第4章　无约束优化方法

求解无约束优化问题

$$\min f(X)$$

的数值迭代解法，称为无约束优化方法。

方法的基本问题是：　选择搜索方向

不同的搜索方向，构成不同的无约束优化算法。

方法分 ：　导数法和模式法两类

导数法：利用梯度和二阶导数构造搜索方向

如梯度法、牛顿法、

变尺度法、共扼梯度法等

模式法：利用某些点上的函数值构造搜索方向

如坐标轮换法、鲍威尔法等

4.1 梯度法

梯度法是一种古老的优化算法，它的迭代方向是由负梯度构成的，也称最速下降法。

梯度法的迭代算式是

$$S^k = -\nabla f\left(X^k\right)$$

$$X^{k+1} = X^k + a_k S^k$$

或

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

式中，α_k 为最优步长因子，由一维搜索确定，即

$$\min f\left(X^k - \alpha \nabla f\left(X^k\right)\right) \quad \rightarrow \quad \alpha_k$$

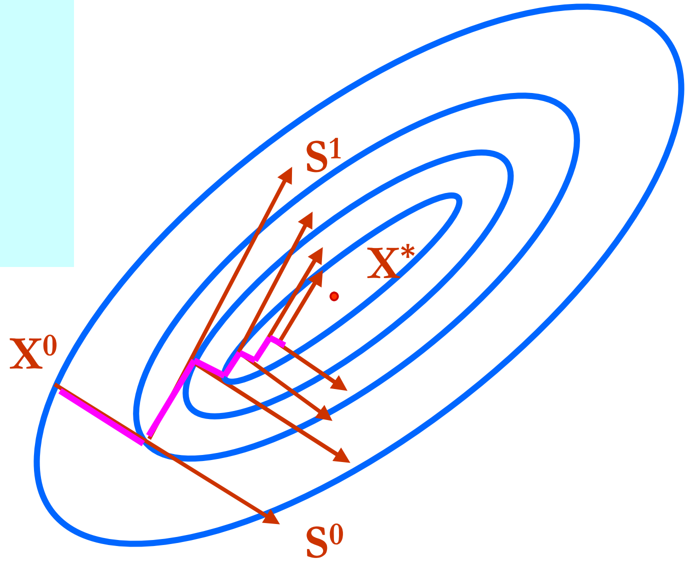
根据极值的必要条件和复合函数的求导公式，有

$$\nabla f\left(X^{k+1}\right) = -\left[\nabla f\left(X^k - \alpha_k \nabla f\left(X^k\right)\right)\right]^T \nabla f\left(X^k\right) = 0$$

$$\left[\nabla f\left(X^{k+1}\right)\right]^{T}\nabla f\left(X^{k}\right)=0$$

此式表明：

相邻两迭代点的梯度是彼此正交的，相邻的搜索方向相互垂直，如图所示。



由图可以看出：

① 向极小点的逼近路径是一条曲折的锯齿形路线，而且越接近极小点，前进速度越慢。

② 离极小点较远时，一次迭代得到的函数下降量较大。

因此，许多收敛性较好的算法，第一步迭代都采用梯度法。

梯度法的迭代步骤如下：

（1） 给定初始点X^0和收敛精度ε， 置k=0；

（2） 计算梯度，并构造搜索方向

$$S^k = -\nabla f\left(X^k\right)$$

（3） 一维搜索，求新的迭代点

$$\min f\left(X^k - \alpha \nabla f\left(X^k\right)\right) \quad \rightarrow \quad \alpha_k$$

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

（4） 收敛判断：若满足

$$\left\|\nabla f\left(X^{k+1}\right)\right\| \le \varepsilon$$

则令 $X^*=X^{k+1}$，$f(X^*)=f(X^{k+1})$终止计算；

否则，令k=k+1，转（2）继续迭代。

例4-1 用梯度法求解无约束最优化问题

$$\min f(X)=x_1^2+2x_2^2-2x_1x_2-4x_1$$

已知 $X^0=[1,\ 1]^T,\quad \varepsilon=0.1$

解（1）第一次迭代

求 $\nabla f(X)=\begin{bmatrix}2x_1-2x_2-4\\-2x_1+4x_2\end{bmatrix},\quad \nabla f\left(X^0\right)=\begin{bmatrix}-4\\2\end{bmatrix}$

