# 《最优化方法 (第二版)》部分习题参考答案

## 目录

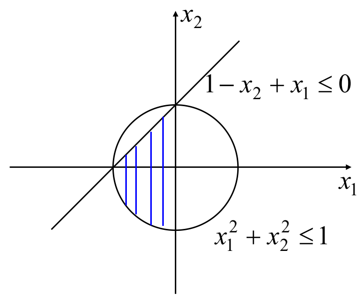
# 第一章　基本概念

1. 考虑由约束

$$x_1^2+x_2^2\leqslant 1,\quad 1-x_2+x_1\geqslant 0,\quad x_1\leqslant 0$$

确定的可行域 $\mathcal{F}$. 判定点 $x^{(1)}=\left(-\dfrac{1}{2},\dfrac{1}{2}\right)^{\mathrm{T}},x^{(2)}=(-1,1)^{\mathrm{T}},x^{(3)}=(-1,0)^{\mathrm{T}},x^{(4)}=\left(0,-\dfrac{1}{2}\right)^{\mathrm{T}}$ 和 $x^{(5)}=\left(-\dfrac{1}{2},-\dfrac{1}{2}\right)^{\mathrm{T}}$ 是否是可行点? 如果是可行点，是内点还是边界点? 是哪个约束的边界点?

**解**: 画出可行域 $\mathcal{F}$，图如下



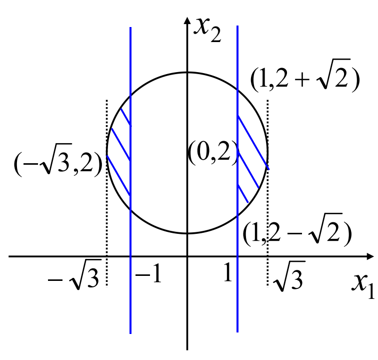

则 $x^{(1)}$ 是可行点，是 $1-x_2+x_1\geqslant 0$ 的边界点；
$x^{(2)}$ 不是可行点；
$x^{(3)}$ 是可行点，是 $x_1^2+x_2^2\leqslant 1$ 和 $1-x_2+x_1\geqslant 0$ 的边界点；
$x^{(4)}$ 是可行点，是 $x_1\leqslant 0$ 的边界点；
$x^{(5)}$ 是可行点，也是内点.

2. 考虑下述约束最优化问题

$$\begin{cases}\min & x_1\\ \text{s.t.} & x_1^2+(x_2-2)^2\leqslant 3,\\ & x_1^2\geqslant 1,\end{cases}$$

画出问题的可行域和目标函数的等位线，并由此确定问题的所有局部最优解和全局最优解.

**解**: 可行域和等位线如下




等位线：$f(x_1)=k$；局部最优解：$x_1=-\sqrt{3},x_2=2$ 和 $x_1=1,x_2\in[2-\sqrt{2},2+\sqrt{2}]$；
全局最优解：$x_1=-\sqrt{3},x_2=2$.

3. 设 $\mathcal{F}=\{x\,|c_i(x)\geqslant 0, i=1,2,\cdots,m\}$，其中所有 $c_i(x)(i=1,2,\cdots,m)$ 都连续. 证明点 $\bar{x}\in\mathcal{F}$ 是内点当且仅当对某一 $\varepsilon>0$ 有

$$\{x\,|||x-\bar{x}||<\varepsilon\}\subset\mathcal{F}.$$

**证**:

$$\bar{x}\text{是内点}\iff c_i(\bar{x})>0\iff\forall\delta>0,U(A,\delta)\subset\mathcal{F}\iff\forall\varepsilon>0,\{x\,|||x-\bar{x}||<\varepsilon\}\subset\mathcal{F}.$$

4. (1) 证明有限个凸集的交集仍然是凸集.

   (2) 设 $D_1=\{x|x_1+x_2\leqslant 1,x_1\geqslant 0\},D_2=\{x|x_1-x_2\geqslant 0,x_1\leqslant 0\}$. 令 $D=D_1\cup D_2$. 证明 $D_1,D_2$ 均是凸集，但 $D$ 却不是凸的，由此得出凸集的并集未必是凸集.

**证**:

(1) 设 $C=\bigcap\limits_{i=1}^{n}C_i$，其中 $C_i(i=1,2,\cdots,n)$ 均是凸集. 取 $x_1,x_2\in C$，则对于 $\forall i$ 都有 $x_1,x_2\in C_i$. 取 $\mu_1,\mu_2\in[0,1]$ 且 $\mu_1+\mu_2=1$，则有 $\mu_1x_1+\mu_2x_2\in C_i$. 因此有 $\mu_1x_1+\mu_2x_2\in\bigcap\limits_{i=1}^{n}C_i=C$，即 $C$ 为凸集. 说明有限个凸集的交集仍然是凸集.

(2) 证明 $D_1$ 是凸集:
$\forall x,y\in D_1,x=(x_1,x_2),y=(y_1,y_2),\lambda\in[0,1],\lambda x+(1-\lambda)y=(\lambda x_1+(1-\lambda)y_1,\lambda x_2+(1-\lambda)y_2)$,

$$\begin{aligned}\lambda x_1+(1-\lambda)y_1+\lambda x_2+(1-\lambda)y_2&=\lambda(x_1+x_2)+(1-\lambda)(y_1+y_2)\\&\leqslant\lambda+1-\lambda=1\end{aligned}$$

$\lambda x_1+(1-\lambda)y_1\geqslant 0\Rightarrow\lambda x+(1-\lambda)y\in D_1\Rightarrow D_1$ 为凸集.
证明 $D_2$ 是凸集:
$\forall x,y\in D_2,x=(x_1,x_2),y=(y_1,y_2),\lambda\in[0,1],\lambda x+(1-\lambda)y=(\lambda x_1+(1-\lambda)y_1,\lambda x_2+(1-\lambda)y_2)$,

$$\begin{aligned}\lambda x_1+(1-\lambda)y_1-\lambda x_2-(1-\lambda)y_2&=\lambda(x_1-x_2)+(1-\lambda)(y_1-y_2)\\&\geqslant 0\end{aligned}$$

$\lambda x_1+(1-\lambda)y_1\leqslant 0\Rightarrow\lambda x+(1-\lambda)y\in D_2\Rightarrow D_2$ 为凸集.
证明 $D$ 不是凸集:
取 $x=(1,0)\in D,y=(0,-1)\in D,\lambda=\dfrac{1}{2}$，则 $\lambda x+(1-\lambda)y=\left(\dfrac{1}{2},-\dfrac{1}{2}\right)\notin D\Rightarrow D$ 不是凸集.

5. 设 $u,v\in\mathbf{R}^n$，对于 2-范数 $\|\cdot\|$，如果

$$\|u+v\|=\|u\|+\|v\|,$$

则 $u=0$ 或 $v=\lambda u$，对某个实数 $\lambda$.

**证**: 向量的 2-范数的定义为

$$\|x\|_2=\left(\sum_{i=1}^{n}x_i^2\right)^{\frac{1}{2}}.$$

设 $u=(u_1,u_2,\cdots,u_n),v=(v_1,v_2,\cdots,v_n)$. 对 $\|u+v\|=\|u\|+\|v\|$ 两边平方，可得

$$\begin{aligned}\|u+v\|^2&=\sum_{i=1}^{n}(u_i+v_i)^2\\&=\sum_{i=1}^{n}(u_i+2u_iv_i+v_i^2)\\&=\sum_{i=1}^{n}u_i^2+2\sum_{i=1}^{n}u_iv_i+\sum_{i=1}^{n}v_i^2\\&=\|u\|^2+2\|u\|\|v\|+\|v\|^2\\&=\sum_{i=1}^{n}u_i^2+\sum_{i=1}^{n}v_i^2+\sqrt{\sum_{i=1}^{n}u_i^2\cdot\sum_{i=1}^{n}v_i^2}\end{aligned}$$

即

$$\left(\sum_{i=1}^{n}u_iv_i\right)^2=\sum_{i=1}^{n}u_i^2\cdot\sum_{i=1}^{n}v_i^2,$$

而由柯西不等式

$$\left(\sum_{i=1}^{n}u_iv_i\right)^2\leqslant\sum_{i=1}^{n}u_i^2\cdot\sum_{i=1}^{n}v_i^2,$$

当且仅当 $\dfrac{v_1}{u_1}=\dfrac{v_2}{u_2}=\cdots=\dfrac{v_n}{u_n}=\lambda(\lambda\neq 0)$ 或 $u_i,v_i$ 其中一方全为 0 时，取到等号，即 $u=0$ 或 $v=\lambda u$.

6. 设 $f(x)$ 为定义在凸集 $D\subset\mathbf{R}^n$ 上的凸函数，$\alpha$ 为一个给定的实数，称集合

$$\mathcal{T}=\{x|f(x)\leqslant\alpha\}$$

为函数 $f(x)$ 关于实数 $\alpha$ 的水平集. 证明对任意实数 $\alpha$，集合 $\mathcal{T}$ 是凸集.

**证**: 对于 $\forall x_1,x_2\in\mathcal{T}$，根据 $\mathcal{T}$ 的定义则有 $f(x_1)\leqslant\alpha,f(x_2)\leqslant\alpha$. 由于 $D$ 是凸集，则对于 $\forall\lambda\in[0,1]$，必有

$$\lambda x_1+(1-\lambda)x_2\in D$$

又由于 $f(x)$ 是 $D$ 上的凸函数，则有

$$\begin{aligned}f\left(\lambda x_1+(1-\lambda)x_2\right)&\leqslant\lambda f(x_1)+(1-\lambda)f(x_2)\\&\leqslant\lambda\alpha+(1-\lambda)\alpha=\alpha\end{aligned}$$

因此 $\lambda x_1+(1-\lambda)x_2\in\mathcal{T}$，故对任意实数 $\alpha$，集合 $\mathcal{T}$ 是凸集.

7. 设 $f_i(x)(i=1,2,\cdots,m)$ 是定义在凸集 $D\subset\mathbf{R}^n$ 上的凸函数，证明函数

$$g(x)=\sum_{i=1}^{m}\alpha_if_i(x)$$

也是 $D$ 上的凸函数，其中 $\displaystyle\sum_{i=1}^{m}\alpha_i=1,\alpha_i\geqslant 0(i=1,2,\cdots,m)$，即凸函数的组合还是凸函数.

**证**: 对于 $\forall x,y\in D,\lambda\in[0,1]$，由于 $\sum\limits_{i=1}^{m}\alpha_i=1,\alpha_i\geqslant 0(i=1,2,\cdots,m)$，则有

$$\begin{aligned}
g(\lambda x+(1-\lambda)y)&=\sum_{i=1}^{m}\alpha_i f_i(\lambda x+(1-\lambda)y)\\
&\leqslant\sum_{i=1}^{m}\alpha_i[\lambda f_i(x)+(1-\lambda)f_i(y)]\\
&=\lambda\sum_{i=1}^{m}\alpha_i f_i(x)+(1-\lambda)\sum_{i=1}^{m}\alpha_i f_i(y)\\
&=\lambda g(x)+(1-\lambda)g(y)
\end{aligned}$$

则 $g(x)=\sum\limits_{i=1}^{m}\alpha_i f_i(x)$ 也是 $D$ 上的凸函数.

8. 设 $c_i(x)(i\in E)$ 是线性函数，$c_i(x)(i\in I)$ 是凹函数，证明约束集

$$\Omega=\left\{x\;\middle|\;\begin{array}{ll}c_i(x)=0, & i\in E\\ c_i(x)\geqslant 0, & i\in I\end{array}\right\}$$

是凸集.

**证**: $\forall x_1,x_2\in\Omega,\lambda\in[0,1]$，则可以分成以下三种情况：

(1) 若 $c_i(x_1)=0,c_i(x_2)=0$，则

$$c_i(\lambda x_1+(1-\lambda)x_2)=\lambda c_i(x_1)+(1-\lambda)c_i(x_2)=0\Rightarrow\lambda x_1+(1-\lambda)x_2\in\Omega,$$

故 $\Omega$ 是凸集.

(2) 若 $c_i(x_1)\geqslant 0,c_i(x_2)\geqslant 0$，则

$$-c_i(\lambda x_1+(1-\lambda)x_2)\leqslant-\lambda c_i(x_1)-(1-\lambda)c_i(x_2)\leqslant 0\Rightarrow\lambda x_1+(1-\lambda)x_2\in\Omega,$$

故 $\Omega$ 是凸集.

(3) 若 $c_i(x_1)=0,c_i(x_2)\geqslant 0$(当 $x_1,x_2$ 互换时，与此情况同理)，则

$$-c_i(\lambda x_1+(1-\lambda)x_2)\leqslant-\lambda c_i(x_1)-(1-\lambda)c_i(x_2)=-(1-\lambda)c_i(x_2)\leqslant 0\Rightarrow\lambda x_1+(1-\lambda)x_2\in\Omega,$$

故 $\Omega$ 是凸集.

综上，约束集 $\Omega$ 是凸集.

9. 设多变量函数 $f:\mathbf{R}^n\to\mathbf{R}^1$ 和单变量函数 $\phi:\mathbf{R}^1\to\mathbf{R}^1$ 都是凸函数，证明复合函数 $h(x)=\phi(f(x)),x\in\mathbf{R}^n$ 是凸函数. (补充：函数 $\phi$ 单调非减)

**证**: $\because f$ 是凸函数，$\forall x,y\in\mathbf{R}^n,\lambda\in[0,1]$，

$$\therefore f(\lambda x+(1-\lambda)y\leqslant\lambda f(x)+(1-\lambda)f(y).$$

$$\begin{aligned}h(\lambda x+(1-\lambda)y)&=\phi(f(\lambda x+(1-\lambda)y)\\&\leqslant \phi(\lambda f(x)+(1-\lambda)f(y))\quad (\text{函数}\phi\text{单调非减})\\&\leqslant \lambda\phi(f(x))+(1-\lambda)\phi(f(y))\quad (\text{函数}\phi\text{是凸函数})\\&=\lambda h(x)+(1-\lambda)h(y)\end{aligned}$$

则复合函数 $h(x)=\phi(f(x)), x\in \mathbf{R}^n$ 是凸函数.

