

自己組織化サンプリングの提案とその最適化手法 -多峰性関数への適用-

Self-Organized Sampling as an Optimization Method: An Approach towards Multimodal Functions

比護 貴之 高玉 圭樹
Takayuki Higo Keiki Takadama

東京工業大学 大学院総合理工学研究科 知能システム科学専攻
Department of Computational Intelligence and System Science
Interdisciplinary Graduate School of Science and Engineering
Tokyo Institute of Technology

This paper improves Self-Organized Sampling (SOS) method, which is our proposed optimization method for multimodal functions, by using new parameter instead of the inverse temperature parameter on Boltzman distribution that causes the problem of an appropriate parameter setting. Through intensive simulations, we have found an effectiveness of the proposed parameter. Concretely, SOS with the proposed parameter can optimize both the original objective function and the linearly scaled function using the same parameter value, while SOS with the inverse temperature parameter cannot always optimize both functions using the same parameter value.

Keywords: global optimization, multimodal function, boltzman distribution, annealing, inverse temperature

1. はじめに

近年最適化問題の中でも特に多峰性関数最適化が盛んに研究されている。多峰性関数最適化問題に対する有効なアプローチとして、統計量の遺伝に着目した遺伝的アルゴリズム (Genetic Algorithms: GA)[Kita 1999] と、GA と関連性の強い分布推定アルゴリズム (Estimation of Distribution Algorithms: EDA)[Larranaga 2001] が挙げられる [佐久間 2003]。統計量の遺伝に基づく GA では、理想的な交叉アルゴリズムである SPX [樋口 2001] が提案され有効性が確認されているが、多峰性関数最適化においては問題が残る。一方、EDA では探索点の背景にある確率分布を推定し、その分布に従うサンプルを発生させるという設計指針に基づいたアルゴリズムが研究されている。

これらに対して、著者らの提案する自己組織化サンプリング (Self-Organized Sampling: SOS) は、サンプルの期待評価値とエントロピーを最大化する点で最適なアルゴリズムであり、一部の多峰性関数最適化問題に対して有効性が確認されつつある [比護 2004]。しかし、この手法には SOS のパラメータである逆温度を適切に設定することが難しいという欠点がある。逆温度は探索点の収束度合を決めるパラメータであるため、最適化効率に大きな影響をあたえる。しかしながら、ある逆温度における探索点の収束度合は目的関数に依存する上、見積もることが困難であるため、試行錯誤を要する。

そこで、本研究では逆温度の問題を解決するために、逆温度の代替パラメータと、それをを用いた計算アルゴリズムを提案する。また、この改良により SOS が目的関数の線形スケールングに対して不変なアルゴリズムとなることを示す。

以下 2 章では、SOS とそれをを用いた最適化手法及びその問題点についてまとめる。3 章では逆温度の代替パラメータを提

案し、それを SOS に組み込む方法を説明する。4 章では実験を行い、5 章で考察を行う。そして、6 章で本研究をまとめる。

2. 自己組織化サンプリング (SOS)

2.1 SOS の特徴

SOS の特徴としては次の 3 つが挙げられる。

- 目的関数のボルツマン分布に従うサンプリング。
- 目的関数を階段関数で近似する。
- 確率密度和を一定にするような分割を行う。

これらについて以下で簡単な説明を行う。まず、ボルツマン分布は次のような分布である。

$$p(x) = \frac{\exp(-f(x)\beta)}{Z} \quad (1)$$

$$Z = \int \exp(-f(x)\beta) dx \quad (2)$$

$f(x)$ はエネルギー関数、 β は逆温度、 Z は分配関数と呼ばれる。ボルツマン分布の特徴はエネルギー関数の期待値 $\int f(x)p(x)dx$ を最小化し、エントロピー $-\int p(x) \log p(x) dx$ を最大化する分布であり、そのトレードオフは逆温度で決まる [深尾 1987]。エネルギー関数に最小化問題の目的関数を用いる場合 (最大化問題の場合は目的関数を -1 倍する) ボルツマン分布の特徴は最適化において良い性質と考えられる。

