无约束最优化问题的直接方法

1. 模式搜索法

2. Powell 算法

3. 单纯形替换法

无约束最优化问题

$$\min \quad f(x);$$

直接方法：不用计算导数，只需计 算函数值的方法。

模式搜索法（$Hooke-Jeeves$ 方法,1961 年）

1. 基本思想：算法从初始基点开始，交替实施两种搜索：轴向搜索和模式搜索。轴向搜索依次沿 n 个坐标轴的方向进行，用来确定新的基点和有利于函数值下降的方向。模式搜索则沿着相邻两个基点的连线方向进行，试图使函数值下降更快。

2. 算法分析

令 $e_j=(0,\cdots,0,1,0,\cdots,0)^T$，$j=1,2,\cdots,n$ 表示 n 个坐标轴方向。

给定初始步长 δ，加速因子 α。任取初始点 $x^{(1)}$ 作为第一个基点。

以下用 $x^{(j)}$ 表示第 j 个基点。

在每一轮轴向搜索中，用$y^{(i)}$表示沿第i个坐标轴e_i方向搜索时的出发点。

轴向搜索:

令 $y^{(1)}=x^{(1)}$。

沿e_1方向搜索:

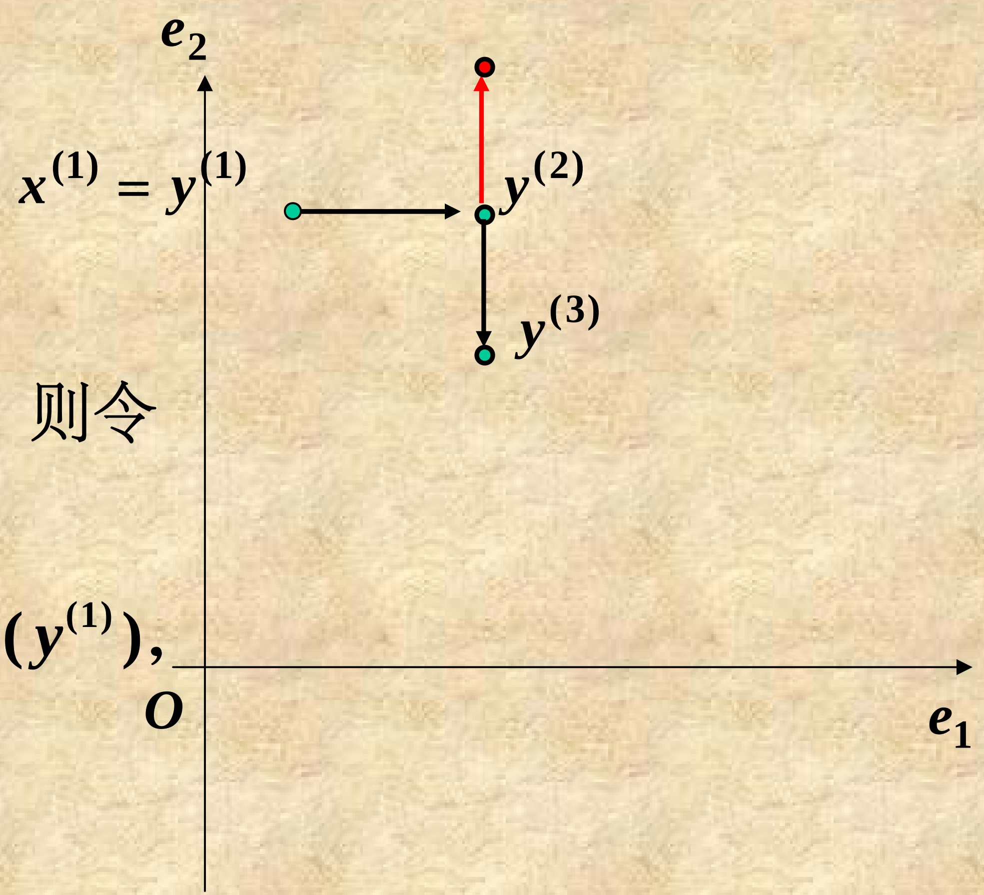
如果 $f(y^{(1)}+\delta e_1)<f(y^{(1)})$，则令

