equal’)
```



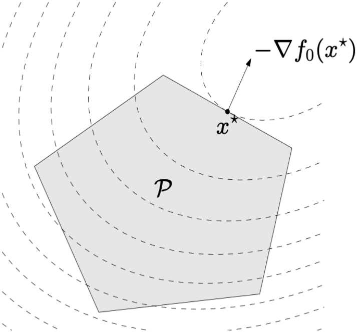
多面体 P 及求得的切比雪夫中心，如图所示。

SMaLL
Statistical Machine Learning Lyceum

二次规划 (QP)

$$\begin{array}{ll} \text{minimize} & (1/2)x^TPx + q^Tx + r \\ \text{subject to} & Gx \preceq h \\ & Ax = b \end{array} \tag{37}$$

- $P \in \mathbf{S}^n_+$, 所以目标是凸二次函数
- 最小化多面体上的凸二次函数



示例

1. 最小二乘回归

$$\text{minimize} \quad \frac{1}{2}\|Ax-b\|_2^2 \tag{38}$$

- 解析解 $x^\star = A^\dagger b$ ($A^\dagger$ 是 A 的伪逆)
- 可以添加线性约束, $e.g., l \preceq x \preceq u$

2. 具有随机成本的线性规划

$$\begin{array}{ll} \text{minimize} & \bar{c}^T x + \gamma x^T \Sigma x = \mathbf{E} c^T x + \gamma \mathbf{var}\,(c^T x) \\ \text{subject to} & Gx \preceq h, \quad Ax = b \end{array} \tag{39}$$

- c 是具有均值 $\bar{c}$ 和协方差 Σ 的随机向量
- 因此, $c^T x$ 是具有平均值 $\bar{c}^T x$ 和方差 $x^T \Sigma x$ 的随机变量
- $\gamma > 0$ 是风险规避参数；控制预期成本和差异（风险）之间的权衡

3. 多面体之间的距离

- 多面体 $\mathcal{P}_1=\{\boldsymbol{x} \mid \boldsymbol{A}_1\boldsymbol{x} \preceq \boldsymbol{b}_1\}$ 和 $\mathcal{P}_2=\{\boldsymbol{x} \mid \boldsymbol{A}_2\boldsymbol{x} \preceq \boldsymbol{b}_2\}$ 在 $\mathbf{R}^n$ 上的欧氏距离定义为

$$\operatorname{dist}(\mathcal{P}_1,\mathcal{P}_2)=\inf\{\|\boldsymbol{x}_1-\boldsymbol{x}_2\|_2 \mid \boldsymbol{x}_1\in\mathcal{P}_1, \boldsymbol{x}_2\in\mathcal{P}_2\}$$

- 如果多面体相交，则距离为零。
- 求解下面的二次规划 $\rightarrow$ $\mathcal{P}_1$ 和 $\mathcal{P}_2$ 之间的距离

$$\begin{array}{ll}\text{minimize}_{\,\boldsymbol{x}_1,\boldsymbol{x}_2\in\mathbf{R}^n} & \|\boldsymbol{x}_1-\boldsymbol{x}_2\|_2^2 \\ \text{subject to} & \boldsymbol{A}_1\boldsymbol{x}_1 \preceq \boldsymbol{b}_1, \boldsymbol{A}_2\boldsymbol{x}_2 \preceq \boldsymbol{b}_2\end{array}$$

- 当且仅当多面体相交时，最优值为零, 否则最优解 $\boldsymbol{x}_1$ 和 $\boldsymbol{x}_2$ 分别是 $\mathcal{P}_1$ 和 $\mathcal{P}_2$ 中最接近的点

例. 二次规划求多面体间的距离

考虑多面体 $\mathcal{P}_1$

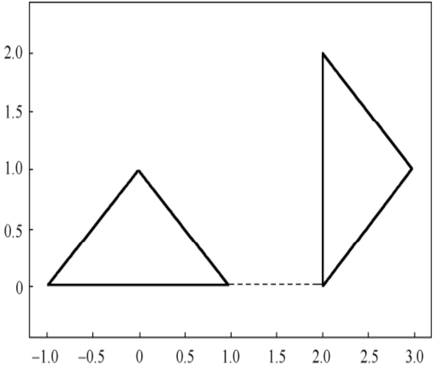
$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ -1 & 0 \end{bmatrix} \cdot \begin{bmatrix} y \\ x \end{bmatrix} \preccurlyeq \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

和多面体 $\mathcal{P}_2$

$$\begin{bmatrix} 0 & -1 \\ -1 & 1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} y \\ x \end{bmatrix} \preccurlyeq \begin{bmatrix} -2 \\ 2 \\ 4 \end{bmatrix}$$

