线性规划问题的内点法

第9章

应用最优化方法及MATLAB实现

刘兴高

线性规划

$$\begin{aligned} \min f_p(\boldsymbol{x}) &= \boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.}\, A\boldsymbol{x} &= \boldsymbol{b} \\ \boldsymbol{x} &\geq 0 \end{aligned}$$

Primal problem

$\boldsymbol{x} \in \mathcal{R}^n$, $\boldsymbol{c} \in \mathcal{R}^2$

$A \in \mathcal{R}^{m\times n}$, $\operatorname{rank} A = m < n$

$\boldsymbol{b} \geq \boldsymbol{0} \in \mathcal{R}^n$

9.1 内点法的相关概念与基本原理

9.1.1 内点法与单纯形法

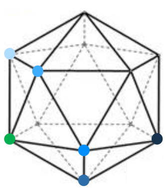
几何观点：线性规划的可行域是凸多面体，其顶点和基可行解一一对应

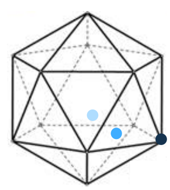
单纯形法的求解过程：从凸多面体某个顶点开始，沿着凸多面体上彼此相邻的顶点前进，最终找到使目标函数取最优值的顶点

内点法的求解过程：从凸多面体内部的某个点出发，逐渐逼近最优解对应的顶点

计算量：内点法的每迭代一次的计算量，比单纯形的大一些；　适合小规模问题

单纯形的迭代次数，比内点法的多得多。　　适合大规模问题





REVIEW

弱对偶定理 与 强对偶定理

$$\min f(\boldsymbol{x})$$
$$\text{s.t.}\ c_i(\boldsymbol{x}) = 0, \quad i \in E$$
$$c_i(\boldsymbol{x}) \geq 0, \quad i \in I$$

Primal problem

$$\max_{\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{\mu}} L(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{\mu})$$
$$\text{s.t.}\ \nabla_x L(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{\mu}) = \boldsymbol{0}$$
$$\boldsymbol{\mu} \geq \boldsymbol{0}$$

Dual problem

定理7.4.2 弱对偶定理 原问题的任意可行点对应的目标函数值都不小于对偶问题的可行点的目标函数值。 $f \geq L$

定理7.4.3 强对偶定理 设正则点$\boldsymbol{x}^*$是原问题的一个极小点，$\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*$是相应的拉格朗日乘子，则$(\boldsymbol{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$是对偶问题的极大点，且极大值$L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) = f(\boldsymbol{x}^*)$。

$$\min f = \max L$$

9.1.2 线性规划问题的对偶问题与对偶间隔

$$\min f_P(\boldsymbol{x}) = \boldsymbol{c}^T\boldsymbol{x} \quad \text{s.t.}\ A\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \geq \boldsymbol{0}$$

Primal problem

$$\max\ f_D(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{\mu}) = \boldsymbol{c}^T\boldsymbol{x} - \boldsymbol{\lambda}^T(A\boldsymbol{x}-\boldsymbol{b}) - \boldsymbol{\mu}^T\boldsymbol{x} \quad \text{s.t.}\ A^T\boldsymbol{\lambda} + \boldsymbol{\mu} = \boldsymbol{c},\ \boldsymbol{\mu} \geq \boldsymbol{0}$$

Dual problem

$$L_p(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{\mu}) = \boldsymbol{c}^T\boldsymbol{x} - \boldsymbol{\lambda}^T(A\boldsymbol{x}-\boldsymbol{b}) - \boldsymbol{\mu}^T\boldsymbol{x}$$

KKT条件

$$\begin{aligned}
&\nabla_x L_p(\boldsymbol{x}^*,\boldsymbol{\lambda}^*,\boldsymbol{\mu}^*) = \boldsymbol{c} - A^T\boldsymbol{\lambda}^* - \boldsymbol{\mu}^* = \boldsymbol{0}\\
&\nabla_\lambda L_p(\boldsymbol{x}^*,\boldsymbol{\lambda}^*,\boldsymbol{\mu}^*) = -(A\boldsymbol{x}^* - \boldsymbol{b}) = \boldsymbol{0}\\
&\nabla_\mu L_p(\boldsymbol{x}^*,\boldsymbol{\lambda}^*,\boldsymbol{\mu}^*) = -\boldsymbol{x}^* \leq \boldsymbol{0}\\
&X^*\boldsymbol{\mu}^* = \boldsymbol{0}\\
&\boldsymbol{\mu}^* \geq \boldsymbol{0}
\end{aligned} \tag{9.1.3}$$

$$X^* = \operatorname{diag}\{x_1^*, x_2^*, \dots, x_n^*\}$$

$$A^T\boldsymbol{\lambda} + \boldsymbol{\mu} = \boldsymbol{c} \rightarrow \boldsymbol{\lambda}^T\boldsymbol{A} + \boldsymbol{\mu}^T = \boldsymbol{c}^T$$

$$\begin{aligned}
L_D(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{\mu}) &= \boldsymbol{c}^T\boldsymbol{x} - \boldsymbol{\lambda}^T(A\boldsymbol{x}-\boldsymbol{b}) - \boldsymbol{\mu}^T\boldsymbol{x}\\
&= (\boldsymbol{c}^T - \boldsymbol{\lambda}^T A - \boldsymbol{\mu}^T)\boldsymbol{x} + \boldsymbol{\lambda}^T\boldsymbol{b}\\
&= \boldsymbol{\lambda}^T\boldsymbol{b}
\end{aligned}$$

$$A^T\boldsymbol{\lambda} + \boldsymbol{\mu} = \boldsymbol{c} \rightarrow \boldsymbol{\mu} = \boldsymbol{c} - A^T\boldsymbol{\lambda} \geq \boldsymbol{0}$$

$$\max L_D(\boldsymbol{\lambda}) = \boldsymbol{\lambda}^T\boldsymbol{b} \quad \text{s.t.}\ \boldsymbol{c} - A^T\boldsymbol{\lambda} \geq \boldsymbol{0}$$

$$\max f_D(\boldsymbol{\lambda}) = \boldsymbol{\lambda}^T\boldsymbol{b} \quad \text{s.t.}\ \boldsymbol{c} - A^T\boldsymbol{\lambda} \geq \boldsymbol{0} \tag{D-P}$$

$$L_{DP}(\boldsymbol{\lambda},\boldsymbol{\eta}) = -\boldsymbol{\lambda}^T\boldsymbol{b} - \boldsymbol{\eta}^T(\boldsymbol{c} - A^T\boldsymbol{\lambda}) = -\boldsymbol{\eta}^T\boldsymbol{c}$$

max → min

$$\max L_{DP}(\boldsymbol{\lambda}) = -\boldsymbol{c}^T\boldsymbol{\eta} \quad \text{s.t.}\ A\boldsymbol{\eta} = \boldsymbol{b},\ \boldsymbol{\eta} \geq \boldsymbol{0} \tag{D-D}$$

KKT条件

$$\begin{aligned}
&\nabla_\lambda L_{DP}(\boldsymbol{\lambda}^*,\boldsymbol{\eta}^*) = A\boldsymbol{\eta}^* - \boldsymbol{b} = \boldsymbol{0}\\
&\nabla_\eta L_{DP}(\boldsymbol{\lambda}^*,\boldsymbol{\eta}^*) = -(\boldsymbol{c} - A^T\boldsymbol{\lambda}^*) \leq \boldsymbol{0}\\
&\Gamma^*(\boldsymbol{c} - A^T\boldsymbol{\lambda}^*) = \boldsymbol{0}\\
&\boldsymbol{\eta}^* \geq \boldsymbol{0}
\end{aligned}$$

$$A^T\boldsymbol{\lambda} + \boldsymbol{\mu} = \boldsymbol{c} \rightarrow \boldsymbol{\mu}^* = \boldsymbol{c} - A^T\boldsymbol{\lambda}^*$$

$$\begin{aligned}
&\boldsymbol{\mu}^* = \boldsymbol{c} - A^T\boldsymbol{\lambda}^*\\
&\nabla_\lambda L_{DP}(\boldsymbol{\lambda}^*,\boldsymbol{\eta}^*) = A\boldsymbol{\eta}^* - \boldsymbol{b} = \boldsymbol{0}\\
&\nabla_\eta L_{DP}(\boldsymbol{\lambda}^*,\boldsymbol{\eta}^*) = -\boldsymbol{\mu}^* \leq \boldsymbol{0}\\
&\Gamma^*\boldsymbol{\mu}^* = \boldsymbol{0}\\
&\boldsymbol{\eta}^* \geq \boldsymbol{0}
\end{aligned} \tag{9.1.8}$$

$$\Gamma^* = \operatorname{diag}\{\eta_1^*, \eta_2^*, \dots, \eta_n^*\}$$

γ

$$f_P(\boldsymbol{x}) = \boldsymbol{c}^T\boldsymbol{x} \geq \boldsymbol{\lambda}^T\boldsymbol{b} = f_D(\boldsymbol{\lambda})$$ 弱对偶定理

$\boldsymbol{\eta}^* = \boldsymbol{x}^*$时， $\boldsymbol{c}^T\boldsymbol{x}^* = \boldsymbol{\lambda}^{*T}\boldsymbol{b}$ (9.1.3) ⟺ (9.1.8) 强对偶定理

原问题与对偶问题的最优解同时达到，
$(\boldsymbol{x}^*,\boldsymbol{\lambda}^*,\boldsymbol{\mu}^*)$称为原问题和对偶问题的原-对偶解

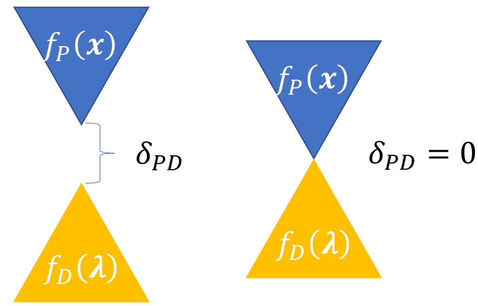
$$f_P(\boldsymbol{x}) = \boldsymbol{c}^T\boldsymbol{x} \geq \boldsymbol{\lambda}^T\boldsymbol{b} = f_D(\boldsymbol{\lambda})$$

对偶间隔为原问题与对偶问题目标函数值的差

$$\delta_{PD} = f_P(\boldsymbol{x}) - f_D(\boldsymbol{\lambda}) = \boldsymbol{c}^T\boldsymbol{x} - \boldsymbol{\lambda}^T\boldsymbol{b}$$

$$A^T\boldsymbol{\lambda} + \boldsymbol{\mu} = \boldsymbol{c}$$
$$\rightarrow \boldsymbol{\mu}^T = \boldsymbol{c}^T - \boldsymbol{\lambda}^T A$$
$$\rightarrow \boldsymbol{\mu}^T\boldsymbol{x} = \boldsymbol{c}^T\boldsymbol{x} - \boldsymbol{\lambda}^T A\boldsymbol{x} = \boldsymbol{c}^T\boldsymbol{x} - \boldsymbol{\lambda}^T\boldsymbol{b}$$

$$\delta_{PD} = \boldsymbol{c}^T\boldsymbol{x} - \boldsymbol{\lambda}^T\boldsymbol{b} = \boldsymbol{\mu}^T\boldsymbol{x}$$



9.1.3 内点与中心路径

原问题的KKT条件

$$\begin{cases} \boldsymbol{c} - A^T\boldsymbol{\lambda}^* - \boldsymbol{\mu}^* = \boldsymbol{0}, \boldsymbol{\mu}^* \geq \boldsymbol{0} \\ A\boldsymbol{x}^* - \boldsymbol{b} = \boldsymbol{0}, \quad \boldsymbol{x}^* \geq \boldsymbol{0} \\ X^*\boldsymbol{\mu}^* = \boldsymbol{0} \end{cases} \quad (9.1.12)$$

$\boldsymbol{x}$是P问题可行域的内点
$\boldsymbol{\lambda}$是D问题可行域的内点
$(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu})$满足

$$\begin{cases} \boldsymbol{c} - A^T\boldsymbol{\lambda} - \boldsymbol{\mu} = \boldsymbol{0}, \boldsymbol{\mu} \geq \boldsymbol{0} \\ A\boldsymbol{x} - \boldsymbol{b} = \boldsymbol{0}, \quad \boldsymbol{x} \geq \boldsymbol{0} \end{cases} \quad (9.1.13)$$

Dual Feasible

Primal Feasible

$\tau \rightarrow 0$

构成方程组

$$\begin{cases} \boldsymbol{c} - A^T\boldsymbol{\lambda} - \boldsymbol{\mu} = \boldsymbol{0}, \boldsymbol{\mu} \geq \boldsymbol{0} \\ A\boldsymbol{x} - \boldsymbol{b} = \boldsymbol{0}, \quad \boldsymbol{x} \geq \boldsymbol{0} \\ X\boldsymbol{\mu} = \tau\boldsymbol{e} \end{cases}$$

增加

$X = \mathrm{diag}\{x_1, x_2, \ldots, x_n\}$
$\boldsymbol{e} = [1,1, \ldots, 1]^T \in \mathcal{R}^n$
$\tau > 0,\ \tau \rightarrow 0$

其解

$$\boldsymbol{z}(\tau) = [\boldsymbol{x}(\tau), \boldsymbol{\lambda}(\tau), \boldsymbol{\mu}(\tau)]$$

$$\begin{cases} \boldsymbol{c} - A^T\boldsymbol{\lambda}^* - \boldsymbol{\mu}^* = \boldsymbol{0}, \boldsymbol{\mu}^* \geq \boldsymbol{0} \\ A\boldsymbol{x}^* - \boldsymbol{b} = \boldsymbol{0}, \qquad \boldsymbol{x}^* \geq \boldsymbol{0} \\ X^*\boldsymbol{\mu}^* = \boldsymbol{0} \end{cases} \quad (9.1.12)$$

$$\begin{cases} \boldsymbol{c} - A^T\boldsymbol{\lambda} - \boldsymbol{\mu} = \boldsymbol{0}, \boldsymbol{\mu} \geq \boldsymbol{0} \\ A\boldsymbol{x} - \boldsymbol{b} = \boldsymbol{0}, \qquad \boldsymbol{x} \geq \boldsymbol{0} \\ X\boldsymbol{\mu} = \tau\boldsymbol{e} \end{cases} \quad (9.1.14)$$

$X = \mathrm{diag}\{x_1, x_2, \dots, x_n\}$

$\boldsymbol{e} = [1,1,\dots,1]^T \in \mathcal{R}^n$

解 $\boldsymbol{z}(\tau) = [\boldsymbol{x}(\tau), \boldsymbol{\lambda}(\tau), \boldsymbol{\mu}(\tau)]$ $\quad \tau > 0$

$\tau \to 0$时，内点序列$\{\boldsymbol{x}(\tau)\}$和$\{\boldsymbol{\lambda}(\tau)\}$分别在原问题和对偶问题的可行域内画出一条轨迹，且轨迹指向最优解所在的凸多面体顶点，分别称为原问题和对偶问题的中心路径

$\tau \to 0$时，$\boldsymbol{z}(\tau) \to$原问题的解$(\boldsymbol{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$，因为(9.1.14) → P问题的KKT条件

同时，从对偶间隔来看，$\delta_{PD} = \boldsymbol{\mu}^T\boldsymbol{x} = n\tau \to 0$

例9.1 画出例8.5中线性规划问题的中心路径。

$$\begin{aligned} \max f(\boldsymbol{x}) &= x_1 + x_2 + 5x_3 \\ \text{s.t.}\ 3x_1 + 2x_2 + \frac{1}{4}x_3 &\leq 6 \\ x_3 &\leq 4 \\ \boldsymbol{x} &\geq 0 \end{aligned}$$

$$\begin{cases} \boldsymbol{c} - A^T\boldsymbol{\lambda} - \boldsymbol{\mu} = \boldsymbol{0}, \boldsymbol{\mu} \geq \boldsymbol{0} \\ A\boldsymbol{x} - \boldsymbol{b} = \boldsymbol{0}, \qquad \boldsymbol{x} \geq \boldsymbol{0} \\ X\boldsymbol{\mu} = \tau\boldsymbol{e} \end{cases} \quad (9.1.14)$$

解：
$$\begin{aligned} \min -f(\boldsymbol{x}) &= -x_1 - x_2 - 5x_3 \\ \text{s.t.}\ 3x_1 + 2x_2 + \frac{1}{4}x_3 + x_4 &= 6 \\ x_3 + x_5 &= 4 \\ \boldsymbol{x} &\geq 0 \end{aligned}$$

$\boldsymbol{c} = [-1 \ \ -1 \ \ -5 \ \ 0 \ \ 0]^T$

$A = \begin{bmatrix} 3 & 2 & 1/4 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix}$

$\boldsymbol{b} = [6 \ \ 4]^T$

$$\begin{aligned} &-1 + 3\lambda_1 - \mu_1 = 0 \\ &-1 + 2\lambda_1 - \mu_2 = 0 \\ &-5 + \frac{1}{4}\lambda_1 + \lambda_2 - \mu_3 = 0 \\ &\lambda_1 - \mu_4 = 0 \\ &\lambda_2 - \mu_5 = 0 \end{aligned}$$

$$\begin{aligned} &3x_1 + 2x_2 + \frac{1}{4}x_3 + x_4 = 6 \\ &x_3 + x_5 = 4 \end{aligned}$$

