Monograph:

Tsang, K.M. and Billings, S.A. (1991) A Prediction-Error Estimation Algorithm for the Reconstruction of Linear and Nonlinear Continuous Time Models from Frequency Response Data. Research Report. Acse Report 421 . Dept of Automatic Control and System Engineering. University of Sheffield

A Prediction-Error Estimation Algorithm For The Reconstruction Of Linear and Nonlinear Continuous Time Models From Frequency Response Data

by

K. M. Tsang†

and

S. A. Billings*

† Department of Electrical Engineering
Hong Kong Polytechnic
Hung Hom
Kowloon
Hong Kong

* Department of Control Engineering
University of Sheffield
Sheffield S1 3JD
UK

Research Report No 421

February 1991

Abstract

A prediction error estimation algorithm is derived for the estimation of complex number systems. The algorithm is applied to reconstruct both linear and nonlinear differential equation models from frequency response data. A simulation study is included to illustrate the algorithm.

1. Introduction

It is well known that the estimation of continuous time models from sampled data records can be problematic. The main difficulty associated with this procedure are the numerical errors which can be induced when the derivatives of the input and output signals are computed. An alternative approach is to estimate a discrete time model from the sampled data signals and to use this to evaluate the system frequency response functions. A continuous time description can then be estimated by curve fitting in the frequency domain. The main advantages of this approach are that no integration or differentiation of data is involved and it can be applied to both linear and nonlinear systems.

A modified orthogonal least squares algorithm coupled with an error reduction ratio test was derived in an earlier publication [1] as one possible solution to the reconstruction problem. In the present paper, a new prediction-error estimation algorithm [2] is derived for complex number systems and this is applied to reconstruct linear and nonlinear continuous time models. A simulation study is included to illustrate the algorithm.

2. Linear and nonlinear frequency response functions

2.1 Linear systems

Consider a frequency response function $H(j\omega)$ obtained either by spectral or parametric estimation methods. In order to reconstruct a continuous time model from the frequency response data, a model of the form

$$\hat{H}(j\omega) = \frac{\theta_{n+2}(j\omega)^{m}+\ldots+\theta_{n+m+1}(j\omega)+\theta_{n+m+2}}{\theta_{1}(j\omega)^{n}+\ldots+\theta_{n}(j\omega)+\theta_{n+1}} \tag{1}$$

is fitted to $H(j\omega)$ where θ are the unknown parameters, n and m are the order of the denominator and numerator respectively. In the present study, a prediction error estimation algorithm is derived as an alternative to the orthogonal algorithm introduced in an earlier study [1].

Define the error function

$$\epsilon(j\omega) = H(j\omega) - \hat{H}(j\omega) \tag{2}$$

The objective of the estimation is to minimise the mean square of the magnitude of the error function $\epsilon(j\omega)$ using a prediction error algorithm. Equation (1) is not linear-in-the-parameters, and therefore linear least squares cannot be applied unless the problem is reformulated [1]. However, prediction-error estimation is a general estimation method which can be readily applied to eqns. (1) and (2). The prediction-error method (PEM) produces an estimate of the parameter vector θ by minimising a loss function [2]. The asymptotic properties of the method are very similar to those of the maximum likelihood estimator which can be shown to produce consistent, asymptotically normally distributed and asymptotically efficient estimates with a covariance matrix that reaches the Cramer-Rao bound asymptotically. The original prediction error algorithm is not applicable to complex number systems, and a modified version is derived below.

For a given choice of parameter vector θ, define the loss function

$$J(\theta) = \frac{1}{N}\sum_{i=1}^{N} \epsilon(j\omega)\epsilon^*(j\omega) \tag{3}$$

where the superscript * denotes the complex conjugate and N is the number of points taken in the error function of eqn.(2). The gradient and Hessian of $J(\theta)$ are given by

$$\begin{aligned}
\frac{\partial J(\theta)}{\partial\theta_k} &= \frac{1}{N}\sum_{i=1}^{N}\left[\epsilon(j\omega_i)\frac{\partial\epsilon^*(j\omega_i)}{\partial\theta_k} + \frac{\partial\epsilon(j\omega_i)}{\partial\theta_k}\epsilon^*(j\omega_i)\right] \\
\frac{\partial^2 J(\theta)}{\partial\theta_k\partial\theta_l} &= \frac{1}{N}\sum_{i=1}^{N}\left[\epsilon(j\omega_i)\frac{\partial^2\epsilon^*(j\omega_i)}{\partial\theta_k\theta_l} + \frac{\partial\epsilon(j\omega_i)}{\partial\theta_l}\frac{\partial\epsilon^*(j\omega_i)}{\partial\theta_k} \right. \\
&\qquad \left. + \frac{\partial\epsilon(j\omega_i)}{\partial\theta_k}\frac{\partial\epsilon^*(j\omega_i)}{\partial\theta_l} + \frac{\partial^2\epsilon(j\omega)}{\partial\theta_k\partial\theta_l}\epsilon^*(j\omega_i)\right] \\
&\qquad k,\ l = 1,2,\ldots,n+m+2.
\end{aligned} \tag{4}$$

