第九章 优化基础

第 31 讲 凸优化

黄定江

DaSE @ ECNU
djhuang@dase.ecnu.edu.cn

凸优化问题的标准形式

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

- $x \in R^n$ 是优化变量
- $f_0 : R^n \to R$ 是目标函数或者损失函数
- $f_i : R^n \to R, \quad i = 1, \cdots, m$ 是不等式约束函数
- $h_i : R^n \to R$ 是等式约束函数

凸优化问题的标准形式

最优值：

$$p^* = \inf\{f_0(x) | f_i(x) \leq 0, i = 1, \cdots, m, h_i(x) = 0, i = 1, \cdots, p\}$$

- $p^* = \pm\infty$，如果问题不可行 (没有 x 满足约束)
- $p^* = -\infty$，问题无下界

最优点与局部最优点

目标函数和约束函数所有有定义点的集合：

$$\mathcal{D} = \bigcap_{i=0}^{m} \mathbf{dom}\, f_i \cap \bigcap_{i=0}^{p} \mathbf{dom}\, h_i$$

如果 $x \in \mathcal{D}$ 且满足约束 $f_i(x) \leq 0, i = 1, \cdots, m, h_i(x) = 0, i = 1, \cdots, p$，则 x **是可行的**，所有可行点的集合称为**可行集**。

最优点与局部最优点

可行的 x 是最优的如果 $f_0(x) = p^*$； X_{opt} 是最优点的集合
可行解 x 为局部最优，如果存在 $R > 0$ 使得

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, i = 1, \cdots, m \quad h_i(x) = 0, i = 1, \cdots, p \\ & \|z - x\|^2 \leq R \end{array}$$

最优点与局部最优点
例子

- $f_0(x) = \frac{1}{x}, \mathbf{dom}\, f_0 = R^{++} : p^* = 0$，**无最优点**
- $f_0(x) = -\log x, \mathbf{dom}\, f_0 = R^{++} : p^* = -\infty$
- $f_0(x) = x\log x, \mathbf{dom}\, f_0 = R^{++} : p^* = -\frac{1}{e}, x = \frac{1}{e}$ **是最优点**
- $f_0(x) = x^3 - 3x, p^* = -\infty$， $x = 1$ **是局部最优**

凸优化问题

标准形式的凸优化问题

$$\begin{aligned} \text{minimize} \quad & f_0(x) \\ \text{subject to} \quad & f_i(x) \le 0, \quad i = 1, \cdots, m \\ & a_i^T x = b_i, \quad i = 1, \cdots, p \end{aligned}$$

- $f_0, \cdots, f_m$ 为凸函数；等式约束是仿射的。

经常写作：

$$\begin{aligned} \text{minimize} \quad & f_0(x) \\ \text{subject to} \quad & f_i(x) \le 0, \quad i = 1, \cdots, m \\ & Ax = b \end{aligned}$$

重要性质：凸优化问题的可行集是凸的。

例 1

$$\begin{aligned} \textit{minimize} \quad & f_0(x) = x_1^2 + x_2^2 \\ \textit{subject to} \quad & f_1(x) = x_1/(1+x_2^2) \le 0 \\ & h_1(x) = (x_1 + x_2)^2 = 0 \end{aligned}$$

- f_0 是凸的；可行集 $\{(x_1, x_2 | x_1 = -x_2 \le 0)\}$ 是凸的
- *(根据定义)* 不是凸优化问题：因为 f_1 不是凸的，h_1 不是仿射的。
- 等价于凸优化问题

$$\begin{aligned} \textit{minimize} \quad & f_0(x) = x_1^2 + x_2^2 \\ \textit{subject to} \quad & f_1(x) = x_1 \le 0 \\ & h_1(x) = x_1 + x_2 = 0 \end{aligned}$$

局部最优和全局最优

凸优化问题中，局部最优点就是 (全局) 最优点。
证明：设 x 是凸优化问题的局部最优解，y 是最优点使得 $f_0(y) < f_0(x)$
x 是局部最优解，存在 $R > O$. 有