之间的距离。

学习实例代码 3.4.1, 基于二次规划求解上述两个多面体之间距离。
代码输出求得的最优解为：[5.30e−04，1.00e+00，5.78e−04，2.00e+00]，表示两个多面体相距最近的两个点近似为 (1,0) 和 (2,0)。



如图展示了这两个多面体及其最相距最近的两个点。

二次约束二次规划 (QCQP)

$$\begin{array}{ll} \text{minimize} & (1/2)x^TP_0x+q_0^Tx+r_0 \\ \text{subject to} & (1/2)x^TP_ix+q_i^Tx+r_i\leq 0, \quad i=1,\ldots,m \\ & Ax=b \end{array} \tag{40}$$

- $P_i\in\mathbf{S}_+^n$; 目标函数和约束函数是凸二次函数
- 如果 $P_1,\ldots,P_m\in\mathbf{S}_{++}^n$, 可行域是 m 个椭球和仿射集 $Ax=b$ 的交集

例. QCQP

$$\max \quad \theta$$

$$\text{s.t.} \quad \|\mu\|_2 \leq 1,$$

$$\mu_i \geq 0, i = 1, \ldots, p,$$

$$\frac{1}{2}\sum_{i=1}^{p} \mu_i f_i(\alpha) - \sum_{i=1}^{l} \alpha_i \geq \theta,$$

$$0 \leq \alpha_i \leq s_i C, \ \ i = 1, \ldots, l,$$

$$\sum_{i=1}^{l} y_i \alpha_i = 0,$$

where $f_i(\alpha) = \sum_{j,k=1}^{l} \alpha_j \alpha_k y_j y_k k^i(\mathbf{x}_j, \mathbf{x}_k), i = 1, \ldots, p.$

Ling Jian, Zhonghang Xia, Xinnan Niu, Xijun Liang, Parimal Samir, Andrew J. Link:L2 基于多核模糊支持向量机的多肽识别数据融合. IEEE ACM Trans. Comput. Biol. Bioinform. 13(4): 804-809

二阶锥规划

$$
\begin{array}{ll}
\text{minimize} & f^T x \\
\text{subject to} & \|A_i x + b_i\|_2 \leq c_i^T x + d_i, \quad i = 1, \ldots, m \\
& Fx = g
\end{array} \tag{41}
$$

$\left(A_i \in \mathbb{R}^{n_i \times n}, F \in \mathbb{R}^{p \times n}\right)$

- 不等式称为二阶锥（SOC）约束:

$$
\left(A_i x + b_i, c_i^T x + d_i\right) \in \mathbb{R}^{n_i+1} \text{ 上的二阶锥} \tag{42}
$$

$\mathbb{R}^{n+1}$ 上的二阶锥: $\{(x, t) \in \mathbb{R}^{n+1} | \|x\|_2 \leq t\}$

- 当 $n_i = 0$, 简化成 LP 问题; 如果 $c_i = 0$, 变为一个 QCQP 问题
- 比 QCQP 和 LP 更普遍

例. 鲁棒线性规划

优化问题中的参数通常是不确定的, *e.g.*, 在 LP 问题中

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i, \quad i = 1, \ldots, m \end{array} \tag{43}$$

在 c, a_i, b_i 中可能存在不确定性

处理不确定性的两种常见方法 (为了简单起见，以 a_i 为例)

- 确定性模型: 约束必须适用于所有 $a_i \in \mathcal{E}_i$

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i \quad \forall a_i \in \mathcal{E}_i, \quad i = 1, \ldots, m \end{array} \tag{44}$$

- 随机模型: a_i 为随机变量；约束为概率约束

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \mathbf{prob}\,(a_i^T x \leq b_i) \geq \eta, \quad i = 1, \ldots, m \end{array} \tag{45}$$

例. 鲁棒线性规划

确定性方法 ← SOCP（二阶锥规划）(自学)

- 选择椭球 $\mathcal{E}_i$：

