例1. 梯度下降法. 最小值

用梯度下降法求下列函数最小值: $f(x)=(x_1-2)^2+2x_2^2$. 初值(2,2). 精度10^{-4}.

解: $x^1=x^0-\lambda\nabla f(x)$.

$\nabla f(x)=\begin{pmatrix}2x_1-4\\4x_2\end{pmatrix}$ 初值 $x^0=\begin{pmatrix}2\\2\end{pmatrix}$ 故 $\nabla f(x)=\begin{pmatrix}0\\8\end{pmatrix}$.

① $x^1=x^0-\lambda\nabla f(x)=x^0-\lambda\begin{pmatrix}0\\8\end{pmatrix}=\begin{pmatrix}2\\2-8\lambda\end{pmatrix}$

$f(x^1)=2(2-8\lambda)^2$

令 $\frac{\partial f(x^1)}{\partial\lambda}=0$. 解出 $\lambda=\frac{1}{4}$. $\therefore x^1=\begin{pmatrix}2\\0\end{pmatrix}$

此时 $\nabla f(x^1)=\begin{pmatrix}0\\0\end{pmatrix}$. $\|\nabla f(x^1)\|<\varepsilon$. 达到精度要求.

故函数在 $x=\begin{pmatrix}2\\0\end{pmatrix}$ 处取到min. 为0.

例2. 共轭梯度下降法.

用共轭梯度下降法求下列函数最小值: $f(x)=(x_1-1)^2+2(x_2-x_1^2)^2$. 初值(0,0). 精度$10^{-4}$.

解: ① $\nabla f(x)=\begin{pmatrix}8x_1^3+2x_1-8x_1x_2-2\\4x_2-4x_1^2\end{pmatrix}$

$S_0=-\nabla f(x)=\begin{pmatrix}2\\0\end{pmatrix}$ 负梯度方向. 起始点(0,0)

$x^1=x^0+\lambda S_0=\begin{pmatrix}2\lambda\\0\end{pmatrix}$

$f(x^1)=(2\lambda-1)^2+32\lambda^4$.

令 $f'(x^1)=0$. 求得 $\lambda=\frac{1}{4}$. 则 $x^1=\begin{pmatrix}0.5\\0\end{pmatrix}$.

② 共轭方向: $S_1=-\nabla f(x^1)+\beta_0 S_0$.

$\nabla f(x^1)=\begin{pmatrix}0\\-1\end{pmatrix}$ $\beta_0=\frac{\|\nabla f(x^1)\|^2}{\|\nabla f(x)\|^2}=\frac{1}{4}$

$\therefore S_1=\begin{pmatrix}\frac{1}{2}\\1\end{pmatrix}$

$x^2=x^1+\lambda S_1=\begin{pmatrix}0.5+0.5\lambda\\ \underbrace{\lambda}\end{pmatrix}$. x中保留λ

$f(x^2)=\cdots\cdots$ 略. 代入即可.

令 $f'(x^2)=0$. 求得 $\lambda=1$. 则 $x^2=\begin{pmatrix}1\\1\end{pmatrix}$.

③ $\nabla f(x^2)=\begin{pmatrix}0\\0\end{pmatrix}$. $\|\nabla f(x^2)\|<\varepsilon=10^{-4}$.

达到精度要求. 迭代结束

综: 在(1,1)处原函数有min. 最小值为0.

共轭梯度下降法:

第1次是负梯度方向.

第2次开始是共轭方向

↑

2个方向叠加形成

例3.(1) $\max z = x_1 + 2x_2 + x_3$

$$\text{s.t.}\begin{cases} 2x_1 - 3x_2 + 2x_3 \le 15 \\ \frac{1}{3}x_1 + x_2 + 5x_3 \le 20. \\ x_1, x_2, x_3 \ge 0. \end{cases}$$

标准的单纯形表，有现成的单位阵.

解：先化作标准形式：

$\max z = x_1 + 2x_2 + x_3$

$$\text{s.t.}\begin{cases} 2x_1 - 3x_2 + 2x_3 + x_4 = 15 \\ \frac{1}{3}x_1 + x_2 + 5x_3 + x_5 = 20 \\ x_1, x_2, x_3 \ge 0.\ x_4, x_5 \ge 0. \end{cases}$$

对应目标函数中系数

c_j：目标函数中系数.

