砂時計型多層パーセプトロンによる非線形次元圧縮と次元数選択

Reduction and Selection of Nonlinear Dimensionality by Multi-Layer Perceptron

高橋 隆史 (PY)
筑波大学 電子・情報工学系 学振特別研究員
Takashi TAKAHASHI
JSPS Research Fellow, Univ. of Tsukuba
email : Takashi@Takahashi.com

徳永 隆治
筑波大学 電子・情報工学系
Ryuji TOKUNAGA
Institute of Information Sciences and Electoronics, Univ. of Tsukuba

Abstract In order to perform efficient dimensionality reduction by multi-layer perceptron, it is crucial to select the appropriate number of units in the bottleneck-layer. In this paper, we report a simple energy function, called "superposed energy," is successfully applicable to nonlinear dimensionality reduction by autoassociative learning to estimate the intrinsic dimensionality of nonlinear data.

1 はじめに

多層パーセプトロン (MLP) による次元圧縮に関して、非線形素子を含む5層などの砂時計型MLPの恒等写像学習によって非線形主成分分析が実現できることが知られている。しかし、このようなMLPで効率の良い次元圧縮を実現するためには、ボトルネックとなる中間層の素子数 (中間次元数) を適切に設定することが必要である。この問題に対して、平岡・吉澤[4]は、中間次元数の異なる砂時計型MLPの中から最適なものを選択するための基準量を提案している。一方われわれは、中間層出力に寄与順に並んだデータ表現を自己組織するためのエネルギー関数 (重畳エネルギー [2, 3]) を検討し、これによって学習後に中間次元数が任意に選択可能となることを示している。本稿では、重畳エネルギーを用いた砂時計型MLPに平岡・吉澤の基準を適用した結果を報告する。

2 MLPによる次元圧縮

2.1 砂時計型MLPによる次元圧縮

入力層と出力層の素子数が N 、中間層の素子数が $K < N$ の砂時計型MLPを考える。このMLPの恒等写像学習は、「ノイズを含む観測値から真値の満たす関係を推測し、観測値を補正する」という問題と考えることができ、以下のような推定問題と等価となる[1]：「N 次元空間中に、次元数 L が未知のある曲面 S^* を考える。S^* 上にデータ集合

$$\left\{ \boldsymbol{z}_t \in R^N \right\}_{t=1}^{T} \tag{1}$$

が与えられ、それにノイズの加わったデータ

$$\{\boldsymbol{x}_t = \boldsymbol{z}_t + \boldsymbol{n}_t\}_{t=1}^{T} \tag{2}$$

が観測される。ただし、$\boldsymbol{n}_t$ は N 次元正規分布 $N(\mathbf{o}, \sigma^2 I)$ に従うノイズとする。このとき、観測値 $\{\boldsymbol{x}_t\}$ から S^* と真の値 $\{\boldsymbol{z}_t\}$ を推定せよ。」
入力 $\boldsymbol{x}_t$ に対する砂時計型MLPの出力を $\boldsymbol{y}_t$ と表すと、このMLPはエネルギー関数

$$E = \frac{1}{T}\sum_{t=1}^{T} ||\boldsymbol{x}_t - \boldsymbol{y}_t||^2 \tag{3}$$

を最小化するように学習を行う。その結果、出力 $\boldsymbol{y}_t$ は $\boldsymbol{x}_t$ を K 次元のある曲面 S でフィッティングして近似したものとなる。式(3)の値を小さくするためには、MLPの中間次元数 K を大きくとればよい。しかし、真に最小化したいのは真値 $\boldsymbol{z}$ に対する推定値 $\boldsymbol{y}$ の誤差

$$E^* = \mathrm{E}\left[||\boldsymbol{y} - \boldsymbol{z}||^2\right] \tag{4}$$

であり、この値は、中間次元数 K が未知の次元数 L に一致して S が S^* の良い近似となったときに小さくなる。したがって、中間次元数 K を適切に選択することが必要となる。

2.2 中間次元数の選択基準

平岡らは、砂時計型MLPにおける中間次元数 K の選択基準として、真の曲面 S^* やMLPの曲面 S が平面で近似できるという仮定のもとで真の誤差 E^* の推定値を導いている[4]。ノイズの分散 σ^2 が既知の場合の推定値 E_{known} は

$$E_{known} = E + (2K - N)\sigma^2 \tag{5}$$

---

keywords: multi-layer perceptron, dimensionality reduction

で与えられる。一方 σ^2 が未知の場合、これを推定して式 (5) に代入することで

$$E_{unknown} = \frac{K}{N-K}E \tag{6}$$

が導かれる。いずれの値も $K = L$ において最小値をとることから、様々な中間次元数の MLP で式 (5) あるいは (6) の値を比較すれば、最適な中間次元数の MLP を選択することが可能となる。

