

约束优化习题讲义

2023 年 5 月 11 日

1 二次规划

Exercise 1 设 x^* 是一般的二次规划问题(122)的局部极小点, 则 x^* 也必是等式约束问题

$$(EQ) \begin{cases} \min & Q(x) = \frac{1}{2}x^\top Gx + c^\top x \\ \text{s.t.} & a_i^\top x = b_i, i \in \mathcal{E} \cup \mathcal{I}(x^*) \end{cases}$$

的局部极小点。反之, 如果 x^* 是一般问题(122)的可行点, 同时是(EQ)的K-T点, 且相应的Lagrange乘子 λ^* 满足 $\lambda_i^* \geq 0, i \in \mathcal{I}(x^*)$, 则 x^* 必是原问题(122)的K-T点。

一般的二次规划为

$$\begin{aligned} \min Q(x) &= \frac{1}{2}x^\top Gx + c^\top x \\ \text{s.t. } a_i^\top x &= b_i, i \in \mathcal{E} = \{1, \dots, m_e\} \\ a_i^\top x &\geq b_i, i \in \mathcal{I} = \{m_e + 1, \dots, m\} \end{aligned} \tag{122}$$

证明. x^* 是(122)的局部极小点, 即存在 x^* 的邻域 U , 使得

$$\begin{aligned} x^* &= \arg \min_x \left\{ Q(x) : \begin{array}{l} a_i^\top x = b_i, i \in \mathcal{E} \\ a_i^\top x \geq b_i, i \in \mathcal{I} \\ x \in U \end{array} \right\} \\ &= \arg \min_x \left\{ Q(x) : \begin{array}{l} a_i^\top x = b_i, i \in \mathcal{E} \cup \mathcal{I}(x^*) \\ a_i^\top x > b_i, i \in \mathcal{I} \setminus \mathcal{I}(x^*) \\ x \in U \end{array} \right\}. \end{aligned}$$

由连续性, 存在 x^* 的邻域 V 使得对 $\forall x \in V, i \in \mathcal{I} \setminus \mathcal{I}(x^*), a_i^\top x > b_i$, 有

$$\begin{aligned} x^* &= \arg \min_x \left\{ \begin{array}{l} Q(x) : \\ a_i^\top x = b_i, i \in \mathcal{E} \cup \mathcal{I}(x^*) \\ a_i^\top x > b_i, i \in \mathcal{I} \setminus \mathcal{I}(x^*) \\ x \in U \cap V \end{array} \right\} \\ &= \arg \min_x \left\{ \begin{array}{l} Q(x) : \\ a_i^\top x = b_i, i \in \mathcal{E} \cup \mathcal{I}(x^*) \\ x \in U \cap V \end{array} \right\} \end{aligned}$$

即 x^* 为(EQ)的局部极小点。

反之, x^* 是(EQ)的K-T点, 即

$$\nabla Q(x^*) - \sum_{i \in \mathcal{E} \cup \mathcal{I}(x^*)} \lambda_i^* a_i^\top = 0$$

且 $\lambda_i^* \geq 0, i \in \mathcal{I}(x^*)$, 则令

$$\bar{\lambda}_i = \begin{cases} \lambda_i^* & i \in \mathcal{E} \cup \mathcal{I}(x^*) \\ 0 & i \in \mathcal{I} \setminus \mathcal{I}(x^*) \end{cases}$$

加上 x^* 为(122)的可行点, 即有(122)的K-T条件:

$$\begin{cases} \nabla Q(x^*) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \bar{\lambda}_i a_i = 0 \\ a_i x^* = 0, i \in \mathcal{E} \\ a_i x^* \geq 0, i \in \mathcal{I} \\ \bar{\lambda}_i \geq 0, i \in \mathcal{I} \\ \bar{\lambda}_i (a_i^\top x - b_i) = 0, i \in \mathcal{I} \end{cases}$$

□

Exercise 2 考虑等式约束问题

$$(EQ1) \begin{cases} \min & \frac{1}{2} s^\top G s + (Gx^{(k)} + c)^\top s \\ \text{s.t.} & a_i^\top s = 0, i \in \mathcal{E}_k \end{cases}$$

求得其解为 $s^{(k)}$, 及其相应的Lagrange乘子 $\lambda_i^{(k)}, i \in \mathcal{E}_k$ 。

若 $s^{(k)} = 0$, 且 $\lambda_i^{(k)} \geq 0, i \in \mathcal{E}_k$ 不成立, 则由 $\lambda_{i_q}^{(k)} = \min_{i \in \mathcal{I}(x^{(k)})} \lambda_i^{(k)} < 0$ 确定 i_q , 那么如下问题

$$(EQ3) \begin{cases} \min & \frac{1}{2} s^\top G s + (Gx^{(k)} + c)^\top s \\ \text{s.t.} & a_i^\top s = 0, i \in \hat{\mathcal{E}} = \mathcal{E}_k \setminus \{i_q\}. \end{cases}$$

