求解对称非线性方程组的两种修正 MHS 无导数型共轭梯度法

沈冬梅[1],杨忠选[2]

(1.南昌工学院 教育学院,江西 南昌 330108;2.华东交通大学 理学院,江西 南昌 330108)

摘　要:针对对称非线性方程组的求解问题,使用近似梯度代替精确梯度,构造了两种修正的 MHS 无导数型共轭梯度法,这两个算法的显著特征是搜索方向均为下降方向,且适用于求解大规模的对称非线性方程组问题。在适当的条件下,充分利用方程组的对称性结构,证明了算法的全局收敛性。数值试验表明这两种算法是求解大规模对称非线性方程组的有效算法。

关键词:对称非线性方程组;全局收敛性;修正 MHS 无导数型共轭梯度法

中图分类号:O224　　**文献标志码**:A　　**文章编号**:2095-4824(2021)06-0106-06

考虑对称非线性方程组(1)的求解问题

$$F(x) = 0, \tag{1}$$

式中: $F: R^n \to R^n$ 为连续可微函数,其雅克比矩阵是对称的,即 $J(x) = J(x)^T$。

对称非线性方程组(1)的求解已被广泛研究。Li 和 Fukushima[1] 提出了利用 Gauss-Newton-based BFGS 的 Derivative-Free(无导数)线性搜索求解对称非线性方程组,并证明了算法的全局收敛性。文献[2-5]分别利用无导数修正 FR 算法、无导数修正 PRP 算法、无导数修正 CD 算法、无导数修正 HS 算法求解对称非线性方程组,并证明了算法的全局收敛性。Zhou 和 Shen[6] 提出了两种近似 PRP 型无导数方法,并用于求解该问题,同时证明其具有全局收敛性。徐瑞昌[7] 利用修正的 BFGS 算法求解对称非线性方程组,并指出算法具有全局收敛性和超线性收敛速度。沈冬梅等[8] 证明了近似 PRP 算法具有超线性收敛速度。Abubakar 等[9] 提出了一个修正的 Dai-Liao 共轭梯度法用于求解对称非线性方程组,并证明了该算法的全局收敛性。Zhou[10] 提出了求解对称线性方程组的无导数 MBFGS 拟牛顿法,并证明该算法的全局收敛性。Sabi'u 等[11] 构造了一个修正的 PRP 共轭梯度法求解大规模的非线性对称方程组,并证明了其全局收敛性。

本文继续研究求解对称非线性方程组的下降无导数共轭梯度算法。这里首先回顾文献[12]提出的两种修正的 HS 算法,分别为修正的 MHS 共轭梯度法算法(简称为 MTTHS 算法)和保守的 MHS 共轭梯度法算法(简称为 CTTHS):

$$x_{k+1} = x_k + \lambda_k d_k$$

$$d_k^{MTTHS} = \begin{cases} - \ f(x_k), & \text{如果 } k = 0 \\ - \ f(x_k) + \beta_k^{MHS} d_{k-1} - \theta_k^M z_{k-1}, & \text{如果 } k > 0 \end{cases}, \tag{2}$$

$$d_k^{CTTHS} = \begin{cases} - \ f(x_k), & \text{如果 } s_{k-1}^T y_{k-1} < \epsilon \| \ f(x_{k-1}) \|^r s_{k-1}^T s_{k-1} \\ - \ f(x_k) + \beta_k^{HS} d_{k-1} - \theta_k y_{k-1}, & \text{否则} \end{cases}, \tag{3}$$

式中: $\ f(x_k)$ 为 $f(x)$ 在 x_k 处的梯度, $s_k = x_{k+1} - x_k = \lambda_k d_k$, $y_{k-1} = \ f(x_k) - \ f(x_{k-1})$,

收稿日期:2021-09-09

基金项目:江西省教育厅科学技术研究项目(GJJ191110)

作者简介:沈冬梅(1988-),女,湖北广水人,南昌工学院教育学院讲师,硕士。
杨忠选(1990-),男,江西九江人,华东交通大学理学院讲师,硕士。

$\beta_k^{MHS} = \dfrac{f(x_k)^T z_{k-1}}{d_{k-1}^T z_{k-1}}$，$\theta_k^M = \dfrac{f(x_k)^T d_{k-1}}{d_{k-1}^T z_{k-1}}$，$\beta_k^{HS} = \dfrac{f(x_k)^T y_{k-1}}{d_{k-1}^T y_{k-1}}$，$\theta_k = \dfrac{f(x_k)^T d_{k-1}}{d_{k-1}^T y_{k-1}}$。

$$z_k = f(x_{k+1}) - f(x_k) + t\| f(x_k)\|^r s_k = y_k + t\| f(x_k)\|^r s_k \tag{4}$$

