

多変数関数の極値を求める 2~3 の方法の比較

— 制約条件のない場合 —

戸田英雄 (ETL), 高山文雄 (ETL)

高沢嘉光 (東大工)

1. まえぶき

多変数関数の極値を求める問題 (あるいは連立の非線型方程式を解く問題) のアプローチは, よく知られているが, 二つの段階に分けられる:

オ1 段階は, オ1 近似を概測すること

オ2 段階は, オ1 近似の解の精度を更によくすること

でオ1 段の一般的な解法はない. オ2 段は一般的な方法が種々あつて, すでに S.S.P. などを使用できる.

ここでは, 非常に簡単なテスト用の問題と, 筆者が最近試みたある連立型の非線型方程式 (Belaga, Pan [4] による多変数の新しい計算^式を作るための方程式) を解くことについて, 2~3 の極値を求める方法の比較を述べる. いづれも制約条件のない場合である.

2. 極値を求める様々な方法 (制約条件のない場合)

オ2次の解法で試みた方法を述べる。テスト用の所与な問題として次のものを同いた。

$$F(x_1, x_2) = (6x_1 + 6x_2 - 5)^4 + (6x_1 - 6x_2 - 1)^2 + (2x_1 - 1)^2 \cdot (3x_2 - 1)^2$$

の極値を与える (x_1, x_2) を求めよ。ただし $-1 < x_1, x_2 < 1$ の範囲である。

$F(x_1, x_2)$ の曲図を表す図は高沢がカーブプロッタで描いたものである。

(角牛法-1) 座標軸 x_h に平行にある規則で降下していく方法

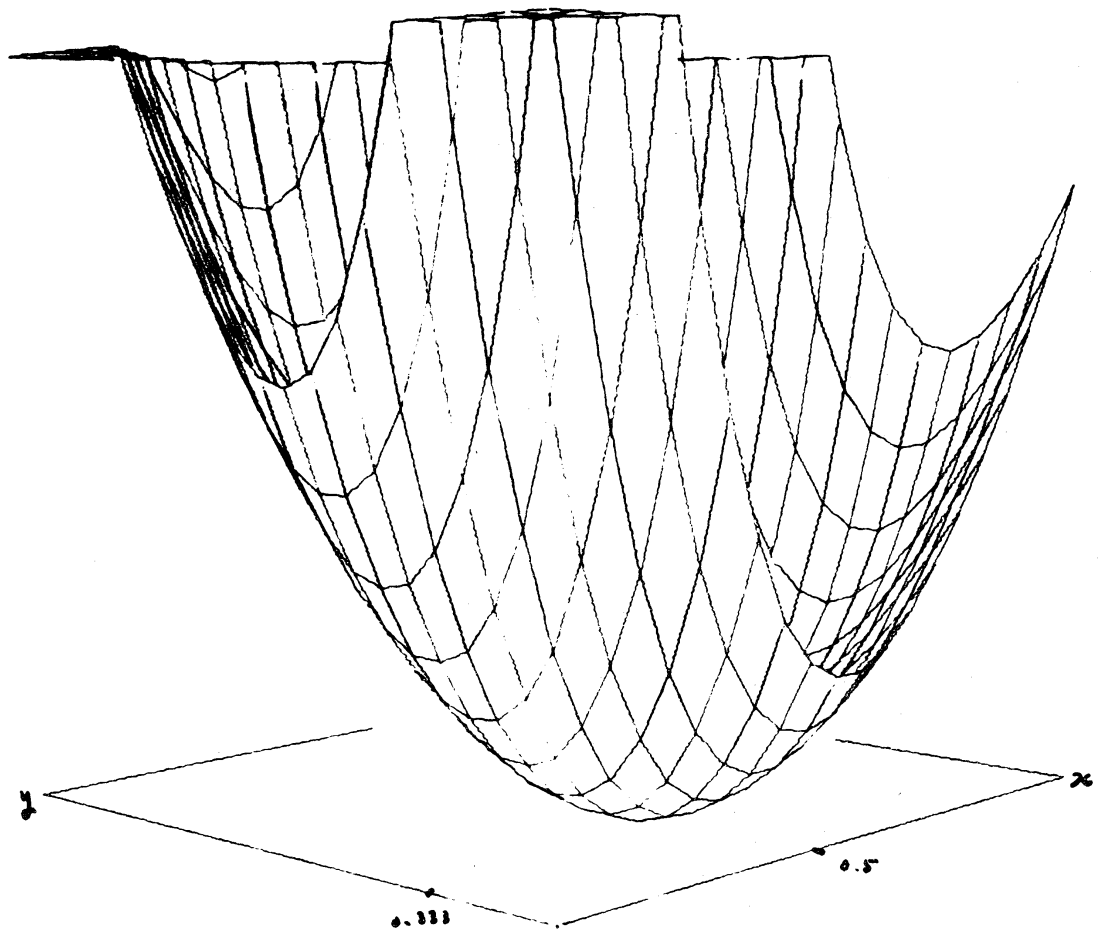
[手順] i) 収束判定用の EPS と刻み巾 h を決める。オ1近似 $(x_1^{(1)}, x_2^{(1)})$ を与えて $F(x_1^{(1)}, x_2^{(1)}) = F_{min}$ とする。

