

多峰性多変数関数の最大値の存在域に関する 検証アルゴリズム†

津田孝夫†† 柴田英明†††

多次元矩形領域で定義された連続関数があり、定義領域の1点で大域的な最大値をもつとする。この論文では、目的関数が定義されている矩形領域の内部に含まれる一つの小矩形領域を考えたとき、最大値がその小矩形中に含まれているか、いないか、を数値的に判定するアルゴリズムを提案している。探索アルゴリズムは通常局所的な方法であり、局所的な方法で得られた候補の解について本論文の手法を適用すれば、大域的な最大値かどうか判定できる。数値的に0か1かの判別に帰着するので、数値計算に高精度は要求されず、モンテカルロ積分法が有効に使える。

1. 序 論

多次元ユークリッド空間の有界領域で定義された連続関数の大域的な最大値 (global maximum) (以下単に最大値という) の存在を数値的に検証する新しい方法について論ずる。探索領域 D は超立方体であると仮定する。 D を多数の小矩形に分割していったとき、どの小矩形で D 上の最大値が起きるかが判定できる。

この方法は、著者による以前の論文¹⁾の拡張であって、この問題についての動機や一般的背景に関してはそこに述べてあるので参照していただきたい。

拘束条件のない最大値探索のアルゴリズムは、山のぼり法であろうと、ランダム・サンプリングであろうと、通常は単に真の解の「候補」を与えるにすぎない。すなわち、得た解が最大値であるという保証はなく、局所的な極大値 (local maximum) であるかもしれない。本論文の方法は、通常のほかのアルゴリズムと異なっていて、当該小領域内に真の解が存在するかどうかを判定する。したがってほかの有効な局所探索法と本論文の方法を結合して用いれば、解を誤認するおそれは少なくなると思われる。

文献 1) では目的関数を $F(\mathbf{x})$ ($\mathbf{x} \in D \triangleq D_1 \times D_2 \times \dots \times D_k$; k = 次元数; $D_i \triangleq \{x_i | 0 \leq x_i \leq 1\}$; D において $0 \leq F(\mathbf{x}) < \infty$) とし、それを指数関数の肩のせて $f(\mathbf{x}) = \exp[F(\mathbf{x})]$ とした上で、さらに次の関数

$$h(\mathbf{x}; \alpha) \triangleq [f(\mathbf{x})]^\alpha / \int_D [f(\xi)]^\alpha d\xi \quad (1)$$

を考えると、これは

$$h(\mathbf{x}; \alpha) \rightarrow \delta(\mathbf{x} - \mathbf{x}_M), \quad (\alpha \rightarrow \infty) \quad (2)$$

という性質を持つ。すなわち目的関数はデルタ関数に変換され、しかもそれは最大値の点 $\mathbf{x} = \mathbf{x}_M$ において無限大になり、ほかのすべての \mathbf{x} (二義的な局所的極大値を与える点も含む) に対してはゼロとなる。計算の途中の過程でオーバーフローが起きるので、この α は理想的に無限大にはできないなどの諸点については、文献 1) で言及した。パラメータ α については、付録 2 も参照のこと。(2) の関係は、「相対つりあげ」の効果とも呼ぶべきもので、1 例を図 1 に示す。ここでは視覚的に表現するため、目的関数は 2 変数関数としてある。この図からわかるように、パラメータ α はかきりに無限大にできなくとも、たんに 1 ケタ増大させるだけで、最大値をつりあげる著しい「整形作用」があることがわかる。

2. 3次元から多次元へ

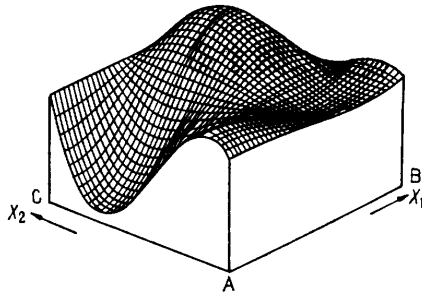
ここで静電気学との類似について述べる。点 \mathbf{x}_M において問題のデルタ関数は無限大になるが、このデルタ関数を空間的な電荷密度と考えると、点 \mathbf{x}_M に 1 個の「点電荷」があるとみなすことができる。そのとき、ガウスの定理 (発散定理) によれば、ある閉曲面の内部に問題の点電荷が存在するかどうかは、点電荷から出ている電気力線の線束 (フラックス) を、その閉曲面に交差している分について積分して総和を求めればわかる。すなわち積分値がゼロであれば、その閉曲面の内部には存在しない。数学的にはこれは、

