数值最优化算法及应用

张鑫航

国防科技大学

版本：1.0

更新：2024 年 7 月 28 日

目录

1 无约束优化

1.1 最优性条件的应用

定理 1.1 设 $\boldsymbol{x}^*$ 为优化问题 $\min_x \in \mathbb{R}^n f(\boldsymbol{x})$ 的最优解，则 $\nabla f(\boldsymbol{x}^*) = \mathbf{0}, \nabla^2 f(\boldsymbol{x}^*)$ 半正定。

定理 1.2 对优化问题 $\min_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x})$ ，设 $\nabla f(\boldsymbol{x}^*) = 0, \nabla^2 f(\boldsymbol{x}^*)$ 正定，则 $\boldsymbol{x}^*$ 是该优化问题的严格最优解，

例 1.1 优化问题的解析解⭐⭐⭐⭐⭐

$$\min_{x>0,y\geqslant 0} f(x,y) = \frac{10}{x} + \frac{(x-y)^2}{2x} + \frac{3y^2}{2x}$$

解 先忽略约束：利用 $\min_{x,y} f(x,y) = \min_x \min_y f(x,y)$，先固定 x，关于 y 做内层优化，再求解关于 x 的外层优化

$$\min_y f(x,y) = \frac{10}{x} + \frac{(x-y)^2}{2x} + \frac{3y^2}{2x}$$

目标函数关于 y 为凸函数，利用最优性条件得最优解

$$f(x,y) = \frac{1}{2x}\left(20 + x^2 - 2xy + 4y^2\right)$$
$$y = \frac{1}{4}x$$

将上述最优解代入目标函数得外层优化问题

$$\min f(x) = \frac{10}{x} + \frac{3}{8}x$$

目标函数关于 x 为凸函数（二阶导数大于 0）。再利用最优性条件得 $x = \frac{4}{3}\sqrt{15}, y = \frac{1}{4}x = \frac{1}{3}\sqrt{15}$

它们满足约束条件，自然为原问题的最优解。

1.2 最速下降算法

例 1.2 利用最速下降方法求⭐⭐⭐⭐⭐

$$\min_{\boldsymbol{x}\in\mathbb{R}^2} f(x_1,x_2) = \frac{1}{3}x_1^2 + \frac{1}{2}x_2^2$$

解 显然，唯一最优解 $\boldsymbol{x}^* = (0,0)$，取初始点 $\boldsymbol{x}_0 = (3,2)$，那么最速下降方法产生的迭代点列 $\boldsymbol{x}_k$

- $k=0$ 时

$$\boldsymbol{g}_0=\nabla f(\boldsymbol{x}^{(0)})=\begin{pmatrix}2\\2\end{pmatrix}$$

$$\begin{aligned}\alpha_0&=\arg\min_{\alpha>0} f(\boldsymbol{x}^{(0)}-\alpha \boldsymbol{g}^{(0)})\\&=\arg\min_{\alpha>0}\frac{1}{3}\left(10\alpha^2-24\alpha+\cdots\right)=\frac{6}{5}\\\boldsymbol{x}^{(1)}=\boldsymbol{x}^{(0)}-\alpha\boldsymbol{g}^{(0)}&=(3,2)-\frac{6}{5}(2,2)=(\frac{3}{5},-\frac{2}{5})\end{aligned}$$

- $k=1,\cdots$

综上，得到迭代序列 $\boldsymbol{x}^{(k)}$

$$\boldsymbol{x}_k=\left(\frac{3}{5^k};(-1)^k\frac{2}{5^k}\right)\xrightarrow{k\to\infty}(0;0)$$

全局收敛，收敛速度线性！有 $\|\boldsymbol{x}_k-\boldsymbol{x}^*\|=\sqrt{13}\left(\frac{1}{5}\right)^k$. 那么

$$\begin{aligned}\frac{\|\boldsymbol{x}_{k+1}-\boldsymbol{x}^*\|}{\|\boldsymbol{x}_k-\boldsymbol{x}^*\|}&=\frac{1}{5}<\frac{1}{5}\sqrt{\frac{3}{2}}=\frac{\lambda_1-\lambda_2}{\lambda_1+\lambda_2}\sqrt{\frac{\lambda_1}{\lambda_2}}\\\frac{f(\boldsymbol{x}_{k+1})-f(\boldsymbol{x}^*)}{f(\boldsymbol{x}_k)-f(\boldsymbol{x}^*)}&=\frac{1}{25}=\left(\frac{\lambda_1-\lambda_2}{\lambda_1+\lambda_2}\right)^2\end{aligned}$$