ii) x_1 軸上を刻み巾 h で $x_1 := x_1 + h$ を繰返して F_{min} の値を求める。

iii) x_2 軸上で同様に $x_2 := x_2 + h$ を繰返して F_{min} の値を求める。

iv) ii) と iii) を繰返して F_{min} の値が動かなくなったら $h := h/2$ とする。

v) $|h| \leq EPS$ ならばやめる。 $|h| > EPS$ のときは ii) に戻る。



(解法-2) 3^n 型実験計画法

山登り法とは通稱で(42率の山に登る話), Box-Wilsonが
 忘答曲面の探検に利用したものである。これはいわゆる最適
 条件を求めるための逐次実験計画法で, 品質管理(1956) PP.431
 ~435に森口氏がこの解説と批判を行い, 最後の一つの案と
 与えている。それによると, 最初から 3^n 型で2次曲面をあて
 はめその中心に向かって進む方法である。

$n=2$ の場合では, ある真のまわりに, 3^2-1 回の観測値
 とおいて, 2次曲面を

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11}(3x_1^2-2) + \beta_{22}(3x_2^2-2) + \beta_{12}x_1x_2$$

であてはめ, $\beta_0, \beta_1, \dots, \beta_{12}$ を最小二乗法のやり方で推定す
 る。データの構造が直交性のある計画となっているので,
 正規方程式は簡単に与えて, 推定された2次曲面の中心は,

$$\begin{bmatrix} 6 & \beta_{11} & \beta_{12} \\ \beta_{12} & 6 & \beta_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -\beta_1 \\ -\beta_2 \end{bmatrix}$$

をいって $\{x_1, x_2\}$ から求められる。はじめ, ある真を
 $\{c_1, c_2\}$ と任意の中心とすれば, 次の中心は

$$c_1 + x_1 \rightarrow c_1, \quad c_2 + x_2 \rightarrow c_2$$

となる。 $n \rightarrow \infty$ として繰返す。

(解法-3) Newton-Raphson 法

関数 $F(X)$ が連続微分可能で、極値 X_{min} のオ n 近似 $X^{(n)}$ とする。

$$X_{min} = X^{(n)} + \delta \quad (3-1)$$

$$\left. \frac{\partial}{\partial x_i} F(X) \right|_{X=X_{min}} = 0 \quad (3-2)$$

(3-2) を $X^{(n)}$ のまわりに T - δ 展開し δ^2 のオ- δ を省略すると、

$$\left. \frac{\partial}{\partial x_i} F(X) \right|_{X=X^{(n)}} + \sum_{j=1}^2 \left. \frac{\partial^2}{\partial x_i \partial x_j} F(X) \right|_{X=X^{(n)}} \delta_j \doteq 0 \quad (3-3)$$

$$\text{ここを、} \quad g_i \equiv \left. \frac{\partial}{\partial x_i} F(X) \right|_{X=X^{(n)}}, \quad G_{ij} \equiv \left. \frac{\partial^2}{\partial x_i \partial x_j} F(X) \right|_{X=X^{(n)}}$$

$$g \equiv [g_i], \quad G \equiv [G_{ij}]$$

と置く。 G は正の定値とすると、

$$(3-2) \quad \text{は} \quad g(X_{min}) = 0$$

$$(3-3) \quad \text{は} \quad g(X^{(n)}) \equiv g^{(n)} = G(X^{(n)} - X_{min})$$

となる。 X : 2 Newton-Raphson の反復法は、

$$X^{(n+1)} = X^{(n)} - G^{-1} \cdot g^{(n)} \quad (3-4)$$

となる。ここを $|G| \neq 0$ のとき G^{-1} とどうにかして求めるかによって、次のような解法に分れる。

(解法 3-1) G から G^{-1} を直接求める法

(解法 3-2) Fletcher, Powell (1963) の方法 (FMFP法) [1]

