本节讲述三种这类算法

（1）外罚函数法

（2）内罚函数法

（3）乘子法

## 4.2.1 外罚函数法

### 1. 经济解释

把目标函数看为价格，约束条件看为某种〝规定〞的范围，则问题可描述为：在规定的范围内买价格最低的东西。若同时制定超范围购买的高〝罚款〞政策，例如关税政策。这样

总代价 = 价格 + 罚款

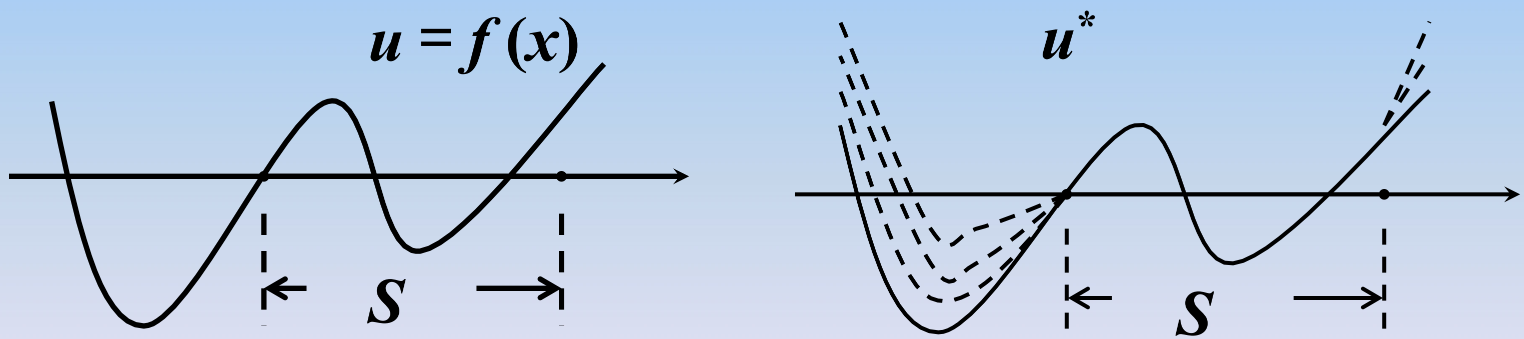
当罚款高到一定程度时，用最小总代价所买到的东西总在规定的范围内。于是约束问题转化成为无约束问题。为达此目的，可逐次加大罚款求极小。

## 2. 图形解释

设 $u = f(x), x \in R^1$ 约束范围为$S$。改造$u$为

$$u^* = \begin{cases} f(x), & x \in S \\ \text{加大}, & x \notin S \end{cases}$$

使 $u^*$的无约束极小点成为$u$ 的约束极小点，如图所示。

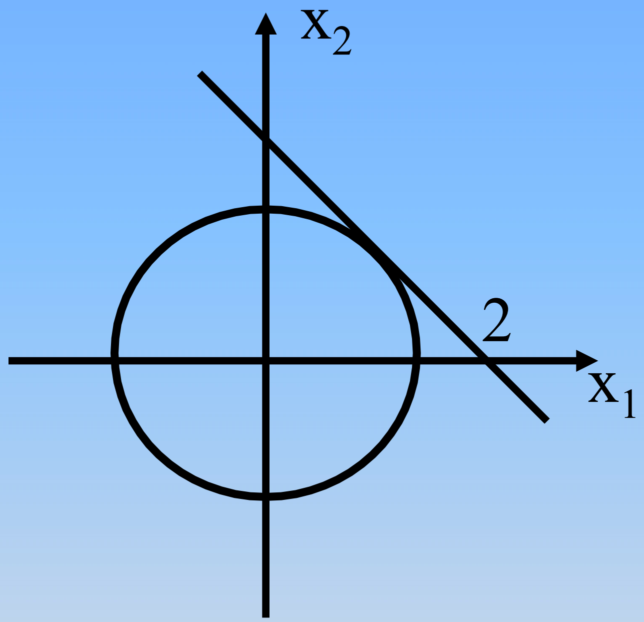



图

例 4.2.1 求解约束问题

$$\min f(x_1,x_2) = x_1^{\ 2} + x_2^{\ 2}$$

$$s.t. \quad x_1 + x_2 - 2 = 0$$




解　其最优解 $x^* = (1,1)^T$

$$F(x_1,x_2)=\begin{cases} x_1^2 + x_2^2 & x_1 + x_2 = 2 \\ +\infty & x_1 + x_2 \neq 2 \end{cases}$$

