第11章　无约束最优化的直接方法

坐标轮换法(Cyclic Coordinate Method）

设在R^n中有n个基，分别为

$$e_1=(1,0,\cdots,0)^T,e_2=(0,1,\cdots,0)^T,\cdots,e_n=(0,0,\cdots,1)^T$$

基本思想：

$$x^{(1)}\Rightarrow x_{11}\xrightarrow{e_1}x_{12}\xrightarrow{e_2}x_{13}\rightarrow\cdots\rightarrow x_{1n}\xrightarrow{e_n}x_{1n+1}$$

$$x_{1n+1}\Rightarrow x_{21}\xrightarrow{e_1}x_{22}\xrightarrow{e_2}x_{23}\rightarrow\cdots\rightarrow x_{2n}\xrightarrow{e_n}x_{2n+1}$$

$$x_{2n+1}\Rightarrow x_{31}\xrightarrow{e_1}x_{32}\xrightarrow{e_2}x_{33}\rightarrow\cdots$$

步 骤：

1.任取$x^{(0)} \in R^n, \varepsilon > 0$，沿$e_1$进行一维搜索：

$$\min f(x^{(0)} + \lambda e_1) = f(x^{(0)} + \lambda_0 e_1),$$

令$x^{(1)} = x^{(0)} + \lambda_0 e_1$, 置$k = 1$.

2.沿e_{k+1}进行一维搜索：

$$\min f(x^{(k)} + \lambda e_{k+1}) = f(x^{(k)} + \lambda_k e_{k+1}),$$

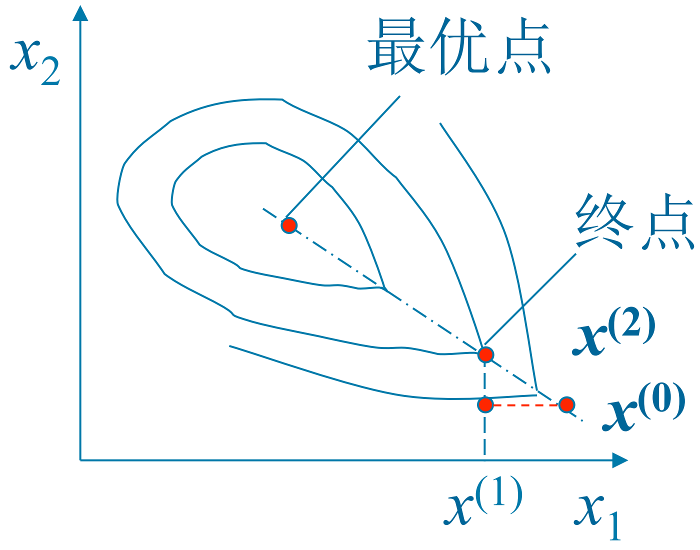
令$x^{(k+1)} = x^{(k)} + \lambda_k e_{k+1}$, 置$k = k+1$，若$k = n$, 则进入3;

否则，返回1.

3.若$\left\| x^{(n)} - x^{(0)} \right\| < \varepsilon$, 则停止计算，得点$x^* = x^{(n)}$;

否则，令$x^{(0)} = x^{(n)}$，返回1。

这种方法简单、直观，但对于山脊形函数或自变量间有大的交互作用不适用。



模式搜索法（步长加速法）

$$\text{两种类型的移动}\begin{cases}\text{探测移动}\\ \text{模式移动}\end{cases}$$

探测移动：

依次沿n个坐标轴进行，用于确定新的基点和有利于函数值下降的方向。

模式移动：

沿相邻两个基点连线方向进行，试图使函数值更快减少。

设目标函数$f(x)$，$x \in R^n$，坐标轴方向分别为

$e_1 = (1,0,\cdots,0)^T, e_2 = (0,1,\cdots,0)^T, \cdots, e_n = (0,0,\cdots,1)^T$，

给定初始点$x^{(1)} \in R^n$，各坐标方向的初始步长$\delta > 0$，

加速因子$\alpha \geq 1$，缩短因子$\beta \in (0,1)$及精度要求$\varepsilon > 0$。

把沿第i个坐标方向搜索得到的点记为$y^{(i+1)}$.

初始点$y^{(1)} = x^{(1)}$.

探测性搜索：

(1) 沿e_1方向进行探测性搜索

若$f(y^{(1)}) > f(y^{(1)} + \delta e_1)$，则称探测成功，

令$y^{(2)} = y^{(1)} + \delta e_1$.

若$f(y^{(1)}) \le f(y^{(1)} + \delta e_1)$，则称探测失败，

这时再沿e_1的反方向$-e_1$进行探测性搜索：

若$f(y^{(1)}) > f(y^{(1)} - \delta e_1)$，则称反方向探测成功，

令$y^{(2)} = y^{(1)} - \delta e_1$.

若$f(y^{(1)}) \le f(y^{(1)} - \delta e_1)$，则称探测失败，

令$y^{(2)} = y^{(1)}$。

(2) 从$y^{(2)}$出发，沿e_2进行类似的探测搜索，得$y^{(3)}$。

(3)重复上述步骤,直到沿n个坐标方向都探测完毕,最后得到$y^{(n+1)}$.

(4) 检验总的探测效果是否满意：

若$f(y^{(n+1)}) < f(x^{(1)})$,则称完成了探测性搜索，以$y^{(n+1)}$为新的基点，记$x^{(2)} = y^{(n+1)}$，进行下一步探索

-----模式搜索。

否则，要将步长缩短为$\beta\delta$，再从$x^{(1)}$开始进行探测性搜索。

模式搜索

从$x^{(2)}$出发，沿方向$x^{(2)}-x^{(1)}$作模式移动，令

$$y^{(1)}=x^{(2)}+\alpha(x^{(2)}-x^{(1)})。$$

以$y^{(1)}$为起点，沿坐标轴方向进行探测搜索，探测完毕得到的点仍记为$y^{(n+1)}$。

若$f(y^{(n+1)})<f(x^{(2)})$，则表明此次模式移动成功，得新的基点$x^{(3)}=y^{(n+1)}$，再沿$x^{(3)}-x^{(2)}$进行模式移动。

若$f(y^{(n+1)})\geq f(x^{(2)})$，则表明此次模式移动失败，退回$x^{(2)}$，减少步长$\delta$，再从$x^{(2)}$出发，沿坐标轴方向进行探测移动，如此下去，直到$\delta<\varepsilon$为此。

