

無制約最適化問題に対するメモリーレス修正 SR1 法の 大域的収束性について

東京理科大学大学院 中山 舜民 (Shummin Nakayama)
Graduate school, Tokyo University of Science
横浜国立大学 成島 康史 (Yasushi Narushima)
Yokohama National University
東京理科大学 矢部 博 (Hiroshi Yabe)
Tokyo University of Science

1 はじめに

本稿では、以下の無制約最小化問題について考える。

$$\min f(x) \quad (1.1)$$

ただし、 $f: \mathbf{R}^n \rightarrow \mathbf{R}$ は十分滑らかな関数とし、その勾配ベクトルを $g := \nabla f(x)$ で表す。一般に、無制約最適化問題 (1.1) を解くために反復法が広く用いられている。反復法は反復式

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

によって与えられる。ここで、 x_k は k 反復目の近似解とし、 $g_k = \nabla f(x_k)$ と表す。また $\alpha_k > 0$ はステップ幅、 d_k を探索方向と呼ばれる。本稿では以下の反復法のアルゴリズムを用いる。

アルゴリズム 1 (反復法のアルゴリズム)

Step0: 初期点 x_0 と $0 < \delta < \sigma < 1$, $\varepsilon > 0$ を与え、 $k = 0$ とする。

Step1: 終了条件 $\|g_k\|_\infty < \varepsilon$ を満たすならばアルゴリズムは停止して x_k を最適解とする。

Step2: 探索方向 d_k を与える。ただし、 $k = 0$ のときは $d_0 = -g_0$ とする。

Step3: Wolfe 条件

$$\begin{aligned} f(x_k) - f(x_k + \alpha_k d_k) &\geq -\delta \alpha_k g_k^T d_k \\ g(x_k + \alpha_k d_k)^T d_k &\geq \sigma g_k^T d_k \end{aligned} \quad (1.2)$$

を満たすようなステップ幅 $\alpha_k > 0$ を求める。

Step4: 点 x_{k+1} を $x_{k+1} = x_k + \alpha_k d_k$ によって更新する。

Step5: $k = k + 1$ として Step1 に戻る。□

Step2 で与えられる探索方向 d_k の選び方として様々な方法が提案されており、最急降下法、ニュートン法、準ニュートン法、非線形共役勾配法などがよく知られている。中でも準ニュートン法は有効な方法として知られている。準ニュートン法の探索方向は

$$d_k = -H_k g_k \quad (1.3)$$

によって与えられる。ただし、 H_k は $\nabla^2 f(x_k)^{-1}$ の近似行列とする。一般に、近似行列 H_k はセカント条件

$$s_{k-1} = H_k y_{k-1} \quad (1.4)$$

を満たすように選ばれることが多い。ここで、 $s_{k-1} = x_k - x_{k-1}$, $y_{k-1} = g_k - g_{k-1}$ とする。通常、 H_k はひとつ前の近似行列 H_{k-1} を用いて更新される。セカント条件 (1.4) を満たすような行列の更新式がいくつか提案されており、Broyden-Fletcher-Goldfarb-Shanno (BFGS) 公式

$$H_k = H_{k-1} - \frac{H_{k-1} y_{k-1} s_{k-1}^T + s_{k-1} y_{k-1}^T H_{k-1}}{s_{k-1}^T y_{k-1}} + \left(1 + \frac{y_{k-1}^T H_{k-1} y_{k-1}}{s_{k-1}^T y_{k-1}}\right) \frac{s_{k-1} s_{k-1}^T}{s_{k-1}^T y_{k-1}} \quad (1.5)$$

や、対称ランクワン (Symmetric Rank-one; SR1) 公式

$$H_k = H_{k-1} + \frac{(s_{k-1} - H_{k-1} y_{k-1})(s_{k-1} - H_{k-1} y_{k-1})^T}{(s_{k-1} - H_{k-1} y_{k-1})^T y_{k-1}} \quad (1.6)$$

などが知られている。BFGS 公式を用いた場合、 H_{k-1} が正定値で $s_{k-1}^T y_{k-1} > 0$ ならば H_k も正定値になる。したがって BFGS 公式を用いた準ニュートン法は、常に降下方向を生成する。ここで、降下方向とは、すべての k に対して方向微係数が負 ($d_k^T g_k < 0$) となることをいう。一方、SR1 公式を用いた場合、近似行列は正定値であるとは限らないため、常に降下方向を生成するとは限らないという欠点がある。そのため、Sun [12] によって修正 SR1 公式

$$H_k = \theta_{k-1} H_{k-1} + \frac{(s_{k-1} - \theta_{k-1} H_{k-1} y_{k-1})(s_{k-1} - \theta_{k-1} H_{k-1} y_{k-1})^T}{(s_{k-1} - \theta_{k-1} H_{k-1} y_{k-1})^T y_{k-1}} \quad (1.7)$$