10. 设 $c_i(x)(i=1,2,\cdots,m)$ 为 $\mathbf{R}^n$ 上的凹函数，$f(x)$ 是 $\mathbf{R}^n$ 上的凸函数，证明函数

$$P(x)=f(x)-\mu\sum_{i=1}^{m}\log c_i(x)$$

在集合 $D=\{x|c_i(x)>0\}$ 上是凸函数. (补充：$\mu>0$)

**证**: 易证：$\log c_i(x)$ 是凹函数，所以 $\forall x,y\in D$，则

$$\begin{aligned}P(\lambda x+(1-\lambda)y)&=f(\lambda x+(1-\lambda)y)-\mu\sum_{i=1}^{m}\log c_i(\lambda x+(1-\lambda)y)\\&\leqslant \lambda f(x)+(1-\lambda)f(y)-\mu\sum_{i=1}^{m}\log(\lambda c_i(x)+(1-\lambda)c_i(y))\\&\leqslant \lambda f(x)+(1-\lambda)f(y)-\lambda\mu\sum_{i=1}^{m}\log c_i(x)-(1-\lambda)\mu\sum_{i=1}^{m}\log c_i(y)\\&=\lambda P(x)+(1-\lambda)P(y)\end{aligned}$$

所以，该函数在集合 $D=\{x|c_i(x)>0\}$ 上是凸函数.

11. 设 $A$ 是 $m\times n$ 矩阵，$b$ 是 $n$ 维向量，利用 Farkas 引理证明下述两组系统中有且仅有一组有解:

$$\begin{aligned}&Ax\leqslant 0, x\geqslant 0, b^{\mathrm{T}}x>0;\\&A^{\mathrm{T}}y=b, y\geqslant 0.\end{aligned}$$

**证**: 先给这个系统标号：

$$\begin{aligned}&Ax\leqslant 0, x\geqslant 0, b^{\mathrm{T}}x>0;\quad (1)\\&A^{\mathrm{T}}y=b, y\geqslant 0;\quad (2)\end{aligned}$$

要证 (1)(2) 中有且仅有一组解，即证 (1) 有解 $\Longleftrightarrow$ (2) 无解。
先证充分性：若 (1) 有解，则说明 $\exists\bar{x}\geqslant 0$ 使得 $A\bar{x}\leqslant 0, b^{\mathrm{T}}\bar{x}>0$. 用反证法证明 (2) 无解，若在 (1) 的条件下，(2) 有解，则 $\exists\bar{y}\geqslant 0$ 使得 $A^{\mathrm{T}}\bar{y}=b$，即 $\bar{y}^{\mathrm{T}}A=b^{\mathrm{T}}$，两边同时右乘 $\bar{x}$，则有

$$\bar{y}^{\mathrm{T}}A\bar{x}=b^{\mathrm{T}}\bar{x},$$

由于 $\bar{y}\geqslant 0, A\bar{x}\leqslant 0$，则 $\bar{y}^{\mathrm{T}}A\bar{x}=b^{\mathrm{T}}\bar{x}\leqslant 0$，与 (1) 的 $b^{\mathrm{T}}\bar{x}>0$ 矛盾，则若 (1) 有解，则 (2) 无解.
再证必要性：若 (2) 无解，令 $D=\{z|z=A^{\mathrm{T}}y, y\geqslant 0\}$，则 $D\subset\mathbf{R}^n$ 为非空闭凸集. 由定理 1.2.6，$\exists a\in\mathbf{R}^n(a\geqslant 0), \beta\in\mathbf{R}$ 使得

$$a^{\mathrm{T}}x\leqslant\beta<a^{\mathrm{T}}b,\quad \forall x\in D,$$

由于 $0 \in D$，则可知 $\beta \geqslant 0, a^{\mathrm{T}} b > 0$. 同时

$$\beta \geqslant a^{\mathrm{T}} x = a^{\mathrm{T}} A^{\mathrm{T}} y = y^{\mathrm{T}} A a, \quad \forall y \geqslant 0,$$

由于 $y \geqslant 0$ 可以任意大，则 $Aa \leqslant 0$，即 $a$ 是 (1) 的解，则若 (2) 无解，则 (1) 有解. 证毕.

12. 设 $x^*$ 是凸规划问题的一个解. 证明如果目标函数严格凸，则 $x^*$ 是唯一全局最优解.

**证**: 先证 $x^*$ 是一个全局最优解：设 $x^*$ 是凸规划问题的一个解 (即局部最优解)，但不是全局最优解，则 $\exists y$ 满足 $f(y) < f(x^*)$. 由可行域的凸性，对于 $\lambda \in [0,1]$，点 $\lambda x^* + (1-\lambda) y$ 都是可行点. 又根据目标函数的凸性有

$$\begin{aligned} f(\lambda x^* + (1-\lambda) y) &\leqslant \lambda f(x^*) + (1-\lambda) f(y) \\ &\leqslant \lambda f(x^*) + (1-\lambda) f(x^*) = f(x^*) \end{aligned}$$

这表明在 $x^*$ 的任意小的邻域内都存在函数值小于 $f(x^*)$ 的可行点，这与 $x^*$ 是局部最优解相矛盾，则 $x^*$ 是一个全局最优解.
再证 $x^*$ 是唯一的：由于目标函数是严格凸的，设 $x^* \neq y^*$ 都是全局最优解，则 $f(x^*) = f(y^*)$. 由严格凸函数的定义，而 $\forall \lambda \in (0,1)$，有

$$f(\lambda x^* + (1-\lambda) y^*) < \leqslant \lambda f(x^*) + (1-\lambda) f(y^*) = f(x^*).$$

这与 $x^*$ 是全局最优解矛盾，则 $x^*$ 是唯一的. 综上，$x^*$ 是唯一全局最优解.

13. 设 $f(x)$ 定义在集合 $D \subset \mathbf{R}^n$ 上且连续可微.

(1) 证明 $f(x)$ 是 $D$ 上严格凸函数的充分必要条件为

$$f(y) > f(x) + \nabla f(x)^{\mathrm{T}} (y - x), \quad \forall x, y \in D, x \neq y.$$

(2) 证明 $f(x)$ 是 $D$ 上一致凸函数的充分必要条件为

$$f(y) \geqslant f(x) + \nabla f(x)^{\mathrm{T}} (y - x) + c \|y - x\|^2, \quad \forall x, y \in D,$$

其中 $c > 0$ 是某个常数. (补充：$c$ 应该 $< 0$)

**证**:

(1) 必要性：因为 $f(x)$ 是严格凸函数，则对 $\forall \lambda \in (0,1)$，有

$$f(\lambda x + (1-\lambda) y) < \lambda f(x) + (1-\lambda) f(y),$$

即

$$f(x + \lambda(y - x)) < f(x) + \lambda(f(y) - f(x)).$$

由泰勒公式，有

$$f(x + \lambda(y - x)) = f(x) + \lambda \nabla f(x)^{\mathrm{T}} (y - x) + o(\|\lambda(y - x)\|).$$

因此得到

$$f(y)-f(x)>\nabla f(x)^{\mathrm{T}}(y-x)+\frac{o(\|\lambda(y-x)\|)}{\lambda}.$$

令 $\lambda\to 0^+$，则

$$f(y)>f(x)+\nabla f(x)^{\mathrm{T}}(y-x),\quad \forall x,y\in D,x\neq y.$$

必要性得证.

充分性：令 $z=\lambda x+(1-\lambda)y$，则

$$f(x)>f(z)+\nabla f(z)^{\mathrm{T}}(x-z),$$

同理

$$f(y)>f(z)+\nabla f(z)^{\mathrm{T}}(y-z),$$

从而

$$\lambda f(x)+(1-\lambda)f(y)-f(z)>\nabla f(z)^{\mathrm{T}}[\lambda(x-z)+(1-\lambda)(y-z)]=0$$

即

$$f(\lambda x+(1-\lambda)y)<\lambda f(x)+(1-\lambda)f(y).$$

充分性得证.

(2) 一致凸函数的定义为：$f(x)$ 是定义在 $D$ 上的实值函数，若存在 $k>0$，对 $\forall x,y\in D,\lambda\in[0,1]$，有

$$f(\lambda x+(1-\lambda)y)\leqslant\lambda f(x)+(1-\lambda)f(y)+k\lambda(1-\lambda)\|x-y\|^2,$$

则称 $f(x)$ 是 $D$ 上的一致凸函数.

必要性：因为 $f(x)$ 是一致凸函数，则

$$f(\lambda x+(1-\lambda)y)\leqslant\lambda f(x)+(1-\lambda)f(y)+k\lambda(1-\lambda)\|x-y\|^2,$$

即

$$f(x+\lambda(y-x))\leqslant f(x)+\lambda(f(y)-f(x))+k\lambda(1-\lambda)\|x-y\|^2.$$

由泰勒公式，有

$$f(x+\lambda(y-x))=f(x)+\lambda\nabla f(x)^{\mathrm{T}}(y-x)+o(\|\lambda(y-x)\|).$$

因此得到

$$f(y)-f(x)\geqslant\nabla f(x)^{\mathrm{T}}(y-x)+\frac{o(\|\lambda(y-x)\|)}{\lambda}-k(1-\lambda)\|x-y\|^2.$$

令 $\lambda\to 0^+$，则

$$f(y)\geqslant f(x)+\nabla f(x)^{\mathrm{T}}(y-x)-k\|x-y\|^2,$$

令 $c=-k<0$，则

$$f(y)\geqslant f(x)+\nabla f(x)^{\mathrm{T}}(y-x)+c\|y-x\|^2,\quad \forall x,y\in D,x\neq y.$$

必要性得证.

充分性：令 $z=\lambda x+(1-\lambda)y$，则

$$f(x)\geqslant f(z)+\nabla f(z)^{\mathrm{T}}(x-z)+c\|x-z\|^2,$$

同理

$$f(y) \geqslant f(z) + \nabla f(z)^{\mathrm{T}}(y-z) + c\|y-z\|^2,$$

从而

$$\begin{aligned} \lambda f(x) + (1-\lambda)f(y) - f(z) &\geqslant \nabla f(z)^{\mathrm{T}}[\lambda(x-z) + (1-\lambda)(y-z)] + c\lambda\|x-z\|^2 + c(1-\lambda)\|y-z\|^2 \\ &\geqslant 0 + c\lambda\|x-z\|^2 + c(1-\lambda)\|y-z\|^2 \\ &\geqslant c\lambda(1-\lambda)\|y-x\|^2 \end{aligned}$$

令 $k=-c>0$，则

$$f(\lambda x + (1-\lambda)y) \leqslant \lambda f(x) + (1-\lambda)f(y) + k\lambda(1-\lambda)\|x-y\|^2.$$

充分性得证.

14. 求出函数

$$f(x) = 2x_1^3 - 3x_1^2 - 6x_1x_2(x_1 - x_2 - 1)$$

的所有稳定点，其中哪一个点是极小值点？哪一个点是极大值点？有没有既不是极大又不是极小的点？

**解**: $\nabla f = \begin{pmatrix} 6x_1^2 - 6x_1 - 12x_1x_2 + 6x_2^2 + 6x_2 \\ x_1(-6x_1 + 12x_2 + 6) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$,

$\nabla^2 f = \begin{pmatrix} 12x_1 - 6 - 12x_2 & -12x_1 + 12x_2 + 6 \\ -12x_1 + 12x_2 + 6 & 12x_1 \end{pmatrix}$，则所有稳定点为

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

依次代入 $\nabla^2 f$，分别使得 $\nabla^2 f$ 为不定，不定，负定，正定，则依次为非极大极小点，非极大极小点，局部极大点，局部极小点.

15. 确定线性函数 $f(x) = 2x_1 - x_2 + 3x_3$ 的所有下降方向. 请问这样的下降方向是否同所在点的位置有关？

**解**: $\nabla f(x)^{\mathrm{T}} = (2,-1,3)$，则 $\nabla f(x)^{\mathrm{T}} s < 0$，设 $s = (s_1, s_2, s_3)^{\mathrm{T}}$，则

$$2s_1 - s_2 + 3s_3 < 0,$$

故函数 $f(x) = 2x_1 - x_2 + 3x_3$ 的所有下降方向为 $\mathcal{D}(x) = \{s | 2s_1 - s_2 + 3s_3 < 0, s = (s_1, s_2, s_3)^{\mathrm{T}}\}$. 这样的下降方向与所在点的位置无关.

16. 考虑约束最优化问题

$$\begin{cases} \min & -x_1 + x_2 \\ \text{s.t.} & x_1 + 2x_2 \leqslant 2, \\ & x_1 \geqslant 0, x_2 \geqslant 0, \end{cases}$$

(1) 确定在点 $x^{(1)} = (0,0)^{\mathrm{T}}, x^{(2)} = (0,1)^{\mathrm{T}}, x^{(3)} = \left(1, \dfrac{1}{2}\right)^{\mathrm{T}}$ 和 $x^{(4)} = (2,0)^{\mathrm{T}}$ 处的可行方向.

(2) 在这些点是否存在可行的下降方向，并由此从中确定最优解.

**解**: 本题需要教材 7.1 节的知识.

先将所求化为标准形式:

$$\begin{cases}\min & -x_1+x_2\\ \text{s.t.} & 2-x_1-2x_2\geqslant 0,\\ & x_1\geqslant 0, x_2\geqslant 0,\end{cases}$$

(1) 对于 $x^{(1)}=(0,0)^{\mathrm{T}}$，有

$$\begin{cases}d^{(1)\,\mathrm{T}}\begin{pmatrix}1\\0\end{pmatrix}\geqslant 0,\\ d^{(1)\,\mathrm{T}}\begin{pmatrix}0\\1\end{pmatrix}\geqslant 0,\end{cases}\quad\Rightarrow d_1^{(1)}\geqslant 0, d_2^{(1)}\geqslant 0, d^{(1)}\neq 0;$$

对于 $x^{(2)}=(0,1)^{\mathrm{T}}$，有

$$\begin{cases}d^{(2)\,\mathrm{T}}\begin{pmatrix}-1\\-2\end{pmatrix}\geqslant 0,\\ d^{(2)\,\mathrm{T}}\begin{pmatrix}1\\0\end{pmatrix}\geqslant 0,\end{cases}\quad\Rightarrow -d_1^{(2)}-2d_2^{(2)}\geqslant 0, d_1^{(2)}\geqslant 0, d^{(2)}\neq 0;$$

对于 $x^{(3)}=\left(1,\dfrac{1}{2}\right)^{\mathrm{T}}$，有

$$\begin{cases}d^{(3)\,\mathrm{T}}\begin{pmatrix}-1\\-2\end{pmatrix}\geqslant 0,\end{cases}\quad\Rightarrow -d_1^{(3)}-2d_2^{(3)}\geqslant 0, d^{(3)}\neq 0;$$

对于 $x^{(4)}=(2,0)^{\mathrm{T}}$，有

$$\begin{cases}d^{(4)\,\mathrm{T}}\begin{pmatrix}-1\\-2\end{pmatrix}\geqslant 0,\\ d^{(4)\,\mathrm{T}}\begin{pmatrix}0\\1\end{pmatrix}\geqslant 0,\end{cases}\quad\Rightarrow -d_1^{(4)}-2d_2^{(4)}\geqslant 0, d_2^{(4)}\geqslant 0, d^{(4)}\neq 0.$$

(2) $\nabla f(x)^{\mathrm{T}}=(-1,2)$，$x^{(1)}$ 存在可行下降方向，所以从 $x^{(2)},x^{(3)},x^{(4)}$ 中寻找最优解，由于 $f(x^{(2)})=1, f(x^{(3)})=-\dfrac{1}{2}, f(x^{(4)})=-2$，所以最优解为 $x^{(4)}=(2,0)^{\mathrm{T}}$.

