

直交計画法を用いた最適化法とその性能比較評価

白石 将[†], 田中 秀俊[†]

実数を変数とした連続な目的関数を最小化する関数最適化問題を考えている。我々は、現在点を繰り返し更新していくことによって準最適な解に到達する反復法の1種であり、2水準直交計画により現在点の近傍点を選択して目的関数値を求め、各変数の+方向と-方向の優劣を判定、その結果を総合して改善方向を定めて現在点を更新するODLSを提案している。このままでは多峰性の目的関数に対応できないため、異なる近傍サイズを利用し、さらに解が悪化する方向への現在点更新を許可することによってローカルミニマムを脱出可能としたODLSを新たに実装した。このODLSを代表的ベンチマーク問題に適用し、SAと比較してはるかに良い性能が得られることを確認した。しかし、タンパク質立体構造予測への適用実験では、平均してSAの方がODLSよりも良い性能を示した。

Experimental Evaluation of Optimization Using Orthogonal Design Local Search

Masashi Shiraishi[†], Hidetoshi Tanaka[†]

ODLS (Orthogonal Design Local Search) is an optimization algorithm which estimates the gradient vector by utilizing experimental design. We revised the ODLS by incorporating the concept of variable neighborhoods to escape from local minima. Experimental evaluations of this algorithm solved representative benchmark problems and protein potential function minimization problems. ODLS performed well for the benchmark problems. But SA outperformed ODLS for the latter problems on average.

1. はじめに

実数を変数とした連続な目的関数を最小化する関数最適化問題を考える。このような最適化問題を解くための手法としては、以下の手順に基づく反復法を利用することが多い。

1. 現在点を初期化して目的関数により評価する。
2. 予め定められた操作を現在点に適用することで近傍点を生成し、目的関数により評価する。
3. 手順2の結果に基づいて現在点を更新し、手順2に戻る。

ここで「現在点」は、対象としている最適化問題に対する各時点における解候補を意味する。目的関数の導関数についての情報を入手できる場合は、現在点を勾配が急であるような方向に更新することによって、効率よく探索を進めることができる。導関数の情報を直接入手できない場合、複数の点における目的関数の値から目的関数の減少方向を見積もるとの方法が考えられ、例えばMDS(Multi-Directional Search)[1]などのアルゴリズムが提案されている。

一方、我々は直交計画法を利用して近傍点生成および現在点更新を行う最適化アルゴリズムODLS(Orthogonal Design Local Search)を提案している[2][3][4]。しかし、もとのODLS[2]はローカルミニマムを脱出する機構を持たないため、多峰性の目的関数の最適化問題には適しないとの問題点があった。そこで、異なる近傍サイズを利用し、さらに

解が悪化する方向への現在点更新を許可することによって、ローカルミニマムからの脱出を可能にする拡張をODLSに施した。そして、この拡張版ODLSを代表的ベンチマーク問題に適用して評価した。さらに、実世界で解くことが求められている困難な最適化問題としてタンパク質立体構造予測を取り上げ、拡張版ODLSの評価実験を行った。本稿ではODLSとその拡張、また評価実験結果について報告する。

2. 最適化アルゴリズム

2.1 基本的なODLS

ODLSでは2水準直交計画により現在点の近傍点を生成してそれらの目的関数値を求め、各変数の+方向と-方向の優劣を判定する。そして、その結果を総合して現在点の改善方向を定めて更新する。

以下、ODLSによる近傍点生成および現在点更新の具体的手順を示す。但し、変数の数を n 、近傍点の数を m 、所定の正の定数を w および d 、現在点を x 、現在点の第 j 変数を x_j (但し $j = 0, 1, \dots, n-1$)、第 i 近傍点を N_i (但し $i = 0, 1, \dots, m-1$)、第 i 近傍点の第 j 変数を N_{ij} 、目的関数を $f(\cdot)$ と表す。

1. 近傍点数を $m = 2^q$ (ここで $2^{q-1} \leq n < 2^q$, q は整数)と定める。 i のグレイコード表現を $[g_0 g_1 \dots g_{q-1}]$, $1 \sim m-1$ からランダムに n 個選んで $0 \sim n-1$ とランダムに対応させる関数を $C(\cdot)$ とする。 $C(j)$ の2進コード表現を $[b_0 b_1 \dots b_{q-1}]$, 2の剰余を mod_2 とする。この時、近傍点 N_i を以下の式により定める。ここで、 L_{ij} は0と1の2水準の直交計画を表す行