1.3 牛顿算法

例 1.3 用牛顿法求解⭐⭐⭐⭐⭐

$$\min f(x)=\sqrt{1+x^2}$$

解 0 为最优解

目标函数导数 $f'(x)=\dfrac{x}{\sqrt{1+x^2}}\quad f''(x)=\dfrac{1}{(1+x^2)^{3/2}}$

迭代过程

$$\begin{aligned}x_{k+1}&=x_k-\frac{f'(x_k)}{f''(x_k)}\\&=x_k-x_k(1+x_k^2)\\&=-x_k^3\end{aligned}$$

- 当 $x_0<1$，算法快速收敛到最优解；
- 当 $x_0\geqslant 1$，算法不收敛

例 1.4 用牛顿算法求解⭐⭐⭐⭐⭐

$$\min f(\boldsymbol{x})=4x_1^2+x_2^2-x_1^2x_2$$

解 最优解 $\boldsymbol{x}^* = (0,0)$

函数鞍 $(2\sqrt{2},4)$, $(-2\sqrt{2},4)$

目标函数梯度信息

$$\nabla f(\boldsymbol{x}) = \begin{pmatrix} 8x_1 - 2x_1x_2 \\ 2x_2 - x_1^2 \end{pmatrix}, \quad \nabla^2 f(\boldsymbol{x}) = \begin{pmatrix} 8-2x_2 & -2x_1 \\ -2-x_1 & 2 \end{pmatrix},$$

取精度 $\varepsilon = 10^{-3}$ 和不同初始点

1.4 线性共轭梯度法

例 1.5 利用共轭梯度法求 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} = \boldsymbol{0}$ 或者说，利用共轭梯度法求 ★★★★★

$$\min x_1^2 + \frac{1}{2}x_2^2 + \frac{1}{2}x_3^2$$

其中，

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

解 取初始点 $\boldsymbol{x}_0 = (1,1,1)^{\mathrm{T}}$，迭代过程：

1. $\boldsymbol{x}_0 = (1,1,1)^{\mathrm{T}}$，$\boldsymbol{g}_0 = \boldsymbol{A}\boldsymbol{x}_0 - \boldsymbol{0} = (2,1,1)^{\mathrm{T}}$，$\beta_{-1} = 0$，$\boldsymbol{d}_0 = -\boldsymbol{g}_0$.

$$\begin{aligned} \alpha &= \arg\min f(\boldsymbol{x}_0 + \alpha \boldsymbol{d}_0) \\ &= (1-2\alpha)^2 + \tfrac{1}{2}(1-\alpha)^2 + \tfrac{1}{2}(1-\alpha)^2 = \tfrac{3}{5} \end{aligned}$$

2. $\boldsymbol{x}_1 = \boldsymbol{x}_0 + \alpha_0 \boldsymbol{d}_0 = \frac{1}{5}(-1,2,2)^{\mathrm{T}}$，$\boldsymbol{g}_1 = \boldsymbol{A}\boldsymbol{x}_1 - \boldsymbol{0} = \frac{1}{5}(-1,2,2)^{\mathrm{T}}$，$\beta_0 = \frac{\boldsymbol{g}_1^{\mathrm{T}}\boldsymbol{g}_1}{\boldsymbol{g}_0^{\mathrm{T}}\boldsymbol{g}_0} = \frac{2}{25}$

$$\boldsymbol{d}_1 = -\boldsymbol{g}_1 + \beta_0 \boldsymbol{d}_0 = -\frac{6}{25}(1,-2,-2)^{\mathrm{T}}$$

$$\begin{aligned} \alpha &= \arg\min f(\boldsymbol{x}_0 + \alpha \boldsymbol{d}_0) \\ &= (1-2\alpha)^2 + \tfrac{1}{2}(1-\alpha)^2 + \tfrac{1}{2}(1-\alpha)^2 = \tfrac{3}{5} \end{aligned}$$