$$\nabla^2 \phi(\mathbf{x}) = -\delta(\mathbf{x} - \mathbf{x}_M) \quad (3)$$

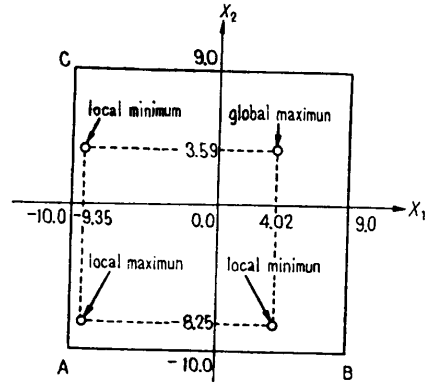
† Computational Verification on the Existence of the Global Maximum of a Multivariable Function by TAKAO TSUDA (Department of Information Science, Kyoto University) and HIDEAKI SHIBATA (Department of Electrical Engineering, Hokkaido University).

†† 京都大学工学部情報工学教室

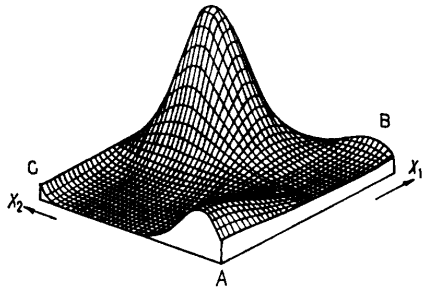
††† 北海道大学大学院工学研究科電気工学専攻 (現在 (株) 日立製作所)



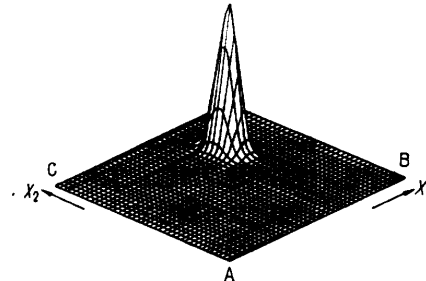
(a) $F(x_1, x_2) - F_0$ ($F_0 = -6.0 \times 10^4$) の関数の形を示す
 (a) Profile of function $F(x_1, x_2)$ minus F_0 ($= -6.0 \times 10^4$).



(b) 正方形 $D = [-10, 9] \times [-10, 9]$ 内の極大と極小の位置を示す
 (b) Location of local maxima and minima of $F(x_1, x_2)$ in D .



(c) $h(x; \alpha)$ のプロフィール ($\alpha = 0.86 \times 10^{-4}$)
 (c) Profile of $h(x; \alpha)$ with $\alpha = 0.86 \times 10^{-4}$.



(d) $h(x; \alpha)$ のプロフィール ($\alpha = 0.86 \times 10^{-3}$)
 (d) Profile of $h(x; \alpha)$ with $\alpha = 0.86 \times 10^{-3}$.

図1 相対つりあげの効果. 目的関数は

$$F(x_1, x_2) = (x_1 + 15x_1 + 1x_1 - 8x_2 + 11x_2 + 5x_2 - 9)/10$$

Fig. 1 Effect of the relative peaking for the objective function

$$F(x_1, x_2) = (x_1 + 15x_1 + 1x_1 - 8x_2 + 11x_2 + 5x_2 - 9)/10.$$

という形のポワソン方程式で表わされ、ガウスの定理により面積分 (交差フラックスの総和)

$$\iint_S \nabla \varphi \cdot dS$$

をとり囲んだ電荷の総量と等しいと置いて、次の関係が成り立つ。すなわち、

$$\begin{aligned} \iint_S \nabla \varphi \cdot dS &= \iiint_D \delta(x - x_M) dx \\ &= \begin{cases} 1 & (x_M \in D \text{ のとき}), \\ 0 & (x_M \notin D \text{ のとき}). \end{cases} \end{aligned} \quad (4)$$

このアナロジは、数学的に一般の k 次元 ($k > 3$) の場合にも成立し、したがって k 次元の場合について、(4)の左辺を D に含まれる小矩形 D' の表面にわたる面積分として計算すれば、 D' に x_M が含まれるかどうか判別できることになる。 D' を矩形とするのは、積分の実行を容易にするためである。 k 次元の場合の(4)の証明については付録1参照のこと。ところで、

