第十二章　可行方向法

Zoutendijk可行方向法

一．线性约束的情形

$$\begin{cases} \min \ f(x) \\ s.t. \quad Ax \geq b \qquad\qquad (1) \\ \qquad\quad Ex = e \end{cases}$$

其中$f(x)$可微，$A_{m\times n}$，$E_{l\times n}$，$x_{n\times 1}$，$b_{m\times 1}$，$e_{l\times 1}$

$S = \{x \mid Ax \geq b, Ex = e\}$ ----可行域

定义： 对 $\min_{x\in E^n} f(x)$，设 $\bar{x}\in E^n$ 是任给一点，$d\neq 0$，若存在 $\delta>0$，使得对任意的 $\lambda\in(0,\delta)$，有 $f(\bar{x}+\lambda d)<f(\bar{x})$，则称 d 为 $f(x)$ 在点 $\bar{x}$ 处的下降方向。

$$F_0=\left\{d \,\middle|\, \nabla f(\bar{x})^T d<0\right\}$$

称为点 $\bar{x}$ 处的下降方向集。

定义： 设集合$S \subset E^n, \bar{x} \in clS, d$为非零向量，若存在数$\delta > 0$，使得对任意$\lambda \in (0, \delta)$，都有

$$\bar{x} + \lambda d \in S$$

则称d为集合S在$\bar{x}$的可行方向。

$$D = \left\{ d \,\middle|\, d \neq 0, \bar{x} \in clS, \exists \delta > 0, \forall \lambda \in (0, \delta), 有 \bar{x} + \lambda d \in S \right\}$$

$\bar{x}$处的可行方向锥。

定理1： 设$\bar{x}$是问题(1)的可行解,在$\bar{x}$点处,有

$$A_1\bar{x}=b_1,\ A_2\bar{x}>b_2$$

其中$A=\begin{bmatrix}A_1\\A_2\end{bmatrix}, b=\begin{bmatrix}b_1\\b_2\end{bmatrix}$,则非零向量$d$是$\bar{x}$处的可行方向

的充要条件是$A_1d\geq 0, Ed=0.$

证明：“$\Rightarrow$”d是$\bar{x}$处的可行方向，则$\exists\delta>0$，对$\forall\lambda\in(0,\delta)$,

有 $\quad A(\bar{x}+\lambda d)\geq b, E(\bar{x}+\lambda d)=e$

$\Rightarrow A_1(\bar{x}+\lambda d)\geq b_1, E\bar{x}+\lambda Ed=e$

$\because \bar{x}$为可行解，且$A_1\bar{x}=b_1$，$E\bar{x}=e$,

$\therefore \lambda A_1 d\geq 0$，$\lambda Ed=0$,

$\because \lambda>0, \therefore A_1 d\geq 0$，$Ed=0$。

$$\boxed{\begin{cases}\min\ f(x)\\ s.t.\quad Ax\geq b\\ \qquad\quad Ex=e\end{cases}}$$

定理1： 设$\bar{x}$是问题(1)的可行解,在$\bar{x}$点处,有

$$A_1\bar{x}=b_1,\ A_2\bar{x}>b_2$$

其中$A=\begin{bmatrix}A_1\\A_2\end{bmatrix}, b=\begin{bmatrix}b_1\\b_2\end{bmatrix}$,则非零向量$d$是$\bar{x}$处的可行方向的充要条件是$A_1d\geq 0, Ed=0.$

$$\boxed{\begin{cases}\min\ f(x)\\ s.t.\quad Ax\geq b\\ \qquad\quad Ex=e\end{cases}}$$

证明： “$\Leftarrow$” 设$A_1d\geq 0$，$Ed=0$

$\because A_2\bar{x}>b_2, \therefore \exists\delta>0$,使得对$\forall\lambda\in(0,\delta)$，有

$A_2(\bar{x}+\lambda d)\geq b_2$

由假设$A_1\bar{x}=b_1, A_1d\geq 0 \therefore A_1(\bar{x}+\lambda d)\geq b_1 \Rightarrow A(\bar{x}+\lambda d)\geq b$

又由$Ed=0$，$E\bar{x}=e\Rightarrow E(\bar{x}+\lambda d)=e$

$\therefore \bar{x}+\lambda d$是(1)的可行解,因此$d$是可行方向.

$$
(2) \quad \begin{cases} \min \nabla f(x)^T d \\ s.t. \quad A_1 d \geq 0 \\ \qquad\quad Ed = 0 \\ \qquad\quad |d_j| \leq 1,\ j = 1,2,\cdots,n \end{cases}
$$

定理2： 在问题(1)中，设x是可行解，在点x处有

$A_1 x = b_1$，$A_2 x > b_2$，其中

$$A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix},$$

$$\boxed{\begin{cases} \min \nabla f(x)^T d \\ s.t. \quad A_1 d \geq 0 \\ \qquad\quad Ed = 0 \\ \qquad\quad |d_j| \leq 1, \ j = 1,2,\cdots,n \end{cases}}$$

则x是KKT点的充要条件是问题(2)的目标函数最优值$=0$。

证明：x是KKT点$\Leftrightarrow$存在向量$w \geq 0, v$，使得

$$\nabla f(x) - A_1^T w - E^T v = 0$$

令$v = p - q, \ p, q \geq 0$，把上式写成

$$(*)\quad \begin{cases} \left(-A_1^T, -E^T, E^T\right)\begin{pmatrix} w \\ p \\ q \end{pmatrix} = -\nabla f(x) \\ \begin{pmatrix} w \\ p \\ q \end{pmatrix} \geq 0 \end{cases}$$

$$\boxed{\begin{cases} \min \nabla f(x)^T d \\ s.t. \quad A_1 d \geq 0 \\ \qquad\quad Ed = 0 \\ \qquad\quad |d_j| \leq 1,\ j = 1,2,\cdots,n \end{cases}}$$

由$Farkas$引理，（*）有解$\Leftrightarrow \begin{pmatrix} -A_1 \\ -E \\ E \end{pmatrix} d \leq 0, \quad -\nabla f(x)^T d > 0$

无解，即$\nabla f(x)^T d < 0, A_1 d \geq 0, Ed = 0$无解，

等价于问题(2)目标函数最优值$= 0$。

步长的确定

设$x^{(k)}$是问题(1)的可行解，$d^{(k)}$是下降可行方向，从$x^{(k)}$出发，沿$d^{(k)}$作一维搜索，得到下一个点$x^{(k+1)}$，即

$$\min f(x^{(k)}+\lambda d^{(k)}) = f(x^{(k)}+\lambda_k d^{(k)})$$

$$x^{(k+1)} = x^{(k)}+\lambda_k d^{(k)}$$

可通过求解下列一维搜索问题来确定λ_k

$$(3) \qquad \begin{cases} \min f(x^{(k)}+\lambda d^{(k)}) \\ s.t. \quad A(x^{(k)}+\lambda d^{(k)}) \geq b \\ \qquad\quad E(x^{(k)}+\lambda d^{(k)}) = e \\ \qquad\quad \lambda \geq 0 \end{cases}$$

对问题(3)化简

(1) $\because d^{(k)}$是可行方向，$x^{(k)}$是可行解$\therefore Ed^{(k)}=0,$

$Ex^{(k)}=e$，因此问题(3)第2个约束是多余的。

(2)考虑第一个约束。在点$x^{(k)}$处，假设

$$A_1x^{(k)}=b_1,\quad A_2x^{(k)}>b_2$$

即有
$$\begin{pmatrix} A_1x^{(k)}+\lambda A_1d^{(k)} \\ A_2x^{(k)}+\lambda A_2d^{(k)} \end{pmatrix} \geq \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

$$\boxed{\begin{cases} \min\ f(x^{(k)}+\lambda d^{(k)}) \\ s.t.\quad A(x^{(k)}+\lambda d^{(k)})\geq b \\ \qquad\ \ E(x^{(k)}+\lambda d^{(k)})=e \\ \qquad\ \ \lambda\geq 0 \end{cases}}$$

$\because d^{(k)}$是可行方向$,\therefore A_1d^{(k)}\geq 0,\lambda\geq 0$

$\Rightarrow A_1x^{(k)}+\lambda A_1d^{(k)}\geq b_1$

$\therefore$第一个约束可简化为

$$A_2x^{(k)}+\lambda A_2d^{(k)}\geq b_2。$$

问题(3) $$\begin{cases} \min\ f(x^{(k)} + \lambda d^{(k)}) \\ s.t. \quad A(x^{(k)} + \lambda d^{(k)}) \geq b \\ \qquad\ \ E(x^{(k)} + \lambda d^{(k)}) = e \\ \qquad\ \ \lambda \geq 0 \end{cases}$$ 等价于

(4) $$\begin{cases} \min\ f(x^{(k)} + \lambda d^{(k)}) \\ s.t. \quad A_2 x^{(k)} + \lambda A_2 d^{(k)} \geq b_2 \\ \qquad\ \ \lambda \geq 0 \end{cases}$$

确定问题(4)的可行域

$A_2x^{(k)}+\lambda A_2d^{(k)}\ge b_2$移项，得

$$\lambda A_2d^{(k)}\ge b_2-A_2x^{(k)}$$

令$\hat{b}=b_2-A_2x^{(k)},\hat{d}=A_2d^{(k)}$,

则(4)的约束可写成$\begin{cases}\lambda\hat{d}\ge\hat{b}\\ \lambda\ge 0\end{cases}$

$$\begin{cases}\min\ f(x^{(k)}+\lambda d^{(k)})\\ s.t.\\ A_2x^{(k)}+\lambda A_2d^{(k)}\ge b_2\\ \qquad\lambda\ge 0\end{cases}$$

$$\therefore\lambda_{\max}=\begin{cases}\min\left\{\dfrac{\hat{b}_i}{\hat{d}_i}\,\middle|\,\hat{d}_i<0\right\} & \hat{d}\not\ge 0\\ \infty & \hat{d}\ge 0\end{cases}$$

问题(4) $\begin{cases} \min\ f(x^{(k)}+\lambda d^{(k)}) \\ s.t.\quad A_2x^{(k)}+\lambda A_2 d^{(k)} \ge b_2 \\ \qquad\quad \lambda \ge 0 \end{cases}$ 等价于

(5) $\begin{cases} \min\ f(x^{(k)}+\lambda d^{(k)}) \\ s.t.\quad 0 \le \lambda \le \lambda_{\max} \end{cases}$

其中$\lambda_{\max} = \begin{cases} \min\left\{ \dfrac{\hat{b}_i}{\hat{d}_i} \mid \hat{d}_i < 0 \right\} & \hat{d} \not\ge 0 \\ \infty & \hat{d} \ge 0 \end{cases}$

问题(1)初始可行解的确定

引入人工变量(向量)ξ和η，解辅助线性规划问题：

$$\begin{cases} \min\left(\sum_{i=1}^{m}\xi_i+\sum_{i=1}^{l}\eta_i\right) \\ s.t. \quad Ax+\xi\geq b \\ \qquad\quad Ex+\eta=e \\ \qquad\quad \xi,\eta\geq 0 \end{cases}$$

如果有最优解$(\bar{x},\bar{\xi},\bar{\eta})=(\bar{x},0,0)$，则$\bar{x}$为(1)的一个可行解。

步 骤

1.给定初始可行点$x^{(1)}$，置$k=1$。

2.在点$x^{(k)}$处把A和b分解成

$$\begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \text{和} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

使得$A_1 x^{(k)} = b_1$，$A_2 x^{(k)} > b_2$。计算$\nabla f(x^{(k)})$。

3.求解线性规划问题

$$\begin{cases} \min \nabla f(x^{(k)})^T d \\ s.t. \quad A_1 d \geq 0 \\ \qquad\quad Ed = 0 \\ \qquad\quad -1 \leq d_i \leq 1 \quad i = 1, 2, \cdots, n \end{cases}$$

得到最优解$d^{(k)}$。

4.如果$\nabla f(x^{(k)})^T d^{(k)} = 0$，则停止计算，$x^{(k)}$为$KKT$点；否则，转5。

5.计算$\hat{b} = b_2 - A_2 x^{(k)}$，$\hat{d} = A_2 d^{(k)}$。

6.若$\hat{d} \geq 0$,则从$x^{(k)}$出发，沿$d^{(k)}$进行一维搜索

$$\min f(x^{(k)} + \lambda d^{(k)}) = f(x^{(k)} + \lambda_k d^{(k)})$$

令$x^{(k+1)} = x^{(k)} + \lambda_k d^{(k)}$，转8;否则，转7。

7.计算$\lambda_{\max} = \min\left\{ \dfrac{\hat{b}_i}{\hat{d}_i} \mid \hat{d}_i < 0 \right\}$,求解

$$\begin{aligned} &\min f(x^{(k)} + \lambda d^{(k)}) \\ &s.t. \quad 0 \le \lambda \le \lambda_{\max} \end{aligned}$$