一般的にボルツマン分布を計算することは困難であり、既存手法にマルコフ連鎖モンテカルロ法 [伊庭 2003] が挙げられるがいくつかの問題がある。そこで、SOS は領域を単体に分割し、単体内部では確率密度が一定となるようなモデルで近似する。このモデルは、ボルツマン分布のエネルギー関数を階段関数とすることで構成できる。

近似においてはその精度が問題となる。そこで、それぞれの単体領域における確率密度の和を一定化するように、単体へ

連絡先: 比護 貴之, 東京工業大学 大学院総合理工学研究科 知能システム科学専攻, 〒 226-8502 神奈川県横浜市緑区長津田町 4259 番地, Tel:045-924-5216, higo@cas.dis.titech.ac.jp

の分割を変えることで、確率密度の高い領域を正確に近似し、確率密度の低い領域を粗く近似する。

2.2 アルゴリズム

まず、初期化について説明する。SOS は線形制約を想定しており、それは凸多面体となる。そこで、凸多面体の頂点に疑似探索点を配置し、内部に探索点を適当に配置する。探索点と疑似探索点の違いは、探索点が最適化の過程でその位置を修正するのに対し、疑似探索点は位置を変えない。そして、探索点と疑似探索点に対してドロネ三角形分割 [今井 2003][譚 2001] を適用する。ドロネ三角形分割により探索点及び疑似探索点の間に枝が定義され、探索領域が単体に分割される。単体というのは d 次元の空間において $d+1$ 個の頂点を持つ多面体を指す。

次に、最適化過程について説明する。最適化過程ではランダムに探索点を選び、それを適切な位置、つまり全ての単体の確率密度和が一定となるような位置へ修正することを繰り返す。ある探索点を適切な位置へ修正するためには、その探索点に隣接する単体から M 個のサンプルを発生させ、そのサンプル情報を用いて次のように更新を行う：(1) 探索点に隣接する単体 s_i の重心 $c(s_i)$ を求める。重心とは頂点の座標を平均したものを指す。(2) サンプルから単体 s_i の平均確率密度 $P(s_i)$ を求める。平均確率密度は次のように計算する。

$$p(s_i) = \sum_{j=1}^M \frac{1}{M} \exp(-f(x_{ij})\beta) \quad (3)$$

ただし、 x_{ij} は単体 s_i の内部から一様に発生させた M 個のサンプルであり、 $f(x)$ は最小化問題の目的関数とする。単体から一様にサンプリングする方法は [樋口 2001] を参照されたい。(3) 探索点を次のように更新する。

$$x = \sum_{s_i \in N_x} \frac{p(s_i)}{\sum_{s_k \in N_x} p(s_k)} c(s_i) \quad (4)$$

ただし、 N_x は探索点 x に隣接する単体の集合である。最後に、アルゴリズムの手順を図 1 に示す。

- 初期化
 1. 線形制約が作る凸多面体の頂点に疑似探索点を配置する。
 2. 凸多面体内部に探索点を配置する。
 3. 探索領域をドロネ三角形分割する。
 4. x_i の隣接単体集合 N_{x_i} を求める。
- 最適化
 1. 探索点 x_p をランダムに選ぶ。
 2. 隣接単体集合 N_{x_p} に含まれる単体の内部から一様に M 個サンプルを発生させる
 3. 探索点 x_p を式 (4) に従って更新する。
 4. 1 に戻る

図 1: SOS アルゴリズム

2.3 アニーリング

アニーリングとは、逆温度の低い状態から高い状態へゆっくり変化させる操作を意味し、SOS では局所解へ陥ることを防ぐ役割をする。逆温度のアニーリングは焼き鈍し法 (Simulated Annealing: SA)[白石 2002] でも登場する。SA では指数的に逆温度を変えるアニーリングがよく利用され、それは次のように書ける。

$$\beta_t = \beta_0 r^t \quad (5)$$

ただし、 t は時間 (関数評価回数) を表す。この方法の有効性は経験的に知られているが、初期値 β_0 と変化率 r をどのように設定すべきかという問題がある。直感的には、ボルツマン分布が一様分布に近い逆温度を初期値とし、ボルツマン分布がある一点のみで確率 1 である分布とみなせるような逆温度を終了値とし、そこから変化率 r を見積もればよいのだが、初期値と終了値は目的関数に依存している上に、見積もるのが困難である。また、初期値と終了値の範囲は 0 から無限大の範囲にあり、あまりにも大きい逆温度を計算機で扱うのは難しい。このような問題が SOS に内在している。