(3)の $\varphi(x)$ は、多次元のグリーン関数と(3)の右辺との間のたたみこみ (convolution) で表わされ、さらにこの(3)の右辺は、与えられた目的関数 $F(x)$ を用いて表現される²⁾から、計算すべき公式を書き下すことができる。

3. 多次元の場合の公式と数値計算上の考察

D を序論で述べた単位超立方体とし、 D' を小矩形として $D' = [x_{11}, x_{12}] \times \dots \times [x_{k1}, x_{k2}]$, $0 \leq x_{i1} \leq x_{i2} \leq 1$ ($i=1, 2, \dots, k$) とし、 D' の表面を S' と表わす。文献1)で仮定したように、ここでも最大値は D の1点 x_M においてのみ起こるとする。以上により、(4)の判定は二者択一で、唯一の必要な数値情報は、 S' にわたる面積分の値が0か1かということである。理論上は、それ以外の中間的な値は生じない。このことから、通常の科学技術計算と異なり、上述の面積分の評価としては10進2ケタの精度があれば十分目的を達

することがわかる。したがって(3)の $\varphi(\mathbf{x})$ の表現に用いるグリーン関数としては、無限遠でゼロとなるものを採用でき、また多重積分自体にはモンテカルロ法³⁾が使える。これらの近似が実際上差支えないことは、後述の実例に関する詳しい計算で検証できる。

小矩形 D' をとった場合、(4)は次のようになる。

$$\begin{aligned} & \int_{S'} \nabla \varphi \cdot dS \\ &= \sum_{i=1}^k \int_{x_{i1}}^{x_{i2}} \dots \int_{x_{i-1,1}}^{x_{i-1,2}} \int_{x_{i+1,1}}^{x_{i+1,2}} \dots \int_{x_{k1}}^{x_{k2}} \\ & \left(\frac{\partial}{\partial x_i} \varphi \Big|_{x_i=x_{i1}} - \frac{\partial}{\partial x_i} \varphi \Big|_{x_i=x_{i2}} \right) dx_1 \dots \\ & dx_{i-1} dx_{i+1} \dots dx_k. \end{aligned} \quad (5)$$

ここでカッコの中はさらに

$$\begin{aligned} & \frac{\partial}{\partial x_i} \varphi \Big|_{x_i=x_{i1}} - \frac{\partial}{\partial x_i} \varphi \Big|_{x_i=x_{i2}} \\ &= \frac{\int_D \exp[\alpha F(\mathbf{t})] \left(\frac{\partial}{\partial x_i} G(\mathbf{x}; \mathbf{t}) \Big|_{x_i=x_{i1}} \right) \exp[\alpha F(\mathbf{t})] dt}{\int_D \exp[\alpha F(\mathbf{t})] dt} \\ & \quad - \frac{\int_D \exp[\alpha F(\mathbf{t})] \left(\frac{\partial}{\partial x_i} G(\mathbf{x}; \mathbf{t}) \Big|_{x_i=x_{i2}} \right) dt}{\int_D \exp[\alpha F(\mathbf{t})] dt} \end{aligned} \quad (6)$$

であって、グリーン関数は次式で与えられる。

$$G(\mathbf{x}; \mathbf{t}) = \frac{\Gamma(k/2)}{2(k-2)\pi^{k/2}} \cdot \frac{1}{\left(\sum_{j=1}^k (x_j - t_j)^2 \right)^{\frac{k-2}{2}}}, \quad (k \geq 3). \quad (7)$$

ここでパラメータ α は、十分大きく、しかしオーバーフローを生じない程度にとる¹⁾。付録2も参照のこと。

(5)は、当該小矩形領域 D' にわたる判別式とも呼ぶべきものである。

数値計算上の諸注意を次にまとめておこう。コーディングの仕方では計算の速さが相当変る。

(I) (5), (6)の多重積分を評価するのにモンテカルロ法が簡便である。またこの方法の性質から、計算量は、次元数の増加に対し指数関数的に増加することはない。したがって高次元の場合が計算できる。

(II) 扱う数値の変域が広いので、二重精度の計算が必要である。しかしこのためにとくに計算の速さが落ちることはない。

(III) (6)の関数 $G(\mathbf{x}; \mathbf{t})$ に現われる特異点 ($\mathbf{x} = \mathbf{t}$) の扱いには、特別な注意が必要である。変数 \mathbf{t} について積分を実行する前に $\mathbf{t} = \mathbf{x}$ となる \mathbf{t} の値はわかっている。そこでいわゆる層別サンプリングの手法³⁾を用いて、超立方体 D の各辺を3等分し、等しい

