ZESZYTY NAUKOWE POLITECHNIKI RZESZOWSKIEJ Nr 271
Budownictwo i Inżynieria Środowiska z. 57 (2/10) 2010

Marzena KŁOS
Politechnika Rzeszowska

PROJEKTOWANIE SZTUCZNYCH SIECI NEURONOWYCH W ZAGADNIENIU IDENTYFIKACJI PARAMETRÓW GEOMETRYCZNYCH ŁUKÓW

W pracy zaproponowano dwie metody projektowania sztucznych sieci neuronowych (SSN) do identyfikacji parametrów geometrycznych łuków. Jedna z metod to powszechnie stosowana metoda walidacji krzyżowej, w której poszukuje się minimum funkcji błędu. Druga to nowoczesna metoda MML – maksimum całkowitej wiarygodności, oparta na podejściu bayesowskim. W celu porównania obu metod przeanalizowano sieci kaskadowe, w których wejście stanowiło zawsze sześć podstawowych częstości drgań własnych i w każdym kroku kaskady otrzymane z uczenia sieci parametry geometryczne łuku. Wyjściem zawsze był tylko jeden parametr geometryczny. Otrzymane wyniki potwierdzają skuteczność stosowania metody MML zamiast metody walidacji krzyżowej bezpośrednio na całym zbiorze danych, bez wielokrotnych powtórzeń.

1. Wprowadzenie

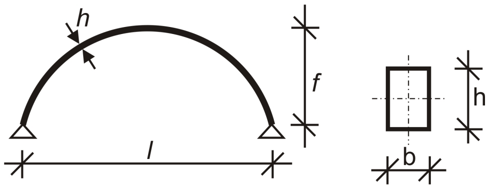
W pracy zaproponowano dwie metody projektowania sztucznych sieci neuronowych (SSN) do identyfikacji parametrów geometrycznych łuków. Projektowanie odnosi się do obliczania optymalnej liczby neuronów H w warstwie ukrytej SSN. Analizowano problem regresji odpowiadającej odwzorowaniu:

$$\{\omega_i\}_{i=1}^{n} \rightarrow \left\{\bar{l}, \bar{f}, \bar{h}\right\} \tag{1}$$

gdzie: ω_i – sześć pierwszych częstości drgań własnych łuku,

$\left\{\bar{l}, \bar{f}, \bar{h}\right\}$ – unormowane wartości parametrów geometrycznych łuku, rys. 1.

Pierwszą zastosowaną metodą jest metoda krzyżowej walidacji. Opiera się ona na obliczaniu wartości minimalnej krzywej walidacji, gdyż krzywa uczenia jest funkcją monotonicznie malejącą parametru H. Drugą zastosowaną metodą projektowania SSN jest metoda maksimum całkowitej wiarygodności, stosowana w podejściu bayesowskim. Kryterium projektowania SSN w tej metodzie jest oparte na obliczeniu maksimum funkcji całkowitej wiarygodności lnCW(H,L), gdzie CW jest prawdopodobieństwem całkowitej wiarygodności z liczebnością zbioru uczącego.



Rys. 1. Łuk dwuprzegubowy

2. Zbiory danych

W celu porównania obu metod projektowano sieci kaskadowe z jednym wyjściem. Uwzględniono sześć przykładów sieci warstwowych z jedną warstwą ukrytą i skalarnym wyjściem o architekturze przedstawionej w tabeli 1.

Tabela 1.

Krok kaskady	sieć	Wejście sieci	Wyjście sieci
1	6-H-1	$\omega_i = \{\overline{\omega}_1, \overline{\omega}_2, ...\overline{\omega}_6\}$	l’
2	7-H-1	ω_i + l’	f’
3	8-H-1	ω_i + l’ + f’	h’
4	9-H-1	ω_i +l’ + f’ + h’	l’’
5	9-H-1	ω_i + l’’+f’+h’	f’’
6	9-H-1	ω_i +l’’+f’’+h’	h’’

gdzie: ω_i – sześć pierwszych częstości drgań własnych unormowanych do przedziału [0.0, 1.0] przy założeniu, że czynnikiem normującym jest ω_{imax} obliczone programem MES ADINA [1]; l, f, h – parametry geometryczne łuku odpowiadające unormowanym do przedziału [0.0, 1.0] wartościom rozpiętości, strzałki i wysokości prostokątnego przekroju poprzecznego łuku, rys. 1.

