

無制約最適化問題に対するセカント条件に基づいた 降下条件を保証する非線形共役勾配法

福島工業高等専門学校 コミュニケーション情報学科
成島 康史
東京理科大学 理学部 数理情報科学科
矢部 博

1 はじめに

本論文では無制約最適化問題：

$$\text{minimize } f(x), \quad x \in \mathbf{R}^n$$

に対する数値解法を取り扱う。ただし、 $f: \mathbf{R}^n \rightarrow \mathbf{R}$ は連続微分可能であるとし、目的関数の勾配 ∇f を g で表すこととする。無制約最適化問題のための反復法は初期点 $x_0 \in \mathbf{R}^n$ から出発し

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

によって点列 $\{x_k\}$ を生成する。ここで、 $\alpha_k > 0$ をステップ幅、 $d_k \in \mathbf{R}^n$ を探索方向と呼ぶ。また、以下では $\nabla f(x_k) \equiv g_k$ と表すこととする。反復法の中でも非線形共役勾配法は行列を利用しないため、大規模な問題に対して効果的な解法であり、近年、盛んに研究されている。非線形共役勾配法の探索方向は

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

によって与えられる。ここで、 β_k は非線形共役勾配法を特徴づけるパラメータであり、通常は目的関数が狭義凸 2 次で直線探索が正確な場合には線形共役勾配法に帰着するように選ばれる。しかしながら、目的関数が狭義凸 2 次でない場合には様々な選択が可能であり、その選択法によって数値的な効率性が大きく異なるため、 β_k の選択法の研究が盛んに行われている。よく知られた公式としては Fletcher-Reeves (FR), Hestenes-Stiefel (HS), Polak-Ribière (PR), Dai-Yuan (DY) などがあり、それぞれ以下によって与えられる。

$$\beta_k^{FR} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2}, \quad \beta_k^{HS} = \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}},$$
$$\beta_k^{PR} = \frac{g_k^T y_{k-1}}{\|g_{k-1}\|^2}, \quad \beta_k^{DY} = \frac{\|g_k\|^2}{d_{k-1}^T y_{k-1}}.$$

ただし、 $y_{k-1} = g_k - g_{k-1}$ である。これらの β_k を用いた非線形共役勾配法に対する大域的な収束性は多くの研究者によって議論されており、それらの結果は Hager と Zhang のサーベイ論文 [11] で詳しく解説されている。

近年では、非線形共役勾配法の収束を加速するために、目的関数の2次の情報を取り込むような非線形共役勾配法が提案されており、多くの研究者から注目を集めている。目的関数の2次の情報を取り込むために、Dai と Liao [2] は準ニュートン法の考えであるセカント条件に基づいた非線形共役勾配法を提案し、その大域的な収束性を示している。その後、修正を加えたセカント条件を用いることによって Dai-Liao 法の変種がいくつか提案されている [5, 12, 17, 24]。これらのセカント条件に基づいた方法は、上記のよく知られた方法よりも効率的であるという数値実験結果が報告されている。しかしながら、セカント条件に基づいた方法は、大域的収束性では十分な降下条件:

$$g_k^T d_k \leq -\bar{c} \|g_k\|^2 \quad \text{for some } \bar{c} > 0 \text{ and all } k \quad (2)$$

を仮定しているのも関わらず、実際には、必ずしも降下方向を生成しないという弱点がある。

一方、降下方向を保証するような非線形共役勾配法も研究されている。Hager と Zhang [8] は β_k^{HS} を修正することにより、自動的に十分な降下条件 (2) を満たす非線形共役勾配法を提案している。その後、Yu ら [18] は Hager-Zhang の修正法を用いて β_k^{PR} を修正することにより十分な降下条件 (2) を満たす修正 PR 法を提案している。また、Yuan [19] は Yu らの方法の変種を提案するとともに、 β_k^{HS} や β_k^{PR} 以外の β_k にも Hager-Zhang の修正法が適用できることを示唆している。一方、Zhang ら [21-23] は非線形共役勾配法の探索方向 (1) を修正することにより、 $g_k^T d_k = -\|g_k\|^2$ を常に満たすような修正 FR, PR, HS 法を提案しており、その後、Narushima ら [14] は Zhang らの修正 FR, PR, HS 法を含み、常に $g_k^T d_k = -\|g_k\|^2$ を満たすような非線形3項共役勾配法の族を提案している。Sugiki ら [16] はセカント条件に基づいた β_k を Narushima らの非線形3項共役勾配法に適用し、常に $g_k^T d_k = -\|g_k\|^2$ を満たすようなセカント条件に基づいた非線形3項共役勾配法を提案している。