サイズの 3^k 個の小立方体に分割する。このうち $\mathbf{t} = \mathbf{x}$ という点が含まれる小立方体 ((6)の分子の被積分関数で、(5)の積分変数の1つの値に対し、そのような小立方体はたかだか2個存在する)にはほかの通常の小立方体に割当てるサンプル点数 N_1 に比し、十分多く N_2 個のサンプル点を割当てる。たとえば、 N_2 は N_1 の10倍にとる。また(5)の和記号の右側の各積分については、それぞれ S' の部分表面上で N_s 個のサンプル点をとることにする。

上記(I)および本項(III)に関連して、計算量について後述第5章で論ずる。

(IV) 実際にモンテカルロ計算を実行する際、(6)の右辺の分子は2つの積分の差として計算してはいけない。すなわち、

$$\int_D \exp[\alpha F(\mathbf{t})] \frac{\partial G}{\partial x_i} \Big|_{x_i=x_{i1}} dt - \int_D \exp[\alpha F(\mathbf{t})] \frac{\partial G}{\partial x_i} \Big|_{x_i=x_{i2}} dt \quad (8)$$

のように分解してはいけない。第一の理由は、平均の差をとるのでなく、差の平均をとらないと、誤差はきわめて大きくなるからである。すなわち(8)の形では独立にゆらぐ2つの和の引き算となり、単独の和に比し分散は極端に増幅されてしまう。第二の理由は、(8)の2項を(6)のように結合してモンテカルロ式に計算すると、サンプリングのゆらぎによる誤差を一定値に固定した場合、 $\exp[\alpha F(\mathbf{t})]$ の評価回数は(8)の場合の50%ですむからである。

(V) 目的関数の微分は必要でない。ただ関数 G の微分形が使われる。

(VI) 指数関数の計算や乱数発生を関数呼出しの形で行うと、とくにこれらの計算が一番深いループで頻繁に行われるため、リンケージだけで相当時間的に不経済になる。コンパイラによっては組み込み関数をオプション指定によりインライン(in-line)化できるが、これを用いると計算時間がかかなり短縮される。

4. 計算例

文献1)と同じ例題について考える。すなわち、

$$\begin{aligned} F(\mathbf{x}) &= F(x_1, x_2, \dots, x_5) \\ &= F_1(x_1) F_2(x_2) \dots F_5(x_5), \end{aligned} \quad (9)$$

ただし、

$$\left. \begin{aligned} F_1(x_1) &= x_1(x_1+13)(x_1-15)/100, \\ F_2(x_2) &= (x_2+15)(x_2+1)(x_2-8)/100, \\ F_3(x_3) &= (x_3+9)(x_3-2)(x_3-9)/100, \end{aligned} \right\} \quad (10)$$

$$\left. \begin{aligned} F_4(x_4) &= (x_4 + 11)(x_4 + 5)(x_4 - 9)/100, \\ F_5(x_5) &= (x_5 + 9)(x_5 - 9)(x_5 - 10)/100 \end{aligned} \right\}$$

とし、探索域 D を

$$-10 \leq x_i \leq 10 \quad (i=1, 2, \dots, 5) \quad (11)$$

とする。区間 $[-10, 10]$ はプログラムで線形変換され、 $[0, 1]$ となる。 D 内には 16 個の極大値があり、最大値はそのうちの 1 つの点

$$\begin{aligned} x &= (x_1, x_2, \dots, x_5) \\ &= (8.7564, -9.3582, -4.5720, 3.5921, \\ &\quad -2.8400) \end{aligned}$$

で起こる。(5)の $(k-1)$ 次元の積分は、面 S' 上にサンプル点をランダムに N_s 個とって実行し、(6)の右辺は(5)との関係でパラメータとして現われる各 x につき $3^k \cdot N_1$ 個のサンプル点を使って評価する。前述の特別扱いする小立方体においてとる N_2 個のサンプル点からの寄与は、全体としての計算量では無視できる。したがって今の例の場合、 $k=5, N_1=4$ であって、最大値が当該部分領域に入っているかどうか 1 回判定するのに、約 $5 \times 100 \times 1,000$ 回の関数評価を行った。これに要する計算時間は約 40 秒 (CPU 時間) であった (京都大学大型計算機センタ M200 を使用)。

