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

四. 共轭梯度法

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

- 共轭方向及其性质
- 二次函数共轭梯度法的迭代原理
- 二次函数共轭梯度法的迭代步骤
- 一般函数的共轭梯度法
- ***PRP***算法的迭代步骤
- 共轭梯度法的注释

1. 共轭方向及其性质

定义3-13 设Q是n阶对称正定矩阵，若向量组

$p^{(1)}, p^{(2)}, \cdots, p^{(m)} \in R^n$ 满足:

$$p^{(i)T} Q p^{(j)} \begin{cases} = 0 & i \neq j \\ \neq 0 & i = j \end{cases} \quad i, j = 1,2,\cdots,m(m \leq n) \quad (3-24)$$

则称该向量组Q共轭(Q正交)。

当 $Q = E$,(3-24)就是通常的正交条件:

$$p^{(i)T} p^{(j)} \begin{cases} = 0 & i \neq j \\ \neq 0 & i = j \end{cases} \quad i, j = 1, 2, \cdots, m$$

则向量组 $p^{(1)}, p^{(2)}, \cdots, p^{(m)}$ 正交。

1. 共轭方向及其性质

定义3-13 设Q是n阶对称正定矩阵，若向量组

$$p^{(1)}, p^{(2)}, \cdots, p^{(m)} \in R^n \text{ 满足：}$$

$$p^{(i)T} Q p^{(j)} \begin{cases} = 0 & i \neq j \\ \neq 0 & i = j \end{cases} \quad i, j = 1,2,\cdots,m(m \leq n) \quad (3-24)$$

则称该向量组Q共轭(Q正交)。

共轭方向的性质：

定理3-13 设 $p^{(1)}, p^{(2)}, \cdots, p^{(n)}$对于对称正定矩阵$Q$共轭，则 $p^{(1)}, p^{(2)}, \cdots, p^{(n)}$线性无关。

$$\lambda_k = -\frac{g^{(k)^T} p^{(k)}}{p^{(k)^T} Q p^{(k)}}$$

1. 共轭方向及其性质

定理3-14

设 $p^{(1)}, p^{(2)}, \cdots, p^{(n)}$ 对于对称正定矩阵Q共轭，则从任意一点$X^{(1)}$出发, 依次以 $p^{(1)}, p^{(2)}, \cdots, p^{(n)}$ 为搜索方向的下述算法：

$$\begin{cases} \min\limits_{\lambda\ge 0} f(X^{(k)}+\lambda p^{(k)}) = f(X^{(k)}+\lambda_k p^{(k)}) \\ X^{(k+1)} = X^{(k)} + \lambda_k p^{(k)} \end{cases} \qquad \begin{matrix} k=1 \\ k=2 \\ k=n \end{matrix}$$

经n次一维搜索收敛于 $\min\limits_{X\in R^n} f(X) = \frac{1}{2} X^T Q X + b^T X + c$ 的最优解 X^*。

分析： $X^{(1)} \quad X^{(2)} \quad X^{(3)} \cdots\cdots \quad X^{(n)} \quad X^{(n+1)} = X^*$

$p^{(1)} \quad p^{(2)} \quad p^{(3)} \cdots\cdots \quad p^{(n)}$

结论： 共轭方向法具有二次终止性.

$$\lambda_k = -\frac{g^{(k)^T} p^{(k)}}{p^{(k)^T} Q p^{(k)}}$$

1. 共轭方向及其性质

定理3-14

设 $p^{(1)}, p^{(2)}, \cdots, p^{(n)}$ 对于对称正定矩阵Q共轭，则从任意一点$X^{(1)}$出发, 依次以 $p^{(1)}, p^{(2)}, \cdots, p^{(n)}$ 为搜索方向的下述算法：

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

经n次一维搜索收敛于 $\min\limits_{X \in R^n} f(X) = \frac{1}{2} X^T Q X + b^T X + c$ 的最优解 X^*。

推论：

$$\begin{array}{ccccc} X^{(1)} & X^{(2)} & X^{(3)} \cdots\cdots & X^{(k)} & X^{(k+1)} \\ p^{(1)} & p^{(2)} & p^{(3)} \cdots\cdots & p^{(k)} & g^{(k+1)} \end{array}$$

则 $g^{(k+1)}$与 $p^{(1)}, p^{(2)}, \cdots, p^{(k)}$ 的任意线性组合都正交。

$$\mu_1 p^{(1)} + \mu_2 p^{(2)} + \cdots + \mu_k p^{(k)}$$

四. 共轭梯度法

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

- 共轭方向及其性质
- 二次函数共轭梯度法的迭代原理
- 二次函数共轭梯度法的迭代步骤
- 一般函数的共轭梯度法
- ***PRP***算法的迭代步骤
- 共轭梯度法的注释

