2012 年 8 月 计 算 数 学 第 34 卷第 3 期
Aug., 2012 MATHEMATICA NUMERICA SINICA Vol.34, No.3

一类新型 DL 共轭梯度法研究*1)

邓松海 万 中
(中南大学数学与统计学院, 长沙 410083)

摘 要

提出了求解无约束优化问题的新型 DL 共轭梯度方法. 同已有方法不同之处在于, 该方法构造了一种修正的 Armijo 线搜索规则, 它不仅能给出当前迭代步步长, 而且还能同时确定计算下一步搜索方向时需要用到的共轭参数值. 在较弱的条件下, 建立了算法的全局收敛性理论. 数值试验表明, 新型共轭梯度算法比同类方法具有更好的计算效率.

关键词: 无约束规划; 共轭梯度; 全局收敛; 非精确线性搜索; 下降算法

MR (2000) 主题分类: 90C25, 90C30

1. 引 言

共轭梯度法在求解大规模无约束优化问题上具有优越性 (低运算成本和内存占用少), 因而近年来得到优化界同仁的广泛研究. 最新研究成果参看 [1-11], [15-23] 及其后所列参考文献. 然而, 正如 [4] 指出, 使用该类方法求解无约束优化问题时仍存在许多挑战. 实际上, N. Andrei 在 [4] 中就算法的初始方向, 步长, 基于已知目标函数值来计算的共轭参数值, 线搜索精度的选择等各方面的问题, 共提出了 14 个公开问题.

值得注意的是, 文献 [8] 基于如下推广的共轭条件:

$$d_k^T y_{k-1} = -t g_k^T s_{k-1}, \tag{1.1}$$

提出了共轭参数的新的选取方式:

$$\beta_k^{DL+} = \max\left\{\frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}}, 0\right\} - t\frac{g_k^T s_{k-1}}{d_{k-1}^T y_{k-1}}, \tag{1.2}$$

其中 d_{k-1} 是 $k-1$ 次迭代的搜索方向, g_k 是 k 次迭代目标函数的梯度向量, $y_{k-1} \triangleq g_k - g_{k-1}$, $s_{k-1} \triangleq x_k - x_{k-1}$, $t \geq 0$ 是已知常数, β_k^{DL+} 就是共轭参数.

同已往选取共轭参数的方式不同, 本文首先构造了一种修正的 Armijo 线搜索规则, 该规则不仅能给出当前迭代步步长, 而且还能同时得到计算下一步搜索方向时需要用到的共轭参数值. 因为该参数值计算式与 (1.2) 相似, 所以我们把由此确定搜索方向的方法称之为新型 DL 共轭梯度法, 并据此设计了求解无约束优化问题的算法. 本文首先将证明这类新型共轭梯度算法的适定性, 然后再在较弱的条件下建立了算法的全局收敛性理论. 通过数值试验, 我们把新算法同已有类似方法进行计算效率比较. 结果表明本文提出的新型 DL 共轭梯度算法具有更好的数值表现.

* 2012 年 2 月 22 日收到.
1) 基金项目: 国家自然科学基金资助 (基金号: 71071162, 70921001).

2. 新的共轭参数的选取和相应的算法

无约束优化问题的数学模型如下:

$$\min f(x), x \in R^n, \tag{2.1}$$

其中 $f: R^n \to R$ 是连续可微函数.

用 $g: R^n \to R^n$ 表示 f 的梯度函数. 设 x_0 是 (2.1) 的任意初始近似解. 则求解问题 (2.1) 的标准共轭梯度法是指由如下计算过程生成近似解序列 $\{x_k\}$:

$$x_{k+1} = x_k + \alpha_k d_k, \quad k = 0, 1, \ldots, \tag{2.2}$$

其中 α_k 由某线搜索规则确定的步长, d_k 是由下式确定的搜索方向:

$$d_k = \begin{cases} -g_k, & \text{如果} \;\; k = 0, \\ -g_k + \beta_k d_{k-1}, & \text{如果} \;\; k > 0. \end{cases} \tag{2.3}$$

(2.3) 式中 β_k 称作共轭参数, 其选取方式既影响算法收敛性分析, 也直接决定了算法的数值效率, 因而具有重要的研究意义. β_k 的经典选取方式可参看 [14].

