

誤差逆伝播学習法による自己組織化ロバスト主成分分析

樋口勇夫*

Self-Organized Robust Principal Component Analysis by Back-Propagation Learning

Isao HIGUCHI

The purpose of this study is the suggestion of a self-organized back-propagation algorithm for robust principal component analysis. The self-organizing algorithm that discriminates the influence of data automatically is applied to learning of a sandglass type neural network.

Key words: robust principal component analysis, neural networks, back-propagation, self-organizing rule

1. はじめに

生物の脳をモデルにしたニューラルネットワークの学習理論は、新しいデータ解析の手法として注目され、研究されてきた。特に情報の特徴抽出などの面でその有用性が証明されており、様々な新手法が提案されてきた。また、このようなネットワークモデルは入力信号を出力信号へと変換する情報処理装置とみなすことが出来る。このような装置の入出力関係は統計処理における元データと統計量との関係に酷似しており、ニューラルネットワークによる統計解析の研究が進められてきた。

一方、多変量解析の中で代表的な手法である主成分分析は、多次元データの特徴を抽出する方法として古くから使われてきた。そのためニューラルネットワークモデルによる応用も古くからなされており、大きく2種類の方法が提起されている。一つは線形ニューロンのヘップ学習に基づくものであり^{1, 2)}、もう一つは砂時計型と呼ばれる多層ネットワークの誤差逆伝播学習法則^{3, 4)}に基づくものである。

しかし、主成分分析は少数の外れ値の存在によって分析結果が大きく違ってしまふ、という問題があった。そのためロバストな主成分分析の研究も進められてきた。従来の主成分分析はデータ分散共分散行列の固有値問題に帰着されるため、分散行列のロバストな推定量を用いることで、ロバストな主成分分析を行うというのが主な方法である。

ニューラルネットワークによる主成分分析では分散行列は陽な形で現れてこない。そのためニューラルネットワ

ークでは影響の大きいデータを自動判別する、自己組織化というアプローチがとられ、線形ニューロンのヘップ学習法則に基づくアルゴリズムに対しては自己組織化アルゴリズム⁵⁾が提起された。筆者らは Higuchi and Eguchi⁶⁾の中で自己組織化による主成分分析がある重みつき分散行列の固有値問題に帰着されることを示し、その行列のロバスト性を影響関数の計算で示すことにより自己組織化アルゴリズムの優位性を理論的に証明した。

もう一方の砂時計型ネットワークについても外れ値の問題は存在する。本稿では砂時計型ネットワークに対して自己組織化アルゴリズムを適用し、その有用性を調べる。

2. ニューラルネットワークと学習

ニューラルネットワークはニューロンと呼ばれる情報処理素子が相互に結合して構成される。ニューロンは入力ベクトル \mathbf{x} と出力 \mathbf{y} をもつ。もっと単純な線形ニューロンでは入出力関係は、

$$\mathbf{y} = \mathbf{w}^T \mathbf{x}$$

となる。ここで \mathbf{w} は結合比重と呼ばれるパラメータである。ニューラルネットワークの学習とは、各ニューロンの結合比重を修正していくことである。

データベクトルを $\{\mathbf{x}_i\}, (i=1, L, N)$ とすると、

* 広島大学大学院工学研究科 複雑システム工学専攻

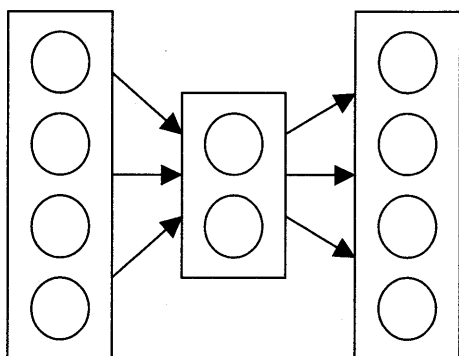
線形ニューロンによる主成分分析の学習アルゴリズムは、

$$\mathbf{w}(t+1) = \mathbf{w}(t) + \alpha(t) \sum_{i=1}^N \frac{\partial}{\partial \mathbf{w}} r(\mathbf{x}_i - \bar{\mathbf{x}}, \mathbf{w}(t)),$$

$$r(\mathbf{x}, \mathbf{w}) = \|\mathbf{x}\|^2 - \frac{(\mathbf{w}^T \mathbf{x})^2}{\|\mathbf{w}\|^2}$$

