第三章 非线性规划

第四节 无约束优化问题的解法

- 最速下降法
- *Newton*法
- 拟*Newton*法
- 共轭梯度法

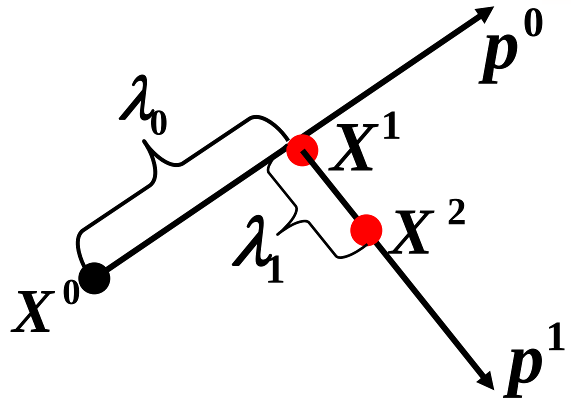
复习

最速下降法迭代原理：

$$\min_{X\in R^n} f(X)$$

λ_0——最优步长

λ_1——最优步长



$$X^0,\quad p^0=-\nabla f(X^0),\ \min_{\lambda\ge 0} f(X^0+\lambda p^0)=f(X^0+\lambda_0 p^0),\ X^1=X^0+\lambda_0 p^0$$

$$X^1,\quad p^1=-\nabla f(X^1),\ \min_{\lambda\ge 0} f(X^1+\lambda p^1)=f(X^1+\lambda_1 p^1),\ X^2=X^1+\lambda_1 p^1$$

$$X^k,\quad p^k=-\nabla f(X^k),\ \min_{\lambda\ge 0} f(X^k+\lambda p^k)=f(X^k+\lambda_k p^k), X^{k+1}=X^k+\lambda_k p^k$$

得到一个点列：$X^0, X^1, \cdots, X^{(k)}, \cdots,$

可以证明：$f(X^{(k)})$严格 $\downarrow$, $\therefore\ X^{(k)} \xrightarrow[k\to\infty]{} X^*$ $(Th3-10)$,

线性收敛 $(Th3-11)$

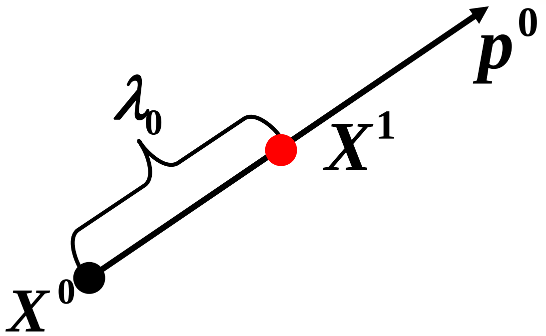
结论： 最速下降法的任何两个相邻搜索方向正交：$p^{(k+1)^T} p^{(k)}=0$

复习

最速下降法迭代原理：

$$\min_{X\in R^n} f(X)= x_1^4 + x_2^2 + 2$$

$$\nabla f(X) = (4x_1^3, 2x_2)^T$$



$$X^0,\quad p^0 = -\nabla f(X^0),\ \min_{\lambda\ge 0} f(X^0+\lambda p^0) = f(X^0+\lambda_0 p^0),\ X^1 = X^0+\lambda_0 p^0$$

$$X^0 = (1,1)^T,\ p^0 = -\nabla f(X^0) = -(4,2)^T$$

$$X^0+\lambda p^0 = \begin{pmatrix}1\\1\end{pmatrix} - \lambda\begin{pmatrix}4\\2\end{pmatrix} = \begin{pmatrix}1-4\lambda\\1-2\lambda\end{pmatrix}$$

$$f(X^0+\lambda p^0) = (1-4\lambda)^4 + (1-2\lambda)^2 + 2 = F(\lambda)$$

$$\min_{\lambda\ge 0} f(X^0+\lambda p^0) = F(\lambda)$$

λ	$F(\lambda)$
0	4
0.5	3
1	84

一维搜索找极小点 λ_0： 1)确定[0 , 1], 精度0.1

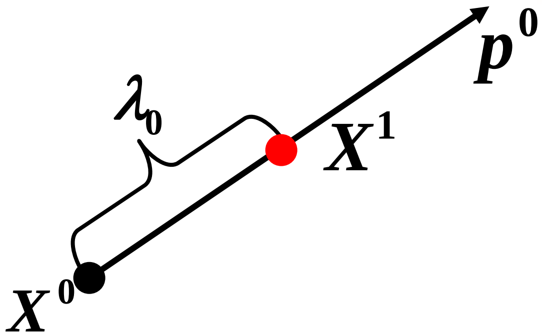
2)用0.618法得到 $\lambda_0 = 0.34375$

最速下降法迭代原理：

复习

$$\min_{X\in R^n} f(X)= x_1^4 + x_2^2 + 2$$

$$\nabla f(X) = (4x_1^3, 2x_2)^T$$



$$X^0,\quad p^0 = -\nabla f(X^0),\ \min_{\lambda\ge 0} f(X^0+\lambda p^0) = f(X^0+\lambda_0 p^0),\ X^1 = X^0+\lambda_0 p^0$$

$$X^0 = (1,1)^T,\ p^0 = -\nabla f(X^0) = -(4,2)^T$$

$$X^0+\lambda p^0 = \begin{pmatrix}1\\1\end{pmatrix} - \lambda\begin{pmatrix}4\\2\end{pmatrix} = \begin{pmatrix}1-4\lambda\\1-2\lambda\end{pmatrix}$$

$$f(X^0+\lambda p^0) = (1-4\lambda)^4 + (1-2\lambda)^2 + 2 = F(\lambda)$$

$$\min_{\lambda\ge 0} f(X^0+\lambda p^0) = F(\lambda) \qquad \lambda_0 = 0.34375$$

$$X^1 = (1,1)^T - 0.34375(4,2)^T = (-0.375, 0.3125)^T$$

$$f(X_1) = 2.11743 < f(X_0) = 4$$

第三章 非线性规划

第四节 无约束优化问题的解法

- 最速下降法
- *Newton*法
- 拟*Newton*法
- 共轭梯度法

二. *Newton*法

$$(NP)\quad \min_{X\in R^n} f(X)$$

- *Newton*法的迭代原理
- *Newton*法的收敛结论
- *Newton*法注释
- *Newton*法的优缺点

1. *Newton*法的迭代原理

设 $f(X)$在上R^n具有连续的二阶偏导数，

即 $G(X)=\nabla^2 f(X)$ 连续，且$G(X^*)$正定。

在X^*的邻域内，$G(X)$也正定。

求解 $\min\limits_{X\in R^n} f(X)$ 的局部最优解 X^*

$$\nabla^2 f(X)=\nabla(\nabla f(X))=\nabla\begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}=\begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

