共轭梯度法(CG)

寇彩霞

Email: koucx@bupt.edu.cn

北京邮电大学理学院 主楼-816

引入

- 解线性方程组 $Ax = b(A > 0) \quad \Leftrightarrow \quad \min \frac{1}{2}x^T Ax - bx(A > 0)$
- Gauss消去法$O(n^3)$
- 若A是特殊矩阵，如对角阵，好解。
- CG法的本质就是，使得A变成对角阵来求解。

引入

- 解线性方程组 $Ax = b(A > 0) \quad \Leftrightarrow \quad \min \frac{1}{2}x^TAx - bx(A > 0)$
- Gauss消去法$O(n^3)$
- 若A是特殊矩阵，如对角阵，好解。
- CG法的本质就是，使得A变成对角阵来求解。

- 解线性方程组 $Ax = b(A > 0) \quad \Leftrightarrow \quad \min \frac{1}{2}x^T Ax - bx(A > 0)$
- Gauss消去法$O(n^3)$
- 若A是特殊矩阵，如对角阵，好解。
- CG法的本质就是，使得A变成对角阵来求解。

- 解线性方程组 $Ax = b(A > 0) \quad \Leftrightarrow \quad \min \frac{1}{2}x^T Ax - bx(A > 0)$
- Gauss消去法$O(n^3)$
- 若A是特殊矩阵，如对角阵，好解。
- CG法的本质就是，使得A变成对角阵来求解。

共轭

定义

设G是$n \times n$对称正定矩阵，d_1, d_2是n维非零向量. 如果

$$d_1^T G d_2 = 0, \tag{4.3.1}$$

则称向量d_1和d_2是G–共轭的，简称共轭的.
设$d_1, \cdots, d_m$是R^n中任一组非零向量，如果

$$d_i^T G d_j = 0, (i \neq j) \tag{4.3.2}$$

则称$d_1, \cdots, d_m$是G–共轭的，简称共轭的.

显然，如果$d_1, \cdots, d_m$是G–共轭的，则它们是线性无关的. 如果$G = I$，则共轭性就是通常的直交性.

例子: 有了共轭多方便:)

$$\min \frac{1}{2}x^T Hx + b^T x$$

- 假设我们现在已经有n个相互$H-$共轭的方向$p_1, p_2, \cdots, p_n$.
- 因为共轭$\Rightarrow$线性无关，故至多有n个共轭方向，并可作为基。
- 设最优点$x^* = \sum_{i=1}^n a_i p_i$,

$$f(x^*) = \frac{1}{2}x^{*T}Hx^* + b^T x^* = \sum_{i=1}^{n} \frac{1}{2}(a_i^2 p_i^T Hp_i + 2a_i b^T p_i)$$

由$\frac{\partial f(a)}{\partial a_i} = 0$得:

$$a_i = -\frac{p_i^T b}{p_i^T Hp_i}, \quad i = 1, 2, \cdots, n.$$

- 思考：若一般的n个非线性方向组成的基，结果怎样呢?

算法4.3.2

- 步1. 给出初始点x_0, 终止误差$\varepsilon > 0$, 令$k := 0$. 计算$g_0 = g(x_0)$和初始下降方向d_0, 使$d_0^T g_0 < 0$.
- 步2. 如果$\|g_k\| \le \varepsilon$, 停止迭代.
- 步3. 计算α_k和x_{k+1}, 使得

$$f(x_k + \alpha_k d_k) = \min_{\alpha \ge 0} f(x_k + \alpha d_k), \tag{4.3.3}$$

$$x_{k+1} = x_k + \alpha_k d_k. \tag{4.3.4}$$

- 步4. 采用某种共轭方向法计算d_{k+1}使得

$$d_{k+1}^T G d_j = 0, j = 0, 1, \cdots, k.$$

- 步5. 令$k := k + 1$, 转步2.