令 $S^0=-\nabla f\left(X^0\right)=\begin{bmatrix}4\\-2\end{bmatrix}$

则有 $X^1=X^0+\alpha_0S^0=\begin{bmatrix}1\\1\end{bmatrix}+\alpha_0\begin{bmatrix}4\\-2\end{bmatrix}=\begin{bmatrix}1+4\alpha_0\\1-2\alpha_0\end{bmatrix}$

$$f\left(X^1\right)=\left(1+4\alpha_0\right)^2+2\left(1-2\alpha_0\right)^2-2\left(1+4\alpha_0\right)\left(1-2\alpha_0\right)-4\left(1+4\alpha_0\right)=f\left(\alpha_0\right)$$

令 $f'\left(\alpha_0\right)=8\left(1+4\alpha_0\right)-8\left(1-2\alpha_0\right)-8\left(1-2\alpha_0\right)+4\left(1+4\alpha_0\right)-16=0$

解得

$$\alpha_0=\frac{1}{4}=0.25\,,\quad X^1=\begin{bmatrix}2\\0.5\end{bmatrix},\quad f\left(X^1\right)=-5.5$$

因 $\left\|\nabla f\left(X^1\right)\right\|=\sqrt{5}>\varepsilon$ 还需继续迭代

（2）第二次迭代

同理有

$$\nabla f\left(X^1\right)=\begin{bmatrix}-1\\-2\end{bmatrix},\quad S^1=\begin{bmatrix}1\\2\end{bmatrix}$$

$$X^2=X^1+\alpha_1 S^1=\begin{bmatrix}2\\0.5\end{bmatrix}+\alpha_1\begin{bmatrix}1\\2\end{bmatrix}=\begin{bmatrix}2+\alpha_1\\0.5+2\alpha_1\end{bmatrix}$$

$$f\left(X^2\right)=\left(2+\alpha_1\right)^2+2\left(0.5+2\alpha_1\right)^2-2\left(2+\alpha_1\right)\left(0.5+2\alpha_1\right)-4\left(2+\alpha_1\right)=f\left(\alpha_1\right)$$

$$f'\left(\alpha_1\right)=2\left(2+\alpha_1\right)+8\left(0.5+2\alpha_1\right)-2\left(0.5+2\alpha_1\right)-4\left(2+\alpha_1\right)-4=0$$

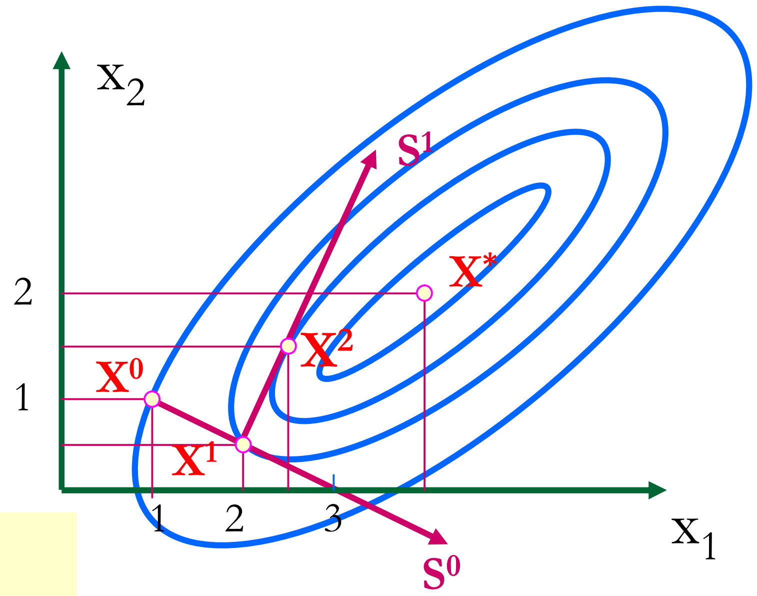
解得 $\alpha_1 = 0.5$

$$X^2 = \begin{bmatrix} 2.5 \\ 1.5 \end{bmatrix}, \quad f\left(X^2\right) = -6.75$$

$$\nabla f\left(X^2\right) = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