3. 接近率の提案

逆温度の問題を解決するために、次のような代替パラメータ γ を提案する。以下では目的関数 $f(x)$ の最大化問題を想定して話を進める。最大化問題ではボルツマン分布の分子が $\exp(f(x)\beta)$ となることに注意されたい。

$$\gamma = \frac{\exp(f(x_u)\beta)}{\exp(f(x_u)\beta) + \exp(f(x_v)\beta)} \quad (6)$$

$$= \frac{1}{1 + \exp(\delta\beta)} \quad (7)$$

ただし、 x_u, x_v は適当に決めるとし、 $\delta = f(x_v) - f(x_u)$ とする。 γ はある 2 点の評価値の比を表している。これは SOS のアルゴリズムにおける探索点の更新において、ある探索点を更新する際に隣接する単体の重心への近付き方の比と解釈できる。その様子を図 2 に示す。以下ではこのパラメータ γ を接近率と呼ぶことにする。

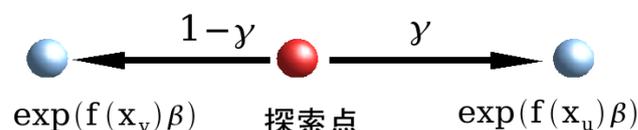


図 2: 接近率の解釈

逆温度は式 (7) から接近率を用いて次のように書ける。

$$\beta = \frac{1}{\delta} \log\left(\frac{1}{\gamma} - 1\right) \quad (8)$$

$f(x_v) < f(x_u)$ を仮定すると、 $\delta < 0$ となり、この時 $\gamma = \frac{1}{2}$ で $\beta = 0$ 、 $\gamma = 1$ で $\beta = \infty$ になることがわかる。従って、 γ を $\frac{1}{2}$ から 1 まで動かすことでアニーリングが行える。また、 $\gamma = 1$ の時 $\beta = \infty$ となることから探索点が収束することは明らかである。つまり、 $\gamma = 1$ のときは、少しでも評価値に差があれば探索点は評価値の高い方へ完全に移動するため、収束することが保証される。ただし、(準)最適解が二つありその評価値が全く同じ場合は例外である。

次にこれをアルゴリズムに組み込む方法を説明する。逆温度 β が必要となるのは式 (4) の計算である。式 (4) は各単体の重心ベクトル $c(s_i)$ を単体の平均確率密度 $p(s_i)$ で平均することを意味しており、 $p(s_1) : p(s_2) : \dots : p(s_n)$ の比がわかればよい。つまり、各 i, j に対して $\frac{p(s_j)}{p(s_i)}$ を計算できればよい。これは次のように計算できる。

$$\frac{p(s_j)}{p(s_i)} = \frac{\exp(\bar{f}_j \beta)}{\exp(\bar{f}_i \beta)} \quad (9)$$

$$= \exp((\bar{f}_j - \bar{f}_i) \beta) \quad (10)$$

$$= \exp\left(\frac{\bar{f}_j - \bar{f}_i}{\delta} \log\left(\frac{1}{\gamma} - 1\right)\right) \quad (11)$$

$$= \left(\frac{1}{\gamma} - 1\right)^{\frac{\bar{f}_j - \bar{f}_i}{\delta}} \quad (12)$$

ただし、 \bar{f} はサンプルから求まる単体の平均評価値である。また、あらかじめ $p(s_i)$ を大きい順に並び替え、最も大きいものを 1 と置き、大きい順に比を求めていけば値は発散しない。ただし、 δ が $\bar{f}_j - \bar{f}_i$ と比べて小さくなり過ぎると $\frac{\bar{f}_j - \bar{f}_i}{\delta}$ が無限大になるので注意されたい。

