修正セカント条件に基づいた準ニュートン法の大域的収束性について

02103870 東京理科大学 *柏 徹 KASHIWA Toru
01702330 東京理科大学 矢部 博 YABE Hiroshi
01012740 東京理科大学 小笠原 英穂 OGASAWARA Hideho

1. はじめに

以下のような無制約最小化問題

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

を解くための準ニュートン法を考える．ただし目的関数 $f : \mathbf{R}^n \to \mathbf{R}$ は十分に滑らかであるとする．

準ニュートン法は，$x_{k+1} = x_k - \alpha_k B_k^{-1} \nabla f(x_k)$ の反復式によって近似解を生成する解法である．ここで B_k は $\nabla^2 f(x_k)$ の近似行列，α_k はステップ幅であり，特に，$B_k = \nabla^2 f(x_k)$ のときニュートン法に相当する．準ニュートン法では，近似行列 B_k を更新して B_{k+1} を作る際にセカント条件が課され，セカント条件を満足する更新公式として BFGS 公式, DFP 公式, Broyden family などいろいろなものが提案されている．その中でもとりわけ BFGS 公式の有効性が認知されている．

他方，Zhang, Deng and Chen [3]，Zhang and Xu [4] は近似行列の精度を高めるために，修正されたセカント条件を提案し，それに基づいた修正 BFGS 公式，修正 DFP 公式の局所的超 1 次収束性を示した．さらに，吉野, 矢部, 小笠原 [2] は Zhang らの研究を拡張して，修正 BFGS 公式，修正 DFP 公式を含むような修正 Broyden family を構築し，その局所的超 1 次収束性を示した．本稿ではその修正 Broyden family の大域的収束性について議論する．

2. 修正セカント条件と修正 Broyden family

Zhang らにより提案された修正セカント条件は以下のとおりである．

$$B_{k+1} s_k = y_k + \frac{\theta_k}{s_k^T u} u, \tag{1}$$

ただし，$s_k = x_{k+1} - x_k$，$y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$，$\theta_k = 6(f(x_k) - f(x_{k+1})) + 3(\nabla f(x_k) + \nabla f(x_{k+1}))^T s_k$ であり，$u \in \mathbf{R}^n$ は $s_k^T u \neq 0$ となる任意のベクトルである．

本稿では，(1) にパラメータ $\rho \geq 0$ を導入した次のセカント条件を考える．

$$B_{k+1} s_k = z_k, \qquad z_k = y_k + \frac{\rho \theta_k}{s_k^T u} u. \tag{2}$$

ここで，$\rho = 1$ とおけば (1) となることに注意する．

修正セカント条件 (2) に基づいた次の Broyden family を修正 Broyden family と呼ぶことにする．

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{z_k z_k^T}{s_k^T z_k} + \phi_k (s_k^T B_k s_k) v_k v_k^T,$$
$$v_k = \frac{z_k}{s_k^T z_k} - \frac{B_k s_k}{s_k^T B_k s_k}. \tag{3}$$

ここで，ϕ_k はスカラーパラメータである．特に，$\rho = 1$ のとき，$\phi_k = 0, \phi_k = 1$ とおいた場合がそれぞれ Zhang らによって扱われた修正 BFGS 公式，修正 DFP 公式である．

3. アルゴリズム

前節で提案した修正 Broyden family を用いた準ニュートン法のアルゴリズムは，以下のとおりである．

Algorithm [QN]

Step 0. 初期点 $x_1 \in \mathbf{R}^n$ と初期行列 $B_1 \in \mathbf{R}^{n \times n}$ を与える．$k := 1$ とおく．

Step 1. 連立 1 次方程式 $B_k d_k = -\nabla f(x_k)$ を解く．

Step 2. 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．

$$x_{k+1} = x_k + \alpha_k d_k. \tag{4}$$

Step 3. 停止条件を満たしているならば，終了する．満たしていなければ，Step 4 へ進む．

Step 4. B_k を修正 Broyden family (3) により更新する．

Step 5. $k := k + 1$ とおき，Step 1 に戻る．

修正セカント条件を用いた準ニュートン法の特徴は，式 (2) に含まれるベクトル u を自由に選べることにある．これによって準ニュートン法の計算効率が高められる可能性がある．ベクトル u の選び方はいろいろと考えられるが，新たに計算するよりも既存のベクトルを再利用する方が実用的であろう．例として，u を s_k,

$\nabla f(x_k)$, $\nabla f(x_{k+1})$, y_k とおくことが Zhang らにより提案されている.

4. 大域的収束性

修正 Broyden family を用いた準ニュートン法の大域的収束性を示すために, 次のような仮定をする. 以下では簡単のために, $g(x) = \nabla f(x)$ とし, $f_k = f(x_k)$, $g_k = g(x_k)$ と表すことにする.

