运筹学第一章习题

2018 年 10 月 25 日

1. 证明凸集表示定理：设$\mathcal{S}=\{x|Ax=b,x\geq 0\}$为非空多面集，则有

   (1) 极点集非空且有限

   (2) 极方向集合为空集的充要条件是$\mathcal{S}$有界

   (3) 若$\mathcal{S}$无界，则存在有限个极方向

   (4) $x\in\mathcal{S}$的充要条件是

   $$x=\sum_{i=1}^{k}\lambda_i x^{(i)}+\sum_{j=1}^{\ell}\mu_j d^{(j)} \tag{1}$$

   其中$\lambda_i\geq 0,i=1,\ldots,k,\sum_{i=1}^{k}\lambda_i=1,\mu_j\geq 0,j=1,\ldots,\ell$。

也可以参考*Linear Programming and Network Flows*

证明. (1) 令y为$\mathcal{S}$中0分量最多的一点，即

$$\#\{i|y_i=0\}\geq\max_{x\in\mathcal{S}}\#\{j|x_j=0\}$$

下证明y为$\mathcal{S}$的极点。

若不然，存在$x^{(1)}\neq x^{(2)}\in\mathcal{S},\lambda\in(0,1)$使得$\lambda x^{(1)}+(1-\lambda)x^{(2)}=y$。由于$x^{(1)},x^{(2)}\geq 0$，对$i$使得$y_i=0$，必有$x_i^{(l)}=0,l=1,2$。令$\alpha=x^{(2)}-x^{(1)}$，易见$A\alpha=0$。令$r$满足

$$|\frac{y_r}{\alpha_r}|=\min\{|\frac{y_j}{\alpha_j}||\alpha_j\neq 0\}$$

则$y'=y-\frac{y_r}{\alpha_r}\alpha\in\mathcal{S}$，且$y'_r=0$；对$i$使得$y_i=0$，亦有$y'_i=0$。这与$y$的定义矛盾。因此$y$是极点。这就证明了极点集非空。

由第3题，极点与可行基解一一对应，而可行基解的数量不多于$\begin{pmatrix} n \\ m \end{pmatrix}$，因此极点集有限。

(2)

(3) 考察

$$\mathcal{S}_0 = \{x|Ax = 0, x \geq 0, \sum_{i=1}^{n} x_i = 1\}$$

及标准化后的方向集

$$\mathcal{D} = \{\frac{d}{\|d\|_1}|d\text{为}\mathcal{S}\text{的方向}\}$$

可以看出$\mathcal{S}_0 = \mathcal{D}$，因此$\mathcal{D}$是多面集。下证$\mathcal{S}$标准化后的极方向与$\mathcal{D}$的极点一一对应。

设d为$\mathcal{D}$的极点，并假设$d = \lambda_1 d^{(1)} + \lambda_2 d^{(2)}$，$\lambda_1, \lambda_2 > 0$。因为

$$d = \lambda_1 \|d^{(1)}\|_1 \frac{d^{(1)}}{\|d^{(1)}\|_1} + \lambda_2 \|d^{(2)}\|_1 \frac{d^{(2)}}{\|d^{(2)}\|_1}$$

我们不妨直接假设$\|d^{(j)}\|_1 = 1, j = 1, 2$。于是

$$1 = \|d\|_1 = \lambda_1 \|d^{(1)}\|_1 + \lambda_2 \|d^{(2)}\|_1 = \lambda_1 + \lambda_2$$

即d是$\mathcal{D}$上两点的凸组合。由d为极点知$d^{(1)} = d^{(2)} = d$，因此d是$\mathcal{S}$的极方向。

设d为$\mathcal{S}$的极方向，并且$\|d\|_1 = 1$。若有$d^{(1)}, d^{(2)} \in \mathcal{D}$及$\lambda \in (0, 1)$使得$d = \lambda d^{(1)} + (1 - \lambda) d^{(2)}$，因为这是两个方向的正线性组合，因此有$d = d^{(1)} = d^{(2)}$，故$d$是$\mathcal{D}$的极点。

若$\mathcal{S}_0$为空集，则任意x满足$Ax = 0$且$\|x\|_1 = 1$都具有负分量。令

$$\delta = -\max\{\sum_{i:x_i<0} x_i | Ax = 0, \|x\|_1 = 1\}$$

因为$\{x|Ax = 0, \|x\|_1\}$是一个有界闭集，因此最大值可以取到并严格大于0。$\mathcal{S}$非空，设$y \in \mathcal{S}$，则

$$\mathcal{S} \subseteq y + \{\lambda x | Ax = 0, \|x\|_1 = 1, \lambda \in [0, \frac{\|y\|_1}{\delta}]\}$$

即$\mathcal{S}$有界。

因此，$\mathcal{S}$无界则有$\mathcal{S}_0$非空。由(1)知$\mathcal{S}_0$有极点，因此$\mathcal{S}$有极方向。进一步，极点个数有限即得极方向个数有限。

若极方向集合非空，则$\mathcal{S}$显然是无界的。

(4) 充分性显然。

必要性：对$\mathcal{S}$中元素的正分量个数进行归纳。

由(1)知正分量个数最少的一点（零分量最多的一点）为极点。

假设正分量个数小于t的点都具有(1)表示。设x的正分量个数等于t，不妨设

$$x = (x_1, \dots, x_t, 0, \dots, 0)$$

若$x_1, \dots, x_t$对应A中的列向量$a_1, \dots, a_t$线性无关，则由$Rank(A) = m$，可以再选$m - t$列构成A的一个可逆子矩阵$B = (a_1, \dots, a_m)$，知x是可行基解，亦为极点。

若$a_1, \dots, a_t$线性相关，不妨设$a_{s+1}, \dots, a_t$为它们的最大线性无关组，则a_1具有表示

$$a_1 = \mu_{s+1} a_{s+1} + \dots + \mu_t a_t$$

在A的其余列中选择$m + s - t$列构成可逆阵B，不妨记为

$$B = (a_{s+1}, \dots, a_k, a_{k+1}, \dots, a_{s+m})$$

令$d_N = (1, 0, \dots, 0)$，$d_B = -B^{-1} N d_N$，构成的d满足$Ad = 0$。于是

$$Ad = a_1 + d_2 a_2 + \dots + d_n a_n = a_1 + d_{s+1} a_{s+1} + \dots + d_{s+m} a_{s+m} = 0$$

即

$$a_1 = -d_{s+1} a_{s+1} - \dots + d_{s+m} a_{s+m}$$

代入上上式得

$$(\mu_{s+1} + d_{s+1}) a_{s+1} + \dots + (\mu_t + d_t) a_t + d_{t+1} a_{t+1} + \dots + d_{s+m} a_{s+m} = 0$$

故$d_{t+1} = \dots = d_{s+m} = 0.$

- 若$d \geq 0$，令

$$\theta = \min_{1 \leq i \leq t} \{ \frac{x_i}{d_i} | d_i > 0 \} = \frac{x_l}{d_l} > 0$$

并令$x' = x - \theta D$，则x'为可行解，$x'_l = 0$。故x'前k个分量中至少有一个为0，而后$n-t$个分量等于0，因此由归纳假设，

$$x' = \sum_{i=1}^{k} \lambda_i x^{(i)} + \sum_{j=1}^{\ell} \mu_j d^{(j)}$$

而$\frac{d}{\|d\|_1}$为$\mathcal{S}_0$上的一点，由讲义的结论2，有界多面集内任意一点可以表示为极点的凸组合，故θd可以表示为极方向的正线性组合。故$x = x' + \theta d$满足假设。

