Estimation of delayed process model parameters in the frequency domain

AIDAN O'DWYER,
School of Control Systems and Electrical Engineering,
Dublin Institute of Technology,
Kevin St., Dublin 8,
IRELAND.

Abstract:- This paper discusses the estimation of the parameters (including the time delay) of a generalised single input, single output (SISO) process model from an appropriate number of arbitrarily specified points on the process frequency response. The method involves combining an analytical approach with a least squares approach using a gradient algorithm, to provide accurate and robust estimates of the parameters.

Keywords:- Estimation, time delay processes, frequency domain, gradient methods. *CSCC'99* Proc.pp..1361-1367

1 Introduction

Identification in the frequency domain often involves the estimation of the process frequency response over an appropriate frequency range followed by the estimation of the model parameters. This paper focuses on delayed model parameter estimation from an appropriate number of arbitrarily specified points on the process frequency response; the frequency domain appears to be intuitively appropriate for the estimation of the delay (in particular), as the delay affects the process phase response, but not its magnitude response. Graphical [1-3], analytical [4-6] and least squares [3, 7-9] approaches have all been considered for the estimation problem. Lilja [8], for example, estimates the non-delay parameters of a first order lag plus time delay (FOLPD) process model by minimising an appropriate cost function; the delay is estimated separately by determining the global minimum of a non-unimodal cost function using a modified Newton-Raphson gradient algorithm. Many such approaches have the disadvantage of separately estimating the non-delay parameters and the delay; this leads to biased estimation of the delay or difficulty in achieving reliable convergence of the delay estimate to its optimum value.

These difficulties motivate an investigation of the possibility of estimating the non-delay and delay parameters together. A two stage approach, combining an analytical approach and a gradient approach, will be defined for the estimation of the parameters of an arbitrary order delayed model. The analytical methods are based on direct calculation of the parameters from the frequency response, using simultaneous equations. These estimates are updated to more accurate model parameters using a gradient algorithm. This two stage approach will rely on the analytical estimates being sufficiently accurate so that unimodality of the cost function (which is a function of the sum of the squares of the sampled errors between the process and model frequency responses) with respect to the parameter estimates, exists from the estimates determined analytically to the final model parameter estimates.

The analytical formulae to estimate the model parameters are developed in Section 2. The gradient approach to estimating the parameters, from the initial estimates of the parameters, is developed in Section 3. Implementation issues and simulation results are discussed in Sections 4 and 5. In Section 6, conclusions are drawn and future work is outlined.

2 Analytical Estimation of the Model Parameters

The estimates of the parameters of an v^{th} order delayed model using an analytical approach are obtained by calculating the non-delay parameters from an appropriate number of simultaneous equations, using data points on the magnitude response; the delay may then be calculated from one data point on the phase response. The transfer function of the v^{th} order delayed model is defined as follows (with $v \geq u$)

$$G_m(s) = \frac{\left(b'_{0m} + b'_{1m}s + b'_{2m}s^2 + \ldots\ldots. + b'_{um}s^u\right)e^{-s\tau_m}}{1 + a_{1m}s + a_{2m}s^2 + \ldots\ldots. + a_{vm}s^v} \quad (1)$$

with a parameter vector

$$x_1 = \left[a_{1m} \quad a_{2m}\ldots\ldots..a_{vm} \quad b'_{0m} \quad b'_{1m} \quad b'_{2m}\ldots\ldots.b'_{um} \quad \tau_m\right]^T$$
$$x_1 \in R^{u+v+2} \quad (2)$$

The numerator and denominator terms of the frequency transfer function, $G_m(j\omega)$, may be written as

$$N_m(j\omega) = \left\{ \sum_{q=0}^{\mathrm{int}\left[\frac{u}{2}\right]} b'_{2qm}(-1)^q \omega^{2q} + j \sum_{q=1}^{\mathrm{int}\left[\frac{u+1}{2}\right]} b'_{(2q-1)m}(-1)^{q-1}\omega^{2q-1} \right\} \quad (3)$$

and

$$D_m(j\omega) = \sum_{r=0}^{\mathrm{int}\left[\frac{v}{2}\right]} a_{2rm}(-1)^r \omega^{2r} + j \sum_{r=1}^{\mathrm{int}\left[\frac{v+1}{2}\right]} a_{(2r-1)m}(-1)^{r-1}\omega^{2r-1} \quad (4)$$

with $a_{0m} = 1$, $\mathrm{int}\left[\frac{v}{2}\right] =$ integer part of $\frac{v}{2}$, $\mathrm{int}\left[\frac{u}{2}\right] =$ integer part of $\frac{u}{2}$.