$$y^{(2)}=y^{(1)}+\delta e_1;$$

否则，如果 $f(y^{(1)}-\delta e_1)<f(y^{(1)})$,

则令 $y^{(2)}=y^{(1)}-\delta e_1;$

否则, 令 $y^{(2)}=y^{(1)}$。



再从$y^{(2)}$出发，仿上沿e_2进行搜索得到$y^{(3)}$，

依次进行搜索，直到得到点 $y^{(n+1)}$。（一轮轴向搜索结束。）

如果 $f(y^{(n+1)}) \ge f(x^{(1)})$,则缩小步长 δ，仍以 x^1为起点进行新的轴向搜索。　否则，进行模式搜索。

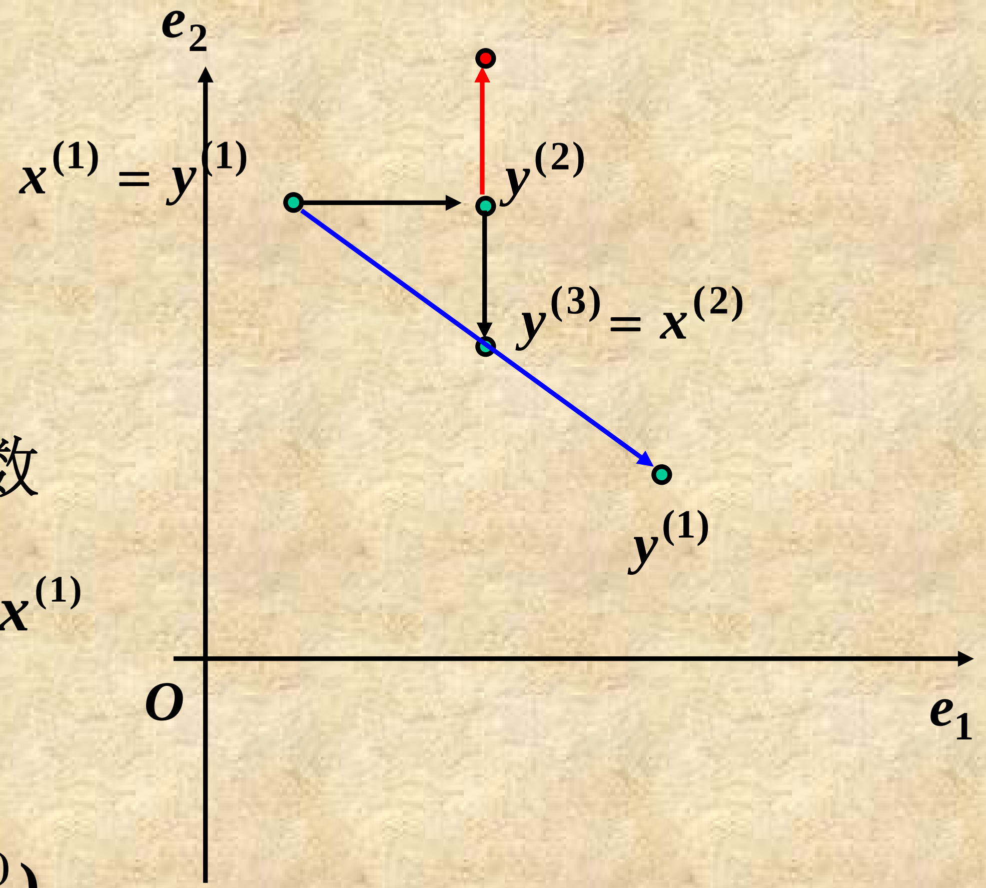
模式搜索:

如果 $f(y^{(n+1)}) < f(x^{(1)})$,

则令 $x^{(2)} = y^{(n+1)}$。

$x^{(2)} - x^{(1)}$方向可能有利于函数值下降，因此下一步沿 $x^{(2)} - x^{(1)}$方向进行模式搜索。

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



如何判断模式搜索是否 有效?

以 $y^{(1)}$ 为起点进行下一轮轴向 搜索，所得的点仍记为 $y^{(n+1)}$。

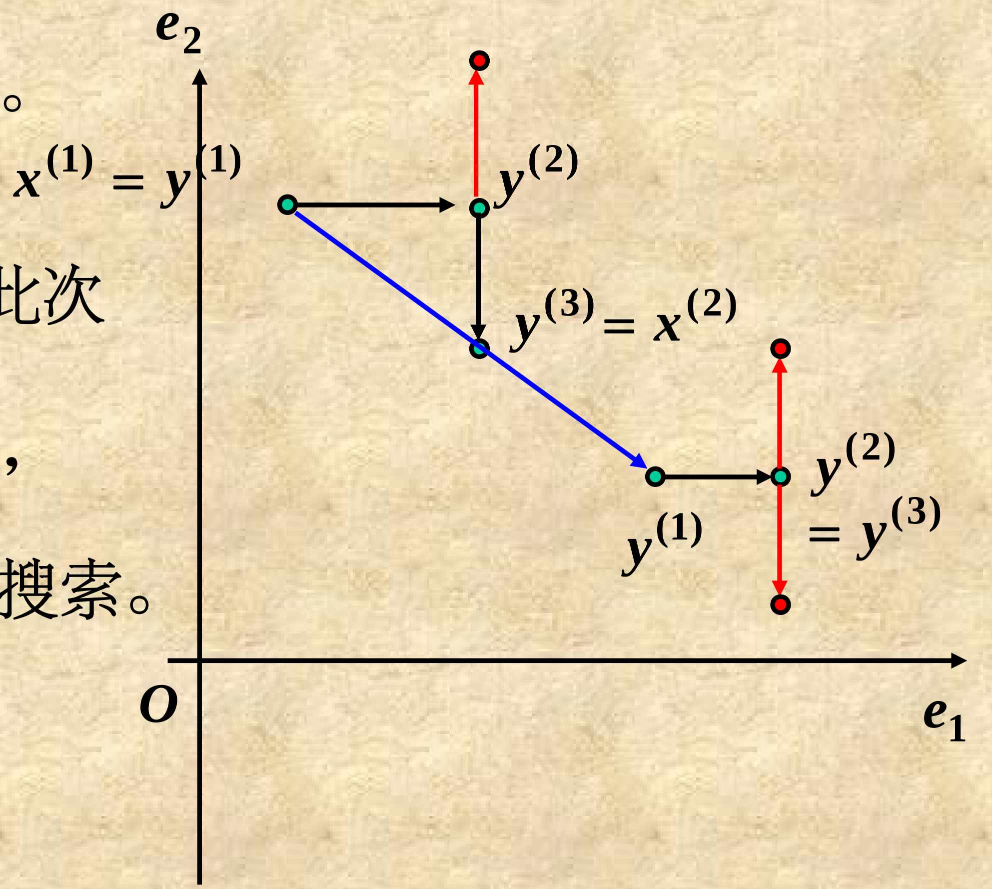
如果 $f(y^{(n+1)}) < f(x^{(2)})$,表明此次模式搜索成功 ，令