- 若d中有分量小于0，则令

$$\theta_1 = \min_{1\le i\le t}\{\frac{x_i}{d_i}|d_i > 0\}$$

$$\theta_2 = \min_{1\le i\le t}\{\frac{x_i}{-d_i}|d_i < 0\}$$

并令$x' = x - \theta_1 d, x'' = x + \theta_2 d$，则$x', x''$的正分量个数都小于$k$，由归纳假设

$$x' = \sum_{i=1}^{k} \lambda_i^{(1)} x^{(i)} + \sum_{j=1}^{\ell} \mu_j^{(1)} d^{(j)}$$

$$x'' = \sum_{i=1}^{k} \lambda_i^{(2)} x^{(i)} + \sum_{j=1}^{\ell} \mu_j^{(2)} d^{(j)}$$

又由

$$x = \frac{\theta_2}{\theta_1+\theta_2}x' + \frac{\theta_1}{\theta_1+\theta_2}x''$$

得

$$x = \sum_{i=1}^{k} \lambda'_i x^{(i)} + \sum_{j=1}^{\ell} \mu'_j d^{(j)}$$

其中

$$\lambda'_i = \frac{\theta_2\lambda_i^{(1)} + \theta_1\lambda_i^{(2)}}{\theta_1+\theta_2} \ge 0$$

$$\sum_i \lambda'_i = \frac{\theta_2}{\theta_1+\theta_2}\sum_i \lambda_i^{(1)} + \frac{\theta_1}{\theta_1+\theta_2}\sum_i \lambda_i^{(2)} = 1$$

$$\mu'_i = \frac{\theta_2\mu_i^{(1)} + \theta_1\mu_i^{(2)}}{\theta_1+\theta_2} \ge 0$$

即对x假设成立。

由数学归纳法原理，证毕。

□

2. 试给出

$$\begin{aligned} &\min -x_1+3x_2 \\ &\text{s.t. } x_1+x_2 \le 8 \\ &\quad x_2 \le 2 \\ &\quad x_1, x_2 \ge 0 \end{aligned}$$

所有极点和可行解。

解 将原问题标准化，

$$\begin{aligned} &\min\ -x_1+3x_2 \\ &\text{s.t. } x_1+2x_2+x_3=8 \\ &\quad x_2+x_4=2 \\ &\quad x_1,x_2,x_3,x_4 \ge 0 \end{aligned}$$

令$A=\begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$，$x=(x_1,x_2,x_3,x_4)^T$，$b=\begin{pmatrix} 8 \\ 2 \end{pmatrix}$，则变成

$$\begin{aligned} &\min\ -x_1+3x_2 \\ &\text{s.t. } Ax=b \\ &\quad x \ge 0 \end{aligned}$$

取B为A的二阶可逆子阵，得到

$$\begin{aligned} x^{(1)} =&(4,2,0,0)^T \\ x^{(2)} =&(8,0,0,2)^T \\ x^{(3)} =&(0,2,4,0)^T \\ x^{(4)} =&(0,4,0,-2)^T \text{ (舍去)} \\ x^{(5)} =&(0,0,8,2)^T \end{aligned}$$

至于解集的极点，应该画图后标出。

3. 证明极点集-可行基解集等价定理。

证明. $\Leftarrow$: 令$x = \begin{pmatrix} x_B \\ x_N \end{pmatrix}$为可行基解，即$x_N = 0$，$B$可逆。设$x^{(1)}, x^{(2)} \in \mathcal{S}, \lambda \in (0,1)$使得$\lambda x^{(1)} + (1-\lambda)x^{(2)} = x$。由于$x^{(1)}, x^{(2)} \geq 0$，而$x_N = 0$，故$x_N^{(1)} = x_N^{(2)} = 0$，于是$x_B^{(1)} = B^{-1} = x_B^{(2)} = x_B$，即$x^{(1)} = x^{(2)} = x$，这说明$x$是一个极点。

$\Rightarrow$: 令x为一个极点。令$x_{p_1}, x_{p_2}, \ldots, x_{p_k}$为$x$中非零分量，对应$A$中$a_{p_1}, \ldots, a_{p_k}$列。

断言：$a_{p_1}, \ldots, a_{p_k}$线性无关。

若$\exists \lambda_1, \ldots, \lambda_k$不全为0使得$\sum_{i=1}^{k} \lambda_i a_{p_i} = 0$，令$x_{p_i}^{(1)} = x_{p_i} + \delta\lambda_i$，$x_{p_i}^{(2)} = x_{p_i} - \delta\lambda_i$，$i = 1, \ldots, k$，$x^{(1)}$与$x^{(2)}$其余分量均为0，那么由于$x_{p_i} > 0$，取$|\delta|$足够小，有$Ax^{(l)} = b, x^{(l)} \geq 0, l = 1, 2$，且$\frac{1}{2}x^{(1)} + \frac{1}{2}x^{(2)} = x$，这与$x$为极点矛盾。这就证明了断言。

由断言可知$k \leq m$。在x的其余分量中选择$n-k$个分量，使得A对应的列向量$a_{p_{k+1}}, \ldots, a_{p_m}$与$a_{p_1}, \ldots, a_{p_k}$线性无关（由Rank$(A)$=$m$保证）。令$B = (a_{p_1}, \ldots, a_{p_m})$，则$B\begin{pmatrix} x_{p_1} \\ \vdots \\ x_{p_m} \end{pmatrix} = b$，而且$x_i = 0, \forall i \notin \{p_1, \ldots, p_m\}$，故$x$是可行基解。

□

若x非可行基解，包含两种情况：$x_N \neq 0$或者B不可逆。

4. 证明

$$x = (x_{B_1}, \ldots, x_{B_{r-1}}, 0, x_{B_{r+1}}, \ldots, x_{B_m}, 0, \ldots, x_p, \ldots, 0)^T$$

是可行基解。

证明. 只需要证明A中对应的列$a_{B_1}, \ldots, a_{B_{r-1}}, a_{B_{r+1}}, \ldots, a_{B_m}, a_p$线性无关。

若它们线性相关，设$a_p = \lambda_1 a_{B_1} + \ldots + \lambda_{r-1} a_{B_{r-1}} + \lambda_{r+1} a_{B_{r+1}} + \ldots + \lambda_m a_{B_m}$。由$B^{-1}a_p = \begin{pmatrix} y_{1p} \\ \vdots \\ y_{mp} \end{pmatrix}$且$y_{rp} > 0$，得到$a_p = \sum_{i=1}^m y_{ip} a_{B_i}$，综合以上两式得到

$$(\lambda_1 - y_{ip}) a_{B_1} + \ldots + (\lambda_{r-1} - y_{r-1,p}) a_{B_{r-1}} - y_{rp} a_{B_r}$$

$$+(\lambda_{r+1} - y_{r+1,p}) a_{B_{r+1}} + \ldots + (\lambda_m - y_{mp}) a_{B_m} = 0$$

其中$y_{rp} \neq 0$，与B可逆矛盾。 □

5. 先列出单纯形法的计算步骤，然后求解线性规划问题：

$$\begin{aligned} \min\ & -4x_1 - x_2 \\ \text{s.t.}\ & -x_1 + 2x_2 \leq 4 \\ & 2x_1 + 3x_2 \\ & x_1 - x_2 \leq 3 \\ & x_1, x_2 \geq 0 \end{aligned}$$

解 最优解$(\frac{21}{5}, \frac{6}{5})^T$，最优值$-18$，过程略。

6. 证明在执行主元消去法前后两个不同可行基下的判别系数和目标函数值有如下关系

$$\begin{aligned} (z_j - c_j)^{new} =& (z_j - c_j) - \frac{y_{rj}}{y_{rp}} (z_p - c_p) \\ (c_B^T B^{-1} b)^{new} =& c_B^T B^{-1} b - \frac{\bar{b}_r}{y_{rp}} (z_p - c_p) \end{aligned}$$