$$\mathcal{E}_i = \{\bar{a}_i + P_i u \mid \|u\|_2 \le 1\} \quad (\bar{a}_i \in \mathbb{R}^n, \quad P_i \in \mathbb{R}^{n\times n}) \tag{46}$$

球心是 $\bar{a}_i$, 由奇异值/向量 P_i 确定半轴

- 鲁棒线性规划问题

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \le b_i \quad \forall a_i \in \mathcal{E}_i, \quad i = 1, \ldots, m \end{array} \tag{47}$$

等价于 SOCP（二阶锥规划）

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \bar{a}_i^T x + \|P_i^T x\|_2 \le b_i, \quad i = 1, \ldots, m \end{array} \tag{48}$$

由于 $\sup_{\|u\|_2 \le 1} (\bar{a}_i + P_i u)^T x = \bar{a}_i^T x + \|P_i^T x\|_2 \Big)$

基于 SOCP（二阶锥规划）的随机方法

- 假定 a_i 为高斯分布，均值为 $\bar{a}_i$, 方差为 Σ_i $(a_i \sim \mathcal{N}(\bar{a}_i, \Sigma_i))$
- $a_i^T x$ 是一个高斯 $r.v.$ ，均值为 $\bar{a}_i^T x$, 方差为 $x^T \Sigma_i x$;

$$\textbf{prob}\left(a_i^T x \leq b_i\right) = \Phi\left(\frac{b_i - \bar{a}_i^T x}{\left\|\Sigma_i^{1/2} x\right\|_2}\right) \tag{49}$$

其中 $\Phi(x) = (1/\sqrt{2\pi}) \int_{-\infty}^{x} e^{-t^2/2} dt$ 是 $\mathcal{N}(0,1)$ 的分布函数

- 鲁棒线性规划

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \textbf{prob}\left(a_i^T x \leq b_i\right) \geq \eta, \quad i = 1, \ldots, m \end{array} \tag{50}$$

$\eta \geq 1/2$, 等价于 SOCP 问题

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \bar{a}_i^T x + \Phi^{-1}(\eta) \left\|\Sigma_i^{1/2} x\right\|_2 \leq b_i, \quad i = 1, \ldots, m \end{array} \tag{51}$$

例. 二阶锥规划问题

$$
\begin{aligned}
\text{minimize} \quad & -2x_1 + x_2 + 5x_3 \\
\text{subject to} \quad & \left\| \begin{bmatrix} -13x_1 + 3x_2 + 5x_3 - 3 \\ -12x_1 + 12x_2 - 6x_3 - 2 \end{bmatrix} \right\|_2 \leqslant -12x_1 - 6x_2 + 5x_3 - 12 \\
& \left\| \begin{bmatrix} -3x_1 + 6x_2 + 2x_3 \\ x_1 + 9x_2 + 2x_3 + 3 \\ -x_1 - 19x_2 + 3x_3 - 42 \end{bmatrix} \right\|_2 \leqslant -3x_1 + 6x_2 - 10x_3 + 27
\end{aligned}
$$

学习实例代码 3.4.2, 调用 Cvxopt 包求解上述问题, 所求得的最优解为 $[-5.01\mathrm{e}+00,\ -5.77\mathrm{e}+00, -8.52\mathrm{e}+00]$

几何规划

单项式函数

$$f(x) = cx_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}, \qquad \mathbf{dom}\ f = \mathbb{R}^n_{++} \tag{52}$$

$c > 0$; a_i 可以为任意实数

多项式函数: 单项式求和

$$f(x) = \sum_{k=1}^{K} c_k x_1^{a_{1k}} x_2^{a_{2k}} \cdots x_n^{a_{nk}}, \qquad \mathbf{dom}\ f = \mathbb{R}^n_{++} \tag{53}$$

