

入札価格決定支援システムにおいて高い汎化能力を持つ改善BP計算法

韓, 敏

大連理工大学電子情報工程学院 | 九州大学大学院システム情報科学研究所 : 訪問助教授

金江, 春植

九州大学大学院システム情報科学研究所電気電子システム工学部門

和田, 清

九州大学大学院システム情報科学研究所電気電子システム工学部門

<https://doi.org/10.15017/1515836>

出版情報 : 九州大学大学院システム情報科学紀要. 8 (1), pp. 73-76, 2003-03-26. 九州大学大学院システム情報科学研究所

バージョン :

権利関係 :

入札価格決定支援システムにおいて高い汎化能力を持つ改善 BP 計算法

韓 敏* · 金江春植** · 和田清**

An Improved BP Algorithm with Better Generalization Ability in Bidding System

Min HAN, Shunshoku KANAE and Kiyoshi WADA

(Received December 20, 2002)

Abstract: According to the demand of practical bidding procedure in the construction field, the paper adapts the K-L information distance to the learning error function of the neural network and improves the traditional BP algorithm. The bidding system's learning ability and generalization ability to various learning algorithms are compared. Simulations illustrate that the improved BP algorithm makes the whole bidding system have better generalization ability and guarantees its validity in the practical application.

Keywords: BP algorithm, Generalization, Bidding system

1. はじめに

最近、ニューラルネットワークはモデル化、パターン認識等の多くの分野に広く応用されている¹⁾。しかし、システムの汎化能力の問題はいまだに未解決のままである^{2),3)}。システムの汎化能力とは、学習後のネットワークが未学習のサンプル（学習データと同一分布にある）に対しても正確に応答する能力のことである⁴⁾。学習速度と並んで、汎化能力はニューラルネットワークを評価する重要な指標である^{5),6),7)}。しかし、長い間ニューラルネットワーク学習法の研究は学習時間をいかに短縮するか、また学習時にローカルミニマムに陥る問題をいかに解決するかに重点がおかれており⁸⁾、汎化能力についてはあまり考慮されていない。

実際の応用においては、汎化能力はニューラルネットワークの有用性に関わる大事な事柄である。本論文では、建築業界における入札価格決定支援システムについて考えるが、これは競争入札に参加する建設業者が市場の動向、自分の会社の負荷、可能な競争相手等の種々の要素を考慮して、最適な入札価格を決定するシステムである。建設工事の入札価格決定においては、各工事の競争入札情勢が異なっている。競争入札への参加に当たって、現在の入札情勢を代表する最新のデータに基づいて、これらのデータを自動的に分析・計算し、落札の前提で一番大きい利益を追求できるような適切な値段を決めるシステムが求められる^{9),10)}。したがって、入札価格決定支援

システムの汎化能力は実際の応用において、最も重要であり、さらに研究改善をする必要がある¹¹⁾。

本論文では、ニューラルネットワークの確率表現を用い、Kullback-Leibler (K-L) 情報距離とニューラルネットワークの汎化能力との関係に基づいて、新しいニューラルネットワーク学習誤差関数を構成し、従来の BP 計算法を改善した。またこの方法と他のニューラルネットワーク学習法を比較したシミュレーション結果を与え、改善した学習誤差関数による学習ネットワークはより良い汎化能力を持つことを示した。

2. ネットワーク汎化能力を高めるために用いた改善学習誤差関数

2.1 基本原理

ネットワークの学習に用いられるサンプルデータを $D = \{(x_i, y_j), i = 1, 2, \dots, m\}$ とする。いろいろな原因によってサンプルへのノイズの影響は避けられない。そこでシステムの出力は正規分布 $N(0, \sigma_0^2)$ に従うノイズ n_0 の影響を受けるとする。したがって、各サンプルの出力は $y_i = g(x_i) + n_{0i}$ で表せる。ここで、 n_{0i} は i 番目のサンプルのノイズである。 y_i からノイズ n_{0i} を分離するのは不可能であるから、ニューラルネットワークの出力 $f(x, \omega)$ にランダム変数 n を加えたもの $f(x, \omega) + n$ を持って、実際の出力 $g(x) + n_0$ を近似し、ニューラルネットワーク $f(x, \omega)$ が真の関数 $g(x)$ を近似するようにする。また、 $g(x)$ は重み ω_0 のニューラルネットワーク $g(x, \omega_0)$ で厳密に表現できるとする。

ニューラルネットワーク $f(x, \omega)$ の出力に印加したノイズ n の分散を σ^2 とし、その確率を $p_{\omega, \sigma}(x, y)$ で表記する。これはランダム変数 (x, y) の入力出力空間 $X \times Y$

平成14年12月20日受付

* 大連理工大学電子与信息工程学院, 九州大学大学院システム情報科学研究院訪問助教授

** 電気電子システム工学部門

における密度関数を表す¹²⁾. 具体的に数式で表すと

$$p_{\omega, \sigma}(x, y) \equiv p(x)p(y|x, \omega, \sigma) \\ = p(x) \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{1}{2\sigma^2}(y - f(x, \omega))^2\right) \quad (1)$$

となる. ただし, $p(x)$ は入力変数 x の入力空間上の確率密度関数であり.