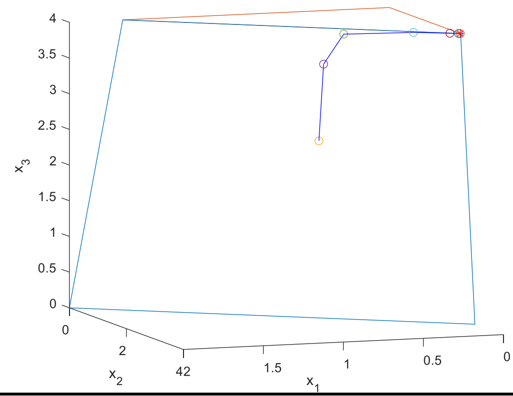
$\mu_i x_i = \tau, \qquad i = 1,2,\dots,5$

Example_9_1_XinggaoLiu.m

$$\begin{cases} -1+\dfrac{3\tau}{6-3x_1-2x_2-x_3/4}-\dfrac{\tau}{x_1}=0 \\ -1+\dfrac{2\tau}{6-3x_1-2x_2-x_3/4}-\dfrac{\tau}{x_2}=0 \\ -1+\dfrac{\tau}{4(6-3x_1-2x_2-x_3/4)}+\dfrac{\tau}{4-x_3}-\dfrac{\tau}{x_3}=0 \end{cases}$$

K=7

x_min =
0.0020
2.4959
3.9998



9.1.4 内点法的基本原理

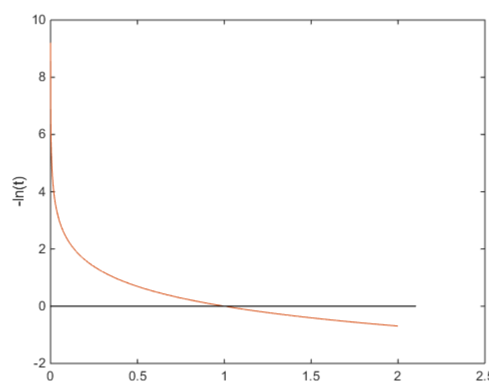
障碍罚函数法，内点罚函数法

$$\min f(\boldsymbol{x})=\boldsymbol{c}^T\boldsymbol{x}-\tau\sum_{j=1}^{n}\ln x_j$$

$$\text{s.t.}\ A\boldsymbol{x}=\boldsymbol{b}$$

每次迭代都获得中心路径上的点，
并趋于最优解点

其KKT条件就是(9.1.14)



现代内点法中，最成功的是 原-对偶路径跟踪法
每次迭代的点不一定在中心路径上，但是，能够围绕或跟踪中心路径直至找到最优解点

$$L(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{\mu}) = \boldsymbol{c}^T\boldsymbol{x} - \tau\sum_{j=1}^{n}\ln x_j - \boldsymbol{\lambda}^T(A\boldsymbol{x}-\boldsymbol{b}) - \boldsymbol{\mu}^T\boldsymbol{x}$$

$$\boldsymbol{c}^T - \sum_{j=1}^{n}\frac{\tau}{x_j} - \boldsymbol{\lambda}^T A = \boldsymbol{0}$$

$$\nabla_x L(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{\mu}) = \boldsymbol{c} - \sum_{j=1}^{n}\frac{\tau}{x_j} - A^T\boldsymbol{\lambda} - \boldsymbol{\mu} = \boldsymbol{0} \longrightarrow \boldsymbol{c} - \sum_{j=1}^{n}\frac{\tau}{x_j} - A^T\boldsymbol{\lambda} = \boldsymbol{0}$$

$A\boldsymbol{x} - \boldsymbol{b} = \boldsymbol{0}$

$\boldsymbol{x} > \boldsymbol{0}$

$\boldsymbol{\mu} \geq \boldsymbol{0}$

$\mu_i x_i = 0, i = 1, \dots, n \longrightarrow \boldsymbol{\mu} = \boldsymbol{0}$

$A\boldsymbol{x} - \boldsymbol{b} = \boldsymbol{0}$

$$\sum_{j=1}^{n}\frac{\tau}{x_j} = \tau X^{-1}\boldsymbol{e} \qquad X^{-1} = \begin{bmatrix} \frac{1}{x_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{x_n} \end{bmatrix}$$

$$\boldsymbol{c} - \sum_{j=1}^{n}\frac{\tau}{x_j} - A^T\boldsymbol{\lambda} = \boldsymbol{0} \longrightarrow \boldsymbol{c}^T - \sum_{j=1}^{n}\frac{\tau}{x_j} - \boldsymbol{\lambda}^T A = 0$$

$A\boldsymbol{x} - \boldsymbol{b} = \boldsymbol{0}$

$$\longrightarrow \boldsymbol{c}^T\boldsymbol{x} - \left[\sum_{j=1}^{n}\frac{\tau}{x_j}\right]^T \boldsymbol{x} - \boldsymbol{\lambda}^T A\boldsymbol{x} = 0$$

$$\begin{aligned}\left[\sum_{j=1}^{n}\frac{\tau}{x_j}\right]^T \boldsymbol{x} &= [\tau X^{-1}\boldsymbol{e}\,]^T\boldsymbol{x} \\ &= \tau\boldsymbol{e}^T[X^{-1}\,]^T\boldsymbol{x} \\ &= \tau\boldsymbol{e}^T \begin{bmatrix} \frac{1}{x_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{x_n} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \\ &= \tau\boldsymbol{e}^T\boldsymbol{e} \\ &= n\tau \end{aligned}$$

$$\longrightarrow \boldsymbol{c}^T\boldsymbol{x} - n\tau - \boldsymbol{\lambda}^T\boldsymbol{b} = 0$$

$$\boldsymbol{c}^T\boldsymbol{x} - \boldsymbol{\lambda}^T\boldsymbol{b} = n\tau$$

对偶间隔 $\delta_{PD} = \boldsymbol{c}^T\boldsymbol{x} - \boldsymbol{\lambda}^T\boldsymbol{b}$

$= n\tau$

9.2 原-对偶可行路径跟踪法

9.2.1 问题形式

原问题

$$\begin{aligned} \min f_P(\boldsymbol{x}) &= \boldsymbol{c}^T\boldsymbol{x} & \boldsymbol{x} \in \mathcal{R}^n, \boldsymbol{c} \in \mathcal{R}^n \\ \text{s.t.}\, A_E\boldsymbol{x} &= \boldsymbol{b}_E & A_E \in \mathcal{R}^{m_1\times n}, \boldsymbol{b}_E \in \mathcal{R}^{m_1} \\ A_I\boldsymbol{x} &\geq \boldsymbol{b}_I & A_I \in \mathcal{R}^{m_2\times n}, \boldsymbol{b}_I \in \mathcal{R}^{m_2} \end{aligned} \qquad \operatorname{rank} A_E = m_1 < n$$

(9.2.1)

引入松弛变量

$$\begin{aligned} \min f_P(\boldsymbol{x}) &= \boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.}\, A_E\boldsymbol{x} &= \boldsymbol{b}_E \\ A_I\boldsymbol{x} - \boldsymbol{y} &= \boldsymbol{b}_I \qquad \boldsymbol{y} \in \mathcal{R}^{m_2} \\ \boldsymbol{y} &\geq \boldsymbol{0} \end{aligned}$$

(9.2.4)

变量为$\begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix}$

$$\begin{aligned} \min f_P(\boldsymbol{x}) &= \boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.}\, A_E\boldsymbol{x} &= \boldsymbol{b}_E \\ A_I\boldsymbol{x} &\geq \boldsymbol{b}_I \end{aligned} \longrightarrow \begin{aligned} \min f_P(\boldsymbol{x}) &= \boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.}\, A_E\boldsymbol{x} &= \boldsymbol{b}_E \\ A_I\boldsymbol{x} - \boldsymbol{y} &= \boldsymbol{b}_I \\ \boldsymbol{y} &\geq \boldsymbol{0} \end{aligned}$$

$$L_P(\,\boldsymbol{x},\boldsymbol{y},\boldsymbol{\lambda},\boldsymbol{\mu}\,) = \boldsymbol{c}^T\boldsymbol{x} - \boldsymbol{\lambda}_E^T(A_E\boldsymbol{x} - \boldsymbol{b}_E) - \boldsymbol{\lambda}_I^T(A_I\boldsymbol{x} - \boldsymbol{b}_I - \boldsymbol{y}\,) - \boldsymbol{\mu}^T\boldsymbol{y}$$

精确KKT条件

$$\begin{aligned} &\nabla_{\boldsymbol{x}} L_P(\,\boldsymbol{x},\boldsymbol{y},\boldsymbol{\lambda},\boldsymbol{\mu}\,) = \boldsymbol{c} - A_E^T\boldsymbol{\lambda}_E - A_I^T\boldsymbol{\lambda}_I \ = \boldsymbol{0} \\ &\nabla_{\boldsymbol{y}} L_P(\,\boldsymbol{x},\boldsymbol{y},\boldsymbol{\lambda},\boldsymbol{\mu}\,) = \boldsymbol{\lambda}_I - \boldsymbol{\mu} = \boldsymbol{0} \\ &\nabla_{\boldsymbol{\lambda}_E} L_P(\,\boldsymbol{x},\boldsymbol{y},\boldsymbol{\lambda},\boldsymbol{\mu}\,) = -A_E\boldsymbol{x} + \boldsymbol{b}_E = \boldsymbol{0} \\ &\nabla_{\boldsymbol{\lambda}_I} L_P(\,\boldsymbol{x},\boldsymbol{y},\boldsymbol{\lambda},\boldsymbol{\mu}\,) = -A_I\boldsymbol{x} + \boldsymbol{b}_I + \boldsymbol{y} = \boldsymbol{0} \\ &\nabla_{\boldsymbol{\mu}} L_P(\,\boldsymbol{x},\boldsymbol{y},\boldsymbol{\lambda},\boldsymbol{\mu}\,) = -\boldsymbol{y} \leq \boldsymbol{0} \\ &M\boldsymbol{y} = 0 \\ &\boldsymbol{\mu} \geq 0 \end{aligned} \qquad \begin{aligned} &\boldsymbol{\lambda}_I = \boldsymbol{\mu} \geq \boldsymbol{0} \\ &\boldsymbol{c} - A_E^T\boldsymbol{\lambda}_E - A_I^T\boldsymbol{\mu} = \boldsymbol{0}, \boldsymbol{\mu} \geq \boldsymbol{0} \\ &A_E\boldsymbol{x} - \boldsymbol{b}_E = \boldsymbol{0} \\ &A_I\boldsymbol{x} - \boldsymbol{b}_I - \boldsymbol{y} = \boldsymbol{0}, \boldsymbol{y} \geq \boldsymbol{0} \\ &M\boldsymbol{y} = 0 \end{aligned}$$

(9.2.5)

$$M = \operatorname{diag}\{\mu_1, \mu_2, \dots, \mu_{m_2}\}$$

$$\boldsymbol{\lambda} = \begin{bmatrix} \boldsymbol{\lambda}_E \\ \boldsymbol{\lambda}_I \end{bmatrix}$$

$$f_D(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\lambda},\boldsymbol{\mu}) = L_P(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\lambda},\boldsymbol{\mu})$$

$$f_D(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\lambda},\boldsymbol{\mu}) = \boldsymbol{c}^T\boldsymbol{x} - \boldsymbol{\lambda}_E^T(A_E\boldsymbol{x} - \boldsymbol{b}_E) - \boldsymbol{\lambda}_I^T(A_I\boldsymbol{x} - \boldsymbol{b}_I - \boldsymbol{y}) - \boldsymbol{\mu}^T\boldsymbol{y}$$

$$= [\boldsymbol{c}^T - \boldsymbol{\lambda}_E^T A_E - \boldsymbol{\lambda}_I^T A_I]\boldsymbol{x} + (\boldsymbol{\lambda}_I^T - \boldsymbol{\mu}^T)\boldsymbol{y} + \boldsymbol{\lambda}_E^T\boldsymbol{b}_E + \boldsymbol{\lambda}_I^T\boldsymbol{b}_I$$

$$= [\boldsymbol{c} - A_E^T\boldsymbol{\lambda}_E - A_I^T\boldsymbol{\lambda}_I]^T\boldsymbol{x} + [\boldsymbol{\lambda}_I - \boldsymbol{\mu}]^{\boldsymbol{T}}\boldsymbol{y} + \boldsymbol{\lambda}_E^T\boldsymbol{b}_E + \boldsymbol{\lambda}_I^T\boldsymbol{b}_I$$

$$\boldsymbol{c} - A_E^T\boldsymbol{\lambda}_E - A_I^T\boldsymbol{\lambda}_I = \boldsymbol{0}$$
$$\boldsymbol{\lambda}_I - \boldsymbol{\mu} = \boldsymbol{0}$$

$$= \boldsymbol{\lambda}_E^T\boldsymbol{b}_E + \boldsymbol{\lambda}_I^T\boldsymbol{b}_I$$

Let $\boldsymbol{\lambda} = \boldsymbol{\lambda}_E$

$$= \boldsymbol{\lambda}^T\boldsymbol{b}_E + \boldsymbol{\mu}^T\boldsymbol{b}_I$$

$$\begin{aligned} & f_D(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\lambda},\boldsymbol{\mu}) = L_P(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\lambda},\boldsymbol{\mu}) \\ & \text{s.t.}\ \boldsymbol{c} - A_E^T\boldsymbol{\lambda}_E - A_I^T\boldsymbol{\lambda}_I = \boldsymbol{0} \\ & \quad \boldsymbol{\lambda}_I - \boldsymbol{\mu} = \boldsymbol{0} \\ & \quad \boldsymbol{\mu} \geq \boldsymbol{0} \end{aligned}$$

令$\boldsymbol{\lambda} = \boldsymbol{\lambda}_E$，代入$\boldsymbol{\lambda}_I = \boldsymbol{\mu}$

$$\begin{aligned} & \max f_D(\boldsymbol{\lambda},\boldsymbol{\mu}) = \boldsymbol{\lambda}^T\boldsymbol{b}_E + \boldsymbol{\mu}^T\boldsymbol{b}_I \\ & \text{s.t.}\ \boldsymbol{\mu} \geq \boldsymbol{0} \end{aligned}$$

9.2.2 原-对偶可行路径跟踪法的基本原理

1.原问题（9.2.1）与其对偶问题的关系

$$\begin{aligned} & \max f_D(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\lambda},\boldsymbol{\mu}) = \boldsymbol{\lambda}^T\boldsymbol{b}_E + \boldsymbol{\mu}^T\boldsymbol{b}_I \\ & \text{s.t.}\ \boldsymbol{c} - A_E^T\boldsymbol{\lambda}_E - A_I^T\boldsymbol{\lambda}_I = \boldsymbol{0} \\ & \qquad \boldsymbol{\lambda}_I - \boldsymbol{\mu} = \boldsymbol{0} \\ & \qquad \boldsymbol{\mu} \geq \boldsymbol{0} \end{aligned}$$

令$\boldsymbol{\lambda} = \boldsymbol{\lambda}_E$，代入$\boldsymbol{\lambda}_I = \boldsymbol{\mu}$

$$\begin{aligned} & \max f_D(\boldsymbol{\lambda},\boldsymbol{\mu}) = \boldsymbol{\lambda}^T\boldsymbol{b}_E + \boldsymbol{\mu}^T\boldsymbol{b}_I \\ & \text{s.t.}\ \boldsymbol{\mu} \geq \boldsymbol{0} \end{aligned}$$

设点$(\boldsymbol{x},\boldsymbol{y})$是原问题可行域的内点，点$(\boldsymbol{\lambda},\boldsymbol{\mu})$是对偶问题可行域的内点，
则中心路径上的点$\boldsymbol{z}(\tau) = [\boldsymbol{x}(\tau),\boldsymbol{y}(\tau),\boldsymbol{\lambda}(\tau),\boldsymbol{\mu}(\tau)]$满足

扰动KKT条件

$$\begin{aligned} & \boldsymbol{c} - A_E^T\boldsymbol{\lambda} - A_I^T\boldsymbol{\mu} = \boldsymbol{0}, \boldsymbol{\mu} \geq \boldsymbol{0} && \text{DF} \\ & A_E\boldsymbol{x} - \boldsymbol{b}_E = \boldsymbol{0} && \\ & A_I\boldsymbol{x} - \boldsymbol{y} = \boldsymbol{b}_I, \boldsymbol{y} \geq \boldsymbol{0} && \text{PF} \\ & MY\boldsymbol{e} = \tau\boldsymbol{e} && \end{aligned} \quad (9.2.8)$$

$$Y = \text{diag}\{y_1, y_2, \ldots, y_{m_2}\} \quad \boldsymbol{e} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