为了使共轭梯度方法具备谱方法 (如拟牛顿法) 的特征, 文献 [8] 利用割线方程导出了如下确定 β_k 的方法 (称为 DL 方法):

$$\beta_k^{DL} = \frac{g_k^T(y_{k-1} - t s_{k-1})}{d_{k-1}^T y_{k-1}}. \tag{2.4}$$

便于能够就一般光滑极小化问题建立此类方法的全局收敛性, 文献 [8] 进一步提出用 (1.2) 式中的 β_k^{DL+} 代替 (2.4) 式中的 β_k^{DL} 计算共轭参数值.

不难看出, 当 $t = \dfrac{2\|y_k\|^2}{s_k^T y_k}$ 时, (2.4) 式可写成

$$\beta_k = \frac{1}{d_k^T y_k} \left(y_k - 2d_k \frac{\|y_k\|^2}{d_k^T y_k} \right)^T g_{k+1}. \tag{2.5}$$

它实际上就是 [12] 中给出的共轭参数的选取公式. 最近, 文献 [1] 中构造了如下一类共轭梯度方法:

$$d_k = -g_k + \lambda_k \left(I - \frac{g_k g_k^T}{\|g_k\|^2} \right) \overline{d_k} \tag{2.6}$$

其中 λ_k 是给定标量, I 是 n 阶单位矩阵, 而 $\overline{d_k}$ 是满足一定条件的已知方向. 该文进一步指出, 其他许多共轭梯度方法 (如 [23]) 都可以看成它的特殊情形.

受 (2.6) 启发, 我们令 $\overline{d_k} = d_{k-1}$, 再将 (2.6) 式两边左乘 y_{k-1}^T, 即

$$y_{k-1}^T d_k = -y_{k-1}^T g_k + \lambda_k y_{k-1}^T \left(I - \frac{g_k g_k^T}{\|g_k\|^2} \right) d_{k-1}.$$

从而得到 λ_k 的如下选取方式:

$$
\begin{aligned}
\lambda_k &= \frac{y_{k-1}^T g_k + y_{k-1}^T d_k}{y_{k-1}^T\left(I-\dfrac{g_k g_k^T}{\|g_k\|^2}\right)d_{k-1}} \\
&= \frac{g_k^T y_{k-1} + d_k^T y_{k-1}}{\left(\left(I-\dfrac{g_k g_k^T}{\|g_k\|^2}\right)d_{k-1}\right)^T y_{k-1}} \\
&= \frac{g_k^T y_{k-1} + d_k^T y_{k-1}}{d_{k-1}^T\left(I-\dfrac{g_k g_k^T}{\|g_k\|^2}\right)y_{k-1}}. \qquad (2.7)
\end{aligned}
$$

在 (2.7) 中引入参数 t, 把它写成:

$$
\lambda_k = \frac{g_k^T(y_{k-1}-ts_{k-1}) + tg_k^T s_{k-1} + d_k^T y_{k-1}}{d_{k-1}^T\left(I-\dfrac{g_k g_k^T}{\|g_k\|^2}\right)y_{k-1}}. \qquad (2.8)
$$

利用 [8] 中的新的共轭条件 (1.1), 并令 $t=1$, 则得到本文选取共轭参数的方法:

$$
\beta_k^N = \frac{g_k^T(y_{k-1}-s_{k-1})}{d_{k-1}^T\left(I-\dfrac{g_k g_k^T}{\|g_k\|^2}\right)y_{k-1}}. \qquad (2.9)
$$

以 β_k^N 为基础, 我们将提出了新型 DL 共轭梯度方法.