2. 二次函数共轭梯度法的迭代原理

$$\lambda_k = -\frac{g^{(k)^T} p^{(k)}}{p^{(k)^T} Q p^{(k)}}$$

求 $\min\limits_{X \in R^n} f(X) = \frac{1}{2} X^T Q X + b^T X + c$ 的最优解 X^*。Q是对称正定矩阵。

$X^{(1)}$　　$p^{(1)} = -g^{(1)}$　　　　$X^{(2)} = X^{(1)} + \lambda_1 p^{(1)}$

$X^{(2)}$　　$p^{(2)} = -g^{(2)} + \beta_1 p^{(1)}$　　$\beta_1 = \dfrac{g^{(2)^T} Q p^{(1)}}{p^{(1)^T} Q p^{(1)}}$　　$X^{(3)} = X^{(2)} + \lambda_2 p^{(2)}$

确定 β_1：

$$p^{(2)^T} Q p^{(1)} = -g^{(2)^T} Q p^{(1)} + \beta_1 p^{(1)^T} Q p^{(1)} = 0$$

$$\Downarrow$$

$$\beta_1 = \frac{g^{(2)^T} Q p^{(1)}}{p^{(1)^T} Q p^{(1)}}$$

2. 二次函数共轭梯度法的迭代原理

$$\lambda_k = -\frac{g^{(k)^T} p^{(k)}}{p^{(k)^T} Q p^{(k)}}$$

求 $\min\limits_{X \in R^n} f(X) = \frac{1}{2} X^T Q X + b^T X + c$ 的最优解 X^*。Q是对称正定矩阵。

$X^{(1)} \quad p^{(1)} = -g^{(1)}$ $\qquad X^{(2)} = X^{(1)} + \lambda_1 p^{(1)}$

$X^{(2)} \quad p^{(2)} = -g^{(2)} + \beta_1 p^{(1)} \qquad \beta_1 = \frac{g^{(2)^T} Q p^{(1)}}{p^{(1)^T} Q p^{(1)}}$ $\qquad X^{(3)} = X^{(2)} + \lambda_2 p^{(2)}$

$X^{(3)} \quad p^{(3)} = -g^{(3)} + \beta_2 p^{(2)}$

确定 β_2： $\quad p^{(3)^T} Q p^{(2)} = -g^{(3)^T} Q p^{(2)} + \beta_2 p^{(2)^T} Q p^{(2)} = 0$

$$\Downarrow$$

$$\beta_2 = \frac{g^{(3)^T} Q p^{(2)}}{p^{(2)^T} Q p^{(2)}}$$

2. 二次函数共轭梯度法的迭代原理

$$\lambda_k = -\frac{g^{(k)^T} p^{(k)}}{p^{(k)^T} Q p^{(k)}}$$

求 $\min\limits_{X\in R^n} f(X)=\frac{1}{2}X^TQX+b^TX+c$ 的最优解 X^*。Q是对称正定矩阵。

$X^{(1)}$ $\quad p^{(1)}=-g^{(1)}$ $\qquad\qquad X^{(2)}=X^{(1)}+\lambda_1 p^{(1)}$

$X^{(2)}$ $\quad p^{(2)}=-g^{(2)}+\beta_1 p^{(1)}$ $\quad \beta_1=\frac{g^{(2)^T}Qp^{(1)}}{p^{(1)^T}Qp^{(1)}}$ $\quad X^{(3)}=X^{(2)}+\lambda_2 p^{(2)}$

$X^{(3)}$ $\quad p^{(3)}=-g^{(3)}+\beta_2 p^{(2)}$ $\quad \beta_2=\frac{g^{(3)^T}Qp^{(2)}}{p^{(2)^T}Qp^{(2)}}$ $\quad X^{(4)}=X^{(3)}+\lambda_3 p^{(3)}$

已知 $p^{(3)}$与$p^{(2)}$，$p^{(2)}$与$p^{(1)}$都Q共轭，$p^{(3)}$与$p^{(1)}$是否Q共轭?

2. 二次函数共轭梯度法的迭代原理

$$p^{(1)} = -g^{(1)} \qquad p^{(2)} = -g^{(2)} + \beta_1 p^{(1)} \qquad p^{(3)} = -g^{(3)} + \beta_2 p^{(2)}$$

