

機械学習に基づくマクロ経済変動の数理モデリング

齊木 吉隆 (一橋大学)

概要

近年、機械学習の一種で、リカーレントネットワークをもったリザーバーコンピューティングが決定論的構造を背後にもつ時系列やそのリャプノフ指数などの予測に有効であることが報告されている。我々はこの機械学習手法を用いて比較的平穏な流体のエネルギー変数のモデルを構成し、予想を成功させている (Nakai and Saiki, Physical Review E, 2018)。この知見を経済時系列データに応用して経済予測に取り組むために必要な決定論的な構造の抽出や、機械学習により得られた力学系モデルの妥当性を検証するために、決定論的に時系列データがどの程度学習できるか解明した。

1 共同研究に関する情報

1.1 共同研究を実施した拠点名

京都大学, 東京大学, 大阪大学

1.2 共同研究分野

- 超大規模数値計算系応用分野
- 超大規模データ処理系応用分野

1.3 参加研究者の役割分担

本共同研究は 4 人体制でおこなわれている。

- 研究代表者 齊木吉隆: 背後にある数学的構造の解明、リザーバーコンピューティングによる数理モデル構築
- 副代表者 小林幹 (立正大学 経済学部): 背後にある数学的構造の解明
- その他の共同研究者 中井拳吾 (東京海洋大学 学術研究院 流通情報工学部門): トレーニングデータの作成、リザーバーコンピューティングによる数理モデル構築
- その他の共同研究者 武藤誠 (一橋大学 大学院経営管理研究科): トレーニングデー

データの作成

各側面においてスーパーコンピュータが活用される。

2 研究の目的と意義

機械学習を用いたデータ駆動型研究が近年急速に発展している。本研究では、この手法を経済データに適用することにより、新たな知見の蓄積を目指す。特に、集計結果が得られるまでに長い時間を要する景気指標を、迅速かつ精緻に推測することを目的とする。こうした知見の蓄積は政府の景気対策の有効性向上につながるばかりでなく、企業経営にとっても有益な情報となるため、経済安定と経済発展に貢献することが期待される。研究代表者は最近の研究で、リザーバーコンピューティングと呼ばれるリカーレントニューラルネットワークを用いた機械学習手法と力学系理論の知見を用いて、流体乱流に関するマクロ統計量の時系列データからマクロ統計量の時間発展を捉えるモデルを構築することに成功した。同種の手法を応用することに

より、経済変動に関する定量的な時間発展モデルの構築とそれを用いた将来予測をおこなう。本研究は、力学系理論、数理モデリング、機械学習、経済学、大規模計算にまたがる真の学際的・融合的研究であると言える。それゆえ、いずれの分野の知見を欠いても十分満足のいく結果が得られることはないという意味で極めてチャレンジングな数理科学研究である。

研究代表者は機械学習を用いた流体乱流に関するマクロ変数に着目した数理モデリングに成功しているが、それはあくまで決定論的な古典力学に立脚した流体力学変数のマクロ集約量であり、時々刻々支配方程式が変化するようなことはない。一方、一国の経済変動は多くの経済主体（政府、企業、家計、諸外国等）による集団現象である上に、決定論的な要素と確率論的な要素の双方に由来する部分と、政策や制度の変更等、構造変化に由来する部分が入り混じっており、それが定量的なモデリングを困難にさせてきた。高周波の振る舞いは主に確率的なものであり、低周波の振る舞いは主に構造変化等によるものであると考えられる。中周波の振る舞いは確率論的な振る舞いが集約されたマクロ的な振る舞いを含む決定論的なものが主であり、かなり再帰性の高いダイナミクスを示していると期待できる。本研究では中周波の振る舞いを描写する時系列データを学習データとして用いた機械学習をおこない、これまで存在しなかった定量的な決定論的経済変動モデルの構築を目指す。経済学分野で機械学習が適用された事例はほとんどないことから、極めてチャレンジングな研究課題であるといえる。

3 当拠点公募型研究として実施した意義

リザーバーコンピューティングに基づく物理ダイナミクスの予測に関しては、アメリカのメ

リーランド大学のグループがこれまでリードしてきたが、リザーバーコンピューティングを用いた流体乱流の数理モデリングならびにそれを用いた予測に、我々、研究代表者と中井拳吾氏のグループがはじめて成功した。拠点公募型共同研究として数学・経済学・大規模計算の学際的なチームを編成して多くのスーパーコンピュータ資源を用いた研究を推進することによって、マクロ経済変動に関するリザーバーコンピューティングを用いた数理モデリングにおいても我が国が世界をリードすることが期待される。

