2001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

局所的エラーバウンド条件のもとでの Inexact Levenberg-Marquardt 法の収束性

京都大学情報学研究科 *檀 寛成 DAN Hiroshige
京都大学情報学研究科 山下 信雄 YAMASHITA Nobuo
京都大学情報学研究科 福島 雅夫 FUKUSHIMA Masao

1 序論

本発表では，非線形方程式

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

の解法を考える．ここで，F は $\Re^n$ から $\Re^m$ への連続微分可能な関数とする．Levenberg-Marquardt 法 (LMM) [2] は，次の線形方程式の解 $\hat{d}^k$ を用いて，$x^{k+1}:=x^k+\hat{d}^k$ として点列 $\{x^k\}$ を生成する手法である．

$$\left(F'\left(x^k\right)^T F'\left(x^k\right)+\mu_k I\right) d=-F'\left(x^k\right)^T F\left(x^k\right) \tag{2}$$

ここで μ_k は正の定数であり，I は単位行列とする．このとき，$F'\left(x^k\right)^T F'\left(x^k\right)+\mu_k I$ は正定値行列になるので，解 $\hat{d}^k$ は必ずただ一つ存在する．

最近，山下と福島 [3] は，解における F' の正則性の代わりに $\|F(x)\|$ が局所的エラーバウンドになるという仮定のもとで，LMM が局所的に二次収束することを示した．ここで，$\|F(x)\|$ が局所的エラーバウンドになるという条件は，解 x^* において $F'(x^*)$ が正則になるという条件よりも緩い条件であることに注意する．ところで，[3] で提案された LMM では (2) を厳密に解くことが必要であった．しかし，大規模問題においては各反復で厳密な解を求めるよりも，適当な規範を満たす近似解を求めて次の反復に移る inexact 法がしばしば有効である．そこで，本発表では，(2) の近似解 d^k を用いた Inexact Levenberg-Marquardt 法 (ILMM) を考える．ここで r^k を

$$r^k:=\left(F'\left(x^k\right)^T F'\left(x^k\right)+\mu_k I\right) d^k+F'\left(x^k\right)^T F\left(x^k\right)$$

とする．r^k は近似解 d^k によって生じた残差ベクトルである．

本発表では，[3] と同様の解析を行うことで，$\|F(x)\|$ が解の近傍で局所的エラーバウンドを与えるとき，$\left\|r^k\right\|$ がある条件を満たせば，ILMM によって生成される点列が解集合に超一次収束することを示す．さらに，Armijo のステップサイズルールを組み合わせたアルゴリズムを提案し，そのアルゴリズムが大域的収束性を持つことを示す．

2 局所的収束性

まず，関数 θ^k を次のように定義する．

$$\theta^k(d)=\left\|F'\left(x^k\right)d+F\left(x^k\right)\right\|^2+\mu_k\|d\|^2$$

ここで，$\theta^k(d)$ は狭義凸 2 次関数であるので，制約なし最小化問題

$$\min\ \theta^k(d) \tag{3}$$

の最適性の 1 次の必要十分条件は (2) で与えられる．よって，(2) と (3) は等価な問題となる．この最小化問題の性質を用いて，ILMM の収束性についての解析を行う．

本発表では次のことを仮定する．

仮定 1 次の (i), (ii) を満たすような (1) の解 x^* が存在する．

(i) 以下の不等式が成立する定数 $b_1\in(0,1), c_1\in(0,\infty)$ が存在する．

$$\|F'(y)(x-y)-(F(x)-F(y))\| \le c_1\|x-y\|^2 \quad \forall x,y\in B(x^*,b_1)$$

ただし，$B(x^*,b_1):=\{x \mid \|x-x^*\|\le b_1\}$ である．

(ii) $\|F(x)\|$ は $B(x^*,b_1)$ 上で (1) に対するエラーバウンドとなる．つまり，次の不等式が成立する正の定数 c_2 が存在する．

$$c_2\,\mathrm{dist}(x,X^*)\le\|F(x)\| \quad \forall x\in B(x^*,b_1)$$

ここで，X^* は (1) の解集合を表す．

本発表においては仮定 1 (ii) が最も本質的な役割を果たす．また，仮定 1 (i) は，F' がリプシッツ連続であれば成立する．

本発表では μ_k を次のように更新する．

仮定 2

$$\mu_k=\left\|F(x^k)\right\|^\delta$$

ただし，δ は $0<\delta\le 2$ となる定数である．

これらの仮定のもと，次の定理が成立する．

定理 1 仮定 1, 2 が成立し，初期点 x^0 が x^* の十分近くにあるものとする．さらに，$\{x^k\}$ を ILMM で生成された点列とする．このとき，各反復 k において，

$$\|r^k\| / \mu_k = o\left(\mathrm{dist}\left(x^k, X\right)\right)$$

が満たされていれば，$\{\mathrm{dist}\left(x^k, X^*\right)\}$ は 0 に超一次収束する．さらに，点列 $\{x^k\}$ はある解 $\hat{x}$ に収束する．特に，

$$\|r^k\| / \mu_k = O\left(\mathrm{dist}\left(x^k, X\right)^2\right)$$

が満たされていれば，$\{\mathrm{dist}\left(x^k, X^*\right)\}$ は 0 に二次収束する． □

3 大域的収束するアルゴリズム

本節では，ILMM に Armijo のステップサイズルールを組み合わせたアルゴリズムを提案し，そのアルゴリズムが大域的収束することを示す．本節で提案するアルゴリズムは以下の通りである．ただし，メリット関数として $\phi(x) = \frac{1}{2}\|F(x)\|^2$ を用いる．