17. 设

$$q(x)=\frac{1}{2}x^{\mathrm{T}}Bx+b^{\mathrm{T}}x+c$$

其中，$x\in\mathbf{R}^n, B\in\mathbf{R}^{n\times n}$ 为对称矩阵，$b\in\mathbf{R}^n, c\in\mathbf{R}$. 证明 $q(x)$ 有唯一极小点的充分必要条件是矩阵 $B$ 正定，并给出这个极小解.

**证**: 充分性:

$$\begin{cases}\nabla q(x)=Bx+b=0, & x=-B^{-1}b\\ \nabla^2 q(x)=B, & B\text{对称正定}\end{cases}\Rightarrow q(x)\text{有唯一极小点}.$$

必要性：$q(x)$ 有唯一极小点 $x^* \Rightarrow \nabla q(x^*) = 0, B \geqslant 0$. 假设 $B$ 有特征值 0，即 $\exists u \neq 0, Bu = 0$，则

$$\begin{aligned}
q(x^*+u) &= \frac{1}{2}(x^*+u)^{\mathrm{T}}B(x^*+u) + b^{\mathrm{T}}(x^*+u) + c \\
&= \frac{1}{2}x^{*\mathrm{T}}Bx^* + b^{\mathrm{T}}x^* + c + b^{\mathrm{T}}u \\
&= q(x^*) + b^{\mathrm{T}}u
\end{aligned}$$

(1) 如果 $b^{\mathrm{T}}u = 0$，与唯一极小矛盾；

(2) 如果 $b^{\mathrm{T}}u < 0$，$q(x^*+u) < q(x^*)$，与极小矛盾；

(3) 如果 $b^{\mathrm{T}}u > 0$，$q(x^*-u) = q(x^*) - b^{\mathrm{T}}u < q(x^*)$，与极小矛盾.

则必要性得证.
必要性证明或用

$$\begin{aligned}
q(x) - q(x^*) &= \nabla q(x^*)(x-x^*) + \frac{1}{2}(x-x^*)^{\mathrm{T}}B(x-x^*) \\
&= \frac{1}{2}(x-x^*)^{\mathrm{T}}B(x-x^*) > 0 \quad (\forall x^* \neq x, x-x^* \neq 0)
\end{aligned}$$

18. 考虑问题

$$\begin{cases} \min & \dfrac{1}{2}x^{\mathrm{T}}Bx + b^{\mathrm{T}}x + c \\ \text{s.t.} & \|x\| \leqslant \rho, \end{cases}$$

$\rho > 0$ 为一个给定的常数，证明

(1) 如果 $B$ 正定，且 $\|B^{-1}b\| \leqslant \rho$，则问题的最优解为

$$x^* = -B^{-1}b;$$

(2) 如果 (1) 的条件不满足，则问题的最优解是下述方程组对某一 $\mu > 0$ 的解

$$\begin{cases} (B+\mu I)x = -b, \\ \|x\| = \rho, \end{cases}$$

其中 $\mu > 0$ 使得矩阵 $B + \mu I$ 至少正半定.

**证**:

(1) 由 17 题可知，该问题的最优解为 $x^* = -B^{-1}b$;

(2) **略**.

19. 叙述 KKT 条件和本书介绍的三个约束规范条件.
**解**: (均在教材 7.1 节中) KKT 条件：设 $x^*$ 是问题

$$\begin{aligned}
\min \quad & f(x) \\
\text{s.t.} \quad & c_i(x) = 0, i = 1, 2, \cdots, m', \\
& c_i(x) \geqslant 0, i = m'+1, m'+2, \cdots, m, \\
& x \in \mathbf{R}^n, m' \leqslant m.
\end{aligned}$$

的局部极小点，设 $f(x), c_i(x)(i=1,\cdots,m)$ 在 $x^*$ 的领域内一阶连续可微. 如果约束规范条件

$$\mathcal{SFD}(x^*,X)=\mathcal{LFD}(x^*,X)$$

成立，则存在 $\lambda_i^*(i=1,2,\cdots,m)$ 满足 KKT 条件：

$$\begin{aligned}
&\nabla f(x^*)=\sum_{i=1}^{m}\lambda_i^*\nabla c_i(x^*),\\
&c_i(x^*)=0, i\in\mathcal{E},\\
&c_i(x^*)\geqslant 0, i\in\mathcal{I},\\
&\lambda_i^*\geqslant 0, i\in\mathcal{I},\\
&\lambda_i^*c_i(x^*)=0, i\in\mathcal{I}.
\end{aligned}$$

三个约束规范条件：

(1) 约束规范条件 (CQ)：$\mathcal{SFD}(x^*,X)=\mathcal{LFD}(x^*,X)$.

(2) 线性函数约束规范条件 (LFCQ)：所有的约束函数 $c_i(x)(x\in\mathcal{E}\cap\mathcal{I}(x^*))$ 都是线性函数.

(3) 线性无关约束规范条件 (LICQ)：约束函数的梯度 $\nabla c_i(x)(x\in\mathcal{E}\cap\mathcal{I}(x^*))$ 线性无关.

20. 考虑下述三个数列：

$$\begin{aligned}
u_k&=\frac{1}{c^{2-k}}(c>0),\\
v_k&=\frac{1}{k^k},\\
w_k&=a^{p^k}(0<a<1,p>1),
\end{aligned}$$

证明它们分别具有线性，超线性和阶数为 $p$ 的收敛率. (补充：$c<1$)

**证**: 易得：$u^*=v^*=w^*=0$，则

$$\lim_{k\to\infty}\frac{|\varepsilon_{k+1}|}{|\varepsilon_k|^1}=\lim_{k\to\infty}\frac{c^{k+1-2}}{c^{k-2}}=c\Rightarrow \text{线性收敛}$$

$$\lim_{k\to\infty}\frac{\left|\frac{1}{(k+1)^{k+1}}-0\right|}{\left|\frac{1}{k^k}\right|^1}=\lim_{k\to\infty}\frac{\left(\frac{k}{k+1}\right)^{k+1}\cdot\frac{1}{k}}{\left(1-\frac{1}{k+1}\right)^{k+1}}=0\Rightarrow \text{超线性收敛}$$

$$\lim_{k\to\infty}\frac{|a^{p^{k+1}}-0|}{|a^{p^k}|^p}=1\Rightarrow \text{阶数为}p$$

21. 设 $a_0=b$，考虑数列 $a_{k+1}=\frac{1}{2}\left(a_k+\frac{b}{a_k}\right)$，证明数列 $\{a_k\}$ 的极限为 $a^*=\sqrt{b}$，且收敛率是二次的. (补充：$b>0$)

**证**: 先证收敛：$a_k>0, a_{k+1}=\frac{1}{2}\left(a_k+\frac{b}{a_k}\right)\geqslant\sqrt{b}>0$，则

$$\frac{a_{k+1}}{a_k}=\frac{1}{2}+\frac{1}{2}\cdot\frac{b}{a_k^2}\leqslant 1\Rightarrow\sqrt{b}\leqslant a_{k+1}\leqslant a_k\leqslant b\Rightarrow\{a_k\}\text{单调有界},$$

则 $\lim\limits_{k\to\infty}a_{k+1}=a^*=\lim\limits_{k\to\infty}\frac{1}{2}\left(a_k+\frac{b}{a_k}\right)=\frac{1}{2}\left(a^*+\frac{b}{a^*}\right)\Rightarrow a^*=\sqrt{b}$.

再证二阶收敛：

$$\lim_{k\to\infty}\frac{|a_{k+1}-\sqrt{b}|}{|a_k-\sqrt{b}|^2}=\frac{1}{2\sqrt{b}}\Rightarrow\{a_k\}\text{收敛率是二次的}.$$

**第一章：补充题目**

1. 证明：Rosenbrock 函数 (见附录)

$$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

不是 $\mathbf{R}^2$ 上的凸函数，但是具有唯一的全局极小点.

**证**: 法一：取 $x = (0,0)^{\mathrm{T}}, y = (1,1)^{\mathrm{T}}, \lambda = \dfrac{1}{2}$，则

$$\begin{aligned} f\left(\frac{1}{2}x + \frac{1}{2}y\right) &= f\left(\left(\frac{1}{2}, \frac{1}{2}\right)^{\mathrm{T}}\right) = \frac{13}{2}, \\ f(x) &= 1, \\ f(y) &= 0, \\ \frac{1}{2}f(x) + \frac{1}{2}f(y) &= \frac{1}{2} < \frac{13}{2} \end{aligned}$$

故 Rosenbrock 函数不是 $\mathbf{R}^2$ 上的凸函数.

法二：利用 Hessen 矩阵进行判断，

$$\begin{aligned} \frac{\partial f}{\partial x_1} &= 2x_1 - 400x_1(-x_1^2 + x_2) - 2, \frac{\partial f}{\partial x_2} = -200x_1^2 + 200x_2, \\ \frac{\partial^2 f}{\partial x_1^2} &= 1200x_1^2 - 400x_2 + 2, \frac{\partial^2 f}{\partial x_2^2} = 200, \\ \frac{\partial^2 f}{\partial x_1 x_2} &= -400x_1 \end{aligned}$$

故

$$\nabla^2 f(x) = \begin{pmatrix} 1200x_1^2 - 400x_2 + 2 & -400x_1 \\ -400x_1 & 200 \end{pmatrix}, |\nabla^2 f(x)| = 80000x_1^2 - 80000x_2 + 400,$$

故 Rosenbrock 函数在 $\mathbf{R}^2$ 上不是恒为半正定的，Rosenbrock 函数不是 $\mathbf{R}^2$ 上的凸函数.

由于

$$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2 \geqslant 0,$$

当且仅当 $x_2 - x_1^2 = 0, 1 - x_1 = 0$，即 $x_1 = x_2 = 1$ 时取到等号，即最小值，故 Rosenbrock 函数在 $\mathbf{R}^2$ 上具有唯一的全局极小点.

# 第二章　线性规划

1. 用图解法确定下述线性规划问题的最优解.

$$
(1)\ \begin{cases}
\min & 8x_1+5x_2 \\
\text{s.t.} & -x_1+x_2\geqslant 0, \\
 & 6x_1+11x_2\geqslant 66, \\
 & 2x_1+x_2\geqslant 10, \\
 & x_1\geqslant 0, x_2\geqslant 0.
\end{cases}
$$

$$
(2)\ \begin{cases}
\max & 2x_1+3x_2 \\
\text{s.t.} & -x_1+2x_2\geqslant 12, \\
 & -2x_1+x_2\geqslant 3, \\
 & x_1\geqslant 0, x_2\geqslant 0.
\end{cases}
$$

$$
(3)\ \begin{cases}
\min & 4x_1-3x_2 \\
\text{s.t.} & -11x_1+10x_2\leqslant 20, \\
 & x_1\geqslant 0, 0\leqslant x_2\leqslant 6.
\end{cases}
$$

**解:**

(1) $\min=\dfrac{89}{2}, x_1=\dfrac{11}{4}, x_2=\dfrac{9}{2}$.

(2) 无界解.

(3) $\min=-6, x_1=0, x_2=2$.

2. 确定由下述约束条件所形成的可行域的所有顶点，并计算函数 $f(x)$ 在这些顶点的函数值.

$$
(1)\quad \begin{aligned}
& f(x)=3x_1+2x_2, \\
& x_1+x_2\leqslant 3, \\
& x_1-2x_2\leqslant 4, \\
& 0\leqslant x_1\leqslant 4, \\
& 0\leqslant x_2\leqslant 1.
\end{aligned}
$$

$$
(2)\quad \begin{aligned}
& f(x)=-4x_1-3x_2, \\
& 2x_1+x_2\leqslant 10, \\
& x_1+x_2\leqslant 10, \\
& 5x_1+3x_2\leqslant 20, \\
& x_1\geqslant 0, x_2\geqslant 0.
\end{aligned}
$$

**解:**

(1) $f[(0,0)]=0, f[(0,1)]=2, f[(2,1)]=8, f[(3,0)]=9$.

(2) $f[(0,0)]=0, f[(0,10)]=-30, f[(5,0)]=-20$.

3. 将下列线性规划问题转化成标准形.

(1) $$\begin{cases} \max & 2x_1 - x_2 - 3x_3 \\ \text{s.t.} & 2x_1 - x_2 + 2x_3 \leqslant 2, \\ & -x_1 + 2x_2 - 3x_3 \leqslant -2, \\ & x_1 \geqslant 0, x_2 \geqslant 0, x_3 \geqslant 0. \end{cases}$$

(2) $$\begin{cases} \min & 2x_1 - x_2 - 3x_3 - x_4 \\ \text{s.t.} & x_1 + 3x_2 + 2x_3 + x_4 \leqslant 15, \\ & 2x_2 + x_3 + x_4 = 10, \\ & -x_1 + 5x_2 - 2x_3 + 2x_4 \geqslant 3, \\ & x_1 \geqslant 0, x_2 \geqslant 0, x_3 \geqslant 0, x_4 \geqslant 0. \end{cases}$$

(3) $$\begin{cases} \max & 3x_1 - 13x_2 \\ \text{s.t.} & 5x_1 + 3x_2 \leqslant 18, \\ & x_1 + 2x_2 \geqslant 9, \\ & 2x_1 + 5x_2 \geqslant 3, \\ & 2x_1 + x_2 \geqslant 2.5, \\ & x_1 \geqslant 0. \end{cases}$$

解:

(1) $$\begin{cases} \min & 2y_1 - 2y_2 \\ \text{s.t.} & 2y_1 - y_2 \geqslant 2, \\ & -y_1 + 2y_2 \geqslant -1, \\ & 2y_1 - 3y_2 \geqslant -3, \\ & y_1 \leqslant 0, y_2 \leqslant 0. \end{cases}$$

(2) $$\begin{cases} \max & 15y_1 + 10y_2 + 3y_3 \\ \text{s.t.} & y_1 - y_3 \leqslant 2, \\ & 3y_1 + 2y_2 + 5y_3 \leqslant -1, \\ & 2y_1 + y_2 - 2y_3 \leqslant -3, \\ & y_1 + y_2 + 2y_3 \leqslant -1, \\ & y_1 \leqslant 0, y_3 \geqslant 0. \end{cases}$$

(3) $$\begin{cases} \min & 18y_1 + 9y_2 + 3y_3 + 2.5y_4 \\ \text{s.t.} & 5y_1 + y_2 + 2y_3 + 2y_4 \geqslant 3, \\ & 3y_1 + 2y_2 + 5y_3 + y_4 = -13, \\ & y_1 \geqslant 0, y_2 \leqslant 0, y_3 \leqslant 0, y_4 \leqslant 0. \end{cases}$$

4. 用单纯形表的方法求下列线性规划问题的最优解.