接下来, 我们构造一种新型线搜索规则, 该规则不但能给出当前迭代步步长, 而且同时确定下一步搜索方向时需要用到的共轭参数值 (2.9). 就是说, 新的线搜索规则在确定当前迭代点处的步长的同时也给出了下一步的搜索方向. 该线搜索规则如下: 给定 $0<\delta<1$, $\eta>0$, $0<c<1$, 选取步长 $\alpha_k \in \max\left\{\rho^j, j=0,1,2,\cdots,\right\}$, 使得不等式

$$
f(x_k+\alpha_k d_k) \le f_k + \delta\alpha_k g_k^T d_k \qquad (2.10)
$$

和

$$
g(x_k+\alpha_k d_k)^T d(x_k+\alpha_k d_k) \le -c\|g(x_k+\alpha_k d_k)\|^2 \qquad (2.11)
$$

成立, 其中 $d(x_k+\alpha_k d_k) = -g(x_k+\alpha_k d_k) + \tilde{\beta}_{k+1}d_k$. 当条件

$$
d_k^T\left(I-\frac{g(x_k+\alpha_k d_k)g(x_k+\alpha_k d_k)^T}{\|g(x_k+\alpha_k d_k)\|^2}\right)y_k > \eta\|g_k\|^2
$$

成立时,

$$
\tilde{\beta}_{k+1} = \frac{g(x_k+\alpha_k d_k)^T(y_k-s_k)}{d_k^T\left(I-\dfrac{g(x_k+\alpha_k d_k)g(x_k+\alpha_k d_k)^T}{\|g(x_k+\alpha_k d_k)\|^2}\right)y_k}. \qquad (2.12)
$$

否则, $\tilde{\beta}_{k+1}=0$.

我们先证明上述线搜索规则是适定的.

定理 1. 设 $f: R^n \to R$ 是连续可微的函数, 其梯度函数 g 李普希兹连续. 又设 d_k, g_k 分别是 f 在第 k 次迭代 x_k 处的下降方向和梯度. 则必存在非负整数 j_0, $\alpha=\rho^{j_0}$ 使得 (2.10) 式和 (2.11) 式成立.

证明. 事实上, 令

$$t_k=\frac{(1-c)\eta\|g_k\|^2}{(L+1)\|d_k\|^2},$$

其中 L 是李普希兹常数. 当 $\alpha<t_k$ 时, 若

$$d_k^T\left(I-\frac{g(x_k+\alpha d_k)g(x_k+\alpha d_k)^T}{\|g(x_k+\alpha_k d_k\|^2}\right)y_k\le\eta\|g_k\|^2,$$

则 $\tilde{\beta}_{k+1}=0$, $g(x_k+\alpha d_k)^Td(x_k+\alpha d_k)=-\|g(x_k+\alpha d_k)\|^2<-c\|g(x_k+\alpha d_k)\|^2$. 若

$$d_k^T\left(I-\frac{g(x_k+\alpha d_k)g(x_k+\alpha d_k)^T}{\|g(x_k+\alpha_k d_k\|^2}\right)y_k>\eta\|g_k\|^2,$$

则

$$\begin{aligned}
&g(x_k+\alpha d_k)^Td(x_k+\alpha d_k)\\
=\;&-\|g(x_k+\alpha d_k)\|^2+\frac{g(x_k+\alpha d_k)^T(y_k-s_k)}{d_k^T\left(I-\dfrac{g(x_k+\alpha d_k)g(x_k+\alpha d_k)^T}{\|g(x_k+\alpha d_k)\|^2}\right)y_k}g(x_k+\alpha d_k)^Td_k\\
\le\;&-\|g(x_k+\alpha d_k)\|^2+\frac{|g(x_k+\alpha d_k)^T(y_k-s_k)|}{d_k^T\left(I-\dfrac{g(x_k+\alpha d_k)g(x_k+\alpha d_k)^T}{\|g(x_k+\alpha d_k)\|^2}\right)y_k}|g(x_k+\alpha d_k)^Td_k|\\
<\;&-\|g(x_k+\alpha d_k)\|^2+\frac{\|g(x_k+\alpha d_k)\|^2(L+1)\alpha\|d_k\|^2}{\eta\|g_k\|^2}\\
<\;&-c\|g(x_k+\alpha d_k)\|^2.
\end{aligned}$$

故只要 $\alpha<t_k$, 就有 $g(x_k+\alpha d_k)^Td(x_k+\alpha d_k)<-c\|g(x_k+\alpha d_k)\|^2$.

