Journal of ELECTRICAL ENGINEERING, VOL. 53, NO. 11-12, 2002, 324–327

COMMUNICATIONS

NON–LINEAR CONTROL USING PARAMETER ESTIMATION FROM FORWARD NEURAL MODEL

Anna Jadlovská *

A neural network application for system identification and control of non-linear process is described in this paper. The non-linear identification is mostly using feed-forward neural networks as a useful mathematical tool to build a non-parametric model between the input and output of a real non-linear process. The possibility of on-line estimation of the actual parameters from an off-line trained neural model of the non-linear process using the gain matrix is considered in this paper. This linearization technique is used in the algorithm of on-line tuning of the controller parameters based on a pole placement control design for the non-linear SISO process.

K e y w o r d s: neural model, NARMAX structure, non-linear parameter estimation, instantaneous linearization, non-linear control

1 INTRODUCTION

The purpose of this paper is to show how a feed-forward neural network (Multi Layer Perceptron — MLP) can be used for modeling and control of the non-linear process. When the mathematical model of the process cannot be derived with an analytical method, then the only way is using the relationship between the input and output of the process. Fitting the model from the data is known as an identification of the process. For linear processes this technique is generally well known and described in [7].

For processes which are complex or difficult to model the non-linear identification can use a feed-forward neural network — MLP as a useful mathematical tool to build the non-parametric model between the input and the output of the real non-linear process [2], [3], [5] and [9]. We will consider the possibility of an on-line estimation of the actual parameters from an off-line trained neural model of the non-linear process using the gain matrix introduced later.

We know that a model obtained through linearization around the operating point can be considered valid only in a certain regime around this point. The character of the nonlinearities and the size of the operating range will then determine whether it is sufficient to use a single linear model, or if it is necessary to obtain more by linearization around a large set of operating points.

In this paper a linearization technique called instantaneous linearization will be applied to neural network models. The linearization is carried out at each sampling instant and will be used to perform on-line tuning of the controller parameters based on a pole placement control design in the control structure, the functional behaviour of which is similar to a gain scheduling control, [1] and [4]. The advantage of using instantaneous linearization is that the controller parameters can be changed in response to process changes.

2 NON–LINEAR SYSTEM IDENTIFICATION

In this part we will discuss some basic aspects of non-linear system identification using from among numerous neural networks structures only Multi-Layer Perceptron — MLP (a feed-forward neural network) with respect to model based neural control, where the control law is based upon the neural model.

We will use in this paper a feed-forward neural network MLP with a single hidden layer. This structure is shown in matrix notation in Fig. 1, [8].

Matrix $\mathbf{W}_1$ represents the input weights, matrix $\mathbf{W}_2$ represents the output weights, $\mathbf{F}$ represents a vector function containing the non-linear (tanh) neuron functions. The "1" shown in Fig. 1 together with the last column in $\mathbf{W}_1$ gives the offset in the network. The net input is represented by vector $\mathbf{Z}_{\text{in}}$ and the net output is represented by vector $\hat{\mathbf{Z}}_{\text{out}}$. The mismatch between the desired output $\mathbf{Z}_{\text{out}}$ and $\hat{\mathbf{Z}}_{out}$ is the prediction error $\mathbf{E}$.

The output from MLP can be written as

$$\mathbf{X}_2 = \mathbf{W}_2\,\mathbf{F}\left(\mathbf{W}_1\begin{Bmatrix}\mathbf{Y}_0\\ 1\end{Bmatrix}\right). \tag{1}$$

From a trained MLP (by Back-Propagation Error Algorithm — first-order gradient method, Gauss-Newton algorithm — second-order gradient method) which has m_0 inputs and m_2 outputs a gain matrix $\mathbf{M}$ can be found

* Department of Cybernetics and Artificial Intelligence, Faculty of Electrical Engineering and Informatics, Technical University of Košice, Letná 9, 042 00 Košice, e-mail: Anna.Jadlovska@tuke.sk

ISSN 1335-3632

by differentiating with respect to the input vector of the network.