证明. $B^{new} = (a_{B_1}, \ldots, a_{B_{r-1}}, a_p, a_{B_{r+1}}, \ldots, a_{B_m})$，于是

$$B^{-1} B^{new} = (e_1, \ldots, e_{r-1}, y_p, e_{r+1}, \ldots, e_m)$$

$$\begin{aligned}(z_j-c_j)^{new}-(z_j-c_j)=&(c_B^{new})^T(B^{new})^{-1}a_j-c_B^TB^{-1}a_j\\=&((c_B^{new})^T-c_B^T)(B^{new})^{-1}a_j\\&+c_B^T((B^{new})^{-1}-B^{-1})a_j\\=&(0,\dots,0,c_p-c_{B_r},0,\dots,0)(B^{new})^{-1}a_j\\&+c_B^T(0,\dots,0,e_r-y_p,0,\dots,0)(B^{new})^{-1}a_j\\=&(0,\dots,0,c_p-c_{B_r},0,\dots,0)(B^{new})^{-1}a_j\\&+(0,\dots,0,c_{B_r}-c_B^Ty_p,0,\dots,0)(B^{new})^{-1}a_j\end{aligned}$$

注意到$(B^{new})^{-1}$的第r行与B^{-1}的第r行只相差$\frac{\det B}{\det B^{new}}$倍（伴随矩阵），即

$$(B^{new})^{-1}_r=\frac{\det B}{\det B^{new}}(B^{-1})_r$$

而

$$\frac{\det B}{\det B^{new}}=\frac{\det B^{-1}\det B}{\det B^{-1}\det B^{new}}=\frac{\det B^{-1}B}{\det B^{-1}B^{new}}=\frac{1}{y_{rp}}$$

故

$$\begin{aligned}=&(c_p-c_{B_r})\frac{1}{y_{rp}}(B^{-1}_{r1},\dots,B^{-1}_{rm})a_j\\&+(c_{B_r}-c_B^Ty_p)\frac{1}{y_{rp}}(B^{-1}_{r1},\dots,B^{-1}_{rm})a_j\\=&(c_p-z_p)\frac{y_{rj}}{y_{rp}}\end{aligned}$$

而$(c_B^TB^{-1}b)^{new}=c^Tx^{new}$其中$x_{B_r}=0,x_p=\Delta=\frac{\bar b_r}{y_{rp}}$，于是由讲义式(19)知$c^Tx^{new}=z_0+(c_p-z_p)\Delta=c_B^TB^{-1}b-\frac{\bar b_r}{y_{rp}}(z_p-c_p)$. □

7. 证明对偶定理。

证明.

$$(LP)\quad\begin{array}{ll}\min & c^Tx\\ \text{s.t.} & Ax\ge b\\ & x\ge 0\end{array}\qquad\qquad (DP)\quad\begin{array}{ll}\max & b^Tw\\ \text{s.t.} & A^Tw\le c\\ & w\ge 0\end{array}$$

先将(LP)标准化

$$(LP')\quad\begin{array}{ll}\min & c^Tx\\ \text{s.t.} & (A\ \ -I)\begin{pmatrix}x\\ y\end{pmatrix}=b\\ & x,y\ge 0\end{array}$$

令$\begin{pmatrix} x^* \\ y^* \end{pmatrix}$为$(LP')$的最优解，并考虑$\begin{pmatrix} x^* \\ y^* \end{pmatrix}$对应的可行基矩阵分解$B, N$。对任意$j \in N$，有$c_B^T B^{-1} a_j - c_j \leq 0$；对任意$j \in B$有$c_B^T B^{-1} a_j = c_j$。（对$j > n$，记$c_j = 0,\ a_j = -e_{j-n}$）总之有

$$c_B^T B^{-1} a_j - c_j \leq 0$$

令$w = (c_B^T B^{-1})^T$，则对$1 \leq j \leq n$

$$w^T a_j - c_j \leq 0 \Leftrightarrow a_j^T w \leq c_j$$

$$\Rightarrow A^T w \leq c$$

对$n + 1 \leq j \leq m$

$$w^T a_j \leq 0 \Leftrightarrow -w_j \leq 0$$

$$\Rightarrow w \geq 0$$

因此w是(DP)的一个可行解。又

$$b^T w = w^T b = c_B^T B^{-1} b = (c^T\ 0) \begin{pmatrix} x^* \\ y^* \end{pmatrix} = c^T x^*$$

由弱对偶定理知w为(DP)的最优解。

由(DP)最优解推出(LP)最优解的过程类似，不再赘述。 $\square$

Exercise 2.1: 依据对偶理论，给出最短路径问题的一种求解方法。

最短路径问题的对偶问题为：

$$
\begin{aligned}
&\max\ y_s - y_t \\
&\text{s.t. } y_i - y_j \le c_{ij}\ \forall (i,j) \in E
\end{aligned}
$$

对于基变量x_{ij}，$y_i - y_j = c_{ij}$，即$\sigma_{ij} = y_i - y_j - c_{ij} = 0$.

对于非基变量x_{ij}而言，若$\sigma_{ij} = y_i - y_j - c_{ij} \le 0$. 则已经对偶可行；若$\sigma_{ij} = y_i - y_j - c_{ij} > 0$, 则非对偶可行，需要引进基。

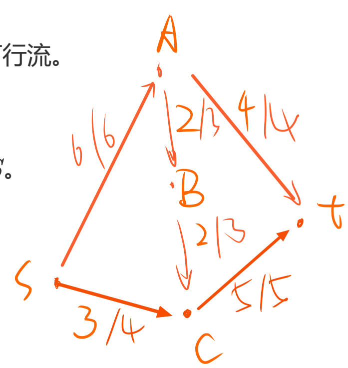
求解步骤类似于单纯形法求解运输模型：

1. 选取一组路径，作为初始可行基解。
2. 检验当前解是否可以改进，如果可改进，则引入一个非基变量进行步3，否则停止。
3. 当把步2中挑选的变量引进时，确定哪个路径应当由基解中退出。
4. 调整其他基本路径的流量（满足可行性），返回步2。

Exercise 2.2: 证明最大流最小割定理：任意一个流网络的最大流量等于该网络的最小割的容量。

由Exercise 2.3知若可行流x为最大流，则不存在关于x的流扩充路，否则可以找到更大的可行流。

已知不存在关于x的流扩充路，寻找截集(S,T)，使得$x(s)=U(S,T)$：



令$s\in S$，若$i\in S$，且$x(i,j)<u(i,j)$，则令$j\in S$；若$i\in S$，且$\cancel{x(i,j)}>0$，则令$j\in S$。

$x(j,i)$

因为不存在关于x的流扩充路，所以$t\notin S$。令$T=V/S$，

$$
x(i,j) = \begin{cases} u(i,j) & (i,j)\in(S,T) \\ 0 & (i,j)\in(T,S) \end{cases}
$$

$S=\{s,C,A,B\}$

$f^*=9$

则有$x(s)=\tilde{U}(S,T)$。

$T=\{t\}$

下证$\tilde{U}(S,T)$是最小割：

由割集定义，$x(s)\le U(S,T)$，$U(S,T)$为V的分割。所以$\max x(s)\le \min U(S,T)$，又有$x(s)=\tilde{U}(S,T)$，所以$\tilde{U}(S,T)$为最小割。

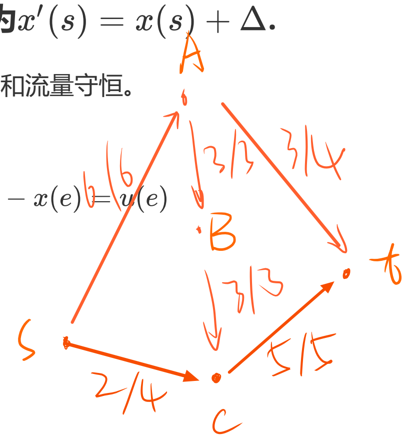
综上，任一流网络的最大流量等于该网络最小割的容量。

Exercise 2.3: 验证(63)式给出的x'是原网络$\mathcal{N}$的可行流，并且其流值为$x'(s)=x(s)+\Delta$.