今回我々は、Hager and Zhang の方法に倣い、セカント条件に基づいた非線形共役勾配法のパラメータ β_k を修正することにより、自動的に降下方向を生成するセカント条件に基づいた非線形共役勾配法を提案し、その大域的収束性を議論する。

2 降下方向を生成するセカント条件に基づいた非線形共役勾配法

この節では自動的に降下方向を生成するセカント条件に基づいた非線形共役勾配法を提案する。まず、そのためにセカント条件に基づいた非線形共役勾配法の紹介を行う。

線形共役勾配法の探索方向は共役性条件 $d_i^T A d_j = 0$ ($i \neq j$) を満たすように生成される。ここで、 A は狭義凸2次関数のヘッセ行列とする。一方、目的関数が2次関数でない場合にはヘッセ行列は一意には決定しないので、 $d_i^T A d_j = 0$ ($i \neq j$) を満たすように探索方向を選択することはできない。したがって、平均値定理から、定数 $\tau \in (0, 1)$ が存在して $d_k^T y_{k-1} = \alpha_{k-1} d_k^T \nabla^2 f(x_{k-1} + \tau \alpha_{k-1} d_{k-1}) d_{k-1}$ が成立するので、 $d_i^T A d_j = 0$ の代わりに $d_k^T \nabla^2 f(x_{k-1} + \tau \alpha_{k-1} d_{k-1}) d_{k-1} = 0$ を用いて

$$d_k^T y_{k-1} = 0 \quad (3)$$

を非線形共役勾配法における共役性条件と定義する. ここで, 共役性条件 (3) に非線形共役勾配法の探索方向 (1) を代入して β_k を求めると β_k^{HS} が得られることを注意しておく.

Perry [15] は準ニュートン法の探索方向の式 $B_k d_k = -g_k$ とセカント条件 $B_k s_{k-1} = y_{k-1}$ を利用して関係式

$$d_k^T y_{k-1} = d_k^T (B_k s_{k-1}) = (B_k d_k)^T s_{k-1} = -g_k^T s_{k-1}$$

を導いた. ただし, $s_{k-1} = x_k - x_{k-1}$ であり, B_k は $\nabla^2 f(x_k)$ の対称な近似行列とする. この関係式 $d_k^T y_{k-1} = -g_k^T s_{k-1}$ を Perry の共役性条件とよんでいる. なお, 正確な直線探索の場合は $g_k^T s_{k-1} = 0$ より, Perry の共役性条件は通常の共役性条件 (3) に帰着される. 2001 年に Dai と Liao [2] は Perry の共役性条件に非負パラメータ t を導入した条件 $d_k^T y_{k-1} = -t g_k^T s_{k-1}$ を提案し, この条件に (1) を代入して, 以下の β_k を提案している.

$$\beta_k^{DL} = \frac{g_k^T (y_{k-1} - t s_{k-1})}{d_{k-1}^T y_{k-1}}.$$

ここで, 正確な直線探索の場合には β_k^{DL} は β_k^{HS} と一致することを注意しておく. Dai と Liao は目的関数が一様凸関数である場合について, β_k^{DL} を用いた非線形共役勾配法の大域的収束性を証明し, さらに, 一般の目的関数における大域的収束性を保証するため β_k^{DL} を

$$\beta_k^{DL+} = \max \left\{ 0, \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}} \right\} - t \frac{g_k^T s_{k-1}}{d_{k-1}^T y_{k-1}}$$

と修正し, 一般の目的関数に対して下記の大域的収束性を証明している. Dai と Liao の研究の後, セカント条件の種類を変更することにより, いくつかの非線形共役勾配法が提案されている.

Yabe と Takano [17] は Zhang ら [20] によって提案された修正セカント条件:

$$B_k s_{k-1} = z_{k-1}^{YT}, \quad z_{k-1}^{YT} = y_{k-1} + \phi_k \left(\frac{\theta_{k-1}}{s_{k-1}^T u_{k-1}} u_{k-1} \right)$$

に基づいた β_k を提案した. ただし, $\phi_k \geq 0$ はスカラーであり,

$$\theta_{k-1} = 6(f(x_{k-1}) - f(x_k)) + 3(g_{k-1} + g_k)^T s_{k-1}$$

とし, $u_{k-1} \in \mathbf{R}^n$ は $s_{k-1}^T u_{k-1} \neq 0$ を満たす任意のベクトルとする. 一方, Zhou と Zhang [24] は, Li と Fukushima [13] によって提案された MBFGS セカント条件

$$B_k s_{k-1} = z_{k-1}^{ZZ}, \quad z_{k-1}^{ZZ} = y_{k-1} + \zeta \|g_k\|^q s_{k-1}$$

