

局所線形ウェーブレットニューラルネットワークのベイズ的設計法

A Bayesian Local Linear Wavelet Neural Network

小林 邦和 (P)[†], 大林 正直[†], 呉本 堯[†]

Kunikazu Kobayashi(P), Masanao Obayashi, and Takashi Kuremoto

[†] 山口大学 大学院理工学研究科

koba@yamaguchi-u.ac.jp

Abstract— This paper proposes a Bayesian design method for a local linear wavelet neural network, which has developed to overcome the curse of dimensionality. The performance of the proposed method is evaluated through computer simulation.

Keywords— Local Linear Model, Wavelet Neural Network, Network Design, Parameter Learning, Bayesian Method

1 はじめに

近年、信号解析の分野で注目を集めているウェーブレット [1] とニューラルネットワークを融合させたウェーブレットニューラルネットワーク (WNN) に関する研究が数多く行われている [2, 3, 4]. WNN の最大の特徴は、時間的にも周波数的にも局在化したマザーウェーブレットと呼ばれる基底関数を用いているところである. また局在化した基底関数の位置と解像度 (局在度) は、それぞれ移動パラメータと解像パラメータにより、自由に調節できるため、任意精度の局所性の同定が実現できる.

一方、WNN は次元数の増加に伴い、必要な中間ユニット数が爆発的に多くなるといういわゆる次元の呪いという問題を抱えている. T.Wang らは、このような問題を解消するために、局所線形モデルを WNN に導入した局所線形ウェーブレットニューラルネットワーク (LLWNN) を提案している [5]. それ以降、LLWNN に関しては、主にパラメータ学習の面から研究が行われているが [6], ネットワーク設計に関する研究はほとんど行われていない.

本論文では、ベイズ法 [7] を用いて、LLWNN のネットワーク設計とパラメータ学習を行う設計法を提案する [8]. 提案法では、C.C.Holmes らの枠組み [9] に従って、LLWNN の定式化を行っている. 関数近似問題を用いて、提案法の有効性を検証する.

2 局所線形ウェーブレットニューラルネットワーク

3 層階層型 LLWNN の構成を図 1 に示す. LLWNN は入力ベクトル $\mathbf{x} = \{x_1, x_2, \dots, x_N\} \in R^N$ が与えられた時、ウェーブレットの基底関数 Ψ (中間ユニット) の出力を局所線形パラメータ c_i で重み付け加算した出

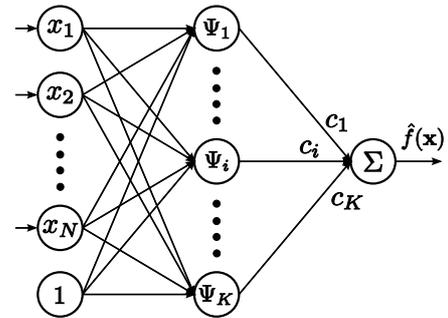


図 1: 局所線形ウェーブレットニューラルネットワークの構成

力 $\hat{f}(\mathbf{x})$ に変換を行う. つまり、LLWNN の出力 \hat{f} は、次式で表される [5].

$$\begin{aligned} \hat{f}(\mathbf{x}) &= \sum_{i=1}^K c_i \Psi_i(\mathbf{x}) \\ &= \sum_{i=1}^K (w_{i0} + w_{i1}x_1 + \dots + w_{iN}x_N) \Psi_i(\mathbf{x}) \quad (1) \end{aligned}$$

式 (1) において、基底関数 Ψ_i は次式で定義される.

$$\Psi_i(\mathbf{x}) = |a_i|^{-\frac{1}{2}} \psi \left(\frac{\mathbf{x} - \mathbf{b}_i}{a_i} \right) \quad (2)$$

ここで、 $a_i \in R^N$ は解像パラメータ、 $\mathbf{b}_i \in R^N$ は移動パラメータ、また基底関数 ψ は、アドミッシブル条件を満足するものとする [1].

LLWNN では、局所線形モデルを用いることにより、通常の WNN と比較して、入力空間の広い範囲をカバーできるため、中間層のユニット数と全体のパラメータ数を抑えることができ、次元の呪いを回避することができる. また局所線形モデルは、高次元において、学習サンプル数が少ない時に良い補完を行うことも示されている.

一方、ネットワーク設計に関しては、LLWNN に関するほとんどの先行研究では、経験的に中間層ユニット数を決めている [5, 6]. すなわちネットワーク構造は、WNN のパラメータ学習の期間中、固定のままである.

3 LLWNN のベイズ的設計法

本論文では, LLWNN のネットワーク設計に対して, ベイズ法を用いた手法を提案する. 提案法は, C.C.Holmes らの枠組み [9] に従って, LLWNN の定式化を行っている.