(1) $\begin{cases} \max & 5x_1+6x_2+4x_3 \\ \text{s.t.} & 2x_1+2x_2 \leqslant 5, \\ & 5x_1+3x_2+4x_3 \leqslant 15, \\ & x_1+x_2 \leqslant 10, \\ & x_1 \geqslant 0, x_2 \geqslant 0, x_3 \geqslant 0. \end{cases}$

(2) $\begin{cases} \min & 3x_1-2x_2+x_3 \\ \text{s.t.} & 2x_1+2x_2-x_3 \leqslant 2, \\ & -3x_1-x_2+2x_3 \leqslant -2, \\ & x_1 \geqslant 0, x_2 \geqslant 0, x_3 \geqslant 0. \end{cases}$

(3) $\begin{cases} \max & 4x_1+2x_2+10x_3+x_4+2x_5 \\ \text{s.t.} & 5x_1+3x_2+2x_3+x_4-x_5 \leqslant 32, \\ & x_1+2x_2+3x_3+4x_4+4x_5 \leqslant 10, \\ & x_i \geqslant 0, i=1,2,3,4,5. \end{cases}$

**解**:

(1) $\max = \dfrac{89}{2}, x_1 = 0, x_2 = \dfrac{5}{2}, x_3 = \dfrac{15}{8}$.

(2) $\min = \dfrac{1}{2}, x_1 = \dfrac{1}{2}, x_2 = \dfrac{1}{2}, x_3 = 0$.

(3) $\max = \dfrac{484}{13}, x_1 = \dfrac{76}{13}, x_2 = 0, x_3 = \dfrac{18}{13}, x_4 = 0, x_5 = 0$.

5. 叙述单纯形法的最优解形式和无界解形式.

**解**: 见教材 65-66 页的式 (2.2.14) 和 (2.2.15).

6. 考察下述单纯形表

| 基变量 | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | 右端项 |
|---|---|---|---|---|---|---|
| $-f$ | 0 | $b$ | $e$ | 0 | 0 | $-9$ |
| $x_1$ | 1 | $c$ | 1 | 0 | 0 | 3 |
| $x_4$ | 0 | $d$ | $-1$ | 1 | 0 | 2 |
| $x_5$ | 0 | $-1$ | 1 | 0 | 1 | 4 |

问：参数 $b, c, d, e$ 满足什么条件才能使得

(1) 单纯形表是最优解形式?

(2) 单纯形表是无界解形式?

**解**:

(1) $b \geqslant 0, e \geqslant 0, c \in \mathbf{R}, d \in \mathbf{R}$.

(2) $b<0, c\leqslant 0, d\leqslant 0, e\in \mathbf{R}$.

7. 用修正单纯形方法求下列问题的解.

(1) $$\begin{cases} \min & -3x_1+2x_2-x_3-x_4 \\ \text{s.t.} & 2x_1+x_2+x_3+3x_4\leqslant 20, \\ & x_1+x_3+2x_4=10, \\ & -2x_1-x_2+2x_3+5x_4\geqslant 3, \\ & x_1\geqslant 0, x_2\geqslant 0, x_3\geqslant 0, x_4\geqslant 0. \end{cases}$$

(2) $$\begin{cases} \max & 2x_1-x_2-3x_3 \\ \text{s.t.} & 2x_1-x_2+2x_3\leqslant 2, \\ & -x_1+2x_2-3x_3\leqslant -2, \\ & x_1\geqslant 0, x_2\geqslant 0, x_3\geqslant 0. \end{cases}$$

**解**:

(1) $\min=-\dfrac{37}{2}, x_1=\dfrac{17}{4}, x_2=0, x_3=\dfrac{23}{4}, x_4=0$.

(2) $\max=-8, x_1=0, x_2=2, x_3=2$.

8. 考察下列线性规划问题

$$\begin{cases} \min & -\frac{3}{4}x_1+150x_2-\frac{1}{50}x_3+6x_4 \\ \text{s.t.} & \frac{1}{4}x_1-60x_2-\frac{1}{25}x_3=9x_4\leqslant 0, \\ & \frac{1}{2}x_1-90x_2-\frac{1}{50}x_3+3x_4\leqslant 0, \\ & x_3\leqslant 1, \\ & x_1\geqslant 0, x_2\geqslant 0, x_3\geqslant 0, x_4\geqslant 0. \end{cases}$$

(1) 引入松弛变量将其化成标准形后用单纯形方法进行迭代看是否出现循环的现象 (在确定出基变量时如有多于一个指标使 $b_j^{(k)}/\hat{a}_{jp}^{(k)}$ 取得最小值，取其中最先出现的变量作为出基变量).

(2) 把 Bland 方法结合进单纯形方法进行单纯形迭代看能否求得问题的最优解.

**略**.

9. 给出下列线性规划问题的对偶

(1) $$\begin{cases} \max & x_1 \\ \text{s.t.} & -4x_1+3x_3\geqslant 0, \\ & 2x_2+5x_3\leqslant 0, \\ & 6x_1+7x_2=0, \\ & x_2\geqslant 0, x_3\geqslant 0. \end{cases}$$

(2) $\begin{cases} \min & 2x_1-5x_2+3x_3-6x_4 \\ \text{s.t.} & -8x_1+7x_2+3x_3+4x_4=11, \\ & 7x_1+6x_2+2x_3+3x_4 \geqslant 23, \\ & 3x_1+2x_2+4x_3+7x_4 \leqslant 12, \\ & x_1 \leqslant 0, x_3 \geqslant 0, x_4 \geqslant 0. \end{cases}$

(3) $\begin{cases} \min & c^{\mathrm{T}}x \\ \text{s.t.} & Ax=b, \\ & l \leqslant x \leqslant u, \end{cases}$

其中 $l$ 和 $u$ 是 $x$ 的上下界向量.

(4) $\begin{cases} \min & c^{\mathrm{T}}x \\ \text{s.t.} & b_1 \leqslant Ax \leqslant b_2, \\ & x \geqslant 0, \end{cases}$

其中 $b_1$ 和 $b_2$ 是 $Ax \in \mathbf{R}^m$ 的上下界向量.

**解**:

(1) $\begin{cases} \min & 0 \\ \text{s.t.} & -4y_1+6y_3=1, \\ & 2y_2+7y_3 \geqslant 0, \\ & 3y_1+5y_2 \geqslant 0, \\ & y_1 \leqslant 0, y_2 \geqslant 0, y_3\text{无限制}. \end{cases}$

(2) $\begin{cases} \max & 11y_1+23y_2+12y_3 \\ \text{s.t.} & -8y_1+7y_2+3y_3 \geqslant 2, \\ & 7y_1+6y_2+2y_3=-5, \\ & 3y_1+2y_2+4y_3 \leqslant 3, \\ & 4y_1+3y_2+7y_3 \leqslant -6, \\ & y_1\text{无限制}, y_2 \geqslant 0, y_3 \leqslant 0. \end{cases}$

(3) 设 $c=(c_1,c_2,\cdots,c_n)^{\mathrm{T}}, A=(a_1,a_2,\cdots,a_n)$, 其中 $a_i$ 是 $m\times 1$ 的向量, 则对偶问题为 ($y$ 是 $(m+2)\times 1$

的向量)

$$\begin{cases} \max & (b^{\mathrm{T}}, 0, 0)y \\ \text{s.t.} & (a_1^{\mathrm{T}}, 0, 0)y = c_1, \\ & (a_2^{\mathrm{T}}, 0, 0)y = c_2, \\ & \vdots \\ & (a_{n-2}^{\mathrm{T}}, 0, 0)y = c_{n-2}, \\ & (a_{n-1}^{\mathrm{T}}, 1, 0)y = c_{n-1}, \\ & (a_n^{\mathrm{T}}, 0, 1)y = c_n, \\ & y_1, y_2, \cdots, y_m \text{无限制}, y_{m+1} \geqslant 0, y_{m+2} \leqslant 0. \end{cases}$$

(4) 设 $c = (c_1, c_2, \cdots, c_n)^{\mathrm{T}}, A = (a_1, a_2, \cdots, a_n)$，其中 $a_i$ 是 $m \times 1$ 的向量，则对偶问题为 ($y$ 是 $2m \times 1$ 的向量)

$$\begin{cases} \max & (b_1^{\mathrm{T}}, b_2^{\mathrm{T}})y \\ \text{s.t.} & (a_1^{\mathrm{T}}, a_1^{\mathrm{T}})y \leqslant c_1, \\ & (a_2^{\mathrm{T}}, a_2^{\mathrm{T}})y \leqslant c_2, \\ & \vdots \\ & (a_n^{\mathrm{T}}, a_n^{\mathrm{T}})y \leqslant c_m, \\ & y_1, y_2, \cdots, y_m \geqslant 0, y_{m+1}, y_{m+2} \cdots, y_{2m} \leqslant 0. \end{cases}$$

10. 求下列线性规划问题的对偶

$$\begin{aligned} \min \quad & c^{\mathrm{T}}x \\ \text{s.t.} \quad & Ax = b \\ & Bx \leqslant a, \\ & x \geqslant 0. \end{aligned}$$

**解**: 设 $c = (c_1, c_2, \cdots, c_n)^{\mathrm{T}}, A = (A_1, A_2, \cdots, A_n), B = (B_1, B_2, \cdots, B_n), b = (b_1, b_2, \cdots, b_m)^{\mathrm{T}}, a = (a_1, a_2, \cdots, a_m)^{\mathrm{T}}$，其中 $A_i$ 和 $B_i$ 都是 $m \times 1$ 的向量，则对偶问题为 ($y$ 是 $2m \times 1$ 的向量)

$$\begin{cases} \max & (b^{\mathrm{T}}, a^{\mathrm{T}})y \\ \text{s.t.} & (A_1^{\mathrm{T}}, B_1^{\mathrm{T}})y \leqslant c_1, \\ & (A_2^{\mathrm{T}}, B_2^{\mathrm{T}})y \leqslant c_2, \\ & \vdots \\ & (A_n^{\mathrm{T}}, B_n^{\mathrm{T}})y \leqslant c_m, \\ & y_1, y_2, \cdots, y_m \text{无限制}, y_{m+1}, y_{m+2} \cdots, y_{2m} \leqslant 0. \end{cases}$$

11. 用对偶单纯形方法求下列线性规划问题的解.

(1) $\begin{cases} \min & 9x_1+5x_2+3x_3 \\ \text{s.t.} & 3x_1+2x_2-3x_3 \geqslant 3, \\ & 2x_1+x_3 \geqslant 5, \\ & x_1 \geqslant 0, x_2 \geqslant 0, x_3 \geqslant 0. \end{cases}$

(2) $\begin{cases} \min & 3x_1+2x_2+4x_3 \\ \text{s.t.} & 2x_1-x_2 \geqslant 5, \\ & 2x_2-x_3 \geqslant 10, \\ & x_1 \geqslant 0, x_2 \geqslant 0, x_3 \geqslant 0. \end{cases}$

(3) $\begin{cases} \max & -6x_1-5x_2-7x_3-2x_4 \\ \text{s.t.} & -3x_1-2x_2+5x_3+4x_4 \geqslant 15, \\ & 6x_1-2x_2+2x_3+7x_4 \leqslant 35, \\ & -5x_1-4x_2-3x_3+2x_4 \geqslant 20, \\ & x_1 \geqslant 0, x_2 \geqslant 0, x_3 \geqslant 0, x_4 \geqslant 0. \end{cases}$

**解**:

(1) $\min = 21, x_1 = 2, x_2 = 0, x_3 = 1.$

(2) $\min = 25, x_1 = 5, x_2 = 5, x_3 = 0.$

(3) 无可行解.

12. 将 Klee-Minty 问题转换成标准形，并对 $m=3$ 的情况用单纯形方法求解 (要求取最负的价值系数的变量为入基变量). 如果在第一次迭代取第三个约束的松弛变量为入基变量，$x_3$ 为入基变量，结果又如何?
**略**.

13. (1) 设 $\mu_{k+1} = \left(1-\dfrac{\gamma}{\sqrt{n}}\right)\mu_k, 0<\gamma<1$. 给定 $\mu_0>0, \varepsilon>0$ 确定使 $\mu_k \leqslant \varepsilon$ 成立所需要的迭代次数.
(2) 估计原始对偶内点算法的每次迭代的运算工作量.

**略**.

14. 用原始对偶内点算法求下述线性规划问题的解

$$\begin{cases} \max & x_1+3x_2 \\ \text{s.t.} & x_1+4x_2 \leqslant 10, \\ & x_1+2x_2 \leqslant 1, \\ & x_1 \geqslant 0, x_2 \geqslant 0. \end{cases}$$

**解**: $\max = \dfrac{3}{2}, x_1 = 0, x_2 = \dfrac{1}{2}.$

# 第三章　线性搜索与信赖域方法

1. 推导 0.618 法的迭代公式.

**解**: 由 Fibonacci 法迭代公式：

$$\lambda_k = a_k + \left(1 - \frac{F_{n-k}}{F_{n-k+1}}\right)(b_k - a_k), \quad \mu_k = a_k + \frac{F_{n-k}}{F_{n-k+1}}(b_k - a_k),$$

由式 (3.2.20) 可知：$\tau = \lim\limits_{n\to\infty} \dfrac{F_{n-k}}{F_{n-k+1}} = \dfrac{\sqrt{5}-1}{2}$，取 $\tau \approx 0.618$，则可得 0.618 法的迭代公式：

$$\lambda_k = a_k + 0.382(b_k - a_k), \quad \mu_k = a_k + 0.618(b_k - a_k).$$

2. 分别用 0.618 法，二次插值法，Goldstein 不精确线性搜索方法求函数

$$f(x) = (\sin x)^6 \tan(1-x) \mathrm{e}^{30x}$$

在区间 $[0,1]$ 上的极大值.