因 $\left\|\nabla f\left(X^2\right)\right\| = \sqrt{5} > \varepsilon$

可知 X^2 也不是极小点，还应继续进行迭代。



此问题的最优解是

$X^*=[4，2]^T$，$f（X^*）=-8$

用梯度法求解时，需要经过相当多次迭代才能得到一个近似的最优解。其中前两次的迭代路线如图所示。

4.2 牛顿法

牛顿法的搜索方向由目标函数的负梯度和二阶导数矩阵构造。

4. 2.1 基本牛顿法

将函数 f (X)在点X^k处展成泰勒二次式：

$$f(X) \approx f\left(X^{(k)}\right) + \left[\nabla f\left(X^{(k)}\right)\right]^T \left[X - X^k\right] + \frac{1}{2}\left[X - X^k\right]^T \nabla^2 f\left(X^k\right)\left[X - X^k\right]$$

对上式求梯度，并设Ｘ（k+1）是函数的极小点，则有

$$\nabla f\left(X^{k+1}\right) = \nabla f\left(X^k\right) + \nabla^2 f\left(X^k\right)\left[X^{k+1} - X^k\right] = 0$$

由此解得 $$X^{k+1} = X^k - \left[\nabla^2 f\left(X^k\right)\right]^{-1} \nabla f\left(X^k\right)$$

令 $$S^k = -\left[\nabla^2 f\left(X^k\right)\right]^{-1} \nabla f\left(X^k\right)$$ 则有 $$X^{k+1} = X^k + S^k$$

由此构成的算法称基本牛顿法，S^k称牛顿方向。

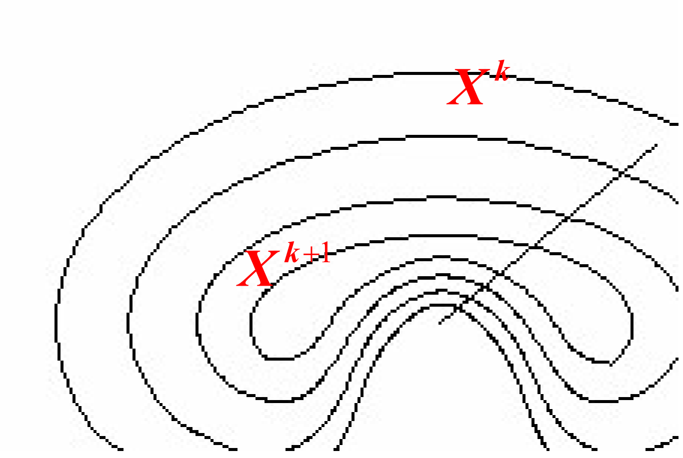
分析可知：

(1) 对于正定二次函数，X^{k+1}是精确极小点，方向 S^k 是直指函数的极小点。

(2) 用基本牛顿法求解正定二次函数时，无论从哪个初始点出发，计算所得牛顿方向直指极小点，而且步长等于1。

(3) 对于一般非线性函数，点X^{k+1}只是原函数的一个近似极小点。故将此点作为下一个迭代X^{k+1}。

(4) 但是对于非正定函数，由上式得到的点X^{k+1}，不能始终保持函数的下降性，基本牛顿法有可能失败。如图所示，



原因：步长因子总是等于1，

改进方法：引入步长因子和一维搜索。

由此构成如下阻尼牛顿法。

4.2.2 阻尼牛顿法

令

$$
\begin{aligned}
& S^k = -\left[\nabla^2 f\left(X^k\right)\right]^{-1} \nabla f\left(X^k\right) \\
& \min \quad f\left(X^k + \alpha S^k\right) \quad \rightarrow \quad \alpha_k \\
& X^{k+1} = X^k + \alpha_k S^k
\end{aligned}
$$

