doi:10.16055/j.issn.1672-058X.2015.0005.005

一类修正 DY 共轭梯度法及其全局收敛性

马 文 亚

(重庆师范大学 数学学院,重庆 401331)

摘 要:对 DY 共轭梯度方法进行修正,使得修正的共轭梯度方法(MDY^*)在 Wolfe 线搜索下满足充分下降条件和全局收敛性.

关键词:修正 DY 共轭梯度法;Wolfe 线搜索;充分下降性;全局收敛性

中图分类号:O224 **文献标识码**:A **文章编号**:1672-058X(2015)05-0017-03

1 基础知识

考虑无约束优化问题:

$$\min f(x),x \in \mathbf{R}^n \tag{1}$$

其中 $f:\mathbf{R}^n \to \mathbf{R}$ 是一个连续可微的函数.

共轭梯度法是求解问题(1)的具有如下的迭代格式一类方法:

$$x_{k+1} = x_k + \alpha_k d_k, k = 0,1,\cdots \tag{2}$$

其中 α_k 是某种线搜索下的步长,d_k 满足

$$d_k = \begin{cases} -g_k, & k = 0 \\ -g_k + \beta_k d_{k-1}, & k \geq 1 \end{cases} \tag{3}$$

其中 β_k 是共轭参数.$g_0 \leq \varepsilon$ 的不同取法对应于不同的非线性共轭梯度法.著名的 HS[1],DY[2] 方法的参数 β_k 分别为

$$\beta_k^{HS} = \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}}, \beta_k^{DY} = \frac{g_k{}^2}{d_{k-1}^T y_{k-1}}$$

其中,$\|\cdot\|$ 为欧几里得范数,$y_{k-1} = g_k - g_{k-1}$.

其中 HS 方法被公认为是数值效果较好的方法,而 DY 方法则有较好的收敛性.最近,文献[3]对文献[4]的 MHS 公式进行改进,得到如下新的 β_k 公式:

$$\beta_k^{new} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|d_{k-1}\|} | g_k^T d_{k-1} |}{d_{k-1}^T y_{k-1}} \tag{4}$$

并证明了由新公式产生的算法在 Wolfe 线搜索条件下满足充分下降性和全局收敛性.

Dai 和 Wen[5] 对文献[6]的 NHS 公式进行改进,得到如下新的 β_k 公式:

收稿日期:2014-08-24;**修回日期**:2014-10-11.

作者简介:马文亚(1988-),女,陕西杨凌人,硕士研究生,从事最优化计算方法及理论研究.

$$\beta_k^{\mathrm{DHS}}=\frac{\|g_k\|^2-\frac{\|g_k\|}{\|g_{k-1}\|}|g_k^{\mathrm{T}}g_{k-1}|}{d_{k-1}^{\mathrm{T}}y_{k-1}+\mu|g_k^{\mathrm{T}}d_{k-1}|},\mu>1 \tag{5}$$

并证明了由新公式产生的算法不仅每步迭代都可以产生一个充分下降方向,且在 Wolfe 线搜索条件下全局收敛.

受文献[3][5]启发,此处提出如下新的 β_k 修正公式:

$$\beta_k^{\mathrm{MDY}^*}=\frac{\|g_k\|^2-\frac{\|g_k\|}{\|d_{k-1}\|}|g_k^{\mathrm{T}}d_{k-1}|}{d_{k-1}^{\mathrm{T}}y_{k-1}+\mu|g_k^{\mathrm{T}}d_{k-1}|} \tag{6}$$

其中参数 $\mu>1$.易见,在精确线搜索下,$\beta_k^{\mathrm{MDY}^*}=\beta_k^{\mathrm{DY}}$.

在共轭梯度法的许多理论分析和数值实现中,常常使用 Wolfe 线搜索,其要求 α_k 满足

$$f(x_k+\alpha_kd_k)\leqslant f(x_k)+\delta\alpha_kg_k^{\mathrm{T}}d_k \tag{7}$$

$$g(x_k+\alpha_kd_k)^{\mathrm{T}}d_k\geqslant\sigma g_k^{\mathrm{T}}d_k \tag{8}$$

其中 $0<\delta\leqslant\sigma<1$.

2 算 法

MDY* 算法:

初始步:给定初始点 $x_0\in\mathbf{R}^n$,$\delta\in(0,1)$,$\sigma\in(0,1)$,$\mu>1$,$\varepsilon>0$,令 $d_0=-g_0$,$k:=0$;

步骤 1:若$g_0\leqslant\varepsilon$,则停止;

步骤 2:由 Wolfe 线搜索准则计算步长 α_k;

步骤 3:令 $x_{k+1}=x_k+\alpha_kd_k$,若$\|g_{k+1}\|\leqslant\varepsilon$,则停止;

步骤 4:由式(6)计算 β_{k+1},由式(2)计算 d_{k+1};

步骤 5:令 $k:=k+1$,转步骤 2.

3 MDY* 算法的收敛性

为了证明新算法的全局收敛性,首先给出两个假设:

(A) 水平集 $\Omega=\{x\in\mathbf{R}^n|f(x)\leqslant f(x_0)\}$有下界,其中 $x_0\in\mathbf{R}^n$ 为初始点.

