

時系列データに対する確率微分方程式モデルの 統計的係数決定公式と軌道の予測可能性について

中野直人 (Naoto Nakano)*

Advanced Institute for Materials Research, Tohoku University

稲津將 (Masaru Inatsu)

Graduate School of Science, Hokkaido University

楠岡誠一郎 (Seiichiro Kusuoka)

Graduate School of Science, Tohoku University

齊木吉隆 (Yoshitaka Saiki)

Graduate School of Commerce and Management, Hitotsubashi University

1 序

状態が時々刻々変化していく現象に対して、現在の状態から未来の状態をどれだけ予測できるかはその現象を表すモデルと現在の状態を表すデータの正確さにかかっている。とはいえ、実際はいわゆるモデル化誤差や観測誤差が発生するため、モデルとデータを用いた解析では何らかの不確定な要素が常に入り込むものである。そのためそのようなデータ解析では、理想的な「正確さ」を追求できるとは限らないため、確率論的枠組みが用いられることが多い。たとえば経済学 [6]、気象学 [1, 5]、生物モデル [13]、材料科学 [14] などの幅広い研究分野で確率論的モデルによる研究が進められている。確率論的モデルと一口に言ってもアプローチの仕方は多岐にわたり、例えば回帰分析のような予測モデルや隠れマルコフモデルなど様々な手法が開発されているが、ここでは確率論的枠組みの中でも時間連続的に変化するデータに対して用いられる確率微分方程式モデルに焦点を当てることにする

*e-mail: n.nakano@wpi-aimr.tohoku.ac.jp

[3, 4, 7, 9, 10].

確率微分方程式によるモデリングは、時系列データが相空間内の状態の変化（軌道）として表されるものと解釈し、状態の変化に対して決定論的要因と確率論的要因が両存すると解釈される現象のデータに対して用いられる。しばしば、スケールの大きな変動（遅い変数）が決定論的寄与を、微細なゆらぎ（速い変数）が確率論的寄与を表すようなモデルが多く、Langevin 方程式がその最もシンプルでかつ重要な例である。確率微分方程式による解析というと、与えられた方程式を数学的に解析したり数値シミュレーションしたりすることでその現象の理解を深めるモデル主体のアプローチもあるが、対象の時系列データが何かの確率微分方程式に従って時間発展したものであるとみなし、その係数をデータから推定するという逆問題のアプローチも存在する。すなわち、長さ K の N 次元の時系列データ

$$\{\mathbf{X}_t = (X_t^1, X_t^2, \dots, X_t^N)^T \in \mathbb{R}^N \mid t \in \mathbb{T}\}, \quad \mathbb{T} = \{t = j\Delta t \mid j = 1, 2, \dots, K\}$$

は相空間 $\mathbb{R}^N$ 上のマルコフ型確率微分方程式

$$d\mathbf{X}_t = \mathbf{A}(\mathbf{X}_t)dt + \mathbf{S}(\mathbf{X}_t)d\mathbf{W}_t \quad (1.1)$$

の解軌道から Δt おきにサンプルされた点であると仮定し、そのデータから係数 $\mathbf{A}$ と $\mathbf{S}$ を決定して解の挙動や統計的性質を調べるというものである。ここで、 $\mathbf{x} = (x^1, x^2, \dots, x^N)^T$ は相空間の位置とし、 $\mathbf{A}(\mathbf{x}) = (A^1(\mathbf{x}), A^2(\mathbf{x}), \dots, A^N(\mathbf{x}))^T$ はドリフトベクトル、 $\mathbf{W}_t = (W_t^1, W_t^2, \dots, W_t^M)^T$ は独立な M 個の Winner 過程を成分に持つ確率過程であり、 $\mathbf{T}$ は転置を表す。 $\mathbf{S}(\mathbf{x}) = (S^{ij}(\mathbf{x}))$ ($i = 1, 2, \dots, N, j = 1, 2, \dots, M$) は乗法的ノイズの係数行列であり、系にかかるランダム性の強さが状態に依存して変化しうる場合を考慮している。ここでいう系のランダム性とは、データの持つ測定誤差や Langevin 方程式のように確率論的ノイズとして入るランダム外力だけを指すのではない。高次元力学を低次元相空間に射影したときは、縮約された高次の力学はその低次元空間での力学として表現できなくなるため、確率論的に扱われることになる。このようなシステム外の影響も含めたゆらぎを表す項として乗法的ノイズ項を採用するのである。

ここで、(1.1) から導かれる Fokker–Planck 方程式は、 $p(\mathbf{x}, t)$ を (1.1) の解の遷移確率密度函数とすると、以下のように表される。

$$\frac{\partial p}{\partial t}(\mathbf{x}, t) = - \sum_{i=1}^N \frac{\partial}{\partial x^i} (A^i(\mathbf{x})p(\mathbf{x}, t)) + \sum_{i,j=1}^N \frac{\partial^2}{\partial x^i \partial x^j} (B^{ij}(\mathbf{x})p(\mathbf{x}, t)). \quad (1.2)$$

ここで $\mathbb{B}(\mathbf{x}) = (B^{ij}(\mathbf{x}))$ は拡散行列とよばれ、 $\mathbb{B} = \frac{1}{2}\mathbb{S}\mathbb{S}^T$ で表される非負対称行列である。確率論的には $\mathbb{S}(\mathbf{x})$ ではなく $\mathbb{B}(\mathbf{x})$ が (1.1) の解の拡散をつかさどる行列であることが知られている。この Fokker-Planck 方程式 (1.2) の係数は、確率微分方程式 (1.1) の解と以下のような関係があることが知られている：

$$A^i(\mathbf{x}) = \lim_{t \rightarrow 0} \frac{E_{\mathbf{x}}[X_t^i - x^i]}{t}, \quad B^{ij}(\mathbf{x}) = \lim_{t \rightarrow 0} \frac{E_{\mathbf{x}}[(X_t^i - x^i)(X_t^j - x^j)]}{2t}. \quad (1.3)$$

ただし、 $E_{\mathbf{x}}$ は、 $\mathbf{x}$ を初期値とする (1.1) の解が導く確率測度に対する期待値である。(1.3) は、相空間の位置 $\mathbf{x}$ を出発する (1.1) の解からその位置における (1.2) の係数の値を決定することができることを示している。

データ解析の観点では、この解と係数の関係式を利用して時系列データから統計的に $\mathbf{A}(\mathbf{x})$ と $\mathbb{B}(\mathbf{x})$ の近似値を計算しようという研究手法が提案されてきた [1, 11]. すなわち、相空間をセルに分割し、各セルに滞在する時系列データを用いて

$$\begin{cases} \bar{A}^i(\bar{\mathbf{x}}_m) = \frac{\langle X_{t+\Delta t}^i - X_t^i \rangle}{\Delta t}, \\ \bar{B}^{ij}(\bar{\mathbf{x}}_m) = \frac{\langle (X_{t+\Delta t}^i - X_t^i)(X_{t+\Delta t}^j - X_t^j) \rangle}{2\Delta t} \end{cases} \quad (1.4)$$

と計算するのである (Figure 1). ただし、 $\bar{\mathbf{x}}_m$ は相空間を分割したときの m 番目のセル、 $\bar{\mathbf{A}}(\bar{\mathbf{x}}_m) = (\bar{A}^1(\bar{\mathbf{x}}_m), \bar{A}^2(\bar{\mathbf{x}}_m), \dots, \bar{A}^N(\bar{\mathbf{x}}_m))^T$ と $\bar{\mathbb{B}}(\bar{\mathbf{x}}_m) = (\bar{B}^{ij}(\bar{\mathbf{x}}_m))$ はそれぞれセル $\bar{\mathbf{x}}_m$ におけるドリフトベクトルと拡散行列の計算値である。変数の上に付くバー (̄) は離散化されたセルやその上で定義される変数に対して用いることにする。また、 $\langle \cdot \rangle$ は各セルに属する時系列データに対するアンサンブル平均を意味し、例えばドリフトベクトルの計算では、セル $\bar{\mathbf{x}}_m$ の滞在時刻を $\mathbb{T}_m = \{t \in \mathbb{T} \mid \mathbf{X}_t \in \bar{\mathbf{x}}_m\}$ とおき、 $\{t \in \mathbb{T}_m \mid t < K\Delta t\}$ が空でない限り

$$\bar{A}^i(\bar{\mathbf{x}}_m) = \frac{\langle X_{t+\Delta t}^i - X_t^i \rangle}{\Delta t} = \frac{1}{\#\mathbb{T}_m - 1} \sum_{t \in \mathbb{T}_m, t < K\Delta t} \frac{X_{t+\Delta t}^i - X_t^i}{\Delta t}$$

とおこなう。ここでは時系列データにエルゴード性を仮定し、データ長が十分あり且つその軌道がセル $\bar{\mathbf{x}}_m$ を十分多く通過するのであれば、そのセルに含まれるデータに対するアンサンブルを計算することで、(1.3) での期待値の計算の代わりにすることができるとしている。この手法を用いれば時系列データだけから確率微分方程式モデルを構成することができるわけである。