で与えられる。

砂時計型ネットワークではニューロンは次のような階層型構造を持つネットワークになる。



左から入力層、中間層、出力層であり、入力層と出力層のニューロンの個数は等しく、中間層はそれよりも少ない。このネットワークで主成分分析を行う場合、入力層のデータが中間層で低次元に縮約されることになる。中間層から出力層への流れでは、縮約されたデータを元の次元に復元する。主成分分析では、元の次元に復元するときだけ再現されているのが望ましい。従って、入力データを教師信号とした誤差逆伝播法で主成分分析のアルゴリズムを構成することができる。この場合のアルゴリズムは、

$$\mathbf{w}(t+1) = \mathbf{w}(t) + \alpha(t) \delta(t) \mathbf{w}(t)^T \mathbf{x}_i,$$

$$\delta(t) = -(\mathbf{x}_i - \mathbf{z}(t)),$$

$$\mathbf{z}(t) = \mathbf{w}(t) (\mathbf{w}(t)^T \mathbf{x}_i),$$

で与えられる。

3. 自己組織化アルゴリズム

主成分分析は外れ値の影響を受けやすく、ロバストではない。線形ニューロンの学習法則において、ロバストな主成分分析を行うため、Xu and Yuille⁵⁾ はデータの主成分ベクトルからの残差、

$$r(\mathbf{x} - \bar{\mathbf{x}}, \mathbf{w}) = \|\mathbf{x} - \bar{\mathbf{x}}\|^2 - \frac{(\mathbf{w}^T (\mathbf{x} - \bar{\mathbf{x}}))^2}{\|\mathbf{w}\|^2}$$

の大きさを元に、データが外れ値かどうかを判別し、外れ値ならば自動的にそのウェイトを小さくする自己組織化の

アルゴリズムを提起した。そのアルゴリズムは次のようになる。

$$\mathbf{w}(t+1) = \mathbf{w}(t) + \alpha(t) \sum_{i=1}^N \gamma_i(t) \frac{\partial}{\partial \mathbf{w}} r(\mathbf{x}_i - \bar{\mathbf{x}}, \mathbf{w}(t)),$$

$$r(\mathbf{x}, \mathbf{w}) = \|\mathbf{x}\|^2 - \frac{(\mathbf{w}^T \mathbf{x})^2}{\|\mathbf{w}\|^2},$$

$$\gamma_i(t) = \frac{1}{1 + \exp[\beta(t) \{r(\mathbf{x}_i - \bar{\mathbf{x}}, \mathbf{w}(t)) - \eta(t)\}]}$$

Higuchi and Eguchi⁶⁾ ではこのアルゴリズムが、重みつき分散行列

$$S_\gamma = \sum_{i=1}^N \gamma_i(\mathbf{x}, \mathbf{w}) (\mathbf{x} - \bar{\mathbf{x}}) (\mathbf{x} - \bar{\mathbf{x}})^T,$$

$$\gamma_i(\mathbf{x}, \mathbf{w}) = \frac{1}{1 + \exp[\beta \{r(\mathbf{x}, \mathbf{w}) - \eta\}]}$$

の固有ベクトルに収束することを示した。

この考え方を砂時計型ネットワークの学習について当てはめ、それぞれのデータに対して上と同様の重みをつけた学習にすれば自己組織化ロバスト主成分分析アルゴリズムが構成できる。従って、砂時計型ネットワークにおけるアルゴリズムは、

$$\mathbf{w}(t+1) = \mathbf{w}(t) + \alpha(t) \gamma(t) \delta(t) \mathbf{w}(t)^T \mathbf{x}_i,$$

$$\delta(t) = -(\mathbf{x}_i(t) - \mathbf{z}(t)),$$

$$\mathbf{z}(t) = \mathbf{w}(t) (\mathbf{w}(t)^T \mathbf{x}_i),$$

$$\gamma(t) = \frac{1}{1 + \exp[\beta(t) \{r(\mathbf{x}_i - \bar{\mathbf{x}}, \mathbf{w}(t)) - \eta(t)\}]}$$