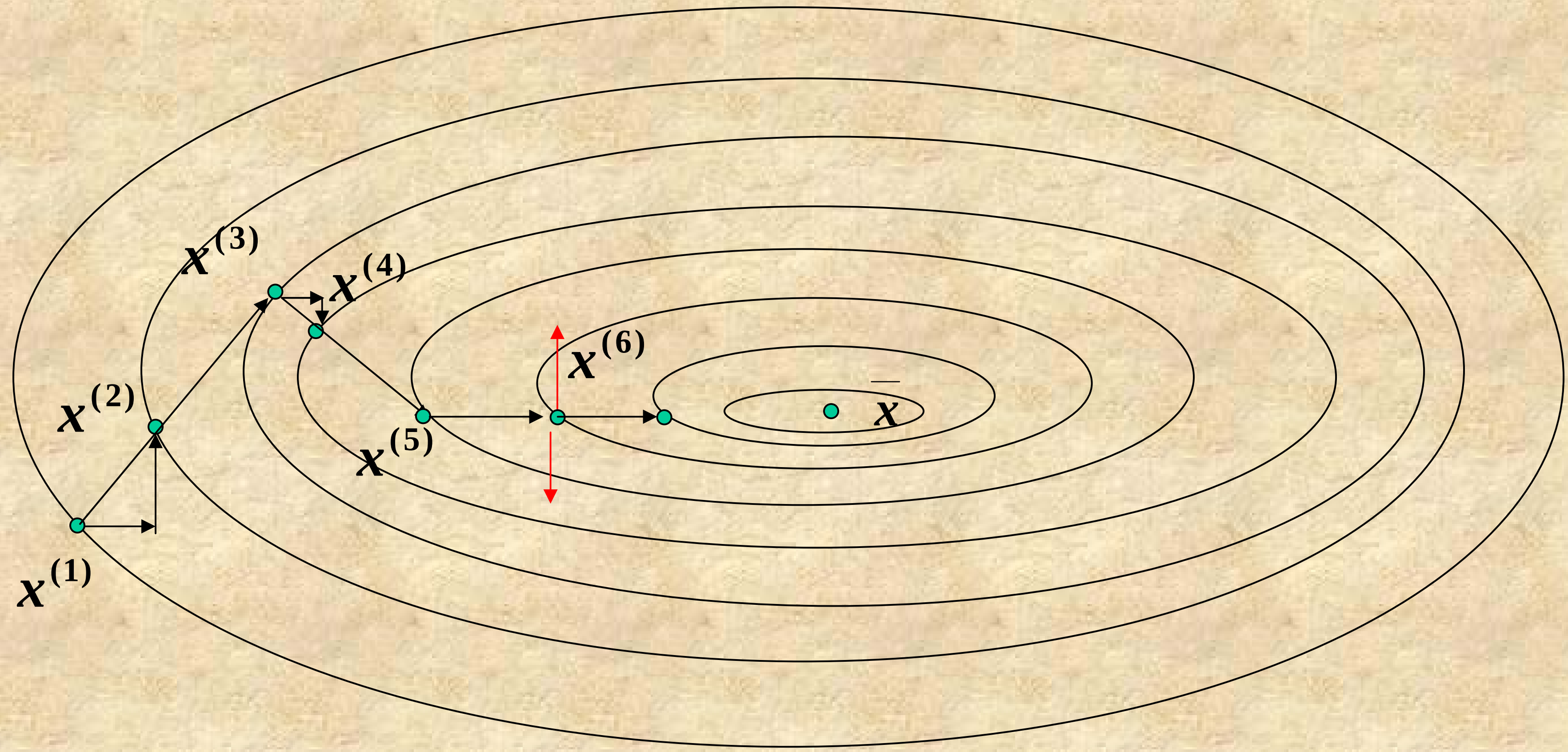
$x^{(3)} = y^{(n+1)}$。仿上继续进行迭代。

如果 $f(y^{(n+1)}) \geq f(x^{(2)})$,表明此次

模式搜索失败，返回基 点 $x^{(2)}$,

缩小步长，进行下一轮 轴向搜索。





$x^{(3)}$
$x^{(4)}$
$x^{(6)}$
$x^{(2)}$
$\bar{x}$
$x^{(5)}$
$x^{(1)}$

模式搜索法:

