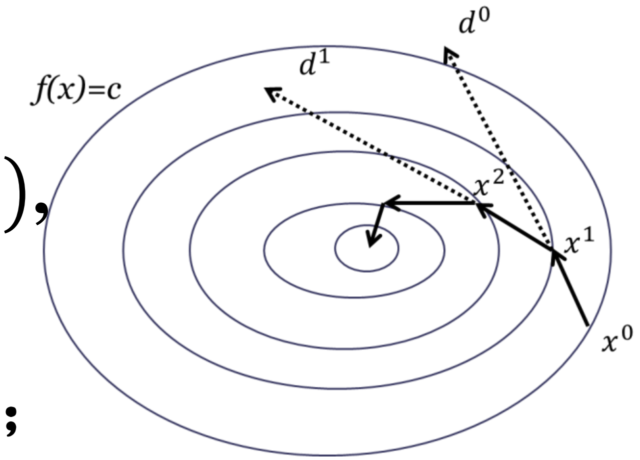
最优化理论与方法

研究生学位课

陈军华（副教授、副主任）

- **求解框架**
  - 1、从起点$\boldsymbol{x}^0$开始，计算$\boldsymbol{f}(\boldsymbol{x}^0),\nabla\boldsymbol{f}(\boldsymbol{x}^0),\nabla^2\boldsymbol{f}(\boldsymbol{x}^0)$;



  - 2、选择一个从$\boldsymbol{x}^0$出发的寻优方向 $\boldsymbol{d}^0$;
  - 3、沿着方向 $\boldsymbol{d}^0$找到新点（值） $\boldsymbol{x}^1$，计算$\boldsymbol{f}(\boldsymbol{x}^1),\nabla\boldsymbol{f}(\boldsymbol{x}^1),\nabla^2\boldsymbol{f}(\boldsymbol{x}^1)$;
  - 4、重复以上步骤，产生新点$\boldsymbol{x}^2,\boldsymbol{x}^3,\ldots,$在每步迭代中取得更优的值。
- **关键问题**
  - 哪个方向？步长？算法能保证到达最优吗？多少步？

- **梯度法收敛性分析**
  - 使用最速下降法，收敛性为线性：

$$\lim_{p\to\infty}\frac{f\left(x^{p+1}\right)-f\left(x^{*}\right)}{f\left(x^{p}\right)-f\left(x^{*}\right)}=a<1$$

求得局部最优，与初始点有关。

迭代次数也与初始点有关，初始点离最优值越接近越好。

二阶收敛梯度搜索算法技术

- **二次函数**
  - 可用正式表示：

$$f(x_1,\dots,x_n)=\frac{1}{2}\sum_{j=1}^{n}\sum_{i=1}^{n}x_j a_{ij} x_i+\sum_{i=1}^{n}b_i x_i+c$$

  - 向量矩阵的形式为：

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

c如果A为正定阵，则$f(x)$为正定二次函数。

二次函数的梯度： $\nabla f(x)=Ax+b.$

若A正定，则 $f(x)$有唯一最优解：

$$x^*=-A^{-1}b$$

- **二次函数**
  - $\boldsymbol{f(x) = c + b^T x + \frac{1}{2} x^T A x}$，若$\boldsymbol{A > 0}$，则$\boldsymbol{f(x)}$有全局最优解，且为：$\boldsymbol{x^* = -A^{-1} b}$
  - 证明：

- **二次函数**
  - $f(x) = c + b^T x + \frac{1}{2}x^T A x$，若$A > 0$，则$f(x)$有全局最优解，且为：$x^* = -A^{-1}b$
  - 证明：构造一个函数$\varphi(x) = \frac{1}{2}\left(x + A^{-1}b\right)^T A(x + A^{-1}b)$

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

$$= \frac{1}{2}x^T A x + b^T x + \frac{1}{2}b^T A^{-1} b$$

$$= \frac{1}{2}x^T A x + b^T x + c - c + \frac{1}{2}b^T A^{-1} b$$

$$= f(x) - f(x^*)$$

由于$A > 0$,当$x \neq x^* = -A^{-1}b$ 时，$\varphi(x) > 0$. 即证。

- **共轭方向**
  - 设向量 $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{E}^n$， 若$\boldsymbol{u}, \boldsymbol{v}$ 内积为0，即 $\boldsymbol{u}^T\boldsymbol{v} = (\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{0}$ ，则称 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 正交。
  - 再设$\boldsymbol{A}$为正定矩阵，若 $\boldsymbol{u}$ 和 $\mathbf{A}\boldsymbol{v}$ 正交，即 $\boldsymbol{u}^T\boldsymbol{A}\boldsymbol{v} = (\boldsymbol{u}, \boldsymbol{A}\boldsymbol{v}) = \boldsymbol{0}$ ，则称 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 为$\boldsymbol{A}$共轭。
  - 推广到向量组的共轭

  设$\boldsymbol{A}$为正定矩阵，若非零向量组： $\boldsymbol{u}^1, \boldsymbol{u}^2, \dots \boldsymbol{u}^n \in \boldsymbol{E}^n$，满足条件：$(\boldsymbol{u}^i)^T\boldsymbol{A}\boldsymbol{u}^j = \boldsymbol{0}$ $(\boldsymbol{i} \neq \boldsymbol{j}, \boldsymbol{i}, \boldsymbol{j} = \mathbf{1}, \mathbf{2} \dots \boldsymbol{n})$，则称该向量组为A共轭。当A为单位阵时，向量组为正交。

- **共轭**

共轭：按照一定规律相配的一对。

- **共轭方向**
  - 定理：设 $\boldsymbol{A}$ 为正定矩阵，若非零向量组： $\boldsymbol{u}^1, \boldsymbol{u}^2, \ldots \boldsymbol{u}^n \in \boldsymbol{E}^n$ ，为A共轭，则该向量组线性无关。

- **共轭方向**
  - 定理：设 A 为正定矩阵，若非零向量组： $u^1, u^2, \dots u^n \in E^n$ ，为A共轭，则该向量组线性无关。
  - 证明：若线性相关，则存在不全为0的数$a = (a_1, a_n, \dots a_n)$使得：

$$a_1 u^1 + a_2 u^2 +, \dots, + a_n u^n = 0$$

用$(u^i)^T A$左乘上式：

$$(u^i)^T A a_1 u^1 + (u^i)^T a_2 u^2 +, \dots, + (u^i)^T a_n u^n = 0$$

即： $a_i (u^i)^T A u^i = 0$

A为正定矩阵$u^i \neq 0$，$(u^i)^T A u^i > 0$

故：$a_i = 0$ 矛盾，即证。

- **共轭方向几何意义**
  - 设有二次函数

$$f(x)=\frac{1}{2}(x-\bar{x})^T A(x-\bar{x})$$

A为$n\times n$对称正定矩阵，$\bar{x}$是一个定点。

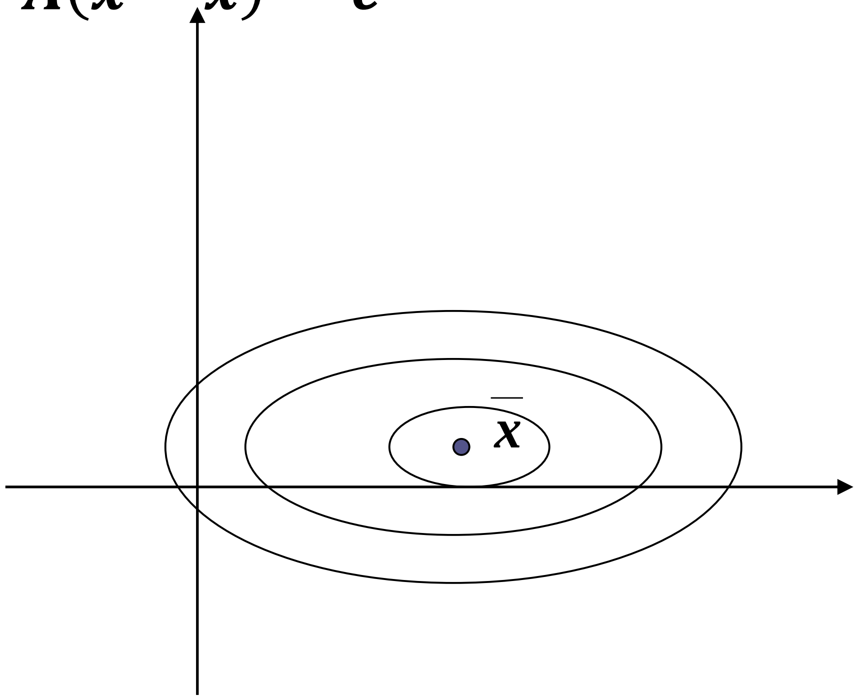
则函数$f(x)$的等值面：$\frac{1}{2}(x-\bar{x})^T A(x-\bar{x})=c$