一方で、(1.4) は (1.3) から自然に導かれる時空離散化を用いており、統計的係数決定公式として理に適うように思えるのだが、実はドリフトベクトルのノルムが

大きい領域で拡散行列の計算誤差が大きくなる傾向となることが今回わかった。特に、拡散行列のノルムは系の本質的なノイズの強さを表す量であるが、(1.4)から計算される量ではノルムが必ず過大評価となってしまう。これにより、データから評価した相空間の各位置におけるノイズの強さが実際の相空間における時系列の軌道の振る舞いとは直感的に異なってしまう例があり (§3 参照)、実用上好ましくない。また、ドリフトベクトルの計算時にもセルのサイズとデータサンプル幅に制限があることもわかっており (§2.1 参照)、時系列データから確率微分方程式の係数を算出するときには幾つか払うべき注意点があることがわかる。本研究では、それらの注意点を数学的に示し、確率微分方程式モデリングにおける新しい係数決定公式を提示することが目的である。

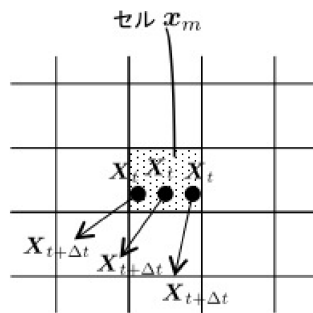


Figure 1: 相空間分割セルの概念図

本稿の構成は以下の通りである。まず §2.1 ではドリフトベクトルの計算手法に対する注意点を例を挙げながら説明し、その問題点を除去した公式を紹介する (2.1)。§2.2 では拡散行列に対する新しい係数決定公式 (2.2) を示し、その有効性を簡単な例を通して議論する。最後に §3 において Lorenz system と Rössler system の数値解の平面射影データに対して、旧公式と新公式の結果を比較することで新公式の有効性を数値的に検証する。

2 統計的係数決定公式

2.1 ドリフトベクトルの計算式に関する注意点

ドリフトベクトルの計算においては、時間スケールすなわちデータのサンプル間隔 Δt と空間スケールすなわちセルのサイズ Δx のバランスが大切になる。すなわち、

平均移動距離より比較的大きいビンサイズを取るときに (1.1) は直感と異なる結果を与えることがある。また、ドリフトの計算式から対象の時系列にふさわしい Δt と Δx の取り方がわかる。これを順を追って説明する。

まず用語を定義する。時系列データがあるセルに入ってから最初に出るまでをそのセルにおける“一連のイベント”とよぶことにする。セル $\bar{x}_m$ に X_t で入って $X_{t+k\Delta t}$ で出た場合、すなわち $X_{t-\Delta t} \notin \bar{x}_m$ かつ $X_{t+j\Delta t} \in \bar{x}_m$ ($j = 0, 1, \dots, k-1$) かつ $X_{t+k\Delta t} \notin \bar{x}_m$ のとき、その一連のイベントは長さが k であるという。

Example 2.1

二次元相空間 ($N = 2$) 上のセル $\bar{x}_m = [a, a + \Delta x) \times [b, b + \Delta x)$ の近傍を

$$\begin{aligned} X_{t_1+j\Delta t} &= \left(a + \frac{j\Delta x}{4}, b + \frac{j\Delta x}{8} \right) \quad (j = 0, 1, \dots, 4), \\ X_{t_2+j\Delta t} &= \left(a + \frac{j\Delta x}{4}, b + \frac{(j+4)\Delta x}{8} \right) \quad (j = 0, 1, \dots, 4) \end{aligned}$$

というように長さ 4 の 2 つの一連のイベントが $\bar{x}_m$ を通過するとする。ここでセル $\bar{x}_m$ に滞在したデータ点は合計 8 点で $\mathbb{T}_m = \{t_1 + j\Delta t, t_2 + j\Delta t \mid j = 0, 1, 2, 3\}$ であることに注意する。このときドリフトベクトルは

$$\left\{ \begin{aligned} A^1(\bar{x}) &= \frac{1}{8} \sum_{t \in \mathbb{T}_m} \frac{X_{t+\Delta t}^1 - X_t^1}{\Delta t} \\ &= \frac{1}{8} \left(\sum_{j=0}^3 \frac{X_{t_1+(j+1)\Delta t}^1 - X_{t_1+j\Delta t}^1}{\Delta t} + \sum_{j=0}^3 \frac{X_{t_2+(j+1)\Delta t}^1 - X_{t_2+j\Delta t}^1}{\Delta t} \right) \\ &= \frac{1}{8} \left(\frac{X_{t_1+4\Delta t}^1 - X_{t_1}^1}{\Delta t} + \frac{X_{t_2+4\Delta t}^1 - X_{t_2}^1}{\Delta t} \right) = \frac{\Delta x}{4\Delta t}, \\ A^2(\bar{x}) &= \frac{1}{8} \sum_{t \in \mathbb{T}_m} \frac{X_{t+\Delta t}^2 - X_t^2}{\Delta t} \\ &= \frac{1}{8} \left(\sum_{j=0}^3 \frac{X_{t_1+(j+1)\Delta t}^2 - X_{t_1+j\Delta t}^2}{\Delta t} + \sum_{j=0}^3 \frac{X_{t_2+(j+1)\Delta t}^2 - X_{t_2+j\Delta t}^2}{\Delta t} \right) \\ &= \frac{1}{8} \left(\frac{X_{t_1+4\Delta t}^2 - X_{t_1}^2}{\Delta t} + \frac{X_{t_2+4\Delta t}^2 - X_{t_2}^2}{\Delta t} \right) = \frac{\Delta x}{8\Delta t} \end{aligned} \right.$$

と求められる。ドリフトの計算式では和の線形性のために、単位時間 (Δt) あたりに進む速度 (すなわち一連のイベントの長さの逆数) は結果に関係せず、一連のイベントにおける最初の点から最後の点までの各成分の (符合込みの) 直線距離だけが計算に利いていることがわかる。この例のように一つのセル内をほぼ平行にほ

ば同速度で軌道が通過するような決定論的な時系列データの場合では特に問題は起きない。

Example 2.2

次に、同じセル内で軌道が交差するとき、例えば

$$\begin{aligned} \mathbf{X}_{t_1+j\Delta t} &= \left(a + \frac{j\Delta x}{4}, b + \frac{\Delta x}{2} \right) \quad (j = 0, 1, \dots, 4), \\ \mathbf{X}_{t_2+j\Delta t} &= \left(a + \frac{\Delta x}{2}, b + \frac{j\Delta x}{8} \right) \quad (j = 0, 1, \dots, 8) \end{aligned}$$

というような2つの一連のイベントが $\bar{\mathbf{x}}_m$ を通過するときを考える ($\mathbb{T}_m = \{t_1 + j\Delta t \ (j = 0, 1, 2, 3), t_2 + k\Delta t \ (k = 0, 1, \dots, 7)\}$). この場合では、時刻 t_1 から x 軸に平行に通過するときの一連のイベントの方が t_2 から y 軸に平行に通過するときより2倍速い。一方で、ドリフトの計算結果は

$$\left\{ \begin{aligned} A^1(\bar{\mathbf{x}}) &= \frac{1}{12} \sum_{t \in \mathbb{T}_m} \frac{X_{t+\Delta t}^1 - X_t^1}{\Delta t} \\ &= \frac{1}{12} \left(\sum_{j=0}^3 \frac{X_{t_1+(j+1)\Delta t}^1 - X_{t_1+j\Delta t}^1}{\Delta t} + \sum_{j=0}^7 \frac{X_{t_2+(j+1)\Delta t}^1 - X_{t_2+j\Delta t}^1}{\Delta t} \right) \\ &= \frac{1}{12} \frac{X_{t_1+4\Delta t}^1 - X_{t_1}^1}{\Delta t} = \frac{\Delta x}{12\Delta t}, \\ A^2(\bar{\mathbf{x}}) &= \frac{1}{12} \sum_{t \in \mathbb{T}_m} \frac{X_{t+\Delta t}^2 - X_t^2}{\Delta t} \\ &= \frac{1}{12} \left(\sum_{j=0}^3 \frac{X_{t_1+(j+1)\Delta t}^2 - X_{t_1+j\Delta t}^2}{\Delta t} + \sum_{j=0}^7 \frac{X_{t_2+(j+1)\Delta t}^2 - X_{t_2+j\Delta t}^2}{\Delta t} \right) \\ &= \frac{1}{12} \frac{X_{t_2+8\Delta t}^1 - X_{t_2}^1}{\Delta t} = \frac{\Delta x}{12\Delta t} \end{aligned} \right.$$

となり、ドリフトの x 成分も y 成分も同じ値という実際の時系列データから受ける直感とは異なる結果が出される。やはりセル内で進んだ距離のみが計算式に寄与するため、元の時系列データの速さは計算には加味されないことがわかる。このような同一のセル内での軌道交差は、本研究で対象とするような確率論的ノイズによって摂動を受ける軌道や高次元力学の軌道を低次元相空間に射影した場合で起こり得る。したがって、時系列データの平均移動距離より大きいセルサイズを用いるときは適切にドリフトが計算されているか注意が必要となる。

