

平成 22年 5月 20日現在

研究種目： 若手研究(B)
 研究期間： 2008～2009
 課題番号： 20700217
 研究課題名（和文） ネットワーク構造の推定を基盤とした複雑システムの理解とその応用
 研究課題名（英文） Estimating Network Structure of Multivariate Complex Systems and Its Applications
 研究代表者
 鈴木 智也（SUZUKI TOMOYA）
 茨城大学・工学部・准教授
 研究者番号： 70408649

研究成果の概要（和文）：多変量システムを効果的に予測するために、まず観測データのみを用いて要素間の相互作用およびダイナミクスを同定し、次にその知見を活かして最適な予測モデルを構築した。その適用例として経済市場に注目し、為替取引価格の変動を予測した。さらに実システムは構造が動的に変化する可能性があるため、時間経過に伴い予測モデルの再構築を繰り返す効果も検証した。特にこの効果を確認することで、システムが動的に構造変化するか否かを判定できる可能性を示した。

研究成果の概要（英文）：Although we can often observe time-series data of many elements, these elements do not always interact with each other. This study proposes a scheme to estimate the interdependency among observed elements only by time-series data, which is useful for selecting essential elements to optimize multivariate prediction model. Because this estimation is a sort of combinatorial optimization problems, we applied the genetic algorithm as a method to moderate this problem. Through some simulations, we confirmed performance of our method, which can identify interaction of multivariate system and can improve its prediction accuracy. Especially, our method can be applied to predict real foreign-exchange markets even if system has nonstational property and its structure changes dynamically.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2008年度	1,800,000	540,000	2,340,000
2009年度	1,200,000	360,000	1,560,000
年度			
年度			
年度			
総計	3,000,000	900,000	3,900,000

研究分野：総合領域

科研費の分科・細目：情報学・感性情報学・ソフトコンピューティング

キーワード：複雑系, 複雑ネットワーク, カオス, 経済物理学

1. 研究開始当初の背景

自然界には多数の要素が相互作用し、複雑な現象を生み出す力学系が存在する。特に、経済システムや脳などの生体システムに関する研究は従来から注目されている。近年においては観測技術の向上により、関与する要素の振る舞いを同時に観測する事が可能になってきた。さらに計算機の性能向上に伴い、膨大な観測データを解析することで複雑システムの仕組みを議論する試みが活発化している。

大規模な要素で構成される非線形システムを理解するためには、ネットワークの概念が有効である。各要素に機能するダイナミクスが分かったとしても、全体のネットワーク構造(各要素の関わり合い方)によっては系全体の振る舞いが大きく変わってくる。近年、ネットワーク構造を特徴づける指標として、各ノード間の平均経路長やクラスター係数などが有効であり、これらを駆使して、ソーシャルネットワークやスケールフリーネットワークなどといった実社会で頻繁に見られるネットワークを特徴づけることが可能となった。現在では、この分野は複雑ネットワーク科学と認知され、理系文系の垣根を越えた幅広い研究テーマに応用され、多くの新しい知見を生み出している。

2. 研究の目的

自然界の多くのシステムは、複数の要素が複雑に相互作用する多変量システムである。システム全体の振舞いは複雑になり、その結果、予測や制御が困難になる。本研究では、この多変量システムを効果的に予測する手法を議論する。

もしシステムを構成する1つの要素*i*を予測対象とする場合、観測できる全要素の振舞いを予測のための情報として利用することができる。しかし、どの要素が要素*i*に影響を及ぼしているのか同定することは難しい。もし要素*i*と関連しない要素までも予測に利用してしまうと、予測モデルは複雑になり、情報量基準の観点や予測精度の観点から適切であるとは言えない。つまり、要素*i*と直接的に相互作用する要素*ij*の振舞いのみを予測の手がかりにする必要があり、そのためには、全変数から要素*ij*を同定する方法が必要となる。このようにシステム全体の因果構造を同定できれば、予測に限らず、システム全体の理解に役立つ。