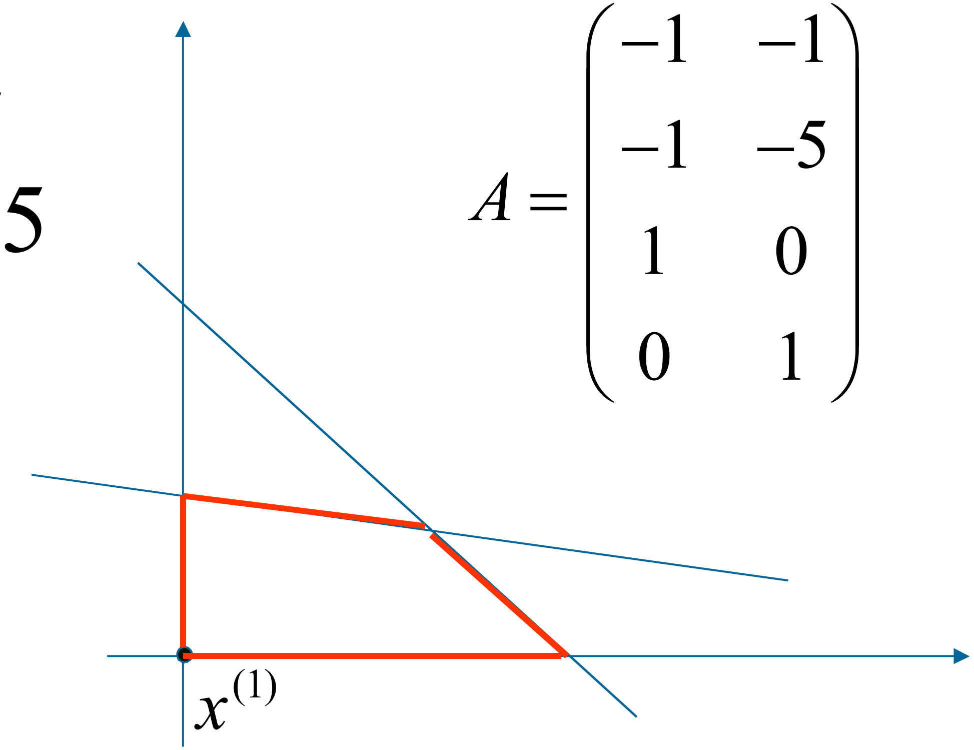
得最优解λ_k，令$x^{(k+1)} = x^{(k)} + \lambda_k d^{(k)}$，转8。

8.确定$x^{(k+1)}$的起作用约束，修改A_1, A_2及b_1, b_2,置$k := k+1$，返回2。

用Zoutendijk可行方向法求解

$$\begin{cases} \min\ f(x) = 2x_1^2 + 2x_2^2 - 2x_1x_2 - 4x_1 - 6x_2 \\ s.t. \quad -x_1 - x_2 \geq -2 \\ \qquad -x_1 - 5x_2 \geq -5 \\ \qquad x_1 \geq 0 \\ \qquad x_2 \geq 0 \end{cases}$$

$$A = \begin{pmatrix} -1 & -1 \\ -1 & -5 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$



取$x^{(1)} = (0, 0)^T$,

$$\nabla f(x) = (4x_1 - 2x_2 - 4, 4x_2 - 2x_1 - 6)^T$$

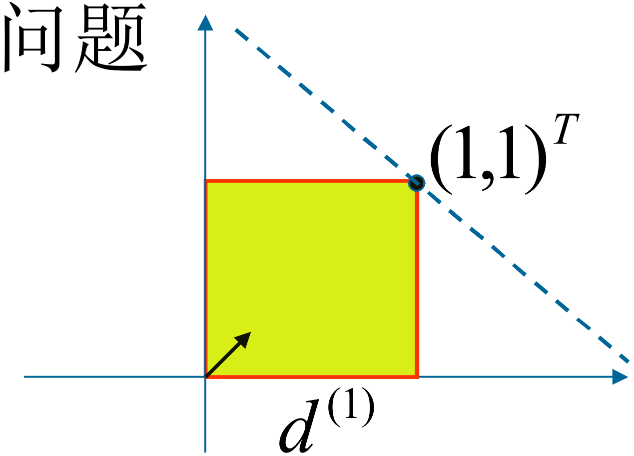
第一次迭代

$$A_1=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, A_2=\begin{pmatrix}-1 & -1\\ -1 & -5\end{pmatrix}, b_1=\begin{pmatrix}0\\ 0\end{pmatrix}, b_2=\begin{pmatrix}-2\\ -5\end{pmatrix},$$

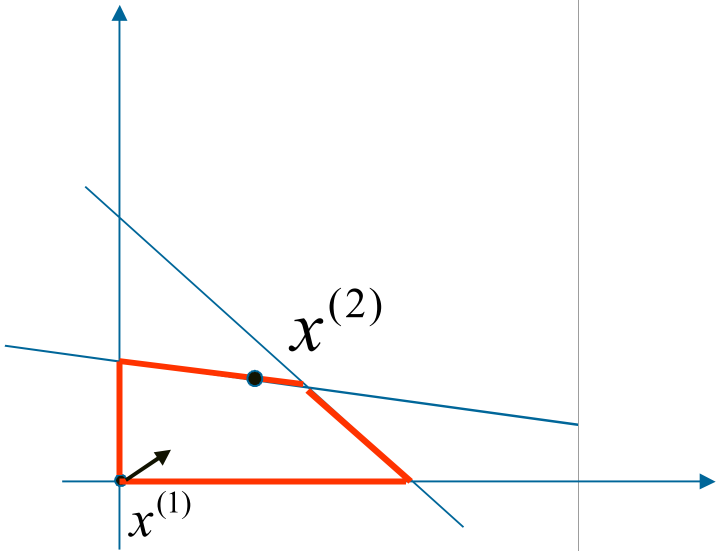
$$\nabla f(x^{(1)})=(-4,-6)^T$$

求搜索方向$d^{(1)}$，即解线性规划问题

$$\begin{cases}\min -4d_1-6d_2\\ s.t.\quad d_1\geq 0\\ \qquad\quad d_2\geq 0\\ \qquad\quad -1\leq d_i\leq 1\quad i=1,2\end{cases}$$



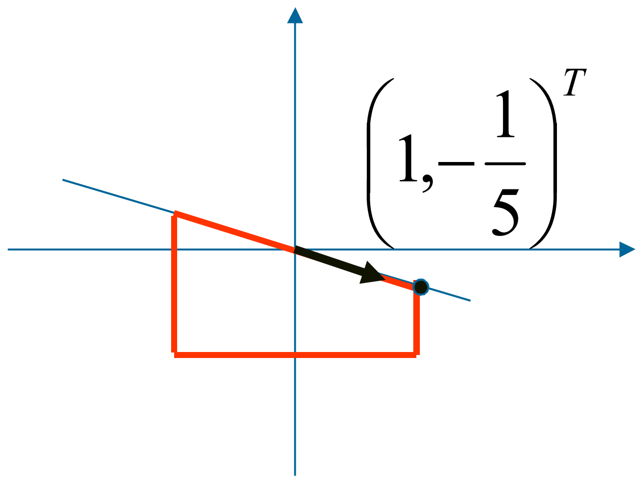
用单纯形法求解，得最优解$d^{(1)}=(1,1)^T$


$x^{(2)}$
$x^{(1)}$

第二次迭代

$$A_1=(-1,-5),A_2=\begin{pmatrix}-1&-1\\1&0\\0&1\end{pmatrix},b_1=(-5),b_2=\begin{pmatrix}-2\\0\\0\end{pmatrix},$$

$$\nabla f(x^{(2)})=\left(-\frac{7}{3},-\frac{13}{3}\right)^T$$



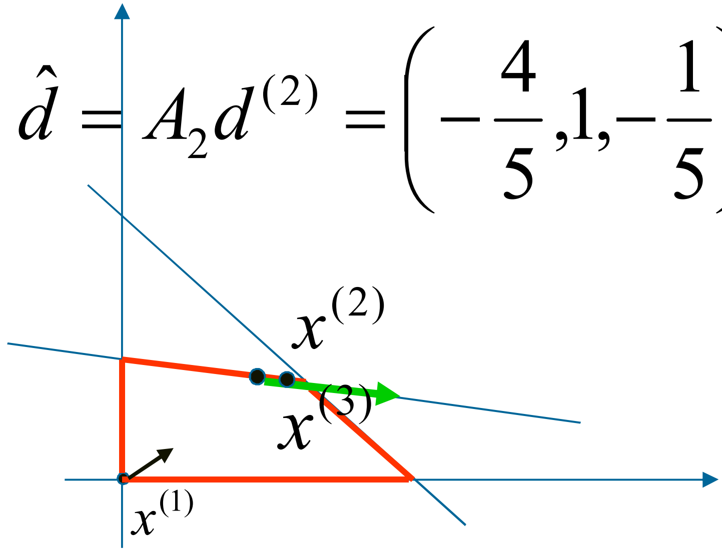
求搜索方向$d^{(2)}$，即解线性规划问题

$$\begin{cases}\min -\frac{7}{3}d_1-\frac{13}{3}d_2\\ s.t.\quad -d_1-5d_2\ge 0\\ \qquad -1\le d_i\le 1\quad i=1,2\end{cases}$$

$\Rightarrow$ 得最优解$d^{(2)}=\left(1,-\frac{1}{5}\right)^T$

$$\nabla f(x^{(2)})^T d^{(2)} = -\frac{22}{15} \neq 0,$$

令 $\hat{b} = b_2 - A_2 x^{(2)} = \left(-\frac{1}{3}, -\frac{5}{6}, -\frac{5}{6}\right)^T$, $\quad \hat{d} = A_2 d^{(2)} = \left(-\frac{4}{5}, 1, -\frac{1}{5}\right)$



$$\lambda_{\max} = \min\left\{ \frac{-1/3}{-4/5}, \frac{-5/6}{-1/5} \right\} = \frac{5}{12}$$

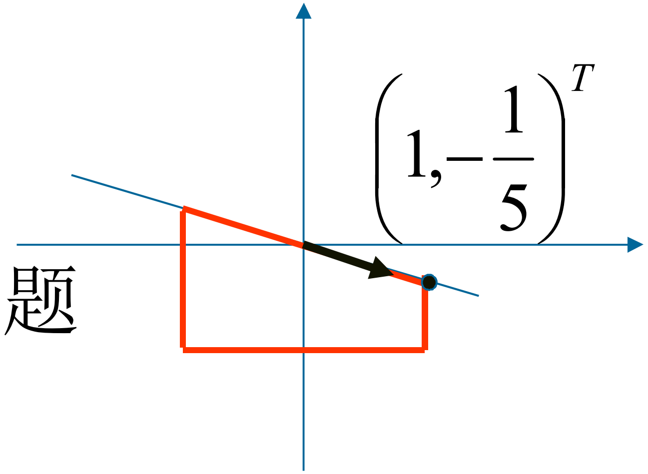
求解一维搜索问题 $\begin{cases} \min f(x^{(1)} + \lambda d^{(1)}) = \dfrac{62}{25}\lambda^2 - \dfrac{22}{15}\lambda - \dfrac{125}{18} \\ s.t. \quad 0 \leq \lambda \leq \dfrac{5}{12} \end{cases}$

得 $\lambda_2 = \frac{55}{186} \Rightarrow x^{(3)} = x^{(2)} + \lambda_2 d^{(2)} = \left(\frac{35}{31}, \frac{24}{31}\right)^T.$

第三次 迭代

$$A_1=(-1,-5),A_2=\begin{pmatrix}-1 & -1\\ 1 & 0\\ 0 & 1\end{pmatrix},b_1=(-5),b_2=\begin{pmatrix}-2\\ 0\\ 0\end{pmatrix},$$

$$\nabla f(x^{(3)})=\left(-\frac{32}{31},-\frac{160}{31}\right)^T$$



求搜索方向$d^{(3)}$，即解线性规划问题

$$\begin{cases}\min -\dfrac{32}{31}d_1-\dfrac{160}{31}d_2\\ s.t.\quad -d_1-5d_2\geq 0\\ \qquad -1\leq d_i\leq 1\quad i=1,2\end{cases}\Rightarrow \text{得最优解} d^{(3)}=\left(1,-\frac{1}{5}\right)^T$$

$\nabla f(x^{(3)})^T d^{(3)} = 0$,迭代停止。,

得KKT点$x^{(3)} = \left(\frac{35}{31}, \frac{24}{31} \right)^T = x^*$.

由于$f(x)$是凸函数，原问题为凸规划，

所以x^*为最优解。

二.非线性约束的情形

$$(A)\qquad \begin{cases} \min\ f(x) \\ s.t.\quad g_i(x) \geq 0,\ i = 1,2,\cdots,m \end{cases}$$

其中$f(x), g_i(x)$均为可微函数。

问题:

如何求可行下降方向?