The gain matrix $\mathbf{M}$ can be calculated from (1)

$$\mathbf{M} = \frac{d\hat{\mathbf{Z}}_{out}}{d\mathbf{Z}_{in}^{\top}} = \frac{d\mathbf{X}_2}{d\mathbf{Y}_0^{\top}} = \frac{d\mathbf{X}_2}{d\mathbf{Y}_1^{\top}} \frac{d\mathbf{Y}_1}{d\mathbf{X}_1^{\top}} \frac{d\mathbf{X}_1}{d\mathbf{Y}_0^{\top}} = \mathbf{W}_2\, \mathbf{F}'(\mathbf{X}_1)\, \mathbf{W}_1^* \tag{2}$$

where $\mathbf{W}_1^* \doteq \mathbf{W}_1$ (excludig last column).

The above mentioned gain matrix $\mathbf{M}$ allows an on-line estimation of the actual model parameters from an off-line trained neural model — MLP of the non-linear process.

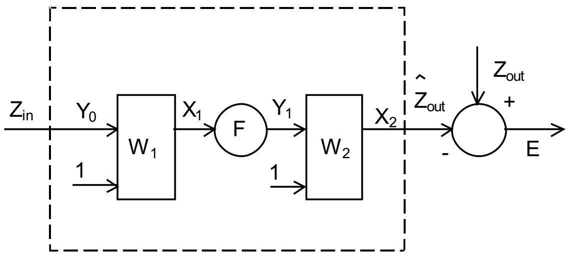

Fig. 1. Matrix block diagram of an MLP.

3 NARMAX MODEL OF NON–LINEAR PROCESS

In this paper we will apply the idea of using input-output parametric non-linear models ARMAX (NARMAX) in non-linear system identification by neural networks, [2] and [6].

We know from linear identification, [7], that the linear ARMAX model is written

$$A(q^{-1})y(k) = B(q^{-1})u(k) + C(q^{-1})e(k) \tag{3}$$

where $y(k)$, $u(k)$ and $e(k)$ are output, input and noise, respectively (for SISO process they are all scalars); A, B and C are polynomials in the backward time shift operator q^{-1}. The noise $e(k)$ is normal, white and with zero mean.

The polynomials A, B and C are

$$\begin{aligned} A(q^{-1}) &= 1 + a_1 q^{-1} + \dots + a_p q^{-p} \\ B(q^{-1}) &= \phantom{1 + {}} b_1 q^{-1} + \dots + b_m q^{-m} \\ C(q^{-1}) &= 1 + c_1 q^{-1} + \dots + c_p q^{-p} \end{aligned} \tag{4}$$

where p and m denote the numbers of delayed outputs and inputs, respectively.

The output $y(k)$ is calculated by a difference equation

$$\begin{aligned} y(k) = {} & -a_1 y(k-1) - \dots - a_p y(k-p) + b_1 u(k-1) + \dots \\ & + b_m u(k-m) + e(k) + c_1 e(k-1) + \dots + c_p e(k-p) \end{aligned} \tag{5}$$

from which the optimal one step predictor is found

$$\begin{aligned} \hat{y}(k) = {} & -a_1 y(k-1) - \dots - a_p y(k-p) + b_1 u(k-1) \\ & + \dots + b_m u(k-m) + c_1 e(k-1) + \dots + c_p e(k-p) \\ y(k) = {} & \hat{y}(k) + e(k)\,. \end{aligned} \tag{6}$$

With inspiration from this linear ARMAX model the non-linear ARMAX (NARMAX) model can be defined as

$$\begin{aligned} \hat{\mathbf{Y}}(k) &= \mathbf{F}\big(\mathbf{Y}(k-1), \dots, \mathbf{Y}(k-p), \mathbf{U}(k-1), \dots, \\ &\qquad \mathbf{U}(k-m), \mathbf{E}(k-1), \dots, \mathbf{E}(k-p), \boldsymbol{\theta}\big) \\ \mathbf{Y}(k) &= \hat{\mathbf{Y}}(k) + \mathbf{E}(k) \end{aligned} \tag{7}$$

where $\mathbf{F}$ is non-linear vector function, $\boldsymbol{\theta}$ represents the parameters and $\mathbf{E}(k)$ is the prediction error. Here p and m denote the number of delayed outputs and inputs.

The neural NARMAX model with input and output vectors

$$\begin{aligned} \mathbf{Z}_{in}(k) &= \big\{\mathbf{Y}(k-1), \dots, \mathbf{Y}(k-p), \mathbf{U}(k-1), \\ &\qquad \dots, \mathbf{U}(k-m), \mathbf{E}(k-1), \dots, \mathbf{E}(k-p)\big\}^{\top} \\ \hat{\mathbf{Z}}_{out}(k) &= \hat{\mathbf{Y}}(k) \end{aligned} \tag{8}$$

is shown in Fig. 2. which is a recurrent network.

