二次规划求解实例（起作用集）

2018年6月17日　　9:03

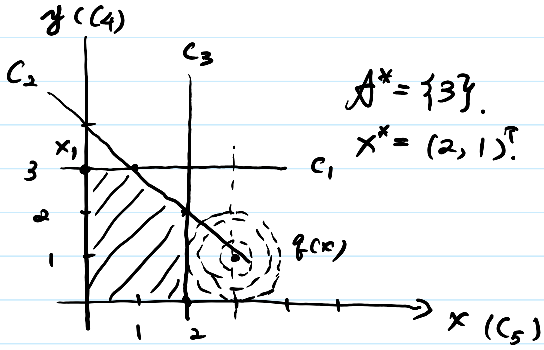
例 8.3

$$\begin{cases} \min \ \frac{1}{2}[(x-3)^2+(y-1)^2]-5 \\ \text{s.t.} \quad 3-y\geq 0, \\ \qquad 4-x-y\geq 0 \\ \qquad 2-x\geq 0 \\ \qquad x\geq 0,\ y\geq 0 \end{cases}$$

← 这一约束 $x_i\geq 0$, $i=1,2,\cdots,n$，几乎总是有的，可以通过适当的变量替换和松弛得到。

即 $\forall\ x_i\in\mathbb{R}$，$x_i=x_i^+-x_i^-$，$x_i^+, x_i^-\geq 0$.

对小规模低维问题，可先分析一下可行域情况：先选取一个初值，这里可以选④的顶点或边界。这里显然一个合理的选择是 $\mathcal{A}_1=\{3,5\}$，$x_1=(2,0)$。但我们不能假设总是能通过绘图得到有用的信息，所以我们按书上的做法，仅从 $\mathcal{A}_1$ 方便构建的角度，直接观察约束，取 $\mathcal{A}_1=\{1,4\}$，$x_1=(0,3)$。这样对应子问题为：



$\mathcal{A}^*=\{3\}$.
$x^*=(2,1)^T$.

$$\begin{cases} \min \ \frac{1}{2}[(x-3)^2+(y-1)^2]-5 \\ \text{s.t.} \quad y=3,\ x=0 \end{cases}$$

$$\Rightarrow \min \ \frac{1}{2}[9+4]-5$$

对一个常数不可能有任何改进，故 $d_1=(0,0)^T$。此时检查 λ_k：

$$\nabla_x \mathcal{L}(x,\lambda)=\begin{bmatrix} x-3 \\ y-1 \end{bmatrix}-\lambda_1\begin{bmatrix} 0 \\ -1 \end{bmatrix}-\lambda_2\begin{bmatrix} -1 \\ -1 \end{bmatrix}-\lambda_3\begin{bmatrix} -1 \\ 0 \end{bmatrix}-\lambda_4\begin{bmatrix} 1 \\ 0 \end{bmatrix}-\lambda_5\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

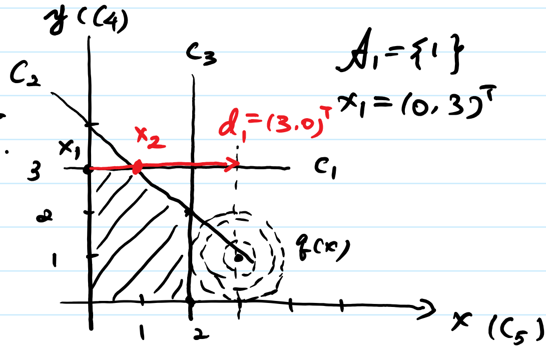
代入 $\lambda_2=\lambda_3=\lambda_5=0$，$x=0$，$y=3$，得：

$$\begin{cases} -3-\lambda_4=0 \\ 2+\lambda_1=0 \end{cases} \Rightarrow \begin{cases} \lambda_1=-2 \\ \lambda_4=-3 \end{cases} \Rightarrow \lambda_k=\{-2,0,0,-3,0\}.$$

不满足 KKT。这里 λ_4 更小，我们选择去掉这个约束。于是 $\mathcal{A}_1=\{1\}$。问题变为：

$$\begin{cases} \min \ \frac{1}{2}[(x-3)^2+(y-1)^2]-5 \\ \text{s.t.} \quad y=3 \end{cases}$$

$$\Rightarrow \quad \min \ \frac{1}{2}(x-3)^2-3 \quad \Rightarrow \quad d=(3,0)^T.$$