因为π之外的边无流量变化，所以为验证新的流是可行流，需要对π验证两点：流量限制和流量守恒。

流量限制：对于由Rule1生成的边：

$$
0\le x(e)\le x(e)+\Delta\le x(e)+min\{u(e)-x(e)|e\in\pi\}\le x(e)+u(e)-x(e)=u(e)
$$



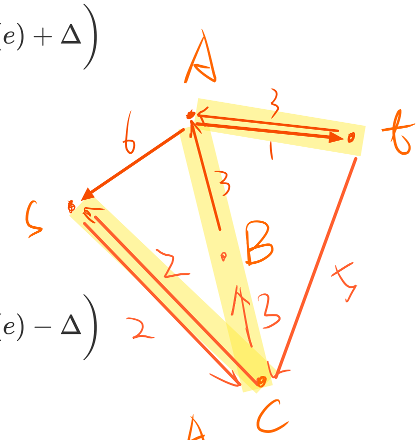
对于由Rule2生成的边：

$$0 = x(e) - x(e) \le x(e) - min\{x(e)|e \in \pi\} \le x(e) - \Delta \le x(e) \le u(e)$$

流量守恒：分类讨论。$\forall v \in \pi, v \neq s, v \neq t$，考虑生成$v$在$\pi$上入边和出边的规则。 $s \xrightarrow{1} C \xrightarrow{2} B \xrightarrow{2} A \xrightarrow{1} t$

入边和出边均由Rule1生成

$$\begin{aligned}\sum_{e \in Out(v)} x'(e) - \sum_{e \in In(v)} x'(e) &= \left(\sum_{e \in Out(v)} x(e) + \Delta\right) - \left(\sum_{e \in In(v)} x(e) + \Delta\right) \\ &= \sum_{e \in Out(v)} x(e) - \sum_{e \in In(v)} x(e) \\ &= 0\end{aligned}$$

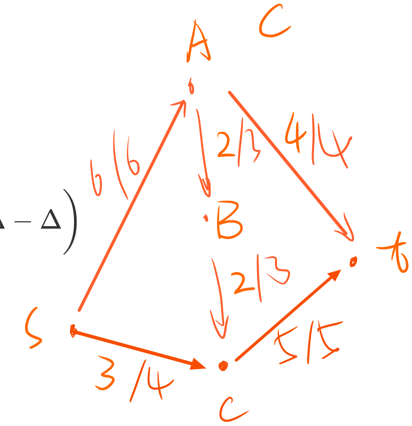


入边和出边均有Rule2生成 B.

$$\begin{aligned}\sum_{e \in Out(v)} x'(e) - \sum_{e \in In(v)} x'(e) &= \left(\sum_{e \in Out(v)} x(e) - \Delta\right) - \left(\sum_{e \in In(v)} x(e) - \Delta\right) \\ &= \sum_{e \in Out(v)} x(e) - \sum_{e \in In(v)} x(e) \\ &= 0\end{aligned}$$

说明在N中是入边

入边由Rule1生成，出边由Rule2生成 C



$$\begin{aligned}\sum_{e \in Out(v)} x'(e) - \sum_{e \in In(v)} x'(e) &= \sum_{e \in Out(v)} x(e) - \left(\sum_{e \in In(v)} x(e) + \Delta - \Delta\right) \\ &= \sum_{e \in Out(v)} x(e) - \sum_{e \in In(v)} x(e) \\ &= 0\end{aligned}$$

入边由Rule2生成，出边由Rule1生成 A

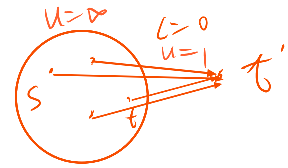
$$\begin{aligned}\sum_{e \in Out(v)} x'(e) - \sum_{e \in In(v)} x'(e) &= \left(\sum_{e \in Out(v)} x(e) + \Delta - \Delta\right) - \sum_{e \in In(v)} x(e) \\ &= \sum_{e \in Out(v)} x(e) - \sum_{e \in In(v)} x(e) \\ &= 0\end{aligned}$$

所以新的流是可行流，下面验证$x'(s) = x(s) + \Delta$.

$$\begin{aligned}x'(s) &= \sum_{e \in Out(s)} x'(e) - \sum_{e \in In(s)} x'(e) \\ &= \left(\sum_{e \in Out(s)} x(e) + \Delta\right) - \sum_{e \in In(s)} x(e) \text{ (Rule1)} \\ &\text{ or } \sum_{e \in Out(s)} x(e) - \left(\sum_{e \in In(s)} x(e) - \Delta\right) \text{ (Rule 2)} \\ &= x(s) + \Delta\end{aligned}$$

Exercise 2.4: 通过构造说明最小成本流问题作为其特殊情况包含：最短路问题和最大流问题。

最短路问题：考虑最短路问题网络$\mathcal{N}=(G,s,t,c)$, 引入哑元终点$\tilde{t}$和新边



$$(v,\tilde{t})\ s.t.\ c(v,\tilde{t})=0\ u(v,\tilde{t})=1\ \forall v\in V$$

则最短路径问题的求解化为求解流值$f^*=n$的最大流问题，其中n为顶点数。 $\pi=(s, v_{i_1}, v_{i_k}, t)$

因为连接$\tilde{t}$与其他顶点的边容量均为1，且流值为$n=|V|$，所以由s到$\tilde{t}$的最小成本流使得s到各个顶点的成本最小，又因为原网络中各边容量为∞，所以为最短路径。

最大流问题：引入哑元始点$\tilde{s}$，添加新边： $N(G, s, t, u)$

$$(\tilde{s},s)\ s.t.\ c(\tilde{s},s)=0\ u(\tilde{s},s)=\infty$$
$$(\tilde{s},t)\ s.t.\ c(\tilde{s},t)=1\ u(\tilde{s},t)=\infty$$

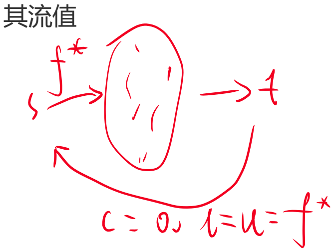
令原网络中$c(e)=0,\forall e\in E$，令流值$f^*=\sum_{e\in E}u(e)$, 源点为$\tilde{s}$, 汇点为t，则最大流问题转化为求解$\tilde{s}$到t的最小成本流问题。

这是因为为使成本最小，尽量避免使流量直接从$\tilde{s}$到t，而是流经s到t，在原网络流量达到最大流后剩下的流量从$\tilde{s}$到t。

Exercise 2.5: 证明最小成本循环流问题与最小成本流问题具有等价的模型化能力。

最小成本流⇒最小成本循环流：考虑网络$\mathcal{N}=(G,s,t,c,u)$上的最小成本流问题，其流值

$$f^*=\sum_{e\in Out(s)}x(e)-\sum_{e\in In(s)}x(e)$$



添加汇点到源点的新边(t,s)并且令

$$\tilde{c}(t,s)=0,\tilde{l}(t,s)=\tilde{u}(t,s)=f^*.$$

在原网络中，令

$$\tilde{c}(e)=c(e),\tilde{l}(e)=0,\tilde{u}(e)=u(e)$$

则转化为最小成本循环流问题。

最小成本循环流⇒最小成本流：考虑网络$\mathcal{N}=(G,c,l,u)$上的最小成本循环流问题，$\forall v\in V$，令

$$f^*(v)=\sum_{e\in In(v)}l(e)-\sum_{e\in Out(v)}l(e)$$

对于$f^*(v)>0$，则令v为源点，$S=\{v|f^*(v)>0\}$; 对于$f^*(v)<0$，则令v为汇点，$T=\{v|f^*(v)<0\}$.