几何规划 (GP)

$$\begin{array}{lll} \text{minimize} & f_0(x) & \\ \text{subject to} & f_i(x) \leq 1, & i = 1, \ldots, m \\ & h_i(x) = 1, & i = 1, \ldots, p \end{array} \tag{54}$$

f_i 为多项式, h_i 为单项式

凸形几何规划

将变量更改为 $y_i = \log x_i$, 并取目标和约束条件的对数

- 单项式 $f(x) = cx_1^{a_1} \cdots x_n^{a_n}$ 变为

$$\log f(e^{y_1}, \ldots, e^{y_n}) = a^T y + b \quad (b = \log c) \tag{55}$$

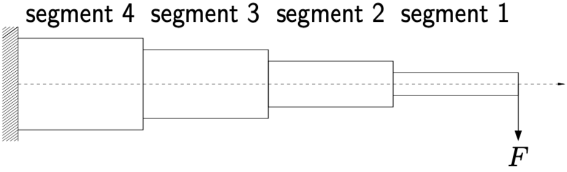
- 多项式 $f(x) = \sum_{k=1}^K c_k x_1^{a_{1k}} x_2^{a_{2k}} \cdots x_n^{a_{nk}}$ 变为

$$\log f(e^{y_1}, \ldots, e^{y_n}) = \log\left(\sum_{k=1}^K e^{a_k^T y + b_k}\right) \quad (b_k = \log c_k) \tag{56}$$

- 几何规划转变为凸问题

$$\begin{array}{ll} \text{minimize} & \log\left(\sum_{k=1}^K \exp\left(a_{0k}^T y + b_{0k}\right)\right) \\ \text{subject to} & \log\left(\sum_{k=1}^K \exp\left(a_{ik}^T y + b_{ik}\right)\right) \le 0, \quad i = 1, \ldots, m \\ & Gy + d = 0 \end{array} \tag{57}$$

悬臂梁设计



- 具有单位长度的线段 N ，大小 $w_i \times h_i$ 为矩形截面
- 给定右端施加的垂直力 F

设计问题

minimize	总重量
subject to	w_i, h_i 上下界
	纵横比 (aspect ratio)h_i/w_i的上下界
	每段应力的上下界
	每段末尾的垂直偏转 (vertical deflection) 力的上界

(58)

悬臂梁设计

目标函数和约束条件

- 总重量 $w_1h_1+\cdots+w_Nh_N$ 是多项式
- 纵横比 h_i/w_i 和反纵横比 w_i/h_i 是单项式
- 线段 i 中的最大应力由一个单项式给出 $6iF/\left(w_ih_i^2\right)$
- 线段 i 右端中心轴的垂直偏转 y_i 和斜率 v_i 递归定义为

$$\begin{aligned} v_i &= 12(i-1/2)\frac{F}{Ew_ih_i^3}+v_{i+1} \\ y_i &= 6(i-1/3)\frac{F}{Ew_ih_i^3}+v_{i+1}+y_{i+1} \end{aligned} \tag{59}$$

$i=N,N-1,\ldots,1,\ v_{N+1}=y_{N+1}=0$($E$ 是杨氏模量)

v_i 和 y_i 是 w,h 的多项式函数

悬臂梁设计

构造一个 GP 问题

$$
\begin{aligned}
\text{minimize} \quad & w_1 h_1 + \cdots + w_N h_N \\
\text{subject to} \quad & w_{\max}^{-1} w_i \leq 1, \quad w_{\min} w_i^{-1} \leq 1, \quad i = 1, \ldots, N \\
& h_{\max}^{-1} h_i \leq 1, \quad h_{\min} h_i^{-1} \leq 1, \quad i = 1, \ldots, N \\
& S_{\max}^{-1} w_i^{-1} h_i \leq 1, \quad S_{\min} w_i h_i^{-1} \leq 1, \quad i = 1, \ldots, N \\
& 6iF\sigma_{\max}^{-1} w_i^{-1} h_i^{-2} \leq 1, \quad i = 1, \ldots, N \\
& y_{\max}^{-1} y_1 \leq 1
\end{aligned} \tag{60}
$$

注意

- $w_{\min} \leq w_i \leq w_{\max}$, $h_{\min} \leq h_i \leq h_{\max} \rightarrow$

$$
w_{\min}/w_i \leq 1, \quad w_i/w_{\max} \leq 1, \quad h_{\min}/h_i \leq 1, \quad h_i/h_{\max} \leq 1 \tag{61}