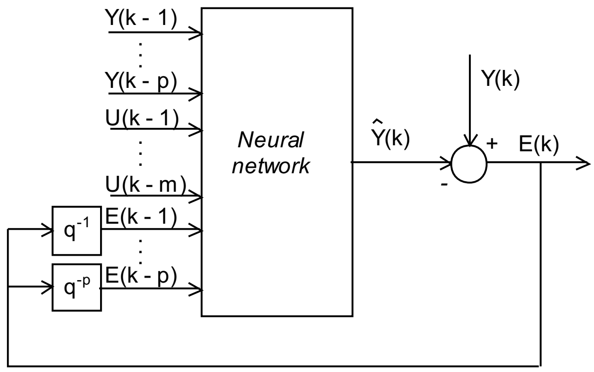

Fig. 2. Input-output neural NARMAX model.

After training the neural network MLP the actual gain matrix $\mathbf{M}(k)$ can be on-line estimated, which is calculated by (2) and for NARMAX model by (9):

$$\begin{aligned} \mathbf{M}(k) &= \frac{d\,\hat{\mathbf{Z}}_{out}(k)}{d\,\mathbf{Z}_{in}^{\top}(k)} = \frac{d\,\hat{\mathbf{Y}}(k)}{d\,\{\mathbf{Y}(k-1)\,\dots\,\mathbf{E}(k-p)\}} = \\ & \big\{-\hat{a}_1(k)\,\dots\,-\hat{a}_p(k)\;\hat{b}_1(k)\,\dots\,\hat{b}_m(k)\;\hat{c}_1(k)\,\dots\,\hat{c}_p(k)\big\} \end{aligned} \tag{9}$$

where $\hat{a}_i(k)$ for $i = 1, \dots, p$, $\hat{b}_i(k)$ for $i = 1, \dots, m$, $\hat{c}_i(k)$ for $i = 1, \dots, p$ are estimated parameters of neural NARMAX model for step k.

Because neural NARMAX model in Fig. 2 contains feedback loops around MLP, we will apply for training this recurrent network a second order Recursive Prediction Error Method (RPEM) using Gauss-Newton search direction, [2], [4] and [7].

4 NON–LINEAR CONTROL USING PARAMETER ESTIMATION

The model of non-linear process and the training method have been considered generally for the multivariable case in parts 2 and 3. Next we will think about the control design for a non-linear SISO process using the neural NARMAX model.

A trained neural NARMAX model representing the model of the non-linear process we use for an on-line estimation of actual process parameters by gain matrix $\mathbf{M}(k)$. This linearization technique called instantaneous linearization allows on-line tuning of the controller parameters using the pole-placement control strategy. This control concept, which is well known from the linear control theory, will be implemented with RST — controller and is described in [1] and [4].

An example of the control structure using the estimation of process parameters from neural NARMAX model which is applied to on-line tuning of the parameters of RST-controller by the pole-placement design is illustrated in Fig. 3.

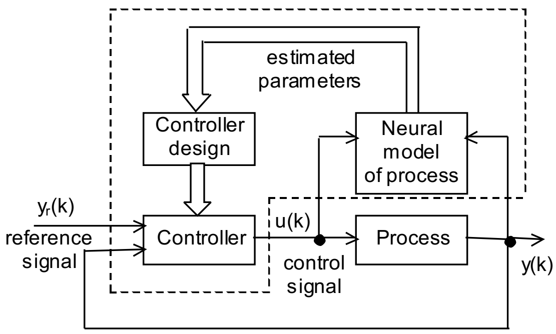

Fig. 3. Control scheme using a trained neural network for parameter estimation.

We can see from Fig. 3 that the structural equivalence is evident with the indirect self-tuning controller. In the self-tuning controller a recursive estimation algorithm is used for identification of a new linear model at each sampling instant. In this paper a linear model is extracted from a more complex non-linear neural network NARMAX model by the gain matrix. The feature of the linearization is that essentially any linear control design can be incorporated in the design block of the controller. This has a number of advantages, such as stabilization of a large class of processes, easy tuning and compensation for disturbances.