θ ← $\theta_L = \min\left\{\frac{b_i}{a_{ik}} \mid a_{ik} > 0\right\}$. 用于确定离基

	x_b	C_b	b \ c_j	1 x_1	2 x_2	1 x_3	0 x_4	0 x_5	θ
①	x_4	0	15	2	$-3 \to 0$	2	1	0	\ ($-3<0$. 不用算)
②	x_5	0	20	$\frac{1}{3}$	1	5	0	1	20 ($20 \div 1 = 20$) (小，出基 x_5)
			σ	1	2 (大. x_2入基)	1	0	0	

检验数 $\sigma_j = c_j - \sum c_i a_{ij}$　　　基变量 σ 必为0.

x_1: $\sigma_1 = 1 - (0\times2 + 0\times\frac{1}{3}) = 1$

	x_b	C_b	b \ c_j	1 x_1	2 x_2	1 x_3	0 x_4	0 x_5	θ	
③	x_4	0	75	$3 \to 1$	0	17	1	3	25 (小. x_4出)	③=①+3×② (行列式变换)
④	x_2	2	20	$\frac{1}{3} \to 0$	1	5	0	1	60	④=②
			σ	$\frac{1}{3}$	0	-9	0	-2		

大. x_1入基.

	x_b	C_b	b	x_1	x_2	x_3	x_4	x_5	θ	
⑤	x_1	1	25	1	0	$\frac{17}{3}$	$\frac{1}{3}$	1		⑤=③÷3
⑥	x_2	2	$\frac{35}{3}$	0	1	$-\frac{2}{3}$	$-\frac{1}{3}$	0		⑥=④−⑤÷3
			σ	0	0	$-\frac{10}{3}$ <0	$-\frac{110}{9}$ <0	-25 <0		

已找到最优解. 结束.

说明有唯一最优解. 当 $x = (25, \frac{35}{3}, 0)$ 时. $\max z = \frac{105}{3}$.

例3(2). $\max z = x_1 + x_2 - 2x_3$

$$\text{s.t.}\begin{cases} 3x_1 + x_2 - x_3 \le 5 \\ x_1 - 4x_2 + x_3 \ge 7 \\ x_1, x_2, x_3 \ge 0. \end{cases}$$

无现成单位阵. 要添加人工变量.

▲注意：$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ 不是单位阵！

解：标准化. $\max z = x_1 + x_2 - 2x_3$

$$\text{s.t.}\begin{cases} 3x_1 + x_2 - x_3 + x_4 = 5 \\ x_1 - 4x_2 + x_3 \quad - x_5 = 7 \\ x_1, x_2, x_3, x_4, x_5 \ge 0. \end{cases}$$

无现成单位阵. 要添加人工变量. 可以用2种方法 $\begin{cases}\text{大M法} \\ \text{两阶段法.}\end{cases}$

法1：大M法. $\max z = x_1 + x_2 - 2x_3 - \underline{\underline{M}} a_1$ （M是无穷大）.

	x_b	c_b	c_j / b	1 / x_1	1 / x_2	-2 / x_3	0 / x_4	0 / x_5	-M / a_1	θ	
①	x_4	0	5	3→1	1	-1	1	0	0	$\frac{5}{3}$	（小. x_4 出基）
②	x_5	-M	7	1→0	-4	1	0	-1	1	7	
			σ	M+1	1-4M	M-2	0	-M	0		

$1-0\times3-1\times(-M)=M+1$

（大, x_1 进基）

	x_b	c_b	b	x_1	x_2	x_3	x_4	x_5	a_1	θ	
①÷3=③	x_1	1	$\frac{5}{3}$	1	$\frac{1}{3}$	$-\frac{1}{3}$→0	$\frac{1}{3}$	0	0	负	
②-③=④	a_1	-M	$\frac{16}{3}$	0	$-\frac{13}{3}$	$\frac{4}{3}$→1	$-\frac{1}{3}$	-1	1	4	a_1 出基
			σ	0	$\frac{2}{3}$-13M	$\underline{\frac{4}{3}M-\frac{5}{3}}$	$\frac{1}{3}-\frac{M}{3}$	-M	0		

（大, x_3 入基）

	x_b	c_b	b	x_1	x_2	x_3	x_4	x_5	a_1	θ
③-⑥×$\frac{1}{3}$=⑤	x_1	1	3	1	$-\frac{3}{4}$	0	$+\frac{1}{4}$	$-\frac{1}{4}$	$\frac{1}{4}$	
④×$\frac{3}{4}$=⑥	x_3	-2	4	0	$-\frac{13}{4}$	1	$-\frac{1}{4}$	$-\frac{3}{4}$	$\frac{3}{4}$	
			σ	0	$-\frac{19}{4}$	0	$-\frac{3}{4}$	$-\frac{5}{4}$	$-M+\frac{5}{4}$	

检验数σ皆小于等于0. 迭代结束. 有唯一的最优解.