$$

- $S_{\min} \leq h_i/w_i \leq S_{\max} \leftarrow$

$$
S_{\min} w_i/h_i \leq 1, \quad h_i/(w_i S_{\max}) \leq 1 \tag{62}
$$

最小化非负矩阵的谱半径

erron-Frobenius 特征值 $\lambda_{\text{pf}}(A)$

- 存在对每个元素都为正的矩阵 $A \in \mathbb{R}^{n\times n}$
- 实际的正特征值 A, 等于光谱半径 $\max_i |\lambda_i(A)|$
- 确定 A^k 的渐近增长（衰减）率:$A^k \sim \lambda_{\text{pf}}^k$ as $k \to \infty$
- 替代性特征: $\lambda_{\text{pf}}(A) = \inf\{\lambda \mid Av \preceq \lambda v \text{ for some } v \succ 0\}$

最小化多项式矩阵的谱半径

- 最小化 $\lambda_{\text{pf}}(A(x))$, 其中 $A(x)_{ij}$ 中元素是 x 的多项式
- 等效几何规划:

$$\begin{array}{ll} \text{minimize} & \lambda \\ \text{subject to} & \sum_{j=1}^n A(x)_{ij} v_j / (\lambda v_i) \leq 1, \quad i = 1, \dots, n \end{array} \tag{63}$$

变量为 λ, v, x

广义不等式约束

具有广义不等式约束的凸问题

$$
\begin{array}{ll}
\text{minimize} & f_0(x) \\
\text{subject to} & f_i(x) \preceq_{K_i} 0, \quad i=1,\ldots,m \\
& Ax = b
\end{array} \tag{64}
$$

- $f_0 : \mathbb{R}^n \to \mathbb{R}$ 为凸函数; $f_i : \mathbb{R}^n \to \mathbb{R}^{k_i} K_i$ -凸， K_i: proper cone (正常锥)
- 与标准凸问题相同的性质（凸可行集，局部最优解是全局最优的, etc.)

锥形问题: 具有仿射目标函数和约束条件的特例

$$
\begin{array}{ll}
\text{minimize} & c^T x \\
\text{subject to} & Fx + g \preceq_K 0 \\
& Ax = b
\end{array} \tag{65}
$$

将线性规划 ($K = \mathbb{R}^m_+$) 推广到非多面体锥

半定规划 (SDP)

$$
\begin{array}{ll}
\text{minimize} & c^T x \\
\text{subject to} & x_1 F_1 + x_2 F_2 + \cdots + x_n F_n + G \preceq 0 \\
 & Ax = b
\end{array} \tag{66}
$$

其中 $F_i, G \in \mathbf{S}^k$

- 不等式约束称为线性矩阵不等式 (LMI)
- 包括具有多个 LMI 约束的问题: 例如,

$$
x_1 \hat{F}_1 + \cdots + x_n \hat{F}_n + \hat{G} \preceq 0, \quad x_1 \tilde{F}_1 + \cdots + x_n \tilde{F}_n + \tilde{G} \preceq 0 \tag{67}
$$

等价于单线矩阵不等式 SLMI

$$
x_1 \begin{bmatrix} \hat{F}_1 & 0 \\ 0 & \tilde{F}_1 \end{bmatrix} + x_2 \begin{bmatrix} \hat{F}_2 & 0 \\ 0 & \tilde{F}_2 \end{bmatrix} + \cdots + x_n \begin{bmatrix} \hat{F}_n & 0 \\ 0 & \tilde{F}_n \end{bmatrix} + \begin{bmatrix} \hat{G} & 0 \\ 0 & \tilde{G} \end{bmatrix} \preceq 0 \tag{68}
$$

特征值最小化

$$\text{minimize} \quad \lambda_{\max}(A(x)) \tag{69}$$

其中 $A(x) = A_0 + x_1 A_1 + \cdots + x_n A_n$ (给定 $A_i \in \mathbf{S}^k$)

等价于 SDP 问题