に基づいた β_k を提案した. ただし, $\zeta > 0$ と $q > 0$ は定数である. さらに, Ford ら [5] は多段セカント条件 [4]:

$$B_k h_{k-1}^{F1} = z_{k-1}^{F1}, \quad h_{k-1}^{F1} = s_{k-1} - \xi_{k-1} s_{k-2}, \quad z_{k-1}^{F1} = y_{k-1} - \xi_{k-1} y_{k-2}$$

に基づいた β_k を提案した. ただし, $\xi_{k-1} = \frac{\delta_{k-1}^2}{1+2\delta_{k-1}}$, $\delta_{k-1} = \eta_k \frac{\|s_{k-1}\|}{\|s_{k-2}\|}$ であり, $\eta_k \geq 0$ である. また Ford らは z_{k-1} を修正した多段セカント条件:

$$B_k h_{k-1}^{F2} = z_{k-1}^{F2}, \quad h_{k-1}^{F2} = s_{k-1} - \xi_{k-1} s_{k-2}, \quad z_{k-1}^{F2} = y_{k-1} - t \xi_{k-1} y_{k-2}$$

に基づく β_k も提案している.

上記のセカント条件は一般に $B_k h_{k-1} = z_{k-1}$ と表せ, Dai-Liao に倣うことにより統一的な公式

$$\beta_k = \frac{g_k^T (z_{k-1} - th_{k-1})}{d_{k-1}^T z_{k-1}} \quad (4)$$

を得ることができる. 上で紹介した研究では, 上記の β_k やその修正版の

$$\beta_k^+ = \max \left\{ 0, \frac{g_k^T z_{k-1}}{d_{k-1}^T z_{k-1}} \right\} - t \frac{g_k^T h_{k-1}}{d_{k-1}^T z_{k-1}}$$

を提案し, 大域的な収束性を議論している. 本論文では, 必ずしも $d_{k-1}^T z_{k-1} \neq 0$ とは限らないため, (4) の代わりに

$$\beta_k^{Secant} = g_k^T (z_{k-1} - th_{k-1}) (d_{k-1}^T z_{k-1})^\dagger \quad (5)$$

を用いることとする. ここで, \dagger は $a \neq 0$ ならば $a^\dagger = 1/a$, $a = 0$ ならば $a^\dagger = 0$ の一般化逆数とする.

次に, 統一的な公式 (5) に対し Hager-Zhang 法の修正と同様の修正を行うことによりセカント条件に基づき, かつ, 十分な降下条件 (2) を満たす非線形共役勾配法を提案する.

Hager と Zhang [8, 11] は β_k^{HS} を修正した

$$\beta_k^{HZ} = \frac{1}{d_{k-1}^T y_{k-1}} g_k^T \left(y_{k-1} - \lambda d_{k-1} \frac{\|y_{k-1}\|^2}{d_{k-1}^T y_{k-1}} \right) = \beta_k^{HS} - \lambda \left(\frac{\|y_{k-1}\|}{d_{k-1}^T y_{k-1}} \right)^2 g_k^T d_{k-1} \quad (6)$$

を提案している. ここで $\lambda > 1/4$ はパラメータである. Hager と Zhang は (6) を用いた非線形共役勾配法が十分な降下条件 (2) を $\bar{c} = 1 - 1/(4\lambda)$ で満たすことを示している.

Hager-Zhang の修正法に倣い, (5) を修正することにより以下の β_k を得る:

$$\begin{aligned} \beta_k^{DS} &= \beta_k^{Secant} - \lambda \|z_{k-1} - th_{k-1}\|^2 g_k^T d_{k-1} \{(d_{k-1}^T z_{k-1})^2\}^\dagger \\ &= g_k^T (z_{k-1} - th_{k-1}) (d_{k-1}^T z_{k-1})^\dagger - \lambda \|z_{k-1} - th_{k-1}\|^2 g_k^T d_{k-1} \{(d_{k-1}^T z_{k-1})^2\}^\dagger, \end{aligned} \quad (7)$$

ここで, 非線形共役勾配法 (1), (7) は常に $g_k^T d_k \leq -(1 - 1/(4\lambda)) \|g_k\|^2$ を満たす, つまり, 常に十分な降下条件 (2) を $\bar{c} = 1 - 1/(4\lambda)$ として満たしている. また, 大域的な収束性を議論するために β_k^{DS} を非負となるよう補正した

$$\beta_k^{DS+} = \max\{0, \beta_k^{DS}\} \quad (8)$$

も考慮しアルゴリズムを構築することとする. ここで, (8) を用いた非線形共役勾配法も十分な降下条件 (2) を $\bar{c} = 1 - 1/(4\lambda)$ で満たすことを注意しておく. ここで, 直線探索に Wolfe 条件を用いた提案法のアルゴリズムは以下のように与えられる.