最後に、便宜上 $\delta = f(x_u) - f(x_v)$ と置き直せば (符号を反転させれば) 式 (12) は次のように書ける。

$$\frac{p(s_j)}{p(s_i)} = \left(\frac{1}{\gamma} - 1\right)^{\frac{\bar{f}_i - \bar{f}_j}{\delta}} \quad (13)$$

以後 δ はこの定義を用いることにする。

4. 実験

4.1 問題

次のような関数を用いる。

$$f(x) = 0.5(\cos((x - c) * 400) - 1) + 10g(x) \quad (14)$$

$$g(x) = \begin{cases} \frac{x}{c} & x < c \\ \frac{1-x}{1-c} & c < x \end{cases} \quad (15)$$

ただし、 $c = 0.3$ で、 $0 < x < 1$ とする。この関数は図 3 のようになる。本実験ではこの関数の最大化問題を考える。

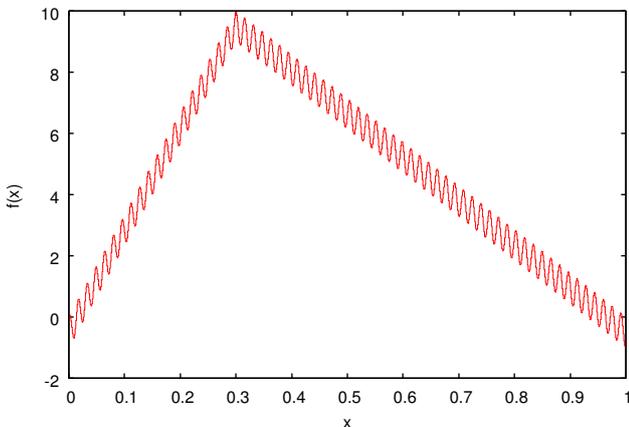


図 3: テスト用の関数

4.2 設定

先ほど定義した関数に対して、探索点数 15、サンプル数 1 の SOS を適用し 6 万回のサンプリングを行う。ただし、サンプル数というのは単体内部から一度に発生させるサンプルの数である。この条件のもと以下で示す 4 つの実験を行う。

実験 1 逆温度を 10^{-2} から 10^{10} まで式 (5) で変化させる。

実験 2 目的関数を 10^{-6} 倍した関数 $g(x) = 10^{-6} f(x)$ に対して、逆温度を 10^{-2} から 10^{10} まで式 (5) で変化させる。

実験 3 接近率を 0.5 から 1 まで式 (16) で変化させる。

実験 4 目的関数を 10^{-6} 倍した関数 $g(x) = 10^{-6} f(x)$ に対して、接近率を 0.5 から 1 まで式 (16) で変化させる。

$$\gamma_t = 0.5 + \frac{0.5}{60000} t \quad (16)$$

接近率における δ の値は、探索点を更新する際に平均確率密度の 1 番大きい単体と、2 番目の単体の平均評価値 $\bar{f}_{fst}, \bar{f}_{sec}$ を使って次のように更新する。

$$\delta_{t+1} = \delta_t + \alpha((\bar{f}_{fst} - \bar{f}_{sec}) - \delta_t) \quad (17)$$

ただし、 $\alpha = 0.05$ とする。

4.3 結果

実験 1、2 の結果を図 4、5 に示す。図の縦軸は探索空間、横軸は逆温度を表し、最適化過程でサンプリングした点をプロットする。実験 1 では、逆温度が 10^{-2} から既に収束が始まり、 10^2 でほぼ収束しているのがわかる。一方、実験 2 では逆温度が 10^4 辺りから収束が始まり、 10^7 でほぼ収束しているのがわかる。実験 1 と 2 で収束の始まりと終わりの逆温度が異なるのは、目的関数を 10^{-6} 倍しているためである。従って、逆温度の初期値や終了値は関数のスケールに依存すると言える。

