サイジング効果付き記憶制限準ニュートン法

02401973 静岡大学工学部 ＊根岸建彦 NEGISHI Tatsuhiko
01792180 静岡大学工学部 八巻直一 YAMAKI Naokazu
01702330 東京理科大学理学部 矢部 博 YABE Hiroshi

1 概要

非線形関数の最小化のための数値解法として, 最急降下法の「大域的収束性」とニュートン法の「局所的に速い収束性」というそれぞれの長所をあわせ持つ準ニュートン法 (quasi-Newton method) がよく知られている. 準ニュートン法は無制約問題に対する数値解法の中では, 現在もっとも有力な方法の１つである. 近年, 準ニュートン法を大規模な問題に適用できるような工夫として, 記憶制限準ニュートン法が注目されている. 本研究の目的は, 記憶制限準ニュートン法にたいして, サイジングが自然に適用される更新公式をつくることである.

2 非線形計画問題

対象とする非線形計画問題は以下の問題である.

$$\min f(x)$$

ただし,$x \in R^n$,$f : R^n \to R$ である.

本研究では, このような非線形計画問題の内,x の次元が大きい, 大規模な問題を考える.

3 準ニュートン法

準ニュートン法のアルゴリズムは, 以下のとおりである.

- **STEP0)** 初期点 x_1 を与え, $H_1 = I$ とする. $k = 1$ とおく.
- **STEP1)** 探索方向 d_k を, 次のように求める.

$$d_k = -H_k g_k, \quad (g_k = \nabla f(x_k))$$

- **STEP2)** 停止判定をする.
- **STEP3)** ステップ幅 α_k を決定する.
- **STEP4)** x を次のように更新する.

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

- **STEP5)** H_k を更新し, $k = k + 1$ として STEP1 へ

記憶制限準ニュートン法は, 上のアルゴリズムにおいて, H_k を行列として保持せず, ベクトル演算によって求めるものである.

3.1 セカント条件の拡張

$$s_k = x_{k+1} - x_k, \quad y_k = g_{k+1} - g_k$$

とすると, セカント条件とは,

$$s_k = H_{k+1} y_k$$

である. ここでは, セカント条件を拡張して,

$$\begin{aligned} S_{k+1} &= [s_1, s_2, \cdots, s_k], \\ Y_{k+1} &= [y_1, y_2, \cdots, y_k] \end{aligned}$$

と定め,

$$S_{k+1} = H_{k+1} Y_{k+1}$$

とする. 記憶制限準ニュートン法は, 過去 t ステップの情報を記憶する. すなわち,

$$s_{k-t}, s_{k-t+1}, \cdots, s_{k-1}$$

および

$$y_{k-t}, y_{k-t+1}, \cdots, y_{k-1}$$

を保持する. したがって, 過去 t ステップに対して拡張されたセカント条件を考慮することが出来る.

3.2 拡張セカント条件を満たす公式

Yamaki and Yabe[3],Yabe and Yamaki[4] は, 近似行列 H_{k+1} の生成において, 拡張セカント条件を満たす公式を提案している. 公式は以下のとおりである.

© 日本オペレーションズ・リサーチ学会. 無断複写・複製・転載を禁ず.

$$
\begin{aligned}
H_k &= P_k + R_k \\
v_k &= P_k y_k - \frac{y_k^T P_k y_k}{y_k^T u_k} u_k \qquad (1) \\
u_k &= s_k - R_k y_k \\
R_k &= S_k (Y_k^T S_k)^{-1} S_k^T \\
P_{k+1} &= w_k [P_k - \frac{P_k y_k y_k^T P_k}{y_k^T P_k y_k} + \frac{1}{y_k^T P_k y_k} v_k v_k^T] \\
P_1 &= I, \quad R_1 = 0
\end{aligned}
$$

3.3 サイジング

Oren[1] は，目的関数を 2 次モデルと仮定した場合，近似行列 H_k の固有値を目的関数のヘッセ行列 $Q = \nabla^2 f(x_k)$ の逆行列の固有値に単調に近づけることを提案している．このことは，

$$\tilde{H}_k = Q^{\frac{1}{2}} H_k Q^{-\frac{1}{2}}$$

とすると，次のスペクトル条件数 κ_{k+1} を単調に減少させることと同値である．

$$\kappa_k = \| \tilde{H}_k \| \cdot \| \tilde{H}_k^{-1} \|, \quad \kappa_k > \kappa_{k+1}$$

