修正セカント条件に基づいた非線形共役勾配法の大域的収束性について

02302850　東京理科大学　＊高野　正博　TAKANO Masahiro
01702330　東京理科大学　　矢部　博　　YABE Hiroshi

1. はじめに

以下のような無制約最適化問題を考える。

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

ただし $f : R^n \to R$ は滑らかな関数で、その勾配ベクトルを g とおく。この問題に対する共役勾配法のアルゴリズムは以下のとおりである。

〔共役勾配法〕

Step1. 初期点 x_0 と初期探索方向 $d_0 = -g_0$ を与える。$k = 0$ とおく。(ただし $g_k = g(x_k)$)

Step2. 直線探索によりステップ幅 α_k を計算して、点 x_k を更新する。

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

Step3. 収束判定をする。

Step4. β_{k+1} を計算して、探索方向 d_k を更新する。

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

Step5. $k \leftarrow k+1$ として Step2 へ戻る。

共役勾配法は行列を保存する必要がないため、大規模な問題を解くのに有効な方法である。(2) の β の選び方によっていろいろな種類の共役勾配法が考えられるが、最もよく知られているのは、Fletcher-Reeves(FR)、Polak-Ribiere(PR)、Hestenes-Stiefel(HS) の方法である。本研究では、Dai and Liao [1] の研究にならって準ニュートン法の考えを取り入れ、さらに関数値の情報を持った β を提案し、その大域的収束性を示す。

2. Dai and Liao の共役勾配法

$s_{k-1} = x_k - x_{k-1}$、$y_{k-1} = g_k - g_{k-1}$ としたとき、準ニュートン法におけるセカント条件は次のように表せる。

$$H_k y_{k-1} = s_{k-1} \tag{3}$$

ただし H_k はヘッセ行列の逆行列 $\nabla^2 f(x_k)^{-1}$ を近似する正定値対称行列である。このとき準ニュートン法の探索方向 d_k は

$$d_k = -H_k g_k$$

と表せるので、

$$d_k^T y_{k-1} = -g_k^T (H_k y_{k-1}) = -g_k^T s_{k-1}$$

が成り立つ。Dai and Liao [1] はセカント条件を考慮して、共役性条件を以下のように定義した。

$$d_k^T y_{k-1} = -t g_k^T s_{k-1} \quad (t \geq 0)$$

ここで $t = 0$ の場合が従来の共役性を表し、$t = 1$ の場合がセカント条件に対応する。この条件を満たすような探索方向 d_k を生成するために、彼らは (2) より次のような β_k を提案した。

$$\beta_k = \frac{g_k^T (y_{k-1} - t s_{k-1})}{d_{k-1}^T y_{k-1}} \tag{4}$$

さらに (4) の修正として

$$\beta_k^+ = \max\left\{\frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}}, 0\right\} - t\frac{g_k^T s_{k-1}}{d_{k-1}^T y_{k-1}} \tag{5}$$

を提案し、(5) を用いた共役勾配法の大域的収束性を示した。

3. 修正セカント条件に基づいた共役勾配法

Zhang and Xu [4] はセカント条件 (3) を拡張して、以下のような修正セカント条件を提案した。

$$H_k \hat{y}_{k-1} = s_{k-1} \tag{6}$$

ただし

$$\begin{cases} \hat{y}_{k-1} &= y_{k-1} + \dfrac{\theta_{k-1}}{s_{k-1}^T u_{k-1}} u_{k-1} \quad (s_{k-1}^T u_{k-1} \neq 0) \\ \theta_{k-1} &= 6(f_{k-1} - f_k) + 3(g_{k-1} + g_k)^T s_{k-1} \end{cases}$$

ここでは、この修正セカント条件に基づいた共役勾配法を生成するために、パラメータ $\rho \geq 0$ を用いて z_k を次のように定義する。

$$z_{k-1} = y_{k-1} + \left(\frac{\rho\theta_{k-1}}{s_{k-1}^T u_{k-1}}\right) u_{k-1}$$

ここで $\rho = 0$ の場合が従来のセカント条件 (3) であり、$\rho = 1$ の場合が修正セカント条件 (6) である。この z_{k-1}

を用いて、新たな共役性条件を Dai and Liao [1] の考え方に基づいて

$$d_k^T z_{k-1} = -t g_k^T s_{k-1} \quad (t \geq 0) \tag{7}$$

と定義する。そして (7) を満たす方向ベクトルを生成するために、(2) より次の β_k を提案する。

$$\beta_k^{new} = \frac{g_k^T (z_{k-1} - t s_{k-1})}{d_{k-1}^T z_{k-1}} \tag{8}$$

$$\beta_k^{new+} = \max\left\{\frac{g_k^T z_{k-1}}{d_{k-1}^T z_{k-1}}, 0\right\} - t\frac{g_k^T s_{k-1}}{d_{k-1}^T z_{k-1}} \tag{9}$$