$x=(3,0,4)$ 时有最优解 $\max z=-5$

法2：两阶段法.

① $\max z = -a_1$

x_b	c_b	c_j / b	x_1 (0)	x_2 (0)	x_3 (0)	x_4 (0)	x_5 (0)	a_1 (−1)	θ
x_4	0	5	3 →1	1	−1	1	0	0	$\frac{5}{3}$ → (小, x_4出基)
a_1	−1	7	1 →0	−4	1	0	−1	1	7
		σ	1	−4	1	0	−1	0	

(大, x_1入基)

x_b	c_b	b	x_1	x_2	x_3	x_4	x_5	a_1
x_1	0	3	1	$-\frac{3}{4}$	0	$\frac{1}{4}$	$-\frac{1}{4}$	$\frac{1}{4}$
x_3	0	4	0	$-\frac{13}{4}$	1	$-\frac{1}{4}$	$-\frac{3}{4}$	$\frac{3}{4}$
		σ	0	0	0	0	0	−1

✓

(?)

② $a_1 = 0$.

x_b	c_b	c_j / b	x_1 (1)	x_2 (1)	x_3 (−2)	x_4 (0)	x_5 (0)
x_1	1	3	1	$-\frac{3}{4}$	0	$+\frac{1}{4}$	$-\frac{1}{4}$
x_3	−2	4	0	$-\frac{13}{4}$	1	$-\frac{1}{4}$	$-\frac{3}{4}$
		σ	0	$-\frac{19}{4}$	0	$-\frac{3}{4}$	$-\frac{5}{4}$
				<0		<0	<0

迭代结束. 存在唯一的最优解. $x=(3, 0, 4)$时. 取到最优解 $\max z = -5$

DP作业. 方案设计.

例4. 动态规划方案设计.

用户	利润 \ 设备数目	0	1	2	3	4	5	6
A	$g_1(x)$	0	4	9	12	14	16	19
B	$g_2(x)$	0	3	8	11	15	17	18
C	$g_3(x)$	0	5	10	12	14	16	17

$f_k(x)$: 数量x的设备给前k个用户得利益最大

$g_k(y)$: 数量y的设备给第k个用户的得利润

Step1: 求 $f_1(x)$. 此时 $f_1(x)=g_1(x)$

设备	0	1	2	3	4	5	6
$f_1(x)=g_1(x)$	0	4	9	12	14	16	19
最优策略	0	1	2	3	4	5	6

Step2: 求 $f_2(x)$. 此时要考虑用户A.B之间如何分配使利润max.

① $f_2(6)=\max\limits_{y=0,1..6}\{g_2(y)+f_1(6-y)\}=\max\begin{Bmatrix} g_2(0)+f_1(6)\\ g_2(1)+f_1(5)\\ g_2(2)+f_1(4)\\ g_2(3)+f_1(3)\\ g_2(4)+f_1(2)\\ g_2(5)+f_1(1)\\ g_2(6)+f_1(0)\end{Bmatrix}=\max\begin{Bmatrix}0+19\\3+16\\8+14\\11+12\\15+9\\17+4\\18+0\end{Bmatrix}=23$

(A, B)=(3,3)时 最大利润为:23.

② $f_2(5)=\max\limits_{y=0,1..5}\{g_2(y)+f_1(5-y)\}=\max\begin{Bmatrix} g_2(0)+f_1(5)\\ g_2(1)+f_1(4)\\ g_2(2)+f_1(3)\\ g_2(3)+f_1(2)\\ g_2(4)+f_1(1)\\ g_2(5)+f_1(0)\end{Bmatrix}=\max\begin{Bmatrix}0+16\\3+14\\8+12\\11+9\\15+4\\17+0\end{Bmatrix}=20$