3. $\boldsymbol{x}_2 = \boldsymbol{x}_1 + \alpha_1 \boldsymbol{d}_1 = \boldsymbol{0}$，$\|\boldsymbol{g}_2\| = 0$，终止。

k	$\boldsymbol{x}_k$	$\boldsymbol{g}_k$	β_{k-1}	$\boldsymbol{d}_k$	α_k
0	$(1,1,1)^{\mathrm{T}}$	$(2,1,1)^{\mathrm{T}}$	0	$-(2,1,1)^{\mathrm{T}}$	$\frac{3}{5}$
1	$\frac{1}{5}(-1,2,2)^{\mathrm{T}}$	$\frac{1}{5}(-2,2,2)^{\mathrm{T}}$	$\frac{2}{25}$	$-\frac{6}{25}(1,-2,-2)^{\mathrm{T}}$	$\frac{5}{6}$
2	$(0,0,0)^{\mathrm{T}}$	$(0,0,0)^{\mathrm{T}}$			

2 凸优化

2.1 临近点算子

例 2.1 求 $f_2(x) = \begin{cases} 0, & x \neq 0 \\ -\lambda, & x = 0 \end{cases}$ 的临近点算子 $(\lambda > 0)$。★★★★★

解

$$\begin{aligned}
\operatorname{prox}_{f_2}(x) &= \arg\min_{y\in\mathbb{R}} \left\{ f_2(y) + \frac{1}{2}(y-x)^2 \right\} \\
&= \arg\min_{y\in\mathbb{R}} \left\{ \min_{y=0} \left\{ -\lambda + \frac{x^2}{2} \right\}, \min_{y\neq 0} \left\{ \frac{1}{2}(y-x)^2 \right\} \right\} \\
&= \arg\min_{y\in\mathbb{R}} \left\{ -\lambda + \frac{x^2}{2}, 0 \right\} \\
&= \begin{cases} \{0\}, & |x| < \sqrt{2\lambda} \\ \{x\}, & |x| > \sqrt{2\lambda} \\ \{0, x\}, & |x| = \sqrt{2\lambda} \end{cases}
\end{aligned}$$

例 2.2 求 $f_3(x) = \begin{cases} 0, & x \neq 0 \\ \lambda, & x = 0 \end{cases}$ 的临近点算子 $(\lambda > 0)$。★★★★★

解

$$\operatorname{prox}_{f_3}(x) = \begin{cases} \{x\}, & x \neq 0 \\ \emptyset, & x = 0 \end{cases}$$

2.2 交替极小化算法

例 2.3 反例: 连续不可微.$\min \Psi(x_1, x_2) = |3x_1 + 4x_2| + |x_1 - 2x_2|$★★★★★

解 连续凸函数, 水平集有界, 且对任一分量有唯一最优解. 对任意 $\alpha > 0$,

$$\Psi(-4\alpha, t) = |4t - 12\alpha| + |2t + 4\alpha| = \begin{cases} -6t + 8\alpha, & t < -2\alpha, \\ -2t + 16\alpha, & -2\alpha \leqslant t \leqslant 3\alpha \\ 6t - 8\alpha, & t > 3\alpha, \end{cases}$$

$$\Psi(t,3\alpha)=|3t+12\alpha|+|t-6\alpha|=\begin{cases}-4t-6\alpha, & t<-4\alpha,\\ 2t+18\alpha, & -4\alpha\leqslant t\leqslant 6\alpha,\\ 4t+6\alpha, & t>6\alpha,\end{cases}$$

对任意 $\alpha\leqslant 0$

$$-4\alpha=\arg\min_{x_1\in\mathbb{R}}\Psi(x_1,3\alpha),$$
$$3\alpha=\arg\min_{x_2\in\mathbb{R}}\Psi(-4\alpha,x_2)$$

交替极小化方法

$$\min\Psi(x_1,x_2)=|3x_1+4x_2|+|x_1-2x_2|$$

若 x_1 非零，则在首次迭代后，算法滞留在 $(-4\alpha,3\alpha)$ 点（聚点）。

$\boldsymbol{x}=\mathbf{0}$ 为函数的唯一最小值点，$(-4\alpha,3\alpha)$ 既不是该函数的最小值点，也不是其稳定点。

例 2.4 反例：子问题最优解不唯一⭐⭐⭐⭐⭐

解

$$f(x,y,z)=-xy-yz-zx+[x-1]_+^2+[-x-1]_+^2+[y-1]_+^2+[-y-1]_+^2+[z-1]_+^2+[-z-1]_+^2.$$