Therefore, from equations (3) and (4), the magnitudes of the numerator and denominator terms may be written as

$$|N_m(j\omega)| = \sqrt{\left[\sum_{q=0}^{\mathrm{int}\left[\frac{u}{2}\right]} (-1)^q b'_{2qm}\omega^{2q}\right]^2 + \left[\sum_{q=1}^{\mathrm{int}\left[\frac{u+1}{2}\right]} (-1)^{q-1} b'_{(2q-1)m}\omega^{2q-1}\right]^2} \quad (5)$$

$$|D_m(j\omega)| = \sqrt{\left[\sum_{r=0}^{\mathrm{int}\left[\frac{v}{2}\right]} (-1)^r a_{2rm}\omega^{2r}\right]^2 + \left[\sum_{r=1}^{\mathrm{int}\left[\frac{v+1}{2}\right]} (-1)^{r-1} a_{(2r-1)m}\omega^{2r-1}\right]^2} \quad (6)$$

Now, $|G_m(j\omega)|^2$ may be written as

$$|G_m(j\omega)|^2 = \frac{\left(d_{0m} + d_{1m}\omega^2 + d_{2m}\omega^2 + \ldots\ldots. + d_{um}\omega^{2u}\right)}{1 + c_{1m}\omega^2 + c_{2m}\omega^4 + \ldots\ldots. + c_{vm}\omega^{2v}} \quad (7)$$

with

$$d_{(j-1)m} = b'^{\,2}_{(j-1)m} + 2\sum_{q=1}^{\mathrm{int}\left[\frac{j-1}{2}\right]} b'_{(2q-1)m}b'_{(2j-2q-1)m} - 2\sum_{q=1}^{\mathrm{int}\left[\frac{j}{2}\right]} b'_{(2q-2)m}b'_{(2j-2q)m} \quad (8)$$

$$d_{jm} = b'^{\,2}_{jm} + 2\sum_{q=1}^{\mathrm{int}\left[\frac{j}{2}\right]} b'_{(2q-2)m}b'_{(2j-2q+2)m} - 2\sum_{q=1}^{\mathrm{int}\left[\frac{j}{2}\right]} b'_{(2q-1)m}b'_{(2j-2q+1)m} \quad (9)$$

and

$$c_{(k-1)m} = a^2_{(k-1)m} + 2\sum_{r=1}^{\mathrm{int}\left[\frac{k-1}{2}\right]} a_{(2r-1)m}a_{(2k-2r-1)m} - 2\sum_{r=1}^{\mathrm{int}\left[\frac{k}{2}\right]} a_{(2r-2)m}a_{(2k-2r)m} \quad (10)$$

$$c_{km} = a^2_{km} + 2\sum_{r=1}^{\mathrm{int}\left[\frac{k}{2}\right]} a_{(2r-2)m}a_{(2k-2r+2)m} - 2\sum_{r=1}^{\mathrm{int}\left[\frac{k}{2}\right]} a_{(2r-1)m}a_{(2k-2r+1)m} \quad (11)$$

and with j and k even, $j \in [0,u], k \in [2,v]$. A minimum of u+v+1 data points on the magnitude response are required to estimate the parameters. If just u+v+1 data points are taken, the vector of magnitude response values squared is

$$F_1 = \left[|G_p(j\omega_1)|^2 \ldots\ldots |G_p(j\omega_{u+v+1})|^2\right]^T, F \in \Re^{u+v+1} \quad (12)$$

with $|G_p(j\omega)|$ = process magnitude at frequency ω. Then, from equations (7) and (12), it may be deduced that

$$\begin{bmatrix} d_{0m} \\ d_{1m} \\ . \\ . \\ d_{um} \\ c_{1m} \\ . \\ . \\ c_{vm} \end{bmatrix} = A^{-1} \begin{bmatrix} |G_p(j\omega_1)|^2 \\ |G_p(j\omega_2)|^2 \\ . \\ . \\ |G_p(j\omega_u)|^2 \\ |G_p(j\omega_{u+1})|^2 \\ . \\ . \\ |G_p(j\omega_{u+v+1})|^2 \end{bmatrix} \quad (13)$$

with A given in Appendix 1. The non-delay parameters of the model may subsequently be calculated from equations (8) and (9). The model delay may be calculated (using equations (1), (3) and (4)) to be