1. *Newton*法的迭代原理 $X^{(k)} \longrightarrow X^{(k+1)}$

在$X^{(k)}$处用$f(X)$的正定二次函数来近似$f(X)$:

一元函数泰勒公式:

$$f(x) = f(x_k) + f'(x_k)(x - x_k) + \frac{1}{2} f''(x_k)(x - x_k)^2 + o(x - x_k)^2$$

$$f(x) \cong f(x_k) + f'(x_k)(x - x_k) + \frac{1}{2} f''(x_k)(x - x_k)^2 = \varphi(x)$$

$$f(X) \cong f(X^{(k)}) + g(X^{(k)})^T (X - X^{(k)}) + \frac{1}{2}(X - X^{(k)})^T G(X^{(k)})(X - X^{(k)}) = \varphi(X)$$

$$g(X^{(k)}) = \nabla f(X^{(k)}) \qquad G(X^{(k)}) = \nabla^2 f(X^{(k)})$$

1. *Newton*法的迭

$$f(X)=\frac{1}{2}X^TQX+b^T$$

用 $\varphi(X)$ 近似 $f(X)$

用 $\varphi(X)$ 的极小点 X^{k+1}

近似 $f(X)$ 的极小点 X^*

在$X^{(k)}$处用正定二次函数来近似$f(X$

$$f(x)\cong f(x_k)+f'(x_k)(x-x_k)+\frac{1}{2}f''(x_k)(x$$

$$f(X)\cong f(X^{(k)})+g(X^{(k)})^T(X-X^{(k)})+\frac{1}{2}(X-X^{(k)})^TG(X^{(k)})(X-X^{(k)})=\varphi(X)$$

$\because G(X^*)$正定，且$G(X)$连续 $\therefore$ 在X^*附近，$G(X^{(k)})$也正定.

$\therefore \varphi(X)$为正定二次函数．求$\varphi(X)$的极小点：

$$\nabla\varphi(X)=G(X^{(k)})(X-X^{(k)})+g(X^{(k)})=0 \longrightarrow$$

$$G(X^{(k)})(X-X^{(k)})=-g(X^{(k)}) \longrightarrow X-X^{(k)}=-G(X^{(k)})^{-1}g(X^{(k)})$$

$$X=X^{(k)}-G(X^{(k)})^{-1}g(X^{(k)})$$

*Newton*迭代公式

$$X^{(k+1)}=X^{(k)}-G(X^{(k)})^{-1}g(X^{(k)})$$

二. *Newton*法

$$(NP)\ \min_{X\in R^n} f(X)$$

- *Newton*法的迭代原理
- *Newton*法的收敛结论
- *Newton*法注释
- *Newton*法的优缺点

2. *Newton*法的收敛结论

$$g(X^{(k)}) = \nabla f(X^{(k)})$$

$$G(X^{(k)}) = \nabla^2 f(X^{(k)})$$

$$X^{(k+1)} = X^{(k)} - G(X^{(k)})^{-1} g(X^{(k)})$$

*Newton*法产生的点列：$X^{(0)}, X^{(1)}, X^{(2)}, \cdots, X^{(k)}, \cdots$

$X^{(0)}$

$$X^{(1)} = X^{(0)} - G(X^{(0)})^{-1} g(X^{(0)})$$

$$X^{(2)} = X^{(1)} - G(X^{(1)})^{-1} g(X^{(1)})$$

$$X^{(3)} = X^{(2)} - G(X^{(2)})^{-1} g(X^{(2)})$$

收敛结论：

当$X^{(0)}$充分靠近X^*时，$X^{(k)} \xrightarrow[k\to\infty]{\text{二阶收敛}} X^*$

因此, *Newton*法往往与最速下降法结合使用, 前期使用最速下降法, 后期使用*Newton*法.

$$X^{(k+1)} = X^{(k)} - G(X^{(k)})^{-1} g(X^{(k)})$$

例3-10 用Newton法求 $f(X) = x_1^2 + 4x_2^2$ 的极小点，

$$X^{(0)} = (1,1)^T, \varepsilon = 10^{-4} \quad X^* = (0,0)^T$$

解：

$$X^{(1)} = X^{(0)} - G(X^{(0)})^{-1} g(X^{(0)})$$

$$g(X) = \nabla f(X) = \begin{pmatrix} 2x_1 \\ 8x_2 \end{pmatrix} \quad G(X) = \nabla^2 f(X) = \begin{pmatrix} 2 & 0 \\ 0 & 8 \end{pmatrix}$$

$$g(X^{(0)}) = \nabla f(X^{(0)}) = \begin{pmatrix} 2 \\ 8 \end{pmatrix} \quad G(X^{(0)}) = \nabla^2 f(X^{(0)}) = \begin{pmatrix} 2 & 0 \\ 0 & 8 \end{pmatrix}$$

$$X^{(1)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 2 & 0 \\ 0 & 8 \end{pmatrix}^{-1} \begin{pmatrix} 2 \\ 8 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 1/2 & 0 \\ 0 & 1/8 \end{pmatrix} \begin{pmatrix} 2 \\ 8 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = X^*$$

二. *Newton*法

$$(NP)\ \min_{X\in R^n} f(X)$$

- *Newton*法的迭代原理
- *Newton*法的收敛结论
- *Newton*法注释
- *Newton*法的优缺点

3. *Newton*法注释

$$X^{(k+1)}=X^{(k)}-G(X^{(k)})^{-1}g(X^{(k)})$$

$$X^{(k+1)}=X^{(k)}+\lambda_k p^{(k)}$$

1^0 由于*Newton*迭代公式中没有使用一维

即$X^{(k+1)}=X^{(k)}+p^{(k)},\quad p^{(k)}=-G(X^{(k)})^{-1}g(X^{(k)})$

即 $\lambda_k=1,\therefore$ 不能保证$\{f(X^{(k)})\}$严格$\downarrow$

$\therefore$ 不能保证当 $X^{(0)}$ 远离极小点X^*时算法收敛。

复习：要想保证$\{f(X^{(k)})\}$严格↓，需要保证两条：

1）$p^{(k)}$是$f(X)$在 $X^{(k)}$处的下降方向，即 $\nabla f(X^{(k)})^T p^{(k)}<0$

2）$\min\limits_{\lambda\ge 0} f(X^k+\lambda p^k)=f(X^k+\lambda_k p^k)$，$X^{k+1}=X^k+\lambda_k p^k \quad f(X^{k+1})<f(X^k)$

$p^{(k)}=-G(X^{(k)})^{-1}g(X^{(k)})$是$f(X)$在 $X^{(k)}$处的下降方向：

$$\nabla f(X^{(k)})^T p^{(k)}=-g(X^{(k)})^T G(X^{(k)})^{-1}g(X^{(k)})<0$$

$\because G(X^*)$正定，且$G(X)$连续$\therefore$ 在X^*附近，$G(X^{(k)})$也正定.