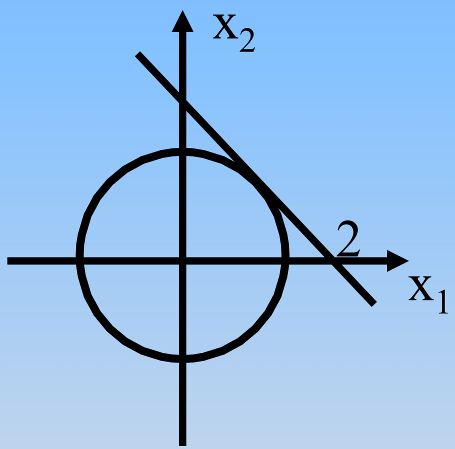
函数 $F(x_1,x_2)$的性态极坏，无法用有效的无约束优化算法求解。

$$\begin{aligned} \text{Min } f(x_1,x_2) &= x_1^2 + x_2^2 \\ s.t. \quad x_1 + x_2 - 2 &= 0 \end{aligned}$$

考虑：

$$P(x_1,x_2,\sigma) = x_1^2 + x_2^2 + \sigma(x_1 + x_2 - 2)^2$$

其中 $\sigma$ 是很大的正数。




求解无约束问题，其最优解的解析式：

$$x_1^{(\sigma)} = x_2^{(\sigma)} = \frac{2\sigma}{2\sigma+1}$$

当 $\sigma \to +\infty$ 时，有

首先考虑仅含等式约束的优化问题：

$$\min f(x)$$

$$s.t.\quad c_i(x)=0 \quad i \in E=\{1,2,\square,l\}$$

构造

$$P(x,\sigma)=f(x)+\sigma\sum_{i=1}^{l}|c_i(x)|^{\beta}, \qquad \beta \ge 1$$

$$\tilde{P}(x)=\sum_{i=1}^{l}|c_i(x)|^{\beta}, \quad \beta \ge 1$$

当 $x$ 为可行解时，当 $P(x,\sigma)$ $f(x)$，不受惩罚；

$x$ 不是可行解时，$P(x,\sigma)=f(x)+\sigma\tilde{P}(x)$ 罚因子

$\sigma$

因此当 $\sigma$ 充分大时，要使 $P(x,\sigma)$ 取极小值，

$P(x,\sigma)$

其次考虑含不等式约束的最优化问题:

$$\min f(x)$$
$$s.t.\quad c_i(x)\ge 0 \quad i\in I=\{1,2,\cdots,m\}$$

构造函数

$$P(x,\sigma)=f(x)+\sigma\tilde{P}(x),\quad \sigma>0$$

$$\tilde{P}(x)=\begin{cases}0 & c_i(x)\ge 0\\ \sum_{i=1}^{m}|c_i(x)|^{\alpha} \quad \alpha\ge 1, & c_i(x)<0\end{cases}$$

或 $\tilde{P}(x)=\sum_{i=1}^{m}|\min(0,c_i(x))|^{\alpha}=\sum_{i=1}^{m}\left(\frac{|c_i(x)|-c_i(x)}{2}\right)^{\alpha}$

当 $x$ 为可行解时，$c_i(x) \ge 0,\ P(x,\sigma) = f(x)$ 不受惩罚；

当 $x$ 不是可行解时，

$$c_i(x) < 0,\ \tilde{P}(x) > 0,\ P(x,\sigma) = f(x) + \sigma \tilde{P}(x)$$

$\sigma$ 越大，惩罚越重。

因此当 $\sigma$ 充分大时，要使 $P(x,\sigma)$ 取极小值，

$P(x,\sigma)$

考虑一般约束最优化问题：

$$\begin{aligned}
&\min f(x) \\
&s.t. \quad c_i(x)=0 \quad i\in E=\{1,2,\cdots,l\} \\
&\qquad\ \ c_i(x)\ge 0 \quad i\in I=\{l+1,\cdots,m\}
\end{aligned}$$

可行域为

$$D=\{x\in R^n \mid c_i(x)=0, i\in E, c_i(x)\ge 0, i\in I\}$$

构造函数

$$P(x,\sigma)=f(x)+\sigma\tilde{P}(x), \quad \sigma>0$$

其中

$$\tilde{P}(x)=\sum_{i=1}^{l}|c_i(x)|^{\beta}+\sum_{i=l+1}^{m}|\min(0,c_i(x))|^{\alpha} \quad \alpha\ge 1, \beta\ge 1$$