依次两两固定 y,z,x （注：$[x]_+=\frac{x+|x|}{2}=\max\{x,0\}$）

$$f(x;y,z)=-x(y+z)+\left(\frac{(x-1)+|x-1|}{2}\right)^2+\left(\frac{-(x+1)+|x+1|}{2}\right)^2+a$$
$$=\begin{cases}-x(y+z)+(x+1)^2+a & x<-1\\ -x(y+z)+0+a & -1<x<1\\ -x(y+z)+(x-1)^2+a & x>1\end{cases}$$

$$f_x'(x;y,z)=\begin{cases}2(x+1)-(y+z) & x<-1\\ -(y+z) & -1<x<1\\ 2(x-1)-(y+z) & x>1\end{cases}$$

根据 $f_x'(x;y,z)=0$，得到

$x<-1$ 时，$x=-1+\frac{y+z}{2}$，($y+z<0$ 时取到)

$x>1$ 时，$x=1+\frac{y+z}{2}$，($y+z>0$ 时取到) 故而，有

$$\arg\min_x f(x,y,z)=\begin{cases}\operatorname{sgn}(y+z)\left(1+\frac{1}{2}|y+z|\right), & y+z\neq 0,\\ [-1,1], & y+z=0.\end{cases}$$

同理可得，

$$\arg\min_y f(x,y,z)=\begin{cases}\operatorname{sgn}(x+z)\left(1+\frac{1}{2}|x+z|\right), & x+z\neq 0,\\ [-1,1], & x+z=0,\end{cases}$$

$$\arg\min_z f(x,y,z)=\begin{cases}\operatorname{sgn}(x+y)(1+\frac{1}{2}|x+y|), & x+y\neq 0,\\ [-1,1], & x+y=0.\end{cases}$$

3 约束优化最优性条件

3.1 单约束优化

例 3.1 设 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m\in\mathbb{R}^n$ 满足 $\sum\limits_{i=1}^m \boldsymbol{a}_i\neq\boldsymbol{0}$ 求解优化问题★★★★★

$$\min\{\sum_{i=1}^m\|\boldsymbol{x}-\boldsymbol{a}_i\|^2 \mid \boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1\}$$

解 令

$$\begin{aligned}L(\boldsymbol{x},\lambda)&=\sum_{i=1}^m\|\boldsymbol{x}-\boldsymbol{a}_i\|^2-\lambda(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}-1)\\&=\sum_{i=1}^m\sum_{j=1}^n(\boldsymbol{x}_j-(\boldsymbol{a}_i)_j)^2-\lambda\left(\sum_{j=1}^n\boldsymbol{x}_j^2-1\right)\end{aligned}$$

对 x_k 求偏导，得到

$$\frac{\partial L}{\partial x_k}=\sum_{i=1}^m 2(x_k-(\boldsymbol{a}_i)_k)-\lambda\cdot 2x_k$$

故而有

$$\nabla_x L=\begin{pmatrix}\sum_{i=1}^m 2(x_1-(\boldsymbol{a}_i)_1)-\lambda 2x_1\\ \sum_{i=1}^m 2(x_2-(\boldsymbol{a}_i)_2)-\lambda 2x_2\\ \vdots\\ \sum_{i=1}^m 2(x_n-(\boldsymbol{a}_i)_n)-\lambda 2x_n\end{pmatrix}$$

所以，有

$$\begin{cases}\sum\limits_{i=1}^m 2(\boldsymbol{x}-\boldsymbol{a_i})=2\lambda\boldsymbol{x}\\ \boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1\end{cases}$$

设方程的解为 $\boldsymbol{x}^*$，则 $\boldsymbol{x}^*$ 与 $\sum\limits_{i=1}^m\boldsymbol{a}_i$ 同向或者反向。

若 $\sum\limits_{i=1}^m\boldsymbol{a}_i\neq\boldsymbol{0}$，则

$$\boldsymbol{x}^*=\frac{\sum_{i=1}^m\boldsymbol{a}_i}{\|\sum_{i=1}^m\boldsymbol{a}_i\|}$$

