数理解析研究所講究録
第1931巻 2015年 67-78

微分不可能な関数を含む非線形方程式系に対する PRP型平滑化スケーリング共役勾配法について

横浜国立大学 成島 康史
Yokohama National University Yasushi Narushima

鈴与シンワート株式会社 大谷 亮介
Suzuyo Shinwart Corp. Ryousuke Ootani

東京理科大学 矢部 博
Tokyo University of Science Hiroshi Yabe

概 要

微分不可能な関数を含む非線形方程式系の求解問題は, 実社会において生じる基本的な問題の一つである. そのような問題に対して平滑化ニュートン法は効果的な反復法としてよく知られているが, 行列を保存する必要があるため大規模な問題を解くことに適していないという欠点がある. 本論文では, 微分不可能な関数を含む方程式系に対して平滑化手法とスケーリング共役勾配法を組み合わせて, 陽に行列を使用しないようなアルゴリズムを提案する. そして, 提案したアルゴリズムの大域的収束性を証明し, 最後に数値計算結果を報告する.

1 はじめに

関数 $F : \boldsymbol{R}^n \to \boldsymbol{R}^n$ を必ずしも微分可能とは限らない連続関数とする. 本論文では, 次の非線形方程式系:

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

に対する数値解法を考える. ただし, (1) は少なくとも1つの解を持つものとする. このような方程式系の求解問題は実社会において生じる基本的な問題の一つで, 例えば相補性問題や変分不等式問題を再定式化した際に表れる [3, 12–14]. 取り扱う方程式に微分不可能な関数が含まれる場合, 広く使われている勾配を用いた方法は使用できなくなる. このため, いくつかの手法が提案されているが, その一つに平滑化手法が挙げられる.

定義 1.1. 関数 $\tilde{F} : \boldsymbol{R} \times \boldsymbol{R}^n \to \boldsymbol{R}^n$ が $\boldsymbol{R}_{++} \times \boldsymbol{R}^n$ において連続的微分可能で, かつすべての $x \in \boldsymbol{R}^n$ に対して

$$\lim_{t \to +0} \tilde{F}(t, x) = F(x)$$

を満たすとき, $\tilde{F}$ を F の平滑化関数と呼ぶ. ただし, $\boldsymbol{R}_{++} = \{t \in \boldsymbol{R} | t > 0\}$ である.

ここで, 上で定義した平滑化関数を用いて関数 $H : \boldsymbol{R} \times \boldsymbol{R}^n \to \boldsymbol{R}^{1+n}$ を

$$H(t,x) = \begin{pmatrix} t \\ \tilde{F}(t,x) \end{pmatrix}$$

とすると, $H(t,x) = 0$ ならば $F(x) = 0$ が成立する. したがって, 最小二乗法の考えに基づいたメリット関数

$$\Psi(t,x) = \frac{1}{2}\|H(t,x)\|^2 = \frac{1}{2}\{t^2 + \|\tilde{F}(t,x)\|^2\} \tag{2}$$

を定義して, 次の無制約最小化問題

$$\min \Psi(t,x) \tag{3}$$

を解くことは, (1) の解を求めることと同値となる. ただし, ベクトルのノルム $\|\cdot\|$ は 2 ノルムである.

ここで注意しておかなければならないことは, $\Psi(t,x)$ は $\boldsymbol{R}_{++} \times \boldsymbol{R}^n$ 上では連続的微分可能だが, それ以外の領域 ($t \leq 0$ となる場合) では連続的微分可能とは限らないことである. 問題 (3) をニュートン法やそれに類似する反復法で解く方法は既に多くの研究者によって研究されており, それらは論文 [13] にまとめられている. しかし, ニュートン法は行列を保存しなければならないため, 大規模問題に適していないという弱点がある. よって本論文では, 行列を陽に使用しないような数値解法である共役勾配法に着目する.

無制約最適化問題に対する共役勾配法は行列を保存する必要がないため, 大規模問題に対する数値解法として近年注目されている. 一般の (微分可能な) 無制約最適化問題:

$$\min \quad f(z), \qquad (\text{ただし}, z \in \boldsymbol{R}^l \text{ とし } f : \boldsymbol{R}^l \to \boldsymbol{R} \text{ は微分可能とする})$$

を解くための共役勾配法は, 初期点 $z_0 \in \boldsymbol{R}^l$ が与えられたとき,

$$z_{k+1} = z_k + \alpha_k d_k,$$

$$d_k = \begin{cases} -\nabla f(z_k) & k = 0, \\ -\nabla f(z_k) + \beta_k d_{k-1} & k \geq 1, \end{cases}$$

によって点列 $\{z_k\}$ を生成する解法である. ただし, $\alpha_k > 0$ はステップ幅, d_k は探索方向とよばれ, β_k は共役勾配法を特徴付けるパラメータである. 共役勾配法は β_k のとりかたによって計算効率が変わるため, 多くの研究者 [2, 4–7, 11] が効果的なパラメータ β_k の選び方について研究している. 一方で, 共役勾配法には必ずしも十分な降下方向を生成するとは限らないという欠点がある. ここで, 探索方向 d_k が十分な降下条件を満たすとは, ある正の定数 $\bar{c}$ が存在し, 全ての $k \geq 0$ に対し,

$$\nabla f(z_k)^T d_k \leq -\bar{c}\|\nabla f(z_k)\|^2$$

を満たすことをいう. この欠点を解消するため, 直線探索の方法によらずに十分な降下方向を生成する三項共役勾配法やスケーリング共役勾配法が近年盛んに研究されている (例えば, [1, 8, 10, 15–17] 参照). そうした研究の一つとして, Narushima, Yabe and Ford [10] はいくつかの三項共役勾配法やスケーリング共役勾配法を含むような族を提案しており,

