

関数空間におけるメトロポリス法のための局所的提案分布

Local Proposal Distributions for Metropolis Algorithm in Function Space

嶋本 義己†
Yoshiki Shimamoto

末松 伸朗†
Nobuo Suemotsu

岩田 一貴†
Kazunori Iwata

1 はじめに

近年、マルコフ連鎖モンテカルロ (MCMC) 法による関数のノンパラメトリック推定が行われるようになってきている [1]. しかし、実際上高次元空間のマルコフ連鎖となるため、複雑な関数を推定しようとする精度が落ちるという問題が生じる. そこで本稿では、候補の散布度に加えて、MCMC 法に局所的提案分布を導入して変化を与える領域を調整できるようにすることにより、複雑な関数に対する推定精度を向上させる手法を提案する. また、回帰問題においてその有効性を検証したので、その結果を報告する.

2 ガウス過程 [2]

確率的に定まる関数 $f(x)$ が平均関数 $m(x)$, 共分散関数 $k(x, x')$ のガウス過程に従うとき、

$$f(x) \sim \mathcal{GP}(m(x), k(x, x'))$$

と書く. このとき、任意の N , 任意の $\mathbf{x}=[x_1, \dots, x_N]^T$ に対して、 $\mathbf{f} = [f(x_1), \dots, f(x_N)]^T$ は平均 $\mathbf{m}=[m(x_1), \dots, m(x_N)]^T$, 共分散行列 $K(\mathbf{x}, \mathbf{x})$ の正規分布に従う. ここで、 $K(\mathbf{x}, \mathbf{x}) = [k(x_i, x_j)]_{N \times N}$ である. 本研究では二乗指数型共分散関数

$$k(x, x') = v \exp\left(-\frac{(x-x')^2}{2l^2}\right)$$

を用いる. ここで、 v は分散パラメータ、 l はスケールパラメータである.

3 MCMC 法による関数推定

データ D に対する回帰問題において、区間 $[a, b]$ 上に定義される関数 $f(\cdot)$ を M 個の標本点により、 $\mathbf{f} = [f(s_1), \dots, f(s_M)]^T$ と表現し、これを MCMC 法で推定を行うことを考える. 本研究ではメトロポリス法を用いる. $f(\cdot)$ がガウス過程事前分布に従うと仮定すると、その最も基本的な方法では、候補を

$$\mathbf{f}' = \mathbf{f} + \beta \boldsymbol{\delta} \quad (\boldsymbol{\delta} \sim \mathcal{N}(\mathbf{0}, K(\mathbf{s}, \mathbf{s}))) \quad (1)$$

により生成し、確率 $\alpha(\mathbf{f}, \mathbf{f}')$ で受理するという操作を繰り返す. ただし、 β は候補の散布度を調整するパラメータであり、 $\mathcal{N}(\mathbf{0}, K(\mathbf{s}, \mathbf{s}))$ は $f(\cdot)$ の事前ガウス過程により定まる \mathbf{f} の事前正規分布である. また、 $P(\mathbf{f}) = P(D|\mathbf{f})\mathcal{N}(\mathbf{f}|\mathbf{0}, K(\mathbf{s}, \mathbf{s}))$ として、 $\alpha(\mathbf{f}, \mathbf{f}') = \min\{1, P(\mathbf{f}')/P(\mathbf{f})\}$ である.

† 広島市立大学大学院情報科学研究科

4 提案手法

前述したメトロポリス法では、高次元ベクトル \mathbf{f} 全体を一度に変更することによって次の候補を生成している. そのため、複雑な関数を推定すると精度が落ちるという問題が起こる. この問題は、式 (1) における候補を調整するパラメータ β を調整することである程度改善することが期待できるが、必ずしもそれで上手くいくというわけではない. β を大きくすると候補の受理率が低下し、推定精度が落ち、逆に、 β を小さくすると関数 \mathbf{f} に与える変化量が小さくなるため、目標の分布に収束するまでに時間が掛かってしまう. そこで本稿では、この問題を軽減するため、 $[a, b]$ 内にウィンドウを設定し、 \mathbf{f} をウィンドウに限定して更新を繰り返す手法を提案する.