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

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

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

$$f_0(z) \leq (1-\theta)f_0(x) + \theta f_0(y) < f_0(x)$$

与 x 是局部最优解矛盾。

可微函数 f_0 的最优性准则

x 是最优解，当且仅当 x 可行，且

$$\nabla f_0(x)^T(y-x) \geq 0, \qquad \text{对所有可行的}y.$$

如果 $\nabla f_0(x) \neq 0$,$-\nabla f_0(x)$ 在 x 处定义了可行集 X 的一个支撑超平面。

- **无约束问题**: x 是最优点当且仅当

$$x \in \textbf{dom}\, f_0, \qquad \nabla f_0(x) = 0$$

- **只含等式约束的问题**:

$$\text{maximize} \quad f_0(x) \qquad \text{subject to} \quad Ax = b$$

x 是最优点当且仅当存在 v，使得

$$x \in \textbf{dom}\, f_0, \qquad Ax = b, \qquad \nabla f_0(x) + A^T v = 0$$

- **非负象限中的极小化**

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

x 是最优点当且仅当

$$x \in \textbf{dom}\, f_0, \qquad x \succeq 0, \qquad \begin{cases} \nabla f_0(x)_i \geq 0 & x_i = 0 \\ \nabla f_0(x)_i = 0 & x_i > 0 \end{cases}$$

等价的凸问题

如果从一个问题的解，容易得到另一个问题的解，且反之亦然，称两问题 (非正式定义) 等价

- **消除等式约束**

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

等价于

$$
\begin{array}{ll}
\text{minimize} & f_0(Fz + x_0) \\
\text{subject to} & f_i(Fz + x_0) \le 0,
\end{array}
$$

其中 F 和 x_0 满足：

$$
Ax = b \iff x = Fz + x_0
$$

- **引入等式约束**

$$
\begin{aligned}
\text{minimize} \quad & f_0(A_0x+b_0) \\
\text{subject to} \quad & f_i(A_ix+b_i) \le 0, \quad i=1,\cdots,m
\end{aligned}
$$

等价于

$$
\begin{aligned}
\text{minimize} \quad & f_0(y_0) \\
\text{subject to} \quad & f_i(y_i) \le 0, \quad i=1,\cdots,m \\
& y_i = A_ix+b_i, \quad i=1,\cdots,m
\end{aligned}
$$

- **引入松弛变量**

$$\begin{aligned} \text{minimize} \quad & f_0(x) \\ \text{subject to} \quad & a_i^T x \le b_i, \quad i = 1, \cdots, m \end{aligned}$$

等价于

$$\begin{aligned} \text{minimize} \quad & f_0(x) \\ \text{subject to} \quad & a_i^T x + s_i = b_i, \quad i = 1, \cdots, m \\ & s_i \ge 0, \quad i = 1, \cdots, m \end{aligned}$$

线性规划

$$
\begin{aligned}
\text{minimize} \quad & c^T x + d \\
\text{subject to} \quad & Gx \leq h \\
& Ax = b
\end{aligned}
$$

- 目标函数和约束函数都是仿射的凸优化问题
- 可行集是一个多面体

例子

食谱问题：选择 n 种食物的质量 $x_1, \cdots, x_m$

- 每单位食物 j 的费用为 c_j，包含营养 i 的数量为 a_{ij}
- 健康饮食所需营养 i 的质量至少为 b_i

为了设计出一份最便宜的健康食谱，

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

例子

分片线性极小化

$$\text{minimize} \quad \max_{i=1,\cdots,m} (a_i^T x + b_i)$$

等价于一个线性规划问题：

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

例子

多面体的 Chebyshev 中心

多面体 $\mathcal{P} = \{x | a_i^T x \le b_i, \quad i = 1, \cdots, m\}$ 的 Chebyshev 中心是最大的内切球球心 $\mathcal{B} = \{x_c + u | \|u\|_2 \le r\}$

- 对于所有的 $x \in B, \quad a_i^T x \le b_i$ 当且仅当 $\sup\{a_i^T (x_c + u) | \|u\|_2 \le r\} = a_i^T x_c + r\|a_i\|_2 \le b_i$
- 因此 x_c, r 可以通过解决一个 LP 问题被确定下来

$$
\begin{array}{ll}
\text{maxmize} & r \\
\text{subject to} & a_i^T x_c + r\|a_i\|_2 \le b_i, \quad i = 1, \cdots, m
\end{array}
$$

线性分式规划

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & Gx \leq h \\ & Ax = b \end{array}$$

线性分式规划

$$f_0(x) = \frac{c^T x + d}{e^T x + f}, \qquad \mathbf{dom}\, f_0 = \{x | e^T x + f \geq 0\}$$

可以转换为等价的线性规划

$$\begin{array}{ll} \text{minimize} & c^T y + dz \\ \text{subject to} & Gy - hz \leq 0 \\ & Ay - bz = 0 \\ & e^T y + fz = 1 \\ & z \geq 0 \end{array}$$

二次优化问题

$$\begin{aligned} \text{minimize} \quad & (1/2)x^T P x + q^T x + r \\ \text{subject to} \quad & Gx \leq h \\ & Ax = b \end{aligned}$$