その探索方向は

$$d_k = \begin{cases} -\nabla f(z_k) & k=0, \\ -\nabla f(z_k) + \beta_k (\nabla f(z_k)^T p_k)^\dagger \{(\nabla f(z_k)^T p_k) d_{k-1} - (\nabla f(z_k)^T d_{k-1}) p_k\} & k \geq 1, \end{cases} \tag{4}$$

で与えられる．ただし, $p_k \in \boldsymbol{R}^l$ は任意のベクトルで, 記号 $\dagger$ は

$$a^\dagger = \begin{cases} \dfrac{1}{a} & a \neq 0, \\ 0 & a = 0, \end{cases}$$

で定義される．このとき, (4) の探索方向 d_k に左から $\nabla f(z_k)^T$ をかけると $\nabla f(z_k)^T d_k = -\|\nabla f(z_k)\|^2$ が得られる．これは探索方向が十分な降下条件を満たすことを意味している．通常, パラメータベクトル p_k として, $p_k = \nabla f(z_k)$, または, $p_k = \widehat{y}_{k-1} \equiv \nabla f(z_k) - \nabla f(z_{k-1})$ が選ばれることが多い．例えば, β_k として, Polak-Ribièreの選択法 [11] $\left(\beta_k^{PR} = \frac{\nabla f(z_k)^T \widehat{y}_{k-1}}{\|\nabla f(z_{k-1})\|^2}\right)$ を適用し, $p_k = \widehat{y}_{k-1}$ とすると, 式 (4) は $\nabla f(z_k)^T \widehat{y}_{k-1} \neq 0$ のとき,

$$d_k = \begin{cases} -\nabla f(z_k) & k=0, \\ -\nabla f(z_k) + \beta_k^{PR} d_{k-1} - \dfrac{\nabla f(z_k)^T d_{k-1}}{\|\nabla f(z_{k-1})\|^2} \widehat{y}_{k-1} & k \geq 1, \end{cases} \tag{5}$$

となる．これは, Zhang ら [16] によって提案された三項共役勾配法と一致する．一方, $p_k = \nabla f(z_k)$ とした場合には, (4) の探索方向 d_k は

$$d_k = \begin{cases} -\nabla f(z_k) & k=0, \\ -\left(1 + \beta_k \dfrac{\nabla f(z_k)^T d_{k-1}}{\|\nabla f(z_k)\|^2}\right) \nabla f(z_k) + \beta_k d_{k-1} & k \geq 1, \end{cases} \tag{6}$$

と表すことができる．例えば, $\beta_k = \beta_k^{PR}$ とすると, Cheng [1] のスケーリング共役勾配法と一致する．

一方, 方程式 (1) を解くために, 平滑化関数を含んだ無制約最適化問題 (3) に対する共役勾配法も研究されており, Narushima [9] は Zhang らの三項共役勾配法 (5) に倣い平滑化三項共役勾配法を提案した．この方法では探索方向はメリット関数 Ψ の降下方向を常に生成している．さらに Narushima は直線探索において, Armijo 条件の変種を用いたアルゴリズムを構築し, その大域的収束性を証明している．

探索方向 (4) において, $p_k = \widehat{y}_{k-1}$ とした場合に, 三項共役勾配法 (5) と一致するためには, $\nabla f(z_k)^T \widehat{y}_{k-1} \neq 0$ という条件が必要だったのに対し, $p_k = \nabla f(z_k)$ とした場合には, こうした条件を課すことなくスケーリング共役勾配法 (6) と一致する．したがって, p_k として $\nabla f(z_k)$ を選択する方が自然であるように思われる．そこで, 今回, 我々はスケーリング共役勾配法 (6) に基づき, 無制約最適化問題 (3) を解くような平滑化スケーリング共役勾配法を提案する．

本論文は以下のように構成される．第 2 節では, スケーリング共役勾配法 (6) に基づき, メリット関数 Ψ の降下方向となるような探索方向を提案する．さらに, 直線探索において Armijo 条件の変種を用いたアルゴリズムを与え, その大域的収束性を示す．第 3 節では数値実験の結果を報告する．

2 平滑化スケーリング共役勾配法とその大域的収束性

まず本節で必要となる事項を導入しておく. 平滑化関数 $\tilde{F}$ は $\boldsymbol{R}_{++} \times \boldsymbol{R}^n$ 上で連続的微分可能なので, $\nabla\tilde{F}$ が存在し,

$$\nabla\tilde{F}(t,x) = \begin{pmatrix} \nabla_t\tilde{F}(t,x) \\ \nabla_x\tilde{F}(t,x) \end{pmatrix}$$

となり, (2) から $\nabla\Psi$ は

$$\nabla\Psi(t,x) = \begin{pmatrix} \nabla_t\Psi(t,x) \\ \nabla_x\Psi(t,x) \end{pmatrix} = \begin{pmatrix} t + \nabla_t\tilde{F}(t,x)\tilde{F}(t,x) \\ \nabla_x\tilde{F}(t,x)\tilde{F}(t,x) \end{pmatrix}$$

となる. ここで, $\nabla_t\tilde{F}(t,x)$ は n 次元行ベクトル, $\nabla_x\tilde{F}(t,x)$ は n 次正方行列となるため, $\nabla_t\tilde{F}(t,x)\tilde{F}(t,x)$ はスカラーとなり, $\nabla_x\tilde{F}(t,x)\tilde{F}(t,x)$ は n 次元列ベクトルとなる. したがって, $\nabla\Psi(t,x)$ は $1+n$ 次元列ベクトルとなる. また, $(t,x) \in \boldsymbol{R} \times \boldsymbol{R}^n$ を $(t,x^T)^T$ の代わりに表記し, $v = (t,x) \in \boldsymbol{R}^{1+n}$ と表す.