アルゴリズム CGDS.

Step 0. 初期点 x_0 , 及び初期探索方向 $d_0 = -g_0$ を与える. $k = 0$ とおいて, Step 2 へ進む.

Step 1. (7)(又は(8))により β_k を計算し, (1)により d_k を計算する.

Step 2. 直線探索により Wolfe 条件:

$$\begin{aligned} f(x_k + \alpha_k d_k) - f(x_k) &\leq \sigma_1 \alpha_k g_k^T d_k, \\ g(x_k + \alpha_k d_k)^T d_k &\geq \sigma_2 g_k^T d_k \end{aligned} \quad (9)$$

を満たすステップ幅 α_k を計算する. ただし, $0 < \sigma_1 < \sigma_2 < 1$ とする.

Step 3. $x_{k+1} = x_k + \alpha_k d_k$ として点列を更新する.

Step 4. 停止条件を満たしていれば終了する.

Step 5. $k \leftarrow k + 1$ として Step 1 へ戻る.

アルゴリズム CGDS では一般的な公式 (7) を用いていたが, 2 節の前半で紹介した具体的な z_{k-1} と h_{k-1} を用いることにより以下の β_k を考えることができる:

$$\begin{aligned} \beta_k^{DSDL} &= g_k^T (y_{k-1} - ts_{k-1})(d_{k-1}^T y_{k-1})^\dagger - \lambda \|y_{k-1} - ts_{k-1}\|^2 g_k^T d_{k-1} \{(d_{k-1}^T y_{k-1})^2\}^\dagger, \\ \beta_k^{DSYT} &= g_k^T (z_{k-1}^{YT} - ts_{k-1})(d_{k-1}^T z_{k-1}^{YT})^\dagger - \lambda \|z_{k-1}^{YT} - ts_{k-1}\|^2 g_k^T d_{k-1} \{(d_{k-1}^T z_{k-1}^{YT})^2\}^\dagger, \\ \beta_k^{DSZZ} &= g_k^T (z_{k-1}^{ZZ} - ts_{k-1})(d_{k-1}^T z_{k-1}^{ZZ})^\dagger - \lambda \|z_{k-1}^{ZZ} - ts_{k-1}\|^2 g_k^T d_{k-1} \{(d_{k-1}^T z_{k-1}^{ZZ})^2\}^\dagger, \\ \beta_k^{DSF1} &= g_k^T (z_{k-1}^{F1} - th_{k-1}^{F1})(d_{k-1}^T z_{k-1}^{F1})^\dagger - \lambda \|z_{k-1}^{F1} - th_{k-1}^{F1}\|^2 g_k^T d_{k-1} \{(d_{k-1}^T z_{k-1}^{F1})^2\}^\dagger, \\ \beta_k^{DSF2} &= g_k^T (z_{k-1}^{F2} - th_{k-1}^{F2})(d_{k-1}^T z_{k-1}^{F2})^\dagger - \lambda \|z_{k-1}^{F2} - th_{k-1}^{F2}\|^2 g_k^T d_{k-1} \{(d_{k-1}^T z_{k-1}^{F2})^2\}^\dagger. \end{aligned}$$

また, (8) に対応する β_k として

$$\begin{aligned} \beta_k^{DSDL+} &= \max\{0, \beta_k^{DSDL}\}, & \beta_k^{DSYT+} &= \max\{0, \tilde{\beta}_k^{DSYT}\}, \\ \beta_k^{DSF1+} &= \max\{0, \beta_k^{DSF1}\}, & \beta_k^{DSF2+} &= \max\{0, \beta_k^{DSF2}\} \end{aligned}$$

も考えることができる. ここで $\tilde{\beta}_k^{DSYT}$ は z_k^{YT} を

$$z_{k-1} = z_{k-1}^{YT+} \equiv y_{k-1} + \phi_k \left(\frac{\max\{0, \theta_{k-1}\}}{s_{k-1}^T u_{k-1}} u_{k-1} \right)$$

に置き換えたものとする.

3 大域的収束性

前節で提案したアルゴリズム CGDS の大域的収束性を示すために, 目的関数に以下の仮定をする.

仮定 1. 初期点における準位集合 $\mathcal{L} = \{x | f(x) \leq f(x_0)\}$ は有界であり, その近傍 \mathcal{N} において f は連続微分可能で, かつ, g はリプシッツ連続である. つまり, 正の定数 L が存在し, 任意の $x, y \in \mathcal{N}$ に対し

$$\|g(x) - g(y)\| \leq L \|x - y\|$$

が成立する.

