

利用分解校正矩阵确定搜索方向的 BFGS 算法

柳力

(吉林市广播电视大学教学处, 吉林 吉林 132002)

摘要: 本文把正定矩阵关于向量的等内积分解算法应用于改进 BFGS 算法中搜索方向的计算. 通过建立不依赖于搜索方式的用分解矩阵表达的校正公式, 给出了用 Hesse 近似矩阵的等内积分解矩阵确定搜索方向的 BFGS 算法.

关键词: BFGS 算法; 校正矩阵; 等内积分解; 搜索方向; 算法

MR(2010) 主题分类号: 90C30 中图分类号: O221.2

文献标识码: A 文章编号: 0255-7797(2016)05-1035-05

1 引言

利用正定矩阵关于向量的等内积分解算法^[1], 文 [2] 提出了由校正矩阵的等内积分解矩阵确定搜索方向的拟 Newton 算法. 即对于无约束最优化问题

$$\min f(x), x \in R^n, \quad (1.1)$$

其中 $f: R^n \rightarrow R^1$ 是连续可微函数, 通过建立 Hesse 近似矩阵 B 关于目标函数梯度向量的等内积分解矩阵的校正公式, 计算搜索方向的一种新算法.

以拟 Newton 算法中最有代表性的 BFGS 算法为例, 当记 x^k 为迭代点, $s_k = x^{k+1} - x^k$, $y_k = g_{k+1} - g_k$, $g_k = \nabla f(x^k)$, 则对 B 和对 H 的 BFGS 校正公式分别为

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{s_k^T y_k}, \quad (1.2)$$

$$H_{k+1} = \left(I - \frac{s_k y_k^T}{s_k^T y_k}\right) H_k \left(I - \frac{y_k s_k^T}{s_k^T y_k}\right) + \frac{s_k s_k^T}{s_k^T y_k}. \quad (1.3)$$

在求迭代点 x^{k+1} 处的搜索方向 d^{k+1} 时, 一般是解方程组

$$B_{k+1} d^{k+1} = -g_{k+1}, \quad (1.4)$$

而不是用 (1.3) 式直接计算.

$$d^{k+1} = -H_{k+1} g_{k+1}. \quad (1.5)$$

采用 (1.4) 式的计算量显然大于用 (1.5) 式, “舍近求远”的原因是, 由于计算误差的存在, 实际计算时用 (1.3) 式可能出现 H_{k+1} 不正定甚至奇异的情况. 而 (1.2) 式可以化为

$$B_{k+1} = B_k + ((d^k)^T g_k)^{-1} g_k g_k^T + (\alpha_k (d^k)^T y_k)^{-1} y_k y_k^T. \quad (1.6)$$

*收稿日期: 2014-02-22 接收日期: 2014-07-31

基金项目: 吉林省教育厅“十二五”科学技术研究项目资助 (2014598).

作者简介: 柳力 (1958-), 男, 吉林吉林, 副教授, 主要研究方向: 数学规划.

据此, Gill 与 Murray 提出了利用 Cholesky 分解技术解方程组 (1.4) 的方法. 即先对 B_{k+1} 作 Cholesky 分解: $B_{k+1} = L_{k+1}D_{k+1}L_{k+1}^T$, 其中 L_{k+1} 为单位下三角阵, D_{k+1} 为对角阵, 再依次解方程组

$$L_{k+1}v = -g_{k+1}, \quad L_{k+1}^T d^{k+1} = D_{k+1}^{-1}v. \quad (1.7)$$

同时利用 (1.6) 式, 给出由 L_k, D_k 求 L_{k+1}, D_{k+1} 的迭代公式, 使实际计算时不必每次迭代都进行一次 B_k 的 Cholesky 分解, 具体算法详见文 [3-5].