运算

$$MY\boldsymbol{e}=\begin{bmatrix}\mu_1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \mu_m\end{bmatrix}\begin{bmatrix}y_1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & y_m\end{bmatrix}\begin{bmatrix}1\\ \vdots\\ 1\end{bmatrix}=\begin{bmatrix}\mu_1 y_1\\ \vdots\\ \mu_m y_m\end{bmatrix}=\boldsymbol{\mu}\circ\boldsymbol{y}$$

$$\begin{bmatrix}\mu_1 y_1\\ \vdots\\ \mu_m y_m\end{bmatrix}=MY\boldsymbol{e}=\tau\boldsymbol{e}=\begin{bmatrix}\tau\\ \vdots\\ \tau\end{bmatrix}$$

$$\boldsymbol{\mu}^T\boldsymbol{y}=\sum_{i=1}^{m_2}\mu_i y_i=m_2\tau$$

$\tau\to 0$时，扰动KKT条件 $\to$ 精确KKT条件

$$\boldsymbol{z}(\tau)=[\boldsymbol{x}(\tau),\boldsymbol{y}(\tau),\boldsymbol{\lambda}(\tau),\boldsymbol{\mu}(\tau)]\ \to$$ 原-对偶问题的最优解

从对偶间隔的角度看，

$$\begin{aligned}\delta_{PD}&=f_P(\boldsymbol{x})-f_D(\boldsymbol{\lambda},\boldsymbol{\mu})\\ &=\boldsymbol{c}^T\boldsymbol{x}-\boldsymbol{\lambda}^T\boldsymbol{b}_E-\boldsymbol{\mu}^T\boldsymbol{b}_I\\ &=(A_E^T\boldsymbol{\lambda}_E-A_I^T\boldsymbol{\lambda}_I)^T\boldsymbol{x}-\boldsymbol{\lambda}^T(A_E\boldsymbol{x})-\boldsymbol{\mu}^T(A_I\boldsymbol{x}-\boldsymbol{y})\\ &=\boldsymbol{\mu}^T\boldsymbol{y}\\ &=m_2\tau\to 0\end{aligned}$$

2.扰动KKT条件的线性化及求解

(1) 对当前迭代点$\boldsymbol{z}^{(k)}=\left[\boldsymbol{x}^{(k)},\boldsymbol{y}^{(k)},\boldsymbol{\lambda}^{(k)},\boldsymbol{\mu}^{(k)}\right]^T$做适当的扰动$\boldsymbol{\delta}_{\boldsymbol{z}}^{(k)}=\left[\boldsymbol{\delta}_x^{(k)},\boldsymbol{\delta}_y^{(k)},\boldsymbol{\delta}_\lambda^{(k)},\boldsymbol{\delta}_\mu^{(k)}\right]^T$，得到下一个迭代点$\boldsymbol{z}^{(k+1)}=\left[\boldsymbol{x}^{(k+1)},\boldsymbol{y}^{(k+1)},\boldsymbol{\lambda}^{(k+1)},\boldsymbol{\mu}^{(k+1)}\right]^T$

$$\boldsymbol{z}^{(k+1)}=\boldsymbol{z}^{(k)}+\boldsymbol{\delta}_{\boldsymbol{z}}^{(k)}$$

(2) 在可行路径跟踪法中，当前点$\boldsymbol{z}^{(k)}$是可行域的内点，满足

$$\begin{aligned}&\boldsymbol{c}-A_E^T\boldsymbol{\lambda}^{(k)}-A_I^T\boldsymbol{\mu}^{(k)}=\mathbf{0},\qquad \boldsymbol{\mu}^{(k)}\geq\mathbf{0}\\&A_E\boldsymbol{x}^{(k)}-\boldsymbol{b}_E=\mathbf{0}\\&A_I\boldsymbol{x}^{(k)}-\boldsymbol{b}_I-\boldsymbol{y}^{(k)}=\mathbf{0},\qquad \boldsymbol{y}^{(k)}\geq\mathbf{0}\end{aligned}$$

将点$\boldsymbol{z}^{(k+1)}$带入扰动KKT条件，略去关于扰动量的二次项$\boldsymbol{\delta}_y^{(k)}\circ\boldsymbol{\delta}_\mu^{(k)}$，得

$$\begin{aligned}&\boldsymbol{c}-A_E^T\left(\boldsymbol{\lambda}^{(k)}+\boldsymbol{\delta}_\lambda^{(k)}\right)-A_I^T\left(\boldsymbol{\mu}^{(k)}+\boldsymbol{\delta}_\mu^{(k)}\right)=\mathbf{0}\\&A_E\left(\boldsymbol{x}^{(k)}+\boldsymbol{\delta}_x^{(k)}\right)-\boldsymbol{b}_E=\mathbf{0}\\&A_I\left(\boldsymbol{x}^{(k)}+\boldsymbol{\delta}_x^{(k)}\right)-\left(\boldsymbol{y}^{(k)}+\boldsymbol{\delta}_y^{(k)}\right)=\boldsymbol{b}_I\\&M^{(k+1)}Y^{(k+1)}\boldsymbol{e}\approx M^{(k)}Y^{(k)}\boldsymbol{e}+M^{(k)}\boldsymbol{\delta}_y^{(k)}+Y^{(k)}\boldsymbol{\delta}_\mu^{(k)}=\tau^{(k+1)}\boldsymbol{e}\end{aligned}$$

化简这四个方程

前三个方程

$$\begin{aligned}&\boldsymbol{c}-A_E^T\left(\boldsymbol{\lambda}^{(k)}+\boldsymbol{\delta}_\lambda^{(k)}\right)-A_I^T\left(\boldsymbol{\mu}^{(k)}+\boldsymbol{\delta}_\mu^{(k)}\right)=\mathbf{0}\\&A_E\left(\boldsymbol{x}^{(k)}+\boldsymbol{\delta}_x^{(k)}\right)-\boldsymbol{b}_E=\mathbf{0}\\&A_I\left(\boldsymbol{x}^{(k)}+\boldsymbol{\delta}_x^{(k)}\right)-\left(\boldsymbol{y}^{(k)}+\boldsymbol{\delta}_y^{(k)}\right)=\boldsymbol{b}_I\end{aligned}$$

$$\begin{aligned}&\boldsymbol{c}-A_E^T\boldsymbol{\lambda}^{(k)}-A_I^T\boldsymbol{\mu}^{(k)}=\mathbf{0}\\&A_E\boldsymbol{x}^{(k)}-\boldsymbol{b}_E=\mathbf{0}\\&A_I\boldsymbol{x}^{(k)}-\boldsymbol{b}_I-\boldsymbol{y}^{(k)}=\mathbf{0}\end{aligned}$$

$$\begin{aligned}&A_E^T\boldsymbol{\delta}_\lambda^{(k)}+A_I^T\boldsymbol{\delta}_\mu^{(k)}=\mathbf{0}\\&A_E\boldsymbol{\delta}_x^{(k)}=\mathbf{0}\\&A_I\boldsymbol{\delta}_x^{(k)}-\boldsymbol{\delta}_y^{(k)}=\mathbf{0}\end{aligned}$$

第四个方程

$$M^{(k+1)}Y^{(k+1)}\boldsymbol{e} = \left(M^{(k)} + \Delta M^{(k)}\right)\left(Y^{(k)} + \Delta Y^{(k)}\right)\boldsymbol{e}$$

$$= M^{(k)}Y^{(k)}\boldsymbol{e} + \Delta M^{(k)}Y^{(k)}\boldsymbol{e} + M^{(k)}\Delta Y^{(k)}\boldsymbol{e} + \Delta M^{(k)}\Delta Y^{(k)}\boldsymbol{e}$$

$$= M^{(k)}Y^{(k)}\boldsymbol{e} + Y^{(k)}\Delta M^{(k)}\boldsymbol{e} + M^{(k)}\Delta Y^{(k)}\boldsymbol{e} + \Delta M^{(k)}\Delta Y^{(k)}\boldsymbol{e}$$

$$= M^{(k)}Y^{(k)}\boldsymbol{e} + Y^{(k)}\boldsymbol{\delta}_{\mu}^{(k)} + M^{(k)}\boldsymbol{\delta}_{y}^{(k)} + \boldsymbol{\delta}_{\mu}^{(k)}\circ\boldsymbol{\delta}_{y}^{(k)}$$

略去

$$\approx M^{(k)}Y^{(k)}\boldsymbol{e} + Y^{(k)}\boldsymbol{\delta}_{\mu}^{(k)} + M^{(k)}\boldsymbol{\delta}_{y}^{(k)}$$

略去关于扰动量的二次项$\boldsymbol{\delta}_{y}^{(k)}\circ\boldsymbol{\delta}_{\mu}^{(k)}$后

$$M^{(k+1)}Y^{(k+1)}\boldsymbol{e} \approx M^{(k)}Y^{(k)}\boldsymbol{e} + M^{(k)}\boldsymbol{\delta}_{y}^{(k)} + Y^{(k)}\boldsymbol{\delta}_{\mu}^{(k)} = \tau^{(k+1)}\boldsymbol{e}$$

$$M^{(k)}\boldsymbol{\delta}_{y}^{(k)} + Y^{(k)}\boldsymbol{\delta}_{\mu}^{(k)} = \tau^{(k+1)}\boldsymbol{e} - M^{(k)}Y^{(k)}\boldsymbol{e}$$

化简后，得方程组

$$A_E^T\boldsymbol{\delta}_{\lambda}^{(k)} + A_I^T\boldsymbol{\delta}_{\mu}^{(k)} = \boldsymbol{0}$$
$$A_E\boldsymbol{\delta}_{x}^{(k)} = \boldsymbol{0}$$
$$A_I\boldsymbol{\delta}_{x}^{(k)} - \boldsymbol{\delta}_{y}^{(k)} = \boldsymbol{0}$$
$$M^{(k)}\boldsymbol{\delta}_{y}^{(k)} + Y^{(k)}\boldsymbol{\delta}_{\mu}^{(k)} = \tau^{(k+1)}\boldsymbol{e} - M^{(k)}Y^{(k)}\boldsymbol{e}$$

$$\boldsymbol{\delta}_{z}^{(k)} = \left[\boldsymbol{\delta}_{x}^{(k)}, \boldsymbol{\delta}_{y}^{(k)}, \boldsymbol{\delta}_{\lambda}^{(k)}, \boldsymbol{\delta}_{\mu}^{(k)}\right]^T$$

方程组的矩阵式

$$\begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} & A_E^T & A_I^T \\ \boldsymbol{0} & M^{(k)} & \boldsymbol{0} & Y^{(k)} \\ A_E & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ A_I & -I & \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}\boldsymbol{\delta}_{z}^{(k)} = \begin{bmatrix} \boldsymbol{0} \\ \tau^{(k+1)}\boldsymbol{e} - M^{(k)}Y^{(k)}\boldsymbol{e} \\ \boldsymbol{0} \\ \boldsymbol{0} \end{bmatrix}$$

左乘 $-\left(Y^{(k)}\right)^{-1}$

$$\begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} & A_E^T & A_I^T \\ \boldsymbol{0} & -\left(Y^{(k)}\right)^{-1}M^{(k)} & \boldsymbol{0} & -I \\ A_E & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ A_I & -I & \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}\boldsymbol{\delta}_{z}^{(k)} = \begin{bmatrix} \boldsymbol{0} \\ -\tau^{(k+1)}\left(Y^{(k)}\right)^{-1}\boldsymbol{e} + \left(Y^{(k)}\right)^{-1}M^{(k)}Y^{(k)}\boldsymbol{e} \\ \boldsymbol{0} \\ \boldsymbol{0} \end{bmatrix}$$

其中，矩阵运算

$$\left(Y^{(k)}\right)^{-1}M^{(k)}Y^{(k)}\boldsymbol{e}=\begin{bmatrix} y_1^{(k)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & y_m^{(k)} \end{bmatrix}^{-1}\begin{bmatrix} \mu_1^{(1)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mu_m^{(m)} \end{bmatrix}\begin{bmatrix} y_1^{(k)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & y_m^{(k)} \end{bmatrix}\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

$$=\begin{bmatrix} \frac{1}{y_1^{(k)}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{y_m^{(k)}} \end{bmatrix}\begin{bmatrix} \mu_1^{(1)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mu_m^{(m)} \end{bmatrix}\begin{bmatrix} y_1^{(k)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & y_m^{(k)} \end{bmatrix}\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

$$=\begin{bmatrix} \frac{\mu_1^{(1)}}{y_1^{(k)}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{\mu_m^{(m)}}{y_m^{(k)}} \end{bmatrix}\begin{bmatrix} y_1^{(k)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & y_m^{(k)} \end{bmatrix}\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

$$=\begin{bmatrix} \mu_1^{(1)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mu_m^{(m)} \end{bmatrix}\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

$$=\boldsymbol{\mu}^{(k)}$$

$$\boldsymbol{\mu}^{(k)}=\begin{bmatrix} \mu_1^{(1)} \\ \vdots \\ \mu_m^{(m)} \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{0} & \mathbf{0} & A_E^T & A_I^T \\ \mathbf{0} & -\left(Y^{(k)}\right)^{-1}M^{(k)} & \mathbf{0} & -I \\ A_E & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ A_I & -I & \mathbf{0} & \mathbf{0} \end{bmatrix}\boldsymbol{\delta}_z^{(k)}=\begin{bmatrix} \mathbf{0} \\ -\tau^{(k+1)}\left(Y^{(k)}\right)^{-1}\boldsymbol{e}+\left(Y^{(k)}\right)^{-1}M^{(k)}Y^{(k)}\boldsymbol{e} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}=\begin{bmatrix} \mathbf{0} \\ -\boldsymbol{v}_y^{(k)} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}$$

$$\left(Y^{(k)}\right)^{-1}M^{(k)}Y^{(k)}\boldsymbol{e}=\boldsymbol{\mu}^{(k)}$$

$$\boldsymbol{\delta}_z^{(k)}=\begin{bmatrix} \boldsymbol{\delta}_x^{(k)} \\ \boldsymbol{\delta}_y^{(k)} \\ \boldsymbol{\delta}_\lambda^{(k)} \\ \boldsymbol{\delta}_\mu^{(k)} \end{bmatrix}$$

$$\begin{aligned}\boldsymbol{v}_y^{(k)}&=-\tau^{(k+1)}\left(Y^{(k)}\right)^{-1}\boldsymbol{e}+\left(Y^{(k)}\right)^{-1}M^{(k)}Y^{(k)}\boldsymbol{e}\\&=\boldsymbol{\mu}^{(k)}-\tau^{(k+1)}\left(Y^{(k)}\right)^{-1}\boldsymbol{e}\end{aligned}$$

$$\begin{bmatrix} \mathbf{0} & \mathbf{0} & A_E^T & A_I^T \\ \mathbf{0} & -\left(Y^{(k)}\right)^{-1}M^{(k)} & \mathbf{0} & -I \\ A_E & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ A_I & -I & \mathbf{0} & \mathbf{0} \end{bmatrix}\boldsymbol{\delta}_z^{(k)}=\begin{bmatrix} \mathbf{0} \\ -\boldsymbol{v}_y^{(k)} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \quad (9.2.15)$$

$$\begin{aligned}\boldsymbol{\delta}_y^{(k)}&=A_I\boldsymbol{\delta}_x^{(k)}\\ \boldsymbol{\delta}_\mu^{(k)}&=\boldsymbol{v}_y^{(k)}-\left(Y^{(k)}\right)^{-1}M^{(k)}\boldsymbol{\delta}_y^{(k)}\end{aligned}$$

(3) 简化方程组 $$\begin{bmatrix} \mathbf{0} & \mathbf{0} & A_E^T & A_I^T \\ \mathbf{0} & -\left(Y^{(k)}\right)^{-1}M^{(k)} & \mathbf{0} & -I \\ A_E & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ A_I & -I & \mathbf{0} & \mathbf{0} \end{bmatrix} \boldsymbol{\delta}_z^{(k)} = \begin{bmatrix} \mathbf{0} \\ -\boldsymbol{v}_y^{(k)} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}$$

$\mathbf{0}$	$\mathbf{0}$	A_E^T	A_I^T	$\mathbf{0}$
$\mathbf{0}$	$-\left(Y^{(k)}\right)^{-1}M^{(k)}$	$\mathbf{0}$	$-I$	$-\boldsymbol{v}_y^{(k)}$
A_E	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$
A_I	$-I$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$

$-\left(M^{(k)}\right)^{-1}Y^{(k)} \times ② \Longrightarrow ②$

$\mathbf{0}$	$\mathbf{0}$	A_E^T	A_I^T	$\mathbf{0}$
$\mathbf{0}$	I	$\mathbf{0}$	$\left(M^{(k)}\right)^{-1}Y^{(k)}$	$\left(M^{(k)}\right)^{-1}Y^{(k)}\boldsymbol{v}_y^{(k)}$
A_E	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$
A_I	$-I$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$

$\mathbf{0}$	$\mathbf{0}$	A_E^T	A_I^T	$\mathbf{0}$
$\mathbf{0}$	I	$\mathbf{0}$	$\left(M^{(k)}\right)^{-1}Y^{(k)}$	$\left(M^{(k)}\right)^{-1}Y^{(k)}\boldsymbol{v}_y^{(k)}$
A_E	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$
A_I	$-I$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$

