

遺伝的アルゴリズムを用いた基底関数構築と 非線形時系列予測

塚田 将太^{1,a)} 佐藤 仁樹^{2,b)}

受付日 2013年3月4日, 採録日 2013年9月13日

概要: 時系列データの相関に基づき遺伝的アルゴリズムにより抽出された基底関数を用いて, 非線形時系列の予測精度を改善した. まず, 遺伝的アルゴリズムを用いて, 時系列データと相関が高い基底関数を抽出した. 次に, 時系列データを基底関数により変換し, 重回帰分析を用いて予測を行った. その結果, 予測に必要な説明変数の種類や数を特定できない場合でも, 目的変数に関する可能性がある観測データをできるだけ多く説明変数として用いることにより, 従来手法と比べて高い予測精度が得られた.

キーワード: 非線形, 予測, 関数近似, 重回帰分析, 遺伝的アルゴリズム

Basis Function Construction with Genetic Algorithm for Nonlinear Time Series Prediction

SHOUTA TSUKADA^{1,a)} HIDEKI SATOH^{2,b)}

Received: March 4, 2013, Accepted: September 13, 2013

Abstract: The accuracy of nonlinear time series prediction has been improved using basis functions extracted by a genetic algorithm based on the correlation in the time series data. First, basis functions related to the time series data were extracted using the genetic algorithm. Next, the time series data was transformed using the basis functions, and the target time series was predicted using the multiple regression analysis. As a result, we obtained a higher accuracy compared with conventional methods using as many predictor variables that may be related to the criterion variable as possible, even if the predictor variables that are necessary to predict the criterion variable cannot be identified.

Keywords: nonlinear, prediction, function approximation, multiple regression analysis, genetic algorithm

1. はじめに

我々の生活において, 予測は非常に重要な役割を持っている. 身近な例では, 天気, 地震, 台風の経路などの自然現象の予測があげられる. また, 電力の需要予測は, 安定な電力供給にとって非常に重要である. このように, 我々の生活は様々な現象の予測によって支えられている. しか

し, 現実の予測問題は, 予測対象が複雑な挙動を示す非線形の予測問題であり, 自己回帰 (AR) モデルに基づく線形予測 [1] では, 高い精度が得られなかった. そのため, 非線形な現象を予測するための様々な手法が提案されてきた.

生物の神経回路系をモデル化したニューラルネットワークは, モデル化困難な対象に対して有効に機能し, 汎化性能が優れている. そのため, 様々な対象に適用されている [2]. RBF (radial basis function) [3] を用いたニューラルネットワークによる予測手法では, 入力データの空間的特徴が RBF の超楕円性により効率良く抽出される [4]. また, ウェーブレット変換を用いた予測では, 時系列を時間域に局在した周波数成分の和の形に分解する. その結果, 周波数の情報を時刻ごとに解析できるため, ホワイトノイ

¹ 日立アイ・エヌ・エス・ソフトウェア株式会社
Hitachi INS Software, Ltd., Yokohama, Kanagawa 220-0011, Japan

² 公立はこだて未来大学
Future University Hakodate, Hakodate, Hokkaido 041-8655, Japan

^{a)} s-tsukada@ins-hitachi.co.jp

^{b)} jamisato@m.ieice.org

ズや低周波のノイズが重畳した時系列の予測に有効である [5]. サポートベクタマシンを回帰問題に適用したサポートベクタ回帰 (SVR) は, 問題に適したパラメータを用いることにより, 非線形の問題に対して非常に高い汎化能力を発揮する. そのため, パターン探索法 [6] や正則化 [7] など様々な手法によって, SVR のパラメータ推定が行われている [8].

これらの手法より, 予測対象となる目的変数とその変化を決める説明変数の関係が複雑な非線形性を持つ場合でも, 高い精度で予測が可能となった. しかし, 現実の予測問題では, 目的変数の変化を決めるメカニズムが非常に複雑であり, 予測に必要な説明変数の種類や数を必ずしも特定できるとは限らない. これは, 目的変数と説明変数の関係が一般に非線形であるため, 目的変数と観測データの線形な相関関数では目的変数の値を決定する説明変数を抽出できないためである. さらに, 目的変数と説明変数の非線形性を考慮した場合, 説明変数の候補となる観測データの種類に対して, 評価すべき非線形関数の組合せが指数関数的に増加する. そのため, すべての観測データの非線形な組合せと目的変数の相関を評価できない.

そのため, 予測精度を上げるためには, 目的変数に関係する可能性のある観測データをできるだけ多く説明変数として用いることが望ましい. しかし, 状態変数の次元に対して必要なニューロンや基底の数が指数関数的に増加するという問題 (次元の呪い) が発生する [3]. さらに, 説明変数として用いられる観測データが目的変数と独立な場合, それらが予測精度に悪影響を与えるノイズとして働く場合がある. このように, 従来は, 真に必要な説明変数をシステマチックに選択できないだけでなく, 十分な数の説明変数を使えなかった.

これらの問題を解決するために, 我々は遺伝的アルゴリズム (GA) [9] を用いて, 時系列データと相関が高い基底関数を抽出し, 重回帰分析により予測を行う手法を開発した. 本論文では, 提案したアルゴリズムを用いて, 様々な時系列の予測を行い, 本手法の有効性を示す.

2. 予測におけるスパース学習

d_x 次元のベクトル $\mathbf{x}_t = (x_{1;t}, x_{2;t}, \dots, x_{d_x;t})^T$ を未知の関数 $f(\mathbf{x}_t)$ の入力, y_{t+p} をその出力, t を時刻, p を正の整数とする. \mathbf{x}_t を説明変数, y_{t+p} を目的変数として, 学習データ $(y_{1+p}, \mathbf{x}_1), (y_{2+p}, \mathbf{x}_2), \dots$ から $f(\mathbf{x}_t)$ の近似関数 $\tilde{f}(\mathbf{x}_t)$ を構築する問題を関数近似問題という. また, $\tilde{f}(\mathbf{x}_t)$ を構築し, \mathbf{x}_t から y_{t+p} の近似値 \tilde{y}_{t+p} を推定する問題を予測問題という.

$f(\mathbf{x}_t)$ の近似関数 $\tilde{f}(\mathbf{x}_t)$ は, 次式で表される.

$$\tilde{y}_{t+p} = \tilde{f}(\mathbf{x}_t)$$

$$= \sum_{i=0}^N \alpha_i \phi_i(\mathbf{x}_t) \quad (1)$$

ここで, N は展開次数, $\{\phi_i(\mathbf{x}_t)\}$ は基底関数, $\{\alpha_i\}$ は学習データから求められる係数である. $\{\phi_i(\mathbf{x}_t)\}$ が正規直交基底 (付録 A.1 参照) の場合, 式 (1) は Fourier 級数展開であり, $\{\alpha_i\}$ は Fourier 係数となる. この場合, $\{\alpha_i\}$ は, 重回帰分析 (multiple regression analysis (MRA)) [10] により得られる.