步骤：

1.给定初始点$x^{(1)}$，n个坐标方向$e_1,e_2,\cdots,e_n$，初始步长δ，加速因子$\alpha \ge 1$，缩减率$\beta \in (0,1)$，允许误差$\varepsilon > 0$，置$y^{(1)} = x^{(1)}$，$k = 1, j = 1$。

2.如果$f(y^{(j)} + \delta e_j) < f(y^{(j)})$，则令$y^{(j+1)} = y^{(j)} + \delta e_j$，转4；否则转3。

3.如果$f(y^{(j)} - \delta e_j) < f(y^{(j)})$，则令$y^{(j+1)} = y^{(j)} - \delta e_j$，转4；否则令$y^{(j+1)} = y^{(j)}$,转4。

4.如果$j < n$，则置$j = j + 1$，转2；否则，转5。

5.若$f(y^{(n+1)}) < f(x^{(k)})$，则转6；否则，转7。

6.置$x^{(k+1)} = y^{(n+1)}$，令$y^{(1)} = x^{(k+1)} + \alpha(x^{(k+1)} - x^{(k)})$，置$k = k+1$，$j = 1$，转2。

7.如果$\delta \le \varepsilon$，则停止迭代，得点$x^{(k)}$；否则，置$\delta = \beta\delta$，$y^{(1)} = x^{(k)}$，置$k = k+1, j = 1$，转2。

$$\min f(x) = (1 - x_1)^2 + 5(x_2 - x_1^2)^2$$

$$x^{(1)} = (2,0)^T, e_1 = (1,0)^T, e_2 = (0,1)^T, \delta = \beta = \frac{1}{2}, \alpha = 1.$$

i	$y^{(i)}$	$f(y^{(i)})$	$y^{(i)} + \delta e_i$	$f(y^{(i)} + \delta e_i)$	$y^{(i)} - \delta e_i$	$f(y^{(i)} - \delta e_i)$
1	$(2,0)^T$	81	$\left(\frac{5}{2},0\right)^T$	$197\frac{9}{16}$(失败)	$\left(\frac{3}{2},0\right)^T$	$25\frac{9}{16}$(成功)
2	$\left(\frac{3}{2},0\right)^T$	$25\frac{9}{16}$	$\left(\frac{3}{2},\frac{1}{2}\right)^T$	$15\frac{9}{16}$(成功)		
3	$\left(\frac{3}{2},\frac{1}{2}\right)^T$	$15\frac{9}{16}$				

$\because f(y^{(3)}) < f(x^{(1)}) = f(y^{(1)}), \therefore$ 第一轮探测结束，得基点

$$x^{(2)} = y^{(3)} = \left(\frac{3}{2}, \frac{1}{2}\right)^T。$$

沿$x^{(2)}-x^{(1)}$方向进行模式移动，令

$$y^{(1)}=x^{(2)}+\alpha\left(x^{(2)}-x^{(1)}\right)=(1,1)^T.$$

i	$y^{(i)}$	$f(y^{(i)})$	$y^{(i)}+\delta e_i$	$f(y^{(i)}+\delta e_i)$	$y^{(i)}-\delta e_i$	$f(y^{(i)}-\delta e_i)$
1	$(1,1)^T$	0	$\left(\frac{3}{2},1\right)^T$	$8\frac{1}{16}$(失败)	$\left(\frac{1}{2},1\right)^T$	$3\frac{1}{10}$(失败)
2	$(1,1)^T$	0	$\left(1,\frac{3}{2}\right)^T$	$1\frac{1}{4}$(失败)	$\left(1,\frac{1}{2}\right)^T$	$1\frac{1}{4}$(失败)
3	$(1,1)^T$					

$\because f(y^{(3)})=0<f(x^{(2)}),\therefore$ 模式移动成功，得基点

$$x^{(3)}=y^{(3)}=(1,1)^T。$$

从$x^{(3)}$出发，沿$x^{(3)}-x^{(2)}$方向进行模式移动，令

$$y^{(1)}=x^{(3)}+\alpha\left(x^{(3)}-x^{(2)}\right)=\left(\frac{1}{2},\frac{3}{2}\right)^T.$$

i	$y^{(i)}$	$f(y^{(i)})$	$y^{(i)}+\delta e_i$	$f(y^{(i)}+\delta e_i)$	$y^{(i)}-\delta e_i$	$f(y^{(i)}-\delta e_i)$
1	$\left(\frac{1}{2},\frac{3}{2}\right)^T$	8.06	$\left(1,\frac{3}{2}\right)^T$	1.25(成功)		
2	$\left(1,\frac{3}{2}\right)^T$	1.25	$(1,2)^T$	5(失败)	$(1,1)^T$	0(成功)
3	$(1,1)^T$	0				

$\because f(y^{(3)})=0=f(x^{(3)}),\therefore$ 模式移动失败，退回基点$x^{(3)}$，减少步长$\delta=\beta\delta=\dfrac{1}{4}$。