3. *Newton*法注释

$$X^{(k+1)}=X^{(k)}-G(X^{(k)})^{-1}g(X^{(k)})$$

$$X^{(k+1)}=X^{(k)}+\lambda_k p^{(k)}$$

1^0 由于*Newton*迭代公式中没有使用一维

即$X^{(k+1)}=X^{(k)}+p^{(k)},\quad p^{(k)}=-G(X^{(k)})^{-1}g(X^{(k)})$

即 $\lambda_k=1$,∴ 不能保证$\{f(X^{(k)})\}$严格↓

∴ 不能保证当 $X^{(0)}$ 远离极小点X^*时算法收敛。

2^0 为了保证在 $X^{(0)}$ 远离极小点的地方算法的收敛性，

取 $p^{(k)}=-G(X^{(k)})^{-1}g(X^{(k)})$ 为搜索方向，

作一维搜索：$\min\limits_{\lambda\geq 0} f(X^{(k)}+\lambda p^{(k)})=f(X^{(k)}+\lambda_k p^{(k)})$,

$$X^{(k+1)}=X^{(k)}+\lambda_k p^{(k)}$$

这样得到的算法称为阻尼*Newton*法.(P158)

3. *Newton*法注释

3^0 *Newton*法的搜索方向： $p^{(k)}=-G(X^{(k)})^{-1}g(X^{(k)})$,

最速下降法的搜索方向：$p^{(k)}=-g(X^{(k)})$

*Newton*法的搜索方向与最速下降法相比较，不仅利用了梯度的信息，还利用了*Hesse*矩阵的信息，从而提高了收敛速度(二阶收敛)。

4^0 对于正定二次目标函数，使用*Newton*法只需迭代一次就可求出它的极小点。所以*Newton*法具有二次终止性。

$$X^{(k+1)} = X^{(k)} - G(X^{(k)})^{-1} g(X^{(k)})$$

例3-10 用Newton法求 $f(X) = x_1^2 + 4x_2^2$ 的极小点，

$$X^{(0)} = (1,1)^T, \varepsilon = 10^{-4} \quad X^* = (0,0)^T$$

解：

$$X^{(1)} = X^{(0)} - G(X^{(0)})^{-1} g(X^{(0)})$$

$$g(X) = \nabla f(X) = \begin{pmatrix} 2x_1 \\ 8x_2 \end{pmatrix} \quad G(X) = \nabla^2 f(X) = \begin{pmatrix} 2 & 0 \\ 0 & 8 \end{pmatrix}$$

$$g(X^{(0)}) = \nabla f(X^{(0)}) = \begin{pmatrix} 2 \\ 8 \end{pmatrix} \quad G(X^{(0)}) = \nabla^2 f(X^{(0)}) = \begin{pmatrix} 2 & 0 \\ 0 & 8 \end{pmatrix}$$

$$X^{(1)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 2 & 0 \\ 0 & 8 \end{pmatrix}^{-1} \begin{pmatrix} 2 \\ 8 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 1/2 & 0 \\ 0 & 1/8 \end{pmatrix} \begin{pmatrix} 2 \\ 8 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = X^*$$

1. *Newton*法的迭代原理 $X^{(k)} \longrightarrow X^{(k+1)}$

在$X^{(k)}$处用正定二次函数来近似$f(X)$:

$$f(x) \cong f(x_k) + f'(x_k)(x-x_k) + \frac{1}{2} f''(x_k)(x-x_k)^2 = \varphi(x)$$

$$f(X) \cong f(X^{(k)}) + g(X^{(k)})^T (X-X^{(k)}) + \frac{1}{2}(X-X^{(k)})^T G(X^{(k)})(X-X^{(k)}) = \varphi(X)$$

$\because G(X^*)$正定,且$G(X)$连续$\therefore$ 在X^*附近,$G(X^{(k)})$也正定.

$\therefore \varphi(X)$为正定二次函数．求$\varphi(X)$的极小点：

$$\nabla\varphi(X) = G(X^{(k)})(X-X^{(k)}) + g(X^{(k)}) = 0 \longrightarrow$$

$$G(X^{(k)})(X-X^{(k)}) = -g(X^{(k)}) \longrightarrow X - X^{(k)} = -G(X^{(k)})^{-1} g(X^{(k)})$$

$$X = X^{(k)} - G(X^{(k)})^{-1} g(X^{(k)})$$

*Newton*迭代公式

$$X^{(k+1)} = X^{(k)} - G(X^{(k)})^{-1} g(X^{(k)})$$

二. *Newton*法

$$(NP)\ \min_{X\in R^n} f(X)$$

- *Newton*法的迭代原理
- *Newton*法的收敛结论
- *Newton*法注释
- *Newton*法的优缺点

4. *Newton*法的优缺点

$$X^{(k+1)} = X^{(k)} - G(X^{(k)})^{-1} g(X^{(k)})$$

优点： $X^{(k)} \xrightarrow[k\to\infty]{\text{二阶收敛}} X^{*}$

*Newton*法最突出的优点是收敛速度快.

缺点： 每次迭代都要计算目标函数的*Hesse*阵的逆阵,当维数n较大时,计算量迅速增加,这就抵消了*Newton*法的优点.

二. *Newton*法

$$(NP)\ \min_{X \in R^n} f(X)$$

- *Newton*法的迭代原理
- *Newton*法的收敛结论
- *Newton*法注释
- *Newton*法的优缺点

第三章 非线性规划

第四节 无约束优化问题的解法

- 最速下降法
- *Newton*法
- 拟*Newton*法
- 共轭梯度法

三. 拟*Newton*法

*Newton*迭代公式: $$X^{(k+1)} = X^{(k)} - G(X^{(k)})^{-1} g(X^{(k)})$$

*Newton*法搜索方向: $$p^{(k)} = -G(X^{(k)})^{-1} g(X^{(k)})$$

拟*Newton*法搜索方向: $$p^{(k)} = -H^{(k)} g(X^{(k)})$$

$H^{(k)}$是 $G(X^{(k)})^{-1}$ 的近似

拟*Newton*法是一大类算法的总称,其中使用最多的是**DFP**方法和**BFGS**方法.

