日本オペレーションズ・リサーチ学会
2021年 秋季研究発表会

メモリーレス準ニュートン法に基づく 非厳密ニュートン型近接DCアルゴリズム

05000392 中央大学 *中山 舜民 NAKAYAMA Shummin
01406032 慶應義塾大学 成島 康史 NARUSHIMA Yasushi
01702330 東京理科大学 矢部 博 YABE Hiroshi

1. はじめに

本稿では，次の最適化問題を扱う：

$$\min_{x\in\mathbb{R}^n} f(x)=g(x)+h(x). \tag{1}$$

ただし，g は L-平滑な関数とし，h は二つの凸関数 h_1 と h_2 の差で表される DC 関数 (Difference of two Convex functions)：

$$h(x)=h_1(x)-h_2(x)$$

であるとする．以降，∇g は g の勾配，∂h は h の劣微分とする．また，凸関数 h_1 に対して正定値対称行列 A を重み行列とした近接写像を

$$\mathrm{Prox}_{h_1}^{A}(\bar{x})=\operatorname*{argmin}_{x\in\mathbb{R}^n}\left\{h_1(x)+\frac{1}{2}\|x-\bar{x}\|_A^2\right\} \tag{2}$$

で定義する．ただし，$\|x\|_A\equiv\sqrt{x^TAx}$ は重み付きノルムである．ここで，$I\in\mathbb{R}^{n\times n}$ を単位行列とし，$A=I$ のときは重み付き近接写像と重み付きノルムの A の表記を省略することとする．

問題 (1) を解くための方法として近接 DCA(DC Algorithm) が知られている．近接 DCA は，任意の初期点 $x_0\in\mathbb{R}^n$ に対し，k 回目の反復点 x_k での h_2 の劣勾配 $\xi_k\in\partial h_2(x_k)$ を用いて，

$$\begin{aligned} x_{k+1}&=\operatorname*{argmin}_{x\in\mathbb{R}^n}\Big\{h_1(x)+g(x_k)-h_2(x_k)\\ &\qquad+(\nabla g(x_k)-\xi_k)^T(x-x_k)+\frac{L}{2}\|x-x_k\|^2\Big\}\\ &=\mathrm{Prox}_{\frac{1}{L}h_1}\left(x_k-\frac{1}{L}(\nabla g(x_k)-\xi_k)\right) \end{aligned} \tag{3}$$

によって点列を生成する方法である．$h_2(x)=0$ ($\xi_k=0$) の場合，近接 DCA は通常の近接勾配法と一致することから，近接勾配法を拡張した方法となっている．また，近接 DCA は以下で定義される臨界点に収束することが知られている．

定義 1. 以下の条件を満たす点 x^* を臨界点と呼ぶ．

$$0\in\nabla g(x^*)+\partial h_1(x^*)-\partial h_2(x^*).$$

Nesterov の加速法を加えた近接勾配法と同様に，近接 DCA においても外挿付き近接 DCA [1] が有効な方法として知られている．また，近接勾配法において，Nesterov の加速法とは別のアプローチとしてニュートン型近接勾配法が知られている．通常，ニュートン型近接勾配法は重み付き近接写像 (2) の計算に手間がかかるため，Nakayama et al. [2] は非厳密に重み付き近接写像を計算する非厳密ニュートン型近接勾配法を考え，重み行列としてメモリーレス準ニュートン法を取り入れた非厳密ニュートン型近接勾配法を提案した．本研究では，非厳密ニュートン型近接勾配法を拡張した非厳密ニュートン型近接 DCA の提案と，重み行列としてメモリーレス準ニュートン法の更新式を使用した方法を提案する．

2. 提案手法

2.1. 非厳密ニュートン型近接 DCA

提案手法の反復式は，以下で与えられる：

$$x_{k+1}=x_k+\eta_k d_k,\ d_k=x_k^+-x_k. \tag{4}$$

ただし，$\eta_k>0$ はステップ幅であり，探索方向に含まれている x_k^+ は以下の最適化問題の非厳密な解とする：

$$\begin{aligned} &\operatorname*{argmin}_{x\in\mathbb{R}^n}\Big\{h_1(x)+g(x_k)-h_2(x_k)\\ &\qquad+(\nabla g(x_k)-\xi_k)^T(x-x_k)+\frac{1}{2}\|x-x_k\|_{B_k}^2\Big\}\\ &=\mathrm{Prox}_{h_1}^{B_k}(x_k-H_k(\nabla g(x_k)-\xi_k)). \end{aligned} \tag{5}$$

ここで，B_k はヘッセ行列 $\nabla^2 g(x_k)$ の正定値対称な近似行列であり[1]，$H_k=B_k^{-1}$ とする．また，(5) を非厳密に計算する基準としてパラメータ $\theta_k\in[\bar{\theta},1]$ を与え，

$$\|r_k\|_{H_k}\le(1-\theta_k)\|x_k^+-x_k\|_{B_k}, \tag{6}$$

$$r_k\in\nabla g(x_k)-\xi_k+B_k(x_k^+-x_k)+\partial h_1(x_k^+) \tag{7}$$

を満たす x_k^+ を (5) の近似解として採用する．ここで r_k は勾配残差である．

以上をまとめたアルゴリズムは以下の通りである．

[1] $B_k=L\times I,\ \eta_k=1$ のとき反復式 (4)–(5) は近接 DCA (3) の反復式と一致する．

アルゴリズム 1 (非厳密ニュートン型近接 DCA).