的解 \hat{s} 是原问题在当前点 $x^{(k)}$ 处的可行方向, 即 $a_{i_q}^\top \hat{s} \geq 0$ 。

证明. (EQ1)的K-T条件为

$$\begin{cases} Gs^{(k)} + Gx^{(k)} + c - \sum_{i \in \mathcal{E}_k} \lambda_i^{(k)} a_i = 0 \\ a_i^\top s^{(k)} = 0, i \in \mathcal{E}_k \end{cases}$$

对第一式, 由于 $s^{(k)} = 0$, 等价于

$$Gx^{(k)} + c - \sum_{i \in \mathcal{E}_k} \lambda_i^{(k)} a_i = 0 \quad (1)$$

左乘上 \hat{s}^\top , 得

$$\begin{aligned} \hat{s}^\top (Gx^{(k)} + c) - \sum_{i \in \mathcal{E}_k} \lambda_i^{(k)} \hat{s}^\top a_i &= 0 \\ \Rightarrow \hat{s}^\top (Gx^{(k)} + c) &= \lambda_{i_q}^{(k)} a_{i_q}^\top \hat{s} \quad (\text{由(EQ3)的约束条件}) \end{aligned}$$

只要证明 $\hat{s}^\top (Gx^{(k)} + c) \leq 0$ 即可。

反证: 若 $\hat{s}^\top (Gx^{(k)} + c) > 0$, 则取 $\tilde{s} = -\hat{s}$ 。 \tilde{s} 满足(EQ3)的约束条件, 且

$$\begin{aligned} &\frac{1}{2} \tilde{s}^\top G \tilde{s} + (Gx^{(k)} + c)^\top \tilde{s} \\ &= \frac{1}{2} \hat{s}^\top G \hat{s} - (Gx^{(k)} + c)^\top \hat{s} \\ &< \frac{1}{2} \hat{s}^\top G \hat{s} + (Gx^{(k)} + c)^\top \hat{s} \end{aligned}$$

与 \hat{s} 为(EQ3)的解矛盾。证毕。 □

另一个证明. (EQ3)的K-T条件为

$$\begin{cases} G\hat{s} + (Gx^{(k)} + c) - \sum_{i \in \hat{\mathcal{E}}} \hat{\lambda}_i a_i = 0 \\ a_i^\top \hat{s} = 0, i \in \hat{\mathcal{E}} \end{cases} \quad (2)$$

(1)与(2)的第一式作差, 得

$$\begin{aligned} G\hat{s} + \lambda_{i_q}^{(k)} a_{i_q} + \sum_{i \in \hat{\mathcal{E}}} (\lambda_i^{(k)} - \hat{\lambda}_i) a_i &= 0 \\ \Rightarrow \hat{s}^\top G \hat{s} + \lambda_{i_q}^{(k)} a_{i_q}^\top \hat{s} &= 0 \end{aligned}$$

只要证明 $\hat{s}^\top G \hat{s} \geq 0$ 。

若 $\hat{s}^\top G \hat{s} < 0$, 则(EQ3)无下界, 与 \hat{s} 为其解矛盾, 故得证。 □

2 非线性约束最优化

Exercise 4 证明(125)中定义的 $\psi(x, \lambda)$ 是关于Lagrange-Newton法的下降函数。

证明.

$$\begin{aligned}\nabla\psi(x, \lambda) &= \begin{pmatrix} \nabla_x\psi(x, \lambda) \\ \nabla_\lambda\psi(x, \lambda) \end{pmatrix} \\ &= \begin{pmatrix} 2W(x, \lambda)(\nabla f(x) - A(x)^\top \lambda) + A(x)^\top c(x) \\ -2A(x)(\nabla f(x) - A(x)^\top \lambda) \end{pmatrix} \\ &= 2 \begin{pmatrix} W(x, \lambda) & -A(x)^\top \\ -A(x) & \end{pmatrix} \begin{pmatrix} \nabla f(x) - A(x)^\top \lambda \\ -c(x) \end{pmatrix}\end{aligned}$$

由式(124),

$$\begin{pmatrix} W(x, \lambda) & -A(x)^\top \\ -A(x) & \end{pmatrix} \begin{pmatrix} \delta_x \\ \delta_y \end{pmatrix} = - \begin{pmatrix} \nabla f(x) - A(x)^\top \lambda \\ -c(x) \end{pmatrix}$$