(1) 给定初始点 $x^{(1)} \in R^n$，初始步长 δ，加速因子 $\alpha \geq 1$，缩减率 $\beta \in (0,1)$，精度 $\varepsilon > 0$。令 $y^{(1)} = x^{(1)}, k = 1, j = 1$。

(2) 轴向搜索:

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

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

否则，令 $y^{(j+1)} = y^{(j)}$。

(3) 若 $j < n$，则令 $j := j + 1$，转(2)。

如果 $f(y^{(n+1)}) < f(x^{(k)})$，转(4)；否则，转(5)。

(4) 模式搜索:令 $x^{(k+1)} = y^{(n+1)}, y^{(1)} = x^{(k+1)} + \alpha(x^{(k+1)} - x^{(k)})$。

令 $k := k + 1, j = 1$，转（2）。

(5) 如果 $\delta \le \varepsilon$,停止，得到点 $x^{(k)}$; 否则，令 $\delta := \beta\delta$,

$y^{(1)} = x^{(k)}, x^{(k+1)} = x^{(k)}$。令 $k := k+1, j = 1$,转（2）。

注：将轴向搜索和模式搜索中的固定步长改为用一维搜索确定的最优步长，算法仍然有效。

例1. 用模式搜索法求解问题

$$\min \quad f(x)=x_1^2+x_2^2 。$$

取初始点 $x^{(1)}=(1,1)^T$,初始步长 $\delta=0.25$,加速因子 $\alpha=1$,

缩减率 $\beta=0.2$。

解: 第1轮迭代:

令 $y^{(1)}=x^{(1)}=(1,1)^T$, 则 $f(y^{(1)})=2$。

$\because f(y^{(1)}+\delta e_1)=2.5625>f(y^{(1)})$,

$f(y^{(1)}-\delta e_1)=1.5625<f(y^{(1)})$,

$\therefore\ y^{(2)}=y^{(1)}-\delta e_1=(0.75,1)^T$ 。

$\because\ f(y^{(2)}+\delta e_2)=2.125> f(y^{(2)}),$

$f(y^{(2)}-\delta e_2)=1.125< f(y^{(2)}),$

$\therefore\ y^{(3)}=y^{(2)}-\delta e_2=(0.75,0.75)^T$ 。

$\because\ f(y^{(3)})< f(x^{(1)}),$

$\therefore$ 令 $x^{(2)}=y^{(3)}$。

取加速方向 $d^{(1)}=x^{(2)}-x^{(1)}=(-0.25,-0.25)^T$ 。

模式搜索:

$$y^{(1)}=x^{(2)}+\alpha d^{(1)}=(0.5,0.5)^T \text{。}$$

二. Powell 方法

基本思想:

Powell 方法主要由 基本搜索、加速搜索和 调整搜索方向三个部分组成。基本搜索包 括从基点出发沿着已知 的 n 个搜索方向进行一维搜索，确定一 个新基点。加速搜索是指沿着相邻 的两个基点的连线方向进行一维搜索，使函数值下降更快。最后用基点连线方向代替已知的 n 个搜索方向之一，构成 新的搜索方向组，进行下一轮迭代。

原始 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 $\varepsilon > 0$，令 $k = 1$。

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

$d^{(k,1)}, d^{(k,2)}, \cdots, d^{(k,n)}$ 进行搜索，即令

$$\begin{cases} x^{(k,j)} = x^{(k,j-1)} + \lambda_j d^{(k,j)} \\ \lambda_j : f(x^{(k,j-1)} + \lambda_j d^{(k,j)}) = \min\limits_{\lambda} f(x^{(k,j-1)} + \lambda d^{(k,j)}) \end{cases}$$

得到点 $x^{(k,1)}, x^{(k,2)}, \cdots, x^{(k,n)}$。

令 $d^{(k,n+1)} = x^{(k,n)} - x^{(k,0)}$, 从 $x^{(k,n)}$ 出发沿 $d^{(k,n+1)}$ 进行一维搜索得到点 $x^{(k)}$。