而文 [2] 提出的改进算法, 则是运用了正定矩阵关于向量的等内积分解算法, 通过求 B_{k+1} 关于 $g_{k+1} = \nabla f(x^{k+1})$ 的等内积分解矩阵. 即满足条件

$$B_{k+1} = C_{k+1}C_{k+1}^T, \quad C_{k+1}^{-1}g_{k+1} = a_{k+1}e \quad (1.8)$$

的矩阵 C_{k+1}^{-1} , 其中 $a_{k+1} \in R^1, e = (1, 1, \dots, 1)^T \in R^n$, 来计算点 x^{k+1} 处的搜索方向

$$d^{k+1} = -a_{k+1}(d_1^{k+1} \dots d_n^{k+1})^T, \quad (1.9)$$

其中 $d_j^{k+1} (j = 1, 2, \dots, n)$ 是矩阵 C_{k+1}^{-1} 第 j 列元素之和. 具体作法是替代校正公式 (1.2), 直接建立 B 关于目标函数梯度向量的分解矩阵 C_k^{-1} 到 C_{k+1}^{-1} 的校正公式, 在计算出校正矩阵 C_{k+1}^{-1} 的同时, 利用 (1.9) 式求出搜索方向 d^{k+1} .

显然文 [2] 给出了求搜索方向的另一种路径, 但文 [2] 中的相关结论, 都是在精确线搜索条件下给出的, 所以理论意义大于实际应用. 在拟Newton算法中BFGS算法是最有代表性的一个算法, 带Wolfe搜索的BFGS算法具有全局收敛和超线性收敛性 (参见文 [4]), 与拟Newton算法中的其它算法比较其收敛速度和数值稳定性都是最好的. 本文就以BFGS算法为研究对象, 通过建立不依赖搜索方式的用分解矩阵表达的校正公式, 给出用Hesse近似矩阵的等内积分解矩阵确定搜索方向BFGS新算法.

2 用分解矩阵形式表达的 BFGS 校正公式

当记 $u_k = C_k^{-1}y_k, v_k = C_k^{-1}g_{k+1}, w_k = C_k^{-1}g_k$. 由文 [2] 定理 2.2, 在精确线搜索条件下对于BFGS校正公式 (1.2). 若已求得 B_k 关于 g_k 的等内积分解矩阵 C_k^{-1} , 则 B_{k+1} 关于 g_{k+1} 的等内积分解矩阵

$$C_{k+1}^{-1} = Q_k C_k^{-1} \left[I - \frac{(y_k - \sqrt{\lambda_k} g_k) s_k^T}{s_k^T y_k} \right], \quad (2.1)$$

其中 $\lambda_k = \frac{s_k^T y_k}{\|w_k\|^2}, Q_k = I - \frac{2}{\|\sigma_k\|^2} \sigma_k \sigma_k^T, \sigma_k = w_k + t_k v_k, t_k = \frac{\|w_k\|}{\|v_k\|}$. 由文 [2] 定理 2.3, 初始校正矩阵 $B_0 = I$ 关于 g_0 的等内积分解矩阵

$$C_0^{-1} = \begin{cases} I, & g_0 = -e, \\ I - \frac{2}{\|\sigma_0\|^2} \sigma_0 \sigma_0^T, & g_0 \neq -e, \end{cases} \quad (2.2)$$

其中 $\sigma_0 = \sqrt{n}g_0 + \|g_0\|e$.

由 (2.1) 式定义的 C_{k+1}^{-1} 只有在精确线搜索条件下. 即 $v_k^T w_k = 0$ 时, 有 $B_{k+1} = C_{k+1}C_{k+1}^T$. 若迭代点 x^k 是经过非精确线搜索到 x^{k+1} , 则由 (2.1) 式定义的矩阵 C_{k+1}^{-1} 显

然不再是 (1.2) 式定义的 B_{k+1} 的分解矩阵. 下面给出的结论则与搜索方式无关, 仅依赖于 BFGS 校正矩阵的定义式.

