1. 设 $\boldsymbol{x}^{(0)} = (x_1^{(0)}, x_2^{(0)}, \cdots, x_n^{(0)})^T$ 是 $\boldsymbol{Ax} = \boldsymbol{b}$ 的一个解，其中 $\boldsymbol{A} = (\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_n)$ 是 $m \times n$ 矩阵，$\boldsymbol{A}$ 的秩为 m。证明 $\boldsymbol{x}^{(0)}$ 是基本解的充要条件为 $\boldsymbol{x}^{(0)}$ 的非零分量 $x_{i_1}^{(0)}, x_{i_2}^{(0)}, \cdots, x_{i_s}^{(0)}$ 对应的列 $\boldsymbol{p}_{i_1}, \boldsymbol{p}_{i_2}, \cdots, \boldsymbol{p}_{i_s}$ 线性无关。

解. 必要性：设 $\boldsymbol{x}^{(0)} = \begin{bmatrix} \boldsymbol{B}^{-1}\boldsymbol{b} \\ \boldsymbol{0} \end{bmatrix}$ 是基本解，$\boldsymbol{B} = \begin{bmatrix} \boldsymbol{p}_{B_1} & \boldsymbol{p}_{B_2} & \cdots & \boldsymbol{p}_{B_m} \end{bmatrix}$，则 $x^{(0)}$ 非零分量对应的列 $\left\{\boldsymbol{p}_{i_1}, \boldsymbol{p}_{i_2}, \cdots, \boldsymbol{p}_{i_s}\right\} \subset \left\{\boldsymbol{p}_{B_1}, \boldsymbol{p}_{B_2}, \cdots \boldsymbol{p}_{B_m}\right\}$。由于 $\boldsymbol{p}_{B_1}, \boldsymbol{p}_{B_2}, \cdots \boldsymbol{p}_{B_m}$ 线性无关，因此 $\boldsymbol{p}_{i_1}, \boldsymbol{p}_{i_2}, \cdots, \boldsymbol{p}_{i_s}$ 线性无关。

充分性：设 $\boldsymbol{x}^{(0)}$ 的非零分量对应的列 $\boldsymbol{p}_{i_1}, \boldsymbol{p}_{i_2}, \cdots, \boldsymbol{p}_{i_s}$ 线性无关，矩阵 $\boldsymbol{A}$ 的秩为 m，因此，$s \le m$，可以将 $\boldsymbol{p}_{i_1}, \boldsymbol{p}_{i_2}, \cdots, \boldsymbol{p}_{i_s}$ 扩充成一组基。于是 $\boldsymbol{x}^{(0)}$ 可记作：$\begin{bmatrix} \boldsymbol{x}_B^{(0)} \\ \boldsymbol{x}_N^{(0)} \end{bmatrix} = \begin{bmatrix} \boldsymbol{B}^{-1}\boldsymbol{b} \\ \boldsymbol{0} \end{bmatrix}$，$\boldsymbol{x}^{(0)}$ 是基本解。

2. 设 $S = \{\boldsymbol{x} | \boldsymbol{Ax} \ge \boldsymbol{b}\}$，其中 $\boldsymbol{A}$ 是 $m \times n$ 矩阵，$m > n$，$\boldsymbol{A}$ 的秩为 n。证明 $\boldsymbol{x}^{(0)}$ 是 S 的极点的充要条件是 $\boldsymbol{A}$ 和 $\boldsymbol{b}$ 可作如下分解：

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_1 \\ \boldsymbol{A}_2 \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \end{bmatrix}$$

其中 $\boldsymbol{A}_1$ 有 n 个行，且 $\boldsymbol{A}_1$ 的秩为 n，$\boldsymbol{b}_1$ 是 n 维列向量，使得 $\boldsymbol{A}_1\boldsymbol{x}^{(0)} = \boldsymbol{b}_1, \boldsymbol{A}_2\boldsymbol{x}^0 \ge \boldsymbol{b}_2$。

解. 必要性：设 $x^{(0)}$ 是 S 的极点，用反证法，设 A, b 在点 $x^{(0)}$ 分解如下：

$$A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}, b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, A_1 x^{(0)} = b_1, A_2 x^{(0)} > b_2$$

A_1 的秩 $R(A_1) < n$。$A_1 x = b_1$ 的同解线性方程组记作

$$\hat{A}_1 x = \hat{b_1}$$

$\hat{A}_1$ 是行满秩矩阵，$R(\hat{A}_1) = R(A_1) < n$。不妨假设 $\hat{A}_1$ 的前 $R(\hat{A}_1)$ 个列线性无关，记作 $\hat{A}_1 = \begin{bmatrix} B & N \end{bmatrix}$，其中 B 是可逆矩阵，相应地记

$$x = \begin{bmatrix} x_B \\ x_N \end{bmatrix}, x_B = B^{-1}\hat{b_1} - B^{-1} N x_N$$