相应地，最优值为

$$
\begin{aligned}
f(\boldsymbol{x}^*) &= \sum_{i=1}^{m} \|\boldsymbol{x}^* - \boldsymbol{a}_i\|^2 \\
&= \sum_{i=1}^{m} \langle \boldsymbol{x}^* - \boldsymbol{a}_i, \boldsymbol{x}^* - \boldsymbol{a}_i \rangle \\
&= \sum_{i=1}^{m} \|\boldsymbol{x}^*\|^2 - 2\left\langle \frac{\sum_{i=1}^{m} \boldsymbol{a}_i}{\|\sum_{i=1}^{m} \boldsymbol{a}_i\|}, \sum_{i=1}^{m} \boldsymbol{a}_i \right\rangle + \sum_{i=1}^{m} \|\boldsymbol{a}_i\|^2 \\
&= m - 2\|\sum_{i=1}^{m} \boldsymbol{a}_i\| + \sum_{i=1}^{m} \|\boldsymbol{a}_i\|^2
\end{aligned}
$$

3.2 线性规划的对偶

例 3.2 线性规划的对偶⭐⭐⭐⭐⭐

$$
\begin{aligned}
\min \quad & \boldsymbol{c}^{\mathrm{T}} \boldsymbol{x} \\
\text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\
& \boldsymbol{x} \geqslant \boldsymbol{0}
\end{aligned}
$$

对偶

$$
\max_{\boldsymbol{u} \geq 0, \boldsymbol{v}} \min_{\boldsymbol{x} \in \mathbb{R}^n} L(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v})
$$

$$
\begin{aligned}
L(\boldsymbol{x}; \boldsymbol{u}, \boldsymbol{v}) &= \boldsymbol{c}^{\mathrm{T}} \boldsymbol{x} - \boldsymbol{u}^{\mathrm{T}} \boldsymbol{x} - \boldsymbol{v}^{\mathrm{T}} (\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}) \\
&= \left(\boldsymbol{c} - \boldsymbol{u} - \boldsymbol{A}^{\mathrm{T}} \boldsymbol{v}\right)^{\mathrm{T}} \boldsymbol{x} + \boldsymbol{v}^{\mathrm{T}} \boldsymbol{b}, \quad \boldsymbol{u} \geqslant \boldsymbol{0}.
\end{aligned}
$$

$$
\begin{aligned}
\min_{\boldsymbol{x} \in \mathbb{R}^n} L(\boldsymbol{x}; \boldsymbol{u}, \boldsymbol{v}) &= \min_{\boldsymbol{x} \in \mathbb{R}^n} \left(\boldsymbol{c} - \boldsymbol{u} - \boldsymbol{A}^{\mathrm{T}} \boldsymbol{v}\right)^{\mathrm{T}} \boldsymbol{x} + \boldsymbol{v}^{\mathrm{T}} \boldsymbol{b} \\
&= \begin{cases} -\infty, & \boldsymbol{c} - \boldsymbol{u} - \boldsymbol{A}^{\mathrm{T}} \boldsymbol{v} \neq \boldsymbol{0}, \\ \boldsymbol{v}^{\mathrm{T}} \boldsymbol{b}, & \boldsymbol{c} - \boldsymbol{u} - \boldsymbol{A}^{\mathrm{T}} \boldsymbol{v} = \boldsymbol{0}, \end{cases}
\end{aligned}
$$

$$
\begin{aligned}
\max \quad & (\boldsymbol{c} - \boldsymbol{u} - \boldsymbol{A}^{\mathrm{T}} \boldsymbol{v})^{\mathrm{T}} \boldsymbol{x} + \boldsymbol{b}^{\mathrm{T}} \boldsymbol{v} \\
\text{s.t.} \quad & \nabla_x L(\boldsymbol{x}; \boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{c} - \boldsymbol{u} - \boldsymbol{A}^{\mathrm{T}} \boldsymbol{v} = \boldsymbol{0} \\
& \boldsymbol{u} \geqslant 0
\end{aligned}
$$

化简得到

$$
\begin{aligned}
\max \quad & \boldsymbol{v}^{\mathrm{T}} \boldsymbol{b} \\
\text{s.t.} \quad & \boldsymbol{A}^{\mathrm{T}} \boldsymbol{v} \leqslant \boldsymbol{c}
\end{aligned}
$$