下证满足 (2.11) 的 α 总有一个满足 (2.10). 用反证法. 假设 $\forall\alpha<t_k$, 有 $f(x_k+\alpha d_k)>f(x_k)+\delta\alpha g_k^Td_k$ 成立. 根据中值定理可得:

$$g(x_k+\theta\alpha d_k)^T\alpha d_k>\delta\alpha g_k^Td_k,\quad 0<\theta<1,$$

即

$$g(x_k+\theta\alpha d_k)^Td_k>\delta g_k^Td_k.$$

令 $\alpha\to 0$, 得 $g_k^Td_k>\delta g_k^Td_k$, 即 $(1-\delta)g_k^Td_k>0$, 从而知 $g_k^Td_k>0$. 这与题设 d_k 是下降方向矛盾.

若记

$$\beta_k = \frac{g_k^T(y_{k-1} - s_{k-1})}{d_{k-1}^T H_k y_{k-1}}, \tag{2.13}$$

其中

$$H_k = I - \frac{g_k g_k^T}{\|g_k\|^2}$$

是 R^n 上对超平面 $g_k^T x = b$ (b 是任意给定常数) 的投影算子. 进一步令 $\overline{y_{k-1}} = H_k y_{k-1}$, 则 β_k 可写成:

$$\beta_k = \frac{g_k^T(y_{k-1} - s_{k-1})}{d_{k-1}^T \overline{y_{k-1}}}. \tag{2.14}$$

因此, 基于新的线搜索规则, 本文选取共轭参数的方式为：

$$\tilde{\beta}_{k+1} = \begin{cases} \beta_{k+1}, & \text{若 } d_k^T H_{k+1} y_k > \eta \|g_k\|^2, \\ 0, & \text{否则,} \end{cases} \tag{2.15}$$

从而给出了下一步的搜索方向:

$$d_{k+1} = -g(x_{k+1}) + \tilde{\beta}_{k+1} d_k, \quad k \geq 0. \tag{2.16}$$

注 1. 由于 $d_0 = -g(x_0)$ 是充分下降方向, 根据定理 1 的证明过程, 采用数学归纳法可知: $\forall k \geq 0$, d_k 是充分下降方向.

注 2. 不难证明, 投影算子 $H_k = I - \dfrac{g_k g_k^T}{\|g_k\|^2}$ 是对称半正定的, 且 $\|H_k\| = 1$. 因为 H_k 在 (2.14) 中出现在分母位置, 为避免分母为零, 我们采用公式 (2.15) 和 (2.16) 确定搜索方向.

注 3. 因为 (2.14) 在形式上与 β^{DL} 相似, 故我们把本文提出的方法称作新型 DL 共轭梯度方法. 事实上, 当采用精确线搜索时有条件 $g_k^T d_{k-1} = 0$ 成立, 从而

$$d_{k-1}^T \overline{y_{k-1}} = d_{k-1}^T \left(I - \frac{g_k g_k^T}{\|g_k\|^2} \right) y_{k-1} = d_{k-1}^T y_{k-1} - \frac{\left(d_{k-1}^T g_k\right)\left(g_k^T y_{k-1}\right)}{\|g_k\|^2} = d_{k-1}^T y_{k-1},$$

即公式 (2.14) 与 (2.4) 相同 ($t = 1$). 因此, 类似于 DL 方法, 本文提出的新型的 DL 共轭梯度方法同样具有拟牛顿方法的特征.

注 4. 鉴于 (2.9) 是依据新的共轭条件 (1.1) 得出, 我们称本文研究的方法为新型 DL 共轭梯度方法.

至此, 求解无约束优化问题 (2.1) 的新型 DL 共轭梯度算法可描述如下:

算法 1. (新型 DL 共轭梯度算法)

步 0. 给定常数 $0 < \delta < 1$, $\eta > 0$, $0 < c < 1$, $\rho \in (0,1)$, $\epsilon > 0$, 选取初始点 $x_0 \in R^n$, 计算 $d_0 = -g(x_0)$. 置 $k := 0$.