この仮定の下で一般的な公式 β_k^{DS} を用いたアルゴリズム CGDS の大域的な収束性のための十分条件として、以下の定理を得る。

定理 1. 仮定 1 が成立しているとする。正の定数 c_1 と c_2 が存在して、すべての k に対して z_{k-1} と h_{k-1} が

$$\begin{aligned} \|z_{k-1} - th_{k-1}\| &\leq c_1 \|s_{k-1}\| \\ \alpha_{k-1} \|d_{k-1}\|^2 |d_{k-1}^T z_{k-1}|^\dagger &\leq c_2 \end{aligned}$$

を満たすならば、アルゴリズム CGDS は $\liminf_{k \rightarrow \infty} \|g_k\| = 0$ を満たす。

定理 1 を用いることで前述した具体的な β_k を用いたアルゴリズム CGDS の一様凸関数に対する大域的収束性が得られる。ここで、一様凸関数とは、ある正の定数 μ が存在して、任意の $x, y \in \mathbf{R}^n$ と $\lambda \in (0, 1)$ に対して $f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y) - \frac{1}{2}\mu(1-\lambda)\lambda\|x-y\|^2$ を満たすような関数である。

定理 2. 仮定 1 が成立しているとし、目的関数 f が一様凸関数であるとする。また、 x^* を無制約最適化問題の唯一解とする。このとき、以下が成立する。

1. β_k^{DSDL} を用いたアルゴリズム CGDS は $\lim_{k \rightarrow \infty} \|x_k - x^*\| = 0$ を満たす。
2. ある正の定数 \bar{m} と $\bar{\phi}$ が存在し、 $|s_{k-1}^T u_{k-1}| \geq \bar{m} \|s_{k-1}\| \|u_{k-1}\|$ および $0 \leq \phi_k \leq \bar{\phi}$ を満たすならば、 β_k^{DSYT} を用いたアルゴリズム CGDS は $\lim_{k \rightarrow \infty} \|x_k - x^*\| = 0$ を満たす。
3. $2\mu - \bar{\eta}L > 0$ であるような正の定数 $\bar{\eta}$ に対し、 η_k が $0 \leq \eta_k \leq \bar{\eta}$ を満たすならば、 β_k^{DSF1} を用いたアルゴリズム CGDS は $\lim_{k \rightarrow \infty} \|x_k - x^*\| = 0$ を満たす。
4. $2\mu - t\bar{\eta}L > 0$ であるような正の定数 $\bar{\eta}$ に対し、 η_k が $0 \leq \eta_k \leq \bar{\eta}$ を満たすならば、 β_k^{DSF2} を用いたアルゴリズム CGDS は $\lim_{k \rightarrow \infty} \|x_k - x^*\| = 0$ を満たす。

定理 2 では β_k^{DSDL} , β_k^{DSYT} , β_k^{DSF1} , β_k^{DSF2} を用いた非線形共役勾配法の一様凸関数に対する大域的収束性を保証しているが、 β_k^{DSZZ} を用いた非線形共役勾配法に関しては、定理 1 から一般の非線形関数に対する大域的収束性が証明できる。

定理 3. 仮定 1 が成立しているとする。このとき β_k^{DSZZ} を用いた Algorithm 1 は $\liminf_{k \rightarrow \infty} \|g_k\| = 0$ の意味で大域的に収束する。

次に、 β_k^{DSDL} , β_k^{DSYT} , β_k^{DSF1} , β_k^{DSF2} を用いた非線形共役勾配法の一般の非線形関数に対する大域的収束性を考える。今回、Property* (例えば [6] を参照) を用いた収束性の議論を行うため、 $\beta_k \geq 0$ とする必要がある。そのため、 β_k を非負に補正した β_k^{DSDL+} , β_k^{DSYT+} , β_k^{DSF1+} , β_k^{DSF2+} を用いたアルゴリズム CGDS の大域的収束性を考える。

定理 4. 仮定 1 が成立しているとする。このとき、以下が成立する。

1. β_k^{DSDL+} を用いたアルゴリズム CGDS は $\liminf_{k \rightarrow \infty} \|g_k\| = 0$ の意味で大域的に収束する。

