A Regularized Newton Method without Line Search for Unconstrained Optimization

上田　健詞

（京都大学大学院情報学研究科数理工学専攻　現所属・三菱電機㈱）

指導教員　山下信雄 准教授

1. はじめに

本研究では，数理計画の基本的問題の一つである以下の制約無し最小化問題を考える．

$$\underset{x\in\mathbb{R}^n}{\text{minimize}}\ f(x) \tag{1}$$

ここで，目的関数 f は $\mathbb{R}^n$ から $\mathbb{R}$ への2階連続的微分可能な関数とする．

制約無し最小化問題に対する効率的な解法として，f のヘッセ行列を利用したニュートン型手法がある．その中でも，信頼領域法によって大域的収束性を保証したニュートン法（以下，信頼領域法と略す）[1]は理論的に良い性質を持ち，かつ数値的に安定しているため，多くの数理計画ソルバーに組み込まれている．信頼領域法では，現在の反復点 x_k と信頼半径 Δ_k が与えられたとき，以下の非凸な部分問題の解 $\bar{d}_k(\Delta_k)$ を用いて，次の反復点を $x_{k+1}=x_k+\bar{d}_k(\Delta_k)$ とする．

$$\begin{aligned}&\underset{d\in\mathbb{R}^n}{\text{minimize}}\ f(x_k)+\nabla f(x_k)^T d+\frac{1}{2}d^T\nabla^2 f(x_k)d\\ &\text{subject to}\quad \|d\|\leq\Delta_k\end{aligned} \tag{2}$$

この部分問題(2)を解くためにいろいろな手法が提案されているが，ヘッセ行列 $\nabla^2 f(x_k)$ によっては解を得るために多くの時間がかかることが知られている．そのため，信頼領域法と同じような振る舞いをし，かつ部分問題が容易に解けるアルゴリズムが求められている．

本研究では，そのようなアルゴリズムとして，直線探索を用いない正則化ニュートン法を提案する．提案アルゴリズムでは，大域的収束性を保証するため，各反復でステップサイズ（直線探索）の代わりに正則化パラメータを調整する．

2. 提案アルゴリズム

現在の反復点 x_k と正のパラメータ ν に対して，2次モデル関数 $m_k(d,\nu)$ を

$$m_k(d,\nu):=f(x_k)+\nabla f(x_k)^T d+\frac{1}{2}d^T(\nabla^2 f(x_k)+\mu_k(\nu)I)d$$

と定義する．ただし，

$$\mu_k(\nu):=c\max(0,-\lambda_{\min}(\nabla^2 f(x_k)))+\nu\|\nabla f(x_k)\|^{\delta}$$

であり，c, δ は $c>1$, $\delta\geq 0$ を満たす定数，$\lambda_{\min}(\nabla^2 f(x_k))$ は $\nabla^2 f(x_k)$ の最小固有値である．ベクトル $d_k(\nu)$ を以下の部分問題の解とする．

$$\underset{d\in\mathbb{R}^n}{\text{minimize}}\ m_k(d,\nu) \tag{3}$$

$\|\nabla f(x_k)\|\neq 0$ のとき，正則化パラメータ $\mu_k(\nu)$ のおかげで $\nabla^2 f(x_k)+\mu_k(\nu)I$ は正定値となる．よって，部分問題(3)は制約無しの凸2次計画問題となるため，$d_k(\nu)=-(\nabla^2 f(x_k)+\mu_k(\nu)I)^{-1}\nabla f(x_k)$ となる．信頼領域法では探索方向 $\bar{d}_k(\Delta_k)$ を求めるために非凸の部分問題(2)を解かなければならなかったが，$d_k(\nu)$ は線形方程式を1回解くだけで求まる．

既存の正則化ニュートン法では，ある固定された $\bar{\nu}$ に対する探索方向 $d_k(\bar{\nu})$ を用いている．そして，ステップサイズ α をうまく調整することにより，次の反復点 $x_{k+1}=x_k+\alpha d_k(\bar{\nu})$ が $f(x_{k+1})<f(x_k)$ となるようにする．本研究では，ステップサイズは $\alpha=1$ と固定したまま（つまり直線探索は用いないで），ν をうまく調整することにより，目的関数値を減少させる反復点を $x_{k+1}=x_k+d_k(\nu)$ と生成する正則化ニュートン法を提案する．

提案アルゴリズムでは，適切な ν を見つけるために信頼領域法のアイデアを用いる．モデル関数値の減少量に対する目的関数値の減少量の割合 $\rho_k(\nu)$ を

$$\rho_k(\nu):=\frac{f(x_k)-f(x_k+d_k(\nu))}{f(x_k)-m_k(d_k(\nu),\nu)}$$

と定義する．$\rho_k(\nu)$ が1に近いとき，すなわち，2次モデル関数 m_k による f の近似がうまくいっているとき，探索方向 $d_k(\nu)$ を採用し，パラメータ ν の値を小さくする．そうでないときにはパラメータ ν の値を大きくし，探索方向 $d_k(\nu)$ を求めなおす．提案アルゴリズムの詳細については本稿では割愛する．