学習データを $\{(y_{i+p}, \mathbf{x}_i)\}$, $k(\mathbf{x}, \mathbf{x}_i)$ をカーネル関数とする. $\phi_i(\mathbf{x}_t)$ がカーネル関数, すなわち, $\phi_i(\mathbf{x}_t) = k(\mathbf{x}_t, \mathbf{x}_i)$ の場合, 式 (1) の係数 $\{\alpha_i\}$ を学習データから決める代表的な手法として, サポートベクタマシン (SVM) があげられる [11]. SVM はパターン認識や予測の分野において, 最も優れた手法の 1 つであり, 線形問題のみでなく, 非線形問題に対する能力も高い. また, 式 (1) における N は学習データの数に対応し, 説明変数 \mathbf{x}_t が高次元の場合でも高い性能が得られる. SVM では, 係数 $\{\alpha_i\}$ に対する正則化項として l_2 ノルムを用いることにより, 汎化能力を上げている. また, l_1 ノルムを正則化項として用いることにより, 汎化能力を維持したまま, スパースな解 $\{\alpha_i\}$ を得ることができる [12].

一般に, スパースな解を得ることにより, 問題に潜む構造を明らかにできる. 目的変数と説明変数の関係が線形の場合には, 高次元の説明変数に対してスパースな解を得る方法が提案された [13]. これにより, たとえば, 目的変数が d 番目と d' 番目の説明変数に大きく影響を受けるといような構造が陽に得られるため, d 番目と d' 番目の説明変数の観測態勢を強化することにより, 予測精度を改善できる. また, 必要に応じて, d 番目と d' 番目の説明変数を制御することにより, 目的変数を望ましい方向に導ける. このように, 目的変数と説明変数の関係を明らかにすることにより, 予測精度を改善するだけでなく, 予測問題に関連する様々な問題の解決に貢献できる.

しかし, 予測問題では, 目的変数と説明変数が線形の関係とは限らない. そのため, たとえば, 目的変数が d 番目の説明変数の 2 乗に反比例し, d' 番目の説明変数に比例するといった目的変数と説明変数の非線形な関係を明らかにすることが望まれる. 基底関数 $\{\phi_i(\mathbf{x}_t)\}$ をカーネル関数とした場合, 目的変数と説明変数の陽な関係が得られるとは限らない. 一方, 基底関数 $\{\phi_i(\mathbf{x}_t)\}$ が正規直交基底の場合, 学習データからスパースな解 $\{\alpha_i\}$ が得られれば, 目的変数と説明変数の非線形な関係が陽に得られる. 特に, 基底関数が Legendre 関数 (付録 A.1 参照) の場合, 目的変数を説明変数の冪乗で近似できるため, 目的変数と説明変数の高次の関係が明確になる.

しかし, 通常の Fourier 級数の正規直交基底 (付録 A.1 参照) を用いた場合, 説明変数 \mathbf{x}_t の次元 d_x の増加に従い,

必要となる基底および Fourier 係数 α_i の数が指数関数的に増加する．そのため，高次元の説明変数 \mathbf{x}_t を用いた場合，Fourier 係数 $\{\alpha_i\}$ の最適化が困難となる．たとえば，十分な精度を得るために x_d に対する展開次数 N_d ^{*1} を 32 とすると， $d_x = 4$ の場合， N は式 (A.10) より 1,185,920 になる．そのため， $d_x \geq 4$ の対象を，式 (1) でモデル化し，その係数 $\{\alpha_i\}$ を最適化することは困難である．

この問題を解決するために，GA のエリート戦略に基づく染色体の選択手法を用いて正規直交基底を構築し，関数近似を行う手法 [14] が提案された．本論文では，この手法に一点交叉と突然変異の機能を付加し，予測問題に適用する．

3. 遺伝的アルゴリズムを用いた基底関数構築と予測

遺伝的アルゴリズム (GA) は，最適化問題を解決するための進化的アルゴリズムの 1 つである [9]． i 番目の個体は染色体に対応するベクトル \mathbf{g}_i を持ち，個体の優劣は \mathbf{g}_i から計算される適応度 J_i で表される．ここで，染色体 \mathbf{g}_i の d 番目の遺伝子を $g_{i;d}$ とすると，本論文では， g_{\max} を $g_{i;d}$ の値の最大値， $g_{i;d}$ を $g_{i;d} \in \{0, 1, \dots, g_{\max}\}$ で定義した．GA は，優秀な個体を得るために，選択，交叉，突然変異を繰り返すことにより，染色体 \mathbf{g}_i を改良するアルゴリズムである．GA は解空間を大域的に探索できるため，局所解に陥りにくいという特徴がある．

本論文では，エリート戦略に従い染色体を選択した．個体数を b_{\max} とし，個体の適応度が高い順に \mathbf{g}_i をソートする．その後， $\mathbf{g}_1, \dots, \mathbf{g}_N$ の個体を残し， $\mathbf{g}_{N+1}, \dots, \mathbf{g}_{b_{\max}}$ の個体を廃棄する．ここで， N は 2 章で定義された展開時数であり， b_{\max} は $N < b_{\max}$ のように設定される．次に， $\mathbf{g}_1, \dots, \mathbf{g}_N$ からランダムに 2 つのベクトル $\mathbf{g}_i, \mathbf{g}_j$ を選び， \mathbf{g}_i と \mathbf{g}_j を交叉させる．交叉する点はランダムに決められる．上記の交叉により生成されたベクトルが， $\mathbf{g}_1, \dots, \mathbf{g}_N$ のいずれかと等しい場合，交叉により生成されたベクトルを乱数により生成されたベクトルに置き換える．その後，確率 p_M で遺伝子は突然変異を起こす．突然変異を起こした遺伝子の値は， $0, 1, 2, \dots, g_{\max}$ の一様乱数で決められる．

基底関数 $\phi_i(\mathbf{x}_t)$ のインデックスベクトルを \mathbf{g}_i とする (付録 A.1 参照)．GA により改良された染色体 \mathbf{g}_i を $\phi_i(\mathbf{x}_t)$ のインデックスベクトルとすることにより，式 (1) で表された近似関数の精度を改良する． \mathbf{g}_i の d 番目の要素 $g_{i;d}$ は， $x_{d;t}$ の基底関数 $K_d(x_{d;t}, g_{i;d})$ を式 (A.4) または式 (A.5) に従って規定する． $\phi_i(\mathbf{x}_t)$ は， $K_d(x_{d;t}, g_{i;d})$ の積として，式 (A.3) および式 (A.7) により得られる．このように，インデックスベクトルは，基底関数を指定するために使用され，

^{*1} N_d は x_d に対する展開次数であり， x_d に対する精度を決める． N_d の値が大きいほど x_d の精度は高くなる．詳細は付録 A.1 を参照．

予測精度に大きく影響する．そのため，目的変数 y_{t+p} と相関が高い基底関数を与えるインデックスベクトルを，無数にあるインデックスの組合せから抽出できれば，予測精度を改善できる．本論文では，適応度 J_i を次式により定義し，この問題を GA を用いて解決する．

$$J_i = |E[y_{t+p}\phi_i(\mathbf{x}_t)] - E[y_{t+p}]E[\phi_i(\mathbf{x}_t)]| \quad (2)$$