本研究課題では時系列データの機械学習によるモデリングを行った。時系列データの力学系構造を精度良く学習するためには、大量の時系列データを取得し大規模な計算資源を使うことが必要不可欠である。

このためには、JHPCN の枠組みを最大限に活用し、長時間の連続計算が可能な京都大学の system A での長時間データの学習と、非常に高い並列計算を得意とする東京大学の Oakforest-PACS での力学系解析、また、GPU 計算機である大阪大学の OCTOPUS を使い分けて計算することによって研究目標の達成が実現できた。また、並列化計算の実績や指導経験豊富な情報基盤センターと連携することでより効率よく研究を進めることができた。

4 前年度までに得られた研究成果の概要

新規課題のため該当しない。

5 今年度の研究成果の詳細

5.1 都道府県別鉱工業生産指数の同期解析

以下では、都道府県別鉱工業生産指数に対する結果の概要を紹介する。まず、都道府県別鉱工業生産指数の時系列の例を図 1 で示す。スペク

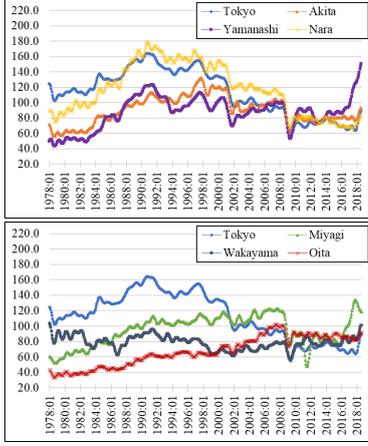


図 1 Seasonally-adjusted Time Series Comparison of IIP Data for Tokyo and Three Other Prefectures (Note: Part(a) compares Tokyo with Akita, Yamanashi, and Nara Prefectures. Part(b) compares Tokyo with Miyagi, Wakayama, and Oita Prefectures. Data for the former (latter) prefectures show the greatest (least) synchronization with Tokyo.)

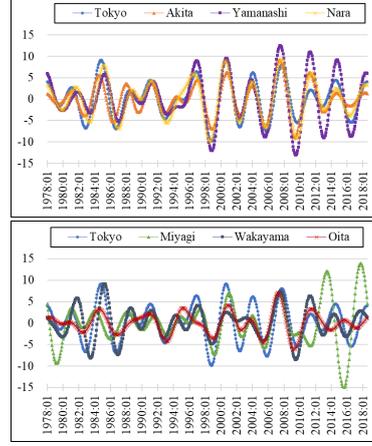


図 2 Comparison of Band-pass-filtered Time Series of IIP Data for Tokyo and the Sampled Prefectures (Note: In Part(a) Akita, Yamanashi, and Nara Prefectures are most closely synchronized with Tokyo. In Part(b) Miyagi, Wakayama, and Oita Prefectures are least synchronized with Tokyo. The upper (lower) cutoff frequency of the band-pass filter is 35 (81) months.)

トル解析に基づいて、およそ3年から7年程度のタイムスケールに着目し、その周波数を取り出すバンドパスフィルターをかけた時系列データを図2に示す。これらの時系列間の同期レベルを評価するためにヒルベルト変換を用いて各時刻における位相を定める。ヒルベルト変換とは、時系列 s_t に対して、

$$s_t^H = \frac{1}{\pi} P.V. \int_{-\infty}^{\infty} \frac{s_{\tau}}{t - \tau} d\tau, \quad (1)$$

で定められるものである。ここで、 $P.V.$ とはコーシーの主値積分である。 s_t を実軸として、そのヒルベルト変換 s_t^H を虚軸にもつ $\hat{s}_t = s_t + is_t^H$ の偏角によって、時刻 t における時系列 s_t の位相を定める。東京都の鉱工業生産指数に対する \hat{s}_t の動きを図3に示す。時刻 t における2時系列の位相差 ψ_t の時間変化の様子を調べることで時系列間の同期レベルを評価した。次の同期指標 $\gamma_t^2 =$

