直交計画法を用いた局所探索法の改良

3D－05

田中秀俊†，白石将†，川上かおり†，青山功†，佐藤裕幸†
†三菱電機（株）情報総研

1 はじめに

関数最適化問題で導関数が入手困難な場合、局所探索法（Local Search, LS）が有効である。LSは初期値を与えて探索を開始し、探索中は、ある点（現在点）の周囲の点（近傍点）の中から良い点を選んで次の現在点とする。導関数が入手困難な場合は、ランダムに近傍点を選択して現在点と比較する方法がよく採用され、遺伝的アルゴリズム、シミュレーテッドアニーリング（Simulated Annealing, SA）、タブーサーチなどはこれに該当する。一方、直交計画局所探索法（Orthogonal Design Local Search, ODLS）[1]のように、直交計画法を用いて近傍点選択と次の現在点の合成とを行うことにより、勾配を見積もる方法もある。これは近傍点集合に対して変数毎に評価値の平均を比較するため、変数が多い場合には有効である。ODLSの既報[2,3,4,5]における合成方法は、近傍点から合成した点をそのまま次の現在点として評価していた。勾配見積もりという立場からは、その方向へ線形探索する方法も当然考えられる。さらにその線形探索において、目的関数に乱数を加え、関数値の悪化を許容させることも考えられる。本稿ではこれらの試みの効果について、SAと比較した数値実験を示し報告する。

2 ODLSの線形探索

ODLSでは、各変数の＋方向と－方向のどちらが評価関数をよくする方向かを、2水準直交計画の枠組で判断して、それを総合して勾配推定とする。n変数の問題の近傍点集合は、n次元の格子点を考えたときの近傍全体3^n-1個の中からガロア体に基づいて選択した約n点、正確には2^q点ただし$2^{q-1} \leq n < 2^q$、qは正整数、となる。他の局所探索法が近傍点集合を1点ないしはたかだか数十点とすることが多いのに対し、ODLSは例えば1000次元の場合1024点を要する。この多量の関数評価回数に加え、512点ずつの＋と－、計2000通りの平均を算出して1000組の比較をするという大きなオーバーヘッドを抱える。しかし、$3^{1000} \fallingdotseq 10^{478}$もの本来の近傍点の中からランダムにたかだか数十点選択する方法に比べ、着実によい勾配方向が得られると期待できる。またオーバーヘッド部分についても、並列性を利用して簡単に高速化可能である。

既報の比較実験[2,3,4,5]によると、ODLSはSAに比べ、1回の改善幅に対する前述のオーバーヘッドが大きいため、収束速度が遅いという欠点があり、総合的にはSAに劣っていた。そのODLSでは、現在点xのi番目の近傍点集合を$x+\delta_i$とすると、δ_iの成分は＋dもしくは－dとしていた。同様に勾配推定による次の現在点を$x+\lambda$とすると、λの成分は$+d_s$、0もしくは$-d_s$とし、$d_s=d$としていた。本稿では、上記の近傍点集合のdと勾配推定のd_sを異なる値にすること、勾配推定のd_sについて線形探索を用いて最適化すること、の2点について改良を試みた。

また、ODLSは勾配推定をベースとするため、局所脱出の機能がないという欠点があった。そこで、目的関数に一様乱数を加えることにより、線形探索途上で関数値の悪化を許容させることとした。

3 数値実験対象および条件

本稿では比較実験の対象として、ベンチマーク用として有名な、Rastrigin関数R(x)とGriewank関数G(x)の最小化問題を用いた。それぞれの数式を以下に示す。xは整数、どちらの関数も最小値は0である。

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

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

ODLSのdの値は100、d_sは1とした。一様乱数の幅はどちらも5とした。SAはフリーソフト[6]を用い、冷却スケジュールは30.0から0.05までのcauchy冷却、ステップ幅はcauchy分布を用いた。

Improvement of Orthogonal Design Local Search
Hidetoshi Tanaka, Masashi Shiraishi, Kaori Kawakami, Isao Aoyama, Hiroyuki Sato.
Mitsubishi Electric Corporation.
5-1-1 Ofuna, Kamakura, Kanagawa 247-8501, Japan