(B) f 在水平集 Ω 的一个领域 N 内连续可微,且其梯度 g 满足 Lipschitz 连续,即存在常数 $L>0$,使得

$$\|g(x)-g(y)\|\leqslant Lx-y,\forall x,y\in N \tag{9}$$

引理 1[7] 假设(A)(B)成立,设序列$\{g_k\}$和$\{d_k\}$由 MDY* 算法产生,并设步长 α_k 由 Wolfe 线搜索式(7)、(8)计算,则

$$\sum_{k=0}^{\infty}\frac{(g_k^{\mathrm{T}}d_k)^2}{\|d_k\|^2}<+\infty \tag{10}$$

引理 2 假设(A)(B)成立,设序列$\{g_k\}$和$\{d_k\}$由 MDY* 算法产生,并设步长 α_k 由 Wolfe 线搜索式(7)、(8)计算,则对所有的 $k\geqslant0$,有

$$g_k^{\mathrm{T}}d_k\leqslant-\left(1-\frac{1}{u}\right)\|g_k\|^2,u>1 \tag{11}$$

证明 用归纳法证明.

当 $k=0$ 时,有 $g_0^{\mathrm{T}}d_0=-\|g_0\|^2\leqslant-\left(1-\dfrac{1}{u}\right)\|g_0\|^2$,因此式(11)成立.

假设 $n=k-1(k\geqslant1)$ 时,式(11)成立,则由式(8)得

$$g_{k-1}^{\mathrm{T}}y_{k-1}\geqslant(1-\sigma)(-g_{k-1}^{\mathrm{T}}d_{k-1})\geqslant(1-\sigma)(1-\frac{1}{u})\|g_{k-1}\|^2>0 \tag{12}$$

由式(6)得

$$0\leqslant\beta_k^{\mathrm{MDY}^*}\leqslant\frac{\|g_k\|^2}{u\,|g_k^{\mathrm{T}}d_{k-1}|} \tag{13}$$

因此,有

$$g_k^{\mathrm{T}}d_k\leqslant-\|g_k\|^2+|\beta_k^{\mathrm{MDY}^*}|\,|g_k^{\mathrm{T}}d_{k-1}|\leqslant-\|g_k\|^2+\frac{\|g_k\|^2}{u\,|g_k^{\mathrm{T}}d_{k-1}|}|g_k^{\mathrm{T}}d_{k-1}|\leqslant-\left(1-\frac{1}{u}\right)\|g_k\|^2 \tag{14}$$

定理 1　假设(A)(B)成立,设序列$\{g_k\}$和$\{d_k\}$由 MDY*算法产生,并设步长 α_k 由 Wolfe 线搜索式(7)(8)计算,假设存在一个正常数 α^* 满足 $\alpha_k\geqslant\alpha^*$,则

$$\lim_{k\to\infty}\|g_k\|=0 \tag{15}$$

证明　由假设(A)(B)知,存在一个常数 $M>0$,使得

$$\|\alpha_k d_k\|=\|s_k\|\leqslant M \tag{16}$$

由式(16)和 $\alpha_k\geqslant\alpha^*$,有

$$\|d_k\|\leqslant\frac{M}{\alpha^*}$$

由式(10)、(11)、(16),知式(14)成立.

参考文献:

[1] HESTENES M R,STIEFEL E.Methods of Conjugate Gradients for Solving Linear Systems[M].NBS,1952

[2] DAI Y H,YUAN Y X.A Nonlinear Conjugate Gradient Method with a Strong Global Convergence Property[J].SIAM Journal on Optimization,1999,10(1):177-182

[3] 江羡珍,马国栋,简金宝.Wolfe 线搜索下的一个新的全局收敛性共轭梯度法[J].工程数学学报,2011,28(6):779-786

[4] YAO S W,WEI Z X,HUANG H.A Note about WYL′s Conjugate Gradient Method and Its Applications[J].Applied Mathematics and Computation,2007(2):381-388

[5] DAI Z F,WEN F H.Another Improved Wei-Yao-Liu Nonlinear Conjugate Gradient Method with Sufficient Descent Property[J].Applied Mathematics and Computation,2012(14):7421-7430

[6] ZHANG L. An Improved Wei-Yao-Liu Nonlinear Conjugate Gradient Method for Optimization Computation [J]. Applied Mathematics and Computation,2009(6):2269-2274

[7] ZOUTENDIJK G. Nonlinear Programming, Computational Methods, in Integer and Nonlinear Programming[M]. North-Holland, Amsterdam,1970

A Class of Modified DY Conjugate Gradient Method and Its Global Convergence

MA Wen-ya

(College of Mathematics Science, Chongqing Normal University, Chongqing 401331, China)

Abstract: Dai-Yuan conjugate gradient method is modified to make it satisfy sufficient descent condition and global convergence with Wolfe line search.

Key words: modified Dai-Yuan conjugate gradient method; Wolfe line search; sufficient descent condition; global convergence