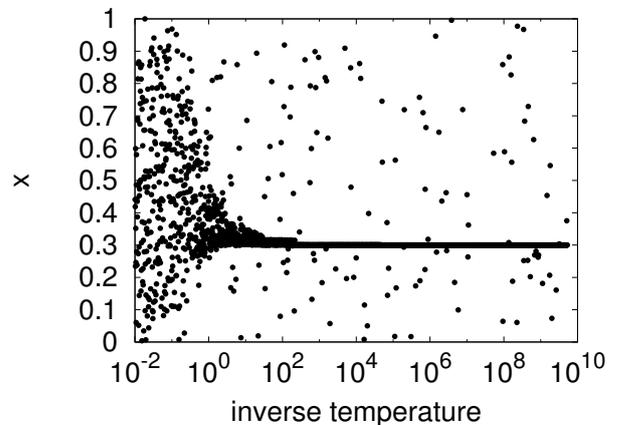
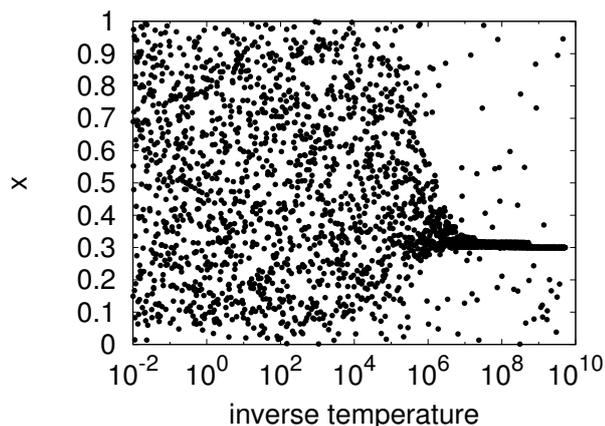
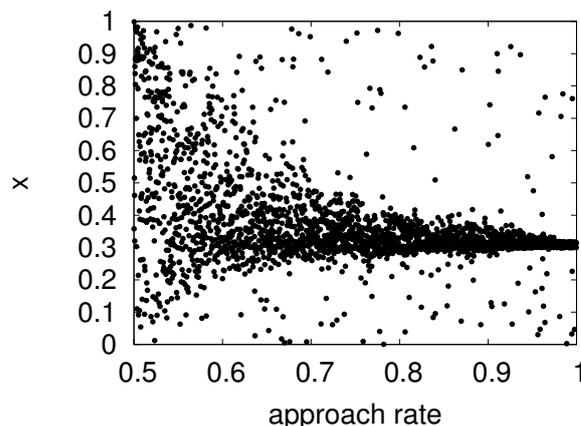


図 4: 逆温度を用いた場合。

次に実験 3、4 の結果を図 6、7 に示す。横軸が接近率になっていることに注意されたい。実験 3、4 を比較すると結果に大きな差は見られないので、目的関数のスケールに依存しないと予想される。接近率が 0.5 の付近では一様な探索をしており、

図 5: 目的関数を 10^{-6} 倍し、逆温度を用いた場合。図 7: 目的関数を 10^{-6} 倍し、接近率を用いた場合。

接近率を 1 に近づけるに従い徐々に $x=0.3$ 付近に収束していきることがわかる。これより、目的関数によらず接近率を 0.5 から 1 へゆっくり変化させることで、アニーリングが実現でき、接近率は逆温度に比べて利便性が高いことがわかる。