アルゴリズム 1

Step 0: パラメータ $\alpha \in (0,1), \beta \in (0,1), \gamma \in (0,1), \rho \in (0,1), \delta \in (0,2], p > 0$ と初期点 x^0 を決める．$\mu_0 = \|F(x^0)\|^\delta$ とする．$k := 0$ とする．

Step 1: x^k が終了条件を満たしていたら計算を終了．

Step 2: 線形方程式 (2) の近似解 d^k を求める．もし，

$$\|F\left(x^k + d^k\right)\| \le \gamma \|F\left(x^k\right)\|$$

ならば，$x^{k+1} := x^k + d^k$ として Step 4 へ．

Step 3: もし

$$\nabla\phi\left(x^k\right)^T d^k \le -\rho \|d^k\|^p \tag{4}$$

でなければ $d^k = -\nabla\phi\left(x^k\right)$ とする．次の不等式を満たす最小の非負の整数 m を求め，$x^{k+1} := x^k + \beta^m d^k$ とする．

$$\phi\left(x^k + \beta^m d^k\right) - \phi\left(x^k\right) \le \alpha\beta^m \nabla\phi\left(x^k\right)^T d^k$$

Step 4: $\mu_{k+1} = \|F(x^{k+1})\|^\delta$ とする．$k := k+1$ として Step 1 へ． □

アルゴリズム 1 の Step 3 においては，探索方向 d^k が (4) を満たしているかどうかを調べている．これは，d^k として，(2) の厳密解ではなく近似解を用いているために，d^k がメリット関数 ϕ の必ずしもよい降下方向にならない可能性があるためである．そのような場合には，$d^k = -\nabla\phi\left(x^k\right)$ とすることで，d^k がメリット関数の降下方向になり，Step 3 の直線探索が well defined となることを保証している．

アルゴリズム 1 に対して，次の定理が成立する．なお，証明には [1] と同様の手法を用いている．

定理 2 $\{x^k\}$ を アルゴリズム 1 で生成された点列とし，探索方向 d^k は，r^k が

$$\|r^k\| \le \min\left\{\eta \left\|F'\left(x^k\right)^T F\left(x^k\right)\right\|, \nu_k \left\|F'\left(x^k\right)^T F\left(x^k\right)\right\|^\delta\right\} \tag{5}$$

を満たすように定められているものとする．ただし，$\eta \in (0,1), \nu_k = o\left(\mathrm{dist}\left(x^k, X^*\right)\right)$ であるものとする．このとき，$\{x^k\}$ の任意の集積点は ϕ の停留点となる．さらに，$\{x^k\}$ の集積点が (1) の解を少なくとも一つ含むとする．その解 x^* において仮定 1 がみたされていれば，$\{\mathrm{dist}\left(x^k, X^*\right)\}$ は 0 に超一次収束する．特に，$\delta = 2$ かつ $\nu_k = O\left(\mathrm{dist}\left(x^k, X^*\right)^2\right)$ であれば，$\{\mathrm{dist}\left(x^k, X^*\right)\}$ は 0 に二次収束する． □

ここで，近似基準 (5) における ν_k は

$$\nu_k = \|F\left(x^k\right)\|^\tau$$

とすることができる．ただし，τ は $\tau > 1$ となる定数である．このとき，十分大きい k に対して $\nu_k = o\left(\mathrm{dist}\left(x^k, X^*\right)\right)$ となる．特に，$\tau \ge 2$ とすれば，十分大きい k に対して $\nu_k = O\left(\mathrm{dist}\left(x^k, X^*\right)^2\right)$ となる．

4 まとめ

本発表では，局所的エラーバウンド性が成立するときの ILMM の収束性について考察し，探索方向を計算するときの誤差が十分小さければ，解集合に超一次収束することを示した．このことは，特に大規模な問題を解くときに大きな利点となる．なお，数値実験の結果を当日に発表する予定である．

参考文献

[1] F. Facchinei and C. Kanzow, A nonsmooth inexact Newton Method for the solution of large-scale nonlinear complementarity problems, *Mathematical Programming* 76 (1997) 493-512.

[2] R. Fletcher, *Practical Methods of Optimization*, second ed., John Wiley & Sons, New York, (1987).

[3] N. Yamashita and M. Fukushima, On the rate of convergence of the Levenberg-Marquardt method, Technical Report 2000-008, Department of Applied Mathematics and Physics, Kyoto University (November 2000).