Weighted Complex Orthogonal Estimator for Identifying Linear and Nonlinear Continuous Time Models from Generalised Frequency Response Functions

A.K.Swain S.A. Billings

Department of Automatic Control
and Systems Engineering,
University of Sheffield
Post Box -600, S1 3JD
U.K

Research Report No: 601

October, 1995

Weighted Complex Orthogonal Estimator for Identifying Linear and Nonlinear Continuous Time Models from Generalised Frequency Response Functions

A.K.SWAIN and S.A.BILLINGS

Department of Automatic Control and Systems Engineering

University of Sheffield

P.O.Box 600, Mappin Street,S1,3JD,U.K.

Abstract

A new weighted orthogonal least squares algorithm is derived to estimate linear and nonlinear continuous time differential equation models from complex frequency response data. The algorithm combines the properties and advantages of both weighted and orthogonal least squares algorithms. A weighted complex orthogonal estimator, obtained by combining the proposed algorithm with the modified error reduction ratio test provides an effective and robust way of detecting the correct model structure or determining which terms to include in the model and identifying the unknown parameters. Since the estimation procedure does not involve any numerical differentiation of the noisy data, the performace of the estimator under the influence of significant noise is quite satisfactory. The proposed estimator has been successfully applied to a variety of both linear and nonlinear systems.

1 INTRODUCTION

The literature on system identification is dominated by algorithms which fit discrete parametric models from real time domain measurements but little work has been reported when the data are complex. Complex data, though rare in the real world is commonly found in the fields of communication and signal processing, and the frequency domain analysis of a system often involves the manipulation of complex data. System identification in the frequency domain based on complex valued frequency response data has not been widely studied. The current emphasis on control synthesis and design tools hinge upon parametric models which in turn have tended to inhibit parallel developments of parametric frequency domain identification techniques which come under the nonparametric system identification frame work.

But frequency and time domain methods give complimentary perspectives of many important problems in linear system and control theory(Ljung and Glover,1981). Occasionally the two methods have been viewed as rivals, particularly on issues of implementation and application to real systems. It is known that the frequency response function and the power spectral density give good insight into properties of the data and provide information about the system behaviour and properties.

With the advent of modern spectral analysers and high performance data acquisition equipment, it is of practical interest to find techniques which identify linear parametric models from frequency response data (i.e.to synthesise the transfer function).But the determination of a linear transfer function as the ratio of two complex polynomials from the discrete frequency response data is in essence a rational approximation problem and the direct solution requires nonlinear least squares based techniques. Several approaches have been suggested based around a linear least squares framework to try and avoid the nonlinear least squares problem (Levy,1959, Payne,1970, Lawrence and Rogers,1979, Stahl,1984, Whitfield,1986). These techniques have been put into a unified frame work by Whitfield(1987) where it was recommended that a constrained optimisation approach was preferred to the least squares based approaches when the data is corrupted by noise.

When the system is nonlinear higher order or generalised frequency response functions have to be considered. These can be estimated by extending the traditional FFT, windowing and smoothing conceptes to multidimensions (Kim and Power,1988) or by fitting a nonlinear parametric model and then mapping this into frequency domain. The latter approach developed by Billings and Tsang(1989), Peyton Jones and Billings (1989) will be used in the present study. Whatever method of estimating the higher order frequency response functions is used, parameterising these estimates is much more difficult than in linear case and very few authors have considered this problem (Tsang and Billings,1992a, Tsang and Billings,1992b). The main difficulty in the nonlinear case is the enormous combinatorial possibilities of terms which could be included in the model. This is not a major problem if the system is linear because the total number of possible terms is relatively small and an acceptable combination of terms can usually be found without much difficulty. In the nonlinear case the number of possible terms increases rapidly as the degree of nonlinearity and the order of the input output dynamics are increased. But estimating nonlinear differential equation models based on the higher order frequency response functions offers a powerful method of reconstructing continuous time models without computing derivatives and this warrants a detailed study of this approach.

In the present study a new algorithm is derived that combines the features of classical weighted least squares with orthogonal least squares and this is extended to the case where the data can be complex. The algorithm is used to estimate nonlinear continuous time models from frequency response data without computing derivatives. With no *a priori* information, the reformulated error reduction ratio provides a measure of the contribution of each candidate model term to the variance (energy) of the dependent variable. This is used to construct nonlinear differential equation models term by term. The most significant term is selected initially and included in the model, then the next most significant term and so on. The orthogonality property of the estimator coupled with the modified error reduction ratio test therefore provides vital information regarding which terms to include in the model. The weighting property of the estimator reduces the sensitivity of the estimates to the choice of frequency range, the number of data samples and the effects of noise. Examples showing how the new procedure can be used to construct both linear and nonlinear differential equation models from estimated frequency response functions are included.

2 Weighted Complex Orthogonal Estimator

Consider a system which can be modeled as

$$z(j\omega) = \sum_{i=1}^{M} \theta_i p_i(j\omega) + \xi(j\omega) \tag{1}$$

where $\theta_i, i = 1, \ldots\ldots M$ are the real unknown deterministic parameters of the system associated with the complex regressors $p_i(j\omega), i = 1, \ldots. M$. $z(j\omega)$ is a complex dependent variable or the term to regress upon and $\xi(j\omega)$ represents the modeling error.

If 'N' measurements of $z(j\omega)$ and $p_i(j\omega)$ are available at $\omega_i, i = 1, \ldots.. N$ the complex system of equation(1) can be represented in matrix form as

$$Z = P\theta + \Xi \tag{2}$$

where

$$Z = \begin{bmatrix} z(j\omega_1) \\ \vdots \\ z(j\omega_N) \end{bmatrix}, \theta = \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_M \end{bmatrix}, \Xi = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_N \end{bmatrix} \tag{3}$$

and

$$P = \begin{bmatrix} p_1(j\omega_1) & p_2(j\omega_1) & \dots & p_M(j\omega_1) \\ p_1(j\omega_2) & p_2(j\omega_2) & \dots & p_M(j\omega_2) \\ \vdots & \vdots & \vdots & \vdots \\ p_1(j\omega_N) & p_2(j\omega_N) & \dots & p_M(j\omega_N) \end{bmatrix} \tag{4}$$

is an $N \times M$ complex matrix of known regressors. The matrix 'P' will be referred to as the Information Matrix.

However,before any attempt is made to obtain an estimate of 'θ', the complex variables involved in eqn(2) should be partitioned into real and imaginary parts;other wise 'θ' could be complex. Thus eqn(2) is represented as

$$Z = \begin{bmatrix} Z_R \\ -- \\ Z_I \end{bmatrix} = \begin{bmatrix} P_R \\ -- \\ P_I \end{bmatrix} \theta + \begin{bmatrix} \Xi_R \\ -- \\ \Xi_I \end{bmatrix} \tag{5}$$

where the subscript'R' denotes the real part and 'I' denotes the imaginary part of the component, $Z = 2N \times 1$ real vector, $P = 2N \times M$ real constant information matrix, $\theta = M \times 1$ real parameter vector and $\Xi = 2N \times 1$ real vector.

In a general algebraic context eqn(5) is an overdetermined set of equations for the unknowns $\theta_1, \theta_2, ...\theta_M$ and thus has a unique least squares solution provided the rank of the information matrix 'P' is M. The weighted least squares solution is given by (Deutsch,1965)

$$\hat{\theta} = (P^T Q P)^{-1} P^T Q Z \tag{6}$$

When the weighting matrix 'Q' is a unit diagonal,eqn.(6) gives the standard least squares solution of the parameter vector θ.

By reformulating the problem such that each parameter in 'θ' can be estimated independently using an orthogonal algorithm,considerable advantages can be achieved. This algorithm is derived in Appendix-A.

The implementation procedure for the weighted complex orthogonal estimator using forward regression is summarised below.

Step-1:-

- Define the model size M and consider all the regressors $p_i(\omega), i = 1, ...M$ as possible candidates to qualify for $w_1(\omega)$; the first term of the auxiliary model.