ここで， $E[\cdot]$ は平均を表す．式 (2) は，目的変数 y_{t+p} とベクトル \mathbf{g}_i によって定められる基底 $\phi_i(\mathbf{x}_t)$ との間の共分散の絶対値であり， J_i の値が大きいほど $\phi_i(\mathbf{x}_t)$ が y_{t+p} に与える影響が大きくなる，すなわち，染色体 \mathbf{g}_i が優れていることを表す．したがって，GA により，目的関数 y_{t+p} との相関が高い基底関数を構築できる．そのアルゴリズムを Algorithm 1 に示す．ここで， g_{\max} は x_d ($d = 1, 2, \dots, d_x$) に対する基底関数の最大次数である． g_{\max} が大きいほど複雑な基底関数を構築できるため，近似精度が高くなる．しかし，探索空間が広がるため GA の収束が遅くなる．また，式 (1) から分かるように，展開時数 N の増加に従い，計算量が増加する一方で近似精度が高くなる．本論文では $b_{\max} = 2N$ とした．

Algorithm 1 : GA を用いた基底関数構築

- Step (1-1) パラメータ (N, p_M, g_{\max} , 基底関数) を与える．
- Step (1-2) $\mathbf{g}_1, \dots, \mathbf{g}_{b_{\max}}$ の要素に 0 から g_{\max} までの整数を乱数で与え，初期化する．
- Step (1-3) $\mathbf{g}_1, \dots, \mathbf{g}_{b_{\max}}$ から適応度 $J_1, \dots, J_{b_{\max}}$ を式 (2) より計算する．
- Step (1-4) 適応度の値により， $J_1, \dots, J_{b_{\max}}$ および $\mathbf{g}_1, \dots, \mathbf{g}_{b_{\max}}$ をソートする．
($J_i > J_{i+1}$)
- Step (1-5) $\{J_i\}$ の値が収束すれば終了．収束していない場合，Step (1-6) へ．
- Step (1-6) $\mathbf{g}_1, \dots, \mathbf{g}_N$ を残し， $\mathbf{g}_{N+1}, \dots, \mathbf{g}_{b_{\max}}$ を廃棄する．
- Step (1-7) $n = N$ とする．
- Step (1-8) $\mathbf{g}_1, \dots, \mathbf{g}_N$ から，2 つのベクトル $\mathbf{g}_i, \mathbf{g}_j$ をランダムに選ぶ．
- Step (1-9) \mathbf{g}_i と \mathbf{g}_j を交叉させ， $\mathbf{g}_{n+1}, \mathbf{g}_{n+2}$ を作成する．交叉ポイントは 2 から $d_x - 1$ までの乱数とする．
- Step (1-10) 以下の条件を満たさない場合，以下の条件を満たすように， \mathbf{g}_{n+1} および \mathbf{g}_{n+2} を乱数で生成する．

$$\begin{pmatrix} \mathbf{g}_{n+1} \notin \{\mathbf{g}_1, \dots, \mathbf{g}_N\} \\ \mathbf{g}_{n+2} \notin \{\mathbf{g}_1, \dots, \mathbf{g}_N\} \\ \mathbf{g}_{n+1} \neq \mathbf{g}_{n+2} \end{pmatrix}$$

Step (1-11) $n + 2 \geq b_{\max}$ ならば，Step (1-12) へ．

$n + 2 < b_{\max}$ ならば, $n = n + 2$ として Step (1-8) へ.

Step (1-12) \mathbf{g}_i ($N + 1 \leq i \leq b_{\max}$) の各要素を確率 p_M の突然変異により変更する.

Step (1-13) \mathbf{g}_i ($N + 1 \leq i \leq b_{\max}$) $\in \{\mathbf{g}_1, \dots, \mathbf{g}_N\}$ の場合, $\mathbf{g}_i \notin \{\mathbf{g}_1, \dots, \mathbf{g}_N\}$ となるように \mathbf{g}_i を乱数で生成する.

Step (1-14) Step (1-3) へ戻る.

上記 GA に基づく Algorithm 1 および重回帰分析 (MRA) を用いて予測を行う手順を Algorithm 2 に示す (以下では MRAGA と略す). ここで, 学習データとは目的変数 y_{t+p} および説明変数 \mathbf{x}_t の時系列であり, インデックスベクトル $\{\mathbf{g}_i\}$ および Fourier 係数 $\{\alpha_i\}$ を推定するために用いられる. 学習データにより推定された $\{\mathbf{g}_i\}$ と $\{\alpha_i\}$ を用いて, 学習データとは異なる評価データに対して次式で定義される予測誤差 ε を評価する.

$$\varepsilon = E[(y_{t+p} - \tilde{y}_{t+p})^2] \quad (3)$$

Algorithm 2 : MRAGA による予測

Step (2-1) 学習データから Algorithm 1 により $\{\mathbf{g}_i\}$ を選択する.

Step (2-2) 学習データ $\{y_{t+p}, \mathbf{x}_t\}$ を $\{y_{t+p}, \phi(\mathbf{x}_t)\}$ に変換し, 次式により Fourier 係数 $\{\alpha_i\}$ を推定する.

$$\begin{cases} \boldsymbol{\alpha} = COV_{\phi\phi}^{-1} \mathbf{cov}_{y\phi} \\ \alpha_0 = E[y_{t+p}] - \boldsymbol{\alpha}^T E[\phi(\mathbf{x}_t)] \end{cases} \quad (4)$$

Step (2-3) Step (2-1) で選択された $\{\mathbf{g}_i\}$ を用いて, 評価データ $\{\mathbf{x}_t\}$ を $\{\phi(\mathbf{x}_t)\}$ に変換する.

Step (2-4) Step (2-2) で推定された $\{\alpha_i\}$ と, Step (2-3) で変換された評価データ $\{\phi(\mathbf{x}_t)\}$ から, 予測値 \tilde{y}_{t+p} を計算する.

ここで, 式 (4) は重回帰分析であり, $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_N)^T$, $\phi(\mathbf{x}_t) = (\phi_1(\mathbf{x}_t), \dots, \phi_N(\mathbf{x}_t))^T$, $E[y_{t+p}]$ は $\{y_{t+p}\}$ の平均, $E[\phi(\mathbf{x}_t)]$ は $\{\phi(\mathbf{x}_t)\}$ の平均, $\mathbf{cov}_{y\phi}$ は $\{y_{t+p}\}$ と $\{\phi_i(\mathbf{x}_t)\}$ の共分散ベクトル, $COV_{\phi\phi}$ は $\{\phi_i(\mathbf{x}_t)\}$ と $\{\phi_j(\mathbf{x}_t)\}$ の共分散行列である.