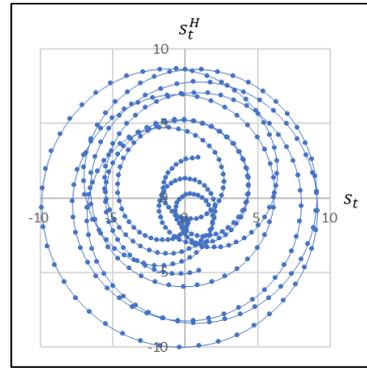


図3 Time Series of Trajectory $P_t(s_t, s_t^H)$ on the complex plane (Note: The variable s_t denotes the band-pass-filtered IIP of Tokyo and s_t^H its Hilbert-transformed time series.)

$\left(\frac{1}{W} \sum_{i=t-p}^{t+p} \cos \psi_i\right)^2 + \left(\frac{1}{W} \sum_{i=t-p}^{t+p} \sin \psi_i\right)^2$, (但し、 $p = (W - 1)/2$, $0 < p < t$) (M. Rosenblum, A. Pikovsky, J. Kurths, C. Schafer and P. A. Tass (2001)) が1に近いときに時刻 t 付近において2時系列間の同期レベルが高く、0

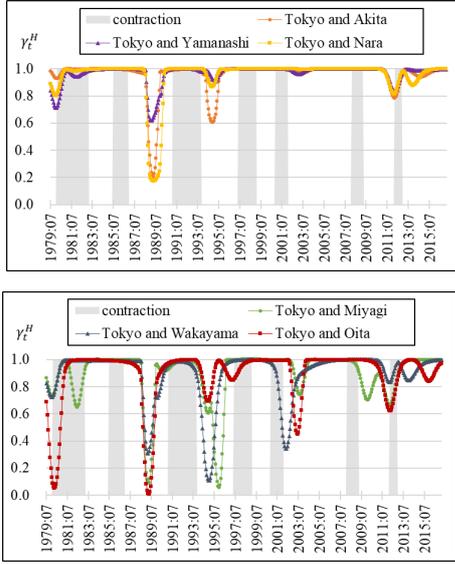


図4 Time Evolution of Synchronization Index γ_t^2 between Tokyo and the Sampled Prefectures (Note: In Part(a) prefectures Akita, Yamanashi, and Nara are most closely synchronized with Tokyo. In Part(b) prefectures Miyagi, Wakayama, and Oita are least synchronized with Tokyo.)

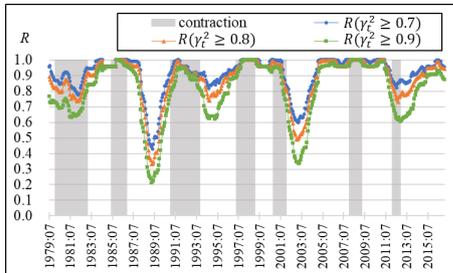


図5 Time Evolution of the Ratio $R(\gamma_t^2 \geq r)$ for $r = 0.7, 0.8,$ and 0.9 (Note: By $R(\gamma_t^2 \geq r)$ we denote the ratio of 2-tuples among 47 Japan's prefectures for which γ_t^2 shows a value greater than or equal to the threshold r at each time t .)

に近いときに同期レベルが低いと見なせる。図4（上）には東京と同期レベルが高い3県と東京との間の同期指標の時間変化を示している。図4（下）には逆に東京と同期レベルが低い3県と東京との間の同期指標の時間変化を示して

いる。図4ならびに次の図5の背景のグレー部分は内閣府により発表された景気後退期間に対応しており、白部分は景気上昇期間に対応している。景気後退期は同期レベルが高いことが見出された。図5では、47都道府県の ${}_{47}C_2$ 種類すべての組み合わせに対して同期指標を計算し、各時間において何種類がしきい値を越す同期レベルをもっていたかを示している。この解析においても図4のいくつかの組み合わせで確認される結果、すなわち今回着目したタイムスケールでの都道府県別鉱工業生産指数は都道府県間で比較的高い同期レベルを維持する傾向があり、景気上昇期には同期レベルが下がる傾向があることが確認される。

5.2 機械学習モデルの力学系解析

決定論的な時系列データとしてローレンツ方程式の学習を行った。ローレンツ方程式はあるパラメータのもとでカオスな振る舞いをすることが知られている。学習により得た機械学習モデルの力学系的性質を解析し、その一部の結果は論文（学術論文 [c]）として発表した。その結果の概要を以下に述べる。