证明： $p^{(3)}$与$p^{(1)}$是否Q共轭

$$p^{(3)^T} Q p^{(1)} = -g^{(3)^T} Q p^{(1)} + \beta_2 p^{(2)^T} Q p^{(1)}$$

$$= -g^{(3)^T} \underline{Q p^{(1)}}$$

2. 二次函数共轭梯度法的迭代原理

$$Qp^{(1)} = ?$$

$$f(X) = \frac{1}{2}X^T Q X + b^T X + c \qquad g(X) = \nabla f(X) = QX + b$$

$$g^{(k)} = QX^{(k)} + b$$

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

$$g^{(k+1)} = QX^{(k+1)} + b$$

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

$$= g^{(k)} + \lambda_k Q p^{(k)} \quad \Rightarrow \quad Qp^{(k)} = \frac{1}{\lambda_k}(g^{(k+1)} - g^{(k)})$$

$$Qp^{(1)} = \frac{1}{\lambda_1}(g^{(2)} - g^{(1)})$$

2. 二次函数共轭梯度法的迭代原理

$$p^{(1)} = -g^{(1)} \qquad p^{(2)} = -g^{(2)} + \beta_1 p^{(1)} \qquad p^{(3)} = -g^{(3)} + \beta_2 p^{(2)}$$

$$Qp^{(1)} = \frac{1}{\lambda_1}(g^{(2)} - g^{(1)})$$

证明： $p^{(3)}$与$p^{(1)}$是否Q共轭

$$p^{(3)^T} Qp^{(1)} = -g^{(3)^T} Qp^{(1)} + \beta_2 p^{(2)^T} Qp^{(1)}$$

$$= -g^{(3)^T} \underline{Qp^{(1)}}$$

$$= -\frac{1}{\lambda_1} g^{(3)^T} (g^{(2)} - g^{(1)}) = -\frac{1}{\lambda_1} (\underline{g^{(3)^T} g^{(2)}} - \underline{g^{(3)^T} g^{(1)}})$$

因为$g^{(1)}$与$g^{(2)}$都是$p^{(1)}$, $p^{(2)}$的线性组合, 由定理**3-14**推论，

1. 共轭方向及其性质

$$\lambda_k = -\frac{g^{(k)^T} p^{(k)}}{p^{(k)^T} Q p^{(k)}}$$

定理3-14

设 $p^{(1)}, p^{(2)}, \cdots, p^{(n)}$ 对于对称正定矩阵Q共轭，则从任意一点$X^{(1)}$出发, 依次以 $p^{(1)}, p^{(2)}, \cdots, p^{(n)}$ 为搜索方向的下述算法：

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

经n次一维搜索收敛于 $\min\limits_{X \in R^n} f(X) = \frac{1}{2} X^T Q X + b^T X + c$ 的最优解 X^*。

推论： $X^{(1)} \quad X^{(2)} \quad X^{(3)} \cdots\cdots \quad X^{(k)}$

$p^{(1)} \quad p^{(2)} \quad g^{(3)} \cdots\cdots \quad p^{(k)}$

$$p^{(1)} = -g^{(1)}$$

$$p^{(2)} = -g^{(2)} + \beta_1 p^{(1)}$$

$$g^{(3)^T} g^{(1)} = 0$$

$$g^{(3)^T} g^{(2)} = 0$$

$g^{(k+1)}$与 $p^{(1)}, p^{(2)}, \cdots, p^{(k)}$ 的任意线性组合都正交。

$$\mu_1 p^{(1)} + \mu_2 p^{(2)} + \cdots + \mu_k p^{(k)}$$

2. 二次函数共轭梯度法的迭代原理

$$p^{(1)} = -g^{(1)} \qquad p^{(2)} = -g^{(2)} + \beta_1 p^{(1)} \qquad p^{(3)} = -g^{(3)} + \beta_2 p^{(2)}$$

$$\boxed{Qp^{(1)} = \frac{1}{\lambda_1}(g^{(2)} - g^{(1)})}$$

证明： $p^{(3)}$与$p^{(1)}$是否Q共轭

$$p^{(3)^T} Qp^{(1)} = -g^{(3)^T} Qp^{(1)} + \beta_2 p^{(2)^T} Qp^{(1)}$$

$$= -g^{(3)^T} \underline{Qp^{(1)}}$$

$$= -\frac{1}{\lambda_1} g^{(3)^T} (g^{(2)} - g^{(1)}) = -\frac{1}{\lambda_1} (\underbrace{g^{(3)^T} g^{(2)}}_{=0} - \underbrace{g^{(3)^T} g^{(1)}}_{=0}) = 0$$