で与えられる。

4. 数値実験

数値実験により、このアルゴリズムの有用性を確認する。5次元のデータを1次元の主成分に縮約する場合を考える。まず、90個のデータは正規分布 $N(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ に従うものとし、10個の外れ値のうち、5個は $N(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$ 、5個は $N(-\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$ に従うものとする。ただし、

$$\begin{aligned}\mu_0 &= (0, 0, 0, 0, 0)^T, \\ \Sigma_0 &= \text{diag}(25, 16, 9, 4, 1), \\ \mu_1 &= (0, 0, 0, 0, 20)^T, \\ \Sigma_1 &= \text{diag}(0.25, 0.25, 0.25, 9, 25)\end{aligned}$$

とする。90個のデータについては平均 \bar{x}_0 と第1主成分

ベクトル p_0 は、

$$\begin{aligned}\bar{x}_0 &= (-0.4399, 0.0715, -0.044, 0.0427, -0.1441)^T, \\ p_0 &= (0.8632, 0.5007, -0.0567, -0.021, -0.0239)^T,\end{aligned}$$

となった。外れ値を加えた全てのデータでの平均 \bar{x} と第1主成分ベクトル p は、

$$\begin{aligned}\bar{x} &= (-0.3907, 0.0697, -0.0542, 0.0878, -0.3267)^T, \\ p &= (0.0450, 0.0385, 0.0111, -0.0462, -0.9971)^T,\end{aligned}$$

であり、外れ値の影響で第1主成分ベクトルの方向が、約 85° 変化していることがわかる。誤差逆伝播法による学習では、

$$p_{bp} = (-0.0490, -0.0333, -0.0115, 0.0370, 0.9975)^T,$$

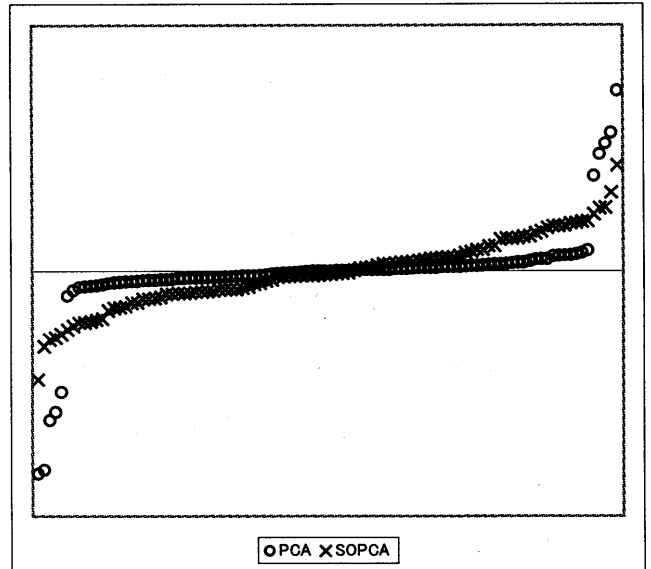
であり、従来の主成分分析とほぼ等しい結果になっている。

一方、 $\beta = 0.1, \eta = 30$ として、自己組織化アルゴリズムを適用すると、

$$p_{so} = (0.7094, 0.7010, -0.0650, -0.0162, -0.0294)^T$$

となった。これは、外れ値を含まない90個のデータによる主成分ベクトル p_0 と比べて、方向が 14° しか違わない。

つまり、単なる誤差逆伝播学習よりも、外れ値の影響を受けていないことがわかる。それぞれの学習法で得られた第1主成分を大きさの順に並べてプロットしてみると、以下のグラフのようになる。



これで見るとわかるように、従来の主成分分析では、10個の外れ値に付いてははっきりと差がでていたものの、90個のデータに付いてはほとんど差異がみとめられなくなっている。それに対して、自己組織化アルゴリズムで求めたものは、全てのデータについて同程度の差異が認められ、第1主成分による縮約が有用なものとして確認できる。

5. 考察

数値実験の結果、誤差逆伝播法においても、自己組織化を適用することによってロバスト主成分分析が可能であることがわかった。また、誤差逆伝播法アルゴリズムを全てのデータについて一括処理する場合、すなわち、

$$w(t+1) = w(t) + \alpha(t) \sum_{i=1}^N \gamma_i(t) \delta_i(t) w(t)^T x_i,$$

$$\delta_i(t) = -(x_i(t) - z_i(t)),$$

$$z_i(t) = w(t) (w(t)^T x_i),$$

$$\gamma_i(t) = \frac{1}{1 + \exp[\beta(t) \{r(x_i - \bar{x}, w(t)) - \eta(t)\}]}$$