当 $x\in D$ 时 $\tilde{P}(x)=0$

$x\notin D$ 时 $\tilde{P}(x)>0$

$$P(x,\sigma) = f(x) + \sigma \tilde{P}(x)$$

称为约束问题的增广目标函数，

$\alpha \quad \beta$

求解约束问题转化为求增广目标函数的系列无约束极小

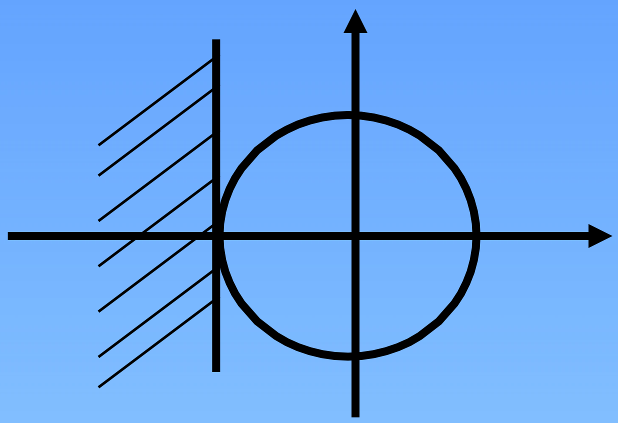
即求解

$$\min P(x,\sigma_k)$$

其中 $\{\sigma_k\}$ 为正的数列且 $\sigma_k \to +\infty$.

例 4.2.2：求解约束问题

$$\min f(x) = x_1^2 + x_2^2$$
$$s.t. \quad x_1 + 1 \le 0$$



解　最优解 $x^* = (-1,0)^T$

$$f(x^*) = 1$$

$$P(x,\sigma) = x_1^2 + x_2^2 + \sigma[\min(0,-x_1-1)]^2$$
$$= \begin{cases} x_1^2 + x_2^2 & x_1 + 1 \le 0 \\ x_1^2 + x_2^2 + \sigma(x_1+1)^2 & x_1 + 1 > 0 \end{cases}$$

$$\frac{\partial P}{\partial x_1} = \begin{cases} 2x_1 & x_1 < -1 \\ 2x_1 + 2\sigma(x_1+1) & x_1 > -1 \end{cases}$$

$$\frac{\partial P}{\partial x_2} = 2x_2$$

令 $\dfrac{\partial P}{\partial x_1}$ $\dfrac{\partial P}{\partial x_2}$ $=0$

得 $x_1(\sigma) = -\dfrac{\sigma}{\sigma+1}$

它是 $\min P(x,\sigma)$ 的最优解，最优值

$$P(x,\sigma) = \left(-\frac{\sigma}{\sigma+1}\right)^2 + \sigma\left(\frac{1}{\sigma+1}\right)^2 = \frac{\sigma}{\sigma+1}$$

当 $\sigma \to +\infty$ 时，故 $x_1(\sigma) \to -1,\ x_2(\sigma) \to 0$

由例 4.2.1 和例 4.2.2 可以看出：

当 $\sigma \to +\infty$ 时 $P(x,\sigma)$ 的最优解 $x(\sigma)$ 趋向于极限 $x^*$，$x^*$ 即为原约束问题的最优解。

$x(\sigma)$ 往往不满足约束条件，在例 4.2.1 中，

$$x_1(\sigma)+x_2(\sigma)=\frac{4\sigma}{2\sigma+1}\neq 2,$$

在例 4.2.2 中，

$$x_1(\sigma)=-\frac{\sigma}{\sigma+1}>-1,$$

$x(\sigma)$ 都是从可行域外部趋向于最优解 $x^*$ 的。因此称 $P(x,\sigma)$ 为外罚函数，而称这种解法为外罚函数法。

**例** 用外点法求

$$\begin{cases} \min f(\mathrm{x}) = \frac{1}{3}(x_1+1)^3 + x_2, \\ s.t. \quad 1 - x_1 \le 0, \\ \qquad -x_2 \le 0. \end{cases}$$