引入哑元始点$\tilde{s}$和哑元终点$\tilde{t}$，并加入新边

$$(\tilde{s},v)\ s.t.\ \tilde{c}(\tilde{s},v)=0\ \tilde{u}(\tilde{s},v)=f^*(v)\ \forall v\in S$$

$$(v,\tilde{t})\ s.t.\ \tilde{c}(v,\tilde{t})=0\ \tilde{u}(v,\tilde{t})=-f^*(v)\forall v\in T$$

在原网络中，令

$$\tilde{u}(e) = u(e) - l(e)\ \tilde{c}(e) = c(e)$$

$$f^*(\tilde{s}) = \sum_{v \in S} f^*(v)$$

$$\tilde{x}(e) = x(e) - l(e)$$

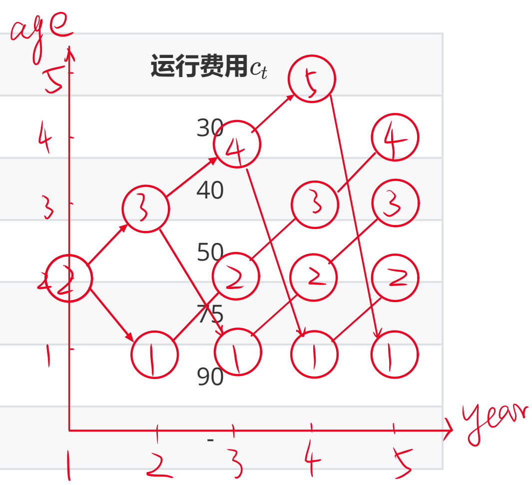
则网络$\mathcal{N} = (G, c, l, u)$上的最小成本循环流问题转化为$\tilde{\mathcal{N}} = (\tilde{G}, \tilde{s}, \tilde{t}, \tilde{c}, \tilde{U})$上的最小成本流问题，$\mathcal{N} = (G, c, l, u)$的最小成本 = $\tilde{\mathcal{N}} = (\tilde{G}, \tilde{s}, \tilde{t}, \tilde{c}, \tilde{U})$的最小成本 + $\sum_{e \in E} c(e) l(e)$。

Exercise 3.1 设现有一台2年龄的设备，另规定5年龄的设备必须更换。在规划期购置新设备的成本分别是

$$(p_1, p_2, p_3, p_4, p_5) = (100, 105, 110, 115, 120)$$

试构建如下设备更新的动态规划模型并求其最优更新策略。

设备年龄t	残值v_t	运行费用c_t
0	-	30
1	50	40
2	25	50
3	10	75
4	5	90
5	2	-



$$f_i(t) = max \begin{cases} -c_t + f_{i+1}(t+1) & \text{K} \\ -c_0 + v_t - p_i + f_{i+1}(1) & \text{R} \end{cases}$$

$$f_5(4) = max \begin{cases} -c_4 + f_6(5) & = -90 + 0 & = -90 & \text{K} \\ -c_0 + v_4 - p_5 + f_6(1) & = -30 + 5 - 120 + 0 & = -145 & \text{R} \end{cases}$$

$$f_5(3) = max \begin{cases} -c_3 + f_6(4) & = -75 + 0 & = -75 & \text{K} \\ -c_0 + v_3 - p_5 + f_6(1) & = -30 + 10 - 120 + 0 & = -140 & \text{R} \end{cases}$$

$$f_5(2) = max \begin{cases} -c_2 + f_6(3) & = -50 + 0 & = -50 & \text{K} \\ -c_0 + v_2 - p_5 + f_6(1) & = -30 + 25 - 120 + 0 & = -125 & \text{R} \end{cases}$$

$$f_5(1) = max \begin{cases} -c_1 + f_6(2) & = -40 + 0 & = -40 & \text{K} \\ -c_0 + v_1 - p_5 + f_6(1) & = -30 + 50 - 120 + 0 & = -100 & \text{R} \end{cases}$$

$$f_4(5) = -c_0 + v_5 - p_4 + f_5(1) = -30 + 2 - 115 + (-40) = -183 \quad R$$

$$f_4(3)=max\begin{cases}-c_3+f_5(4) & =-75+(-90) & =-165 & \mathrm{K}\\ -c_0+v_3-p_4+f_5(1) & =-30+10-115+(-40) & =-175 & \mathrm{R}\end{cases}$$

$$f_4(2)=max\begin{cases}-c_2+f_5(3) & =-50+(-75) & =-125 & \mathrm{K}\\ -c_0+v_2-p_4+f_5(1) & =-30+25-115+(-40) & =-160 & \mathrm{R}\end{cases}$$

$$f_4(1)=max\begin{cases}-c_1+f_5(2) & =-40+(-50) & =-90 & \mathrm{K}\\ -c_0+v_1-p_4+f_5(1) & =-30+50-115+(-40) & =-135 & \mathrm{R}\end{cases}$$

$$f_3(4)=max\begin{cases}-c_4+f_4(5) & =-90+(-183) & =-273 & \mathrm{K}\\ -c_0+v_4-p_3+f_4(1) & =-30+5-110+(-90) & =-225 & \mathrm{R}\end{cases}$$

$$f_3(2)=max\begin{cases}-c_2+f_4(3) & =-50+(-165) & =-215 & \mathrm{K}\\ -c_0+v_2-p_3+f_4(1) & =-30+25-110+(-90) & =-205 & \mathrm{R}\end{cases}$$

$$f_3(1)=max\begin{cases}-c_1+f_4(2) & =-40+(-125) & =-165 & \mathrm{K}\\ -c_0+v_1-p_3+f_4(1) & =-30+50-110+(-90) & =-180 & \mathrm{R}\end{cases}$$

$$f_2(3)=max\begin{cases}-c_3+f_3(4) & =-75+(-225) & =-300 & \mathrm{K}\\ -c_0+v_3-p_2+f_3(1) & =-30+10-105+(-165) & =-290 & \mathrm{R}\end{cases}$$

$$f_2(1)=max\begin{cases}-c_1+f_3(2) & =-40+(-205) & =-245 & \mathrm{K}\\ -c_0+v_1-p_2+f_3(1) & =-30+50-105+(-165) & =-250 & \mathrm{R}\end{cases}$$

$$f_1(2)=max\begin{cases}-c_2+f_2(3) & =-50+(-290) & =-340 & \mathrm{K}\\ -c_0+v_2-p_1+f_2(1) & =-30+25-100+(-245) & =-350 & \mathrm{R}\end{cases}$$

若第六年不出售设备，则最优更新策略为KRKKK，收益为-340；若考虑第六年卖出设备，则最优策略不变，收益为-335。

Exercise 1.写出基于Wolfe-Powell准则的非精确一维搜索算法中插值多项式$p^{(1)}(t), p^{(2)}(t)$的具体表达式。

设$p^{(1)}(t) = At^2 + Bt + C$，则有

$$\begin{cases} A\alpha_1^2 + B\alpha_1 + C & = \varphi_1 \\ 2A\alpha_1 + B & = \varphi_1' \\ A\alpha^2 + B\alpha + C & = \varphi \end{cases}$$