是以$\bar{x}$为中心的椭球面。

$\because \nabla f(\bar{x})=A(\bar{x}-\bar{x})=0,$

而 $\nabla^2 f(\bar{x})=A>0$

$\therefore \bar{x}$是$f(x)$的极小点。



设 $x^{(0)}$ 是在某个等值面上的一点，$d^{(1)}$ 是 R^n 中的一个方向，

$x^{(0)}$ 沿着 $d^{(1)}$ 以最优步长搜索得到点 $x^{(1)}$。

则 $d^{(1)}$ 是点 $x^{(1)}$ 所在等值面的切向量。

该等值面在点 $x^{(1)}$ 处的法向量为

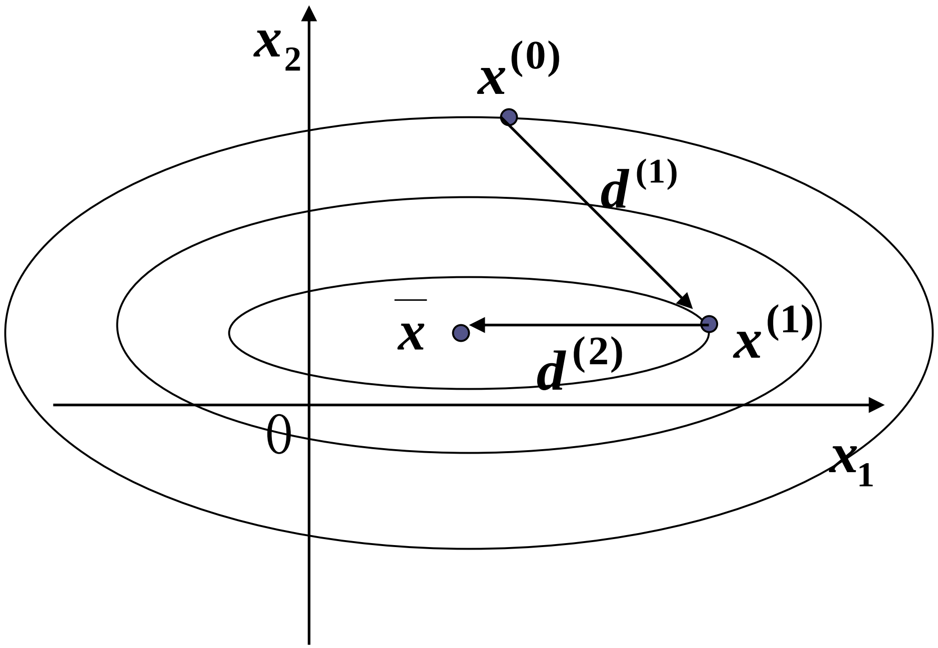
$$\nabla f(x^{(1)}) = A(x^{(1)} - \bar{x}).$$

则 $d^{(1)}$ 与 $\nabla f(x^{(1)})$ 正交，

即 $d^{(1)T}\nabla f(x^{(1)}) = 0,$

令 $d^{(2)} = \bar{x} - x^{(1)},$

所以 $d^{(1)T} A d^{(2)} = 0,$



即等值面上一点处的切 向量与由这一点指向极 小点的向量关于 A 共轭。

- **共轭方向搜索的二阶收敛性**
  - 对于多元正定二次目标函数，从任意初始点出发，如果经过有限次迭代就能够求得极小点，那么这种算法具有二次终止性.
  - 例如Newton法对于二次函数只须经过一次迭代就可以求得极小点，因此是二次终止的；而最速下降法不具有二次终止性；共轭方向法（包括共轭梯度法，变尺度法等）是二次终止的.
  - 一般来说，具有二次终止性的算法，在用于一般函数时，收敛速度较快.

- 共轭方向搜索的二阶收敛性
  - 定理：对于一寻优算法，若每一步寻优方向都相互共轭（即沿着共轭方向），则此算法是二阶收敛的。

证明此定理，可设$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{c} + \boldsymbol{b}^T\boldsymbol{x} + \frac{1}{2}\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}$ 为正定二次函数，从任一初始点$\boldsymbol{x}^0$出发，相继以 $\boldsymbol{u}^1, \boldsymbol{u}^2, \dots \boldsymbol{u}^n$为搜索方向，其中 $\boldsymbol{u}^1, \boldsymbol{u}^2, \dots \boldsymbol{u}^n$为A共轭。试证明经至多 $\boldsymbol{n}$次迭代后可达最优点：$\boldsymbol{x}^* = -\boldsymbol{A}^{-1}\boldsymbol{b}$

迭代规则是：
$$\begin{cases} \min\limits_{\lambda} \boldsymbol{f}\left(\boldsymbol{x}^i + \lambda \boldsymbol{u}^{i+1}\right) \\ \boldsymbol{x}^{i+1} = \left(\boldsymbol{x}^i + \lambda_{i+1}\boldsymbol{u}^{i+1}\right) \end{cases}$$

- **共轭方向搜索的二阶收敛性**
  - 证明：设各次搜索点分别为： $\boldsymbol{x}^1, \boldsymbol{x}^2, \ldots \boldsymbol{x}^n$

则： $\boldsymbol{x}^1 = \boldsymbol{x}^0 + \lambda_1 \boldsymbol{u}^1$

$$\boldsymbol{x}^2 = \boldsymbol{x}^1 + \lambda_2 \boldsymbol{u}^2 = \boldsymbol{x}^0 + \lambda_1 \boldsymbol{u}^1 + \lambda_2 \boldsymbol{u}^2$$

....

$$\boldsymbol{x}^i = \boldsymbol{x}^0 + \sum_{k=1}^{i} \lambda_k \boldsymbol{u}^k$$

考虑 $\boldsymbol{x}^{i+1} = \boldsymbol{x}^i + \lambda_{i+1} \boldsymbol{u}^{i+1}$，根据迭代规则步长$\lambda_{i+1}$要满足：

$\min_{\lambda} \boldsymbol{f}\left(\boldsymbol{x}^i + \lambda \boldsymbol{u}^{i+1}\right)$，根据最速下降的条件可知相邻两方向正交：

$$\left(\boldsymbol{u}^{i+1}, \nabla \boldsymbol{f}\left(\boldsymbol{x}^i + \lambda_{i+1} \boldsymbol{u}^{i+1}\right)\right) = \mathbf{0}$$