$\nabla^2 f(x_k)$ が正定値であるとき，任意の $\nu \in (0, \infty)$ に対して $d_k(\nu)=\bar{d}_k(\Delta)$ を満たす Δ が存在するため，提案アルゴリズムは（特に解のそばで）信頼領域法と同じような振る舞いをすると期待できる．

3. 提案アルゴリズムの収束性

提案アルゴリズムの大域的収束性，超一次収束性は以下の定理で与えられる．

定理 1．（大域的収束性） 提案アルゴリズムによって生成された点列 $\{x_k\}$ は有界であるとする．このとき，$\{x_k\}$ は f の停留点へ収束する．

定理 2．（超一次収束性） 以下の(a)-(d)が成り立つとする．(a)$0<\delta<1$．(b)問題(1)の局所的最適解 x^* が存在する．(c)x^* の十分近くで $\nabla^2 f$ は Lipschitz 連続である．(d)$\nabla^2 f(x^*)$ は正定値である．このとき，初期点 x_0 を x^* の十分近くに選べば，$\{x_k\}$ は x^* に $1+\delta$ 次収束する．

さらに，提案アルゴリズムでは，初期点から停留点を求めるまでに必要な反復回数の上限を見積もることができる．このような見積もりは信頼領域法では与えられていない．

定理 3．（反復回数の上限） 以下の(a)-(c)が成り立つとする．(a)$\delta \leq \frac{1}{2}$．(b)提案アルゴリズムによって生成された点列 $\{x_k\}$ は有界である．(c)$\nabla^2 f$ は Lipschitz 連続である．このとき，与えられた正の定数 ϵ に対して，$\|\nabla f(x_J)\| \leq \epsilon$ となる最初の反復 J は $J=O(\epsilon^{-2})$ となる．

4. 数値実験

提案アルゴリズムと信頼領域法を関数評価回数と線形方程式を解いた回数の観点から比較する．これらの評価尺度を用いる理由は，ニュートン型手法の計算時間の大部分は目的関数の評価と線形方程式の求解で占められているためである．テスト問題は非線形最適化の分野でよく用いられているテスト問題集 CUTEr の問題を 121 問使用した．評価方法として，文献[2]で提案されている分布関数を用いる．アルゴリズム s で問題 p を解いたときの評価尺度（関数評価回数あるいは線形方程式を解いた回数）を $t_{p,s}$ とする．アルゴリズム s_1 と s_2 を比較する場合，s_1 に対する分布関数 $F_{s_1}(\tau)$（ただし，$\tau \geq 1$）は次式で定義される．

$$F_{s_1}(\tau) := \frac{t_{p,s_1} \leq \tau t_{p,s_2}\text{ が成り立つテスト問題 } p \text{ の総数}}{\text{テスト問題の総数}}$$

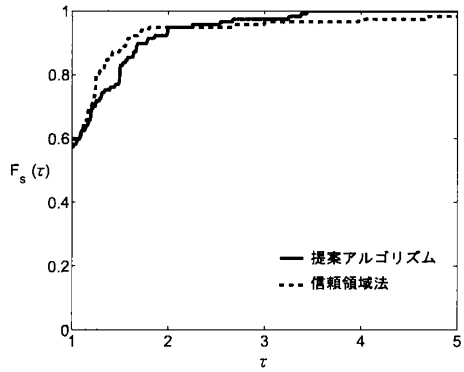

図 1　関数評価回数の比較

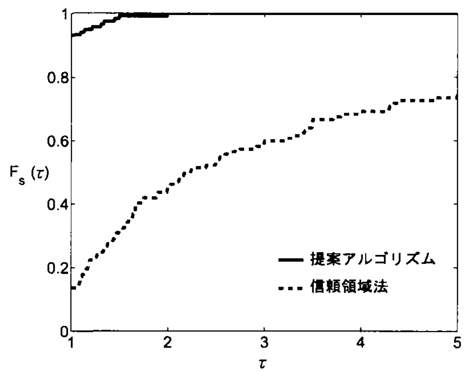

図 2　線形方程式を解いた回数の比較

s_2 に対する分布関数も同様に定義される．他のアルゴリズムの分布関数値に比べて分布関数値が 1 に近いアルゴリズムは多くの問題に対して良い振る舞いをしていると考えられる．

関数評価回数に対する比較を図 1，線形方程式を解いた回数に対する比較を図 2 に示す．

図 1 より，関数評価回数の点では提案アルゴリズムは信頼領域法と同程度であることがわかる．提案アルゴリズムや信頼領域法では関数評価回数を反復回数とみなすことができるので，この結果は反復回数が同程度であることを意味している．一方，図 2 より，線形方程式を解いた回数は提案アルゴリズムの方が信頼領域法よりもはるかに少ない．この理由として，信頼領域法では各反復で非凸な部分問題(2)を解く際に何回も線形方程式を解かなければならないことが挙げられる．以上より，提案アルゴリズムの部分問題(3)は信頼領域法の部分問題(2)よりも簡単に解くことができ，その結果，全体の計算時間も少なくなることがわかる．

参考文献

[1] A. R. Conn, N. I. M. Gould and P. L. Toint, *Trust-Region Methods*, SIAM, Philadelphia, 2000.

[2] E. D. Dolan and J. J. Moré, *Benchmarking optimization software with performance profiles*, Mathematical Programming, 91, 201-213 (2002).