2. ある正の定数 \bar{m} と $\bar{\phi}$ が存在し, $|s_{k-1}^T u_{k-1}| \geq \bar{m} \|s_{k-1}\| \|u_{k-1}\|$ および $0 \leq \phi_k \leq \bar{\phi}$ を満たすとする. このとき, β_k^{DSYT+} を用いたアルゴリズム CGDS は $\liminf_{k \rightarrow \infty} \|g_k\| = 0$ の意味で大域的に収束する.
3. ある正の定数 $\bar{\eta}$ と φ が存在して, $0 \leq \eta_k \leq \bar{\eta}$ と $\varphi \geq \max\{|g_{k-1}^T d_{k-1}|, |g_k^T d_{k-1}|\} |d_{k-1}^T z_{k-1}^{F1}|^\dagger$ を満たすとする. このとき, β_k^{DSF1+} を用いたアルゴリズム CGDS は $\liminf_{k \rightarrow \infty} \|g_k\| = 0$ の意味で大域的に収束する.
4. ある正の定数 $\bar{\eta}$ と φ が存在して, $0 \leq \eta_k \leq \bar{\eta}$ と $\varphi \geq \max\{|g_{k-1}^T d_{k-1}|, |g_k^T d_{k-1}|\} |d_{k-1}^T z_{k-1}^{F2}|^\dagger$ を満たすとする. このとき, β_k^{DSF1+} を用いたアルゴリズム CGDS は $\liminf_{k \rightarrow \infty} \|g_k\| = 0$ の意味で大域的に収束する.

4 数値実験

この節ではアルゴリズム CGDS の数値実験結果を報告する. テスト問題は CUTEr [1, 7] の中から 70 個の問題を選択して実験を行った. 表 1 では実験した関数とその次元を表示している.

表 1: テスト問題

ARWHEAD	5000	DIXMAAND	9000	FLETCHCR	10000	PENALTY1	10000
BDEXP	5000	DIXMAANE	9000	FMINSRF2	5625	POWELLSG	20000
BDQRTIC	5000	DIXMAANF	9000	FMINSURF	5625	POWER	20000
BIGGSB1	5000	DIXMAANG	9000	FREUROTH	5000	QUARTC	10000
BOX	7500	DIXMAANH	9000	GENHUMPS	5000	SCHMVETT	5000
BROYDN7D	5000	DIXMAANI	9000	GENROSE	5000	SINQUAD	10000
BROYDN7D	10000	DIXMAANJ	9000	GENROSE	10000	SPARSINE	5000
BRYBND	10000	DIXMAANK	3000	LIARWHD	10000	SPARSQUR	10000
CHAINWOO	4000	DIXMAANL	9000	MODBEALE	10000	SROSENBR	10000
CHAINWOO	10000	DIXON3DQ	10000	MOREBV	5000	TESTQUAD	5000
COSINE	10000	DQDRTIC	5000	MOREBV	10000	TOINTGSS	10000
CRAGGLVY	5000	DQRTIC	5000	NONCVXU2	5000	TQUARTIC	10000
CURLY10	10000	EDENSCH	10000	NONDIA	10000	TRIDIA	10000
CURLY20	10000	EG2	1000	NONDQUAR	5000	VAREIGVL	5000
CURLY30	5000	ENGVAL1	10000	NONDQUAR	10000	WOODS	4000
DIXMAANA	9000	EXTROSNB	1000	NONSCOMP	5000	WOODS	10000
DIXMAANB	9000	EXTROSNB	10000	OSCIPATH	10000		
DIXMAANC	9000	FLETCHCR	1000	PENALTY1	1000		

今回実験した方法は下記の 6 種類である:

CG-DESCENT	:	Hager-Zhang 法のソフトウェア [8-10],
DSDL+	:	アルゴリズム CGDS, β_k^{DSDL+} , $(\lambda, t) = (2, 0.3)$,
DSYT+	:	アルゴリズム CGDS, β_k^{DSYT+} , $(\lambda, t, \phi_k) = (2, 0.3, 0.3)$, $u_k = y_k$,
DSZZ+	:	アルゴリズム CGDS, $\beta_k^{DSZZ+} \equiv \max\{0, \beta_k^{DSZZ}\}$, $(\lambda, t) = (2, 0.3)$,
DSF1+	:	アルゴリズム CGDS, β_k^{DSF1+} , $(\lambda, t, \eta_k) = (2, 0.3, 0.3)$,
DSF2+	:	アルゴリズム CGDS, β_k^{DSF2+} , $(\lambda, t, \eta_k) = (2, 0.3, 0.3)$.