- For i=1,....M,calculate

$$
\begin{aligned}
w_1^{(i)}(\omega) &= p_i(\omega) \\
\hat{g}_1{}^{(i)} &= \frac{\left\langle Z(\omega), v_1^{(i)} \right\rangle}{\left\langle v_1^{(i)}, w_1^{(i)} \right\rangle}, \quad where \quad v_1^{(i)} = w_1^{T(i)} Q \\
ERR_1^{(i)} &= \frac{(g_1^{(i)})^2 \left\langle v_1^{(i)}(\omega), w_1^{(i)}(\omega) \right\rangle}{\langle (Z(\omega)Q), Z(\omega) \rangle}
\end{aligned}
\tag{7}
$$

- Find the maximum of $ERR_1^{(i)}$,for instance if $ERR_1^{(j)} = max\{ERR_1^{(i)}, 1 \leq i \leq M\}$, then the first term selected should be the jth term of eqn.(1) i.e. $w_1(\omega) = p_j(\omega)$

<u>Step-2</u>:-

- Consider all the $p_i(\omega)$,i=1,...M, $i \neq j$ as possible candidatesfor $w_2(\omega)$. For i=1,....M and $i \neq j$,calculate

$$
\begin{aligned}
&w_2^{(i)}(\omega) = p_i(\omega) - \alpha_{12}^{(i)} w_1(\omega) \\
&\hat{g}_2{}^{(i)} = \frac{\left\langle Z(\omega), v_2^{(i)} \right\rangle}{\left\langle v_2^{(i)}, w_2^{(i)} \right\rangle}, where \quad v_2^{(i)} = w_2^{T(i)} Q \\
&ERR_2^{(i)} = \frac{(g_2^{(i)})^2 \left\langle v_2^{(i)}(\omega), w_2^{(i)}(\omega) \right\rangle}{\langle (Z(\omega)Q), Z(\omega) \rangle} \\
&where \quad \alpha_{12}^{(i)} = \frac{\langle v_1(\omega), p_i(\omega) \rangle}{\langle v_1(\omega), w_1(\omega) \rangle}
\end{aligned}
\tag{8}
$$

- Find the maximum of $ERR_2^{(i)}$,say if $ERR_2^{(k)} = max\{ERR_2^{(i)}, 1 \leq i \leq M, i \neq j\}$,then the second term becomes $w_2(\omega) = w_2^{(k)} = p_k(\omega) - \alpha_{12} w_1(\omega)$

- Continue the process of selecting the terms until the sum of the error resuction ratio (ERR) reaches a value close to 100%

- Recover the unknown parameters of the original system using eqn.(A.20).

3 Model Validation

Model validation is a crucial issue in black box identification where model validity tests are generally used to attach a quality tag to a model. Note however that,the correctness of a model can only be definitely proved if the real process is known, in which case we do not need identification. Hence it follows that we can only try to show that the model is invalid. Failure to invalidate the model is then used to infer that the model is correct. From this perspective,the term model validation can therefore be interpreted as model invalidation.

In time domain identification,the measured data are assumed to be generated by some finite dimensional system corrupted by noise. If the model order is increased above the true order,the model will start to fit to the noise. Several techniques are available to validate the model and the correlation based techniques initially proposed by Billings and Voon (1986) and improved by Billings and Zhu (1993) provide simple and effective ways of testing the validity of the model.

Cross validation is another way of validating the model where a given data set is divided into two disjoint sets; one set is used for identification and the other for model validation. The model validation tests based on correlation techniques are applied to time domain models and can not be used in frequency domain identification. In the frequency domain, cross validation is an efficient means of validating the model. This is easily performed by dividing the frequency response measurements into two disjoint sets; an estimation set and a test or validation set. A natural division is to take every other frequency point as the estimation data and the rest as the test data. A model is then estimated using only the estimation data and the quality of the model is assessed by comparing the estimated model with the test data set.

4 Identification of Continuous Time Models from Frequency Response Data

Before the problem of estimating the parameters of a continuous time model of a nonlinear system can be formulated, the concept of Generalised Frequency Response Functions (GFRF) must be introduced.

It is well known that the input-output behaviour of a wide class of nonlinear systems can be represented by the Volterra series

$$y(t) = \sum_{n=1}^{N} y_n(t) \tag{9}$$

where $y_n(t)$,the nth order output of the system is given by

$$y_n(t) = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} h_n(\tau_1, ...\tau_n) \prod_{i=1}^{n} u(t - \tau_i) d\tau_i, n > 0; \tag{10}$$

and $h_n(\tau_1, ...\tau_n)$ is known as the nth order Volterra Kernel or generalised impulse response function of order n. The multi-dimensional ($n > 1$) Fourier Transform of the 'nth' order impulse response yields the nth order transfer function or the Generalised Frequency Response Function (GFRF)

$$H_n(j\omega_1, ...j\omega_n) = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} h_n(\tau_1, ..., \tau_n) exp(-j(\omega_1\tau_1 + ...\omega_n\tau_n)) d\tau_1....d\tau_n \tag{11}$$

Note that when $n = 1$, the above expression gives the linear transfer function. Since the objective is to determine the structure and estimate the parameters of a nonlinear differential equation model from generalised frequency response functions, the first step is to compute the generalised frequency response functions and show how this problem can be solved using the weighted orthogonal estimator.

4.1 Problem Formulation

Consider a system whose dynamics are described by the nonlinear differential equation

$$\sum_{m=1}^{N_l}\sum_{p=0}^{m}\sum_{l_1,l_{p+q}=0}^{L} c_{p,q}(l_1,..l_{p+q})\prod_{i=1}^{p} D^{l_i}y(t)\prod_{i=p}^{p+q} D^{l_i}u(t) = 0 \tag{12}$$

where 'L' is the maximum order of differential; $p+q=m$, and $m=1,....N_l$ corresponds to various orders of nonlinearity. The operator 'D'is the differential operator. Once 'm' takes a specific number,all the mth-order terms, each of which contains a pth-order factor in '$D^{l_i}y(t)$' and qth-order factor in '$D^{l_i}u(t)$', subject to $p+q=m$, are included and each term is multiplied by a coefficient $c_{p,q}(l_1,...l_{p+q})$ while the multiple summation over the $l_i(l_i=0...L)$ generates all the possible permutations of differentiations. For example, the general linear differential equation is included in the above form by setting the order of nonlinearity 'm' as 1 to give:

$$\sum_{l_1=0}^{L} c_{1,0}(l_1)D^{l_1}y(t) + \sum_{l_1=0}^{L} c_{0,1}(l_1)D^{l_1}u(t) = 0. \tag{13}$$

As an example the differential equation

$$\frac{d^2y}{dt^2} + a_1\frac{dy}{dt} + a_2y + b_1\frac{du}{dt} + b_2u + c_1y^2 + d_1u^2 + c_2y^3 + d_2(\frac{d^2y}{dt^2})^2u = 0 \tag{14}$$

would be represented by the model of eqn.(12) with the following definitions
$c_{1,0}(0)=a_2$, $c_{1,0}(1)=a_1$, $c_{1,0}(2)=1.0$, $c_{0,1}(0)=b_2$, $c_{0,1}(1)=b_1$, $c_{2,0}(0,0)=c_1$, $c_{0,2}(0,0)=d_1$, $c_{3,0}(0,0,0)=c_2$, $c_{2,1}(2,2,0)=d_2$

The frequency domain equivalent of eqn(12) is based on the generalised or higher order frequency response functions which are given by mapping the time-domain representation into the frequency domain (Billings and Peyton-Jones,1990)

$$-\left[\sum_{l_1=0}^{L} c_{1,0}(l_1)(j\omega_1+....j\omega_n)^{l_1}\right] H_n^{asym}(j\omega_1,...j\omega_n)$$

$$= \sum_{l_1,l_n=0}^{L} c_{0,n}(l_1,...l_n)(j\omega_1)^{l_1}.....(j\omega_n)^{l_n}$$

$$+\sum_{q=1}^{n-1}\sum_{p=1}^{n-q}\sum_{l_1,l_n=0}^{L} c_{p,q}(l_1,...l_{p+q})(j\omega_{n-q+1})^{l_{p+1}}...(j\omega_n)^{l_{p+q}}H_{n-q,p}(j\omega_1,..j\omega_{n-q})$$

$$+\sum_{p=2}^{n}\sum_{l_1,l_p=0}^{L} c_{p,0}(l_1,....l_p)H_{n,p}(j\omega_1,...j\omega_n). \qquad (15)$$

where the recursive relation is given by

$$H^{asym}_{n,p}(.) = \sum_{i=1}^{n-p+1} H_i(j\omega_1,...j\omega_i)H_{n-i,p-1}(j\omega_{i+1},..j\omega_n)(j\omega_1+...+j\omega_i)^{l_p}. \qquad (16)$$

The recursion finishes with $p = 1$ and $H_{n,1}(j\omega_1,....j\omega_n)$ has the property

$$H_{n,1}(j\omega_1,....j\omega_n) = H_n(j\omega_1,...j\omega_n)(j\omega_1+....+j\omega_n)^{l_1}. \qquad (17)$$

Eqn.(15-17) for the GFRF will be used to derive the regression equations for estimating the unknown parameters $c_{p,q}(.)$ in the continuous time model of eqn(12).

Note that, the computation of the n-th order frequency response functions is a recursive procedure where each lower order of GFRF contains no effects from higher order terms. This offers a distinct advantage since the parameters corresponding to different nonlinearities or terms can be estimated independently ,begining with first order and then building up to include the nonlinear terms.