式中：$r \geqslant 0$ 及 $t > 0$ 均为常数。

这两种算法的显著特点是在不依赖于任何线性搜索的条件下均能产生下降方向，即 $f(x_k)^T d_k = -\|g_k\|^2$，并在 Wolfe 线性搜索下均具有全局收敛性。本文的想法来自于文献[10,12]，提出了求解(1)式的 MTTHS 无导数型共轭梯度法和 CTTHS 无导数型共轭梯度法。在较弱的条件下证明其全局收敛性，并通过数值实验对所提算法进行测试。

1 修正的 MHS 无导数型共轭梯度算法

对称非线性方程组(1)的求解可转化为如下无约束极小化问题：

$$\min \quad f(x) = \frac{1}{2}\| F(x)\|^2, \quad x \in R^n \tag{5}$$

然而，由(5)式定义的 $f(x)$ 的梯度为 $f(x) = J(x)^T F(x) = J(x)F(x)$，其雅克比矩阵的计算量较大，显然不适合计算大规模问题。为避免计算复杂的梯度，文献[1]提出了一种近似梯度，即

$$g_k = \frac{F(x_k + \lambda_{k-1} F_k) - F(x_k)}{\lambda_{k-1}}, \tag{6}$$

显然，

$$g_k \approx f(x_k) = F'(x_k)F(x_k) = \lim_{\lambda \to 0} \frac{F(x + \lambda F(x)) - F(x)}{\lambda}。 \tag{7}$$

本文利用近似梯度函数 g_k 的计算方式，建立求解对称非线性方程组的 MTTHS 和 CTTHS 无导数型共轭梯度法。这里我们通过以下两个过程确定 MTTHS 无导数型共轭梯度法的搜索方向和搜索步长。

过程 1 确定 MTTHS 无导数型共轭梯度法搜索方向：

将(6)式运用到 MTTHS 共轭梯度法中，把最速下降方向 $-f(x_{k-1})$、$-f(x_k)$ 对应的替换为 $-g_{k-1}$、$-g_k$，得到 MTTHS 无导数型共轭梯度法的搜索方向，即

$$d_k = \begin{cases} -g_k, & k = 0 \\ -g_k + \beta_k^{MHS} d_{k-1} - \theta_k^M z_{k-1}, & k > 0 \end{cases} \tag{8}$$

式中：$\beta_k^{MHS} = \dfrac{g_k^T z_{k-1}}{d_{k-1}^T z_{k-1}}$，$\theta_k = \dfrac{g_k^T d_{k-1}}{d_{k-1}^T z_{k-1}}$，$y_{k-1} = g_k - g_{k-1}$。

通过计算，易有 $g_k^T d_k = -\|g_k\|^2$，当 $\lambda > 0$ 且充分小时，$g_k \approx f(x_k)$，$k \geqslant 1$。

而且，由(7)式知 $\lim\limits_{\lambda \to 0} g_k(\lambda) = f(x_k)$，于是

$$\bar{d}_k = \lim_{\lambda \to 0} d_k(\lambda) = \begin{cases} -f(x_k), & k = 0 \\ -f(x_k) + \bar{\beta}_k^{MHS} d_{k-1} - \bar{\theta}_k^M \bar{z}_{k-1}, & k > 0 \end{cases}$$

式中：$\bar{\beta}_k^{MHS} = \dfrac{g_k^T z_{k-1}}{d_{k-1}^T z_{k-1}}$，$\bar{\theta}_k^M = \dfrac{g_k^T d_{k-1}}{d_{k-1}^T z_{k-1}}$，$\bar{y}_{k-1} = g_k - g_{k-1}$。

从而有，$f(x_k)^T \bar{d}_k = -\| f(x_k)\|^2$。

这表明当 λ 充分小时，$d_k(\lambda)$ 是问题(5)中函数 $f(x)$ 在 x_k 处的下降方向。

过程 2 确定 MTTHS 无导数型共轭梯度法的步长因子：

对于搜索方向 $d_k(\lambda)$ 中的步长 λ，借用文献[13]的方式，令 $\lambda = \max\{\rho^j, j = 0, 1, 2 \cdots\}$ 满足下列不等式：

$$f(x_k + \lambda d_k) \leqslant f(x_k) - \sigma_1 \|\lambda F(x_k)\|^2 - \sigma_2 \|\lambda d_k(\lambda)\|^2 + \eta_k f(x_k) \tag{9}$$