$② + ④ \Longrightarrow ④$

$\mathbf{0}$	$\mathbf{0}$	A_E^T	A_I^T	$\mathbf{0}$
$\mathbf{0}$	I	$\mathbf{0}$	$\left(M^{(k)}\right)^{-1}Y^{(k)}$	$\left(M^{(k)}\right)^{-1}Y^{(k)}\boldsymbol{v}_y^{(k)}$
A_E	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$
A_I	$\mathbf{0}$	$\mathbf{0}$	$\left(M^{(k)}\right)^{-1}Y^{(k)}$	$\left(M^{(k)}\right)^{-1}Y^{(k)}\boldsymbol{v}_y^{(k)}$

$H^{(k)} = A_I^T\left(Y^{(k)}\right)^{-1}M^{(k)}A_I \qquad \boldsymbol{p}^{(k)} = -A_I^T\boldsymbol{v}_y^{(k)}$

$A_I^T\left(Y^{(k)}\right)^{-1}M^{(k)} \times ④ - ① \Longrightarrow ④$

$\mathbf{0}$	$\mathbf{0}$	A_E^T	A_I^T	$\mathbf{0}$
$\mathbf{0}$	I	$\mathbf{0}$	$\left(M^{(k)}\right)^{-1}Y^{(k)}$	$\left(M^{(k)}\right)^{-1}Y^{(k)}\boldsymbol{v}_y^{(k)}$
A_E	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$
$H^{(k)}$	$\mathbf{0}$	$-A_E^T$	$\mathbf{0}$	$-\boldsymbol{p}^{(k)}$

$\mathbf{0}$	$\mathbf{0}$	A_E^T	A_I^T	$\mathbf{0}$
$\mathbf{0}$	I	$\mathbf{0}$	$\left(M^{(k)}\right)^{-1}Y^{(k)}$	$\left(M^{(k)}\right)^{-1}Y^{(k)}\boldsymbol{v}_y^{(k)}$
A_E	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$
$H^{(k)}$	$\mathbf{0}$	$-A_E^T$	$\mathbf{0}$	$-\boldsymbol{p}^{(k)}$

$\times(-1)$

$$\boldsymbol{\delta}_z^{(k)} = \begin{bmatrix} \boldsymbol{\delta}_x^{(k)} \\ \boldsymbol{\delta}_y^{(k)} \\ \boldsymbol{\delta}_\lambda^{(k)} \\ \boldsymbol{\delta}_\mu^{(k)} \end{bmatrix}$$

$\mathbf{0}$	$\mathbf{0}$	A_E^T	A_I^T	$\mathbf{0}$
$\mathbf{0}$	I	$\mathbf{0}$	$\left(M^{(k)}\right)^{-1}Y^{(k)}$	$\left(M^{(k)}\right)^{-1}Y^{(k)}\boldsymbol{v}_y^{(k)}$
$-A_E$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$
$H^{(k)}$	$\mathbf{0}$	$-A_E^T$	$\mathbf{0}$	$-\boldsymbol{p}^{(k)}$

$H^{(k)} = A_I^T\left(Y^{(k)}\right)^{-1}M^{(k)}A_I \qquad \boldsymbol{p}^{(k)} = -A_I^T\boldsymbol{v}_y^{(k)}$

$\Downarrow$

$$\begin{bmatrix} H^{(k)} & -A_E^T \\ -A_E & \mathbf{0} \end{bmatrix}\begin{bmatrix} \boldsymbol{\delta}_x^{(k)} \\ \boldsymbol{\delta}_\lambda^{(k)} \end{bmatrix} = \begin{bmatrix} -\boldsymbol{p}^{(k)} \\ \mathbf{0} \end{bmatrix} \quad \Longrightarrow \quad \begin{aligned} & -A_E\boldsymbol{\delta}_x^{(k)} = \mathbf{0} \\ & H^{(k)}\boldsymbol{\delta}_x^{(k)} - A_E^T\boldsymbol{\delta}_\lambda^{(k)} = -\boldsymbol{p}^{(k)} \end{aligned}$$

二次规划与方程组之间的关系

$$\begin{aligned} & \min q(\boldsymbol{u}) = \frac{1}{2}\boldsymbol{u}^TQ\boldsymbol{u} + \boldsymbol{u}^T\boldsymbol{p} \\ & \text{s.t. } A\boldsymbol{u} = \mathbf{0} \end{aligned} \quad \longleftrightarrow \quad \begin{aligned} & \min q(\boldsymbol{u}) = \frac{1}{2}\boldsymbol{u}^TQ\boldsymbol{u} + \boldsymbol{u}^T\boldsymbol{p} \\ & \text{s.t. } -A\boldsymbol{u} = \mathbf{0} \end{aligned}$$

$\updownarrow \qquad\qquad \downarrow$

$$\begin{aligned} & -A\boldsymbol{u} = \mathbf{0} \\ & Q\boldsymbol{u} - A^T\boldsymbol{\lambda} = -\boldsymbol{p} \end{aligned} \quad \longleftrightarrow \quad \begin{aligned} & L(\boldsymbol{u},\boldsymbol{\lambda}) = \frac{1}{2}\boldsymbol{u}^TQ\boldsymbol{u} + \boldsymbol{u}^T\boldsymbol{p} + \boldsymbol{\lambda}^T(-A\boldsymbol{u}) \\ & \nabla L(\boldsymbol{u},\boldsymbol{\lambda}) = Q\boldsymbol{u} + \boldsymbol{p} - A^T\boldsymbol{\lambda} = \mathbf{0} \\ & -A\boldsymbol{u} = \mathbf{0} \end{aligned}$$

$$\begin{aligned} & -A_E\boldsymbol{\delta}_x^{(k)} = \mathbf{0} \\ & H^{(k)}\boldsymbol{\delta}_x^{(k)} - A_E^T\boldsymbol{\delta}_\lambda^{(k)} = -\boldsymbol{p}^{(k)} \end{aligned} \quad \longleftrightarrow \quad \begin{aligned} & \min \frac{1}{2}\left[\boldsymbol{\delta}_x^{(k)}\right]^T H^{(k)}\boldsymbol{\delta}_x^{(k)} + \left[\boldsymbol{\delta}_x^{(k)}\right]^T\boldsymbol{p}^{(k)} \\ & \text{s.t. } A_E\boldsymbol{\delta}_x^{(k)} = \mathbf{0} \end{aligned}$$

(4) 求解二次规划问题，得到$\boldsymbol{\delta}_x^{(k)}$和$\boldsymbol{\delta}_\lambda^{(k)}$

$$\min \frac{1}{2}\left[\boldsymbol{\delta}_x^{(k)}\right]^T H^{(k)}\boldsymbol{\delta}_x^{(k)} + \left[\boldsymbol{\delta}_x^{(k)}\right]^T \boldsymbol{p}^{(k)}$$
$$\text{s.t.}\ A_E\boldsymbol{\delta}_x^{(k)} = \mathbf{0}$$

调用MATLAB函数quadprog

可以同时得到，$\boldsymbol{\delta}_x^{(k)}$和$\boldsymbol{\delta}_\lambda^{(k)}$
计算稳定性好，效率高

直接解

$$-A_E\boldsymbol{\delta}_x^{(k)} = \mathbf{0}$$
$$H^{(k)}\boldsymbol{\delta}_x^{(k)} - A_E^T\boldsymbol{\delta}_\lambda^{(k)} = -\boldsymbol{p}^{(k)}$$

$\Longrightarrow$

$$\boldsymbol{\delta}_x^{(k)} - \left(H^{(k)}\right)^{-1}A_E^T\boldsymbol{\delta}_\lambda^{(k)} = -\left(H^{(k)}\right)^{-1}\boldsymbol{p}^{(k)}$$

$$\underbrace{A_E\boldsymbol{\delta}_x^{(k)}}_{0} - A_E\left(H^{(k)}\right)^{-1}A_E^T\boldsymbol{\delta}_\lambda^{(k)} = -A_E\left(H^{(k)}\right)^{-1}\boldsymbol{p}^{(k)}$$

$$\left[A_E\left(H^{(k)}\right)^{-1}A_E^T\right]\boldsymbol{\delta}_\lambda^{(k)} = A_E\left(H^{(k)}\right)^{-1}\boldsymbol{p}^{(k)}$$

解线性方程组，得 $\boldsymbol{\delta}_\lambda^{(k)}$

$$\boldsymbol{\delta}_x^{(k)} = \left(H^{(k)}\right)^{-1}A_E^T\boldsymbol{\delta}_\lambda^{(k)} - \left(H^{(k)}\right)^{-1}\boldsymbol{p}^{(k)}$$

一般不用此方法，
因，矩阵求拟不稳定

(5) 求解$\boldsymbol{\delta}_y^{(k)}$和$\boldsymbol{\delta}_\mu^{(k)}$

$$\boldsymbol{\delta}_y^{(k)} = A_I\boldsymbol{\delta}_x^{(k)}$$
$$\boldsymbol{\delta}_\mu^{(k)} = \boldsymbol{v}_y^{(k)} - \left(Y^{(k)}\right)^{-1}M^{(k)}\boldsymbol{\delta}_y^{(k)}$$

这样，新的迭代点为

$$\begin{cases} \boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha_P^{(k)}\boldsymbol{\delta}_x^{(k)} \\ \boldsymbol{y}^{(k+1)} = \boldsymbol{y}^{(k)} + \alpha_P^{(k)}\boldsymbol{\delta}_y^{(k)} \\ \boldsymbol{\lambda}^{(k+1)} = \boldsymbol{\lambda}^{(k)} + \alpha_D^{(k)}\boldsymbol{\delta}_\lambda^{(k)} \\ \boldsymbol{\mu}^{(k+1)} = \boldsymbol{\mu}^{(k)} + \alpha_D^{(k)}\boldsymbol{\delta}_\mu^{(k)} \end{cases}$$

3.步长求解

原-对偶可行路径跟踪法要求所有迭代点$(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\lambda},\boldsymbol{\mu})$都满足$\boldsymbol{y}\geq\boldsymbol{0}$和$\boldsymbol{\mu}\geq\boldsymbol{0}$

$$\begin{cases}\boldsymbol{y}^{(k)}+\alpha_P^{(k)}\boldsymbol{\delta}_y^{(k)}>0\\ \boldsymbol{\mu}^{(k)}+\alpha_D^{(k)}\boldsymbol{\delta}_\mu^{(k)}>0\end{cases}$$

$$\begin{cases}\alpha_{P,min}^{(k)}=\min\left\{-\dfrac{\left(\boldsymbol{y}^{(k)}\right)_i}{\left(\boldsymbol{\delta}_y^{(k)}\right)_i}\middle|\left(\boldsymbol{\delta}_y^{(k)}\right)_i<0,i=1,2,\dots,m_2\right\}\\ \alpha_P^{(k)}=\min\left\{1,c\cdot\alpha_{P,min}^{(k)}\right\}\end{cases}$$

$$\begin{cases}\alpha_{D,min}^{(k)}=\min\left\{-\dfrac{\left(\boldsymbol{\mu}^{(k)}\right)_i}{\left(\boldsymbol{\delta}_{\boldsymbol{\mu}}^{(k)}\right)_i}\middle|\left(\boldsymbol{\delta}_{\boldsymbol{\mu}}^{(k)}\right)_i<0,i=1,2,\dots,m_2\right\}\\ \alpha_D^{(k)}=\min\left\{1,c\cdot\alpha_{D,min}^{(k)}\right\}\end{cases}$$

通常

$$c=1-10^{-3}$$
$$1-10^{-3}\leq c\leq 1-10^{-6}$$

也有采用

$$\alpha^{(k)}=\min\left\{\alpha_P^{(k)},\alpha_D^{(k)}\right\}$$

4.中心参数的更新公式

中心参数：缩减因子$\tau\to 0.$

需要满足：

①保证下一个迭代点$\boldsymbol{z}^{(k+1)}=\left[\boldsymbol{x}^{(k+1)},\boldsymbol{y}^{(k+1)},\boldsymbol{\lambda}^{(k+1)},\boldsymbol{\mu}^{(k+1)}\right]^T$仍满足$\boldsymbol{y}^{(k+1)}>0$和$\boldsymbol{\mu}^{(k+1)}>0$

②使得对偶间隔δ_{PD}越来越小

③使得迭代点离中心轨迹越来越近

研究成果：

$$\sigma^{(k)}=\frac{m_2}{m_2+\rho},\qquad \rho>\sqrt{m_2}$$

$$\tau^{(k+1)}=\sigma^{(k)}\tau^{(k)}=\frac{m_2\tau^{(k)}}{m_2+\rho}$$

9.2.3 原-对偶可行路径跟踪法的计算步骤

步骤1： 输入参数$\boldsymbol{c}, A_E, \boldsymbol{b}_E, A_I, \boldsymbol{b}_I$，选定初始点$\boldsymbol{z}^{(0)} = \left(\boldsymbol{x}^{(0)}, \boldsymbol{y}^{(0)}, \boldsymbol{\lambda}^{(0)}, \boldsymbol{\mu}^{(0)}\right)$
设定精度tol，令$k = 0$

步骤2： 计算缩减因子 $\sigma^{(k)} = \dfrac{m_2}{m_2 + \rho}$, $\rho > \sqrt{m_2}$

$$\tau^{(k+1)} = \sigma^{(k)}\tau^{(k)} = \frac{m_2 \tau^{(k)}}{m_2 + \rho}$$

求解$\boldsymbol{\delta}_x^{(k)}$和$\boldsymbol{\delta}_\lambda^{(k)}$

$$\min \frac{1}{2}\left[\boldsymbol{\delta}_x^{(k)}\right]^T H^{(k)} \boldsymbol{\delta}_x^{(k)} + \left[\boldsymbol{\delta}_x^{(k)}\right]^T \boldsymbol{p}^{(k)}$$
$$\text{s.t.}\, A_E \boldsymbol{\delta}_x^{(k)} = \boldsymbol{0}$$

$$H^{(k)} = A_I^T \left(Y^{(k)}\right)^{-1} M^{(k)} A_I$$
$$\boldsymbol{p}^{(k)} = -A_I^T \boldsymbol{v}_y^{(k)}$$

调用MATLAB函数quadprog
可以同时得到，$\boldsymbol{\delta}_x^{(k)}$和$\boldsymbol{\delta}_\lambda^{(k)}$
计算稳定性好，效率高

求解$\boldsymbol{\delta}_y^{(k)}$和$\boldsymbol{\delta}_\mu^{(k)}$

$$\boldsymbol{\delta}_y^{(k)} = A_I \boldsymbol{\delta}_x^{(k)}$$
$$\boldsymbol{\delta}_\mu^{(k)} = \boldsymbol{v}_y^{(k)} - \left(Y^{(k)}\right)^{-1} M^{(k)} \boldsymbol{\delta}_y^{(k)}$$

步骤3： 计算步长

$$\begin{cases} \alpha_{P,min}^{(k)} = \min\left\{-\dfrac{\left(\boldsymbol{y}^{(k)}\right)_i}{\left(\boldsymbol{\delta}_y^{(k)}\right)_i} \,\middle|\, \left(\boldsymbol{\delta}_y^{(k)}\right)_i < 0, i = 1,2,\dots,m_2\right\} \\ \alpha_P^{(k)} = \min\left\{1, c \cdot \alpha_{P,min}^{(k)}\right\} \end{cases}$$

$$\begin{cases} \alpha_{D,min}^{(k)} = \min\left\{-\dfrac{\left(\boldsymbol{\mu}^{(k)}\right)_i}{\left(\boldsymbol{\delta}_{\boldsymbol{\mu}}^{(k)}\right)_i} \,\middle|\, \left(\boldsymbol{\delta}_{\boldsymbol{\mu}}^{(k)}\right)_i < 0, i = 1,2,\dots,m_2\right\} \\ \alpha_D^{(k)} = \min\left\{1, c \cdot \alpha_{D,min}^{(k)}\right\} \end{cases}$$

通常
$c = 1 - 10^{-3}$

步骤4： 计算新的迭代点

$$\begin{cases} \boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha_P^{(k)} \boldsymbol{\delta}_x^{(k)} \\ \boldsymbol{y}^{(k+1)} = \boldsymbol{y}^{(k)} + \alpha_P^{(k)} \boldsymbol{\delta}_y^{(k)} \\ \boldsymbol{\lambda}^{(k+1)} = \boldsymbol{\lambda}^{(k)} + \alpha_D^{(k)} \boldsymbol{\delta}_\lambda^{(k)} \\ \boldsymbol{\mu}^{(k+1)} = \boldsymbol{\mu}^{(k)} + \alpha_D^{(k)} \boldsymbol{\delta}_\mu^{(k)} \end{cases}$$

步骤5： 计算新的对偶间隔 $\delta_{PD}^{(k+1)} = \left[\boldsymbol{\mu}^{(k+1)}\right]^T \boldsymbol{y}^{(k+1)}$

步骤6： 如果$\delta_{PD}^{(k+1)} < tol$，迭代终止；
$f\left(\boldsymbol{x}^{(k+1)}\right) = \boldsymbol{c}^T \boldsymbol{x}^{(k+1)}$为目标函数极小值， $\boldsymbol{z}^{(k+1)}$为原-对偶解。
否则， $k = k + 1$， 转到步骤2