(3) 若 $\| x^{(k)} - x^{(k-1)} \| < \varepsilon$，停止，得到点 $x^{(k)}$; 否则，令

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

令 $k := k + 1$, 返回（2）。

例2. 用原始 **Powell** 方法求解下述问题：

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

初始点为 $x^{(0)}=(2,1)^T$，初始搜索方向为 $d^{(1,1)}=(1,0)^T$，$d^{(1,2)}=(0,1)^T$。

解：第一轮迭代：

令 $x^{(1,0)}=x^{(0)}$。从 $x^{(1,0)}$出发沿着方向 $d^{(1,1)}$作一维搜索。

$$\min_{\lambda} \quad f(x^{(1,0)}+\lambda d^{(1,1)})$$

$$\because \quad x^{(1,0)}+\lambda d^{(1,1)}=(2,1)^T+\lambda(1,0)^T=(2+\lambda,1)^T$$

记 $\varphi(\lambda)=f(x^{(1,0)}+\lambda d^{(1,1)})=(3+\lambda)^2+(1+\lambda)^2$，

令 $\dfrac{d\varphi}{d\lambda}=2(3+\lambda)+2(1+\lambda)=0$，

解得 $\lambda_1=-2$。

$$\therefore \quad x^{(1,1)}=x^{(1,0)}+\lambda_1 d^{(1,1)}=(0,1)^T。$$

再从 $x^{(1,1)}$出发，沿着方向 $d^{(1,2)}$作一维搜索，即求解

$$\min_{\lambda} \quad f(x^{(1,1)}+\lambda d^{(1,2)}).$$

解得 $\lambda_2=-1$， 所以$x^{(1,2)}=x^{(1,1)}+\lambda_2 d^{(1,2)}=(0,0)^T$。

令方向$d^{(1,3)}=x^{(1,2)}-x^{(1,0)}=(-2,-1)^T$。

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

解得 $\lambda_3=-\dfrac{2}{13}$。 所以 $x^{(1)}=x^{(1,2)}+\lambda_3 d^{(1,3)}=(\dfrac{4}{13},\dfrac{2}{13})^T$。

第二轮搜索：

初始点$x^{(2,0)}=x^1=(\dfrac{4}{13},\dfrac{2}{13})^T$,

搜索方向为：

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

求解 $\min\limits_{\lambda} \quad f(x^{(2,0)}+\lambda d^{(2,1)})$。

解得 $\lambda_1=\dfrac{-6}{13}$，所以 $x^{(2,1)}=x^{(2,0)}+\lambda_1 d^{(2,1)}=(\dfrac{4}{13},\dfrac{-4}{13})^T$。

求解 $\min\limits_{\lambda} \quad f(x^{(2,1)}+\lambda d^{(2,2)})$。

解得 $\lambda_2=\dfrac{-18}{169}$，所以 $x^{(2,2)}=x^{(2,1)}+\lambda_2 d^{(2,2)}=(\dfrac{88}{169},\dfrac{-34}{169})^T$。

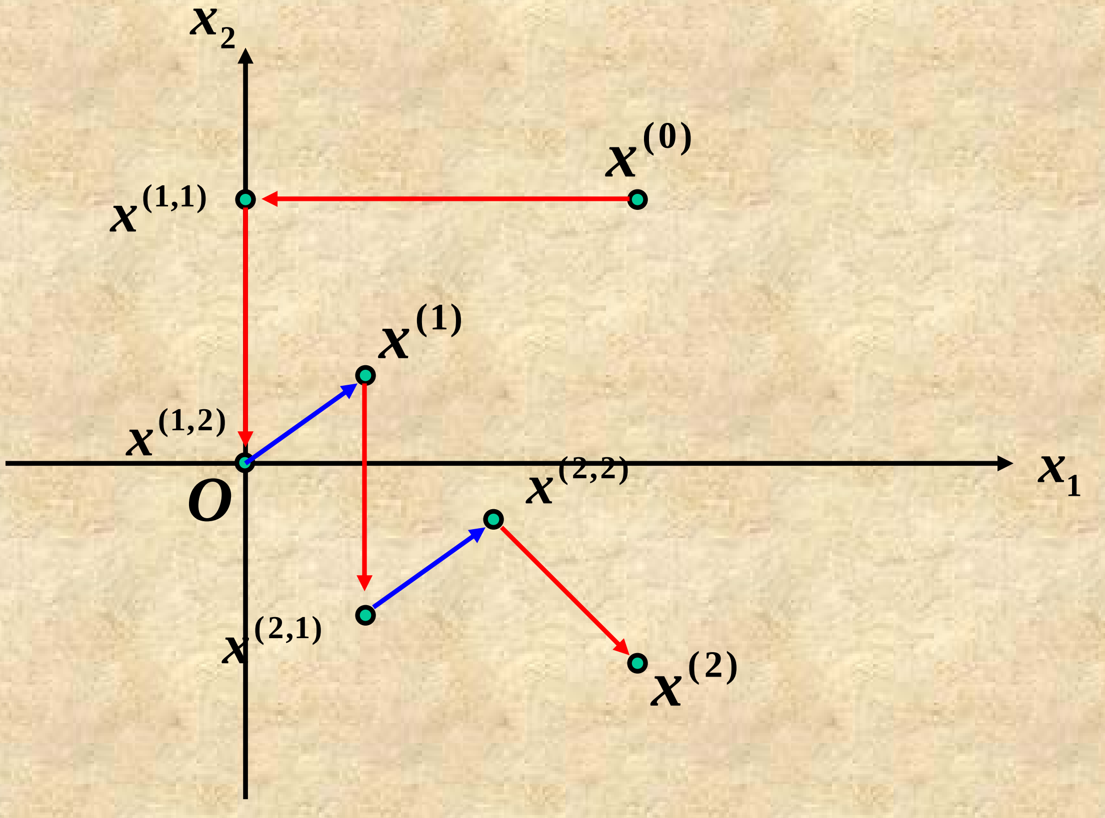
令方向 $d^{(2,3)}=x^{(2,2)}-x^{(2,0)}=(\dfrac{36}{169},\dfrac{-60}{169})^T$。