引理 2.1 设 $u, v \in R^n$, 若 $\beta \neq 0, \gamma \neq 0, v^T u = \beta^{-1} + \gamma^{-1}$, 则矩阵 $I - \beta uv^T$ 非奇异, 并且

$$(I - \beta uv^T)^{-1} = I - \gamma uv^T. \quad (2.3)$$

引理 2.2 对于 BFGS 校正公式 (1.2), 若 B_k 关于 g_k 的等内积分解矩阵为 C_k^{-1} . 记

$$N_k = (I + \frac{\|w_k\|}{\sqrt{s_k^T y_k s_k^T g_k}} y_k s_k^T - \frac{g_k s_k^T}{s_k^T g_k}) C_k, \quad (2.4)$$

则 $B_{k+1} = N_k N_k^T$, 并且

$$N_k^{-1} = C_k^{-1} (I - \frac{y_k s_k^T}{s_k^T y_k} + \frac{g_k s_k^T}{\|w_k\| \sqrt{s_k^T y_k}}). \quad (2.5)$$

证 注意到 $s_k = x^{k+1} - x^k = \alpha_k d^k$, 其中 α_k 为步长, 于是 $B_k s_k = -\alpha_k g_k$. 直接计算得

$$\begin{aligned} N_k N_k^T &= (I + \frac{\|w_k\|}{\sqrt{s_k^T y_k s_k^T g_k}} y_k s_k^T - \frac{g_k s_k^T}{s_k^T g_k}) B_k (I + \frac{\|w_k\|}{\sqrt{s_k^T y_k s_k^T g_k}} s_k y_k^T - \frac{s_k g_k^T}{s_k^T g_k}) \\ &= (B_k - \frac{\alpha_k \|w_k\|}{\sqrt{s_k^T y_k s_k^T g_k}} y_k g_k^T + \frac{\alpha_k g_k g_k^T}{s_k^T g_k}) (I + \frac{\|w_k\|}{\sqrt{s_k^T y_k s_k^T g_k}} s_k y_k^T - \frac{s_k g_k^T}{s_k^T g_k}) \\ &= B_k + \frac{\alpha_k g_k g_k^T}{s_k^T g_k} - \frac{\alpha_k \|w_k\|^2}{s_k^T y_k s_k^T g_k} y_k y_k^T \\ &= B_k + \frac{\alpha_k g_k g_k^T}{s_k^T g_k} + \frac{y_k y_k^T}{s_k^T y_k} \\ &= B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{s_k^T y_k}. \end{aligned}$$

所以 $B_{k+1} = N_k N_k^T$. 由于

$$I + \frac{\|w_k\|}{\sqrt{s_k^T y_k s_k^T g_k}} y_k s_k^T - \frac{g_k s_k^T}{s_k^T g_k} = I - (-\frac{\|w_k\|}{\sqrt{s_k^T y_k s_k^T g_k}} y_k + \frac{1}{s_k^T g_k} g_k) s_k^T,$$

由引理 2.1, 易得

$$(I + \frac{\|w_k\|}{\sqrt{s_k^T y_k s_k^T g_k}} y_k s_k^T - \frac{g_k s_k^T}{s_k^T g_k})^{-1} = I - \frac{y_k s_k^T}{s_k^T y_k} + \frac{g_k s_k^T}{\|w_k\| \sqrt{s_k^T y_k}},$$

从而有 $N_k^{-1} = C_k^{-1} (I - \frac{y_k s_k^T}{s_k^T y_k} + \frac{g_k s_k^T}{\|w_k\| \sqrt{s_k^T y_k}})$. 证毕.