Example 2.3

また、このドリフトの計算ではセルサイズ Δx より小さい変動は算定に加味していない。例えば、

$$\mathbf{X}_{t+j\Delta t} = \left(a + \frac{j\Delta x}{4}, b + \frac{(2-|j-2|)\Delta x}{4} \right) \quad (j = 0, 1, \dots, 4)$$

という同じセル内で軌道が大きく方向を変えるような一連のイベントを考える。このときは、

$$\begin{cases} A^1(\bar{\mathbf{x}}) = \frac{\Delta x}{4\Delta t} \\ A^2(\bar{\mathbf{x}}) = 0 \end{cases}$$

となり、ドリフトの計算値はセルのサイズが最小の解像度となっていることがわかる。すなわち得られる確率微分方程式モデルはそのスケールでの粗視化がされているものであり、したがって Δx としては着目する現象の最小スケールを取るべきである。

とはいえ、 Δt 当たりの平均移動距離は、実験や測定の都合上（もしくは別の目的があつて）着目する現象の最小スケールであるとは限らず、比較的小さい場合がある。この場合は前述の Example 2.2 で挙げた問題点が出てくるため、それに合わせて計算手法を工夫する必要がある。ここでは3つの手法を挙げることにする。

1. 時系列データの平均移動距離がセルサイズと同等になるくらいにデータを等間隔に間引く。すなわち、時系列データを $\mathbb{T} = \{t = \tau j \Delta t \mid j = 1, 2, \dots, [K/\tau]\}$ と τ 毎にサブサンプルするのである。同一セルに2点連続して滞在しないように出来れば Example 2.2 のような問題は起こらない。また、Example 2.3 の空間解像度の限界からデータの間引きによる時間解像度の劣化は評価には大きく影響しない。ただし、各セルにおけるアンサンブル点はほぼ $1/\tau$ となるので、この手法はデータ数が十分ある場合に有効である。
2. ドリフトベクトルの計算式において、次のステップのデータとの差を計算するのではなく、平均移動距離がやはり Δx と同等になるくらいの時間サンプル幅を用いて計算する。すなわち、 $\langle X_{t+\tau\Delta t}^i - X_t^i \rangle \approx \Delta x$ となるような $\tau \in \mathbb{N}$ を取り、 $\{t \in \mathbb{T}_m \mid t + \tau\Delta t \leq K\Delta t\}$ が空でない限り

$$\bar{A}^i(\bar{\mathbf{x}}_m) = \frac{\langle X_{t+\tau\Delta t}^i - X_t^i \rangle}{\tau\Delta t} = \frac{1}{\#\mathbb{T}_m - \tau} \sum_{\substack{t \in \mathbb{T}_m \\ t + \tau\Delta t \leq K\Delta t}} \frac{X_{t+\tau\Delta t}^i - X_t^i}{\tau\Delta t} \quad (2.1)$$

とおこなう。データは間引かずに差の計算だけ τ ステップ先のデータとおこなうため、各セルにおけるアンサンブル数は時間サンプル幅の変更前とほぼ変わらない点が上記1とは異なる。この手法でも Example 2.2 のような影響を減らすことができる。

3. また、各一連のイベントの平均に対するアンサンブル平均を用いて計算することもできる。すなわち $\bar{x}_m$ に滞在するデータ点が

$$\mathbf{X}_{t_j+k\Delta t} \quad (j = 1, 2, \dots, J, \quad k = 0, 1, \dots, k_j - 1)$$

という長さがそれぞれ k_j ($j = 1, 2, \dots, J$) である J 本の一連のイベントで成り立っているとき、

$$\bar{A}^i(\bar{x}_m) = \frac{1}{J} \sum_{j=1}^J \frac{1}{k_j} \sum_{k=0}^{k_j-1} \frac{X_{t_j+(k+1)\Delta t}^i - X_{t_j+k\Delta t}^i}{\Delta t}$$

によってドリフトを求めるのである。これも Example 2.2 のような問題点は解消されている。上記2つの手法のような時系列の時間間隔を一定に保つ計算法とは異なって各イベントに着目するため、相空間の場所によって平均移動距離が大きく変化する場合にも使える。

いずれにせよ、平均移動速度 $\approx \Delta x / \Delta t$ となるように時系列のサンプル間隔とセルサイズを取るのが理想である。

しばしば時系列データの時間解像度は対象の実験・現象や測定手法に依存するものであり、こちらが自由に設定できるとは限らない。仮に時間解像度に合わせて空間解像度を取るとすると、着目したい空間スケールより小さいセルサイズになる場合があり、そうすると解析の目的とは合致しなくなる。また、気象データなど繰り返し観測を行うことができない場合では計算に必要なだけのデータ長を必ずしも取れないため、各セルにおけるアンサンブルメンバーを多く確保するために粗いセルサイズで相空間を分割せねばならないこともある。このようにデータセットの性質に由来するセルサイズの問題は様々であるが、それらを打破してより適切なドリフトベクトルを計算するには上記の2の手法が良いと思う。セルサイズは着目する空間解像度を採用し、平均移動速度が $\Delta x / \Delta t$ と釣り合うように τ を採用すれば、Example 2.2 などの影響は小さくなり、かつアンサンブルメンバーの数も減らさずに済む。

2.2 拡散行列の新計算公式

(1.4) による係数の計算公式ではドリフトの強い領域で誤差が大きくなることがわかっていて、特に誤差が顕著に現れる拡散行列に対してその評価式からドリフトの影響を除去するには、

$$\bar{B}_{\text{new}}^{ij}(\bar{x}_m) = \frac{\langle (X_{t+\tau\Delta t}^i - X_t^i)(X_{t+\tau\Delta t}^j - X_t^j) \rangle - \langle X_{t+\tau\Delta t}^i - X_t^i \rangle \langle X_{t+\tau\Delta t}^j - X_t^j \rangle}{2\tau\Delta t} \quad (2.2)$$

と、二次変動から一次変動に関する二次式を除去する計算式を用いるのが良い。ただし、 τ は、§2.1 で注意した、 $|X_{t+\tau\Delta t} - X_t| \approx \Delta x / \Delta t$ となるような τ である。これの有効性を具体的な確率微分方程式の解を例にとって説明する。

Example 2.4

幾何ブラウン運動

$$\begin{cases} dX_t = aX_t dt + bX_t dW_t, \\ X_0 = x \end{cases} \quad (2.3)$$

の場合に、ドリフトと拡散係数の近似式として、それぞれ

$$\begin{cases} \frac{E_x[X_t - x]}{t}, \\ \frac{E_x[(X_t - x)^2] - E_x[X_t - x]^2}{2t} \end{cases} \quad (2.4)$$

を用いることが有効であることを示そう。ここで $a, b \in \mathbb{R}$ を非ランダムな定数とし、 W_t を Wiener 過程とする。(2.3) のドリフト係数は $A(x) = ax$ で拡散係数は $B(x) = (bx)^2/2$ である。この解は変数分離や伊藤の公式を用いることで計算することができ、

$$X_t = xe^{(a - \frac{b^2}{2})t + bW_t}$$

で与えられる。

ここで、 $E_x[\exp(bW_t)] = \exp(b^2t/2)$ より、解の一次変動は

$$\frac{E_x[X_t - x]}{t} = \frac{x(e^{at} - 1)}{t} = ax + \mathcal{O}(t) \quad (t \rightarrow 0) \quad (2.5)$$

となる。誤差項 $\mathcal{O}(t)$ には x も含まれているため、 $E_x[X_t - x]/t$ はドリフト ax との誤差が原点から遠いほど大きくなる。しかし実際は

$$\mathcal{O}(t) = ax \sum_{n=1}^{\infty} \frac{(at)^n}{(n+1)!}$$

であるため、ドリフト ax との相対誤差は相空間の位置によらず一様なので問題ない。

一方、二次変動を計算すると

$$\begin{aligned}
 \frac{E_x [(X_t - x)^2]}{2t} &= \frac{x^2}{2t} E_x \left[e^{(2a-b^2)t+2bW_t} - 2e^{(a-\frac{b^2}{2})t+bW_t} + 1 \right] \\
 &= \frac{x^2 \left\{ e^{(2a+b^2)t} - 2e^{at} + 1 \right\}}{2t} = \frac{x^2 e^{2at} (e^{b^2t} - 1) + E_x [X_t - x]^2}{2t} \\
 &= \frac{(bx)^2}{2} + \frac{t}{2} \left(\frac{E_x [X_t - x]}{t} \right)^2 + \mathcal{O}(t) \quad (t \rightarrow 0) \\
 &= \frac{(bx)^2}{2} + \frac{A(x)^2}{2} t + \mathcal{O}(t) \quad (t \rightarrow 0)
 \end{aligned} \tag{2.6}$$

を得る。(2.6) の第2項は $\mathcal{O}(t)$ に含まれる項ではあるが、 $t > 0$ である限りは拡散係数の計算式に一次変動の二乗が陽に現れた形となり、拡散係数の値 $(bx)^2/2$ より概ね過剰評価となることがわかる。また、その第2項はドリフトの大きい領域で誤差への寄与が強くなるため、誤差が t の大きさだけではコントロールされずに相空間の場所によって変わるというのは好ましくない。したがって、二次変動から一次変動の二乗を引いておくと、

$$\frac{E_x [(X_t - x)^2] - E_x [X_t - x]^2}{2t} = \frac{x^2 e^{2at} (e^{b^2t} - 1)}{2t} = \frac{(bx)^2}{2} + \mathcal{O}(t) \quad (t \rightarrow 0) \tag{2.7}$$