(A, B)=(2,3)时 或(3,2) 最大利润为:20

③ $f_2(4)=\max\limits_{y=0,1..4}\{g_2(y)+f_1(4-y)\}=\max\begin{Bmatrix} g_2(0)+f_1(4)\\ g_2(1)+f_1(3)\\ g_2(2)+f_1(2)\\ g_2(3)+f_1(1)\\ g_2(4)+f_1(0)\end{Bmatrix}=\max\begin{Bmatrix}0+14\\3+12\\8+9\\11+4\\15+0\end{Bmatrix}=17$

(A, B)=(2,2)时 最大利润为17

④ $f_2(3)=\max\limits_{y=0,1..3}\{g_2(y)+f_1(3-y)\}=\max\begin{Bmatrix} g_2(0)+f_1(3)\\ g_2(1)+f_1(2)\\ g_2(2)+f_1(1)\\ g_2(3)+f_1(0)\end{Bmatrix}=\max\begin{Bmatrix}0+12\\3+9\\8+4\\11+0\end{Bmatrix}=12$

(A, B)=(3,0)或 (2,1)或 (1,2) 时 最大利润为:12

⑤ $f_2(2)=\max\limits_{y=0,1,2}\{g_2(y)+f_1(2-y)\}=\max\left\{\begin{matrix} g_2(0)+f_1(2) \\ g_2(1)+f_1(1) \\ g_2(2)+f_1(0)\end{matrix}\right\}=\max\left\{\begin{matrix}0+9\\3+4\\8+0\end{matrix}\right\}=9$　　$(A,B)=(2,0)$时 最大利润为:9

⑥ $f_2(1)=\max\limits_{y=0,1}\{g_2(y)+f_1(1-y)\}=\max\left\{\begin{matrix} g_2(0)+f_1(1) \\ g_2(1)+f_1(0)\end{matrix}\right\}=\max\left\{\begin{matrix}0+4\\3+0\end{matrix}\right\}=4$　　$(A,B)=(1,0)$时 最大利润为:4

⑦ $f_2(0)=0$　综上，有如下总结:

设备 / $f_2(x)$	0	1	2	3	4	5	6
利润	0	4	9	12	17	20	33
最优策略 (A,B)	(0,0)	(1,0)	(2,0)	(3,0) (1,2) (2,1)	(2,2)	(2,3) (3,2)	(3,3)

Step 3: 把用户C纳入考量，计算 $f_3(6)$

$$f_3(6)=\max_{y=0,1,\dots,6}\{g_3(y)+f_2(6-y)\}=\max\left\{\begin{matrix} g_3(0)+f_2(6)\\ g_3(1)+f_2(5)\\ g_3(2)+f_2(4)\\ g_3(3)+f_2(3)\\ g_3(4)+f_2(2)\\ g_3(5)+f_2(1)\\ g_3(6)+f_2(0)\end{matrix}\right\}=\max\left\{\begin{matrix}0+33\\5+20\\10+17\\12+12\\14+9\\16+4\\17+0\end{matrix}\right\}=33$$

$(A,B,C)=(3,3,0)$ 最大利润为:33

综上，要总利润最大，6台设备的分配方式为：A用户3台，B用户3台，C用户0台。

获得最大利润为：33（万元）

例6

1、共轭梯度法 $\min f(x)=\frac{3}{2}x_1^2+\frac{1}{2}x_2^2-x_1x_2-2x_1$，初始点 $x^{(1)}=(-2,4)^T$

解：$\frac{\partial f}{\partial x_1}=3x_1-x_2-2=f_{x_1}(x)$

$\frac{\partial f}{\partial x_2}=x_2-x_1=f_{x_2}(x)$

初始点 $x^{(1)}=\begin{bmatrix}-2\\4\end{bmatrix}$，故 $\nabla f(x^{(1)})=\begin{bmatrix}-12\\6\end{bmatrix}$，$S^{(1)}=-\nabla f(x^{(1)})=\begin{bmatrix}12\\-6\end{bmatrix}$

于是：$x^{(2)}=x^{(1)}+\alpha S^{(1)}=\begin{bmatrix}-2\\4\end{bmatrix}+\alpha\begin{bmatrix}12\\-6\end{bmatrix}=\begin{bmatrix}-2+12\alpha\\4-6\alpha\end{bmatrix}$

代入原函数：$f(\alpha)=\frac{3}{2}(-2+12\alpha)^2+\frac{1}{2}(4-6\alpha)^2-(-2+12\alpha)(4-6\alpha)-2(-2+12\alpha)$