各座標軸を 2 分割し、それを繰返した計算過程を表 1 に示す。これにより、どの小矩形に最大値が含まれ

るか一目瞭然であろう。すなわち、最初の 5 回の 2 分割 (#2 から #6 まで) により、最大値の存在可能な空間の体積を $1/2^5 (=1/32)$ に縮小している。数値的詳細は別として、この分割過程をわかりやすくするため別の仕方で示したのが図 2 である。0 を書きこんだ小正方形では最大値が存在しえず、1 を書きこんだ小正方形ではそれが存在可能である。さて、この 2 分割をさらに 5 回 (表 1 および図 2 の #8 から #12 まで) 行くと、この存在可能な領域の体積は、最初の $1/4^5 (=1/1,024)$ 倍に縮まる。

何回か 2 分割を繰返し、存在可能な空間の広さを縮めてから、それを改めて目的関数の定義域 D にとりさらに 2 分割を繰返すことが考えられる。パラメータ α は設定しなおす (一般に前処理¹⁾での観測値 $F_{\max} - F_{\min}$ は小さくなるので、 α としてはより大きくなる)。これを行ったのが表 2 で、表 1 の #8 から #12 に対応するものである。このようにアルゴリズムを再帰的に使用すると、被積分関数は相対的になめらかになるため、判別式の計算の結果得られる S の値の 2 乗平均 (root mean square) 誤差は減少する。

ところで、2 分割の分割面と大域的 maximum の位置が一致した場合どうなるであろうか。点電荷から発する電気力線のアナログが成り立つものなら、1 であるは

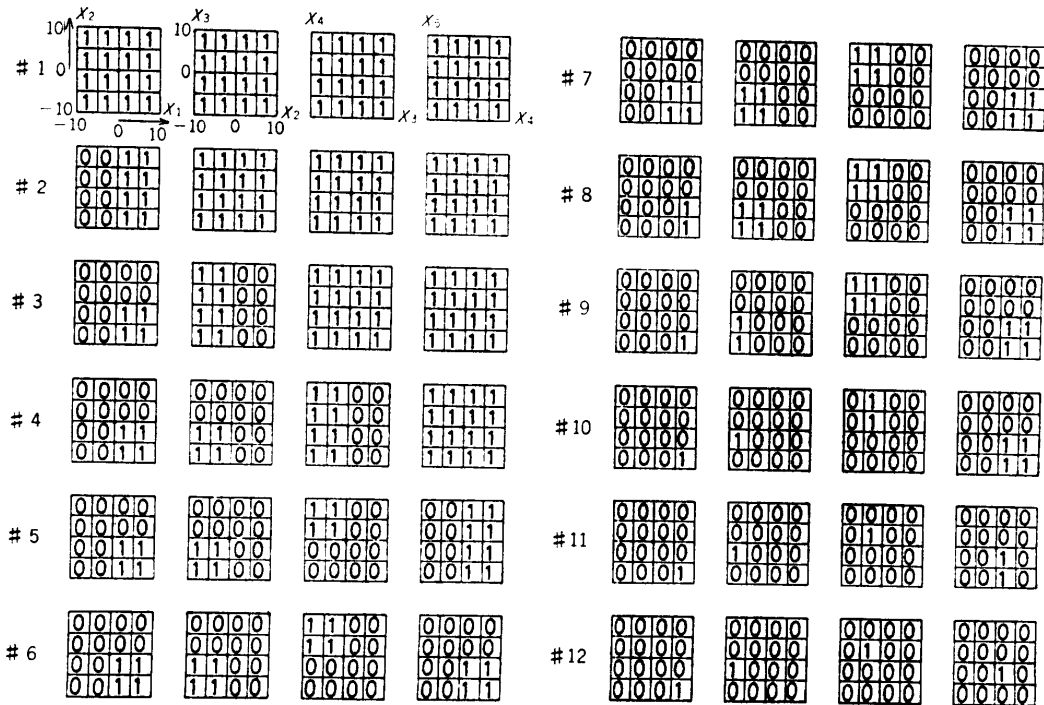


図 2 2 分割をくり返し、最大値の存在領域を縮小する過程 (表 1 に対応)

Fig. 2 Process of narrowing down subdomains by repeated bisections.