4.2 Estimation of Linear Terms ; Synthesis of Linear Transfer Functions

The first order frequency response is only related to the linear input-output terms. Setting $n = 1$ in eqn(15) gives

$$-\left[\sum_{l_1=0}^{L} c_{1,0}(l_1)(j\omega_1)^{l_1}\right] H_1(j\omega_1) = \sum_{l_1=0}^{L} c_{0,1}(l_1)(j\omega_1)^{l_1}. \qquad (18)$$

Notice that multiplying by a constant on both sides has no effect on the form of the equation except all the parameters will be changed. This suggests that all the parameters can not be uniquely estimated. Hence it is necessary to fix a parameter before the estimation begins. Without loss of generality it is assumed that the parameter corresponding to the linear output term $c_{1,0}(0)$ is unity. Moving all other terms to the right hand side gives

$$-H_1(j\omega_1) = \sum_{l_1=1,l_1\neq 0}^{L} c_{1,0}(l_1)(j\omega_1)^{l_1}H_1(j\omega_1) + \sum_{l_1=0}^{L} c_{0,1}(l_1)(j\omega_1)^{l_1}. \qquad (19)$$

which is a linear in the parameters expression. For each value of frequency 'ω' two equations are obtained; one corresponding to the real part and the other to the imaginary part.

However,for clarity, compare eqn.(19) with eqn.(1) and relate the parameters of both the equations.

So that from equation(1)

$$z(j\omega) = \theta_1 p_1(j\omega) + \theta_2 p_2(j\omega) + \ldots + \theta_M p_M(j\omega) \tag{20}$$

where $M = 2L+1$ and

$z(j\omega) = -H_1(j\omega_1)$

$\theta_1 = c_{1,0}(1), \qquad p_1(j\omega) = (j\omega_1)H_1(j\omega_1)$

$\theta_2 = c_{1,0}(2), \qquad p_2(j\omega) = (j\omega_1)^2 H_1(j\omega_1)$

$\vdots$

$\theta_L = c_{1,0}(L), \qquad p_L(j\omega) = (j\omega_1)^L H_1(j\omega_1)$

$\theta_{L+1} = c_{0,1}(0), \qquad p_{L+1}(j\omega) = p_M(j\omega) = (j\omega_1)^0 = 1$

$\theta_{L+2} = c_{0,1}(1), \qquad p_{L+2}(j\omega) = (j\omega_1)$

$\theta_{L+3} = c_{0,1}(2), \qquad p_{L+3}(j\omega) = (j\omega_1)^2$

$\vdots$

$\theta_{2L+1} = c_{0,1}(L), \qquad p_{2L+1}(j\omega) = p_M(j\omega) = (j\omega_1)^L$

The proposed estimation algorithm can now be applied to eqn.(19) to identify the unknown parameters by replacing the frequency response function $H_1(j\omega_1)$ by estimates of this function.

4.3 Estimation of Second Order Nonlinearities

To estimate the parameters associated with second order nonlinear terms, set $n = 2$ in eqn.(15) so that

$$\begin{aligned} -\left[\sum_{l_1=0}^{L} c_{1,0}(l_1)(j\omega_1 + j\omega_2)^{l_1}\right] H_2^{asym}(j\omega_1, j\omega_2) = & \sum_{l_1,l_2=0}^{L} c_{0,2}(l_1,l_2)(j\omega_1)^{l_1}(j\omega_2)^{l_2} \\ & + \sum_{l_1,l_2=0}^{L} c_{1,1}(l_1,l_2)(j\omega_2)^{l_2} H_{1,1}(j\omega_1) \\ & + \sum_{l_1,l_2=0}^{L} c_{2,0}(l_1,l_2) H_{2,2}(j\omega_1, j\omega_2). \end{aligned} \tag{21}$$

By the recursive relation

$$H_{1,1}(j\omega_1) = H_1(j\omega_1)(j\omega_1)^{l_1}. \tag{22}$$

$$\begin{aligned} H_{2,2}^{asym}(j\omega_1, j\omega_2) &= H_1(j\omega_1)H_{1,1}(j\omega_2)(j\omega_1)^{l_2} \\ &= H_1(j\omega_1)H_1(j\omega_2)(j\omega_2)^{l_1}(j\omega_1)^{l_2}. \end{aligned} \tag{23}$$

So that finally a linear regression equation is obtained as

$$
\begin{aligned}
-\left[\sum_{l_1=0}^{L} c_{1,0}(l_1)(j\omega_1 + j\omega_2)^{l_1}\right] H_2^{asym}(j\omega_1, j\omega_2) = \sum_{l_1,l_2=0}^{L} c_{0,2}(l_1,l_2)(j\omega_1)^{l_1}(j\omega_2)^{l_2} \\
+ \sum_{l_1,l_2=0}^{L} c_{1,1}(l_1,l_2)(j\omega_2)^{l_2} H_1(j\omega_1)(j\omega_1)^{l_1} \\
+ \sum_{l_1,l_2=0}^{L} c_{2,0}(l_1,l_2) H_1(j\omega_1) H_1(j\omega_2)(j\omega_1)^{l_2}(j\omega_2)^{l_1}.
\end{aligned} \tag{24}
$$

Notice that the parameters $c_{1,0}(l_1), l_1 = 0, ...L$ on the left hand side have been estimated as linear terms initially (section-4.2) while all the parameters on the right hand side of the equation can be estimated from eqn(24) by replacing the first and second order frequency response functions by their estimates and applying the new estimator.

4.4 Estimation of Higher Order Nonlinearities

When dealing with higher order terms (n=3,4,..),some parameters which have been obtained in previous stages (for orders less than the 'nth') appear on the right hand side of the equation. These are associated with lower order pure output and cross-product terms but not pure input terms.Moving these terms to the left hand side gives

$$
\begin{aligned}
-\left[\sum_{l_1=0}^{L} c_{1,0}(l_1)(j\omega_1 + ...j\omega_n)^{l_1}\right] H_n^{asym}(j\omega_1, ...j\omega_n) \\
-\sum_{q=1}^{n-2}\sum_{p=1}^{n-q-1}\sum_{l_1,l_n=0}^{L} c_{p,q}(l_1, ...l_{p+q})(j\omega_{n-q+1})^{l_{p+1}}...(j\omega_n)^{l_{p+q}} H_{n-q,p}(j\omega_1, ..j\omega_{n-q}) \\
-\sum_{p=2}^{n-1}\sum_{l_1,l_p=0}^{L} c_{p,0}(l_1, ...l_p) H_{n,p}(j\omega_1, ...j\omega_n) \\
= \sum_{l_1,l_n=0}^{L} c_{0,n}(l_1, ...l_n)(j\omega_1)^{l_1}...(j\omega_n)^{l_n} \\
+\sum_{q=1}^{n-1}\sum_{l_1,l_n=0}^{L} c_{p,q}(l_1, ...l_n)(j\omega_{n-q+1})^{l_{p+1}}..(j\omega_n)^{l_{p+q}} H_{n-q,n-q}(j\omega_1, ..j\omega_{n-q}) \\
+ \sum_{l_1,l_n=0}^{L} c_{n,0}(l_1, ...l_p) H_{n,n}(j\omega_1, ..j\omega_n).
\end{aligned} \tag{25}
$$

and the recursive relation

$$
H_{n,p}^{asym}(.) = \sum_{i=1}^{n-p+1} H_i(j\omega_1, ..j\omega_i) H_{n-i,p-1}(j\omega_{i+1}, ..j\omega_n)(j\omega_1 + ..j\omega_i)^{l_p}. \tag{26}
$$

Equation.(21-26) holds for asymmetric frequency response data with a particular frequency permutation and is not necessarily unique. Hence it is essential to use symmetric frequency response data during estimation. For a given asymmetric GFRF $H_n(j\omega_1,...j\omega_n)$, the symmetrised GFRF is obtained by summing the asymmetric function over all possible permutations of the arguments and divided by the number to give

$$H_n^{sym}(j\omega_1,...j\omega_n) = \frac{1}{n!} \sum_{allpermutationsof\,\omega_1,..\omega_n} H_n^{asym}(j\omega_1,...j\omega_n) \tag{27}$$

For example, in the second order case, the symmetric version of eqn.(24) can be written as

$$\begin{aligned} -\left[\sum_{l_1=0}^{L} c_{1,0}(l_1)(j\omega_1+j\omega_2)^{l_1}\right] H_2^{sym}(j\omega_1,j\omega_2) = & \sum_{l_1,l_2=0}^{L} c_{0,2}(l_1,l_2)\left[(j\omega_1)^{l_1}(j\omega_2)^{l_2}+(j\omega_2)^{l_1}(j\omega_1)^{l_2}\right]/2 \\ & + \sum_{l_1,l_2=0}^{L} c_{1,1}(l_1,l_2)\left[(j\omega_2)^{l_2}H_1(j\omega_1)(j\omega_1)^{l_1}+(j\omega_1)^{l_2}H_1(j\omega_2)(j\omega_2)^{l_1}\right]/2 \\ & + \sum_{l_1,l_2=0}^{L} c_{2,0}(l_1,l_2)H_1(j\omega_1)H_1(j\omega_2)\left[(j\omega_1)^{l_2}(j\omega_2)^{l_1}+(j\omega_2)^{l_1}(j\omega_1)^{l_2}\right]/2 \end{aligned} \tag{28}$$

For higher order nonlinearities, the symmetric version of eqn.(25-26) can be implemented accordingly and will be used in the estimation of continuous time models.