4.1 提案アルゴリズム

提案するメトロポリスアルゴリズムをアルゴリズム 1 にまとめる. 4 行目でランダムに定めた位置にウィンドウを生成し、5 行目で、生成したウィンドウに限定して \mathbf{f} の更新を行っている. ウィンドウの具体的な生成方法については次節で説明する.

アルゴリズム 1 提案アルゴリズム

- 1: 初期値 $\mathbf{f}^{(0)}$ を設定する
- 2: **for** $t = 1 : T$ **do**
- 3: 正規分布 $\mathcal{N}(\mathbf{0}, K(\mathbf{s}, \mathbf{s}))$ より $\boldsymbol{\delta}$ を引き出す
- 4: 区間 $[a_t, b_t]$ をランダムに決め、その区間内にウィンドウを持ち、区間外はゼロとなるような重みベクトル \mathbf{w}_t を生成
- 5: $\mathbf{f}' \leftarrow \mathbf{f}^{(t-1)} + \beta \mathbf{w}_t \circ \boldsymbol{\delta}$
// $\mathbf{w}_t \circ \boldsymbol{\delta}$ は要素ごとの積
- 6: 確率 $\alpha(\mathbf{f}^{(t-1)}, \mathbf{f}')$ で $\mathbf{f}^{(t)} \leftarrow \mathbf{f}'$
- 7: **end for**

4.2 ウィンドウの生成方法

まず、生成するウィンドウ t の区間 $[a_t, b_t]$ を $[a, b]$ 内からランダムに決める. ただし、区間 $[a_t, b_t]$ の幅は一定となるように定める. そして、 $[a_t, b_t]$ 内にウィンドウ関数を用いてウィンドウを生成する. ウィンドウ関数は、連続で滑らかな単峰の関数を用いる. 用いることができるウィンドウ関数の例として、正規分布型の密度関数 (式 (2)) やベータ分布型の密度関数 (式 (3)) がある. ただし、 $\xi_t(x) = (x - a_t)/(b_t - a_t)$ である. また、 c はウィンドウ関数のパラメータ、 C はウィンドウ関数の最大値を 1 とする定数である. 本稿では、式 (2)、式 (3) により生成されるウィンドウをそれぞれガウス窓、ベータ窓と呼ぶことにする. ガウス窓、ベータ窓の例を図 1、図 2 にそれぞれ

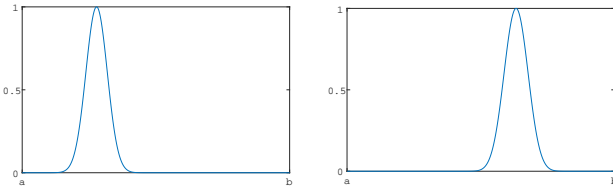


図 1 ガウス窓の例

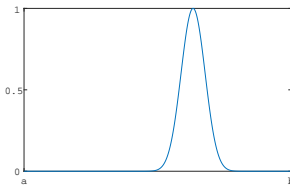


図 2 ベータ窓の例

れ示す.

$$w_t(x|\mu, \sigma) = \begin{cases} C \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) & \text{if } x \in [a_t, b_t] \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

$$w_t(x|c) = \begin{cases} C \xi_t(x)^{c-1} (1 - \xi_t(x))^{c-1} & \text{if } x \in [a_t, b_t] \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

5 実験

データ $D = \{(x_n, y_n) | n = 1, \dots, N\}$ に対する回帰問題を考える. ここで, $y_n = f(x_n) + \varepsilon_n$ であり, ε_n は平均 0, 分散 σ^2 の正規分布に従うノイズである. ガウス過程事前分布の下で未知関数 $f(\cdot)$ を推定するのに MCMC 法は必要ないが, あえて MCMC 法を行うことを考える. これにより得られた推定結果と通常の高ス過程回帰によって得られた結果との誤差を求め, 基本的なメトロポリス法を用いる手法 (基本手法) と提案手法の性能の比較を行う. また, 提案手法については, ベータ窓 (式 (3)) を用いる方法 (提案手法 1) と高ス窓 (式 (2)) を用いる方法 (提案手法 2), それぞれについて実験を行う.