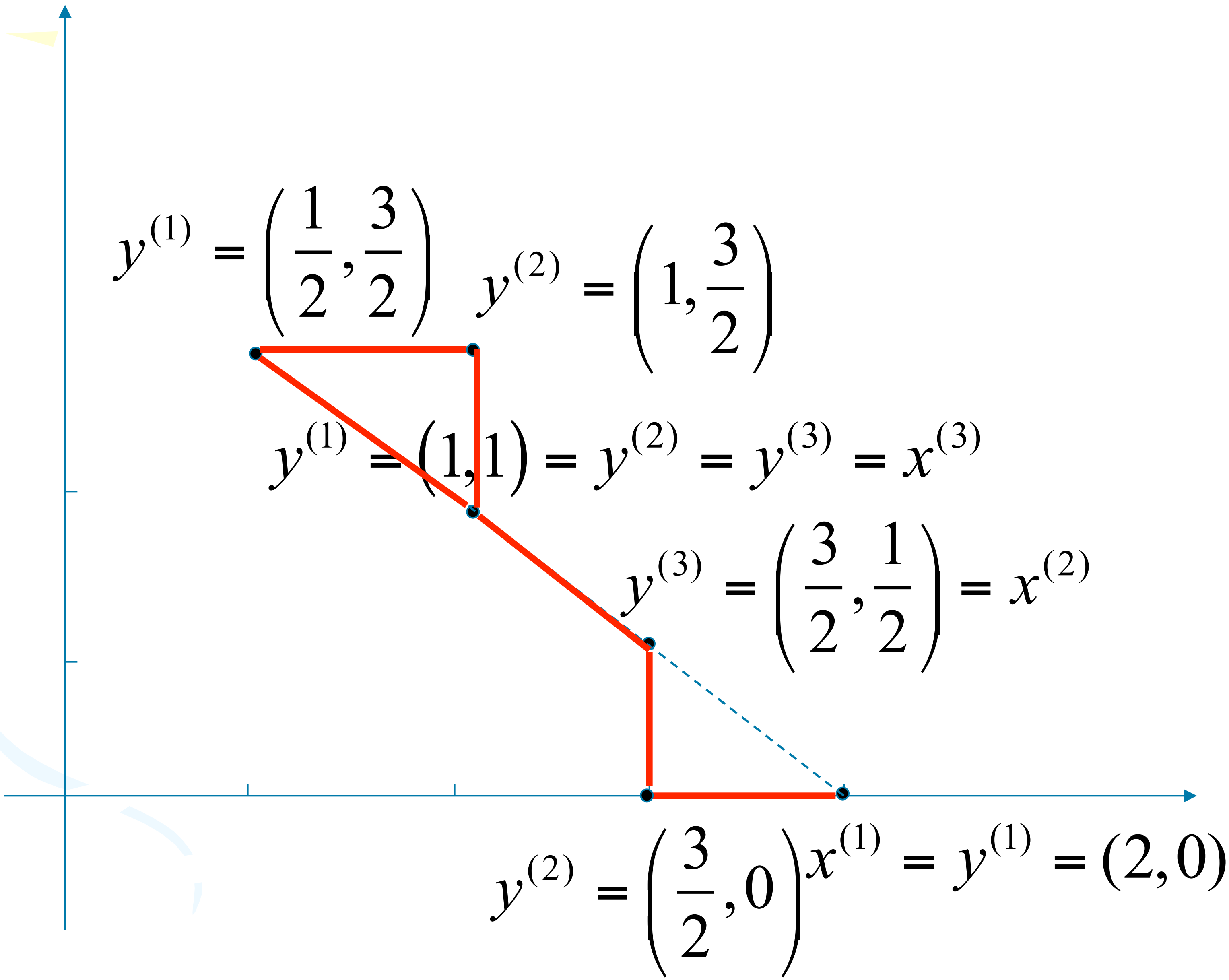
令$y^{(1)}=x^{(3)}=(1,1)^T$，从新探测。

i	$y^{(i)}$	$f(y^{(i)})$	$y^{(i)}+\delta e_i$	$f(y^{(i)}+\delta e_i)$	$y^{(i)}-\delta e_i$	$f(y^{(i)}-\delta e_i)$
1	$(1,1)^T$	0	$\left(\frac{5}{4},1\right)^T$	0.38(失败)	$\left(\frac{3}{4},1\right)^T$	1.02(失败)
2	$(1,1)^T$	0	$\left(1,\frac{5}{4}\right)^T$	0.31(失败)	$\left(1,\frac{3}{4}\right)^T$	0.31(失败)
3	$(1,1)^T$	0				

$\because f(y^{(3)})=0=f(x^{(3)}),\therefore$探测移动失败，退回基点

$x^{(3)}$，减少步长$\delta=\beta\delta=\dfrac{1}{8}$，再从$y^{(1)}$开始，探测仍然

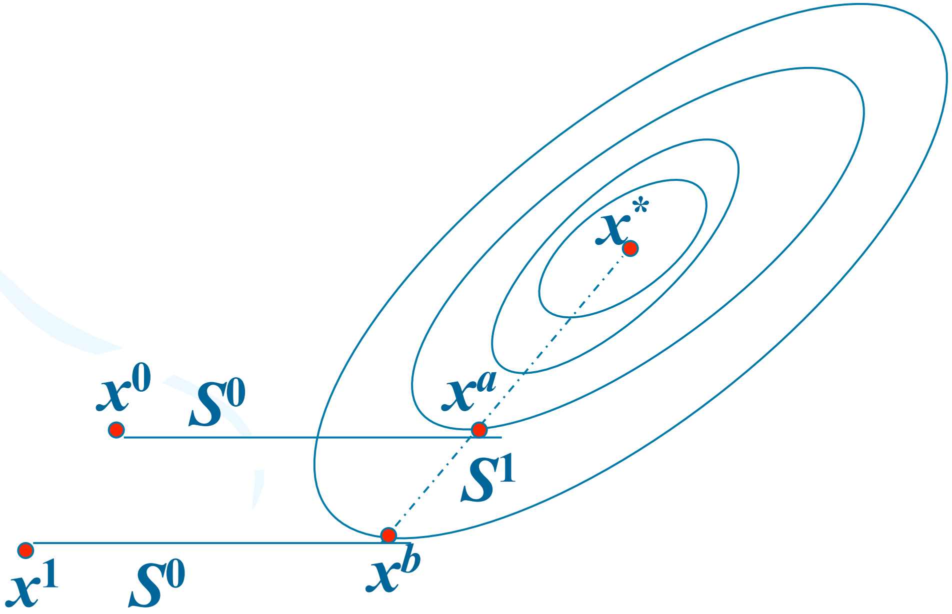
失败，且$\delta<\varepsilon$，所以$x^{(3)}=(1,1)^T$为原问题的最优解。


$y^{(1)} = \left(\frac{1}{2}, \frac{3}{2}\right)$
$y^{(2)} = \left(1, \frac{3}{2}\right)$
$y^{(1)} = (1,1) = y^{(2)} = y^{(3)} = x^{(3)}$
$y^{(3)} = \left(\frac{3}{2}, \frac{1}{2}\right) = x^{(2)}$
$y^{(2)} = \left(\frac{3}{2}, 0\right)$
$x^{(1)} = y^{(1)} = (2,0)$