式中：$0 < \rho < 1$，$\sigma_1 > 0$ 和 $\sigma_2 > 0$，正数序列 $\{\eta_k\}$ 满足 $\sum\limits_{k=0}^{\infty} \eta_k \leqslant \eta < \infty$。

基于搜索方向和步长的确定，建立求解(5)的 MTTHS 型无导数算法，步骤如下：

算法 1 （MTTHS 无导数型共轭梯度法）

步骤 1 给定初始点 $x_0 \in R^n$，参数 $0 < \rho < 1$，$\sigma_1 > 0$ 和 $\sigma_2 > 0$，令 $k := 0$；

步骤 2 利用(8)式计算搜索方向 d_k；

步骤 3 通过(9)式计算步长因子 λ_k；

步骤 4 令 $x_{k+1} = x_k + \lambda_k d_k$；

步骤 5 令 $k := k + 1$，转步骤 2。

类似于 MTTHS 型无导数共轭梯度法搜索方向和步长的确定方式，可得到 CTTHS 型无导数共轭梯度法算法。

算法 2 （CTTHS 无导数型共轭梯度法）

步骤 1 给定初始点 $x_0 \in R^n$，参数 $0 < \rho < 1$，$\sigma_1 > 0$ 和 $\sigma_2 > 0$，$\varepsilon_1 > 0$，$r > 0$，令 $k := 0$；

步骤 2 按照下式计算搜索方向 d_k：

$$d_k = \begin{cases} -g_k, & \text{如果 } s_{k-1}^T y_{k-1} < \varepsilon_1 \|g_{k-1}\|^r s_{k-1}^T s_{k-1} \\ -g_k + \beta_k^{HS} d_{k-1} - \theta_k y_{k-1}, & \text{否则} \end{cases} \quad (10)$$

式中：$s_{k-1} = x_k - x_{k-1}$，$\beta_k^{HS} = \dfrac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}}$，$\theta_k = \dfrac{g_k^T d_{k-1}}{d_{k-1}^T y_{k-1}}$；

步骤 3　利用(9)式计算搜索步长 λ_k；

步骤 4　令 $x_{k+1} = x_k + \lambda_k d_k$；

步骤 5　令 $k := k+1$，转步骤 2。

2　算法的全局收敛性

假设 A

1) 水平集 $\Omega = \{x \in R^n \mid f(x) \leqslant e^{\frac{\eta}{2}} f(x_0)\}$ 有界；

2) $F(x) = 0$ 为对称非线性方程组；

3)在 Ω 的邻域 N 内，$F(x)$ 是连续可微的且其梯度是 Lipschitz 连续的，即存在常数 $\overline{L} > 0$ 使得

$\|F'(x) - F'(y)\| \leqslant \overline{L}\|x - y\|$, $\forall x, y \in N$。

由假设 A 知，对 $x \in N$ 有，$\|F(x)\| \leqslant \gamma_1$，$\|F'(x)\| \leqslant \gamma_2$，$\|\nabla f(x)\| \leqslant \gamma_3$。

由假设 A 知，对 $x, y \in N$ 有，

$$\|\nabla f(x) - \nabla f(y)\| \leqslant L_1\|x - y\|,$$
$$\|F(x) - F(y)\| \leqslant L_2\|x - y\|。$$

引理 1[13]　设序列 $\{x_k\}$ 由 MTTHS 无导数型共轭梯度法或 CTTHS 无导数型共轭梯度法产生，则 $\forall k \geqslant 0$，函数值序列 $\{\|F_k\|\}$ 收敛且序列 $\{x_k\} \subset \Omega$。

引理 2[13]　设序列 $\{x_k\}$ 由 MTTHS 无导数型共轭梯度法或 CTTHS 无导数型共轭梯度法产生，则

$$\sum_{k=0}^{\infty} \|\lambda_k F(x_k)\|^2 < \infty,\ \sum_{k=0}^{\infty} \|\lambda_k d_k\|^2 < \infty。$$