このような動機から、各要素の振舞いの相関係数や偏相関係数などを見積り、類似性をもとに因果の有無を判別する方法がある。しかし自然界の多くのシステムは非線形性を有するので、常に線形統計量で因果推定を行えるとは考え難い。そこで本研究では、遺伝的アルゴリズム(以下GA)を用いて、要素*i*を最適に予測できるように使用する情報を厳選する。さらに、この予測精度の最適化によって、要素*i*と相互作用する変数*ij*を特定する。

この手法の有用性を検証すべく、複雑系を模擬する数理モデルを用いてシミュレーションを行い、因果構造の同定精度を評価する。さらに、同定された因果構造を踏まえて予測をする場合と単純に観測しうる全変数を用いて予測をする場合の予測精度を比較する。これによって、予測に用いる情報を厳選する効果を検証する。最後に、実際の多変量システムとして実際の為替取引市場を予測対象とし、本手法の応用について検討する。また実際のシステムは、構造が動的に変化する可能性があるため、予測モデルを動的に最適化し、その有用性の検証を行う。

3. 研究の方法

多変量システムを構成する*i*番目の要素の振舞いを $x_i(t)$ とすると、システム全体の振舞いは以下のように記述される。

本研究では、 $V(t)$ の将来変動を予測するために、近傍の振舞いを参考にすることで、局所線形近似的に非線形予測を行う。そのため

$$V(t) = \{x_1(t), x_2(t), \dots, x_i(t), \dots, x_N(t)\} \quad (1)$$

$V(t)$ の近傍点 $V(kn)$ を $t-L \leq kn < t$ より探し、その1ステップ後の振舞いを平均化することで、 $V(t+1)$ の予測値(次式)を得る。

しかし、一般的に式(1)を適切に構成することは容易ではない。観測した変数が、システムに内在する変数であるのかは不明であるし、

$$\tilde{V}(t+1) = \langle V(k_n + 1) \rangle_n \quad (2)$$

もし内在する変数だとしても予測対象とする変数と直接的に関連しているとは限らない。例えば、変数*i*が変数*i1*, *i2*, *i3*の影響を受けていれば、変数間の因果構造は次式のように書ける。

$$x_i(t+1) = F(x_i(t), x_{i_1}(t), x_{i_2}(t), x_{i_3}(t)). \quad (3)$$

ここで、 F は関数、 i は目的変数、 $i=1\sim 3$ は説明変数である。この場合は、式(1)を

$$V(t) = \{x_i(t), x_{i_1}(t), x_{i_2}(t), x_{i_3}(t)\} \quad (4)$$

に修正し、式(2)を適用する方が優れた予測精度を期待できる。

本研究では、このような因果構造を観測データのみから推定するために、GAを用いる。GAで用いる遺伝子型を $g_i = \{11001 \dots\}$ のように表現し、 $g_i(j) = 0$ であれば変数 i の予測の際に、式(4)の $V(t)$ として変数 j を用いないことを意味し、 $g_i(j) = 1$ であれば変数 j を用いることを意味している。さらに、各遺伝子型の適合度を求めるために、各 $V(t)$ を用いて、過去に観測されたデータを予測することで予測精度を算出し、その予測精度を適合度とした。予測精度の算出においては、真値と予測値との相関係数を用いた。また本研究では、遺伝子型数を30、世代数を100、交叉確率を0.9、突然変異確率を0.1、エリート個体率を0.1に設定し、GAのアルゴリズムによって各遺伝子型を進化させた。

この手法の有用性を検証するために、次式の数理モデルを用いてシミュレーションを行った。

$$x_i(t+1) = F \left((1-\epsilon)x_i(t) + \frac{\epsilon}{N_i} \sum_{j \in \{i_j(p)\}} x_j(t) \right) \quad (5)$$

これは複雑系を模擬したカオス結合系として広く研究されている数理モデルである。 N_i は要素 i と結合している要素数、 ϵ はその結合強度を表す。ただし各要素の結合パターンはレギュラーグラフだけでなく、WSモデルを導入することで、スモールワールドネットワークやランダムネットワークに変更可能にした。WSモデルにおけるノードのランダム再結合変数を p とすると、 $\{i_j(p)\}$ は要素 i と結合する変数 j の集合を表している。また関数 F として、広く研究されているロジスティック写像 $F(x) = 1-ax^2$ を採用した。