where the derivatives of the error function $\frac{\partial\epsilon(j\omega)}{\partial\theta_k}$, $k=1,\ldots,n+m+2$ can be obtained by differentiating eqn.(2). Notice that $\frac{\partial\epsilon^*(j\omega)}{\partial\theta_k} = \left(\frac{\partial\epsilon(j\omega)}{\partial\theta_k}\right)^*$. For example, if

$$\hat{H}(j\omega) = \frac{\theta_2}{\theta_1 j\omega + 1}$$

the error function becomes

$$\epsilon(j\omega) = H(j\omega) - \frac{\theta_2}{\theta_1 j\omega + 1}$$

and the error function derivatives are given by

$$\frac{\partial\epsilon(j\omega)}{\partial\theta_1} = \frac{\theta_2 j\omega}{(\theta_1 j\omega+1)^2}\ ; \quad \frac{\partial\epsilon(j\omega)}{\partial\theta_2} = -\frac{1}{\theta_1 j\omega+1}\ ;$$

$$\frac{\partial^2\epsilon(j\omega)}{\partial\theta_1^2} = \frac{2\theta_2\omega^2}{(\theta_1 j\omega+1)^3}\ ; \quad \frac{\partial^2\epsilon(j\omega)}{\partial\theta_1\partial\theta_2} = \frac{j\omega}{(\theta_1 j\omega+1)^2}\ ; \quad \frac{\partial^2\epsilon(j\omega)}{\partial\theta_2^2} = 0$$

The prediction-error estimate is expressed by

$$\theta_i(k+1) = \theta_i(k) - \mu^k H^{-1}\nabla J \tag{5}$$

The minimisation of the loss function $J(\theta)$ can be performed very efficiently using Newton's method. The procedure consists of the following steps:

a.) set $k=0$ and choose an initial value θ^0 of the parameter vector.

b.) Evaluate the gradient vector $\nabla J = \frac{\partial J}{\partial \theta}$ and the Hessian matrix $H = \frac{\partial^2 J}{\partial \theta^2}$ at θ^k.

c.) Calculate the direction vector $d^k = -H^{-1}\nabla J$.

d.) Perform a linear search to find the scalar μ^k such that

$$J(\theta^k + \mu^k d^k) = \min_{\mu} J(\theta^k + \mu d^k)$$

e.) Set $\theta^{k+1} = \theta^k + \mu^k d^k$.

f.) If $J(\theta^k)-J(\theta^{k+1}) <$ a small tolerance, stop the algorithm. Otherwise set $k=k+1$ and go to b.)

2.2 Nonlinear Systems

This algorithm equally applies for the reconstruction of nonlinear continuous time systems. Consider the nonlinear differential equation

$$\begin{aligned} F^l\left[\frac{d^n y(t)}{dt},\ldots,y(t),\frac{d^m u(t)}{dt},\ldots,u(t)\right] = \\ \theta_1\frac{d^n y(t)}{dt} + \ldots + \theta_n y(t) + \theta_{n+1}y(t) + \theta_{n+2}\frac{d^m u(t)}{dt} + \ldots + \theta_{n+m+1}\frac{du(t)}{dt} + \theta_{n+m+2}u(t) \\ + \theta_{1,1}\left(\frac{d^n y(t)}{dt}\right)^2 + \theta_{1,2}\frac{d^n y(t)}{dt}\frac{d^{n-1}y(t)}{dt} + \ldots + \theta_{n+m+2,n+m+2}u^2(t) \\ \vdots \\ + \theta_{1,1,\ldots,1}\left(\frac{d^n y(t)}{dt}\right)^l + \ldots = 0 \end{aligned} \tag{6}$$

where l is the degree of nonlinearity, n is the order of dynamics in the output $y(t)$, m is the order of dynamics in the input $u(t)$ and $F^l[.]$ is a polynomial nonlinear function. The symmetric nonlinear frequency response functions of eqn.(6) can be obtained by probing the system equation with exponential inputs [3,4] and the first, second and third order frequency response functions are given by [1]

$$\hat{H}_1(j\omega) = \frac{-(\theta_{n+2}(j\omega)^m + \ldots + \theta_{n+m+2})}{(\theta_1(j\omega)^n + \ldots + \theta_n(j\omega) + \theta_{n+1})}$$