由于$\nabla \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}$ 代入上式可得：

- **共轭方向搜索的二阶收敛性**

$$\lambda_{i+1} = -\frac{(u^{i+1},Ax^{i}+b)}{(u^{i+1},Au^{i+1})} = -\frac{(u^{i+1},Ax^{0}+b)}{(u^{i+1},Au^{i+1})}$$

$$x^n = x^0 + \sum_{k=1}^{n} \lambda_k u^k \quad = x^0 - \sum_{k=1}^{n} \frac{(u^k, Ax^0 + b)}{(u^k, Au^k)} u^k$$

$$= x^0 - \sum_{k=1}^{n} \frac{(u^k, Ax^0)}{(u^k, Au^k)} u^k - \sum_{k=1}^{n} \frac{(u^k, A[A^{-1}b])}{(u^k, Au^k)} u^k$$

$$= x^0 - x^0 - A^{-1}b = -A^{-1}b = x^*$$

- **共轭方向搜索的二阶收敛性**

  1）和初始点无关；

  2）和共轭方向的顺序无关；

  3）迭代次数最多为n，即与维数有关。

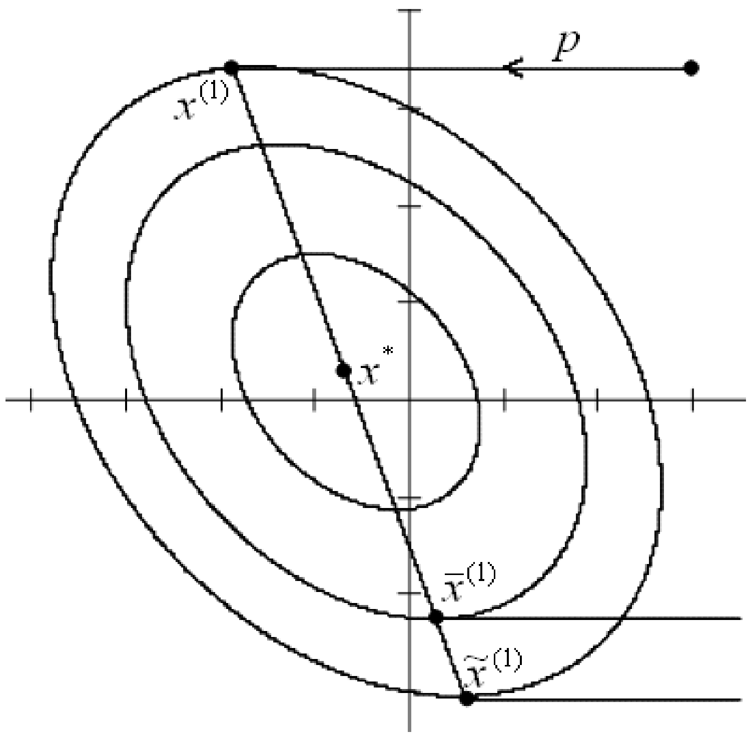


图 5-2 二维正定二次函数的二次终结性

- **例 4_6**

  用共轭方向的寻优原理，求

  $$\boldsymbol{min f(x_1, x_2) = 1 + x_1 - x_2} + x_1^2 + 2x_2^2$$

- **例 4_6**

  用共轭方向的寻优原理，求

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

解：$A = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$，构造为 A 的共轭向量，

$u^1 = [1,0]^T, u^2 = [0,1]^T$，可验证$(u^1, Au^2) = 0$

初始点为$x^0 = [0,0]^T$, 根据 $\lambda_i = -\frac{(u^i, Ax^0 + b)}{(u^i, Au^i)}$，可得：

$$\lambda_1 = -\frac{1}{2}, \lambda_2 = \frac{1}{4}$$

$$x^2 = x^0 + \lambda_1 u^1 + \lambda_2 u^2 = [-\frac{1}{2}, \frac{1}{4}]^T = x^*$$

- **共轭梯度算法（Conjungate-drirection）**

以梯度法的求解框架，产生一组共轭向量。

在生成共轭向量的同时，考虑算法的效率和计算的复杂度，尽量不用二阶梯度。

算法基本思想是把共轭性与最速下降方法相结合，利用已知点处的梯度构造一组共轭方向，并沿这组方向进行搜索，求出目标函数的极小点。

- **Fletcher-Reeves共轭梯度算法**
  - 1：选取初始点$\boldsymbol{x}^0$，初始方向$\boldsymbol{v}^0 = -\nabla \boldsymbol{f}(\boldsymbol{x}^0)$
  - 2：$for\ i = 1, \dots, n$

    2.1：$\boldsymbol{x}^i = \boldsymbol{x}^{i-1} + \boldsymbol{\lambda}_{i-1}\boldsymbol{v}^{i-1}, \boldsymbol{\lambda}_{i-1}$取$\boldsymbol{min f}(\boldsymbol{x}^{i-1} + \boldsymbol{\lambda}\boldsymbol{v}^{i-1})$

    2.2：$\boldsymbol{v}^i = -\nabla \boldsymbol{f}(\boldsymbol{x}^i) + \boxed{\frac{\left\|\nabla \boldsymbol{f}(\boldsymbol{x}^i)\right\|^2}{\left\|\nabla \boldsymbol{f}(\boldsymbol{x}^{i-1})\right\|^2}\boldsymbol{v}^{i-1}}$

  3：$\boldsymbol{x}^0 = \boldsymbol{x}^n$.

- **Fletcher-Reeves共轭梯度–共轭向量组构造**

这里用正定二次函数说明：$f(x)=c+b^T x+\frac{1}{2}x^T Ax\,,A>0$,初始点 x^0，如何构造一组共轭向量$v^0,v^1,\dots,v^{n-1}$

根据算法：$v^0=-\nabla f\left(x^0\right)$

沿$x^0+\lambda v^0$射线进行搜索，$x^1=x^0+\lambda v^0$,

$$\nabla f(x^1)^T v^0=-\nabla f(x^1)^T\nabla f\left(x^0\right)=0$$

即x^0,x^1两点的梯度正交，在此正交系构成的二维空间中寻找下一个方向 v^1，则v^1= $-\nabla f(x^1)\ \ +a_0\nabla f\left(x^0\right)$

要使v^0，v^1为A共轭，则有： $v^1\mathrm{A}v^0=0$

- **Fletcher-Reeves共轭梯度-共轭向量组构造**

对于二次函数有：

$$\nabla f(x^{k+1}) \quad -\nabla f\left(x^{k}\right)=Ax^{k+1}+b-Ax^{k}-b$$

$$=A(x^{k+1}-x^{k})=A(x^{k}+\lambda_{k}v^{k}-x^{k})$$

$$=\lambda_{k}\,Av^{k}$$

由于 $v^{1}Av^{0}=0$ ， 即 $v^{1}\lambda_{0}Av^{0}=0$, 展开可得：

$$v^{1}\lambda_{0}Av^{0}=[-\nabla f\left(x^{1}\right)+a_{0}\,\nabla f(x^{0})]^{T}[\nabla f\left(x^{1}\right)-\nabla f\left(x^{0}\right)]$$

$$=-\nabla f\left(x^{1}\right)^{T}\nabla f\left(x^{1}\right)+a_{0}\,\nabla f(x^{0})^{T}\nabla f\left(x^{1}\right)+\nabla f\left(x^{1}\right)^{T}\nabla f\left(x^{0}\right)-$$

$$a_{0}\,\nabla f\left(x^{0}\right)^{T}\nabla f\left(x^{0}\right)=0$$

即 $a_{0}\,\nabla f\left(x^{0}\right)^{T}\nabla f\left(x^{0}\right)=-\nabla f\left(x^{1}\right)^{T}\nabla f\left(x^{1}\right)$

$$a_{0}=-\frac{\nabla f\left(x^{1}\right)^{T}\nabla f\left(x^{1}\right)}{\nabla f(x^{0})^{T}\nabla f(x^{0})}=-\frac{\left\|\nabla f\left(x^{1}\right)\right\|^{2}}{\|\nabla f(x^{0})\|^{2}}$$

- **Fletcher-Reeves共轭梯度–共轭向量组构造**

沿$x^1+\lambda v^1$射线进行搜索，$x^2=x^1+\lambda v^1$，

求得：$\nabla f(x^2)$

考查式子：$-\nabla f(x^0)^T\left(\nabla f\left(x^2\right)-\nabla f\left(x^1\right)\right)$

$$=-\nabla f\left(x^0\right)^T\left(\lambda_1 A v^1\right)$$
$$=-\lambda_1 v^0\ A v^1$$

由于v^0,v^1共轭，故 $-\nabla f(x^0)^T\left(\nabla f\left(x^2\right)-\nabla f\left(x^1\right)\right)=0$,即：

$$\nabla f(x^0)^T\nabla f\left(x^2\right)=0$$

即 x^0,x^1,x^2 三点的梯度构成三维正交系，寻找下一个方向 v^2，

则$v^2=-\nabla f(x^1)\quad+a_0\nabla f\left(x^0\right)+a_1\nabla f\left(x^1\right)$