$$P(x,\sigma_k) = f(x) + \sigma_k(\min(0, c_1(x))^2 + \min(0, c_2(x))^2)$$

$$= \frac{1}{3}(x_1+1)^3 + x_2 + \sigma_k(\min(0, x_1 - 1)^2 + \min(0, x_2)^2)$$

| $\sigma_k$ | $x_1$ | $x_2$ | $P(x,\sigma_k)$ | $f(x)$ |
|---|---|---|---|---|
| 0.001 | -0.93775 | -500.00000 | -249.9962 | -500.0000 |
| 0.01 | -0.80975 | -50.0000 | -24.965 | -49.9977 |
| 0.1 | -0.45969 | -5.0000 | -2.2344 | -4.9474 |
| 1 | 0.23607 | -0.50000 | 0.9631 | 0.1295 |
| 10 | 0.83216 | -0.05000 | 2.3068 | 2.0001 |
| 100 | 0.98039 | -0.00500 | 2.6249 | 2.5840 |
| 1000 | 0.99800 | -0.00050 | 2.6624 | 2.6582 |
| 10000 | 0.99963 | -0.00005 | 2.6655 | 2.6652 |
| $\infty$ | 1 | 0 | 8/3 | 8/3 |

通过求解一系列无约束最优化问题来求解约束最优化问题的方法又称序列无约束极小化技术（SUMT），故外罚函数法又称SUMT外点法。

## 2 算法4.2.1

已知约束问题，取控制误差$\varepsilon>0$和罚因子的放大系数$c>1$（可取$\varepsilon=10^{-4}$，$c=10$）。

Step 1 给定初始点$x_0$（可以不是可行点）和初始罚因子$\sigma_1$（可取$\sigma_1=1$）。令$k=1$.

Step 2 以$x_{k-1}$为初始点求无约束问题：

$$\min P(x,\sigma_k)=f(x)+\sigma_k P(x)$$

其中$P(x)$如前定义，得最优解

$$x_k=x(\sigma_k)$$

Step 3 若$\sigma_k P(x)<\varepsilon$，则以$x_k$为问题的近似最优解，停止。否则令

$$\sigma_{k+1}=c\sigma_k \quad k=k+1$$

,转 step2.

# 3.收敛性

引理 4.2.1 对于由 SUMT 外点法产生的点列 $\{x_k\}$ $k \geq 1$，总有

$$P(x_{k+1},\sigma_{k+1}) \geq P(x_k,\sigma_k)$$

$$\tilde{P}(x_k) \geq \tilde{P}(x_{k+1})$$

$$f(x_{k+1}) \geq f(x_k)$$

## 定理 4.2.2

设一般约束问题

$$
\begin{aligned}
&\min f(x) \\
&s.t.\ c_i(x) = 0 \quad i \in E = \{1, 2, \cdots, l\} \\
&\qquad c_i(x) \ge 0 \quad i \in I = \{l+1, \cdots, m\}
\end{aligned}
$$