9.2.6 实例测试

例9.2 用原-对偶可行路径跟踪法求解

$$\begin{gathered} \max f(x) = x_1 + x_2 + 5x_3 \\ \text{s.t.}\, 3x_1 + 2x_2 + \frac{1}{4}x_3 \le 6 \\ x_3 \le 4 \\ \boldsymbol{x} \ge \boldsymbol{0} \end{gathered}$$

$$\begin{gathered} \min -f(x) = -x_1 - x_2 - 5x_3 \\ \text{s.t.}\, -3x_1 - 2x_2 - \frac{1}{4}x_3 \ge -6 \\ -x_3 \ge -4 \\ \boldsymbol{x} \ge \boldsymbol{0} \end{gathered}$$

初始点$\boldsymbol{x}^{(0)} = (0.612, 0.9269, 2.0349)$, $tol = 1 \times 10^{-4}$

example_9_2_XinggaoLiu.m

x_optimal =
0.0000
2.4999
4.0000

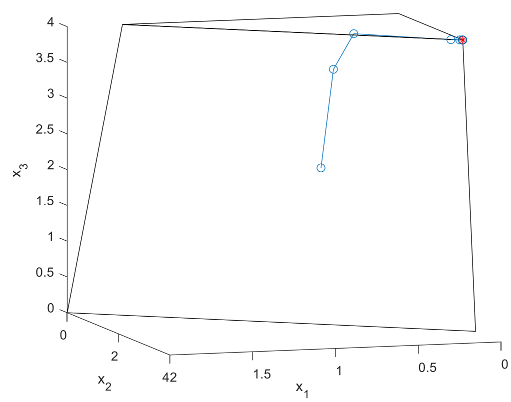
y_optimal =
0.0000
0.0000
0.0000
2.5000
4.0000

f_optimal = 22.4999

k = 8

lamda_optimal = [ ]

mu_optimal =
0.5000
4.8750
0.5000
0.0000
0.0000



9.3 原-对偶非可行路径跟踪法

9.3.1 原-对偶非可行路径跟踪法的基本原理

1. 扰动KKT条件的线性化及求解

(1) 对当前迭代点$\boldsymbol{z}^{(k)}=\left[\boldsymbol{x}^{(k)},\boldsymbol{y}^{(k)},\boldsymbol{\lambda}^{(k)},\boldsymbol{\mu}^{(k)}\right]^T$做适当的扰动$\boldsymbol{\delta}_{\boldsymbol{z}}^{(k)}=\left[\boldsymbol{\delta}_x^{(k)},\boldsymbol{\delta}_y^{(k)},\boldsymbol{\delta}_\lambda^{(k)},\boldsymbol{\delta}_\mu^{(k)}\right]^T$，得到下一个迭代点$\boldsymbol{z}^{(k+1)}=\left[\boldsymbol{x}^{(k+1)},\boldsymbol{y}^{(k+1)},\boldsymbol{\lambda}^{(k+1)},\boldsymbol{\mu}^{(k+1)}\right]^T$

$$\boldsymbol{z}^{(k+1)}=\boldsymbol{z}^{(k)}+\boldsymbol{\delta}_{\boldsymbol{z}}^{(k)}$$

(2) 将点$\boldsymbol{z}^{(k+1)}$带入扰动KKT条件，略去关于扰动量的二次项，得

$$\begin{aligned}
&\boldsymbol{c}-A_E^T\left(\boldsymbol{\lambda}^{(k)}+\boldsymbol{\delta}_\lambda^{(k)}\right)-A_I^T\left(\boldsymbol{\mu}^{(k)}+\boldsymbol{\delta}_\mu^{(k)}\right)=\boldsymbol{0}\\
&A_E\left(\boldsymbol{x}^{(k)}+\boldsymbol{\delta}_x^{(k)}\right)-\boldsymbol{b}_E=\boldsymbol{0}\\
&A_I\left(\boldsymbol{x}^{(k)}+\boldsymbol{\delta}_x^{(k)}\right)-\left(\boldsymbol{y}^{(k)}+\boldsymbol{\delta}_y^{(k)}\right)=\boldsymbol{b}_I\\
&M^{(k+1)}Y^{(k+1)}\boldsymbol{e}\approx M^{(k)}Y^{(k)}\boldsymbol{e}+M^{(k)}\boldsymbol{\delta}_y^{(k)}+Y^{(k)}\boldsymbol{\delta}_\mu^{(k)}=\tau^{(k+1)}\boldsymbol{e}
\end{aligned}$$

预备：矩阵运算

$$
\begin{aligned}
\left(Y^{(k)}\right)^{-1}M^{(k)}Y^{(k)}\boldsymbol{e} &= \begin{bmatrix} y_1^{(k)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & y_m^{(k)} \end{bmatrix}^{-1} \begin{bmatrix} \mu_1^{(1)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mu_m^{(m)} \end{bmatrix} \begin{bmatrix} y_1^{(k)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & y_m^{(k)} \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{y_1^{(k)}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{y_m^{(k)}} \end{bmatrix} \begin{bmatrix} \mu_1^{(1)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mu_m^{(m)} \end{bmatrix} \begin{bmatrix} y_1^{(k)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & y_m^{(k)} \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \\
&= \begin{bmatrix} \frac{\mu_1^{(1)}}{y_1^{(k)}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{\mu_m^{(m)}}{y_m^{(k)}} \end{bmatrix} \begin{bmatrix} y_1^{(k)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & y_m^{(k)} \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \\
&= \begin{bmatrix} \mu_1^{(1)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mu_m^{(m)} \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \\
&= \boldsymbol{\mu}^{(k)}
\end{aligned}
$$

$$
\boldsymbol{\mu}^{(k)} = \begin{bmatrix} \mu_1^{(1)} \\ \vdots \\ \mu_m^{(m)} \end{bmatrix}
$$

注意：当前点$\boldsymbol{z}^{(k)}$不是可行域的内点

含有$\boldsymbol{\delta}$的项放在左边

$$
\begin{aligned}
\boldsymbol{c} - A_E^T\left(\boldsymbol{\lambda}^{(k)} + \boldsymbol{\delta}_\lambda^{(k)}\right) - A_I^T\left(\boldsymbol{\mu}^{(k)} + \boldsymbol{\delta}_\mu^{(k)}\right) &= \boldsymbol{0} \\
A_E\left(\boldsymbol{x}^{(k)} + \boldsymbol{\delta}_x^{(k)}\right) - \boldsymbol{b}_E &= \boldsymbol{0} \\
A_I\left(\boldsymbol{x}^{(k)} + \boldsymbol{\delta}_x^{(k)}\right) - \left(\boldsymbol{y}^{(k)} + \boldsymbol{\delta}_y^{(k)}\right) &= \boldsymbol{b}_I \\
M^{(k)}Y^{(k)}\boldsymbol{e} + M^{(k)}\boldsymbol{\delta}_y^{(k)} + Y^{(k)}\boldsymbol{\delta}_\mu^{(k)} &= \tau^{(k+1)}\boldsymbol{e}
\end{aligned}
\quad \Longrightarrow \quad
\begin{aligned}
A_E^T\boldsymbol{\delta}_\lambda^{(k)} + A_I^T\boldsymbol{\delta}_\mu^{(k)} &= \boldsymbol{c} - A_E^T\boldsymbol{\lambda}^{(k)} - A_I^T\boldsymbol{\mu}^{(k)} \\
A_E\boldsymbol{\delta}_x^{(k)} &= \boldsymbol{b}_E - A_E\boldsymbol{x}^{(k)} \\
A_I\boldsymbol{\delta}_x^{(k)} - \boldsymbol{\delta}_y^{(k)} &= \boldsymbol{b}_I - A_I\boldsymbol{x}^{(k)} + \boldsymbol{y}^{(k)} \\
M^{(k)}\boldsymbol{\delta}_y^{(k)} + Y^{(k)}\boldsymbol{\delta}_\mu^{(k)} &= \tau^{(k+1)}\boldsymbol{e} - M^{(k)}Y^{(k)}\boldsymbol{e}
\end{aligned}
$$

对第4个方程，左乘$\left(Y^{(k)}\right)^{-1}$

$$
\left(Y^{(k)}\right)^{-1}M^{(k)}\boldsymbol{\delta}_y^{(k)} + \boldsymbol{\delta}_\mu^{(k)} = \tau^{(k+1)}\left(Y^{(k)}\right)^{-1}\boldsymbol{e} - \left(Y^{(k)}\right)^{-1}M^{(k)}Y^{(k)}\boldsymbol{e} = \tau^{(k+1)}\left(Y^{(k)}\right)^{-1}\boldsymbol{e} - \boldsymbol{\mu}^{(k)}
$$

$$
\Longrightarrow \quad
\begin{aligned}
A_E^T\boldsymbol{\delta}_\lambda^{(k)} + A_I^T\boldsymbol{\delta}_\mu^{(k)} &= \boldsymbol{c} - A_E^T\boldsymbol{\lambda}^{(k)} - A_I^T\boldsymbol{\mu}^{(k)} &&\triangleq \boldsymbol{v}_x^{(k)} \\
A_E\boldsymbol{\delta}_x^{(k)} &= \boldsymbol{b}_E - A_E\boldsymbol{x}^{(k)} &&\triangleq -\boldsymbol{v}_\lambda^{(k)} \\
A_I\boldsymbol{\delta}_x^{(k)} - \boldsymbol{\delta}_y^{(k)} &= \boldsymbol{b}_I - A_I\boldsymbol{x}^{(k)} + \boldsymbol{y}^{(k)} &&\triangleq \boldsymbol{v}_\mu^{(k)} \\
\left(Y^{(k)}\right)^{-1}M^{(k)}\boldsymbol{\delta}_y^{(k)} + \boldsymbol{\delta}_\mu^{(k)} &= \tau^{(k+1)}\left(Y^{(k)}\right)^{-1}\boldsymbol{e} - \boldsymbol{\mu}^{(k)} &&\triangleq -\boldsymbol{v}_y^{(k)}
\end{aligned}
\quad \Longrightarrow \quad
\begin{aligned}
A_E^T\boldsymbol{\delta}_\lambda^{(k)} + A_I^T\boldsymbol{\delta}_\mu^{(k)} &= \boldsymbol{v}_x^{(k)} \\
-\left(Y^{(k)}\right)^{-1}M^{(k)}\boldsymbol{\delta}_y^{(k)} - \boldsymbol{\delta}_\mu^{(k)} &= -\boldsymbol{v}_y^{(k)} \\
A_E\boldsymbol{\delta}_x^{(k)} &= -\boldsymbol{v}_\lambda^{(k)} \\
A_I\boldsymbol{\delta}_x^{(k)} - \boldsymbol{\delta}_y^{(k)} &= \boldsymbol{v}_\mu^{(k)}
\end{aligned}
$$

$$\begin{aligned} A_E^T\boldsymbol{\delta}_\lambda^{(k)} + A_I^T\boldsymbol{\delta}_\mu^{(k)} &= \boldsymbol{v}_x^{(k)} \\ -\left(Y^{(k)}\right)^{-1}M^{(k)}\boldsymbol{\delta}_y^{(k)} - \boldsymbol{\delta}_\mu^{(k)} &= -\boldsymbol{v}_y^{(k)} \\ A_E\boldsymbol{\delta}_x^{(k)} &= -\boldsymbol{v}_\lambda^{(k)} \\ A_I\boldsymbol{\delta}_x^{(k)} - \boldsymbol{\delta}_y^{(k)} &= \boldsymbol{v}_\mu^{(k)} \end{aligned} \Rightarrow \begin{bmatrix} \mathbf{0} & \mathbf{0} & A_E^T & A_I^T \\ \mathbf{0} & -\left(Y^{(k)}\right)^{-1}M^{(k)} & \mathbf{0} & -I \\ A_E & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ A_I & -I & \mathbf{0} & \mathbf{0} \end{bmatrix} \boldsymbol{\delta}_z^{(k)} = \begin{bmatrix} \boldsymbol{v}_x^{(k)} \\ -\boldsymbol{v}_y^{(k)} \\ -\boldsymbol{v}_\lambda^{(k)} \\ \boldsymbol{v}_\mu^{(k)} \end{bmatrix}$$

注意：$\left(\boldsymbol{x}^{(k+1)}, \boldsymbol{y}^{(k+1)}\right)$不要求是原问题的可行点，$\left(\boldsymbol{\lambda}^{(k+1)}, \boldsymbol{\mu}^{(k+1)}\right)$也不要求是对偶问题的可行点

$$\begin{bmatrix} \boldsymbol{v}_x^{(k)} \\ -\boldsymbol{v}_y^{(k)} \\ -\boldsymbol{v}_\lambda^{(k)} \\ \boldsymbol{v}_\mu^{(k)} \end{bmatrix} = \begin{bmatrix} \boldsymbol{c} - A_E^T\boldsymbol{\lambda}^{(k)} - A_I^T\boldsymbol{\mu}^{(k)} \\ \boldsymbol{\mu}^{(k)} - \tau^{(k+1)}\left(Y^{(k)}\right)^{-1}\boldsymbol{e} \\ \boldsymbol{b}_E - A_E\boldsymbol{x}^{(k)} \\ \boldsymbol{b}_I - A_I\boldsymbol{x}^{(k)} + \boldsymbol{y}^{(k)} \end{bmatrix}$$

当$\boldsymbol{v}_\lambda^{(k)} = \boldsymbol{v}_\mu^{(k)} = 0$时，$\boldsymbol{x}^{(k)}$是原问题的可行点

(3) 简化方程组，求解扰动向量

与9.2.2类似，可得

$$\Longrightarrow \begin{bmatrix} A_I^T\left(Y^{(k)}\right)^{-1}M^{(k)}A_I & -A_E^T \\ -A_E & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\delta}_x^{(k)} \\ \boldsymbol{\delta}_\lambda^{(k)} \end{bmatrix} = \begin{bmatrix} A_I^T\left[\boldsymbol{v}_y^{(k)} + \left(Y^{(k)}\right)^{-1}M^{(k)}\boldsymbol{v}_\mu^{(k)}\right] - \boldsymbol{v}_x^{(k)} \\ \boldsymbol{v}_\lambda^{(k)} \end{bmatrix}$$

$$\Downarrow$$

$$\begin{bmatrix} H^{(k)} & -A_E^T \\ -A_E & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\delta}_x^{(k)} \\ \boldsymbol{\delta}_\lambda^{(k)} \end{bmatrix} = \begin{bmatrix} -\boldsymbol{p}^{(k)} \\ \boldsymbol{v}_\lambda^{(k)} \end{bmatrix} \Longrightarrow \begin{aligned} H^{(k)}\boldsymbol{\delta}_x^{(k)} - A_E^T\boldsymbol{\delta}_\lambda^{(k)} &= -\boldsymbol{p}^{(k)} \\ -A_E\boldsymbol{\delta}_x^{(k)} &= \boldsymbol{v}_\lambda^{(k)} \end{aligned}$$

$$H^{(k)} = A_I^T\left(Y^{(k)}\right)^{-1}M^{(k)}A_I$$

$$\boldsymbol{p}^{(k)} = \boldsymbol{v}_x^{(k)} - A_I^T\left[\boldsymbol{v}_y^{(k)} + \left(Y^{(k)}\right)^{-1}M^{(k)}\boldsymbol{v}_\mu^{(k)}\right]$$

(4) 求解二次规划问题，得到$\boldsymbol{\delta}_x^{(k)}$和$\boldsymbol{\delta}_\lambda^{(k)}$

$$\min \frac{1}{2}\left[\boldsymbol{\delta}_x^{(k)}\right]^T H^{(k)} \boldsymbol{\delta}_x^{(k)} + \left[\boldsymbol{\delta}_x^{(k)}\right]^T \boldsymbol{p}^{(k)}$$
$$\text{s.t.}\, A_E \boldsymbol{\delta}_x^{(k)} = -\boldsymbol{v}_\lambda^{(k)}$$

调用MATLAB函数quadprog

可以同时得到，$\boldsymbol{\delta}_x^{(k)}$和$\boldsymbol{\delta}_\lambda^{(k)}$
计算稳定性好，效率高

直接解

$$H^{(k)} \boldsymbol{\delta}_x^{(k)} - A_E^T \boldsymbol{\delta}_\lambda^{(k)} = -\boldsymbol{p}^{(k)}$$
$$-A_E \boldsymbol{\delta}_x^{(k)} = \boldsymbol{v}_\lambda^{(k)}$$

$$\Longrightarrow$$

$$\boldsymbol{\delta}_x^{(k)} - \left(H^{(k)}\right)^{-1} A_E^T \boldsymbol{\delta}_\lambda^{(k)} = -\left(H^{(k)}\right)^{-1} \boldsymbol{p}^{(k)}$$
$$A_E \boldsymbol{\delta}_x^{(k)} - A_E\left(H^{(k)}\right)^{-1} A_E^T \boldsymbol{\delta}_\lambda^{(k)} = -A_E\left(H^{(k)}\right)^{-1} \boldsymbol{p}^{(k)}$$
$$\left[A_E\left(H^{(k)}\right)^{-1} A_E^T\right] \boldsymbol{\delta}_\lambda^{(k)} = A_E\left(H^{(k)}\right)^{-1} \boldsymbol{p}^{(k)} - \boldsymbol{v}_\lambda^{(k)}$$