利用梯度产生共轭方向的具体步骤

$$\min \frac{1}{2}x^T Hx + b^T x$$

记号：$x_1, x_2, \cdots, x_n$ 迭代点；$g_1 = \nabla f(x_1)$...; 精确线搜索$\alpha_k^* = \frac{g_k^T g_k}{g_k^T H g_k}$;

- $x_1 = 0, d_1 = -g_1$,
- $x_2 = x_1 + \alpha_1^* d_1$, 计算$g_2$; 此时满足$g_2^T d_1 = 0, g_2^T g_1 = 0$;
- 确定搜索方向d_2: 设$d_2 = -g_2 + \beta d_1$,
  要求$d_2^T H d_1 = 0 \Rightarrow \beta = \frac{g_2^T H d_1}{d_1^T H d_1}$.
- $x_3 = x_2 + \alpha_2^* d_2$, 计算$g_3$; 此时满足$g_3^T d_2 = 0, g_3^T d_1 = 0, g_3^T g_2 = 0, g_3^T g_1 = 0$;
- 确定搜索方向d_3: 设$d_3 = -g_3 + \beta_2 d_2 + \beta_1 d_1$,
  要求$d_3^T H d_2 = 0 \Rightarrow \beta_2 = \frac{g_3^T H d_2}{d_2^T H d_2}$.
  要求$d_3^T H d_1 = 0 \Rightarrow \beta_1 = \frac{g_3^T H d_1}{d_1^T H d_1} = 0$.

$$\min \frac{1}{2}x^THx + b^Tx$$

记号：$x_1, x_2, \cdots, x_n$ 迭代点；$g_1 = \nabla f(x_1)...$; 精确线搜索$\alpha_k^* = \frac{g_k^T g_k}{g_k^T H g_k}$;

- $x_1 = 0, d_1 = -g_1$,
- $x_2 = x_1 + \alpha_1^* d_1$, 计算$g_2$; 此时满足$g_2^T d_1 = 0, g_2^T g_1 = 0$;
- 确定搜索方向d_2: 设$d_2 = -g_2 + \beta d_1$,
  要求$d_2^T H d_1 = 0 \Rightarrow \beta = \frac{g_2^T H d_1}{d_1^T H d_1}$.
- $x_3 = x_2 + \alpha_2^* d_2$, 计算$g_3$; 此时满足$g_3^T d_2 = 0, g_3^T d_1 = 0, g_3^T g_2 = 0, g_3^T g_1 = 0$;
- 确定搜索方向d_3: 设$d_3 = -g_3 + \beta_2 d_2 + \beta_1 d_1$,
  要求$d_3^T H d_2 = 0 \Rightarrow \beta_2 = \frac{g_3^T H d_2}{d_2^T H d_2}$.
  要求$d_3^T H d_1 = 0 \Rightarrow \beta_1 = \frac{g_3^T H d_1}{d_1^T H d_1} = 0$.

利用梯度产生共轭方向的具体步骤

$$\min \frac{1}{2}x^T Hx + b^T x$$

记号：$x_1, x_2, \cdots, x_n$ 迭代点；$g_1 = \nabla f(x_1)...$；精确线搜索$\alpha_k^* = \frac{g_k^T g_k}{g_k^T Hg_k}$；

- $x_1 = 0, d_1 = -g_1$，
- $x_2 = x_1 + \alpha_1^* d_1$，计算$g_2$；此时满足$g_2^T d_1 = 0, g_2^T g_1 = 0$；
- 确定搜索方向d_2：设$d_2 = -g_2 + \beta d_1$，
  要求$d_2^T Hd_1 = 0 \Rightarrow \beta = \frac{g_2^T Hd_1}{d_1^T Hd_1}$.
- $x_3 = x_2 + \alpha_2^* d_2$，计算$g_3$；此时满足$g_3^T d_2 = 0, g_3^T d_1 = 0, g_3^T g_2 = 0, g_3^T g_1 = 0$；
- 确定搜索方向d_3：设$d_3 = -g_3 + \beta_2 d_2 + \beta_1 d_1$，
  要求$d_3^T Hd_2 = 0 \Rightarrow \beta_2 = \frac{g_3^T Hd_2}{d_2^T Hd_2}$.
  要求$d_3^T Hd_1 = 0 \Rightarrow \beta_1 = \frac{g_3^T Hd_1}{d_1^T Hd_1} = 0$.