Note that since the higher order frequency response functions do not affect the lower order frequency response functions, all the parameters in the left hand side of eqn(25) that correspond to the coefficients of nonlinear differential terms of order less than 'n', have been estimated in previous stages (section-4.2,4.3), so that the unknown parameters associated with the nth-order nonlinearity can be estimated . The whole estimation procedure is therefore progressive starting from first order terms and building to higher order terms. The estimation results obtained in the previous stages must be used in later estimation but obviously not vica versa.

Before the complex orthogonal estimator is applied it is appropriate to address certain issues associated with the least squares estimator and provide a set of guide lines that should be followed during the implementation phase.

4.5 Guide Lines for Implementing the Estimator

In this section an intuitive approach is followed to explain the factors that affect the performance of the estimator by drawing analogies with the time domain algorithm.

Note that the proposed Weighted Complex Orthogonal Estimator is derived by combining the conventional weighted least squares algorithm with the orthogonal least square(OLS) estimator of Billings et al(1988) and extending this to the complex case. In the past few years, properties of the original OLS estimator and later versions such as the Fast Orthogonal Estimator(Billings et al,1988, Billings et al,1989, Zhu and Billings,1994) have been extensively studied and applied to a variety of situations by Billings and his coworkers. It is known that the properties of the OLS estimator are identical to those of conventional least squares with the difference that the former is equipped with a powerful structure selection criteria to aid in obtaining a parsimonious model.

During the implementation of the estimator to the problems posed in the present chapter, the complex data contained in the regressors are partitioned into real and imaginary components. The estimator can not distinguish whether the data in the information matrix 'P' is obtained from input-output records of a persistently excited system or if it is the real and imaginary component of complex frequency response data. But irrespective of the nature or source of the data, it is well known that in order to obtain an accurate estimate of the system parameters, the data must be representative of the underlying system. In time domain identification this is achieved by exciting the system by a persistently exciting input and sampling the signals at an appropriate rate.

Analogously the performance of the complex orthogonal estimator depends critically on the choice of the frequency range and the number of data points. The frequency range should therefore be selected with great care. It has been observed that when the errors are weighted uniformly over the frequency range of interest it becomes extremely difficult to find the correct frequency span that will give acceptable estimates. The weighted feature of the new algorithm offers a possible solution to overcome the sensitivity of the accuracy of the estimator to the choice of frequency range. This can be achieved by a suitable choice of the weighting matrix 'Q'. One of the possible choices is to set 'Q' in eqn(6) to be (Lancaster and Salkauska,1986)

$$Q = diag[q_1(\omega), q_2(\omega), ...q_N(\omega)] \tag{29}$$

where '$q_i(\omega)$' is a monotonically decreasing exponential of the form

$$q_i = exp(-\omega_i^2/\lambda) \tag{30}$$

The weighting parameter 'λ' controls the distribuition of the weights for each frequency 'ω' as shown in Fig-1.

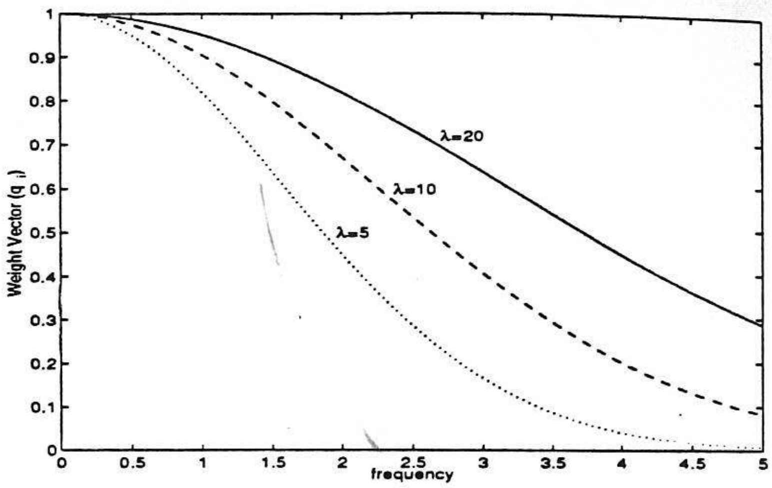

Fig.1 Variation of Weighting Vector with λ

The basis of selecting the tuning parameter 'λ' is intuitive. Note that the terms 'q_i' are positive weights which can be chosen in order to reflect information known about the accuracy of the measurements of the dependent variable at a particular frequency. Since it is known that the measurements of a signal in the the higher frequency range are usually more prone to the effects of measurement noise, the errors in this range should be given comparitively less weight than the lower frequency ranges. This is achieved by proper choice of the parameter 'λ' and will be discussed in more detail in the simulated results. Fig-1 illustrates how the influence of different frequency ranges is affected by different choices of 'λ'.

Under ideal conditions,the total number of noise free data points needed to identify the parameters associated with the 'L th' order transfer of eqn(19) is $2L + 1$. However in practice,the data will rarely be noise free and the exact model structure will not be known *a priori*. Extensive simulations on a large number of systems leads us to conjecture that the performance of the estimator does not critically depend on the number of data points provided this is within an optimum range of around 30 to 100. If the errors are uniformly weighted, the accuracy of the estimates strongly depends on the choice of the frequency range. Since it is difficult even for an experienced user to precisely know the exact frequency range that is suitable for a particular system,it is advisable to specify the frequency range relatively arbitrarily and vary the weighting parameter 'λ' interactively to obtain the final estimates. This will be illustrated in the simulated examples.

4.6 Normalisation of Regressors

In this section the problem of possible numerical ill conditioning of the algorithm is highlighted and a procedure is adopted to overcome this. It is known that when estimating the parameters of discrete time linear or nonlinear models using either of the least squares based algorithms,the information matrix 'P' consists of lagged input-output data. In general if the input is persistently exciting the matrix 'P^TP' is well conditioned (i.e. the condition number of this matrix which measures the sensitivity or variability of a least square solution is small).

However,the situation is not the same when the information matrix 'P' formed by taking discrete frequency response data is analysed. For example,the information matrix for the system of equation(19) is given as

$$P = \begin{bmatrix} (j\omega_1)^L H_1(j\omega_1) & (j\omega_1)^{L-1} H_1(j\omega_1) & \cdots & (j\omega_1)^L & 1 \\ (j\omega_2)^L H_1(j\omega_2) & (j\omega_2)^{L-1} H_1(j\omega_2) & \cdots & (j\omega_2)^L & 1 \\ (j\omega_3)^L H_1(j\omega_3) & (j\omega_3)^{L-1} H_1(j\omega_3) & \cdots & (j\omega_3)^L & 1 \\ \vdots & \vdots & \vdots & \vdots & \\ (j\omega_{N-1})^L H_1(j\omega_{N-1}) & (j\omega_{N-1})^{L-1} H_1(j\omega_{N-1}) & \cdots & (j\omega_{N-1})^L & 1 \\ (j\omega_N)^L H_1(j\omega_N) & (j\omega_N)^{L-1} H_1(j\omega_N) & \cdots & (j\omega_N)^L & 1 \end{bmatrix} \tag{31}$$

Note that,the last column vector of the information matrix corresponding to the pure input term contains a one, whereas the first column vector of the matrix related to the Lth-order derivative of the output contains the terms $(j\omega_k)^L H(j\omega_k)$,for $k = 1, \ldots. N$. The ratio of the norms of these two vectors at relatively high values of 'L' is quite large which at the outset indicates that the condition number of the matrix 'P^TP' will be very large making the matrix highly ill-conditioned. It has been found in several cases that the condition number of the matrix 'P^TP' can become in the range of 10^{27}. Although the present estimator does not estimate the parameters by explicitly computing the inverse of the matrix 'P^TP'; numerical problems of this nature need to be avoided.

It is therefore recommended that the information matrix should be normalised prior to applying the weighted complex orthogonal estimator. For example, consider the model of eqn.(1)

$$\begin{aligned} z(j\omega) &= \textstyle\sum_{i=1}^{M} \theta_i p_i(j\omega) \\ &= \theta_1 p_1(j\omega) + \theta_2 p_2(j\omega) + \ldots. + \theta_M p_M(j\omega) \\ &= \theta_1 \|p_1(j\omega)\| \tfrac{p_1(j\omega)}{\|p_1(j\omega)\|} + \theta_2 \|p_2(j\omega\| \tfrac{p_2(j\omega)}{\|p_2(j\omega)\|} + \ldots\ldots \theta_M \|p_M(j\omega)\| \tfrac{p_M(j\omega)}{\|p_M(j\omega)\|} \\ &= \theta_1^{nr} p_1^{nr} + \theta_2^{nr} p_2^{nr} + \ldots. + \theta_M^{nr} p_M^{nr} \end{aligned} \tag{32}$$

where the superscript 'nr' denotes normalised term. Thus we get the orthogonal version of the normalised equation and estimate the normalised parameters $\theta_1^{nr}, \theta_2^{nr}, \ldots.. \theta_M^{nr}$. The original parameters of the model are then recovered directly from the normalised parameters.

