数理解析研究所講究録
748巻 1991年 43-52

Max型関数の最適化アルゴリズムによる非線形方程式の解法

古川長太（九大 理）

§1 序

非線形方程式の解法に従来よく使われるニュートン法，乃至修正ニュートン法は，方程式の左辺が微分不可能項をもつ場合には適さない。本報告では微分不可能項をもつ方程式としてある意味で代表的であり，応用上も意味があると思われるMax型関数で記述される非線形方程式をとり上げ，それに対する解法並びにアルゴリズムについて紹介する。ここで言うMax型関数とは，複数個のC^1級関数の点ごとに最大値をとった関数のことである。この様な関数は通常の意味では微分可能ではないが，至る所で片側方向微分可能であって，その片側方向微分はMax型関数を構成するもとのC^1級関数の微分（勾配ベクトル）を使って表される。このことに着目して，Max型関数の片側方向微分を全面的に活用して（ということは，もとのC^1級関数の勾配ベクトルを全面的に活用して）解を求めるアルゴリズムを組み立てる。

直お，本報告で紹介するアルゴリズムには筆者が既に

開発したMax型関数の最大化アルゴリズムが重要な基礎を担っているが，それについての説明は割愛する。詳細は末尾にあげた文献に書かれている。

§2　定式化

定義　1

f_i $(i=1,2,\cdots,\ell)$ を $\mathbb{R}^n$ 上で定義された C^1 級実数値関数とする。関数 F を

$$F(x)=\max_{1\leq i\leq \ell} f_i(x)\ ,\quad x\in\mathbb{R}^n \tag{2.1}$$

で定義し，これをMax型関数と呼ぶ。

定義　2

(2.1) で与えられた関数 F に対して，点 x ごとに index の集合 $A(x)$ を次で定義する。即ち $x\in\mathbb{R}^n$ に対し

$$A(x)=\{i\in\{1,2,\cdots,\ell\}\mid F(x)=f_i(x)\}. \tag{2.2}$$

定義　3

f を $\mathbb{R}^n$ 上で定義された実数値関数とする。$x,\ d\in\mathbb{R}^n$ に対して極限

$$f'(x;d)=\lim_{\lambda\downarrow 0}\frac{f(x+\lambda d)-f(x)}{\lambda}$$

が有限確定するとき，f は点 x において d 方向に片側方向微分可能であるといい，このとき $f'(x\,;d)$ を f の x における d 方向の片側方向微分と呼ぶ。

$\mathbb{R}^n$ 上で定義された微分可能な関数 f に対して，f の点 x における勾配ベクトルを $\nabla f(x)$ と書く。ただし勾配ベクトルは行ベクトルとする。一般に $\mathbb{R}^n$ におけるベクトルは特に断りがなければ列ベクトルとする。

<u>補題 1</u>

F を (2.1) で表されるMax型関数とする。このとき F は至る所でどの方向にも片側方向微分可能で，F の点 x における d 方向の片側方向微分は

$$F'(x\,;d) = \max_{i \in A(x)} \nabla f_i(x)\, d \tag{2.3}$$

で与えられる。ここに $A(x)$ は (2.2) で定義された index の集合である。

次にここで取り扱う非線形連立程式を定式化する。$F_1, F_2, \cdots, F_m$ を定義 1で述べた意味の $\mathbb{R}^n$ 上のMax型関数とする。即ち f_{ij}，$i=1,2,\cdots,m$，$j=1,2,\cdots,p_i$，をすべて $\mathbb{R}^n$ 上の C^1 級実数値関数とし，各 i $(1 \leq i \leq m)$ につき F_i は

$$F_i(x) = \max_{1\leq j\leq p_i} f_{ij}(x), \qquad x\in\mathbb{R}^n$$

で与えられるものとする。

本報告で取り扱う非線形連立方程式は

$$F_i(x) = 0, \quad i=1,2,\cdots,m \qquad (2.4)$$

である。次のセクションでその解を求めるためのアルゴリズムを与える。

§3 アルゴリズム

非線形連立方程式(2.4)を解くためのアルゴリズムは，初期の段階から現在に至る迄数回の改良を行い，Max型関数の直線探索に特有の桁落ちを可能な限り回避する手当てを補す迄に至っている。それらをすべ記述すると大変に長くなり，また本質が見えにくくもなるので，本稿ではそのアルゴリズムを要約したものを以下に紹介する。