$$\min \frac{1}{2}x^T Hx + b^T x$$

记号：$x_1, x_2, \cdots, x_n$ 迭代点；$g_1 = \nabla f(x_1)\ldots$; 精确线搜索$\alpha_k^* = \frac{g_k^T g_k}{g_k^T H g_k}$;

- $x_1 = 0, d_1 = -g_1$,
- $x_2 = x_1 + \alpha_1^* d_1$, 计算$g_2$; 此时满足$g_2^T d_1 = 0, g_2^T g_1 = 0$;
- 确定搜索方向d_2: 设$d_2 = -g_2 + \beta d_1$,
  要求$d_2^T H d_1 = 0 \Rightarrow \beta = \frac{g_2^T H d_1}{d_1^T H d_1}$.
- $x_3 = x_2 + \alpha_2^* d_2$, 计算$g_3$; 此时满足$g_3^T d_2 = 0, g_3^T d_1 = 0, g_3^T g_2 = 0, g_3^T g_1 = 0$;
- 确定搜索方向d_3: 设$d_3 = -g_3 + \beta_2 d_2 + \beta_1 d_1$,
  要求$d_3^T H d_2 = 0 \Rightarrow \beta_2 = \frac{g_3^T H d_2}{d_2^T H d_2}$.
  要求$d_3^T H d_1 = 0 \Rightarrow \beta_1 = \frac{g_3^T H d_1}{d_1^T H d_1} = 0$.

利用梯度产生共轭方向的具体步骤

$$\min \frac{1}{2}x^T Hx + b^T x$$

记号：$x_1, x_2, \cdots, x_n$ 迭代点；$g_1 = \nabla f(x_1)...$；精确线搜索$\alpha_k^* = \frac{g_k^T g_k}{g_k^T H g_k}$；

- $x_1 = 0, d_1 = -g_1$,
- $x_2 = x_1 + \alpha_1^* d_1$, 计算$g_2$; 此时满足$g_2^T d_1 = 0, g_2^T g_1 = 0$;
- 确定搜索方向d_2: 设$d_2 = -g_2 + \beta d_1$,
  要求$d_2^T H d_1 = 0 \Rightarrow \beta = \frac{g_2^T H d_1}{d_1^T H d_1}$.
- $x_3 = x_2 + \alpha_2^* d_2$, 计算$g_3$; 此时满足$g_3^T d_2 = 0, g_3^T d_1 = 0, g_3^T g_2 = 0, g_3^T g_1 = 0$;
- 确定搜索方向d_3: 设$d_3 = -g_3 + \beta_2 d_2 + \beta_1 d_1$,
  要求$d_3^T H d_2 = 0 \Rightarrow \beta_2 = \frac{g_3^T H d_2}{d_2^T H d_2}$.
  要求$d_3^T H d_1 = 0 \Rightarrow \beta_1 = \frac{g_3^T H d_1}{d_1^T H d_1} = 0$.

利用梯度产生共轭方向的具体步骤

- 一般情形，已有$d_1, d_2, \cdots, d_k$, 迭代到x_{k+1}; 此时满足：

$$g_{k+1}^T d_i = 0, g_{k+1}^T g_i = 0, \quad i = 1, 2, \cdots, k;$$

- 确定下一个搜索方向d_{k+1}，要求与前k个搜索方向共轭：
  - 设 $d_{k+1} = -g_{k+1} + \beta_k d_k + \sum_{i=1}^{k-1} \beta_i d_i$
  - 要求 $d_{k+1}^T H d_k = 0 \Rightarrow \beta_k = \frac{g_{k+1}^T H d_k}{d_k^T H d_k}$;
  - 要求 $d_{k+1}^T H d_i = 0 \Rightarrow \beta_k = \frac{g_{k+1}^T H d_i}{d_i^T H d_i} = 0 \quad i = 1, \cdots, k;$

  故：$d_{k+1} = -g_{k+1} + \beta_k d_k$.
- ...