が提案されている。この公式はパラメータ θ_{k-1} がある範囲を満たすならば更新された近似行列 H_k が正定値になることが知られている。ここで、BFGS 公式を用いた準ニュートン法を BFGS 法と呼ぶこととし、他の公式も同様とする。

準ニュートン法は有効な数値解法として知られているが、行列の保存を必要とするため、大規模な問題には適さないという欠点がある。そのため、行列を陽に使用しない方法として、記憶制限準ニュートン法 [7, 10] やメモリーレス準ニュートン法 [11] が知られている。本論文では、メモリーレス準ニュートン法に着目する。メモリーレス準ニュートン法の基本的な考え方は、近似行列の更新公式において、 H_{k-1} の代わりに単位行列を用いることである。Shanno [11] は BFGS 公式に基づいたメモリーレス準ニュートン法を提案している。また、Modarres ら [8] は修正 SR1 公式を用いたメモリーレス準ニュートン法を提案している。以降では、BFGS 公式を用いたメモリーレス準ニュートン法をメモリーレス BFGS 法と呼ぶこととし、他の公式も同様とする。

修正 SR1 公式 (1.7) では、近似行列が正定値となる範囲でパラメータ θ_{k-1} をうまく選ぶことで、探索方向をコントロールすることができる。今回我々は、メモリーレス修正 SR1 法が十分な降下条件を満たし、さらに大域的に収束するようなパラメータ θ_{k-1} の選択法を考える。ここで、十分な降下条件とは、ある正定数 c_1 が存在して、すべての k に対して

$$d_k^T g_k \leq -c_1 \|g_k\|^2 \quad (1.8)$$

を満たすことである。ただし、 $\|\cdot\|$ は l_2 ノルムとする。

2 メモリーレス準ニュートン法

この節ではメモリーレス BFGS 法とメモリーレス修正 SR1 法を紹介する. メモリーレス準ニュートン法とは, 準ニュートン法の更新式の H_{k-1} を各反復で常に単位行列 I で置きかえることで得られる探索方向を用いた反復法である. 具体的に, BFGS 公式 (1.5) を用いたメモリーレス準ニュートン法の探索方向は, (1.3) と (1.5) より

$$d_k = -g_k + \frac{g_k^T s_{k-1}}{s_{k-1}^T y_{k-1}} y_{k-1} - \left(\left(1 + \frac{y_{k-1}^T y_{k-1}}{s_{k-1}^T y_{k-1}} \right) \frac{g_k^T s_{k-1}}{s_{k-1}^T y_{k-1}} - \frac{y_{k-1}^T g_k}{s_{k-1}^T y_{k-1}} \right) s_{k-1} \quad (2.1)$$

で与えられる. メモリーレス BFGS 法は正確な直線探索の場合には, $g_k^T s_{k-1} = 0$ となることから, 探索方向が

$$d_k = -g_k + \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}} d_{k-1}$$

となり, Hestenes-Stiefel 公式を用いた非線形共役勾配法の探索方向と一致することが知られている.

次に, メモリーレス修正 SR1 法を紹介する前に, 修正 SR1 公式 (1.7) の詳しく説明する. 修正 SR1 公式は Sun [12] によって SR1 公式が常に正定値な近似行列を生成するように改良された公式である. 修正 SR1 法の近似行列が正定値になるための必要十分条件を示した命題を与える [12].

命題 2.1 $\theta_{k-1} > 0$, $s_{k-1}^T y_{k-1} > 0$, 行列 H_{k-1} は正定値対称行列であるとし, 行列 H_k は (1.7) によって更新された行列とする. このとき H_k が正定値対称行列になる必要十分条件は

$$\theta_{k-1} \notin \left[\frac{s_{k-1}^T y_{k-1}}{y_{k-1}^T H_{k-1} y_{k-1}}, \frac{s_{k-1}^T H_{k-1}^{-1} s_{k-1}}{s_{k-1}^T y_{k-1}} \right] \quad (2.2)$$

が成り立つことである.