步 1. 如果 $\|g_k\| \leq \epsilon$, 则算法终止. 否则, 转入步 2.

步 2. 按新的线搜索规则 (2.10), (2.11) 和 (2.12) 计算步长 α_k 和共轭参数 $\tilde{\beta}_{k+1}$.

步 3. 令 $x_{k+1} := x_k + \alpha_k d_k$, 由 (2.15), (2.16) 计算 d_{k+1}. 置 $k := k + 1$. 转步 1.

注 4. 由定理 1 和注 1 知, 算法 1 是适定的.

3. 全局收敛性

本节中我们研究算法 1 的全局收敛性.

首先作如下一般性假设:

假设 1. 水平集 $\Omega=\{x\in R^n\mid f(x)\le f(x_0)\}$ 有界.

假设 2. 在 Ω 的某个邻域 N 内, f 连续可微, 其梯度函数 g 李普希兹连续, 即存在正常数 $L>0$ 使得

$$\|g(x)-g(y)\|\le L\|x-y\|,\quad \forall x,y\in N. \tag{3.1}$$

因为 $\{f(x_k)\}$ 是下降的, 故由算法 1 生成的序列 $\{\,x_k\,\}$ 包含于由假设 1 确定的有界水平集中, 因而存在收敛子列. 为简单起见, 不失一般性, 设 $\{\,x_k\,\}$ 本身收敛. 另一方面, 由假设 2 知, 存在常数 $\gamma_1>0$ 使得

$$\|g(x)\|\le\gamma_1,\quad \forall x\in\Omega. \tag{3.2}$$

因此, 序列 $\{\,g_k\,\}$ 有界.

下面的引理说明步长 α_k 在每步迭代中不会太小.

引理 1. 设 $\{\alpha_k\}$ 是由算法 1 产生的步长序列. 当假设 2 成立时, 必存在常数 $m>0$, 不等式

$$\alpha_k\ge m\frac{\|g_k\|^2}{\|d_k\|^2} \tag{3.3}$$

恒成立.

证明. 由 α_k 的取法可知, $\alpha_k\rho^{-1}$ 必不能同时满足 (2.10)(2.11), 即

$$f(x_k+\alpha_k\rho^{-1}d_k)>f_k+\delta\alpha_k\rho^{-1}g_k^Td_k, \tag{3.4}$$

和

$$g(x_k+\alpha_k\rho^{-1})^Td(x_k+\alpha_k\rho^{-1})>-c\|g(x_k+\alpha_k\rho^{-1}d_k)\|^2 \tag{3.5}$$

两者中至少一个成立.

若 (3.4) 成立. 由中值定理,

$$g(x_k+\theta\alpha_k\rho^{-1}d_k)^T\alpha_k\rho^{-1}d_k=f(x_k+\alpha_k\rho^{-1}d_k)-f(x_k)>\delta\alpha_k\rho^{-1}g_k^Td_k, 0<\theta<1,$$

则有

$$g(x_k+\theta\alpha_k\rho^{-1}d_k)^Td_k>\delta g_k^Td_k.$$

于是

$$L\alpha_k\rho^{-1}\|d_k\|^2>(g(x_k+\theta\alpha_k\rho^{-1}d_k)-g(x_k))^Td_k>(\delta-1)g_k^Td_k>(1-\delta)c\|g_k\|^2.$$

由此得

$$\alpha_k>\frac{c(1-\delta)\rho\|g_k\|^2}{L\|d_k\|^2}.$$

若 (3.5) 成立, 则由定理 1 的证明可知,

$$\alpha_k\rho^{-1} > t_k = \frac{(1-c)\eta\|g_k\|^2}{(L+1)\|d_k\|^2},$$

于是

$$\alpha_k > \frac{(1-c)\eta\rho\|g_k\|^2}{(L+1)\|d_k\|^2}.$$

记

$$m = \min\left\{\frac{c(1-\delta)\rho}{L}, \frac{(1-c)\eta\rho}{L+1}\right\}.$$

则引理得证.