[†]三菱電機(株)情報技術総合研究所
Information Technology R&D Center, Mitsubishi
Electric Corporation

列の要素となる．

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

$$N_{ij} = x_j + 2w(L_{ij} - 0.5) \quad (2)$$

つまり，現在点の各変数について $-w$ と $+w$ の 2 種類の移動を施すことにより，近傍点が生成される．

2. 各近傍点 N_i の目的関数値 $f(N_i)$ を算出する．
3. 各変数について $-w$ 移動と $+w$ 移動のそれぞれの目的関数値平均を下式により算出する．

$$\mu_j^- = \frac{2}{m} \sum_{i=0}^{m-1} (1 - L_{ij}) f(N_i) \quad (3)$$

$$\mu_j^+ = \frac{2}{m} \sum_{i=0}^{m-1} L_{ij} f(N_i) \quad (4)$$

4. 下式に従って現在点の各変数についての更新ベクトル成分 α_j を算出する[‡]．ここで ϵ は十分小さい実数定数，また k は所定の整数係数である．

$$\Delta\mu = \max_j |\mu_j^- - \mu_j^+| \quad (5)$$

$$\alpha_j = \left\lfloor \frac{(\mu_j^- - \mu_j^+)(k + 1 - \epsilon)}{\Delta\mu} \right\rfloor \quad (6)$$

この時， α_j は $-k, \dots, -1, 0, 1, \dots, k$ のいずれかの値を取る．

5. 各変数について，算出された更新ベクトル成分 α_j の d 倍を現在点に加えることにより，現在点を更新する．

上記手順 1 にて近傍点生成の方法にランダム性が含まれているので，同じ初期点を与えて ODLS を複数回適用しても，一般に探索経路は異なるものになる．ODLS は複数の近傍点の目的関数値の平均に基づいて処理を行うために目的関数算出にノイズが含まれる場合にも性能が良いとの利点がある [2]．

2.2 ODLS の拡張

上述の基本的な ODLS は目的関数値が小さくなる方向に探索を進めるだけであるため，ローカルミニマムから脱出する機能を持たない．そのような機能を持たせるための方策として様々な方法が考えられる．例えば，文献 [3][4] では，現在点と近傍点との間隔を示す数 w (上記手順 1 参照) をランダムに決定する機構を組み込むことにより，ODLS をローカルミニマムから脱出可能に拡張した場合についての代表的ベンチマーク問題 (具体的には，後述の Rastrigin 関数および Griewank 関数最小化) への適用結果が示されている．これらの手法では，現在点の更新幅

[‡]従来は $\mu_j^- - \mu_j^+$ の符号に基づき α_j の値は -1 か $+1$ のみとする方式 [2][3] を用いていたが，現在点更新をより効率良く行うため本文中の方式 [4] を採用した．

を示す数 d (上記手順 5 参照) についても定数ではなく，線形探索ないしは二分探索によって最適な値を求めるように工夫されている．そして評価実験の結果，SA (Simulated Annealing) [5] と比較して ODLS が良い性能を示すことが述べられている．

しかし，上記手法を後述のタンパク質立体構造予測問題に適用したところ，ODLS では現在点の更新が進まず，良い性能が出なかった．そこで，ローカルミニマムからの脱出機能を若干異なる方式で持たせることにした．本稿の評価実験で利用した，拡張版 ODLS の特徴を以下に示す．

- 異なる大きさの w に対応する複数の近傍を用意し，デフォルトでは最小サイズの近傍を利用する．
- 上記手順 5 において $d = w$ として現在点から生成される点を定める (この点を「改善方向点」と呼ぶ)．
- 近傍点および改善方向点のうち，目的関数値最小の点を現在点の更新先候補とする．この点に現在点を更新するか否かは，目的関数値の差によって決定する (目的関数値が悪化する方向への更新も許す)．
- 一定の最適化ループ回数の間，現在点が別の点に更新されない場合，近傍の大きさを 1 段階大きくする．
- 現在点が別の点に更新された場合，近傍の大きさを最小サイズに戻す．

