

確率微分方程式のモデル推定の双対性に基づいた数値計算法の提案 An estimation method based on duality in stochastic differential equation

荒井 祐樹[†]
Yuuki Arai

大久保 潤[†]
Jun Ohkubo

1. はじめに

時系列データにおいて、次の時刻の状態を予測する方法に、データを確率的モデルに従って生成されているものと仮定し、そのモデルを説明できるパラメータを推定するという方法がある。その確率的モデルの中でも確率微分方程式をモデルとしている現象は、金融市場や交通流、人間の心拍動を始めとした社会科学や自然科学の様々な分野でよく見られる [1]。

確率微分方程式のモデルのパラメータ推定は、条件付きモーメントに基づいて行うことができ、その際に Fokker-Planck 方程式を用いた時間発展を利用することができる [2]。しかし、この方法で単純に数値計算を行う場合、計算時間が多大になってしまうおそれがある。本研究では、Fokker-Planck 方程式の双対性を利用した、大幅な計算時間の削減方法を提案する。

2. 確率微分方程式のモデル推定

2.1. 確率微分方程式

確率微分方程式は、揺らぎに相当する Wiener 過程 $W(t)$ を用いて、

$$dx = \mu(x(t), t)dt + \sigma(x(t), t)dW(t) \quad (1)$$

のように表される。 $\mu(x(t), t)$ は時系列データの平均的な動きに関する drift 係数、 $\sigma(x(t), t)$ は分散に関する diffusion 係数である。なお、各係数は定常過程においては座標 $x(t)$ のみに依存する関数となる。時刻 t における確率過程 $x(t)$ の確率密度関数を $P(x, t)$ としたとき、演算子記号 \hat{L} を用いた、

$$\begin{cases} \frac{\partial P(x, t)}{\partial t} = \hat{L}(x, t)P(x, t) \\ \hat{L}(x, t) = -\frac{\partial}{\partial x}D^{(1)}(x, t) + \frac{\partial^2}{\partial x^2}D^{(2)}(x, t) \end{cases} \quad (2)$$

によって Fokker-Planck 方程式を表す。 $D^{(1)}(x, t)$ 、 $D^{(2)}(x, t)$ は係数であり、それぞれ、確率微分方程式における式 (1) の drift 係数 $\mu(x(t), t)$ と、diffusion 係数 $\sigma(x(t), t)$ とは

$$\mu(x(t), t) = D^{(1)}(x, t) \quad (4)$$

$$\sigma(x(t), t) = \sqrt{2D^{(2)}(x, t)} \quad (5)$$

の関係が成り立つ [1]。

2.2. モデル推定の方法

先行研究 [2] では、データから計算した条件付きモーメントとモデルから計算した条件付きモーメントが一致するようにパラメータを探索している。ここでの問

題点として、最適化を行う際にモデルのパラメータを変更する度に条件付きモーメントを計算し直す必要があることがあげられる。つまり、条件付きモーメントの計算を高速化することが重要となる。

2.3. 条件付きモーメントの計算

Fokker-Planck 方程式を用いて、条件付きモーメントを求める。 x を確率変数、 x_0 をある定数としたとき、「確率過程 x について、時刻 t から時刻 $t + \tau$ で、 x_0 から x' に変動する確率」という条件付き確率密度関数を用いると、 n 次の条件付きモーメントは、

$$M_\tau^{(n)}(x_0) = \int_{-\infty}^{\infty} (x' - x_0)^n e^{\hat{L}(x')\tau} \delta(x' - x_0) dx' \quad (6)$$

で得られる。しかし、この形の初期状態は Dirac のデルタ関数となっているため、条件付きモーメントの数値計算が困難になってしまう。そこで、元の Fokker-Planck 方程式を式変形することで得られる、adjoint Fokker-Planck 方程式が導入された [2]。

$$\begin{cases} \frac{\partial \tilde{p}(x, t)}{\partial t} = \hat{L}^\dagger(x, t)\tilde{p}(x, t) \\ \hat{L}^\dagger(x, t) = D^{(1)}(x, t)\frac{\partial}{\partial x} + D^{(2)}(x, t)\frac{\partial^2}{\partial x^2} \end{cases} \quad (7)$$