となり、その誤差項はドリフトの大きさの多寡によらず、空間一様に $t > 0$ の大きさに支配されるものになっている。

以上により、(2.4) がドリフトと拡散係数の近似として有効であることがわかった。

注意 この例からわかるように、(2.5) や (2.7) という $t \rightarrow 0$ の極限を取らない量は、相対誤差は空間一様に取れるものの、概ねドリフトの強さのパラメータ a によって絶対誤差・相対誤差の多寡が左右される傾向があることがわかる。

Example 2.5

N 次元の非斉次線形確率微分方程式

$$\begin{cases} d\mathbf{X}_t = (\mathbf{A}_0 + \mathbf{L}\mathbf{X}_t)dt + (\mathbf{S}_0 + \boldsymbol{\sigma}\mathbf{X}_t)d\mathbf{W}_t, \\ \mathbf{X}_0 = \mathbf{x} \end{cases} \tag{2.8}$$

について、ドリフトベクトルと拡散行列の各成分の近似式として、それぞれ

$$\begin{cases} \frac{E_x [X_t^i - x^i]}{t}, \\ \frac{E_x [(X_t^i - x^i)(X_t^j - x^j)] - E_x [X_t^i - x^i] E_x [X_t^j - x^j]}{2t} \end{cases} \tag{2.9}$$

を用いることが有効であることを示す。ここで、 $\mathbf{A}_0 = (A_0^1, A_0^2, \dots, A_0^N)^T$ は N 次元の定ベクトル、 $\mathbb{L} = (L^{ij})$ 、 $\mathbb{S}_0 = (S_0^{ij})$ はそれぞれサイズ $N \times N$ 、 $N \times M$ の定行列、 $\sigma = (\sigma^{ijk})$ は $N \times M \times N$ の三階の定テンソルであり、 $\mathbf{A}_0, \mathbb{L}, \mathbb{S}_0, \sigma$ は全て非ランダムである。また、テンソルの縮約は $[\sigma \mathbf{X}_t]^{ij} = \sum_{k=1}^N \sigma^{ijk} X_t^k$ ($i = 1, 2, \dots, N$, $j = 1, 2, \dots, M$) とする。(2.8) の解の陰的公式は 1 階の非斉次常微分方程式系と同様に計算することで

$$\mathbf{X}_t = \mathbf{D}(t) + \mathbf{Z}_t,$$

$$\mathbf{D}(t) = e^{\mathbb{L}t} \mathbf{x} + \int_0^t e^{\mathbb{L}(t-s)} \mathbf{A}_0 ds, \quad \mathbf{Z}_t = \int_0^t e^{\mathbb{L}(t-s)} \mathbb{S}(\mathbf{X}_s) d\mathbf{W}_s,$$

$$\mathbb{S}(\mathbf{X}_t) = \mathbb{S}_0 + \sigma \mathbf{X}_t$$

を得る。この解は決定論的な項 $\mathbf{D}(t)$ と確率論的な項 $\mathbf{Z}_t$ の和として与えられる。ここで

$$\int_0^t e^{-\mathbb{L}s} \mathbb{S}(\mathbf{X}_s) d\mathbf{W}_s$$

はマルチンゲールなので、

$$\mathbb{E}_{\mathbf{x}}[\mathbf{Z}_t] = e^{\mathbb{L}t} \mathbb{E}_{\mathbf{x}} \left[\int_0^t e^{-\mathbb{L}s} \mathbb{S}(\mathbf{X}_s) d\mathbf{W}_s \right] = \mathbf{0}$$

となり、

$$\begin{aligned} \frac{\mathbb{E}_{\mathbf{x}}[\mathbf{X}_t - \mathbf{x}]}{t} &= \frac{\mathbf{D}(t) - \mathbf{x}}{t} = \frac{(e^{\mathbb{L}t} - \mathbb{I})\mathbf{x}}{t} + \frac{1}{t} \int_0^t e^{\mathbb{L}(t-s)} \mathbf{A}_0 ds \\ &= \mathbf{A}_0 + \mathbb{L}\mathbf{x} + \mathcal{O}(t) \quad (t \rightarrow 0) \end{aligned} \quad (2.10)$$

が従う。この場合でも絶対誤差は $\mathbf{x}$ に依存するが、

$$\frac{|\mathbb{E}_{\mathbf{x}}[\mathbf{X}_t - \mathbf{x}] / t - (\mathbf{A}_0 + \mathbb{L}\mathbf{x})|}{|\mathbf{A}_0 + \mathbb{L}\mathbf{x}|} \leq c(\mathbf{A}_0, \mathbb{L}, t)$$

となるため、相対誤差は Example 2.4 と同様に空間一様に抑えられる。

また、

$$\begin{aligned} (X_t^i - x^i)(X_t^j - x^j) &= (D^i(t) - x^i)(D^j(t) - x^j) + (D^i(t) - x^i)Z_t^j \\ &\quad + Z_t^i(D^j(t) - x^j) + Z_t^i Z_t^j \end{aligned}$$

であるため、

$$\frac{\mathbb{E}_{\mathbf{x}}[(X_t^i - x^i)(X_t^j - x^j)]}{2t} = \frac{(D^i(t) - x^i)(D^j(t) - x^j) + \mathbb{E}_{\mathbf{x}}[Z_t^i Z_t^j]}{2t}$$

$$\begin{aligned}
&= \frac{\mathbb{E}_{\mathbf{x}} [X_t^i - x^i] \mathbb{E}_{\mathbf{x}} [X_t^j - x^j] + \mathbb{E}_{\mathbf{x}} [Z_t^i Z_t^j]}{2t} \\
&= \frac{t}{2} \cdot \frac{\mathbb{E}_{\mathbf{x}} [X_t^i - x^i]}{t} \cdot \frac{\mathbb{E}_{\mathbf{x}} [X_t^j - x^j]}{t} + \frac{\mathbb{E}_{\mathbf{x}} [Z_t^i Z_t^j]}{2t} \quad (2.11)
\end{aligned}$$

が成り立つ。ここで伊藤の公式を用いると、右辺第2項は

$$\begin{aligned}
\mathbb{E}_{\mathbf{x}} [Z_t^i Z_t^j] &= \int_0^t \mathbb{E}_{\mathbf{x}} [2B^{ij}(\mathbf{X}_s)] ds \\
&\quad + \int_0^t ds \int_0^s \left[2\mathbb{L}e^{\mathbb{L}(s-\tau)} \mathbb{E}_{\mathbf{x}} [\mathbb{B}(\mathbf{X}_\tau)] e^{\mathbb{L}^T(s-\tau)} \right]^{ij} d\tau \\
&\quad + \int_0^t ds \int_0^s \left[2\mathbb{L}e^{\mathbb{L}(s-\tau)} \mathbb{E}_{\mathbf{x}} [\mathbb{B}(\mathbf{X}_\tau)] e^{\mathbb{L}^T(s-\tau)} \right]^{ji} d\tau \\
&= 2tB^{ij}(\mathbf{x}) + \int_0^t 2\mathbb{E}_{\mathbf{x}} [B^{ij}(\mathbf{X}_s) - B^{ij}(\mathbf{x})] ds \\
&\quad + \int_0^t ds \int_0^s \left[2\mathbb{L}e^{\mathbb{L}(s-\tau)} \mathbb{E}_{\mathbf{x}} [\mathbb{B}(\mathbf{X}_\tau)] e^{\mathbb{L}^T(s-\tau)} \right]^{ij} d\tau \\
&\quad + \int_0^t ds \int_0^s \left[2\mathbb{L}e^{\mathbb{L}(s-\tau)} \mathbb{E}_{\mathbf{x}} [\mathbb{B}(\mathbf{X}_\tau)] e^{\mathbb{L}^T(s-\tau)} \right]^{ji} d\tau \\
&= 2tB^{ij}(\mathbf{x}) + \mathcal{O}(t^2) \quad (t \rightarrow 0) \quad (2.12)
\end{aligned}$$

となる。ただし、

$$B^{ij}(\mathbf{x}) = \frac{1}{2} \sum_{l=1}^M S^{il}(\mathbf{x}) S^{jl}(\mathbf{x}) = \frac{1}{2} \sum_{l=1}^M \left(\sum_{m=1}^N S_0^{il} + \sigma^{ilm} x^m \right) \left(\sum_{n=1}^N S_0^{jl} + \sigma^{jln} x^n \right)$$

である。したがって解の二次変動量 (2.11) は、決定論的成分 (右辺第1項) と確率論的成分 (右辺第2項) の和で表されることがわかる。だからこの場合も、二次変動量から一次変動量の二次式を引いておくと良い：

$$\frac{\mathbb{E}_{\mathbf{x}} [(X_t^i - x^i)(X_t^j - x^j)] - \mathbb{E}_{\mathbf{x}} [X_t^i - x^i] \mathbb{E}_{\mathbf{x}} [X_t^j - x^j]}{2t} = B^{ij}(\mathbf{x}) + \mathcal{O}(t). \quad (2.13)$$