Wolfe 条件 (1.2) の下では $s_{k-1}^T y_{k-1} > 0$ が成立するので, (2.2) かつ $\theta_{k-1} > 0$ の範囲で θ_{k-1} を選べば修正 SR1 法の探索方向が降下条件を満たすことを示している. 次に, 修正 SR1 公式を用いたメモリーレス準ニュートン法の探索方向を考えると, (1.3) と (1.7) より

$$d_k = -\theta_{k-1} g_k - \frac{(s_{k-1} - \theta_{k-1} y_{k-1})^T g_k}{(s_{k-1} - \theta_{k-1} y_{k-1})^T y_{k-1}} (s_{k-1} - \theta_{k-1} y_{k-1}) \quad (2.3)$$

で与えられる. 命題 2.1 より, この探索方向が降下条件を満たすためのパラメータ $\theta_{k-1} > 0$ の条件は

$$\theta_{k-1} \notin \left[\frac{s_{k-1}^T y_{k-1}}{y_{k-1}^T y_{k-1}}, \frac{s_{k-1}^T s_{k-1}}{s_{k-1}^T y_{k-1}} \right] \quad (2.4)$$

を満たすことである. (2.4) を満たすパラメータ θ_{k-1} を用いた探索方向 (2.3) を用いた反復法をメモリーレス修正 SR1 法と呼ぶ.

3 メモリーレス修正SR1法のパラメータ θ_{k-1} の選び方

この節ではメモリーレス修正SR1法の大域的収束性が保証されるようなパラメータ θ_{k-1} の選択法について考える. 今回我々は, θ_{k-1} が $\theta_{k-1} > 0$ と (2.4) を満たすべき範囲として

$$0 < \theta_{k-1} < \frac{s_{k-1}^T y_{k-1}}{y_{k-1}^T y_{k-1}} \quad (3.1)$$

の場合について考える (Wolfe 条件より $s_{k-1}^T y_{k-1} > 0$ に注意). もし, (3.1) を満たすならば, $(s_{k-1} - \theta_{k-1} y_{k-1})^T y_{k-1} > 0$ となるので, 方向微係数は

$$d_k^T g_k = -\theta_{k-1} \|g_k\|^2 - \frac{((s_{k-1} - \theta_{k-1} y_{k-1})^T g_k)^2}{(s_{k-1} - \theta_{k-1} y_{k-1})^T y_{k-1}} \leq -\theta_{k-1} \|g_k\|^2$$

となる. よってメモリーレス修正SR1法は降下条件を満たす. さらに, もし正の定数 c が存在して, すべての k に対して $\theta_{k-1} > c$ を満たせば, 十分な降下条件 (1.8) を満たすことが分かる. 今回, 我々は (3.1) を満たすようなパラメータ θ_{k-1} の選択法を2種類提案する.

3.1 H_k の条件数を最小にするようなパラメータ θ_{k-1}

通常, 修正SR1法の場合, パラメータ θ_{k-1} として Wolkowicz [14] によって提案された

$$\theta_{k-1} = \frac{s_{k-1}^T H_{k-1}^{-1} s_{k-1}}{s_{k-1}^T y_{k-1}} \pm \sqrt{\left(\frac{s_{k-1}^T H_{k-1}^{-1} s_{k-1}}{s_{k-1}^T y_{k-1}} \right)^2 - \frac{s_{k-1}^T H_{k-1}^{-1} s_{k-1}}{y_{k-1}^T H_{k-1} y_{k-1}}} \quad (3.2)$$

が選択されることが多い. このパラメータ θ_{k-1} は $H_{k-1}^{-1/2} H_k H_{k-1}^{-1/2}$ の条件数を最小にするようなパラメータである. このパラメータに倣って, メモリーレス修正SR1法のパラメータとして

$$\theta_{k-1}^{(1)} = \frac{s_{k-1}^T s_{k-1}}{s_{k-1}^T y_{k-1}} - \sqrt{\left(\frac{s_{k-1}^T s_{k-1}}{s_{k-1}^T y_{k-1}} \right)^2 - \frac{s_{k-1}^T s_{k-1}}{y_{k-1}^T y_{k-1}}} \quad (3.3)$$

を選ぶ. このパラメータは s_{k-1} と y_{k-1} が互いに線形独立のとき, (3.1) を満たすことを注意しておく. ただし, s_{k-1} と y_{k-1} が線形従属のときには

$$\theta_{k-1}^{(1)} = \frac{s_{k-1}^T s_{k-1}}{s_{k-1}^T y_{k-1}} - \sqrt{\left(\frac{s_{k-1}^T s_{k-1}}{s_{k-1}^T y_{k-1}} \right)^2 - \frac{s_{k-1}^T s_{k-1}}{y_{k-1}^T y_{k-1}}} = \frac{s_{k-1}^T s_{k-1}}{s_{k-1}^T y_{k-1}} = \frac{s_{k-1}^T y_{k-1}}{y_{k-1}^T y_{k-1}}$$