サイジングとは近似行列 H_k をそのまま更新するのではなく，適当な正の数 w_k を H_k にかけてから更新することである．w_k は $w_k H_k$ の固有値の分布がヘッセ行列 $\nabla^2 f(x_k)$ の固有値の分布に近づくように選ばれる．このことによって，計算効率を高めることが期待できる．Yabe and Yamaki[4] では，(1) に対する w_k の選び方として，

$$
\begin{aligned}
w_1 &= \frac{(1-\psi_1) s_1^T y_1}{y_1^T H_1 y_1} + \frac{\psi_1 s_1^T g_1}{g_1^T H_1 y_1}, \quad \psi_1 \in [0,1] \\
w_k &= \frac{\psi_k^1 s_k^T y_k}{y_k^T H_k y_k} + \frac{\psi_k^2 s_k^T g_k}{g_k^T H_k y_k} + \psi_k^3 \\
& \sum_{i=1}^{3} \psi_k^i = 1, \quad \psi_k^1, \psi_k^2, \psi_k^3 \geq 0, \quad \text{for} \quad k \geq 2
\end{aligned}
$$

を提案している．したがって，$k = 1$ の場合には $H_1 = I$ なので，簡単な計算によって w_1 が得られ，$k > 1$ の場合は $w_k = 1$ でよいことがわかる．

3.4 記憶制限付き準ニュートン法

Nocedal[2] は，通常の準ニュートン法では，H_k を保存するのに大量の記憶領域を必要とするのに対して，H_k を保存しないで探索方向 d_k を数本のベクトルの積形式で直接計算することによって，記憶領域を低減できることを示した．これを記憶制限準ニュートン法という．Nocedal と同様の変形をおこなうと，式 (1) は次のような積形式で表すことができる．

$$P_k = (I - \frac{y_{k-1} u_{k-1}^T}{y_{k-1}^T u_{k-1}})^T P_{k-1} (I - \frac{y_{k-1} u_{k-1}^T}{y_{k-1}^T u_{k-1}})$$

ここで，

$$Z_{k-1} = (I - \frac{y_{k-1} u_{k-1}^T}{y_{k-1}^T u_{k-1}})$$

とおくと，P_k は次のようにあらわされる．

$$
\begin{aligned}
P_k &= w_1 Z_{k-1}^T P_{k-1} Z_{k-1} \\
&= w_1 Z_{k-1}^T \cdots Z_1^T P_1 Z_1 \cdots Z_{k-1}
\end{aligned}
$$

このとき Z_{k-1} から Z_1 まで用いて計算するのではなく，s, u, y を t 個保存して，Z_{k-1} から Z_{k-t} まで用いて計算すれば，メモリは大幅に縮減できる．すなわち，$P_{k-t} = I$ とし，

$$P_k = w_1 Z_{k-1}^T \cdots Z_{k-t}^T Z_{k-t} \cdots Z_{k-1}$$

のようにする．探索方向 d_k は

$$
\begin{aligned}
d_k &= -H_k g_k \\
&= -(P_k + R_k) g_k \\
&= -w_1 Z_{k-1}^T \cdots Z_{k-t}^T Z_{k-t} \cdots Z_{k-1} g_k \\
& \quad -R_k g_k
\end{aligned}
$$

である．

上の式において，右辺の各項は g_k を右から掛けることからはじめると，順次 $t \times t$ の行列ないし，n 次元ベクトルが生成されるだけで計算できる．したがって，n が大きい場合，準ニュートン法が $n \times n$ の行列を記憶しなければならないのに比べて，格段に記憶容量が節減できる．

参考文献

[1] Oren,S.S.,Self-scaling variable metric(SSVM) algorithms, Part2; Management Science,20,863-874(1974)

[2] Nocedal,J,Updating quasi-Newton matrices with limited storage, Mathematics of Computation,35.773-782(1980)

[3] Yamaki,N and Yabe,H, A family of the quasi-Newton methods, TRU Mathematics,16-1,49-54(1980)

[4] Yabe,H and Yamaki,N, Some properties of Oren's SSVM type algorithm for unconstrained minimization, TRU Mathematics 16-2,103-111(1980)

© 日本オペレーションズ・リサーチ学会．無断複写・複製・転載を禁ず．