由引理 2 知，$\|\lambda_k d_k\|$ 有界，即 $\|s_k\|$ 有界。

引理 3[13]　设假设 A 的条件成立，序列 $\{x_k\}$ 由 MTTHS 无导数型共轭梯度法或 CTTHS 无导数型共轭梯度法产生，则有

$$\lim_{k\to\infty} \|g_k - g_{k-1}\| = 0。\quad (11)$$

由引理 3 知，存在一个正常数 M_1，使得 $\|g_k - g_{k-1}\| \leqslant M_1$。由(4)式和假设 A 容易得到下面的引理 4。

引理 4　设假设 A 成立，$\{x_k\}$ 由 MTTHS 无导数型共轭梯度法产生，则对任意的 $k = 0, 1, 2, \cdots$，序列 $\{\|z_k\|\}$ 有界。

证明　由 g_k 的定义知 $\|g_k\| = \left\|\dfrac{F(x_k + \lambda_{k-1}F(x_k)) - F(x_k)}{\lambda_{k-1}}\right\| \leqslant$

$L_2 F(x_k) \leqslant \gamma_1 L_2 = L$。

由 z_k 的定义得

$\|z_k\| \leqslant \|g(x_{k+1}) - g(x_k)\| + \|t\|g(x_k)\|^r s_k\|$

$\leqslant \|g(x_{k+1}) - g(x_k)\| + t\|g(x_k)\|^r\|s_k\|$，

而 $\|g(x_{k+1}) - g(x_k)\| \leqslant M_1$，$\|s_k\|$ 有界。

因此，我们有

$$\|z_k\| \leqslant M_1 + tL^r\|s_k\| \leqslant \tilde{\eta}。$$

下面的引理表明搜索方向 d_k 有界。

引理 5　设 $\{x_k\}$ 由 MTTHS 无导数型共轭梯度法产生，若存在一个常数 a 和常量 $\varepsilon > 0$，使得对 $\forall k \in \Omega$，有 $\|\nabla f(x)\| \geqslant \varepsilon$，$\|d_{k-1}^T z_{k-1}\| \geqslant a\varepsilon$ 成立，那么存在一个常数 $M > 0$ 使得

$$\|d_k\| \leqslant M,\ \forall k \geqslant 0。\quad (12)$$

证明　由 d_k 的定义易有

$$\|d_k\| \leqslant \|g_k\| + 2\frac{\|g_k\|\|z_{k-1}\|}{d_{k-1}^T z_{k-1}}\|d_{k-1}\|$$

$$\leqslant L + 2\frac{\|g_k\|\|z_{k-1}\|}{d_{k-1}^T z_{k-1}}\|d_{k-1}\| \leqslant L + \frac{2L\tilde{\eta}}{a\varepsilon}\|d_{k-1}\|$$

$$\leqslant L + \eta_1\|d_{k-1}\| \leqslant L(1 + \eta_1 + \eta_1^2 + \cdots + \eta_1^{k-k_0-1}) + \eta_1^{k-k_0}\|d_{k_0}\| \leqslant \frac{L}{1-\eta_1} + \|d_{k_0}\| \leqslant M$$

式中：$\eta_1 = \dfrac{2L\tilde{\eta}}{a\varepsilon}$，$M = \max\{\|d_1\|, \|d_2\|, \cdots, \|d_{k_0}\|, \dfrac{L}{1-\eta_1} + \|d_{k_0}\|\}$。

下面的定理表明 MTTHS 无导数型共轭梯度算法具有全局收敛性。

定理 1　若假设 A 成立，$\{x_k\}$ 由 MTTHS 无导数型共轭梯度法产生，且 $\|d_{k-1}^T z_{k-1}\| \geqslant a\varepsilon$，则有

$$\liminf_{k\to\infty} \|\nabla f(x_k)\| = 0。\quad (13)$$

证明(反证法)　假设(13)式不成立，则存在一个正常数 τ 使得

$$\|\nabla f(x_k)\| \geqslant \tau,\ \forall k \geqslant 0。\quad (14)$$

由 $\nabla f(x_k) = F'(x_k)^T F(x_k)$ 和(14)式知，存在常数 $\tau_1 > 0$，使得

$$\|F(x_k)\| \geqslant \tau_1。$$

1)若 $\limsup_{k\to\infty} \lambda_k > 0$，由引理 2 知，有

$$\liminf_{k\to\infty} \|F(x_k)\| = 0。$$

由上式和引理 1 知 $\lim_{k\to\infty} \|F(x_k)\| = 0$，这与 $\|F(x_k)\| \geqslant \tau_1$ 矛盾。