3.3 混合约束优化的对偶

例 3.3 求下述优化问题的对偶☆☆☆☆☆

$$
\begin{aligned}
\min \quad & f(x)=x_1^2+x_2^2 \\
\text{s.t.} \quad & x_1+x_2-4\geq 0 \\
& x\geq 0
\end{aligned}
$$

解 Lagrange 函数

$$
L(x,\lambda)=x_1^2+x_2^2-\lambda(x_1+x_2-4),\quad x\geq 0,\quad \lambda\geq 0
$$

Lagrange 对偶 $\max\limits_{\lambda\geqslant 0}\min\limits_{x\geqslant 0} L(\lambda,\boldsymbol{x})$

$$
\begin{aligned}
\theta(\lambda) &= \min_{x\geq 0} L(x,\lambda) \\
&= \min_{x\geq 0}\{x_1^2+x_2^2-\lambda(x_1+x_2-4)\} \\
&= \left(\min_{x_1\geq 0}\{x_1^2-\lambda x_1\}+\min_{x_2\geq 0}\{x_2^2-\lambda x_2\}+4\lambda\right) \\
&= -\frac{1}{4}\lambda^2-\frac{1}{4}\lambda^2+4\lambda \\
&= -\frac{1}{2}\lambda^2+4\lambda
\end{aligned}
$$

其中 $x_1^*=x_2^*=\dfrac{1}{2}\lambda$

$$
\max_{\lambda\geq 0}\theta(\lambda)=-\frac{1}{2}\lambda^2+4\lambda=8
$$

例 3.4 求解无约束优化问题☆☆☆☆☆

$$
\min_{\boldsymbol{x}\in\mathbb{R}^n}\frac{1}{2}\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|^2+\lambda\|\boldsymbol{x}\|_1
$$

解 令 $\boldsymbol{y}=\boldsymbol{x}$ 得

$$
\begin{aligned}
\min \quad & \frac{1}{2}\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|^2+\lambda\|\boldsymbol{y}\|_1 \\
\text{s.t.} \quad & \boldsymbol{y}-\boldsymbol{x}=0
\end{aligned}
$$

对偶

$$
\max_{\boldsymbol{z}\in\mathbb{R}^n}\min_{\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^n}=\frac{1}{2}\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|^2+\lambda\|\boldsymbol{y}\|_1+\boldsymbol{z}^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{y}).
$$

变量可分离

$$
\boxed{\min_{\boldsymbol{x}\in\mathbb{R}^n}\left(\frac{1}{2}\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|^2+\boldsymbol{z}^{\mathrm{T}}\boldsymbol{x}\right)}+\boxed{\min_{\boldsymbol{y}\in\mathbb{R}^n}\left(\lambda\|\boldsymbol{y}\|_1-\boldsymbol{z}^{\mathrm{T}}\boldsymbol{y}\right)}
$$

对于 $\min\limits_{\boldsymbol{y}\in\mathbb{R}^n}\left(\lambda\|\boldsymbol{y}\|_1-\boldsymbol{z}^{\mathrm{T}}\boldsymbol{y}\right)$ 有

$$\min_{\boldsymbol{y}\in\mathbb{R}^n}\lambda\|\boldsymbol{y}\|_1-\boldsymbol{z}^{\mathrm{T}}\boldsymbol{y}=\begin{cases}0, & \|\boldsymbol{z}\|_\infty\leqslant\lambda,\\ -\infty, & \|\boldsymbol{z}\|_\infty>\lambda.\end{cases}$$

综上，将 $\boldsymbol{x}^*=\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)^{-1}\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}-\boldsymbol{z}\right)$ 代入得到

$$\begin{aligned}&\max_{\boldsymbol{z}\in\mathbb{R}^n}\min_{\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^n}L(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})\\&=\max_{\|\boldsymbol{z}\|_\infty\leqslant\lambda}\frac{1}{2}\|\boldsymbol{A}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}-\boldsymbol{z})-\boldsymbol{b}\|^2+\boldsymbol{z}^{\mathrm{T}}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}-\boldsymbol{z})\\&=\max_{\|\boldsymbol{z}\|_\infty\leqslant\lambda}-\frac{1}{2}\boldsymbol{z}^{\mathrm{T}}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{z}+\boldsymbol{z}^{\mathrm{T}}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}\\&=\min_{\|\boldsymbol{z}\|_\infty\leqslant\lambda}\frac{1}{2}\boldsymbol{z}^{\mathrm{T}}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{z}-\boldsymbol{z}^{\mathrm{T}}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}\end{aligned}$$

- 原问题为线性约束的凸优化问题, 强对偶定理成立.
- 从而，原问题等价地化为一带简单约束的光滑优化问题.