利用梯度产生共轭方向的具体步骤

- 一般情形，已有$d_1, d_2, \cdots, d_k$, 迭代到x_{k+1}; 此时满足：

$$g_{k+1}^T d_i = 0, g_{k+1}^T g_i = 0, \quad i = 1, 2, \cdots, k;$$

- 确定下一个搜索方向d_{k+1}，要求与前k个搜索方向共轭：
  - 设 $d_{k+1} = -g_{k+1} + \beta_k d_k + \sum_{i=1}^{k-1} \beta_i d_i$
  - 要求 $d_{k+1}^T H d_k = 0 \Rightarrow \beta_k = \frac{g_{k+1}^T H d_k}{d_k^T H d_k}$;
  - 要求 $d_{k+1}^T H d_i = 0 \Rightarrow \beta_k = \frac{g_{k+1}^T H d_i}{d_i^T H d_i} = 0 \quad i = 1, \cdots, k;$

  故：$d_{k+1} = -g_{k+1} + \beta_k d_k$.

- ...

- 一般情形，已有$d_1, d_2, \cdots, d_k$，迭代到x_{k+1}；此时满足：

$$g_{k+1}^T d_i = 0, g_{k+1}^T g_i = 0, \quad i = 1, 2, \cdots, k;$$

- 确定下一个搜索方向d_{k+1}，要求与前k个搜索方向共轭：
  - 设 $d_{k+1} = -g_{k+1} + \beta_k d_k + \sum_{i=1}^{k-1} \beta_i d_i$
  - 要求 $d_{k+1}^T H d_k = 0 \Rightarrow \beta_k = \frac{g_{k+1}^T H d_k}{d_k^T H d_k}$;
  - 要求 $d_{k+1}^T H d_i = 0 \Rightarrow \beta_k = \frac{g_{k+1}^T H d_i}{d_i^T H d_i} = 0 \quad i = 1, \cdots, k;$

  故：$d_{k+1} = -g_{k+1} + \beta_k d_k$.

- ...

利用梯度产生共轭方向的具体步骤

- 一般情形，已有$d_1, d_2, \cdots, d_k$，迭代到x_{k+1}；此时满足：

$$g_{k+1}^T d_i = 0, g_{k+1}^T g_i = 0, \quad i = 1, 2, \cdots, k;$$

- 确定下一个搜索方向d_{k+1}，要求与前k个搜索方向共轭：
  - 设 $d_{k+1} = -g_{k+1} + \beta_k d_k + \sum_{i=1}^{k-1} \beta_i d_i$
  - 要求 $d_{k+1}^T H d_k = 0 \Rightarrow \beta_k = \frac{g_{k+1}^T H d_k}{d_k^T H d_k}$;
  - 要求 $d_{k+1}^T H d_i = 0 \Rightarrow \beta_k = \frac{g_{k+1}^T H d_i}{d_i^T H d_i} = 0 \quad i = 1, \cdots, k;$

  故：$d_{k+1} = -g_{k+1} + \beta_k d_k$.

- ...

- 一般情形，已有$d_1, d_2, \cdots, d_k$，迭代到x_{k+1}；此时满足：

$$g_{k+1}^T d_i = 0, g_{k+1}^T g_i = 0, \quad i = 1, 2, \cdots, k;$$

- 确定下一个搜索方向d_{k+1}，要求与前k个搜索方向共轭：
  - 设 $d_{k+1} = -g_{k+1} + \beta_k d_k + \sum_{i=1}^{k-1} \beta_i d_i$
  - 要求 $d_{k+1}^T H d_k = 0 \Rightarrow \beta_k = \frac{g_{k+1}^T H d_k}{d_k^T H d_k}$；
  - 要求 $d_{k+1}^T H d_i = 0 \Rightarrow \beta_k = \frac{g_{k+1}^T H d_i}{d_i^T H d_i} = 0 \quad i = 1, \cdots, k$；