$$\tau_m = \frac{1}{\omega}\{-\phi_p(j\omega) + \tan^{-1}\left[\frac{\sum_{q=1}^{\mathrm{int}\left[\frac{u+1}{2}\right]} (-1)^{q-1} b'_{(2q-1)m}\omega^{2q-1}}{\sum_{q=0}^{\mathrm{int}\left[\frac{u}{2}\right]} (-1)^q b'_{2qm}\omega^{2q}}\right]\}$$
$$-\frac{1}{\omega}\tan^{-1}\left[\frac{\sum_{r=1}^{\mathrm{int}\left[\frac{v+1}{2}\right]} (-1)^{r-1} a_{(2r-1)m}\omega^{2r-1}}{\sum_{r=0}^{\mathrm{int}\left[\frac{v}{2}\right]} (-1)^r a_{2rm}\omega^{2r}}\right] \quad (14)$$

with $\phi_p(j\omega)$ = process phase at frequency ω. A less computationally intense alternative to the procedure defined is to estimate the parameters of a v^{th} order model, with no numerator parameters, and a repeated pole [4]. The lower computational intensity of this procedure is traded off against poorer accuracy of the parameters estimated.

3 Gradient estimation of the model parameters

An alternative parameter vector of the v^{th} order model (to equation (2)) is

$$x_2 = [K_m \quad a_{1m} \quad a_{2m} \ldots\ldots a_{vm} \quad b_{1m} \quad b_{2m} \ldots\ldots b_{um} \quad \tau_m]^T , x_1 \in R^{u+v+2} \qquad (15)$$

The phase contributions of the numerator and denominator terms, respectively, may be calculated, from $G_m(j\omega)$, using equation (15), to be

$$\phi_m^N(j\omega) = \tan^{-1} \frac{\sum_{q=1}^{\text{int}\left[\frac{u+1}{2}\right]} b_{(2q-1)m}(-1)^{q-1}\omega^{2q-1}}{\sum_{q=0}^{\text{int}\left[\frac{u}{2}\right]} b_{2qm}(-1)^q \omega^{2q}} \qquad (16)$$

$$\phi_m^D(j\omega) = -\tan^{-1} \frac{\sum_{r=1}^{\text{int}\left[\frac{v+1}{2}\right]} a_{(2r-1)m}(-1)^{r-1}\omega^{2r-1}}{\sum_{r=0}^{\text{int}\left[\frac{v}{2}\right]} a_{2rm}(-1)^r \omega^{2r}} \qquad (17)$$

If u+v+1 frequency response data points are taken for parameter estimation, the vector of frequency response values is

$$F_2 = \left[|G_p(j\omega_1)| \ldots\ldots |G_p(j\omega_{u+v+1})| \quad \phi_p(j\omega_1) \ldots\ldots \phi_p(j\omega_{u+v+1})\right]^T \quad F_2 \in \Re^{2u+2v+2} \qquad (18)$$

The error vector is:

$$e = [e_1 \quad e_2 \ldots\ldots e_{u+v+1} \quad e_{u+v+2} \quad e_{u+v+3} \ldots\ldots e_{2u+2v+2}]^T,$$
$$e_n = \frac{|N_m(j\omega)|}{|D_m(j\omega)|} - |G_p(j\omega)| \;,\; 1 \le n \le u+v+1 \qquad (19)$$

and

$$e_n = \phi_m^N(j\omega_{n1}) + \phi_m^D(j\omega_{n1}) - \omega_{n1}\tau_m - \phi_p(j\omega_{n1}),$$
$$u+v+1 < n \le 2u+2v+2 \,,\, n1 = n-u-v-1 \qquad (20)$$

The cost function, J, is formulated as

$$J = 0.5 e^T P e \qquad (21)$$

with

$$P = \text{diag}\left[\frac{1}{|G_p(j\omega_1)|} \ldots\ldots \frac{1}{|G_p(j\omega_{u+v+1})|} \quad \frac{1}{\omega_1} \ldots\ldots \frac{1}{\omega_{u+v+1}}\right] \qquad (22)$$

The normalising matrix, P, is used to increase the range of parameters over which unimodality of the cost function exists. The cost function, J (using equations (19) to (22)) may be calculated to be

$$J = 0.5 \sum_{n=1}^{u+v+1} \left[\left(\frac{|N_m(j\omega_n)|}{|D_m(j\omega_n)|} - |G_p(j\omega_n)|\right)^2 \frac{1}{|G_p(j\omega_n)|}\right]$$
$$+0.5 \sum_{n=1}^{u+v+1} \left[\frac{1}{\omega_n}(\phi_m^N(j\omega_n) + \phi_m^D(j\omega_n) - \omega_n\tau_m - \phi_p(j\omega_n))^2\right] \qquad (23)$$