和外罚函数无约束问题

$$\min P(x, \sigma) = f(x) + \sigma_k P(x)$$

$$P(x) = \sum_{i=1}^{l} |c_i(x)|$$

$x^*$ $x_k$ $\forall k \ge 1$ $\sigma_k$

$\sigma_{k+1} \ge \sigma_k$ $\sigma_k \to +\infty$

则由SUMT外点法产生的点列$\{x_k\}$ 的任何聚点 $x$ 般约束问题的整体最优解。

## 4.2.2 内罚函数法

1. 考虑不等式约束问题：

$$\min f(x), x \in R^n$$
$$s.t. \quad c_i(x) \geq 0, i = 1, \cdots, m$$

当 x 从可行域

$$D = \{x \in R^n \mid c_i(x) \geq 0, i = 1, \cdots, m\}$$

的内部趋近于边界时，则至少有一个

$c_i(x)$ 趋于零，

因此，可构造如下的增广目标函数：

$$B(x,r) = f(x) + r\tilde{B}(x)$$

其中令

$$\tilde{E}(\ ) = \sum_{i=1}^{m} \frac{1}{c_i(x)}$$

或者

$$\tilde{B}(x) = -\sum_{i=1}^{m} \ln(c_i(x))$$

称为内罚函数或障碍函数（Barrier function），参数 r>0 仍称为罚因子。

我们取正的数列$\{r_k\}$，且 $r_k \to 0$，则求解不等式约束问题转化为求解系列无约束问题，即

$$\min B(x, r_k) = f(x) + r_k B(x)$$

其中$B(x)$如前面定义的内罚函数，

$$B(x) = \sum_{i=1}^{m} \frac{1}{c_i(x)}$$

或者

$$B(x) = -\sum_{i=1}^{m} \ln(c_i(x))$$

称为内罚函数法或 SUMT 内点法。

# 算法 4.2.2 内罚函数法

已知不等式约束问题，其可行域的内点集 $D_0 \neq \phi$ 控制误差

$\varepsilon > 0$ 罚因子的缩小系数 0<c<1（可取

$\varepsilon = 10^{-4}, c = 0.1$

Step 1 选定初始点 $x_0 \in D_0$，给定 $r_1 > 0$（取 $r_1 = 10$）令 k=1。

Step 2 以 $x_{k-1}$ 为初始点，求解无约束问题

$$\min B(x, r_k) = f(x) + r_k \bar{B}(x)$$

得最优解 $x_k = x(r_k)$。

Step 3 若 $r_k \bar{B}(x_k) < \varepsilon$，则 $x_k$ 为不等式约束问题的近似最优解，停止。否则，令

$r_{k+1} = cr_k, k = k+1$ 转 Step 2.

例 4.2.4　用内点法求解

$$\min f(x_1,x_2)=\frac{1}{3}(x_1+1)^3+x_2$$

$$s.t.\quad 1-x_1\le 0$$

$$x_2\ge 0$$

解　增广目标函数为

$$B(x_1,x_2,r)=\frac{1}{3}(x_1+1)^3+x_2+r\left(\frac{1}{x_1-1}+\frac{1}{x_2}\right)$$

增广目标函数也可取为

$$B(x_1,x_2,r)=\frac{1}{3}(x_1+1)^3+x_2-r(\ln(x_1-1)+\ln x_2)$$

令

$$\frac{\partial B}{\partial x_1} = (x_1 + 1)^2 - \frac{r}{(x_1 - 1)^2} = 0,$$

$$\frac{\partial B}{\partial x_2} = 1 - \frac{r}{x_2^{\ 2}} = 0,$$

所以$x(r) = (\sqrt{1+\sqrt{r}}, \sqrt{r})^T,$

当$r \xrightarrow{\text{时，得}} 0$

$$x^* = (1,0)^T, f^* = 8/3$$

SUMT 内点法在实际计算时应该采用系列无约束最优化方法。

取 $x_0 = (3,4)^T, r_1 = 10, c = 0.1, \varepsilon = 0.1,$ 迭代结果如表 4－2

表 4－2

| k | $\gamma_k$ | $x_{k+1}^T$ | $f_{k+1}$ | $r_k B(x_{k+1})$ |
|---|---|---|---|---|
| 1 | 10 | (2.0402,3.1623) | 12.5290 | [illegible] |
| 2 | $10^{-1}$ | | | |
| 3 | $10^{-2}$ | (1.1473,0.3162) | 3.6165 | 0.9951 |
| 4 | $10^{-3}$ | (1.0488,0.1000) | 2.9667 | 0.3049 |
| 5 | $10^{-4}$ | (1.0157,0.0316) | 2.7616 | 0.0953 |

所以最优解 $x^* \approx x_5 = (1.0016, 0.0316)^T$，最优值

$f^* \approx f_5 = 2.7046$

## 2. 收敛性

引理 4.2.3　对于由 SUMT 内点法产生的点列 $\{x_k\}(k \geq 1)$

总有

$$B(x_{k+1}, r_{k+1}) \leq B(x_k, r_k),$$

即增广目标函数 $B(x_k, r_k)$ 关于 k 单调减少的。

定理 4.2.4

设不等式约束问题的可行域 D 的内点集 $D_0=\{x\in R^n \mid c_i(x)>0, i\in I\}$ 非空，$f(x)$ 在 D 上存在极小点 $x^*$，对严格单减的正数列 $\{r_k\}: r_{k+1}<r_k$ 且 $r_k\to 0$，

则由 SUMT 内点法产生的点列 $\{x_k\}$ 的任何聚点 $\bar{x}$ 是不等式约束问题的最优解

与外点法的收敛定理一样,本定理中的最优解均指整体最优解.对于局部最优解,也有类似的定理.