4 実験結果

図１に、今回改良したODLSと上記のSAとの比較の一例を示す。これはR(x)およびG(x)の最小化問題について、関数の評価回数を５万回に限定した場合にそれぞれのアルゴリズムで得られた解の値である。G10, G100, G1000, G2000はG(x)、R10, R100, R1000, R2000はR(x)の、それぞれ変数の数を10, 100, 1000, 2000としたもので、探索開始点を－512～＋511の間の一様乱数を成分とする点として、探索開始点を変えて10回試行した平均値を示している。図１によると、ODLSはR(x)で変数の数が1000と2000、SAはG(x)で2000、R(x)で1000と2000で、最適解0に到達せず終了している。

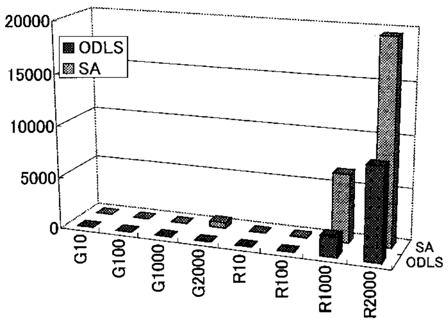

図１　解の近似度の比較（評価回数５万回）

5 考察

局所解の多いR(x)、G(x)ともに、５万回限定の比較でODLSはSAを平均的に上回っている。ODLSに持たせた局所解回避の機能は十分な効果が得られていると考えられる。ただし、ODLSではオーバーヘッドが大きいため、同じ評価回数でも実際の計算時間は、例えば2000変数の場合SA約80秒、ODLS約2700秒（PentiumIII 933MHz）と30倍以上である。

SAでは悪化のケースで初めて乱数が活躍するが、今回のODLSでは近傍点評価から線形探索まで一貫して乱数を用いた。近傍点評価に乱数が入っていても、変数の数が十分多ければ勾配推定に大きな影響はないと考えられる。しかし、線形探索でも常に乱数を用いるため、元の目的関数が探索方向に沿ったある区間において単調減少であっても、その区間の途中で停止するおそれがある。これは現状での局所解脱出機能の副作用であり、この回避法は今後検討する。

R(x)、G(x)は、２次関数 x^2 に周期関数を付加することによって多量の局所解を持つ関数に仕上げている点は共通している。差異はその周期および振幅にある。R(x)は、G(x)に比べて大振幅、長周期を持つ周期関数が付加されている。このような性質を持つ関数に対して、d=100のような大きなdで勾配を推定する際は、間接的に次の２次関数 x^2 の性質が利用されている。

$$f(x+d) > f(x-d) \Rightarrow x > 0$$
$$f(x+d) < f(x-d) \Rightarrow x < 0$$

これは、勾配推定が大局的に見てほぼ常に正しい方向を示すということを意味する。dが小さい場合にはR(x)やG(x)の周期関数部分が効いて、成り立つことが稀だが、dが十分大きい場合には成り立ちやすい。この性質が崩れるような目的関数でどの程度悪い結果になるかは、確認する必要がある。

6 まとめ

直交計画を用いた勾配推定による関数最適化手法に、線形探索と乱数化の２点の改良を施した。その結果、ベンチマーク問題であるRastrigin関数、Griewank関数の変数の数が多い場合に、同一評価回数という条件下で、シミュレーテッドアニーリングよりも平均的によい近似解が得られることが分かった。

参考文献

[1] 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”, 第61回情処全国大会, 1D-06, 2000-9.
[2] 田中，他: “並列局所探索法の比較評価”, 第62回情処全国大会, 2R-04, 2001-3.
[3] 田中，他: “分子ポテンシャル最小化問題に関する並列局所探索法の比較評価”, 数理モデル化と問題解決研究会, MPS-33-12, 2001-3.
[4] 白石，他: “タンパク質立体構造予測問題に関する最適化アルゴリズムの比較評価”, 第63回情処全国大会, 2P-01, 2001-9.
[5] 青山，他: “多峰性関数に対する最適化アルゴリズムの探索性能比較”, 第63回情処全国大会, 2P-04, 2001-9.
[6] SA C++ package, http://www.taygeta.com/