Zhou と Zhang [24] に倣い, DSZZ+においては $\zeta = 0.001$ とし, $\|g_k\| \geq 1.0$ のとき $q = 1.0$ とし, それ以外の場合は $q = 3.0$ とした. 比較の対象として β_k^{HZ} を用いた非線形共役勾配法のソフトウェアである CG-DESCENT [8-10] を選び, そのほかの方法は CG-DESCENT のコードをもとにコーディングを行った. 直線探索は CG-DESCENT に倣い, $\sigma_1 = 10^{-4}$ かつ $\sigma_2 = 0.1$ とした近似 Wolfe 条件:(9) と

$$-(1 - 2\sigma_1)g_k^T d_k \geq g(x_k + \alpha_k d_k)^T d_k \geq \sigma_2 g_k^T d_k$$

を満たすようなステップ幅 α_k を用いている. 終了条件は

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

としている. ここで, $\|\cdot\|_\infty$ は無限大ノルムを表している.

図 1-3 では, それぞれ, 反復回数, 関数評価回数, 勾配ベクトルの評価回数を評価基準とした各方法のパフォーマンスプロファイル [3] を表している. 各図においてそれぞれの線が各方法のパフォーマンスプロファイルであり, 他の方法よりも上に位置するほど, 相対的に効率が良いと判断できる. 図 1-3 から DSF1+ と DSF2+ は CG-DESCENT よりも効率が良いことと DSDL+ が CG-DESCENT とほぼ同程度であることが見て取れる. また, DSYT+ と DSZZ+ については CG-DESCENT より多少下回っているが, それほど大きな差はないことが分かる. これらの結果から提案法は数値的にも有効な方法であると結論付けることができるだろう. 今後の課題としては各方法に含まれるパラメータの効果的な選択法などがあげられる.