真のシステム $y = g(x, \omega_0) + n_0$ の確率を $p_{\omega_0, \sigma_0}(x, y)$ で表記する. ここで, n_0 の分散 σ_0^2 は未知である. 真のシステムとニューラルネットワークの Kullback-Leibler (K-L) 情報距離を次のように定義する¹³⁾.

$$I(p_{\omega_0, \sigma_0}(x, y), p_{\omega, \sigma}(x, y)) \\ = \int p_{\omega_0, \sigma_0}(x, y) \ln \frac{p_{\omega_0, \sigma_0}(x, y)}{p_{\omega, \sigma}(x, y)} dy dx \\ \equiv J(\omega, \sigma) - H_{\omega_0} \quad (2)$$

ここで

$$J(\omega, \sigma) \\ = \int p_{\omega_0, \sigma_0}(x, y) \frac{1}{2\sigma^2} (y - f(x, \omega))^2 dy dx + \ln \sigma \quad (3)$$

であり, $H_{\omega_0} = J(\omega_0, \sigma_0)$ は重み ω および分散 σ とは無関係である. 明らかに, K-L 情報距離は常に非負であり, $\omega = \omega_0$, かつ $\sigma = \sigma_0$ の条件を満たす場合, K-L 情報距離は零となる. このときに限ってニューラルネットワーク $f(x, \omega)$ は正確に真のシステム $g(x)$ を実現する¹⁴⁾.

K-L 距離は二つの確率記述の一致する程度を表す量であり, これをニューラルネットワークの学習を評価する誤差関数にすることができる. この場合, ニューラルネットワーク学習問題は次のような最適化問題に転じる.

$$\min_{\omega, \sigma} \{I(p_{\omega_0, \sigma_0}(x, y), p_{\omega, \sigma}(x, y)), \omega \in W, \sigma \geq 0\} \quad (4)$$

H_{ω_0} と重み ω および分散 σ とは無関係なのでニューラルネットワークの学習問題はさらに以下のように記述できる.

$$\min_{\omega, \sigma} \{J(\omega, \sigma), \omega \in W, \sigma \geq 0\} \quad (5)$$

すなわち, 関数 $J(\omega, \sigma)$ がニューラルネットワークの学習誤差関数となる.

2.2 学習誤差関数の実現

以上の確率公式に基づいて次のようなニューラルネットワークの汎化誤差公式を導出することができる.

$$E_g(\omega) = \int p_{\omega_0, \sigma_0}(x, y) (y - f(x, \omega))^2 dy dx \quad (6)$$

ここで, y はシステムの出力観測値 (ノイズを含む) である. これに対して, 学習誤差関数 $J(\omega, \sigma)$ は次のように書き直せる.

$$J(\omega, \sigma) = \int p_{\omega_0, \sigma_0}(x, y) \frac{(y - f(x, \omega))^2}{2\sigma^2} dy dx + \ln \sigma \\ = \frac{1}{2\sigma^2} E_g(\omega) + \ln \sigma \quad (7)$$

最適化理論により, 最適点に到達したとき, $\sigma \equiv E_g(\omega)$ となり, そのときの J の値は

$$\min_{\omega, \sigma} J(\omega, \sigma) = \frac{1}{2} \min_{\omega} \ln E_g(\omega) + \frac{1}{2} \quad (8)$$

となる. この式は, 関数 $J(\omega, \sigma)$ と汎化誤差 $E_g(\omega)$ の間に一致性が存在することを示している. したがって, $J(\omega, \sigma)$ を学習誤差関数とするときは, 過学習が生じないことが分かる. しかし, 学習誤差関数 $J(\omega, \sigma)$ の表現 (7) において, $E_g(\omega)$ とノイズ n の分散 σ^2 は未知であるから, サンプルデータから, この二つの関数を再構成する必要がある.