となり, (3.1) を満たさない. さらに, このとき

$$(s_{k-1} - \theta_{k-1} y_{k-1})^T y_{k-1} = 0 \quad (3.4)$$

となってしまうため, メモリーレス修正SR1法の探索方向 (2.3) が定義できない. そのため, 我々はセーフガードとして

$$\|s_{k-1}\|^2 \|y_{k-1}\|^2 \geq (1 + \mu)(s_{k-1}^T y_{k-1})^2, \quad \mu > 0 \quad (3.5)$$

を満たさない場合に、 $\theta_{k-1}^{(1)}$ を用いたメモリーレス修正 SR1 法の探索方向をスケーリング付き最急降下方向

$$d_k = -\theta_{k-1}^{(1)} g_k \quad (3.6)$$

に切り替えることとする。探索方向の切り替えにより、(3.4) の状況でも探索方向が定義できる。SR1 公式の分母が小さくなりすぎた場合に、探索方向を切り替えることは、よく用いられる手法である。詳しくは文献 [6] を参照されたい。

ここで、メモリーレス修正 SR1 法の先行研究である文献 [8] では

$$d_k = -\hat{\theta}_{k-1}^{(1)} g_k - \frac{(s_{k-1} - \hat{\theta}_{k-1}^{(1)} \hat{y}_{k-1})^T g_k}{(s_{k-1} - \hat{\theta}_{k-1}^{(1)} \hat{y}_{k-1})^T \hat{y}_{k-1}} (s_{k-1} - \hat{\theta}_{k-1}^{(1)} \hat{y}_{k-1}),$$

$$\hat{\theta}_{k-1}^{(1)} = \frac{s_{k-1}^T s_{k-1}}{s_{k-1}^T \hat{y}_{k-1}} - \sqrt{\left(\frac{s_{k-1}^T s_{k-1}}{s_{k-1}^T \hat{y}_{k-1}} \right)^2 - \frac{s_{k-1}^T s_{k-1}}{\hat{y}_{k-1}^T \hat{y}_{k-1}}} \quad (3.7)$$

を探索方向とした反復法を使用している。ただし、

$$\hat{y}_{k-1} = y_{k-1} + \frac{\psi_k}{s_{k-1}^T u_{k-1}} u_{k-1}, \quad \psi_k = 2[f(x_{k-1}) - f(x_k)] + (g_k + g_{k-1})^T s_{k-1}$$

とする。ここで、 $s_{k-1}^T y_{k-1} \neq 0$ を満たす任意の $u_{k-1} \in R^n$ とする。また、文献 [8] では $u_{k-1} = s_{k-1}$ としている。彼らは、(3.3) と同様の考え方にに基づき、パラメータ θ_{k-1} として (3.7) を用いているが、彼らの方法はリスタートを用いてないため、 s_{k-1} と $\hat{y}_{k-1}$ が線形従属になってしまった場合には探索方向が定義できないという欠点を残している。

3.2 探索方向のリスタートを必要としないパラメータ θ_{k-1}

次に、 $\theta_{k-1}^{(1)}$ とは異なるパラメータ θ_{k-1} の選択法を考える。(3.1) を満たすようなパラメータ θ_{k-1} として

$$\theta_{k-1}^{(2)} = \rho_k \frac{s_{k-1}^T y_{k-1}}{y_{k-1}^T y_{k-1}}, \quad 0 < \rho_k < 1 \quad (3.8)$$

を提案する。(3.8) を用いたメモリーレス修正 SR1 法の探索方向は

$$d_k = -\theta_{k-1}^{(2)} g_k - \frac{(s_{k-1} - \theta_{k-1}^{(2)} y_{k-1})^T g_k}{(s_{k-1} - \theta_{k-1}^{(2)} y_{k-1})^T y_{k-1}} (s_{k-1} - \theta_{k-1}^{(2)} y_{k-1})$$

$$= -\theta_{k-1}^{(2)} g_k - \frac{(s_{k-1} - \theta_{k-1}^{(2)} y_{k-1})^T g_k}{(1 - \rho_k) s_{k-1}^T y_{k-1}} (s_{k-1} - \theta_{k-1}^{(2)} y_{k-1})$$