- 特点：收敛速度比较慢，但编制程序比较简单，对变量不多的问题可以使用，而且确实是一种可靠的方法，另外，此法可用于求解非线性目标规划问题。

Powell方法（方向加速法）

$$f(x) = \frac{1}{2} x^T A x + b^T x + c$$



定理：设$f(x)=\frac{1}{2}x^T Ax+b^T x+c$，$A$对称正定，任意给定方向$d\in E^n$和点$x^{(0)},x^{(1)}\in E^n(x^{(0)}\neq x^{(1)})$,从$x^{(0)}$出发沿$d$做一维搜索得极小点$x^{(a)}$，从$x^{(1)}$出发沿$d$做一维搜索得极小点$x^{(b)}$，则$x^{(b)}-x^{(a)}$与$d$是$A$共轭的。

证明：由一维搜索的性质，得

$$\left.\begin{aligned}\nabla f(x^{(a)})^T d=0\\ \nabla f(x^{(b)})^T d=0\end{aligned}\right\}\Rightarrow\begin{cases}(Ax^{(a)}+b)^T d=0\\ (Ax^{(b)}+b)^T d=0\end{cases}$$

$$\Rightarrow (x^{(b)}-x^{(a)})^T Ad=0$$

$\therefore x^{(b)}-x^{(a)}$与$d$是$A$共轭的。

定理：

设$f(x)=\frac{1}{2}x^T A x+b^T x+c, A_{n\times n}$对称正定，如果从$x^{(0)}$出发，在沿着$k(k\le n)$个共轭方向$d^{(1)}$，$d^{(2)},\cdots$，$d^{(k)}$进行一维搜索后找到$x^{(a)}$；从$x^{(1)}$出发，在沿着相同方向进行一维搜索后找到$x^{(b)}$，则$(x^{(b)}-x^{(a)})$对$k$个方向共轭，即

$$(x^{(b)}-x^{(a)})^T A d^{(i)}=0,\quad i=1,2,\cdots,k.$$

证明：由于$d^{(1)}$，$d^{(2)},\cdots$，$d^{(k)}$是一组关于A共轭的向量，$x^{(a)}$为由$x^{(0)}$出发，在沿着$d^{(1)}$，$d^{(2)},\cdots$，$d^{(k)}$进行一维搜索后得到的点，所以有

$$\nabla f\left(x^{(a)}\right)^T d^{(i)} = 0, i = 1, 2, \cdots, k.$$

同理，有

$$\nabla f\left(x^{(b)}\right)^T d^{(i)} = 0, i = 1, 2, \cdots, k.$$

$$\Rightarrow \left(\nabla f\left(x^{(b)}\right) - \nabla f\left(x^{(a)}\right)\right)^T d^{(i)} = 0$$

$$\because \nabla f\left(x\right) = Ax + b$$

$$\therefore \left(x^{(b)} - x^{(a)}\right)^T A d^{(i)} = 0, i = 1, 2, \cdots, k.$$

Powell算法的基本思想

整个过程分成若干个循环，每个循环有$n+1$个一维搜索，即先经过n个线性无关方向的一维搜索，把所得的点与此循环开始点连接起来，再沿着连线的方向进行第$n+1$次一维搜索，然后用连线的方向代替原来n个方向中的一个，再开始下一个循环。

步骤：

1.给定初始点$x^{(0)}$，n个线性无关的方向

$$d^{(1,1)},d^{(1,2)},\cdots,d^{(1,n)},$$

允许误差$\varepsilon>0$，置$k=1$。

2.置$x^{(k,0)}=x^{(k-1)}$，从$x^{(k,0)}$出发，依次沿方向

$$d^{(k,1)},d^{(k,2)},\cdots,d^{(k,n)}$$

进行一维搜索，得到点$x^{(k,1)},x^{(k,2)},\cdots,x^{(k,n)}$，

再从$x^{(k,n)}$出发，沿方向$d^{(k,n+1)}=x^{(k,n)}-x^{(k,0)}$

作一维搜索，得到点$x^{(k)}$。

3.若$\left\|x^{(k)}-x^{(k-1)}\right\|<\varepsilon$，则停止计算，得点$x^{(k)}$；否则令

$$d^{(k+1,j)}=d^{(k,j+1)},j=1,2,\cdots,n$$

置$k:=k+1$，返回2。

例 $\min f(x)=\frac{3}{2}x_1^2+\frac{1}{2}x_2^2-x_1x_2-2x_1$

初点$x^{(0)}=(-2,4)^T, d^{(1,1)}=(1,0)^T, d^{(1,2)}=(0,1)^T$

第一轮迭代： 从$x^{(0)}$出发，$x^{(1,0)}=x^{(0)}=(-2,4)^T$

先沿$d^{(1,1)}$作一维搜索:$\min f(x^{(1,0)}+\lambda d^{(1,1)})$

在$\lambda=4$达到极小，所以$x^{(1,1)}=x^{(1,0)}+4d^{(1,1)}=(2,4)^T$.

从$x^{(1,1)}$出发，沿$d^{(1,2)}$作一维搜索:$\min f(x^{(1,1)}+\lambda d^{(1,2)})$

在$\lambda=$-2达到极小，所以$x^{(1,2)}=x^{(1,1)}-2d^{(1,2)}=(2,2)^T$.

从$x^{(1,2)}$出发，沿$d^{(1,3)}=x^{(1,2)}-x^{(1,0)}=(4,-2)^T$作一维搜索：

$$\min f(x^{(1,2)}+\lambda d^{(1,3)})$$

在$\lambda=-\frac{2}{17}$达到极小，所以$x^{(1)}=x^{(1,2)}-\frac{2}{17}d^{(1,3)}=\left(\frac{26}{17},\frac{38}{17}\right)^T$.

第二轮迭代：

初点$x^{(2,0)}=x^{(1)}=\left(\frac{26}{17},\frac{38}{17}\right)^T, d^{(2,1)}=(0,1)^T, d^{(2,2)}=(4,-2)^T$

先沿$d^{(2,1)}$作一维搜索：$\min f(x^{(2,0)}+\lambda d^{(2,1)})$

在$\lambda=-\frac{12}{17}$达到极小，所以$x^{(2,1)}=\left(\frac{26}{17},\frac{26}{17}\right)^T$.

从$x^{(2,1)}$出发，沿$d^{(2,2)}$作一维搜索：$\min f(x^{(2,1)}+\lambda d^{(2,2)})$

在$\lambda=-\frac{18}{289}$达到极小，所以$x^{(2,2)}=\left(\frac{370}{289},\frac{478}{289}\right)^T$.