5 Simulated Examples

In the present section the effectiveness of the new estimator is demonstrated using three different types of systems. The first example represents a linear mechanical oscillator possessing three dominant modes with varying degrees of damping. The second example refers to the modified Van der Pol equation and the third example represents the dynamics of a nonlinear system exhibiting chaotic behaviour.

5.1 Example-1

Consider a mechanical oscillator with three degrees of freedom shown in Fig-2.

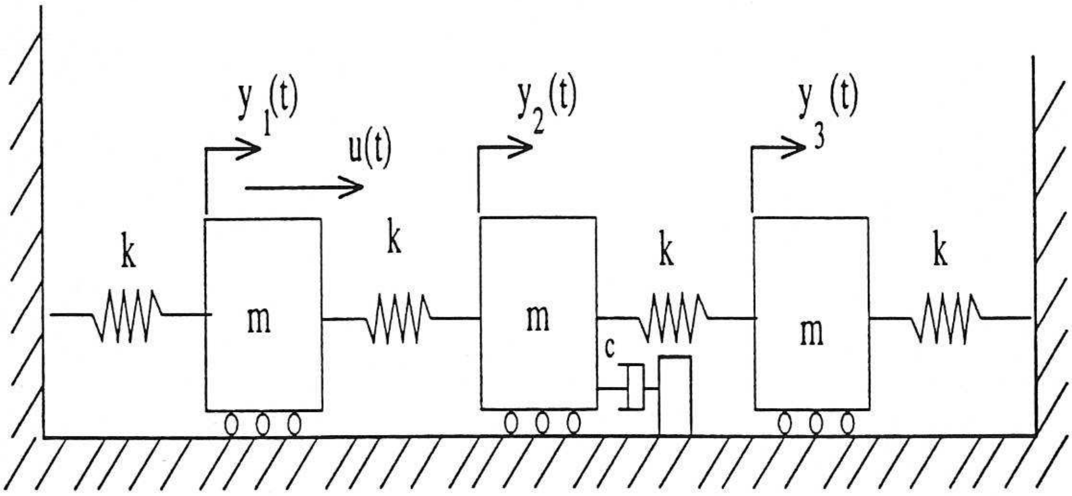

Fig.2 A Three Degree freedom Mechanical Oscillator

The equations of motion for three masses are given by

$$\begin{aligned} m\ddot{y}_1 &= k(y_2 - 2y_1) + u \\ m\ddot{y}_2 &= k(y_3 - 2y_2 + y_1) - c\dot{y}_2 \\ m\ddot{y}_3 &= k(-2y_3 + y_2) \end{aligned} \tag{33}$$

where $y_1(t), y_2(t)$ and $y_3(t)$ represent the displacements of the three masses from static equilibrium as a result of the force $u(t)$ acting on m. The constants 'c' and 'k' represent the viscous damping and stiffness of the system.

For a steady state response to harmonic excitation defined by $u(t) = e^{j\omega t}$, the linear transfer function between y_1 and u is given as

$$H_1(j\omega) = \frac{m^2s^4 + mcs^3 + 4mks^2 + 2cks + 3k^2}{m^3s^6 + m^2cs^5 + 6m^2ks^4 + 4mcks^3 + 10mk^2s^2 + 4ck^2s + 4k^3} \tag{34}$$

When the parameters of the system are $m = 1kg, c = 0.3\ Ns/m, k = 1\ N/m$, the transfer function is given by (Newland,1989)

$$H_1(j\omega) = \frac{s^4 + 0.3s^3 + 4s^2 + 0.6s + 3}{s^6 + 0.3s^5 + 6s^4 + 1.2s^3 + 10s^2 + 1.2s + 4} \tag{35}$$

The magnitude and phase response of the system are shown in the Fig-3.

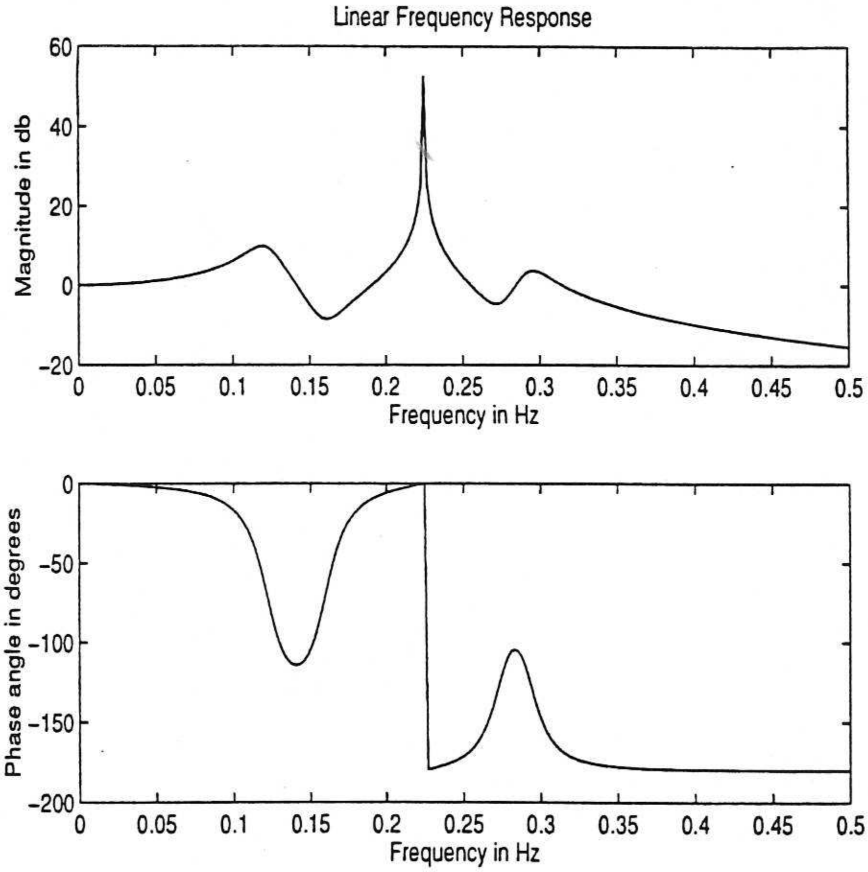

Fig.3 The Gain and Phase Response of system of Example-1

Initially,a set of 100 equally spaced frequency response data were generated from eqn(35) in the frequency range 0 -2.0Hz and these were used to estimate the parameters of the transfer function. An overparameterized (8th order) transfer function was initially specified for the estimation to test the capability of the proposed algorithm to detect the correct model structure. It was found that when the weighting matrix 'Q' was specified to be a unity diagonal matrix,implying that the errors are weighted uniformly over the entire frequency range,the estimator failed to detect the correct structure from the specified 8th order model. However, when the range of frequency was fixed from 0-0.5 Hz,the estimator identified both the correct model terms and parameters. It is found that when the errors are weighted uniformly the performance of the estimator strongly depends on the choice of the frequency range and the number of data points. However,the sensitivity of the estimator to the above factors was considerably reduced when the errors were weighted nonuniformly using the parameter 'λ'. Parameter 'λ' was chosen interactively and fixed at a value of 5.0. Table-1 shows the terms that were selected as the iteration proceeded together with the error reduction ratios. The parameters of the model are shown in Table-2.

Table-1 : Parameters of the Auxiliary model for Example-1 $\lambda = 5.0$

iteration	selected Terms (p_i)	ERR	Par.of Aux.model
1	d^2y/dt^2	89.6036	2.0191
2	u	5.1892	-2.1138
3	d^y/dt^4	0.0976	0.0774
4	d^2u/dt^2	2.8879	-1.4754
5	du/dt	0.1495	0.2318
6	d^3u/dt^3	0.1127	0.0326
7	d^4u/dt^4	0.0838	-0.0088
8	d^6y/dt^6	1.7397	0.9830
9	dy/dt	0.0054	0.0228
10	d^3y/dt^3	0.0113	0.1113
11	d^5y/dt^5	0.1190	0.3018

Since in practice the frequency response data could be corrupted by noise,the performance of the estimator was studied by adding various levels of noise to the frequency response data. At any frequency response data point $H'(j\omega_k)$,the noise corrupted data was taken as

$$H(j\omega_k) = H'(j\omega_k) + \xi(j\omega_k) \tag{36}$$

where each component of the noise term is assumed to have the form

$$\xi(j\omega_k) = \xi_R(\omega_k) + j\xi(\omega_k) \tag{37}$$

$\xi_R(\omega_k)$ and $\xi_I(\omega_k)$ are zero mean independent noise sequences with equal variance; and

$$E(\xi(j\omega_m)\xi(j\omega_l)^T) = 0 \quad \forall m \neq l \tag{38}$$

The results of the estimation for different levels of noise are summarised in Table-2 from which it is evident that as the noise level in the data were increased, the weighting parameter 'λ' was decreased to obtain the correct model structure and parameters. This is in accordance with the fact that since the accuracy of the frequency response estimates in the higher frequency range is poorer compared to the lower frequency range under the effect of significant measurement noise, the errors in higher frequency ranges were given less weight to provide better estimates. Note that ,although the results given here are for specific values of 'λ', simulations show that the exact value of 'λ' used is not crucial. The determination of the model structure and the parameter estimates are much the same for the range of values around that given for 'λ' in Table-2. Thus the algorithm does not seem to be sensitive to a reasonable range of values for 'λ'.