アルゴリズムの中に現れる β, μ, ε, τ はすべてその手続きに関わるパラメータであって，そのうち ε は停止規則を決定するものである。

アルゴリズム

step 0 : $\beta\in(0,1)$, $\mu\in(0,1)$, $\varepsilon\in(0,1)$, $0<\tau\leq+\infty$

m, n ; 正整数

$p_1, p_2, \cdots, p_m$; 正整数

$\bar{x}^0 \in \mathbb{R}^n$, $\nu = 0$.

step 1 : 各 i $(1 \leq i \leq m)$ につき $F_i(\bar{x}^\nu) = \max_{1 \leq j \leq p_i} f_{ij}(\bar{x}^\nu)$ を計算し,

$$\Gamma^\nu = \{ i \in \{1, 2, \cdots, m\} \mid F_i(\bar{x}^\nu) < -\varepsilon \}$$

$$\Lambda^\nu = \{ i \in \{1, 2, \cdots, m\} \mid F_i(\bar{x}^\nu) > \varepsilon \}$$

を求める.

$\Gamma^\nu \neq \Phi$ なら step 2 へ,

$\Gamma^\nu = \Phi$ かつ $\begin{cases} \Lambda^\nu = \Phi & \text{なら ストップ,} \\ \Lambda^\nu \neq \Phi & \text{なら step 11 へ.} \end{cases}$

step 2 : $x^0 = \bar{x}^\nu$, $k = 0$ とおいて step 3 へ.

step 3 : $F^\nu(x^k) = \max_{i \in \Gamma^\nu} F_i(x^k)$ を計算.

$A^\nu(x^k; \tau) = \{ i \in \Gamma^\nu \mid F^\nu(x^k) - F_i(x^k) \leq \tau \}$ を求める.

step 4 : 各 i $(1 \leq i \leq m)$ につき

$D_i(x^k) = \{ j \in \{1, 2, \cdots, p_i\} \mid F_i(x^k) = f_{ij}(x^k) \}$ を求め,

$$\| \nabla f_{i_k j_k}(x^k) \|^2 + f_{i_k j_k}(x^k) = \max_{i \in A^\nu(x^k; \tau)} \max_{j \in D_i(x^k)} \{ \| \nabla f_{ij}(x^k) \|^2 + f_{ij}(x^k) \}$$

をみたす $i_k \in A^\nu(x^k; \tau)$ と $j_k \in D_{i_k}(x^k)$ とを求める.

$\| \nabla f_{i_k j_k}(x^k) \| \leq \varepsilon$ ならストップ,

$\| \nabla f_{i_k j_k}(x^k) \| > \varepsilon$ なら $d^k = \nabla f_{i_k j_k}^T(x^k)$ とおいて step 5 へ.

step 5 : $A^{\nu}(x^k) = \{ i \in \Gamma^{\nu} \mid F^{\nu}(x^k) = F_i(x^k) \}$を求める.

step 6 : $i_k \in A^{\nu}(x^k)$ なら step 7 へ,

$i_k \notin A^{\nu}(x^k)$ かつ $\max_{i \in A^{\nu}(x^k)} \max_{j \in D_i(x^k)} \nabla f_{ij}(x^k) d^k > 0$

ならば step 8 へ,

$i_k \notin A^{\nu}(x^k)$ かつ $\max_{i \in A^{\nu}(x^k)} \max_{j \in D_i(x^k)} \nabla f_{ij}(x^k) d^k \leq 0$

ならば step 10 へ.

step 7 : $\varepsilon \geq F^{\nu}(x^k + (\beta)^l d^k) \geq F^{\nu}(x^k) + \mu(\beta)^l \| d^k \|^2$

をみたす最小の非負整数 $l = l(k)$ を求めて step 9 へ.

step 8 : $f_{i_k j_k}(x^k + (\beta)^s d^k) \geq f_{i_k j_k}(x^k) + \mu(\beta)^s \| d^k \|^2$

をみたす最小の非負整数 $s = s(k)$ を求めて,

$\varepsilon \geq F^{\nu}(x^k + (\beta)^t d^k) > F^{\nu}(x^k)$

をみたす最小の非負整数 $t = t(k)$ を求める.