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & A(x) \preceq tI \end{array} \tag{70}$$

- 变量 $x \in \mathbb{R}^n, t \in \mathbb{R}$
- 来自于

$$\lambda_{\max}(A) \leq t \quad \Longleftrightarrow \quad A \preceq tI \tag{71}$$

矩阵范数最小化

$$\text{minimize} \quad \|A(x)\|_2 = \left(\lambda_{\max}\left(A(x)^T A(x)\right)\right)^{1/2} \tag{72}$$

其中 $A(x) = A_0 + x_1 A_1 + \cdots + x_n A_n$ ($A_i \in \mathbb{R}^{p \times q}$ 给定) $\|A\|_2 = A$ 的最大奇异值 $\leftrightarrow$ SDP

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \begin{bmatrix} tI & A(x) \\ A(x)^T & tI \end{bmatrix} \succeq 0 \end{array} \tag{73}$$

- 变量 $x \in \mathbb{R}^n, t \in \mathbb{R}$
- 约束:

$$\begin{aligned} \|A\|_2 \le t \Longleftrightarrow A^T A \preceq t^2 I, \quad t \ge 0 \\ \Longleftrightarrow \begin{bmatrix} tI & A \\ A^T & tI \end{bmatrix} \succeq 0 \end{aligned} \tag{74}$$

例. 最大割问题的半定规划松弛

- 设 G 是一个无向图, 顶点集合为 $\mathbf{N} = 1, 2, \cdots, n$, 边的集合为 $\mathbf{E}$
- 设 $w_{ij} = w_{ji}$ 为边 (i, j) 上的权值, 其中 $(i, j) \in \mathbf{E}$ 。假设对于所有的 $(i, j) \in \mathbf{E}$, 都有 $w_{ij} > 0$
- 最大割问题是指确定顶点集合 $\mathbf{N}$ 的一个子集 $\mathbf{S}$, 使连接顶点子集 $\mathbf{S}$ 到它的余集 $\bar{\mathbf{S}}$ (其中 $\bar{\mathbf{S}} := N\backslash\mathbf{S}$) 的所有的边权值之和最大
- 若 $j \in \mathbf{S}$, 则令 $x_j = -1$, 否则，令 $x_j = 1$, 可以将最大割问题描述为如下的整数规划问题:

$$
\begin{array}{ll}
\text{maximize} & \frac{1}{4}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\,(1 - x_i x_j) \\
\text{subject to} & x_j \in \{-1, 1\}, j = 1, 2, \cdots, n
\end{array}
$$

例. 最大割问题的半定规划松弛

- $$\begin{array}{ll} \text{maximize} & \frac{1}{4}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\left(1 - x_i x_j\right) \\ \text{subject to} & x_j \in \{-1, 1\}, j = 1, 2, \cdots, n \end{array}$$
- 令 $\boldsymbol{Y} = \mathbf{x}\mathbf{x}^{\mathrm{T}}$, 其中 $\mathrm{Y}_{ij} = x_i x_j, i = 1, 2, \cdots, n, j = 1, 2, \cdots, n$ 。
- 令 $\boldsymbol{W}$ 为权矩阵, 即其第 i 行, 第 j 列的元素为 $\boldsymbol{c} \in \mathbf{R}^n, i = 1, 2, \cdots, n, j = 1, 2, \cdots, n$, 则最大割问题 $\Leftrightarrow$
  $$\begin{array}{ll} \text{maximize} & \frac{1}{4}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} - \boldsymbol{W}\cdot\boldsymbol{Y} \\ \text{subject to} & x_j \in \{-1, 1\}, j = 1, 2, \cdots, n \\ & \boldsymbol{Y} = \mathbf{x}\mathbf{x} \end{array}$$