$$\hat{H}_2(j\omega_1,j\omega_2) = \frac{A}{2!\left(\theta_1(j\omega_1+j\omega_2)^n+\ldots+\theta_n(j\omega_1+j\omega_2)+1\right)}$$

$$A = -\sum_{\substack{\text{all possible } i_1,i_2 \\ i_1\le n+1,\ i_2\le n+1}} \theta_{i_1,i_2} H_1(j\omega_1)H_1(j\omega_2)\left[(j\omega_1)^{n+1-i_1}(j\omega_2)^{n+1-i_2}+(j\omega_1)^{n+1-i_2}(j\omega_2)^{n+1-i_1}\right]$$
$$-\sum_{\substack{\text{all possible } i_1,i_2 \\ i_1\le n+1,\ i_2> n+1}} \theta_{i_1,i_2}\left[(j\omega_1)^{n+1-i_1}(j\omega_2)^{n+m+2-i_2}H_1(j\omega_1)+(j\omega_1)^{n+m+2-i_2}(j\omega_2)^{n+1-i_1}H_1(j\omega_2)\right]$$
$$-\sum_{\substack{\text{all possible } i_1,i_2 \\ i_1> n+1,\ i_2> n+1}} \theta_{i_1,i_2}\left[(j\omega_1)^{n+m+2-i_1}(j\omega_2)^{n+m+2-i_2}+(j\omega_1)^{n+m+2-i_2}(j\omega_2)^{n+m+2-i_1}\right]$$

and

$$\hat{H}_3(j\omega_1,j\omega_2,j\omega_3) = \frac{B}{3!\left\{\theta_1(j\omega_1+j\omega_2+j\omega_3)^n+\ldots+\theta_n(j\omega_1+j\omega_2+j\omega_3)+1\right\}}$$

$$B = -\sum_{\substack{\text{all possible} \\ i_1,i_2,i_3\le n+1}} \theta_{i_1,i_2,i_3} \sum_{\substack{\text{all possible} \\ k_1,k_2,k_3=1,2,3 \\ k_1\ne k_2\ne k_3}} (j\omega_{k_1})^{n+1-i_1}(j\omega_{k_2})^{n+1-i_2}(j\omega_{k_3})^{n+1-i_3}\times H_1(j\omega_{k_1})H_1(j\omega_{k_2})H_1(j\omega_{k_3})$$
$$-\sum_{\substack{\text{all possible} \\ i_1,i_2\le n+1,\ i_3>n+1}} \theta_{i_1,i_2,i_3} \sum_{\substack{\text{all possible} \\ k_1,k_2,k_3=1,2,3 \\ k_1\ne k_2\ne k_3}} (j\omega_{k_1})^{n+1-i_1}(j\omega_{k_2})^{n+1-i_2}(j\omega_{k_3})^{n+m+2-i_3}\times H_1(j\omega_{k_1})H_1(j\omega_{k_2})$$
$$-\sum_{\substack{\text{all possible} \\ i_1\le n+1,\ i_2,i_3>n+1}} \theta_{i_1,i_2,i_3} \sum_{\substack{\text{all possible} \\ k_1,k_2,k_3=1,2,3 \\ k_1\ne k_2\ne k_3}} (j\omega_{k_1})^{n+1-i_1}(j\omega_{k_2})^{n+m+2-i_2}(j\omega_{k_3})^{n+m+2-i_3}H_1(j\omega_{k_1})$$
$$-\sum_{\substack{\text{all possible} \\ i_1,i_2,i_3>n+1}} \theta_{i_1,i_2,i_3} \sum_{\substack{\text{all possible} \\ k_1,k_2,k_3=1,2,3 \\ k_1\ne k_2\ne k_3}} (j\omega_{k_1})^{n+m+2-i_1}(j\omega_{k_2})^{n+m+2-i_2}(j\omega_{k_3})^{n+m+2-i_3}$$

$$-\sum_{\substack{\text{all possible} \\ i_1,i_2\le n+1}} 2!\theta_{i_1,i_2} \sum_{\substack{\text{all possible} \\ k_1,k_2,k_3=1,2,3 \\ k_1\ne k_2\ne k_3}} (j\omega_{k_1}+j\omega_{k_2})^{n+1-i_1}(j\omega_{k_3})^{n+1-i_2}\times H_2(j\omega_{k_1},j\omega_{k_2})H_1(j\omega_{k_3})$$
$$-\sum_{\substack{\text{all possible} \\ i_1\le n+1,\ i_2>n+1}} 2!\theta_{i_1,i_2} \sum_{\substack{\text{all possible} \\ k_1,k_2,k_3=1,2,3 \\ k_1\ne k_2\ne k_3}} (j\omega_{k_1}+j\omega_{k_2})^{n+1-i_1}(j\omega_{k_3})^{n+m+2-i_2}H_2(j\omega_{k_1},j\omega_{k_2})$$