3.4 等式约束二次规划—消去法

3.5 等式约束二次规划–Lagrange 方法

例 3.5 求解以下无约束二次规划★★★★☆

$$\begin{aligned}\min\quad & Q(\boldsymbol{x})=x_1^2-x_2^2-x_3^2\\ \text{s.t.}\quad & x_1+x_2+x_3=1\\ & x_2-x_3=1\end{aligned}$$

解

$$\boldsymbol{x}_B=\begin{pmatrix}x_1\\x_2\end{pmatrix},\ \boldsymbol{x}_N=x_3$$

则 $x_2=x_3+1,x_1=-2x_3$ 问题转化为

$$\begin{aligned}&\min Q(x_3)=4x_3^2-(x_3+1)^2-x_3^2\\&=2x_3^2-x_3-1\end{aligned}$$

得到

$$x_3^*=\frac{1}{2}\quad \boldsymbol{x}^*=\begin{pmatrix}-1\\3/2\\1/2\end{pmatrix}$$

例 3.6 用 Lagrange 法求解如下问题⭐⭐⭐⭐⭐

$$\begin{aligned}
\min \quad & Q(\boldsymbol{x}) = 3x_1^2 + 2x_1x_2 + x_1x_3 + 2.5x_2^2 + 2x_2x_3 + 2x_3^2 - 8x_1 - 3x_2 - 3x_3 \\
\text{s.t.} \quad & x_1 + x_3 = 3 \\
& x_2 + x_3 = 0
\end{aligned}$$

解 由目标函数及约束条件

$$\boldsymbol{G} = \begin{pmatrix} 6 & 2 & 1 \\ 2 & 5 & 2 \\ 1 & 2 & 4 \end{pmatrix}, \boldsymbol{g} = \begin{pmatrix} -8 \\ -3 \\ -3 \end{pmatrix}, \boldsymbol{A} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \boldsymbol{b} = \begin{pmatrix} 3 \\ 0 \end{pmatrix}$$

KKT 系统

$$\begin{pmatrix} 6 & 2 & 1 & -1 & 0 \\ 2 & 5 & 2 & 0 & -1 \\ 1 & 2 & 4 & -1 & -1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 8 \\ 3 \\ 3 \\ -3 \\ 0 \end{pmatrix}$$

解得

$$\boldsymbol{x}^* = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}, \quad \boldsymbol{\lambda}^* = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$$

3.6 二次规划有效集方法

例 3.7 求解二次规划问题⭐⭐⭐⭐⭐

$$\begin{aligned}
\min \quad & Q(\boldsymbol{x}) = x_1^2 + x_2^2 - 2x_1 - 4x_2 \\
\text{s.t.} \quad & -x_1 - x_2 + 1 \geqslant 0 \\
& x_1, x_2 \geqslant 0
\end{aligned}$$

解 $k = 0$:

取初始点

$$\boldsymbol{x}^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, S_0 = \mathcal{A}(\boldsymbol{x}^{(0)}) = \{2, 3\}$$

求解等式约束优化子问题

$$\begin{aligned}
\min \quad & d_1^2 + d_2^2 - 2d_1 - 4d_2 \\
\text{s.t.} \quad & d_1 = 0 \\
& d_2 = 0
\end{aligned}$$

KKT 条件

$$\begin{cases} 2d_1 - 2 - \lambda_1 = 0 \\ 2d_2 - 4 - \lambda_2 = 0 \\ d_1 = 0 \\ d_2 = 0 \end{cases}$$

得最优解和相应的 Lagrange 乘子

$$\boldsymbol{d}^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \boldsymbol{\lambda}^{(0)} = \begin{pmatrix} -2 \\ -4 \end{pmatrix}$$