シミュレーションでは、まず各要素の時系列データ $x(t)$ を元に、カオス結合系の因果構造の同定を行う。もしGAで最適化された遺伝子型が $g^* = \{11001 \dots\}$ で、その適合度(予測精度)が e である時、要素 j から要素 i への因果を表す推定行列 $M_{i \leftarrow j}$ の i 行を $\epsilon \times g_i^*$ で置き換える。以上を全ての要素 ($i = 1 \sim N$) について行えば、 $i \leftarrow j$ の因果を有向かつ重み付きで推定できる。また、 $M_{i,j} = (M_{i \leftarrow j} + M_{j \leftarrow i})/2$ とすることで双方向の因果推定に対応できる。本研究では、式(5)の数

理モデルを解析対象とするので双方向の因果を推定した。推定精度の算出には、次式を用いた。

$$E = \frac{|\{\tilde{i}_j\} \cup \{i_j\}|}{|\{\tilde{i}_j\}|} \quad (6)$$

ここで、 $\{i_j\}$ は正解の因果構造に対応し、 $\{\tilde{i}_j\}$ は推定された因果構造を意味している。ただし、 $|\{\tilde{i}_j\}| \neq |\{i_j\}|$ となるように、 $M_{i,j}$ 中の上位の要素より $\{\tilde{i}_j\}$ を決定した。なお先攻研究では、相関係数、偏相関係数、相互情報量、移動エントロピーを用いて各要素間の振舞いの類似度を評価することで $M_{i,j}$ を算出し、因果構造の同定を行っている。本研究では、式(5)中の p と ϵ を変化させながら、これらの手法とGAによる手法の比較を行った。

図1にシミュレーション結果を示す。ただし、数理モデルにおける要素数 N を30、非線形予測に用いる学習データ長 L を100に設定した場合である。これらのパラメータを可変させた場合においても同様の傾向が得られた。傾向としては、数理モデルのパラメータである結合強度 ϵ と結合のグラフ構造 p が変わっても、概ねGAを用いた因果構造推定法が最良であることが分かる。さらに図1(c)に示すように、たとえGAを用いた推定法が最良でない場合でも、最良手法との推定誤差は僅かであるので、因果構造の同定にとってGAは有用であると言える。

次に、GAで同定した因果構造によって式(4)の $V(t)$ を構成して予測した場合と、単純に全変数を用いた $V(t)$ で予測した場合の予測精度を比較する。なお、前者の予測精度を G_i 、後者の予測精度を A_i とする。図2(a)(b)によれば、GAを用いて予測に用いる情報を厳選した方が、予測精度が向上することが分かる。さらに、 $\{G_i\} = \{A_i\}$ の帰無仮説に対してウィルコクソン符号付順位検定を行ったところ、有意水準 $\alpha = 1[\%]$ で帰無仮説を棄却でき、 $\{G_i\} > \{A_i\}$ を積極的に主張することができる(図2(c))。

実際の多変量システムとして、1996年の外国為替市場における25種の取引価格を予測する。各取引価格は30分毎に記録され、時系列データを構成している。さらに比較対象として、式(5)の数理モデルも予測する。ただし本節では、実システムは一般的に予測が難しいことから、1ステップ後の変動が上昇するか、または下降するかを予測した。つまり $x'_i(t+1) - x_i(t) > 0$ であれば上昇、 $x'_i(t+1) - x_i(t) < 0$ であれば下降と予測し、その的中率を予測精度 P_i とした。また、GAを用いず全変数を用

いて予測する方法を“方法1”とし、GAを用いて予測モデルを最適化する方法を“方法2”とする。ただし、最適化は最初の1回のみ行われ、最適化された予測モデルは時間が経過しても継続使用される。

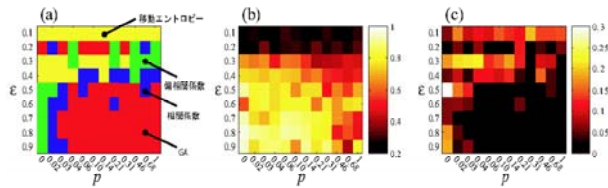


図 1: (a) 最良の同定精度 E が得られた推定手法. (b) 図 (a) で示した最良の方法で得られた同定精度 E . (c) 図 (b) で示した同定精度と GA を用いた場合の同定精度の差.

