1D–06　直交計画法によるノイズつき多次元関数の勾配推定

田中 秀俊, 白石 將, 佐藤 裕幸, 中島 克人

新情報処理開発機構 並列応用三菱研究室
三菱電機 (株) 情報技術総合研究所

1　はじめに

関数最適化問題のうち、変数の数が多いものを考える。例えば、二面角および原子間距離を残基毎に3つずつ変数とする、長さ数百残基のタンパク質の折り畳み問題では、1000変数規模のポテンシャル関数を扱う。このような最適化問題は、しばしば局所探索法 (Local Search, LS) でアプローチされる。LSは遺伝的アルゴリズム [1]、シミュレーテッドアニーリング [2]、タブーサーチ [3] などを含む概念であり、探索アルゴリズムの共通基盤概念と位置づけることができる。LSは (1) 初期値を与える段階、(2) 探索のある時点における「現在値」について、現在値自身およびその適当な近傍の関数値を局所的に探索する段階、(3) 現在値を更新する段階、(4) 探索終了を判定する段階の大きく4段階からなり、このうち (2) 局所探索段階と (3) 現在値更新段階の2段を反復単位として探索を進める方法である [4]。

この第2段階で、現在値でのポテンシャル関数の導関数値（勾配）を見積りたい。それには、差分近似という方法もある。前進差分近似では、注目する変数毎に所与の刻み幅だけプラスし他の変数を固定して得た地点の関数値と、現在の探索点の関数値とから微分係数を近似する。中心差分近似では、注目する変数毎に所与の刻み幅だけプラスおよびマイナスに変化させ、他の変数を固定して得た地点の関数値から微分係数を近似する [5]。これら差分近似を、n 変数の作る格子点上の局所探索段階にほぼそのまま適用すると、現在値の近傍 2^n 点の中から、前進差分では n 個、中心差分では $2n$ 個の範囲の探索を行うことになる。

本稿では直交計画法を用いることにより、この差分近似よりも効率的かつ頑健な方法がとれることを示し、数値実験を通じてその有効性を示す。

2　ODLS

局所探索段階は、各変数についてプラスが良いかマイナスが良いかを決定する2値変数の問題と見なせる [6]。2値変数 n 個のとる空間 2^n の中から少ない実験数で最適組合せを見積もる方法として、2水準直交計画法がある [7]。この場合の2水準直交計画法とは、前述のような格子点近傍 2^n の中から、ガロア体に基づいて選択した約 n 点を選び、その関数値を求め、分散分析の手法を用いて最適化効果の高い変数とその方向とを選び、それを総合して最適化方向とする方法である。変数の数を n、刻み幅を d とすると、以下の手順となる。

(1) 近傍点数は $m=2^q$、ここで $2^{q-1}\leq n<2^q$、q は整数である。2水準 n 変数直交計画における第 i 実験、第 j 因子の水準 L_{ij} は、i のグレーコード表現 $(g_0g_1\ldots g_{q-1})$ と j の二進コード表現 $(b_0b_1\ldots b_{q-1})$ により、

$$L_{ij}=\mathrm{mod}_2\sum_{s=0}^{q-1} g_s b_{q-s-1} \qquad (1)$$

となる。L_{ij} は0もしくは1の2水準となる。

(2) 2水準を $+d$ と $-d$ に割り当て、変数 n 個を水準が全部0の第0因子を除く $m-1$ 個の列にランダムに割り当て、m 個の近傍点集合を設定。

(3) 近傍点集合の各点 N_j で関数値 $f(N_j)$ を算出。

(4) 各変数について $+d$ と $-d$ の関数値平均を算出。(以下は0を $-d$ に割り当てた場合)

$$\begin{aligned}\mu_i^- &= \frac{2}{m}\sum_{j=0}^{m-1}(1-L_{ij})f(N_j) \qquad (2)\\ \mu_i^+ &= \frac{2}{m}\sum_{j=0}^{m-1}L_{ij}f(N_j)\end{aligned}$$

(5) 判断用変数 B を用い、変数 i の次現在値として $+d$ か $-d$ か0を以下のように選択 (最大化問題)。

$$e_i = \begin{cases} +d & (\mu_i^- + B < \mu_i^+) \\ -d & (\mu_i^+ + B < \mu_i^-) \\ 0 & (\text{otherwise}) \end{cases} \qquad (3)$$

(6) 結果を総合して次の現在値とする。

本稿ではこの方法を ODLS（Orthogonal Design of experiment based Local Search）と呼ぶ。2水準直交計画法で得た最適化方向は、問題の考慮していない変数に関しても頑健であり、また前記のような差分近似類似の方法で決定する場合に比べ、関数値のゆらぎに強いと考えられる。

Gradient Approximation of Noisy Multivariate Function Using Orthogonal Design of Experiment
Hidetoshi Tanaka, Masashi Shiraishi, Hiroyuki Sato, Katsuto Nakajima
Real World Computing Partnership, Parallel Application Mitsubishi Laboratory, Mitsubishi Electric Corp.