三. 拟*Newton*法----*DFP*法（Davidon-Fletcher-Powll）

$$(NP)\quad \min_{X\in R^n} f(X)$$

- *DFP*法的搜索方向
- *DFP*法的迭代步骤
- *DFP*法举例
- *DFP*法的性质
- *DFP*法的收敛结论

$$f(X)\cong f(X^{(k)})+g(X^{(k)})^T(X-X^{(k)})+\frac{1}{2}(X-X^{(k)})^T G(X^{(k)})(X-X^{(k)})=\varphi(X)$$

问题： 如何确定 $H^{(k)}$?

$$H^{(k)}\rightarrow H^{(k+1)}\ ?$$

$H^{(k)}$是 $G(X^{(k)})^{-1}$的近似

$H^{(k+1)}$是$G(X^{(k+1)})^{-1}$的近似

$$f(X)\cong f(X^{(k+1)})+g(X^{(k+1)})^T(X-X^{(k+1)})+\frac{1}{2}(X-X^{(k+1)})^T G(X^{(k+1)})(X-X^{(k+1)})$$

$$\nabla f(X)\cong g(X^{(k+1)})+G(X^{(k+1)})(X-X^{(k+1)}$$

$$\nabla f(X^{(k)})\cong g(X^{(k+1)})+G(X^{(k+1)})(X^{(k)}-X$$

$\because G(X^*)$正定,且$G(X)$连续

$\therefore$ 在X^*附近,$G(X^{(k+1)})$也正定.

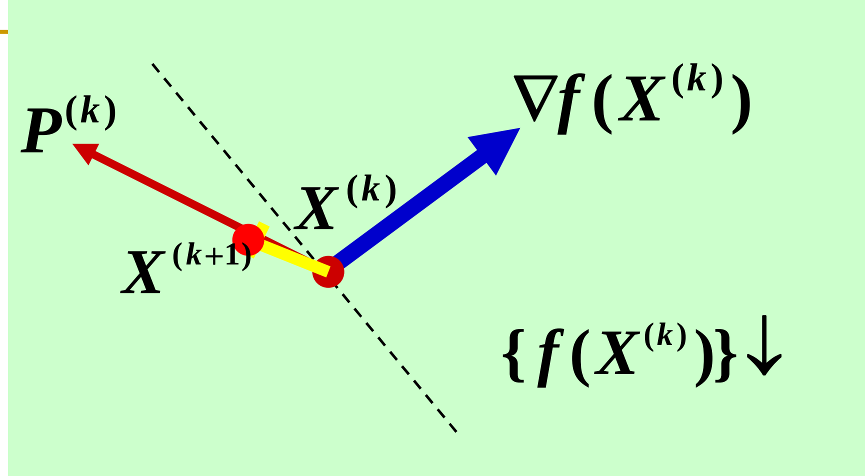
$$g^{(k)}\cong g^{(k+1)}+G^{(k+1)}(X^{(k)}-X^{(k+1)})$$

$$g^{(k+1)}-g^{(k)}\cong G^{(k+1)}(X^{(k+1)}-X^{(k)})\quad (G^{(k+1)}\text{正定})$$

$$G^{(k+1)^{-1}}\underbrace{(g^{(k+1)}-g^{(k)})}_{Y^{(k)}}\cong \underbrace{X^{(k+1)}-X^{(k)}}_{S^{(k)}}$$

$$G^{(k+1)^{-1}}Y^{(k)}\approx S^{(k)}\qquad \boxed{H^{(k+1)}Y^{(k)}=S^{(k)}}\ (3-35)$$ **称为拟*Newton*方程**

1. DFP法的搜索方向



问题： 如何确定 $H^{(k)}$?

$$H^{(k)} \to H^{(k+1)}?$$

1^0 $H^{(k+1)}$满足拟*Newton*方程

2^0 为使搜索方向 $p^{(k)} = -H^{(k)} g^{(k)}$ 是下降方向，要求$H^{(k)}$对称正定。

结论： 当 $\nabla f(X^{(k)})^T p^{(k)} < 0$ 时，$p^{(k)}$是$f(X)$在$X^{(k)}$处的下降方向。

证明： 当 $H^{(k)}$ 对称正定时，$g^{(k)^T} p^{(k)} = -g^{(k)^T} H^{(k)} g^{(k)} < 0,$

所以 $p^{(k)}$ 是 $f(X)$ 在 $X^{(k)}$ 处的下降方向。

三. 拟Newton法

复习

$$S^{(k-1)} = X^{(k)} - X^{(k-1)}$$

$$Y^{(k-1)} = g^{(k)} - g^{(k-1)}$$

迭代原理：

Newton迭代公式： $X^{(k+1)} = X^{(k)} - G(X^{(k)})^{-1} g(X^{(k)})$

Newton法搜索方向： $p^{(k)} = -G(X^{(k)})^{-1} g(X^{(k)})$

拟Newton法搜索方向： $p^{(k)} = -H^{(k)} g(X^{(k)})$

$H^{(k)}$是 $G(X^{(k)})^{-1}$的近似

要求： $H^{(k)} Y^{(k-1)} = S^{(k-1)}$

1^0 $H^{(k)}$满足拟Newton方程; (保证算法收敛快 – 超线性收敛)

2^0 $H^{(k)}$对称正定; $g^{(k)^T} p^{(k)} = -g^{(k)^T} H^{(k)} g^{(k)} < 0,$

(保证搜索方向 $p^{(k)}$是下降方向 $\rightarrow f(X^{(k)}) \downarrow \rightarrow X^{(k)} \rightarrow X^*$)

结论： 当$\nabla f(X^{(k)})^T p^{(k)} < 0$时, $p^{(k)}$是$f(X)$在$X^{(k)}$处的下降方向

1. DFP法的搜索方向

$$p^{(k)} = -H^{(k)} g^{(k)}$$

$H^{(k)}$是$G(X^{(k)})^{-1}$的近似

$H^{(k+1)}$是$G(X^{(k+1)})^{-1}$的近似

$$H^{(k+1)} Y^{(k)} = S^{(k)}$$

问题： 如何确定 $H^{(k)}$?

$$H^{(k)} \to H^{(k+1)} ?$$

1⁰ $H^{(k+1)}$满足拟*Newton*方程

2⁰ 为使搜索方向 $p^{(k)} = -H^{(k)} g^{(k)}$ 是下降方向，要求$H^{(k)}$对称正定。

3⁰ 要求 $H^{(k+1)} - H^{(k)} = \Delta H^{(k)}$ 越简单越好，即秩越小越好。

*DFP*方法如何确定 $\Delta H^{(k)}$：

$$H^{(k+1)} - H^{(k)} = \Delta H^{(k)} \underset{令}{=} \alpha_k U^{(k)} U^{(k)^T} + \beta_k V^{(k)} V^{(k)^T}$$

$$H^{(k+1)} = H^{(k)} + \alpha_k U^{(k)} U^{(k)^T} + \beta_k V^{(k)} V^{(k)^T}$$

其中 $U^{(k)}, V^{(k)}$是n维待定列向量，α_k, β_k是待定常数。

可以证明： $\Delta H^{(k)}$ 是一个秩2对称矩阵。

1. DFP法的搜索方向

$$H^{(k+1)}Y^{(k)}=S^{(k)}$$

$$Y^{(k)}=g^{(k+1)}-g^{(k)}$$

$$S^{(k)}=X^{(k+1)}-X^{(k)}$$

*DFP*方法如何确定 $\Delta H^{(k)}$:

$$H^{(k+1)}=H^{(k)}+\alpha_k U^{(k)}U^{(k)^T}+\beta_k V^{(k)}V^{(k)^T}$$

$$H^{(k+1)}Y^{(k)}=H^{(k)}Y^{(k)}+\alpha_k U^{(k)}\underline{U^{(k)^T}Y^{(k)}}+\beta_k V^{(k)}\underline{V^{(k)^T}Y^{(k)}}$$