以下では, 非線形方程式 (1) に対する平滑化スケーリング共役勾配法のアルゴリズムを提案する. 平滑化スケーリング共役勾配法のアルゴリズムでは, k 回目の反復解を $v_k = (t_k, x_k)$, 探索方向を $d_k \in \boldsymbol{R}^{1+n}$, ステップ幅を $\alpha_k \in (0,1]$ としたとき, $k+1$ 回目の反復解は

$$v_{k+1} = v_k + \alpha_k d_k \tag{9}$$

と表される. ここで, 簡単のために $\tilde{F}(v_k)$ を

$$\tilde{F}_k = \tilde{F}(v_k)$$

と省略し, 他の関数についても同様の省略記号を用いる. 第 1 節でも触れたように Ψ は $\boldsymbol{R}_{++} \times \boldsymbol{R}^n$ 以外の領域では連続的微分可能性が保証されていないので, 全ての k に対して $t_k > 0$ となるようにアルゴリズムを構築する必要がある. そのために次の集合を定義する.

定義 2.1. $\bar{\gamma} \in (0,1)$, $\bar{t} \in (0,1]$ となる正の定数 $\bar{t}$, $\bar{\gamma}$ に対して, 関数 $\gamma : \boldsymbol{R}^{1+n} \to \boldsymbol{R}_+$ と集合 Ω を

$$\gamma(v) = \bar{\gamma}\min\{1, \Psi(v)\}, \quad \Omega = \{v | t \geq \bar{t}\gamma(v)\} \tag{10}$$

と定義する. ただし, $\boldsymbol{R}_+ = \{t \in \boldsymbol{R} | t \geq 0\}$ である.

もし $\{v_k\} \subset \Omega$ かつ $\Psi_k > 0$ が成り立つならば, 全ての $k \geq 0$ に対し $t_k > 0$ が成り立つ. また, $\{v_k\} \subset \Omega$ が成り立ち, さらに t_k が 0 に近づく場合には, Ψ_k も 0 に近づくことがいえる. したがって, 以降では $\{v_k\} \subset \Omega$ となるように点列を生成するようなアルゴリズムを構築する.

次に, 探索方向 d_k を

$$d_k = \begin{pmatrix} \bar{t}\gamma_k - t_k \\ \tilde{d}_k \end{pmatrix} \tag{11}$$

と定める. ここで, $\bar{t}\gamma_k - t_k \in \boldsymbol{R}$ は変数 t に関する探索方向, $\tilde{d}_k \in \boldsymbol{R}^n$ は変数 x に関する探索方向である. また, 定数 $\eta \in (0,1)$ に対して, $\tilde{d}_k$ はスケーリング付き共役勾配法 (6) に基いて, 以下のように定める.

Case 1. $\nabla_x\Psi_k = 0$ の場合:

$$\tilde{d}_k = 0. \tag{12}$$

Case 2. $\nabla_x\Psi_k \neq 0$ かつ $\eta\|\nabla_x\Psi_k\|^2 \geq (\bar{t}\gamma_k - t_k)\nabla_t\tilde{F}_k\tilde{F}_k$ の場合:

$$\tilde{d}_k = \begin{cases} -\nabla_x\Psi_k, & k=0. \\ -\left(1+\beta_k\dfrac{\nabla_x\Psi_k^T\tilde{d}_{k-1}}{\|\nabla_x\Psi_k\|^2}\right)\nabla_x\Psi_k + \beta_k\tilde{d}_{k-1}, & k\geq 1. \end{cases} \tag{13}$$

Case 3. $\nabla_x\Psi_k \neq 0$ かつ $\eta\|\nabla_x\Psi_k\|^2 < (\bar{t}\gamma_k - t_k)\nabla_t\tilde{F}_k\tilde{F}_k$ の場合:

$$\tilde{d}_k = \begin{cases} -\theta_k\nabla_x\Psi_k, & k=0. \\ -\left(\theta_k+\beta_k\dfrac{\nabla_x\Psi_k^T\tilde{d}_{k-1}}{\|\nabla_x\Psi_k\|^2}\right)\nabla_x\Psi_k + \beta_k\tilde{d}_{k-1}, & k\geq 1. \end{cases} \tag{14}$$

ただし, $\gamma_k = \gamma(v_k)$ であり, $\theta_k \in \boldsymbol{R}$, $\beta_k \in \boldsymbol{R}$, $y_{k-1} \in \boldsymbol{R}^n$ はそれぞれ

$$\begin{aligned} \theta_k &= 1 + \frac{(\bar{t}\gamma_k - t_k)\nabla_t\tilde{F}_k\tilde{F}_k}{\|\nabla_x\Psi_k\|^2}, \\ \beta_k &= \frac{\nabla_x\Psi_k^T y_{k-1}}{\|\nabla\Psi_{k-1}\|^2}, \\ y_{k-1} &= \nabla_x\Psi_k - \nabla_x\Psi_{k-1} \end{aligned} \tag{15}$$

とする. ここで, Polak-Ribière の選択法 [11] に対応する β_k の分母が $\|\nabla_x\Psi_{k-1}\|^2$ ではなくて $\|\nabla\Psi_{k-1}\|^2$ であることに注意されたい.

$\{v_k\} \subset \Omega$ を保証するために, 任意の $k \geq 0$ に対して $0 < t_{k+1} \leq t_k$ と $\Psi_{k+1} < \Psi_k$ が成り立つことが必要である.

命題 2.1. ([9, Proposition 2.1]) $v_k \in \Omega$, かつ $t_k > 0$ であるとき, 任意の $\alpha \in (0,1]$ に対して, $0 < t_k + \alpha(\bar{t}\gamma_k - t_k) \leq t_k$ が成立する.

命題 2.1 の α をステップ幅 $\alpha_k \in (0,1]$ とすれば, $0 < t_{k+1} \leq t_k$ が成り立つことを意味している. 次に $\Psi_{k+1} < \Psi_k$ が成り立つことを示すために, 次の補題を与える.