2)若 $\limsup\limits_{k\to\infty} \lambda_k = 0$，则由 $\lambda_k \geqslant 0$ 知

$\lim\limits_{k\to\infty} \lambda_k = 0$。

由上式和 MTTHS 型无导数共轭梯度算法知 $\lambda'_k = \frac{\lambda_k}{\lambda}$ 不满足(9)式，即

$$f(x_k + \lambda'_k d_k) - f(x_k) > -\sigma_1 \| \lambda'_k F(x_k) \|^2 - \sigma_2 \| \lambda'_k d_k \|^2 + \eta_k f(x_k)$$
$$> -\sigma_1 \| \lambda'_k F(x_k) \|^2 - \sigma_2 \| \lambda'_k d_k \|^2,$$

这意味着

$$\frac{f(x_k + \lambda'_k d_k) - f(x_k)}{\lambda'_k} > -\sigma_1 \lambda'_k f(x_k) - \sigma_2 \lambda'_k \| d_k \|^2。\tag{15}$$

另一方面，由中值定理，存在一个常数 $t_k \in (0,1)$，使得

$$\frac{f(x_k + \lambda'_k d_k) - f(x_k)}{\lambda'_k} = f(x_k + t_k \lambda'_k d_k)^T d_k \tag{16}$$
$$= f(x_k + t_k \lambda'_k d_k)^T (-g_k + \beta_k^{MHS} d_{k-1} - \theta_k^M z_{k-1})$$
$$= - f(x_k + t_k \lambda'_k d_k)^T g_k + f(x_k + t_k \lambda'_k d_k)^T \beta_k^{MHS} d_{k-1} - f(x_k + t_k \lambda'_k d_k)^T \theta_k^M z_{k-1}。$$

因为 $|\beta_k^{MHS}| = \left| \frac{g_k^T z_{k-1}}{d_{k-1}^T z_{k-1}} \right| \leqslant \frac{\| g_k^T \| \| z_{k-1} \|}{a\varepsilon} = \frac{\gamma \| z_{k-1} \|}{a\varepsilon}$，而

$$\| z_k \| \leqslant \| g(x_{k+1}) - g(x_k) \| + t \| g(x_k) \|^r \| s_k \|$$
$$\leqslant \| g(x_{k+1}) - g(x_k) \| + t \| g(x_k) \|^r \| \lambda_{k-1} d_{k-1} \|$$

由引理 2 和引理 3 知 $\lim\limits_{k\to\infty} \| g_k - g_{k-1} \| = 0$，$\lim\limits_{k\to\infty} \| \lambda_{k-1} d_{k-1} \| = 0$。

因而 $\lim\limits_{k\to\infty} z_k = 0$，继而 $\lim\limits_{k\to\infty} \beta_k^{MHS} = 0$。

因为 $|\theta_k^M| = \left| \frac{g_k^T d_{k-1}}{d_{k-1}^T z_{k-1}} \right| \leqslant \frac{\| g_k^T \| \| d_{k-1} \|}{d_{k-1}^T z_{k-1}} \leqslant \frac{LM}{a\varepsilon}$，故 θ_k^M 有界。

又因为 $\| f(x_k + t_k \lambda'_k d_k)^T \|$，$\| d_{k-1} \|$ 也有界，故

$$\lim_{k\to\infty} inf \quad f(x_k + t_k \lambda'_k d_k)^T \beta_k^{MHS} d_{k-1} = 0, \tag{17}$$
$$\lim_{k\to\infty} inf \quad f(x_k + t_k \lambda'_k d_k)^T \theta_k^M z_{k-1} = 0。\tag{18}$$

又因为 $\{x_k\} \subset \Omega$ 有界，不妨假设 $x_k \to x^*$，由(6)、(8)和(17)、(18)式，我们有

$$\lim_{k\to\infty} d_k = -\lim_{k\to\infty} g_k + \lim_{k\to\infty} \beta_k^{MHS} d_{k-1} - \lim_{k\to\infty} \theta_k^M z_{k-1} = - f(x^*)。\tag{19}$$

另一方面，我们有 $\lim\limits_{k\to\infty} f(x_k + t_k \lambda'_k d_k) = f(x^*)$。 (20)

由(15)-(16)式和(19)-(20)式得到 $-f(x^*)^T f(x^*) = -\| f(x^*) \|^2 \geqslant 0$。

这意味着 $f(x^*) = 0$，这与 $\| f(x_k) \| \geqslant \tau$ 矛盾，故假设不成立，原命题成立，即证。

下面的定理表明 CTTHS 无导数型共轭梯度算法具有全局收敛性。

定理 2 若假设 A 成立，$\{x_k\}$ 由 CTTHS 无导数型共轭梯度法产生，且 $\| d_{k-1}^T y_{k-1} \| \geqslant b\varepsilon$，则有