沿$d^{(2,3)}=x^{(2,2)}-x^{(2,0)}=\left(-\frac{72}{289},-\frac{168}{289}\right)$作一维搜索：

$$\min f(x^{(2,2)}+\lambda d^{(2,3)})$$

在$\lambda=\frac{9}{8}$达到极小，所以$x^{(2)}=(1,1)^T$.

第三轮迭代：

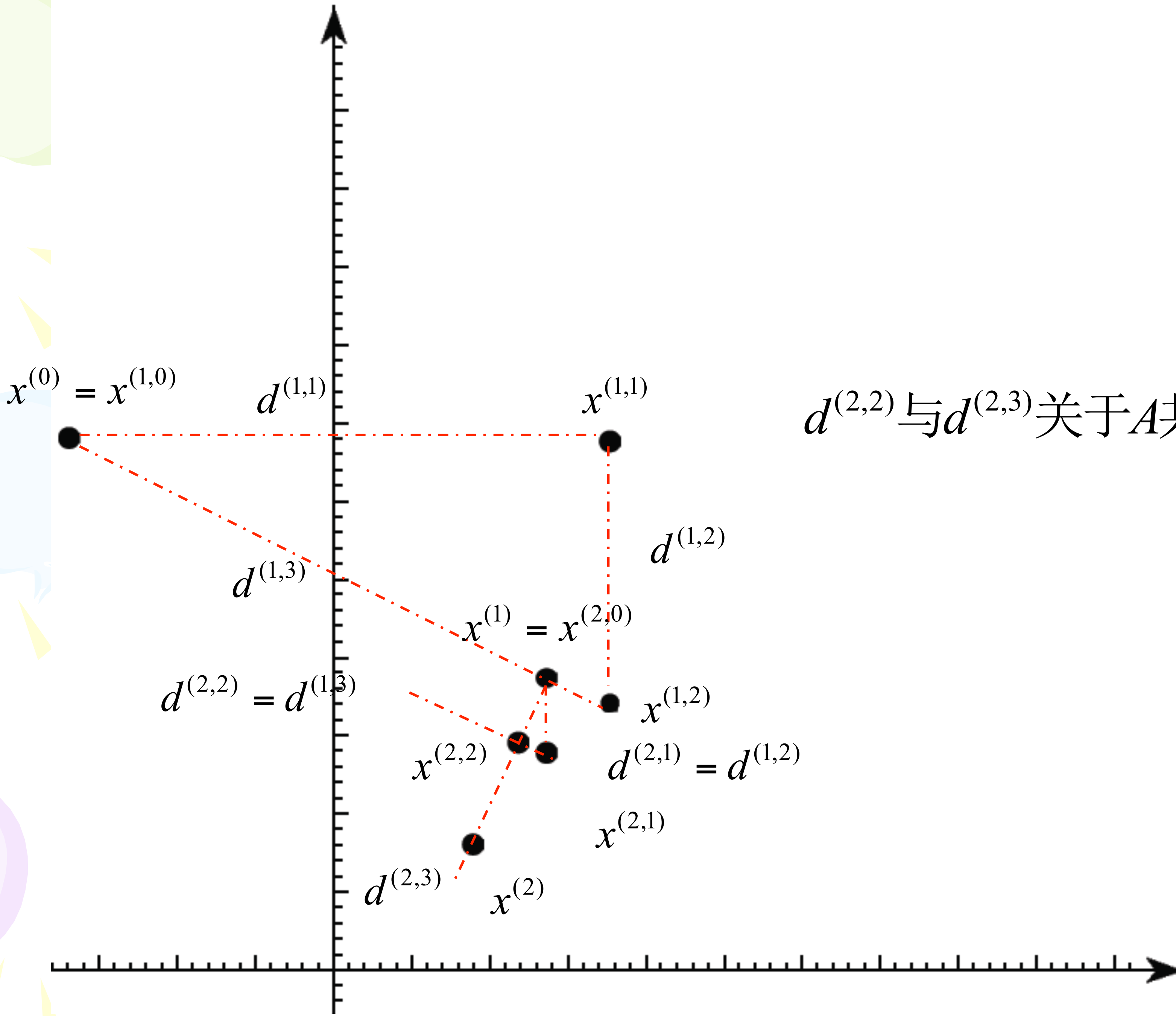
初点$x^{(3,0)}=x^{(2)}=(1,1)^T$,

$$d^{(3,1)}=(4,-2)^T, d^{(3,2)}=d^{(2,3)}=\left(-\frac{72}{289},-\frac{168}{289}\right)^T$$

此时，$x^{(3,1)}=x^{(3,2)}=x^{(3,3)}, x^{(3)}=x^{(3,0)}$,

终止。

最优解：$x^*=(1,1)^T$.



$d^{(2,2)}$与$d^{(2,3)}$关于A共轭

Powell算法的二次终止性

定理：对二次正定函数

$$f(x)=\frac{1}{2}x^TAx+b^Tx+c$$

如果每轮迭代中前n个方向均线性无关，那么Powell方法至多经n轮迭代达到极小点。

证明：第一轮迭代中，初始点$x^{(1,0)}=x^{(0)}$,搜索方向

$$d^{(1,1)},d^{(1,2)},\mathrm{L}\ \ d^{(1,n)},d^{(1,n+1)}=x^{(1,n)}-x^{(1,0)},$$

依次沿$n+1$个方向搜索，最后得点$x^{(1)}$.

第二轮迭代中，初始点$x^{(2,0)}=x^{(1)}$,搜索方向

$$d^{(2,1)},d^{(2,2)},\mathrm{L}\ \ d^{(2,n)},d^{(2,n+1)}=x^{(2,n)}-x^{(2,0)},$$

前n个方向依次为

$$d^{(1,2)},\mathrm{L}\ \ d^{(1,n)},d^{(1,n+1)}=x^{(1,n)}-x^{(1,0)},$$

设这n个方向线性无关，第二轮迭代最后得点$x^{(2)}$.

点$x^{(2,0)}(=x^{(1)})$和点$x^{(2,n)}$是从不同点出发

沿同一方向$d^{(2,n)}(=d^{(1,n+1)})$进行搜索得到.

由定理，方向

$$d^{(2,n)}(=d^{(1,n+1)})\text{与}d^{(2,n+1)}=x^{(2,n)}-x^{(2,0)}$$

关于A共轭.