因为$g^{(1)}$与$g^{(2)}$都是$p^{(1)}, p^{(2)}$的线性组合, 由定理3-14推论,

所以 $p^{(3)}$与$p^{(1)}$ Q共轭。

2. 二次函数共轭梯度法的迭代原理

$$\lambda_k = -\frac{g^{(k)^T} p^{(k)}}{p^{(k)^T} Q p^{(k)}}$$

求 $\min\limits_{X\in R^n} f(X)=\frac{1}{2}X^TQX+b^TX+c$ 的最优解 X^*。Q是对称正定矩阵。

$$X^{(1)} \quad p^{(1)}=-g^{(1)} \qquad X^{(2)}=X^{(1)}+\lambda_1 p^{(1)}$$

$$X^{(2)} \quad p^{(2)}=-g^{(2)}+\beta_1 p^{(1)} \qquad \beta_1=\frac{g^{(2)^T}Qp^{(1)}}{p^{(1)^T}Qp^{(1)}} \qquad X^{(3)}=X^{(2)}+\lambda_2 p^{(2)}$$

$$X^{(3)} \quad p^{(3)}=-g^{(3)}+\beta_2 p^{(2)} \qquad \beta_2=\frac{g^{(3)^T}Qp^{(2)}}{p^{(2)^T}Qp^{(2)}} \qquad X^{(4)}=X^{(3)}+\lambda_3 p^{(3)}$$

$$X^{(k)} \quad p^{(k)}=-g^{(k)}+\beta_{k-1} p^{(k-1)} \qquad \beta_{k-1}=\frac{g^{(k)^T}Qp^{(k-1)}}{p^{(k-1)^T}Qp^{(k-1)}} \qquad X^{(k+1)}=X^{(k)}+\lambda_k p^{(k)}$$

$$\vdots \qquad\qquad \vdots$$

$$X^{(n)} \quad p^{(n)}=-g^{(n)}+\beta_{n-1} p^{(n-1)} \qquad \beta_{n-1}=\frac{g^{(n)T}Qp^{(n-1)}}{p^{(n-1)T}Qp^{(n-1)}} \qquad X^{(n+1)}=X^{(n)}+\lambda_n p^{(n)}$$

共轭梯度法是共轭方向法，具有二次终止性。

四. 共轭梯度法

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

- 共轭方向及其性质
- 二次函数共轭梯度法的迭代原理
- 二次函数共轭梯度法的迭代步骤
- 一般函数的共轭梯度法
- ***PRP***算法的迭代步骤
- 共轭梯度法的注释

3. 二次函数共轭梯度法的迭代步骤

1^0 给定初始点 $X^{(1)}$，容许误差 $\varepsilon > 0$，计算 $p^{(1)} = -g^{(1)}$，$k := 1$

2^0 一维搜索：$\min\limits_{\lambda \geq 0} f(X^{(k)} + \lambda p^{(k)}) = f(X^{(k)} + \lambda_k p^{(k)})$

或 $\lambda_k = \dfrac{g^{(k)T} g^{(k)}}{p^{(k)T} Q p^{(k)}}$（二次函数）　　$X^{(k+1)} = X^{(k)} + \lambda_k p^{(k)}$

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

否则，转 4^0

4^0 计算：$\beta_k = \dfrac{g^{(k+1)T} Q p^{(k)}}{p^{(k)T} Q p^{(k)}}$　　$p^{(k+1)} = -g^{(k+1)} + \beta_k p^{(k)}$

5^0 令 $k := k+1$，转 2^0

$$\lambda_k = -\frac{g^{(k)T} p^{(k)}}{p^{(k)T} Q p^{(k)}}$$

3. 二次函数共轭梯度法的迭代步骤

$$\lambda_k = -\frac{g^{(k)^T} p^{(k)}}{p^{(k)^T} Q p^{(k)}}$$

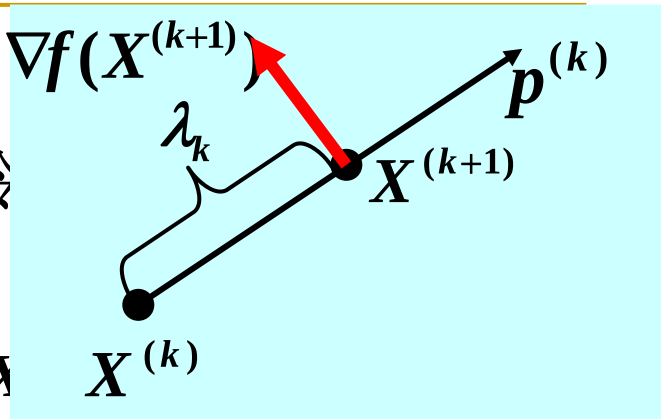
$$p^{(k)} = -g^{(k)} + \beta_{k-1} p^{(k-1)}$$

证明：$\lambda_k = \dfrac{g^{(k)T} g^{(k)}}{p^{(k)T} Q p^{(k)}}$

$$\lambda_k = \frac{-g^{(k)^T} p^{(k)}}{p^{(k)^T} Q p^{(k)}} = \frac{-g^{(k)^T}(-g^{(k)} + \beta_{k-1} p^{(k-1)})}{p^{(k)^T} Q p^{(k)}}$$

$$= \frac{g^{(k)^T} g^{(k)} - \beta_{k-1} \underline{g^{(k)^T} p^{(k-1)}}}{p^{(k)^T} Q p^{(k)}} \quad (3-11)$$