Przyjęto równomierny podział przedziałów wartości dyskretnych podanych zmiennych geometrycznych: l = [0.5; 0.125, 1.0], gdzie skok wartości wynosi 0.125, a przyjęty czynnik normujący l_{max} = 2.0 m, f = [0.167; 0.167; 1.0], czynnik normujący f_{max} = 0.6 m, h = [0.25; 0.25; 1.0], czynnik normujący h_{max} = = 0.04 m. Obliczenia wykonano programem ADINA, przyjmując jednoosiowe, zakrzywione elementy trójwęzłowe w długości łuku osi odpowiadające kątowi rozwarcia łuku l_{elem} / R = 1DEG. Przyjęto stałe materiałowe jak dla stali konstrukcyjnej, tj. E = 200 GPa, ν = 0.3, ρ = 7800 kg/m^3, R – promień łuku.

Łącznie wygenerowano 120 par wzorców:

$$x = \{\overline{\omega}_1, \overline{\omega}_2, ...\overline{\omega}_6\} \rightarrow t = \{\overline{l}, \overline{f}, \overline{h}\} \tag{2}$$

3. Metoda walidacji krzyżowej

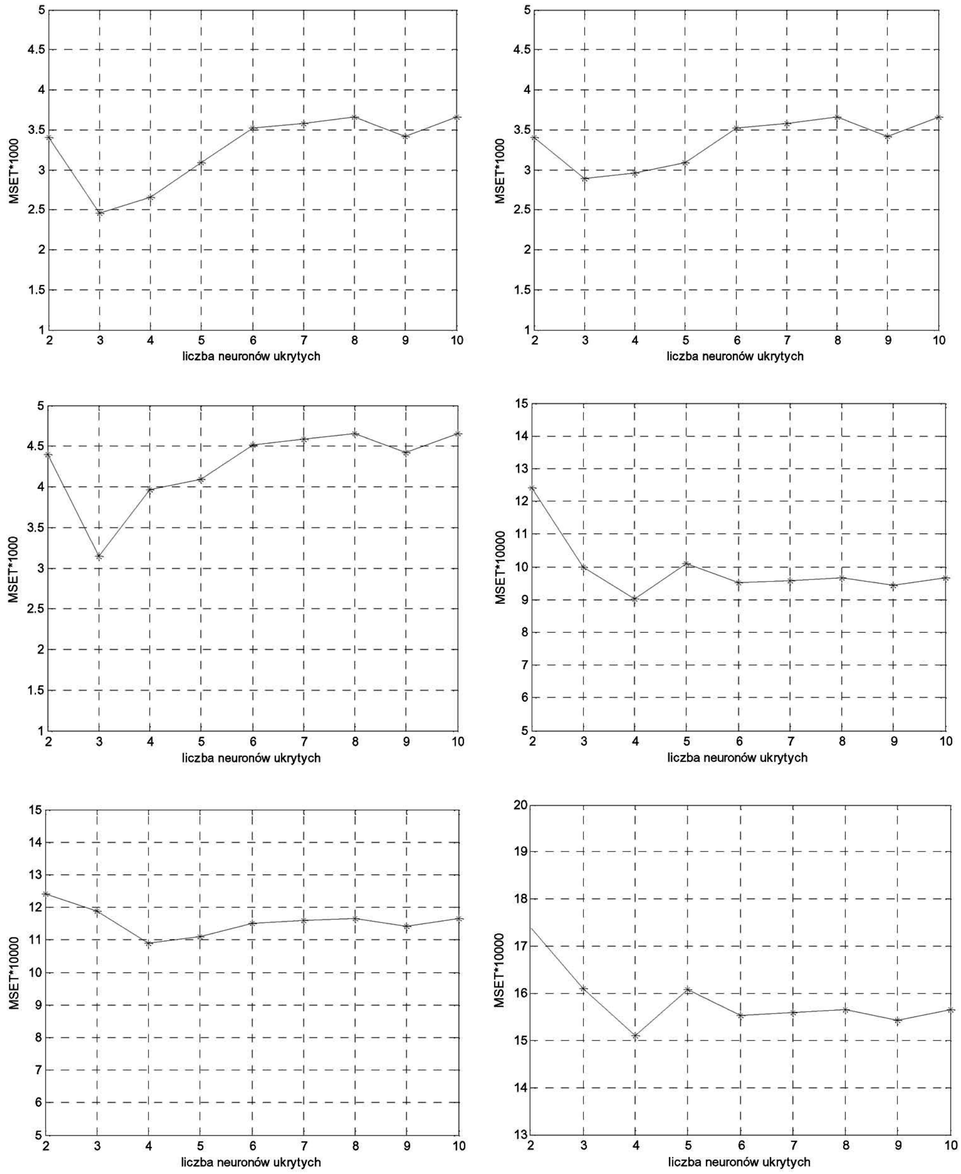
Zbiór wszystkich wzorców rozdzielono losowo na dwa równoliczne zbiory: uczący $L = \{x^p, t^p\}_{p=1}^{l=60}$ i testujący $T = \{x^p, t^p\}_{p=1}^{T=60}$. Dla ustalonej liczby H neuronów w warstwie ukrytej sieć uczono na zbiorze danych uczących L. Nauczoną sieć weryfikuje się za pomocą zbioru testującego T. Wyniki obliczeń neuronowych szacowano za pomocą wzoru MSE (*Mean – Square – Error*):