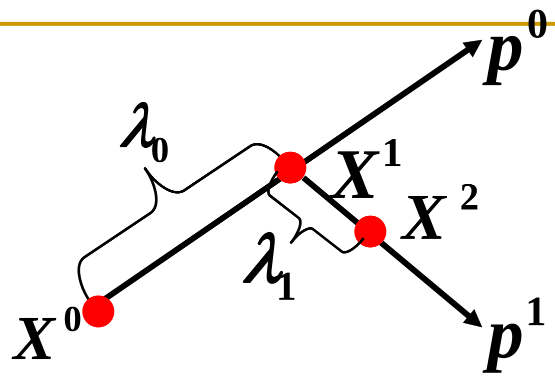
$$=H^{(k)}Y^{(k)}+\underbrace{\alpha_k(U^{(k)^T}Y^{(k)})}_{=1}\underbrace{U^{(k)}}_{S^{(k)}}+\underbrace{\beta_k(V^{(k)^T}Y^{(k)})}_{=-1}\underbrace{V^{(k)}}_{H^{(k)}Y^{(k)}}$$

$$\alpha_k=\frac{1}{U^{(k)T}Y^{(k)}}=\frac{1}{S^{(k)T}Y^{(k)}}\qquad \beta_k=-\frac{1}{V^{(k)^T}Y^{(k)}}=-\frac{1}{Y^{(k)^T}H^{(k)}Y^{(k)}}$$

$$H^{(k+1)}=H^{(k)}-\frac{H^{(k)}Y^{(k)}Y^{(k)^T}H^{(k)}}{Y^{(k)^T}H^{(k)}Y^{(k)}}+\frac{S^{(k)}S^{(k)^T}}{S^{(k)^T}Y^{(k)}}\qquad(3-38)$$

*DFP*公式

1. DFP法的搜索方向



$$p^{(k)} = -H^{(k)}g^{(k)}$$

$$H^{(k+1)} = H^{(k)} - \frac{H^{(k)}Y^{(k)}Y^{(k)^T}H^{(k)}}{Y^{(k)^T}H^{(k)}Y^{(k)}} + \frac{S^{(k)}S^{(k)^T}}{S^{(k)^T}Y^{(k)}} \quad (3-38)$$

***DFP*公式**

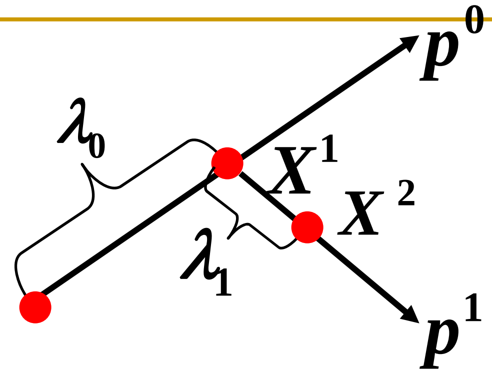
$X^{(0)}$, $H^{(0)} = E$　$p^{(0)} = -H^{(0)}g^{(0)} = -g^{(0)}$　一维搜索得 $X^{(1)}$

$X^{(1)}$,　$p^{(1)} = -H^{(1)}g^{(1)}$　一维搜索得 $X^{(2)}$

$$S^{(0)} = X^{(1)} - X^{(0)} \quad Y^{(0)} = g^{(1)} - g^{(0)}$$

$$H^{(1)} = H^{(0)} - \frac{H^{(0)}Y^{(0)}Y^{(0)^T}H^{(0)}}{Y^{(0)^T}H^{(0)}Y^{(0)}} + \frac{S^{(0)}S^{(0)^T}}{S^{(0)^T}Y^{(0)}}$$

1. DFP法的搜索方向



$$p^{(k)} = -H^{(k)} g^{(k)}$$

$$H^{(k+1)} = H^{(k)} - \frac{H^{(k)} Y^{(k)} Y^{(k)^T} H^{(k)}}{Y^{(k)^T} H^{(k)} Y^{(k)}} + \frac{S^{(k)} S^{(k)^T}}{S^{(k)^T} Y^{(k)}} \quad (3-38)$$

***DFP*公式**

$X^{(0)},\ H^{(0)} = E \quad p^{(0)} = -H^{(0)} g^{(0)} = -g^{(0)}$ 一维搜索得 $X^{(1)}$

$X^{(1)}, \quad p^{(1)} = -H^{(1)} g^{(1)}$ 一维搜索得 $X^{(2)}$

$X^{(2)}, \quad p^{(2)} = -H^{(2)} g^{(2)}$ 一维搜索得 $X^{(3)}$

$$S^{(1)} = X^{(2)} - X^{(1)} \quad Y^{(1)} = g^{(2)} - g^{(1)}$$

$$H^{(2)} = H^{(1)} - \frac{H^{(1)} Y^{(1)} Y^{(1)^T} H^{(1)}}{Y^{(1)^T} H^{(1)} Y^{(1)}} + \frac{S^{(1)} S^{(1)^T}}{S^{(1)^T} Y^{(1)}}$$

三. 拟*Newton*法 ---- *DFP*法

$$(NP)\ \min_{X \in R^n} f(X)$$

- *DFP*法的搜索方向
- *DFP*法的迭代步骤
- *DFP*法举例
- *DFP*法的性质
- *DFP*法的收敛结论

2. DFP法的迭代步骤:

1^0 给定初始点 $X^{(1)}$,容许误差 $\varepsilon > 0$;

2^0 检验 $\left\|g(X^{(1)})\right\| < \varepsilon$? 若满足,则迭代终止,取 $X^* = X^{(1)}$,否则转 3^0

3^0 取 $H^{(1)} = E$,令 $k=1$ $\qquad$ 4^0 令 $p^{(k)} = -H^{(k)} g^{(k)}$

5^0 一维搜索: $\min\limits_{\lambda \geq 0} f(X^{(k)} + \lambda p^{(k)}) = f(X^{(k)} + \lambda_k p^{(k)})$

6^0 令 $X^{(k+1)} = X^{(k)} + \lambda_k p^{(k)}$

7^0 检验 $\left\|g(X^{(k+1)})\right\| < \varepsilon$? 若满足,则迭代终止,取 $X^* = X^{(k+1)}$,