2.3 重畳エネルギー

式 (3) をエネルギー関数とする従来の恒等写像学習では、様々な中間次元数の MLP で学習を行い、そのなかから最適な中間次元数のものを選択する必要がある。これに対して、重畳エネルギー [2, 3] を用いると、一つの MLP の学習で中間次元数を任意に選択可能な内部表現を自己組織することができる。重畳エネルギーは、一部の中間層素子だけを用いたときの出力誤差を考え、さまざまな素子数での誤差を加え合わせたものである。 K 個の中間層素子のうち 1 番目から i 番目 ($i = 1, 2, \ldots, K$) までを用いて得られる恒等写像の近似写像を f_i と表す。このとき、入力 $\boldsymbol{x}$ に対する出力 $f_i(\boldsymbol{x})$ の平均二乗誤差は

$$E_i = \frac{1}{T}\sum_{t=1}^{T} ||\boldsymbol{x}_t - f_i(\boldsymbol{x}_t)||^2 \tag{7}$$

となる。重畳エネルギーは、この E_i の重みつき和

$$F = \sum_{i=1}^{K} \beta_i E_i \tag{8}$$

で与えられる。ただし、 β_i は適当な正の係数である。このエネルギーを用いると中間層素子が寄与順にならぶため、学習後にそれぞれの中間次元数 i で式 (5) および (6) の値を計算し、最適な次元数を選択することが可能となる。このとき、ノイズの分散が未知の場合については、式 (6) を修正して $E_{unknown}$ の推定精度を向上させること、およびこの値を利用して式 (8) の β_i を適応的に変更することができる [5] が、これらについては紙面の都合上省略する。

3 実験

カオス的時系列データの解析に用いた例を示す。未知の力学系から観測ノイズを含む一次元時系列データが得られるものとして、この時系列からアトラクタの再構成に必要な最小の埋め込み次元数を推定する。対象は Rössler 方程式とした。また、MLP への入力は上記時系列を 8 次元の遅延座標系に埋め込んだ $N = 8$ 次元のデータとし、中間次元数を $K = 7$ として重畳エネルギーによる恒等写像学習を行った。

図1に、ノイズを含まない真の時系列データに対する誤差その推定値 $E_{unknown}$ の値を示す。比較のため、主成分分析で同様の次元圧縮を行った場合の誤差も示している。線形の主成分分析では次元圧縮が不可能であるが、MLP では誤差の小さい最適な次元数の選択が可能となっていることがわかる。

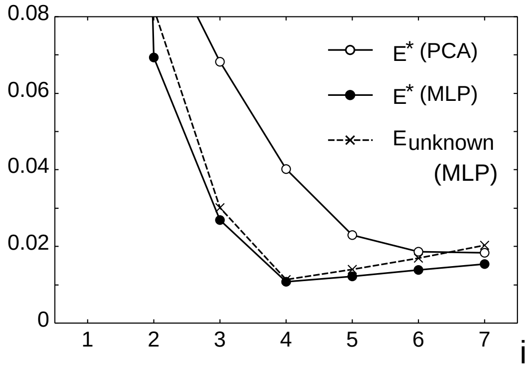

図 1: カオス的時系列データの埋め込み次元推定 :横軸 ;次元数 i 、縦軸 ;真の誤差とその推定値

4 むすび

砂時計型 MLP による次元圧縮において最適な中間次元数を選択するという問題に対して、重畳エネルギー関数と平岡 ·吉澤の基準を用いる手法を検討し、計算機実験によってその有効性を示した。しかし、重畳エネルギーを用いる学習は従来の MLP の学習と比較して計算コストが大きくなる。従来の MLP と同等あるいはそれよりも少ないコストで計算可能な手法を検討することが今後の課題となる。

本研究の一部は、文部省科学研究費 (特別研究員奨励費) の補助のもとに行われたものである。

参考文献

[1] 平岡和幸, 吉澤修治, “データ補正の教師無し学習,” 信学技報, NC97-4, 1997.

[2] 高橋隆史, 徳永隆治, “KL 変換を実現する 3 層線形パーセプトロンの教師付き学習則,” 信学論, vol.J80–D–II, no.5, pp.1267–1275, 1997.

[3] 高橋隆史, 徳永隆治, “重畳エネルギー関数による多層パーセプトロンの冗長性削減,” 信学論, vol.J80-D-II, no.9, pp.2532–2540, 1997.

[4] 平岡和幸, 吉澤修治, “砂時計型ネットの中間次元数選択について,” 信学技報, NC97-139, 1998.

[5] 高橋隆史, 徳永隆治, “重畳エネルギー関数を用いた砂時計型 MLP による非線形主成分分析,” 信学技報, NC99-27, 1999.