Keywords: Multiple modeling, Identification, SOM, Time-variant plants, Neural networks

Abstract

A method for Multiple identification based on the self-organizing map neural networks is presented and some of its properties is investigated. Inputs to the NN are instantaneous parameters and so the reference vectors of the networks outputs are the parameter estimation of the multiple models.

1 Introduction

Identification is one of the most important and interested fields of study in the control systems. There are many methods to identify the parameters of the model of a plant such as LSE, Kalman filtering [wal4], Back Propagation Neural Networks [pha3] and others. One of the problems in the estimation of the parameters is to estimate them when the parameters change in a wide domain. In this case the estimated parameters and consequently the identified model, can cover just part of the plant dynamics. So as the plant changes the parameters must be estimated again. But the speed of estimation is a limiting factor. Multiple modeling is the key to solve this problem. In multiple modeling there are more than one estimated models for the plant, so each one valid in a distinct part of the plant parameters space. In each time their outputs compare to the plant output and the model with the nearest output is chosen as the best model [nar2, wal4].

Here we introduce a method to simultaneously estimate the parameters of a number of models for a plant, comparing them and selecting the best one by Self-Organizing Map Neural Networks (SOMNN). SOM is a neural network to self organically map and classify the properties of a collection of objects to a number of classes. The model is a multiple model, that is, the plant is modeled by some linear adaptive models, where each model is valid in a part of the plant parameters space. SOMNN make, adapt and compare them and introduce the best model.

In SOM there are two types of outputs. The first type is the outputs of the output-neurons (activation of the neurons) that show the degree of belonging of the inputs to the classes defined by the neurons. The other type of outputs have not been called outputs but reference vectors or weights of the branches from the input nodes to the output ones. SOM adjust them so that each vector is similar to a class of inputs. Here the similarity means the closure of the input values to the reference vectors. This concept is the key concept in our method. If the inputs are instantaneous parameters estimation, each set of reference vectors is a linear local parameters estimation of the plant model.

The paper is divided into 5 sections. In the next section the identification method and in the 3rd section the mathematical behavior of the method are presented. The 4th section shows the simulation results and some discussion about the results. The last section concludes the paper.

2 Identification algorithm

In this section the method of estimating the parameters is presented. Suppose the nonlinear plant $P$ defined as

$$y(t) = f(y, u, t)$$

This plant is to be modeled by $N$ local linear model (fig 2)

$$y(t) = a_1^i u(t) + \ldots + a_{m}^i u(t - m + 1) +$$

$$+ b_{1}^i y(t - 1) + \ldots + b_{n}^i y(t - n), \quad i = 1, \ldots, N$$

which can be rewritten as

$$y = \varphi \theta^i$$

$$\varphi = (u(t), \ldots, u(t - m + 1), y(t - 1), \ldots, y(t - n))$$

$$\theta^i = (a_1^i, \ldots, a_{m}^i, b_{1}^i, \ldots, b_{n}^i)^T, \quad i = 1, \ldots, N$$

Here there are $n+m$ parameters for each model which must be estimated. If the values of $\varphi, y$ are given for $k \geq t_{md}$ points $(\Phi, Y)$, where $t_{md}=n+m+\max(n,m)-1$. It is possible to compute the $n+m$ instantaneous parameters $\tilde{\theta}$ by solving the
If $k$ is not enough large the solution is not valid for the whole of the plant dynamic space because of the existence of noise, nonlinearity, time-variation parameters and so on. Here an SOM neural networks is established to find the best estimation set for the parameters of the model. The number of inputs of SOM is equal to number of parameters estimation and it has $N$ output neurons, where $N$ is the desired number of models. As plant runs, instantaneous models are found and their parameters are given to the SOM. SOM learn each set of parameters just for one or a few iterations. So step by step the neurons classify to $N$ parameters sets that as will be shown in the next section, approximate the space of the parameters of the local linear models (2) of the plant.

In each iteration, the reference vector of the winner neuron is the best local model at that time. So if for example at time $t=t_1$ the output of $k$th neuron, $\alpha_k$, is the maximum output, i.e.

$$\alpha^* = \alpha_k = \max(\alpha_j(t_1)), \quad j = 1, \ldots, N$$

where $\alpha^*$ stand for the output of the winner neuron, then the parameters set of the best model will be

$$(\hat{\theta}_1(t_1), \hat{\theta}_2(t_1), \ldots, \hat{\theta}_{n+m}(t_1)) = (w_{i1}(t_1), w_{i2}(t_1), \ldots, w_{i(n+m)}(t_1)) = (w_i(t_1))$$

where $w_j$ is the element of the reference vector from the $i$th input to the $j$th neuron.

3 Analyzing the identification algorithm

The method presented in the previous section is based on two theories: Identification and SOMNN. So approach to analysis of this method has two parts; First to show how the identification problem relates to SOM theory if the above algorithm is used and then using the theory of identification and SOMNN to find the properties of algorithm. Here a simple case is studied to avoid the formulation complexity.