假设第k轮迭代的搜索方向为

$$d^{(k,1)},d^{(k,2)},\cdots\ d^{(k,n)},d^{(k,n+1)}=x^{(k,n)}-x^{(k,0)},$$

前n个方向线性无关, 且后k个方向

$$d^{(k,n-k+2)},\cdots\ d^{(k,n)},d^{(k,n+1)}=x^{(k,n)}-x^{(k,0)}$$

是A共轭的，经过本轮迭代后得点$x^{(k)}$.

在第$k+1$轮迭代中，$x^{(k+1,0)}=x^{(k)}$, 搜索方向为

$$d^{(k+1,1)},d^{(k+1,2)},\cdots d^{(k+1,n)},d^{(k+1,n+1)}=x^{(k+1,n)}-x^{(k+1,0)},$$

它们分别等于

$$d^{(k,2)},\cdots d^{(k,n+1)},d^{(k+1,n+1)}=x^{(k+1,n)}-x^{(k+1,0)},$$

其中前n个方向线性无关.

由于点$x^{(k+1,n)}$和点$x^{(k+1,0)}$是从不同点出发沿同一组共轭方向进行进行一维搜索得到，因此方向组$d^{(k+1,n-k+1)},\cdots,d^{(k+1,n)},d^{(k+1,n+1)}$是$A$共轭的.

所以，在完成n个阶段的迭代后，必能得到n个A共轭的方向，由此，Powell方法具有二次终止性.

$$\min f(x) = (x_1 - x_2 + x_3)^2 + (-x_1 + x_2 + x_3)^2 + (x_1 + x_2 - x_3)^2$$

最优解为：$x^* = (0, 0, 0)^T$

取$x^{(0)} = \left(\frac{1}{2}, 1, \frac{1}{2}\right)^T, d^{(1,1)} = (1,0,0)^T,$

$d^{(1,2)} = (0,1,0)^T, d^{(1,3)} = (0,0,1)^T$

第一轮迭代：

i	$d^{(1,i)}$	$x^{(1,i)}$	$f(x^{(1,i)})$
0		$\left(\frac{1}{2},1,\frac{1}{2}\right)^T$	2
1	$d^{(1,1)}$	$\left(\frac{1}{2},1,\frac{1}{2}\right)^T$	2
2	$d^{(1,2)}$	$\left(\frac{1}{2},\frac{1}{3},\frac{1}{2}\right)^T$	$\frac{2}{3}$
3	$d^{(1,3)}$	$\left(\frac{1}{2},\frac{1}{3},\frac{5}{18}\right)^T$	$\frac{42}{81}$
4	$\left(0,-\frac{2}{3},-\frac{2}{9}\right)^T$	$\left(\frac{1}{2},\frac{1}{4},\frac{1}{4}\right)^T$	

第二轮迭代：

$$d^{(2,1)}=(0,1,0)^T, d^{(2,2)}=(0,0,1)^T,$$

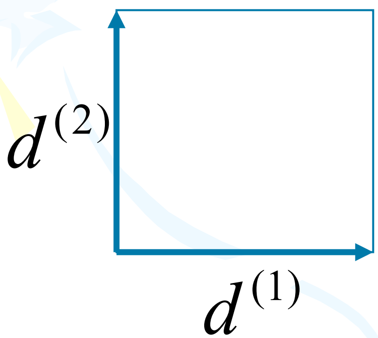
$$d^{(2,3)}=\left(0,-\frac{2}{3},-\frac{2}{9}\right)^T$$

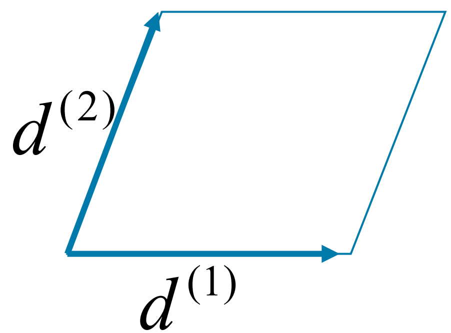
这三个搜索方向的第一个分量都是**0**，所以沿这些方向进行一维搜索得到的点的第一个分量将保持为**1/2**，这样就达不到真正的最优解，原因是$d^{(2,1)}$，$d^{(2,2)}$，$d^{(2,3)}$线性相关。

正交程度与共轭程度

- 正交程度
- 考虑特殊正定二次函数

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





$$\delta=\left\|d^{(1)}\times d^{(2)}\right\|=\left|\det\left(d^{(1)},d^{(2)}\right)\right|$$

定义：设$d^{(1)},d^{(2)},\cdots,d^{(n)}$是$R^n$中的$n$个向量，其正交程度定义为

$$\delta\left(d^{(1)},d^{(2)},\cdots,d^{(n)}\right)=\begin{cases}0, & \exists d^{(i)}=0,\\ \dfrac{\left|\det\left(d^{(1)},d^{(2)},\cdots,d^{(n)}\right)\right|}{\prod_{i=1}^{n}\left\|d^{(i)}\right\|}, & \text{其它}.\end{cases}$$

定理1 $\delta\left(d^{(1)}, d^{(2)}, \cdots, d^{(n)}\right) \leq 1$, 且等式成立当且仅当$d^{(1)}, d^{(2)}, \cdots, d^{(n)}$是$n$个非零的正交向量.

证明：不妨假设$d^{(i)} \neq 0, d^{(1)}, d^{(2)}, \cdots, d^{(n)}$线性无关且$\left\| d^{(i)} \right\| = 1, i = 1, 2, \cdots, n.$

则$\delta\left(d^{(1)}, d^{(2)}, \cdots, d^{(n)}\right) = \left|\det\left(d^{(1)}, d^{(2)}, \cdots, d^{(n)}\right)\right|.$

令$Q = \left(d^{(1)}, d^{(2)}, \cdots, d^{(n)}\right),$

则$Q^T Q$是正定对称矩阵，所以，有

$$\left|\det(Q)\right|^2 = \det\left(Q^T Q\right) \leq \prod_{i=1}^{n}\left(d^{(i)}\right)^T d^{(i)} = 1,$$

等式成立当且仅当$Q^T Q$为对角阵，即$d^{(1)}, d^{(2)}, \cdots, d^{(n)}$正交.