さらに $\mathbb{E}_{\mathbf{x}} [X_t^i] = D^i(t)$ と $\mathbb{E}_{\mathbf{x}} [X_t^i X_t^j] = D^i(t) D^j(t) + \mathbb{E}_{\mathbf{x}} [Z_t^i Z_t^j]$ より

$$\begin{aligned}
\mathbb{E}_{\mathbf{x}} [B^{ij}(\mathbf{X}_t)] &= \sum_{l=1}^M \sum_{m,n=1}^N \sigma^{ilm} \sigma^{jln} \mathbb{E}_{\mathbf{x}} [Z_t^m Z_t^n] + F^{ij}(t), \\
F^{ij}(t) &= \sum_{l=1}^M \sum_{m,n=1}^N \sigma^{ilm} \sigma^{jln} D^m(t) D^n(t) + \sum_{l=1}^M \sum_{m=1}^N \left(\sigma^{ilm} S_0^{jl} + \sigma^{jlm} S_0^{il} \right) D^m(t) + B_0^{ij}
\end{aligned}$$

が従う。これと (2.12) より、 $\mathbb{E}_{\mathbf{x}} [Z_t^i Z_t^j]$ は

$$\mathbb{E}_{\mathbf{x}} [Z_t^i Z_t^j] = \int_0^t \sum_{l=1}^M \sum_{m,n=1}^N 2\sigma^{ilm} \sigma^{jln} \mathbb{E}_{\mathbf{x}} [Z_s^m Z_s^n] ds$$

$$\begin{aligned}
& + \int_0^t ds \int_0^s \sum_{l=1}^M \sum_{m,n,p,q=1}^N 2 [\mathbb{L}e^{\mathbb{L}(s-\tau)}]^{ip} \sigma^{plm} \sigma^{qln} E_{\mathbf{x}} [Z_{\tau}^m Z_{\tau}^n] \left[e^{\mathbb{L}^T(s-\tau)} \right]^{qj} d\tau \\
& + \int_0^t ds \int_0^s \sum_{l=1}^M \sum_{m,n,p,q=1}^N 2 [\mathbb{L}e^{\mathbb{L}(s-\tau)}]^{jp} \sigma^{plm} \sigma^{qln} E_{\mathbf{x}} [Z_{\tau}^m Z_{\tau}^n] \left[e^{\mathbb{L}^T(s-\tau)} \right]^{qi} d\tau \\
& + G^{ij}(t), \\
G^{ij}(t) &= \int_0^t F^{ij}(s) ds + \int_0^t ds \int_0^s \left[2\mathbb{L}e^{\mathbb{L}(s-\tau)} \mathbb{F}(\tau) e^{\mathbb{L}^T(s-\tau)} \right]^{ij} d\tau \\
& + \int_0^t ds \int_0^s \left[2\mathbb{L}e^{\mathbb{L}(s-\tau)} \mathbb{F}(\tau) e^{\mathbb{L}^T(s-\tau)} \right]^{ji} d\tau
\end{aligned}$$

という二階の非斉次線形積分方程式を満たすことがわかり、解の評価から

$$\max_{i,j=1,2,\dots,N} \left| \frac{E_{\mathbf{x}} [Z_t^i Z_t^j] / 2t - B^{ij}(\mathbf{x})}{B^{ij}(\mathbf{x})} \right| \leq c(\mathbf{A}_0, \mathbb{L}, \mathbb{S}_0, \boldsymbol{\sigma}, t)$$

が従う。すなわち、 $E_{\mathbf{x}} [Z_t^i Z_t^j] / 2t$ の $B^{ij}(\mathbf{x})$ に対する相対誤差は相空間において一様に抑えられる。

以上により、(2.9) がドリフトベクトルと拡散行列の近似式として有効であることがわかった。

Example 2.6

最後に、一般のマルコフ型の確率微分方程式

$$\begin{cases} d\mathbf{X}_t = \mathbf{A}(\mathbf{X}_t)dt + \mathbb{S}(\mathbf{X}_t)d\mathbf{W}_t, \\ \mathbf{X}_0 = \mathbf{x} \end{cases} \quad (2.14)$$

の場合を考えよう。このときも (2.9) がドリフトベクトルと拡散行列の近似式として有効であることを示す。

解と係数の関係 (1.3) は一般の確率微分方程式に対して成り立ち、その等式は $t \rightarrow 0$ への極限を取ることで得られる。したがって、係数の近似計算では解は初期時刻付近だけを考えれば良く、 $t \ll 1$ のときに限定してよい。 $\mathbf{A}(\mathbf{x})$ と $\mathbb{S}(\mathbf{x})$ に C^1 級程度の十分な滑らかさがあれば、 t を十分小さく取ることによって $|\mathbf{X}_t - \mathbf{x}| \ll 1$ も同様に成立するようにできる。このとき、

$$\begin{cases} A^i(\mathbf{X}_t) \approx A^i(\mathbf{x}) + \sum_{k=1}^N \frac{\partial A^i}{\partial x^k}(\mathbf{x})(X_t^k - x^k), \\ S^{ij}(\mathbf{X}_t) \approx S^{ij}(\mathbf{x}) + \sum_{k=1}^N \frac{\partial S^{ij}}{\partial x^k}(\mathbf{x})(X_t^k - x^k) \end{cases}$$

と係数を一次近似すると, (2.14) は $\hat{Y}_t = X_t - x$ に対する以下の方程式に線形化される.

$$\begin{cases} d\hat{Y}_t^i = \left(A^i(x) + \sum_{k=1}^N \frac{\partial A^i}{\partial x^k}(x) \hat{Y}_t^k \right) dt + \sum_{j=1}^M \left(S^{ij}(x) + \sum_{k=1}^N \frac{\partial S^{ij}}{\partial x^k}(x) \hat{Y}_t^k \right) dW_t^j, \\ \hat{Y}_0^i = 0 \quad (i = 1, 2, \dots, N) \end{cases} \quad (2.15)$$

これは $\hat{Y}_t$ に関する非斉次線形確率微分方程式であるため, 係数の近似計算は $A_0 = A(x)$, $L = \nabla A(x)$, $S_0 = S(x)$, $\sigma = \nabla S(x)$ として Example 2.5 に帰着される.

X_t と $\hat{Y}_t$ をそれぞれ (2.14) と (2.15) の解とすると, (2.10) から

$$\frac{E_x[X_t^i - x^i]}{t} \approx \frac{\hat{E}_0[\hat{Y}_t^i]}{t} = A^i(x) + \sum_{k=1}^N \frac{\partial A^i}{\partial x^k}(x) \cdot 0 + \mathcal{O}(t) = A^i(x) + \mathcal{O}(t) \quad (t \rightarrow 0)$$

が近似的に成り立つ. ただし, $\hat{E}_0[\cdot]$ は, 初期値 0 の (2.15) の解が導く確率測度についての期待値である. (2.15) の初期条件により, 上式における $\hat{Y}_t^k$ の一次の項は消えることに注意する. 同様に (2.13) から

$$\begin{aligned} & \frac{E_x[(X_t^i - x^i)(X_t^j - x^j)] - E_x[X_t^i - x^i] E_x[X_t^j - x^j]}{2t} \\ & \approx \frac{\hat{E}_0[\hat{Y}_t^i \hat{Y}_t^j] - \hat{E}_0[\hat{Y}_t^i] \hat{E}_0[\hat{Y}_t^j]}{2t} \\ & = \frac{1}{2} \sum_{l=1}^M \left(S^{il}(x) + \sum_{k=1}^N \frac{\partial S^{il}}{\partial x^k}(x) \cdot 0 \right) \left(S^{jl}(x) + \sum_{k=1}^N \frac{\partial S^{jl}}{\partial x^k}(x) \cdot 0 \right) + \mathcal{O}(t) \\ & = B^{ij}(x) + \mathcal{O}(t) \quad (t \rightarrow 0) \end{aligned}$$

が近似的に成り立つ. したがって, 一般のマルコフ型確率微分方程式 (2.14) に対しても, $t \rightarrow 0$ の極限を取ることで得られる確率微分方程式の係数の公式 (1.3) に対する $t > 0$ における近似式としては (2.9) が有効である.