$\liminf\limits_{k\to\infty} \| f(x_k) \| = 0$。

证明 证明方法类似于定理 6。假设结论不成立，则存在常数 $\tilde{\tau} > 0$，使得 $\| f(x_k) \| \geqslant \tilde{\tau}$。

由于 $f(x_k) = F'(x_k)^T F(x_k)$，则存在常数 $\tilde{\tau}_1 > 0$，使得

$\| F(x_k) \| \geqslant \tilde{\tau}_1$。

1)若 $\limsup\limits_{k\to\infty} \lambda_k > 0$，由引理 3 知，有 $\liminf\limits_{k\to\infty} \| F(x_k) \| = 0$。由上式和引理 2 知 $\lim\limits_{k\to\infty} \| F(x_k) \| = 0$，

这与 $\| F(x_k) \| \geqslant \tilde{\tau}_1$ 矛盾。

2)若 $\limsup\limits_{k\to\infty} \lambda_k = 0$，则由 $\lambda_k \geqslant 0$ 知 $\lim\limits_{k\to\infty} \lambda_k = 0$

由上式和算法知 $\lambda'_k = \frac{\lambda_k}{\lambda}$ 不满足算法步骤 2 的式子，即

$$f(x_k + \lambda'_k d_k) - f(x_k) > -\sigma_1 \| \lambda'_k F(x_k) \|^2 - \sigma_2 \| \lambda'_k d_k \|^2 + \eta_k f(x_k)$$
$$> -\sigma_1 \| \lambda'_k F(x_k) \|^2 - \sigma_2 \| \lambda'_k d_k \|^2$$

这意味着

$$\frac{f(x_k + \lambda'_k d_k) - f(x_k)}{\lambda'_k} \geqslant -\sigma_1 \lambda'_k f(x_k) - \sigma_2 \lambda'_k \| d_k \|^2 \tag{21}$$

另一方面，由中值定理，存在一个常数

$$t_k \in (0,1), \quad \frac{f(x_k + \lambda'_k d_k) - f(x_k)}{\lambda'_k} = f(x_k + t_k \lambda'_k d_k)^T d_k。\tag{22}$$

由引理 2 和引理 3 知 $\lim\limits_{k\to\infty} \| g_k - g_{k-1} \| = 0$，$\lim\limits_{k\to\infty} \| \lambda_{k-1} d_{k-1} \| = 0$，

因而 $\lim\limits_{k\to\infty} y_k = 0$，继而 $\lim\limits_{k\to\infty} \beta_k^{HS} = 0$。

因为 $\| \theta_k \| = \left\| \frac{g_k^T d_{k-1}}{d_{k-1}^T y_{k-1}} \right\| \leqslant \frac{\| g_k^T \| \| d_{k-1} \|}{d_{k-1}^T y_{k-1}} \leqslant \frac{L \| d_{k-1} \|}{b\varepsilon}$，故 $\| \theta_k \|$ 有界。

下面将证明 $\| d_k \|$ 有界。若 $d_k = -g_k$，显然

$\|d_k\|$ 有界。若 $d_k \neq -g_k$，由(10)式 d_k 的定义有，

$$\|d_k\| \leqslant \|g_k\| + 2\frac{\|g_k\|\|y_{k-1}\|}{d_{k-1}^T y_{k-1}}\|d_{k-1}\|$$

$$\leqslant L + 2\frac{L\|g_k - g_{k-1}\|}{b\epsilon}\|d_{k-1}\|$$

$$\leqslant L + 2\frac{LM_1}{b\epsilon}\|d_{k-1}\| \leqslant L + \bar{\eta}\|d_{k-1}\|$$

$$\leqslant L(1+\bar{\eta}+\bar{\eta}^2+\cdots+\bar{\eta}^{k-k_0-1}) + \bar{\eta}^{k-k_0}\|d_{k_0}\|$$

$$\leqslant \frac{L}{1-\bar{\eta}} + \|d_{k_0}\| \leqslant \bar{M},$$

式中：$\bar{\eta} = \frac{2LM_1}{b\epsilon}$，$\bar{M} = \max\{\|d_1\|, \|d_2\|, \cdots, \|d_{k_0}\|, \frac{L}{1-\bar{\eta}} + \|d_{k_0}\|\}$。