考虑一般正定二次函数

$$f(x)=\frac{1}{2}x^TAx+b^Tx+c$$

设$d^{(1)},d^{(2)},\cdots,d^{(n)}$为一组关于$A$共轭的向量

令 $\bar{d}^{(i)}=\sqrt{A}d^{(i)},i=1,2,\cdots,n$

则正交方向对应于共轭方向

定义：设A为n阶正定对称矩阵，$d^{(1)},d^{(2)},\cdots,d^{(n)}$是$R^n$中的$n$个向量，其$A$共轭程度定义为

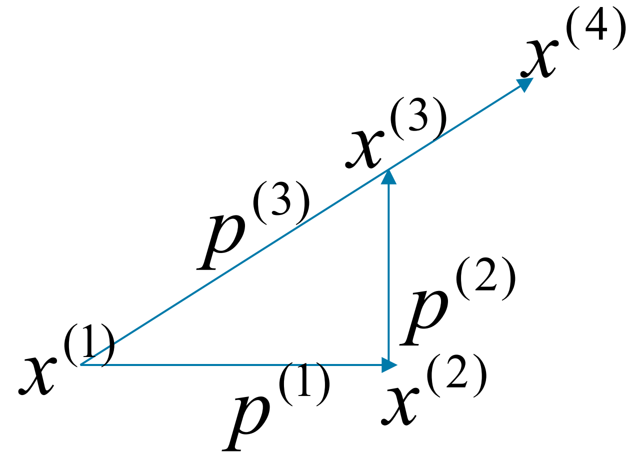
$$\Delta\left(d^{(1)},d^{(2)},\cdots,d^{(n)}\right)$$

$$=\begin{cases} 0, & \exists d^{(i)}=0, \\ \dfrac{\left|\det\left(\sqrt{A}\right)\right|\left|\det\left(d^{(1)},d^{(2)},\cdots,d^{(n)}\right)\right|}{\prod_{i=1}^{n}\sqrt{\left(d^{(i)}\right)^T A d^{(i)}}}, & \text{其它}. \end{cases}$$

定理2　$\Delta\left(d^{(1)},d^{(2)},\cdots,d^{(n)}\right)\le 1$,

且等式成立当且仅当$d^{(1)},d^{(2)},\cdots,d^{(n)}$是$n$个非零的$A$共轭向量.

$$f(x)=\frac{1}{2}x^T Ax+b^T x+c$$



问题：如何从$p^{(1)},p^{(2)}$和

$p^{(3)}=x^{(3)}-x^{(1)}$中，找出

两个共轭程度最大的方向？

分别取$p^{(1)},p^{(2)},p^{(3)}$上的方向$d^{(1)},d^{(2)},d^{(3)}$，

满足 $d^{(i)^T}Ad^{(i)}=1, i=1,2,3.$

$$\Delta=\Delta\left(d^{(1)},d^{(2)}\right)=\left|\det\left(\sqrt{A}\right)\right|\left|\det\left(d^{(1)},d^{(2)}\right)\right|$$

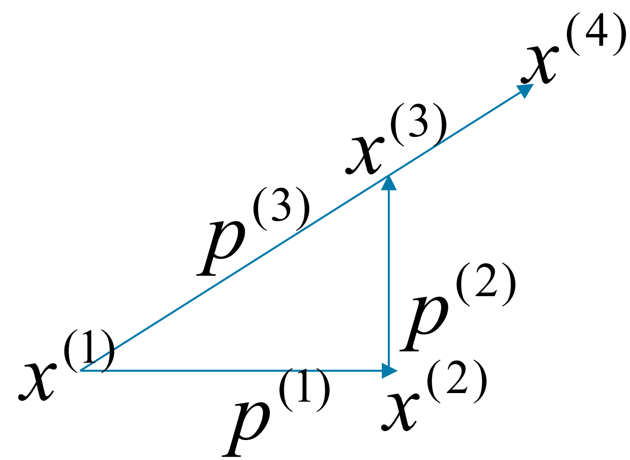
$$\Delta_1=\Delta\left(d^{(2)},d^{(3)}\right)=\left|\det\left(\sqrt{A}\right)\right|\left|\det\left(d^{(2)},d^{(3)}\right)\right|$$

$$\Delta_2=\Delta\left(d^{(3)},d^{(1)}\right)=\left|\det\left(\sqrt{A}\right)\right|\left|\det\left(d^{(3)},d^{(1)}\right)\right|$$

由于$x^{(3)}$是从$x^{(1)}$出发，沿方向$d^{(1)},d^{(2)}$进行一维搜索得到的，所以有

$$x^{(3)}-x^{(1)}=x^{(3)}-x^{(2)}+x^{(2)}-x^{(1)}=\lambda_1 d^{(1)}+\lambda_2 d^{(2)}$$

$$\because x^{(3)}-x^{(1)}=\alpha_3 d^{(3)}$$



$$\Rightarrow d^{(3)}=\frac{\lambda_1}{\alpha_3}d^{(1)}+\frac{\lambda_2}{\alpha_3}d^{(2)}$$

$$\Delta_1 = \Delta\left(d^{(2)}, d^{(3)}\right) = \left|\det\left(\sqrt{A}\right)\right| \left|\det\left(d^{(2)}, \frac{\lambda_1}{\alpha_3} d^{(1)} + \frac{\lambda_2}{\alpha_3} d^{(2)}\right)\right|$$

$$= \left|\frac{\lambda_1}{\alpha_3}\right| \Delta\left(d^{(2)}, d^{(1)}\right) = \left|\frac{\lambda_1}{\alpha_3}\right| \Delta$$

$$\Delta_2 = \Delta\left(d^{(3)}, d^{(1)}\right) = \left|\det\left(\sqrt{A}\right)\right| \left|\det\left(\frac{\lambda_1}{\alpha_3} d^{(1)} + \frac{\lambda_2}{\alpha_3} d^{(2)}, d^{(1)}\right)\right|$$

$$= \left|\frac{\lambda_2}{\alpha_3}\right| \Delta\left(d^{(2)}, d^{(1)}\right) = \left|\frac{\lambda_2}{\alpha_3}\right| \Delta$$

求λ_1, λ_2和α_3的值

由于$x^{(2)}$是从$x^{(1)}$出发，经一维搜索得到，

$$\therefore x^{(2)} = x^{(1)} + \lambda_1 d^{(1)} \text{且} \nabla f\left(x^{(2)}\right)^T d^{(1)} = 0.$$