と表せる。よって、Wolfe 条件 (1.2) の下では $s_{k-1}^T y_{k-1} > 0$ となるので、(3.8) を用いた場合は常に降下方向を生成し、リスタートを必要としない。具体的な ρ_k の選び方として、 $1/2$ や $1/4$ といった定数や、各反復で値を調節するといった意味で $\cos \langle s_{k-1}, y_{k-1} \rangle$ などを選ぶことが出来る。ただし、 $\cos \langle s_{k-1}, y_{k-1} \rangle = 1$ となることがあるため、その場合に限り、スケーリング付き最急降下方向に切り替える必要がある。ここで、

$$\cos \langle s_{k-1}, y_{k-1} \rangle = \frac{s_{k-1}^T y_{k-1}}{\|s_{k-1}\| \|y_{k-1}\|}$$

とする。

4 大域的収束性

この節では、 $\theta_{k-1}^{(1)}$ や $\theta_{k-1}^{(2)}$ を用いたメモリーレス修正 SR1 法の大域的収束性を示す。そのために、目的関数に以下の仮定をする。

仮定 4.1 初期点 x_0 における準位集合 $\mathcal{L} = \{x | f(x) \leq f(x_0)\}$ は有界であるとする。すなわち、正の定数 $\hat{a}$ が存在して、

$$\|x\| \leq \hat{a}, \quad \forall x \in \mathcal{L} \quad (4.1)$$

が成り立つ。

仮定 4.2 目的関数 f は $\mathcal{L}$ の開凸近傍 $\mathcal{N}$ において連続的微分可能とし、勾配 g はリプシッツ連続であるとする。すなわち、正の定数 L が存在して

$$\|g(u) - g(v)\| \leq L\|u - v\|, \quad \forall u, v \in \mathcal{N} \quad (4.2)$$

が成り立つ。

仮定 4.3 目的関数 f は集合 $\mathcal{N}$ 上で一様凸関数であるとする。すなわち、正の定数 m が存在して

$$(\nabla f(u) - \nabla f(v))^T(u - v) \geq m\|u - v\|^2, \quad \forall u, v \in \mathcal{N} \quad (4.3)$$

が成り立つ。

以上の仮定の下で、メモリーレス修正 SR1 法の大域的収束性を証明する。まず、証明に必要な補題をいくつか示す。ただし、以降ではメモリーレス修正 SR1 法とは (2.3) および $\theta_{k-1}^{(1)}$ 又は $\theta_{k-1}^{(2)}$ を用いたアルゴリズム 1 を指すこととする。

補題 4.1 仮定 4.1–4.3 が成り立つとする。このとき、メモリーレス修正 SR1 法は十分な降下条件を満たす。

補題 4.2 仮定 4.1–4.3 が成り立つとし、 d_k をメモリーレス修正 SR1 法の探索方向とする。このとき、正の定数 c_2 が存在して、全ての k に対して

$$\|d_k\| \leq c_2\|g_k\| \quad (4.4)$$

が成り立つ。

以上の補題を用いることで、以下の収束定理を得る。

定理 4.1 仮定 4.1–4.3 が成り立つならば、メモリーレス修正 SR1 法で生成される点列 $\{x_k\}$ は

$$\lim_{k \rightarrow \infty} \|g_k\| = 0 \quad (4.5)$$

の意味で大域的に収束する。

5 メモリーレス修正 SR1' 法

前節で提案したパラメータを用いたメモリーレス修正 SR1 法の数値実験を行った際に、 θ_{k-1} がかなり小さくなることがしばしば見られた。 θ_{k-1} が小さくなりすぎた場合、メモリーレス修正 SR1 法の探索方向は

$$d_k = -\theta_{k-1}g_k - \frac{(s_{k-1} - \theta_{k-1}y_{k-1})^T g_k}{(s_{k-1} - \theta_{k-1}y_{k-1})^T y_{k-1}}(s_{k-1} - \theta_{k-1}y_{k-1}) \approx -\frac{s_{k-1}^T g_k}{s_{k-1}^T y_{k-1}}s_{k-1}$$

となり、一つ前の探索方向の情報しか持たないということが起こる。そのため、点列が停滞してしまうことがある。この節では、このような状況を防ぐための修正法を考える。

メモリーレス修正 SR1 法 (2.3) の探索方向は

$$\begin{aligned} d_k &= -\theta_{k-1}g_k - \frac{(s_{k-1} - \theta_{k-1}y_{k-1})^T g_k}{(s_{k-1} - \theta_{k-1}y_{k-1})^T y_{k-1}}(s_{k-1} - \theta_{k-1}y_{k-1}) \\ &= \theta_{k-1} \left(-g_k - \frac{(s_{k-1} - \theta_{k-1}y_{k-1})^T g_k}{\theta_{k-1}(s_{k-1} - \theta_{k-1}y_{k-1})^T y_{k-1}}(s_{k-1} - \theta_{k-1}y_{k-1}) \right) \end{aligned}$$