3 数値実験

ODLS の挙動を、ノイズ R を含む以下の関数の最大化問題を例として述べる。

$$\text{maximize} \quad f^1(x) \quad = \quad -\sum_{i=1}^{n} [(x_i - 10)^2 + R] \quad (4)$$

ここで $n = 100$、R は整数乱数 ($0 \le R \le 1$、もしくは $0 \le R \le 99$). 刻み $d = 1$、初期値は整数乱数 ($-50 \le x_i < 50$) で決める.

3.1 差分近似

差分近似では各変数の進むべき方向をわずか 2 点の関数値で決める。ノイズを除去するために、同じ点で関数値を k 回とって平均する方法もある。その場合 $n = 100$、$k = 5$ なら、関数値の計算は 1 反復あたり 1000 回実施される。図 1 は差分近似を用いた局所探索法の、反復回数と f^1 の値を、$k = 1$(R1S1) と、$k = 5$(R1S5) と、$k = 10$ (R1S10) について示している。ここで R は 0 もしくは 1 を確率 1/2 でとる乱数である。

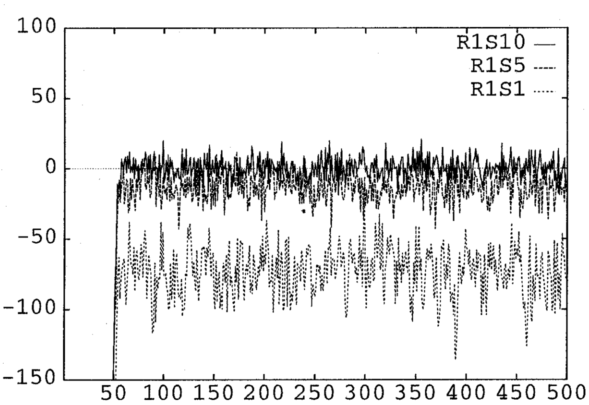

図 1: 差分近似を用いた局所探索法

3.2 判断用変数

ODLS は判断用変数 B を用いて、± 0 を選ぶ幅を決めている。図 2 は反復回数と f^1 を B を変えて示したものである。($B = 0$ (B0), $B = 10$ (B10), $B = 50$ (B50), $B = 100$ (B100), and $B = 150$ (B150))。ここで R は 0 から 99 までを一様にとる整数乱数である。

4 考察

目的関数が簡単な二次関数でノイズもない場合は、差分近似による方法と ODLS とで反復回数は一致する。しかし、差分近似による方が一反復あたりの計算量は少なくてすむ。変数の数を n として、差分近似によれば n 回の比較で一反復なのに対し、ODLS ではおおよそ $n/2$ 個の実験の総和を $2n$ 回とってから n 回の比較を要する。

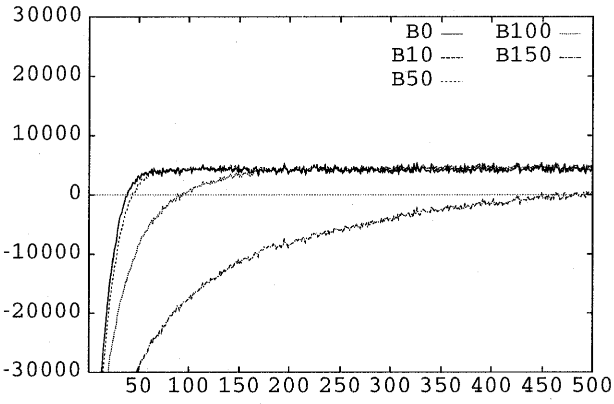

図 2: ODLS の判断用変数

しかし、目的関数が同じ二次関数でもノイズを入れた場合は、ODLS の精度は差分近似による方法の精度を大きく上回る。これは反復回数によらない。ODLS では各変数の増減についてそれぞれおおよそ $n/2$ 回の実験を行って平均を算出しており、この効果でノイズによらないロバストな探索を実行できる。この効果は差分近似において関数値算出を反復し平均しても得られるが、性能上同等な精度を得ようとすると、探索での一反復あたりの実験数が ODLS に比べて非常に大きくなってしまう。

実際問題、目的関数を全ての変数を考慮したものにはできない。ODLS は、考慮しない変数による目的関数の揺れがある場合、効力を発揮すると言える。

5 まとめ

提案した 2 水準直交計画法を用いた局所探索法の有効性を、ノイズ付き二次関数の最適化問題を通じて示した。この問題は、制御できない変数の値の振れをノイズとしてモデル化したものであり、一般性が高い結果と言える。

参考文献

[1] Goldberg, D.E. *Genetic Algorithms in Search, Optimization and Machine Learning.* Addison-Wesley, (1989).

[2] Kirkpatrick, S. et al. Optimization by Simulated Annealing. *Science* 220:671-680. (1983).

[3] Glover, F., and Laguna, M. *Tabu Search.* Kluwer Academic Publishers, (1997).

[4] 柳浦, 茨木. 組合せ最適化問題に対するメタ戦略について. 信学会 D-I, J83:3-25, (2000).

[5] Nemhauser, G.L., et al. 伊理他監訳. 最適化ハンドブック. 朝倉書店, (1995).

[6] Tanaka, H. et al. A Parallel Search Algorithm for Discrete Space and Its Application to Protein Folding Problem. *Proc. of 1998 RWC Symposium.* RWC TR-98001, (1998).

[7] 高橋. 組合せ理論とその応用. 岩波全書, (1979).