由此建立的算法就是阻尼牛顿法，α_k 称阻尼因子

牛顿法的特点：

(1) 构造搜索方向利用了函数的所有的一阶导数和二阶导数，产生的搜索方向直指极小点；

(2) 对于正定二次函数，从任意初始点出发，一次迭代即可得到极小点。对于一般函数，迭代的次数比其它算法都要少。

(3) 每次迭代都需要计算函数的二阶导数矩阵及其逆矩阵。计算量大，计算时间长，计算速度较慢。

(4) 计算中，用差分代替微分。所得梯度和二阶导数矩阵，以及求得的牛顿方向都存在较大误差。

例4-2 用牛顿法求解例4-1

$$\min f(X) = x_1^2 + 2x_2^2 - 2x_1x_2 - 4x_1$$

已知 $X^0 = [1,\ 1]^T,\quad \varepsilon = 0.1$

解 1） 用基本牛顿法

计算 $\nabla f(X) = \begin{bmatrix} 2x_1 - 2x_2 - 4 \\ -2x_1 + 4x_2 \end{bmatrix}$

$$\nabla^2 f(X) = \begin{bmatrix} 2 & -2 \\ -2 & 4 \end{bmatrix}, \qquad \nabla f\left(X^0\right) = \begin{bmatrix} -4 \\ 2 \end{bmatrix}$$

$$\left[\nabla^2 f(X)\right]^{-1} = \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$$

$$S^0 = -\left[\nabla^2 f\left(X^0\right)\right]^{-1} \nabla f\left(X^0\right) = -\begin{bmatrix} 1 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}\begin{bmatrix} -4 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

$$X^1 = X^0 + S^0 = \begin{bmatrix}1\\1\end{bmatrix} + \begin{bmatrix}3\\1\end{bmatrix} = \begin{bmatrix}4\\2\end{bmatrix},\quad f\left(X^1\right) = -8$$

因 $\nabla f\left(X^1\right) = \begin{bmatrix}0\\0\end{bmatrix},\quad \left\|\nabla f\left(X^1\right)\right\| = 0$

故最优解是： $X^* = X^1 = \begin{bmatrix}4\\2\end{bmatrix},\quad f\left(X^*\right) = -8$

2） 用阻尼牛顿法

$$X^1 = X^0 + \alpha_0 S^0 = \begin{bmatrix}1\\1\end{bmatrix} + \alpha_0\begin{bmatrix}3\\1\end{bmatrix} = \begin{bmatrix}1+3\alpha_0\\1+\alpha_0\end{bmatrix}$$

代入原函数

$$f\left(X^1\right) = \left(1+3\alpha_0\right)^2 + 2\left(1+\alpha_0\right)^2 - 2\left(1+3\alpha_0\right)\left(1+\alpha_0\right) - 4\left(1+3\alpha_0\right) = f\left(\alpha_0\right)$$

对 α 求导

$$f'\left(\alpha_0\right) = 6\left(1+3\alpha_0\right) + 4\left(1+\alpha_0\right) - 6\left(1+\alpha_0\right) - 2\left(1+3\alpha_0\right) - 12 = 0$$

解得 $\alpha_0 = 1$

$$X^1 = X^0 + \alpha_0 S^0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$$

因 $f\left(X^1\right) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \left\| \nabla f\left(X^1\right) \right\| = 0$

最优解为 $X^* = X^1 = \begin{bmatrix} 4 \\ 2 \end{bmatrix}, \ f\left(X^*\right) = -8$

可以看出：

此解与基本牛顿法的计算结果完全相同。

两种方法都只进行了一次迭代，这是因为标函数是正定二次函数，基本牛顿法的局限性还没有显露出来的缘故。

基本牛顿法不需一维搜索，因此计算速度较快。

4.2.3 变尺度法

基本思想：

（1） 用简单矩阵代替二阶导数矩阵的逆矩阵 $\left[\nabla^2 f\left(X^k\right)\right]^{-1}$

（2） 用坐标变换简化目标函数

引入矩阵变换U，令 $X - X^k = UY$

代入式泰勒展开式得

$$\phi(Y) = \frac{1}{2} Y^T U^T \nabla^2 f\left(X^k\right) UY + \left[\nabla f\left(X^k\right)\right]^T UY + f\left(X^k\right)$$