一. 最速下降法　3. 迭代步骤



1^0 取初始点 $X^{(0)}$, 容许误差 (精度) $\varepsilon > 0$, 令

2^0 计算 $p^{(k)} = -\nabla f(X^{(k)})$

3^0 检验 $\left\| p^{(k)} \right\| \le \varepsilon$? 若是迭代终止, 取 $X^* = X$

4^0 求最优步长 λ_k, $\min\limits_{\lambda \ge 0} f(X^{(k)} + \lambda p^{(k)}) = f(X^{(k)} + \lambda_k p^{(k)})$ (一维搜索)

5^0 令 $X^{(k+1)} = X^{(k)} + \lambda_k p^{(k)}$, 令 $k := k+1$, 转 2^0

注释:

2^0 结论：一维搜索最优解的梯度 $\nabla f(X^{(k+1)})$ 与搜索方向 $p^{(k)}$ 正交

$$\because \min_{\lambda \ge 0} \underbrace{f(X^{(k)} + \lambda p^{(k)})}_{\varphi(\lambda)} = \underbrace{f(X^{(k)} + \lambda_k p^{(k)})}_{\varphi(\lambda_k)} \text{ 且 } \varphi'(\lambda) = \nabla f(X^{(k)} + \lambda p^{(k)})^T p^{(k)}$$

$$\therefore 0 = \varphi'(\lambda_k) = \nabla f(X^{(k+1)})^T p^{(k)} \quad (3-11)$$

$$\boxed{g^{(k+1)^T} p^{(k)} = 0}$$

3. 二次函数共轭梯度法的迭代步骤

$$\lambda_k = -\frac{g^{(k)^T} p^{(k)}}{p^{(k)^T} Q p^{(k)}}$$

$$p^{(k)} = -g^{(k)} + \beta_{k-1} p^{(k-1)}$$

$$g^{(k+1)^T} p^{(k)} = 0$$

证明： $\lambda_k = \dfrac{g^{(k)T} g^{(k)}}{p^{(k)T} Q p^{(k)}}$

$$\lambda_k = \frac{-g^{(k)^T} p^{(k)}}{p^{(k)^T} Q p^{(k)}} = \frac{-g^{(k)^T}(-g^{(k)} + \beta_{k-1} p^{(k-1)})}{p^{(k)^T} Q p^{(k)}}$$

$$= \frac{g^{(k)^T} g^{(k)} - \beta_{k-1} \underline{g^{(k)^T} p^{(k-1)}}}{p^{(k)^T} Q p^{(k)}} \quad \underline{g^{(k)^T} p^{(k-1)}} = 0 \quad (3-11)$$

$$= \frac{g^{(k)^T} g^{(k)}}{p^{(k)^T} Q p^{(k)}}$$

3. 二次函数共轭梯度法的迭代步骤

$$\lambda_k = -\frac{g^{(k)T} p^{(k)}}{p^{(k)T} Q p^{(k)}}$$

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

2^0一维搜索 : $\min\limits_{\lambda \geq 0} f(X^{(k)} + \lambda p^{(k)}) = f(X^{(k)} + \lambda_k p^{(k)})$

或$\lambda_k = \dfrac{g^{(k)T} g^{(k)}}{p^{(k)T} Q p^{(k)}}$(二次函数)　　$X^{(k+1)} = X^{(k)} + \lambda_k p^{(k)}$

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

否则,转4^0

4^0计算:　$\beta_k = \dfrac{g^{(k+1)T} Q p^{(k)}}{p^{(k)T} Q p^{(k)}}$　　$p^{(k+1)} = -g^{(k+1)} + \beta_k p^{(k)}$

5^0令$k := k + 1$, 转2^0

例3-12 求解 $\min f(X)=x_1^2+4x_2^2$ 取 $X^{(1)}=(1,1)^T, \varepsilon=0.01$

$$=\frac{1}{2}X^TQX$$

解：

$$Q=\begin{pmatrix}2&0\\0&8\end{pmatrix}\quad g(X)=\begin{pmatrix}2x_1\\8x_2\end{pmatrix}\quad p^{(1)}=-g^{(1)}=-\begin{pmatrix}2\\8\end{pmatrix}$$

$$k=1\qquad \lambda_1=\frac{g^{(1)T}g^{(1)}}{p^{(1)T}Qp^{(1)}}=0.13$$

$$X^{(2)}=X^{(1)}+\lambda_1 p^{(1)}=\begin{pmatrix}0.73\\-0.04\end{pmatrix}\qquad g^{(2)}=\begin{pmatrix}1.47\\-0.36\end{pmatrix}\qquad \left\|g^{(2)}\right\|=1.52\text{太大}$$

$$\beta_1=\frac{g^{(2)T}Qp^{(1)}}{p^{(1)T}Qp^{(1)}}=0.03\qquad p^{(2)}=-g^{(2)}+\beta_1 p^{(1)}=\begin{pmatrix}-1.54\\0.09\end{pmatrix}\qquad k=2$$