表 1 2分割のくり返し過程 (パラメータについては英訳説明参照)

Table 1 Repetition of bisections
 ($D=[-10, 10]^2$; $\alpha=0.3 \times 10^{-2}$, $F_0=-0.165 \times 10^4$; for F_0 , see ref. (1); $N_S=100$, $N_1=4$, $N_2=40$; RMS error of S (#8 through #12)=0.111; S is the computed value of the discriminant given by eq. (5)).

case	x_{1l}	x_{1R}	x_{2l}	x_{2R}	x_{3l}	x_{3R}	x_{4l}	x_{4R}	x_{5l}	x_{5R}	S
# 1	-10	10	-10	10	-10	10	-10	10	-10	10	0.927
# 2	-10	0	-10	10	-10	10	-10	10	-10	10	0.027
# 3	-10	10	-10	0	-10	10	-10	10	-10	10	1.053
# 4	-10	10	-10	10	-10	0	-10	10	-10	10	1.075
# 5	-10	10	-10	10	-10	10	-10	0	-10	10	0.025
# 6	-10	10	-10	10	-10	10	-10	10	-10	0	0.983
# 7	0	10	-10	0	-10	0	0	10	-10	0	0.913
# 8	0	5	-10	0	-10	0	0	10	-10	0	0.002
# 9	0	10	-10	-5	-10	0	0	10	-10	0	1.039
# 10	0	10	-10	0	-10	-5	0	10	-10	0	0.147
# 11	0	10	-10	0	-10	0	0	5	-10	0	0.964
# 12	0	10	-10	0	-10	0	0	10	-10	-5	-0.194

表 2 アルゴリズムの再帰的使用 (パラメータについては英訳説明参照)

Table 2 Recursive use of the algorithm
 (new $D=[0, 10] \times [-10, 0] \times [-10, 0] \times [0 \times 10] \times [-10, 0]$; $\alpha=0.45 \times 10^{-2}$, $F_0=-0.119 \times 10^4$; $N_S=100$, $N_1=4$, $N_2=40$; RMS error of S (#1 through #5)=0.061).

case	x_{1l}	x_{1R}	x_{2l}	x_{2R}	x_{3l}	x_{3R}	x_{4l}	x_{4R}	x_{5l}	x_{5R}	S
# 1	5	10	-10	0	-10	0	0	10	-10	0	0.942
# 2	0	10	-5	0	-10	0	0	10	-10	0	0.018
# 3	0	10	-10	0	-5	0	0	10	-10	0	0.992
# 4	0	10	-10	0	-10	0	5	10	-10	0	-0.093
# 5	0	10	-10	0	-10	0	0	10	-5	0	0.922

表 3 2分割の平面と大域的最大値の起きる点が一致した場合 (N_S, N_1, N_2 は表 1 と同じ)
 Table 3 Point of the global maximum falling on planes of bisection (The sizes of samples N_S, N_1, N_2 are the same as in table 1).

x_{1l}	x_{1R}	x_{2l}	x_{2R}	x_{3l}	x_{3R}	x_{4l}	x_{4R}	x_{5l}	x_{5R}	S
-10	10	-10	10	-10	10	-10	10	-2.84	10	0.625
-10	10	-10	10	-10	10	-10	10	-10	-2.84	0.393
0	8.76	-10	10	-10	10	-10	10	-10	10	0.727
8.76	10	-10	10	-10	10	-10	10	-10	10	0.361
-10	10	-10	10	-10	-4.57	-10	10	-10	10	0.857
-10	10	-10	10	-4.57	10	-10	10	-10	10	0.409
-10	10	-10	10	-10	10	-10	3.59	-10	10	0.892
-10	10	-10	10	-10	10	3.59	10	-10	10	0.238

ずのフラックスの総量は、2つに分割されるであろう。実際そのようになることが表 3 に示されている。4つの場合のうち、最初のものは、大域的最大値の起きる点を $x_1 x_2 x_3 x_4$ 平面で2分割した場合である。さらに幾通りもの例について計算したところ、この“フラックスの保存”が成り立つことがわかる。したがって、もし判別式の計算値 S が (1 になるべきところ)、隣接する2つの分割された領域に分配されてしまうことがわかれば、適当な座標の平行移動を行い、同様の計算を続行すればよい。

以上のように本論文で述べたアルゴリズムにはモンテカルロ法が適しており、計算結果が0か1にならなければ単にサンプル点の数(すなわちサンプル・サイズ)を増すか、またはサンプル点の点集合(すなわちサンプル自体)を入れかえるかにより、計算値はどちらかに近くなる。これが実現されない場合は表3について述べた現象が起きている(その対策については言及した)か、等しい高さの最大値が2つ以上存在して仮定が成り立たないためである。また積分の方法として、比較的単純な手法を使ったが、その理由はモンテ

カルロ積分法も単純な方が強靱 (robust) であり, 工夫をこらした複雑なものは必ずしもそうでないことが, 多数の計算例から判明したからである.