Assumption 4.1

1. 目的関数 f は 2 回連続微分可能である.
2. レベル集合 $\mathcal{D} = \{ x \mid f(x) \le f(x_1) \}$ は凸集合で, 任意の $p \in \mathbf{R}^n$, $x \in \mathcal{D}$ に対し,
$$m\|p\|^2 \le p^T \nabla^2 f(x) p \le M\|p\|^2$$
なる正の定数 m, M が存在する.
3. すべての k に対し, (2) で選ばれたベクトル u は $|s_k^T u| \ge \mu \|s_k\| \|u\|$ を満たす. ただし, $\mu \in (0,1]$ は k に依存しない定数である.
4. (2) のパラメータ ρ は $0 \le \rho < \frac{M}{3(M-m)}$ を満たす. ($m = M$ のときは $0 \le \rho$ とみなす.)

(4) のステップ幅 α_k を直線探索で求める際に, 次のような Wolfe 条件を課す. ただし, $0 < \sigma_1 < \sigma_2 < 1$ とする.

$$f(x_k + \alpha_k d_k) \le f_k + \sigma_1 \alpha_k g_k^T d_k, \qquad (5)$$
$$g(x_k + \alpha_k d_k)^T d_k \ge \sigma_2 g_k^T d_k. \qquad (6)$$

また, 最急降下方向 $-g_k$ と s_k がなす角を ψ_k とする. すなわち,

$$\cos \psi_k = \frac{-g_k^T s_k}{\|g_k\| \|s_k\|}.$$

ここで, 大域的最適解を x^* とし, α_k が (5)–(6) を満たすような反復 (4) を考える. このとき Assumption 4.1 を仮定すると, 以下の Lemma が成り立つ.

Lemma 4.2 以下の 2 つの不等式が成り立つ.

$$c_1 \|g_k\| \cos \psi_k \le \|s_k\| \le c_2 \|g_k\| \cos \psi_k,$$
$$f_{k+1} - f_* \le [1 - \sigma_1 m c_1 \cos^2 \psi_k](f_k - f_*).$$

ただし, c_1, c_2 はある正の定数であり, $f_* = f(x^*)$ である.

Lemma 4.3 ステップ幅 α_k は次の式を満たす.

$$c_1 \frac{s_k^T B_k s_k}{\|s_k\|^2} \le \alpha_k \le c_2 \frac{s_k^T B_k s_k}{\|s_k\|^2}.$$

Lemma 4.4 以下の 4 つの不等式が成り立つ.

$$\begin{aligned} \frac{\|z_k\|^2}{s_k^T z_k} &\le M_1, \\ \frac{s_k^T B_k s_k}{s_k^T z_k} &\le \frac{c_3 \alpha_k}{1 - \sigma_2}, \\ \frac{\|B_k s_k\|^2}{s_k^T B_k s_k} &\ge \frac{\alpha_k}{c_2 \cos^2 \psi_k}, \\ \frac{|z_k^T B_k s_k|}{s_k^T z_k} &\le \frac{\alpha_k M_2}{c_1 m_1 \cos \psi_k}. \end{aligned}$$

ただし, c_3, m_1, M_1, M_2 はある正の定数である.

Lemma 4.5 $\phi_k \in [0,1]$ のとき, すべての $k \ge 1$ に対して

$$\prod_{j=1}^{k} \alpha_j \ge c_4^k$$

を満たすある正の定数 c_4 が存在する.

Lemma 4.2 – Lemma 4.5 を利用すれば, 一様凸な目的関数に対して, 修正 Broyden family を用いた準ニュートン法の大域的収束性が次の定理で示される.

Theorem 4.6 点列 $\{x_k\}$ は Algorithm [QN] によって生成されるものとする. ただし, $\phi_k \in [0, \delta]$ $(0 \le \delta < 1)$ とし, α_k には Wolfe 条件 (5)–(6) を課す. このとき, 任意の正定値対称な初期行列 B_1 に対して, 点列 $\{x_k\}$ は x^* に収束する.

参考文献

[1] R. Byrd, J. Nocedal and Y. Yuan, Global convergence of a class of quasi-Newton methods on convex problems, *SIAM J. on Numerical Analysis*, 24 (1987), pp.1171–1190.

[2] 吉野 雅之, 矢部 博, 小笠原 英穂, 修正セカント条件に基づいた準ニュートン法の局所的超 1 次収束性について, 最適化：モデリングとアルゴリズム *16*, 統計数理研究所共同研究レポート 161, pp.9–19, 2003.

[3] J.Z. Zhang, N.Y. Deng and L.H. Chen, New quasi-Newton equation and related methods for unconstrained optimization, *J. of Optimization Theory and Applications*, 102 (1999), pp.147–167.

[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.