$$MSE = \frac{1}{V}\sum_{p=1}^{V}(t_i^p - y_i^p) \tag{3}$$

gdzie: V = L, T – odpowiednio liczba wzorców uczących i testujących,
M – liczba wyjść,
t_i^p, y_i^p – składowe wektora wyjścia odpowiednio dla wzorców znanych i obliczanych za pomocą SSN.

Do uczenia nadzorowanego zastosowano metodę Levenberga-Marquardta [3]. Obliczenia wykonano dla liczby epok S = 100. Obliczenia wykonano dla liczby ukrytych neuronów $H \in [12,...10]$. Dla każdej ustalonej wartości H obliczenia powtórzono niezależnie 25 razy dla wylosowanych wartości początkowych parametrów sieci (wagi synaptyczne i biasy). Następnie obliczano wartości średnie błędu MSE z 25-ciu zakończonych iteracji. Obliczenia dla tej samej wartości H powtarzano dla 25-ciu losowań zbioru uczącego. Dla tej samej liczby losowań obliczono wartości średnie błędu uczenia MSEL. Równocześnie, po zakończeniu każdego uczenia obliczano błąd testowania MSET. Projektowanie wykonano z zastosowaniem symulatora MATLAB NN Toolbox (procedura trainlm). Wyniki projektowania sieci przedstawia sześć kolejnych rysunków dla sieci przykładowych z tabeli 1. Począwszy od kroku pierwszego kaskady, do ostatniego – szóstego. Rysunek 2. przedstawia krzywe błędów testowania w zależności od zmieniającego się parametru H.

Okazuje się, że przy zwiększaniu liczby H neuronów ukrytych błąd testowania MSET monotonicznie maleje w kolejnym cyklu kaskady. Dla pierwszych trzech kroków kaskady optymalna sieć to sieć z H = 3, dla kolejnych kroków optymalne jest H = 4.



Rys. 2.

4. Metoda maksimum całkowitej wiarygodności

Standardowo wyniki obliczeń neuronowych szacowano za pomocą wzoru (3).

Przy zastosowaniu regularyzacji Bayesa, po wprowadzeniu dwóch hiperparametrów: α, β, obliczenia neuronowe szacuje się za pomocą wzoru na rozszerzoną funkcję błędu [5]:

$$F = \beta E_D + \alpha E_W \tag{4}$$

gdzie E_W jest sumą błędu średniokwadratowego dla wag.

W podejściu Bayesowskim dane są przyjmowane jako funkcje gęstości dla wag w postaci reguły Bayesa [2, 7]:

$$p(w / D, \alpha, \beta) = \frac{p(D / w, \beta) p(w / \alpha)}{p(D / \alpha, \beta)} \tag{5}$$

gdzie: D – zbiór danych,
w – wektor wag,
p(w/α) – wcześniejsza gęstość (*priori density*),
p(D/w,β – funkcja wiarygodności(*likelihood function*)
p(D/α,β) – współczynnik normalizujący (gwarantuje, że całkowite prawdopodobieństwo jest równe 1). Ten rozkład prawdopodobieństwa jest określany jako całkowita wiarygodność (*marginal likelihood*):