以上要するに, 大域的最大値の確認は, 各種の最適化問題と関連して重要であり, ほかに適当な代案がないことを考えれば, 本論文で述べた方法は十分存在理由があると結論する.

5. 計算量の評価

判別式(5)の値を S とすると, S は

$$S = \sum_{i=1}^k I_i$$

の形で表わされ, I_i は小矩形 D' のあい向いあう2面 (D' の表面 S' の部分表面) についてとる面積分

$$I_i = \frac{1}{N_i} \sum_{j=1}^{N_i} J_{ij}$$

と書ける. すなわち I_i は $(k-1)$ 重積分で, 上記の部分表面上にランダムに N_i 個のサンプル点をとって crude Monte Carlo で評価する. 一方 J_{ij} はさらに(6)の関係から, 第3章(III)で述べたように k 重積分として stratified Monte Carlo により

$$J_{ij} = \sum_{l=1}^{3^k} K_{ijl}$$

の形で求められる. そこで $\sigma^2(\ast)$, $\sigma(\ast)$ を量 \ast の分散および標準誤差とすると,

$$\sigma^2(S) \leq k \sigma^2(I), \quad \sigma^2(I) = \max_i \{\sigma^2(I_i)\};$$

$$\sigma^2(I_i) \leq \sigma^2(J)/N_i, \quad \sigma^2(J) = \max_{i,j} \{\sigma^2(J_{ij})\};$$

$$\begin{aligned} \sigma^2(J_{ij}) &= \sum_{l=1}^{3^k-2} \sigma^2(K_{ijl}) + \sum_{l=2}^2 \sigma^2(K_{ijl}) \\ &= \sum_{l=1}^{3^k-2} c'_{ijl}/N_1 + \sum_{l=2}^2 c''_{ijl}/N_2, \end{aligned}$$

ただし最後の式の右辺の第2の和 $\sum_{l=2}^2$ は特異点 $t=x$ を含む小立方体についての和, 第1の和はそれ以外的小立方体についての和である. c'_{ijl} , c''_{ijl} はサンプル点数に無関係な定数. ここで

$$\max_{(i,j,l)} c'_{ijl} = c', \quad \max_{(i,j,l)} c''_{ijl} = c''$$

とすると

$$\sigma^2(J_{ij}) \leq (3^k-2)c'/N_1 + 2c''/N_2$$

となるが, stratify する目的はもともと $c'/N_1 \sim c''/N_2$ とすることにあるから, いまこれを仮定すると結局

$$\sigma^2(S) \leq k 3^k c' / (N_1 N_2) \quad (11)$$

となる. そこでいまかりに, [(6)の右辺の分子の積分の被積分関数]/[(6)の右辺の分母の積分] \triangleq integrand と書くと, 3^k 個の小立方体(層)内では crude

Monte Carlo を使うから, モンテカルロ計算の性質³⁾から,

$$c' \leq \max \left\{ \Delta \int (\text{integrand})^2 dt - \left(\int \text{integrand} dt \right)^2 \right\}$$

と表わされる. ここで右辺の2つの積分は, 同一の小立方体にわたるもので, \max はすべての小立方体(各々の体積 $\Delta=3^{-k}$) についての最大値である. これにより c' は $O(\Delta^2)$ であり, ある正の定数 c を用いて

$$c' \leq c \Delta^2$$

となる. サンプル点数(サンプル・サイズ)が N であり, $N \cong k N_1 \cdot 3^k N_2$ (ただし $3^k \gg N_2$ を用いた) であることに注意して, (11)から標準誤差は

$$\sigma(S) \leq k \sqrt{c/N} \quad (12)$$

となる. この係数 c は目的関数に依存し, 目的関数自体の次元数依存性などの構造を指定しない限り評価できない. しかし与えられた目的関数に対し判別式の評価精度は, サンプル・サイズ N の平方根の逆数に比例する. これをもとに適応的にサンプル・サイズを加減する. 表1の#1は, 既知の S の値 (=1) に対する確認で, この程度のサンプル・サイズでよいことを本計算実行の前に確かめている.

単純なランダム・サンプリングでは観測する最良値の漸次的更新により解を得る. 本論文の方法ではそのような情報も包含するが, 対象関数の任意に詳しい「構造的」情報を与える手段が与えられている.