補題 2.2. $\nabla_x\Psi_k \neq 0$ とする. このとき, 任意の定数 $\lambda \in \boldsymbol{R}$, $\beta \in \boldsymbol{R}$ に対して $\tilde{d}_k$ を

$$\tilde{d}_k = -\left(\lambda + \beta\frac{\nabla_x\Psi_k^T\tilde{d}_{k-1}}{\|\nabla_x\Psi_k\|^2}\right)\nabla_x\Psi_k + \beta\tilde{d}_{k-1}$$

と定めると, 関係式

$$\nabla_x\Psi_k^T\tilde{d}_k = -\lambda\|\nabla_x\Psi_k\|^2$$

が成り立つ.

次の命題は, 探索方向 d_k がメリット関数 Ψ の降下方向になることを示している.

命題 2.3. $\nabla_x \tilde{F}_k$ が正則のとき, $v_k \in \Omega$ と $0 < t_k \leq 1$ に対して,

$$\nabla \Psi_k^T d_k \leq -(1-\eta)\|\nabla_x \Psi_k\|^2 + t_k(\bar{t}\gamma_k - t_k) < 0.$$

が成り立つ.

ここで, $0 < t_{k+1} \leq t_k$ と $\Psi_{k+1} < \Psi_k$ を満たす平滑化スケーリング共役勾配法 (smoothing scaled conjugate gradient method: SSCG) のアルゴリズムを与える.

アルゴリズム SSCG.

Step 0. $\bar{t} \in (0,1]$, $\bar{\gamma} \in (0,1)$, $\eta \in (0,1)$, $\sigma \in (0,1)$, $\delta \in (0,1)$ を与える. $t_0 = \bar{t}$ とし, 初期点 $v_0 = (t_0, x_0) \in \Omega$ を与え, $k = 0$ とする.

Step 1. $\|F(x_k)\| = 0$ ならば反復終了. x_k を解とする.

Step 2. パラメータ θ_k, β_k を (15) で計算し, 探索方向 d_k を (11)–(14) により求める.

Step 3. 次の式を満たす最も小さな非負の整数 l を求める:

$$\Psi(v_k + \sigma^l d_k) \leq \Psi(v_k) - \delta\|\sigma^l d_k\|^2.$$

そして, ステップ幅を $\alpha_k = \sigma^l$ とする.

Step 4. v_{k+1} を (9) により更新する.

Step 5. $k = k+1$ とし, Step 1 へ戻る

Step 0 では初期点を $v_0 \in \Omega$ となるように選ばなければならない. ここで, (10) の γ_k の定義より, 任意の定数 $\bar{t} \in (0,1]$ に対して $\bar{t}\gamma(v_0) = \bar{t}\bar{\gamma}\min\{1, \Psi(v_0)\} \leq \bar{t}$ が成り立つ. さらに, $t_0 = \bar{t}$ とすれば, $t_0 \geq \bar{t}\gamma_0$ が成立する. したがって, $t_0 = \bar{t}$ とおけば $v_0 \in \Omega$ が成立する.

アルゴリズム SSCG の Step 3 においてバックトラック法を用いているが, そのかわりに 2 次補間を用いることも可能である. 今回, 我々はバックトラック法を用いたアルゴリズム SSCG の大域的収束性のみを議論するが, 2 次補間を用いたアルゴリズム SSCG の大域的収束性も同様の方法で証明できることを注意しておく.

次にアルゴリズム SSCG の大域的収束性を証明する. そのために, $\tilde{F}$ に関して以下の条件を課すこととする.

仮定 2.1.

A1. 任意の正の実数 t に対して, 次式が成り立つ:

$$\lim_{\|x\|\to\infty} \|\tilde{F}(t,x)\| = \infty.$$

A2. 任意の $v \in (\boldsymbol{R}_{++} \times \boldsymbol{R}^n) \cap \Omega$ に対して, $\nabla_x \tilde{F}(v)$ は正則である.

仮定 2.1 が成り立つ場合, 次の命題が成り立つ.

命題 2.4. 仮定 2.1 が成り立つならば, 任意の $k \geq 0$ に対して, $v_k \in \Omega$ となる.

次にアルゴリズム SSCG の大域的収束性を考える．もし，$F_k = 0$ となれば，(1) の解が得られたこととなり，アルゴリズム SSCG は有限回で終了する．よって以降では，一般性を失うことなく，全ての $k \geq 0$ に対して $F_k \neq 0$ を仮定する．また，$\{v_k\} \subset \Omega$ かつ $\lim_{k\to\infty} t_k = 0$ ならば，任意の $k > 0$ に対して

$$t_k \geq \bar{t}\bar{\gamma} \min\{1, \Psi_k\} > 0$$

が成り立つので，$\lim_{k\to\infty} \Psi_k = 0$ を得る．これは第 1 節で述べたように，(1) の解が得られたこととなる．後述する定理 2.7 でこのことを背理法で証明するために，$\lim_{k\to\infty} t_k \neq 0$ となる場合を考える．このとき，以下の補題が成立する．

補題 2.5. ([9, Lemma 3.1]) 仮定 2.1 を満たすとき，もし $\lim_{k\to\infty} t_k \neq 0$ ならば，

$$\liminf_{k\to\infty} \|\nabla\Psi(v_k)\| \neq 0$$

が成り立つ．

補題 2.6. 仮定 2.1 と $\lim_{k\to\infty} t_k \neq 0$ が成り立つとする．このとき，アルゴリズム SSCG で生成される探索方向の列 $\{d_k\}$ は有界となる．