を考える。ここで w を固定し、データ $\{x_i\}$ に対して重み

γ_i によって変換されるデータ $\{x'_i\}$ を次のように定義する。

る。

$$x'_i = \sqrt{\gamma_i} x_i$$

このとき、自己組織化誤差逆伝播アルゴリズムは、

$$\begin{aligned} \mathbf{w}(t+1) \\ = \mathbf{w}(t) + \alpha(t) \sum_{i=1}^N \sqrt{\gamma_i(t)} \delta_i(t) \mathbf{w}(t)^T \left(\sqrt{\gamma_i(t)} \mathbf{x}_i \right) \end{aligned}$$

となるが、

$$\begin{aligned} \sqrt{\gamma_i} z_i(t) &= \mathbf{w}(t) \left(\mathbf{w}(t)^T \sqrt{\gamma_i} \mathbf{x}_i \right) \\ &= \mathbf{w}(t) \left(\mathbf{w}(t)^T \mathbf{x}'_i \right) = z'_i(t), \\ \sqrt{\gamma_i} \delta_i(t) &= - \left(\sqrt{\gamma_i} \mathbf{x}_i(t) - \sqrt{\gamma_i} z_i(t) \right) \\ &= - \left(\mathbf{x}'_i(t) - z'_i(t) \right) = \delta'_i(t) \end{aligned}$$

であるから、

$$\begin{aligned} \mathbf{w}(t+1) \\ = \mathbf{w}(t) + \alpha(t) \sum_{i=1}^N \sqrt{\gamma_i(t)} \delta_i(t) \mathbf{w}(t)^T \left(\sqrt{\gamma_i(t)} \mathbf{x}_i \right) \\ = \mathbf{w}(t) + \alpha(t) \sum_{i=1}^N \delta'_i(t) \mathbf{w}(t)^T \mathbf{x}'_i \end{aligned}$$

となり、変換されたデータ $\{\mathbf{x}'_i\}$ における主成分分析にな

っていることがわかる。変換されたデータ $\{\mathbf{x}'_i\}$ について、分散行列を計算すると、

$$\begin{aligned} S' &= \sum_{i=1}^N \mathbf{x}'_i (\mathbf{x}'_i)^T \\ &= \sum_{i=1}^N \left(\sqrt{\gamma_i} \mathbf{x}_i \right) \left(\sqrt{\gamma_i} \mathbf{x}_i \right)^T \\ &= \sum_{i=1}^N \gamma_i \mathbf{x}_i \mathbf{x}_i^T = S_\gamma \end{aligned}$$

となる。従って、重みつき分散行列 S_γ の固有ベクトルが

このアルゴリズムの不動点であることがわかる。

6. まとめ

自己組織化アルゴリズムは砂時計型ネットワークによる主成分分析をロバスト化する手段としても有効であり、線形ニューロンの場合と同様に重みつき分散行列の固有値問題に帰着される可能性があることがわかった。理論的にロバスト性を示していくためには、さらに収束性についても調査していく必要がある。この研究では1次元に縮約するという最も簡単な場合について調べたが、実用上意味があるのは主成分部分空間への縮約の場合であるため、その検討をしていく必要がある。

参考文献

- 1) Amari, S. -I. (1977). Neural theory of association and concept formation. *Biological Cybernetics*, **26**, 175-185.
- 2) Oja, E. (1982). A simplified neuron model as a principal component analyzer. *Journal of Mathematical Biology*, **15**, 267-273.
- 3) Baldi, P. and Hornik, K. (1989). Neural networks and principal component analysis: Learning from examples without local minima. *Neural Networks*, **2**, 53-58.
- 4) Baldi, P. and Hornik, K. (1995). Learning in Linear Neural Networks: A survey. *IEEE transactions on Neural Networks*, **6**, 837-858.
- 5) Xu, L. and Yuille, A. (1995). Robust principal component analysis by self-organizing rules based on statistical physics approach. *IEEE transactions on Neural Networks*, **6**, 131-143.
- 6) Higuchi, I. and Eguchi, S. (1998). The influence function of principal component analysis by self-organizing rule. *Neural Computation*, **10**, 1435-1444.

平成 16 年 8 月 31 日 受理