约束优化习题讲义

2021 年 6 月 3 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

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

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

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$\lambda_i^* \geqslant 0,\ i \in \mathcal{I}(x^*)$，则$x^*$必是原问题(122)的K-T点。

一般的二次规划为

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

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

证明. x^*是(122)的局部极小点，即存在x^*的邻域U，使得

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

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

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

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

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

加上x^*为(122)的可行点，即有(122)的K-T条件：

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

□

Excercise 2 考虑等式约束问题

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

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$\lambda_i^{(k)} \geqslant 0$，$i \in \mathcal{E}_k$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

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

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

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

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

左乘上$\hat{s}^\top$，得

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

$$\Rightarrow\ \hat{s}^\top(Gx^{(k)} + c) = \lambda_{i_q}^{(k)} a_{i_q}^\top \hat{s} \quad (\text{由(EQ3)的约束条件})$$

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

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

$\hat{\mathcal{E}} = \mathcal{E}_k \setminus \{i_q\}$

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

(1)与(2)的第一式作差，得

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

$\hat{s}^\top$

$$\Rightarrow\ \hat{s}^\top G\hat{s} + \lambda_{i_q}^{(k)} a_{i_q}^\top \hat{s} = 0$$

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

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。**

$$\psi(x,\lambda) = \|\nabla f(x) - A(x)^T\lambda\|^2 + \|c(x)\|^2.$$

证明.

$W(x,\lambda) = \nabla^2 f(x) - \sum_{i=1}^{m} \lambda_i \nabla^2 c_i(x)$

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

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

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

□

Excercise 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。**

$\varepsilon > \min_{x\in\mathbb{R}^n} \|c(x)_-\|$

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

(0) 任选初始点$x^{(0)}$，给定初始罚因子$\sigma_0 > 0$及$\beta > 1, \varepsilon > 0$. 令$k := 0$.

(1) 以$x^{(k)}$作为初始迭代点求解无约束问题的极小点，即

$$x(\sigma_k) = \arg\min_{x\in\mathbb{R}^n} P_{\sigma_k}(x).$$

(2) 若$\|c(x(\sigma_k))_-\| < \varepsilon$，则停止迭代并取$x(\sigma_k)$为原约束问题的近似最优解；否则，置$x^{(k+1)} = x(\sigma_k)$，$\sigma_{k+1} = \beta\sigma_k$，令$k := k+1$返回第(1)步。

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

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

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

$k \to \infty$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

$x(\sigma) = \underset{x\in R^n}{\arg\min}\, P_\sigma(x)$

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

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

$$P_{\sigma_k}(x(\sigma_k)) \leqslant P_{\sigma_{k+1}}(x(\sigma_{k+1})), \tag{3}$$

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

$$f(x(\sigma_k)) \leqslant f(x(\sigma_{k+1})) \tag{5}$$

证明. (1) 由$x(\sigma_k)$的定义，

$x(\sigma_k) = \underset{x\in R^n}{\arg\min}\, P_{\sigma_k}(x)$

$\sigma_{k+1} = \beta\sigma_k > \sigma_k$

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

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

$$f(x(\sigma_k)) + \sigma_k\|c(x(\sigma_k))_-\|^2 \leqslant f(x(\sigma_{k+1})) + \sigma_k\|c(x(\sigma_{k+1}))_-\|^2$$

$$f(x(\sigma_{k+1})) + \sigma_{k+1}\|c(x(\sigma_{k+1}))_-\|^2 \leqslant f(x(\sigma_k)) + \sigma_{k+1}\|c(x(\sigma_k))_-\|^2$$

两式相加，得

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

(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

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

的最优解。

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

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

□

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。

* 各种情况下的最优性条件.
无约束，带约束.
一阶 二阶

考虑等式约束问题

$$\min\ f(x)$$
$$\text{s.t. } c(x)=0$$

其中$c(x)=(x_1(x),\ldots,c_{m_e}(x))^\top$。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。 $\nabla P(x^*,\lambda^*,\sigma)=0$ & $\nabla^2 P(x^*,\lambda^*,\sigma)>0$

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

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

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

$$\nabla_xL(x^*,\lambda^*)=0,\ \ c(x^*)=0$$

且对所有$d\in Ker\,A(x^*)$均有

$$d^\top\nabla^2_{xx}L(x^*,\lambda^*)d>0$$

则$\nabla_xP(x^*,\lambda^*,\sigma)=\nabla_xL(x^*,\lambda^*)+\sigma A(x^*)^\top c(x^*)=0$对所有$\sigma>0$成立。

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

反证，假设对任意正整数k，都有存在方向d_k，满足

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

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

$$d_k^T\left(\nabla_{xx}^2 L(x^*,\lambda^*)+kA(x^*)^TA(x^*)\right)d_k\le 0$$

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

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

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

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

□