リザーバーコンピューティング. 力学系 $\frac{d\phi}{dt} = \mathbf{f}(\phi)$ に関し、その変数 $\mathbf{u} = \mathbf{h}_1(\phi) \in \mathbb{R}^M$ と $\mathbf{s} = \mathbf{h}_2(\phi) \in \mathbb{R}^P$ について考える。ただし、時系列データ \mathbf{u} は \mathbf{s} の時系列を予測する時刻にデータはないとする。リザーバーベクトル $\mathbf{r} \in \mathbb{R}^N$ ($N \gg M$) は次の式で支配されている。

$$\mathbf{r}(t + \Delta t) = (1 - \alpha)\mathbf{r}(t) + \alpha \tanh(\mathbf{A}\mathbf{r}(t) + \mathbf{W}_{\text{in}}\mathbf{u}(t)), \quad (2)$$

ただし、 Δt は時間幅を表す。また、行列 \mathbf{A} , \mathbf{W}_{in} は $N \times N$, $N \times M$ 行列である。パラメータ α ($0 < \alpha \leq 1$) は \mathbf{r} の力学系の非線形性を表し、式 (2) の Δt に依存して決める。リザーバーベクトルは $(0, 1]$ の一様乱数から定めた $\mathbf{r}(-\tau)$ を初期値とする。ただし、 $\tau/\Delta t (\gg 1)$

はトランジェント時間に相当する。

行列 \mathbf{W}_{in} の各列はただ一つのために $[-\sigma, \sigma]$ の一様乱数から値を与え、それ以外は 0 とする。行列 \mathbf{A} は $D \times N$ 個の非零成分を持つ疎行列とする。 $D \times N$ 個の非零成分は $[-1, 1]$ の一様乱数から値を与える。さらに、行列 \mathbf{A} については最大固有値が ρ になるように行列のすべての要素を調整する。

リザーバーベクトル \mathbf{r} は $\hat{\mathbf{s}}(t) = \mathbf{W}_{\text{out}}\mathbf{r}(t) + \mathbf{c}$ により、得たいベクトルに変換する。これにより $L = T/\Delta t$ ステップのリザーバーベクトル $\{\mathbf{r}(l\Delta t)\}_{l=1}^L$ を得る。

$0 < t \leq T$ (学習時間と呼ぶ) での時系列 $\hat{\mathbf{s}}$ が既知の時系列 \mathbf{s} の近似になるように \mathbf{W}_{out} と \mathbf{c} を決定する。この決定がリザーバーコンピューティングの学習に相当する。これは次の二次形式が最小化するように \mathbf{W}_{out} と \mathbf{c} を決定することに対応する。

$$\sum_{l=1}^L \|(\mathbf{W}_{\text{out}}\mathbf{r}(l\Delta t) + \mathbf{c}) - \mathbf{s}(l\Delta t)\|^2 + \beta[\text{Tr}(\mathbf{W}_{\text{out}}\mathbf{W}_{\text{out}}^T)], \quad (3)$$

ただし \mathbf{q} に対して $\|\mathbf{q}\|^2 = \mathbf{q}^T\mathbf{q}$ とし、 $\beta(\geq 0)$ を含む項は \mathbf{W}_{out} の過剰適合を避けるために導入している。この決定が妥当であれば、 $t > T$ (予測時間と呼ぶ) での $\hat{\mathbf{s}}(t)$ は $\mathbf{s}(t)$ を再現できるはずである。二次形式 (3) を最小化する解を $\mathbf{W}_{\text{out}}^*$ と \mathbf{c}^* とすると予測時間では

$$\hat{\mathbf{s}}(t) = \mathbf{W}_{\text{out}}^*\mathbf{r}(t) + \mathbf{c}^*, \quad (4)$$

として時系列 $\hat{\mathbf{s}}$ を得る。 $\mathbf{W}_{\text{out}}^*$ と \mathbf{c}^* は次のように書くことができることが知られている:

$$\mathbf{W}_{\text{out}}^* = \delta\mathbf{S}\delta\mathbf{R}^T(\delta\mathbf{R}\delta\mathbf{R}^T + \beta\mathbf{I})^{-1}, \quad (5)$$

$$\mathbf{c}^* = -[\mathbf{W}_{\text{out}}^*\bar{\mathbf{r}} - \bar{\mathbf{s}}], \quad (6)$$