引理 2.3 若 $0 \neq \delta_k \in R^n$, 记 $Q_k = I - \frac{2\sigma_k \sigma_k^T}{\|\sigma_k\|^2}$, 其中 $\sigma_k = \delta_k + \frac{\|\delta_k\|}{\|w_k\|} w_k$, 则

$$Q_k \delta_k = -\frac{\|\delta_k\|}{\|w_k\|} w_k.$$

证 直接计算可知 $2\sigma_k^T \delta_k = \|\sigma_k\|^2$, 故

$$Q_k \delta_k = \delta_k - \frac{2\sigma_k^T \delta_k}{\|\sigma_k\|^2} \sigma_k = \delta_k - \sigma_k = -\frac{\|\delta_k\|}{\|w_k\|} w_k.$$

证毕.

当记

$$\delta_k = N_k^{-1} g_{k+1} = p_k v_k + q_k w_k, \quad (2.6)$$

其中 $p_k = -\frac{s_k^T g_k}{s_k^T y_k}$, $q_k = \frac{(\|w_k\| + \sqrt{s_k^T y_k}) s_k^T g_{k+1}}{\|w_k\| s_k^T y_k}$, 则有如下结论.

定理 2.1 对于 BFGS 校正公式 (1.2), 若已求得 B_k 关于 g_k 的等内积分解矩阵 C_k^{-1} , 则 B_{k+1} 关于 g_{k+1} 的等内积分解矩阵

$$C_{k+1}^{-1} = Q_k C_k^{-1} \left(I - \frac{y_k s_k^T}{s_k^T y_k} + \frac{g_k s_k^T}{\|w_k\| \sqrt{s_k^T y_k}} \right), \quad (2.7)$$

其中 $Q_k = I - \frac{2\sigma_k \sigma_k^T}{\|\sigma_k\|^2}$, $\sigma_k = \delta_k + \frac{\|\delta_k\|}{\|w_k\|} w_k$.

证 注意到 Q_k 为 Householder 矩阵, $Q_k Q_k^T = I$, 所以

$$C_{k+1} C_{k+1}^T = \left(I + \frac{\|w_k\|}{\sqrt{s_k^T y_k s_k^T g_k}} y_k s_k^T - \frac{g_k s_k^T}{s_k^T g_k} \right) B_k \left(I + \frac{\|w_k\|}{\sqrt{s_k^T y_k s_k^T g_k}} s_k y_k^T - \frac{s_k g_k^T}{s_k^T g_k} \right).$$

故由引理 2.2 知 $B_{k+1} = C_{k+1} C_{k+1}^T$. 再由引理 2.3 得

$$C_{k+1}^{-1} g_{k+1} = Q_k \delta_k = -\frac{\|\delta_k\|}{\|w_k\|} w_k = a_{k+1} e, \quad (2.8)$$

其中 $a_{k+1} = -\frac{\|\delta_k\|}{\|w_k\|} a_k$. 故由 (2.7) 式定义的 C_{k+1}^{-1} 为 B_{k+1} 关于 g_{k+1} 的等内积分解矩阵.

从上述结论的证明过程可以发现, 由 (2.7) 式定义的 C_{k+1}^{-1} 是 (1.2) 式定义的 B_{k+1} 关于 g_{k+1} 的等内积分解矩阵, 这个结论与搜索方式是无关的.

推论 2.1 在精确线搜索条件下, (2.7) 式与 (2.1) 式等价, 即 (2.1) 式为 (2.7) 式的特例.

证 注意到在精确线搜索条件下 $s_k^T g_{k+1} = 0$, $s_k^T g_k = -s_k^T y_k$, 所以由 (2.6) 式得 $\delta_k = N_k^{-1} g_{k+1} = v_k$, 从而 $\sigma_k = \delta_k + \frac{\|\delta_k\|}{\|w_k\|} w_k = \frac{\|v_k\|}{\|w_k\|} (w_k + t_k v_k) = \frac{\|v_k\|}{\|w_k\|} \tilde{\sigma}_k$, 其中 $\tilde{\sigma}_k = w_k + t_k v_k$. 故 $Q_k = I - \frac{2\sigma_k \sigma_k^T}{\|\sigma_k\|^2} = I - \frac{2\tilde{\sigma}_k \tilde{\sigma}_k^T}{\|\tilde{\sigma}_k\|^2}$. 又在精确线搜索条件下有 $\|w_k\| \sqrt{s_k^T y_k} = \frac{1}{\sqrt{\lambda_k}} s_k^T y_k$. 利用以上结果, 由 (2.7) 式可直接推出 (2.1) 式. 所以在精确线搜索条件下 (2.7) 式与 (2.1) 式等价. 证毕.