4. 性能評価

4.1 GA を用いた基底関数構築の評価

本節では, 式 (5) で生成される時系列を用いて, 3 章の GA を用いた基底関数構築法 (Algorithm 1) を評価する.

$$y_{t+1} = 0.39K_1(x_{1;t}, 1)K_2(x_{2;t}, 2) + 0.35K_3(x_{3;t}, 2) + 0.31K_2(x_{2;t}, 1)K_3(x_{3;t}, 1) + 2.04 \quad (5)$$

ここで, $d_x = 4$, $x_{d;t}$ を $0 \leq x_{d;t} \leq 1$ の一様乱数, $\{K_d(\cdot)\}$ を Legendre 関数を用いた 1 次元正規直交基底 (付録 A.1

表 1 Algorithm 1 により得られた J_i および \mathbf{g}_i
Table 1 J_i and \mathbf{g}_i obtained using Algorithm 1.

i	J_i	\mathbf{g}_i
1	3.90×10^{-1}	$(1, 2, 0, 0)^T$
2	3.47×10^{-1}	$(0, 0, 2, 0)^T$
3	3.13×10^{-1}	$(0, 1, 1, 0)^T$
4	2.30×10^{-2}	$(2, 1, 2, 2)^T$
5	2.13×10^{-2}	$(2, 2, 2, 0)^T$
\vdots	\vdots	\vdots

参照) とした.

本節では, Algorithm 1 の動作を検証するために, 学習のために十分な量のデータを観測可能であると仮定し, $\{\mathbf{x}_t, y_{t+1} \mid 0 \leq t < 10^4\}$ を学習データ, $\{\mathbf{x}_t, y_{t+1} \mid 10^4 \leq t < 2 \times 10^4\}$ を評価データとした. また, 時系列を生成するシステムの構造 (基底関数) が既知であると仮定し, MRAGA の基底関数を式 (5) と同様に Legendre 関数を用いて定義した.

Algorithm 1 により, J_i および \mathbf{g}_i が表 1 のように得られた. $\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3$ を式 (A.3) に代入すると, \mathbf{g}_1 は式 (5) の第 1 項, \mathbf{g}_2 は第 2 項, \mathbf{g}_3 は第 3 項を与える. また, J_1, J_2, J_3 の値が J_4, J_5, \dots よりも 10 倍以上大きいことが分かる. これは, 式 (5) を近似するために最も適切な基底関数が, Algorithm 1 により選択されたことを表す. J_1, J_2, J_3 の値は, 各々式 (5) の第 1 項, 第 2 項, 第 3 項の係数と一致する. これは $\{K_d(x_{d;t}, g_{i;d})\}$ が正規直交基底であるため, 式 (2) より明らかである. $N = 3$ として Algorithm 2 で式 (5) の y_{t+1} を予測した結果, 十分小さい予測誤差 ($\varepsilon = 3.6 \times 10^{-2}$) が得られた.

比較のため, Algorithm 1 の GA による基底関数構築を用いずに, 付録 A.1 記載の通常の Fourier 級数の正規直交基底を $\{\phi_i(\mathbf{x}_t)\}$ として, 式 (1) の近似関数 $\tilde{f}(\mathbf{x}_t)$ により予測を行った. 式 (1) の Fourier 係数 $\{\alpha_i\}$ は, Algorithm 2 の Step (2-2) と同様に, 学習データ $\{y_{t+p}, \mathbf{x}_t\}$ を $\{y_{t+p}, \phi(\mathbf{x}_t)\}$ に変換し, 重回帰分析 (式 (4)) により推定された. 次に, Algorithm 2 の Step (2-4) と同様に, 学習データから推定された $\{\alpha_i\}$ と, $\{\phi_i(\mathbf{x}_t)\}$ により変換された評価データ $\{\phi(\mathbf{x}_t)\}$ から, 式 (1) により予測値 \tilde{y}_{t+p} を計算した. 以下では, 重回帰分析 (MRA) のみを用いて Fourier 係数 $\{\alpha_i\}$ を計算する上記手法を MRA と呼ぶ.

MRA で用いる基底関数を MRAGA と同様に Legendre 関数を用いて定義した場合でも, $N = 3$ の MRAGA と同様の近似誤差を得るには, N を 80 以上に設定しなければならなかった. 以上の結果より, 時系列を生成するシステムの基底関数と MRAGA の基底関数が一致し, かつ十分な量の学習データがある場合には, MRAGA の展開次数は MRA より十分小さく設定できることが分かる [14].

4.2 カオス的時系列の予測

本節では、式 (6) で生成される時系列の予測誤差を調べ、MRAGA を評価する。

$$y_{t+1} = 0.3\hat{x}_{1;t+1} + 0.5\hat{x}_{2;t+1}\hat{x}_{3;t+1} + 0.2\hat{x}_{1;t+1}\hat{x}_{3;t+1} \quad (6)$$

$\hat{x}_{1;t}$, $\hat{x}_{2;t}$, $\hat{x}_{3;t}$ は、外乱を含むロジスティック写像 [15] であり、次式で定義される。

$$\begin{cases} \hat{x}_{1;t+1} = h(a_1\hat{x}_{1;t}(1 - \hat{x}_{1;t}) + bw_{1;t}) \\ \hat{x}_{2;t+1} = h(a_2\hat{x}_{2;t}(1 - \hat{x}_{2;t}) + bw_{2;t}) \\ \hat{x}_{3;t+1} = h(a_3\hat{x}_{3;t}(1 - \hat{x}_{3;t}) + bw_{3;t}) \end{cases} \quad (7)$$

ここで、 $h(\cdot)$ は値域を $[0, 1]$ に制限する関数 ($h(\cdot) \stackrel{\text{def}}{=} \max[\min[\cdot, 1], 0]$)、 $a_1 = 3.8823$, $a_2 = 3.777$, $a_3 = 3.951$, b は定数、 $w_{1;t}$, $w_{2;t}$, $w_{3;t}$ は外乱を表し振幅 1 平均 0 の一様乱数で定義される。本節では、十分な量の学習データを観測できると仮定し、 $\{\mathbf{x}_t, y_{t+p} \mid 0 \leq t < 10^3\}$ を学習データ、 $\{\mathbf{x}_t, y_{t+p} \mid 10^3 \leq t < 2 \times 10^3\}$ を評価データとした。時系列を生成するシステム (式 (6), (7)) は、Legendre 関数 (式 (A.5)) の組合せの形に変形できる。すなわち、 $b = 0$ の場合、Legendre 関数を基底関数とすることにより時系列を生成するシステムを誤差なく表せる。そこで、本節では、時系列を生成するシステムの構造 (基底関数) は未知であると仮定し、MRAGA の基底関数を三角関数 (式 (A.4)) を用いて定義した。

MRAGA の近似誤差は、展開次数 N および g_{\max} の増加に従い小さくなる。しかし、その一方で、計算量や必要なメモリが増加する。また、GA の探索空間が広くなり、収束が遅くなる。そこで、 g_{\max} を次式で定義し、 $b = 0$, $p_M = 10^{-3}$ として、 N と \hat{c} が予測精度に与える影響を調べた。

$$g_{\max} = [\hat{c}N^{\frac{1}{4x}} + 0.5] - 1 \quad (8)$$

ここで、 $[\cdot]$ はガウス記号である。また、突然変異率 p_M の値は、GA の収束が早くかつ評価データの予測誤差が小さくなるように最適化された。 N および \hat{c} と予測誤差の関係を図 1 に示す。 $\hat{c} = 2^{1/2}$ および $\hat{c} = 2$ の場合、展開次数 N の増加に従い予測誤差が減少し、 $N > 64$ で $\hat{c} = 2^{1/2}$ の予測誤差が最小となった。一方、 $\hat{c} = 4$ の場合 $N > 64$ で予測誤差が増加する。これは GA の探索空間が広くなり、高い適応度の染色体が得られなかったためである。