要使v^0，v^1，v^2为A共轭，则有： $v^2\mathrm{A}v^0=0$，$v^2\mathrm{A}v^1=0$

- **Fletcher-Reeves共轭梯度−共轭向量组构造**

$\lambda_0 v^2 A v^0$

$= [-\nabla f(x^2) + a_0 \nabla f(x^0) + a_1 \nabla f(x^1)]^T [\nabla f(x^1) - \nabla f(x^0)] = 0$

$\lambda_1 v^2 A v^1$

$= [-\nabla f(x^2) + a_0 \nabla f(x^0) + a_1 \nabla f(x^1)]^T [\nabla f(x^2) - \nabla f(x^1)] = 0$

求得： $a_1 = -\dfrac{\nabla f(x^2)^T \nabla f(x^2)}{\nabla f(x^1)^T \nabla f(x^1)} = -\dfrac{\|\nabla f(x^2)\|^2}{\|\nabla f(x^1)\|^2}$

$$a_0 = a_1 \cdot \frac{\nabla f(x^1)^T \nabla f(x^1)}{\nabla f(x^0)^T \nabla f(x^0)} = a_1 \cdot \frac{\|\nabla f(x^1)\|^2}{\|\nabla f(x^0)\|^2}$$

……

- **例 4_7**

  用FR共轭梯度法求：

  $$\boldsymbol{min f(x_1, x_2)} = \frac{\mathbf{3}}{\mathbf{2}} x_1^2 + \frac{1}{2} x_2^2 - x_1 x_2 - 2x_1,\ x^0 = (-2,4)$$

- **例 4_7**

用FR共轭梯度法求：

$$\boldsymbol{min f(x_1, x_2)} = \frac{\mathbf{3}}{\mathbf{2}}x_1^2 + \frac{1}{2}x_2^2 - x_1x_2 - 2x_1,\ x^0 = (-2,4)$$

解：第1次迭代

$$\nabla \boldsymbol{f}(x^0) = (\mathbf{3}x_1 - x_2 - 2, x_2 - x_1)^{\boldsymbol{T}} = (\mathbf{-12,6})^{\boldsymbol{T}}$$

$$v^0 = -\nabla \boldsymbol{f}(x^0) = (\mathbf{12}, \mathbf{-6})^{\boldsymbol{T}}$$

$$x^1 = x^0 + \lambda v^0 = (\mathbf{-2,4})^{\boldsymbol{T}} + (\mathbf{12}\lambda, \mathbf{-6}\lambda)^{\boldsymbol{T}} = (\mathbf{-2} + \mathbf{12}\lambda, \mathbf{4} - \mathbf{6}\lambda)^{\boldsymbol{T}}$$

$$f(x^1) = \frac{\mathbf{3}}{\mathbf{2}}(\mathbf{-2} + \mathbf{12}\lambda)^2 + \frac{1}{2}(\mathbf{4} - \mathbf{6}\lambda)^2 - (\mathbf{-2} + \mathbf{12}\lambda)(\mathbf{4} - \mathbf{6}\lambda) - 2(\mathbf{-2} + \mathbf{12}\lambda)$$

$$f_\lambda'(x^1) = 612\lambda - 180 = 0$$

$$\lambda = \frac{5}{17},\ x^1 = \left(\frac{26}{17}, \frac{38}{17}\right)^{\boldsymbol{T}}$$

- **例 4_7**

第1次迭代

$$x^1 = (\frac{26}{17}, \frac{38}{17})^T$$

$$\nabla f(x^1) = (3x_1 - x_2 - 2, x_2 - x_1)^T = (\frac{6}{17}, \frac{12}{17})^T$$

$$v^1 = -\nabla f(x^1) + \frac{\|\nabla f(x^1)\|^2}{\|\nabla f(x^0)\|^2} d^0$$

$$= -\left(\frac{6}{17}, \frac{12}{17}\right)^T + \frac{\left\|\left(\frac{6}{17}, \frac{12}{17}\right)^T\right\|^2}{\|(-12,6)^T\|^2} \cdot (12, -6)^T = (-\frac{90}{289}, -\frac{210}{289})^T$$

- **例 4_7**

第2次迭代

$$x^2 = x^1 + \lambda v^1 = (\frac{26}{17}, \frac{38}{17})^{\boldsymbol{T}} + (-\frac{90\lambda}{289}, -\frac{210\lambda}{289})^{\boldsymbol{T}}$$

$$\lambda = \frac{17}{10}$$

$$x^2 = (1,1)^{\boldsymbol{T}}$$

$$\boldsymbol{\nabla f}(x^2) = 0$$

已达最优。

- **例 4_8**

  用FR共轭梯度法求：

$$\boldsymbol{minf}(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{A}\boldsymbol{x}$$

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, x^0 = (-10,5)$$

$$x^* = (0,0)$$

- 例 4_8

$$\nabla f(x) = \binom{x_1 + x_2}{x_1 + 2x_2}$$

$$v^0 = -\nabla f(x) = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$$

$$f(x^0 + \lambda v^0) = \frac{1}{2}(50 - 50\lambda + 25\lambda^2) \rightarrow \lambda_0 = 0.9996$$

$$x^1 = x^0 + \lambda_0 v^0 = [-5.002,5]^T$$

$$v^1 = -\nabla f(x^1) + \frac{\left\|\nabla f(x^1)\right\|^2}{\|\nabla f(x^0)\|^2} v^0 = [4.9980, -4.9980]^T$$

$$f(x^1 + \lambda v^1) = \frac{1}{2}(26.55 - 38.95\lambda + 14.45\lambda^2) \rightarrow \lambda_1 = 1.0005$$

$$x^2 = x^1 + \lambda_1 v^1 = [-0.0016, -0.0004]^T$$

非二次函数的共轭梯度法

非二次函数的共轭梯度法

- 将二次函数的共轭梯度法推广到一般函数，实际上是基于对函数的二次逼近。
- 设$f(x)$是定义在某一凸集R上的二阶连续可微的严格凸函数，则在R内有唯一极小点x^*，在R内任取初始近似点 x^0

  取：$v^0 = -\nabla f(x^0)$为初始搜索方向

  则：$x^1 = x^0 + \lambda v^0$
- 对$f(x^0 + \lambda v^0)$在x^0点作泰勒展开得：

$$f\left(x^0 + \lambda v^0\right) = f\left(x^0\right) + \lambda \nabla f(x^0)^T \cdot v^0 + \frac{1}{2}\lambda^2 v^0 H v^0 + o(\lambda^2)$$

$$\frac{df\left(x^0 + \lambda v^0\right)}{d\lambda} = \nabla f(x^0)^T \cdot v^0 + \frac{2}{2}\lambda (v^0)^T H v^0 = 0$$