参考文献

- 1) Tsuda, T.: Interactive Search for the Global Maximum of a Function of Many Variables, IEEE Trans. on Automatic Control, Vol. AC-22, No. 3, pp. 399-405 (1977).
- 2) 熊ノ郷準: 偏微分方程式, 5.5節 (p. 140), 共立出版 (1978).
- 3) 津田孝夫: モンテカルロ法とシミュレーション (改訂版), 第4章, 培風館 (1977).

付 録 1

矩形領域 D および D' について, 式(4)の性質は次のようにして示せる. すなわち D' ($\subset D$) について,

$$\begin{aligned} & \int_{D'} \nabla^2 \varphi dx_1 dx_2 \dots dx_k \\ &= \sum_{i=1}^k \int_{x_{i,l}}^{x_{i,u}} \dots \int_{x_{i-1,l}}^{x_{i-1,u}} \int_{x_{i+1,l}}^{x_{i+1,u}} \dots \int_{x_{k,l}}^{x_{k,u}} dx_1 \dots \end{aligned}$$

$$\begin{aligned} & dx_{i-1} dx_{i+1} \cdots dx_k \int_{x_{i1}}^{x_{in}} \frac{\partial^2 \varphi}{\partial x_i^2} dx_i \\ &= \sum_{i=1}^k \left(\prod_{j(\neq i)} \int_{x_{j1}}^{x_{jn}} dx_j \right) \left(\frac{\partial \varphi}{\partial x_i} \Big|_{x_i=x_{in}} - \frac{\partial \varphi}{\partial x_i} \Big|_{x_i=x_{i1}} \right) \\ &= - \int_{S'} \nabla \varphi \cdot dS, \end{aligned}$$

ただし、 S' は D' の表面であり、 dS は面要素 dS ($\subset S'$) に対し法線方向で大きさが dS の内向きのベクトルである。(3)を k 次元について考えると、以上により次式を得る。

$$\int_{S'} \nabla \varphi \cdot dS = \begin{cases} 1 & (\mathbf{x}_M \in D' \text{ のとき}), \\ 0 & (\mathbf{x}_M \notin D' \text{ のとき}). \end{cases}$$

付 録 2

パラメータ α の定め方については、文献 1) で前処理による方法が述べてある。ここでは追加説明をしておく。 \mathbf{x}_M を最大値の点、 \mathbf{x}'_M を二義的な局所的極大値 (第 2 峰) の点とすると、相対つりあげの効果は十分小さな正数 ε に対し、

$$\frac{h(\mathbf{x}'_M; \alpha)}{h(\mathbf{x}_M; \alpha)} = \exp[\alpha(F(\mathbf{x}'_M) - F(\mathbf{x}))] < \varepsilon \quad (\text{A1})$$

とすることである。そこで

$$F(\mathbf{x}'_M) = F(\mathbf{x}_M) \delta \quad (\text{A2})$$

すなわち、最高峰に対し第 2 峰以下はたかだか 100 δ %

の高さしかないとする、

$$\varepsilon \ll \delta < 1 \quad (\text{A3})$$

なる大小関係が相対つりあげにより実現されるようパラメータ α を選ぶ。 δ を目的関数における第 2 峰のランキングと呼ぶ。(A1), (A2) から

$$\alpha > \frac{1}{F(\mathbf{x}_M)} \left| \frac{\log \varepsilon}{1 - \delta} \right|$$

となるが、前処理¹⁾で見いだされる F_{\max} について $F(\mathbf{x}_M) > F_{\max}$ であるから、

$$\alpha > \frac{1}{F_{\max}} \left| \frac{\log \varepsilon}{1 - \delta} \right|$$

とすれば十分である。

第 4 章で述べた例の場合、 $F_{\max} \sim 2 \times 10^4$ (実測) で、いまかりに $\delta \sim 0.9$ (仮定) とする。この δ は、いわば統計的検定における危険率のようなもので、目的関数に対しほとんど何らの制限をつけていない本論文の場合、 δ の値は仮定せざるをえない。このようにとると、 $\varepsilon = 10^{-2}$ に対し、 $\alpha > 0.23 \times 10^{-2}$ となる。一方、第 4 章で述べた例 (表 1) では、パラメータ α は文献 1) の前処理法できめているが、その値は $\alpha = 0.3 \times 10^{-2}$ であって、妥当な値である。

(昭和 55 年 4 月 11 日受付)

(昭和 55 年 7 月 17 日採録)