定理： 设x是问题(A)的可行解，$I=\{i \mid g_i(x)=0\}$是在x处起作用约束下标集，又设$f(x)$，$g_i(x)(i\in I)$在x处可微，$g_i(x)(i\notin I)$在x处连续，如果 $\nabla f(x)^T d<0$，$\nabla g_i(x)^T d>0(i\in I)$，则$d$是可行下降方向。

证明：$\because \nabla f(x)^T d<0, \therefore d$为下降方向。

$\because \nabla g_i(x)^T d>0(i\in I)$, 即$-\nabla g_i(x)^T d<0(i\in I)$

$\therefore \exists \delta_1>0$, 使得对$\forall \lambda\in(0,\delta_1)$, 当$i\in I$时，有

$-g_i(x+\lambda d)<-g_i(x)=0 \Rightarrow g_i(x+\lambda d)>0$。

当$i\notin I$时, 有$g_i(x)>0, \because g_i(x)$连续,

$\therefore \exists \delta_2>0$, 对$\forall \lambda\in(0,\delta_2)$, 有$g_i(x+\lambda d)\geq 0$。

取$\delta=\min(\delta_1,\delta_2)$，则当$\lambda\in(0,\delta)$时，有

$g_i(x+\lambda d)\geq 0 \Rightarrow d$为可行方向。

求可行点x处的下降可行方向

$$
(B)\quad \begin{cases} \min z \\ s.t. \quad \nabla f(x)^T d - z \le 0 \\ \qquad\quad \nabla g_i(x)^T d + z \ge 0, \quad i \in I \\ \qquad\quad -1 \le d_i \le 1, \quad i = 1,2,\cdots,n \end{cases}
$$

设问题的最优解为$(\bar{z}, \bar{d})$，若$\bar{z} < 0$，则$\bar{d}$为在x处的可行下降方向。

定理： 设x是问题(A)的可行解，$I=\{i \mid g_i(x)=0\}$，则x是$Fritz\ John$点的充要条件是问题(B)的目标函数最优值$=0$。

证明：问题(B)的目标函数最优值$=0$的充要条件是不等式组

$$(C)\quad \begin{cases} \nabla f(x)^T d<0 \\ \nabla g_i(x)^T d>0 \qquad i\in I \end{cases}$$

无解。

$$\begin{cases} \min\ z \\ s.t.\quad \nabla f(x)^T d-z\le 0 \\ \nabla g_i(x)^T d+z\ge 0, i\in I \\ \quad -1\le d_i\le 1, \\ \quad i=1,2,\cdots,n \end{cases}$$

设问题(B)的目标函数最优值为$\bar{z}$.

"$\Rightarrow$"设$\bar{z}=0$，而系统(C)有解$\bar{d}$，即

$$\nabla f(x)^T \bar{d} < 0 \text{且} \nabla g_i(x)^T \bar{d} > 0 \ \forall i \in I$$

$$\begin{cases} \nabla f(x)^T d < 0 \\ \nabla g_i(x)^T d > 0 \ i \in I \end{cases}$$

显然对任意$\lambda > 0, \lambda\bar{d}$也是系统$(C)$的解

即$\nabla f(x)^T \lambda\bar{d} < 0$且$\nabla g_i(x)^T \lambda\bar{d} > 0 \ \forall i \in I$,

令 $-1 \le \lambda\bar{d}_i \le 1, i = 1,2,\cdots,n$

设$\hat{z} = \max\left\{\nabla f(x)^T \lambda\bar{d}, -\nabla g_i(x)^T \lambda\bar{d}, \forall i \in I\right\}$

显然$\hat{z} < 0$, 且有

$$\nabla f(x)^T \lambda\bar{d} \le \hat{z}, \quad -\nabla g_i(x)^T \lambda\bar{d} \le \hat{z}, \forall i \in I$$

$$\Rightarrow \begin{cases} \nabla f(x)^T \lambda\bar{d} - \hat{z} \le 0 \\ \nabla g_i(x)^T \lambda\bar{d} + \hat{z} \ge 0, \forall i \in I \\ -1 \le \lambda\bar{d}_i \le 1, i = 1,2,\cdots,n \end{cases}$$

$$\begin{cases} \min z \\ s.t. \quad \nabla f(x)^T d - z \le 0 \\ \nabla g_i(x)^T d + z \ge 0, i \in I \\ \quad -1 \le d_i \le 1, \\ \quad i = 1,2,\cdots,n \end{cases}$$

$\therefore (\hat{z}, \lambda\bar{d})$是问题$(B)$的一个可行解，

$\therefore \bar{z} \le \hat{z} < 0$，与$\bar{z}=0$矛盾。

"$\Leftarrow$"设系统(C)无解，即对任意的d，有 $\begin{cases}\nabla f(x)^T d < 0 \\ \nabla g_i(x)^T d > 0\ i \in I\end{cases}$

$$\nabla f(x)^T d \geq 0 \text{ 或 } \nabla g_i(x)^T d \leq 0 (i \in I)$$

不妨设$\nabla g_i(x)^T d \leq 0 (i \in I)$ $\Rightarrow \bar{z} \geq -\nabla g_i(x)^T d \geq 0$

又零向量是问题(B)的一个可行解，$\therefore \bar{z} \leq 0 \Rightarrow \bar{z} = 0$

$$\begin{cases}\min\ z \\ s.t.\quad \nabla f(x)^T d - z \leq 0 \\ \nabla g_i(x)^T d + z \geq 0, i \in I \\ \quad -1 \leq d_i \leq 1, \\ \quad i = 1,2,\cdots,n\end{cases}$$

$$\therefore \bar{z} = 0 \Leftrightarrow \begin{cases}\nabla f(x)^T d < 0 \\ \nabla g_i(x)^T d > 0 (i \in I)\end{cases} \text{无解}$$

$$\Leftrightarrow \begin{cases}\nabla f(x)^T d < 0 \\ -\nabla g_i(x)^T d < 0 (i \in I)\end{cases} \text{无解}$$

$\Leftrightarrow$ 存在不全为0的数w_0，$w_i \geq 0 (i \in I)$使得

$$w_0 \nabla f(x) - \sum_{i \in I} w_i \nabla g_i(x) = 0$$

即x是$Fritz\ John$点。

为确定步长λ_k，仍需要求解一维搜索问题

$$\begin{aligned} &\min\ f(x^{(k)}+\lambda d^{(k)}) \\ &s.t.\quad 0\le\lambda\le\lambda_{\max} \end{aligned}$$

其中

$$\lambda_{\max}=\sup\left\{\lambda \,\middle|\, g_i(x^{(k)}+\lambda d^{(k)})\ge 0, i=1,2,\cdots,m\right\}.$$

步 骤

1.给定初始可行点$x^{(1)}$，允许误差$\varepsilon > 0$，置$k = 1$。

2.令$I = \{i \mid g_i(x^{(k)}) = 0\}$。求解线性规划问题

$$\begin{cases} \min z \\ s.t. \quad \nabla f(x^{(k)})^T d - z \le 0 \\ \qquad \nabla g_i(x^{(k)})^T d + z \ge 0 \quad i \in I \\ \qquad -1 \le d_i \le 1 \quad i = 1,2,\cdots,n \end{cases}$$

得到最优解$(z_k, d^{(k)})$。若$|z_k| \le \varepsilon$，则停止计算，$x^{(k)}$是*Fritz John*点；否则，转3。

3.求解一维搜索问题

$$\min f(x^{(k)}+\lambda d^{(k)})$$
$$s.t. \quad 0 \le \lambda \le \lambda_{\max}$$

其中

$$\lambda_{\max}=\sup\left\{\lambda \middle| g_i(x^{(k)}+\lambda d^{(k)}) \ge 0, i=1,2,\cdots,m\right\}$$

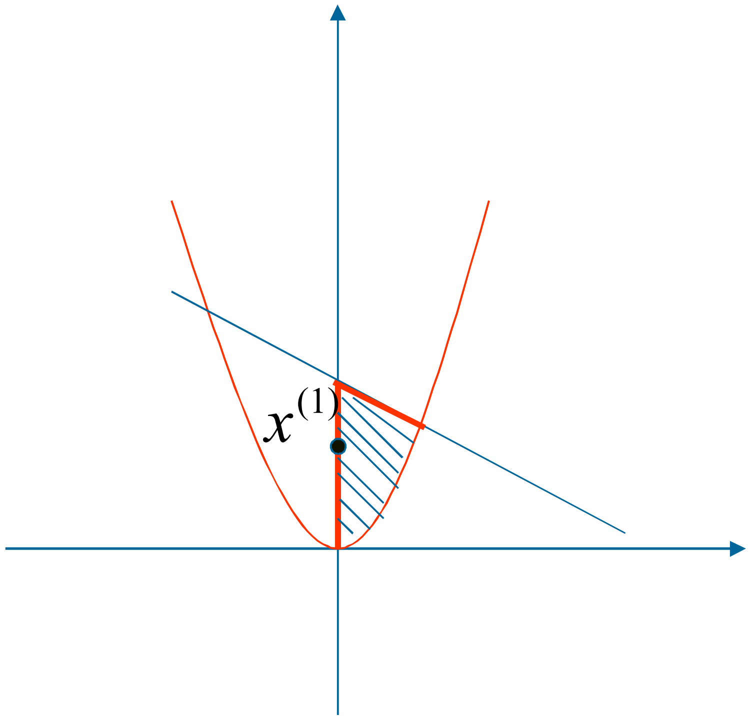
得最优解λ_k。

4.令$x^{(k+1)}=x^{(k)}+\lambda_k d^{(k)}$，置$k:=k+1$，返回2。

用Zoutendijk可行方向法求解

$$\min\ f(x) = 2x_1^2 + 2x_2^2 - 2x_1x_2 - 4x_1 - 6x_2$$

$$\begin{aligned} s.t.\quad & -x_1 - 5x_2 \geq -5 \\ & -2x_1^2 + x_2 \geq 0 \\ & x_1 \geq 0 \\ & x_2 \geq 0 \end{aligned}$$



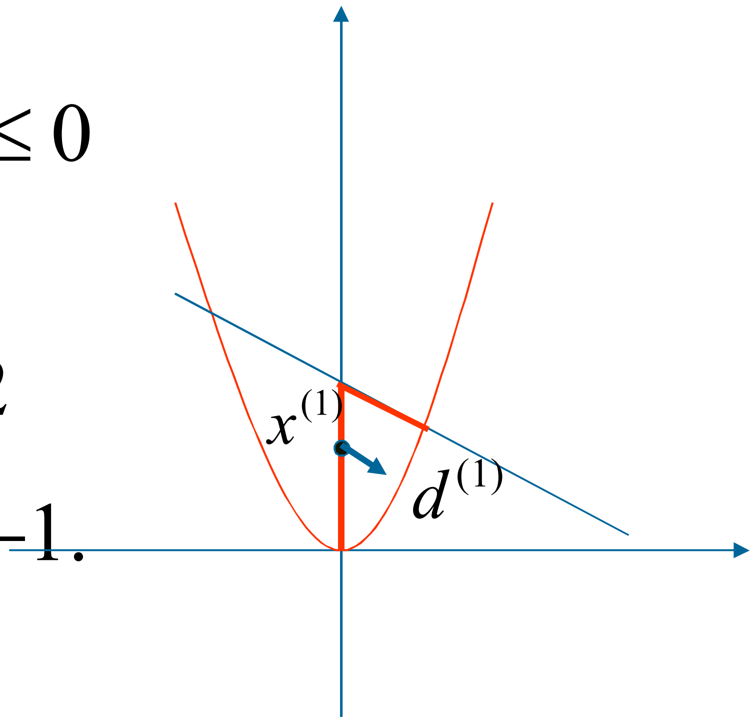
$$x^{(1)} = (0, 0.75)^T$$

$$\nabla f(x) = (4x_1 - 2x_2 - 4,\ 4x_2 - 2x_1 - 6)^T$$

第一次迭代

$$\nabla f(x^{(1)}) = (-5.5, -3)^T, I = \{3\}, \nabla g_3(x^{(1)}) = (1,0)^T$$

求解 $\begin{cases} \min z \\ s.t. \quad -5.5d_1 - 3d_2 - z \le 0 \\ \qquad d_1 + z \ge 0 \\ \qquad -1 \le d_i \le 1, i = 1,2 \end{cases}$



得最优解$d^{(1)} = (1,-1)^T, z_{\min} = -1.$

进行一维搜索

一维搜索：$x^{(1)}+\lambda d^{(1)}=(\lambda,0.75-\lambda)^T$

$$f(x^{(1)}+\lambda d^{(1)})=6\lambda^2-2.5\lambda-3.375$$

$$\lambda_{\max}=\sup\left\{\lambda\left|\begin{array}{l}-\lambda-5(0.75-\lambda)\geq-5\\-2\lambda^2+0.75-\lambda\geq 0\\\lambda\geq 0\\0.75-\lambda\geq 0\end{array}\right.\right\}$$

$$\Rightarrow\lambda_{\max}=\frac{-1+\sqrt{7}}{4}\approx 0.4114$$

求解 $\min\ 6\lambda^2-2.5\lambda-3.375$

$$s.t.\ 0\leq\lambda\leq 0.4114$$

得$\lambda_1=0.2083\Rightarrow x^{(2)}=x^{(1)}+\lambda_1 d^{(1)}=(0.2083,\,0.5417)^T$.