$l(k) = \max(s(k), t(k))$ とおいて step 9 へ.

step 9 : $F^{\nu}(x^k + (\beta)^{l(k)} d^k) < -\varepsilon$ ならば

$x^{k+1} = x^k + (\beta)^{l(k)} d^k$, $k = k+1$ とおいて step 3 へ,

$F^{\nu}(x^k + (\beta)^{l(k)} d^k) \geq -\varepsilon$ ならば

$\bar{x}^{\nu+1} = x^k + (\beta)^{l(k)} d^k$, $\nu = \nu + 1$ とおいて step 1 へ.

step 10 : $\| \nabla f_{I_k J_k}(x^k) \|^2 = \max_{i \in A^{\nu}(x^k)} \max_{j \in D_i(x^k)} \| \nabla f_{ij}(x^k) \|^2$

をみたす $I_k \in A^{\nu}(x^k)$ と $J_k \in D_{I_k}(x^k)$ を求め

$\| \nabla f_{I_k J_k}(x^k) \| \leq \varepsilon$ ならストップ,

$\| \nabla f_{I_k J_k}(x^k) \| > \varepsilon$ なら $d^k = \nabla f_{I_k J_k}^{T}(x^k)$ とおいて step 7 へ.

step 11 : 各 $i\ (1 \le i \le m)$ につき $F_i(z^\nu) = \max_{1 \le j \le p_i} f_{ij}(z^\nu)$ を計算し

$\Lambda^\nu = \{ i \in \{1, 2, \cdots, m\} \mid F_i(z^\nu) > \varepsilon \}$,

$\Gamma^\nu = \{ i \in \{1, 2, \cdots, m\} \mid F_i(z^\nu) < -\varepsilon \}$

を求める.

$\Lambda^\nu \ne \Phi$ なら step 12 へ,

$\Lambda^\nu = \Phi$ かつ $\begin{cases} \Gamma^\nu = \Phi \text{ ならストップ}。 \\ \Gamma^\nu \ne \Phi \text{ なら step 1 へ}. \end{cases}$

step 12 : $x^0 = z^\nu$, $k = 0$ とおいて step 13 へ.

step 13 : $G^\nu(x^k) = \min_{i \in \Lambda^\nu} F_i(x^k)$ を計算

$B^\nu(x^k; \zeta) = \{ i \in \Lambda^\nu \mid G^\nu(x^k) - F_i(x^k) \ge -\zeta \}$ を求める.

step 14 : 各 $i\ (1 \le i \le m)$ につき

$D_i(x^k) = \{ j \in \{1, 2, \cdots, p_i\} \mid F_i(x^k) = f_{ij}(x^k) \}$ を求め,

$$\| \nabla f_{i_k j_k}(x^k) \|^2 - f_{i_k j_k}(x^k) = \max_{i \in B^\nu(x^k; \zeta)} \max_{j \in D_i(x^k)} \left\{ \| \nabla f_{ij}(x^k) \|^2 - f_{ij}(x^k) \right\}$$

をみたす $i_k \in B^\nu(x^k; \zeta)$ と $j_k \in D_{i_k}(x^k)$ を求め

$\| \nabla f_{i_k j_k}(x^k) \| \le \varepsilon$ ならストップ。

$\| \nabla f_{i_k j_k}(x^k) \| > \varepsilon$ なら $d^k = -\nabla f^T_{i_k j_k}(x^k)$ とおいて step 15 へ.

step 15 : $B^\nu(x^k) = \{ i \in \Lambda^\nu \mid G^\nu(x^k) = F_i(x^k) \}$ を求める.