**解**: 由于这三种方法都是求解函数的极小值，故我们不妨设

$$g(x) = -(\sin x)^6 \tan(1-x) \mathrm{e}^{30x}.$$

0.618 法：(取精度为 0.0001)

| 迭代次数 $k$ | $[a_k, b_k]$ | $[\lambda_k, \mu_k]$ | $g(\lambda_k)$ | $g(\mu_k)$ |
|---|---|---|---|---|
| 0 | $[0.00000, 1.00000]$ | $[0.38200, 0.61800]$ | $-180.93169$ | $-1712603.84319$ |
| 1 | $[0.38200, 1.00000]$ | $[0.61800, 0.76392]$ | $-1712603.84319$ | $-236591928.32087$ |
| 2 | $[0.61800, 1.00000]$ | $[0.76392, 0.85408]$ | $-236591928.32087$ | $-3621963454.74789$ |
| 3 | $[0.76392, 1.00000]$ | $[0.85408, 0.90982]$ | $-3621963454.74789$ | $-15629095526.75970$ |
| 4 | $[0.85408, 1.00000]$ | $[0.90982, 0.94426]$ | $-15629095526.75970$ | $-31645853714.38405$ |
| 5 | $[0.90982, 1.00000]$ | $[0.94426, 0.96555]$ | $-31645853714.38405$ | $-40525854255.03307$ |
| 6 | $[0.94426, 1.00000]$ | $[0.96555, 0.97871]$ | $-40525854255.03307$ | $-39217975674.21036$ |
| 7 | $[0.94426, 0.97871]$ | $[0.95742, 0.96555]$ | $-37941310649.54459$ | $-40525854255.03307$ |
| 8 | $[0.95742, 0.97871]$ | $[0.96555, 0.97057]$ | $-40525854255.03307$ | $-41085789962.33502$ |
| 9 | $[0.96555, 0.97871]$ | $[0.97057, 0.97368]$ | $-41085789962.33502$ | $-40851559012.30843$ |
| 10 | $[0.96555, 0.97368]$ | $[0.96866, 0.97057]$ | $-40993501131.26535$ | $-41085789962.33502$ |
| 11 | $[0.96866, 0.97368]$ | $[0.97057, 0.97176]$ | $-41085789962.33502$ | $-41056235078.63431$ |
| 12 | $[0.96866, 0.97176]$ | $[0.96984, 0.97057]$ | $-41070108314.89887$ | $-41085789962.33502$ |
| 13 | $[0.96984, 0.97176]$ | $[0.97057, 0.97103]$ | $-41085789962.33502$ | $-41082739735.01510$ |
| 14 | $[0.96984, 0.97103]$ | $[0.97030, 0.97057]$ | $-41082767355.81810$ | $-41085789962.33502$ |
| 15 | $[0.97030, 0.97103]$ | $[0.97057, 0.97075]$ | $-41085789962.33502$ | $-41085803940.89211$ |
| 16 | $[0.97057, 0.97103]$ | $[0.97075, 0.97085]$ | $-41085803940.89211$ | $-41085091885.42854$ |
| 17 | $[0.97057, 0.97085]$ | $[0.97068, 0.97075]$ | $-41085973087.67970$ | $-41085803940.89211$ |

故 $f(x)$ 的极大点为 0.97071，极大值为 $-4.1086\times 10^{10}$.
二次插值法：(取精度为 0.0001，初始迭代点为 0.1、0.5、1)

| $k$ | $a_1$ | $a_2$ | $a_3$ | $\bar{a}$ | $g_1$ | $g_2$ | $g_3$ | $\bar{g}$ |
|---|---|---|---|---|---|---|---|---|
| 0 | 0.10000 | 0.50000 | 1 | 0.55000 | $-0.00003$ | $-21685.89733$ | 0 | $-144313.48798$ |
| 1 | 0.50000 | 0.55000 | 1 | 0.74609 | $-21685.89733$ | $-144313.48798$ | 0 | $-133409650.84449$ |
| 2 | 0.55000 | 0.74609 | 1 | 0.77494 | $-144313.48798$ | $-133409650.84449$ | 0 | $-335472520.81192$ |
| 3 | 0.74609 | 0.77494 | 1 | 0.86519 | $-133409650.84449$ | $-335472520.81192$ | 0 | $-4941583365.98999$ |
| 4 | 0.77494 | 0.86519 | 1 | 0.88556 | $-335472520.81192$ | $-4941583365.98999$ | 0 | $-8543117362.08340$ |
| 5 | 0.86519 | 0.88556 | 1 | 0.92277 | $-4941583365.98999$ | $-8543117362.08340$ | 0 | $-20938281955.02223$ |
| 6 | 0.88556 | 0.92277 | 1 | 0.93571 | $-8543117362.08340$ | $-20938281955.02223$ | 0 | $-27213528026.34990$ |
| 7 | 0.92277 | 0.93571 | 1 | 0.94986 | $-20938281955.02223$ | $-27213528026.34990$ | 0 | $-34493283849.16349$ |
| 8 | 0.93571 | 0.94986 | 1 | 0.95654 | $-27213528026.34990$ | $-34493283849.16349$ | 0 | $-37577724676.75349$ |
| 9 | 0.94986 | 0.95654 | 1 | 0.96193 | $-34493283849.16349$ | $-37577724676.75349$ | 0 | $-39577028005.21829$ |
| 10 | 0.95654 | 0.96193 | 1 | 0.96495 | $-37577724676.75349$ | $-39577028005.21829$ | 0 | $-40395052418.74240$ |
| 11 | 0.96193 | 0.96495 | 1 | 0.96706 | $-39577028005.21829$ | $-40395052418.74240$ | 0 | $-40798680632.87956$ |
| 12 | 0.96495 | 0.96706 | 1 | 0.96834 | $-40395052418.74240$ | $-40798680632.87956$ | 0 | $-40963298550.04148$ |
| 13 | 0.96706 | 0.96834 | 1 | 0.96919 | $-40798680632.87956$ | $-40963298550.04148$ | 0 | $-41035621691.04285$ |
| 14 | 0.96834 | 0.96919 | 1 | 0.96972 | $-40963298550.04148$ | $-41035621691.04285$ | 0 | $-41065099592.32568$ |
| 15 | 0.96919 | 0.96972 | 1 | 0.97006 | $-41035621691.04285$ | $-41065099592.32568$ | 0 | $-41077456061.09961$ |
| 16 | 0.96972 | 0.97006 | 1 | 0.97028 | $-41065099592.32568$ | $-41077456061.09961$ | 0 | $-41082487976.77145$ |
| 17 | 0.97006 | 0.97028 | 1 | 0.97042 | $-41077456061.09961$ | $-41082487976.77145$ | 0 | $-41084558359.15994$ |
| 18 | 0.97028 | 0.97042 | 1 | 0.97051 | $-41082487976.77145$ | $-41084558359.15994$ | 0 | $-41085400853.61729$ |

故 $f(x)$ 的极大点为 0.97051，极大值为 $-4.1085\times 10^{10}$.
Goldstein 不精确线性搜索方法略.

3. 写出 Fibonacci 法的计算过程和 C 程序 (或 MATLAB,FORTRAN 程序).

**解**: Fibonacci 法的计算过程：

步 1: 给定初始搜索区间 $[a_1,b_1]$ 和精度要求 $\delta>0$，计算满足 $F_n\geqslant \dfrac{b_1-a_1}{\delta}$ 的最小的 $n$；

步 2: $\lambda_1:=a_1+\dfrac{F_{n-2}}{F_n}(b_1-a_1)$，$\mu_1:=a_1+\dfrac{F_{n-1}}{F_n}(b_1-a_1)$
计算 $\varphi(\lambda_1),\varphi(\mu_1)$，置 $k:=1$；

步 3: 若 $k=n-1$，则转步 4，否则转步 5；

步 4: 停止计算，输出 $\lambda_k$(或 $\mu_k$，因为 $\lambda_k=\mu_k$ 此时恰好成立). 也可再计算一个函数值 $\varphi\left(\dfrac{a_k+b_k}{2}\pm\varepsilon\right)$，比较后输出 $\lambda_k$ 或 $\dfrac{a_k+b_k}{2}\pm\varepsilon$；

步 5: 若 $\varphi(\lambda_k)>\varphi(\mu_k)$，转步 6，否则转步 7；

步 6: 令 $a_{k+1}:=\lambda_k,b_{k+1}:=b_k,\lambda_{k+1}:=\mu_k,\varphi(\lambda_{k+1}):=\varphi(\mu_k),\mu_{k+1}:=a_{k+1}+\dfrac{F_{n-k-1}}{F_{n-k}}(b_{k+1}-a_{k+1})$，计算 $\varphi(\mu_{k+1})$，令 $k:=k+1$，转步 3；

步 7: 令 $a_{k+1} := a_k, b_{k+1} := \mu_k, \mu_{k+1} := \lambda_k, \varphi(\mu_{k+1}) := \varphi(\lambda_k), \lambda_{k+1} := a_{k+1} + \frac{F_{n-k-2}}{F_{n-k}}(b_{k+1} - a_{k+1})$，计算 $\varphi(\lambda_{k+1})$，令 $k := k+1$，转步 3.

注：在步 6、步 7 中，当 $k = n-2$，无需计算 $\varphi(\mu_{k+1})$，$\varphi(\lambda_{k+1})$. 因为此时 $\mu_{k+1} = \lambda_{k+1}$，但由于舍入误差，它们未必精确相等.

Matlab 程序为：

```
function [xmin, fval, k] = Fibonacci(f, a, b, delta)
    % f是所求函数表达式，a是区间左端点
    % b是区间右端点，delta是精度
    % xmin是极小点，fval是极小值，k是迭代步数
    format long
    Fn = (b - a) / delta; F = []; F(1) = 1; F(2) = 1; i = 1;
    while F(i) < Fn
        F(i + 2) = F(i) + F(i + 1); i = i + 1;
    end
    n = i;
    lambda = a + F(n - 2) / F(n) * (b - a); mu = a + F(n - 1) / F(n) * (b - a);
    x = symvar(f);
    phi_lambda = subs(f, x, lambda); phi_mu = subs(f, x, mu);
    k = 0;
    fprintf('k \t [a_k, b_k] \t [lambda_k, mu_k] \t phi(lambda_k) \t phi(mu_k)\n');
    fprintf('%d \t [%.5f,%.5f] \t [%.5f,%.5f] \t %.5f \t %.5f\n', k, a, b,
        lambda, mu, phi_lambda, phi_mu);
    k = 1;
    while true
        if k == n - 1
            xmin = vpa(lambda, 5); fval = vpa(subs(f, x, lambda), 5);
            break;
        else
            if phi_lambda > phi_mu
                a = lambda; b = b; lambda = mu; phi_lambda = phi_mu;
                if k == n - 2
                    xmin = vpa(lambda, 5);
                    fval = vpa(subs(f, x, lambda), 5);
                    break;
                else
                    mu = a + F(n - k - 1) / F(n - k) * (b - a);
                    phi_mu = subs(f, x, mu); k = k + 1;
```

```
                    fprintf('%d \t [%.5f,%.5f] \t [%.5f,%.5f] \t %.5f \t %.5f\n',
                        k, a, b, lambda, mu, phi_lambda, phi_mu);
                end
            else
                a = a; b = mu; mu = lambda; phi_mu = phi_lambda;
                if k == n - 2
                    xmin = vpa(lambda, 5);
                    fval = vpa(subs(f, x, lambda), 5);
                    break;
                else
                    lambda = a + F(n - k - 2) / F(n - k) * (b - a);
                    phi_lambda = subs(f, x, lambda); k = k + 1;
                    fprintf('%d \t [%.5f,%.5f] \t [%.5f,%.5f] \t %.5f \t %.5f\n',
                        k, a, b, lambda, mu, phi_lambda, phi_mu);
                end
            end
        end
    end
end
```

4. 设 $\varphi(t)=\mathrm{e}^{-t}+\mathrm{e}^{t}$，区间为 $[-1,1]$.

(1) 用 0.618 法极小化 $\varphi(t)$.

(2) 用 Fibonacci 法极小化 $\varphi(t)$.

**解:**

(1) 取精度为 0.01：

| 迭代次数 $k$ | $[a_k,b_k]$ | $[\lambda_k,\mu_k]$ | $\varphi(\lambda_k)$ | $\varphi(\mu_k)$ |
|---|---|---|---|---|
| 0 | $[-1.00000,1.00000]$ | $[-0.23600,0.23600]$ | 2.05595 | 2.05595 |
| 1 | $[-1.00000,0.23600]$ | $[-0.52785,-0.23600]$ | 2.28515 | 2.05595 |
| 2 | $[-0.52785,0.23600]$ | $[-0.23600,-0.05579]$ | 2.05595 | 2.00311 |
| 3 | $[-0.23600,0.23600]$ | $[-0.05579,0.05570]$ | 2.00311 | 2.00310 |
| 4 | $[-0.05579,0.23600]$ | $[0.05570,0.12454]$ | 2.00310 | 2.01553 |
| 5 | $[-0.05579,0.12454]$ | $[0.01309,0.05570]$ | 2.00017 | 2.00310 |
| 6 | $[-0.05579,0.05570]$ | $[-0.01320,0.01309]$ | 2.00017 | 2.00017 |
| 7 | $[-0.01320,0.05570]$ | $[0.01309,0.02938]$ | 2.00017 | 2.00086 |
| 8 | $[-0.01320,0.02938]$ | $[0.00306,0.01309]$ | 2.00001 | 2.00017 |
| 9 | $[-0.01320,0.01309]$ | $[-0.00316,0.00306]$ | 2.00001 | 2.00001 |

故 $\varphi(x)$ 的极小点为 $-0.000046968$，极小值为 2.