$=306\alpha^2-180\alpha+26$

求导 $f'(\alpha)=612\alpha-180=0$ 时，$\alpha=\frac{180}{612}=\frac{5}{17}$，故 $x^{(2)}=\begin{bmatrix}\frac{26}{17}\\\frac{38}{17}\end{bmatrix}$

则 $\nabla f(x^{(2)})=\begin{bmatrix}3\times\frac{26}{17}-\frac{38}{17}-2\\\frac{38}{17}-\frac{26}{17}\end{bmatrix}=\begin{bmatrix}\frac{6}{17}\\\frac{12}{17}\end{bmatrix}$

第二次迭代，求共轭方向：

$\beta_0=\frac{\|\nabla f(x^{(2)})\|^2}{\|\nabla f(x^{(1)})\|^2}=\frac{(\frac{6}{17})^2+(\frac{12}{17})^2}{12^2+(-6)^2}=\frac{1}{289}$

$S^{(2)}=-\nabla f(x^{(2)})+\beta_0S^{(1)}=-\begin{bmatrix}\frac{6}{17}\\\frac{12}{17}\end{bmatrix}+\frac{1}{289}\begin{bmatrix}12\\-6\end{bmatrix}=\begin{bmatrix}-\frac{90}{289}\\-\frac{210}{289}\end{bmatrix}$

$x^{(3)}=x^{(2)}+\alpha S^{(2)}=\begin{bmatrix}\frac{26}{17}\\\frac{38}{17}\end{bmatrix}+\alpha\begin{bmatrix}-\frac{90}{289}\\-\frac{210}{289}\end{bmatrix}=\begin{bmatrix}\frac{26}{17}-\frac{90}{289}\alpha\\\frac{38}{17}-\frac{210}{289}\alpha\end{bmatrix}$，代入原函数：

$f(x^{(3)})=\varphi(\alpha)=\frac{3}{2}(\frac{26}{17}-\frac{90}{289}\alpha)^2+\frac{1}{2}(\frac{38}{17}-\frac{210}{289}\alpha)^2\cancel{+2}-(\frac{26}{17}-\frac{90}{289}\alpha)(\frac{38}{17}-\frac{210}{289}\alpha)-2(\frac{26}{17}-\frac{90}{289}\alpha)$

$\varphi'(\alpha)=\frac{3}{2}\cdot(\frac{90}{289})^2\cdot2\cdot\alpha-\frac{3}{2}\times\frac{90}{289}\times2+\frac{1}{2}\cdot(\frac{210}{289})^2\cdot2\cdot\alpha+\frac{1}{2}\cdot\frac{210}{289}\cdot2-$

$(-\frac{26}{17}\times\frac{210}{289}-\frac{90}{289}\times\frac{38}{17}+\frac{90\times210}{289^2}\cdot2\alpha)+2\cdot\frac{90}{289}$

$=\alpha\left(3\cdot(\frac{90}{289})^2+(\frac{210}{289})^2-\frac{90\times210}{289^2}\cdot2\right)-3\cdot\frac{90}{289}\cdot\frac{26}{17}+\frac{210}{289}\cdot\frac{38}{17}+\frac{26}{17}\times\frac{210}{289}+\frac{90}{289}\times\frac{38}{17}=0$.

解得 $\alpha=\frac{17}{10}$

故 $x^{(3)}=\begin{bmatrix}1\\1\end{bmatrix}$ 则 $\nabla f(x^{(3)})=\begin{bmatrix}3-1-2\\1-1\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$ $\therefore\|\nabla f(x^{(3)})\|=0$

故当 $x_1=1$，$x_2=1$ 时，$f(x)$ 取到最小值，$f(x)_{\min}=-1$

2、证明向量 $a_1=(1,0)^T$ 与向量 $a_2=(3,-2)^T$ 关于矩阵 $A=\begin{bmatrix}2&3\\3&5\end{bmatrix}$ 共轭.

证明:

$$a_1^T A a_2=[1,0]\begin{bmatrix}2&3\\3&5\end{bmatrix}\begin{bmatrix}3\\-2\end{bmatrix}$$

$$=[2\ \ 3]\begin{bmatrix}3\\-2\end{bmatrix}=2\times3+3\times(-2)=0.$$ 故 $a_1^T A a_2=0$

所以 a_1、a_2 关于 A 共轭. Q.E.D.