3 利用分解校正矩阵确定搜索方向的 BFGS 算法

显然, 把由 (2.2) 式和 (2.7) 式为分解校正矩阵, 由 (1.9) 式计算搜索方向的步骤替换到 BFGS 算法中对应的步骤里, 就可以构成利用分解校正矩阵确定搜索方向的 BFGS 新算法. 由于分解校正矩阵的迭代运算仅是非奇异矩阵到非奇异矩阵的迭代运算, 因此计算精度的要求应低于原 BFGS 算法中对称正定矩阵到对称正定矩阵的迭代运算.

新的 BFGS 算法:

步 1 取初始点 x^0 , $\varepsilon \geq 0$, 计算 $g_0 = g(x^0)$. 若 $\|g_0\| \leq \varepsilon$, 则停止运算; 否则转步 2.

步 2 令 $k = 0$; 对初始矩阵 $B_0 = I$ 作关于 g_0 的等内积分解, 求出分解矩阵 C_0^{-1} 和 $a_0 = -\frac{\|g_0\|}{\sqrt{n}}$.

步 3 求搜索方向 $d^k: d^k = -a_k(d_1^k, \dots, d_n^k)^T$. 其中 $d_j^k, (j = 1, 2, \dots, n)$ 是 C_k^{-1} 的第 j 列元素之和.

步 4 一维搜索: 由线性搜索求出步长因子 α_k .

步 5 令 $x^{k+1} = x^k + \alpha_k d^k$, 计算 $g_{k+1} = g(x^{k+1})$, 若 $\|g_{k+1}\| \leq \varepsilon$, 则停止计算; 否则转步 6.

步 6 由公式 (2.7) 求出 B_{k+1} 关于 g_{k+1} 的等内积分解矩阵 C_{k+1}^{-1} 和系数 $a_{k+1} = -\frac{\|g_{k+1}\|}{\|w_k\|} a_k$.

步 7 令 $k = k + 1$, 转步 3.

参 考 文 献

- [1] 柳力. 一个矩阵分解算法的推广及应用 [J]. 数学的实践与认识, 2011, 41(21): 239–243.
- [2] 柳力. 由校正矩阵的等内积分解矩阵确定搜索方向的拟牛顿算法 [J]. 数学的实践与认识, 2013, 43(10): 214–219.
- [3] Gill P E, Murray W. Quasi-Newton methods for unconstrained optimization[J]. Insti. Math. Appl., 1972,(9): 91–108.
- [4] 席少霖. 非线性最优化方法 [M]. 北京: 高等教育出版社, 1992.
- [5] 邓乃扬等. 无约束最优化计算方法 [M]. 北京: 科学出版社, 1982.

BFGS ALGORITHM BY USING THE DECOMPOSITION MATRIX OF THE CORRECTION MATRIX TO OBTAIN THE SEARCH DIRECTION

LIU Li

(Department of Teaching, TV University of Jilin City, Jilin 132002, China)

Abstract: In this paper, the equal inner product decomposition algorithm of positive definite matrix is applied to improve the search direction calculation in BFGS algorithm. By setting up the correction matrixes of both independent search mode and decomposition matrixes expression, BFGS algorithm is put forward, in which search directions are obtained by using equal inner product decomposition matrixes of the Hesse approximate matrixes.

Keywords: BFGS algorithm; correction matrix; equal inner product decomposition; search direction; algorithm

2010 MR Subject Classification: 90C30