\hat{c} の値が GA の収束速度に与える影響を評価するために、MRAGA の学習回数と予測誤差の関係を図 2 に示す。ここで、展開次数 $N = 256$ である。図 2 より、学習回数の増加に従い、予測誤差が減少することが分かる。また、 \hat{c} の値の増加に従い探索空間が広がり、収束速度が遅くなる。

次に、予測誤差の比較のために、サポートベクタマシン (SVM) の非線形識別機能を用いて予測を行った (付録

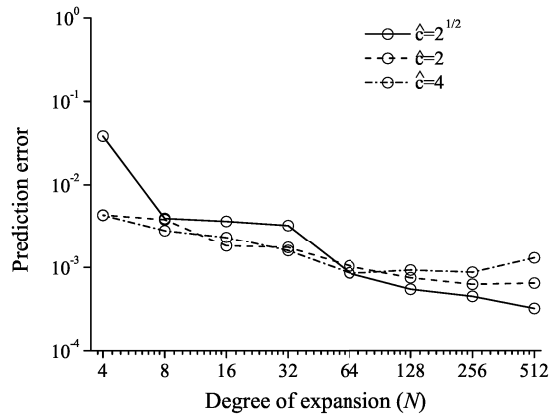


図 1 展開次数 N に対する予測誤差 (MRAGA)

Fig. 1 Effect of degree of expansion (N) on prediction error (MRAGA).

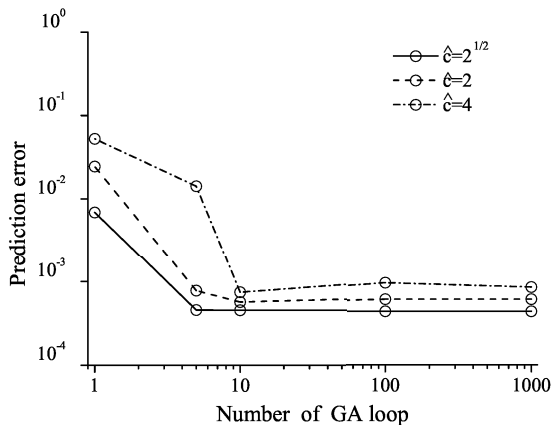


図 2 GA の学習回数に対する予測誤差 (MRAGA)

Fig. 2 Effect of number of GA loop on prediction error (MRAGA).

A.2 参照). SVM はパターン認識の分野において、最も優れた手法の 1 つであり、線形のパターン認識問題のみでなく、非線形問題に対するパターン認識能力も高い。そこで、LIBSVM [16] の識別機能を時系列予測に適用した。SVM のパラメータ (Lagrange 乗数の上限値 C , ガウシアンカーネルにおけるガウス関数の広がりを表す σ , および、クラス数 N_c) を $C = 10^3$, $\sigma = 10^1$, $N_c = 20$ とした。これらの値は、 $b = 0$ の際に評価データの予測誤差が最小になるように最適化された。

MRAGA と従来手法を比較するために、ロジスティック写像の外乱の振幅 b が予測誤差に与える影響を評価した。ここで、従来手法として前述の SVM および 4.1 節で述べた MRA を用いる。MRA の基底関数を、MRAGA と同様に三角関数 (式 (A.4)) を用いて定義した。その結果を図 3 に示す。ここで、 $b = 0$ において、MRAGA, MRA, および SVM の予測誤差ができるだけ等しくなるように MRAGA のパラメータを、 $N = 256$, $\hat{c} = 2$, $p_M = 10^{-3}$, MRA の

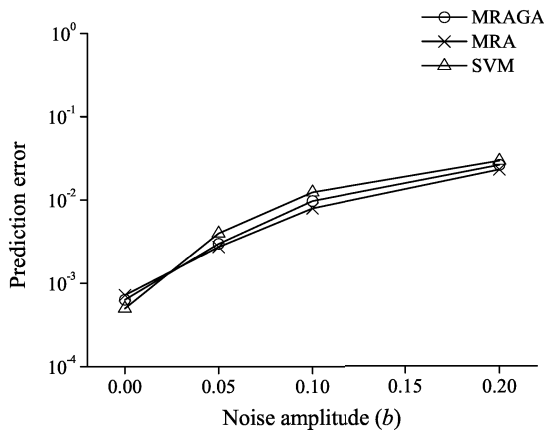


図 3 外乱の振幅 b に対する予測誤差

Fig. 3 Effect of noise amplitude (b) on prediction error.

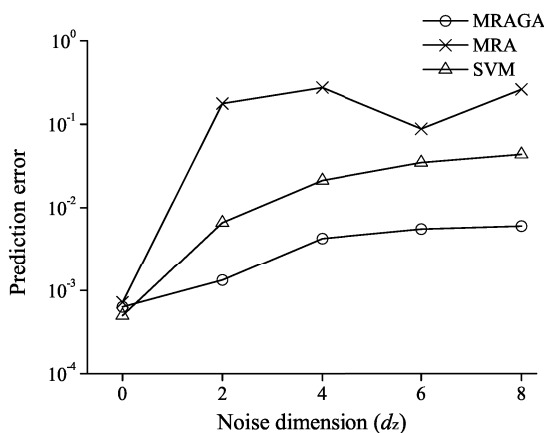


図 4 目的変数と独立な変数の次元 d_z に対する予測誤差

Fig. 4 Effect of noise dimension (d_z) on prediction error.

パラメータを $N \simeq 256$ とした*2. 図 3 より, 各々に有意な差はなく, b の増加に従い予測誤差が増加することが分かる.

最後に, 説明変数を $\hat{x}_t = (\hat{x}_{1;t}, \hat{x}_{2;t}, \hat{x}_{3;t}, z_{1;t}, \dots, z_{d_z;t})^T$ で定義し, 説明変数に含まれる時系列 $z_{1;t}, \dots, z_{d_z;t}$ の影響を評価した. ここで, $z_{1;t}, \dots, z_{d_z;t}$ は目的変数と独立な平均 0.5 振幅 0.5 の一様乱数, $b = 0$ である. また, $d_z = 0$ において, MRAGA, MRA, および SVM の予測誤差ができるだけ等しくなるように, MRAGA のパラメータを $N = 256$, $\hat{c} = 2$, $p_M = 10^{-3}$, MRA のパラメータを $N \simeq 256$ *2 とした. その結果を図 4 に示す.

目的変数と独立な時系列の次元数 d_z の増加に従い, 各手法の予測誤差は増加する. しかし, MRAGA の予測誤差増加率は, SVM および MRA と比較して小さい. これは, 予測対象である目的変数と独立な説明変数を含むことなく, 予測に有効な基底関数の組合せが, Algorithm 1 によって適切に選択されたためである.