以上の準備のもとで，アルゴリズム SSCG の大域的収束性に関する次の定理を与える．

定理 2.7. 仮定 2.1 が成り立つならば，アルゴリズム SSCG によって生成される点列 $\{v_k\}$ は少なくとも一つの集積点を持ち，$\lim_{k\to\infty} t_k = 0$ が成り立つ．さらに，任意の集積点 v^* において $H(v^*) = 0$ が満たされる．したがって，$v^* = (0, x^*)$ となる x^* は (1) の解である．

3 数値実験

この節ではアルゴリズム SSCG を用いた数値実験の結果について記述する．今回，我々は Matlab R2006b でコーディングを行い，以下のスペックの計算機:

CPU	:	Intel(R) Core(TM)2 Duo CPU P8400 @ 2.26GHz, 2.26GHz,
メモリ	:	2.99 GB RAM,

を用いて数値実験を行った．また，関数 $F(x) = [F_1(x), \dots, F_n(x)]^T$ の成分 $F_i(x)$ をそれぞれ

$$
\begin{aligned}
\mathrm{P1}: F_i(x) &= \begin{cases} e^{\sqrt{x_i^2+x_{i+1}^2}} - 1, & i \text{ は奇数} \\ x_{i-1} - x_i, & i \text{ は偶数} \end{cases} \\
\mathrm{P2}: F_i(x) &= \begin{cases} e^{\sqrt{x_i^2+x_{i+1}^2}} - 1, & i \text{ は奇数} \\ \min\{x_{i-1}, x_i\}, & i \text{ は偶数} \end{cases} \\
\mathrm{P3}: F_i(x) &= \begin{cases} \max\{0, x_i + x_{i+1}^2 + 2\} - 2, & i \text{ は奇数} \\ \sqrt{x_{i-1}^2 + x_i^2}, & i \text{ は偶数} \end{cases} \\
\mathrm{P4}: F_i(x) &= \begin{cases} e^{\sqrt{x_i^2+x_{i+1}^2}} - 1, & i \text{ は奇数} \\ \max\{x_{i-1}, x_i\}, & i \text{ は偶数} \end{cases} \\
\mathrm{P5}: F_i(x) &= \begin{cases} e^{|\max\{x_i, x_{i+1}\}|} - 1, & i \text{ は奇数} \\ \min\{x_{i-1}, x_i\}, & i \text{ は偶数} \end{cases} \\
\mathrm{P6}: F_i(x) &= n - 1 + e^{|x_i|} - \sum_{j=1}^{n} \cos x_j,
\end{aligned}
$$

と定め，非線形方程式 (1) の求解問題 P1–P6 を扱った．ここで，問題 P1–P6 の真の解はいずれも $x^* = [0, \dots, 0]^T$ である．また，問題 P1–P6 の関数 $F(x)$ は微分不可能な点を持つ関数を含み，解でも微分不可能となっている．今回の実験では，微分不可能な点を持つ関数 f に対し，次の平滑化関数 $\tilde{f}$ を用いた:

$$
\begin{aligned}
f(\alpha) = \sqrt{\alpha} \quad &\longrightarrow \quad \tilde{f}(t, \alpha) = \sqrt{\alpha + t^2}, \\
f(\alpha) = |\alpha| \quad &\longrightarrow \quad \tilde{f}(t, \alpha) = \sqrt{\alpha^2 + t^2}, \\
f(\alpha) = \max\{\alpha, \beta\} \quad &\longrightarrow \quad \tilde{f}(t, \alpha, \beta) = \frac{1}{2}\left(\alpha + \beta + \sqrt{(\alpha - \beta)^2 + t^2}\right), \\
f(\alpha) = \min\{\alpha, \beta\} \quad &\longrightarrow \quad \tilde{f}(t, \alpha, \beta) = \frac{1}{2}\left(\alpha + \beta - \sqrt{(\alpha - \beta)^2 + t^2}\right).
\end{aligned}
$$

今回我々は，これらの平滑化関数 $\tilde{f}$ に基づいたアルゴリズム:

SSCG	:	アルゴリズム SSCG
SCG	:	Narushima [9] が提案した平滑化三項共役勾配法
SSCG-q	:	2 次補間を用いたアルゴリズム SSCG
SCG-q	:	Narushima [9] が提案した 2 次補間を用いた平滑化三項共役勾配法
SNewton	:	平滑化ニュートン法 [13,14]

の数値実験を行い，それらの結果を比較した．ここで，SSCG-q と SCG-q はそれぞれ SSCG と SCG においてバックトラック法の代わりに 2 次補間を用いたアルゴリズムを表している．また，SNewton は非線形方程式 $H(v) = 0$ を解くために，$\bar{v} = (\bar{t}, 0, ..., 0) \in \boldsymbol{R}^{1+n}$ に対しニュートン方程式:

$$H(v_k) + \nabla H(v_k)^T d_k = \gamma(v_k)\bar{v}$$

を解いて探索方向 d_k を決定している．問題 P1–P5 において $\nabla H(v)$ は疎行列となるため，SNewton 法で探索方向 d_k を定める際にその疎性を利用している．一方，問題 P6 に関して

は $\nabla H(v)$ が密行列となっている.

アルゴリズムの停止条件は

$$\|F(x_k)\| \leq 10^{-5}$$

とした. なお, 探索方向 $\tilde{d}_k$ を計算する際の Case 1 の場合, 実際の数値実験では条件 $\nabla_x \Psi_k = 0$ の代わりに $\|\nabla_x \Psi_k\| < 10^{-15}$ を用いた. また, パラメータの値を $\bar{t} = \min\{0.1, 1/n\}$, $\sigma = 0.5$, $\eta = 0.1$, $\delta = 0.1$ とし, $\bar{\gamma}$ は 0.99 と 0.5 の 2 通りで実験を行った. 今回, 我々は 100 個の初期点を $[-1,1]^n$ からランダムにとり, 実験を行った. 表 1 では $\bar{\gamma} = 0.99$ の場合の結果が, また表 2 では $\bar{\gamma} = 0.5$ の場合の結果がまとめられている. ただし, 表中の数値は「平均反復回数 / 平均関数評価回数 / 平均 CPU タイム (秒)」を意味する.