否则若 $k = n$,令 $X^{(1)} := X^{(k+1)}$,转 3^0

若 $k < n$,令 $Y^{(k)} = g^{(k+1)} - g^{(k)}$, $S^{(k)} = X^{(k+1)} - X^{(k)}$,

$$\text{计算} H^{(k+1)} = H^{(k)} - \frac{H^{(k)} Y^{(k)} Y^{(k)T} H^{(k)}}{Y^{(k)T} H^{(k)} Y^{(k)}} + \frac{S^{(k)} S^{(k)T}}{S^{(k)T} Y^{(k)}}, \text{转} 8^0$$

8^0 令 $k := k+1$,转 4^0

$n=3$

$X^{(1)}$, $H^{(1)}=E$ $k:=1$ $p^{(1)}=-H^{(1)}g^{(1)}=-g^{(1)}$ $X^{(2)}=X^{(1)}+\lambda_1 p^{(1)}$

$X^{(2)}$, $k=1<n=3$ $p^{(2)}=-H^{(2)}g^{(2)}$ $X^{(3)}=X^{(2)}+\lambda_2 p^{(2)}$

$S^{(1)}=X^{(2)}-X^{(1)}$

$$H^{(2)}=H^{(1)}-\frac{H^{(1)}Y^{(1)}Y^{(1)T}H^{(1)}}{Y^{(1)T}H^{(1)}Y^{(1)}}+\frac{S^{(1)}S^{(1)T}}{S^{(1)T}Y^{(1)}}$$

$Y^{(1)}=g^{(2)}-g^{(1)}$

$k:=k+1=2$ 转4^0

$X^{(3)}$, $k=2<n=3$ $p^{(3)}=-H^{(3)}g^{(3)}$ $X^{(4)}=X^{(3)}+\lambda_3 p^{(3)}$

$S^{(2)}=X^{(3)}-X^{(2)}$

$$H^{(3)}=H^{(2)}-\frac{H^{(2)}Y^{(2)}Y^{(2)T}H^{(2)}}{Y^{(2)T}H^{(2)}Y^{(2)}}+\frac{S^{(2)}S^{(2)T}}{S^{(2)T}Y^{(2)}}$$

$Y^{(2)}=g^{(3)}-g^{(2)}$

$k:=k+1=3$ 转4^0

$X^{(4)}$, $k=3=n=3$ $X^{(1)}:=X^{(4)}$转3^0 $k:=1$ $p^{(1)}=-H^{(1)}g^{(1)}=-g^{(1)}$

2^0 $\|g(X^{(1)})\|<\varepsilon$?若是, $X^*=X^{(1)}$, 否则转 3^0 3^0 $H^{(1)}=E, k=1$ 4^0 $p^{(k)}=-H^{(k)}g^{(k)}$

5^0 $\min\limits_{\lambda\ge 0} f(X^{(k)}+\lambda p^{(k)})=f(X^{(k)}+\lambda_k p^{(k)})$ 6^0 $X^{(k+1)}=X^{(k)}+\lambda_k p^{(k)}$

7^0 $\|g(X^{(k+1)})\|<\varepsilon$? 若是, $X^*=X^{(k+1)}$, 否则若$k=n$, 令$X^{(1)}:=X^{(k+1)}$, 转3^0

若$k<n$, 令$Y^{(k)}=g^{(k+1)}-g^{(k)}$, $S^{(k)}=X^{(k+1)}-X^{(k)}$, $H^{(k+1)}=H^{(k)}-\dfrac{H^{(k)}Y^{(k)}Y^{(k)T}H^{(k)}}{Y^{(k)T}H^{(k)}Y^{(k)}}+\dfrac{S^{(k)}S^{(k)T}}{S^{(k)T}Y^{(k)}}$, 转$8^0$

8^0 令$k:=k+1$, 转4^0

2. DFP法的迭代步骤：

3^0 取 $H^{(1)} = E$, 令 $k=1$ $\quad$ 4^0 令 $p^{(k)} = -H^{(k)} g^{(k)}$

5^0 一维搜索 $\min\limits_{\lambda \ge 0} f(X^{(k)} + \lambda p^{(k)}) = f(X^{(k)} + \lambda_k p^{(k)})$ $\quad n = 3$

6^0 令 $X^{(k+1)} = X^{(k)} + \lambda_k p^{(k)}$

7^0 检验 $\|g(X^{(k+1)})\| < \varepsilon$?

否则若 $\underline{k = n}$, 令 $X^{(1)} := X^{(k+1)}$, 转 3^0

若 $\underline{k < n}$, $H^{(k)} \xrightarrow{DFP公式} H^{(k+1)}$, 8^0

8^0 令 $k := k+1$, 转 4^0

$$X^{(1)},\ p^{(1)} = -H^{(1)} g^{(1)} = -g^{(1)}$$
$$X^{(2)},\ p^{(2)} = -H^{(2)} g^{(2)}$$
$$X^{(3)},\ p^{(3)} = -H^{(3)} g^{(3)}$$

$$X^{(4)},\ p^{(4)} = -H^{(4)} g^{(4)} = -g^{(4)}$$
$$X^{(5)},\ p^{(5)} = -H^{(5)} g^{(5)}$$
$$X^{(6)},\ p^{(6)} = -H^{(6)} g^{(6)}$$

每n次迭代中的第一步取负梯度方向为其搜索方向,这种做法简称为 **“n步重新开始”**.这是为了减少舍入误差的影响(舍入误差可能导致某个 $H^{(k)}$ 不可逆或不正定)

三. 拟*Newton*法 ---- *DFP*法

$$(NP)\ \min_{X \in R^n} f(X)$$

- ✓ *DFP*法的搜索方向
- ✓ *DFP*法的迭代步骤
- ➡ *DFP*法举例
- *DFP*法的性质
- *DFP*法的收敛结论

3. DFP法举例

$n=2$

例3-13　用*DFP*方法求解 $\min f(X)=x_1^2+4x_2^2=\dfrac{1}{2}X^TQX$，

取 $X^{(1)}=(1,1)^T$ 为初始点，$\varepsilon=0.01$

解：$Q=\begin{pmatrix}2&0\\0&8\end{pmatrix}$　$g(X)=\begin{pmatrix}2x_1\\8x_2\end{pmatrix}$　$g^{(1)}=\begin{pmatrix}2\\8\end{pmatrix}$　$\left\|g^{(1)}\right\|=\sqrt{68}\not<\varepsilon$

2^0 $\|g(X^{(1)})\| < \varepsilon$?若是,$X^* = X^{(1)}$, 否则转 3^0 3^0 $H^{(1)} = E, k = 1$ 4^0 $p^{(k)} = -H^{(k)} g^{(k)}$ $X^{(1)} = (1,1)^T$

5^0 $\min_{\lambda \ge 0} f(X^{(k)} + \lambda p^{(k)}) = f(X^{(k)} + \lambda_k p^{(k)})$ 6^0 $X^{(k+1)} = X^{(k)} + \lambda_k p^{(k)}$