$A_1 x = b_1$ 的解为

$$x = \begin{bmatrix} x_B \\ x_N \end{bmatrix} = \begin{bmatrix} B^{-1}\hat{b_1} - B^{-1} N x_N \\ x_N \end{bmatrix} \tag{1}$$

其中，x_N 是自由未知量，是 $n - R(A_1)$ 维向量，S 的极点

$$x^{(0)} = \begin{bmatrix} x_B^{(0)} \\ x_N^{(0)} \end{bmatrix} = \begin{bmatrix} B^{-1}\hat{b_1} - B^{-1} N x_N^{(0)} \\ x_N^{(0)} \end{bmatrix} \tag{2}$$

由于 $A_2x^{(0)} > b_2$，则存在 $x_N^{(0)}$ 的 δ 邻域 $N_s(x_N^{(0)})$，使得当 $x_N \in N_s(x_N^{(0)})$ 时，解 (1) 同时满足 $A_1x = b_1$ 和 $A_2x \geq b_2$，在过 $x_N^{(0)}$ 的直线上取不同点 $x_N^{(1)}, x_N^{(2)} \in N_s(x_N^{(0)})$，使 $\lambda x_N^{(1)} + (1-\lambda)x_N^{(2)} = x_N^{(0)}, \lambda \in (0,1)$，代入 (2) 式得到

$$\begin{aligned}\boldsymbol{x}^{(0)} &= \left[\begin{array}{c}\boldsymbol{B}^{-1}\hat{\boldsymbol{b}}_1 - \boldsymbol{B}^{-1}\boldsymbol{N}\left(\lambda\boldsymbol{x}_N^{(1)} + (1-\lambda)\boldsymbol{x}_N^{(2)}\right)\\ \lambda\boldsymbol{x}_N^{(1)} + (1-\lambda)\boldsymbol{x}_N^{(2)}\end{array}\right]\\ &= \lambda\left[\begin{array}{c}\boldsymbol{B}^{-1}\hat{\boldsymbol{b}}_1 - \boldsymbol{B}^{-1}\boldsymbol{N}\boldsymbol{x}_N^{(1)}\\ \boldsymbol{x}_N^{(1)}\end{array}\right] + (1-\lambda)\left[\begin{array}{c}\mathrm{B}^{-1}\hat{\boldsymbol{b}}_1 - \boldsymbol{B}^{-1}\boldsymbol{N}\boldsymbol{x}_N^{(2)}\\ \boldsymbol{x}_N^{(2)}\end{array}\right]\end{aligned}$$

这样，可将 $x^{(0)}$ 表示成集合 S 中两个不同点的凸组合，矛盾。

再证充分性。设在点 $x^{(0)}, A, b$ 可作如下分解（其中 A_1 是 n 阶方阵）

$$A = \begin{bmatrix}A_1\\ A_2\end{bmatrix}, b = \begin{bmatrix}b_1\\ b_2\end{bmatrix}, A_1x^{(0)} = b_1, A_2x^{(0)} \geq b_2, R(A_1) = n$$

又设存在 $x_{(1)}, x_{(2)} \in S$ 使得

$$x^{(0)} = \lambda x^{(1)} + (1-\lambda)x^{(2)}, \lambda \in (0,1) \tag{3}$$

用可逆矩阵 A_1 乘 (3) 式两端，得

$$A_1x^{(0)} = \lambda A_1x^{(1)} + (1-\lambda)A_1x^{(2)} \tag{4}$$

由于 $A_1x^{(0)} = b_1, A_1x^{(1)} \geq b_1, A_1x^{(2)} \geq b_1, \lambda > 0, (1-\lambda) > 0$，代入 (4) 式，得

$$b_1 = A_1x^{(0)} = \lambda A_1x^{(1)} + (1-\lambda)A_1x^{(2)} \geq \lambda b_1 + (1-\lambda)b_1 = b_1$$

因此

$$\lambda A_1x^{(1)} + (1-\lambda)A_1x^{(2)} = \lambda b_1 + (1-\lambda)b_1$$

整理得

$$\lambda(A_1x^{(1)} - b_1) + (1-\lambda)(A_1x^{(2)} - b_1) = 0$$

由于 $\lambda > 0, 1-\lambda > 0, A_1x^{(1)} - b_1 \geq 0, A_1x^{(2)} - b_1 \geq 0$，因此 $A_1x^{(0)} = A_1x^{(1)} = A_1x^{(2)} = b_1$，左乘 A_1^{-1}，得 $x_{(0)} = x_{(1)} = x_{(2)}$，因此 $x^{(0)}$ 是极点。