表 1: $\bar{t} = \min\{0.1, 1/n\}$, $\bar{\gamma} = 0.99$, $\sigma = 0.5$, $\eta = 0.1$, $\delta = 0.1$

	n	SSCG	SCG	SSCG-q	SCG-q	SNewton
P1	2	11.88/26.65/0.00505	12.26/27.97/0.00463	8.39/16.03/0.00206	8.58/16.44/0.00379	5.11/5.89/0.00311
	10	16.43/38.99/0.00863	15.71/35.23/0.00608	10.19/20.6/0.00377	10.24/19.21/0.00597	7.36/10.67/0.00191
	100	18.86/48.59/0.03174	17.23/42.7/0.01422	12.05/22.48/0.01756	10.51/20.78/0.0146	9.21/18.23/0.01706
	1000	21.7/59.01/0.71849	19.74/54.55/0.65787	12.39/24.18/0.39865	9.76/19.67/0.31488	10.85/26.29/0.3749
	2000	22.65/63.45/2.84457	20.58/60.95/2.59395	13.58/27.91/1.659	9.28/19.45/1.13736	11.4/29.47/1.38131
P2	2	10.47/19.64/0.00441	8.69/14.38/0.00426	9.23/16.61/0.00458	8.72/13.81/0.00378	5.77/6.9/0.00142
	10	16.68/31.79/0.00398	12.5/20.33/0.004	15.24/25.83/0.01024	12.59/18.55/0.00838	74.63/1243.9/0.1806
	100	20.03/40.33/0.01731	16.19/27.47/0.01359	17.79/30.02/0.02821	15.46/21.47/0.02179	18.35/71.06/0.04564
	1000	22.96/48.67/0.78036	18.95/37.98/0.64151	22.28/34.97/0.7343	17.27/24.06/0.56727	25.58/92.88/1.02443
	2000	23.68/49.95/2.94125	19.51/40.65/2.4247	23.16/36.32/2.81777	16.83/23.65/2.0453	43.47/110.28/5.65803
P3	2	10.5/18.77/0.00606	6.82/12.74/0.00431	6.74/12.34/0.0048	5.53/10.54/0.00288	75.67/1260.65/0.15005
	10	18.25/31.87/0.00608	8.54/15.81/0.004	10.36/18.57/0.00656	7.33/13.75/0.0064	267.42/4476.28/0.5901
	100	24.2/41.92/0.01893	10.28/19.09/0.00893	12.26/21.58/0.01879	8.35/15.35/0.01323	21.21/69.83/0.04856
	1000	22.96/42.36/0.81581	12.17/24.32/0.43305	14.64/25.31/0.51226	9.03/16.03/0.31588	128.19/900.27/6.13621
	2000	21.33/39.8/2.83986	12.64/26.19/1.70132	14.75/25.52/1.97289	9.02/16.02/1.20524	265.7/2202.34/47.40725
P4	2	9.98/18.58/0.00409	8.35/13.49/0.00379	9.05/15.94/0.0038	8.36/12.99/0.00238	5.51/6.8/0.00219
	10	16.87/31.9/0.01395	12.52/20.37/0.00544	15.97/26.75/0.0115	12.76/18.7/0.00884	8.91/15.63/0.00496
	100	19.71/39.22/0.02991	16.12/27.5/0.01424	17.47/29.48/0.02139	15.67/21.49/0.02037	14.63/57.94/0.03527
	1000	22.65/47.58/0.77009	19.16/37.29/0.65003	22.12/34.69/0.73537	17.23/24.17/0.56914	17.44/62.26/0.70476
	2000	22.99/49.37/2.86452	19.92/41.61/2.49439	23.13/36.18/2.82311	16.33/23.14/1.99106	28.98/78.85/3.78426
P5	2	10.53/13.63/0.00381	4.1/6.69/0.00173	9.6/12.57/0.0035	4.09/6.66/0.00126	5.03/5.33/0.0014
	10	11.21/14.96/0.00667	5.24/8.27/0.00128	10.52/14.01/0.00434	5.09/8.08/0.00434	8.04/11.26/0.00295
	100	11.82/16.07/0.01728	6.35/10.1/0.00512	13.49/17.4/0.02087	6.3/9.8/0.01133	9.19/13.95/0.01973
	1000	13.83/18.37/0.45492	7.13/11.14/0.23707	13.9/18.31/0.4577	6.88/10.64/0.23018	13.61/30.59/0.49539
	2000	14.59/19.35/1.77167	7.56/11.87/0.92847	14.03/18.75/1.71476	7.05/10.95/0.86438	14.63/33.96/1.82068
P6	2	11.41/17.33/0.00471	5.44/10.64/0.00203	6.9/10.6/0.00393	4.34/7.89/0.00374	5.04/5.04/0.0014
	10	13.27/25.34/0.01023	7.25/18.38/0.00272	8.8/14.31/0.00899	6.21/11.5/0.0031	6.51/7.03/0.00341
	100	13.53/52.71/0.03702	9.42/44.17/0.01974	9.99/23.75/0.02561	8.33/20.23/0.02386	9.09/9.13/0.02595
	1000	13/75.18/1.81292	9.23/57.42/1.29833	12.43/41.33/1.68422	9.26/31.12/1.25563	18.23/104.99/6.2342
	2000	13.14/87.56/7.1568	9.68/68.17/5.29583	13.73/45.63/7.23465	9.57/32.81/5.05438	13.05/13.18/22.93099