解得

$$\begin{cases} A = \frac{\varphi - \varphi_1 + \varphi_1'(\alpha_1 - \alpha)}{(\alpha - \alpha_1)^2} \\ B = \frac{\varphi_1'(\alpha^2 - \alpha_1)^2 + 2\alpha_1(\varphi_1 - \varphi)}{(\alpha - \alpha_1)^2} \\ C = \frac{\varphi_1\alpha(\alpha - 2\alpha_1) + \varphi_1'\alpha\alpha_1(\alpha_1 - \alpha) + \varphi\alpha_1^2}{(\alpha_1 - \alpha)^2} \end{cases}$$

设$p^{(2)}(t) = \tilde{A}t^2 + \tilde{B}t + \tilde{C}$，则有

$$\begin{cases} \tilde{A}\alpha^2 + \tilde{B}\alpha + C & = \varphi \\ 2\tilde{A}\alpha_1 + \tilde{B} & = \varphi_1' \\ 2\tilde{A}\alpha + \tilde{B} & = \varphi' \end{cases}$$

解得

$$\begin{cases} \tilde{A} = \frac{\varphi' - \varphi_1'}{2\alpha - 2\alpha_1} \\ \tilde{B} = \frac{\varphi_1'\alpha - \varphi'\alpha_1}{\alpha - \alpha_1} \\ \tilde{C} = \frac{\varphi(2\alpha - 2\alpha_1) + \varphi'(2\alpha\alpha_1 - \alpha^2) - \varphi_1'\alpha^2}{2\alpha - 2\alpha_1} \end{cases}$$

Exercise 2.证明基于Goldstein准则的非精确一维搜索算法的全局收敛性。

Goldstein准则：

$$\begin{cases} \varphi(\alpha) & \leq \varphi(0) + \rho\alpha\varphi'(0) & (92) \\ \varphi(\alpha) & \geq \varphi(0) + (1-\rho)\alpha\varphi'(0) & (93) \end{cases}$$

设$\forall k, g^{(k)} = \nabla f(x^k) \neq 0$和$f(x^{(k)})$有下界，则$f(x^{(k)}) - f(x^{(k+1)}) \to 0$，由(92)式，$-g^{(k)}s^{(k)} \to 0$。

（反证法）若$g^{(k)} \to 0$不成立，则$\exists \epsilon > 0$和子列$\{x^{(k)}\}_{k \in K} s.t. \|g^{(k)}\| \geq \epsilon$从而由

$$-g^{(k)^\top} s^{(k)} = \|g^{(k)}\| \|s^{(k)}\| \cos\theta_k \geq \epsilon \|s^{(k)}\| \sin\mu$$

以及$\theta_k \leq \frac{\pi}{2} - \mu, \forall k$得$\|s^{(k)}\| \to 0.$

又有

$$f(x^{(k)} + s^{(k)}) = f(x^{(k)}) + \nabla f(x^{(k)})^\top s^{(k)} + o(\|s^{(k)}\|)(\text{泰勒展开})$$

得

$$\lim_{k\to\infty}\frac{f(x^{(k)})-f(x^{(k+1)})}{-\nabla f^{(k)}{}^\top s^{(k)}}=1$$

由(93)式得

$$\frac{f(x^{(k)})-f(x^{(k+1)})}{-\nabla f^{(k)}{}^\top s^{(k)}}\le 1-\rho<1$$

矛盾！

所以$g^{(k)}=\nabla f(x^{(k)})\to 0, k\to\infty$.

Exercise 3.将非线性方程组求根$F(x)=0$的牛顿迭代，用于求最优化问题$\min_{x\in\mathbb{R}^n} f(x)$，给出相应的迭代格式并说明理由。

牛顿迭代法的原理是取函数$F(x)$在$x^{(k)}$处泰勒展开的线性部分，作为函数的一阶近似，令其等于0并得到迭代格式，即：

$$F(x^{(k)})+DF(x^{(k)})(x-x^{(k)})=0$$

其中$DF(x^{(k)})$为F在$x^{(k)}$的Jacobian，其形式为

$$DF(x)=\begin{bmatrix}\frac{\partial F_1(x)}{\partial x_1} & \frac{\partial F_1(x)}{\partial x_2} & \dots & \frac{\partial F_1(x)}{\partial x_n}\\ \frac{\partial F_2(x)}{\partial x_1} & \frac{\partial F_2(x)}{\partial x_2} & \dots & \frac{\partial F_n(x)}{\partial x_n}\\ \vdots & \vdots & \ddots & \vdots\\ \frac{\partial F_n(x)}{\partial x_1} & \frac{\partial F_n(x)}{\partial x_2} & \dots & \frac{\partial F_n(x)}{\partial x_n}\end{bmatrix}$$

解得迭代格式为

$$x^{(k+1)}=x^{(k)}-DF(x^{(k)})^{-1}F(x^{(k)})$$

同理，对于无约束最优化问题$min_{x\in\mathbb{R}^n f(x)}$, 类似地求解$\nabla f(x)=0$即可，得到迭代格式为

$$x^{(k+1)}=x^{(k)}-(D(\nabla f(x)))^{-1}\nabla f(x^{(k)})=x^{(k)}-(\nabla^2 f(x))^{-1}\nabla f(x^{(k)})$$

Exercise 4.证明对称秩一牛顿法具有遗传性和二次终止性

对于二次函数$f(x)=\frac{1}{2}x^\top Gx+c^\top x$,

$$\nabla f(x)=Gx+c,\nabla^2 f(x)=G'$$

由牛顿法$y^{(k)}=G_k s^{(k)}$, 正割条件为$H_{k+1}y^{(k)}=s^{(k)}$

遗传性：使用归纳法。

1. $k=1$时，由正割条件，$H_1 y^{(0)}=s^{(0)}$。

2. 假设遗传性对于k成立，即$H_k y^{(l)} = s^{(l)}, l = 0, 1, \ldots, k-1$.
3. 对于$k+1$, 只须证明$l < k$的情况（$l = k$时，由正割条件直接可得结论成立），即

$$H_{k+1} y^{(l)} = s^{(l)}, l = 0, 1, \ldots, k-1$$

 由对称秩一校正公式

$$H_{k+1} y^{(l)} = H_k y^{(l)} + \frac{(s^{(k)} - H_k y^{(k)})(s^{(k)} - H_k y^{(k)})^\top y^{(l)}}{(s^{(k)} - H_k y^{(k)})^\top y^{(k)}}$$

 其中

$$\begin{aligned}
(s^{(k)} - H_k y^{(k)})^\top y^{(l)} &= s^{(k)\top} y^{(l)} - y^{(k)\top} H_k y^{(l)} \\
&= s^{(k)\top} y^{(l)} - y^{(k)\top} s^{(l)} \\
&= s^{(k)\top} G s^{(l)} - s^{(k)\top} G s^{(l)} \\
&= 0
\end{aligned}$$

 故$H_{k+1} y^{(l)} = H_k y^{(l)} = s^{(l)}, l = 0, 1, \ldots, k-1$.

二次终止性：（这里需要假定$s_0, s_1, \ldots, s_{n-1}$线性无关）

由遗传性知

$$\begin{gathered}
H_n y^{(l)} = s^{(l)}, l = 0, 1, \ldots, n-1 \\
\Downarrow \\
H_n G s^{(l)} = s^{(l)}, l = 0, 1, \ldots, n-1 \\
\Downarrow \\
(H_n G - I) s^{(l)} = 0, l = 0, 1, \ldots, n-1
\end{gathered}$$

因为$s_0, s_1, \ldots, s_{n-1}$线性无关，所以$(s_0, s_1, \ldots, s_{n-1})$可逆，所以

$$H_n G - I = 0 \Rightarrow H_n = G^{-1} = (\nabla^2 f(x))^{-1}$$

由迭代格式

$$x^{(n+1)} = x^{(n)} - H_n \nabla f(x^{(n)}) = x^{(n)} - G^{-1} \nabla f(x^{(n)})$$

又因为

$$\begin{gathered}
\nabla f(x^{(n+1)}) - \nabla f(x^{(n)}) = G(x^{(n+1)} - x^{(n)}) \\
\Downarrow \\
x^{(n+1)} = x^{(n)} - G^{-1} \nabla f(x^{(n)}) + G^{-1} \nabla f(x^{(n+1)})
\end{gathered}$$

所以

$$G^{-1} \nabla f(x^{(n+1)}) = 0 \Rightarrow \nabla f(x^{(n+1)}) = 0$$

即有限步终止且$H_n = [\nabla^2 f(x^*)]^{-1}$.