まず, n 個の入力データ $\mathbf{x}_i \in R^N$ と出力データ $y_i \in R$ から成るデータ集合 $D = \{(\mathbf{x}_i, y_i) | i = 1 \sim n\}$ を定義する. \mathbf{x}_i と y_i の間には, $y_i = f(\mathbf{x}_i) + \epsilon_i$ の関係が成り立つと仮定する. なお, ϵ_i はノイズ項を表わす. この時, 入出力特性 $f(\cdot)$ を式 (1) の LLWNN でモデル化する. また, モデル M_k とモデル空間 M をそれぞれ $M_k = \{a_1, \mathbf{b}_1, \dots, a_k, \mathbf{b}_k\}$, $M = \{k, M_k, W\}$ で定義する. なお, $W = (\mathbf{w}_1, \dots, \mathbf{w}_K)$ は重み行列を表わし, $\mathbf{w}_i = (w_{i0}, w_{i1}, \dots, w_{iN})'$ (\mathbf{w}' は \mathbf{w} の転置) である. また, LLWNN の未知パラメータ θ_i を $\theta_i = \{a_i, \mathbf{b}_i, \mathbf{w}_i\}$ で定義する. 予測分布は LLWNN の予測出力 $\hat{f}_M(\mathbf{x})$ を事後確率 $p(W|M_k, D)$ と $p(M_k|D)$ で平均をとることにより得られる.

$$p(y|\mathbf{x}, D) = \sum_k \iint \hat{f}_M(\mathbf{x}) p(W|M_k, D) p(M_k|D) dM_k dW \quad (3)$$

本論文では, 式 (3) における積分の計算を可逆ジャンプ MCMC 法 [10] を用いて解く.

4 計算機シミュレーション

提案法の性能を検証するために計算機シミュレーションを行った. シミュレーションでは, T.Wang らが用いた次式の関数近似問題を用いて性能を評価する.

$$f(x_1, x_2) = 0.5 \{ \sin(\pi x_1) \cos(\pi x_2) + 1.0 \} \quad (4)$$

ここで, 入力変数 x_i の範囲は, $x_i \in [-1.0, 1.0]$ とする. 学習データは $[-1.0, 1.0] \times [-1.0, 1.0]$ の領域を 7×7 の格子状に区切った 49 個の点 (x_1, x_2) とそれに対応する出力 $f(x_1, x_2)$ とする. またテストデータは同じ領域を 20×20 の格子状に区切った 400 個のデータを用いる.

式 (2) における関数 ψ は, $\psi(x) = -x \exp(-x^2/2)$ とする. 性能評価はネットワークサイズ (中間層ユニット数), フリーパラメータ数, 及び平方平均二乗誤差 (RMSE) を用いて行う.

シミュレーション結果を表 1 に示す. 表 1 では, 提案法 BLLWNN の性能を比較するため, 通常の WNN と T.Wang らの LLWNN のシミュレーション結果を掲載している [5]. ここで, WNN と LLWNN に関しては, その中間層ユニット数はシミュレーションを通して経験的に決められたものであるが, 提案法に関しては, 可逆ジャンプ MCMC 法により求めた中間層ユニット数を示している. 表 1 より, BLLWNN は WNN と比較して, ネットワークサイズ, パラメータ数, 及び RMSE

表 1: シミュレーション結果

手法	ユニット数	パラメータ数	RMSE
WNN	8	40	1.65×10^{-2}
LLWNN	4	28	1.58×10^{-2}
BLLWNN	4	28	1.56×10^{-2}

の面から優れていることが分かる. 更に, BLLWNN は LLWNN と比較して, 同じネットワークサイズ, パラメータ数のネットワークで, RMSE が同程度の性能を示していることが分かる. 従って, BLLWNN は, 従来の WNN や LLWNN より, ネットワーク設計とパラメータ学習が同時に行えるという点で優れていると言える.

5 まとめ

本論文では, LLWNN のネットワーク設計とパラメータ学習を同時に行うベイズ的設計法を提案した. 関数近似問題を用いた計算機シミュレーションを通して, 提案法の有効性を確認した.

参考文献

- [1] C. K. Chui (1992) "An introduction to wavelets.", Academic Press.
- [2] Q. Zhang and A. Benveniste (1992) "Wavelet networks.", IEEE Transactions on Neural Networks, **3**, 6, 889–898.
- [3] Y. C. Pati and P. S. Krishnaprasad (1993) "Analysis and synthesis of feedforward neural networks using discrete affine wavelet transformations.", IEEE Trans. on Neural Networks, **4**, 1, 73–85.
- [4] 小林邦和, 鳥岡豊士, 吉田信夫 (1994) "ネットワーク最適化機能をもつウェーブレットニューラルネットワーク.", 電子情報通信学会論文誌, **J77-D-II**, **10**, 2121–2129.
- [5] T. Wang and Y. Sugai (2002) "A local linear adaptive wavelet neural network.", 電気学会論文誌, **122-C**, **2**, 277–284.
- [6] Y. Chen, J. Dong, B. Yang, and Y. Zhang (2004) "A local linear wavelet neural network.", Proc. of the 5th World Congress on Intelligent Control and Automation, 1954–1957.
- [7] D. MacKay (2003) "Information theory, inference, and learning algorithms.", Cambridge University Press.
- [8] 小林邦和, 大林正直, 呉本堯 (2008) "局所線形モデルを導入したウェーブレットニューラルネットワークのベイズ的設計法.", 電気学会 電子・情報・システム部門大会, **GS5-7** (印刷中).
- [9] C. C. Holmes and B. K. Mallick (1998) "Bayesian radial basis functions of variable dimension.", Neural Computation, **10**, 1217–1233.
- [10] P. J. Green (1995) "Reversible jump Markov chain Monte Carlo computation and Bayesian model determination." Biometrika, **82**, 711–732.