继续进行迭代，经过4次迭代后，得

$$x^{(5)}=(0.6302, 0.8740)^T,$$

$$\nabla f(x^{(5)})=(-3.2272, -3.7644)^T, I=\varnothing.$$

求解 $\begin{cases} \min z \\ s.t. \quad -3.2272d_1-3.7644d_2-z\le 0 \\ \qquad\quad -1\le d_i\le 1, \quad i=1,2 \end{cases}$

得$z=0,\ f(x^{(5)})=-6.5403$

本题最优解为：

$$x^*=(0.658872,\ 0.868224)^T$$

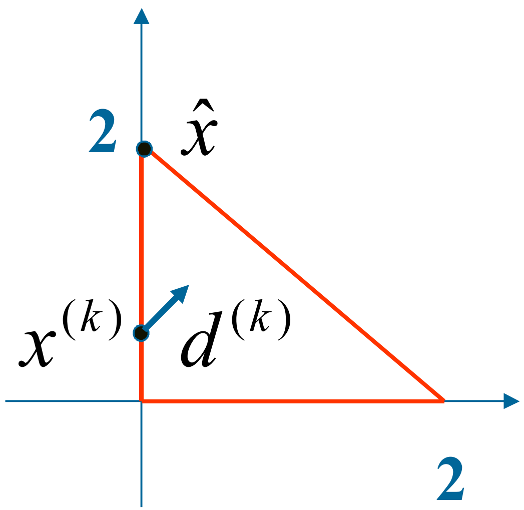
$$f(x^*)=-6.613086$$

Zoutendijk算法的收敛问题

Zoutendijk算法映射$A=MD$，其中D是确定方向的映射，M是一维搜索。

结论：D和M不一定是闭映射。

$$\min -2x_1 - x_2$$
$$s.t. \quad -x_1 - x_2 + 2 \ge 0$$
$$x_1, x_2 \ge 0$$



考虑序列$\{x^{(k)}\}$,其中$x^{(k)} = \left(0, 2-\dfrac{1}{k}\right)^T$

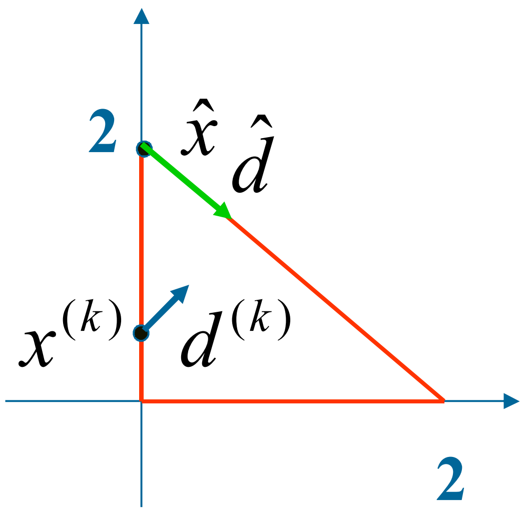
求$x^{(k)}$的下降可行方向

$$\begin{cases} \min -2d_1 - d_2 \\ s.t. \quad 0 \le d_1 \le 1 \\ \qquad -1 \le d_2 \le 1 \end{cases} \Rightarrow d^{(k)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$x^{(k)} = \left(0, 2-\frac{1}{k}\right)^T \to (0,2)^T = \hat{x}$$

求$\hat{x}$的下降可行方向，即求$D(\hat{x})$.

$$\begin{cases} \min -2d_1 - d_2 \\ s.t. \quad -d_1 - d_2 \geq 0 \\ \qquad 0 \leq d_1 \leq 1 \\ \qquad -1 \leq d_2 \leq 1 \end{cases} \Rightarrow \hat{d} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$



$$D(\hat{x}) = \left\{ \left(\hat{x}, \hat{d} \right) \right\}$$

当$x^{(k)} \to \hat{x}$时，$\left(x^{(k)}, d^{(k)}\right) \to (\hat{x}, d)$，其中$d = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

显然$(\hat{x}, d) \notin D(\hat{x})$,因此$D$在$\hat{x}$处不是闭的.

Topkis-Veinott修正的可行方向法

改进措施: x处的下降可行方向d由如下线性规划问题来确定.

$$\begin{cases} \min z \\ s.t. \quad \nabla f(x)^T d - z \le 0 \\ \qquad \nabla g_i(x)^T d + z \ge -g_i(x) \quad i = 1,2,\cdots,m \\ \qquad -1 \le d_i \le 1 \quad i = 1,2,\cdots,n \end{cases}$$

用Topkis-Veinott修正可行方向法求解

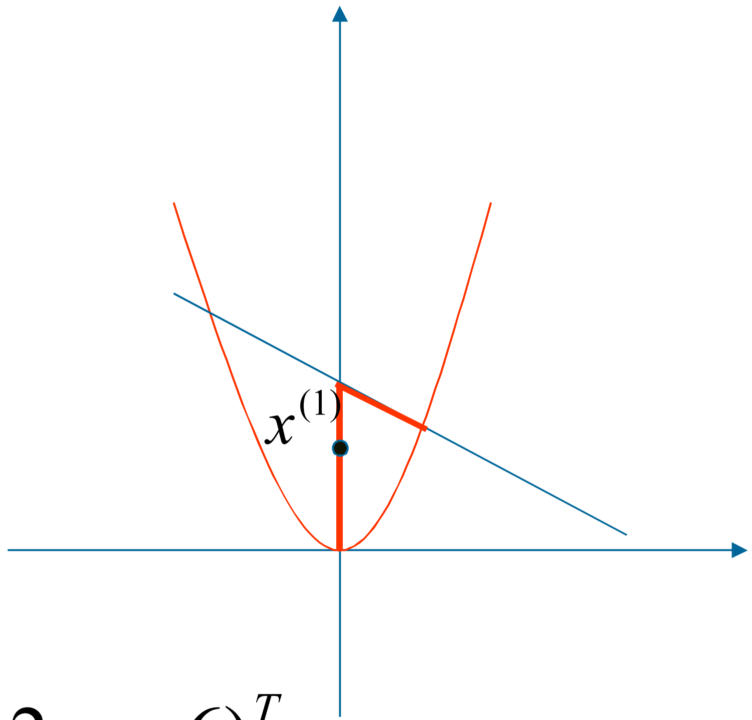
$$\min\ f(x) = 2x_1^2 + 2x_2^2 - 2x_1x_2 - 4x_1 - 6x_2$$

$$s.t.\quad -x_1 - 5x_2 \geq -5$$

$$-2x_1^2 + x_2 \geq 0$$

$$x_1 \geq 0$$

$$x_2 \geq 0$$



$$x^{(1)} = (0, 0.75)^T$$

$$\nabla f(x) = (4x_1 - 2x_2 - 4,\ 4x_2 - 2x_1 - 6)^T$$

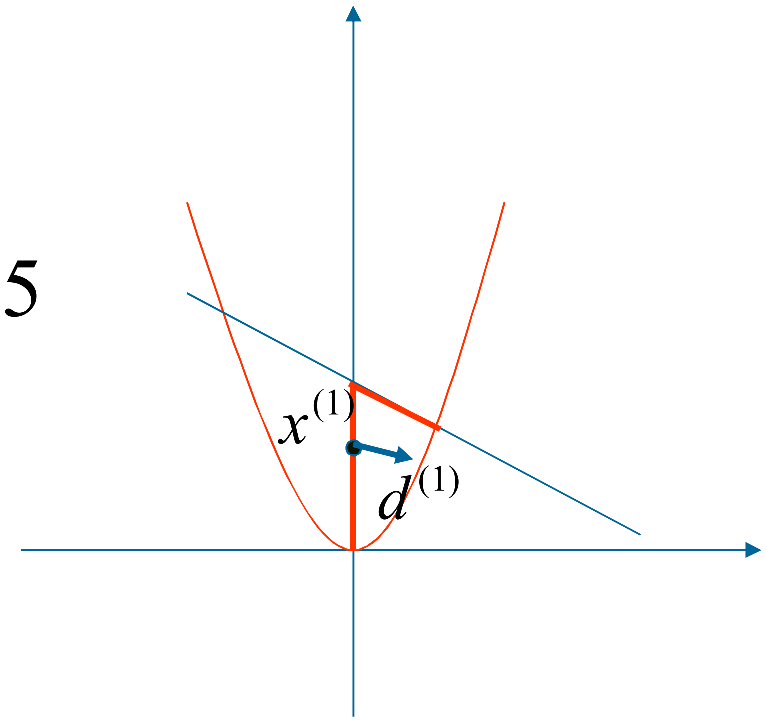
$$\nabla g_1(x) = (-1, -5)^T, \nabla g_2(x) = (-4x_1, 1)^T$$

$$\nabla g_3(x) = (1, 0)^T, \nabla g_4(x) = (0, 1)^T$$

第一次迭代

$\nabla f(x^{(1)})=(-5.5,-3)^T$，求解

$$\begin{cases}\min z \\ s.t.\quad -5.5d_1-3d_2-z\le 0 \\ \qquad -d_1-5d_2+z\ge -1.25 \\ \qquad d_2+z\ge -0.75 \\ \qquad d_1+z\ge 0 \\ \qquad d_2+z\ge -0.75 \\ \qquad -1\le d_i\le 1, i=1,2\end{cases}$$



得最优解$d^{(1)}=(0.7143,-0.03571)^T, z_{\min}=-0.7143.$

进行一维搜索

一维搜索：$x^{(1)}+\lambda d^{(1)}=(0.7143\lambda,0.75-0.03571\lambda)^T$

$$f(x^{(1)}+\lambda d^{(1)})=1.074\lambda^2-3.82152\lambda-3.375$$

$$\lambda_{\max}=\sup\left\{\lambda\left|\begin{array}{l}-\lambda-5(0.75-\lambda)\geq-5\\ -2\lambda^2+0.75-\lambda\geq 0\\ \lambda\geq 0\\ 0.75-\lambda\geq 0\end{array}\right.\right\}$$

$$\Rightarrow\lambda_{\max}=\frac{-1+\sqrt{7}}{4}\approx 0.4114$$

求解 $\min 1.074\lambda^2-3.82152\lambda-3.375$

$$s.t.\ 0\leq\lambda\leq 0.4114$$

得$\lambda_1=0.4114\Rightarrow x^{(2)}=x^{(1)}+\lambda_1 d^{(1)}=(0.29,0.7353)^T$.

继续进行迭代，经过5次迭代后，得

$$x^{(6)} = (0.6548, 0.8575)^T,$$

$$f(x^{(6)}) = -6.5590$$

本题最优解为：

$$x^* = (0.658872,\ 0.868224)^T$$

$$f(x^*) = -6.613086$$

Rosen梯度投影法

$$\begin{cases} \min \; f(x) \\ s.t. \quad Ax \geq b \\ \qquad\quad Ex = e \end{cases} \qquad (1)$$

其中$f(x)$可微，$A_{m\times n}$，$E_{l\times n}$，$x_{n\times 1}$，$b_{m\times 1}$，$e_{l\times 1}$

$S = \{x \mid Ax \geq b, Ex = e\}$ ---- 可行域

基本思想: 当迭代点在可行域内部时,沿迭代点的负梯度方向搜索;当迭代点达到约束区域的边界上时,用起作用约束系数矩阵的行向量生成一个子空间，沿负梯度方向在该子空间的正交补中的投影方向搜索.

设B是一个$m\times n$阶矩阵，且$r(B)=m.$

设B的m个行向量为$\alpha_1^T,\alpha_2^T,\cdots,\alpha_m^T$，由$B$的$m$个行向量生成的子空间记为$V_B$，即

$$V_B=span\left(\alpha_1^T,\alpha_2^T,\cdots,\alpha_m^T\right)$$

V_B的正交子空间记为$V_B^\perp$,则

$$V_B^\perp=\left\{x\mid Bx=0\right\}.$$

结论：$E^n=V_B\oplus V_B^\perp$

定义： 设U是E^n的子空间，对$\forall x \in E^n$，可唯一的分解为

$$x = y + z$$

其中$y \in U, z \in U^\perp$。

由E^n到U的这种映射记为Q，即$Qx = y$，称Q为由E^n到U的正交投影矩阵。由E^n到$U^\perp$的这种映射记为P，即$Px = z$，称P为由E^n到$U^\perp$的正交投影矩阵。

定理： 设$\alpha_1,\alpha_2,\cdots,\alpha_r$是子空间$U$的一组基，记

$$M=\begin{bmatrix}\alpha_1^{\ T}\\ \alpha_2^{\ T}\\ \vdots\\ \alpha_r^{\ T}\end{bmatrix}_{r\times n}$$

则

(1)由E^n到U的正交投影矩阵Q可表示为

$$Q=M^T(MM^T)^{-1}M。$$

(2)由E^n到$U^{\perp}$的正交投影矩阵P可表示为

$$P=I-M^T(MM^T)^{-1}M。$$

证明：(1) $\forall x \in E^n$可唯一分解为

$$x = y + z \quad y \in U, z \in U^{\perp}$$

则$y = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_r\alpha_r$，且$\alpha_i^T z = 0$

记$K = (k_1, k_2, \cdots, k_r)^T$，则$y = M^T K$

$$M = \begin{bmatrix} \alpha_1^T \\ \alpha_2^T \\ \vdots \\ \alpha_r^T \end{bmatrix}_{r\times n}$$

$\Rightarrow x = M^T K + z$

两边左乘M，得$Mx = MM^T K + Mz$

$\because \alpha_i^T z = 0(i = 1,2,\cdots,r) \quad \therefore Mz = 0$

$\therefore Mx = MM^T K$

$\because MM^T$可逆$\therefore K = (MM^T)^{-1} Mx$

$\because y = Qx \quad \therefore Qx = M^T (MM^T)^{-1} Mx \Rightarrow Q = M^T (MM^T)^{-1} M$

(2) 因为P是E^n到$U^\perp$的正交投影矩阵

所以对于$x = y + z \quad y \in U, z \in U^\perp$

有$Px = z \Rightarrow Ix = Qx + Px$

$\Rightarrow I = P + Q$

$$\boxed{\begin{aligned} Q &= M^T (MM^T)^{-1} M \\ P &= I - M^T (MM^T)^{-1} M \end{aligned}}$$

推论1: 若Q是正交投影矩阵，则$Q^T = Q$，$QQ = Q$.