表 2: $\bar{t} = \min\{0.1, 1/n\}$, $\bar{\gamma} = 0.5$, $\sigma = 0.5$, $\eta = 0.1$, $\delta = 0.1$

	n	SSCG	SCG	SSCG-q	SCG-q	SNewton
P1	2	11.84/26.62/0.01075	11.83/27.23/0.0047	8.55/16.12/0.00733	8.48/16.11/0.00406	5.15/5.8/0.00171
	10	16.39/38.95/0.01365	15.28/34.52/0.00464	9.81/19.4/0.00698	10.04/19.14/0.00651	6.52/9.05/0.00295
	100	18.76/48.38/0.03	16.91/41.75/0.01327	12.11/22.52/0.02081	10.56/20.67/0.01877	8.37/16.59/0.01519
	1000	22.49/61.01/0.72792	19.81/55/0.64088	12.52/24.42/0.3918	9.82/19.72/0.30913	9.85/23.81/0.33301
	2000	22.71/63.76/2.75337	21.01/61.81/2.55149	12.95/26.28/1.5364	10.14/20.55/1.20133	10.23/26.38/1.19105
P2	2	10.22/19.48/0.00395	8.53/14.29/0.00299	9.1/16.53/0.00238	8.71/13.66/0.00458	5.81/8.09/0.0014
	10	16.17/30.85/0.00555	12.51/20.25/0.00304	15.51/26.2/0.00977	12.84/18.72/0.00823	26.29/334.9/0.05166
	100	20.04/40.54/0.01688	16.25/27.75/0.01248	17.42/29.61/0.02757	15.62/21.95/0.02412	15.22/59.02/0.04067
	1000	23.09/49.28/0.76333	18.57/37.15/0.61259	19.87/32.97/0.64658	17.05/24.15/0.54986	27.18/77.63/1.03414
	2000	24.07/51.36/2.92178	19.57/40.55/2.37991	19.54/32.78/2.34629	17.22/24.41/2.06417	79.11/126.46/9.77683
P3	2	9.99/17.94/0.00654	6.88/12.86/0.00512	6.57/12.17/0.00463	5.58/10.55/0.00432	16.26/189.8/0.01958
	10	20.28/35.34/0.00719	9.14/16.85/0.00272	10.89/19.34/0.00799	7.91/14.49/0.00575	38.67/522.43/0.07264
	100	25.54/44.47/0.01893	11.04/20.31/0.00751	13.54/23.66/0.02047	8.77/15.71/0.01373	21.26/70.54/0.0468
	1000	27.14/48.41/0.92896	12.07/23.29/0.41816	15.05/26.07/0.51612	9.88/16.89/0.3376	130.72/928.78/6.11176
	2000	24.65/44.97/3.24429	13.81/28.24/1.83338	14.99/25.99/1.95271	9.96/16.99/1.29791	270.26/2245.33/47.41049
P4	2	9.94/18.54/0.00474	8.35/13.6/0.00299	8.96/16.02/0.00253	8.45/13.22/0.00398	5.44/6.73/0.00094
	10	16.58/31.68/0.01061	12.84/20.7/0.00447	15.61/26.26/0.00875	12.48/18.34/0.0095	7.69/12.89/0.00328
	100	19.82/39.96/0.02912	15.92/27.36/0.01392	17.51/29.57/0.02386	15.86/22.11/0.02143	12.97/53.11/0.03334
	1000	23.05/48.4/0.77214	19/37.66/0.63713	19.31/32.35/0.63757	16.96/24.16/0.55414	16.79/53.73/0.66729
	2000	24.02/51.3/2.93996	19.62/40.89/2.40622	19.57/32.86/2.37125	16.91/24.13/2.03401	37.33/73.73/4.70553
P5	2	9.54/12.71/0.00363	4.14/6.72/0.00157	9.1/12.1/0.00427	4.08/6.62/0.00205	4.77/5.05/0.00141
	10	13.58/17.58/0.00775	5.24/8.32/0.00176	12.94/16.58/0.00962	5.15/8.18/0.00434	6.36/7.2/0.00217
	100	13.91/18.25/0.02459	6.58/10.33/0.00512	13.77/17.9/0.01935	6.52/10.15/0.01132	7.31/8.81/0.01397
	1000	20.54/25.21/0.65585	7.16/11.27/0.23298	17.41/21.97/0.55755	6.98/10.72/0.22835	9.44/14.46/0.3276
	2000	21.77/26.67/2.58336	7.62/11.84/0.92272	18.7/23.53/2.22476	7.06/10.94/0.84961	11.44/21.73/1.34525
P6	2	10.45/15.7/0.00993	4.92/9.66/0.00457	6.68/10.3/0.00539	4.23/7.73/0.00458	4.96/4.97/0.00142
	10	12.01/24.49/0.01401	7.18/18.51/0.00927	8.88/14.36/0.00829	6.06/11.42/0.00656	6.53/6.61/0.00336
	100	14.26/54.14/0.03071	9.12/43.06/0.0199	10.3/24.22/0.02081	8.27/20.02/0.01609	9.3/10.24/0.0222
	1000	13.32/76.28/1.87961	9.59/60.44/1.36849	12.92/42.56/1.76305	9.72/32.51/1.3299	61.45/813.09/22.70642
	2000	12.81/87.85/7.10101	9.94/69.65/5.53362	13.57/45.23/7.25007	9.55/32.67/5.10698	41.49/495.33/80.18035

まず, SSCG, SCG, SSCG-q, SCG-q と SNewton を比較すると, 次元数が $n = 1000, 2000$ の場合には SSCG, SCG, SSCG-q, SCG-q の方が反復回数, 関数評価回数, CPU タイム共に良い結果となることが多かった. また, 問題 P3, P6 に関して SNewton はパラメータによって数値実験結果が著しく悪くなることが起こったが, SSCG, SCG, SSCG-q, SCG-q は大きなぶれが無かった. このことから, SSCG, SCG, SSCG-q, SCG-q は次元が大きな問題, 特にヤコビ行列が密であるような問題に対して有効なアルゴリズムであると言える.