(2) 取精度为 0.01：

| 迭代次数 $k$ | $[a_k, b_k]$ | $[\lambda_k, \mu_k]$ | $\varphi(\lambda_k)$ | $\varphi(\mu_k)$ |
|---|---|---|---|---|
| 0 | $[-1.00000, 1.00000]$ | $[-0.23605, 0.23605]$ | 2.05598 | 2.05598 |
| 2 | $[-1.00000, 0.23605]$ | $[-0.52790, -0.23605]$ | 2.28521 | 2.05598 |
| 3 | $[-0.52790, 0.23605]$ | $[-0.23605, -0.05579]$ | 2.05598 | 2.00311 |
| 4 | $[-0.23605, 0.23605]$ | $[-0.05579, 0.05579]$ | 2.00311 | 2.00311 |
| 5 | $[-0.23605, 0.05579]$ | $[-0.12446, -0.05579]$ | 2.01551 | 2.00311 |
| 6 | $[-0.12446, 0.05579]$ | $[-0.05579, -0.01288]$ | 2.00311 | 2.00017 |
| 7 | $[-0.05579, 0.05579]$ | $[-0.01288, 0.01288]$ | 2.00017 | 2.00017 |
| 8 | $[-0.05579, 0.01288]$ | $[-0.03004, -0.01288]$ | 2.00090 | 2.00017 |
| 9 | $[-0.03004, 0.01288]$ | $[-0.01288, -0.00429]$ | 2.00017 | 2.00002 |
| 10 | $[-0.01288, 0.01288]$ | $[-0.00429, 0.00429]$ | 2.00002 | 2.00002 |
| 11 | $[-0.01288, 0.00429]$ | $[-0.00429, -0.00429]$ | 2.00002 | 2.00002 |

故 $\varphi(x)$ 的极小点为 $-0.0042918$，极小值为 2.

5. 用二分法解

$$\begin{aligned} \min \quad & x^2 + 2x \\ \text{s.t.} \quad & -3 \leqslant x \leqslant 6. \end{aligned}$$

取最后区间长度为 $\delta = 0.2$. (解 $x^* = -0.961$).

**解**:

| $k$ | $[a_k, b_k]$ |
|---|---|
| 1 | $[-3.00000, 6.00000]$ |
| 2 | $[-3.00000, 1.50000]$ |
| 3 | $[-3.00000, -0.75000]$ |
| 4 | $[-1.87500, -0.75000]$ |
| 5 | $[-1.31250, -0.75000]$ |
| 6 | $[-1.03125, -0.75000]$ |
| 7 | $[-1.03125, -0.89063]$ |

故 $x^2 + 2x$ 在 $[-3, 6]$ 上的极小点为 $-0.96094$，极小值为 $-0.99847$.

6. 设 $\varphi(t) = 1 - t\mathrm{e}^{-t^2}$，区间为 $[0, 1]$. 试用三点二次插值法极小化 $\varphi(t)$.

**解**: 取精度为 0.0001，初始迭代点为 0、0.5、1

| $k$ | $a_1$ | $a_2$ | $a_3$ | $\bar{a}$ | $\varphi_1$ | $\varphi_2$ | $\varphi_3$ | $\bar{\varphi}$ |
|---|---|---|---|---|---|---|---|---|
| 0 | 0.00000 | 0.50000 | 1.00000 | 0.72381 | 1.00000 | 0.61060 | 0.63212 | 0.57136 |
| 1 | 0.50000 | 0.72381 | 1.00000 | 0.72278 | 0.61060 | 0.57136 | 0.63212 | 0.57133 |
| 2 | 0.50000 | 0.72278 | 0.72381 | 0.70822 | 0.61060 | 0.57133 | 0.57136 | 0.57112 |
| 3 | 0.50000 | 0.70822 | 0.72278 | 0.70769 | 0.61060 | 0.57112 | 0.57133 | 0.57112 |
| 4 | 0.50000 | 0.70769 | 0.70822 | 0.70716 | 0.61060 | 0.57112 | 0.57112 | 0.57112 |
| 5 | 0.50000 | 0.70716 | 0.70769 | 0.70713 | 0.61060 | 0.57112 | 0.57112 | 0.57112 |

故 $\varphi(t)$ 的极小点为 0.70713，极小值为 0.57112.

7. 设 $\varphi(t) = -2t^3 + 21t^2 - 60t + 50$.

   (1) 用 Goldstein 方法极小化 $\varphi(t), t_0 = 0.5, \rho = 0.1$.

   (2) 用 Wolfe 方法极小化 $\varphi(t), t_0 = 0.5, \rho = 0.1, \sigma = 0.8$.

**略**.

8. 设 $f(x) = x_1^4 + x_2^2 + x_2^2$，给定当前点 $x_k = (1,1)^{\mathrm{T}}$，方向 $d_k = (-3,-1)^{\mathrm{T}}$，并设 $\rho = 1, t = 0.5$.

   (1) 试分别用 Goldstein 方法和 Wolfe 方法求新点 $x_{k+1}$.

   (2) 分别以 $\alpha = 1, \alpha = 0.5, \alpha = 0.1$ 说明哪些 $\alpha$ 满足 Wolfe 准则，哪些 $\alpha$ 不满足 Wolfe 准则.

**略**.

9. 设 $f(x) = \dfrac{1}{2}x_1^2 + x_2^2$. 设初始点 $x_0 = (1,1)^{\mathrm{T}}$. 对于 $(1)\Delta_0 = 1$, $(2)\Delta_0 = \dfrac{5}{4}$,

   (i) 用折线法计算下一个迭代点 $x_1$;

   (ii) 用双折线法计算下一个迭代点 $x_1$.

**略**.

10. 证明：

   (1) Cauchy 点满足 (3.8.4).

   (2) 子问题 (3.7.2) 的精确极小点满足 (3.8.4).

**略**.

# 第四章　无约束最优化方法

1. 设 $f(x)=\dfrac{3}{2}x_1^2+\dfrac{1}{2}x_2^2-x_1x_2-2x_1$，设初始点 $x^{(0)}=(-2,4)^{\mathrm{T}}$. 试用最速下降法和牛顿法极小化 $f(x)$.
**解**: 最速下降法：$\nabla f(x)=(3x_1-x_2-2,x_2-x_1)^{\mathrm{T}}$，则

| $k$ | $\nabla f(x^{(k)})$ | $d^{(k)}$ | $\alpha_k$ | $x^{(k+1)}$ |
|---|---|---|---|---|
| 0 | $(-12,6)^{\mathrm{T}}$ | $(12,-6)^{\mathrm{T}}$ | $\dfrac{5}{17}$ | $\left(\dfrac{26}{17},\dfrac{38}{17}\right)^{\mathrm{T}}$ |
| 1 | $\left(\dfrac{6}{17},\dfrac{12}{17}\right)^{\mathrm{T}}$ | $\left(-\dfrac{6}{17},-\dfrac{12}{17}\right)^{\mathrm{T}}$ | $\dfrac{5}{3}$ | $\left(\dfrac{17-1}{17},\dfrac{17+1}{17}\right)^{\mathrm{T}}$ |
| 2 | $\left(-\dfrac{4}{17},\dfrac{2}{17}\right)^{\mathrm{T}}$ | $\left(\dfrac{4}{17},-\dfrac{2}{17}\right)^{\mathrm{T}}$ | $\dfrac{5}{17}$ | $\left(\dfrac{17^2+3}{17^2},\dfrac{17^2+7}{17^2}\right)^{\mathrm{T}}$ |
| 3 | $\left(\dfrac{2}{289},\dfrac{4}{289}\right)^{\mathrm{T}}$ | $\left(-\dfrac{2}{289},-\dfrac{4}{289}\right)^{\mathrm{T}}$ | $\cdots$ | $\left(\dfrac{3\times17^2-1}{3\times17^2},\dfrac{3\times17^2+1}{3\times17^2}\right)^{\mathrm{T}}$ |
| $\cdots$ | $\cdots$ | $\cdots$ | $\cdots$ | $\cdots$ |

逐步迭代，取迭代次数为无穷大时，极小值点 $x^*=(1,1)^{\mathrm{T}},\min f(x)=-1$. 最后，通过最优性条件验证该点是极小值点：

$$\nabla f(x^*)=(0,0)^{\mathrm{T}},\quad G=\begin{pmatrix}3&-1\\-1&1\end{pmatrix}$$

牛顿法：由最速下降法知：$G$ 是正定的，且 $G^{-1}=\begin{pmatrix}0.5&0.5\\0.5&1.5\end{pmatrix}$.
迭代：$x^{(1)}=x^{(0)}-G(x^{(0)})^{-1}\nabla f(x^{(0)})=(1,1)^{\mathrm{T}}$,
计算得：$\nabla f(x^{(1)})=(0,0)^{\mathrm{T}}$. 由最优性条件可知，迭代终止，得最优解：$x^*=x^{(1)}=(0,0)^{\mathrm{T}}$.
若均各迭代一步，牛顿法选用带步长因子的，则解答如下：
将 $f(x)$ 写成 $f(x)=\dfrac{1}{2}x^{\mathrm{T}}Gx-b^{\mathrm{T}}x$ 的形式，有

$$G=\begin{pmatrix}3&-1\\-1&1\end{pmatrix},b=\begin{pmatrix}2\\0\end{pmatrix}$$

最速下降法:设 $x^{(0)}=(-2,4)^{\mathrm{T}}$,则 $g_0=Gx^{(0)}-b=(-12,6)^{\mathrm{T}}$,$d_0=-g_0=(12,-6)^{\mathrm{T}}$,$\alpha_0=-\dfrac{g_0^{\mathrm{T}}d_0}{d_0^{\mathrm{T}}Gd_0}=\dfrac{5}{17}$,
$x^{(1)}=x^{(0)}+\alpha_0d_0=\left(\dfrac{26}{17},\dfrac{38}{17}\right)^{\mathrm{T}}$.
牛顿法: 设 $x^{(0)}=(-2,4)^{\mathrm{T}}$，则 $g_0=Gx^{(0)}-b=(-12,6)^{\mathrm{T}}$，$d_0=-G^{-1}g_0=(3,-3)^{\mathrm{T}}$，$\alpha_0=-\dfrac{g_0^{\mathrm{T}}d_0}{d_0^{\mathrm{T}}Gd_0}=1$,
$x^{(1)}=x^{(0)}+\alpha_0d_0=(1,1)^{\mathrm{T}}$，由于 $G$ 是正定二次函数，一步迭代达到最优解，$(1,1)^{\mathrm{T}}$ 就是极小点.

2. 设 (1) $f(x)=\dfrac{1}{2}(x_1^2+9x_2^2)$, (2) $f(x)=\dfrac{1}{2}(x_1^2+10^4x_2^2)$. 试讨论最速下降法的收敛速度.
**解**: 显然 (1), (2) 都是正定二次函数，故一定都是线性收敛速度，而

$$C_1=\frac{\sqrt{9}(9-1)}{9+1}=2.4,C_2=\frac{\sqrt{10^4}(10^4-1)}{10^4+1}=99.98>C_1,$$

所以 (1) 的收敛速度快于 (2).

3. 设 $f(x)=\frac{1}{2}x^{\mathrm{T}}x+\frac{1}{4}\sigma(x^{\mathrm{T}}Ax)$，其中

$$A=\begin{bmatrix} 5 & 1 & 0 & \frac{1}{2} \\ 1 & 4 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 3 & 0 \\ \frac{1}{2} & 0 & 0 & 2 \end{bmatrix}$$

设

(1) $x_0=(\cos 70°,\sin 70°,\cos 70°,\sin 70°)^{\mathrm{T}}$,

(2) $x_0=(\cos 50°,\sin 50°,\cos 50°,\sin 50°)^{\mathrm{T}}$,

试在 $\sigma=1$ 和 $\sigma=10^4$ 的情况下讨论基本牛顿法、线性搜索牛顿法和牛顿型信赖域方法的数值结果和收敛性态.

**略**.

4. 证明：当极小化正定二次函数时，共轭梯度法 FR 公式，PRP 公式和 Dixon 公式是等价的.

**证**: 先证 FR 与 PRP 公式等价，即 $g_k^{\mathrm{T}}(g_k-g_{k-1})=g_k^{\mathrm{T}}g_k$，即证 $g_k^{\mathrm{T}}g_{k-1}=0$.

$\because d_{k-1}=-g_{k-1}+\beta_{k-2}d_{k-2}, \therefore g_{k-1}=-d_{k-1}+\beta_{k-2}d_{k-2}$.

$$\begin{aligned} g_k^{\mathrm{T}}g_{k-1} &= g_k^{\mathrm{T}}(-d_{k-1}+\beta_{k-2}d_{k-2}) \\ &= -g_k^{\mathrm{T}}d_{k-1}+\beta_{k-2}g_k^{\mathrm{T}}d_{k-2} \end{aligned}$$

由精确线搜索：$g_k^{\mathrm{T}}d_{k-1}=g_k^{\mathrm{T}}d_{k-2}=0$，所以 $g_k^{\mathrm{T}}g_{k-1}=0$，FR 与 PRP 公式等价.

再证 FR 与 Dixon 公式等价，即 $g_{k-1}^{\mathrm{T}}g_{k-1}=-d_{k-1}^{\mathrm{T}}g_{k-1}$.

$\because d_{k-1}=-g_{k-1}+\beta_{k-2}d_{k-2}, \therefore -d_{k-1}^{\mathrm{T}}g_{k-1}=g_{k-1}^{\mathrm{T}}g_{k-1}-\beta_{k-2}d_{k-2}^{\mathrm{T}}g_{k-1}$.

由精确线搜索：$d_{k-2}^{\mathrm{T}}g_{k-1}=g_{k-1}^{\mathrm{T}}d_{k-2}=0$，所以 $g_{k-1}^{\mathrm{T}}g_{k-1}=-d_{k-1}^{\mathrm{T}}g_{k-1}$，FR 与 Dixon 公式等价.

所以，当极小化正定二次函数时，共轭梯度法 FR 公式，PRP 公式和 Dixon 公式是等价的.

5. 用 FR 共扼梯度法解极小化问题

$$\min f(x_1,x_2)=1+x_1-x_2+x_1^2+2x_2^2.$$

**解**: 初始点取 $(0,0)^{\mathrm{T}}$，化为标准形式 $f(x)=\frac{1}{2}x^{\mathrm{T}}Gx-b^{\mathrm{T}}x-c$，其中 $G=\begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}, b=\begin{pmatrix} -1 \\ 1 \end{pmatrix}, g(x)=\nabla f(x)=(1+2x_1,4x_2-1)^{\mathrm{T}}$，则

(1) $r_0=g_0=(1,-1)^{\mathrm{T}}, d_0=-r_0=(-1,1)^{\mathrm{T}}, \alpha_0=\frac{r_0^{\mathrm{T}}r_0}{d_0^{\mathrm{T}}Gd_0}=\frac{1}{3}, x^1=x^0+\alpha_0d_0=\left(-\frac{1}{3},\frac{1}{3}\right)^{\mathrm{T}} \Rightarrow r_1=g_1=\left(\frac{1}{3},\frac{1}{3}\right)^{\mathrm{T}}$.