- **非二次函数的共轭梯度法**

$$\lambda^0 = -\frac{\nabla f\left(x^0\right)^T \cdot v^0}{(v^0)^T H v^0}$$

即为近似最优解。

现构造向量$v^1 = -\nabla f\left(x^1\right) + \beta_0 v^0$ 使其满足共轭条件：

$$(v^1)^T H v^0 = 0$$

即：

$$\left(-\nabla f\left(x^1\right) + \beta_0 v^0\right)^T H\left(v^0\right) = 0$$

$$\beta_0 = \frac{\nabla f\left(x^1\right)^T H \cdot v^0}{(v^0)^T H v^0}$$

- **非二次函数的共轭梯度法**

这就确定了v^1，进而确定x^2

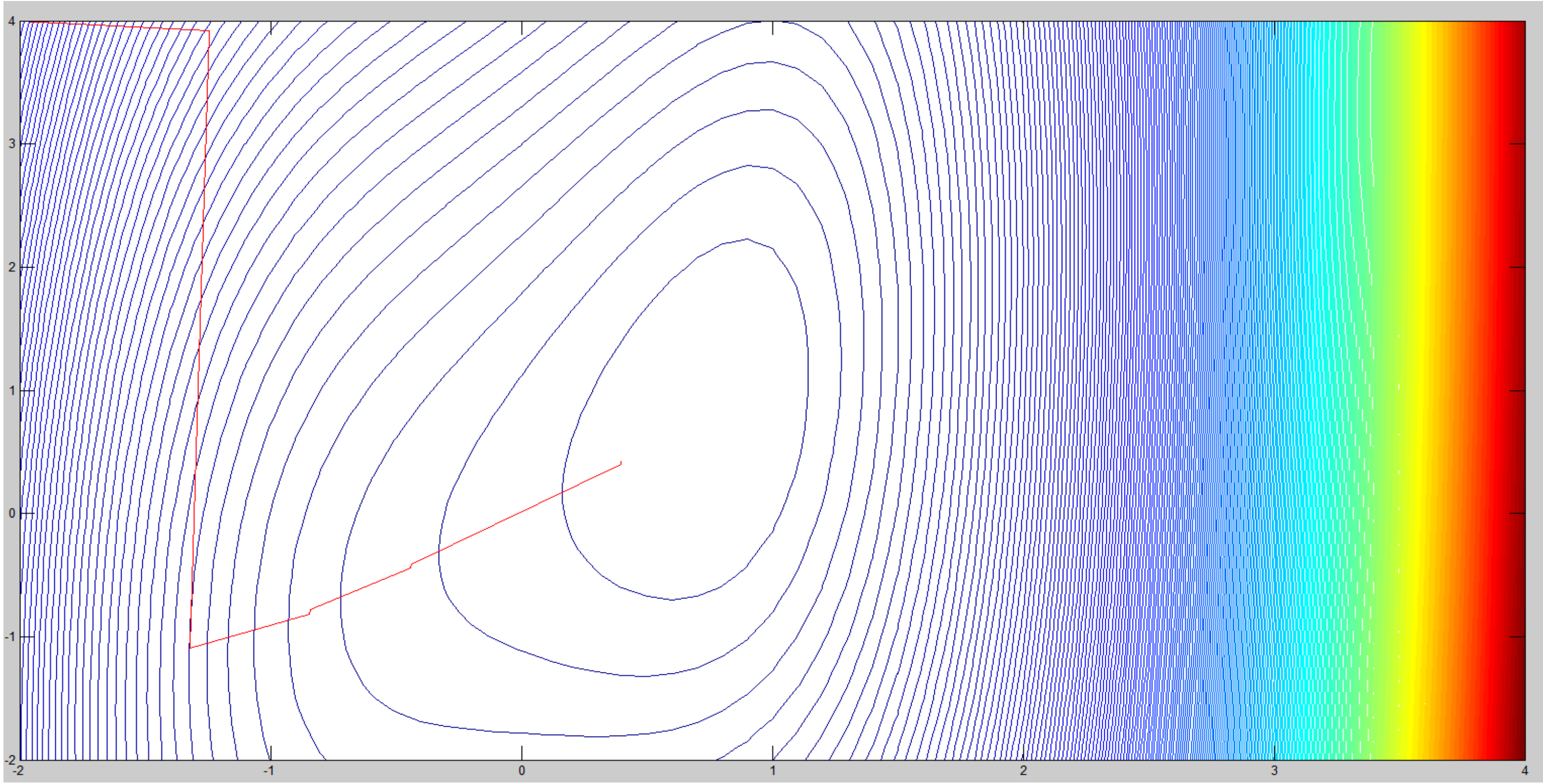
如此，可连续构造方向并继续迭代，若至k步，有：

$$\begin{cases} \boxed{\begin{aligned} x^{k+1} &= x^k + \lambda^k v^k \\ \lambda^k &= -\frac{\nabla f\left(x^k\right)^T \cdot v^k}{(v^k)^T H(x^k) v^k} \end{aligned}} \\ \boxed{\begin{aligned} v^{k+1} &= -\nabla f\left(x^{k+1}\right) + \beta_k v^k \\ \beta_k &= \frac{\nabla f\left(x^{k+1}\right)^T H(x^k) v^k}{(v^k)^T v^k} \end{aligned}} \end{cases}$$

注意：由于推导中应用了近似公式，严格说来，各方向是不共轭的，这就是说，n步收敛是不可能的。

- **非二次函数的共轭梯度法**

  - 作业：设计算法并实现 对非二次（高次）函数的寻优。
  - 求：$\boldsymbol{min f}(\boldsymbol{x_1}, \boldsymbol{x_2}) = \frac{\mathbf{3}}{\mathbf{2}} x_1^4 + \frac{1}{2} x_2^2 - x_1 x_2 - 2x_1, x^0 = (-2,4)$,精度 **0.01**



- **共轭梯度法特点**
  - 具有全局收敛性
  - 收敛速度快（二次收敛）。对于正定二次函数，具有二次终结性。
  - 算法结构简单，计算量小，存储量小。（如FR法只要3个n维向量的存储空间）
  - 适用性：特别适用于大规模优化问题的求解。

  - 共轭梯度法需要计算梯度，这个过程很耗时。

- **共轭梯度法**
  - 共轭梯度法需要计算梯度，这个过程很耗时。
  - **Powell** 设计了通过一维搜索找到共轭方向的方法。

- **Powell 算法思想**
  - 该方法主要由 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。
  - 最后用基点连续方向代替已知的n个搜索方向之一，构成新的搜索方向组，进行下一轮迭代。

- **Powell 算法**
  - 设有一组线性独立的向量$\{v^i, i = 1, \ldots, n\}$, x^0为初始点。算法步骤为:
  - 1: *for* i=1,…,n，找到 $\min_{\lambda} f(y^{i-1} + \lambda v^i)$,的 λ值，递推下一点 $y^i = y^{i-1} + \lambda v^i$
  - 2: *for* i=1,…,n，使 $v^i = v^{i+1}$
  - 3: $v^n = y^n - y^0$
  - 4: 找到 $\min_{\lambda} f(x^n + \lambda(x^n - x^0))$,的 λ值，替换$x^0 = y^n + \lambda(y^n - y^0)$
  - 5: 达不到精度回到step1

- **例 4_9**

用powell方法求 $minf(x)=(x_1+x_2)^2+(x_1-1)^2$,初始点

$x^0=\begin{bmatrix}2\\1\end{bmatrix}$，初始搜索方向 $d^{(1,1)}=(1,0)^T$, $d^{(1,2)}=(0,1)^T$

- **例 4_9**

用powell方法求 $\boldsymbol{minf}(\boldsymbol{x}) = (\boldsymbol{x}_1 + \boldsymbol{x}_2)^2 + (\boldsymbol{x}_1 - 1)^2$,初始点

$\boldsymbol{x}^0 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$，初始搜索方向 $\boldsymbol{d}^{(1,1)} = (1,0)^T$, $\boldsymbol{d}^{(1,2)} = (0,1)^T$

解：第一轮迭代：

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

$$\min_{\lambda} \boldsymbol{f}\left(\boldsymbol{y}^{(1,1)} + \lambda \boldsymbol{d}^{(1,1)}\right)$$

$$\boldsymbol{y}^{(1,1)} + \lambda \boldsymbol{d}^{(1,1)} = (2,1)^T + \lambda(1,0)^T = (2 + \lambda, 1)^T$$

$$\boldsymbol{\varphi}(\lambda) = \boldsymbol{f}\left(\boldsymbol{y}^{(1,1)} + \lambda \boldsymbol{d}^{(1,1)}\right) = (3 + \lambda)^2 + (1 + \lambda)^2$$

$$\frac{d\boldsymbol{\varphi}}{d\lambda} = 2(3 + \lambda) + 2(1 + \lambda) = 0, \lambda_1 = -2$$

$$\boldsymbol{y}^{(1,2)} = \boldsymbol{x}^{(1,1)} + \lambda_1 \boldsymbol{d}^{(1,1)} = (0,1)^T$$