故 $\|d_k\|$ 有界。

又因为 $\{x_k\} \subset \Omega$ 有界，不失一般性，设 $x_k \to x^*$，因为 $k \to \infty$ 时，$\beta_k^{HS} \to 0$，$y_k \to 0$，且 $\|\theta_k\|$，$\|d_{k-1}\|$ 有界，因而易有

$$\lim_{k\to\infty} d_k = -\lim_{k\to\infty} g_k + \lim_{k\to\infty}\beta_k^{HS} d_{k-1} - \lim_{k\to\infty}\theta_k y_{k-1} = -f(x^*)。 \tag{23}$$

另一方面，我们有

$$\lim_{k\to\infty} f(x_k + t_k \lambda'_k d_k) = f(x^*), \tag{24}$$

由(21)－(24)式，我们得到

$$-f(x^*)^T f(x^*) = -\|f(x^*)\|^2 \geqslant 0$$

这意味着 $f(x^*) = 0$，与 $\|f(x_k)\| \geqslant \tau$ 矛盾。故假设不成立，原命题成立，即证。

3 数值试验

为考察 MTTHS 无导数型共轭梯度法和 CTTHS 无导数型共轭梯度法的数值表现，用 Matlab 编程如下对称非线性方程组问题进行测试，见表 1。数值结果表明，本文提出的 MTTHS 无导数型共轭梯度法与 CTTHS 无导数型共轭梯度法均具有良好的数值效果，主要表现在搜索方向是下降方向，且一如既往的具有占用内存低、迭代次数少，计算速度快的优势，同时也反映出修正的 MHS 方法的计算表现与近似 PRP 算法 1 很类似(该算法的搜索方向不一定均为下降方向)，是求解大规模对称非线性方程组的有效算法。

测试问题 1：$F(x) = Ax + f(x)$，其 $f(x) = (e^{x_1} - 1, \cdots, e^{x_n} - 1)^T$，矩阵 A 由文献[13]给出。

测试问题 2：$F_1(x) = x_1(x_1^2 + x_2^2) - 1$，$F_i(x) = x_i(x_{i-1}^2 + 2x_i^2 + x_{i+1}^2) - 1, i = 2, 3, \cdots, n-1$，$F_n(x) = x_n(x_{n-1}^2 + x_n^2)$。

选取的参数初始值如下：$\sigma_1 = \sigma_2 = 10^{-4}$，$\lambda_0 = 0.01$，$r = 0.2$，$\epsilon_1 = 10^{-6}$，$t = 5$；$\eta_k = \frac{1}{(k+1)^2}$，初始值 $x_0 = (0.1, 0.1, \cdots 0.1)$，算法的终止准则为 $\|F(x_k)\| \leqslant 10^{-3}$。

数值试验表 1 中的记号意义如下。Dim：测试问题的维数，Iter：算法迭代次数，Time：算法迭代所需时间，$\|F_k\|$：$\|F(x_k)\|$ 的最终值。

4 结束语

针对求解对称非线性方程组问题，本文利用近似梯度代替精确梯度，构造了 MTTHS 无导数型和 CTTHS 无导数型的共轭梯度法。这两个算法的搜索方向均为下降方向，且由于利用近似梯度代替精确梯度，避免了梯度的复杂计算，从而适

表 1 算法的数值结果

		MTTHS 无导数型共轭梯度法			CTTHS 无导数型共轭梯度法			近似 PRP 算法 1[13]		
问题	Dim	Iter	Time	$\|F_k\|$	Iter	Time	$\|F_k\|$	Iter	Time	$\|F_k\|$
P1	10	22	0.0360	8.6993e-04	37	0.1244	9.0636e-04	27	0.0379	9.9566e-04
	50	37	0.0479	9.6175e-04	47	0.0836	9.7129e-04	26	0.0404	8.5321e-04
	100	36	0.1070	9.2375e-04	50	0.1051	9.7589e-04	28	0.0622	5.2109e-04
	500	39	0.8226	9.4633e-04	48	0.7747	9.8409e-04	36	0.8107	8.9261e-04
	1000	40	2.4634	9.0051e-04	51	3.6949	8.8978e-04	31	1.8118	7.9903e-04
	2000	42	7.5006	9.4035e-04	53	10.1751	9.3731e-04	33	6.4984	7.9339e-04
P2	10	43	0.0147	8.3085e-04	114	0.0156	9.6187e-04	42	0.0130	9.7633e-04
	50	51	0.0154	9.8781e-04	117	0.0210	9.9710e-04	45	0.0147	9.5281e-04
	100	46	0.0187	9.4747e-04	117	0.0261	9.9692e-04	43	0.0141	8.9002e-04
	500	54	0.0171	8.7601e-04	118	0.0211	9.8169e-04	43	0.0155	8.5609e-04
	1000	50	0.0198	9.2948e-04	118	0.0317	9.8173e-04	46	0.0199	8.8368e-04
	2000	51	0.0244	9.8693e-04	118	0.0388	9.8187e-04	52	0.0255	9.1798e-04
	5000	51	0.0633	8.9799e-04	119	0.0998	9.3286e-04	54	0.0845	8.9205e-04