解线性方程组，得 $\boldsymbol{\delta}_\lambda^{(k)}$

$$\boldsymbol{\delta}_x^{(k)} = \left(H^{(k)}\right)^{-1} A_E^T \boldsymbol{\delta}_\lambda^{(k)} - \left(H^{(k)}\right)^{-1} \boldsymbol{p}^{(k)}$$

3.步长和中心参数计算 与（9.2.2）一样 $\boldsymbol{y} \geq \boldsymbol{0}$, $\boldsymbol{\mu} \geq \boldsymbol{0}$

$$\begin{cases} \alpha_{P,min}^{(k)} = \min\left\{ -\dfrac{\left(\boldsymbol{y}^{(k)}\right)_i}{\left(\boldsymbol{\delta}_y^{(k)}\right)_i} \,\middle|\, \left(\boldsymbol{\delta}_y^{(k)}\right)_i < 0, i = 1,2,\ldots,m_2 \right\} \\ \alpha_P^{(k)} = \min\left\{1, c \cdot \alpha_{P,min}^{(k)}\right\} \end{cases}$$

$$\begin{cases} \alpha_{D,min}^{(k)} = \min\left\{ -\dfrac{\left(\boldsymbol{\mu}^{(k)}\right)_i}{\left(\boldsymbol{\delta}_{\boldsymbol{\mu}}^{(k)}\right)_i} \,\middle|\, \left(\boldsymbol{\delta}_{\boldsymbol{\mu}}^{(k)}\right)_i < 0, i = 1,2,\ldots,m_2 \right\} \\ \alpha_D^{(k)} = \min\left\{1, c \cdot \alpha_{D,min}^{(k)}\right\} \end{cases}$$

通常

$$c = 1 - 10^{-3}$$
$$1 - 10^{-3} \leq c \leq 1 - 10^{-6}$$

$$\sigma^{(k)} = \frac{m_2}{m_2 + \rho}, \qquad \rho > \sqrt{m_2}$$
$$\tau^{(k+1)} = \sigma^{(k)} \tau^{(k)} = \frac{m_2 \tau^{(k)}}{m_2 + \rho}$$

9.3.2 原-对偶非可行路径跟踪法的计算步骤

步骤1： 输入参数$\boldsymbol{c}, A_E, \boldsymbol{b}_E, A_I, \boldsymbol{b}_I$，选定初始点$\boldsymbol{z}^{(0)} = \left(\boldsymbol{x}^{(0)}, \boldsymbol{y}^{(0)}, \boldsymbol{\lambda}^{(0)}, \boldsymbol{\mu}^{(0)}\right)$
设定精度tol，令$k = 0$

步骤2： 计算缩减因子 $$\sigma^{(k)} = \frac{m_2}{m_2 + \rho}, \qquad \rho > \sqrt{m_2}$$

$$\tau^{(k+1)} = \sigma^{(k)} \tau^{(k)} = \frac{m_2 \tau^{(k)}}{m_2 + \rho}$$

求解$\boldsymbol{\delta}_x^{(k)}$ 和$\boldsymbol{\delta}_\lambda^{(k)}$

$$\min \frac{1}{2}\left[\boldsymbol{\delta}_x^{(k)}\right]^T H^{(k)} \boldsymbol{\delta}_x^{(k)} + \left[\boldsymbol{\delta}_x^{(k)}\right]^T \boldsymbol{p}^{(k)}$$
$$\text{s.t.}\, A_E \boldsymbol{\delta}_x^{(k)} = -\boldsymbol{v}_\lambda^{(k)}$$

调用MATLAB函数quadprog
可以同时得到，$\boldsymbol{\delta}_x^{(k)}$和$\boldsymbol{\delta}_\lambda^{(k)}$
计算稳定性好，效率高

求解$\boldsymbol{\delta}_y^{(k)}$和$\boldsymbol{\delta}_\mu^{(k)}$

$$\boldsymbol{\delta}_y^{(k)} = A_I \boldsymbol{\delta}_x^{(k)} - \boldsymbol{v}_\mu^{(k)}$$
$$\boldsymbol{\delta}_\mu^{(k)} = \boldsymbol{v}_y^{(k)} - \left(Y^{(k)}\right)^{-1} M^{(k)} \boldsymbol{\delta}_y^{(k)}$$

步骤3： 计算步长

$$\begin{cases} \alpha_{P,min}^{(k)} = \min\left\{ -\dfrac{\left(\boldsymbol{y}^{(k)}\right)_i}{\left(\boldsymbol{\delta}_y^{(k)}\right)_i} \middle| \left(\boldsymbol{\delta}_y^{(k)}\right)_i < 0, i = 1,2,\dots,m_2 \right\} \\ \alpha_P^{(k)} = \min\left\{1, c \cdot \alpha_{P,min}^{(k)}\right\} \end{cases}$$

$$\begin{cases} \alpha_{D,min}^{(k)} = \min\left\{ -\dfrac{\left(\boldsymbol{\mu}^{(k)}\right)_i}{\left(\boldsymbol{\delta}_{\boldsymbol{\mu}}^{(k)}\right)_i} \middle| \left(\boldsymbol{\delta}_{\boldsymbol{\mu}}^{(k)}\right)_i < 0, i = 1,2,\dots,m_2 \right\} \\ \alpha_D^{(k)} = \min\left\{1, c \cdot \alpha_{D,min}^{(k)}\right\} \end{cases}$$

通常
$c = 1 - 10^{-3}$

步骤4：计算新的迭代点

$$\begin{cases} \boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha_P^{(k)} \boldsymbol{\delta}_x^{(k)} \\ \boldsymbol{y}^{(k+1)} = \boldsymbol{y}^{(k)} + \alpha_P^{(k)} \boldsymbol{\delta}_y^{(k)} \\ \boldsymbol{\lambda}^{(k+1)} = \boldsymbol{\lambda}^{(k)} + \alpha_D^{(k)} \boldsymbol{\delta}_\lambda^{(k)} \\ \boldsymbol{\mu}^{(k+1)} = \boldsymbol{\mu}^{(k)} + \alpha_D^{(k)} \boldsymbol{\delta}_\mu^{(k)} \end{cases}$$

步骤5：计算新的对偶间隔 $\delta_{PD}^{(k+1)} = \left[\boldsymbol{\mu}^{(k+1)}\right]^T \boldsymbol{y}^{(k+1)}$

步骤6：如果$\delta_{PD}^{(k+1)} < tol$，$f\left(\boldsymbol{x}^{(k+1)}\right) = \boldsymbol{c}^T \boldsymbol{x}^{(k+1)}$，$\boldsymbol{z}^{(k+1)}$为目标函数极小值和原-对偶解；迭代终止。

否则，$k = k + 1$，转到步骤2

9.3.5 实例测试

例9.5 用原-对偶可行路径跟踪法求解

$$\begin{gathered} \max f(x) = x_1 + x_2 + 5x_3 \\ \text{s.t.}\, 3x_1 + 2x_2 + \frac{1}{4}x_3 \le 6 \\ x_3 \le 4 \\ \boldsymbol{x} \ge \boldsymbol{0} \end{gathered}$$

$$\begin{gathered} \min -f(x) = -x_1 - x_2 - 5x_3 \\ \text{s.t.} -3x_1 - 2x_2 - \frac{1}{4}x_3 \ge -6 \\ -x_3 \ge -4 \\ \boldsymbol{x} \ge \boldsymbol{0} \end{gathered}$$

初始点$\boldsymbol{x}^{(0)} = (2.5, 2.5, 3)$, $tol = 1 \times 10^{-4}$

example_9_5_XinggaoLiu.m

x_optimal =
0.0000
2.5000
4.0000

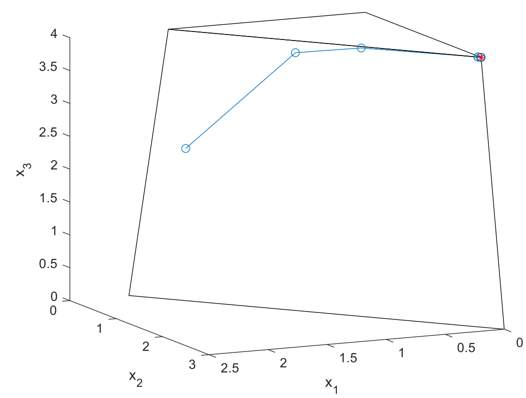
y_optimal =
0.0000
0.0000
0.0000
2.5000
4.0000

f_optimal=22.5000

k = 8

lamda_optimal =
空的 0×1 double 列向量

mu_optimal =
0.5000
4.8750
0.5000
0.0000
0.0000



9.4 带预测校正的原-对偶路径跟踪法

9.4.1 基本原理

Mehrotra方法借鉴常微分法方程数值解法中的预测校正思想:

对当前迭代点$\boldsymbol{z}^{(k)}=\left[\boldsymbol{x}^{(k)},\boldsymbol{y}^{(k)},\boldsymbol{\lambda}^{(k)},\boldsymbol{\mu}^{(k)}\right]^T$，做适当的扰动$\boldsymbol{\delta}_{\boldsymbol{z}}^{(k)}=\boldsymbol{\delta}_{\text{pre}}^{(k)}+\boldsymbol{\delta}_{\text{cor}}^{(k)}$，

得到下一个迭代点 $\boldsymbol{z}^{(k+1)}=\left[\boldsymbol{x}^{(k+1)},\boldsymbol{y}^{(k+1)},\boldsymbol{\lambda}^{(k+1)},\boldsymbol{\mu}^{(k+1)}\right]^T$

$$\boldsymbol{z}^{(k+1)}=\boldsymbol{z}^{(k)}+\boldsymbol{\delta}_{\boldsymbol{z}}^{(k)}$$

$$\boldsymbol{z}^{(k+1)}=\begin{bmatrix}\boldsymbol{x}^{(k+1)}\\ \boldsymbol{y}^{(k+1)}\\ \boldsymbol{\lambda}^{(k+1)}\\ \boldsymbol{\mu}^{(k+1)}\end{bmatrix}=\begin{bmatrix}\boldsymbol{x}^{(k)}\\ \boldsymbol{y}^{(k)}\\ \boldsymbol{\lambda}^{(k)}\\ \boldsymbol{\mu}^{(k)}\end{bmatrix}+\begin{bmatrix}\boldsymbol{\delta}_{x,\text{pre}}^{(k)}+\boldsymbol{\delta}_{x,\text{cor}}^{(k)}\\ \boldsymbol{\delta}_{y,\text{pre}}^{(k)}+\boldsymbol{\delta}_{y,\text{cor}}^{(k)}\\ \boldsymbol{\delta}_{\lambda,\text{pre}}^{(k)}+\boldsymbol{\delta}_{\lambda,\text{cor}}^{(k)}\\ \boldsymbol{\delta}_{\mu,\text{pre}}^{(k)}+\boldsymbol{\delta}_{\mu,\text{cor}}^{(k)}\end{bmatrix}$$

预测方向,仿射方向

$$\boldsymbol{\delta}_{\text{pre}}^{(k)}=\begin{bmatrix}\boldsymbol{\delta}_{x,\text{pre}}^{(k)}\\ \boldsymbol{\delta}_{y,\text{pre}}^{(k)}\\ \boldsymbol{\delta}_{\lambda,\text{pre}}^{(k)}\\ \boldsymbol{\delta}_{\mu,\text{pre}}^{(k)}\end{bmatrix}$$

校正方向

$$\boldsymbol{\delta}_{\text{cor}}^{(k)}=\begin{bmatrix}\boldsymbol{\delta}_{x,\text{cor}}^{(k)}\\ \boldsymbol{\delta}_{y,\text{cor}}^{(k)}\\ \boldsymbol{\delta}_{\lambda,\text{cor}}^{(k)}\\ \boldsymbol{\delta}_{\mu,\text{cor}}^{(k)}\end{bmatrix}$$

校正方向补偿线性化的误差，
使得搜索方向靠近中心路径

(1) 扰动KKT条件的展开与分解

前面的方法中，最后一个条件，略去了扰动的二次项

$$M^{(k+1)}Y^{(k+1)}\boldsymbol{e} \approx M^{(k)}Y^{(k)}\boldsymbol{e} + M^{(k)}\boldsymbol{\delta}_y^{(k)} + Y^{(k)}\boldsymbol{\delta}_\mu^{(k)}$$

保留

事实上 $M^{(k+1)}Y^{(k+1)}\boldsymbol{e} = M^{(k)}Y^{(k)}\boldsymbol{e} + M^{(k)}\boldsymbol{\delta}_y^{(k)} + Y^{(k)}\boldsymbol{\delta}_\mu^{(k)} + \Delta Y^{(k)}\boldsymbol{\delta}_\mu^{(k)}$

$$\begin{aligned}
&\boldsymbol{c} - A_E^T\left(\boldsymbol{\lambda}^{(k)} + \boldsymbol{\delta}_\lambda^{(k)}\right) - A_I^T\left(\boldsymbol{\mu}^{(k)} + \boldsymbol{\delta}_\mu^{(k)}\right) = \boldsymbol{0}\\
&A_E\left(\boldsymbol{x}^{(k)} + \boldsymbol{\delta}_x^{(k)}\right) - \boldsymbol{b}_E = \boldsymbol{0}\\
&A_I\left(\boldsymbol{x}^{(k)} + \boldsymbol{\delta}_x^{(k)}\right) - \left(\boldsymbol{y}^{(k)} + \boldsymbol{\delta}_y^{(k)}\right) = \boldsymbol{b}_I\\
&M^{(k)}Y^{(k)}\boldsymbol{e} + M^{(k)}\boldsymbol{\delta}_y^{(k)} + Y^{(k)}\boldsymbol{\delta}_\mu^{(k)} + \Delta Y^{(k)}\boldsymbol{\delta}_\mu^{(k)} = \tau^{(k+1)}\boldsymbol{e}
\end{aligned}$$

$$\begin{aligned}
A_E^T\boldsymbol{\delta}_\lambda^{(k)} + A_I^T\boldsymbol{\delta}_\mu^{(k)} &= \boldsymbol{c} - A_E^T\boldsymbol{\lambda}^{(k)} - A_I^T\boldsymbol{\mu}^{(k)}\\
A_E\boldsymbol{\delta}_x^{(k)} &= \boldsymbol{b}_E - A_E\boldsymbol{x}^{(k)}\\
A_I\boldsymbol{\delta}_x^{(k)} - \boldsymbol{\delta}_y^{(k)} &= \boldsymbol{b}_I - A_I\boldsymbol{x}^{(k)} + \boldsymbol{y}^{(k)}\\
M^{(k)}\boldsymbol{\delta}_y^{(k)} + Y^{(k)}\boldsymbol{\delta}_\mu^{(k)} &= M^{(k)}Y^{(k)}\boldsymbol{e} - \Delta Y^{(k)}\boldsymbol{\delta}_\mu^{(k)} - \tau^{(k+1)}\boldsymbol{e}
\end{aligned}$$

(1) 扰动KKT条件的展开与分解

$$\begin{aligned}
A_E^T\boldsymbol{\delta}_\lambda^{(k)} + A_I^T\boldsymbol{\delta}_\mu^{(k)} &= \boldsymbol{c} - A_E^T\boldsymbol{\lambda}^{(k)} - A_I^T\boldsymbol{\mu}^{(k)}\\
A_E\boldsymbol{\delta}_x^{(k)} &= \boldsymbol{b}_E - A_E\boldsymbol{x}^{(k)}\\
A_I\boldsymbol{\delta}_x^{(k)} - \boldsymbol{\delta}_y^{(k)} &= \boldsymbol{b}_I - A_I\boldsymbol{x}^{(k)} + \boldsymbol{y}^{(k)}\\
M^{(k)}\boldsymbol{\delta}_y^{(k)} + Y^{(k)}\boldsymbol{\delta}_\mu^{(k)} &= M^{(k)}Y^{(k)}\boldsymbol{e} - \Delta Y^{(k)}\boldsymbol{\delta}_\mu^{(k)} - \tau^{(k+1)}\boldsymbol{e}
\end{aligned}$$