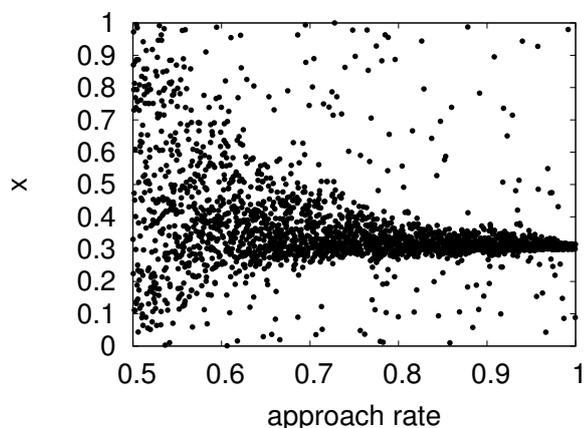


図 6: 接近率を用いた場合。

5. 考察

5.1 線形スケールに対する不変性

実験より、接近率を用いた SOS は目的関数の線形スケールリング、すなわち目的関数 $f(x)$ を $af(x) + b$ と置き直しても、もとの目的関数に適用した場合と同じように動作すると予想される。これは式 (13) に、 $\delta = f(x_u) - f(x_v)$ を代入し、関数 $f(x)$ を $af(x) + b$ と置き換えることで値が変わらないことから証明できる。より一般的に $\delta_t = \sum_{i=1}^t w_i (f(x_{ut}) - f(x_{vt}))$ を考えよう。ただし、 $\sum_t w_t = 1$ である。これは、アルゴリズムの過程で得られた x_{ut} 、 x_{vt} に対して、 δ を計算し、 w で加重平均したものである。これを式 (13) に代入し目的関数を置き換えれば同様に値が変わることがわかる。そして、本実験で用いた δ の更新式 (17) もこれに含まれる。

5.2 問題点

本論文の実験では接近率を線形的に変化させたが、この方法が最適である保証はない。従って、どのように接近率を調整するかは今後の課題となるが、No Free Lunch Theorem

[Wolpert 1995] のような結論も出ており慎重な議論が必要である。

6. おわりに

本論文では SOS で扱うボルツマン分布に表れる逆温度が応用上問題となることに着目し、代替パラメータである接近率を提案した。そして、逆温度と接近率の比較実験によりその有効性を示すとともに、接近率を用いた SOS が目的関数の線形スケールリングに対して不変であることを示した。

今後の課題としては、初期化における三角形分割の計算量が次元の指数オーダーとなるため、これに代わるアルゴリズムを提案することが先決である。また、接近率のアニーリングを調整するアルゴリズムの提案も課題の一つである。

謝辞

本研究の一部は文部科学省の科学研究費補助金 (基盤研究 (B), 課題番号 17360424) の支援によって行われた。

参考文献

- [深尾 1987] 深尾 毅: 分散熱力学システム論, 昭晃堂 (1987).
- [比護 2004] 比護 貴之, 高玉 圭樹: “適応的探索領域に基づく最適化手法の提案: 多峰性関数へのアプローチ”, 計測自動制御学会, システム・情報部門 学術講演会 2004, 95/100 (2004).
- [樋口 2001] 樋口 隆英, 筒井 茂義, 山村 雅幸: “実数値 GA におけるシンプレクス交叉の提案”, 人工知能学会論文誌, 16, 1Q, 147/155 (2001).
- [伊庭 2003] 伊庭 幸人: ベイズ統計と統計物理, 岩波書店 (2003).
- [今井 2003] 今井 浩, 今井 圭子: 計算幾何学, 共立出版 (2003).
- [Kita 1999] Kita, H. and Yamamura, M.: “A Functional Specialization Hypothesis for Designing Genetic Algorithms”, *Proc.1999 IEEE International Conference on Systems, Man, and Cybernetics*, 579/584 (1999).
- [Larranaga 2001] Larranaga P., Lozano, J.A.: *Estimation of Distribution Algorithms: A New Tool for Evolutionary Optimization*. Kluwer Academic Publishers, Boston (2001).
- [佐久間 2003] 佐久間 淳, 小林 重信: “確率分布推定に基づく実数値 GA の新展開”, 人工知能学会誌, Vol.18, No.5, 479/486 (2003).
- [白石 2002] 白石 洋一: 組合せ最適化アルゴリズムの最新手法, 丸善株式会社 (2002).
- [譚 2001] 譚 学厚, 平田 富夫: 計算幾何学入門, 森北出版 (2001).
- [Wolpert 1995] Wolpert, D.H. and Macready, W.G.: “No Free Lunch Theorems for Search”, *Technical Report SFI-TR-95-02-010, Santa Fe, NM* (1995).