推论2: 若Q是正交投影矩阵，则Q是半正定的.

定理1 设x是问题(1)的可行解，在点x处，有$A_1x=b_1$，$A_2x>b_2$，其中

$$A=\begin{bmatrix}A_1\\A_2\end{bmatrix}\quad b=\begin{bmatrix}b_1\\b_2\end{bmatrix}$$

$$\boxed{\begin{cases}\min\ f(x)\\ s.t.\quad Ax\ge b\\ \qquad\quad Ex=e\end{cases}\qquad (1)}$$

又设$M=\begin{bmatrix}A_1\\E\end{bmatrix}$为满秩矩阵，令

$$P=I-M^T\left(MM^T\right)^{-1}M$$

若$P\nabla f(x)\neq 0$，令$d=-P\nabla f(x)$，则d为下降可行方向。

证明：$\because P$为正交投影矩阵，且$P\nabla f(x)\neq 0$

$$\therefore \nabla f(x)^T d=-\nabla f(x)^T P\nabla f(x)$$

$$=-\left\|P\nabla f(x)\right\|^2<0$$

$$\boxed{M=\begin{bmatrix}A_1\\E\end{bmatrix}}$$

所以d为下降方向。又

$$Md=-MP\nabla f(x)$$

$$=-M\left(I-M^T(MM^T)^{-1}M\right)\nabla f(x)$$

$$=(-M+M)\nabla f(x)=0$$

即$A_1 d=0, Ed=0\Rightarrow d$是可行方向。

定理2: 设x是问题(1)的可行解，在点x处，有$A_1 x = b_1$，$A_2 x > b_2$，其中

$$A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \quad b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

又设$M = \begin{bmatrix} A_1 \\ E \end{bmatrix}$为满秩矩阵，令

$$P = I - M^T \left(MM^T \right)^{-1} M$$

$$W = \left(MM^T \right)^{-1} M \nabla f(x) = \begin{bmatrix} u \\ v \end{bmatrix}$$

其中u和v分别对应于A_1和E。设$P \nabla f(x) = 0$，则

1. 若$u \geq 0$, 则x为KKT点；

2. 如果u中含有负分量，不妨设$u_j < 0$，这时从A_1中去掉u_j对应的行，得到$\hat{A}_1$。令

$$\hat{M} = \begin{bmatrix} \hat{A}_1 \\ E \end{bmatrix} \qquad \hat{P} = I - \hat{M}^T \left(\hat{M}\hat{M}^T \right)^{-1} \hat{M}$$

$$d = -\hat{P}\nabla f(x)$$

则d为x处的下降可行方向。

$$\boxed{W = \left(MM^T \right)^{-1} M\nabla f(x) = \begin{bmatrix} u \\ v \end{bmatrix}}$$

证明:(1)设$u \geq 0$，则

$$0 = P\nabla f(x) = \left[I - M^T(MM^T)^{-1}M\right]\nabla f(x)$$

$$= \nabla f(x) - M^T(MM^T)^{-1}M\nabla f(x)$$

$$= \nabla f(x) - \begin{pmatrix} A_1^T & E^T \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix}$$

$$\boxed{M = \begin{bmatrix} A_1 \\ E \end{bmatrix}}$$

$$= \nabla f(x) - A_1^T u - E^T v$$

因为$u \geq 0$，所以x为KKT点。

$$\boxed{W = \left(MM^T\right)^{-1} M\nabla f(x) = \begin{bmatrix} u \\ v \end{bmatrix}}$$

2. 如果u中含有负分量，不妨设$u_j < 0$，这时从A_1中去掉u_j对应的行，得到$\hat{A}_1$。令

$$\hat{M} = \begin{bmatrix} \hat{A}_1 \\ E \end{bmatrix} \qquad \hat{P} = I - \hat{M}^T \left(\hat{M}\hat{M}^T \right)^{-1} \hat{M}$$

$$d = -\hat{P}\nabla f(x)$$

则d为x处的下降可行方向。

$$\boxed{W = \left(MM^T \right)^{-1} M\nabla f(x) = \begin{bmatrix} u \\ v \end{bmatrix}}$$

(2)　设$u_j < 0$。若$\hat{P}\nabla f(x) = 0$,则有

$$0 = \hat{P}\nabla f(x) = (I - \hat{M}^T(\hat{M}\hat{M}^T)^{-1}\hat{M})\nabla f(x)$$

$$= \nabla f(x) - \hat{M}^T\hat{W} \qquad (a)$$

其中$\hat{W} = (\hat{M}\hat{M}^T)^{-1}\hat{M}\nabla f(x)$.设$A_1$中对应$u_j$的行向量为$r_j$,

$$\because A_1^T u + E^T v = \hat{A}_1^T\hat{u} + u_j r_j^T + E^T v = \hat{M}^T\bar{W} + u_j r_j^T$$

$$\therefore 0 = \nabla f(x) - A_1^T u - E^T v$$

$$= \nabla f(x) - \hat{M}^T\bar{W} - u_j r_j^T \qquad (b)$$

$$\Rightarrow 0 = \hat{M}^T(\bar{W} - \hat{W}) + u_j r_j^T$$

$\Rightarrow M$的行向量组线性相关，与M行满秩矛盾。

因此必有$\hat{P}\nabla f(x) \neq 0$。

由于$\hat{P}$为投影矩阵，$\hat{P}\nabla f(x) \neq 0$，则

$$\nabla f(x)^T d = -\nabla f(x)^T \hat{P}\nabla f(x)$$

$$= -\left\|\hat{P}\nabla f(x)\right\|^2 < 0$$

$\therefore d = -\hat{P}\nabla f(x)$是下降方向。

由于$\hat{M}d = -\hat{M}\hat{P}\nabla f(x)$

$$= -\hat{M}(I - \hat{M}^T(\hat{M}\hat{M}^T)^{-1}\hat{M})\nabla f(x)$$

$$= -(\hat{M} - \hat{M})\nabla f(x) = 0$$

因此 $\hat{A}_1 d = 0 \quad Ed = 0$

(b)两边左乘$r_j\hat{P}$，得到 $\boxed{0=\nabla f(x)-\hat{M}^T\overline{W}-u_j r_j^T \quad (b)}$

$$r_j\hat{P}\nabla f(x)-r_j\hat{P}\hat{M}^T\overline{W}-u_j r_j\hat{P}r_j^T=0$$

因为$\hat{P}\hat{M}^T=\left(I-\hat{M}^T(\hat{M}\hat{M}^T)^{-1}\hat{M}\right)\hat{M}^T=0,$

$d=-\hat{P}\nabla f(x)$

$\Rightarrow \quad r_j d+u_j r_j\hat{P}r_j^T=0$

由于$\hat{P}$半正定，$r_j\hat{P}r_j^T\geq 0$及$u_j<0$，因此

$$r_j d=-u_j r_j\hat{P}r_j^T\geq 0$$

因此有 $A_1 d\geq 0,\quad Ed=0$

$\Rightarrow \quad d$为可行方向。

$\Rightarrow \quad d$为x处的下降可行方向。

步 骤

1.给定初始可行点$x^{(1)}$，置$k=1$。

2.在点$x^{(k)}$处把A和b分解成

$$\begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \text{和} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

使得$A_1 x^{(k)} = b_1$，$A_2 x^{(k)} > b_2$。

3.令$M = \begin{bmatrix} A_1 \\ E \end{bmatrix}$，如果$M$是空的，则令$P = I$；否则令

$$P = I - M^T (MM^T)^{-1} M。$$

4.令$d^{(k)}=-P\nabla f(x^{(k)})$，若$d^{(k)}\neq 0$，则转6；否则，转5。

5.若M是空的，则停止计算，得$x^{(k)}$；否则令

$$W=(MM^{T})^{-1}M\nabla f(x^{(k)})=\begin{bmatrix}u\\ v\end{bmatrix}。$$

若$u\geq 0$,则停止计算，$x^{(k)}$为KKT点；如果u包含负分量，则选择一个负分量，如u_j，去掉A_1对应u_j的行，转3。

6.计算$\hat{b}=b_2-A_2x^{(k)}$，$\hat{d}=A_2d^{(k)}$，转7。

7.若$\hat{d}\geq 0$,则从$x^{(k)}$出发，沿$d^{(k)}$进行一维搜索

$$\min f(x^{(k)}+\lambda d^{(k)})=f(x^{(k)}+\lambda_k d^{(k)})$$

令$x^{(k+1)}=x^{(k)}+\lambda_k d^{(k)}$，转9；否则，转8。

8.计算$\lambda_{\max} = \min\left\{\frac{\hat{b}_i}{\hat{d}_i} \mid \hat{d}_i < 0\right\}$,求解

$$\begin{aligned} &\min f(x^{(k)} + \lambda d^{(k)}) \\ &s.t. \quad 0 \le \lambda \le \lambda_{\max} \end{aligned}$$

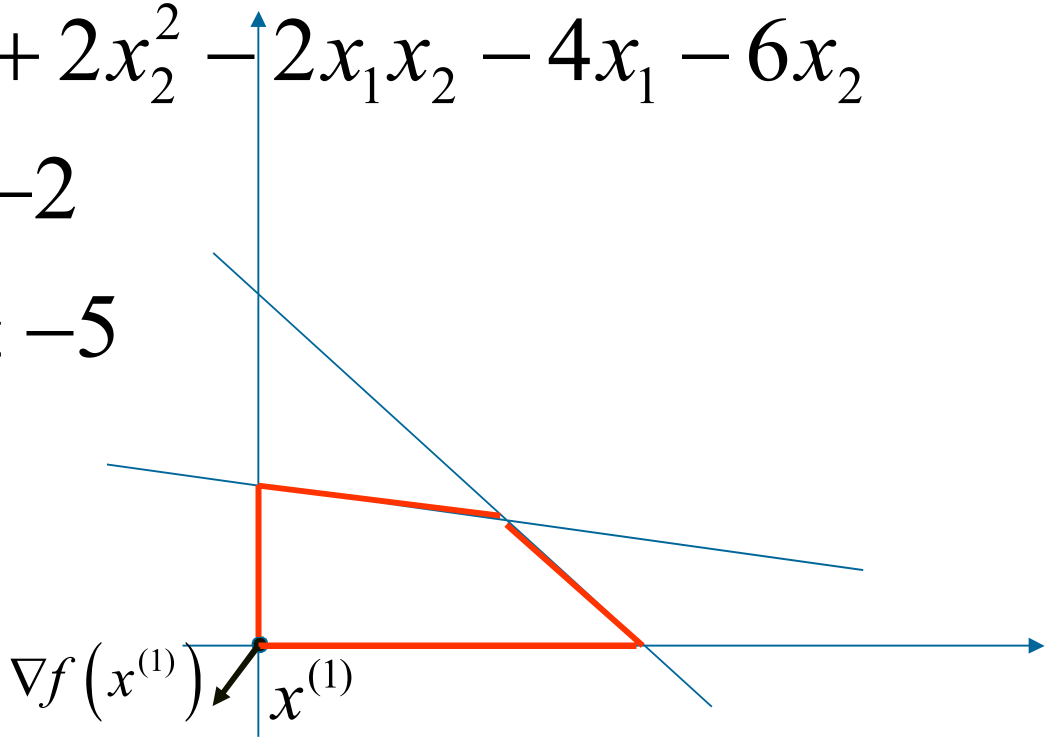
得最优解λ_k，令$x^{(k+1)} = x^{(k)} + \lambda_k d^{(k)}$，转9。

9.确定$x^{(k+1)}$的起作用约束，修改A_1, A_2及b_1, b_2,置$k := k+1$，返回3。

用Rosen梯度投影法求解

$$\min\ f(x)=2x_1^2+2x_2^2-2x_1x_2-4x_1-6x_2$$

$$
\begin{aligned}
s.t.\quad & -x_1-x_2\geq -2\\
& -x_1-5x_2\geq -5\\
& x_1\geq 0\\
& x_2\geq 0
\end{aligned}
$$