  故：$d_{k+1} = -g_{k+1} + \beta_k d_k$.
- …

利用梯度产生共轭方向的具体步骤

- 一般情形，已有$d_1, d_2, \cdots, d_k$, 迭代到x_{k+1}; 此时满足：

$$g_{k+1}^T d_i = 0, g_{k+1}^T g_i = 0, \quad i = 1, 2, \cdots, k;$$

- 确定下一个搜索方向d_{k+1}，要求与前k个搜索方向共轭：
  - 设 $d_{k+1} = -g_{k+1} + \beta_k d_k + \sum_{i=1}^{k-1} \beta_i d_i$
  - 要求 $d_{k+1}^T H d_k = 0 \Rightarrow \beta_k = \frac{g_{k+1}^T H d_k}{d_k^T H d_k}$;
  - 要求 $d_{k+1}^T H d_i = 0 \Rightarrow \beta_k = \frac{g_{k+1}^T H d_i}{d_i^T H d_i} = 0 \quad i = 1, \cdots, k$;

  故：$d_{k+1} = -g_{k+1} + \beta_k d_k$.
- ...

关键是 d_{k+1} 的确定？

- 二次函数时：

$$d_{k+1} = -g_{k+1} + \beta_k d_k \quad \beta_k = \frac{g_{k+1}^T H d_k}{d_k^T H d_k} \tag{1}$$

- 非二次函数的一般非线性函数：如何推广？
-

$$d_{k+1} = -g_{k+1} + \beta_k d_k \tag{2}$$

-

$$\beta_k =? \tag{3}$$

关键是 d_{k+1} 的确定？

- 二次函数时：

$$d_{k+1} = -g_{k+1} + \beta_k d_k \quad \beta_k = \frac{g_{k+1}^T H d_k}{d_k^T H d_k} \tag{1}$$

- 非二次函数的一般非线性函数：如何推广？

- $$d_{k+1} = -g_{k+1} + \beta_k d_k \tag{2}$$

- $$\beta_k = ? \tag{3}$$

关键是 d_{k+1} 的确定？

- 二次函数时：

$$d_{k+1} = -g_{k+1} + \beta_k d_k \quad \beta_k = \frac{g_{k+1}^T H d_k}{d_k^T H d_k} \tag{1}$$

- 非二次函数的一般非线性函数：如何推广？
-

$$d_{k+1} = -g_{k+1} + \beta_k d_k \tag{2}$$

-

$$\beta_k = ? \tag{3}$$

关键是 d_{k+1} 的确定？

- 二次函数时：

$$d_{k+1} = -g_{k+1} + \beta_k d_k \quad \beta_k = \frac{g_{k+1}^T H d_k}{d_k^T H d_k} \tag{1}$$

- 非二次函数的一般非线性函数：如何推广？
-

$$d_{k+1} = -g_{k+1} + \beta_k d_k \tag{2}$$

-

$$\beta_k = ? \tag{3}$$

6种β_k

$\beta_k =$

分子：$g_{k+1}^T(g_{k+1} - g_k)$， $\|g_{k+1}\|^2$

分母：$d_k^T(g_{k+1} - g_k)$， $-d_k^T g_k$， $\|g_k\|^2$

6种β_k

$$\beta_k = \frac{g_{k+1}^T(g_{k+1}-g_k)}{d_k^T(g_{k+1}-g_k)}, \text{(Hestenes-Stiefel 公式1952)} \tag{4}$$

$$\beta_k = \frac{\|g_{k+1}\|^2}{\|g_k\|^2} \text{(Fletcher-Reeves 公式1964)} \tag{5}$$

$$\beta_k = \frac{g_{k+1}^T(g_{k+1}-g_k)}{\|g_k\|^2} \text{(Polak-Ribiére-Polyak 公式1969)} \tag{6}$$

$$\beta_k = \frac{\|g_{k+1}\|^2}{d_k^T(g_{k+1}-g_k)} \text{(Dai-Yuan 公式1999)} \tag{7}$$

$$\beta_k = \frac{g_{k+1}^T(g_{k+1}-g_k)}{-d_k^T g_k} \text{(Conjugate Descent, Fletcher1987)} \tag{8}$$

$$\beta_k = \frac{\|g_{k+1}\|^2}{-d_k^T g_k} \text{(Liu-Storey 公式1991)} \tag{9}$$

二次终止性

- 精确线性搜索条件下，共轭方向法具有二次终止性(对正定二次函数，方法有限步终止.)