と表せるため、くり出した θ_{k-1} の部分を取り外し、方向を変えることなく、新たに探索方向を

$$\hat{d}_k = -g_k - \frac{(s_{k-1} - \theta_{k-1}y_{k-1})^T g_k}{\theta_{k-1}(s_{k-1} - \theta_{k-1}y_{k-1})^T y_{k-1}}(s_{k-1} - \theta_{k-1}y_{k-1}) \quad (5.1)$$

と定義する。くり出した θ_{k-1} 部分の計算は直線探索によって調整されることに注意しよう。この探索方向は

$$\begin{aligned} \hat{d}_k &= -g_k - \frac{(s_{k-1} - \theta_{k-1}y_{k-1})^T g_k}{\theta_{k-1}(s_{k-1} - \theta_{k-1}y_{k-1})^T y_{k-1}}(s_{k-1} - \theta_{k-1}y_{k-1}) \\ &= -g_k - \frac{(s_{k-1} - \theta_{k-1}y_{k-1})^T g_k}{\theta_{k-1}(s_{k-1} - \theta_{k-1}y_{k-1})^T y_{k-1}}s_{k-1} + \frac{(s_{k-1} - \theta_{k-1}y_{k-1})^T g_k}{(s_{k-1} - \theta_{k-1}y_{k-1})^T y_{k-1}}y_{k-1} \end{aligned}$$

と表せるため、常に勾配 g_k と y_{k-1} の情報を持つ。 s_{k-1} の係数の分母に θ_{k-1} があるが、(3.3) や (3.8) を用いたとき、 θ_{k-1} が小さくなる場合には、 $\frac{(s_{k-1} - \theta_{k-1}y_{k-1})^T g_k}{(s_{k-1} - \theta_{k-1}y_{k-1})^T y_{k-1}}$ も同様に小さくなるため、 s_{k-1} の係数が大きくなりすぎることが少ない。そのため、一つ前の探索方向の情報しか持たないことによる点列の停滞を防ぐことができる。

また、(3.1) を満たす範囲でパラメータ $\theta_{k-1} > 0$ を選んだならば、(1.8) 式において $c_1 = 1$ とした十分な降下条件

$$\hat{d}_k^T g_k = -\|g_k\|^2 - \frac{((s_{k-1} - \theta_{k-1}y_{k-1})^T g_k)^2}{\theta_{k-1}(s_{k-1} - \theta_{k-1}y_{k-1})^T y_{k-1}} \leq -\|g_k\|^2 \quad (5.2)$$

を常に満たす。この探索方向と (2.4) を満たすパラメータ θ_{k-1} を用いた反復法をメモリーレス修正 SR1' 法と呼ぶ。

パラメータ θ_{k-1} として、 $\theta_{k-1}^{(1)}$ または $\theta_{k-1}^{(2)}$ を用いたメモリーレス修正 SR1' 法は、前節と同じ仮定の下で、同様の大域的な収束定理を得ることができる。

6 数値実験

本節では下記の方法を数値実験結果を報告する:

MBFGS	: メモリーレス BFGS 法
MMSR1 $\theta^{(1)}$	: $\theta_{k-1}^{(1)}$ を用いたメモリーレス修正 SR1 法
MMSR1 1/2	: $\rho_k = 1/2$ とした $\theta_{k-1}^{(2)}$ を用いたメモリーレス修正 SR1 法
MMSR1 1/4	: $\rho_k = 1/4$ とした $\theta_{k-1}^{(2)}$ を用いたメモリーレス修正 SR1 法
MMSR1 1/8	: $\rho_k = 1/8$ とした $\theta_{k-1}^{(2)}$ を用いたメモリーレス修正 SR1 法
MMSR1 1/32	: $\rho_k = 1/32$ とした $\theta_{k-1}^{(2)}$ を用いたメモリーレス修正 SR1 法
MMSR1 cos	: $\rho_k = \cos\langle s_{k-1}, y_{k-1} \rangle$ とした $\theta_{k-1}^{(2)}$ を用いたメモリーレス修正 SR1 法
MMSR1' $\theta^{(1)}$	: $\theta_{k-1}^{(1)}$ を用いたメモリーレス修正 SR1' 法
MMSR1' 1/4	: $\rho_k = 1/4$ とした $\theta_{k-1}^{(2)}$ を用いたメモリーレス修正 SR1' 法
MMSR1' 1/32	: $\rho_k = 1/32$ とした $\theta_{k-1}^{(2)}$ を用いたメモリーレス修正 SR1' 法
MMSR1' cos	: $\rho_k = \cos\langle s_{k-1}, y_{k-1} \rangle$ とした $\theta_{k-1}^{(2)}$ を用いたメモリーレス修正 SR1' 法