2.3 比較対象のアルゴリズム

本稿の目的は，拡張版 ODLS を最適化問題に適用して，その性能を評価することである．そのため，ODLS の他，比較対象として以下に説明する最適化アルゴリズム MDS [1] および SA を利用した．

MDS MDS は， n 変数の最適化問題について，現在点および他の n 個の点から構成される n 次元単体 (例えば 2 次元単体は三角形となる) を用いる．探索の各反復において，現在点は単体を構成する頂点のうち最も目的関数値が小さい頂点として定められる．そして単体に対して現在点を基準とした反転/拡大/縮小操作を適用して変形する．さらに，変形後の単体の各頂点の目的関数値に基づいて現在点を更新することにより，探索が進められる．

SA SA は目的関数値が大きくなる方向への現在点の更新を許すことにより，ローカルミニマムからの脱出を可能にする．動作手順は以下の通り．

1. 現在点および温度を初期化して，現在点を目的関数で評価する．
2. 現在点に微小でランダムな擾乱を加えることにより近傍点を生成し，近傍点を目的関数で評価する．
3. 近傍点の目的関数値が現在点のそれよりも小さい場合，その近傍点を新たな現在点とする．そうでない場合，目的関数値の差および温度に依存して算出される確率に基づいて，その近傍点を新たな現在点とする．

4. 予め定められた冷却スケジュールに従って温度を下げて手順 2 に進む。

ここで、目的関数値が大きくなる方向への更新確率は、温度が高いほど高くなるように設定される。温度は近傍の範囲を定めるのにも用いられ、温度が高い場合にはより広範囲の近傍が探索される [5]。

本稿では各変数毎に近傍の範囲を 1 次元コーシー分布として定める方法を評価の比較対象として選択し、これに伴って冷却スケジュールも高速なものを使用した。

3. ベンチマーク問題による評価実験

3.1 実験条件

拡張版 ODLS を評価するため、まず代表的なベンチマーク問題である Rastrigin 関数 $R(x)$, Griewank 関数 $G(x)$ の最小化問題への適用実験を実施した。 $R(x)$, $G(x)$ とともに最小値は 0 である。それぞれの数式は以下の通り。但し、 n は変数の数、また $x_i (i = 1, 2, \dots, n)$ は整数変数とする。

$$R(x) = 10n + \sum_{i=1}^n \left\{ \left(\frac{x_i}{100} \right)^2 - 10 \cos \left(\frac{2\pi x_i}{100} \right) \right\} \quad (7)$$

$$G(x) = 1 + \sum_{i=1}^n \frac{x_i^2}{4000} - \prod_{i=1}^n \cos \left(\frac{x_i}{\sqrt{i}} \right) \quad (8)$$

評価実験は、1000 変数および 2000 変数の Rastrigin 関数、Griewank 関数を対象として実施した。

ODLS, MDS, SA の各アルゴリズムが利用できる目的関数評価の回数として、30,000 から 150,000 までの範囲の値を設定した。そして、各最適化試行をそれぞれ 50 回実施し (初期点の各変数は $-512 \sim 511$ の間の一様乱数とした)、各試行で得られた目的関数最小値の平均を算出してアルゴリズムの性能の指標とした。

各最適化アルゴリズムはそれぞれいくつかの調整パラメータを持つが、予備実験によって良い結果が得られたパラメータ値を選択した。

3.2 結果

各目的関数について、目的関数評価回数を横軸に、得られた目的関数最小値の平均を縦軸に取ったグラフを図 1, 図 2 に示す。但し、Griewank 関数についてのグラフは、縦軸を対数スケールとした (図 2)。

4. タンパク質立体構造予測による評価実験

タンパク質はアミノ酸が重合したものであり、天然状態ではエネルギーが最小となるような立体構造を持つと考えられている。そこで、タンパク質立体構造予測は、エネルギー最小化問題として定式化することができる。この問題は解ければ非常に有用であるが、変数の数が多い、またローカルミニマムの数が多いという特徴を持ち、解くことが困難である。