7^0 $\|g(X^{(k+1)})\| < \varepsilon$? 若是,$X^* = X^{(k+1)}$, 否则若$k = n$,令$X^{(1)} := X^{(k+1)}$,转$3^0$

若$k < n$,令$Y^{(k)} = g^{(k+1)} - g^{(k)}, S^{(k)} = X^{(k+1)} - X^{(k)}$, $H^{(k+1)} = H^{(k)} - \dfrac{H^{(k)} Y^{(k)} Y^{(k)T} H^{(k)}}{Y^{(k)T} H^{(k)} Y^{(k)}} + \dfrac{S^{(k)} S^{(k)T}}{S^{(k)T} Y^{(k)}}$,转$8^0$

8^0 令$k := k+1$,转4^0

解: $Q = \begin{pmatrix} 2 & 0 \\ 0 & 8 \end{pmatrix}$ $g(X) = \begin{pmatrix} 2x_1 \\ 8x_2 \end{pmatrix}$ $g^{(1)} = \begin{pmatrix} 2 \\ 8 \end{pmatrix}$ $\|g^{(1)}\| = \sqrt{68} \not< \varepsilon$

$k = 1$ $H^{(1)} = E$, $p^{(1)} = -H^{(1)} g^{(1)} = -g^{(1)} = -\begin{pmatrix} 2 \\ 8 \end{pmatrix}$

所以*DFP*方法与最速下降法具有相同的第一个迭代点。

$X^{(2)} = \begin{pmatrix} 0.73846 \\ -0.04616 \end{pmatrix}$ $g^{(2)} = \begin{pmatrix} 1.47692 \\ -0.36923 \end{pmatrix}$ $\|g^{(2)}\| = 1.52237 \not< \varepsilon$

$\because k = 1 < n = 2 \therefore$ 计算$H^{(2)}$: $S^{(1)} = X^{(2)} - X^{(1)} = \begin{pmatrix} -0.26154 \\ -1.04616 \end{pmatrix}$

$S^{(1)} S^{(1)T} = \begin{pmatrix} 0.06840 & 0.27361 \\ 0.27361 & 1.09445 \end{pmatrix}$ $Y^{(1)} = g^{(2)} - g^{(1)} = \begin{pmatrix} -0.52308 \\ -8.36923 \end{pmatrix}$

$S^{(1)T} Y^{(1)} = 8.89236$

2^0 $\|g(X^{(1)})\|<\varepsilon$?若是, $X^*=X^{(1)}$, 否则转 3^0 3^0 $H^{(1)}=E, k=1$ 4^0 $p^{(k)}=-H^{(k)}g^{(k)}$ $X^{(1)}=(1,1)^T$

5^0 $\min\limits_{\lambda\geq 0} f(X^{(k)}+\lambda p^{(k)})=f(X^{(k)}+\lambda_k p^{(k)})$ 6^0 $X^{(k+1)}=X^{(k)}+\lambda_k p^{(k)}$

7^0 $\|g(X^{(k+1)})\|<\varepsilon$? 若是, $X^*=X^{(k+1)}$, 否则若 $k=n$, 令 $X^{(1)}:=X^{(k+1)}$, 转 3^0

若 $k<n$, 令 $Y^{(k)}=g^{(k+1)}-g^{(k)}, S^{(k)}=X^{(k+1)}-X^{(k)}$, $H^{(k+1)}=H^{(k)}-\dfrac{H^{(k)}Y^{(k)}Y^{(k)T}H^{(k)}}{Y^{(k)T}H^{(k)}Y^{(k)}}+\dfrac{S^{(k)}S^{(k)T}}{S^{(k)T}Y^{(k)}}$, 转 8^0

8^0 令 $k:=k+1$, 转 4^0

解: $Q=\begin{pmatrix}2 & 0\\ 0 & 8\end{pmatrix}$ $g(X)=\begin{pmatrix}2x_1\\ 8x_2\end{pmatrix}$ $g^{(1)}=\begin{pmatrix}2\\ 8\end{pmatrix}$ $\|g^{(1)}\|=\sqrt{68}\not<\varepsilon$

$k=1$ $\because k=1<n=2 \therefore$ 计算 $H^{(2)}$:

$$S^{(1)}S^{(1)^T}=\begin{pmatrix}0.06840 & 0.27361\\ 0.27361 & 1.09445\end{pmatrix} \quad S^{(1)^T}Y^{(1)}=8.89236 \quad Y^{(1)}=g^{(2)}-g^{(1)}=\begin{pmatrix}-0.52308\\ -8.36923\end{pmatrix}$$

$$H^{(1)}Y^{(1)}Y^{(1)^T}H^{(1)}=Y^{(1)}Y^{(1)^T}=\begin{pmatrix}0.27361 & 4.37778\\ 4.37778 & 70.4401\end{pmatrix}$$

$$Y^{(1)^T}H^{(1)}Y^{(1)}=Y^{(1)^T}Y^{(1)}=70.31762$$

$$H^{(2)}=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}-\frac{1}{70.31}\begin{pmatrix}0.27 & 4.37\\ 4.37 & 70.4\end{pmatrix}+\frac{1}{8.89}\begin{pmatrix}0.06 & 0.27\\ 0.27 & 1.09\end{pmatrix}=\begin{pmatrix}1.00 & -0.03\\ -0.03 & 0.12\end{pmatrix}$$

2^0 $\|g(X^{(1)})\|<\varepsilon$?若是,$X^*=X^{(1)}$, 否则转 3^0 3^0 $H^{(1)}=E, k=1$ 4^0 $p^{(k)}=-H^{(k)}g^{(k)}$ $X^{(1)}=(1,1)^T$

5^0 $\min\limits_{\lambda\geq 0} f(X^{(k)}+\lambda p^{(k)})=f(X^{(k)}+\lambda_k p^{(k)})$ 6^0 $X^{(k+1)}=X^{(k)}+\lambda_k p^{(k)}$

$$\lambda_2=-\frac{g^{(2)^T}p^{(2)}}{p^{(2)^T}Qp^{(2)}}$$

7^0 $\|g(X^{(k+1)})\|<\varepsilon$? 若是,$X^*=X^{(k+1)}$, 否则若$k=n$,令$X^{(1)}:=X^{(k+1)}$,转$3^0$

若$k<n$,令$Y^{(k)}=g^{(k+1)}-g^{(k)}, S^{(k)}=X^{(k+1)}-X^{(k)}$, $H^{(k+1)}=H^{(k)}-\frac{H^{(k)}Y^{(k)}Y^{(k)T}H^{(k)}}{Y^{(k)T}H^{(k)}Y^{(k)}}+\frac{S^{(k)}S^{(k)T}}{S^{(k)T}Y^{(k)}}$,转$8^0$