以上の議論により, 時系列データからその公式を用いて係数を定量的に評価するときは, (1.4)₂ ではなく (2.2) を用いるのが適切であることがわかる. ただ, (2.2) は

$$\frac{\langle (X_{t+\tau\Delta t}^i - X_t^i)(X_{t+\tau\Delta t}^j - X_t^j) \rangle - (\langle X_{t+\tau\Delta t}^i \rangle - \langle X_t^i \rangle)(\langle X_{t+\tau\Delta t}^j \rangle - \langle X_t^j \rangle)}{2\tau\Delta t}$$

と表されるため, $X_{t+\tau\Delta t}$ と X_t が大概同じセルに滞在するような場合, すなわち相空間分割セルのサイズと比べて時系列データの移動距離が小さい場合は, 既存の計算方法 (1.4)₂ を用いても結果は変わらないというのは事実である. しかし, §2.1

でのドリフトベクトルの計算時の注意から、 $\mathbf{X}_{t+\tau\Delta t}$ が $\mathbf{X}_t$ とは異なるセルに属するように τ を取るようにしたのであった。したがって、そのような時系列の取り方をする場合は、適切にドリフトの影響を除去しないと相対誤差が一様に抑えられないなどの弊害が生じてくるため、拡散行列の定量的評価としては総じて (2.2) がより適切な計算手法となる。

2.3 拡散行列の統計的計算公式に対する考察

確率微分方程式の本質的なランダム性をつかさどるのが拡散行列であるため、そのノルムを計算すると相空間の中でランダム性の強い部分を調べることができる。非負定値行列 $\mathbb{B}$ に対しては

$$\frac{1}{\sqrt{N}}\text{tr}\mathbb{B} \leq \|\mathbb{B}\| \leq \text{tr}\mathbb{B}$$

となるため、ここでは行列ノルムの代わりに $\text{tr}\mathbb{B}$ を用いてその大きさを評価することにする。ただし $\|\mathbb{B}\|$ はフロベニウスノルムである。すると、(2.2) の統計的係数決定公式では、これに対応する量は

$$\text{tr}\bar{\mathbb{B}}_{\text{new}}(\bar{\mathbf{x}}_m) = \frac{\langle |\mathbf{X}_{t+\tau\Delta t} - \mathbf{X}_t|^2 \rangle - \langle \mathbf{X}_{t+\Delta t} - \mathbf{X}_t \rangle^2}{2\tau\Delta t}$$

であり、これはセル $\bar{\mathbf{x}}_m$ から出発する軌道の分散の時間変化度を表している。すなわちこの量は、連続形として考えると確率微分方程式の拡散行列のノルムの離散近似に対応し、離散形として考えると分割セルを通る軌道の分散の平均変化率に対応しているのである。また、 $\bar{\mathbb{B}}_{\text{new}}(\bar{\mathbf{x}}_m)$ 自体は分割セルを通る軌道の共分散行列の時間変化度であり、連続形における拡散行列が確率微分方程式の各成分のノイズのかけり方の強弱の相関を表すことに対応している。したがって、 $\bar{\mathbb{B}}_{\text{new}}(\bar{\mathbf{x}}_m)$ は連続形と離散形のどちらの視点でも系のランダム性を表現するのに適切であり、この公式によって得られる計算値に対してきちんとした意味付けを与えることができる。この確率微分方程式モデリングは両者をつなぐ関係性を見出すことができるという意味でも重要である。

$$\begin{aligned} B^{ij}(\mathbf{x}) &= \lim_{t \rightarrow 0} \frac{\mathbb{E}_{\mathbf{x}}[(X_t^i - x^i)(X_t^j - x^j)] - \mathbb{E}_{\mathbf{x}}[X_t^i - x^i]\mathbb{E}_{\mathbf{x}}[X_t^j - x^j]}{2t} \\ &\approx \frac{\langle (X_{t+\tau\Delta t}^i - X_t^i)(X_{t+\tau\Delta t}^j - X_t^j) \rangle - \langle X_{t+\tau\Delta t}^i - X_t^i \rangle \langle X_{t+\tau\Delta t}^j - X_t^j \rangle}{2\tau\Delta t} = \bar{B}_{\text{new}}^{ij}(\bar{\mathbf{x}}_m). \end{aligned}$$

$\text{tr}\mathbb{B}$ は軌道の分散の平均変化率を表すが、評価時間となる $\tau\Delta t$ が大きければそれだけ軌道は分散してしまう。このため、 $\text{tr}\mathbb{B}$ を $|\mathbf{A}|^2$ で割っておくと、 $\tau\Delta t$ の間の

変動に対するノイズの相対的寄与を得ることができる。この量はいわゆる Wayland test [12] における並進誤差に対応する。Wayland test はスカラー時系列を遅延座標系を用いて高次元空間に埋め込み、近接軌道群の方向分散を計算して力学の次元の間接推定をする手法であり、短いデータ長の時系列データでも有効であるのが利点である。Wayland test はカオス検定に用いられる手法で用途が異なるものの、近接軌道群の方向分散を求めるという点では本研究の趣旨に近いが、並進誤差の計算では方向分散を計算する点の選択性の影響を排除するために幾つかの点における方向分散の中間値や平均値を取るため、時系列データセット1つに対して決まる量である。確率微分方程式モデリングは、相空間の場所場所で軌道の分散の強弱を許す（そして実際にそのように考えると時系列データに対するより深い理解につながる）という意味で、Wayland test の拡張になっていると言える。スカラーデータに対する埋め込み時系列解析の文脈における確率微分方程式モデルを用いた貢献は今後考えるべき問題であろう。

拡散行列の近似式においては、連続形でも離散形でも二次変動量から一次変動量の二次項を除去するのが本稿におけるポイントであるが、これは確率論的な理由によって生じるのではなく、空間離散化（粗視化）した枠組みで計算することで発生する問題点であることがわかる。例を挙げよう。

$$\mathbf{X}(t) = \begin{pmatrix} X(t) \\ Y(t) \end{pmatrix} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (2.16)$$

という $\mathbf{x} = (x, y)^T$ を出発する角速度一定の円周軌道を考える。ここで $\mathbf{x}_0 = (x_0, y_0)^T$ を中心とする一辺が $\delta > 0$ な正方形 $I = [x_0 - \delta/2, x_0 + \delta/2] \times [y_0 - \delta/2, y_0 + \delta/2]$ 上から、一様に (2.16) の軌道が出発するとして各変動量を計算すると、

$$\begin{aligned} \frac{1}{|I|} \int_I (\mathbf{X}(t) - \mathbf{x}) d\mathbf{x} &= \begin{pmatrix} x_0(\cos t - 1) - y_0 \cos t \\ x_0 \sin t + y_0(\cos t - 1) \end{pmatrix}, \\ \frac{1}{|I|} \int_I |\mathbf{X}(t) - \mathbf{x}|^2 d\mathbf{x} &= (1 - \cos t) \left(2|\mathbf{x}_0|^2 + \frac{\delta^2}{3} \right), \end{aligned} \quad (2.17)$$

$$\frac{1}{|I|} \int_I |\mathbf{X}(t) - \mathbf{x}|^2 d\mathbf{x} - \left| \frac{1}{|I|} \int_I (\mathbf{X}(t) - \mathbf{x}) d\mathbf{x} \right|^2 = (1 - \cos t) \frac{\delta^2}{3} \quad (2.18)$$

を得る。(2.18) は I 上の軌道の分散を計算しているため、 $\delta \rightarrow 0$ の極限で $I \rightarrow \{\mathbf{x}\}$ となることを加味すると 0 に収束することは自然である。一方、(2.17) では $\delta \rightarrow 0$ でも ($t > 0$ かつ $|\mathbf{x}_0| > 0$ である限り) 値は 0 に収束しない。これは、(2.17) が I 上の軌道の 2 次モーメントの平均であることを考えると、やはり自然なことである。

これら二つの量は性質の異なる量であり，目的に応じて計算する量を選ばねばならない．本研究で着目しているような系の相空間におけるランダム性の強さには，当然2次モーメントでなく軌道の分散度を採用すべきであるから，(2.18)のように二次変動量から一次変動量の二乗を引く量を計算する方が好ましいということがわかる．

3 数値的検証

ここでは，3次元非線形常微分方程式系の Lorenz system と Rössler system の数値解をある2次元平面に射影した得た時系列データに対して，(2.1)と(2.2)でドリフトベクトルと拡散テンソルを計算し，統計的係数決定公式の有効性を数値的に検証する．ここで，カオス時系列を用いる理由は二度と同じ軌道を通らない長い時系列を作成できるためである．また，3次元空間の解軌道を2次元平面に射影することで，高次元力学を低次元相空間へ縮約した力学を模擬している．

ここで用いる Lorenz system

$$x' = -px + py, \quad y' = -xz + rx - y, \quad z' = xy - bz, \quad p = 10, \quad r = 28, \quad b = 8/3$$

と Rössler system

$$x' = -y - z, \quad y' = x + ay, \quad z' = b + xz - cz, \quad a = 0.2, \quad b = 0.2, \quad c = 5.7$$

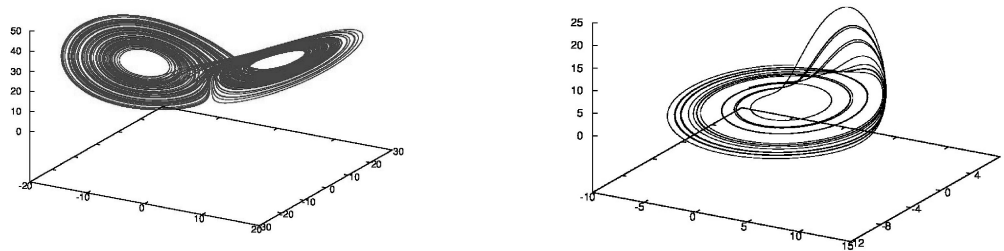


Figure 2: Lorenz system (左) と Rössler system (右) の数値解の軌道．

の数値解は，どちらも初期値 $(1.0, 1.0, 1.0)$ ，時間幅 1.0×10^{-5} を用いて4次 Runge-Kutta 法で時間発展を計算し，初期から $t < 100$ までを遷移過程として打ち切り，

