1995年度日本オペレーションズ・リサーチ学会
春季研究発表会

安定な準ニュートン法の収束性について

01702180　システム計画研究所　八巻直一* 　YAMAKI　Naokazu
01702330　東京理科大学工学部　矢部　博　YABE　Hiroshi

1　安定な準ニュートン法

本稿では，次の無制約最小化問題に対する数値解法を考える.

[UMP]　$f: R^n \to R$ のとき, $f(x)$を$x \in R^n$について最小化せよ.

問題[UMP]を解くための準ニュートン法のアルゴリズムのプロトタイプは，以下の通りである.

[NL]

Step 0. 初期点$x_0 \in R^n$と初期行列$B_0 \in R^{n\times n}$を与える. $k=0$とおく.

Step 1. $\|\nabla f(x_k)\|$が非常に小さいならば停止する. さもなければStep 2へいく.

Step 2. 次の線形方程式を解いて，探索方向d_kを作る

$$B_k d = -\nabla f(x_k). \tag{1.1}$$

Step 3. $x_{k+1} = x_k + \alpha_k d_k$とおく. ただし，$\alpha_k$は適当なステップ幅である.

Step 4. 新しい行列B_{k+1}を作り，$k = k+1$とおいて，Step 1へ行く. □

ここで，行列B_kはヘッセ行列$\nabla^2 f(x_k)$自身，もしくはその近似である. アルゴリズムのStep 4のB_kを変えることによって，いろいろな解法が得られる.

準ニュートン法においては，ヘッセ行列の近似行列の正定値対称性の保証が，アルゴリズムの安定性に本質的に望ましい. B_kが正定値対称であれば，探索ベクトルは$f(x)$の減少方向となり，したがって直線探索では$\alpha_k > 0$となる. このような手法のクラスを，直線探索戦略という. 本稿では，正定値性の保証されたB_kを伴う直線探索戦略に基づく準ニュートン法を，安定な準ニュートン法と定義する.

アルゴリズム[NL]において，セカント条件と正定値対称性を保証するB_{k+1}のクラスとして，次のような更新公式が得られる.

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{s_k^T y_k} + (s_k^T B_k s_k) w_k w_k^T, \tag{1.2}$$

ただし，$B_k = L_k^T L_k$, u_kは$s_k^T L_k^T u_k \neq 0$を満たす任意のベクトルで,

$$w_k = \frac{L_k^T u_k}{s_k^T L_k^T u_k} - \frac{B_k s_k}{s_k^T B_k s_k}. \tag{1.3}$$

また，$s_k = x_{k+1} - x_k$，$y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$である.

2　局所的超一次収束性

ここでは，安定な準ニュートン法のクラス (1.2) の，局所的超一次収束性を示す定理を与える．

非負のパラメータ ϕ_k を持つ Broyden 公式族は (1.2) 含まれ，かつ定理の条件を満たすことに注意する．

$D \subset R^n$ を開凸集合とし，$f(x)$ の局所最適解 x_* を含むものとする．$f(x)$ は以下の仮定を満たすものとする．

(A1) 任意の $x \in D$ に対して，次の条件を満たす正定数 ξ と p が存在する．

$$\|\nabla^2 f(x) - \nabla^2 f(x_*)\| \leq \xi \|x - x_*\|^p \tag{2.1}$$

(A2) $\nabla^2 f$ は x_* で正定値対称である．

また，行列 Q に対して $\|Q\|_F$ は Frobenius ノルム，$\|Q\|_{F,M} = \|M^{-1}QM^{-1}\|_F$ とする．

このとき，以下の結果を得る．

［定理］

仮定 (A1) および (A2) が満たされているとし，B_k は (1.2) によって更新されるものとする．このとき，正の定数 η_1 と η_2 があって，u_k が

$$\|(s_k^T B_k s_k)(w_k w_k^T - v_k v_k^T)\|_{F,M} \leq \eta_1 \frac{\|\hat{s}_k - \hat{B}_k \hat{s}_k\|^2}{\|\hat{s}_k\|^2} + \eta_2 \sigma_k^p \tag{2.2}$$

を満足するものとする．点列 $\{x_k\}$ は

$$x_{k+1} = x_k + s_k \quad , \quad B_k s_k = -\nabla f(x_k). \tag{2.3}$$

によって定義されるものとする．

このとき，任意の $\nu \in (0,1)$ に対して，正定数 ε と δ が存在して，

$$\|x_0 - x_*\| \leq \varepsilon, \qquad x_0 \in D$$

および

$$\|B_0 - \nabla^2 f(x_*)\|_{F,M} \leq \delta,$$

であれば，点列 $\{x_k\}$ は x_* に超一次収束する．

ここに，

$$v_k = \frac{y_k}{s_k^T y_k} - \frac{B_k s_k}{s_k^T B_k s_k}$$

および

$$\sigma_k = \max(\|x_{k+1} - x_*\|, \|x_k - x_*\|),$$

さらに

$$M = \nabla^2 f(x_*)^{\frac{1}{2}}, \quad \hat{B}_k = M^{-1} B_k M^{-1}, \quad \hat{s}_k = M s_k \quad , \quad \hat{B}_{k+1} = M^{-1} B_{k+1} M^{-1}$$

である．

参考文献

[1] A. Stachurski, Superlinear convergence of Broyden's bounded θ-class of methods , *Mathematical Programming*, 20 (1981) 196-212.