- $P \in S^n_+$ ，因此目标函数是凸二次型
- 在多面体上极小化一个凸二次函数

例

最小二乘问题

$$\|Ax - b\|_2^2$$

- 解析解 $x = A^\dagger b$，其中，$A^\dagger$ 是 A 的伪逆。
- 可以增加线性约束，$l \leq x \leq u$

例

带有随机损失的线性规划

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

- c 是随机向量，均值为 $\bar{c}$，协方差 Σ
- 因此，$c^T x$ 是随机变量，均值 $\bar{c}^T x$ 协方差 $x^T \Sigma x$
- $\gamma > 0$ 为风险厌恶参数；权衡期望损失和方差 (风险)

二次约束二次规划 (QCQP)

$$
\begin{array}{ll}
\text{minimize} & (1/2)x^T P_0 x + q_0^T x + r_0 \\
\text{subject to} & (1/2)x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1, \cdots, m \\
 & Ax = b
\end{array}
$$

- $P \in S_+^n$ ，目标和限制函数都是凸二次型
- 如果 $P \in S_{++}^n$，可行域是 m 个椭圆和一个仿射集合的交集

二阶锥规划

$$
\begin{aligned}
\text{minimize} \quad & f^T x \\
\text{subject to} \quad & \|A_i x + b_i\|_2 \le c_i^T x + d_i, \qquad i = 1, \cdots, m \\
& Fx = g
\end{aligned}
$$

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

- 不等式称为为二阶锥 (SOC) 约束:

$$(Ax + b, c^T x + d) \in R^{n_i + l}\text{的二阶锥中}$$

- $c_i = 0, i = 1, \cdots, m$ 时，SOCP 等同于 QCQO。
- $A_i = 0, i = 1, \cdots, m$ 时，SOCP 退化为 LP。

鲁棒线性规划

优化问题中的参数经常是不确定的，e.g. 在线性规划中

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & a_i^T x \leq b_i, \qquad i = 1, \cdots, m \end{aligned}$$

其中的参数 c, a_i, b_i 含有一些不确定性或变化

鲁棒线性规划

两种通用方式处理不确定性 (简化起见，只考虑 a_i)

- 确定性方法：所有的 $a_i \in \mathcal{E}_i$ 必须满足约束

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

- 随机性方法：a_i 是随机变量；以概率 η 满足约束

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \textbf{prob}(a_i^T x \leq b_i) \geq \eta, \qquad i = 1, \cdots, m \end{array}$$

SOCP 确定性方法

例 2

- 给定椭球 $\mathcal{E}_i$:

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

$\bar{a}_i$ 是椭球中心，半轴由 P_i 的奇异值/奇异向量决定。

SOCP 确定性方法

例 3

- 鲁棒线性规划

$$\begin{aligned} \textit{minimize} \quad & c^T x \\ \textit{subject to} \quad & a_i^T x \le b_i, \quad \forall a_i \in \mathcal{E}_i, \qquad i = 1, \cdots, m \end{aligned}$$

等价于 *SOCP*

$$\begin{aligned} \textit{minimize} \quad & c^T x \\ \textit{subject to} \quad & \bar{a}_i^T x + \|p_i^T x\|_2 \le b_i, \qquad i = 1, \cdots, m \end{aligned}$$

由 $\sup_{\|u\|_2}(\bar{a}_i + P_i u)^T x = \bar{a}_i^T x + \|P_i^T x\|_2$ 得到

SOCP 随机性方法

例 4

- 假设 a_i 是服从高斯分布，均值为 $\bar{a}_i$，协方差阵 Σ_i $(a_i \sim \mathcal{N}(\bar{a}_i, \Sigma_i))$
- $a_i^T x$ 服从高斯分布，均值为 $\bar{a}_i^T x$，协方差阵 $x^T \Sigma_i x$；因此

$$\textbf{\textit{prob}}(a_i^T x \leq b_i) = \phi(\frac{b_i - \bar{a}_i^T x}{\| \sum_i^{1/2} x\|_2})$$

其中 $\phi(x) = (1/\sqrt{2\pi}) \int_{-\infty}^{x} e^{-t^2/2} dt$

SOCP 随机性方法

例 5

- 鲁棒线性规划

$$\begin{aligned} \textit{minimize} \quad & c^T x \\ \textit{subject to} \quad & \textbf{prob}(a_i^T x \le b_i) \ge \eta, \quad i = 1, \cdots, m, \end{aligned}$$

其中 $\eta \ge 1/2$，等价于 *SOCP*

$$\begin{aligned} \textit{minimize} \quad & c^T x \\ \textit{subject to} \quad & \bar{a}_i^T x + \phi^{-1}(\eta)\|\Sigma_i^{1/2} x\|_2 \le b_i, \quad i = 1, \cdots, m, \end{aligned}$$