引理 2. 设 d_k 由 (2.12) 确定. 当假设 1 和 2 成立时, 必有:

$$\sum_{k\geq 0}\frac{\|g_k\|^4}{\|d_k\|^2} < \infty. \tag{3.6}$$

证明. 由 (2.10), 引理 1 和假设 1 知:

$$0 < c\delta m\frac{\|g_k\|^4}{\|d_k\|^2} < c\delta\alpha_k\|g_k\|^2 < -\delta\alpha_k g_k^T d_k \leq f_k - f_{k+1}.$$

因而,

$$0 < c\delta m\sum_{k=0}^{n_1}\frac{\|g_k\|^4}{\|d_k\|^2} < f_0 - f_{n_1} < 2M,$$

其中 $f_0 = f(x_0)$. 令 $n_1 \to \infty$, 即得结论.

引理 3. 设步长和搜索方向 α_k, d_k 由 (2.10)-(2.15) 及 (2.3) 确定. 当假设 1 和 2 成立时, 必有:

$$\|d_{k+1}\| \leq (2-c)\|g_{k+1}\|. \tag{3.7}$$

证明. 由线搜索 (2.10), (2.11) 可知, 如果得到的 $d_{k+1} = -g_{k+1}$, 则结论自然成立. 否则由定理 1 的 t_k, 当 $\alpha_k \leq t_k$ 时, 即

$$\alpha_k \leq \frac{(1-c)\eta\|g_k\|^2}{(L+1)\|d_k\|^2}.$$

有

$$\begin{aligned}
\|d_{k+1}\| &< \|g_{k+1}\| + \frac{|g_{k+1}^T(y_k - s_k)|}{d_k^T(I - \frac{g_{k+1}g_{k+1}^T}{\|g_{k+1}\|^2})y_k}\|d_k\| \\
&\leq \|g_{k+1}\| + \frac{\|g_{k+1}\|(L+1)\alpha_k\|d_k\|^2}{\eta\|g_k\|^2} \\
&\leq (2-c)\|g_{k+1}\|.
\end{aligned}$$

定理 2. 若假设 1 和 2 成立, 则

$$\lim_{k\to\infty} inf\|g_k\| = 0. \tag{3.8}$$

证明. 由引理 2 和引理 3 知,

$$\frac{\|g_k\|^2}{(2-c)^2} < \frac{\|g_k\|^4}{\|d_k\|^2} \to 0, k \to +\infty.$$

于是,

$$\lim_{k\to\infty} inf\|g_k\| = 0.$$

4. 数值实验

在本节中, 我们用 [13] 中的 17 个基准测试问题验证算法 1 数值效率, 问题的维数从 2 到 1000. 为说明新型 DL 算法的优点, 我们把它和已有 DL 方法, 典型的共轭梯度算法 (包括 PRP 和 FR 算法) 和谱共轭梯度方法 [17] 进行了比较.

所有代码用 MATLAB 7.0 写成, 并在 PC 2.20 GHz CPU, 760M RAM 及 Windows XP 环境下运行. 相关参数选取如下:

$$\epsilon = 10^{-6}, \quad \rho = 0.75, \quad \delta = 0.4, \quad \eta = 0.01.$$

各算法的数值表现见表 1 和表 2. 表 1 和 2 中的记号的意义是:
DIM: 目标函数的维数;
NI: 迭代次数;
NF: 目标函数计算次数;
NG: 梯度值的计算次数.
this: 本文采取的方法;
prp: 标准 PRP 共轭梯度法;
msprp : 谱共轭梯度法 [17];
fr: Fletcher-Reeves(FR) 法;
dl: Dai-Liao 法 [8](直接引用原数据).

从表 1 和表 2 的运算结果可以看出, 新型 DL 共轭梯度算法具有更好的数值表现: 适合求解高维优化问题, 其迭代次数, 或函数值的计算次数, 或梯度值的计算次数在许多标准测试问题中都优于作比较的方法.

结论. 本文提出了一类求解无约束优化问题的新型 DL 共轭梯度法, 建立了该方法的全局收敛性理论. 数值试验表明, 同已有相似方法相比, 新型共轭梯度算法具有更好的计算效率.