実験コードは CG-DESCENT5.3 [3, 5] を修正して作成した. CG-DESCENT は Hager and Zhang [4] が作成した非線形共役勾配のソフトウェアである. 直線探索などの設定は CG-DESCENT5.3 に倣っている. また, 収束判定条件として

$$\|g_k\|_\infty \leq 10^{-6}$$

を使用しており, 実行時間が 500 秒を超えた場合もアルゴリズムを停止している. テスト問題として CUTEr 問題集 [2] から 135 問を解いて実験を行った. 表 1 では実験を行った方法の紹介をしている.

各方法の比較のために, Dolan and Moré [1] の提案したパフォーマンスプロファイルを用いた. 各方法のパフォーマンスプロファイル $P(\tau)$ の $\tau = \bar{\tau}$ のときの値は, その解法が全て問題の中で最も早く解くことができた方法の求解時間の $\bar{\tau}$ 倍以内に解くことのできた問題の割合を表している. $\tau = 1$ のときの値は, その方法がすべての方法の中で最も早く解けた問題数の割合を表し, τ が十分大きいときは, その方法の解くことができた問題数の割合を意味する. どの τ においても, $P(\tau)$ が 1 に近いほうが好ましく, 複数の数値解法のパフォーマンスプロファイルのグラフを並べたときにはグラフが上に位置するほど効率が良い数値解法と考えられる.

図 1 はパラメータごとによるメモリーレス修正 SR1 法とメモリーレス BFGS 法を比較したパフォーマンスプロファイルである. 図 1 から, $\theta_{k-1}^{(1)}$ を用いた修正 SR1 法はメモリーレス BFGS 法より悪い結果となっているが, $\theta_{k-1}^{(2)}$ を用いたメモリーレス修正 SR1 法は, どの ρ_k を選んだ場合でもメモリーレス BFGS 法より良い結果を得ていることが分かる. さらに, ρ_k を定数としたときは, 値が小さい方がいい結果が得られ, $\rho_k = \cos\langle s_{k-1}, y_{k-1} \rangle$ としたときが一番良い結果が得られている.

図 2 はいくつかのパラメータを用いたメモリーレス修正 SR1 法とメモリーレス修正 SR1' 法を比較したパフォーマンスプロファイルである. 図 2 から, メモリーレス修正 SR1 法よりメモリーレス修正 SR1' 法の方が良いことがわかる. 特に, $\theta_{k-1}^{(2)}$ の ρ を一番小さくとっている $\rho = 1/32$ のとき, メモリーレス修正 SR1 法とメモリーレス修正 SR1' 法との違いが顕著に表れている.