1^0 $p^{(1)}=-g^{(1)},\quad k:=1$

2^0 $\min\limits_{\lambda\ge 0} f(X^{(k)}+\lambda p^{(k)})=f(X^{(k)}+\lambda_k p^{(k)})$ $\lambda_k=\dfrac{g^{(k)T}g^{(k)}}{p^{(k)T}Qp^{(k)}}$ $X^{(k+1)}=X^{(k)}+\lambda_k p^{(k)}$

3^0 $\left\|g^{(k+1)}\right\|<\varepsilon$？若是，$X^*=X^{(k+1)}$否则，转$4^0$ 4^0 $\beta_k=\dfrac{g^{(k+1)T}Qp^{(k)}}{p^{(k)T}Qp^{(k)}}$，$p^{(k+1)}=-g^{(k+1)}+\beta_k p^{(k)}$

5^0 令 $k:=k+1$，转 2^0

例3-12 求解 $\min f(X)=x_1^2+4x_2^2$ 取 $X^{(1)}=(1,1)^T,\varepsilon=0.01$

解：

$$Q=\begin{pmatrix}2&0\\0&8\end{pmatrix}\quad g(X)=\begin{pmatrix}2x_1\\8x_2\end{pmatrix}\quad p^{(1)}=-g^{(1)}=-\begin{pmatrix}2\\8\end{pmatrix}\quad g^{(2)}=\begin{pmatrix}1.47\\-0.36\end{pmatrix}$$

$$X^{(2)}=X^{(1)}+\lambda_1 p^{(1)}=\begin{pmatrix}0.73\\-0.04\end{pmatrix}\qquad p^{(2)}=-g^{(2)}+\beta_1 p^{(1)}=\begin{pmatrix}-1.54\\0.09\end{pmatrix}$$

$$k=2\quad \lambda_2=\frac{g^{(2)T}g^{(2)}}{p^{(2)T}Qp^{(2)}}=0.47\qquad X^{(3)}=X^{(2)}+\lambda_2 p^{(2)}=\begin{pmatrix}0.00\\0.00\end{pmatrix}$$

$$g^{(3)}=\begin{pmatrix}0\\0\end{pmatrix}\qquad \left\|g^{(3)}\right\|=0<\varepsilon$$ 迭代终止，取 $X^*=X^{(3)}$

共轭梯度法具有二次终止性.

1^0 $p^{(1)}=-g^{(1)},\quad k:=1$

2^0 $\min\limits_{\lambda\ge 0} f(X^{(k)}+\lambda p^{(k)})=f(X^{(k)}+\lambda_k p^{(k)})$ $\lambda_k=\dfrac{g^{(k)T}g^{(k)}}{p^{(k)T}Qp^{(k)}}$ $X^{(k+1)}=X^{(k)}+\lambda_k p^{(k)}$

3^0 $\left\|g^{(k+1)}\right\|<\varepsilon$？若是，$X^*=X^{(k+1)}$否则，转$4^0$ 4^0 $\beta_k=\dfrac{g^{(k+1)T}Qp^{(k)}}{p^{(k)T}Qp^{(k)}}$，$p^{(k+1)}=-g^{(k+1)}+\beta_k p^{(k)}$

5^0 令 $k:=k+1$，转 2^0

四、共轭梯度法

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

- 共轭方向及其性质
- 二次函数共轭梯度法的迭代原理
- 二次函数共轭梯度法的迭代步骤
- 一般函数的共轭梯度法
- *PRP*算法的迭代步骤
- 共轭梯度法的注释

4. 一般函数的共轭梯度法

$$Qp^{(k)}=\frac{1}{\lambda_k}\left(g^{(k+1)}-g^{(k)}\right)$$

求 $\min\limits_{X\in R^n} f(X)\cancel{=}\frac{1}{2}X^TQX+b^TX+c$ 的最优解 X^*.