$t = 100$ から 100 ステップおきに 10^8 個データを採取する。このため、データのサンプル時間間隔は $\Delta t = 1.0 \times 10^{-3}$ となる。得られた軌道を Figure 2 に示す。

さらに、得られた 3 次元データに対してそれぞれ主成分分析を行い、上位 2 モードが張る 2 次元部分空間に射影し、これらを解析対象の時系列データとする。射影された軌道は Figure 3 のとおりである。双方とも平面への射影に因る軌道の擬似交差が確認できる。二次元平面上の力学としては、この軌道交差領域では次ステップでの状態の変化が定まらず、二次元の決定論的力学として定式化することができないことを表す。

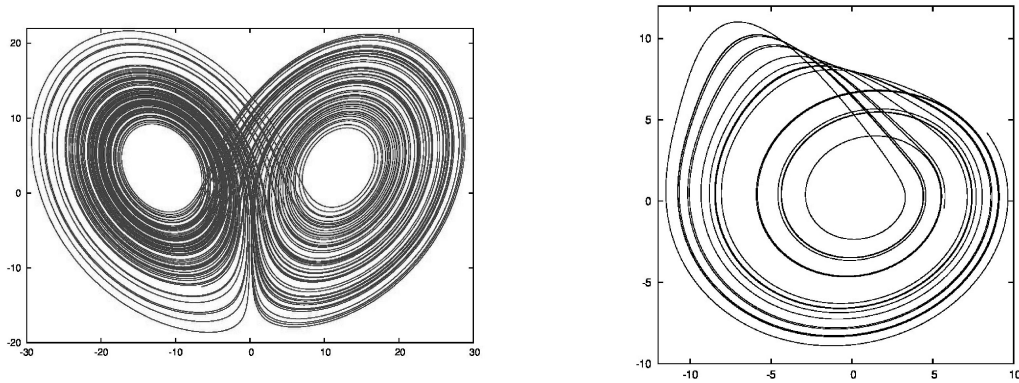


Figure 3: 主成分分析による上位 2 モードが張る部分空間へ射影された Lorenz system (左) と Rössler system (右) の軌道。

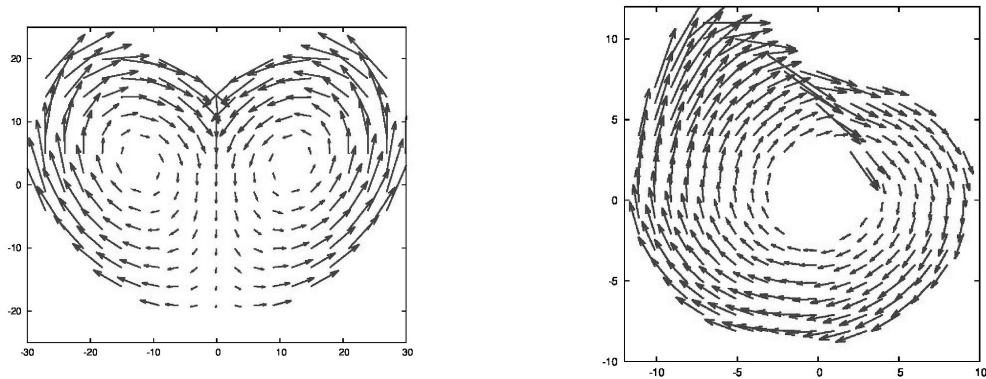


Figure 4: $\bar{A}(\bar{x}_m)$ のベクトルプロット: Lorenz system (左) と Rössler system (右)。

以上の手続きで得られた相空間（射影された平面）上の時系列データに対し、(2.1) と (2.2) の統計的係数決定公式を用いて相空間におけるドリフトと拡散の強さを評価する。Lorenz system では Δt あたりの平均移動距離が 0.09 であるため、 $\tau = 1$ とし、相空間の分割セルのサイズを 0.1 と取る。Rössler system は平均移動距離が 0.008 なので、 $\tau = 10$ として 0.05 の正方形に取る。ドリフトベクトル (Figure

4) は概ね軌道の平均的な挙動を表すことができています。Lorenz system の射影データの場合 (Figure 4 左), 軌道が相空間の原点近傍で交差し且つ交差する軌道の数も左右で拮抗しているため, 平均的にベクトルは第2主成分が負の方向を向く。したがって, 左半平面から右半平面へ, またはその逆に移動するためのドリフトベクトルは得られていない。すなわち, この枠組みではそのような相空間内の移動は2次元平面上の決定論的力学としては表されないことを示している。また, Rössler system の場合 (Figure 4 右) は軌道の向き (時計回り) と軌道の交差部はきちんと抽出できている。また, 元の3次元空間では軌道が3次元的に他の軌道を乗り越えるときに速く遷移するが, その軌道の速さは射影データにもドリフトベクトルの大きさとして表されている。

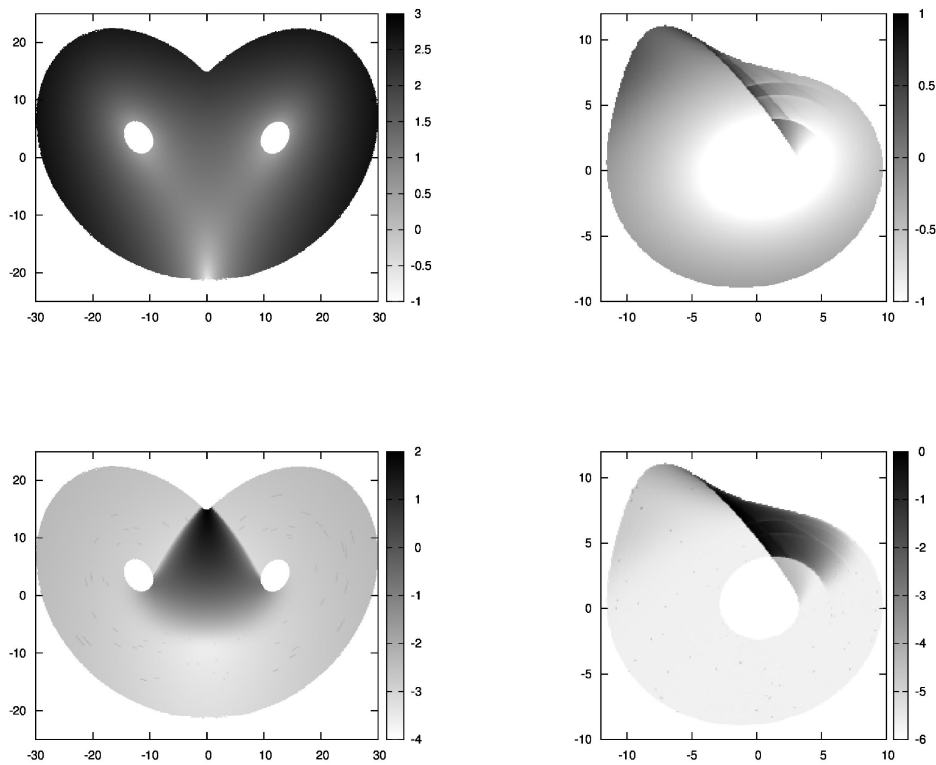


Figure 5: 主成分分析による上位2モードが張る部分空間へ射影されたデータから求めた $\text{tr}\bar{\mathbb{B}}$ (上) と $\text{tr}\bar{\mathbb{B}}_{\text{new}}$ (下) の常用対数のプロット: Lorenz system (左), Rössler system (右). 双方ともに時系列点が100以上滞在したセルのみ表示している。

次に拡散の大きさを表す $\text{tr}\bar{\mathbb{B}}_{\text{new}}$ と従来の公式 (1.4)₂ で求めた $\text{tr}\bar{\mathbb{B}}$ の図を Figure 5 に示す。まず, 従来の公式による $\text{tr}\bar{\mathbb{B}}$ の値では, 軌道が速い (遅い) 部分で値が大

きい（小さい）という情報しか得ることができないことが Lorenz system, Rössler system の双方からわかる (Figure 5 左上, 右上). これは (2.11) にあるように, $\bar{\mathbb{B}}$ の評価にドリフトベクトルの寄与が大きく出ているためである. 一方で, 修正された公式での $\text{tr}\bar{\mathbb{B}}_{\text{new}}$ では Lorenz system, Rössler system のどちらも軌道交差の領域で値が大きくなっており (Figure 5 左下, 右下), そこでの軌道の分散を捉えられていることがわかる. ただ, $\text{tr}\bar{\mathbb{B}}_{\text{new}}$ にも軌道の速さが影響を及ぼさないわけではない.