Suppose the LSE and a linear noise free plant

$$y(t) = \theta \phi(t)$$

In LSE, it is to minimize the measure

$$J = \sum_{t=t_1}^{t_2} (y(t) - \varphi(t) \hat{\theta})^2$$

if $t_2-t_1=t_{\text{meq}}-1$, then for the solution of the equation

$$Y = \Phi \hat{\theta}$$

that is

$$\hat{\theta} = (\Phi^T \Phi)^{-1} \Phi^T Y$$

the measure is satisfied completely, i.e.

$$J = 0$$

So if the plant is a linear time invariant (LTI) one with input and output degrees $m$, $n$ respectively and in the lack of phenomena like noise, $\hat{\theta}$ given by (10) is the optimal parameters. But in a real case, for a good solution it needs $t_2-t_1>>t_{\text{meq}}$. The measure function $J$ in (8), integrates all the information from $t-t_1$ up to $t=t_2$ by its square error summation and so (9) find the best parameters, by means of LS, for the plant.

On the other hand, the reference vectors of the SOM are functions of the inputs given to the network, so

$$W = f(x_{t_1})$$

where $x_{t_i}$ are the inputs to the SOM up to $t=t_1$ and $W=(w_{i1}, w_{i2}, \ldots, w_{iN})$, where $w_i$ is the reference vector of the $i$th neuron. SOMNN tries to decrease the distance between the inputs and the reference vector of the most activated neuron (and increase it for the neighbors of that neuron), i.e.

$$J_{\text{SOM}} = \text{dist}(w^*, x_{t_1})$$

where $w^*$ is the reference vector of the winner neuron for the input $x(t)$ and the distance function is a norm of $(w - x_{t_1})$.

If $\theta(t)$ is the solution of (4), then replacing $x(t)$ by $\theta(t)$ in SOM gives;

$$J_{\text{SOM}} = \text{dist}(w^*, \theta(t))$$

where $\theta(t)$ is the set of the instantaneous parameters up to $t=t_1$. So since $\theta(t)$ are the parameters which minimize (8) for the last $n+m$ steps, the reference vector $w_i(t)$ is the best approximations for the parameters of the $i$th local model (2) of the plant (1) at time $t$ (eq.(6)).

$$\theta'(t) = w'(t) = w^*(t)$$

So (2), (14) and (15) introduce a mathematical description for the $N$ local models of the plant $P$ which is solved by SOMNN. Therefore the properties of the identification, such as convergence, validity of the solutions and so on, depend on the properties of the SOMNN. Although the theory of SOMNN is
an open problem, but in many cases have been well developed [pha3].

4 Simulation results

Suppose the linear time variant noise free plant

\[ F(s) = \frac{a_1}{1 + a_2 q^{-1} + a_3 q^{-2}} \quad (14) \]

where the parameters change randomly in each units of time approximately as normal random variables, all with mean 0.5 and standard deviation 0.2 which they pick up just in the band \([-0.1, 1.1]\). The input to the plant is a uniform noise in the band \([0, 1]\). The SOMNN is a rectangular lattice, which has 3 input and 64 output neurons and with the learning factor \(0.3 \times \alpha^*\) where \(\alpha^*\) is the highest output neurons activation and the neighborhood of the neurons defined as

\[ \text{neighbor} = \max(1, m^\left(\frac{10 - \hat{\theta}}{10}\right)) \quad (15) \]

where \(m\) is the number of neurons, and \(t\) is the time. A window of data in the last 10 steps is considered to find the instantaneous models. Fig. 4 shows the parameters of the plant and the winner models for the steps \([8500, 9000]\). Winner parameters track the plant ones. However naturally there is some error. The SOM has a rectangular structure. During learning, the values of the reference vectors change but the neighborhood of them is fixed, So the cube is reshaped. Fig.3 is the final SOM structure in the 3-dimensional space.

One may expect that the winner parameters or the reference vectors also be nearly normal. But it is not the case since they are functions of the plant parameters and these functions themselves depend on the learning process of SOMNN. So the pdf of the weights are not necessarily the same as the pdf of the plant parameters but it is a monotonic function of them. Histogram of the 3 types of parameters, that is, plant, instantaneous models and whole of reference vectors in the last 1000 steps are demonstrated respectively in the figures 4 to 6. The reference vectors histogram seems to show up an statistic relation with the plant parameters.

5 Conclusion

In this paper a method of identification based on SOMNN was presented. As any other identification method, this one can also have various versions depend on the definition of SOM functions and parameters such as distance function, learning factors and so on. The best matched model is the winner of the SOMNN when the instantaneous model teach to the networks. So in this method for selecting the best matched model there is no need to compute the output of each models and compare them to the plant one so it reduces the number of computation steps. In addition since in each steps the instantaneous model identified by just \(n+m\) input-output data sets so again the number of computation steps is reduced.

Other versions of this method, their mathematical properties and applications specially in Multiple Control are under investigation by the authors.

References

Fig. 2) Parameters of the plant, best model

Fig. 3) The structure of the SOM networks at the step 9000

Fig. 4) Histogram of the plant parameters through the last 1000 steps
Fig. 5) Histogram of the instantaneous parameters through the last 1000 steps.

Fig. 6) Histogram of the reference vectors through the last 1000 steps.