Table-2 : Parameters for Example-1 at Different Levels of Noise

Candidate Terms	True Values	Estimated Values $\lambda = 5.0$	Estimated Values (35.38db) $\lambda = 1.0$	Estimated Values (25.84db) $\lambda = 0.9$	Estimated Values (21.40db) $\lambda = 0.7$
d^6y/dt^6	1.00	1.0033	.9973	0.9920	1.0004
d^5y/dt^5	0.3	0.3018	0.2984	0.2940	0.2916
d^4y/dt^4	6.0	6.0158	5.9920	5.9756	5.9950
d^3y/dt^3	1.2	1.1995	1.2022	1.2024	1.1997
d^2y/dt^2	10.0	10.0185	9.9952	9.9851	9.9915
dy/dt	1.2	1.2028	1.2095	1.2244	1.2268
d^4u/dt^4	1.0	1.0032	0.9935	0.9814	0.9899
d^3u/dt^3	0.3	0.3084	0.2969	0.2853	0.2539
d^2u/dt^2	4.0	4.0084	3.988	3.9663	3.9762
du/dt	0.6	0.6218	0.5916	0.5668	0.5154
u	3.0	3.0007	2.9987	2.9962	2.9954

5.1.1 Parameter Estimation From Frequency Response Estimates

Rather than synthesising the transfer function by generating data from the exact system transfer function (eqn.35), it is possible to obtain data from an estimate of the system transfer function. To further illustrate the effectiveness of the algorithm, the transfer function was estimated directly from input-output data using the functions available in MATLAB . The plot of the estimated transfer function is shown in Fig-4.

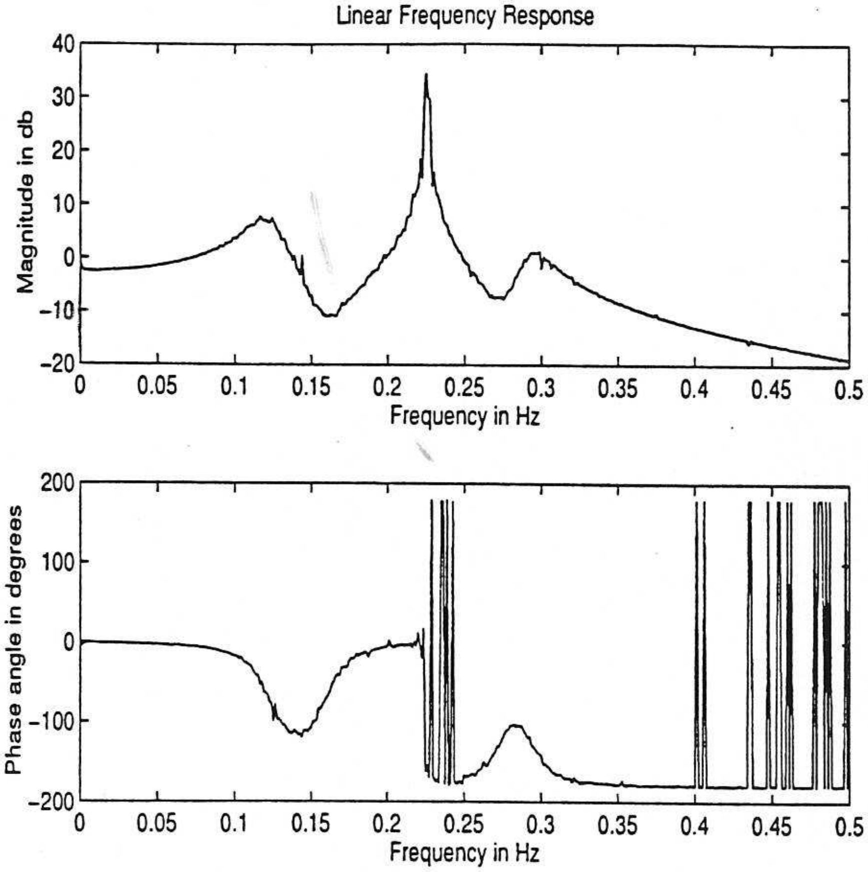

Fig.4 The Gain and Phase of Estimated Transfer Function for Example-1

From the plot it is evident that the transfer function estimate is not accurate in the entire frequency range (compare with Fig-3). However, 100 equally spaced frequency response data were taken in the frequency range 0 -0.37 Hz to estimate the parameters of the transfer function. Table-3 shows the terms that were selected from an 8th order linear model with '$\lambda = 2$', as the iteration proceeds, together with the associated error reduction ratios. Note that the estimator is capable of capturing the correct model structure. From this, it is found that the estimated parameters compare well with the true parametrs.

Table-3a : Parameters of the Auxiliary model of Example-1 based on Estimated Data

iteration	selected Terms (p_i)	ERR	Par.of Aux.model
1	d^2y/dt^2	96.8824	2.0098
2	u	1.5117	-2.2968
3	d^4y/dt^4	0.2536	0.7185
4	d^2u/dt^2	0.8258	-2.0339
5	d^6y/dt^6	0.2137	0.4521
6	d^4u/dt^4	0.2725	-1.0009
7	d^3u/dt^3	0.0009	0.0262
8	du/dt	0.0013	0.1524
9	d^5y/dt^5	0.0001	-0.0009
10	dy/dt	0.0072	0.2293
11	d^3y/dt^3	0.0269	1.2091

5.1.2 Table-3b : Parameters of Example-1 based on Estimated Data

Candidate Terms	Estimated Parameters	True Parameters
d^6y/dt^6	0.9975	1.00
d^5y/dt^5	0.2983	0.3
d^4y/dt^4	5.9788	6.0
d^3y/dt^3	1.2091	1.2
d^2y/dt^2	9.9588	10.0
dy/dt	1.2206	1.2
d^4u/dt^4	0.9759	1.0
d^3u/dt^3	0.2783	0.30
d^2u/dt^2	3.9525	4.0
du/dt	0.5717	0.6
u	3.0204	3.0

6 Synthesis of Nonlinear Transfer Functions

When the system is nonlinear,the traditional method of computing generalised frequency response functions (GFRF) involves extending the classical cross and power spectral density concepts to multi-dimentions and using correlation or FFT based techniques coupled with multidimensional windowing and smoothing (Kim and Power,1988)

An alternative indirect approach is to estimate a nonlinear time domain model (usually a polynomial or rational NARMAX model)(Leontaritis and Billings,1985) from the sampled input output data and then to compute the GFRF of the system from the unbiased process model using the recursive algorithms proposed in (Billings and Tsang,1989,Peyton Jones and Billings,1989). The advantages of the latter approach is that only a small data sets are required and the complexities of multi-dimensional windowing and smoothing are avoided. The prcedure for implementing the above procedure is summarised below.

- Fit a discrete model (often a polynomial or rational NARMAX model) to the input-output data.
- Compute the frequency response functions i.e. $H_1(j\omega_1)$, $H_2(j\omega_1, j\omega_2)$,...$H_n(j\omega_1, ...j\omega_n)$ from the estimated model.
- Estimate the parameters of the nonlinear continuous time model using the new estimator replacing the frequency response functions by their estimates in eqn.(19-26).

The effectiveness of the new estimator in estimating the parameters of nonlinear systems using the above procedure is demonstrated in the follwing two examples.

6.1 Example-2

Consider the following modified Van der Pol Equation

$$\frac{d^2y}{dt^2} + 2\xi\omega_n(1 - y^2(t))\frac{dy}{dt} + \omega_n^2 y(t) = u(t) \tag{39}$$

The system posesses nonlinear damping behaviour and has a stable node at the origin with a domain of attraction which lies within an unstable limit cycle. For $\xi = 0.01$ and $\omega_n = 45\pi rad/sec$,the linear frequency response exhibits a resonant peak at a frequency of 22.5Hz. In order to estimate a discrete model of the system, the system was excited by a unifomly distributed white noise sequence with rms value 10 and input-output data were collected by sampling at 200Hz. A discrete NARMAX model was estimated using structure detection,parameter estimation and model validation to yield the model

$$\begin{aligned} y(k) &= 1.4858y(k-1) - 0.9859y(k-2) \\ &+0.000025u(k-1) + 0.0141y(k-1)y(k-1)y(k-1) \\ &-0.0141y(k-1)y(k-1)y(k-2) + \Psi_\xi(k-1)\hat{\Theta}_\xi + \xi(k) \end{aligned} \tag{40}$$

where the last two terms of eqn.(40) represent the noise model which was included to ensure unbiased process model coefficients. This model can now be used to compute generalised frequency response functions using the procedure of Peyton Jones and Billings (1989). The procedure consists of computing a set of explicit algebraic equations in a recursive framework.