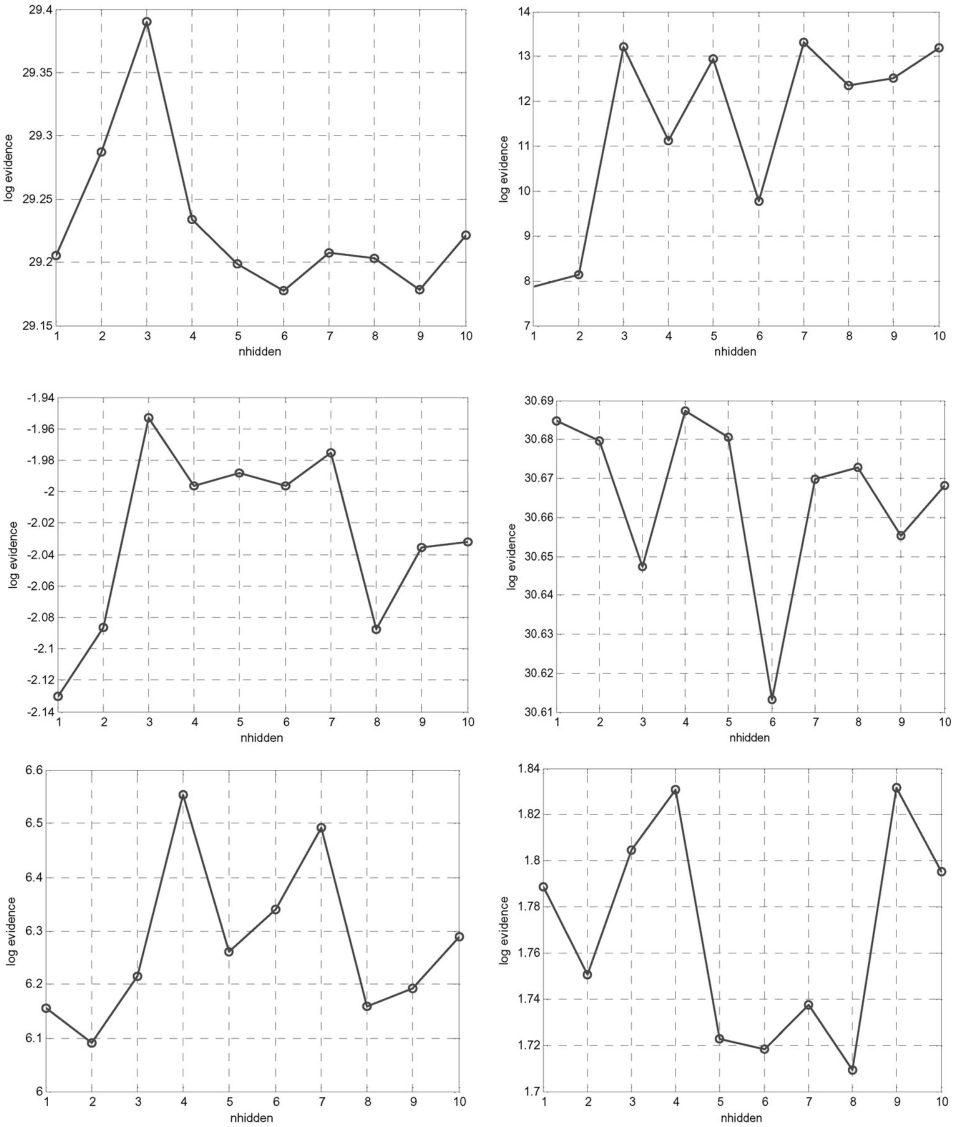
$$p(D / \alpha, \beta) = \int p(D / w, \beta)\, p(w / \alpha) dw \tag{6}$$

Rozwijając wzór (4) na uogólnioną funkcję błędu sieci neuronowej przy rozwiązaniu maximum a posteriori, otrzymanym dla sieci neuronowej uczonej z uwzględnieniem funkcji kary Ew, otrzymuje się:

$$F(w) = \frac{\beta}{2} \sum_{n=1}^{N} \left\{ t^n - y(x^n; w) \right\}^2 + \frac{\alpha}{2} \sum_{i=1}^{W} w_i^2 \tag{7}$$

Po zlogarytmowaniu wyrażenia (7) otrzymuje się wzór na funkcję lnCW, która uwzględnia optymalizację parametru H (liczby neuronów ukrytych). Przy projektowaniu sieci można ograniczyć się tylko do uczenia i całkowicie pominąć testowanie, w szczególności jest to bardzo korzystne przy małym zbiorze uczącym. W pracy projektowanie wykonano z zastosowaniem symulatora NETLAB NN Toolbox (procedura evidence). Zastosowano sieci bayesowskie typu SSN/MAP. Podczas uczenia sieci wprowadzono człon regularyzacyjny, a wartości hiperparametrów ustalano iteracyjnie. Dzięki zastosowaniu wnioskowania bayesowskiego wielkości wyjściowe traktowano jako zmienne probabilistyczne.