若 $\nabla f\left(X^k\right)$ 正定，必存在矩阵U

使 $$U^T \nabla^2 f\left(X^k\right) U = I \qquad (4-12)$$

代入前式得

$$\phi(Y)=\frac{1}{2}Y^TY+\nabla^2 f\left(X^k\right)^T UY+f\left(X^k\right)$$

上式中只包含Y的二次项和一次项，采用坐标平移可消去一次项，使其等值线变为同心椭圆（球）。

用 U 和 U^{-1} 分别左乘和右乘式(4－12)，有

$$U^T\nabla^2 f\left(X^k\right)U=I$$

$$UU^T\nabla^2 f\left(X^k\right)=I$$

$$\left[\nabla^2 f\left(X^k\right)\right]^{-1}=UU^T$$

由此可见，通过坐标变换可以得到 $\left[\nabla^2 f\left(X^k\right)\right]^{-1}$

将上式代入牛顿法的迭代算式得

$$S^k = -UU^T \nabla f\left(X^k\right)$$

$$X^{k+1} = X^k - \alpha_k UU^T \nabla f\left(X^k\right)$$

记矩阵$A = -UU^T$，称变尺度矩阵，则有

$$S^k = -A^k \nabla f\left(X^k\right)$$

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

由以上两式构成的下降迭代解法称变尺度法。

其中 $S^k = -A^k \nabla f\left(X^k\right)$ 称变尺度方向。

经推导，变尺度矩阵可由以下递推公式得到：

$$\boldsymbol{A}^{k+1} = \boldsymbol{A}^{k} + \boldsymbol{E}^{k}$$

其中，**A⁰=I**（单位矩阵）；**E**^{k}称校正矩阵

$$\boldsymbol{E}^{k} = \frac{\Delta \boldsymbol{X}^{k}\left[\Delta \boldsymbol{X}^{k}\right]^{T}}{\left[\Delta \boldsymbol{g}^{k}\right]^{T}\Delta \boldsymbol{X}^{k}} - \frac{\boldsymbol{A}^{k}\Delta \boldsymbol{g}^{k}\left[\Delta \boldsymbol{g}^{k}\right]^{T}\boldsymbol{A}^{k}}{\left[\Delta \boldsymbol{g}^{k}\right]^{T}\boldsymbol{A}^{k}\Delta \boldsymbol{g}^{k}}$$

式中

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

$$\Delta \boldsymbol{g}^{k} = \nabla \boldsymbol{f}\left(\boldsymbol{X}^{k+1}\right) - \nabla \boldsymbol{f}\left(\boldsymbol{X}^{k}\right)$$

4.2.4 共轭梯度法

4.2.4.1 共轭方向

设**H**为正定对称矩阵，有一组非零向量 $S_1, S_2, \cdots, S_n$ 满足

$$S_i^T H S_j = 0 \quad (i \neq j)$$

则称这组向量关于矩阵**H**共轭，或是**H**的一组共轭向量。

当**H**为单位矩阵时，即有

$$S_i^T S_j = 0 \quad (i \neq j)$$

此时称向量 S_n（i=1，2，…，n）相互正交。

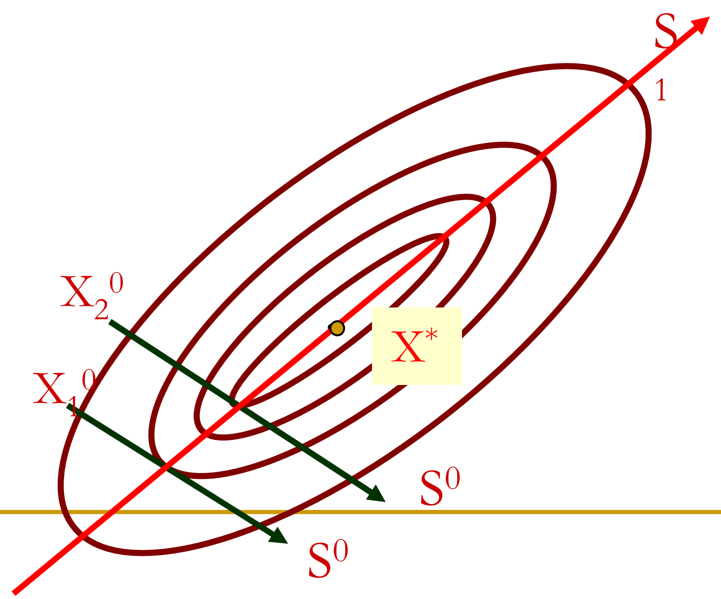
共轭方向具有以下性质：

（1）对于正定二次函数, 从任意初始点X^0出发，依次沿一组（n个）共轭方向一维搜索，最多n次即可以达到极小点。

（2）从任意两点 X_1^0 和 X_2^0 出发，分别沿同一方向 S^0 进行一维搜索，得到两个一维极小点 X_1^1 和 X_2^1，则连接此两点所形成的向量