To estimate the coefficients of the linear terms, 100 equally spaced frequency response data were generated in the frequency range of 0-20.0Hz. To further demonstrate the effectiveness of the proposed estimator in detecting correct model structure, initially an overparameterised model (linear continuous time model having 5th order dynamics of input and output) was specified. The weighted parameter 'λ' was chosen to be 10. The selected terms together with the associated error reduction ratios are shown in Table-4a. It was found that after three iterations the sum of the error reduction ratios equaled 100% indicating that these three terms have captured the linear dynamics associated with the system.

Tabl-4a Parameters of the Linear Auxiliary Model for Example-2

iteration	selected Terms (p_i)	ERR	Par.of Auxiliary Model
1	u	99.9788	-0.0005048
2	d^2y/dt^2	0.0208	0.000466
3	dy/dt	0.0004	0.000019

It is evident from the model of eqn(40) that the system does not posses second order nonlinear terms implying that the system does not have a second order transfer function. Hence the continuous time model of the system will not have any second degree nonlinear terms. To estimate the terms corresponding to third order nonlinearity, 27 equally spaced frequency response data were generated in the frequency range of 0-5Hz. Intially an overparameterised model having third order nonlinear terms was specified. It was found that with the inclusion of a term '$y^2\frac{dy}{dt}$', which contributes 99.923% in the first iteration, the contribution of other terms in the next iterations are insignificant. The final estimated parameters of the system are shown in Table-4b which indicates that the estimates are in good agreement with the true parametrs of the system. Notice that in all the examples no prior knowledge of the model structure or terms to include in the model was assumed. The algorithm selected appropriate model terms and parameter estimates automatically from the data samples provided.

Table-4b : Final Parameters for Example-2

Candidate Terms	True Value	Estimated Value
d^2y/dt^2	1.0	0.9904
dy/dt	2.8286	2.8213
u	1.0	0.99999
$y^2\frac{dy}{dt}$	2.8286	2.8113

6.2 Example-3

Consider the well known Duffing-Holmes equation (Moon,1987) given by

$$\frac{d^2y}{dt^2} + \delta\frac{dy}{dt} - \beta y + y^3 = Acos(\omega(t)) \tag{41}$$

For $\delta = 0.15$,$\beta = 1.0$,A=0.3 and $\omega = 1.0$ rad/sec, the system exhibits chaotic behaviour and settles to a strange attracter. The estimated discrete NARMAX model of this system at a sampling frequency of 4.4446Hz is given by (Aguirre and Billings,1995)

$$\begin{aligned} y(k) &= 0.84725y(k-1) + 0.35713y(k-3) - 0.069431y(k-1)^3 \\ &+0.01278y(k-4) + 0.061319u(k-1) + 0.40325y(k-2) \\ &-0.00024349y(k-1)y(k-2)y(k-5) - 0.46215y(k-5) \\ &+0.096339u(k-3) - 0.15316y(k-2)y(k-3)y(k-4) \\ &-0.0073618y(k-1)^2y(k-5) + 0.071822y(k-1)y(k-3)y(k-4) \\ &+\Theta_\xi + \xi(t) \end{aligned} \tag{42}$$

From the model it is evident that the system has a linear and third order transfer functions. To estimate the parameters of the linear part of the continuous time model of the system, 100 equally spaced frequency response data were generated from eqn(42) in the frequency range 0-2.0Hz with the weighting parameter 'λ' equal to 0.8. To further show the capability of the proposed estimator in selecting the correct model structure, an overparametrised linear model of 5th order was specified. Table-5a shows the selected terms together with the associated error reduction ratios as the iteration proceeds. It was found that after three iterations,the sum of error reduction ratios of the first three terms selected equaled 99.9998% which implies that these three terms are adequate to represent the underlying linear dynamics of the system.

Table-5a : Parameters of the Linear Auxiliary Model for Example-3

iteration	selected Terms (p_i)	ERR	Par.of Aux.model
1	u	94.0888	0.8003
2	d^2y/dt^2	5.5058	-0.9658
3	dy/dt	0.4052	-0.1530

For the estimation of the third order nonlinearities, 27 frequency response data were used in the frequency range of 0-0.2Hz. An overparameterised model of third order nonlinerity was intially specified and it was found that in the first iteration with the selection of the term y^3, the total ERR becomes 99.99%. The contribution of other terms in subsequent iterations were insignificant;thus indicating that the term y^3 captured the nonlinear dynamics of the system. The final estimated parameters of the system are shown in table-5b which compare well with the original system parameters.

Table-5b : Final Parameters for Example-3

Candidate Terms	True Value	Estimated Value
d^2y/dt^2	0.15	0.15297
dy/dt	1.0	0.97946
u	1.0	0.99560
y^3	1.0	0.9942

7 Conclusions

A new complex orthogonal estimation algorithm has been derived to estimate continuous time nonlinear differential equation models from frequency response data. The algorithm includes structure detection procedures which determine which linear and nonlinear terms to include in the estimated continuous time model. This means that given frequency response estimates and no other *a priori* information about a system, parsimonious differential equation models can be determined without the need to compute derivatives. The frequency response estimates can be determined by any procedure which yields unbiased estimates. The effects of weighting were investigated and shown to have a significant effect on the quality of the estimates obtained. Several simulated examples including both linear and nonlinear systems were included to demonstrate the effectiveness of the new procedure.

Acknowledgements

AKS gratefully acknowledges the financial support provided by the Commonwealth Scholarship Commission of the United Kingdom and is thankful to Board of Governors, Regional Engineering College, Rourkela, India for granting study leave. SAB gratefully acknowledges that part of this work was supported by EPSRC grant Ref.GR/J05149.

References

[1] Aguirre,L.A. and Billings,S.A. (1995), "Retrieving Dynamical Invariants from Chaotic Data Using NARMAX Models", Int. J. of Bifurcation and Chaos (in press).

[2] Billings,S.A.,Korenberg,M.J. and Chen,S (1988),"Identification of nonlinear output affine systems using an orthogonal least square algorithm",Int.Jou.of System Science,Vol.19,pp.1559-1568

[3] Billings,S.A. and Peyton Jones,J.C.(1990)."Mapping Nonlinear Integro-Differential Equations into the Frequency Domain',Int.J.Control, vol.52(4), pp.863-879.

[4] Billings,S.A. and Voon,W.S.F. (1986),"Correlation based model validity tests for nonlinear models",Int.J.Control,vol.44,pp.235-244.

[5] Billings,S.A., Chen,S. and Korenberg,M.J. (1989). "Identification of MIMO Nonlinear Systems using a forward regessor orthogonal estimator" Int.J.Control, Vol.49, No.6, pp.2157-2189

[6] Billings,S.A. and Tsang,K.M. (1989), "Spectral Analysis for Nonlinear Systems, Part-I-Parametric Nonlinear Spectral Analysis, "J.Mech. Systems and Signal Processing, vol.3, no.4, pp.319-339.

[7] Billings,S.A. and Zhu,Q.(1993). "A structure detection algorithm for nonlinear dynamic rational model',Int.J.Control, Vol.59,No.6, pp.1439-1463

[8] Deutsch,Ralph(1965)."Estimation Theory", Prentice Hall-Inc, Englewood Cliffs, N.J.

[9] Kim,K.I. and Powers,E.J. (1988),"A digital method of modelling quadratically nonlinear systems with a general random input", IEEE Transactions ASSP, Vol.36, pp.1758-1769.

[10] Lancaster,P. and Kestutis,S. (1986),"Curve and Surface Fitting An Introduction", Academic Press

[11] Lawrence,P.J. and Rogers,G.J. (1979), "Sequential Transfer Function Synthesis from measured data", Proc.IEE,vol.126,no.1,pp.104-106.

[12] Leontaritis,I.J. and Billings,S.A. (1985), "Input-Output parametric models for Nonlinear Systems,Part-I-Deterministic Nonlinear Systems, Part-II-Stochastic Nonlinear Systems,Int.J.Control,vol.41, pp.303-344.

[13] Levy,E.C.(1959), "Complex Curve Fitting", IRE Trans. on Automatic Control,vol.4, pp.37-43.

[14] Ljung,L. and Glover,K.(1981), "Frequency Domain Versus Time Domain Methods in System Identification", Automatica,17,pp.71-86.

[15] Ljung,L (1987), "System Identification-Theory for the User", Prentice Hall,Inc, Englewood Cliffs, New Jersey.

[16] Moon,F.C. (1987), "Chaotic Vibrations-an introduction for applied scientists and Engineers", John Willey and Sons, New York.

[17] Newland,D.E.(1989),"Mechanical vibration analysis and computation",Longman Scientific and Technical Publications.

[18] Payne,P.A. (1970), "An Improved technique for Transfer Function Synthesis from Frequency Response Data", IEEE. Trans, Automatic Control, pp.480-483.