本稿では、タンパク質の立体構造は、二面角 (化学結合により連結した 4 つの原子を A-B-C-D と表した場合、平面 A-B-C と平面 B-C-D とがなす角度)

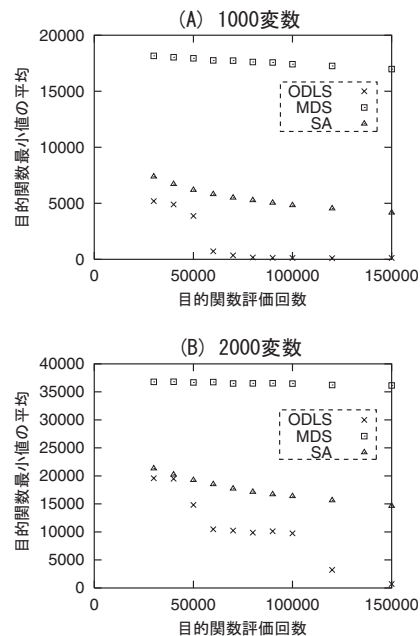


図 1: 評価回数と目的関数の関係 (Rastrigin 関数最小化)

によって一意に定まるものとし、さらに二面角は 1 度刻みで変化し得るものとした。また、エネルギー算出方式としては CHARMM [6] を利用した。

4.1 実験条件

以下の 3 種のタンパク質のエネルギー最小化問題を対象として最適化アルゴリズムを評価した。

- Met-Enkephalin (二面角の数は 23)。
- Oxytocin (二面角の数は 47)。
- HIV Replication Inhibitor (二面角の数は 102)。

各アルゴリズムが利用できるエネルギー評価の回数として、1,000 から 300,000 までの範囲の値を設定した。そして、各最適化試行をそれぞれ 85 回実施し (初期点は予め定めた一定値とした)、各試行で得られたエネルギー最小値の平均を算出してアルゴリズムの性能の指標とした。

各最適化アルゴリズムの調整パラメータは、ベンチマーク問題と同様、予備実験により定めた。

なお、エネルギーの算出、SA の実行には、それぞれフリーソフト [7][8] を使用した。

4.2 結果

各タンパク質について、エネルギー評価回数を横軸に、得られたエネルギー最小値の平均を縦軸に取ったグラフを図 3 に示す。

5. 考察

MDS と ODLS は現在点を 1 回更新するために、変数の数程度の目的関数評価を実施する必要がある。そのため、本稿で実施したベンチマーク問題のように変数の数が多い場合には、現在点の更新回数が少なくなり、いかに正確に目的関数の減少方向を推定できるかが重要になる。Griewank 関数への ODLS 適用に関して探索過程を調査すると、解の改善方向