有

$$\begin{aligned}\nabla\psi(x, \lambda)^\top \begin{pmatrix} \delta_x \\ \delta_y \end{pmatrix} &= -2 \begin{pmatrix} \nabla f(x) - A(x)^\top \lambda \\ -c(x) \end{pmatrix}^\top \begin{pmatrix} \nabla f(x) - A(x)^\top \lambda \\ -c(x) \end{pmatrix} \\ &= -2\psi(x, \lambda).\end{aligned}$$

□

Exercise 5 证明罚函数法求解带误差界近似问题的算法有限终止性。

证明. 反证。假设对所有的 σ_k 都有 $\|c(x(\sigma_k))_-\| \geq \varepsilon$ 。由题意, 存在 \bar{x} 满足 $\|c(\bar{x})_-\| < \varepsilon$ 。由 $x(\sigma_k)$ 的定义有

$$f(\bar{x}) + \sigma_k \|c(\bar{x})_-\|^2 \geq f(x(\sigma_k)) + \sigma_k \|c(x(\sigma_k))_-\|^2.$$

由引理1(3), $f(x(\sigma_k)) \geq f(x(\sigma_1))$, 故

$$f(\bar{x}) + \sigma_k \|c(\bar{x})_-\|^2 \geq f(x(\sigma_k)) + \sigma_k \|c(x(\sigma_k))_-\|^2 \geq f(x(\sigma_1)) + \sigma_k \|c(x(\sigma_k))_-\|^2,$$

整理可得

$$0 > \|c(\bar{x})_-\|^2 - \varepsilon^2 \geq \|c(\bar{x})_-\|^2 - \|c(x(\sigma_k))_-\|^2 \geq \frac{1}{\sigma_k} (f(x(\sigma_1)) - f(\bar{x})).$$

$k \rightarrow \infty$ 时 σ_k 也趋于无穷, 故上式取极限得 $0 > 0$, 矛盾。

□

引理1 考虑简单罚函数

$$P_\sigma(x) = f(x) + \sigma \|c(x)_-\|^2$$

记 $x(\sigma)$ 是无约束问题 $\min_{x \in \mathbb{R}^n} P_\sigma(x)$ 的最优解, 设 $\sigma_{k+1} > \sigma_k > 0$, 则有

$$P_{\sigma_k}(x(\sigma_k)) \leq P_{\sigma_{k+1}}(x(\sigma_{k+1})), \quad (3)$$

$$\|c(x(\sigma_k))_-\| \geq \|c(x(\sigma_{k+1}))_-\|, \quad (4)$$

$$f(x(\sigma_k)) \leq f(x(\sigma_{k+1})) \quad (5)$$

证明. (1) 由 $x(\sigma_k)$ 的定义,

$$\begin{aligned} f(x(\sigma_k)) + \sigma_k \|c(x(\sigma_k))_-\|^2 &\leq f(x(\sigma_{k+1})) + \sigma_k \|c(x(\sigma_{k+1}))_-\|^2 \\ &\leq f(x(\sigma_{k+1})) + \sigma_{k+1} \|c(x(\sigma_{k+1}))_-\|^2 \end{aligned}$$

(2) 由 $x(\sigma_k)$ 和 $x(\sigma_{k+1})$ 的定义:

$$\begin{aligned} f(x(\sigma_k)) + \sigma_k \|c(x(\sigma_k))_-\|^2 &\leq f(x(\sigma_{k+1})) + \sigma_k \|c(x(\sigma_{k+1}))_-\|^2 \\ f(x(\sigma_{k+1})) + \sigma_{k+1} \|c(x(\sigma_{k+1}))_-\|^2 &\leq f(x(\sigma_k)) + \sigma_{k+1} \|c(x(\sigma_k))_-\|^2 \end{aligned}$$

两式相加, 得

$$\begin{aligned} \sigma_k \|c(x(\sigma_k))_-\|^2 + \sigma_{k+1} \|c(x(\sigma_{k+1}))_-\|^2 &\leq \sigma_k \|c(x(\sigma_{k+1}))_-\|^2 + \sigma_{k+1} \|c(x(\sigma_k))_-\|^2 \\ (\sigma_{k+1} - \sigma_k) \|c(x(\sigma_{k+1}))_-\|^2 &\leq (\sigma_{k+1} - \sigma_k) \|c(x(\sigma_k))_-\|^2 \\ \Rightarrow \|c(x(\sigma_{k+1}))_-\|^2 &\leq \|c(x(\sigma_k))_-\|^2 \end{aligned}$$

(2)式得证。

(3) 由(1)、(2)立得。

□

引理3 令 $\delta = \|c(x(\sigma))_-\|$, 则 $x(\sigma)$ 也是约束问题

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & \|c(x)_-\| \leq \delta \end{aligned}$$