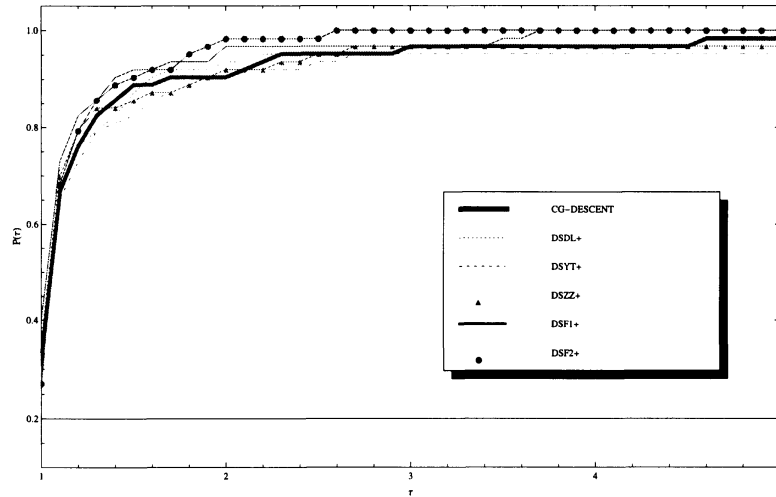


図 1: 反復回数を基準としたパフォーマンスプロファイル

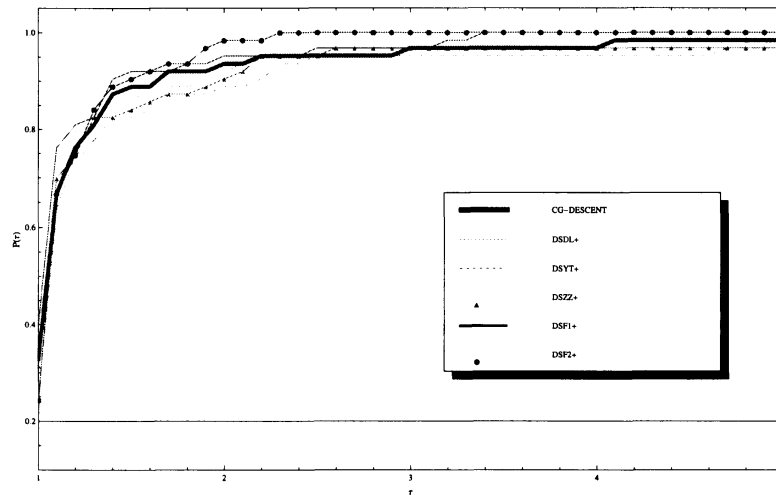


図 2: 関数評価回数を基準としたパフォーマンスプロファイル

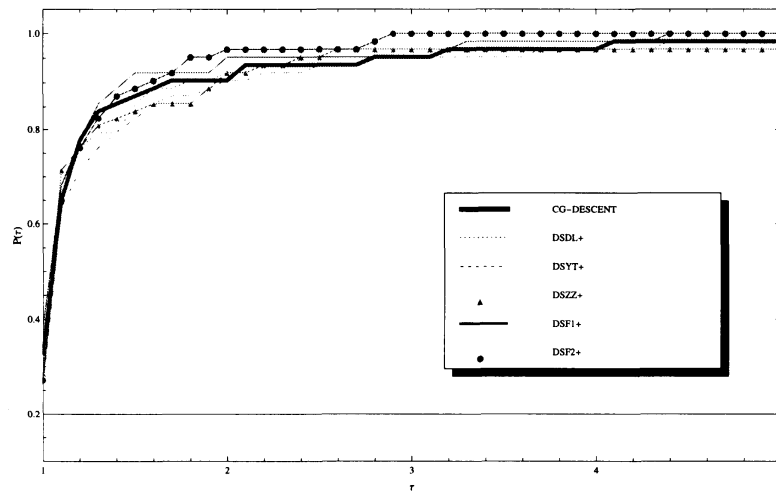


図 3: 勾配ベクトルの評価回数を基準としたパフォーマンスプロファイル

参考文献

- [1] I. Bongartz, A.R. Conn, N.I.M. Gould and P.L. Toint, CUTE: constrained and unconstrained testing environment, *ACM Transactions on Mathematical Software*, **21** (1995) 123–160.
- [2] Y.H. Dai and L.Z. Liao, New conjugacy conditions and related nonlinear conjugate gradient methods, *Applied Mathematics and Optimization*, **43** (2001) 87–101.
- [3] E.D. Dolan and J.J. Moré, Benchmarking optimization software with performance profiles, *Mathematical Programming*, **91** (2002) 201–213.
- [4] J.A. Ford and I.A. Moghrabi, Alternative parameter choices for multi-step quasi-Newton methods, *Optimization Methods and Software*, **2** (1993) 357–370.
- [5] J.A. Ford, Y. Narushima and H. Yabe, Multi-step nonlinear conjugate gradient methods for unconstrained minimization, *Computational Optimization and Applications*, **40** (2008) 191–216.
- [6] J.C. Gilbert and J. Nocedal, Global convergence properties of conjugate gradient methods for optimization, *SIAM Journal on Optimization*, **2** (1992) 21–42.
- [7] N.I.M. Gould, D. Orban and P.L. Toint, CUTER web site, <http://www.cuter.rl.ac.uk/>.
- [8] 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.
- [9] W.W. Hager and H. Zhang, CG_DESCENT Version 1.4 User' Guide, University of Florida, November 2005, <http://www.math.ufl.edu/~hager/>.
- [10] 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.
- [11] W.W. Hager and H. Zhang, A survey of nonlinear conjugate gradient methods, *Pacific Journal of Optimization*, **2** (2006) 35–58.
- [12] M. Kobayashi, Y. Narushima and H. Yabe, Nonlinear conjugate gradient methods with structured secant condition for nonlinear least squares problems, *Journal of Computational and Applied Mathematics*, **234** (2010) 375–397.
- [13] D.H. Li and M. Fukushima, A modified BFGS method and its global convergence in nonconvex minimization, *Journal of Computational and Applied Mathematics*, **129** (2001) 15–35.

- [14] Y. Narushima, H. Yabe and J.A. Ford, A three-term conjugate gradient method with sufficient descent property for unconstrained optimization, *SIAM Journal on Optimization*, **21** (2011) 212-230.
- [15] A. Perry, A modified conjugate gradient algorithm, *Operations Research*, **26** (1978) 1073-1078.
- [16] K. Sugiki, Y. Narushima and H. Yabe, Globally convergent three-term conjugate gradient methods that use secant conditions and generate descent search directions for unconstrained optimization, *Journal of Optimization Theory and Applications*, to appear.
- [17] H. Yabe and M. Takano, Global convergence properties of nonlinear conjugate gradient methods with modified secant condition, *Computational Optimization and Applications*, **28** (2004) 203-225.
- [18] G. Yu, L. Guan and G. Li, Global convergence of modified Polak-Ribière-Polyak conjugate gradient methods with sufficient descent property, *Journal of Industrial and Management Optimization*, **4** (2008) 565-579
- [19] G. Yuan, Modified nonlinear conjugate gradient methods with sufficient descent property for large-scale optimization problems, *Optimization Letters*, **3** (2009) 11-21.
- [20] J.Z. Zhang, N.Y. Deng, and L.H. Chen, New quasi-Newton equation and related methods for unconstrained optimization, *Journal of Optimization Theory and Applications*, **102** (1999) 147-167.
- [21] 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.
- [22] L. Zhang, W. Zhou and D.H. Li, A descent modified Polak-Ribière-Polyak conjugate gradient method and its global convergence, *IMA Journal of Numerical Analysis*, **26** (2006), 629-640.
- [23] 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.
- [24] W. Zhou and L. Zhang, A nonlinear conjugate gradient method based on the MBFGS secant condition, *Optimization Methods and Software*, **21** (2006) 707-714.