$$S^1 = X_1^1 - X_2^1$$

与原方向S^0关于该函数的二阶导数矩阵相共轭。



4.2.4.2 共轭方向的产生

平行搜索法　　向量组合法。

1 .平行搜索法

见后面的鲍威尔法

2． 向量组合法

（1）基向量组合法

取n个基向量e^{i}（i=0，1，…，n-1）和另一个独立向量S^0，令

$$S^1 = e^0 + \beta_0 S^0 \qquad (4\text{-}26)$$

式中，β_0为待定常数。称 S^0 是 e^0 和 S^1 的线性组合。

欲使 **S⁰** 和 **S¹** 关于函数的二阶导数矩阵相共轭，则必须使

$$\left[S^0\right]^T \nabla^2 f\left(X^k\right) S^1 = 0 \qquad (4\text{-}27)$$

将式（**4-26**）代入上式解得

解得

$$\beta_0 = -\frac{\left[S^0\right]^T \nabla^2 f\left(X^k\right) e^0}{\left[S^0\right]^T \nabla^2 f\left(X^k\right) S^0}$$

代入式（**4-26**）得到与 **S⁰** 共轭的向量

$$S^1 = e^0 - \frac{\left[S^0\right]^T \nabla^2 f\left(X^k\right) e^0}{\left[S^0\right]^T \nabla^2 f\left(X^k\right) S^0} \cdot S^0$$

同理，若已求得 **k+1** 个共轭向量 S^0，S^1，…，S^k，令新的向量

$$S^{k+1} = e^k + \sum_{i=0}^{k} \beta_i S^i$$

解得

$$\beta_i = -\frac{\left[S^i\right]^T \nabla^2 f\left(X^k\right) e^i}{\left[S^i\right]^T \nabla^2 f\left(X^k\right) S^i}$$

于是，有

$$S^{k+1} = e^k - \sum_{i=0}^{k} \frac{\left[S^i\right]^T \nabla^2 f\left(X^k\right) e^i}{\left[S^i\right]^T \nabla^2 f\left(X^k\right) S^i} S^i$$

此向量就是与原 **k+1** 个向量相共轭的新向量。

4.2.5 梯度组合法

将函数f (X)展开为如下二次函数

$$f(X)=\frac{1}{2}X^{T}HX+B^{T}X+c$$

从任意点 X^k 出发，沿函数的负梯度方向作一次一维搜索

令 $$S^{k}=-\nabla f\left(X^{k}\right) \tag{4-34}$$

得 $$X^{k+1}=X^{k}+\alpha_{k}S^{k} \tag{4-35}$$

又令 $$S^{k+1}=-\nabla f\left(X^{k+1}\right)+\beta_{k}S^{k} \tag{4-36}$$

把式（**4-34**）和式（**4-36**）代入共轭条件

$$\left[S^{k+1}\right]^{T}HS^{k}=0 \tag{4-37}$$

$$-\left[\nabla f\left(X^{k}\right)\right]^{T}H\left[-\nabla f\left(X^{k+1}\right)+\beta_{k}S^{k}\right]=0$$

由此解得

$$\beta_k = \frac{\left[\nabla f\left(X^k\right)\right]^T H \nabla f\left(X^{k+1}\right)}{\left[\nabla f\left(X^k\right)\right]^T H \nabla f\left(X^k\right)}$$

由于

$$\nabla f\left(X^k\right) = HX^k + B$$

$$\nabla f\left(X^{k+1}\right) = HX^{k+1} + B$$

两式相减并将式（**4-35**）代入得

$$a_k HS^k = \nabla f\left(X^{k+1}\right) - \nabla f\left(X^k\right)$$

将式（**4-36**）和上式的两边分别相乘，并注意式（**4-37**）得

$$\left[-\nabla f\left(X^{k+1}\right) + \beta_k \nabla f\left(X^k\right)\right]^T \left[\nabla f\left(X^{k+1}\right) - \nabla f\left(X^k\right)\right] = 0$$

将上式展开，并注意到相邻两点梯度间的正交关系，整理后得

$$\beta_k = \frac{\left[\nabla f\left(X^{k+1}\right)\right]^T \nabla f\left(X^{k+1}\right)}{\left[\nabla f\left(X^{k}\right)\right]^T \nabla f\left(X^{k}\right)} = \frac{\left\|\nabla f\left(X^{k+1}\right)\right\|^2}{\left\|\nabla f\left(X^{k}\right)\right\|^2}$$