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

解得 $\lambda_3=\dfrac{9}{4}$，

所以极小点为 $x^{(2)}=x^{(2,2)}+\lambda_3 d^{(2,3)}=(1,-1)^T$。

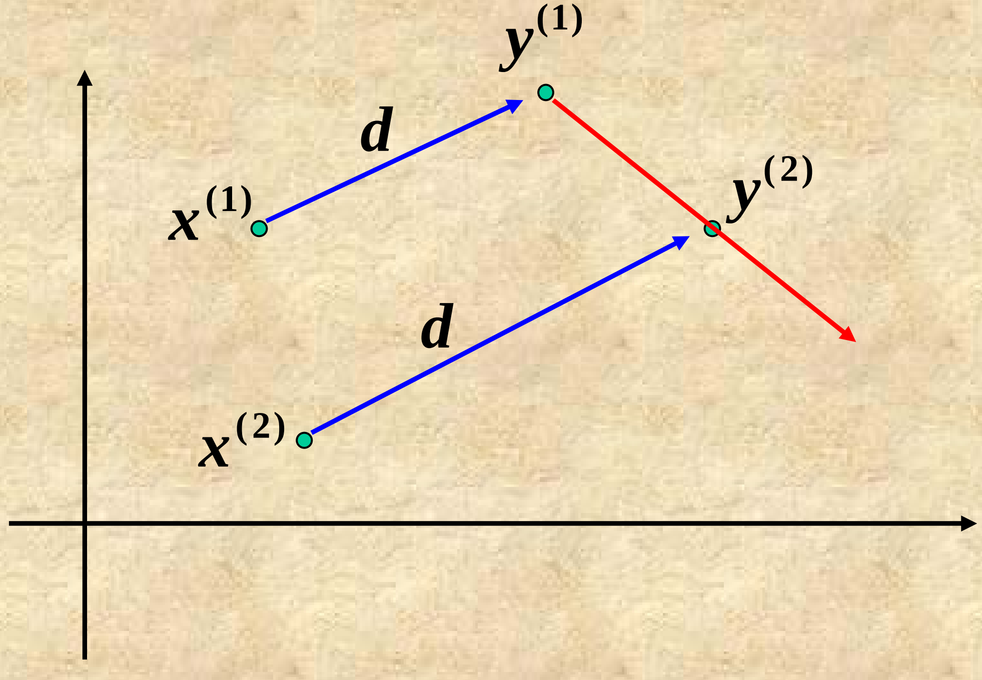
迭代过程



定理 设$f(x)=\frac{1}{2}x^T Ax+b^T x+c$，其中$A$是$n$阶对称正定矩阵。

任意取定方向 d 和点 $x^{(1)}, x^{(2)}$。从$x^{(1)}$出发沿方向 d作一维搜索得极小点 $y^{(1)}$,从$x^{(2)}$出发沿方向d 作一维搜索得极小点 $y^{(2)}$,则有

$y^{(2)}-y^{(1)}$与方向d 关于 A 共轭。



证明：

因为 $y^{(1)}$ 为从 $x^{(1)}$ 出发沿 d 由一维搜索所得，所以有

$$\nabla f(y^{(1)})^T d = 0$$

即
$$(Ay^{(1)} + b)^T d = 0$$

$$\therefore (y^{(1)^T} A + b^T) \cdot d = 0 \qquad (1)$$

同理可得
$$\therefore (y^{(2)^T} A + b^T) \cdot d = 0 \qquad (2)$$

$(1)-(2)$ 可得：$\quad (y^{(1)} - y^{(2)})^T Ad = 0$

所以 $y^{(2)} - y^{(1)}$ 与方向 d 关于 A 共轭。

对例 2的分析:

第一轮搜索方向:

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

第二轮搜索方向:

$$d^{(2,1)}=(0,1)^T, d^{(2,2)}=(-2,-1)^T, d^{(2,3)}=(\frac{36}{169},-\frac{60}{169})^T。$$