5 SIMULATION RESULTS OF NON–LINEAR CONTROL USING POLE–PLACEMENT CONTROL STRATEGY

The idea and results of the estimation process parameters from an off-line trained neural NARMAX model and its using for the tuning parameters of RST — controller designed by the pole-placement strategy (non-linear system control) are presented for non-linear test SISO process, [4]:

$$y(k+1) = \frac{0.95\,y(k) + 0.25\,u(k) + 0.58\,u(k)\,y(k)}{1 + y(k)^2}\,. \quad (10)$$

We consider NARMAX model with 6 inputs ($p = m = 2$) and 8 neurons in the hidden layer. The activation function in the hidden layer is tanh function and in the output layer a linear function is selected. The actual gain matrix $\mathbf{M}(k)$ can be calculated by (2) and the actual values of the estimated $\hat{a}$- and $\hat{b}$-parameters can be obtained from $\mathbf{M}(k)$ by (5). These parameters can be used to calculate the RST-controller parameters.

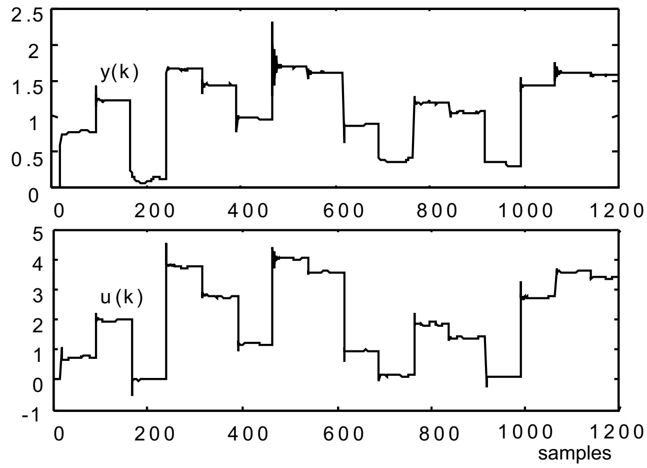

Fig. 4. The input and output training signals.

In Fig. 4 the output and input signals are plotted, which are used to train the neural net in the NARMAX structure by simulation scheme in Fig. 5. The neural NARMAX model is trained with Gauss-Newton algorithm based on RPEM. The model is validated using the time validation test in Fig. 6 [2]. It is clear that the neural model can be accepted by this test.

Presentation of the results of a non-linear control, a control with a RST-controller designed by the pole-placement design using on-line parameter estimation from an off-line trained neural model is illustrated in Fig. 7 while changing one parameter of the non-linear SISO process (10).

The plot in the top panel compares the reference output of the system and the output of the desired closed-loop model. A perfect model-following behavior is achieved, although we can see an oscillating control signal. This example shows the real power of neural modeling using the structure ARMAX known from the theory of linear identification and also the possibility to apply the method of pole-placement known from the linear control theory for the control of non-linear SISO processes in the control structure, which is similar to self-tuning control.

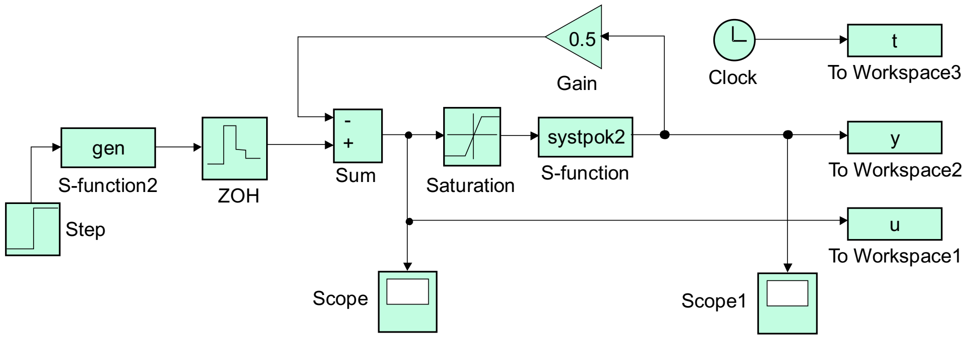

Fig. 5. The simulation scheme for collection of the training data

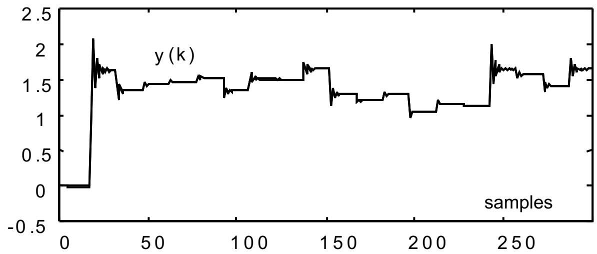

Fig. 6. Validation of the neural model.