表 1 同其他算法的效率比较

Functions	algorithm	DIM	NI	NF	NG
Rosenbrock	this	2	22	500	478
	prp	2	71	1160	71
	msprp	2	32	598	33
	fr	2	74	1428	75
Box three-dimensional	this	3	44	1027	983
	prp	3	321	1840	321
	msprp	3	174	430	175
	fr	3	38	195	39
Powell singular	this	4	61	1020	959
	prp	4	280	3312	280
	msprp	4	136	2007	137
	fr	4	192	3904	193
Biggs EXP6	this	6	200	2463	2463
	prp	6	689	4834	689
	msprp	6	827	3057	828
	fr	6	392	3384	393
Penalty II	this	40	315	5991	5676
	dl	40	1326	4085	1838
	prp	40	441	3927	441
	msprp	40	652	9581	653
	fr	40	227	3198	228
Extended Rosenbrock	this	1000	34	821	821
	dl	1000	23	119	68
	prp	1000	189	3270	189
	msprp	1000	296	6116	297
Penalty I	this	1000	11	233	233
	dl	1000	29	93	69
	prp	1000	21	183	21
	msprp	1000	11	133	12
	fr	1000	13	121	14
trigonometric	this	100	53	471	471
	dl	100	58	106	104
	prp	100	99	102	100
	msprp	100	124	36	124

表 2 同其他算法的效率比较

Functions	algorithm	DIM	NI	NF	NG
Broyden tridiagonal	this	500	33	753	753
	dl	500	34	109	41
	prp	500	101	679	101
	msprp	500	32	472	33
	fr	500	56	767	57
Brown almost-linear	this	100	19	622	622
	prp	100	61	800	61
	msprp	100	28	901	29
	fr	100	13	329	14
Discrete boundary value	this	100	1257	20278	20278
	prp	100	612	4704	612
	msprp	100	1280	3538	1281
	fr	100	8228	99232	8229
linear function-full rank	this	100	6	43	43
	prp	100	29	191	29
	msprp	100	9	18	10
	fr	100	16	22	17
Brown badly scaled	this	2	52	5163	5163
	prp	2	1717	95952	1717
	msprp	2	49	4863	50
	fr	2	-	-	-
Wood	this	4	170	1779	1609
	prp	4	299	3708	299
	msprp	4	188	3742	189
	fr	4	368	7600	369
Helical valley	this	3	47	1065	1065
	prp	3	94	948	94
	msprp	3	31	581	32
	fr	3	57	916	58
Bard	this	3	63	142	87
	prp	3	180	1586	180
	msprp	3	64	1067	65
	fr	3	214	4469	215
Gaussian	this	3	7	112	112
	prp	3	30	150	30
	msprp	3	10	25	11
	fr	3	12	54	13

参 考 文 献

[1] An Xiaomin, Li Donghui, Xiao Yunhai. Sufficient descent directions in unconstrained optimization[J]. Comput. Optim. Appl., 2011, 48: 515-532.

[2] Andrei N. Acceleration of conjugate gradient algorithms for unconstrained optimization[J]. Appl. Math. Comput., 2009, 213: 361-369.

[3] Andrei N. New accelerated conjugate gradient algorithms as a modification of Dai-Yuan's computational scheme for unconstrained optimization[J]. Journal of computational and Applied Mathematics, 2010, 234: 3397-3410.

[4] Andrei N. Open problems in conjugate gradient algorithms for unconstrained optimization[J]. Bulletin of the Malaysian Mathematical Sciences Society, 2011, 34(2): 319-330.

[5] Birgin E, Martínez J M. A spectral conjugate gradient method for unconstrained optimiztion[J]. Appl. Math. Optim., 2001, 43: 117-128.

[6] Dai Yuhong. Convergence of conjugate gradient methods with constant stepsizes[J]. Optimization Methods and Software, 2011, 26(6): 895-909.

[7] Dai Yuhong. A family of hybrid conjugate gradient methods for uncon- strained optimization[J]. Math. Comput., 2003, 72: 1317-1328.