Step 0: 初期点 $x_0 \in \mathbb{R}^n$, $\delta,\ \beta \in (0,1)$, $\bar{\theta} \in (0,1]$ を与え，$k=0$ とする．

Step 1: 終了条件を満たすならばアルゴリズムを停止して x_k を最適解とする．

Step 2: B_k と H_k，$\xi_k \in \partial h_2(x_k)$, $\theta_k \in [\bar{\theta},1]$ を与え，(6)–(7) を満たす $x_k^+ \approx \mathrm{Prox}_{h_1}^{B_k}(x_k - H_k(\nabla g(x_k) - \xi_k))$ を求め，探索方向を $d_k = x_k^+ - x_k$ とする．

Step 3: 以下の条件を満たす最小の非負整数 i を見つけ，ステップ幅を $\eta_k = \beta^i$ とする：

$$f(x_k + \beta^i d_k) \le f(x_k) + \delta \beta^i \lambda_k,$$
$$\lambda_k = (\nabla g(x_k) - \xi_k)^T d_k + h_1(x_k^+) - h_1(x_k).$$

Step 4: 点 x_{k+1} を (4) によって更新する．

Step 5: $k = k+1$ として Step 1 に戻る． □

このアルゴリズムは $h_2(x) = 0$ ($\xi_k = 0$) である場合に [2] のアルゴリズムの枠組みと一致する．次に，収束定理のために目的関数とアルゴリズムにそれぞれ以下の仮定を設ける．

仮定 1. *1.* g は連続的微分可能であり，∇g はリプシッツ連続である．

2. h_1 は下半連続な真凸関数とし，h_2 は連続な凸関数とする．

仮定 2. すべての反復 k において B_k は一様正定値かつ有界である．

上記の仮定の下で以下の定理が成り立つ．

定理 1. 仮定 *1–2* を満たすとし，点列 $\{x_k\}$ はアルゴリズム *1* によって生成されるとする．このとき，目的関数 f が下に有界であれば

$$\lim_{k\to\infty} \|d_k\| = 0$$

が成り立つ．さらに，点列 $\{x_k\}$ が有界であれば点列 $\{x_k\}$ の任意の集積点は (1) の臨界点である．

2.2. メモリーレス BFGS 公式

提案手法では，アルゴリズム 1 の Step 2 の B_k として Nakayama et al. [2] の提案したメモリーレス BFGS 公式を用いる[2]：

$$B_k = I - \frac{s_{k-1}s_{k-1}^T}{s_{k-1}^T s_{k-1}} + \gamma_k \frac{z_{k-1}z_{k-1}^T}{s_{k-1}^T z_{k-1}}. \qquad (8)$$

ここで，$s_{k-1} = x_k - x_{k-1}$，$z_{k-1} = \nabla g(x_k) - \nabla g(x_{k-1}) + \nu_k s_{k-1}$ であり，$\gamma_k > 0$, $\nu_k \ge 0$ はパラメータである．パラメータ γ_k と ν_k を適切に選択することで，仮定 2 が成立することを注意しておく．

次に，アルゴリズム 1 の k 回目の反復における Step 2 で x_k^+ を求める方法を考える．以降では $\bar{x} = x_k - H_k(\nabla g(x_k) - \xi_k)$,

$$u_1 = \sqrt{\frac{\gamma_k}{s_{k-1}^T z_{k-1}}}\, z_{k-1}, \quad u_2 = \frac{s_{k-1}}{\|s_{k-1}\|}$$

と置き，$\alpha = (\alpha_1, \alpha_2)^T \in \mathbb{R}^2$ に対して，$\zeta(\alpha) = \bar{x} - \alpha_1 u_1 + \alpha_2 (I + u_1 u_1^T)^{-1} u_2$ とし，α^* を連立非線形方程式

$$u_1^T(\bar{x} + \alpha_2 (I + u_1 u_1^T)^{-1} u_2 - \mathrm{Prox}_{h_1}(\zeta(\alpha))) + \alpha_1 = 0$$
$$u_2^T(\bar{x} - \mathrm{Prox}_{h_1}(\zeta(\alpha))) + \alpha_2 = 0$$

の唯一解であるとする．このとき，

$$\mathrm{Prox}_{h_1}^{B_k}(\bar{x}) = \mathrm{Prox}_{h_1}(\zeta(\alpha^*))$$

が成立する [3]．よって，一般化ニュートン法を用いて上述の 2 次元の連立非線形方程式を解くことで x_k^+ を容易に計算できる．

3. 数値実験

提案手法（公式 (8) を用いたアルゴリズム 1 ）と外挿付き近接 DCA [1] の比較を行い，提案手法の有効性を検証する．数値実験の結果は当日発表する．

参考文献

[1] B. Wen, X. Chen and T.K. Pong, A proximal difference-of-convex algorithm with extrapolation, *Computational Optimization and Applications*, **69** (2018), 297–324.

[2] S. Nakayama, Y. Narushima and H. Yabe Inexact proximal memoryless quasi-Newton methods based on the Broyden family for minimizing composite functions, *Computational Optimization and Applications*, **79** (2021), 127–154.

[3] 中山舜民，成島康史，矢部博，“メモリーレス BFGS 公式に基づく非厳密ニュートン型近接勾配法における内部反復の改良について”，日本オペレーションズ・リサーチ学会 2021 年春季研究発表会アブストラクト集，1-A-9.

[2] $H_k = B_k^{-1}$ も B_k^{-1} を計算するのではなく，更新公式から計算可能だが，紙面の都合上，その更新公式は省略する．