- 上述问题中第一部分约束等价于 $\mathrm{Y}_{jj} = 1, j = 1, \cdots, n \rightarrow$
  $$\begin{array}{ll} \text{maximize} & \frac{1}{4}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} - \boldsymbol{W}\cdot\boldsymbol{Y} \\ \text{subject to} & \mathrm{Y}_{jj} = 1, j = 1, 2, \cdots, n \\ & \boldsymbol{Y} = \mathbf{x}\mathbf{x}^{\mathrm{T}} \end{array}$$

例. 最大割问题的半定规划松弛

- 矩阵 $\boldsymbol{Y} = \mathbf{x}\mathbf{x}^{\mathrm{T}}$ 是一个秩为 1 的半正定矩阵 → 松弛, 去掉秩为 1 这一限制 → 半定规划松弛问题:

$$\begin{array}{ll} \text{maximize} & \frac{1}{4}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} - \boldsymbol{W}\cdot\boldsymbol{Y} \\ \text{subjectto} & \mathrm{Y}_{jj} = 1, j = 1, 2, \cdots, n \\ & \boldsymbol{Y} \succeq 0 \end{array}$$

- 已知松他问题（RELAX）为最大割问题（MAXCUT) 提供了一个上界, 即:

$$MAXCUT \leqslant RELAX \text{ 。}$$

- 可以证明下面的不等式成立

$$0.8786 RELAX \leqslant \text{ MAXCUT } \leqslant RELAX \text{ 。}$$

- → 半定规划松他问题的最优值与最大割这一 NP 难问题的最优值相差不超过 13%

SMaLL Statistical Machine Learning Lyceum

例. 最大割问题的半定规划松弛

- 实例代码 3.6.1 给出了半定规划求解最大割问题的示例，其中有 5 个结点
- 学习代码，调用 cvxpy 包求解最大割问题的半定规划松弛问题

作业题

1. 设 $A \in \mathbb{R}^{n\times n}$, $c, d \in \mathbb{R}^n$, $d \in \mathbb{R}$, 证明：映射 $\mathcal{A} : x \mapsto \left(Ax + b, c^T x + d\right)$, $\mathbb{R}^n \to \mathbb{R}^{n+1}$ 为仿射变换.
2. 编程题 (使用 Python): 求解鲁棒线性规划的二次规划问题

$$\begin{array}{ll} \text{minimize} & \bar{c}^T x + \gamma x^T \Sigma x = \mathbf{E} c^T x + \gamma \mathbf{var}\left(c^T x\right) \\ \text{subject to} & Gx \preceq h, \quad Ax = b \end{array} \tag{75}$$

c 是均值为 $\bar{c}$ 方差为 Σ 的随机向量, 自行设置 $\bar{c}, \Sigma$ 的值.

作业题

(选作) 3. 编程题 (使用 Python): 求解鲁棒 LP的 SOCP 问题

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \mathbf{prob}\,(a_i^T x \leq b_i) \geq \eta, \quad i = 1, \ldots, m \end{array} \tag{76}$$

其中 $\eta \geq 1/2$, 等价于 SOCP 问题

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \bar{a}_i^T x + \Phi^{-1}(\eta) \left\| \Sigma_i^{1/2} x \right\|_2 \leq b_i, \quad i = 1, \ldots, m \end{array} \tag{77}$$

4. 利用 Python 绘制以下函数的图形

$$\log f(e^{y_1}, \ldots, e^{y_n}) = \log \left(\sum_{k=1}^{K} e^{a_k^T y + b_k} \right) \quad (b_k = \log c_k) \tag{78}$$

with $y \in \mathbb{R}^2$.