step 16: $i_k \in B^\nu(x^k)$ なら step 17 へ,

$i_k \notin B^\nu(x^k)$ かつ $\min\limits_{i \in B^\nu(x^k)} \max\limits_{j \in D_i(x^k)} \nabla f_{ij}(x^k) d^k < 0$

ならば step 18 へ,

$i_k \notin B^\nu(x^k)$ かつ $\min\limits_{i \in B^\nu(x^k)} \max\limits_{j \in D_i(x^k)} \nabla f_{ij}(x^k) d^k \geq 0$

ならば step 20 へ.

step 17: $-\varepsilon \leq G^\nu(x^k + (\beta)^l d^k) \leq G^\nu(x^k) - \mu (\beta)^l \| d^k \|^2$

をみたす最小の非負整数 $l = l(k)$ を求めて step 19 へ.

step 18: $f_{i_k j_k}(x^k + (\beta)^s d^k) \leq f_{i_k j_k}(x^k) - \mu(\beta)^s \| d^k \|^2$

をみたす最小の非負整数 $s = s(k)$ と

$-\varepsilon \leq G^\nu(x^k + (\beta)^t d^k) < G^\nu(x^k)$

をみたす最小の非負整数 $t = t(k)$ を求める.

$l(k) = \max(s(k), t(k))$ とおいて step 19 へ.

step 19: $G^\nu(x^k + (\beta)^{l(k)} d^k) > \varepsilon$ ならば

$x^{k+1} = x^k + (\beta)^{l(k)} d^k$, $k = k+1$ とおいて step 13 へ,

$G^\nu(x^k + (\beta)^{l(k)} d^k) \leq \varepsilon$ ならば

$z^{\nu+1} = x^k + (\beta)^{l(k)} d^k$, $\nu = \nu+1$ とおいて step 11 へ.

step 20: $\| \nabla f_{I_k J_k}(x^k) \|^2 = \max\limits_{i \in B^\nu(x^k)} \max\limits_{j \in D_i(x^k)} \| \nabla f_{ij}(x^k) \|^2$

をみたす $I_k \in B^\nu(x^k)$ と $J_k \in D_{I_k}(x^k)$ を求め,

$\| \nabla f_{I_k J_k}(x^k) \| \leq \varepsilon$ ならストップ,

$\| \nabla f_{I_k J_k}(x^k) \| > \varepsilon$ なら $d^k = -\nabla f_{I_k J_k}^T(x^k)$ とおいて

step 17 へ.

§4 補足と今後の問題点

上述のアルゴリズムで近似解を求めることが出来る，非線形方程式の代表的なタイプとして，次のものが有る。

f_i, g_i $(i=1,2,\cdots,m)$ を $\mathbb{R}^n$ 上で定義された C^1 級実数値関数とする。このとき，連立方程式

$$|f_i(x)| + g_i(x) = 0, \quad i=1,2,\cdots,m, \qquad (4.1)$$

は §2 の (2.4) の型に帰着できる。何となれば

$$F_i(x) = \max\{f_i(x)+g_i(x),\ -f_i(x)+g_i(x)\},\ x\in\mathbb{R}^n$$
$$i=1,2,\cdots,m$$

で F_i $(i=1,2,\cdots,m)$ を定義すると，(4.1) は

$$F_i(x) = 0, \quad i=1,2,\cdots,m$$

と同値であり，かつ各 F_i は §2 で述べた意味の Max 型関数である。従って (4.1) の近似解は §3 のアルゴリズムで求めることが出来る。実際の数値例については，本稿では割愛する。

おわりに，アルゴリズムの step 11 以下については，部分的に Max 型関数の最小化アルゴリズムを挿入した方が効率が良いと思われる個所にも，Max 型関数の最大化の原理を適用している所が有る。この点については数値実験の結果と照合して，現在アルゴリズムを改良中である。

参考文献

[1] 古川長太　Max型関数の最適化アルゴリズムとその応用, 応用数学合同研究集会報告集, (1989).