用于大规模对称非线性方程组问题。在适当的条件下,证明了这两种算法的全局收敛性,这两种算法的搜索方向均为下降方向,且具有较好的数值表现。

[参 考 文 献]

[1] LI D, FUKUSHIMA M. A globally and superlinearly convergent Gauss-Newton-based BFGS method for symmetric nonlinear equations[J]. SIAM Journal on Numerical Analysis, 1999, 37(1): 152-172.

[2] LI D H, WANG X L. A modefied Fletcher-Reeves-type derivative-free method for symmetric nonlinear equations[J]. Numerical Algebra, Control & Optimization, 2011, 1:71-82.

[3] 李灿.求解对称非线性方程组的 MPRP 型 Derivative-free 算法[J].西南大学学报(自然科学版),2014,36(1):67-71.

[4] 邹志伟.求解对称非线性方程组的一种修正共轭梯度法[D].长沙:湖南大学, 2010:1-20.

[5] 廖昌隆.求解对称非线性方程组的一种 Hesteness-Stiefel 共轭梯度型算法[D].长沙:湖南大学,2011:1-20.

[6] ZHOU W, SHEN D M. An inexact PRP conjugate gradient method for symmetric nonlinear equations. Numerical Functional Analysis and Optimization, 2014, 35: 370-388.

[7] 徐瑞昌.求解对称非线性方程组的混合 BFGS 修正算法[D].长沙:湖南大学,2017:1-34.

[8] 沈冬梅,王国威,胡中波.求解对称非线性方程组的近似 PRP 型共轭梯度法的收敛速度分析[J].湖北工程学院学报,2018,38(6):62-66.

[9] ABUBAKAR A B, KUMAM P, AWWAL A M. An inexact conjugate gradient method for symmetric nonlinear equations[J/OL]. Comp and Math Methods. 2019, 1: e1065. https://doi.org/10.1002/cmm4.1065.

[10] ZHOU W J. A modified BFGS type quasi-Newton method with line search for symmetric nonlinear equations problems[J]. Journal of Computational and Applied Mathematics, 2020, 16: 1-8.

[11] SABI'U J, MUANGCHOO K, SHAH A, et al. A modified PRP-CG type derivative-free algorithm with optimal choices for solving Large-scale nonlinear symmetric equations[J/OL]. Symmetry, 2021, 13(2): 234. https://doi.org/10.3390/sym13020234.

[12] ZHANG L, ZHOU W J, LI D H. Some descent three-term conjugate gradient methods and their global convergence[J]. Optimization Methods & Software, 2007, 22(4):697-711.

[13] 沈冬梅.求解对称非线性方程组 PRP 型算法研究[D].长沙:长沙理工大学,2014:1-39.

Two Modified MHS Derivative-Free Type Conjugate Gradient Methods for Solving Symmetric Nonlinear Equations

Shen Dongmei[1], Yang Zhongxuan[2]

(1. *School of Education, Nanchang Institute of Science & Technology, Nanchang, Jiangxi* 330108, *China*;

2. *School of Science, East China Jiaotong University, Nanchang, Jiangxi* 330013, *China*)

Abstract: This paper constructs two modified MHS derivative-free conjugate gradient methods by using approximate gradients instead of exact gradients for solving the problems with symmetric nonlinear equations. The attractive property of these two algorithms is that the search directions are both descending directions, and it is suitable for solving large-scale symmetric nonlinear equations. By sufficiently utilizing the symmetric structure of the nonlinear equations, it is proved that the algorithms have global convergence based on suitable assumptions. Numerical experiments show that the algorithms are effective for solving large-scale symmetric nonlinear equations.

Key Words: symmetric nonlinear equations; global convergence; modified MHS derivative-free type conjugate gradient method

(责任编辑:熊文涛)