この adjoint Fokker-Planck 方程式との双対性によって、式 (6) は以下の手順で変形される。

$$\begin{aligned} M_\tau^{(n)}(x_0) &= \int_{-\infty}^{\infty} (x' - x_0)^n \{e^{\hat{L}(x')\tau} \delta(x' - x_0) dx'\} \\ &= \int_{-\infty}^{\infty} \{e^{\hat{L}^\dagger(x')\tau} (x' - x_0)^n\} \delta(x' - x_0) dx' \\ &= \int_{-\infty}^{\infty} \tilde{p}(x', \tau) \delta(x' - x_0) dx' \\ &= \tilde{p}(x_0, \tau) \end{aligned} \quad (9)$$

この変形によって偏微分演算子を、右側の $\delta(x' - x_0)$ への作用の代わりに左側の $(x' - x_0)^n$ への作用とみなすことができる。式 (9) での $\tilde{p}(x', \tau)$ は、

$$\tilde{p}(x', \tau) = e^{\hat{L}^\dagger(x')\tau} (x' - x_0)^n \quad (10)$$

であり、adjoint Fokker-Planck 方程式によって初期状態を $(x' - x_0)^n$ として時間発展している形式解となっている。これにより、初期条件における Dirac のデルタ関数を避けることができ、数値計算が可能となる。なお、Fokker-Planck 方程式では、確率密度関数を時間発展していたが、adjoint Fokker-Planck 方程式での関数 $\tilde{p}(x', \tau)$ は確率密度関数とは解釈できない。

条件付きモーメントは確率密度関数の座標を引数とする関数である。そのため、 $(x' - x_0)^n$ を初期条件とし

[†]埼玉大学, Saitama University

て時間発展させる場合、 x_0 を様々な値に変えて時間発展をし直す必要がある。その際、全ての座標に関して計算をすることは困難であるため、ある間隔を置いて x_0 を変えながら時間発展を行うが、その場合でも計算コストはかかってしまう。先行研究 [2] では、数値計算方法が明示的に述べられていなかったため、パラメータである係数条件を細かく調べていくときには、莫大な計算量になってしまうことが懸念される。この問題を回避する数値計算の提案法を次の節で解説する。

3. 数値計算における提案法

3.1. 初期状態の展開による計算時間の短縮

パラメータ推定を行う際、初期状態を $(x' - x_0)^n$ として時間発展を行った場合に起きてしまう計算時間の問題を解決するために、 $(x' - x_0)^n$ を先に展開してしまうことを考える。パラメータ推定に用いる条件付きモーメント $M^{(n)}$ は、 $n = 1, n = 2$ の場合のみであり、

$$\begin{aligned}\tilde{p}^{(1)}(x', 0) &= x' - x_0 \\ \tilde{p}^{(2)}(x', 0) &= x'^2 - 2x_0x' + x_0^2\end{aligned}$$

のように展開できる。ここで、

$$\tilde{p}_n(x', \tau) = e^{\hat{L}^\dagger(x')\tau} x'^n$$

として展開した後の x'^m の次数ごとに時間発展を行った $\tilde{p}_n(x', \tau)$ を考える。これを先に解くことで、求める条件付きモーメントの式は、

$$\begin{aligned}M_\tau^{(1)}(x_0) &= \tilde{p}_1(x_0, \tau) - x_0 \\ M_\tau^{(2)}(x_0) &= \tilde{p}_2(x_0, \tau) - 2x_0\tilde{p}_1(x_0, \tau) - x_0^2\end{aligned}$$

で求めることができる。これによって、初期状態に関わる x_0 の値を変えたときに再び時間発展を行う必要がなくなり、 x_0 を引数に用いるだけで別の初期状態による条件付きモーメントの結果を得ることができる。