解： $x^{(1)}=(0,0)^T, I=\{3,4\}$

$$\nabla f(x)=(4x_1-2x_2-4,\ 4x_2-2x_1-6)^T$$

$$A_1=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, A_2=\begin{pmatrix}-1 & -1\\ -1 & -5\end{pmatrix}, b_1=\begin{pmatrix}0\\ 0\end{pmatrix}, b_2=\begin{pmatrix}-2\\ -5\end{pmatrix}$$

$$\Rightarrow d^{(1)}=-P\nabla f(x^{(1)})=\begin{pmatrix}0\\ 0\end{pmatrix}$$

$$\boxed{P=I-M^T(MM^T)^{-1}M}$$

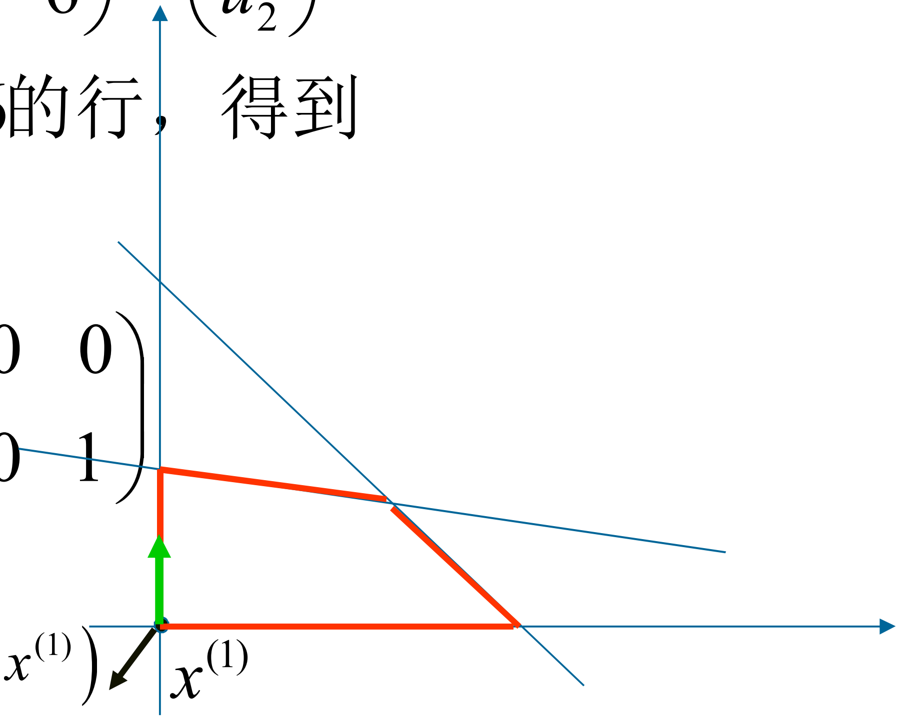
$$令 W=(A_1A_1^T)^{-1}A_1\nabla f(x^{(1)})=\begin{pmatrix}-4\\ -6\end{pmatrix}=\begin{pmatrix}u_1\\ u_2\end{pmatrix}$$

修正A_1，去掉A_1中对应$u_2=-6$的行，得到

$$\hat{A}_1=\begin{pmatrix}1 & 0\end{pmatrix}=M$$

$$\Rightarrow P=I-M^T(MM^T)^{-1}M=\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}$$

$$\Rightarrow d^{(1)}=-P\nabla f(x^{(1)})=\begin{pmatrix}0\\ 6\end{pmatrix}$$



求步长λ_1：

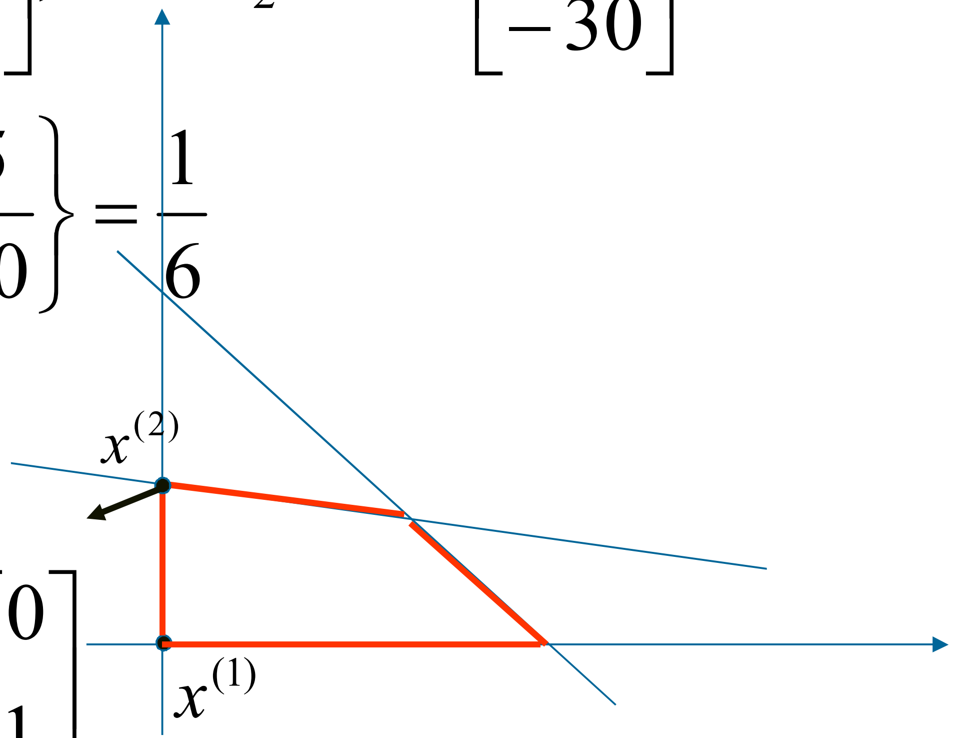
$$\begin{cases} \min f(x^{(1)}+\lambda d^{(1)}) \\ s.t. \quad 0 \le \lambda \le \lambda_{\max} \end{cases}$$

$$\because \hat{b} = b_2 - A_2 x^{(1)} = \begin{bmatrix} -2 \\ -5 \end{bmatrix}, \hat{d} = A_2 d^{(1)} = \begin{bmatrix} -6 \\ -30 \end{bmatrix}$$

$$\therefore \lambda_{\max} = \min\left\{\frac{-2}{-6}, \frac{-5}{-30}\right\} = \frac{1}{6}$$

$$\Rightarrow \lambda_1 = \frac{1}{6}$$

$$\therefore x^{(2)} = x^{(1)} + \lambda_1 d^{(1)} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$



第二次迭代

$$I=\{2,3\},\nabla f(x^{(2)})=\begin{bmatrix}-6\\-2\end{bmatrix}$$

$$A_1=\begin{pmatrix}-1&-5\\1&0\end{pmatrix},A_2=\begin{pmatrix}-1&-1\\0&1\end{pmatrix},b_1=\begin{pmatrix}-5\\0\end{pmatrix},b_2=\begin{pmatrix}-2\\0\end{pmatrix}$$

$$P=I-A_1^T(A_1A_1^T)^{-1}A_1=\begin{pmatrix}0&0\\0&0\end{pmatrix},\quad \therefore d^{(2)}=-P\nabla f(x^{(2)})=0$$

$$\text{令}W=(A_1A_1^T)^{-1}A_1\nabla f(x^{(2)})=\begin{pmatrix}\dfrac{2}{5}\\-\dfrac{28}{5}\end{pmatrix}=\begin{pmatrix}u_1\\u_2\end{pmatrix}$$

从A_1中去掉u_2所对应的第二行，

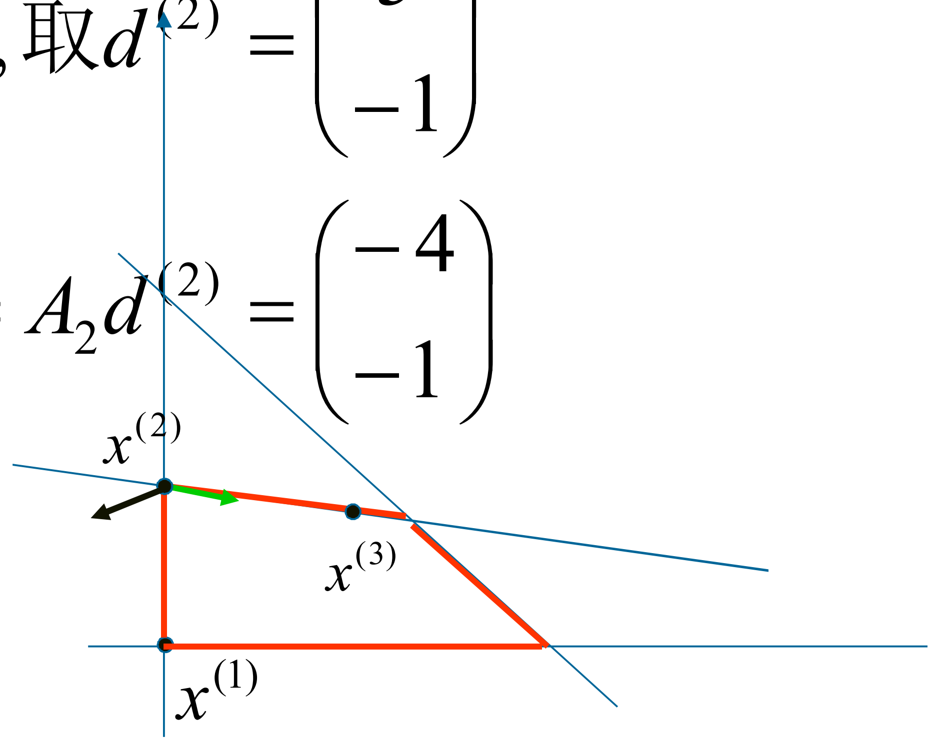
$$\hat{A}_1 = \begin{pmatrix} -1 & -5 \end{pmatrix}$$

$$\hat{P} = I - \hat{A}_1^T(\hat{A}_1\hat{A}_1^T)^{-1}\hat{A}_1 = \begin{pmatrix} \frac{25}{26} & -\frac{5}{26} \\ -\frac{5}{26} & \frac{1}{26} \end{pmatrix}$$

$$d^{(2)} = -\hat{P}\nabla f(x^{(2)}) = \frac{14}{13}\begin{pmatrix} 5 \\ -1 \end{pmatrix}, \text{取} d^{(2)} = \begin{pmatrix} 5 \\ -1 \end{pmatrix}$$

$$\text{则} \hat{b} = b_2 - A_2 x^{(2)} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \hat{d} = A_2 d^{(2)} = \begin{pmatrix} -4 \\ -1 \end{pmatrix}$$

$$\therefore \lambda_{\max} = \min\left\{\frac{-1}{-4}, \frac{-1}{-1}\right\} = \frac{1}{4}$$



$$求步长\lambda_2:\begin{cases}\min\ f(x^{(2)}+\lambda d^{(2)})\\ s.t.\quad 0\le\lambda\le\dfrac{1}{4}\end{cases}$$

$$得\lambda_2=\frac{7}{31},\therefore x^{(3)}=x^{(2)}+\lambda_2 d^{(2)}=\left(\frac{35}{31},\ \frac{24}{31}\right)^T$$

第三次迭代

$$I=\{2\},\nabla f(x^{(3)})=\left(-\frac{32}{31},\ -\frac{160}{31}\right)^T$$

$$A_1=\begin{pmatrix}-1 & -5\end{pmatrix},A_2=\begin{pmatrix}-1 & -1\\ 1 & 0\\ 0 & 1\end{pmatrix},b_1=(-5),b_2=\begin{pmatrix}-2\\ 0\\ 0\end{pmatrix}$$

$$P=I-A_1^T(A_1A_1^T)^{-1}A_1=\frac{1}{26}\begin{pmatrix}25 & -5\\ -5 & 1\end{pmatrix},d^{(3)}=0,$$

$$W=\frac{32}{31}>0,\Rightarrow x^{(3)}是KKT点。$$

Wolfe既约梯度法

$$\min \ f(x)$$

$$s.t. \quad Ax = b, \quad x \geq 0$$

$A_{m\times n}, r(A)=m, b_{m\times 1}, f$是$E^n$上的连续可微函数

基本思想：借鉴求解线性规划的单纯形算法，选择某些变量为基变量，其它的作为非基变量，将基变量用非基变量表示，并从目标函数中消去基变量，得到以非基变量为自变量的简化的目标函数，进而利用此函数的负梯度构造下降可行方向。

既约梯度：简化目标函数关于非基变量的梯度。

$$A=(B,N),\quad x=\begin{pmatrix}x_B\\x_N\end{pmatrix}$$

$$\begin{aligned}\min\ & f(x_B,x_N)\\ s.t.\quad & Bx_B+Nx_N=b\\ & x_B,x_N\ge 0\end{aligned}\quad\Longleftrightarrow\quad\begin{aligned}\min\ & F(x_N)\\ s.t.\quad & x_B,x_N\ge 0\end{aligned}$$

$$x_B=B^{-1}b-B^{-1}Nx_N$$

$$\Rightarrow F(x_N)=f(x_B(x_N),x_N)$$

***f*(*x*)的既约梯度为**

$$r(x_N)=\nabla F(x_N)$$

$$=\nabla_{x_N}f(x)-(B^{-1}N)^T\nabla_{x_B}f(x)$$

确定下降可行方向$d^{(k)}$

$$d^{(k)}=\begin{pmatrix} d_B^{(k)} \\ d_N^{(k)} \end{pmatrix}, d_B^{(k)}\text{和}d_N^{(k)}\text{分别对应基变量和非基变量。}$$

定义$d_N^{(k)}$，使得

$$d_{N_j}^{(k)}=\begin{cases} -x_{N_j}^{(k)} r_j(x_N^{(k)}) & \text{当} r_j(x_N^{(k)})>0 \\ -r_j(x_N^{(k)}) & \text{当} r_j(x_N^{(k)})\le 0 \end{cases}$$

为得到可行方向，应有$Ad^{(k)}=0$

即 $Bd_B^{(k)}+Nd_N^{(k)}=0$

$\therefore$ 取 $d_B^{(k)}=-B^{-1}Nd_N^{(k)}$

$$\Rightarrow \quad d^{(k)}=\begin{pmatrix} -B^{-1}Nd_N^{(k)} \\ d_N^{(k)} \end{pmatrix}$$

$$\boxed{\begin{aligned} r(x_N) &= \nabla F(x_N) \\ &= \nabla_{x_N} f(x)-(B^{-1}N)^T \nabla_{x_B} f(x) \end{aligned}}$$