$\mathcal{A}_1=\{1\}$
$x_1=(0,3)^T$

从图中可以看出 d_1 已违背约束 C_2，C_3。且在 d_1 方向上首先违背的是 C_2。或从代数形式，我们不需要验证

C_1，因为 C_1 在 $\mathcal{A}_1$ 中。

C_4, C_5，因为

$$\left.\begin{array}{l} a_4^T d_1=(1,0)\begin{pmatrix} 3 \\ 0 \end{pmatrix}=3>0, \text{ 在 } d_1 \text{ 正方.} \\ a_5^T\cdot d_1=(0,1)\begin{pmatrix} 3 \\ 0 \end{pmatrix}=0, \text{ 和 } d_1 \text{ 平行} \end{array}\right\rangle \text{ 均不可能阻挡 } d_1.$$

只需验证：

C_2，C_3，因为

c_2，c_3，因为

$$a_2^T\cdot d_1 = (-1,-1)\cdot\begin{pmatrix}3\\0\end{pmatrix} = -3 < 0，\text{在 } d_1 \text{ 前方}.$$

$$a_3^T\cdot d_1 = (-1,0)\cdot\begin{pmatrix}3\\0\end{pmatrix} = -3 < 0，\text{在 } d_1 \text{ 前方}.$$

在两个可能的阻挡中，选"近"的一个，通过计算阻挡的位置：

$$\min\left\{1,\ \begin{array}{l}\dfrac{-4-a_2^Tx_1}{a_2^Td_1} = \left(-4-(-1,-1)\begin{pmatrix}0\\3\end{pmatrix}\right)/-3 = \dfrac{1}{3},\\[2ex] \dfrac{-2-a_3^Tx_1}{a_3^Td_1} = \left(-2-(-1,0)\begin{pmatrix}0\\3\end{pmatrix}\right)/-3 = \dfrac{2}{3}\end{array}\right\} \quad \left(b = \begin{bmatrix}-3\\-4\\-2\\0\\0\end{bmatrix}，\text{注意符号}\right)$$

$= \frac{1}{3} = \alpha_1$. 故 $x_2 = x_1 + \alpha_1\cdot d_1 = \begin{pmatrix}0\\3\end{pmatrix} + \frac{1}{3}\begin{pmatrix}3\\0\end{pmatrix} = \begin{pmatrix}1\\3\end{pmatrix}$.（以上操作，注意几何意义和代数变换的一致性.）

此时，只有 c_2：$4-1-3=0$，活跃. 故 $\mathcal{A}_2 = \{1,2\}$.（原因？）

接下去的过程是机械的. 自己回家完成剩下步骤.

等值优化子问题的求解：

对 x_k，和 $\mathcal{A}_k$，我们需要求解等值约束优化：

$$\begin{aligned}&\min\ q(x)\\ &\text{s.t. } a_i^Tx = b_i,\ i\in\mathcal{A}_k\end{aligned}$$

这里记 $A_k = [a_1, a_2, \cdots, a_m]$. 则约束可记为：

$$A^Tx = b,\quad b = (b_1, b_2, \cdots, b_m)^T.$$

对此问题，我们可以用之前介绍过的零空间方法：即对 $A_k \in \mathbb{R}^{n\times m}$ 做 QR 分解：

$$A_k = Q_k\begin{bmatrix}R_k\\0\end{bmatrix} = \begin{bmatrix}Q_1^{(k)} & Q_2^{(k)}\end{bmatrix}\begin{bmatrix}R_k\\0\end{bmatrix},$$

其中 $Q_k \in \mathbb{R}^{n\times n}$ 是正交阵，$R_k \in \mathbb{R}^{m\times m}$ 是非奇异上三角阵，$Q_1^{(k)} \in \mathbb{R}^{n\times m}$，$Q_2^{(k)} \in \mathbb{R}^{n\times(n-m)}$.

若从 k 步到 $k+1$ 步，增加了一个约束 p，即

$$\mathcal{A}_{k+1} = \mathcal{A}_k \cup \{p\},$$

且

$$A_{k+1} = [A_k\ \ a_p] \qquad (\text{参}(3\text{-}5)),$$

显然 $A_{k+1} \in \mathbb{R}^{n\times(m+1)}$. 现在我们需要计算

$$A_{k+1} = Q_{k+1}\begin{bmatrix}R_{k+1}\\0\end{bmatrix}.$$

然而并没有必要重新计算. 注意到

$$Q_kQ_k^T = \begin{bmatrix}Q_1^{(k)} & Q_2^{(k)}\end{bmatrix}\begin{bmatrix}Q_1^{(k)T}\\Q_2^{(k)T}\end{bmatrix} = Q_1^{(k)}Q_1^{(k)T} + Q_2^{(k)}Q_2^{(k)T} = I_{n\times n}$$