5.1 ガウス過程回帰

未知関数 $f(\cdot)$ を高ス過程事前分布 $\mathcal{GP}(0, k(\cdot, \cdot))$ の下で推定すると, $f(\cdot)$ の事後分布は

$$f(\cdot) | D \sim \mathcal{GP}(\hat{f}(\cdot), k'(\cdot, \cdot)).$$

ここで,

$$\begin{aligned} \hat{f}(x) &= \mathbf{k}(x, \mathbf{x})K(\mathbf{x}, \mathbf{x}) + \sigma^2 I^{-1} \mathbf{y}, \\ k'(x, x') &= k(x, x') - \mathbf{k}(x, \mathbf{x})K(\mathbf{x}, \mathbf{x}) + \sigma^2 I^{-1} \mathbf{k}(\mathbf{x}, x'). \end{aligned} \quad (4)$$

ただし,

$$\begin{aligned} \mathbf{x} &= [x_1, \dots, x_N]^T, \\ \mathbf{y} &= [y_1, \dots, y_N]^T, \\ \mathbf{k}(x, \mathbf{x}) &= [k(x, x_1), \dots, k(x, x_N)] = \mathbf{k}(\mathbf{x}, x)^T. \end{aligned}$$

MCMC 法により得られた関数 $f(\cdot)$ の事後分布の平均は, 式 (4) に漸近的に一致するはずである.

5.2 結果

誤差の評価指標として, RMS (二乗平均平方根) を用いた. 推定対象となる関数を図 3 に示す. 未知関数 $f(\cdot)$ は $M = 720$ 個の標本点により離散化して表し, 候補の散布度を調整するパラメータ β は各手法において最適な値を用いた. それぞれの手法でサンプリングを 10000 回行った.

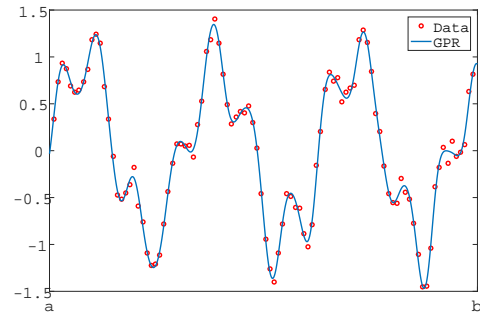


図 3 推定した関数

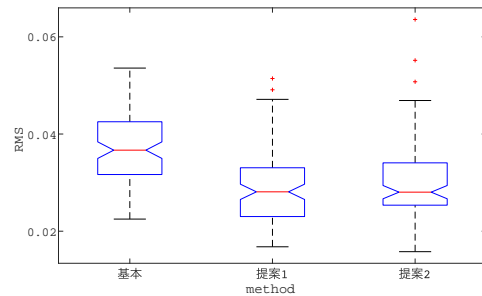


図 4 推定誤差のボックスプロット

表 1 推定誤差

	平均 RMS
基本	3.84×10^{-2}
提案 1 (ベータ窓)	2.69×10^{-2}
提案 2 (高ス窓)	2.82×10^{-2}

各手法により推定された関数と高ス過程回帰により推定された関数の RMS を表 1 に, 各手法を 100 回実行した時の RMS のボックスプロットを図 4 に示す. それらより, 提案手法は基本手法と比べ, 精度が向上していることが分かる. また, 提案手法 1, 2 については推定性能に大きな差がないことが分かる.

6 まとめ

MCMC 法に局所的提案分布を導入することにより, 複雑な関数に対する推定を向上させる手法を提案した. また実験より, ガウス過程回帰問題において提案手法の有効性を示した.

参考文献

- [1] BESKOS, A. and STUART, A. MCMC methods for sampling function space, 6th Int. Cong. on Industrial and Applied Math. (2009).
- [2] MACKAY, D. J. INTRODUCTION TO GAUSSIAN PROCESSES, Technical report, Cambridge University (1998).