的最优解。

证明. 对任意 x 满足 $\|c(x)_-\| \leq \delta$,

$$\begin{aligned} f(x) + \sigma \|c(x)_-\|^2 &\geq f(x(\sigma)) + \sigma \|c(x(\sigma))_-\|^2 \\ f(x) &\geq f(x(\sigma)) + \sigma (\delta^2 - \|c(x)_-\|^2) \geq f(x(\sigma)). \end{aligned}$$

□

Exercise 6 给出约束最优化问题的二阶充分最优性条件，并用于说明增广Lagrange函数的极小点与原问题最优解的等价性。

考虑等式约束问题

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c(x) = 0 \end{aligned}$$

其中 $c(x) = (x_1(x), \dots, c_{m_e}(x))^T$ 。定义增广Lagrange函数

$$P(x, \lambda, \sigma) = L(x, \lambda) + \frac{\sigma}{2} \|c(x)\|^2.$$

本题包括两个命题的证明：

- (1) 设 \bar{x} 是等式约束问题的可行解，且对某个 $\bar{\lambda}$, \bar{x} 满足 $P(x, \bar{\lambda}, \sigma)$ 的极小点二阶充分条件，则 \bar{x} 是该等式约束问题的严格局部最优解。
- (2) 设 x^* 和 λ^* 满足等式约束问题局部最优解的二阶充分条件，则存在 σ_0 使得当 $\sigma > \sigma_0$ 时， x^* 是函数 $P(x, \lambda^*, \sigma)$ 的严格局部极小点。

证明. (1) 设 \bar{x} 满足 $P(x, \bar{\lambda}, \sigma)$ 的极小点二阶充分条件，故存在 $\delta > 0$ ，对任意 x 满足 $0 < \|x - \bar{x}\| < \delta$ 都有

$$P(x, \bar{\lambda}, \sigma) = f(x) + \bar{\lambda}^T c(x) + \frac{\sigma}{2} \|c(x)\|^2 > P(\bar{x}, \bar{\lambda}, \sigma) = f(\bar{x}),$$

因而对满足 $0 < \|x - \bar{x}\| < \delta$ 的可行点 x 均有

$$f(x) = P(x, \bar{\lambda}, \sigma) > P(\bar{x}, \bar{\lambda}, \sigma) = f(\bar{x}),$$

即 \bar{x} 是等式约束问题的严格局部最优解

(2) 原问题局部最优解的二阶充分条件为

$$\nabla_x L(x^*, \lambda^*) = 0, \quad c(x^*) = 0$$

且对所有 $d \in \text{Ker } A(x^*)$ 均有

$$d^T \nabla_{xx}^2 L(x^*, \lambda^*) d > 0$$

则 $\nabla_x P(x^*, \lambda^*, \sigma) = \nabla_x L(x^*, \lambda^*) + \sigma A(x^*)^T c(x^*) = 0$ 对所有 $\sigma > 0$ 成立。

下证存在 σ_0 使得当 $\sigma > \sigma_0$ 时，有 $\nabla_x^2 P(x^*, \lambda^*, \sigma) > 0$ 。

反证，假设对任意正整数 k ，都有存在方向 d_k ，满足

$$\|d_k\| = 1 \text{ 且 } d_k^T \nabla_x^2 P(x^*, \lambda^*, k) d_k \leq 0, \quad (\text{即取 } \sigma_k = k)$$

将 $\nabla_{xx}^2 P(x^*, \lambda^*, \sigma)$ 展开得

$$d_k^T (\nabla_{xx}^2 L(x^*, \lambda^*) + k A(x^*)^T A(x^*)) d_k \leq 0$$

$$\begin{aligned} d_k^T \nabla_{xx}^2 L(x^*, \lambda^*) d_k &\leq -k \|A(x^*) d_k\|^2 \\ -\frac{1}{k} d_k^T \nabla_{xx}^2 L(x^*, \lambda^*) d_k &\geq \|A(x^*) d_k\|^2 \\ \frac{1}{k} \|\nabla_{xx}^2 L(x^*, \lambda^*)\|_2 &\geq \|A(x^*) d_k\|^2 \end{aligned}$$

其中 $\|\nabla_{xx}^2 L(x^*, \lambda^*)\|_2$ 为 $\nabla_{xx}^2 L(x^*, \lambda^*)$ 的谱范数。由于 d_k 属于 \mathbb{R}^n 中的紧集

$$\{d \in \mathbb{R}^n : \|d\| = 1\},$$

因此有聚点 \bar{d} ，满足

$$\lim_{k \rightarrow \infty} \frac{1}{k} \|\nabla_{xx}^2 L(x^*, \lambda^*)\|_2 = 0 \geq \|A(x^*) \bar{d}\|^2$$

因此 $\bar{d} \in \text{Ker } A(x^*)$ ，但 $\bar{d}^T \nabla_{xx}^2 L(x^*, \lambda^*) \bar{d} \leq 0$ 与原问题二阶充分条件矛盾。

□