respectively. The n'th order error function can be defined as

$$\epsilon_n(j\omega_1,\ldots,j\omega_n) = H(j\omega_1,\ldots,j\omega_n) - \hat{H}_n(j\omega_1,\ldots,j\omega_n) \tag{7}$$

where $\hat{H}_n(j\omega_1,\ldots,j\omega_n)$ is one of the nonlinear frequency response function extracted from eqn.(6) and $H_n(j\omega_1,\ldots,j\omega_n)$ is the frequency response function to be evaluated. The loss function for an n'th order error function is given by

$$J_n(\theta) = \frac{1}{N}\sum_{i=1}^{N}\epsilon(j\omega_{1_i},\ldots,j\omega_{n_i})\epsilon^*(j\omega_{1_i},\ldots,j\omega_{n_i}) \tag{8}$$

and the optimisation algorithm can thus be applied to eqn.(8) for the estimation of n'th order nonlinear terms.

3. Simulated example

The estimated discrete NARMAX model for a nonlinear circuit described by the differential equation

$$0.2\,\frac{dy(t)}{dt} + y(t) + 0.16y^2(t) = u(t) \tag{9}$$

was given by [1]

$$\begin{aligned} y(k) = {} & 0.1758y(k-1) + 0.0623u(k) + 0.1616u(k-1) \\ & - 0.03839y^2(k-1) + 0.569y(k-2) + 0.03143u(k-2) \end{aligned} \tag{10}$$

400 equally spaced linear frequency response data in the frequency range -5Hz to 5Hz obtained from eqn.(10) were used for the reconstruction of the linear part of the system model. A continuous time model of the form

$$\frac{\theta_2 j\omega + \theta_3}{\theta_1 j\omega + 1}$$

was initially specified for the optimisation process and the result obtained using the new prediction error estimation algorithm was

$$\frac{0.0002j\omega + 1.0009}{0.1936j\omega + 1}$$

with a loss function of

$$J(\theta) = 9.3461\times10^{-7}$$

Notice that θ_2=0.0002 is insignificant compared with θ_3=1.0009. If θ_2 is excluded from the final estimation, the model becomes

$$\frac{0.9998}{0.1932j\omega + 1} \tag{11}$$

and there is no dramatic increase in the loss function which is now given by

$$J(\theta) = 1.8422\times10^{-6}$$

Equation (11) is very similar to the linear part of the original system eqn.(9). For the estimation of the second order nonlinearity, nonlinear terms $y^2(t)$, $y(t)u(t)$ and $u^2(t)$ were specified for the estimation. 800 equally spaced second order frequency response data were obtained from eqn.(10) and used to reconstruct the second order nonlinearities. The prediction error estimates corresponding to the nonlinear terms $y^2(t)$, $y(t)u(t)$ and $u^2(t)$ were 0.1563, -0.0051 and -0.0003 respectively while the loss function was equal to 1.0768×10^{-7}. Since the coefficients corresponding to the nonlinear terms $y(t)u(t)$ and $u^2(t)$ are small, if they were excluded in the final model, the coefficient corresponding to the nonlinear term $y^2(t)$ becomes 0.1508 while the loss function is given as 7.22291×10^{-7}. There is only a slight increase in the loss function. Combining the linear and second order nonlinear estimates gives the final model

$$0.1932\frac{dy(t)}{dt} + y(t) + 0.1508y^2(t) - 0.9998u(t) \qquad (12)$$

which is comparable to the original system eqn.(9). Notice that a distinct advantage of this estimation procedure is that each degree of nonlinearity can be independently reconstructed thus simplifying the procedure.

4. Conclusions

A prediction error estimation algorithm has been derived for the estimation of complex number systems. The application of the algorithm to the reconstruction of continuous time linear and nonlinear models provides a new method of estimating continuous time models based on sampled data records with no numerical approximation to differentiation and integration.

5. References

1. Tsang K.M. and Billings S.A. [1990]: Reconstruction of linear and nonlinear continuous time models from discrete time sampled-data systems, submitted for publication.
2. Goodwin G.C. and Payne R.L. [1977]: Dynamic system identification: Experiment design and data analysis, Academic Press, New York.
3. Billings S.A. and Tsang K.M. [1989]: Spectral analysis for nonlinear systems, Part I: Parametric nonlinear spectral analysis, Mechanical Systems and Signal Processing 3, 319-339.
4. Bedrosian E. and Rice S.O. [1971]: The output properties of Volterra systems (nonlinear system with memory) driven by harmonic and Gaussian input, Proc. IEEE 59, 1688-1707.