W pracy zastosowano metodę MML (*Maximum of Marginal Likelihood*) do optymalizacji sześciu przykładowych sieci kaskadowych w problemie identyfikacji parametrów geometrycznych łuków. Obliczenia wykonano dla liczby epok S = 100 oraz liczby ukrytych neuronów $H \in [2,...10]$. Przyjęto początkowe hiperparametry regularyzacji $\alpha = 0.01$, $\beta = 100$. Na rysunku 3. pokazano krzywe ln-CW(H;L), obliczone dla wszystkich 120 wzorców, po kolei w każdym kroku



Rys. 3.

kaskady. Z wykresów wynika, że optymalna liczba neuronów ukrytych dla pierwszych 3 kroków wynosi H = 3, a dla dalszych optymalne H = 4.

5. Podsumowanie

W pracy przedstawiono kryterium MML (*Maximum of Marginal Likelihood*) oparte na obliczaniu maksymalnej wartości funkcji lnCW w odniesieniu do zagadnienia identyfikacji parametrów geometrycznych płaskich łuków dla znanych wartości podstawowych częstości drgań własnych. W celu przebadania większej liczby sieci neuronowych zastosowano sieć kaskadową. Sprawdzono dwie metody projektowania sieci neuronowych. W obu przypadkach optymalne sieci dla pierwszych trzech kroków kaskady to sieci z H = 3, 6-3-1, 7-3-1, 8-3-1, dla kolejnych kroków h = 4, a optymalna sieć to 9-4-1. Kryterium MML może zastąpić kryterium minimum błędu walidacji krzyżowej (*Cross-Validation*). Zwłaszcza że kryterium to można stosować dla dowolnego zbioru danych, w szczególności tylko dla zbioru uczącego, co ma znaczenie, gdy zbiór jest małoliczny.

Literatura

1. *ADINA System Online Manuale*. Theory and modeling guide, ADINA R& D Inc., Watertown 2001.
2. Dan Foresee F., Hagan M.T.: *Gauss-Newton Aproximation to Bayesian learning.*
3. Haykin S.: *Neural networks. A comprehensive foundation*. 2nd Ed. Upper Saddle River, NJ: Prentice Hall, 1999.
4. Kłos M.: *Neuronowa analiza zagadnień symulacji częstości drgań własnych i identyfikacji parametrów geometrycznych łuków płaskich*. Zeszyty Naukowe Politechniki Rzeszowskiej, nr 245, 2007, 103-113.
5. MacKay D.J.C.: *Bayesian Interpolation*, Neural Computation, vol. 4, 1992, pp. 415-447.
6. *Neural Network Toolbox for Use with Matlab*. User`s Guide, Version 3.0. The Math-Works, Inc. 2006.
7. Waszczyszyn Z., Słoński M.: *Maksimum całkowitej wiarygodności zamiast walidacji krzyżowej w projektowaniu sztucznych sieci neuronowych*. Zeszyty Naukowe Politechniki Rzeszowskiej, nr 243, 2007, 170-182.

PROJECTING SSN OF THE IDENTIFICATION OF THE SHAPE PARAMETERS FOR THE ARCHES

S u m m a r y

In the work two methods were proposed for design of artificial neural networks (ANN) of the identification of the shape parameters for the arches. One of methods is applied universally method of the Cross-Validation in which we seek the minimum of the function of the mistake.

Second is modern method MML Maximum of Marginal Likelihood taken from Bayesian approach. In the paper the design of ANN is related to searching of an optimal value of the number of neurons H in the hidden layer of network. It is illustrated on six numerical examples. In these problems the input vector always composed of the first six eigenfrequencies and made up the plus in every the step the cascade one the shape parameter. The obtained results enable to formulate a conclusion the criterion MLM can be used instead of the cross-validation method. This conclusion if of practical value, since it permits to design ANNs without formulation of a test set of patterns.

Złożono w Oficynie Wydawniczej w maju 2009 r.