まず, サンプルデータ集合の偏差値

$$E_m(\omega) = \frac{1}{m} \sum_{x \in D} (y_i - f(x_i, \omega))^2 \quad (9)$$

を $E_g(\omega)$ の代わりとする. $f(x, \omega) + n$ を用いてサンプル集合 D を近似するから

$$n_i = y_i - f(\omega, x_i) \quad , (x_i, y_i) \in D$$

は $N(0, \sigma^2)$ に従うノイズサンプルである. したがって,

$$E_2(\omega) \\ = \frac{1}{m} \sum_{i=1}^m \{[y_i - f(x_i, \omega)] - \frac{1}{m} \sum_{j=1}^m [y_j - f(x_j, \omega)]\}^2 \\ \approx \sigma^2 \quad (10)$$

がある. $E_m(\omega)$ と $E_2(\omega)$ を式 (7) に代入すると, 学習誤差関数 $J(\omega, \sigma)$ は

$$J(\omega) = \frac{E_m(\omega)}{2E_2(\omega)} + \frac{1}{2} \ln E_2(\omega) \quad (11)$$

となる. このとき, ニューラルネットワークの学習問題は $J(\omega)$ を目標関数とする無制約の最小化問題となる. $J(\omega)$ の最適化については種々の成熟した方法がある. それらの手法により最適化を行った後に, 式 (8) により

ニューラルネットワークの汎化誤差 $E_g(\omega)$ を計算できる.

$$E_g(\omega) = \exp(2 \min\{J(\omega, \sigma) - 1\})$$

$$\approx \exp(2J(\omega) - 1) \quad (12)$$

3. 各学習法による入札価格データに対する学習結果の比較

前節で導出した学習誤差関数 $J(\omega)$ (式(11)) によって従来の BP 計算法を改善し, 改善した方法でニューラルネットワークの入札価格決定支援システムを学習させた. 図 1 に本論文で用いた入札価格決定支援システムのニューラルネットワークモデルの構造を示した. 一つの隠れ層を持つ多入力-出力の階層型ネットワークである. 主成分分析理論に基づき, ネットワークの入力変数を次のように選択した: 市場条件, 競争相手の数, 間接費用率, 現在の仕事量, 労働力, 工事量, 場所, 工期, 工事複雑度¹⁵⁾. ネットワークの出力変数は標準価格に対する入札価格のパーセンテージである. 従来の BP 計算法, モーメント BP 計算法と本論文の改善 BP 計算法でネットワークの学習を行った^{16),17)}. 図 2 と図 3 に三種の方法で計算した学習誤差と汎化誤差の曲線を与えた. また, 表 1 には各種計算法の最終学習結果と汎化誤差を示した.

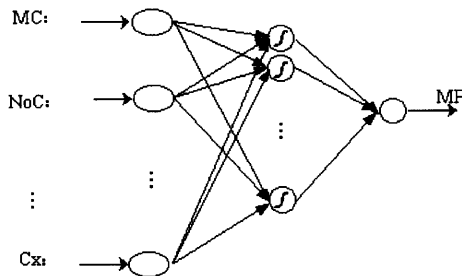


Fig.1 The structure of the neural network.

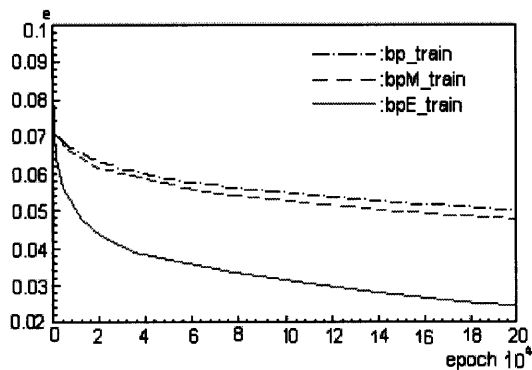


Fig.2 Training error curves of three algorithms.

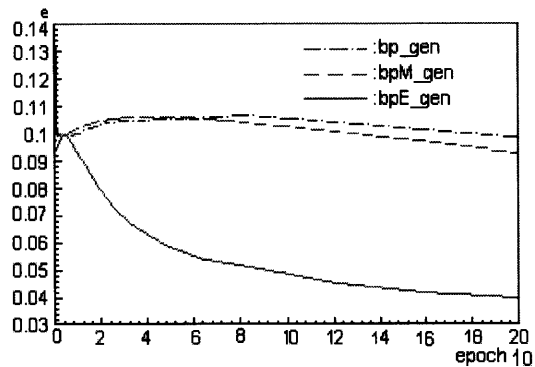


Fig.3 Generalization errors of three algorithms.