[19] Peyton Jones,J.C. and Billings,S.A.(1989), "A recursive Algorithm for computing the frequency response of a class of Nonlinear Difference Equation Models",Int.J.Control,vol.50, no.5,pp.1925-1940.

[20] Stahl,H. (1984), "Transfer Function Synthesis using frequency response data", Int.J.Control, vol.39,no.3,pp.541-550.

[21] Tsang,K.M. and Billings,S.A.(1992a),"An Orthogonal Estimation Algorithm for Complex Number Systems", Int.J.Sys.Science, vol.23,no.6,pp.1011-1018.

[22] Tsang,K.M. and Billings,S.A.(1992b), "Reconstruction of Linear and Nonlinear Continuous Time Models from Discrete Time Sampled Data Systems", Mech.Systems and Signal Processing,vol.6(1),pp.69-84.

[23] Whitfield,A.H.(1986), "Transfer Function Synthesis Using Frequency Response Data",Int.J.Control, vol.43,no.5,pp.1413-1426.

[24] Whitfield,A.H.(1987), "Asymptotic behaviour of Transfer Function Synthesis Methods", Int.J.Control, vol.43. pp.1413-1426.

[25] Zhu,Q. and Billings,S.A.(1994), "Fast Orthogonal Identification of Nonlinear Stochastic Models and Radial Basis Function Neural Network", Int. J. of Control (in press).

A APPENDIX

A.1 Weighted Orthogonal Least Squares

The weighted complex orthogonal estimator is derived under the assumption that the modelling error

$$\Xi(j\omega) = \Xi_R(\omega) + j\Xi_I(\omega) \tag{A.1}$$

in eqn(2) is a complex white noise sequence of zero mean and variance σ^2 and the matrix P^TQP is positive definite.

Consider the complex system of eqn(2) which after partitioning,as shown in eqn.(5), is represented as

$$Z = P\theta + \Xi \tag{A.2}$$

Introducing a weighting matrix 'Q' (normally a symmetric positive definite matrix or a diagonal matrix) the matrix $P^T QP$ can be shown to be symmetric and positive definite,and hence can be expressed as

$$P^T QP = T^T DT \tag{A.3}$$

where 'D' is a diagonal matrix and 'T' is an upper triangular matrix with unity diagonal elements such that

$$T = \begin{bmatrix} 1 & \alpha_{12} & \alpha_{13} & \dots & \alpha_{1M} \\ 0 & 1 & \alpha_{23} & \dots & \alpha_{2M} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ & & & 1 & \alpha_{M-1M} \\ 0 & & & & 1 \end{bmatrix} \tag{A.4}$$

Since $T^T T = I$ the model of eqn(A.2) can be transformed to an auxiliary model parameterised with respect to vector 'g' as

$$\begin{aligned} Z &= P\theta + \Xi \\ &= P(T^{-1}T)\theta + \Xi \\ &= Wg + \Xi \end{aligned} \tag{A.5}$$

where

$$W = PT^{-1} \quad \text{, and} \quad g = T\theta \tag{A.6}$$

The properties of the matrix W are such that $W^T QW$ is orthogonal because

$$\begin{aligned} W^T QW &= (PT^{-1})^T Q(PT^{-1}) = (T^{-1})^T P^T QPT^{-1} \\ &= (T^T)^{-1} T^T DTT^{-1} = D \end{aligned} \tag{A.7}$$

Further let

$$V = W^T Q \tag{A.8}$$

Therefore

$$VW = D \tag{A.9}$$

Now premultiplying $W^T Q$ and post multiplying by T in eqn.(A.6),we have

$$\begin{aligned} W^T QWT &= W^T QPT^{-1}T \\ &= W^T QP \end{aligned} \tag{A.10}$$

Hence

$$\begin{aligned} T &= (W^T QW)^{-1} W^T QP \\ &= (VW)^{-1} VP \\ &= D^{-1} VP \end{aligned} \tag{A.11}$$

The regressors of the auxiliary model of eqn.(A.5) can be obtained recursively from

$$W = P - W(T - I) \tag{A.12}$$

that is

$$\begin{aligned} w_1(\omega) &= p_1(\omega) \\ w_i(\omega) &= p_i(\omega) - \sum_{k=1}^{i-1} \alpha_{ki} w_k(\omega) \quad for\ k < i \end{aligned} \tag{A.13}$$

where from eqn.(A.4,A.11)

$$\begin{aligned} \alpha_{ki} &= \frac{\langle (w_k^T(\omega)Q), p_i(\omega)\rangle}{\langle (w_k^T(\omega)Q), w_k(\omega)\rangle} \\ &= \frac{\langle v_k(\omega), p_i(\omega)\rangle}{\langle v_k(\omega), w_k(\omega)\rangle} \quad for\ \ k = 1.....i-1 \end{aligned} \tag{A.14}$$

and ' $\langle .,. \rangle$ ' denotes the dot product of the vectors. The auxiliary parameter vector

$$g = \begin{bmatrix} g_1 \\ \vdots \\ g_M \end{bmatrix} \tag{A.15}$$

satisfies the equation

$$g = D^{-1}W^TQZ - D^{-1}W^TQ\Xi \tag{A.16}$$

so that the estimated 'g' is given by

$$\hat{g} = D^{-1}W^TQZ \tag{A.17}$$

or

$$\hat{g}_i = \frac{\langle Z(\omega), v_i(\omega)\rangle}{\langle v_i(\omega), w_i(\omega)\rangle} \quad for\ i = 1, ...M \tag{A.18}$$

Inspection of eqn(A.18) indicates that if $v_i(\omega)$ and hence $w_i(\omega)$ is not related to the output $Z(\omega)$, the estimated parameter $\hat{g}_i$ will be small, since the average value $\langle Z(\omega), v_i(\omega)\rangle$ will not be significant. Hence the orthogonal parameter '$\hat{g}_i$' could be used as an indicator of the significance of the terms '$w_i(\omega)$'. If the magnitude of the estimated orthogonal parameter '$\hat{g}_i$' is less than a certain threshold,the associated term '$w_i(\omega)$' should be regarded as insignificant. However,the value of the threshold depends critically on the problem under consideration. Hence it is desirable to have a structure selection criteria based on the Error Reduction Ratio which is derived section A.2

Once the parameters $g_i, i = 1, ...M$ are estimated,the original system parameters $\theta_i, i = 1, ...M$ can easily be recovered according to the formula

$$\hat{\theta} = \hat{g} - (T - I)\hat{\theta} \tag{A.19}$$

that is

$$\begin{aligned} \hat{\theta}_M &= \hat{g}_M \\ \hat{\theta}_k &= \hat{g}_k - \sum_{i=k+1}^{M} \alpha_{ki}\hat{\theta}_i, \quad for\ k = 1, ...M-1 \end{aligned} \tag{A.20}$$

Therefore by using the above equations,the unknown parameters $\theta_i, i = 1, ...M$ can be estimated step by step.

A.2 Error Reduction Ratio for the Weighted Complex Orthogonal Estimator

The proportion of the variance of the dependent variable explained by $w_i(\omega)$ is

$$\frac{g_i^2 \langle v_i(\omega), w_i(\omega)\rangle}{\langle (z^T(\omega)Q), z(\omega)\rangle} \tag{A.21}$$

This is because from eqn(A.5)

$$\begin{aligned} Z^TQZ &= g^TW^TQWg + \Xi^TQ\Xi + g^TW^TQ\Xi + \Xi^TQWg \\ &= g^TVWg + \Xi^TQ\Xi + g^TV\Xi + \Xi^TQWg \end{aligned} \tag{A.22}$$

The assumption on noise properties ensure that $E[V\Xi] = 0$, and $E[\Xi^TQW] = 0$; hence

$$\left\langle (Z^T(\omega)Q), Z(\omega)\right\rangle = \sum_{i=1}^{M} g_i^2 \langle v_i(\omega), w_i(\omega)\rangle + \langle (\Xi(\omega)Q), \Xi(\omega)\rangle \tag{A.23}$$

When the model is empty, i.e. no term is selected ($M = 0$), the weighted mean square of the magnitude of the prediction errors will achieve a maximum value of

$$\langle (\Xi(\omega)Q), \Xi(\omega)\rangle\,|_{M=0} = \left\langle (Z^T(\omega)Q), Z(\omega)\right\rangle \tag{A.24}$$

Once significant terms have been included,the mean square value of the prediction error $\Xi(\omega)$ will be reduced accordingly. The reduction resulting by including the term '$g_iw_i(\omega)$' can be expressed as a percentage reduction in the total mean square of the prediction errors by defining

$$ERR_i = \frac{g_i^2 \langle v_i(\omega), w_i(\omega)\rangle}{\langle (z^T(\omega)Q), z(\omega)\rangle} \times 100\%, \quad i = 1, ...M \tag{A.25}$$

This can be used as a selection tool to determine which terms are to be included in the final model. However the value of the ERR_i may depend upon the order in which the term $p_i(\omega)$ (eqn.(1)) enter into the equation. This problem can be overcome by using the forward regession procedure of Billings et al(1988) described in section-2.