左乘$-\left(Y^{(k)}\right)^{-1}$

$$\begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} & A_E^T & A_I^T\\ \boldsymbol{0} & -\left(Y^{(k)}\right)^{-1}M^{(k)} & \boldsymbol{0} & -I\\ A_E & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\\ A_I & -I & \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}\boldsymbol{\delta}_z^{(k)} = \begin{bmatrix} \boldsymbol{v}_x^{(k)}\\ \boldsymbol{\mu}^{(k)} + \left(Y^{(k)}\right)^{-1}\Delta Y^{(k)}\boldsymbol{\delta}_\mu^{(k)} - \tau^{(k+1)}\left(Y^{(k)}\right)^{-1}\boldsymbol{e}\\ -\boldsymbol{v}_\lambda^{(k)}\\ \boldsymbol{v}_\mu^{(k)}\end{bmatrix}$$

$$\begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} & A_E^T & A_I^T\\ \boldsymbol{0} & -\left(Y^{(k)}\right)^{-1}M^{(k)} & \boldsymbol{0} & -I\\ A_E & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\\ A_I & -I & \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}\left(\boldsymbol{\delta}_{\text{pre}}^{(k)} + \boldsymbol{\delta}_{\text{cor}}^{(k)}\right) = \begin{bmatrix} \boldsymbol{v}_x^{(k)}\\ \boldsymbol{\mu}^{(k)}\\ -\boldsymbol{v}_\lambda^{(k)}\\ \boldsymbol{v}_\mu^{(k)}\end{bmatrix} + \begin{bmatrix} 0\\ \left(Y^{(k)}\right)^{-1}\Delta Y^{(k)}\boldsymbol{\delta}_\mu^{(k)} - \tau^{(k+1)}\left(Y^{(k)}\right)^{-1}\boldsymbol{e}\\ 0\\ 0\end{bmatrix}$$

$$\begin{bmatrix} \mathbf{0} & \mathbf{0} & A_E^T & A_I^T \\ \mathbf{0} & -\left(Y^{(k)}\right)^{-1}M^{(k)} & \mathbf{0} & -I \\ A_E & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ A_I & -I & \mathbf{0} & \mathbf{0} \end{bmatrix}\left(\boldsymbol{\delta}_{\text{pre}}^{(k)} + \boldsymbol{\delta}_{\text{cor}}^{(k)}\right) = \begin{bmatrix} \boldsymbol{v}_x^{(k)} \\ \boldsymbol{\mu}^{(k)} \\ -\boldsymbol{v}_\lambda^{(k)} \\ \boldsymbol{v}_\mu^{(k)} \end{bmatrix} + \begin{bmatrix} 0 \\ \left(Y^{(k)}\right)^{-1}\Delta Y^{(k)}\boldsymbol{\delta}_\mu^{(k)} - \tau^{(k+1)}\left(Y^{(k)}\right)^{-1}\boldsymbol{e} \\ 0 \\ 0 \end{bmatrix}$$

分成两个方程组

$$\begin{bmatrix} \mathbf{0} & \mathbf{0} & A_E^T & A_I^T \\ \mathbf{0} & -\left(Y^{(k)}\right)^{-1}M^{(k)} & \mathbf{0} & -I \\ A_E & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ A_I & -I & \mathbf{0} & \mathbf{0} \end{bmatrix}\begin{bmatrix} \boldsymbol{\delta}_{x,\text{pre}}^{(k)} \\ \boldsymbol{\delta}_{y,\text{pre}}^{(k)} \\ \boldsymbol{\delta}_{\lambda,\text{pre}}^{(k)} \\ \boldsymbol{\delta}_{\mu,\text{pre}}^{(k)} \end{bmatrix} = \begin{bmatrix} \boldsymbol{v}_x^{(k)} \\ \boldsymbol{\mu}^{(k)} \\ -\boldsymbol{v}_\lambda^{(k)} \\ \boldsymbol{v}_\mu^{(k)} \end{bmatrix} \implies \boldsymbol{\delta}_{\text{pre}}^{(k)}$$

?

$$\begin{bmatrix} \mathbf{0} & \mathbf{0} & A_E^T & A_I^T \\ \mathbf{0} & -\left(Y^{(k)}\right)^{-1}M^{(k)} & \mathbf{0} & -I \\ A_E & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ A_I & -I & \mathbf{0} & \mathbf{0} \end{bmatrix}\begin{bmatrix} \boldsymbol{\delta}_{x,\text{cor}}^{(k)} \\ \boldsymbol{\delta}_{y,\text{cor}}^{(k)} \\ \boldsymbol{\delta}_{\lambda,\text{cor}}^{(k)} \\ \boldsymbol{\delta}_{\mu,\text{cor}}^{(k)} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \left(Y^{(k)}\right)^{-1}\Delta Y^{(k)}\boldsymbol{\delta}_\mu^{(k)} - \tau^{(k+1)}\left(Y^{(k)}\right)^{-1}\boldsymbol{e} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ -\boldsymbol{v}_y^{(k)} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \implies \boldsymbol{\delta}_{\text{cor}}^{(k)}$$

取 $\Delta Y^{(k)} \approx \Delta Y_{\text{pre}}^{(k)}$, $\boldsymbol{\delta}_\mu^{(k)} \approx \boldsymbol{\delta}_{\mu,\text{pre}}^{(k)}$

$$-\boldsymbol{v}_y^{(k)} = \left(Y^{(k)}\right)^{-1}\Delta Y_{\text{pre}}^{(k)}\boldsymbol{\delta}_{\mu,\text{pre}}^{(k)} - \tau^{(k+1)}\left(Y^{(k)}\right)^{-1}\boldsymbol{e}$$

预测方向、校正方向的方程组

$$\begin{bmatrix} \mathbf{0} & \mathbf{0} & A_E^T & A_I^T \\ \mathbf{0} & -\left(Y^{(k)}\right)^{-1}M^{(k)} & \mathbf{0} & -I \\ A_E & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ A_I & -I & \mathbf{0} & \mathbf{0} \end{bmatrix}\begin{bmatrix} \boldsymbol{\delta}_{x,\text{pre}}^{(k)} \\ \boldsymbol{\delta}_{y,\text{pre}}^{(k)} \\ \boldsymbol{\delta}_{\lambda,\text{pre}}^{(k)} \\ \boldsymbol{\delta}_{\mu,\text{pre}}^{(k)} \end{bmatrix} = \begin{bmatrix} \boldsymbol{v}_x^{(k)} \\ \boldsymbol{\mu}^{(k)} \\ -\boldsymbol{v}_\lambda^{(k)} \\ \boldsymbol{v}_\mu^{(k)} \end{bmatrix} \quad (9.4.4) \implies \boldsymbol{\delta}_{\text{pre}}^{(k)}$$

预测方向

$$\begin{bmatrix} \mathbf{0} & \mathbf{0} & A_E^T & A_I^T \\ \mathbf{0} & -\left(Y^{(k)}\right)^{-1}M^{(k)} & \mathbf{0} & -I \\ A_E & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ A_I & -I & \mathbf{0} & \mathbf{0} \end{bmatrix}\begin{bmatrix} \boldsymbol{\delta}_{x,\text{cor}}^{(k)} \\ \boldsymbol{\delta}_{y,\text{cor}}^{(k)} \\ \boldsymbol{\delta}_{\lambda,\text{cor}}^{(k)} \\ \boldsymbol{\delta}_{\mu,\text{cor}}^{(k)} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ -\boldsymbol{v}_y^{(k)} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \quad (9.4.7) \implies \boldsymbol{\delta}_{\text{cor}}^{(k)}$$

校正方向

$$-\boldsymbol{v}_y^{(k)} = \left(Y^{(k)}\right)^{-1}\Delta Y_{\text{pre}}^{(k)}\boldsymbol{\delta}_{\mu,\text{pre}}^{(k)} - \tau^{(k+1)}\left(Y^{(k)}\right)^{-1}\boldsymbol{e}$$

(2) 简化方程组，求解扰动向量

与9.2.2类似，可得预测方向

$$\begin{bmatrix} A_I^T\left(Y^{(k)}\right)^{-1}M^{(k)}A_I & -A_E^T \\ -A_E & \mathbf{0} \end{bmatrix}\begin{bmatrix} \boldsymbol{\delta}_{x,\mathrm{pre}}^{(k)} \\ \boldsymbol{\delta}_{\lambda,\mathrm{pre}}^{(k)} \end{bmatrix} = \begin{bmatrix} A_I^T\left[-\boldsymbol{\mu}^{(k)} + \left(Y^{(k)}\right)^{-1}M^{(k)}\boldsymbol{v}_\mu^{(k)}\right] - \boldsymbol{v}_x^{(k)} \\ \boldsymbol{v}_\lambda^{(k)} \end{bmatrix}$$

$$\begin{bmatrix} H^{(k)} & -A_E^T \\ -A_E & \mathbf{0} \end{bmatrix}\begin{bmatrix} \boldsymbol{\delta}_{x,\mathrm{pre}}^{(k)} \\ \boldsymbol{\delta}_{\lambda,\mathrm{pre}}^{(k)} \end{bmatrix} = \begin{bmatrix} -\boldsymbol{p}_{\mathrm{pre}}^{(k)} \\ \boldsymbol{v}_\lambda^{(k)} \end{bmatrix}$$

$$H^{(k)} = A_I^T\left(Y^{(k)}\right)^{-1}M^{(k)}A_I$$

$$\boldsymbol{p}_{\mathrm{pre}}^{(k)} = \boldsymbol{v}_x^{(k)} - A_I^T\left[-\boldsymbol{\mu}^{(k)} + \left(Y^{(k)}\right)^{-1}M^{(k)}\boldsymbol{v}_\mu^{(k)}\right]$$

求解$\boldsymbol{\delta}_{x,\mathrm{pre}}^{(k)}$ 和$\boldsymbol{\delta}_{\lambda,\mathrm{pre}}^{(k)}$

$$\min \frac{1}{2}\left[\boldsymbol{\delta}_{x,\mathrm{pre}}^{(k)}\right]^T H^{(k)}\boldsymbol{\delta}_{x,\mathrm{pre}}^{(k)} + \left[\boldsymbol{\delta}_{x,\mathrm{pre}}^{(k)}\right]^T \boldsymbol{p}_{\mathrm{pre}}^{(k)}$$
$$\mathrm{s.t.}\ A_E\boldsymbol{\delta}_{x,\mathrm{pre}}^{(k)} = -\boldsymbol{v}_\lambda^{(k)}$$

调用MATLAB函数quadprog

可以同时得到，$\boldsymbol{\delta}_{x,\mathrm{pre}}^{(k)}$和$\boldsymbol{\delta}_{\lambda,\mathrm{pre}}^{(k)}$

计算稳定性好，效率高

求解$\boldsymbol{\delta}_{y,\mathrm{pre}}^{(k)}$和$\boldsymbol{\delta}_{\mu,\mathrm{pre}}^{(k)}$

$$\boldsymbol{\delta}_{y,\mathrm{pre}}^{(k)} = A_I\boldsymbol{\delta}_{x,\mathrm{pre}}^{(k)} - \boldsymbol{v}_\mu^{(k)}$$
$$\boldsymbol{\delta}_{\mu,\mathrm{pre}}^{(k)} = -\boldsymbol{\mu}^{(k)} - \left(Y^{(k)}\right)^{-1}M^{(k)}\boldsymbol{\delta}_{y,\mathrm{pre}}^{(k)}$$

另一种解方程组的方法

不常用

当$A_E \neq \emptyset$时

$$\begin{bmatrix} H^{(k)} & -A_E^T \\ -A_E & \mathbf{0} \end{bmatrix}\begin{bmatrix} \boldsymbol{\delta}_{x,\mathrm{pre}}^{(k)} \\ \boldsymbol{\delta}_{\lambda,\mathrm{pre}}^{(k)} \end{bmatrix} = \begin{bmatrix} -\boldsymbol{p}_{\mathrm{pre}}^{(k)} \\ \boldsymbol{v}_\lambda^{(k)} \end{bmatrix}$$

$$\begin{matrix} H^{(k)} & -A_E^T & -\boldsymbol{p}_{\mathrm{pre}}^{(k)} \\ -A_E & \mathbf{0} & \boldsymbol{v}_\lambda^{(k)} \end{matrix}$$

$$\begin{matrix} I & -\left(H^{(k)}\right)^{-1}A_E^T & -\left(H^{(k)}\right)^{-1}\boldsymbol{p}_{\mathrm{pre}}^{(k)} \\ -A_E & \mathbf{0} & \boldsymbol{v}_\lambda^{(k)} \end{matrix}$$

$$\boldsymbol{\delta}_{x,\mathrm{pre}}^{(k)} = \left(H^{(k)}\right)^{-1}\left[A_E^T\boldsymbol{\delta}_{\lambda,\mathrm{pre}}^{(k)} - \boldsymbol{p}_{\mathrm{pre}}^{(k)}\right]$$

$$\begin{matrix} A_E & -A_E\left(H^{(k)}\right)^{-1}A_E^T & -A_E\left(H^{(k)}\right)^{-1}\boldsymbol{p}_{\mathrm{pre}}^{(k)} \\ -A_E & \mathbf{0} & \boldsymbol{v}_\lambda^{(k)} \end{matrix}$$

$$\begin{matrix} \mathbf{0} & -A_E\left(H^{(k)}\right)^{-1}A_E^T & \boldsymbol{v}_\lambda^{(k)} - A_E\left(H^{(k)}\right)^{-1}\boldsymbol{p}_{\mathrm{pre}}^{(k)} \\ -A_E & \mathbf{0} & \boldsymbol{v}_\lambda^{(k)} \end{matrix}$$

$$A_E\left(H^{(k)}\right)^{-1}A_E^T\boldsymbol{\delta}_{\lambda,\mathrm{pre}}^{(k)} = A_E\left(H^{(k)}\right)^{-1}\boldsymbol{p}_{\mathrm{pre}}^{(k)} - \boldsymbol{v}_\lambda^{(k)}$$

解线性方程组，得到$\boldsymbol{\delta}_{\lambda,\mathrm{pre}}^{(k)}$

另一种解方程组的方法

不常用

当$A_E=\emptyset$时

$$\begin{bmatrix} H^{(k)} & -A_E^T \\ -A_E & \mathbf{0} \end{bmatrix}\begin{bmatrix} \boldsymbol{\delta}_{x,\mathrm{pre}}^{(k)} \\ \boldsymbol{\delta}_{\lambda,\mathrm{pre}}^{(k)} \end{bmatrix}=\begin{bmatrix} -\boldsymbol{p}_{\mathrm{pre}}^{(k)} \\ \boldsymbol{v}_{\lambda}^{(k)} \end{bmatrix}$$

$$H^{(k)}\boldsymbol{\delta}_{x,\mathrm{pre}}^{(k)}=-\boldsymbol{p}_{\mathrm{pre}}^{(k)}$$

解线性方程组，得到$\boldsymbol{\delta}_{x,\mathrm{pre}}^{(k)}$

$$\boldsymbol{\delta}_{\lambda,\mathrm{pre}}^{(k)}=\emptyset$$

校正方向

$$\begin{bmatrix} A_I^T\left(Y^{(k)}\right)^{-1}M^{(k)}A_I & -A_E^T \\ -A_E & \mathbf{0} \end{bmatrix}\begin{bmatrix} \boldsymbol{\delta}_{x,\mathrm{cor}}^{(k)} \\ \boldsymbol{\delta}_{\lambda,\mathrm{cor}}^{(k)} \end{bmatrix}=\begin{bmatrix} A_I^T\boldsymbol{v}_y^{(k)} \\ \mathbf{0} \end{bmatrix}$$

$$H^{(k)}=A_I^T\left(Y^{(k)}\right)^{-1}M^{(k)}A_I$$

$$\boldsymbol{p}_{\mathrm{cor}}^{(k)}=-A_I^T\boldsymbol{v}_y^{(k)}$$

$$\begin{bmatrix} H^{(k)} & -A_E^T \\ -A_E & \mathbf{0} \end{bmatrix}\begin{bmatrix} \boldsymbol{\delta}_{x,\mathrm{cor}}^{(k)} \\ \boldsymbol{\delta}_{\lambda,\mathrm{cor}}^{(k)} \end{bmatrix}=\begin{bmatrix} A_I^T\boldsymbol{v}_y^{(k)} \\ \mathbf{0} \end{bmatrix}$$

求解$\boldsymbol{\delta}_{x,\mathrm{cor}}^{(k)}$和$\boldsymbol{\delta}_{\lambda,\mathrm{cor}}^{(k)}$

$$\min \frac{1}{2}\left[\boldsymbol{\delta}_{x,\mathrm{cor}}^{(k)}\right]^T H^{(k)}\boldsymbol{\delta}_{x,\mathrm{cor}}^{(k)}+\left[\boldsymbol{\delta}_{x,\mathrm{cor}}^{(k)}\right]^T\boldsymbol{p}_{\mathrm{cor}}^{(k)}$$

$$\mathrm{s.t.}\ A_E\boldsymbol{\delta}_{x,\mathrm{cor}}^{(k)}=-\boldsymbol{v}_{\lambda}^{(k)}$$

调用MATLAB函数quadprog

可以同时得到，$\boldsymbol{\delta}_{x,\mathrm{cor}}^{(k)}$和$\boldsymbol{\delta}_{\lambda,\mathrm{cor}}^{(k)}$
计算稳定性好，效率高

求解$\boldsymbol{\delta}_{y,\mathrm{cor}}^{(k)}$和$\boldsymbol{\delta}_{\mu,\mathrm{cor}}^{(k)}$

$$\boldsymbol{\delta}_{y,\mathrm{cor}}^{(k)}=A_I\boldsymbol{\delta}_{x,\mathrm{cor}}^{(k)}$$

$$\boldsymbol{\delta}_{\mu,\mathrm{cor}}^{(k)}=\boldsymbol{v}_y^{(k)}-\left(Y^{(k)}\right)^{-1}M^{(k)}\boldsymbol{\delta}_{y,\mathrm{cor}}^{(k)}$$