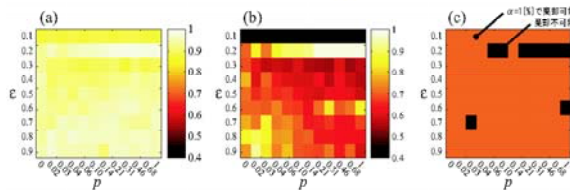


図 2: (a) GA で同定した因果構造で $V(t)$ を構成し、予測した時の予測精度 $\langle G_i \rangle_i$. (b) 全変数を用いて $V(t)$ を構成し、予測した時の予測精度 $\langle A_i \rangle_i$. (c) $\{G_i\} > \{A_i\}$ に関する有意差検定.

しかし実際のシステムを対象とする場合、外部の影響によってシステムの構造が変化する可能性を考慮する必要がある。つまり、予測開始時刻が変わる毎に因果構造を推定し直し、再度予測モデルを最適化する必要がある。この動的最適化による予測法を“予測法3”とし、有用性の評価を行った。

予測結果を表1, 2に示す。 $\langle P_i \rangle_i$ は各要素の予測精度の平均値、 R は予測法1に対して予測精度が向上した要素の割合を示している。さらに前節と同様にウィルコクソンの符号付順位検定を行った。

結果として、予測法1に対するGAによる最適化の有用性(予測法2)または動的最適化の有用性(予測法3)を確認できる。しかし、外国為替市場においては予測法2の有用性を確認できない。これは外国為替市場の構造が動的変化している事実を示唆している。よって、実システムを対象とする場合は、予測法3のような動的最適化法が効果的である。

表 1: 外国為替取引市場の予測結果.

	$\langle P_i \rangle_i$	R	予測法 1 との有意差検定
予測法 1	56[%]		
予測法 2	56[%]	44[%]	棄却不可
予測法 3	60[%]	96[%]	棄却可能 ($\alpha = 1$ [%])

表 2: 数理モデル (式 (5)) の予測結果.

	$\langle P_i \rangle_i$	R	予測法 1 との有意差検定
予測法 1	84[%]		
予測法 2	88[%]	73[%]	棄却可能 ($\alpha = 1$ [%])
予測法 3	92[%]	80[%]	棄却可能 ($\alpha = 1$ [%])

4. 研究成果

本研究では、多変量システムの予測において、予測に用いる情報を厳選する手法を議論した。その選択基準として、過去のデータに対する予測精度を最良にする予測モデルを見つける必要があるが、これは組合せ最適化問題となるため、その解法として本研究ではGAを採用した。さらに計算機実験を通じて以下を示した。

- 最適化された予測モデルを参照することで、システムの因果構造を同定できる。
- 最適化された予測モデルを用いれば、新規データに対しても精度良く予測できる。
- 実システムのように構造が動的に変化する場合、予測モデルを毎回最適化することで予測精度を向上できる。
- 予測モデルの動的最適化により予測精度が向上すれば、それはシステムが動的に構造変化している証拠となる。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 4 件)

- ①. Tomoya Suzuki, Yuta Ueoka, Haruki Sato, Estimating Structure of Multivariate Systems with Genetic Algorithms for Nonlinear Prediction, Physical Review E, 査読有, Vol. 80, No. 6, 066208/1-066208/6, 2009.
- ②. 鈴木智也, 池田真一, 複雑システムにおけるネットワーク中心性が予測精度に与える影響, 情報処理学会論文誌 数理モデル化と応用, 査読有, Vol. 2, No. 2, pp. 1-9, 2009.
- ③. 鈴木智也, 太田真喜, 時系列データの天底予測のための非線形予測法, 情報処理学会論文誌 数理モデル化と応用, 査読有, Vol. 2, No. 1, pp. 123-132, 2009.
- ④. 鈴木智也, 情報伝達に基づいた有向重み付き複雑ネットワーク解析, 情報処理学会論文誌 数理モデル化と応用, 査読有, Vol. 2, No. 1, pp. 70-79, 2009.