定理 (4.3.3)

设$x_0 \in R^n$是任意初始点. 对于极小化二次函数

$$\min f(x) = \frac{1}{2}x^T G x - b^T x, \tag{4.3.5}$$

共轭方向法至多经n步精确线性搜索终止；且每一x_{i+1}都是$f(x)$在x_0 和方向$d_0, \cdots, d_i$所张成的线性流形$\{x | x = x_0 + \sum_{j=0}^{i} \alpha_j d_j, \forall \alpha_j\}$中的极小点.

算法4.3.8(FR-CG)

- 步1. 初始步：给出$x_0, \varepsilon > 0$, 计算$f_0 = f(x_0), g_0 = \nabla f(x_0)$, 令$d_0 = -g_0, k := 0$.
- 步2. 如果$\|g_k\| \leq \varepsilon$, 停止.
- 步3. 由线性搜索求步长因子α_k, 并令

$$x_{k+1} = x_k + \alpha_k d_k.$$

- 步4.

$$\beta_k = \frac{g_{k+1}^T g_{k+1}}{g_k^T g_k},$$
$$d_{k+1} = -g_{k+1} + \beta_k d_k.$$

- 步5. 令$k := k + 1$, 转步2.

定理 (4.3.9)

假定$f(x)$在有界水平集$L(x_0)=\{x\in R^n|\ f(x)\leq f(x_0)\}$上连续可微，且有下界，那么采用精确线性搜索的*Fletcher-Reeves*共轭梯度法产生的序列$\{x_k\}$ 至少有一个聚点是驻点，即

*(1)*当$\{x_k\}$是有限点列时，其最后一个点是$f(x)$的驻点.

*(2)*当$\{x_k\}$是无穷点列时，它必有聚点，且任一聚点都是$f(x)$的驻点.

收敛速度

- 假定$f(x)$在有界水平集$L(x_0)=\{x|f(x)\leq f(x_0)\}$上连续可微，且有下界，那么采用精确线性搜索的Fletcher-Reeves共轭梯度法产生的迭代序列$\{x_k\}$ 线性收敛到$f(x)$的极小点x^*.
- 在FR共轭梯度算法4.3.8中，为了保证方向是下降的，必须要求

$$g_k^T d_k = -\|g_k\|^2 + \beta_{k-1} g_k^T d_{k-1} < 0.$$

  - 在精确线性搜索情形，由于$g_k^T d_{k-1}=0$, 故$g_k^T d_k<0$是显然的.
  - 在非精确线性搜索情形，可能有$g_k^T d_k>0$的情形发生，从而d_k不是下降方向. 如果采用强Wolfe不精确线性搜索准则(3.5.9) – (3.5.10)产生步长因子α_k, 则d_k一定是下降方向.

收敛速度

- 定理的证明依赖于$d_0 = -g_0$, 即初始搜索方向d_0取最速下降方向。
- 定义

$$\kappa(G) = \|G\|_2 \|G^{-1}\|_2 = \lambda_1/\lambda_n.$$

可以证明

$$\|x_k - x^*\|_G \le \left(\frac{\sqrt{\kappa(G)} - 1}{\sqrt{\kappa(G)} + 1} \right)^k \|x_0 - x^*\|_G. \tag{4.3.39}$$

收敛速度依赖于$\sqrt{\kappa(G)}$.

- 如果能够采取某种措施改善G的条件数，则收敛可以加快. 预处理技术就提供了解决这个问题的一个办法.

- 设C是某个非奇异矩阵，令

$$\hat{x} = Cx. \tag{4.3.40}$$

则二次目标函数(4.3.5)变成

$$\min \hat{f}(\hat{x}) = \frac{1}{2}\hat{x}^T(C^{-T}GC^{-1})\hat{x} - (C^{-T}b)^T\hat{x}. \tag{4.3.41}$$

等价于解线性方程组

$$(C^{-T}GC^{-1})\hat{x} = C^{-T}b. \tag{4.3.42}$$

- 方法的收敛依赖于$C^{-T}GC^{-1}$的条件数而不是G的条件数. 如果能够选择非奇异矩阵C使得$C^{-T}GC^{-1}$的条件数明显好于G的条件数，则收敛速度将明显改善. 根据这一思想，定义$M = C^TC$，并给出预处理共轭梯度法.