(解法 3-3) CG法 (FMC法) [2]

簡単なテスト用の問題ごとの比較表

	内 容	関数値	1階 の関数	2階の 関数	計 算 行 数	
解法 1	座標軸 x_k に平行に降下	要	不要	不要	・499 2	・332 5
解法 2	3^n 型実験計画法	要	不要	不要	・500 001	・333 342
解法 3-1	Newton 法 (直接法)	要	要	要	・499 996	・333 329
3-2	" (FMFP 法)	要	要	不要	・500 001	・333 335
3-3	" (FMCG 法)	要	要	不要	・500 260	・333 593

上記のような簡単な場合 ($n=2$ で極値は 1 つしかない, 極値のある範囲がわかっている) にはどうやって最適解は得られるか。各解法の特徴を次にあげる。

解法 1	関数値だけが必要。プログラムの一番単純。
解法 2	関数値だけが必要。1 に比べて精度はいいがプログラムの少しめんどう。
解法 3-1	正攻法。2 階の導関数まで必要。FORMAC 等の数式処理言語と併用すると便利と思われる。
3-2	SSP (IBM 360 用) として用いられている。
3-3	(3-2) より精度で少し劣る場合がある。

3. ある非線型の連立方程式を解く場合の比較表

筆者は最近次のような方程式を解く必要があって、11 近似の精度をよくするために、解法 3-1, 3-2, 3-3 を比較した結果を示す。

問題: Belaga, Pan [4] に与る多項式計算の新しいアルゴリズムを求める問題である;

$$\text{多項式 } p(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$$

で、 $a_n, a_{n-1}, \dots, a_1, a_0$; x から $p(x)$ を計算するには $2n-1$ 回のかけ算と n 回のたし算が必要となる。

$$p(x) = ((\dots ((a_n x + a_{n-1})x + a_{n-2}) \dots) + a_1)x + a_0$$

でやれば n 回のかけ算と n 回のたし算で出来る。しかしかけ算の回数は、どの位まで減らすことが出来るか? 多くの人間により研究されている。その中で Belaga, Pan のアルゴリズムとして次のものがある。

$$p_0 = x^2, \quad p'_0 = p_0 + x, \quad p_1 = x + \lambda_1$$

$$p_4^{(\Delta)} = (p'_0 + \lambda_{4\Delta-2}) \cdot (p_0 + \lambda_{4\Delta-1}) + \lambda_{4\Delta} \quad \left. \vphantom{p_4^{(\Delta)}} \right\} (\Delta = 1, 2, \dots, k)$$

$$p_{4\Delta+1} = p_{4\Delta-3} \cdot p_4^{(\Delta)} + \lambda_{4\Delta+1}$$

(3-1)

$$p_{4k+3} = p_{4k+1} \cdot (p_0 + \lambda_{4k+2}) + \lambda_{4k+3},$$

$$p(x) = \begin{cases} a_n p_n & (n = 4k+1, 4k+3) \\ a_n x p_{n-1} + a_0 & (n = 4k+2, 4k+4) \end{cases}$$

かけ算の回数は $\left[\frac{n+4}{2} \right]$ になるが、たし算の回数は $n+1$

となる。たとえば $n=9$ の場合

$$\begin{aligned} p(x) &= a_9 x^9 + a_8 x^8 + \dots + a_1 x + a_0 = a_9 \{ x^9 + b_8 x^8 + \dots + b_1 x + b_0 \} \\ &= a_9 p_9 \end{aligned}$$