(2) $\beta_0=\frac{r_1^{\mathrm{T}}r_1}{r_0^{\mathrm{T}}r_0}=\frac{1}{9} \Rightarrow d_1=-r_1+\beta_0d_0=\left(-\frac{4}{9},-\frac{2}{9}\right)^{\mathrm{T}}, \alpha_1=\frac{r_1^{\mathrm{T}}r_1}{d_1^{\mathrm{T}}Gd_1}=\frac{3}{8}, x^2=\left(-\frac{1}{2},\frac{1}{4}\right)^{\mathrm{T}} \Rightarrow r_2=g_2=(0,0)^{\mathrm{T}}$.

所以，$x^2 = \left(-\frac{1}{2}, \frac{1}{4}\right)^{\mathrm{T}}$ 是极小点，$\min f(x) = \frac{5}{8}$.

6. 设 $s \in \mathbf{R}^n, s \neq 0, I$ 是 $n \times n$ 单位矩阵. 证明：

$$\left\| I - \frac{ss^{\mathrm{T}}}{s^{\mathrm{T}}s} \right\| = 1.$$

**略**.

7. 证明逆的秩—校正公式 (Sherman-Morrison 公式)：设 $A \in \mathbf{R}^{n\times n}$ 非奇异，$u, v \in \mathbf{R}^n$ 是任意向量，若 $I + v^{\mathrm{T}}Au \neq 0$，则 $A + uv^{\mathrm{T}}$ 非奇异，且其逆矩阵可以表示为

$$(A + uv^{\mathrm{T}})^{-1} = A^{-1} - \frac{A^{-1}uv^{\mathrm{T}}A^{-1}}{1 + v^{\mathrm{T}}A^{-1}u}.$$

**证**:

$$\begin{aligned}
\left(A^{-1} - \frac{A^{-1}uv^{\mathrm{T}}A^{-1}}{1 + v^{\mathrm{T}}A^{-1}u}\right)(A + uv^{\mathrm{T}}) &= I + A^{-1}uv^{\mathrm{T}} - \frac{A^{-1}uv^{\mathrm{T}}}{1 + v^{\mathrm{T}}A^{-1}u} - \frac{A^{-1}uv^{\mathrm{T}}A^{-1}uv^{\mathrm{T}}}{1 + v^{\mathrm{T}}A^{-1}u} \\
&= I + A^{-1}uv^{\mathrm{T}} - \frac{A^{-1}uv^{\mathrm{T}}}{1 + v^{\mathrm{T}}A^{-1}u} - \frac{(v^{\mathrm{T}}A^{-1}u)A^{-1}uv^{\mathrm{T}}}{1 + v^{\mathrm{T}}A^{-1}u} \\
&= I + \frac{(1 + v^{\mathrm{T}}A^{-1}u)A^{-1}uv^{\mathrm{T}} - A^{-1}uv^{\mathrm{T}} - (v^{\mathrm{T}}A^{-1}u)Auv^{\mathrm{T}}}{1 + v^{\mathrm{T}}A^{-1}u} = I
\end{aligned}$$

证毕.

8. 证明 Wolfe 条件 (3.5.5) 和强 Wolfe 条件 (3.5.9) 意味着 $s_k^{\mathrm{T}}y_k > 0$.
**略**.

9. (1) 证明公式 (4.4.23) 和公式 (4.4.24) 是互逆的.
   (2) 证明公式 (4.4.26) 和公式 (4.4.27) 是互逆的.

**证**:

(1) 利用 $B_{k+1}s_k = y_k$ 和 $H_{k+1}y_k = s_k$，在公式 (4.4.23) 中互换 $s_k$ 和 $y_k$，并用 $B_{k+1}$ 和 $B_k$ 分别取代 $H_{k+1}$ 和 $H_k$，就能得到公式 (4.4.24).

(2) 利用 Sherman-Morrison 公式，以及上一问的思路，即可证明公式 (4.4.26) 和公式 (4.4.27) 是互逆的.

10. 分别用最速下降法、共轭梯度法、牛顿法和拟牛顿法极小化 Rosenbrock 函数 $f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2, x^{(0)} = (-1.2, 1)^{\mathrm{T}}, x^* = (1, 1)^{\mathrm{T}}, f(x^*) = 0$.
**解**: 可使用 Matlab 求解，并且四种方法均可以迭代收敛到最优解 $x^* = (1, 1)^{\mathrm{T}}$.

11. 分别用共轭梯度法和拟牛顿法极小化 Powell 奇异函数

$$f(x) = (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(x_1 - x_4)^4,$$

初始点 $x^{(0)} = (3, -1, 0, 1)^{\mathrm{T}}$，解为 $x^* = (0, 0, 0, 0)^{\mathrm{T}}, f(x^*) = 0$.
**解**: 可使用 Matlab 求解：
共轭梯度法：$x^* = (0.0468, -0.0047, 0.0215, 0.0217)^{\mathrm{T}}$；
拟牛顿法：$x^* = (0.0468, -0.0047, 0.0215, 0.0217)^{\mathrm{T}}$.

12. 证明 (4.3.29) 和 (4.3.32).

**证**: 见书本 151 页定理 4.3.6 证明.

13. 试推导预处理共轭梯度法.

**解**: 见课本 154 页.

14. 推导 (4.4.25).

**解**: 利用 Sherman-Morrison 公式，并两次运用即可得到 (4.4.25).

15. 证明拟牛顿方向是椭球范数 $\|\cdot\|_{B_k}$ 意义下的最速下降方向.

**证**: 首先求方向导数的下界

$$\frac{\mathrm{d}\, f}{\mathrm{d}\, a}=\frac{g^{\mathrm{T}}a}{\|a\|}=-\frac{-g^{\mathrm{T}}G^{-1}Ga}{\|a\|}\geqslant-\frac{\|G^{-1}g\|\times\|a\|}{\|a\|}=-\|G^{-1}g\|,$$

当且仅当 $a=G^{-1}g$ 时，

$$\frac{\mathrm{d}\, f}{\mathrm{d}\, a}=\frac{g^{\mathrm{T}}a}{\|a\|}=-\frac{-g^{\mathrm{T}}G^{-1}Ga}{\|a\|}=-\frac{(-G^{-1}g)^{\mathrm{T}}G(-G^{-1}g)}{\|a\|}=-\frac{\|a\|^2}{\|a\|}=-\|a\|=-\|G^{-1}g\|,$$

方向导数达到下界，则拟牛顿方向是椭球范数 $\|\cdot\|_{B_k}$ 意义下的最速下降方向.

或者见课本 160-161 页.

# 第五章　线性与非线性最小二乘问题

1. 设 $A$ 是 $m\times n$ 矩阵，证明 $A^{\mathrm{T}}A$ 的条件数是 $A$ 的条件数的平方，即 $\kappa(A^{\mathrm{T}}A)=[\kappa(A)]^2$.
   **略**.($A$ 应该是方阵)

2. 设 $f(x)=\dfrac{1}{2}\sum\limits_{i=1}^{5}r_i(x)^2$，其中

$$
\begin{aligned}
r_1(x)&=x_1^2+x_2^2+x_3^2-1,\\
r_2(x)&=x_1^2+x_2^2+(x_3-2)^2-1,\\
r_3(x)&=x_1+x_2+x_3-1,\\
r_4(x)&=x_1+x_2-x_3+1,\\
r_5(x)&=x_1^3+3x_2^2+(5x_3-x_1+1)^2-36,
\end{aligned}
$$

   通过计算 $r(x)$ 的雅可比矩阵确定 $\nabla f(x), M(x)=A(x)^{\mathrm{T}}A(x)$ 以及 $\nabla^2 f(x)$. 在点 $x=(0,0,1)^{\mathrm{T}}$ 是否有 $M(x)=\nabla^2 f(x)$ 成立？为什么？
   **略**.

3. 设 $f(x)=\dfrac{1}{2}\sum\limits_{i=1}^{4}r_i(x)^2, r_i(x)=x_1\mathrm{e}^{-x_2t_i}-y_i, i=1,2,3,4$，其中 $t_1=-1,t_2=0,t_3=1,t_4=2,y_1=2.7,y_2=1,y_3=0.4,y_4=0.1$.
   (1) 以 $x^{(1)}=(1,1)^{\mathrm{T}}$ 为初始点，计算进行一次 Gauss-Newton 迭代所得的点 $x^{(2)}$;
   (2) 在点 $x^{(2)}$ 还是用矩阵 $A_1(x)^{\mathrm{T}}A_1(x)$ 进行一次近似的 Gauss-Newton 迭代得点 $x^{(3)}$. 由这三个点估计方法对此问题的线性收敛速度.

   **略**.

4. 设 $\delta^{(1)}$ 和 $\delta^{(2)}$ 分别是方程组

$$(A^{\mathrm{T}}A+\mu_1 I)\delta=-A^{\mathrm{T}}r, i=1,2$$

   对应于 $\mu_1$ 和 $\mu_2$ 的解，其中 $\mu_1>\mu_2>0, A\in\mathbf{R}^{n\times n}, r\in\mathbf{R}^m$. 证明

$$q(\delta^{(2)})>q(\delta^{(1)}),$$

   这里 $q(\delta)=\dfrac{1}{2}\|A\delta+r\|^2=\dfrac{1}{2}\delta^{\mathrm{T}}A^{\mathrm{T}}A\delta+r^{\mathrm{T}}A^{\mathrm{T}}\delta+\dfrac{1}{2}r^{\mathrm{T}}r$.
   **略**.

5. 设 $f(x)=x_1^4+x_1x_2+(1+x_2)^2, x^{(1)}=(0,0)^{\mathrm{T}}$，确定参数 $\mu$ 的一个下界 $\bar{\mu}$ 使得 $\nabla^2 f(x^{(1)})+\mu I$ 在 $\mu>\bar{\mu}$ 时是正定的. 令 $\mu_1=1$，确定方程组

$$(\nabla^2 f(x^{(1)})+\mu_1 I)\delta=-\nabla f(x^{(1)})$$

   的解 $\delta^{(1)}$，并验证有 $f(x^{(1)}+\delta^{(1)})<f(x^{(1)})$.
   再验证只有当 $\mu\geqslant 0.9$ 时所得的解 $\delta^{(1)}$ 才能使 $f(x)$ 在 $x^{(1)}+\delta^{(1)}$ 处的函数值小于点 $x^{(1)}$ 处的函数值，而最优的下降量在 $\mu=1.2$ 时近似取得.
   **略**.

6. 设 $A \in \mathbf{R}^{m\times m}, \nabla f = b \in \mathbf{R}^n, \mu \geqslant 0, \delta^{(1)}$ 满足下述方程组

$$(A^{\mathrm{T}}A + \mu I)\delta = -b, \|\delta\| \leqslant \Delta.$$

证明：

(1) 如果 $\mu = 0$，则 $\delta^{(1)}$ 是下述信赖域子问题 (TR) 的解；

$$\begin{cases} \min & q(\delta) = f + \nabla f^{\mathrm{T}}\delta + \dfrac{1}{2}\delta^{\mathrm{T}}A^{\mathrm{T}}A\delta \\ \text{s.t.} & \|\delta\| \leqslant \Delta; \end{cases}$$

(2) 如果 $\mu \geqslant 0$ 且有 $\|\delta^{(1)}\| = \Delta$，则 $\delta^{(1)}$ 也是子问题 (TR) 的解；

(3) 如果 $\mu > 0$，则 $\delta^{(1)}$ 还是子问题 (TR) 的唯一解.

**略**.

7. 设 $d \in \mathbf{R}^n$ 为一满足

$$d^{\mathrm{T}}\nabla f \leqslant -c_1\|d\|\|\nabla f\|$$

的下降方向，其中 $c_1 \in (0,1]$. 又设 $\bar{\delta}$ 是问题 6 中信赖域子问题 (TR) 的一个近似解，满足

$$\mathrm{Pred}(\bar{\delta}) = f - q(\bar{\delta}) \geqslant f - \min\{f + \alpha d^{\mathrm{T}}\nabla f + \frac{1}{2}\alpha^2 d^{\mathrm{T}}A^{\mathrm{T}}Ad \mid \|\alpha d\| \leqslant \Delta\}.$$

证明：

$$\mathrm{Pred}(\bar{\delta}) \geqslant \frac{c_1}{2}\|\nabla f\| \min\left\{\Delta, c_1\frac{\|\nabla f\|}{\|A^{\mathrm{T}}A\|}\right\}.$$

**略**.

8. 设

$$q(\delta) = \frac{1}{2}\delta^{\mathrm{T}}G\delta + b^{\mathrm{T}}\delta,$$

其中

$$G = \begin{bmatrix} 14 & 1 \\ 1 & 2 \end{bmatrix}, \qquad b = (6 \quad 2)^{\mathrm{T}}.$$

分别计算 $q(\delta)$ 沿最速下降方向的最优修正，记为 $\delta_{cp}$，以及由方程组

$$D\delta = -b$$

所确定的牛顿修正，记为 $\delta_N$. 验证有 $\|\delta_N\| > \|\delta_{cp}\|$.

令 $\delta(\lambda) = \delta_{cp} + \lambda(\delta_N - \delta_{cp})$，确定 $\lambda \in (0,1)$ 的一个值使得 $\|\delta(\lambda)\| = 1$.

**略**.

9. 考虑线性等式约束的线性最小二乘问题

$$\begin{cases} \min & \dfrac{1}{2}\|Ax - b\|_2^2 \\ \text{s.t.} & Cx = y, \end{cases}$$

其中 $x\in\mathbf{R}^n, A\in\mathbf{R}^{m\times n}, b\in\mathbf{R}^m, C\in\mathbf{R}^{p\times n}$ 和 $y\in\mathbf{R}^p, p<n$. 设 $\operatorname{rank}(A)=n, \operatorname{rank}(C)=p$，证明 $x^*\in\mathbf{R}^n$ 是问题唯一解的充分必要条件为存在 Lagrange 乘子 $\lambda^*$ 使得 $(x^*,\lambda^*)$ 是问题的 Lagrange 方程组的解，并验证 $x^*,\lambda^*$ 由 (5.2.18) 和 (5.2.19) 给出.

**略**.

10. 对于问题 9，如果矩阵 $A$ 不满秩，证明问题的解集合为

$$\mathcal{S}=\{(AE)^+(b-AC^+y)+C^+y+(I-(AE)^+A\xi)\,|\,\xi\in\mathcal{N}(C)\},$$

而目标函数的最优值为

$$\|(I_A(AE)^+)(AC^+y-b)\|,$$

其中 $E=I-C^+C$，$\mathcal{N}(C)$ 是矩阵 $C$ 的零空间，$C^+$ 是矩阵 $C$ 的广义逆，$(AE)^+$ 是矩阵 $AE$ 的广义逆.