4.3 共轭梯度法

把上式代入式（**4-36**）得

$$S^{k+1} = -\nabla f\left(X^{k+1}\right) + \frac{\left\|\nabla f\left(X^{k+1}\right)\right\|^2}{\left\|\nabla f\left(X^{k}\right)\right\|^2} S^k$$

$$X^{k+2} = X^{k+1} + \alpha_{k+1} S^{k+1}$$

可见，只需利用相邻两点的梯度就可以构造一个共轭方向。

以这种方式产生共轭方向并进行迭代运算的算法称共轭梯度法。

例4-4 用共轭梯度法求解例4-1。

解（1） 第一次迭代沿负梯度方向搜索。由例4-1得

$$X^0=\begin{bmatrix}1\\1\end{bmatrix},\ \nabla f\left(X^0\right)=\begin{bmatrix}-4\\2\end{bmatrix},\ S^0=\begin{bmatrix}4\\-2\end{bmatrix}$$

$$X^1=\begin{bmatrix}2\\0.5\end{bmatrix},\ \nabla f\left(X^1\right)=\begin{bmatrix}-1\\-2\end{bmatrix}$$

（2） 第二次迭代。求

$$\beta_0=\frac{\left\|\nabla f\left(X^1\right)\right\|^2}{\left\|\nabla f\left(X^0\right)\right\|^2}=\frac{5}{20}=0.25$$

$$S^1=-\nabla f\left(X^1\right)+\beta_0 S^0=-\begin{bmatrix}-1\\-2\end{bmatrix}+0.25\begin{bmatrix}4\\-2\end{bmatrix}=\begin{bmatrix}2\\1.5\end{bmatrix}$$

$$X^2=X^1+\alpha_1 S^1=\begin{bmatrix}2\\0.5\end{bmatrix}+\alpha_1\begin{bmatrix}2\\1.5\end{bmatrix}=\begin{bmatrix}2+2\alpha_1\\0.5+1.5\alpha_1\end{bmatrix}$$

$$f\left(X^2\right)=\left(2+2\alpha_1\right)^2+2\left(0.5+1.5\alpha_1\right)^2-2(2+2\alpha_1)(0.5+1.5\alpha_1)-4(2+2\alpha_1)=f(\alpha_1)$$

$$f'\left(\alpha_1\right)=4\left(2+2\alpha_1\right)+6\left(0.5+1.5\alpha_1\right)-4(0.5+1.5\alpha_1)-3(2+2\alpha_1)-8=0$$