2^0 一维搜索：$\min\limits_{\lambda\geq 0} f(X^{(k)}+\lambda p^{(k)})=f(X^{(k)}+\lambda_k p^{(k)})$ 一维搜索

或 $\lambda_k\cancel{=}\dfrac{g^{(k)^T}g^{(k)}}{p^{(k)^T}Qp^{(k)}}$ $\qquad X^{(k+1)}=X^{(k)}+\lambda_k p^{(k)}$

3^0 检验 $\left\|g^{(k+1)}\right\|<\varepsilon$? 若满足，迭代终止，取 $X^*=X^{(k+1)}$ 否则，转 4^0

4^0 计算：$\beta_k\cancel{=}\dfrac{g^{(k+1)^T}Qp^{(k)}}{p^{(k)^T}Qp^{(k)}}$ $\qquad p^{(k+1)}=-g^{(k+1)}+\beta_k p^{(k)}$

4. 一般函数的共轭梯度法

$$Qp^{(k)} = \frac{1}{\lambda_k}(g^{(k+1)} - g^{(k)})$$

$$\beta_k = \frac{g^{(k+1)^T} Qp^{(k)}}{p^{(k)^T} Qp^{(k)}}$$

$$= \frac{g^{(k+1)^T}(g^{(k+1)} - g^{(k)})}{p^{(k)^T}(g^{(k+1)} - g^{(k)})} = \frac{g^{(k+1)^T}(g^{(k+1)} - g^{(k)})}{-p^{(k)^T} g^{(k)}}$$

$$g^{(k+1)^T} p^{(k)} = 0$$

下面推导 β_k 的三种形式，它们分别对应三种不同的共轭梯度法。

4. 一般函数的共轭梯度法

$$\beta_k = \frac{g^{(k+1)^T} Q p^{(k)}}{p^{(k)^T} Q p^{(k)}} = \frac{g^{(k+1)^T}(g^{(k+1)} - g^{(k)})}{-p^{(k)^T} g^{(k)}}$$

$$g^{(k+1)^T} p^{(k)} = 0$$

$$p^{(k)} = -g^{(k)} + \beta_{k-1} p^{(k-1)}$$

$$0 = g^{(k+1)^T} p^{(k)} = g^{(k+1)^T}(-g^{(k)} + \beta_{k-1} p^{(k-1)})$$

$$g^{(k+1)^T} p^{(k-1)} = 0$$

$$= -g^{(k+1)^T} g^{(k)} + \beta_{k-1} g^{(k+1)^T} p^{(k-1)}$$

定理3-14推论

1. 共轭方向及其性质

$$\lambda_k = -\frac{g^{(k)^T} p^{(k)}}{p^{(k)^T} Q p^{(k)}}$$

定理3-14 设 $p^{(1)}, p^{(2)}, \cdots, p^{(n)}$ 对于对称正定矩阵Q共轭，则从任意一点$X^{(1)}$出发，依次以 $p^{(1)}, p^{(2)}, \cdots, p^{(n)}$ 为搜索方向的下述算法：

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

经n次一维搜索收敛于 $\min\limits_{X \in R^n} f(X) = \frac{1}{2} X^T Q X + b^T X + c$ 的最优解X^*。

推论：

$$\begin{array}{ccccccc} X^{(1)} & X^{(2)} & X^{(3)} & \cdots\cdots & X^{(k)} & X^{(k+1)} & \\ p^{(1)} & p^{(2)} & p^{(3)} & \cdots\cdots & p^{(k)} & g^{(k+1)} & \end{array}$$

$$g^{(k+1)^T} p^{(k-1)} = 0$$

则 $g^{(k+1)}$与$p^{(1)}, p^{(2)}, \cdots, p^{(k)}$ 的任意线性组合都正交。

$$\mu_1 p^{(1)} + \mu_2 p^{(2)} + \cdots + \mu_k p^{(k)}$$

4. 一般函数的共轭梯度法

$$\beta_k = \frac{g^{(k+1)^T} Q p^{(k)}}{p^{(k)^T} Q p^{(k)}} = \frac{g^{(k+1)^T}(g^{(k+1)} - g^{(k)})}{-p^{(k)^T} g^{(k)}}$$

$$g^{(k+1)^T} p^{(k)} = 0$$

$$p^{(k)} = -g^{(k)} + \beta_{k-1} p^{(k-1)}$$

$$0 = g^{(k+1)^T} p^{(k)} = g^{(k+1)^T}(-g^{(k)} + \beta_{k-1} p^{(k-1)})$$

$$g^{(k+1)^T} p^{(k-1)} = 0$$

$$= -g^{(k+1)^T} g^{(k)} + \beta_{k-1} \underline{g^{(k+1)^T} p^{(k-1)}}$$

定理3-14推论

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

$$g^{(k+1)^T} g^{(k)} = 0$$

$$g^{(k)^T} p^{(k)} = g^{(k)^T}(-g^{(k)} + \beta_{k-1} p^{(k-1)})$$

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

$$= -g^{(k)^T} g^{(k)} + \beta_{k-1} \underline{g^{(k)^T} p^{(k-1)}}$$

$$= -g^{(k)^T} g^{(k)}$$

4. 一般函数的共轭梯度法

$$\beta_k = \frac{g^{(k+1)^T} Q p^{(k)}}{p^{(k)^T} Q p^{(k)}} = \frac{g^{(k+1)^T}(g^{(k+1)} - g^{(k)})}{-p^{(k)^T} g^{(k)}}$$