## 4.2.3　几点说明

1. 方法简单、易懂。

2. 混合罚函数法

当初始点 $x_0$ 给定后，对等式约束和不被 $x_0$ 满足的不等式约束采用外罚函数、而对被 $x_0$ 满足的那些不等式约束采用内罚函数。

即对一般约束问题，引进增广目标函数

$$p(x,r) = f(x) - r\sum_{i\in I_1}\ln(c_i(x)) + \frac{1}{r}\left[\sum_{i\in I_2}(\min(0,c_i(x)))^2 + \sum_{i=1}^{l}(c_i(x))^2\right]$$

其中

$$I_1 = \{i \mid c_i(x_0) > 0, i\in I\},$$

$$I_2 = \{i \mid c_i(x_0) \le 0, i\in E\cup I\},$$

其迭代步骤与内点法相同

3． 罚函数法由于增广目标函数的 Hesse 矩阵

$\nabla^2 P(x,\sigma)$ 和 $\nabla^2 B(x,r)$ 的条件数随 $\sigma$ 的增大和 r 减少而变大，造成在求解系列无约束问题的困难，使得选择罚函数因子 $\sigma$ 和 r 时往往处于进退维谷的境地。

## 4.2.4 乘子法

罚函数法的主要缺点之一，是增广目标函数的病态性质，其原因是由罚因子 $\sigma_k \to \infty$（或 $r_k \to 0$）引起的。

能不能找到 $\lambda^*$，使 $(x^*, \lambda^*)$ 就是 $L(x, \lambda)$ 的极小点？

但是 Lagrange 函数的极小点往往是不存在的。

例 4.2.5 求解约束问题

$$\min f(x) = x_1^2 - 3x_2 - x_2^2$$

$$s.t. \quad x_2 = 0$$

最优解 $x^* = (0, 0)^T$

$$L(x, \lambda) = x_1^2 - 3x_2 - x_2^2 - \lambda x_2$$

$$= x_1^2 - (\lambda + 3)x_2 - x_2^2$$

对于任何 $\lambda$，$L(x, \lambda)$ 关于 $x$ 的极小点是不存在的。

考虑函数 $L(x,\lambda)+\sigma P(x)$

称为增广 Lagrange 函数。通过求解增广 Lagrange 函数的系列无约束问题的解来获得原约束问题的解，这就是下面介绍的乘子法。

## 1．等式约束问题的乘子法

将等式约束问题写成向量形式

$$\begin{aligned} &\min f(x), x \in R^n \\ s.t.\quad &C(x)=0 \end{aligned}$$

其中 $C(x)=(c_1(x),\cdots,c_l(x))^T$

$f(x)$ 和 $c_i(x)(i=1,\cdots,l)$ 是二次连续可微函数。

与如下问题等价：

$$\begin{aligned} &\min L(x,\lambda^*) \\ s.t.\quad &C(x)=0 \end{aligned}$$

增广目标函数（也称为增广 Lagrange 函数）为

$$M(x,\lambda,\sigma)=L(x,\lambda)+\frac{\sigma}{2}C(x)^T C(x)$$

其无约束优化问题为

$$\min M(x,\lambda,\sigma)$$

由最优性条件可得

$$\nabla_x M(x^*,\lambda^*,\sigma)$$

$$=\nabla_x L(x^*,\lambda^*)+\sigma\sum_{i=1}^{l}c_i(x^*)\nabla c_i(x^*)$$

$$=0$$

其中$x^*$是$M(x,\lambda^*,\sigma)$的稳定点。

我们将证明，当$\sigma$适当大时，$x^*$是$M(x,\lambda^*,\sigma)$的极小点，且$\lambda^*$是未知向量。

引理4.2.5 已知矩阵 $A_{n\times n}$ 和 $B_{n\times m}$，则对满足 $B^T x = 0$ 的任意 $x \neq 0$ 都有 $x^T Ax > 0$ 的充分必要条件是：

存在一个数 $\sigma^* > 0$，使得当 $\sigma \geq \sigma^*$，$x \in R^n, x \neq 0$ 时，

$$x^T(A+\sigma BB^T)x > 0$$

证 充分性

因为

$$B^T x = 0$$

故

$$x^T BB^T x = 0$$

因此有

$$x^T Ax > 0$$

必要性　先证明存在一个数 $\sigma^* > 0$ ，对任意的

$x \in R^n$ 有

$$x^T(A+\sigma^* BB^T)x > 0, (x \neq 0)$$

用反证法，假设不成立，

即对任意正整数 k 必存在向量 $x_k$ 且 $\|x_k\| = 1$ 使得

$$x_k^T(A+kBB^T)x_k \le 0$$