**略**.

11. 设

$$C^{\mathrm{T}}=QR=[V_1\ V_2]\begin{bmatrix}R_1^{\mathrm{T}}\\0\end{bmatrix},$$

其中 $Q=[V_1\ V_2]$ 为一 $n\times n$ 正交矩阵，$V_1\in\mathbf{R}^{n\times p}, V_2\in\mathbf{R}^{n\times(n-p)}$. 假定 Lagrange 矩阵

$$\begin{bmatrix}A^{\mathrm{T}}A & -C^{\mathrm{T}}\\ C & 0\end{bmatrix}$$

非奇异，证明：

(1) 矩阵 $C$ 满秩;

(2) 矩阵 $V_2^{\mathrm{T}}A^{\mathrm{T}}AV_2$ 非奇异.

**略**.

12. 证明由 (5.3.4) 定义的 Gauss-Newton 方向是下降方向.

**略**.

13. 设 $r_c=r(x^{(c)}), A_c=A(x^{(c)})$ 满秩，$S_c=S(x^{(c)})=\sum_{i=1}^m r_i((x^{(c)}))\nabla^2 r_i((x^{(c)})), \delta_c^G=-(A_c^{\mathrm{T}}A_c)^{-1}A_c^{\mathrm{T}}r_c$, $\delta_c^N=-(A_c^{\mathrm{T}}A_c+S_C)^{-1}A_c^{\mathrm{T}}r_c$. 证明

$$\delta_c^G-\delta_c^N=(A_c^{\mathrm{T}}A_c)^{-1}S_c\delta_c^{(k)}.$$

设 $x_c$ 充分接近 $f(x)=\dfrac{1}{2}\displaystyle\sum_{i=1}^m r_i(x)^2$ 的最优解 $x^*$，$f(x)$ 的 Hesse 矩阵 $\nabla^2 f(x)=A(x)^{\mathrm{T}}A(x)+S(x)$ 在点 $x^*$ 非奇异，在点 $x^*$ 的领域内 Lipschitz 连续，利用上述关系式证明

$$\|x_c^G-x^*\|\leqslant\|(A_c^{\mathrm{T}}A_c)^{-1}\|\|S_c\|\|x_c-x^*\|+O(\|x_c-x^*\|^2),$$

其中 $x_c^G=x_c+\delta_c^G$.

**略**.

14. 已知 $r \in \mathbf{R}^m, A \in \mathbf{R}^{m\times n}, \mu > 0$. 证明 $\delta = -(A^{\mathrm{T}}A + \mu I)^{-1}A^{\mathrm{T}}r$ 是下述最小二乘问题的解

$$\min \|J\delta + y\|^2,$$

其中

$$J = \begin{bmatrix} A \\ \mu^{\frac{1}{2}} I \end{bmatrix}, \quad y = \begin{bmatrix} r \\ 0 \end{bmatrix}.$$

**略.**

15. 考虑非线性最小二乘问题

$$\min f(x) = \frac{1}{2}\sum_{i=1}^{m} r_i(x)^2,$$

并假定 $r(x)$ 的 Jacobian 矩阵 $A(x)$ 对所有 $x \in \mathbf{R}^n$ 都满秩，用 $\delta^G, \delta^{LM}(\mu)$ 和 $\delta^S$ 分别表示在点 $x$ 的 Gauss-Newton 方向，Levenberg-Marquardt 方向和最速下降方向，即有

$$\begin{aligned}
\delta^G &= (A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}r, \\
\delta^{LM}(\mu) &= (A^{\mathrm{T}}A + \mu I)^{-1}A^{\mathrm{T}}r, \\
\delta^S &= -A^{\mathrm{T}}r.
\end{aligned}$$

证明

$$\begin{aligned}
\lim_{\mu\to 0}\delta^{LM}(\mu) &= \delta^G, \\
\lim_{\mu\to\infty}\frac{\delta^{LM}(\mu)}{\|\delta^{LM}(\mu)\|} &= \frac{\delta^S}{\|\delta^S\|}.
\end{aligned}$$

**略.**

# 第六章　二次规划

1. 解下面二次规划问题并且作图解释几何意义.

$$\begin{aligned}\max\quad & f(x)=2x_1+3x_2+4x_1^2+2x_1x_2+x_2^2\\ \text{s.t.}\quad & x_1-x_2\geqslant 0,\\ & x_1+x_2\leqslant 4,\\ & x_1\leqslant 3.\end{aligned}$$

**解**: $\max=52, x_1=3, x_2=1$.

2. 从一点 $x_0$ 的超平面 $\{x|Ax=b\}$ 寻找最短距离能构成二次规划形式，这里 $A$ 是行满秩矩阵

$$\begin{aligned}\min\quad & f(x)=\frac{1}{2}(x-x_0)^{\mathrm{T}}(x-x_0)\\ \text{s.t.}\quad & Ax=b.\end{aligned}$$

证明最优乘子为

$$\lambda^*=(AA^{\mathrm{T}})^{-1}(b-Ax_0)$$

和最优解为

$$x^*=x_0-A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}(b-Ax_0),$$

并且证明当 $A$ 是行向量时，从 $x_0$ 至 $Ax=b$ 的可行解集的最短距离为

$$\frac{|b-Ax_0|}{\|A\|}.$$

[注意：最优解应为 $x^*=x_0+A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}(b-Ax_0)$]

**证**: 令 $A$ 是 $m\times n$ 的矩阵，由于 $A$ 是行满秩矩阵，则 $\operatorname{rank}(A)=m$，而 $\operatorname{rank}(AA^{\mathrm{T}})=\operatorname{rank}(A)=m$，则 $AA^{\mathrm{T}}$ 是 $m\times m$ 的可逆方阵.

本问题的 Lagrange 函数为 $L(x;\lambda)=\dfrac{1}{2}(x-x_0)^{\mathrm{T}}(x-x_0)-\lambda^{\mathrm{T}}(Ax-b)$，则 KKT 条件为

$$\begin{cases}(x-x_0)=A^{\mathrm{T}}\lambda\\ Ax=b\end{cases}$$

于是，$AA^{\mathrm{T}}\lambda=A(x-x_0)=Ax-Ax_0=b-Ax_0$，而 $AA^{\mathrm{T}}$ 是 $m\times m$ 的可逆方阵，则 $\lambda^*=(AA^{\mathrm{T}})^{-1}(b-Ax_0)$，再代入第一个 KKT 条件，得 $x^*=x_0+A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}(b-Ax_0)$.

当 $A$ 是行向量，即 $m=1$ 时，

$$\begin{aligned}d^2=\|x-x_0\|^2&=(x-x_0)^{\mathrm{T}}(x-x_0)\\ &=(b-Ax_0)^{\mathrm{T}}\cdot\frac{1}{(AA^{\mathrm{T}})^2}AA^{\mathrm{T}}(b-Ax_0)\\ &=\frac{(b-Ax_0)^{\mathrm{T}}(b-Ax_0)}{AA^{\mathrm{T}}}\end{aligned}$$

因此，$d=\dfrac{|b-Ax_0|}{\|A\|}$. 证毕.

3. 证明 (6.2.1) 的一阶必要性条件是 (6.2.2).

   **略**.

4. 写出下面既有等式约束又有不等式约束的二次凸规划的 KKT 条件

$$\begin{aligned} \min \quad & q(x) = \frac{1}{2}x^{\mathrm{T}}Gx + x^{\mathrm{T}}g \\ \text{s.t.} \quad & Ax \geqslant b, \quad \bar{A}x = \bar{b}. \end{aligned}$$

   这里 $G$ 是对称正半定矩阵，使用这些条件导出类似的原始—对偶步.

   **略**.

5. 考虑二次规划

$$\begin{aligned} \min \quad & q(x) = 6x_1 + 4x_2 - 13 - x_1^2 - x_2^2 \\ \text{s.t.} \quad & x_1 + x_2 \leqslant 3, \\ & x_1, x_2 \geqslant 0. \end{aligned}$$

   使用图解法求解，再使用有效约束集方法求解.

   **解**: $\min = -13, x_1 = 0, x_2 = 0.$

6. 用有效集方法求解二次规划问题

$$\begin{aligned} \min \quad & q(x) = x_1^2 + 2x_2^2 - 2x_1 - 6x_2 - 2x_1x_2 - 2 \\ \text{s.t.} \quad & \frac{1}{2}x_1 + \frac{1}{2}x_2 \leqslant 1, \\ & -x_1 + 2x_2 \leqslant 2, \\ & x_1, x_2 \geqslant 0. \end{aligned}$$

   **解**: $\min = -\frac{46}{5}, x_1 = \frac{4}{5}, x_2 = \frac{6}{5}.$

7. 推导 (6.2.24).

   **略**.

# 第七章　约束最优化的理论与方法

1. 对于 $x \in \mathbf{R}$，证明函数 $\min(z,0)^2$ 在 $z=0$ 处有二阶不连续导数.

   **证**: 将这个函数记为 $f(z)$，则 $f(z)$ 可写为

$$f(z)=\begin{cases} z^2, & z=0 \\ 0, & z\neq 0 \end{cases}, \quad \therefore f''(z)=\begin{cases} 1, & z\neq 0 \\ +\infty, & z\to 0^+ \\ -\infty, & z\to 0^- \end{cases}$$

   所以，函数 $\min(z,0)^2$ 在 $z=0$ 处有二阶不连续导数.

2. 考虑问题

$$\begin{aligned} \min \quad & \frac{1}{1+x^2} \\ \text{s.t.} \quad & x \geqslant 1, \end{aligned}$$

   写出 $P(x;\mu)$.

   **解**: 对数障碍函数为

$$P(x;\mu)=\frac{1}{1+x^2}-\mu\log(x-1).$$

   分数障碍函数为

$$P(x;\mu)=\frac{1}{1+x^2}-\frac{\mu}{x-1}.$$

3. 考虑问题

$$\begin{aligned} \min \quad & x \\ \text{s.t.} \quad & x^2 \geqslant 0, \\ & x+1 \geqslant 0, \end{aligned}$$

   解为 $x^*=-1$，写出此问题的 $P(x;\mu)$，寻找局部极小点.

   **解**: $P(x;\mu)=x-\mu\log(x^2)-\mu\log(x+1)$，则

$$\frac{\mathrm{d}P}{\mathrm{d}x}=1-\mu\frac{2x}{x^2}-\mu\frac{1}{x+1}=0,$$

   解得

$$x=\frac{3\mu-1\pm\sqrt{(1-3\mu)^2+8\mu}}{2}(\text{由可行性舍去}-\text{的情况}),$$

   令 $\mu\to 0^+$，则 $x=-1$.

   也可以将 $x^2\geqslant 0$ 舍弃，因为这是个无效约束，则 $P(x;\mu)=x-\mu\log(x+1)$，则

$$\frac{\mathrm{d}P}{\mathrm{d}x}=1-\mu\frac{1}{x+1}=0,$$

   解得

$$x=\mu-1,$$

   令 $\mu\to 0^+$，则 $x=-1$.

4. 假设当前点 $x_k$ 是 $P(x;\mu_k)$ 的精确极小点，写出关于当前点 $x_k$ 对于 $P(x;\mu_{k+1})$ 的牛顿步.

   **略**.

5. 试推导 (7.4.24).

   **略**.

6. 对附录中试验函数 $2.1-2.3$ 分别用罚函数法和序列二次规划方法求解 (用 C,MATLAB 或 FORTRAN 等语言编制程序).

   **略**.

**第七章：补充题目**

1. 考虑约束最优化问题

$$
\begin{aligned}
\min \quad & (x_1-2)^2+x_2^2 \\
\text{s.t.} \quad & x_1^2-x_2 \leqslant 0, \\
& x_1-x_2 \leqslant 0,
\end{aligned}
$$

求出其 KKT 点并判断是否为其全局最优解.

**解**: 本问题的 KKT 条件为

$$
\begin{cases}
\begin{pmatrix} 2x_1-4 \\ 2x_2 \end{pmatrix} = \lambda_1 \begin{pmatrix} -2x_1 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} -1 \\ 1 \end{pmatrix} \\
\lambda_1 \geqslant 0, \lambda_2 \geqslant 0, \\
\lambda_1(x_1^2-x_2)=0, \lambda_2(x_1-x_2)=0.
\end{cases}
$$

利用第三个 KKT 条件，可得以下情况：

(1) $\lambda_1=0, \lambda_2=0, x_1=2, x_2=0$，不满足约束条件，不是 KKT 点.

(2) $\lambda_1=0, x_1-x_2=0 \Rightarrow x_1=x_2=1, \lambda_2=2$，是 KKT 点.

(3) $x_1^2-x_2=0, \lambda_2=0 \Rightarrow 2x_1^3+x_1-2=0$，令 $g(x_1)=2x_1^3+x_1-2$，显然 $g(x_1)$ 单调递增，而 $g(0)=-2<0, g(1)=1>0$，于是 $x_1 \in (0,1), x_2=x_1^2<x_1$，不满足第二个约束，不是 KKT 点.

(4) $x_1^2-x_2=0, x_1=x_2=0 \Rightarrow x^{(1)}=(0,0)^{\mathrm{T}}, x^{(2)}=(1,1)^{\mathrm{T}}$，$x^{(2)}$ 就是情况 (2)，$x^{(1)}$ 对应的 $\lambda_1=-4, \lambda_2=4$ 不满足 KKT 条件，不是 KKT 点.

由于这个问题是个凸规划，所以其 KKT 点就是全局最优解.

2. 考虑约束最优化问题

$$
\begin{aligned}
\min \quad & x_1^2+2x_2^2 \\
\text{s.t.} \quad & x_1+x_2-1 \geqslant 0,
\end{aligned}
$$

利用对数障碍函数法求解.

**解**: 对数障碍函数 $P(x;\mu)=x_1^2+2x_2^2-\mu\log(x_1+x_2-1)$，则

$$
\begin{cases}
\dfrac{\partial P}{\partial x_1} = 2x_1 - \dfrac{\mu}{x_1+x_2-1} = 0 \\
\dfrac{\partial P}{\partial x_2} = 4x_2 - \dfrac{\mu}{x_1+x_2-1} = 0
\end{cases}
$$

解得

$$
\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \dfrac{1+\sqrt{1+3\mu}}{3} \\ \dfrac{1+\sqrt{1+3\mu}}{6} \end{pmatrix}, \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \dfrac{1-\sqrt{1+3\mu}}{3} \\ \dfrac{1-\sqrt{1+3\mu}}{6} \end{pmatrix} \text{(不可行，舍去)}
$$

当 $\mu \to 0^+$ 时，

$$
\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \dfrac{2}{3} \\ \dfrac{1}{3} \end{pmatrix}.
$$