一方, SSCG と SCG, SSCG-q と SCG-q をそれぞれ比較すると, SCG と SCG-q の方が実験結果が良いことが見て取れる。特に, P5 の反復回数に着目すると, SSCG (SSCG-q) は SCG (SCG-q) の約 2 倍以上かかってしまっている. 特に, $\bar{\gamma} = 0.5$ (表 2) で n が大きいときには SSCG (SSCG-q) は SCG (SCG-q) の約 3 倍近くの反復が必要である場合もあることが分かる. この原因を探るため, SSCG の各反復でいくつかのパラメータの数値的な挙動を調べたところ, SSCG が SCG よりも多くの反復を必要とするときには, 反復の途中で t_k が極端に小さくなってしまっている場合が多いことが判明した. このことは Ω の定義から, $\bar{\gamma}$ が小さい時の方が SSCG の実験結果が悪いという事実にも合致する. そこで, $t_0 = \bar{t}$ が少し大きくなるように, $\bar{t} = \min\{0.1, 1/\sqrt{n}\}$ として P5 (n=2000, $\bar{\gamma} = 0.5$) に対

してSSCGを実行した．初期点100個に対してSSCGを実行して，平均値をとったものが表3である．表3から，$\bar{t}$が小さくなりすぎないように選んだ方が実験結果が良いことが見て取れる．ただし，$\bar{t}$を大きくとりすぎると，アルゴリズムが解に収束しないことが多いことを注意しておく．

表 3: $\bar{t}=\min\{0.1, 1/\sqrt{n}\}$, $\bar{\gamma}=0.5$, $\sigma=0.5$, $\eta=0.1$, $\delta=0.1$, $n=2000$

		反復回数	関数評価回数	CPUtime
P5	SSCG	10	22	1.29018

4 おわりに

本論文では，微分不可能な関数を含む非線形方程式系に対する行列を用いない解法として，平滑化スケーリング共役勾配法を提案した．この方法で生成される探索方向は，メリット関数 Ψ の降下方向を必ず生成する．さらに，直線探索においてArmijo条件の変種を用いたアルゴリズムを構築し，その大域的収束性を証明した．また，数値実験によって，平滑化スケーリング共役勾配法が高次元の問題において平滑化ニュートン法 [13,14] と比較して有効であることを確かめた．一方で，平滑化スケーリング付き共役勾配法は平滑化三項共役勾配法 [9] と比較して数値実験結果が期待したほどよくなかった．さらに，平滑化スケーリング付き共役勾配法は t_k の値によって数値実験の結果が大きく左右されることがわかった．

今後の課題としては，SSCGの計算効率向上のために，t_k の値を適当にコントロールする方法を構築することなどがあげられる．

謝辞

本研究の一部は日本学術振興会科学研究費補助金基盤研究 (C)(25330030) および若手研究 (B)(25870239) の支援を受けて行われている．

参考文献

[1] W. Cheng, A two-term PRP-based descent method, *Numerical Functional Analysis and Optimization* 28 (2007) 1217–1230.

[2] Y.H. Dai and Y. Yuan, A nonlinear conjugate gradient method with a strong global convergence property, *SIAM Jounal on Optimization* 10 (1999) 177–182.

[3] F. Facchienei and J.S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems* I, II, Springer, New York, 2003.

[4] R. Fletcher and C.M. Reeves, Function minimization by conjugate gradients, *The Computer Journal* 7 (1964) 149–154.

[5] J.C. Gilbert and J. Nocedal, Global convergence properties of conjugate gradient methods for optimization, *SIAM Journal on Optimization* 2 (1992) 21–42.

[6] W.W. Hager and H. Zhang, A new conjugate gradient method with guaranteed descent and an efficient line search, *SIAM Journal on Optimization* 16 (2005) 170–192.

[7] W.W. Hager and H. Zhang, A survey of nonlinear conjugate gradient methods, *Pacific Journal of Optimization* 2 (2006) 35–58.

[8] W. Nakamura, Y. Narushima and H. Yabe, Nonlinear conjugate gradient methods with sufficient descent properties for unconstrained optimization, *Journal of Industrial and Management Optimization* 9 (2013) 595–619.

[9] Y. Narushima, A smoothing conjugate gradient method for solving systems of nonsmooth equations, *Applied Mathematics and Computation* 219 (2013) 8646–8655.

[10] Y. Narushima, H. Yabe and J.A. Ford, A three-term conjugate gradient method with sufficient descent property for unconstrained optimization, *SIAM Jounal on Optimization* 21 (2011) 212–230.

[11] J. Nocedal and S.J. Wright, *Numerical Optimization, Second Edition*, Springer Series in Operations Research, Springer, New York, 2006.

[12] L. Qi and J. Sun, A nonsmooth version of Newton's method, *Mathematical Programming* 58 (1993) 353–367.

[13] L. Qi and D. Sun, A survey of some nonsmooth equations and smoothing Newton methods, A Eberhard, R. Hill, D. Ralph, B.M. Glover(Eds.), *Progress in Optimization*, Springer, 1999, 121–146.

[14] L. Qi, D. Sun and G. Zhou, A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained variational inequalities, *Mathmatical Programming* 87 (2000) 1–35.

[15] L. Zhang, W. Zhou and D.H. Li, Global convergence of a modified Fletcher-Reeves conjugate gradient method with Armijo-type line search, *Numerische Mathematik* 104 (2006) 561–572.

[16] L. Zhang, W. Zhou and D.H. Li, A desent modified Polak-Ribière-Polyak conjugate gradient method and its global convergence, *IMA Journal of Numerical Analysis* 26 (2006) 629–640.

[17] L. Zhang, W. Zhou and D.H. Li, Some descent three-term conjugate gradient methods and their global convergence, *Optimization Methods and Software* 22 (2007) 697–711.