Exercise 5.利用秩一校正的求逆公式（sherman-Morrison定理），由$H_{k+1}^{(DFP)}$推导$B_{k+1}^{(DFP)}$.

$$(A+uv^\top)^{-1}=A^{-1}-\frac{A^\top uv^\top A^{-1}}{1+v^\top A^{-1}u}$$
$$H_{k+1}^{(DFP)}=H_k+\frac{s^{(k)}s^{(k)}{}^\top}{s^{(k)}{}^\top y^{(k)}}-\frac{H_ky^{(k)}y^{(k)}{}^\top H_k}{y^{(k)}{}^\top H_ky^{(k)}}$$
$$B_{k+1}^{(DFP)}=B_k+(1+\frac{s^{(k)}{}^\top B_ks^{(k)}}{y^{(k)}{}^\top s^{(k)}})\frac{y^{(k)}y^{(k)}{}^\top}{y^{(k)}{}^\top s^{(k)}}-\frac{B_ks^{(k)}y^{(k)}{}^\top+y^{(k)}s^{(k)}{}^\top B_k}{y^{(k)}{}^\top s^{(k)}}$$

为方便书写，忽视所有的角标k, 并令$M=H+\frac{ss^\top}{s^\top y}$，则有

$$M^{-1}=H^{-1}-\frac{H^{-1}ss^\top H^{-1}}{s^\top y+s^\top Bs}=B-\frac{Bss^\top B}{s^\top y+s^\top Bs}\quad(5.1)$$

$$B_{k+1}^{(DFP)}=(H_{k+1}^{(DFP)})^{-1}=M^{-1}+\frac{M^{-1}Hyy^\top HM^{-1}}{y^\top Hy-y^\top HM^{-1}Hy}\quad(5.2)$$

将(5.1)代入(5.2)第二个等号右边第二项并展开，消去分子分母常数和$BH=I$得

$$\begin{aligned}\frac{M^{-1}Hyy^\top HM^{-1}}{y^\top Hy-y^\top HM^{-1}Hy}&=\frac{yy^\top(s^\top y+s^\top Bs)}{y^\top ss^\top y}-\frac{ys^\top B}{s^\top y}-\frac{Bsy^\top}{y^\top s}+\frac{Bss^\top B}{s^\top y+s^\top Bs}\\&=(1+\frac{s^\top Bs}{y^\top s})\frac{yy^\top}{y^\top s}-\frac{ys^\top B+Bsy^\top}{y^\top s}+\frac{Bss^\top B}{s^\top y+s^\top Bs}\quad(5.3)\end{aligned}$$

将(5.1), (5.3)代入(5.2)即可得

$$B_{k+1}^{(DFP)}=B_k+(1+\frac{s^{(k)}{}^\top B_ks^{(k)}}{y^{(k)}{}^\top s^{(k)}})\frac{y^{(k)}y^{(k)}{}^\top}{y^{(k)}{}^\top s^{(k)}}-\frac{B_ks^{(k)}y^{(k)}{}^\top+y^{(k)}s^{(k)}{}^\top B_k}{y^{(k)}{}^\top s^{(k)}}$$

Exercise 6.共轭梯度法性质定理：设目标函数$f(x)=\frac{1}{2}x^\top Gx+c^\top x$,则采用精确一维搜索的共轭梯度法经$m\le n$步迭代后终止，且对所有的$1\le k\le m$成立下列关系：

$$\begin{aligned}&d^{(k)}{}^\top Gd^{(j)}=0,j=0,1,\dots,k-1&&(6.1)\\&g^{(k)}{}^\top g^{(j)}=0,j=0,1,\dots,k-1&&(6.2)\\&d^{(k)}{}^\top g^{(k)}=-g^{(k)}{}^\top g^{(k)}&&(6.3)\\&span\{g^{(0)},g^{(1)},\dots,g^{(k)}\}=span\{g^{(0)},Gg^{(0)},\dots,G^kg^{(0)}\}&&(6.4)\\&span\{d^{(0)},d^{(1)},\dots,d^{(k)}\}=span\{d^{(0)},Gg^{(0)},\dots,G^kg^{(0)}\}&&(6.5)\end{aligned}$$

共轭梯度法步骤中得到的的条件:

$$
\begin{aligned}
&g^{(k+1)^\top} d^{(k)} = 0 && (\text{由精确一维搜索})(6.6)\\
&\alpha_k = \frac{g^{(k)^\top} g^{(k)}}{d^{(k)^\top} G d^{(k)}} && (\text{由精确一维搜索})(6.7)\\
&g^{(k+1)} = g^{(k)} + \alpha_k G d^{(k)} && (\nabla f(x^{(k+1)})\text{直接展开})(6.8)\\
&\beta_k = \frac{g^{(k+1)^\top} g^{(k+1)}}{g^{(k)^\top} g^{(k)}} && (6.9)\\
&d^{(k+1)} = -g^{(k+1)} + \beta_k d^{(k)} && (6.10)
\end{aligned}
$$

证明(6.3):

$$
\begin{aligned}
d^{(k)^\top} g^{(k)} &= -g^{(k)^\top} g^{(k)} + \beta_k d^{(k-1)^\top} g^{(k)} && (by(6.10))\\
&= -g^{(k)^\top} g^{(k)} + 0 && (by(6.6))\\
&= -g^{(k)^\top} g^{(k)}
\end{aligned}
$$

证明(6.1)**与**(6.2):

1. k = 1时直接验证可得结论成立
2. 假设(6.1)与(6.2)对k成立
3. (6.8)式左右两边转置后同右乘$g^{(j)}$得

$$
\begin{aligned}
g^{(k+1)} g^{(j)} &= g^{(k)^\top} g^{(j)} - \alpha_k d^{(k)^\top} G g^{(j)}\\
&= g^{(k)^\top} g^{(j)} - \alpha_k d^{(k)^\top} G(d^{(j)} - \beta_{j-1} d^{(j-1)})\\
&= g^{(k)^\top} g^{(j)} - \alpha_k d^{(k)^\top} G d^{(j)}
\end{aligned}
$$

$j = k$时，将(6.7)式代入即可得上式为0；$j < k$时，由归纳假设得上式为0。

综上，(6.2)成立。

4. 由(6.10)式，

$$
\begin{aligned}
d^{(k+1)^\top} G d^{(j)} &= (-g^{(k+1)} + \beta_k d^{(k)}) G d^{(j)}\\
&= -g^{(k+1)} G d^{(j)} + \beta_k d^{(k)} G d^{(j)}\\
&= g^{(k+1)^\top} (g^{(j)} - g^{(j+1)})/\alpha_k + \beta_k d^{(k)} G d^{(j)}
\end{aligned}
$$

$j = k$时，由$(6.2), (6.7), (6.9)$得上式为0；$j < k$时，由归纳假设得上式为0.