[学会発表] (計 17 件)

- ①. 鈴木智也, 等時間間隔サンプリングによって欠落するシステムの非線形性, 電子情報通信学会 2009 年総合大会, 宮城, 2010 年 3

- 月 16 日.
- ②. 鈴木智也, 等時間間隔サンプリングによって見失う非線形システムの特徴, 第 72 回情報処理学会全国大会, 東京, 2010 年 3 月 10 日.
 - ③. 鈴木智也, 複雑システムの理解と予測のための観測時系列データの最適利用, 第 52 回自動制御連合講演会, 大阪, 2009 年 11 月 21 日.
 - ④. 鈴木智也, 佐藤春樹, 金子彰吾, 非線形経済予測モデルの動的最適化, 日本物理学会 2009 年次大会, 東京, 2009 年 3 月 27 日.
 - ⑤. 三井貴視, 瀬木宏, 鈴木智也, 非線形時系列解析における欠損データが及ぼす影響, 電子情報通信学会 2009 年総合大会, 愛媛, 2009 年 3 月 17 日.
 - ⑥. 鈴木智也, 上岡祐太, 佐藤春樹, 金子彰吾, 多変量予測モデル構築における組合せ最適化問題, 電子情報通信学会 2009 年総合大会, 愛媛, 2009 年 3 月 17 日.
 - ⑦. 鈴木智也, 上岡祐太, 佐藤春樹, 進化的計算手法を用いた多変量システムの因果推定と予測問題への応用, 情報処理学会 2009 年全国大会, 滋賀, 2009 年 3 月 11 日.
 - ⑧. 上岡祐太, 鈴木智也, 山本誠一, 部分的ネットワーク情報を利用した大域的ネットワーク構造の推定, 情報処理学会 2009 年全国大会, 滋賀, 2009 年 3 月 11 日.
 - ⑨. 佐藤春樹, 鈴木智也, 進化的計算手法を用いた多変量予測モデルの動的最適化, 情報処理学会 2009 年全国大会, 滋賀, 2009 年 3 月 10 日.
 - ⑩. 上野佑輔, 鈴木智也, ブートストラップ法を用いた少数データに対する局所線形近似法, 情報処理学会 2009 年全国大会, 滋賀, 2009 年 3 月 10 日.
 - ⑪. 瀬木宏, 三井貴視, 鈴木智也, 欠損を含む時系列データの非線形予測, 情報処理学会 2009 年全国大会, 滋賀, 2009 年 3 月 10 日.
 - ⑫. Shougo Kaneko, Tomoya Suzuki, Dynamical Optimization for Nonlinear Prediction, International Symposium on Nonlinear Circuits and Signal Processing, USA, March 3, 2009.
 - ⑬. Yuta Ueoka, Tomoya Suzuki, Seiichi Yamamoto, Available

- Partial Information to Estimate the Whole Structure of Complex Systems, International Symposium on Nonlinear Circuits and Signal Processing, USA, March 3, 2009.
- ⑭. Tomoya Suzuki, Yuta Ueoka, Haruki Sato, Combinatorial Optimization for Multivariate Nonlinear Prediction, International Symposium on Nonlinear Circuits and Signal Processing, USA, March 1, 2009.
 - ⑮. Yuta Ueoka, Tomoya Suzuki, Tohru Ikeguchi, Yoshihiko Horio, Efficiency of Statistical Measures to Estimate Network Structure of Chaos Coupled Systems, 2008 International Symposium on Nonlinear Theory and its Applications, Hungary, September 8, 2008.
 - ⑯. Tomoya Suzuki, Characterizing Cluster Coefficient in Directed and Weighted Complex Networks on the Basis of Information Flow, 2008 International IEEE Workshop on Nonlinear Dynamics of Electronic Systems, Russia, July 24, 2008.
 - ⑰. Yuta Ueoka, Tomoya Suzuki, Tohru Ikeguchi, Yoshihiko Horio, Estimating Network Structure of Chaos Coupled Systems, 2008 International IEEE Workshop on Nonlinear Dynamics of Electronic Systems, Russia, July 21, 2008.

[その他]

ホームページ等

http://tsuzuki.ise.ibaraki.ac.jp/TS_lab/

6. 研究組織

(1) 研究代表者

鈴木 智也 (SUZUKI TOMOYA)

茨城大学・工学部・准教授

研究者番号：70408649