解得 $\alpha_1=1$

$$X^2=\begin{bmatrix}4\\2\end{bmatrix},\quad f\left(X^2\right)=-8$$

$$\nabla f\left(X^2\right)=\begin{bmatrix}0\\0\end{bmatrix}$$

因 $\left\|\nabla f\left(X^2\right)\right\|=0$

所以 $X^2=\left[4,2\right]^T$ 和 $f\left(X^2\right)=-8$ 就是所求最优解。

4.4 鲍威尔（Powell）法

利用平行搜索构造共轭方向，并沿共轭方向一维搜索的算法。由于方向的产生不需要计算导数，因此属于模式法。

鲍威尔法是模式法中最好的，具有超线性收敛速度。

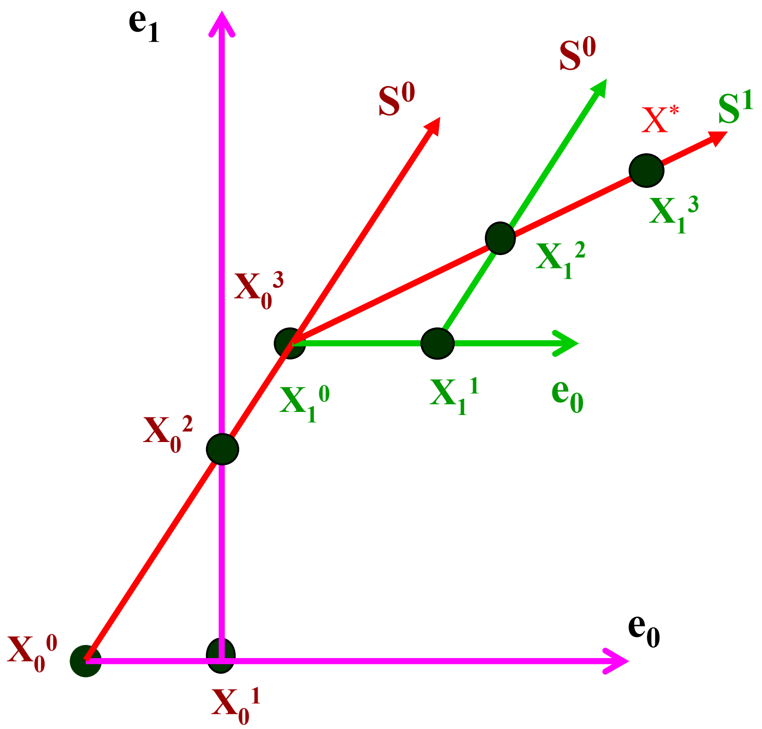
4.4.1 基本思想

从两个初始点出发，沿同一个方向，分别进行两次一維搜索，就可产生一个与原方向共轭的方向。

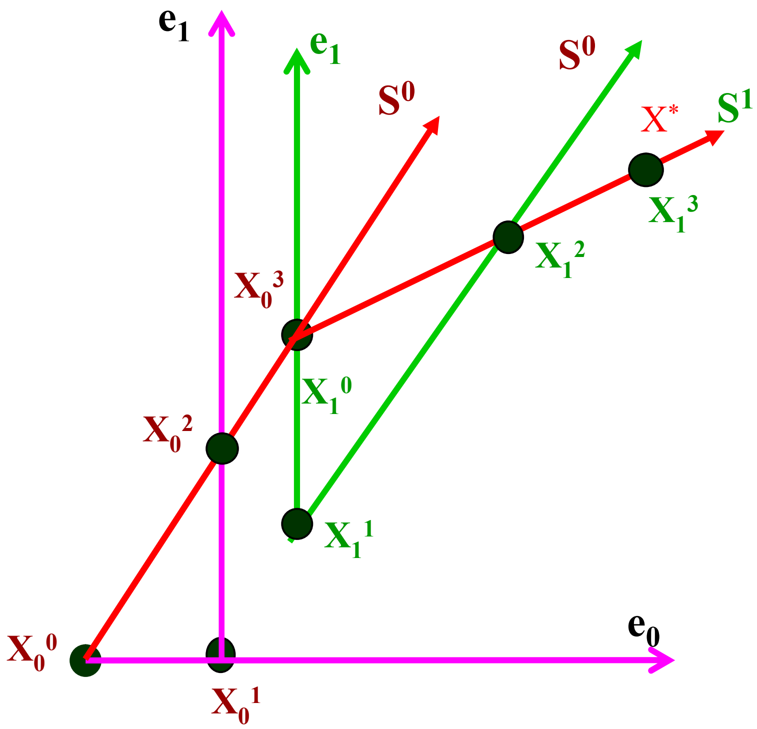
但算法只能给定一个初始点 X^0 。

故实现平行搜索的第二个初始点和一个任意方向，只能借助坐标方向在搜索中逐渐产生。

4.4.2 迭代过程



e_1
S^0
S^0
X^*
S^1
$X_1^{\ 3}$
$X_1^{\ 2}$
$X_0^{\ 3}$
$X_1^{\ 0}$
$X_1^{\ 1}$
e_0
$X_0^{\ 2}$
$X_0^{\ 0}$
e_0
$X_0^{\ 1}$



e_1
e_1
S^0
S^0
X^*
S^1
$X_1^{\ 3}$
$X_0^{\ 3}$
$X_1^{\ 2}$
$X_1^{\ 0}$
$X_0^{\ 2}$
$X_1^{\ 1}$
$X_0^{\ 0}$
e_0
$X_0^{\ 1}$

4.4.3 算法特点

（1）对于正定二元二次函数，经过两轮迭代，6次一維搜索即可达到极小点。

对正定 n 元二次函数，最多经过 n 轮迭代，n (n+1) 次一維搜索就可达到极小点。

（2）不需计算函数的导数。

（3）一維搜索的次数较多，但构成搜索方向的计算极其简单。