定理 设x是可行解，$A=(B,N)$是$m\times n$矩阵，B为m阶可逆矩阵，$x=\left(x_B{}^T,x_N{}^T\right)^T$，$x_B>0$，函数$f$在点$x$处可微，又设

$$d=\begin{pmatrix}-B^{-1}Nd_N\\ d_N\end{pmatrix},$$

其中
$$d_{N_j}=\begin{cases}-x_{N_j}r_j(x_N) & r_j(x_N)>0\\ -r_j(x_N) & r_j(x_N)\le 0\end{cases}$$

如果$d\neq 0$，则d是下降可行方向，而且$d=0$的充要条件是x为KKT点。

证明：根据方向d的定义，

$$d_{N_j} = \begin{cases} -x_{N_j} r_j(x_N), & r_j(x_N) > 0 \\ -r_j(x_N) & r_j(x_N) \le 0 \end{cases}$$

$$Ad = Bd_B + Nd_N$$

$$= B(-B^{-1}Nd_N) + Nd_N = 0$$

注意$x_B > 0$，且当$x_{N_j} = 0$时，有$d_{N_j} \ge 0$，

所以，d为可行方向。由于

$$\nabla f(x)^T d = \nabla_{x_B} f(x)^T d_B + \nabla_{x_N} f(x)^T d_N$$

$$= \nabla_{x_B} f(x)^T (-B^{-1}Nd_N) + \nabla_{x_N} f(x)^T d_N$$

$$= r(x_N)^T d_N$$

当$d_N \ne 0$时，$r(x_N)^T d_N < 0$，因此d是下降方向。

$\Rightarrow d$为下降可行方向。

$$r(x_N) = \nabla F(x_N)$$
$$= \nabla_{x_N} f(x) - (B^{-1}N)^T \nabla_{x_B} f(x)$$

设x是KKT点，则存在乘子$w_B, w_N \geq 0$和v，使得

$$\begin{bmatrix} \nabla_{x_B} f(x) \\ \nabla_{x_N} f(x) \end{bmatrix} - \begin{bmatrix} B^T \\ N^T \end{bmatrix} v - \begin{bmatrix} w_B \\ w_N \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$w_B^T x_B = 0$$

$$w_N^T x_N = 0 \qquad (*)$$

$$\boxed{d_{N_j} = \begin{cases} -x_{N_j} r_j(x_N), & r_j(x_N) > 0 \\ -r_j(x_N) & r_j(x_N) \leq 0 \end{cases}}$$

由于$x_B > 0$,所以$w_B = 0 \Rightarrow v = \left(B^T\right)^{-1} \nabla_{x_B} f(x)$

因此$w_N = \nabla_{x_N} f(x) - \left(B^{-1}N\right)^T \nabla_{x_B} f(x)$

$$= r(x_N) \geq 0$$

由$(*)$，得$r(x_N)^T x_N = 0 \Rightarrow d = 0.$

$$\boxed{\begin{aligned} & r(x_N) = \nabla F(x_N) \\ & = \nabla_{x_N} f(x) - (B^{-1}N)^T \nabla_{x_B} f(x) \end{aligned}}$$

"$\Rightarrow$"已知$d=0$，则$r(x_N)$的分量均非负。令

$$w_N = r(x_N)$$
$$= \nabla_{x_N} f(x) - (B^{-1}N)^T \nabla_{x_B} f(x) \geq 0$$

又令 $w_B = 0, v = (B^T)^{-1} \nabla_{x_B} f(x)$,则有

$$\begin{bmatrix} \nabla_{x_B} f(x) \\ \nabla_{x_N} f(x) \end{bmatrix} - \begin{bmatrix} B^T \\ \mathrm{N}^T \end{bmatrix} v - \begin{bmatrix} w_B \\ w_N \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$w_B{}^T x_B = 0$$
$$w_N{}^T x_N = 0$$

$$\boxed{d_{N_j} = \begin{cases} -x_{N_j} r_j(x_N), & r_j(x_N) > 0 \\ -r_j(x_N) & r_j(x_N) \leq 0 \end{cases}}$$

成立，所以，x是KKT点。

确定步长

为保持$x^{(k+1)} \geq 0$

即　$x_j^{(k+1)} = x_j^{(k)} + \lambda d_j^{(k)} \geq 0,\ j = 1, 2, \cdots, n$

当$d_j^{(k)} \geq 0$时，取$\lambda \geq 0$

当$d_j^{(k)} < 0$时，应取 $\lambda \leq \dfrac{x_j^{(k)}}{-d_j^{(k)}}$

$$\lambda_{\max} = \begin{cases} \infty & \text{当} d^{(k)} \geq 0 \\ \min\left\{ -\dfrac{x_j^{(k)}}{d_j^{(k)}} \mid d_j^{(k)} < 0 \right\} & \text{其他情形} \end{cases}$$

步 骤

1.给定初始可行点$x^{(1)}$，允许误差$\varepsilon > 0$，置$k = 1$。

2.从$x^{(k)}$中选择m个最大分量，它们的下标集记为J_k，

A的第j列记为p_j，令

B是由$\{p_j \mid j \in J_k\}$构成的m阶矩阵

N是由$\{p_j \mid j \notin J_k\}$构成的$m \times (n-m)$阶矩阵

计算$r(x_N) = \nabla_{x_N} f(x_B(x_N), x_N) - (B^{-1}N)^T \nabla_{x_B} f(x_B(x_N), x_N)$

计算
$$d_{N_j}^{(k)} = \begin{cases} -x_{N_j}^{(k)} r_j\left(x_N^{(k)}\right) & r_j\left(x_N^{(k)}\right) > 0 \\ -r_j\left(x_N^{(k)}\right) & r_j\left(x_N^{(k)}\right) \le 0 \end{cases}$$

$d_B^{(k)} = -B^{-1} N d_N^{(k)}$,得搜索方向$d^{(k)} = \begin{pmatrix} d_B^{(k)} \\ d_N^{(k)} \end{pmatrix}$.

3. 若$\left\|d^{(k)}\right\| \le \varepsilon$，则停止计算，得点$x^{(k)}$；否则，转4。

4. 计算
$$\lambda_{\max} = \begin{cases} \infty & d^{(k)} \ge 0 \\ \min\left\{-\dfrac{x_j^{(k)}}{d_j^{(k)}} \mid d_j^{(k)} < 0\right\} & \text{否则} \end{cases}$$

从$x^{(k)}$出发，沿$d^{(k)}$进行一维搜索：

$$\min f(x^{(k)} + \lambda d^{(k)})$$
$$s.t. \quad 0 \le \lambda \le \lambda_{\max}$$

得到最优解λ_k。

5. 令$x^{(k+1)} = x^{(k)} + \lambda_k d^{(k)}$，置$k := k+1$，返回2。

收敛性

定理：设问题

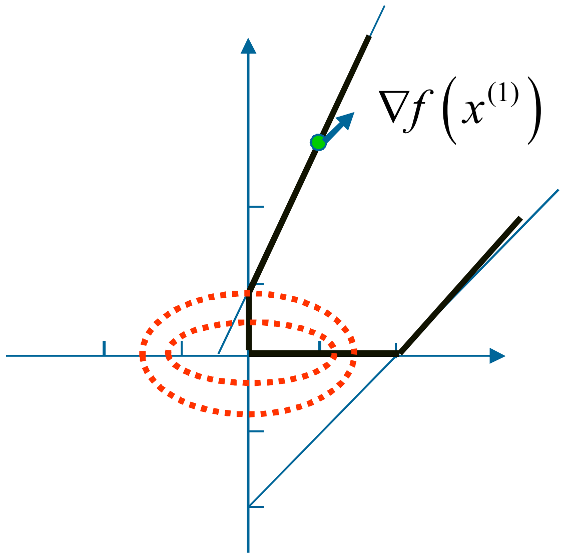
$$\min f(x)$$

$$s.t. \quad Ax = b, \quad x \geq 0$$

中f是E^n上的连续可微函数，若A的任意m列均线性无关且所有基本可行解非退化，则$Wolfe$既约梯度法产生的点列$\left\{x^{(k)}\right\}$任意聚点是KKT点。

例： $\min\ 2x_1^2 + x_2^2$

$$s.t.\quad x_1 - x_2 + x_3 = 2$$
$$-2x_1 + x_2 + x_4 = 1$$
$$x_j \geq 0,\ j = 1,2,3,4$$



初始可行点$x^{(1)} = (1, 3, 4, 0)^T$

$$\nabla f(x) = (4x_1, 2x_2, 0, 0)^T \qquad \nabla f(x^{(1)}) = (4, 6, 0, 0)^T$$

第一次迭代 $\quad J_1 = \{2,3\}$

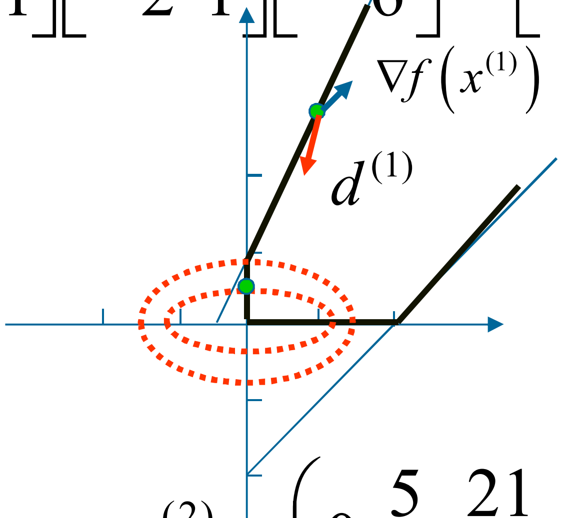
$$x_B = \begin{bmatrix} x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \qquad x_N = \begin{bmatrix} x_1 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$B=\begin{bmatrix}-1 & 1\\ 1 & 0\end{bmatrix}\quad B^{-1}=\begin{bmatrix}0 & 1\\ 1 & 1\end{bmatrix}\quad N=\begin{bmatrix}1 & 0\\ -2 & 1\end{bmatrix}$$

$$r\left(x_N^{(1)}\right)=\begin{bmatrix}4\\ 0\end{bmatrix}-\left(\begin{bmatrix}0 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix}1 & 0\\ -2 & 1\end{bmatrix}\right)^T\begin{bmatrix}6\\ 0\end{bmatrix}=\begin{bmatrix}4\\ 0\end{bmatrix}-\begin{bmatrix}-2 & -1\\ 1 & 1\end{bmatrix}\begin{bmatrix}6\\ 0\end{bmatrix}=\begin{bmatrix}16\\ -6\end{bmatrix}$$

$$d_N^{(1)}=\begin{bmatrix}d_1^{(1)}\\ d_4^{(1)}\end{bmatrix}=\begin{bmatrix}-16\\ 6\end{bmatrix}\quad d_B^{(1)}=\begin{bmatrix}d_2^{(1)}\\ d_3^{(1)}\end{bmatrix}=-\begin{bmatrix}0 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix}1 & 0\\ -2 & 1\end{bmatrix}\begin{bmatrix}-16\\ 6\end{bmatrix}=\begin{bmatrix}-38\\ -22\end{bmatrix}$$

$$\Rightarrow d^{(1)}=(-16,-38,-22,6)^T$$



$$\lambda_{\max}=\min\left\{-\frac{1}{-16},-\frac{3}{-38},-\frac{4}{-22}\right\}=\frac{1}{16}$$

求解问题$\begin{cases}\min\ f(x^{(1)}+\lambda d^{(1)})\\ s.t.\quad 0\le\lambda\le\dfrac{1}{16}\end{cases}$得到$\lambda_1=\dfrac{1}{16},\Rightarrow x^{(2)}=\left(0,\dfrac{5}{8},\dfrac{21}{8},\dfrac{3}{8}\right)^T$

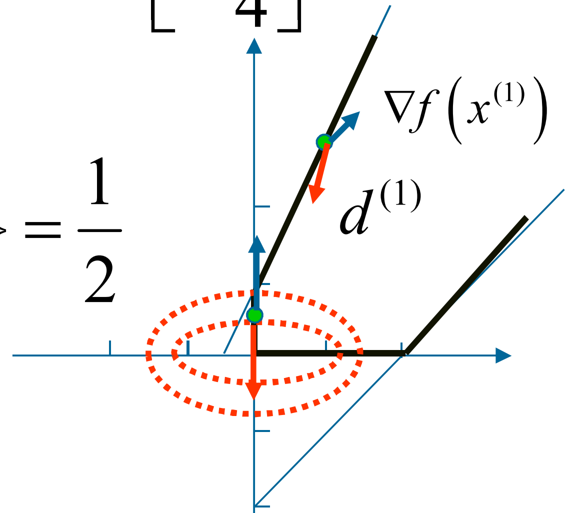
第二次迭代

$$J_2 = \{2,3\}$$

$$\nabla f(x^{(2)}) = \left(0, \frac{5}{4}, 0, 0\right)^T \quad x_B^{(2)} = \begin{bmatrix} x_2 \\ x_3 \end{bmatrix} \quad x_N^{(2)} = \begin{bmatrix} x_1 \\ x_4 \end{bmatrix}$$

$$r(x_N^{(2)}) = \begin{bmatrix} \frac{5}{2} \\ -\frac{5}{4} \end{bmatrix} \quad d_N^{(2)} = \begin{bmatrix} d_1^{(2)} \\ d_4^{(2)} \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{5}{4} \end{bmatrix} \quad d_B^{(2)} = \begin{bmatrix} d_2^{(2)} \\ d_3^{(2)} \end{bmatrix} = \begin{bmatrix} -\frac{5}{4} \\ -\frac{5}{4} \end{bmatrix}$$

$$\Rightarrow d^{(2)} = \left(0, -\frac{5}{4}, -\frac{5}{4}, \frac{5}{4}\right)^T \quad \lambda_{\max} = \min\left\{\frac{\frac{5}{8}}{\frac{5}{4}}, \frac{\frac{21}{8}}{\frac{5}{4}}\right\} = \frac{1}{2}$$



$$\text{求解问题}\begin{cases} \min\ f(x^{(2)} + \lambda d^{(2)}) \\ s.t. \quad 0 \le \lambda \le \frac{1}{2} \end{cases} \Rightarrow \lambda_2 = \frac{1}{2}, \Rightarrow x^{(3)} = (0,0,2,1)^T$$