沿方向$d^{(1,3)}$搜索得到极小点 $x^{(1)}=x^{(2,0)}$,沿方向$d^{(2,2)}$搜索得

极小点$x^{(2,2)}$, 所以由定理可知方向 $d^{(2,3)}=x^{(2,2)}-x^{(2,0)}$和

和方向$d^{(2,2)}$共轭。

$x^{(2)}$是沿共轭方向搜索得到 的，因此必为极小点。

定理 设$f(x)=\frac{1}{2}x^TAx+b^Tx+c$，其中$A$是$n$阶对称正定矩阵。

用原始 *Powell* 法求解下述最优化问题

$$\min \quad f(x)。$$

若迭代已进行了 $m(m\le n)$轮，且每一轮迭代后为 下一轮所确定的前 n个搜索方向 $d^{(k,1)},d^{(k,2)},\cdots,d^{(k,n)}(k\le m)$线性无关，则各轮迭代所产 生的加速方向必构成 A共轭的向量组。

注

1. 原始 *Powell* 算法是一种共轭方向算 法。
2. 原始 *Powell* 算法不能保证各轮迭代 的前 n个搜索方向线性无关。

例3. 用原始 **Powell** 方法求解下述问题：

$$\min \quad f(x)=x_1^2+x_2^2,$$

初始点为 $x^{(0)}=(1,1)^T$，初始搜索方向为 $d^{(1,1)}=(1,-1)^T$，$d^{(1,2)}=(0,1)^T$。

解： 第一轮迭代：令 $x^{(1,0)}=x^{(0)}$。

求解 $\min\limits_{\lambda} \quad f(x^{(1,0)}+\lambda d^{(1,1)})$.

解得 $\lambda_1=0$，所以 $x^{(1,1)}=x^{(1,0)}+\lambda_1 d^{(1,1)}=(1,1)^T$。

求解 $\min\limits_{\lambda} \quad f(x^{(1,1)}+\lambda d^{(1,2)})$.

解得 $\lambda_2=-1$，所以 $x^{(1,2)}=x^{(1,1)}+\lambda_2 d^{(1,2)}=(1,0)^T$。

令方向 $d^{(1,3)}=x^{(1,2)}-x^{(1,0)}=(0,-1)^T$。

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

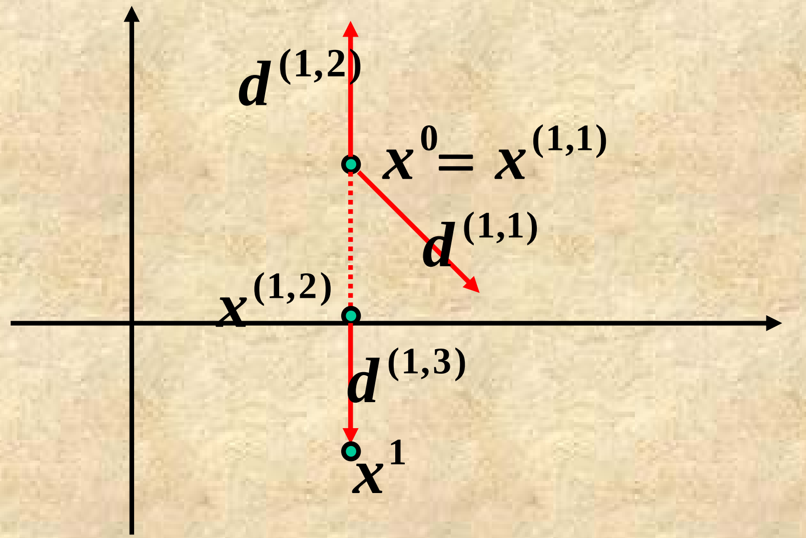
解得 $\lambda_3 = 0$，所以 $x^{(1)} = x^{(1,2)} + \lambda_3 d^{(1,3)} = (1,0)^T$。

第二轮迭代：

搜索方向：$d^{(2,1)} = d^{(1,2)} = (0,1)^T, d^{(2,2)} = d^{(1,3)} = (0,-1)^T$。

$d^{(2,1)}$和$d^{(2,2)}$线性相关，以下迭代得 到的 $x^{(k)} = (1,0)^T (k \geq 2)$,

不能得到最优解　$x^* = (0,0)^T$。



问题：如何确保各轮迭代的前 n 个搜索方向线性无关？

(1) 分析：搜索方向线性相关的原因。

$$\begin{aligned}d^{(k,n+1)} &= x^{(k,n)} - x^{(k,0)} \\ &= (x^{(k,n-1)} + \lambda_n d^{(k,n)}) - x^{(k,0)} \\ &= \cdots \\ &= x^{(k,0)} + \lambda_1 d^{(k,1)} + \lambda_2 d^{(k,2)} + \cdots + \lambda_n d^{(k,n)} - x^{(k,0)} \\ &= \lambda_1 d^{(k,1)} + \lambda_2 d^{(k,2)} + \cdots + \lambda_n d^{(k,n)}\end{aligned}$$

如果 $\lambda_1 = 0$，则 $d^{(k,n+1)}$ 是 $d^{(k,2)}, d^{(k,3)}, \cdots, d^{(k,n)}$ 的线性组合。

令 $d^{(k+1,i)} = d^{(k,i+1)}$，则第 $k+1$ 轮的 n 个搜索方向线性相关。

(2)解决方法:

每次用 $\boldsymbol{d}^{(k,n+1)}$ 换出原来的 $\boldsymbol{n}$ 个搜索方向中的一个，但不一定是 第一个 。

如何确定应换出哪一个 搜索方向?