*2 MRA では, N の値は式 (A.10) で決まるため, 任意の値をとれない. そこで, N の値を 256 より大きくかつできるだけ近い値とした.

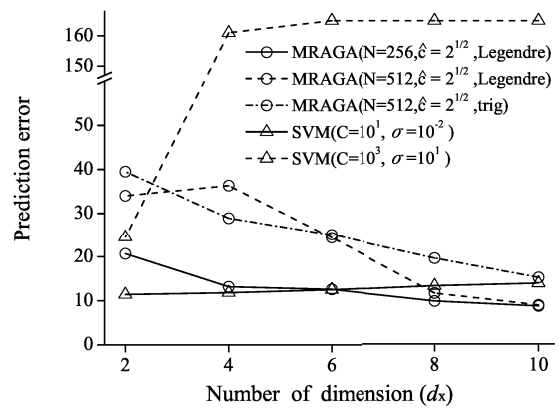


図 5 説明変数の次元数 d_x に対する予測誤差

Fig. 5 Effect of state dimension (d_x) on prediction error.

4.3 気象データの予測

本節では, 時系列を生成するシステムの構造 (基底関数) が未知な実データに対して MRAGA を評価するために, 2000 年から 2011 年までの北海道函館市における 1 日の平均気温 [17] を用いてシミュレーションを行った. 2000 年から 2005 年までのデータを学習データ, 2006 年から 2011 年までのデータを評価データ, t を 2000 年 1 月 1 日または 2006 年 1 月 1 日からの積算の日付, t における 1 日の平均気温を \hat{x}_t として, 目的変数 y_{t+p} を \hat{x}_{t+p} , 説明変数 x_t を次式で定義した.

$$x_t \stackrel{\text{def}}{=} (\hat{x}_t, \hat{x}_{t-1}, \dots, \hat{x}_{t-d_x+1})^T \quad (9)$$

ここで, \hat{x}_t の定義域は $[-15, 30]$ である.

気象データの場合, 目的変数の変化を決めるメカニズムが非常に複雑であり, 予測に必要な説明変数の種類や数を必ずしも特定できるとは限らない. このような場合でも, 目的変数に関係する可能性がある観測データをできるだけ多く説明変数として用いることにより, MRAGA の性能が向上することを示すために, 説明変数の次元数 d_x が MRAGA および SVM の予測誤差に与える影響を評価した. ここで, $p = 3$, SVM のクラス数 N_c を 45 とした. MRAGA および SVM のパラメータとして, 4.2 節で得られた最適なパラメータ (MRAGA: $N = 512$, $\hat{c} = 2^{1/2}$, $p_M = 10^{-3}$, SVM: $C = 10^3, \sigma = 10^1$), および本節のデータに対する最適なパラメータ (MRAGA: $N = 256$, $\hat{c} = 2^{1/2}$, $p_M = 10^{-3}$, SVM: $C = 10^1, \sigma = 10^{-2}$) の 2 種類を用いた. また, 基底関数が予測誤差に与える影響を評価するために, MRAGA の基底関数を三角関数 (式 (A.4)) および Legendre 関数 (式 (A.5)) を用いて定義した.

その結果を図 5 に示す. 図 5 より, SVM のパラメータは予測誤差に大きく影響することが分かる. また, SVM では, どちらのパラメータを用いても d_x の増加に従い予測誤差が増加する. 一方, MRAGA では, d_x の増加に従い予測誤差が減少する. また, 基底関数が Legendre 関数の場合, d_x の増加に従いパラメータの違いによる予測誤差

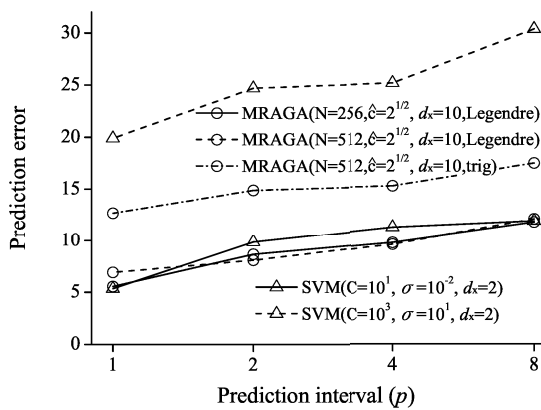


図 6 予測時間差 p に対する予測誤差

Fig. 6 Effect of prediction interval (p) on prediction error.

の差が減少し、 $d_x = 10$ でその差はほぼ 0 になる。しかし、基底関数が三角関数の場合（図の凡例では、trig で表される）、基底関数が Legendre 関数の場合と比較して予測誤差は大きい。SVM の最小予測誤差は $C = 10^1$, $\sigma = 10^{-2}$, $d_x = 2$ で $\varepsilon = 11.4$, MRAGA の最小予測誤差は $N = 256$, $d_x = 10$ で $\varepsilon = 8.9$ であり、MRAGA の予測誤差が小さいことが分かる。

次に、図 5 で最小予測誤差が得られた d_x を用いて、 p が予測誤差に与える影響を評価した。その結果を図 6 に示す。ここで、 p は予測対象（目的変数）と説明変数の時間差を表し、 p が大きいほど長期の予測となる。図 6 より、MRAGA および SVM の両者とも p の増加に従い予測誤差が増加することが分かる。しかし、図 5 と同様に、基底関数として Legendre 関数を用いた場合、MRAGA のパラメータによる予測誤差の差は SVM と比較して小さく、MRAGA の予測誤差は SVM と比較して小さい。また、MRAGA の基底関数として Legendre 関数を用いた場合の方が、三角関数を用いた場合よりも、予測誤差が小さい。これらの結果より、気象データの予測には、基底関数として Legendre 関数が適していることが分かる。

図 3 から図 6 のシミュレーションでは、ある与えられた条件 (b, d_z, p, d_x) においてパラメータ (N, \hat{c}, p_M , 基底関数, C, σ, N_c) を最適化し、その最適化されたパラメータを用いて、様々な条件に対する予測誤差を調べた。その結果、SVM と比較して、MRAGA は、

- (1) 予測対象や条件による予測誤差の変動が少ない,
 - (2) できるだけ多くの説明変数を用いることにより予測誤差を改善できる,
 - (3) 予測誤差が小さい,
- という特徴を持つことが分かった。

パラメータが固定された状態における上記 (1) の特徴は、MRAGA が条件の変化にロバストであることを表しており、MRAGA は他の手法と比べて汎化能力が高いことが予想される。

表 2 最適化されたパラメータに対する予測誤差 (分割法)
Table 2 Prediction errors for optimum parameters (hold out method).

予測方式	予測誤差	最適なパラメータ
MRAGA	1.03×10^1	$\hat{c} = 2, N = 64$
SVM	1.11×10^1	$\sigma = 10^{-3}, C = 10^1$

表 3 最適化されたパラメータに対する予測誤差 (交差検定)
Table 3 Prediction errors for optimum parameters (cross validation).