3.2. 冪級数展開を利用した計算時間の短縮

Adjoint Fokker-Planck 方程式で時間発展する関数 $\tilde{p}(x, t)$ を以下のように冪級数展開する。

$$\tilde{p}_n(x, t) = \sum_{m=0}^{\infty} p_n(m, t)x^m$$

$p_n(m, t)$ は、 $\tilde{p}_n(x, t)$ を x の冪乗で級数展開したときの x^m の項の係数を示し、 m は 0 以上の整数値をとる。

元の確率過程を式 (1) で $\mu(x) = -\gamma x$ 、 $\sigma(x) = D$ とした、Ornstein-Uhlenbeck 過程 (O-U 過程)

$$dx = (-\gamma x)dt + DdW(t) \quad (11)$$

であると仮定する。この時、式 (7)、式 (8) に冪級数展開式を用いると、O-U 過程における冪級数展開の係数の時間発展式が

$$\frac{\partial}{\partial t} p_n(m, t) = -\gamma m p_n(m, t) + D(m+2)(m+1)p_n(m+2, t)$$

として得られ、必要な $p_n(m, t)$ の状態数として $m = 4$ 状態までのみを用いて計算できる。

4. 数値実験

3 節の提案法の有用性を検証するために、最適化を行う際に要する計算速度に関する実験を行った。最適化の目的関数には先行研究 [2] と同様のものを使用し、(A) 初期状態を展開せずに $(x' - x_0)^n$ として時間発展を行う方法、(B) 3.1 節で提案した方法、(C) 3.2 節で提案した方法の、三通りの計算時間を比較した。推定する元データの生成には、時間間隔を $\Delta t = 10^{-3}$ として式 (11) の O-U 過程を離散化したものを用いた。この時間間隔で更新式を 1000 回計算した結果が、時刻 1.0 の間に動いた変位ということになり、時刻 1.0 ごとの結果を 10000 個サンプリングしたものを元データとした。時間発展における実験条件は、時間発展を行う時間を 1.0、その時間幅を 10^{-4} とした。初期条件に関わる x_0 は $[-2.0 : 2.0]$ の範囲で 0.5 刻みで変化させた。また、(A),(B) の時間発展させる関数の x 座標は 0.04 刻みで $[-8.00 : 8.00]$ の範囲で定義し、偏微分の実装には差分法を用いた。以上の条件のもと、二つのパラメータ γ 、 D のそれぞれの値を 0.5 から 1.5 までの間で 0.1 刻みで変化させて目的関数を計算した。

結果として、数値計算を 10 回行った時の平均時間は、(A) が 2 分 29.061 秒、(B) が 15.852 秒、(C) が 3.263 秒であった。(A) の手法では初期条件の x_0 を変更してモーメントの計算を時間発展を繰り返す必要があった。それに対して (B) の手法では、時間発展計算を行った後に初期条件である x_0 を代入する形となるため、時間発展の回数を減らすことができた。また、(A),(B) の手法では定義する x' の境界条件が自明でないため、モーメントとして扱う範囲よりも大きく定義する必要があり、解くべき連立微分方程式の数も多くなってしまった。一方、(C) の手法では状態数を 4 状態に抑え、解く連立微分方程式の数を大幅に削減することができたため、時間発展自体の計算量を削減することができた。

5. まとめと今後の課題

条件付きモーメントの式を展開してから時間発展を行うことで、計算量を抑えることができる。さらに、時間発展を行う関数を冪級数に展開することで、偏微分の数値計算を行わずにモーメントを算出することができ、O-U 過程のモデルにおける時間発展計算量を削減できることが示された。

また、冪級数展開を利用した提案手法については、現時点では O-U 過程でのみその有用性が示されただけである。今後、別の過程への適用や、スペクトル法の考え方に基づいた他の関数系での展開を検討したい。

謝辞

本研究の一部は JSPS 科研費 JP16K00323 の助成を受けたものです。

参考文献

- [1] C. Gardiner, *Stochastic Methods*, Fourth Edition, Springer (2009).
- [2] C. Honisch, R. Friedrich, Estimation of Kramers-Moyal coefficients at low sampling rates, *Physical Review E* 83, 066701 (2011).