ただし、 $\bar{\mathbf{r}} = \sum_{l=1}^L \mathbf{r}(l\Delta t)/L$, $\bar{\mathbf{s}} = \sum_{l=1}^L \mathbf{s}(l\Delta t)/L$ とし、 \mathbf{I} は $N \times N$ の単位行

列、 $\delta\mathbf{R}$ (rsp, $\delta\mathbf{S}$) は l 行が $\mathbf{r}(l\Delta t) - \bar{\mathbf{r}}$ (rsp, $\mathbf{s}(l\Delta t) - \bar{\mathbf{s}}$) となるような行列とする。

ただし、予測時間 $t > T$ での時系列データ \mathbf{u} は利用できないため、 $\mathbf{u}(t)$ として式 (2) で得た $\hat{\mathbf{s}}(t) = \hat{E}(\cdot, t)$ を用いる:

$$\mathbf{r}(t+\Delta t) = (1-\alpha)\mathbf{r}(t) + \alpha \tanh(\mathbf{A}\mathbf{r}(t) + \mathbf{W}_{\text{in}}\hat{\mathbf{s}}(t)).$$

学習に用いる変数 $X(t)$ を対等に扱うために正規化した $\tilde{X}(t)$ を用いる: $\tilde{X}(t) = [X(t) - X_1]/X_2$. ただし、 X_1 は平均値、 X_2 は標準偏差を表す。予測時間の時系列データ $X(t)$ を $\tilde{X}(t)$ から再現するときは、学習時間の正規化に用いた X_1 と X_2 を用いる。この正規化より σ の調整を避けることができる。

ローレンツ方程式. ローレンツ方程式 $\dot{x} = -10x + 10y$, $\dot{y} = -xz + rx - y$, $\dot{z} = xy - 8z/3$ について考える。この方程式系はカオス性がある。そこで、本研究ではパラメータ r として 28 (もしくは 60) について、ローレンツ方程式の時間により得られた時系列データをリザーバーコンピューティングにより学習させて得られた機械学習時間発展モデルが、元の力学系構造をどの程度再現するかを解析する。

機械学習モデルの解析-2 次元射影. 学習により得た機械学習時間発展モデルのアトラクタの大きさや概形について、ローレンツ方程式のものと比較する。3次元では比較が困難であるため、2次元平面に落として考察する。具体的には解軌道が適当な平面を横切るときにプロットして得るポアンカレ断面というものをを用いて比較をする。学習により得た機械学習時間発展モデルから得られた 10^6 時間の時系列データとローレンツ方程式の時間により得られた 10^6 時間の時系列データ、学習に用いた 5000 時間の時系列データに対して、適当な平面として

$x = y$ となる平面にポアンカレ断面を書いたものが図 6 である。

まず、学習に用いた 5000 時間の時系列データにより得たポアンカレ断面はローレンツ方程式の 10^6 時間の時系列データにより得られたポアンカレ断面に比べて小さいことが見て取れる。他方で、機械学習時間発展モデルから得られたポアンカレ断面はローレンツ方程式の 10^6 時間の時系列データにより得られたポアンカレ断面と似ていることがわかる。このことから、ローレンツ方程式の 10^6 時間の時系列データにより得られたポアンカレ断面に比べて学習に用いた 5000 時間の時系列データにより得たポアンカレ断面は小さいにもかかわらず、得られた学習モデルであっても機械学習時間発展モデルから得られたポアンカレ断面はローレンツ方程式の断面を十分に再現していることがわかる。

機械学習モデルの解析-密度分布. 本解析では $r = 28$ と $r = 60$ の二つの場合を調べる。 $r = 60$ 付近では力学系は非双曲的であり、パラメータの微小な変化により力学系の構造が大きく変化しうる一方で、 $r = 28$ 付近では力学系は双曲的で構造安定であり、そのような構造変化は起こらないことが知られている。ここでは機械学習時間発展モデルから得られた 10^6 時間の密度分布とローレンツ方程式の時間により得られた 10^6 時間の密度分布、学習に用いた 5000 時間の密度分布に対して、 x 変数の時系列データの滞在時間密度分布を計算した (図 7)。