で、Belaga-Pan のプロセスによれば、 $b_8, b_7, \dots, b_0$ が与えら

れる。また、 $\lambda_1, \lambda_2, \dots, \lambda_9$ を求めるためには、かけ算は 6 回、

たし算は 10 回となる。

$\lambda_1, \lambda_2, \dots, \lambda_9$ を求めるには、次の連立方程式を解く必要がある。未知数は、 $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \beta_2, \beta_3, \beta_4, \lambda_9$ である。

$$b_8 = 1 + \alpha_1$$

$$b_7 = \alpha_2 + \alpha_1 + \beta_2$$

$$b_6 = \alpha_3 + \alpha_2 + \beta_2 \cdot \alpha_1 + \beta_3$$

$$b_5 = \alpha_4 + \alpha_3 + \beta_2 \cdot \alpha_2 + \beta_3 \cdot \alpha_1 + \beta_4$$

$$b_4 = \alpha_5 + \alpha_4 + \beta_2 \cdot \alpha_3 + \beta_3 \cdot \alpha_2 + \beta_4 \cdot \alpha_1$$

$$b_3 = \alpha_5 + \beta_2 \cdot \alpha_4 + \beta_3 \cdot \alpha_3 + \beta_4 \cdot \alpha_2$$

$$b_2 = \beta_2 \cdot \alpha_5 + \beta_3 \cdot \alpha_4 + \beta_4 \cdot \alpha_3$$

$$b_1 = \beta_3 \cdot \alpha_5 + \beta_4 \cdot \alpha_4$$

$$b_0 = \beta_4 \cdot \alpha_5 + \lambda_9$$

(3-2)

を解けば、 $\beta_2, \beta_3, \beta_4$ から $\lambda_6, \lambda_7, \lambda_8$ が求まり、

$\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5$ から $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5$ と $(\lambda_2 + \lambda_3), \lambda_3, (\lambda_2 \cdot \lambda_3 + \lambda_4)$

が求められる。

数値例として, $b_8 = 9$, $b_7 = 9 \cdot 8 = 9^{(2)}$, $b_6 = 9 \cdot 8 \cdot 7 = 9^{(3)}$, $\dots$, $b_1 = 9^{(8)}$,
 $b_0 = 9!$ にとり ($f(x) = e^x \doteq \frac{1}{9!} (x^9 + b_8 x^8 + \dots + b_0)$ の場合)
 λ_9 の近似値として高木が 16770 と求めた。

(3-2) 式を $\{b_8 - (1 + \alpha_1)\}^2 + \{b_7 - (\alpha_2 + \alpha_1 + \beta_2)\}^2 + \dots + \{b_0 - (b_1 \cdot \alpha_8 + \lambda_9)\}^2$
 の極値を与える真を求める問題として, 解法 (3-1), (3-2),
 (3-3) と比較した。

(3-2) 式と解く	判 定
解法 3-1 Newton-Raphson 法	収 果 し た
3-2 FMFP 法	IER=1 (LIMIT=30 で収果しない)
3-3 FMCG 法	IER=2 (1000 回以上で収果しない)

よい方が近似から出発しないと, どの方法でもうまくいかない。
 (3-2) 式 の ちやうどな問題では偏微分係数が簡単に求まる
 ので, 解法 (3-1) が一番よいのは当然である。

数値例の OUTPUT (解法 3-1) を次にのせる:

4. むすび

多変数関数の極値を与える変を求める問題(制約条件のない場合)を連立型の非線型の方程式を解く問題として考えた。解のより近似を探すのが一番大変な仕事であった。

計算機で問題を解くときの心構えとしてよく知られているが、なかなか実行が難しい。とど、

1) その問題に関する理論のうちから、利用できるものは出来るだけ利用して、問題や解の性質を把握する。(各分野での理論の活用)

2) 問題や解の性質と出来るだけ活用した解法を選択することが大切である。

計算セクタにあるライブラリ(SSP等)を機械的に利用してもうまく行かないことが多いのは当然で、とくに多変数関数の極値探しでは(1),(2)の必要性は大きいと思われる。

最後に、Balaga, Pan による多項式の新しい計算法等について種々教えを頂いた伊理正夫氏に御礼申し上げる。

文献

- [1] R. Fletcher and M.J.D. Powell, "A Rapidly Convergent Descent Method for Minimization", Computer Journal, Vol. 6, (1963), PP. 163-168.
- [2] R. Fletcher and C.M. Reeves, "Function minimization by conjugate gradients", Computer Journal, Vol. 7, (1964), PP. 149-154.
- [3] IBM data center user's guide SSP PP. 1-9.
- [4] V. Ya. Pan, "Methods of computing values of polynomials. Russian Math. Surveys, Vol. 21, PP. 105-136 (1966).
- [5] L.A. Lyusternik, O.A. Chervonenkis and A.R. Yanpal'shii, Handbook for Computing Elementary Functions. Pergamon Press, 1965.
- [6] 伊理正夫, 講義プリント (1970)