- **例 4_9**

从$\boldsymbol{y}^{(1,2)}$出发沿着方向$\boldsymbol{d}^{(1,2)}$作一维搜索。

$$\min_{\lambda} f\left(\boldsymbol{y}^{(1,2)}+\lambda \boldsymbol{d}^{(1,2)}\right)$$

$\lambda_2=-1$，$\boldsymbol{y}^{(1,3)}=\boldsymbol{y}^{(1,2)}+\lambda_2\boldsymbol{d}^{(1,2)}=(0,0)^T$ **(坐标轮换结束)**

创建方向$\boldsymbol{d}^{(1,3)}=\boldsymbol{y}^{(1,3)}-\boldsymbol{y}^{(1,1)}=(-2,-1)^T$

求解 $\min_{\lambda} f\left(\boldsymbol{x}^{(1,2)}+\lambda \boldsymbol{d}^{(1,3)}\right)$

$\lambda_3=-\frac{2}{13}$，$\boldsymbol{x}^1=\boldsymbol{y}^{(1,3)}+\lambda_3\boldsymbol{d}^{(1,3)}=(\frac{4}{13},\frac{2}{13})^T$

进入第2轮搜索：

初始点 $\boldsymbol{y}^{(2,1)}=\boldsymbol{x}^1=(\frac{4}{13},\frac{2}{13})^T$，

搜索方向为：

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

- **例 4_9**

求解 $\min\limits_{\lambda} f\left(y^{(2,1)}+\lambda d^{(2,1)}\right)$

$\lambda_1=\frac{-6}{13},\ \ y^{(2,2)}=y^{(2,1)}+\lambda_1 d^{(2,1)}=(\frac{4}{13},\frac{-4}{13})^T$

求解 $\min\limits_{\lambda} f\left(y^{(2,2)}+\lambda d^{(2,2)}\right)$

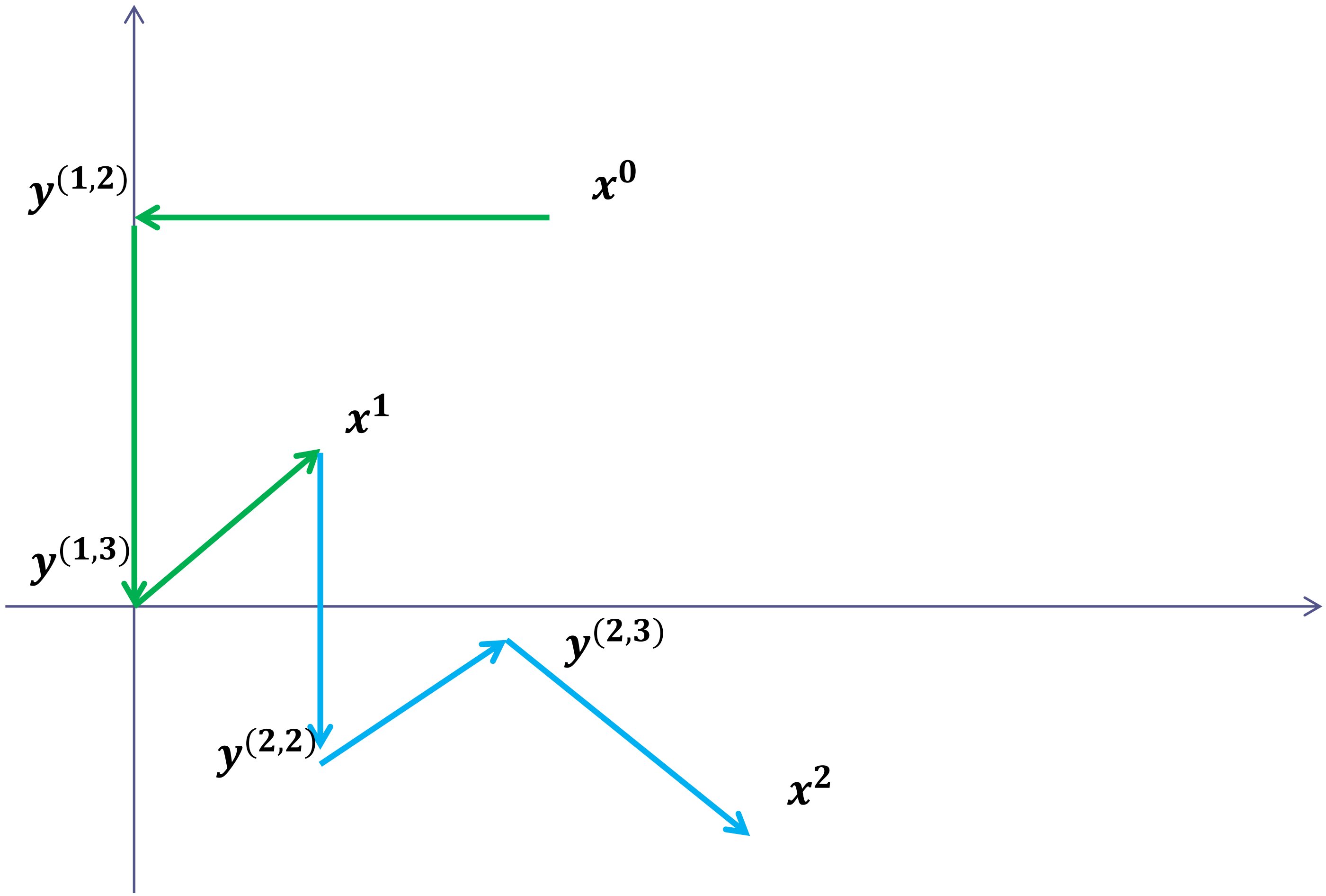
$\lambda_2=\frac{-18}{169},\ \ y^{(2,3)}=y^{(2,2)}+\lambda_2 d^{(2,2)}=(\frac{88}{169},\frac{-34}{169})^T$

创建方向$d^{(2,3)}=y^{(2,3)}-y^{(2,1)}=(\frac{36}{169},\frac{-60}{169})^T$

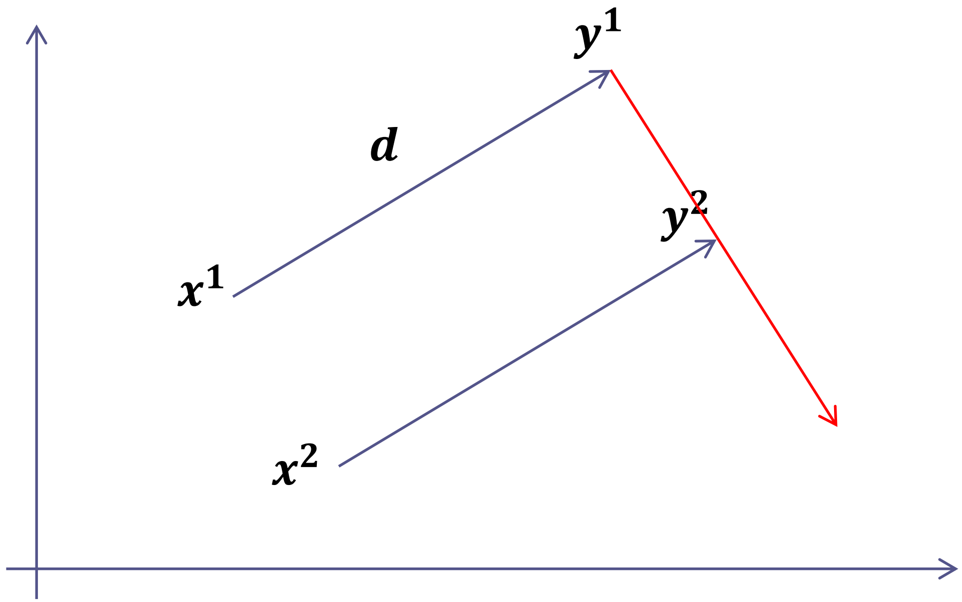
求解 $\min\limits_{\lambda} f\left(y^{(2,3)}+\lambda d^{(2,3)}\right)$