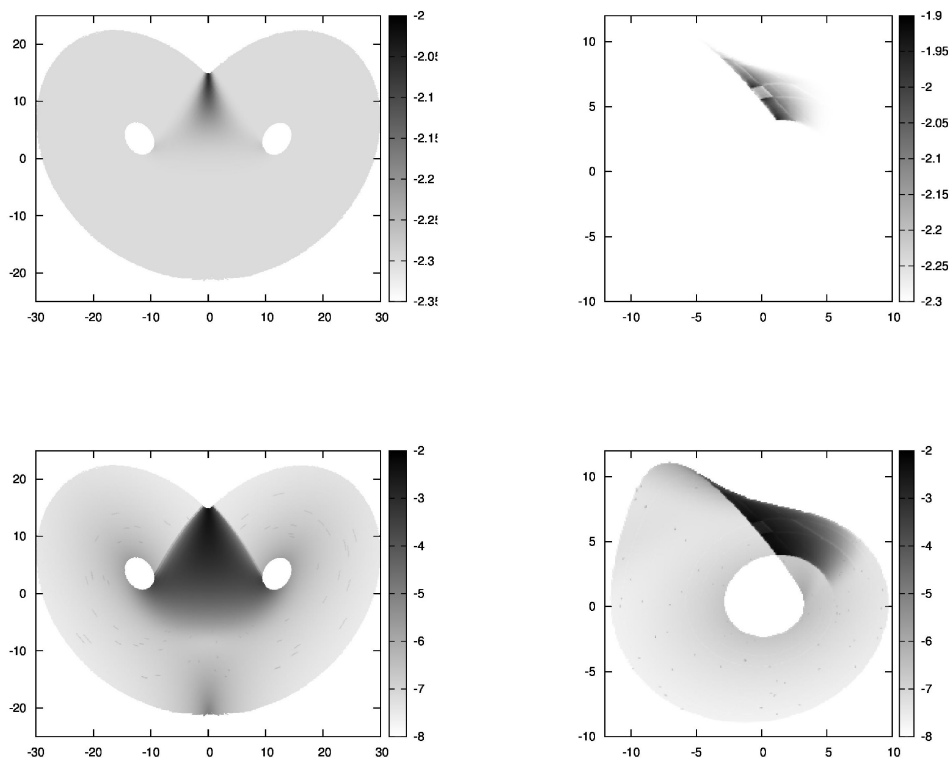


Figure 6: 主成分分析による上位2モードが張る部分空間へ射影されたデータから求めたノイズ指標（下）と従来の公式を用いた $\text{tr}\bar{\mathbb{B}}/|\bar{\mathbf{A}}|^2$ （上）の常用対数のプロット：Lorenz system（左），Rössler system（右）. 双方ともに時系列点が100以上滞在したセルのみ表示している.

特に Lorenz system の方が顕著であるが, 軌道の速い第1主成分の絶対値が大きい領域では $\text{tr}\bar{\mathbb{B}}_{\text{new}}$ の値が大きい傾向が見え, かつ軌道の遅い第1主成分が0付近かつ第2主成分が負の領域では値はとても小さくなっていることがわかる. 軌道の単位時間当たりの移動距離が長い領域では, 仮に決定論性が強くとも軌道の分散は大

きくなり得るからである。これは Example 2.4 の注意で指摘したように、ドリフトの強さがその評価に関わってくることも関係している。したがって、 $\text{tr}\bar{\mathbb{B}}_{\text{new}}$ では軌道の分散を適切に評価する指標であるといえる一方で、時系列の決定論性と確率論性の分離という点に着目する場合はドリフトベクトルの大きさを規格化の方が良い。

Figure 6 は拡散行列のノルムとドリフトベクトルの大きさの二乗の比を示したものである。これは軌道の変動に対する決定論性と確率論性の強さの指標であり、値が大きければ軌道の変動には確率論的な寄与が大きいことを意味する。従来の公式を用いた $\text{tr}\bar{\mathbb{B}}/|\bar{\mathbf{A}}|^2$ の場合は (Figure 6 上), 新しい公式を用いたもの (Figure 6 下) と比較すると Lorenz system, Rössler system 共に値域が狭く、強弱の傾向を掴みづらいことがわかる。Figure 6 (下) の $\text{tr}\bar{\mathbb{B}}_{\text{new}}/|\bar{\mathbf{A}}|^2$ では, Lorenz system, Rössler system 双方とも軌道交差の領域で値が大きくなっているが、軌道の速くかつ決定論性の強い領域 (Lorenz では原点から遠い領域, Rössler では下半平面の原点から離れた領域や軌道交差直前の領域) では小さくなる傾向が見られる。 $\text{tr}\bar{\mathbb{B}}_{\text{new}}/|\bar{\mathbf{A}}|^2$ では軌道の速さによる影響も規格化されるため、軌道の変動のうち確率論性の寄与が強い領域を検出することができている。特に不安定固定点まわりでその値が小さくなっていることも特筆すべきことである。固定点まわりは軌道速度が遅いもののその不安定性によって軌道がばらけるため、相対的に確率論性が高い領域となっているとみなせる。

これらの例で示した通り、新しい公式 $\text{tr}\bar{\mathbb{B}}_{\text{new}}$ を用いると適切に相空間内の拡散 (軌道の分散の平均変化率) の強い領域を抽出することが可能であり、本研究で提案する統計的係数決定公式 (2.1), (2.2) は数学的にも数値的にも機能することがわかった。また、 $\text{tr}\bar{\mathbb{B}}_{\text{new}}$ と $\text{tr}\bar{\mathbb{B}}_{\text{new}}/|\bar{\mathbf{A}}|^2$ を合わせて調べることにより、軌道の分散が大きい部分からさらに確率論的寄与が大きい領域を捉えることができる。これにより時系列データの決定論性と確率論性の強弱が説明付けられるため、本稿で扱った確率微分方程式モデリングの有効性を示している。

4 謝辞

本研究を進めるにあたり、MaeT 数学-気象学連携研究チーム、京都大学大学院理学研究科の坂上貴之教授、京都大学数理解析研究所の山田道夫教授、東北大学原子分子材料科学高等研究機構の小布施祈織助教、京都大学大学院理学研究科の森田英

俊博士研究員には有意義な議論や厳しくも温かいご指摘をいただきましたことをこの場を借りて御礼申し上げます。また、本研究は文部科学省 H22 年度「気候変動適応戦略イニシアチブ」気候変動適応研究推進プログラム「北海道を対象とする総合的ダウンスケール手法の開発と適用」と日本学術振興会科学研究費助成事業（挑戦的萌芽研究：課題番号 25610028）の支援を受けたものです。

References

- [1] J. BERNER: Linking nonlinearity and non-Gaussianity of planetary wave behavior by the Fokker-Planck equation. *J. Atmos. Sci.* **62**, 2098-2117 (2003).
- [2] G. BRANSTATOR AND J. BERNER: Linear and nonlinear signatures in the planetary wave dynamics of an AGCM: phase space tendencies. *J. Atmos. Sci.* **62**, 1792-1811 (2003).
- [3] A. BHADRA, E. L. IONIDES, K. LANERI, M. PASCUAL, M. BOUMA AND R. C. DHIMAN: Malaria in northwest India: Data analysis via partially observed stochastic differential equation models driven by Lévy noise. *J. Am. Stat. Assoc.* **106**, 440-451 (2011).
- [4] S. DITLEVSEN AND A. DE GAETANO: Mixed effects in stochastic differential equation models. *REVSTAT* **3**, 137-153 (2005).
- [5] K. HASSELMANN: Stochastic climate models. Part I. Theory. *Tellus* **28**, 473-485 (1976).
- [6] A. F. IVANOV AND A. V. SWISHCHUK: Optimal control of stochastic differential delay equations with application in economics. *Int. J. Qual. Theory Diff. Eq. Appl.* **2**, 201-213 (2008).
- [7] M. INATSU, N. NAOTO AND H. MUKOUGAWA: Dynamics and predictability of extratropical wintertime low-frequency variability examined by a stochastic differential equation in a low-dimensional system. *J. Atmos. Sci.*, **70**, 939-952 (2013).

- [8] E. N. LORENZ: A study of the predictability of a 28-variable atmospheric model. *Tellus* **17**, 321-333 (1965).
- [9] V. MISRA, W. -B. GONG AND D. TOWSLEY: Stochastic differential equation modeling and analysis of TCP-window size behavior. *J. Atmos. Sci.* **62**, 1792-1811 (2003).
- [10] 中野直人, 稲津將, 向川均, 楠岡誠一郎: 確率微分方程式を用いた気候モデルについて. 京都大学数理解析研究所講究録 **1823**, 79-96 (2013).
- [11] P. SURA, M. NEWMAN, C. PENLAND AND P. D. SARDESHMUKH: Multiplicative noise and non-Gaussianity: A paradigm for atmospheric regimes? *J. Atmos. Sci.* **62**, 1391-1409 (2005).
- [12] R. WAYLAND, D. BROMLEY AND D. PICKETT: Recognizing determinism in a time series. *Phys. Rev. Lett.* **70**, 580-582 (1993).
- [13] D. J. WILKINSON: Stochastic modelling for quantitative description of heterogeneous biological systems. *Nature* **122**, 122-133 (2009).
- [14] N. ZABARAS AND SETHURAMAN SANKARAN: An information-theoretic approach to stochastic materials modeling. *Computing in Science & Engineering* **9**, 50-59 (2007).