$$g^{(k+1)^T} g^{(k)} = 0$$

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

1^0 $\beta_k = \dfrac{g^{(k+1)T} g^{(k+1)}}{g^{(k)T} g^{(k)}}$ (3-32) 称为*FR*公式

2^0 $\beta_k = \dfrac{g^{(k+1)T} g^{(k+1)}}{-p^{(k)T} g^{(k)}}$ (3-33) 称为*DM*公式

3^0 $\beta_k = \dfrac{g^{(k+1)^T}(g^{(k+1)} - g^{(k)})}{g^{(k)^T} g^{(k)}}$ (3-34) 称为*PRP*公式

这三个公式对应的共轭梯度法分别称为***FR*, *DM*** 和 ***PRP***算法.

四. 共轭梯度法

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

- 共轭方向及其性质
- 二次函数共轭梯度法的迭代原理
- 二次函数共轭梯度法的迭代步骤
- 一般函数的共轭梯度法
- *PRP*算法的迭代步骤
- 共轭梯度法的注释

5. PRP算法的迭代步骤

$3^0\ \beta_k = \dfrac{g^{(k+1)^T}(g^{(k+1)} - g^{(k)})}{g^{(k)^T} g^{(k)}}$ *PRP*公式

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

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

3^0 令 $p^{(1)} = -g^{(1)}$, 置 $k := 1$

$4^0\ \min\limits_{\lambda \geq 0} f(X^{(k)} + \lambda p^{(k)}) = f(X^{(k)} + \lambda_k p^{(k)})\quad X^{(k+1)} = X^{(k)} + \lambda_k p^{(k)}$

$5^0\ \|g^{(k+1)}\| < \varepsilon$? 若满足, 迭代终止, 取 $X^* = X^{(k+1)}$ 否则, 转 6^0

6^0 若 $k = n$, 则令 $X^{(1)} := X^{(k+1)}$ 转 3^0

若 $k < n$, 计算 $\beta_k = \dfrac{g^{(k+1)^T}(g^{(k+1)} - g^{(k)})}{g^{(k)^T} g^{(k)}}$, $p^{(k+1)} = -g^{(k+1)} + \beta_k p^{(k)}$

在*PRP*算法中, 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, 加快收敛速度。

四. 共轭梯度法

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

- 共轭方向及其性质
- 二次函数共轭梯度法的迭代原理
- 二次函数共轭梯度法的迭代步骤
- 一般函数的共轭梯度法
- ***PRP***算法的迭代步骤
- 共轭梯度法的注释

6. 共轭梯度法的注释

1^0 $p^{(k)} = -g^{(k)} + \beta_{k-1} p^{(k-1)}$是 $f(X)$在 $X^{(k)}$处的下降方向。

证明：

$$g^{(k)^T} p^{(k)} = g^{(k)^T} (-g^{(k)} + \beta_{k-1} p^{(k-1)}) = -g^{(k)^T} g^{(k)} = -\left\| g^{(k)} \right\|^2 < 0$$

所以共轭梯度法对一般目标函数是下降算法，

因此共轭梯度法是收敛算法。

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

6. 共轭梯度法的注释

1^0 $p^{(k)}=-g^{(k)}+\beta_{k-1}p^{(k-1)}$是 $f(X)$在 $X^{(k)}$处的下降方向。

所以共轭梯度法对一般目标函数是下降算法，

因此共轭梯度法是收敛算法。

2^0 在*PRP*算法中,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,加快收敛速度。

6. 共轭梯度法的注释

3^0 $\quad \because p^{(k)} = -g^{(k)} + \beta_{k-1} p^{(k-1)} \qquad p^{(k)} = -G(X^{(k)})^{-1} g^{(k)}$

所以共轭梯度法是最速下降法的一种改进算法。

当 $\beta_{k-1} = 0$ 时，就变为最速下降法。

共轭梯度法优于最速下降法,但是非“n步重新开始”的共轭梯度法也仅仅具有线性收敛速度.对于“n步重新开始”的*PRP*的算法,可以证明它具有n步二阶收敛速度.

4^0 和*Newton*法相比较,共轭梯度法的另一个优点是:计算机存储量小，因为它不涉及矩阵，仅仅存放向量。所以它适于求解较高维的问题。

四. 共轭梯度法

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

- 共轭方向及其性质
- 二次函数共轭梯度法的迭代原理
- 二次函数共轭梯度法的迭代步骤
- 一般函数的共轭梯度法
- ***PRP***算法的迭代步骤
- 共轭梯度法的注释

作业：P245　16

作业：P155　16