$\lambda_3=\frac{9}{4},\ \ x^2=y^{(2,3)}+\lambda_3 d^{(2,3)}=(1,-1)^T$

- **例 4_9　迭代过程**



- **关于powell算法的共轭方向**
  - 定理：设$\boldsymbol{f}(\boldsymbol{x})=\frac{1}{2}\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b}^T\boldsymbol{x}+\boldsymbol{c}$，$A$为$n$阶正定阵。任取方向$d$和点$\boldsymbol{x}^1$，$\boldsymbol{x}^2$。从$\boldsymbol{x}^1$出发沿方向$d$作一维搜索得极小点 $\boldsymbol{y}^1$，从$\boldsymbol{x}^2$出发沿方向d作一维搜索得极小点 $\boldsymbol{y}^2$，则有$\boldsymbol{y}^2-\boldsymbol{y}^1$与方向$\boldsymbol{d}$关于$A$共轭。



- **关于powell算法的共轭方向**
  - 对例4_5分析共轭方向
  - 第一轮搜索方向：

$$\boldsymbol{d}^{(1,1)}=(1{,}0)^T,\boldsymbol{d}^{(1,2)}=(0{,}1)^T,\quad \boldsymbol{d}^{(1,3)}=(-2,-1)^T$$

  - 第二轮搜索方向：

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

沿方向$\boldsymbol{d}^{(1,3)}$得 $\boldsymbol{x}^{1}$,沿方向$\boldsymbol{d}^{(2,2)}$得 $\boldsymbol{x}^{(2,2)}$,而$\boldsymbol{d}^{(2,3)}=\boldsymbol{x}^{(2,2)}-\boldsymbol{x}^{(2,0)}$,故和方向 $\boldsymbol{d}^{(2,2)}$共轭。

而$\boldsymbol{x}^{2}$是沿共轭方向搜索得到的，因此必为极小点。

- **例 4_10**
  - 求 $\min f(x) = x_1^2 + 2x_2^2, d^1 = [1,-1]^T, d^2 = [1,1]^T,$ $x^0 = [20,20]^T$

- **例 4_10**
  - 求 $\min f(x) = x_1^2 + 2x_2^2, d^1 = [1,-1]^T, d^2 = [1,1]^T,$

$x^0 = [20,20]^T$

解：第一轮迭代：

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

$$\min_{\lambda} f\left(y^{(1,1)} + \lambda d^{(1,1)}\right)$$

$$y^{(1,1)} + \lambda d^{(1,1)} = (20,20)^T + \lambda(1,-1)^T = (20+\lambda, 20+\lambda)^T$$

$$\varphi(\lambda) = f\left(y^{(1,1)} + \lambda d^{(1,1)}\right) = 3 \times (20+\lambda)^2$$

$$\frac{d\varphi}{d\lambda} = 60(20+\lambda) = 0, \lambda_1 = -20$$

$$y^{(1,2)} = x^{(1,1)} + \lambda_1 d^{(1,1)} = (0,0)^T$$

- **例 4_10**

从$\boldsymbol{y}^{(1,2)}$出发沿着方向$\boldsymbol{d}^{(1,2)}$作一维搜索。

$$\min_{\lambda} \boldsymbol{f}\left(\boldsymbol{y}^{(1,2)} + \boldsymbol{\lambda} \boldsymbol{d}^{(1,2)}\right)$$

$\boldsymbol{\lambda_2 = 0}$，$\boldsymbol{y}^{(1,3)} = \boldsymbol{y}^{(1,2)} + \boldsymbol{\lambda_2 d}^{(1,2)} = (\mathbf{0,0})^{\boldsymbol{T}}$ (坐标轮换结束)

创建方向$\boldsymbol{d}^{(1,3)} = \boldsymbol{y}^{(1,3)} - \boldsymbol{y}^{(1,1)} = (\mathbf{-20, -20})^{\boldsymbol{T}}$

求解 $\min_{\lambda} \boldsymbol{f}\left(\boldsymbol{x}^{(1,2)} + \boldsymbol{\lambda} \boldsymbol{d}^{(1,3)}\right)$

$\boldsymbol{\lambda_3 = 0}$，$\boldsymbol{x}^{\mathbf{1}} = \boldsymbol{y}^{(1,3)} + \boldsymbol{\lambda_3 d}^{(1,3)} = (\mathbf{0, 0})^{\boldsymbol{T}}$

- **例4_10**

$$f(x)=x_1^2+2x_2^2 \quad v^1=[1,-1]^T,\quad v^2=[1,1]^T,\quad x^0=[20,20]^T$$

$$\lambda_1=6.66, x^1=x^0+\lambda_1 v^1=[26.66,13.34]^T$$

$$\lambda_2=-17.8, x^2=x^1+\lambda_2 v^2=[8.86,-4.46]^T$$

$$u^1=x^2-x^0=[-11.14,-24.46]^T$$

$$\lambda_3=-0.15, x^3=x^2+\lambda_3 u^1=[10.48,-0.92]^T$$

$$\lambda_4=-3.46, x^4=x^3+\lambda_4 v^2=[7.02,-4.18]^T$$

$$\lambda_5=-0.15, x^5=x^4+\lambda_5 u^1=[8.68,-0.64]^T$$

$$u^2=x^5-x^3=[-1.8,0.28]^T$$

$$(u^1,Au^2)=[-11.14,-24.46]\begin{bmatrix}2 & 0\\ 0 & 4\end{bmatrix}\begin{bmatrix}-1.8\\ 0.28\end{bmatrix}=3\neq 0$$

- **powell算法运行条件**
  - 在此算法中，保持n个搜索方向线性无关十分重要。然而，Powell本人也注意到，他的方法可能选取接近线性相关的搜索方向，特别是变量很多时更是如此。这种可能性会给收敛性带来严重后果。
  - 为解决这个问题，提出了改进的powell算法。
  - 改进的算法，当初始搜索方向线性无关时，能够保证每轮迭代中以搜索方向为列的行列式不为零，因此这些方向是线性无关的。而且随着迭代的延续，搜索方向接近共轭的程度逐渐增加。
  - 改进的powell算法不再具有二次终止性，但是，它的计算效果仍然令人满意。

- **在牛顿法基础上的改进思路**
  - **1：** 采用共轭方向

    共轭梯度算法

    powell算法
  - **2：** 采用线性逼近

    变尺度法（**DFP**算法）

- **变尺度法**
  - 在牛顿法中，对于二次函数，是以方向$\boldsymbol{d}^k=-\boldsymbol{H}^{-1}\nabla\boldsymbol{f}(\boldsymbol{x}^k)$为寻优方向的。这种方向要求二阶逆，效率不高。
  - $\nabla\boldsymbol{f}\left(\boldsymbol{x}^0\right)=\nabla\boldsymbol{f}(\boldsymbol{x}^*)+\boldsymbol{H}(\boldsymbol{x}^*)\left(\boldsymbol{x}^0-\boldsymbol{x}^*\right)$

  $\boldsymbol{x}^*=\boldsymbol{x}^0-\boldsymbol{H}(\boldsymbol{x}^*)^{-1}\nabla\boldsymbol{f}(\boldsymbol{x}^0)$
  - 变尺度法是构造另一个矩阵$\overline{\boldsymbol{H}}(\boldsymbol{x})$,用它来逼近二阶导数矩阵的逆$\boldsymbol{H}^{-1}$，此法也称拟牛顿法。根据构造矩阵的方法不同，派生出不同的变尺度，这里介绍**DFP**法，即**Davidon-Fletcher-Powell**法。

变尺度算法（DFP）

- x^i为搜索点，H_i为H阵的近似。 $g^i = \nabla f(x^i)$
- 1：选取 x^0, H_0
- 2：for i=0,1,…n-1

  $x^{i+1} = x^i + \lambda_i v^i$,其中$v^i = -H_i g^i$，$\lambda_i$为一维搜索最优步长，$\min\limits_{\lambda} f(x^i + \lambda v^i)$
- 3：令 $u^i = x^{i+1} - x^i$ ， $y^i = g^{i+1} - g^i$,

  $$H_{i+1} = H_i + A_i - B_i, A_i = \frac{u^i (u^i)^T}{(u^i)^T y^i}, B_i = \frac{H_i y^i (H_i y^i)^T}{(y^i)^T H_i y^i}$$
- 4：判断精度，不满足则x^0=x^n,转2.