第三次迭代

$$J_3 = \{3,4\}$$

$$\nabla f(x^{(2)}) = (0,0,0,0)^T$$

$$x_B^{(2)} = \begin{bmatrix} x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad x_N^{(2)} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad B^{-1} = B \quad N = \begin{bmatrix} 1 & -1 \\ -2 & 1 \end{bmatrix}$$

$$r(x_N^{(2)}) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -2 & 1 \end{bmatrix} \right)^T \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow d^{(2)} = (0,0,0,0)^T$$

$$x^* = (0,0,2,1)^T$$

广义既约梯度法

$$
\begin{aligned}
&\min \ f(x) \\
&s.t. \quad h_j(x) = 0, \quad j = 1,2,\cdots,m \\
&\qquad\quad l \le x \le u
\end{aligned}
$$

其中$f, h_j(j=1,2,\cdots,m)$是连续可微函数， $x \in E^n$,

$m \le n$, l和u都是n维列向量：

$$l = (l_1,\cdots,l_n)^T, \quad u = (u_1,\cdots,u_n)^T$$

$$\min \ f(x)$$
$$s.t. \quad h(x) = 0$$
$$l \le x \le u$$

其中$h(x) = (h_1(x), \cdots, h_m(x))^T$.

将变量区分为基变量向量x_B和非基变量向量x_N

$$\frac{\partial h}{\partial x} = \begin{pmatrix} \frac{\partial h_1}{\partial x_1} & \cdots & \frac{\partial h_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial h_m}{\partial x_1} & \cdots & \frac{\partial h_m}{\partial x_n} \end{pmatrix} = \left(\frac{\partial h}{\partial x_B}, \quad \frac{\partial h}{\partial x_N} \right)$$

***h*(*x*)的Jacobi**
矩阵

假设$\dfrac{\partial h}{\partial x_B}$可逆，则$x_B$可用$x_N$表示，所以把目标函数化为：

$$F(x_N) = f(x_B(x_N), x_N)$$

求既约梯度 $r(x_N) = \nabla F(x_N)$

$$\mathrm{d}f = \left(\nabla_{x_B} f\right)^T \mathrm{d}x_B + \left(\nabla_{x_N} f\right)^T \mathrm{d}x_N$$

其中

$$\nabla_{x_B} f = \left(\frac{\partial f}{\partial x_1}, \cdots, \frac{\partial f}{\partial x_m}\right)^T$$

$$\nabla_{x_N} f = \left(\frac{\partial f}{\partial x_{m+1}}, \cdots, \frac{\partial f}{\partial x_n}\right)^T$$

$$\mathrm{d}x_B = \left(\mathrm{d}x_1, \cdots, \mathrm{d}x_m\right)^T$$

$$\mathrm{d}x_N = \left(\mathrm{d}x_{m+1}, \cdots, \mathrm{d}x_n\right)^T.$$

由$h(x)=0$，得

$$\begin{cases} \mathrm{d}h_1 = \dfrac{\partial h_1}{\partial x_1}\mathrm{d}x_1 + \cdots + \dfrac{\partial h_1}{\partial x_n}\mathrm{d}x_n = 0 \\ \cdots\cdots \\ \mathrm{d}h_m = \dfrac{\partial h_m}{\partial x_1}\mathrm{d}x_1 + \cdots + \dfrac{\partial h_m}{\partial x_n}\mathrm{d}x_n = 0 \end{cases}$$

$$\Rightarrow \quad \frac{\partial h}{\partial x_B}\mathrm{d}x_B + \frac{\partial h}{\partial x_N}\mathrm{d}x_N = 0$$

$$\Rightarrow \quad \mathrm{d}x_B = -\left(\frac{\partial h}{\partial x_B}\right)^{-1}\frac{\partial h}{\partial x_N}\mathrm{d}x_N$$

$$
\mathrm{d}f = \left(\nabla_{x_B} f\right)^T \mathrm{d}x_B + \left(\nabla_{x_N} f\right)^T \mathrm{d}x_N
$$

$$
= \left[\nabla_{x_N} f - \left(\left(\frac{\partial h}{\partial x_B}\right)^{-1} \frac{\partial h}{\partial x_N}\right)^T \nabla_{x_B} f\right]^T \mathrm{d}x_N
$$

所以，既约梯度为：

$$
r(x_N) = \frac{\mathrm{d}f}{\mathrm{d}x_N} = \nabla_{x_N} f - \left(\left(\frac{\partial h}{\partial x_B}\right)^{-1} \frac{\partial h}{\partial x_N}\right)^T \nabla_{x_B} f.
$$

其中 $\frac{\mathrm{d}f}{\mathrm{d}x_N} = \left(\frac{\mathrm{d}f}{\mathrm{d}x_{m+1}}, \cdots, \frac{\mathrm{d}f}{\mathrm{d}x_n}\right)^T$

确定搜索方向

给定一点$x^{(k)}$，定义$d_N^{(k)}$，使它的分量满足：

$$
d_{N_j}^{(k)}=\begin{cases}0 & \text{当}x_{N_j}^{(k)}=l_{N_j}\text{且}r_j\left(x_N^{(k)}\right)>0 \\ & \text{或当}x_{N_j}^{(k)}=u_{N_j}\text{且}r_j\left(x_N^{(k)}\right)<0 \\ -r_j\left(x_N^{(k)}\right) & \text{其他情形}\end{cases}
$$

确定步长

令 $\hat{x}_N = x_N^{(k)} + \lambda d_N^{(k)}$

使得 $l_N \le \hat{x}_N \le u_N$

求解非线性方程组： $h(y, \hat{x}_N) = 0$

得到$\hat{y}$. 若$\hat{y}$满足

$$f\left(\hat{y}, x_N^{(k)}\right) < f\left(x_B^{(k)}, x_N^{(k)}\right)$$

且 $l_B \le \hat{y} \le u_B$

则得到新的可行点$(\hat{y}, \hat{x}_N)$; 否则减少步长，

重复以上过程。

Frank-Wolfe方 法

$$
(1)\quad \begin{cases} \min\ f(x) \\ s.t.\quad Ax \geq b \\ \qquad\quad Ex = e \end{cases}
$$

$A_{m\times n}, r(A) = m, E_{l\times n}, f(x)$连续可微

令 $S = \{x \mid Ax \geq b, Ex = e, x \in E^n\}$

基本思想：在每次迭代中，将目标函数$f(x)$线性化，通过解线性规划求得下降可行方向，进而沿此方向在可行域内做一维搜索。

任取$x^{(k)} \in S$，在$x^{(k)}$处以$f(x)$的一阶$Taylor$展开式作为$f(x)$的线性逼近函数：

$$f_L(x) = f(x^{(k)}) + \nabla f(x^{(k)})^T (x - x^{(k)})$$

$$= f(x^{(k)}) - \nabla f(x^{(k)})^T x^{(k)} + \nabla f(x^{(k)})^T x$$

求解线性规划问题

$$\begin{aligned} &\min \ f_L(x) \\ &s.t. \quad x \in S \end{aligned} \quad \Longleftrightarrow \quad (2) \quad \begin{cases} \min \nabla f(x^{(k)})^T x \\ s.t. \quad x \in S \end{cases}$$

设问题(2)存在有限最优解$y^{(k)}$.

求解(2)有下列两种情形之一：

1. 若$\nabla f(x^{(k)})^T(y^{(k)}-x^{(k)})=0$，则停止迭代，$x^{(k)}$是(1)的$KKT$点。

2. 若$\nabla f(x^{(k)})^T(y^{(k)}-x^{(k)})\neq 0$，则必有

$$\nabla f(x^{(k)})^T(y^{(k)}-x^{(k)})<0$$

$$\boxed{\begin{cases} \min \nabla f(x^{(k)})^T x \\ s.t. \quad x\in S \end{cases}}$$

$\Rightarrow y^{(k)}-x^{(k)}$为$x^{(k)}$处的下降方向。

$\because S$是凸集，$\therefore$ 对$\forall \lambda\in(0,1)$，有

$$\lambda y^{(k)}+(1-\lambda)x^{(k)}=x^{(k)}+\lambda(y^{(k)}-x^{(k)})\in S$$

$\Rightarrow y^{(k)}-x^{(k)}$为$x^{(k)}$处的可行方向。

$\Rightarrow y^{(k)}-x^{(k)}$为$x^{(k)}$处的下降可行方向。

定理： 设$y^{(k)}$是(2)的最优解，且满足

$\nabla f(x^{(k)})^T(y^{(k)}-x^{(k)})=0$，则$x^{(k)}$是(1)的$KKT$点。

证明：由$\nabla f(x^{(k)})^T(y^{(k)}-x^{(k)})=0$，得

$$\nabla f(x^{(k)})^T y^{(k)}=\nabla f(x^{(k)})^T x^{(k)}$$

$\because y^{(k)}$是(2)的最优解，且$x^{(k)}\in S$

$\therefore x^{(k)}$是(2)的KKT点$\Rightarrow \exists w\geq 0(w\in E^m),v\in E^n$使得

$$\begin{cases}\nabla f(x^{(k)})-A^T w-E^T v=0\\ w^T\left(Ax^{(k)}-b\right)=0\\ Ex^{(k)}=e\end{cases}$$

$$\boxed{\begin{cases}\min \nabla f(x^{(k)})^T x\\ s.t.\quad x\in S\end{cases}}$$

这也是(1)的KKT条件$\therefore x^{(k)}$是(1)的KKT点。

步骤

1.给定初始可行点$x^{(1)}$，允许误差$\varepsilon > 0$，置$k = 1$。

2. 求解线性规划问题

$$\begin{cases} \min \nabla f(x^{(k)})^T x \\ s.t. \quad x \in S \end{cases}$$

得到最优解$y^{(k)}$。

3. 若$\left|\nabla f(x^{(k)})^T (y^{(k)} - x^{(k)})\right| \le \varepsilon$,则停止，得$x^{(k)}$；否则转4。

4. 从$x^{(k)}$出发，沿方向$y^{(k)} - x^{(k)}$进行一维搜索：

$$\begin{cases} \min\ f(x^{(k)} + \lambda(y^{(k)} - x^{(k)})) \\ s.t. \quad 0 \le \lambda \le 1 \end{cases}$$

得最优解λ_k.

5. $x^{(k+1)} = x^{(k)} + \lambda_k (y^{(k)} - x^{(k)})$,置$k := k + 1$，返回2。

例：
$$\begin{cases} \min\ f(x) = 4x_1^2 + (x_2 - 2)^2 \\ s.t. \quad -2 \le x_1 \le 2 \\ \qquad\ \ -1 \le x_2 \le 1 \end{cases}$$

解：取初始点$x^{(1)} = (-2, -1)^T$

第一次迭代 $\nabla f(x^{(1)}) = (-16, -6)^T$

求$\begin{cases} \min \nabla f(x^{(1)})^T x = -16x_1 - 6x_2 \\ s.t. \quad -2 \le x_1 \le 2 \\ \qquad\ \ -1 \le x_2 \le 1 \end{cases}$

得$y^{(1)} = (2, 1)^T$.

从$x^{(1)}$出发，沿$y^{(1)}-x^{(1)}$作一维搜索

$$\begin{cases} \min\ f(x^{(1)}+\lambda(y^{(1)}-x^{(1)})) \\ s.t.\quad 0\le\lambda\le 1 \end{cases}$$

得$\lambda_1=0.56$

$$\therefore x^{(2)}=x^{(1)}+\lambda_1(y^{(1)}-x^{(1)})=(0.24,0.12)^T$$

第二次迭 代

求$\begin{cases} \min \nabla f(x^{(2)})^T x = 1.92x_1 - 3.76x_2 \\ s.t.\quad -2\le x_1\le 2 \\ \qquad\quad -1\le x_2\le 1 \end{cases}$

得$y^{(2)}=(-2,1)^T$.

从$x^{(2)}$出发，沿$y^{(2)}-x^{(2)}$作一维搜索，得$x^{(3)}$，

这样继续下去，得

$$y^{(1)}=y^{(3)}=y^{(5)}=\cdots=(2,1)^T$$

$$y^{(2)}=y^{(4)}=y^{(6)}=\cdots=(-2,1)^T$$

显然$\lim\limits_{k\to\infty} x^{(k)}=(0,1)^T=x^*$.