再开始共轭梯度法

- 对于一般非二次函数，共轭梯度法常常采用再开始技术：每n步以后周期性地采用最速下降方向作为新的搜索方向.
- 当迭代从$f(x)$的非二次区域进入可由二次函数很好地逼近的区域时，重新取最速下降方向作为搜索方向，则其后n次迭代方向接近于共轭方向，从而使方法有较快的收敛速度. (共轭梯度法的二次终止性依赖于取最速下降方向作为初始搜索方向)
- 大型问题，经常进行再开始，例如每隔r步迭代再开始($r < n$或$r \ll n$).

再开始技巧

- 每隔n步再开始.
- $g_k^T d_k > 0$, 即d_k是上升方向时再开始.
- 对于二次函数，相邻两次迭代的梯度相互直交，因此，如果它们偏离直交性较大，例如

$$\frac{g_k^T g_{k-1}}{\|g_k\|^2} \geq v, \tag{4.3.53}$$

这里可取$v = 0.1$, 则进行再开始.

算法4.3.10(再开始FR共轭梯度法)

- 步1. 初始步：给出初始点$x_0, \varepsilon > 0, k := 0$.
- 步2. 计算$g_0 = g(x_0)$. 如果$\|g_0\| \leq \varepsilon$, 停止迭代，输出$x^* = x_0$; 否则令$d_0 = -g_0$.
- 步3. 线性搜索求步长因子α_k, 并令

$$
\begin{aligned}
x_{k+1} &= x_k + \alpha_k d_k. \\
k &:= k + 1.
\end{aligned}
$$

- 步4. 计算$g_k = g(x_k)$.
  若$\frac{g_k^T g_{k-1}}{\|g_k\|^2} \geq 0.1$, 令$x_0 := x_k$, 转步2;
  如果$\|g_k\| \leq \varepsilon$, 停止迭代.
- 步5. 若$k = n$, 令$x_0 := x_k$, 转步2.
- 步6. 计算

$$
\beta = \frac{g_k^T g_k}{g_{k-1}^T g_{k-1}}, \quad d_k = -g_k + \beta d_{k-1}.
$$

- 步7. 如果$d_k^T g_k > 0$, 令$x_0 := x_k$, 转步2；否则转步3.

收敛性结果

- 与共轭梯度法一样，再开始共轭梯度法仍然具有总体收敛性.
- 再开始共轭梯度法至少有线性收敛速度.
- 再开始共轭梯度法产生的迭代点列具有n步二次收敛速度，即

$$\|x_{k+n} - x^*\| = O(\|x_k - x^*\|^2)$$

作业

- 用共轭梯度法求解极小化问题

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

  初始点$x_0 = (-2, 4)^T$.
- 熟悉 Matlab 中可求解无约束最小化的 fminunc 和 optimset 函数. 尝试求解

$$\min f(x) = 3x_1^2 + 2x_1x_2 + x_2^4.$$

- 证明FR-CG方法的全局收敛性（包括采用精确线搜索和非精确线搜索）

例: $\min_{x\in\mathcal{R}^2} f(x) = 100(x_1^2 - x_2)^2 + (x_1 - 1)^2$

- 编写M文件: fun.m 和 gfun.m
- 线搜索:
- CG:

例: $\min_{x \in \mathcal{R}^2} f(x) = 100(x_1^2 - x_2)^2 + (x_1 - 1)^2$

- 编写M文件: fun.m 和 gfun.m
- 线搜索:
- CG:

例: $\min_{x\in\mathcal{R}^2} f(x) = 100(x_1^2 - x_2)^2 + (x_1 - 1)^2$

- 编写M文件: fun.m 和 gfun.m
- 线搜索：
- CG：

参考文献

- W. W. Hager and H. Zhang, *A survey of nonlinear conjugate gradient methods*, Pacific J. Optim. 2 (2006): 35-58.
- Y. H. Dai, *Nonlinear Conjugate Gradient Methods*, Wiley Encyclopedia of Operations Research and Management SciencePublished OnlineFeb 2011, DOI: 10.1002/9780470400531.eorms0183/pdf