- **变尺度法–拟牛顿条件**
  - $\bar{H}(x)$ 构造原则是：它是一阶梯度 $\nabla f(x^k)$的函数，且接近各搜索点的$[\nabla^2 f(x^k)]^{-1}$，对于二次函数，我们有：

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

$$x^{k+1} - x^k = A^{-1}[\nabla f\left(x^{k+1}\right) - \nabla f\left(x^k\right)]$$

仿二次函数有：

$$x^{k+1} - x^k = \bar{H}^{k+1}\left[\nabla f\left(x^{k+1}\right) - \nabla f\left(x^k\right)\right] \qquad (1)$$

上式即为拟牛顿条件。

- **变尺度法–DFP算子**

$$\boldsymbol{x}^{k+1}-\boldsymbol{x}^{k}=\overline{\boldsymbol{H}}^{k+1}[\nabla \boldsymbol{f}(\boldsymbol{x}^{k+1})-\nabla \boldsymbol{f}(\boldsymbol{x}^{k})]$$

令：

$$\boldsymbol{y}^{k}=\Delta \boldsymbol{g}^{k}=\nabla \boldsymbol{f}(\boldsymbol{x}^{k+1})-\nabla \boldsymbol{f}(\boldsymbol{x}^{k})$$

$$\boldsymbol{u}^{k}=\Delta \boldsymbol{X}^{k}=\boldsymbol{x}^{k+1}-\boldsymbol{x}^{k}$$

$\boldsymbol{x}^{k+1}-\boldsymbol{x}^{k}=\overline{\boldsymbol{H}}^{k+1}\boldsymbol{y}^{k}=\boldsymbol{u}^{k}$

现假设$\overline{\boldsymbol{H}}^{k}$已知，则有：

$$\overline{\boldsymbol{H}}^{k+1}=\overline{\boldsymbol{H}}^{k}+\Delta\overline{\boldsymbol{H}}^{k}$$

$\Delta\overline{\boldsymbol{H}}^{k}$为第k次校正矩阵，并应满足拟牛顿条件：

- **变尺度法-DFP算子**

$$u^k = \overline{H}^{k+1} y^k$$

$$u^k = (\overline{H}^k + \Delta\overline{H}^k)\ \ y^k = \overline{H}^k y^k + \Delta\overline{H}^k y^k$$

$$\Delta\overline{H}^k y^k = u^k - \overline{H}^k y^k \qquad (2)$$

设$\Delta\overline{H}^k = u^k \cdot \left(Q^k\right)^T - \overline{H}^k y^k \cdot \left(W^k\right)^T$ **(**Q^k, W^k 为待定向量**)**

则应满足：

$$\left(Q^k\right)^T y^k = 1$$

$$\left(W^k\right)^T u^k = 1$$

$\Delta\overline{H}^k$为对称阵，简单的方法是取：

- **变尺度法-DFP算子**

$$Q^k = \mu_k u^k \qquad W^k = \varepsilon_k \bar{H}^k y^k$$

如此：

$$\left(\mu_k u^k\right)^T y^k = 1$$

$$\left(\varepsilon_k \bar{H}^k y^k\right)^T y^k = 1$$

$$\mu_k = \frac{1}{(u^k)^T y^k}$$

$$\varepsilon_k = \frac{1}{(y^k)^T \bar{H}^k y^k}$$

- **变尺度法-DFP算子**

$$\Delta\bar{H}^k = u^k\cdot\left(Q^k\right)^T - \bar{H}^k y^k\cdot\left(W^k\right)^T$$

$$= u^k\cdot\left(\mu_k u^k\right)^T - \bar{H}^k y^k\cdot\left(\varepsilon_k \bar{H}^k y^k\right)^T$$

$$= \frac{u^k\cdot\left(u^k\right)^T}{\left(u^k\right)^T y^k} - \frac{\bar{H}^k y^k\cdot\left(\bar{H}^k y^k\right)^T}{\left(y^k\right)^T \bar{H}^k y^k}$$

为 n 阶对称阵，由分母不为零。

$$\bar{H}^{k+1} = \bar{H}^k + \Delta\bar{H}^k$$

$$= \bar{H}^k + \frac{u^k\cdot\left(u^k\right)^T}{\left(u^k\right)^T y^k} - \frac{\bar{H}^k y^k\cdot\left(\bar{H}^k y^k\right)^T}{\left(y^k\right)^T \bar{H}^k y^k} \qquad (3)$$

- **例 4_11**

用DFP方法求$\min f(x)=\frac{1}{2}x^TAx$, $A=\begin{bmatrix}3 & -1\\ -1 & 1\end{bmatrix}$, $x^0=\begin{pmatrix}10\\10\end{pmatrix}$,

$$H_0=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$$

- **例 4_11**

用DFP方法求$\min f(x)=\frac{1}{2}x^TAx$, $A=\begin{bmatrix}3 & -1\\ -1 & 1\end{bmatrix}$, $x^0=\begin{pmatrix}10\\10\end{pmatrix}$,

$H_0=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$

解：$\nabla f(x)=\begin{pmatrix}3x_1-x_2\\ -x_1+x_2\end{pmatrix}$,

$\nabla f(x^0)=\begin{pmatrix}20\\0\end{pmatrix}$, $v^0=-H_0\nabla f(x^0)=-\begin{pmatrix}20\\0\end{pmatrix}$

从x^0出发沿着方向v^0作一维搜索。

$$\min_{\lambda} f\left(x^0+\lambda v^0\right)$$

$$\varphi(\lambda)=f\left(x^0+\lambda v^0\right)=100-400\lambda+600\lambda^2$$

$$\frac{d\varphi}{d\lambda}=-400+1200\lambda=0,\lambda_0=0.33$$

$$x^1=x^0+\lambda_0 v^0=(3.34,10)^T$$

- **例 4_11**

$$\nabla f(x^1) = \begin{pmatrix} 0.00 \\ 6.66 \end{pmatrix}, \quad u^0 = \lambda_0 v^0 = -\begin{pmatrix} 6.6 \\ 0 \end{pmatrix}$$

$$y^0 = \nabla f(x^1) - \nabla f(x^0) = -\begin{pmatrix} 20 \\ -6.66 \end{pmatrix}$$

$$A_i = \frac{u^i (u^i)^T}{(u^i)^T y^i}, \quad B_i = \frac{H_i y^i (H_i y^i)^T}{(y^i)^T H_i y^i}$$

$$A_0 = \begin{bmatrix} 0.33 & 0 \\ 0 & 0 \end{bmatrix}, \quad B_0 = -\begin{bmatrix} 0.9 & 0.3 \\ -0.3 & 0.01 \end{bmatrix}$$

$$H_{i+1} = H_i + A_i - B_i$$

$$H_1 = H_0 + A_0 - B_0 = \begin{bmatrix} 0.43 & 0.3 \\ 0.3 & 0.1 \end{bmatrix}$$

- **例 4_11**
  - $\boldsymbol{v}^1 = -\boldsymbol{H}_1 \nabla \boldsymbol{f}(\boldsymbol{x}^1) = -\begin{pmatrix} 2 \\ 6 \end{pmatrix}$

从$\boldsymbol{x}^1$出发沿着方向$\boldsymbol{v}^1$作一维搜索。

$$\min_{\lambda} \boldsymbol{f}(\boldsymbol{x}^1 + \lambda \boldsymbol{v}^1)$$

$$\lambda_1 = 5/3$$

$$\boldsymbol{x}^2 = \boldsymbol{x}^1 + \lambda_1 \boldsymbol{v}^1 = (-0,0)^T$$

$$\boldsymbol{H}_2 = \boldsymbol{H}_1 + \boldsymbol{A}_1 + \boldsymbol{B}_1 = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 1.5 \end{bmatrix}$$

$\boldsymbol{H}_2$ 有何特点？

- **作业**