[8] Dai Yuhong, Liao Lizhi. New conjugacy conditions and related nonlinear conjugate gradient methods[J]. Appl. Math. Optim., 2001, 43: 87-101.

[9] Du Shouqiang, Chen Yuanyuan. Global convergence of a modified spectral FR conjugate gradient method[J]. App. Math. Comput., 2008, 202: 766-770.

[10] Gilbert J C, Nocedal J. Global convergence properties of conjugate gradient methods for optimization[J]. SIAM J. Optim., 1992, 2(1): 21-42.

[11] Grippo L, Lucidi S. A globally convergent version of the Polak-Ribiére conjugate gradient method[J]. Math. Prog., 1997, 78: 375-391.

[12] Hager W W, Zhang H. A new conjugate gradient method with guaranteed descent and an efficient line search[J]. SIAM Journal on Optimization, 2005, 16: 170-192

[13] Jorge J Moré, Burton S Garbow, Kenneth E Hillstrom. Testing unconstrained optimization software[J]. ACM Transactions on Mathematical Software, 1981, 7(1): 17-41.

[14] Nocedal J, Wright S J. Numerical Optimization, Second Edition, Springer Verlag, New York, 2006.

[15] Shi Zhengjun, Shen Jie. Convergence of the Polak-Ribiére-Polyak conjugate gradient method[J]. Nonlinear Analysis, 2007, 66: 1428-1441.

[16] Wan Zhong, Hu Chaomin, Yang Zhanlu. A Spectral PRP Conjugate Gradient Methods for Nonconvex Optimization Problem Based on Modified Line Search[J]. Discrete and Continuous Dynamical Systems: Series B, 2011, 16(4): 1157-1169.

[17] Wan Zhong, Yang Zhanlu, Wang Yalin. New spectral PRP conjugate gradient method for unconstrained optimization[J]. Appl. Math. Letter, 2011, 24: 16-22.

[18] Wei Zengxin, Li Guoyin, Qi Liqun. Global convergence of the Polak-Ribiere-Polyak conjugate gradient method with an Armijo-type inexact line search for nonconvex unconstrained optimization problems[J]. Mathematics of Computation, 2008, 77: 2173-2193.

[19] Yu Gaohang, Guan Lutai, Wei Zengxin. Globally convergent Polak-Ribiére-Polyak conjugate gradient methods under a modified Wolfe line search[J]. Appl. Math. Comput., 2009, 215(8): 3082-3090.
[20] Yuan Gonglin, Lu Xiwen, Wei Zengxin. A conjugate gradient method with descent direction for unconstrained optimization[J]. J. Comput. Appl. Math., 2009, 233(2): 519-530.
[21] Yuan Gonglin. Modified nonlinear conjugate gradient methods with sufficient descent property for large-scale optimization problems[J]. Optimization Letters, 2009, 3: 11-21.
[22] Zhang Li, Zhou Weijun, Li Donghui. A descent modified Polak-Ribiére-Polyak conjugate gradient method and its global convergence[J]. IMA J. Numer. Anal., 2006, 26: 629-640.
[23] Zhang Li, Zhou Weijun, Li Donghui. Global convergence of a modified Fletcher-Reeves conjugate gradient method with Armijo-type line search[J]. Numer. Math., 2006, 104: 561-572.

A NEW DL-TYPE CONJUGATE GRADIENT METHOD FOR NONCONVEX UNCONSTRAINED OPTIMIZATION PROBLEMS

Deng Songhai Wan Zhong
(*School of Mathematics and Statistics, Central South University, Changsha 410083, China*)

Abstract

In this paper, a new DL-type conjugate gradient method is proposed for solving nonconvex unconstrained optimization problems. Different from the existent ones, a new modified Armijo-type line search rule is constructed to give both the steplength and the conjugated parameter being used to determine a search direction in the mean time at each iteration. Under weak conditions, the global convergence of the developed algorithm is established. Numerical experiments show the efficiency of the algorithm, particularly in comparison with the similar ones available in the literature.

Keywords: unconstrained optimization; conjugate gradient; global convergence; inexact line search; descent algorithm

2000 Mathematics Subject Classification: 90C25, 90C30