因为 $\boldsymbol{d}^{(0)} = \mathbf{0}$，新的迭代点为

$$\boldsymbol{x}^{(1)} = \boldsymbol{x}^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

因为 $\lambda_i < 0$，修正指标集

$$S_1 = S_0/\{3\} = \mathcal{A}_0/\{3\} = \{2\}$$

$k = 1$:

求解子问题

$$\begin{aligned} \min \quad & d_1^2 + d_2^2 - 2d_1 - 4d_2 \\ \text{s.t.} \quad & d_1 = 0 \end{aligned}$$

KKT 条件

$$\begin{cases} 2d_1 - 2 - \lambda_1 = 0 \\ 2d_2 - 4 = 0 \\ d_1 = 0 \end{cases}$$

得最优解

$$\boldsymbol{d}^{(1)} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \boldsymbol{\lambda}^{(1)} = \begin{pmatrix} -2 \end{pmatrix}$$

因为 $\boldsymbol{d}^{(1)} \neq \mathbf{0}$，转第三步，计算步长 α，其中 $i \in \mathcal{I}/S_1 = \{1,2,3\}\,/\,\{2\} = \{1,3\}$
因为

$$\boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{d}^{(1)} = \begin{pmatrix} -1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 2 \end{pmatrix} = -2$$

$$\boldsymbol{a}_2^{\mathrm{T}}\boldsymbol{d}^{(1)} = \begin{pmatrix} 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 2 \end{pmatrix} = 2 > 0$$

$$\begin{aligned} \alpha_1 &= \min\left\{1, \frac{b_i - \boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{x}^{(1)}}{\boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{d}^{(1)}} \mid i = 1,3, \boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{d}^{(1)} < 0\right\} \\ &= \min\left\{1, \frac{-1 - \begin{pmatrix} -1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \end{pmatrix}}{-2}\right\} = \frac{1}{2} \end{aligned}$$

从 $\mathcal{I}/S_1=\{1,3\}$ 中取

$$
\begin{aligned}
\boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{x}^{(2)} &= \begin{pmatrix}-1 & -1\end{pmatrix}\cdot\begin{pmatrix}0\\1\end{pmatrix} = -1 = \boldsymbol{b}_1\\
\boldsymbol{a}_3^{\mathrm{T}}\boldsymbol{x}^{(2)} &= \begin{pmatrix}0 & 1\end{pmatrix}\cdot\begin{pmatrix}0\\1\end{pmatrix} = -1 \neq \boldsymbol{b}_3 = 0
\end{aligned}
$$

不等式积极约束 $i=1$ 令

$$
\boldsymbol{x}^{(2)} = \boldsymbol{x}^{(1)} + \alpha_1\boldsymbol{d}^{(1)} = \begin{pmatrix}0\\1\end{pmatrix},\ S_2 = S_1\cup\{1\} = \{1,2\}
$$

$k=3$:

求解子问题

$$
\begin{aligned}
\min\quad & d_1^2+d_2^2-2d_1-2d_2\\
\text{s.t.}\quad & -d_1-d_2=0\\
& d_1=0
\end{aligned}
$$

KKT 条件

$$
\begin{cases}
2d_1-2+\lambda_1=0\\
2d_2-2+\lambda_1-\lambda_2=0\\
d_1+d_2=0\\
d_1=0
\end{cases}
$$

得最优解

$$
\boldsymbol{d}^{(2)}=\begin{pmatrix}0\\0\end{pmatrix},\ \boldsymbol{\lambda}^{(2)}=\begin{pmatrix}2\\0\end{pmatrix}
$$

由 $\boldsymbol{d}^{(2)}=0$，且对于 $\forall i\in S_k\cap\mathcal{I}(\boldsymbol{x}^{(2)})=\{1,2\},\lambda_i^{(2)}\geqslant 0$ 算法停止。原问题最优解和最优 Lagrange 乘子分别为：

$$
\boldsymbol{x}^*=\boldsymbol{x}^{(2)}=\begin{pmatrix}0\\1\end{pmatrix},\quad \boldsymbol{\lambda}^*=\begin{pmatrix}2\\0\\0\end{pmatrix}
$$

$$
f(\boldsymbol{x}^*)=-3
$$