3. 设函数 $f : \mathbb{R}^n \to \mathbb{R}$，矩阵 $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ 对称，向量 $\boldsymbol{b} \in \mathbb{R}^n$。

1. 写出函数 f 是凸函数的定义，并列出至少两个判定函数 f 是凸函数的充要条件；
2. 设 $f(x_1, x_2) = (x_2 - x_1^2)^2, S = \{(x_1, x_2) | -1 < x_1 < 1, -1 < x_2 < 1\}$，判断函数 $f(x_1, x_2)$ 是否为 S 上的凸函数？说明理由；
3. 证明 $f(x) = \frac{1}{2}x^TAx + b^Tx$ 为严格凸函数的充要条件是其 $Hessian$ 阵 A 正定。

解. 1. 函数 f 的定义域 S 是非空凸集，对于 $\forall x_1, x_2 \in S, 0 \le \lambda \le 1$，都有

$$f(\lambda x_1 + (1-\lambda)x_2) \le \lambda f(x_1) + (1-\lambda) f(x_2)$$

充要条件:

(a) 函数 f 的定义域 S 是非空开凸集，函数 $f(X)$ 可微，$\forall x_1, x_2 \in S, f(x_2) \ge f(x_1) + \nabla f(x_1)^T(x_2 - x_1)$。

(b) 函数 f 的定义域 S 是非空开凸集，函数 $f(X)$ 二次可微 $\forall x \in S$, $f(x)$ 在 x 处的 Hessian 矩阵 $\nabla^2 f(x)$ 半正定。

2. $\forall \boldsymbol{x}, \boldsymbol{y} \in S, 0 \le \lambda \le 1$

$$\begin{aligned}
\boldsymbol{z} &= \lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y} \\
z_1 &= \lambda x_1 + (1-\lambda) y_1 \\
z_2 &= \lambda x_2 + (1-\lambda) y_2 \\
-1 = \lambda \cdot (-1) + (1-\lambda) \cdot (-1) \le z_1 &\le \lambda \cdot 1 + (1-\lambda) \cdot 1 = 1 \\
-1 = \lambda \cdot (-1) + (1-\lambda) \cdot (-1) \le z_2 &\le \lambda \cdot 1 + (1-\lambda) \cdot 1 = 1
\end{aligned}$$

$\boldsymbol{z} \in S$，所以集合 S 是一个凸集。

$$\begin{aligned}
\partial_{x_1x_1} f &= 4(3x_1^2 - x_2) \\
\partial_{x_1x_2} f &= -4x_1 \\
\partial_{x_2x_2} f &= 2 \\
\nabla^2 f &= \begin{pmatrix} 4(3x_1^2 - x_2) & -4x_1 \\ -4x_1 & 2 \end{pmatrix}
\end{aligned}$$

$\nabla^2 f$ 矩阵不是半正定矩阵，所以函数 $f(x_1, x_2)$ 不是 S 上的凸函数。

3. 必要性: 对 $\forall \boldsymbol{x} \in E^n, \exists \delta > 0$ 使当 $\lambda \in (0, \delta)$，有 $\bar{\boldsymbol{x}} + \lambda \boldsymbol{x} \in S$。$\Rightarrow f(\bar{\boldsymbol{x}} + \lambda \boldsymbol{x}) > f(\bar{\boldsymbol{x}}) + \lambda \nabla f(\bar{\boldsymbol{x}})^T \boldsymbol{x}$

$$f(\bar{x} + \lambda x) = f(\bar{x}) + \lambda \nabla f(\bar{x})^T x + \frac{1}{2}\lambda^2 x^T \nabla^2 f(\bar{x}) x + o(\|\lambda x\|^2)$$

可得

$$\frac{1}{2}\lambda^2 x^T \nabla^2 f(\bar{x}) x + o(\|\lambda x\|^2) > 0$$

两边除以 λ^2，令 $\lambda \to 0$，得 $x^T \nabla^2 f(\bar{x}) x > 0$，所以 $\nabla^2 f(x)$ 正定。

充分性：$\nabla^2 f(x)$ 在任意点 $x \in S$ 正定，对 $\forall x, \bar{x} \in S$，

$$f(x) = f(\bar{x}) + \nabla f(\bar{x})^T(x - \bar{x}) + \frac{1}{2}(x - \bar{x})\nabla^2 f(\xi)(x - \bar{x})$$

其中 $\xi = \lambda\bar{x} + (1 - \lambda)x, \lambda \in (0, 1)$。因为 S 是凸集，所以 $\xi \in S$，又 $\nabla^2 f(x)$ 正定，得 $\frac{1}{2}(x - \bar{x})^T\nabla^2 f(\xi)(x - \bar{x}) > 0 \Rightarrow f(x) > f(\bar{x}) + \nabla f(\bar{x})^T(x - \bar{x})$。所以函数 $f(x)$ 严格凸。