Model Order Estimations for Noisy Black-box Identifications

Tibor Laczó and László Sragner

One of the principal targets of nonlinear system identification is determining the complexity of an unknown system. If the system is dynamic we have to choose the class of the model and determine the order of the model. The order of the model means how many past inputs and outputs are used in the calculation of the current output. There are several information criteria to qualify a model but they need the setting of the model´s parameters what is the most time consuming step in the identification. We will use the Lipschitz method which estimates the order of the model directly form input-output data [1]. The main idea of the method is that a continuous function´s gradient has a definite maximum and this is a characteristic feature of the function. We calculate the following quotient:

$$L^{(n)} = \left(\prod_{k=1}^{p} \sqrt{n} L^{(n)}(k) \right)^{\frac{1}{p}}$$

Where $L^{(n)}(k)$ is the k-th largest gradient of the function calculated from the known points of the function. This quotient has the property that, if n_0 is the optimal order than $L^{(n_0-1)}$ is much larger than $L^{(n_0)}$ and $L^{(n_0+1)}$ is very close to $L^{(n_0)}$, so we have to find the sharpest breakpoint in the graph of $L^{(n)}$ vs. n. The problem with Lipschitz method is that if noise appears on the inputs and outputs the estimation of the order will be improper or the graph has no sharp breakpoint. To avoid this we use the Errors-in-Variables (EIV) cost function what modifies the input data to eliminate or reduce the noise [2]. The main idea of EIV is that we weight the error of the input and the output according to the reciprocal of the noise variance. For each measurement k the sample variances of the inputs $\hat{\sigma}^2_{u,k}$ and outputs $\hat{\sigma}^2_{y,k}$ the sample covariance matrix $\hat{\sigma}^2_{uy,k}$ and the mean values $\hat{u}_k$ and $\hat{y}_k$ can be calculated or can be estimated [2].

$$C_{EIV} = \frac{M}{N} \sum_{k=1}^{N} \left(\frac{(\hat{y}_k - f_{NN}(u_k, \Theta))^2}{\sigma^2_{y,k}} + \frac{(\hat{u}_k - u_k)^2}{\sigma^2_{u,k}} \right)$$

The greater noise causes that the error counts with smaller weight and the training points with less noise cost more. It can be proved that with noisy data it converges to the true model´s parameters. Because of the modification of the input data, the number of free parameters is increased, therefore the EIV training is very prone to overfitting. To avoid overfitting we first train the network with Backpropagation (BP) and Least Squares (LS) cost function. Our main idea is to combine Lipschitz method and EIV cost function to get more characteristic estimations of the model order. First we estimate the order of the unknown model from its input-output data. The Lipschitz method uses the input-output data as a NARX [3] model. Because of the noise this estimation is probably not characteristic enough. Second we create the model with the estimated order and train it with BP and LS. At a defined error level we stop the training and start using the EIV. The training is controlled with early-stopping to avoid overfitting. If training is finished we get the model and we also get a new input sequence modified by the EIV. On the modified input sequence we use the Lipschitz method again and we compare the results. If the curve sharpened at the estimated order than the order could be correct. We test the method on artificial and real-life industrial problems (LD converter in steel production) [4].

References

[1] X. He, H. Asada, "A new method for identifying orders of input-output models for nonlinear dynamic systems" Proceedings of the American Control Conference, San Francisco, California, June 1993, pp. 2520-2523.

[2] G. Vandersteen, "Identification of Linear and Nonlinear Systems in an Errors-In-Variables Least Squares and Total Least Squares Framework", PhD thesis, Vrije Universiteit Brussel, Belgium, April 1997

[3] K. S. Narendra, K. Pathasarathy, "Identification and Control of Dynamical Systems Using Neural Networks", IEEE Trans. Neural Networks, Vol.1.pp. 4-27,1990

[4] P. Berényi, G. Horváth, B. Pataki, Gy. Strausz, "Hybrid-Neural Modelling of a Complex Industrial Process", IEEE Instrumentation and Measurement Technology Conference, Budapest, 2001. vol. III. pp. 1424-1429.