综上，(6.1)成立。

证明(6.4)

由(6.10)式知，存在可逆方阵

$$
Q = \begin{bmatrix}
-1 & B_0 & 0 & \dots & 0 & 0\\
0 & -1 & B_1 & \dots & 0 & 0\\
0 & 0 & -1 & \dots & 0 & 0\\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\
0 & 0 & 0 & \dots & -1 & B_{k-1}\\
0 & 0 & 0 & \dots & 0 & -1
\end{bmatrix}
$$

使得$(d^{(0)}, d^{(1)}, \dots, d^{(k)})Q = (g^{(0)}, g^{(1)}, \dots, g^{(k)})$，所以

$$span\{g^{(0)},g^{(1)},\dots,g^{(k)}\}=span\{d^{(0)},d^{(1)},\dots,d^{(k)}\}(6.11)$$

1. $k=0$时，直接由定义得结论成立。
2. 假设结论对k成立，即$span\{g^{(0)},g^{(1)},\dots,g^{(k)}\}=span\{g^{(0)},Gg^{(0)},\dots,G^kg^{(0)}\}$.
3. 对于$k+1$，由(6.8)式和归纳假设，
$$g_{k+1}\in span\{g^{(0)},Gg^{(0)},\dots,G^{k+1}g^{(0)}\}$$
$d^{(0)},d^{(1)},\dots,d^{(k)}$是一组共轭方向，由共轭方向法基本定理得
$$g^{(k+1)}\perp span\{d^{(0)},d^{(1)},\dots,d^{(k)}\}(6.12)$$
所以由$(6.11),(6.12)$和归纳假设
$$g^{(k+1)}\notin span\{d^{(0)},d^{(1)},\dots,d^{(k)}\}=span\{g^{(0)},Gg^{(0)},\dots,G^kg^{(0)}\}$$
所以结论对$k+1$成立，即
$$span\{g^{(0)},g^{(1)},\dots,g^{(k+1)}\}=span\{g^{(0)},Gg^{(0)},\dots,G^{k+1}g^{(0)}\}$$

(6.5)**式证明与**(6.4)**式同理。**

Exercise 8.证明折线法（信赖域方法）子问题模型的函数单调性。

$$\begin{aligned}s_C^{(k)}&=-\frac{\|g^{(k)}\|^2}{{g^{(k)}}^\top B_kg^{(k)}}g^{(k)}\\ s_N^{(k)}&=-B_k^{-1}g^{(k)}\\ x^{(k+1)}&=x^{(k)}+s_C^{(k)}+\lambda(s_N^{(k)}-s_C^{(k)})\end{aligned}$$

(1)证明沿着Cauchy点$x_C^{(k+1)}$和牛顿点$x_N^{(k+1)}$的连线，到$x^{(k)}$的距离单调增加.

$$\begin{aligned}L(\lambda)&=\|s_C^{(k)}+\lambda(s_N^{(k)}-s_C^{(k)})\|^2\\ &=(s_C^{(k)}+\lambda(s_N^{(k)}-s_C^{(k)}))^\top(s_C^{(k)}+\lambda(s_N^{(k)}-s_C^{(k)}))\\ &={s_C^{(k)}}^\top s_C^{(k)}+2\lambda {s_C^{(k)}}^\top(s_N^{(k)}-s_C^{(k)}))+\lambda^2(s_N^{(k)}-s_C^{(k)}))^\top(s_N^{(k)}-s_C^{(k)}))\\ L(\lambda)'&=2{s_C^{(k)}}^\top(s_N^{(k)}-s_C^{(k)}))+2\lambda(s_N^{(k)}-s_C^{(k)}))^\top(s_N^{(k)}-s_C^{(k)}))\\ &\geq 2{s_C^{(k)}}^\top(s_N^{(k)}-s_C^{(k)}))\\ &=2\frac{\|g^{(k)}\|^2}{{g^{(k)}}^\top B_kg^{(k)}}g^{(k)}B_k^{-1}g^{(k)}(1-\frac{\|g^{(k)}\|^2g^{(k)}}{{g^{(k)}}^\top B_kg^{(k)}B_k^{-1}g^{(k)}})\end{aligned}$$

其中

$$
\begin{aligned}
1-\frac{\|g^{(k)}\|^2 g^{(k)}}{g^{(k)\top} B_k g^{(k)} B_k^{-1} g^{(k)}} &= 1-\frac{\|g^{(k)}\|^4}{g^{(k)\top} B_k g^{(k)} g^{(k)\top} B_k^{-1} g^{(k)}} \\
&= 1-\frac{((\sqrt{B_k} g^{(k)})^\top ((\sqrt{B_k^{-1}} g^{(k)}))^2}{g^{(k)\top} B_k g^{(k)} g^{(k)\top} B_k^{-1} g^{(k)}} \\
&\geq 1-\frac{\|\sqrt{B_k} g^{(k)}\|^2 \|\sqrt{B_k^{-1}} g^{(k)}\|^2}{g^{(k)\top} B_k g^{(k)} g^{(k)\top} B_k^{-1} g^{(k)}} \\
&= 1-\frac{g^{(k)\top} B_k g^{(k)} g^{(k)\top} B_k^{-1} g^{(k)}}{g^{(k)\top} B_k g^{(k)} g^{(k)\top} B_k^{-1} g^{(k)}} \\
&= 0
\end{aligned}
$$

所以$L(\lambda)' \geq 0, \lambda \in [0,1]$，具有单调性。

(2)证明沿着Cauchy点$x_C^{(k+1)}$和牛顿点$x_N^{(k+1)}$的连线，子问题模型函数值单调减少。

$$
\begin{aligned}
h(\lambda) &= q^{(k)}(s_C^{(k)} + \lambda(s_N^{(k)} - s_C^{(k)})) \\
&= f(x^{(k)}) + g^{(k)\top}(s_C^{(k)} + \lambda(s_N^{(k)} - s_C^{(k)})) \\
&\quad + \frac{1}{2}(s_C^{(k)} + \lambda(s_N^{(k)} - s_C^{(k)}))^\top B_k(s_C^{(k)} + \lambda(s_N^{(k)} - s_C^{(k)})) \\
h(\lambda)' &= g^{(k)\top}(s_N^{(k)} - s_C^{(k)}) + s_C^{(k)\top} B_k(s_N^{(k)} - s_C^{(k)}) + \lambda(s_N^{(k)} - s_C^{(k)})^\top B_k(s_N^{(k)} - s_C^{(k)}) \\
&\leq g^{(k)\top}(s_N^{(k)} - s_C^{(k)}) + s_C^{(k)\top} B_k(s_N^{(k)} - s_C^{(k)}) + (s_N^{(k)} - s_C^{(k)})^\top B_k(s_N^{(k)} - s_C^{(k)}) \\
&= (g^{(k)\top} + s_C^{(k)\top} B_k + (s_N^{(k)} - s_C^{(k)})^\top B_k)(s_N^{(k)} - s_C^{(k)}) \\
&= (g^{(k)\top} + s_N^{(k)\top} B_k)(s_N^{(k)} - s_C^{(k)}) \\
&= 0 \cdot (s_N^{(k)} - s_C^{(k)}) \\
&= 0
\end{aligned}
$$

所以沿着Cauchy点$x_C^{(k+1)}$和牛顿点$x_N^{(k+1)}$的连线，子问题模型函数值单调减少。

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

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

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

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

证明. x^*是(122)的局部极小点，即存在x^*的邻域U，使得

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

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

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

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

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

□

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

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

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

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

(1)与(2)的第一式作差，得

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

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

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。

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

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

□

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。

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

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

$k\to\infty$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

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

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，

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

(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

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

(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。

考虑等式约束问题

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

其中$c(x) = (x_1(x), \dots, c_{m_e}(x))^\top$。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。

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

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

且对所有$d \in Ker\, A(x^*)$均有

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

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

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

反证，假设对任意正整数k，都有存在方向d_k，满足

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

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

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

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

其中$\|\nabla^2_{xx}L(x^*,\lambda^*)\|_2$为$\nabla^2_{xx}L(x^*,\lambda^*)$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

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

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

□