将$f(x)$在$x^{(2)}$点处$Taylor$展开，有

$$f\left(x^{(1)}\right) = f\left(x^{(2)}\right) + \nabla f\left(x^{(2)}\right)^T \left(x^{(1)} - x^{(2)}\right)$$

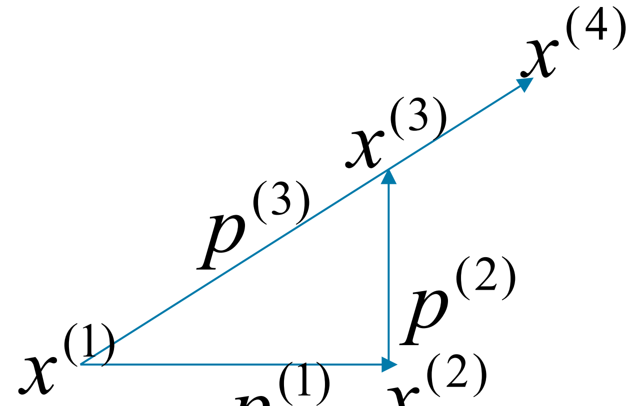
$$+ \frac{1}{2}\left(x^{(1)} - x^{(2)}\right)^T A \left(x^{(1)} - x^{(2)}\right)$$

$$= f\left(x^{(2)}\right) + \nabla f\left(x^{(2)}\right)^T \left(-\lambda_1 d^{(1)}\right) + \frac{1}{2}\lambda_1^2 \left(d^{(1)}\right)^T A d^{(1)}$$

$$= f\left(x^{(2)}\right) + \frac{1}{2}\lambda_1^2$$

$$\Rightarrow \lambda_1^2 = 2\left[f\left(x^{(1)}\right) - f\left(x^{(2)}\right)\right].$$

同理 $\lambda_2^2 = 2\left[f\left(x^{(2)}\right) - f\left(x^{(3)}\right)\right].$



由于$x^{(4)}$是从$x^{(1)}$出发，沿方向$p^{(3)} = x^{(3)} - x^{(1)}$进行一维搜索得到的，所以有

$$x^{(4)} = x^{(1)} + \lambda_3 p^{(3)} = x^{(1)} + \lambda_3 \alpha_3 d^{(3)}$$

其中 $f\left(x^{(1)} + \lambda_3 \alpha_3 d^{(3)}\right) = \min f\left(x^{(1)} + \lambda d^{(3)}\right)$

$$\Rightarrow \alpha_3^2 = \frac{2\left[f\left(x^{(1)}\right) - f\left(x^{(4)}\right)\right]}{\lambda_3^2}.$$

当 $|\lambda_1|>|\alpha_3|$ 时，即

$$2\left[f\left(x^{(1)}\right)-f\left(x^{(2)}\right)\right]>\frac{2\left[f\left(x^{(1)}\right)-f\left(x^{(4)}\right)\right]}{\lambda_3^2}$$

$$\Leftrightarrow\quad |\lambda_3|>\left(\frac{f\left(x^{(1)}\right)-f\left(x^{(4)}\right)}{f\left(x^{(1)}\right)-f\left(x^{(2)}\right)}\right)^{\frac{1}{2}}$$

此时，用$d^{(3)}$替换方向$d^{(1)}$.

当 $|\lambda_3|>\left(\frac{f\left(x^{(1)}\right)-f\left(x^{(4)}\right)}{f\left(x^{(2)}\right)-f\left(x^{(3)}\right)}\right)^{\frac{1}{2}}$

此时，用$d^{(3)}$替换方向$d^{(2)}$.

令 $$\mu = \max\left\{f\left(x^{(i)}\right) - f\left(x^{(i+1)}\right) \mid i = 1,2,\cdots,n\right\}$$
$$= f\left(x^{(m)}\right) - f\left(x^{(m+1)}\right)$$

判别条件为

$$\left|\lambda_{n+1}\right| > \left(\frac{f\left(x^{(1)}\right) - f\left(x^{(n+2)}\right)}{\mu}\right)^{\frac{1}{2}}$$

此时，用$p^{(n+1)}$替换方向$p^{(m)}$.

改进的Powell法步骤：

1.给定初始点$x^{(0)}$，n个线性无关的方向

$$d^{(1,1)},d^{(1,2)},\cdots,d^{(1,n)},$$

允许误差$\varepsilon > 0$，置$k = 1$。

2.置$x^{(k,0)} = x^{(k-1)}$，从$x^{(k,0)}$出发，依次沿方向

$$d^{(k,1)},d^{(k,2)},\cdots,d^{(k,n)}$$

进行一维搜索，得到点$x^{(k,1)},x^{(k,2)},\cdots,x^{(k,n)}$，

求指标m，使得

$$f(x^{(k,m-1)})-f(x^{(k,m)}) = \max_{j=1,\cdots,n}\left\{f(x^{(k,j-1)})-f(x^{(k,j)})\right\}$$

令$d^{(k,n+1)} = x^{(k,n)} - x^{(k,0)}$，若$\left\|x^{(k,n)} - x^{(k,0)}\right\| \le \varepsilon$,则

停止迭代；否则，转3。

3.求λ_{n+1}，使得

$$f(x^{(k,0)}+\lambda_{n+1}d^{(k,n+1)})=\min_{\lambda} f(x^{(k,0)}+\lambda d^{(k,n+1)})$$

令 $$x^{(k+1,0)}=x^{(k)}=x^{(k,0)}+\lambda_{n+1}d^{(k,n+1)}$$

若$\left\|x^{(k)}-x^{(k-1)}\right\|\le\varepsilon$，则停止计算，得点$x^{(k)}$；否则转4。

4.若$\left|\lambda_{n+1}\right|>\left[\dfrac{f(x^{(k,0)})-f(x^{(k+1,0)})}{f(x^{(k,m-1)})-f(x^{(k,m)})}\right]^{\frac{1}{2}}$，则令

$$d^{(k+1,j)}=d^{(k,j)}\quad j=1,\cdots,m-1$$

$$d^{(k+1,j)}=d^{(k,j+1)}\quad j=m,\cdots,n.$$

置$k=k+1$，转2；否则，令$d^{(k+1,j)}=d^{(k,j)}$，$j=1,\cdots,n$，

置$k=k+1$，转2。