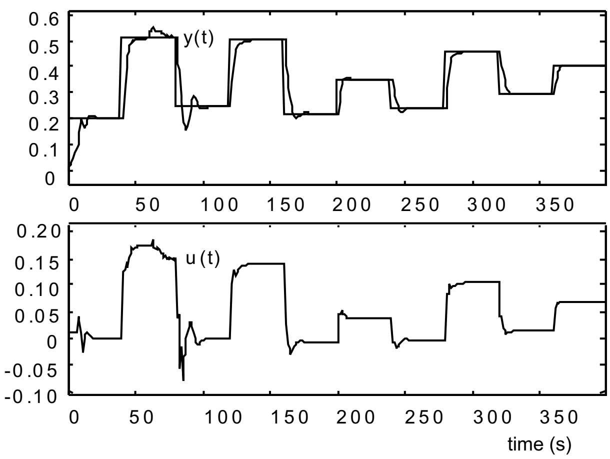

Fig. 7. The RST controller based on actual parameter estimation from neural NARMAX model.

6 CONCLUSION

In this paper neural NARMAX model is trained as a one-step predictor for a non-linear SISO process. After training this NARMAX model can be used in closed control loop for on-line estimation of the process parameters which allow tuning of the controller parameters by the pole-placement method. Practical simulations by language MATLAB/SIMULINK and Neural Toolbox illustrate that this control strategy using the linearization technique by the gain matrix from neural NARMAX model produces excellent performance for the control of non-linear SISO process. But this controller design can be applied only to a non-linear process which does not contain hard non-linearities.

References

[1] ÅSTROM, K., J.—WITTENMARKB. : Computer Controlled Systems, Theory and design, second edition, Prentice-Hall, 1990.

[2] CHEN, S.—BILLINGS, S.—GRANT, P. : Non-Linear System Identification Using Neural Networks, International J. of Control **51** No. 6 (1990), 1191–1214.

[3] JADLOVSKÁ, A. : An Optimal Tracking Neuro-Controller for Nonlinear Dynamic Systems, In: Control System Design, A Proceedings volume from IFAC Conference Bratislava, Slovak Republic, 18-20 June, 2000, Published for IFAC by Pergamon — an Imprint of Elsevier Science, 493-499, 2000, ISBN 00-08 043 546 7.

[4] JADLOVSKÁ, A. : Modeling and Control of Nonlinear Processes Using Neural Networks, PhD thesis, FEI TU Košice, 2001.

[5] JADLOVSKÁ, A.—SARNOVSKÝ, J. : Neural Predictive Control of Non-linear System, In: Proceedings of the 13-th International Scientific-Technical Conference PROCESS CONTROL‘01, Štrbské Pleso, High Tatras, Slovak Republic, CD-ROM, 5 pages, 2001, ISBN 80-227-1542-5.

[6] LEONTARITIS, I.J. : Input-Output Parametric Models for Non-Linear Systems, Part 1 and 2., International J. of Control **41** No. 2 (1985), 303–344.

[7] LJUNG, L. : System Identification, Theory for User, Prentice Hall, 1987.

[8] NAJVÁREK, J. : Matlab and Neural Networks, FEI VUT, Brno, 1996.

[9] NARENDRA, K.S.—LEVIN, A.U. : Identification and Control of Dynamical Systems Using Neural Networks, IEEE Transaction on Neural Networks **1** No. 1 (1990), 4–27.

Received 3 September 2002

Anna Jadlovská (Ing, PhD), was born in Banská Bystrica, Slovakia, on October 29, 1960. She received the Ing degree in Technical Cybernetics from the Faculty of Electrical Engineering of Technical University Košice in 1984 and PhD degree in Automation and Control in 2001 from the same university. Her PhD thesis dealt with modeling and control of non-linear processes using neural networks. Since 1994 she has worked at the Department of Cybernetics and Artificial Intelligence, Faculty of Electrical Engineering and Informatics at TU in Košice as assistant professor. Her main research and teaching activities include the issues of adaptive and optimal control — especially predictive control with constrains for non-linear processes with time-variant parameters using neural networks (Intelligent Control Design). She is the author of articles and contributions published in journals and international conference proceedings. She is also co-author of two monographs.