予測方式	予測誤差	最適なパラメータ
MRAGA	9.50×10^0	$\hat{c} = 1.12 \sim 2, N = 32 \sim 64$
SVM	1.07×10^1	$\sigma = 10^{-2} \sim 10^{-3}, C = 10^0 \sim 10^1$

そこで、MRAGA および SVM の汎化能力を検証するために、気象データを 2000 年～2003 年、2004 年～2007 年、2008 年～2011 年に 3 分割し、2000 年～2003 年のデータを学習データ、2004 年～2007 年のデータを評価データとして、評価データの予測誤差が最も小さくなる最適なパラメータを選定した。ここで、すべてのパラメータ (MRAGA: N, p_M, \hat{c} , 基底関数, SVM: N_c, C, σ) および d_x を最適化することは困難なため、基底関数, p_M, N_c , および d_x を、図 5 において MRAGA と SVM の予測誤差がほぼ一致する点の値に固定した (基底関数を Legendre 関数, $p_M = 10^{-3}$, $N_c = 45$, $d_x = 6$)。また、 $p = 3$ とした。最適なパラメータは、MRAGA では N が 8, 16, 32, ..., 512 の 7 種類、 \hat{c} が 1, ..., 22.6 (g_{\max} が 1, 3, 7, 15, 31 の 5 種類となるように選択), SVM では σ が $10^{-4}, 10^{-3}, \dots, 10^2$ の 7 種類、 C が $10^{-1}, 10^0, \dots, 10^4$ の 6 種類の組合せの中から選ばれた。次に、得られた最適なパラメータを用いて、2004 年～2007 年のデータを学習データ、2008 年～2011 年のデータを評価データとして、予測誤差を測定した (分割法)。その結果を表 2 に示す。同様に、3 分割されたデータを順に入れ替え予測誤差を測定し、その平均を求めた (交差検定)。その結果を表 3 に示す。表 2 および表 3 に示すとおり、MRAGA の予測誤差は SVM より小さく、図 3 から図 6 の評価と同様に、MRAGA の汎化能力が高いことを確認できた。

5. 結論

予測に必要な説明変数の種類や数を特定できない場合でも、目的変数に関係する可能性がある観測データをできるだけ多く説明変数として用いることにより、従来手法と比べて高い予測精度が得られた。これは、目的変数との相関が必ずしも明らかでない場合でも、可能性のある様々な観測値を、提案手法の説明変数に取り入れられることを表す。すなわち、十分な数の説明変数を使えないという問題を解決し、真に必要な説明変数とその基底関数をシステムチックに選択できること意味する。今後、提案手法にお

ける選択, 交差, および突然変異を改良することにより, 収束速度と予測精度を改善していく.

参考文献

[1] 長瀬隆久: AR モデルにダミー変数を加えた時系列予測法, 情報処理学会論文誌, Vol.43, No.10, pp.3247–3250 (2002).

[2] 松葉育雄: ニューラルネットワークによる非線形時系列予測, 日本信頼性学会誌, Vol.28, No.7, pp.442–450 (2006).

[3] Sutton, R.S. and Barto, A.G.: *Reinforcement Learning*, MIT Press, USA (1998).

[4] Menache, I., Mannor, S. and Shimkin, N.: Basis function adaptation in temporal difference reinforcement learning, *Annals of Operations Research*, Vol.134, No.1, pp.215–238 (2005).

[5] 増田直紀: ウェーブレット係数列を用いたカオス時系列の予測, 電子情報通信学会論文誌, Vol.J82-A, No.11, pp.1710–1718 (1999).

[6] Momma, M. and Bennett, K.P.: A Pattern Search method for model selection of Support Vector Regression, *Proc. SIAM International Conference on Data Mining*, pp.261–274 (2002).

[7] Scholkopf, B.: Shrinking the Tube: A New Support Vector Regression Algorithm, *NIPS*, pp.330–336 (1999).

[8] Muller, K.-R., Smola, A.J., Ratsch, G., Scholkopf, B., Kohlmorgen, J. and Vapnik, V.: Predicting Time Series with Support Vector Machines, *Lecture Notes in Computer Science*, Vol.1327, pp.999–1004 (1997).

[9] 三宮信夫, 喜多 一, 玉置 久, 岩本貴司: 遺伝アルゴリズムと最適化, 朝倉書店 (1998).

[10] Johnson, R.A. and Wichern, D.W.: *Applied Multivariate Statistical Analysis, 5th ed.*, Pearson Education (Prentice Hall), USA (2001).

[11] Cristianini, N. and Taylor, J.S.: *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*, Cambridge Univ. Press, UK (2000).

[12] Graepel, T., Herbrich, R., Scholkopf, B., Smola, A., Bartlett, P., Muller, K.R., Obermayer K. and Williamson, R.: Classification on Proximity Data with LP-Machines, *Proc. ICANN'99*, Vol.1, pp.304–309 (1999).

[13] Zhang, T.: Adaptive Forward-Backward Greedy Algorithm for Learning Sparse Representations, *IEEE Trans. Info. Th.*, Vol.57, pp.4689–4708 (2011).

[14] Satoh, H.: Basis Construction with Genetic Algorithm for Function Approximation and Prediction, *IEICE Society Conference*, A-2-15 (2012).

[15] Ott, E.: *Chaos in Dynamical Systems, 2nd ed.*, Cambridge University Press, UK (2002).

[16] Chih-Chung, Chih-Jen Lin: LIBSVM a library for support vector machines (2012), available from (<http://www.csie.ntu.edu.tw/~cjlin/libsvm>).

[17] 気象庁: 気象統計情報 (2013), 入手先 (<http://www.data.jma.go.jp/obd/stats/etrn/index.php>).

[18] Bronshtein, I.N. and Semendyayev, K.A.: *Handbook of Mathematics*, Springer-Verlag, UK (1997).

[19] Papoulis, A.: *The Fourier Integral and its Applications*, McGraw-Hill, USA (1962).

付 録

A.1 関数近似のための正規直交基底

$\mathbf{x} \stackrel{\text{def}}{=} (x_1, \dots, x_{d_x})^T$, $\mathcal{D}_x \stackrel{\text{def}}{=} \{\mathbf{x} | x_{\min d} \leq x_d \leq x_{\max d}, 1 \leq d \leq d_x\}$ とする. 本章では, $\mathbf{x} \in \mathcal{D}_x$ の正規直交基底について述べる. $\mathbf{g} \stackrel{\text{def}}{=} (g_1, \dots, g_{d_x})^T \in \mathcal{Z}$ をインデックスベクトル, \mathcal{Z} を \mathbf{g} の集合, $h(\mathbf{g})$ を \mathbf{g} に対する Fourier 係数とする. 関数 $f(\mathbf{x})$ の Fourier 級数展開は, 次式により定義される [18].

$$f(\mathbf{x}) \simeq \sum_{\mathbf{g} \in \mathcal{Z}} h(\mathbf{g})K(\mathbf{x}, \mathbf{g}) \tag{A.1}$$

$$h(\mathbf{g}) \stackrel{\text{def}}{=} \int_{\mathcal{D}_x} f(\mathbf{x})K^*(\mathbf{x}, \mathbf{g})d\mathbf{x} \tag{A.2}$$