3.步长计算

3.1 沿扰动方向$\boldsymbol{\delta}_{\boldsymbol{z}}^{(k)}$的步长计算

$$\boldsymbol{\delta}_{\boldsymbol{z}}^{(k)} = \boldsymbol{\delta}_{\text{pre}}^{(k)} + \boldsymbol{\delta}_{\text{cor}}^{(k)}$$

$$\begin{cases} \alpha_{P,min}^{(k)} = \min\left\{-\dfrac{\left(\boldsymbol{y}^{(k)}\right)_i}{\left(\boldsymbol{\delta}_{y}^{(k)}\right)_i} \middle| \left(\boldsymbol{\delta}_{y}^{(k)}\right)_i < 0, i = 1,2,\dots,m_2\right\} \\ \alpha_{P}^{(k)} = \min\left\{1, c\cdot\alpha_{P,min}^{(k)}\right\} \end{cases}$$

$$\begin{cases} \alpha_{D,min}^{(k)} = \min\left\{-\dfrac{\left(\boldsymbol{\mu}^{(k)}\right)_i}{\left(\boldsymbol{\delta}_{\boldsymbol{\mu}}^{(k)}\right)_i} \middle| \left(\boldsymbol{\delta}_{\boldsymbol{\mu}}^{(k)}\right)_i < 0, i = 1,2,\dots,m_2\right\} \\ \alpha_{D}^{(k)} = \min\left\{1, c\cdot\alpha_{D,min}^{(k)}\right\} \end{cases}$$

通常

$$c = 1 - 10^{-3}$$
$$1 - 10^{-3} \le c \le 1 - 10^{-6}$$

3.步长计算　与（9.2.2）一样

3.2 沿预测方向$\boldsymbol{\delta}_{\text{pre}}^{(k)}$的步长　用于计算中心参数

$$\begin{cases} \alpha_{P,\text{pre},min}^{(k)} = \min\left\{-\dfrac{\left(\boldsymbol{y}^{(k)}\right)_i}{\left(\boldsymbol{\delta}_{y,\text{pre}}^{(k)}\right)_i} \middle| \left(\boldsymbol{\delta}_{y,\text{pre}}^{(k)}\right)_i < 0, i = 1,2,\dots,m_2\right\} \\ \alpha_{P,\text{pre}}^{(k)} = \min\left\{1, c\cdot\alpha_{P,\text{pre},min}^{(k)}\right\} \end{cases}$$

$$\begin{cases} \alpha_{D,\text{pre},min}^{(k)} = \min\left\{-\dfrac{\left(\boldsymbol{\mu}^{(k)}\right)_i}{\left(\boldsymbol{\delta}_{\boldsymbol{\mu},\text{pre}}^{(k)}\right)_i} \middle| \left(\boldsymbol{\delta}_{\boldsymbol{\mu},\text{pre}}^{(k)}\right)_i < 0, i = 1,2,\dots,m_2\right\} \\ \alpha_{D,\text{pre}}^{(k)} = \min\left\{1, c\cdot\alpha_{D,\text{pre},min}^{(k)}\right\} \end{cases}$$

通常

$$c = 1 - 10^{-3}$$
$$1 - 10^{-3} \le c \le 1 - 10^{-6}$$

4. 中心参数计算

参数τ估计（Mehrotra方法）　　启发式公式，无严格理论

$$\sigma^{(k)} = \left(\frac{\tau_{\text{pre}}^{(k)}}{\tau^{(k)}}\right)^3$$

$$\tau_{\text{pre}}^{(k)} = \frac{1}{m_2}\left[\left(\boldsymbol{\mu}^{(k)} + \alpha_{D,\text{pre}}^{(k)}\boldsymbol{\delta}_{\mu,\text{pre}}^{(k)}\right)^T\left(\boldsymbol{y}^{(k)} + \alpha_{P,\text{pre}}^{(k)}\boldsymbol{\delta}_{y,\text{pre}}^{(k)}\right)\right]$$

$$\tau^{(k)} = \frac{\left(\boldsymbol{\mu}^{(k)}\right)^T\boldsymbol{y}^{(k)}}{m_2}$$

$$\tau^{(k+1)} = \sigma^{(k)}\tau^{(k)}$$

9.4.2 带预测校正的原-对偶路径跟踪法的计算步骤

步骤1：输入参数$\boldsymbol{c}, A_E, \boldsymbol{b}_E, A_I, \boldsymbol{b}_I$，选定初始点$\boldsymbol{z}^{(0)} = \left(\boldsymbol{x}^{(0)}, \boldsymbol{y}^{(0)}, \boldsymbol{\lambda}^{(0)}, \boldsymbol{\mu}^{(0)}\right)$
设定精度tol，令$k = 0$

步骤2：计算预测（仿射）方向$\boldsymbol{\delta}_{\text{pre}}^{(k)}$

①计算参数

$$\begin{bmatrix} \boldsymbol{v}_x^{(k)} \\ \boldsymbol{\mu}^{(k)} \\ -\boldsymbol{v}_\lambda^{(k)} \\ \boldsymbol{v}_\mu^{(k)} \end{bmatrix} = \begin{bmatrix} \boldsymbol{c} - A_E^T\boldsymbol{\lambda}^{(k)} - A_I^T\boldsymbol{\mu}^{(k)} \\ \boldsymbol{\mu}^{(k)} \\ \boldsymbol{b}_E - A_E\boldsymbol{x}^{(k)} \\ \boldsymbol{b}_I - A_I\boldsymbol{x}^{(k)} + \boldsymbol{y}^{(k)} \end{bmatrix}$$

$$H^{(k)} = A_I^T\left(Y^{(k)}\right)^{-1}M^{(k)}A_I$$

$$\boldsymbol{p}_{\text{pre}}^{(k)} = \boldsymbol{v}_x^{(k)} - A_I^T\left[-\boldsymbol{\mu}^{(k)} + \left(Y^{(k)}\right)^{-1}M^{(k)}\boldsymbol{v}_\mu^{(k)}\right]$$

步骤2：计算预测（仿射）方向 $\boldsymbol{\delta}_{\text{pre}}^{(k)} = \left[\boldsymbol{\delta}_{x,\text{pre}}^{(k)} \quad \boldsymbol{\delta}_{y,\text{pre}}^{(k)} \quad \boldsymbol{\delta}_{\lambda,\text{pre}}^{(k)} \quad \boldsymbol{\delta}_{\mu,\text{pre}}^{(k)}\right]^T$

②求解$\boldsymbol{\delta}_{x,\text{pre}}^{(k)}$和$\boldsymbol{\delta}_{\lambda,\text{pre}}^{(k)}$

$$\min \frac{1}{2}\left[\boldsymbol{\delta}_{x,\text{pre}}^{(k)}\right]^T H^{(k)} \boldsymbol{\delta}_{x,\text{pre}}^{(k)} + \left[\boldsymbol{\delta}_{x,\text{pre}}^{(k)}\right]^T \boldsymbol{p}_{\text{pre}}^{(k)}$$
$$\text{s.t.}\ A_E \boldsymbol{\delta}_{x,\text{pre}}^{(k)} = -\boldsymbol{v}_{\lambda}^{(k)}$$

调用MATLAB函数quadprog

可以同时得到，$\boldsymbol{\delta}_{x,\text{pre}}^{(k)}$和$\boldsymbol{\delta}_{\lambda,\text{pre}}^{(k)}$
计算稳定性好，效率高

③求解$\boldsymbol{\delta}_{y,\text{pre}}^{(k)}$和$\boldsymbol{\delta}_{\mu,\text{pre}}^{(k)}$

$$\boldsymbol{\delta}_{y,\text{pre}}^{(k)} = A_I \boldsymbol{\delta}_{x,\text{pre}}^{(k)} - \boldsymbol{v}_{\mu}^{(k)}$$
$$\boldsymbol{\delta}_{\mu,\text{pre}}^{(k)} = -\boldsymbol{\mu}^{(k)} - \left(Y^{(k)}\right)^{-1} M^{(k)} \boldsymbol{\delta}_{y,\text{pre}}^{(k)}$$

步骤3：计算预测（仿射）方向的步长$\alpha_{P,\text{pre}}^{(k)}$和$\alpha_{D,\text{pre}}^{(k)}$

$$\begin{cases} \alpha_{P,\text{pre},min}^{(k)} = \min\left\{-\dfrac{\left(\boldsymbol{y}^{(k)}\right)_i}{\left(\boldsymbol{\delta}_{y,\text{pre}}^{(k)}\right)_i} \middle| \left(\boldsymbol{\delta}_{y,\text{pre}}^{(k)}\right)_i < 0, i = 1,2,\dots,m_2\right\} \\ \alpha_{P,\text{pre}}^{(k)} = \min\left\{1, c\cdot\alpha_{P,\text{pre},min}^{(k)}\right\} \end{cases}$$

通常
$$c = 1 - 10^{-3}$$
$$1 - 10^{-3} \le c \le 1 - 10^{-6}$$

$$\begin{cases} \alpha_{D,\text{pre},min}^{(k)} = \min\left\{-\dfrac{\left(\boldsymbol{\mu}^{(k)}\right)_i}{\left(\boldsymbol{\delta}_{\mu,\text{pre}}^{(k)}\right)_i} \middle| \left(\boldsymbol{\delta}_{\mu,\text{pre}}^{(k)}\right)_i < 0, i = 1,2,\dots,m_2\right\} \\ \alpha_{D,\text{pre}}^{(k)} = \min\left\{1, c\cdot\alpha_{D,\text{pre},min}^{(k)}\right\} \end{cases}$$

步骤4：计算中心参数$\sigma^{(k)}$与缩减因子$\tau^{(k+1)}$

$$\sigma^{(k)} = \left(\frac{\tau_{\text{pre}}^{(k)}}{\tau^{(k)}}\right)^3$$
$$\tau_{\text{pre}}^{(k)} = \frac{1}{m_2}\left[\left(\boldsymbol{\mu}^{(k)} + \alpha_{D,\text{pre}}^{(k)}\boldsymbol{\delta}_{\mu,\text{pre}}^{(k)}\right)^T\left(\boldsymbol{y}^{(k)} + \alpha_{P,\text{pre}}^{(k)}\boldsymbol{\delta}_{y,\text{pre}}^{(k)}\right)\right]$$
$$\tau^{(k+1)} = \sigma^{(k)}\tau^{(k)}$$
$$\tau^{(k)} = \frac{\left(\boldsymbol{\mu}^{(k)}\right)^T \boldsymbol{y}^{(k)}}{m_2}$$

步骤5： 计算校正方向$\boldsymbol{\delta}_{\text{cor}}^{(k)}$，以及搜索方向$\boldsymbol{\delta}_{\boldsymbol{z}}^{(k)}$

①计算参数 $\boldsymbol{v}_y^{(k)} = \tau^{(k+1)}\left(Y^{(k)}\right)^{-1}\boldsymbol{e} - \left(Y^{(k)}\right)^{-1}\Delta Y_{\text{pre}}^{(k)}\boldsymbol{\delta}_{\mu,\text{pre}}^{(k)}$

$H^{(k)} = A_I^T\left(Y^{(k)}\right)^{-1}M^{(k)}A_I$

$\boldsymbol{p}_{\text{cor}}^{(k)} = -A_I^T\boldsymbol{v}_y^{(k)}$

②求解$\boldsymbol{\delta}_{x,\text{cor}}^{(k)}$和$\boldsymbol{\delta}_{\lambda,\text{cor}}^{(k)}$

$$\min\frac{1}{2}\left[\boldsymbol{\delta}_{x,\text{cor}}^{(k)}\right]^T H^{(k)}\boldsymbol{\delta}_{x,\text{cor}}^{(k)} + \left[\boldsymbol{\delta}_{x,\text{cor}}^{(k)}\right]^T\boldsymbol{p}_{\text{cor}}^{(k)}$$
$$\text{s.t.}\ A_E\boldsymbol{\delta}_{x,\text{cor}}^{(k)} = \boldsymbol{0}$$

调用MATLAB函数quadprog

可以同时得到，$\boldsymbol{\delta}_{x,\text{cor}}^{(k)}$和$\boldsymbol{\delta}_{\lambda,\text{cor}}^{(k)}$
计算稳定性好，效率高

注意：if $A_E = [\ \ ]$；$[]\boldsymbol{\delta}_{x,\text{cor}}^{(k)} = []$

③求解$\boldsymbol{\delta}_{y,\text{cor}}^{(k)}$和$\boldsymbol{\delta}_{\mu,\text{cor}}^{(k)}$

$$\boldsymbol{\delta}_{y,\text{cor}}^{(k)} = A_I\boldsymbol{\delta}_{x,\text{cor}}^{(k)}$$
$$\boldsymbol{\delta}_{\mu,\text{cor}}^{(k)} = \boldsymbol{v}_y^{(k)} - \left(Y^{(k)}\right)^{-1}M^{(k)}\boldsymbol{\delta}_{y,\text{cor}}^{(k)}$$

④计算搜索方向 $\boldsymbol{\delta}_{\boldsymbol{z}}^{(k)} = \boldsymbol{\delta}_{\text{pre}}^{(k)} + \boldsymbol{\delta}_{\text{cor}}^{(k)}$

步骤6： 计算搜索方向$\boldsymbol{\delta}_{\boldsymbol{z}}^{(k)}$的步长$\alpha_P^{(k)}$和$\alpha_D^{(k)}$

$$\begin{cases}\alpha_{P,min}^{(k)} = \min\left\{-\dfrac{\left(\boldsymbol{y}^{(k)}\right)_i}{\left(\boldsymbol{\delta}_y^{(k)}\right)_i}\middle|\left(\boldsymbol{\delta}_y^{(k)}\right)_i < 0, i = 1,2,\ldots,m_2\right\} \\ \alpha_P^{(k)} = \min\left\{1, c\cdot\alpha_{P,min}^{(k)}\right\}\end{cases}$$

$$\begin{cases}\alpha_{D,min}^{(k)} = \min\left\{-\dfrac{\left(\boldsymbol{\mu}^{(k)}\right)_i}{\left(\boldsymbol{\delta}_{\boldsymbol{\mu}}^{(k)}\right)_i}\middle|\left(\boldsymbol{\delta}_{\boldsymbol{\mu}}^{(k)}\right)_i < 0, i = 1,2,\ldots,m_2\right\} \\ \alpha_D^{(k)} = \min\left\{1, c\cdot\alpha_{D,min}^{(k)}\right\}\end{cases}$$

通常
$c = 1 - 10^{-3}$

步骤7：计算新的迭代点

$$\begin{cases} \boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha_P^{(k)} \boldsymbol{\delta}_x^{(k)} \\ \boldsymbol{y}^{(k+1)} = \boldsymbol{y}^{(k)} + \alpha_P^{(k)} \boldsymbol{\delta}_y^{(k)} \\ \boldsymbol{\lambda}^{(k+1)} = \boldsymbol{\lambda}^{(k)} + \alpha_D^{(k)} \boldsymbol{\delta}_\lambda^{(k)} \\ \boldsymbol{\mu}^{(k+1)} = \boldsymbol{\mu}^{(k)} + \alpha_D^{(k)} \boldsymbol{\delta}_\mu^{(k)} \end{cases}$$

步骤8：计算新的对偶间隔 $\delta_{PD}^{(k+1)} = \left[\boldsymbol{\mu}^{(k+1)}\right]^T \boldsymbol{y}^{(k+1)}$

步骤9：如果$\delta_{PD}^{(k+1)} < tol$，$f\left(\boldsymbol{x}^{(k+1)}\right) = \boldsymbol{c}^T \boldsymbol{x}^{(k+1)}$， $\boldsymbol{z}^{(k+1)}$为目标函数极小值和原-对偶解；迭代终止。
否则，$k = k + 1$，转到步骤2

9.4.4 实例测试

例9.8 用原-对偶可行路径跟踪法求解

$$\begin{gathered} \max f(x) = x_1 + x_2 + 5x_3 \\ \text{s.t.}\, 3x_1 + 2x_2 + \frac{1}{4}x_3 \le 6 \\ x_3 \le 4 \\ \boldsymbol{x} \ge \boldsymbol{0} \end{gathered}$$

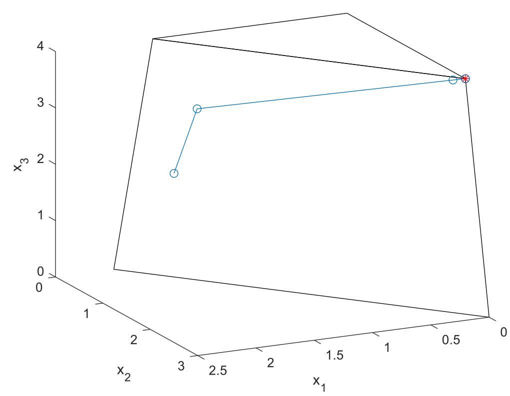
$$\begin{gathered} \min -f(x) = -x_1 - x_2 - 5x_3 \\ \text{s.t.} -3x_1 - 2x_2 - \frac{1}{4}x_3 \ge -6 \\ -x_3 \ge -4 \\ \boldsymbol{x} \ge \boldsymbol{0} \end{gathered}$$

初始点$\boldsymbol{x}^{(0)} = (2.5,2.5,3)$, $tol = 1 \times 10^{-4}$

example_9_8_XinggaoLiu.m

```
x_optimal =
    0.0000
    2.5000
    4.0000

y_optimal =
    0.0000
    0.0000
    0.0000
    2.5000
    4.0000

f_optimal = 22.5000

k =  4
```

```
lamda_optimal =
  空的 0×1 double 列向量

mu_optimal =
    0.5000
    4.8750
    0.5000
    0.0000
    0.0000
```



作业

9-1

9-3

9-4