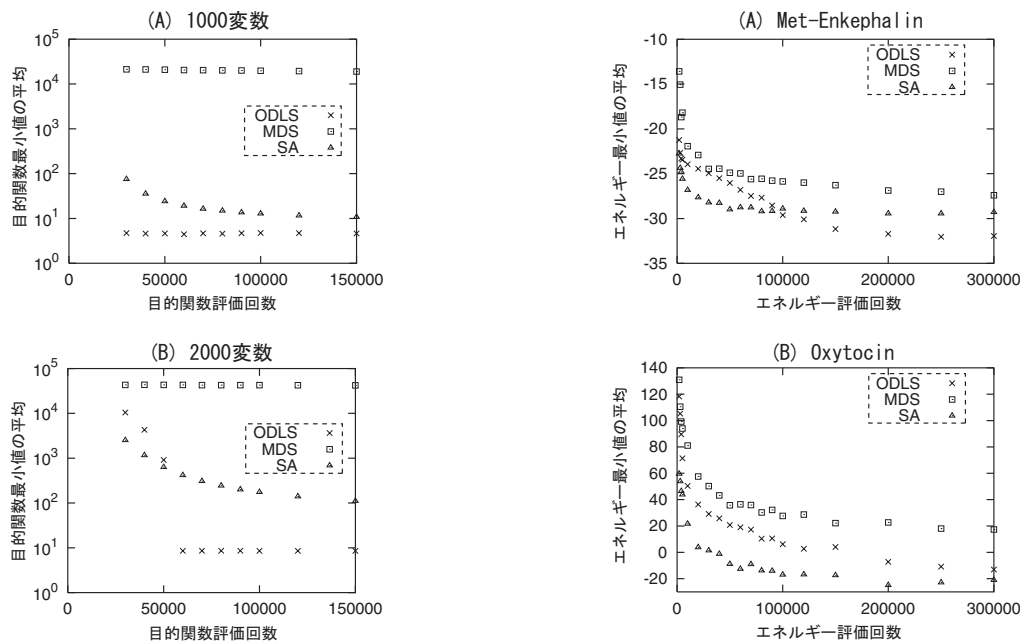


図 2: 評価回数と目的関数の関係 (Griewank 関数最小化)

への移動のみでほぼ収束状態に達していた[§]。従って、変数の数が大きい場合、目的関数の減少方向の推定について、ODLS は MDS よりもはるかに優れていると言える (図 2 参照)。MDS はローカルミニマムからの脱出機能を持たないこともあり、実施したいずれの評価実験についても性能は良くなかった。

Rastrigin 関数への ODLS 適用に関して探索過程を調査すると、一旦ローカルミニマムに陥っても、大きな近傍を利用して脱出できており、結果的に SA よりもはるかに良い性能を示している (図 1 参照)。つまり、ローカルミニマムからの脱出機能は有効に働いている。一方、タンパク質立体構造予測については、全体的には SA の方が ODLS よりも良い性能を示しており、ODLS の目的関数減少方向推定がベンチマーク問題ほどには良好に機能していないことをうかがわせる。ただ、評価回数が大きくなるにつれ、性能面で ODLS が SA に追いつく傾向が見られ、さらなる調査が必要である。

6. おわりに

直交計画を利用する最適化アルゴリズム ODLS について、異なる近傍サイズを利用し、さらに解が悪化する方向への現在点更新を許可することによって拡張し、多峰性関数に対応可能とした。ODLS と SA を代表的なベンチマーク問題とタンパク質立体構造予測に適用して比較評価を実施したところ、ベンチマーク問題では ODLS の性能が圧倒的に良かったが、タンパク質立体構造予測では全体として、SA の方が優れていた。これらの評価結果の違いを解明するため、目的関数の性質と最適化アルゴリズムの挙動について、詳細な調査を実施する必要がある。

[§] 解の改善方向への移動のみで比較的良好な結果に到達したのは、Griewank 関数が大域的には単峰性関数のような形状をしているためであろうと思われる。

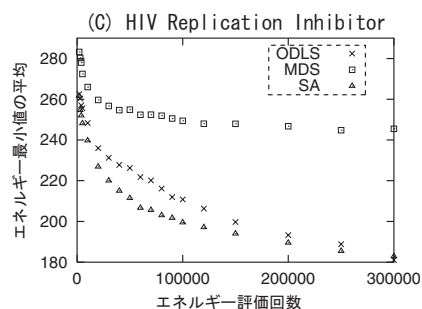


図 3: 評価回数とエネルギーの関係 (タンパク質立体構造予測)

また、ODLS へのローカルミニマムからの脱出機能の組み込み方についても、今後、さらに検討したいと考えている。

参考文献

- [1] V. J. Torczon, Multi-Directional Search: A Direct Search Algorithm for Parallel Machines, Rice University, Houston, Texas 1989.
- [2] 田中他, 直交計画法によるノイズつき多次元関数の勾配推定, 1D-6, 情報処理学会第 61 回全国大会, 2000 年 9 月.
- [3] 田中他, ノイズを用いた局所探索法, 数理モデル化と問題解決研究会, MPS-38-8, 2002 年 3 月.
- [4] 田中, 直交計画法を用いた最適解探索法, 2003 年電子情報通信学会総合大会, 2003 年 3 月.
- [5] L. Ingber, Very Fast Simulated Re-Annealing, Mathematical Computer Modelling, 12, pp. 967-973, 1989.
- [6] B. R. Brooks et al., CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations, Computational Chemistry, vol. 4, pp. 187-217, 1983.
- [7] The TINKER Molecular Modeling Package, <http://dasher.wustl.edu/tinker/>
- [8] SA C++ package by Everett F. Carter, <http://www.taygeta.com/>