而

$$\begin{aligned}A_{k+1} &= [A_k\ \ a_p] = Q_k\begin{bmatrix}R & Q_1^{(k)T}a_p\\0 & Q_2^{(k)T}a_p\end{bmatrix}\\ &= \begin{bmatrix}Q_1^{(k)} & Q_2^{(k)}\end{bmatrix}\begin{bmatrix}R & Q_1^{(k)T}a_p\\0 & Q_2^{(k)T}a_p\end{bmatrix}\\ &= \left[A_p\ \ (Q_1^{(k)}Q_1^{(k)T} + Q_2^{(k)}Q_2^{(k)T})a_p\right] = [A_k\ \ a_p].\end{aligned} \tag{6.22}$$

实际上就是除最后一列，已分解. 我们只需继续分解，即将 $Q_2^{(k)T}a_p$ 通过正交变换为 $\gamma\cdot e_1^{(n-m)}$，比如，求 Householder 矩阵 $\hat{Q} \in \mathbb{R}^{(n-m)\times(n-m)}$，使

$$\hat{Q}(Q_2^{(k)T}a_p) = \begin{bmatrix}\gamma\\0\\\vdots\\0\end{bmatrix} = \gamma\cdot e_1^{(n-m)},$$

由Householder变换定义：

$$\hat{Q}\left(Q_2^{(k)T}a_p\right) = \begin{bmatrix}\gamma\\0\\\vdots\\0\end{bmatrix} = \gamma\cdot e_1^{(n-m)},$$

由Householder变换定义:

$$\gamma = \|Q_2^{(k)T}a_p\|.$$

而完整的QR分解为,

$$A_{k+1} = Q_k\begin{bmatrix}R_k & Q_1^{(k)T}a_p\\ 0 & \hat{Q}^T\begin{bmatrix}\gamma\\0\\\vdots\\0\end{bmatrix}\end{bmatrix} = Q_k\begin{bmatrix}I_{m\times m} & 0\\ 0 & \hat{Q}^T\end{bmatrix}\begin{bmatrix}R_k & Q_1^{(k)T}a_p\\ 0 & \gamma e_1^{(n-m)}\end{bmatrix}$$

令

$$Q_{k+1} = Q_k\begin{bmatrix}I_{m\times m} & 0\\ 0 & \hat{Q}^T\end{bmatrix},\quad R_{k+1} = \begin{bmatrix}R_k & Q_1^{(k)T}a_p\\ 0 & \gamma\end{bmatrix}$$

$$= \left[Q_1^{(k)}\ \ Q_2^{(k)}\right]\begin{bmatrix}I & 0\\ 0 & \hat{Q}^T\end{bmatrix} = \left[Q_1^{(k)}\ \ Q_2^{(k)}\hat{Q}^T\right] = \left[Q_1^{(k+1)}\ \ Q_2^{(k+1)}\right].$$

注意此处, $Q_1^{(k)}\in\mathbb{R}^{n\times m}$, 而 $Q_1^{(k+1)}\in\mathbb{R}^{n\times(m+1)}$. 故 $Q_2^{(k+1)}$ 为 $Q_2^{(k)}\hat{Q}^T$ 的后 $n-m-1$ 列).
而相应地,

$$Z_{k+1} = Q_2^{(k+1)}\in\mathbb{R}^{n\times(n-m-1)}.$$

而对 A_k 中某一列未被移出的情况, 则此时的情况是

$$\underset{\text{q列移出}}{A_k} = Q_k\begin{bmatrix}R_k\\0\end{bmatrix} = \underset{Q_k\text{不变}}{Q_k}\begin{bmatrix}* & * & * & *\\ 0 & * & * & *\\ 0 & 0 & * & *\\ 0 & 0 & 0 & *\\ 0 & 0 & 0 & 0\end{bmatrix},\quad \text{此时 } R_k = \begin{bmatrix}* & * & *\\ 0 & * & *\\ 0 & 0 & *\\ 0 & 0 & *\\ 0 & 0 & 0\end{bmatrix} \leftarrow \text{需要消去.}$$

q列移出.

即 R_k 不再是上三角阵, 我们需要

$$\hat{Q}R_k = \begin{bmatrix}R_{k+1}\\0\end{bmatrix},$$

其中 R_{k+1} 是上三角阵. 对此不难发现, 我们实际上只需消去副对角线上 $m-q$ 个非零元. 对此可用Givens变换逐个消去. 此即 $\hat{Q}$ 的构建方式. 而后, 可令

$$Q_{k+1} = Q_k\begin{bmatrix}\hat{Q}^T & 0\\ 0 & I\end{bmatrix},$$

而 Z_{k+1} 为 Q_{k+1} 的后 $n-m+1$ 列).