学習に用いた 5000 時間の時系列データにより得た密度分布はローレンツ方程式の 10^6 時間の時系列データにより得られた密度分布に比べて激しく振動していることが見て取れる。他方で、機械学習時間発展モデルから得られた密度分布とローレンツ方程式の 10^6 時間の時系列データにより得られた密度分布の差はかなり小さいことが見て取れる。このことから、学習に

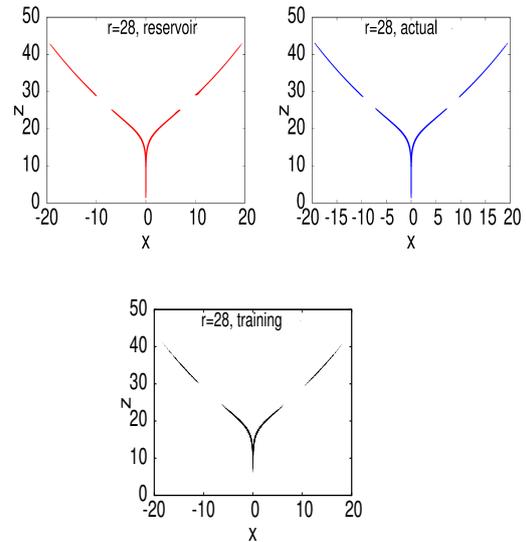


図 6 **Poincaré section-like plots** ($r = 28$). The sets of points (x, z) : along a trajectory of the data-driven model using reservoir computing and along a long trajectory of the actual Lorenz system and along a short trajectory used for the training data are plotted when $|x - y| < 0.05$. The time lengths of the three trajectories are $T = 10^6, 10^6$, and 5000, respectively.

用いた時系列データから得られた密度分布は時間が短いために激しく振動しているにもかかわらず、機械学習時間発展モデルから得られた密度分布はローレンツ方程式の断面を十分に再現していることがわかる。特に、 $r = 60$ の密度分布は特異な構造があるが、この構造も機械学習モデルは再現していることが見て取れる。

機械学習モデルの解析-不動点. ローレンツ方程式には上で見たカオスアトラクタ以外に不動点や周期軌道といった不変集合が存在する。例えば $(x, y, z) = (0, 0, 0)$ は不動点である。ローレンツ方程式には $(x, y, z) = (0, 0, 0)$ に加えて、他に 2 つの不動点 (図 8 の + のところ) がある。学習により得た時間発展モデルの不動点も + の近傍にあり、そこでの固有値も

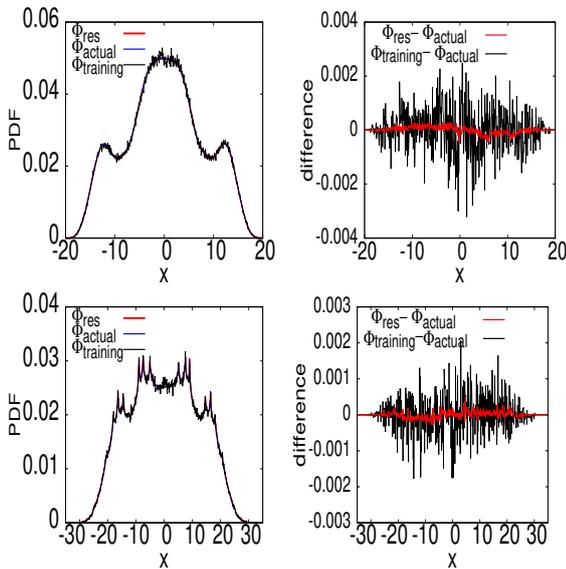


図7 Density distributions of a variable ($r = 28$ (top) and $r = 60$ (bottom)). The density distribution of the x variable calculated from a length $T = 10^6$ trajectory of the data-driven model by reservoir computing (Φ_{res}) is plotted together with that computed from the length $T = 10^6$ long trajectory of the actual Lorenz system (Φ_{actual}) and the length $T = 5000$ short trajectory (Φ_{training}) used as training data for constructing the data-driven model. The difference in the distributions are shown on the right.