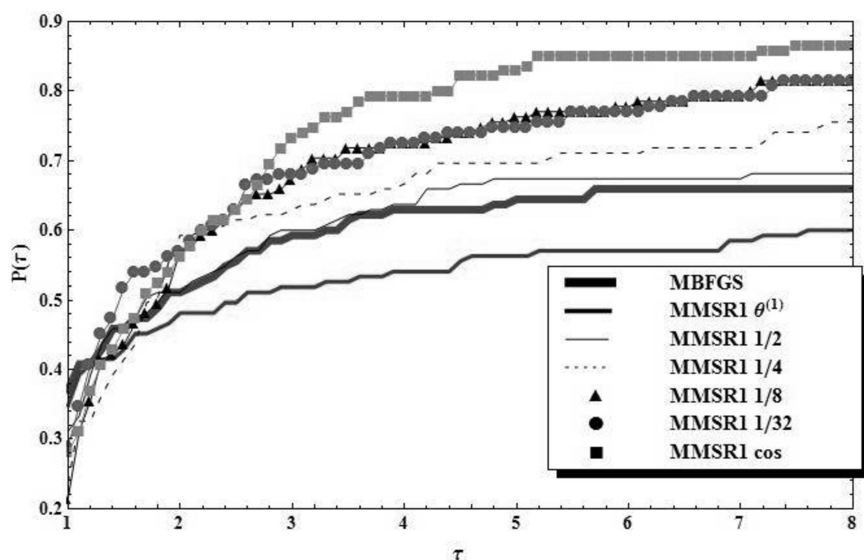


図 1: メモリーレス修正 SR1 法のパラメータごとの比較

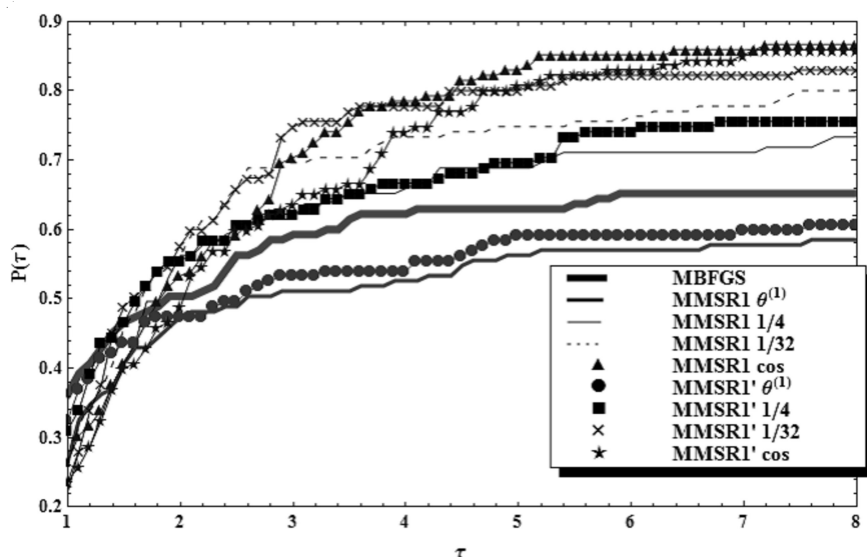


図 2: メモリーレス修正 SR1 法とメモリーレス修正 SR1' 法の比較

7 おわりに

今回我々は、いくつかのパラメータを用いたメモリーレス修正 SR1 法の収束性を示した。また、それらのパラメータを用いたメモリーレス修正 SR1 法の有効性を数値実験により検証した。しかしながら、すべてのテスト問題を解けているわけではない。それは目的関数が一様凸関数という仮定の下でしか大域的収束性が保証されていないことが原因の

可能性がある。よって、今後の課題は、一般関数に対する大域的収束性を証明すること、または大域的収束性を保証するようにメモリーレス修正 SR1 法を改良することである。

参考文献

- [1] E.D. Dolan and J.J. Moré, Benchmarking optimization software with performance profiles, *Mathematical Programming*, **91** (2002), 201-213.
- [2] N.I.M. Gould, D. Orban and P.L. Toint, CUTER and SifDec: A constrained and unconstrained testing environment revisited, *ACM Transactions on Mathematical Software*, **29** (2003), 373-394.
- [3] W.W. Hager, Hager's web page, <http://people.clas.ufl.edu/hager/>, the last access date was August 6, 2015.
- [4] 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.
- [5] W.W. Hager and H. Zhang, Algorithm 851: CG_DESCENT, A conjugate gradient method with guaranteed descent, *ACM Transactions on Mathematical Software*, **32** (2006), 113-137.
- [6] W.J. Leong and M.A. Hassan, A restarting approach for the symmetric rank one update for unconstrained optimization, *Computational Optimization and Applications*, **42** (2009), pp. 327-334.
- [7] D.C. Liu and J. Nocedal, On the limited memory method for large-scale optimization, *Mathematical Programming*, **45** (1989), 503-528.
- [8] F. Modarres, M.A. Hassan and W.J. Leong, Memoryless modified symmetric rank-one method for large-scale unconstrained optimization, *American Journal of Applied Sciences*, **6** (2009), 2054-2059.
- [9] Y. Narushima and H. Yabe, A survey of sufficient descent conjugate gradient method for unconstrained optimization, *SUT Journal of Mathematics*, **50** (2014), 167-203.
- [10] J. Nocedal, Updating quasi-Newton matrices with limited storage, *Mathematics of Computation*, **35** (1980), 773-782.
- [11] D.F. Shanno, Conjugate gradient methods with inexact searches, *Mathematics of Operations Research*, **3** (1978), 244-256.
- [12] L. Sun, An approach to scaling symmetric rank-one update, *Pacific Journal of Optimization*, **2** (2006), 105-118.

- [13] Z. Wei, G. Li and L. Qi, New quasi-Newton methods for unconstrained optimization problems, *Journal of Applied Mathematics and Computational*, **175** (2006) 1156-1188.
- [14] H. Wolkowicz, Measure for symmetric rank-one updates, *Mathematics of Operations Research*, **19** (1994), 815-830.