三種の計算法による学習誤差曲線 (図 2) から良く分かるように, 従来の BP 計算法とモーメント BP 計算法でネットワークを学習させると, 学習誤差は最初の時点では早く低下するが, その後, 学習回数の増加につれて誤差低下速度は緩くなる. 改善後の学習誤差関数 $J(\omega)$ でネットワークの重みを学習すると, 同じ学習回数でも他の二方法と比べて学習誤差はより小さくなり, 降下の速度もより速い. 三種の計算法の汎化誤差 (図 3) を考察すると, 従来 BP 計算法とモーメント BP 計算法でネットワークを学習する時のシステムの汎化誤差は最初の時点で降下速度は速く, その後学習回数の増加につれて上昇する傾向があるが, 後はゆるいスピードで降下する. 改善 BP 計算法の汎化誤差は最初の瞬間上昇の傾向があるが, その後は学習回数の増加につれて誤差は比較的早い速度で降下する. 最終的に誤差は小さくなる. 学習回数が同じ状態で, システムの最終誤差は従来の BP 計算法およびモーメント BP 計算法のそれより小さい (表 1). したがって, 改善 BP 計算法で入札価格データを学習して構築された入札価格決定支援システムは比較的良好な汎化能力を持つ.

Table 1 Training errors and generalization errors of three algorithms.

	final error	generalization error
Revised BP	0.0242	0.0400
Momentum BP	0.0477	0.9850
General BP	0.0502	0.0928

4. 結 言

実際の入札価格の決定プロセスは毎回異なっている. すなわち毎回の入札状況が過去と完全に同等な可能性はない. 本論文はこの入札価格決定分野における特別な要求に基づいて, ニューラルネットワークの学習計算法の

汎化能力に着目し、ニューラルネットワークの汎化能力に影響を与える各種の要素について解析し、汎化能力を高める方法について理論的な検討をした。従来の BP 計算法、モーメント BP 計算法と改善 BP 計算法の 3 つの方法により入札価格データに対して学習を行い、ニューラルネットワークの入札価格決定支援システムの学習能力と汎化能力を検証し、各種計算法のシミュレーション結果を比較し、提案した改善 BP 計算法が入札価格決定支援システムの汎化能力を高めることができることが明らかになった。

参 考 文 献

- 1) Osama, M., Tarek, H. and Paul, F.: Neural networks as tools in construction. *Journal of Construction Engineering and Management, ASCE*, **117**(4), pp606-625, 1991.
- 2) Jiang X.-J. and Tang H.-W.: System analysis for generalization of MFNN. *System Engineering Theory and Practice*, **8**(1), pp36-40, 2000.
- 3) Dong C. and Guo X.: The research and advances on some focus questions. *Control Theory and Applications*, **17**(5), pp691-696, 2000.
- 4) Zhang N. and Yan P.: Neural network and fuzzy control. Tsinghua University Press, Beijing, 1998.
- 5) Dong C.: Approximation and generalization mechanism of multiplayer feed forward neural network. *Control and Decision*, **13**(supplement), pp413-417, 1998.
- 6) Atiya, A., Ji, C.Y.: How initial conditions affect generalization performance in large networks. *IEEE Trans Neural Networks*, **8**(2), pp448-451, 1997.
- 7) Lin Y. and Han M.: Improvement on Generalization Ability of the ANN-Bidding System with Genetic Algorithm. *Proceedings of International Symposium on Knowledge and Systems Sciences*, pp230-235, Dalian, 2001.
- 8) Funahashi, K.I.: On the approximate realization of continuous mappings by neural networks. *Neural Networks*, **2**(3), pp183-192, 1989.
- 9) Li, H., Shen, L.Y. and Love, P.E.D.: ANN-based mark-up Estimation system with self-explanatory capacities. *Journal of Construction Engineering and Management*, **125**(3), pp185-189, 1999.
- 10) Chua, D.K.H., Kog, Y.C., Loh, P.K. and Jaselskis, E.J.: Model for Construction Budget Performance-Neural Network Approach. *Journal of Construction Engineering and Management*, **123**(3), pp214-222, 1997.
- 11) Li, H.: Neural network models for intelligent support of mark-up estimation, *Int.J. Engrg. Constr. and Arch. Mgmt.*, **3**(1), pp69-82, 1996.
- 12) Chen K.: Probability and mathematical statistics, *Shanghai Scientific Press*, Shanghai, 1989.
- 13) Ljung, L.: System Identification - theory for users, *East China Normal University Press*, Shanghai, 1990.
- 14) Li J. and Han Z.: The learning error function of neural network, *Control and Decision*, **15**(1), pp95-97, 2001.
- 15) Han M., Lin Y. and Qing D.: A factor selection approach in bidding process using neural network. *Journal of Dalian University of Technology*, **42**(1), pp110-114, 2002.
- 16) Shi Y. and Li J.: MATLAB language toolbox-TOOLBOX application instructor. *Northwestern Polytechnical University Press*, Xi'an, 1999.
- 17) Tang H. and Qin X.: Practical methods of optimization, *Dalian University of Technology Press*, 2000, Dalian.