假设初始的 $\boldsymbol{n}$ 个搜索方向是单位向量 。

令 $\boldsymbol{d}^{(k,n+1)} = \dfrac{\boldsymbol{x}^{(k,n)} - \boldsymbol{x}^{(k,0)}}{\left\| \boldsymbol{x}^{(k,n)} - \boldsymbol{x}^{(k,0)} \right\|}$。

选择搜索方向 $\boldsymbol{d}^{(k,s)}$,使其 满足

$$|\lambda_s| = \max_{1 \le i \le n} |\lambda_i|,$$

且 $|\det(\boldsymbol{d}^{(k,1)},\cdots,\boldsymbol{d}^{(k,s-1)},\boldsymbol{d}^{(k,n+1)},\boldsymbol{d}^{(k,s+1)},\cdots,\boldsymbol{d}^{(k,n)})| > \varepsilon,$

则用$d^{(k,n+1)}$换出$d^{(k,s)}$，构成第$k+1$次的搜索方向。

否则，令$d^{(k+1,i)}=d^{(k,i)}, i=1,2,\cdots,n$。

(3)行列式的计算

$$\Delta_{k+1}=|\det(d^{(k,1)},\cdots,d^{(k,s-1)},d^{(k,n+1)},d^{(k,s+1)},\cdots,d^{(k,n)})|$$

$$=|\det(d^{(k,1)},\cdots,d^{(k,s-1)},\frac{\lambda_1 d^{(k,1)}+\cdots+\lambda_n d^{(k,n)}}{\left\|x^{(k,n)}-x^{(k,0)}\right\|},d^{(k,s+1)},\cdots,d^{(k,n)})|$$

$$=\frac{|\lambda_s|}{\left\|x^{(k,n)}-x^{(k,0)}\right\|}|\det(d^{(k,1)},\cdots,d^{(k,s-1)},d^{(k,s)},d^{(k,s+1)},\cdots,d^{(k,n)})|$$

$$=\frac{|\lambda_s|}{\left\|x^{(k,n)}-x^{(k,0)}\right\|}\Delta_k$$

改进的 *Powell* 方法：

(1) 给定初始点 $x^{(0)}$, n个线性无关的方向： $d^{(0,i)}=e_i, i=1,\cdots,n$。

允许误差 $\varepsilon>0$，令 $\Delta_0=1, k=0$。

(2) 令 $x^{(k,0)}=x^{(k)}$。 依次沿n个方向进行一维搜索， 即

$$\begin{cases} x^{(k,j)}=x^{(k,j-1)}+t_j d^{(k,j)} \\ t_j: f(x^{(k,j-1)}+t_j d^{(k,j)})=\min_t f(x^{(k,j-1)}+t d^{(k,j)}) \end{cases}$$

(3) 令 $d^{(k,n+1)}=\dfrac{x^{(k,n)}-x^{(k,0)}}{\left\|x^{(k,n)}-x^{(k,0)}\right\|}$。令

$$\begin{cases} x^{(k+1)}=x^{(k,n)}+t_{n+1} d^{(k,n+1)} \\ t_{n+1}: f(x^{(k,n)}+t_{n+1} d^{(k,n+1)})=\min_t f(x^{(k,n)}+t d^{(k,n+1)}) \end{cases}$$

(4)如果$\left\|x^{k+1}-x^{k}\right\| \le \varepsilon$,算法结束，令$x^{*}=x^{(k+1)}$;否则转（5）。

(5)若$k=n-1$,则令$d^{(0,i)}=e_i, i=1,\cdots,n$。令$x^{(0)}=x^{(n)}, k:=0$, 转（2）。

若$k<n-1$,确定s使得$\left|\lambda_s\right|=\max\limits_{1\le i\le n}\left|\lambda_i\right|$。

如果$\dfrac{\left|\lambda_s\right|}{\left\|x^{(k,n)}-x^{(k,0)}\right\|}\Delta_k \le \varepsilon$，则$d^{(k+1,i)}=d^{(k,i)}, i=1,\cdots,n$。

令$\Delta_{k+1}=\Delta_k$,转（2）。

如果$\dfrac{\left|\lambda_s\right|}{\left\|x^{(k,n)}-x^{(k,0)}\right\|}\Delta_k > \varepsilon$,则令$d^{(k+1,i)}=d^{(k,i)}, i=1,\cdots,s-1$;

$d^{(k+1,s)}=d^{(k,n+1)}; d^{(k+1,i)}=d^{(k,i)}, i=s+1,\cdots,n$。

令$\Delta_{k+1}=\dfrac{\left|\lambda_s\right|}{\left\|x^{(k,n)}-x^{(k,0)}\right\|}\Delta_k, k:=k+1$,转（2）。