近似できていることを確認した。

これらのことから、学習によって構成したモデルは学習に用いた時系列データには含まれない不動点をも再現することがわかった。このことは機械学習分野における通説である「外部領域の予測困難である」ということからは一見想像し難い結果であるが、背後に存する力学系構造の学習に成功したことによって不動点が再現できたものと考えられる。

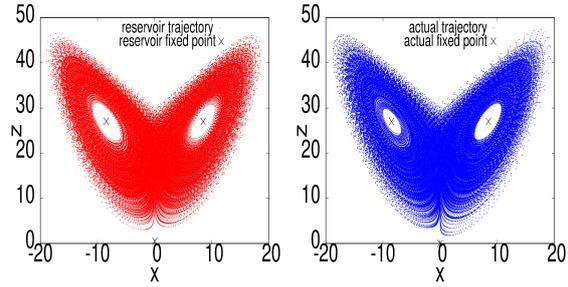


図8 Fixed points. The three fixed points (x^*, y^*, z^*) of the data-driven model (top) and the corresponding unstable fixed points of the actual Lorenz system (bottom) are plotted together with the trajectory points.

6 今年度の進捗状況と今後の展望

実施計画において次の2つの項目

- A. マクロ経済変数の推定モデルの構成
- B. 適切なバンドパスフィルターの構成
- C. トイモデルを用いた学習の妥当性の評価

を予定していた。

項目 [A] については、予備的研究として、まず、LSTM (Long short-term memory) と呼ばれるリカレントニューラルネットワークを、ドル円為替レートなどいくつかの時系列データに適用して予測を試みたが、高々 AR モデル等の手法の予測力と大差がないことを確認し、[B] の目的である経済変動における決定論的構造を抽出することの重要性を確認した。

項目 [B] については、複数の経済時系列から決定論的情報を取り出す試みを行った。経済変数の時系列データは、短いタイムスケールは確率的な要素が支配的であり、長いタイムスケールは構造変化が支配的であると考えられ、その間のタイムスケールに決定論的かつ回帰的なダイナミクスが潜んでいることが期待される。そ

ここで、重要な経済変数である為替レートならびに都道府県別鉱工業生産指数の分析をおこなった。本研究では経済変数の時系列にみられる決定論的構造を取り出すことが重要である。為替レートの決定要因に購買力平価説 PPP がある。本研究では、2 国間の為替レートが両国の物価水準に調整されていく構造が見られるタイムスケールをスペクトル解析で見出し、3 年から 20 年くらいのタイムスケールに購買力平価に起因する決定論的ふるまいが観察されることをヒルベルト変換を用いた同期解析により確認した。また、都道府県別鉱工業生産指数に関しては、半年から 7 年程度のタイムスケールのダイナミクスに決定論的ふるまいが観察されることを確認した。これらの結果を論文 [a,b] としてまとめて査読付き国際誌に投稿した。

項目 [C] については、決定論的な時系列から機械学習力学系モデルを構築し、それが時系列の背後にあるどのような力学系構造まで再現しうるかを明らかにした。本研究では、不動点、周期点のような力学系の不変集合や、その安定性、不安定性ならびに安定多様体、不安定多様体といった各種力学系的性質が再現しうることを示した。結果は、論文 [c] としてまとめて査読付き国際誌に投稿した。

なお、論文 [a,b,c] はいずれもプレプリントアーカイブにもアップロードしている。

7 研究業績一覧（発表予定も含む）

国際会議発表（査読なし）

1. K. Nakai, SOI Asia, Machine Learning for Time Series Data, Keio University, Jan. 2021.
2. K. Nakai, Differential Equations for Data Science 2021, "Dynamical system analysis of a data-driven model constructed by reservoir computing" Kyoto

University, Mar.10th 2021.

国内会議発表（査読なし）

3. K. Nakai, 工学と数学の接点を求めて、「機械学習により得た時間発展モデルの力学系構造」大阪大学, 2020 年 11 月.

その他（特許, プレス発表, 著書等）

- a. M. Muto and Y. Saiki, *Synchronization analysis between exchange rates based on purchasing power parity using the Hilbert transform*, submitted, arxiv 2010.08835.
- b. M. Muto, T. Onozaki and Y. Saiki, *Synchronization of Prefectural Business Cycles in Japan 1978–2018*, submitted, arxiv 2010.08825.
- c. M. Kobayashi, K. Nakai, Y. Saiki, and N. Tsutsumi, *Dynamical system analysis of a data-driven model constructed by reservoir computing*, submitted, arxiv 2102.13475.