4. 大域的収束性

まず探索方向 d_k に関して

$$g_k^T d_k < 0 \tag{10}$$

もしくは

$$g_k^T d_k \leq -c\|g_k\|^2 \qquad (c \geq 0) \tag{11}$$

を仮定する。また z_k に関して次の条件を仮定する。

$$\|z_k\| \leq M\|s_k\| \qquad (M > 0) \tag{12}$$

さらに目的関数に関して次のような一般的な仮定をする。

Assumption 4.1

1. $\mathcal{L} = \{x | f(x) \leq f(x_0)\}$ は有界である。
2. $\mathcal{L}$ の近傍 $\mathcal{N}$ において f は連続微分可能でその勾配はリプシッツ連続である。すなわち

   $$\|\nabla f(x) - \nabla f(\bar{x})\| \leq L\|x - \bar{x}\| \quad (\forall x, \bar{x} \in \mathcal{N})$$

   を満たす $L > 0$ が存在する。

(1) のステップ幅 α_k を直線探索で求める際に、次のような強い Wolfe 条件を課す。

$$\begin{aligned} f(x_k + \alpha_k d_k) - f_k &\leq \delta \alpha_k g_k^T d_k & (13) \\ |g(x_k + \alpha_k d_k)^T d_k| &\leq -\sigma g_k^T d_k & (14) \end{aligned}$$

ただし $0 < \delta < \sigma < 1$ である。強い Wolfe 条件の直線探索を用いた如何なる共役勾配法も、以下の結果を得ることが Dai et al. [2] によって示されている。

Lemma 4.2 *Assumption 4.1* を仮定する。共役勾配法 *(1)*、*(2)* において、d_k は *(10)* を満たし、α_k は *(13)*、*(14)* を満たすとする。このとき

$$\sum_{k \geq 0} \frac{1}{\|d_k\|^2} = \infty$$

ならば

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

が成り立つ。

Lemma4.2 を利用すれば、一様凸な目的関数に対して、(8) を用いた共役勾配法の大域的収束性が次の定理で示される。

Theorem 4.3 *Assumption 4.1* を仮定し、f は一様凸関数とする。*(8)* を用いた共役勾配法において、d_k は *(10)* を、α_k は *(13)*,*(14)* を、また z_k は *(12)* をそれぞれ満たすものとする。このとき正数 $\rho^* \geq \frac{1}{3}$ が存在して、$0 \leq \rho \leq \rho^*$ ならば $\lim_{k \to \infty} \inf \|g_k\| = 0$ が成り立つ。

次に (9) の β_k が持つ性質を述べる。これは Gilbert and Nocedal[3] による Property(*) と同様のものである。

Lemma 4.4 *Assumption 4.1* と *(13)*,*(14)* を仮定し、全ての k に対して $\|g_k\| \geq \gamma$ とする。また z_k は *(12)* を満たすものとして、$0 \leq \rho < \frac{1-\sigma}{3(1+\sigma-2\delta)}$ とする。このとき *(11)* が成り立つならば、全ての k に対して

$$|\beta_k| \leq b$$

かつ

$$\|s_{k-1}\| \leq \xi \Longrightarrow |\beta_k| \leq \frac{1}{b}$$

を満たす $b > 1$、$\xi > 0$ が存在する。

一般の目的関数に対して、(9) を用いた共役勾配法の大域的収束性が次の定理で示される。

Theorem 4.5 *Assumption 4.1* を仮定する。*(9)* を用いた共役勾配法において、d_k は *(11)* を、α_k は *(13)*,*(14)* を、また z_k は *(12)* をそれぞれ満たすものとする。このとき $0 \leq \rho < \frac{1-\sigma}{3(1+\sigma-2\delta)}$ ならば、$\lim_{k \to \infty} \inf \|g_k\| = 0$ が成り立つ。

参考文献

[1] Y. H. Dai and L. Z. Liao, New conjugacy conditions and related nonlinear conjugate gradient methods, Appl. Math. Optim, 43 (2001), pp.87-101.

[2] Y. H. Dai, J. Y. Han, G. H. Liu, D. F. Sun, H. X. Yin and Y. Yuan, Convergence properties of nonlinear conjugate gradient methods, SIAM J. Optimization, 10 (1999), pp. 348-358.

[3] J. C. Gilbert and J. Nocedal, Global convergence properties of conjugate gradient methods for optimization, SIAM J. Optimization, 2 (1992), pp. 21-42.

[4] J. Zhang and C. Xu, Properties and numerical performance of quasi-Newton methods with modified quasi-Newton equations, J. of Computational and Applied Mathematics, 137 (2001), pp. 269-278.