8^0 令$k:=k+1$,转4^0

解： $Q=\begin{pmatrix}2&0\\0&8\end{pmatrix}$ $g(X)=\begin{pmatrix}2x_1\\8x_2\end{pmatrix}$ $g^{(1)}=\begin{pmatrix}2\\8\end{pmatrix}$ $\|g^{(1)}\|=\sqrt{68}\not<\varepsilon$

$k=2$ $X^{(2)}=\begin{pmatrix}0.73846\\-0.04616\end{pmatrix}$ $g^{(2)}=\begin{pmatrix}1.47692\\-0.36923\end{pmatrix}$ $H^{(2)}=\begin{pmatrix}1.00&-0.03\\-0.03&0.12\end{pmatrix}$

$p^{(2)}=-H^{(2)}g^{(2)}=\begin{pmatrix}-1.49\\0.09\end{pmatrix}$ $\lambda_2=-\dfrac{g^{(2)^T}p^{(2)}}{p^{(2)^T}Qp^{(2)}}$ $(3-13)P_{155}$

$X^{(3)}=X^{(2)}+\lambda_2 p^{(2)}=\begin{pmatrix}0\\0\end{pmatrix}$ $g^{(3)}=\begin{pmatrix}0\\0\end{pmatrix}$ $\|g^{(3)}\|=0<\varepsilon\ \therefore X^*=\begin{pmatrix}0\\0\end{pmatrix}$

3. DFP法举例

$n=2$

例3-13　用*DFP*方法求解 $\min f(X)=x_1^2+4x_2^2=\frac{1}{2}X^TQX$，

取 $X^{(1)}=(1,1)^T$ 为初始点，$\varepsilon=0.01$

注释：

$f(X)$ 是二维正定二次函数,两次迭代就求出了极小点,这不是偶然。是因为*DFP*方法具有二次终止性。对于*n*维正定二次函数，最多经过*n*次迭代就可以求到极小点。

三. 拟*Newton*法 ---- *DFP*法

$$(NP)\ \min_{X\in R^n} f(X)$$

- *DFP*法的搜索方向
- *DFP*法的迭代步骤
- *DFP*法举例
- *DFP*法的性质
- *DFP*法的收敛结论

4. DFP法的性质

$$p^{(k)} = -H^{(k)} g^{(k)}$$

$$H^{(k+1)} = H^{(k)} - \frac{H^{(k)} Y^{(k)} Y^{(k)T} H^{(k)}}{Y^{(k)T} H^{(k)} Y^{(k)}} + \frac{S^{(k)} S^{(k)T}}{S^{(k)T} Y^{(k)}}$$

性质1： 若 $H^{(k)}$ 对称正定，则 $H^{(k+1)}$ 也对称正定。

性质2： 对于正定二次函数 $f(X) = \frac{1}{2} X^T Q X + b^T X + c$

当 $H^{(1)} = E$ 时，则

(1) *DFP*方法产生的搜索方向 $p^{(1)}, p^{(2)}, \cdots, p^{(m)}$ 是互相 Q 共轭的。

因此*DFP*方法是一种共轭方向法,具有二次终止性。对于*n*维正定二次函数，最多经过*n*次迭代就可以求到极小点。

4. DFP法的性质

$p^{(k)} = -H^{(k)} g^{(k)}$

$$H^{(k+1)} = H^{(k)} - \frac{H^{(k)} Y^{(k)} Y^{(k)T} H^{(k)}}{Y^{(k)T} H^{(k)} Y^{(k)}} + \frac{S^{(k)} S^{(k)T}}{S^{(k)T} Y^{(k)}}$$

性质1： 若 $H^{(k)}$ 对称正定，则 $H^{(k+1)}$ 也对称正定。

性质2： 对于正定二次函数 $f(X) = \frac{1}{2} X^T Q X + b^T X + c$

当 $H^{(1)} = E$ 时，则

(1) DFP方法产生的搜索方向 $p^{(1)}, p^{(2)}, \cdots, p^{(m)}$ 是互相 Q 共轭的。

(2) 若经过n次迭代才求到极小点，则有：$H^{(n+1)} = Q^{-1}$.

($H^{(k)}$ 是 $G^{-1}(X^{(k)}) = Q^{-1}$ 的近似)

三. 拟*Newton*法 ---- *DFP*法

$$(NP)\ \min_{X\in R^n} f(X)$$

- *DFP*法的搜索方向
- *DFP*法的迭代步骤
- *DFP*法举例
- *DFP*法的性质
- *DFP*法的收敛结论

5. DFP法的收敛结论

结论1： 由于*DFP*方法是一种共轭方向法,因此是一种收敛算法. 即 $X^{(k)} \xrightarrow[k\to\infty]{} X^*$

结论2： 由于 $p^{(k)} = -H^{(k)} g^{(k)}$， $H^{(k)}$是 $G^{-1}(X^{(k)})$ 的近似，

$$\therefore\ X^{(k)} \xrightarrow[k\to\infty]{\text{超线性}} X^*$$

优点： 计算方便,收敛速度快. 被中小型问题广泛使用。

缺点： 由于舍入误差可能导致某个$H^{(k)}$奇异或不正定,这时可以重置 $H^{(k)} = E$

三. 拟*Newton*法 ---- *DFP*法

$$(NP)\ \min_{X\in R^n} f(X)$$

- ***DFP***法的搜索方向
- ***DFP***法的迭代步骤
- ***DFP***法举例
- ***DFP***法的性质
- ***DFP***法的收敛结论

三. 拟*Newton*法 ---- *BFGS*方法

拟***Newton***法中比***DFP***法更好的算法是***BFGS***算法。一般认为***BFGS***算法是目前最有效的算法，不仅对于精确的一维搜索，就是对于满足一定条件的不精确的一维搜索,也具有超线性收敛性。

BFGS算法中 $H^{(k)}$ 的迭代公式是:

$$H^{(k+1)}=H^{(k)}+[(1+\frac{Y^{(k)^T}H^{(k)}Y^{(k)}}{S^{(k)^T}Y^{(k)}})S^{(k)}S^{(k)^T}-S^{(k)}Y^{(k)^T}H^{(k)}+H^{(k)}Y^{(k)}S^{(k)^T}]\frac{1}{S^{(k)^T}Y^{(k)}}$$

—— P_{176}

只要在***DFP***算法中将 $H^{(k)}$ 的计算用上式代替就构成了***BFGS***算法。因为用上式计算，$H^{(k)}$ 不易变为奇异或不正定，所以***BFGS***算法比***DFP***算法具有更好的数值稳定性。

第三章 非线性规划

第四节 无约束优化问题的解法

- 最速下降法
- *Newton*法
- 拟*Newton*法
- 共轭梯度法

作业：P245　15(1)　　18(1)

作业：P155　15(1)　　18(1)