ここで, 上付き添え字 * は複素共役, $\{K(\mathbf{x}, \mathbf{g})\}$ は多次元正規直交基底である. $\{K_d(x_d, g_d)\}$ を 1次元の正規直交基底とすると, $K(\mathbf{x}, \mathbf{g})$ は次式で定義される.

$$K(\mathbf{x}, \mathbf{g}) \stackrel{\text{def}}{=} \prod_{d=1}^{d_x} K_d(x_d, g_d) \tag{A.3}$$

$K_d(x_d, g_d)$ が三角関数を用いた正規直交基底の場合, $K_d(x_d, g_d)$ は次式により定義される [19].

$$K_d(x_d, g_d) \stackrel{\text{def}}{=} \begin{cases} \sqrt{\frac{1}{D_{xd}}} & \text{for } g_d = 0 \\ \sqrt{\frac{1}{D_{xd}}} \sin(\frac{g_d+1}{2}\omega_{0d}(x_d - x_{\min d})) & \text{for } g_d \in \mathcal{N}_o \\ \sqrt{\frac{1}{D_{xd}}} \cos(\frac{g_d}{2}\omega_{0d}(x_d - x_{\min d})) & \text{for } g_d \in \mathcal{N}_e \end{cases} \tag{A.4}$$

ここで, $D_{xd} \stackrel{\text{def}}{=} x_{\max d} - x_{\min d}$, $\omega_{0d} \stackrel{\text{def}}{=} 2\pi/D_{xd}$, \mathcal{N}_o は偶数の集合, \mathcal{N}_e は奇数の集合である.

また, $K_d(x_d, g_d)$ が Legendre 関数を用いた正規直交基底の場合, $K_d(x_d, g_d)$ は次式により定義される.

$$K_d(x_d, g_d) \stackrel{\text{def}}{=} \sqrt{\frac{2g_d+1}{D_{xd}}} P\left(2\frac{x_d - x_{\min d}}{D_{xd}} - 1, g_d\right) \tag{A.5}$$

ここで, $P(\cdot)$ は $x \in [-1, 1]$ の Legendre 多項式であり, 次式で表される.

$$\begin{cases} P(x, 0) = 1, \\ P(x, 1) = x, \\ P(x, 2) = \frac{(3x^2 - 1)}{2}, \\ P(x, 3) = \frac{(5x^3 - 3x)}{2}, \\ \vdots \end{cases} \tag{A.6}$$

基底関数 $\{\phi_i(\cdot)\}$ を次式で定義する.

$$\phi_i(\mathbf{x}) \stackrel{\text{def}}{=} K(\mathbf{x}, \mathbf{g}) \quad (\text{A.7})$$

ここで、 i は基底のインデックスである。

通常の Fourier 級数の場合、 N_d を x_d の展開次数、 $\mathcal{Z}_d \stackrel{\text{def}}{=} \{0, 1, \dots, N_d\}$ 、 \mathcal{Z} を \mathcal{Z}_d の直積により $\mathcal{Z} \stackrel{\text{def}}{=} \mathcal{Z}_1 \times \mathcal{Z}_2 \times \dots \times \mathcal{Z}_{d_x}$ のように与えると、 i は \mathbf{g} を用いて次式で表される。

$$i = \sum_{d=1}^{d_x} g_d \prod_{d'=d+1}^{d_x} (N_{d'} + 1) \quad (\text{A.8})$$

すなわち、通常の Fourier 級数は、Fourier 係数を α_i 、基底関数 $\{\phi_i(\mathbf{x})\}$ を式 (A.3), (A.7), (A.8) で表される正規直交基底で与えることにより、

$$f(\mathbf{x}) \simeq \sum_{i=0}^N \alpha_i \phi_i(\mathbf{x}) \quad (\text{A.9})$$

で定義される。また、 \mathbf{x} の展開次数 N は次式により得られる。

$$N = \prod_{d=1}^{d_x} (N_d + 1) - 1 \quad (\text{A.10})$$

A.2 Support Vector Machine による時系列予測

LIBSVM [16] の非線形 SVM の手法の 1 つであるサポートベクタ分類機 (C-support vector classification (C-SVC)) のパターン認識機能は、多次元ベクトルで表される入力 \mathbf{x} が属するクラスの番号を出力する。その機能を用いて時系列の予測を行うアルゴリズムを、Algorithm 3 に示す。

Algorithm 3 では、予測対象である目的変数 y_{t+p} を量子化数 N_c で量子化し、量子化された値をクラス $class_{t+p} \in \{1, 2, \dots, N_c\}$ に変換する。これにより実数 y_{t+p} の予測問題をパターン認識問題に置き換え、SVM のパターン認識機能を用いて予測を行う。ここで、上記アルゴリズムの SVM で用いられるカーネル関数はガウシアンカーネルである。また、量子化数 N_c は SVM のクラス数と等しい。

SVM の内部計算で用いられる Lagrange 乗数の上限値 C およびガウシアンカーネルにおけるガウス関数の広がりを表す σ は予測精度に大きく影響する。また、量子化数 N_c を大きくすることにより、量子化誤差を削減できる。しかし、 N_c の増加に従い SVM の計算が複雑になるため、予測精度が上がるとは限らない。

Algorithm 3 : SVM を用いた予測

Step (3-1) パラメータ (N_c, C, σ) を与える。

Step (3-2) 学習データの目的変数 y_{t+p} を量子化数 N_c で量子化し、量子化された値をクラス数 $class_{t+p} \in \{1, 2, \dots, N_c\}$ に変換する。

Step (3-3) $class_{t+p}$ を出力、学習データの説明変数 \mathbf{x}_t

を入力として SVM の学習を行う。

Step (3-4) 評価データの説明変数 \mathbf{x}_t を Step (3-3) で構築した SVM に入力し、その出力を、 \mathbf{x}_t に対応するクラス $class_{t+p}$ の推定値 \widetilde{class}_{t+p} とする。

Step (3-5) \widetilde{class}_{t+p} を逆量子化し、予測値 \widetilde{y}_{t+p} とする。



塚田 将太 (正会員)

昭和 63 年生。平成 23 年公立はこだて未来大学システム情報科学部情報アーキテクチャ学科卒業。平成 25 年同大学大学院修士課程修了。強化学習、データマイニング、および非線形時系列予測の研究に従事。同年日立ア

イ・エヌ・エス・ソフトウェア (株) 入社。



佐藤 仁樹

昭和 36 年生。昭和 62 年早稲田大学理工学研究科修士課程修了。同年 (株) 東芝研究開発センター入社。ATM 網のトラフィック制御、およびインターネットの輻輳制御の研究に従事。平成 12 年 4 月より (株) ワイ・アール・

ピー移動通信基盤技術研究所に出向。移動通信網の送信電力制御、輻輳制御、およびインターネット TV 会議システムの研究に従事。平成 14 年より公立はこだて未来大学にて非線形科学および通信システム制御の研究に従事。博士 (情報科学)。