Then, the updated estimate of the parameters at sample (k+1) may be calculated from the estimates at sample k, using the gradient algorithm

$$x_2(k+1) = x_2(k) - \mu \frac{\partial J}{\partial x_2(k)} \qquad (24)$$

with μ = learning rate. The initial values of the parameter estimates are determined using the analytical technique. If $|N_m(j\omega)|$ (equation (5)) is formulated as

$$|N_m(j\omega)| = K_m \sqrt{\left[\sum_{q=0}^{\text{int}\left[\frac{u}{2}\right]} (-1)^q b_{2qm}\omega^{2q}\right]^2}$$
$$+K_m \sqrt{\left[\sum_{q=1}^{\text{int}\left[\frac{u+1}{2}\right]} (-1)^{q-1} b_{(2q-1)m}\omega^{2q-1}\right]^2} \qquad (25)$$

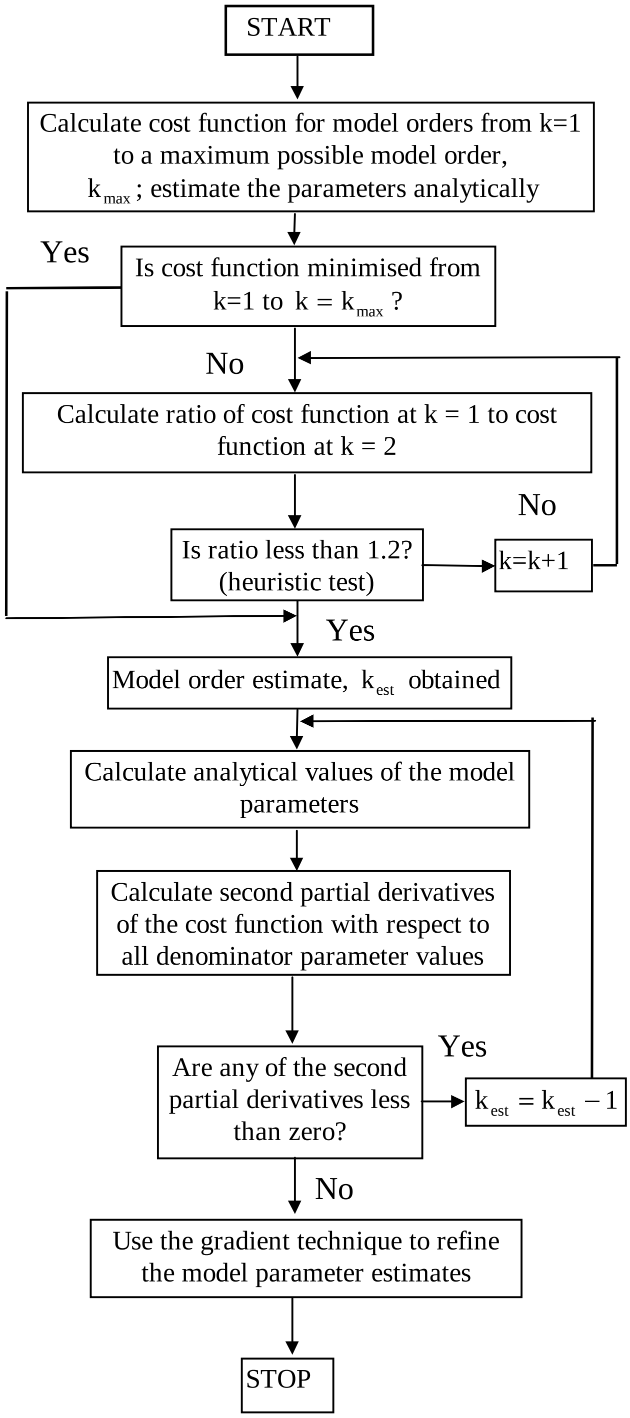
with $b_{0m} = 1$, then it is clear from equations (23) and (25) that the cost function is quadratic in the gain estimate, K_m and the delay estimate, τ_m. The cost function is not, however, quadratic in the estimates of the other numerator and denominator parameter values, as may be deduced from equations (6), (16), (17), (23) and (25). The cost function must be unimodal with respect to each of these parameter values (allowing the delay estimate, gain estimate and other parameter estimates to vary), and must have its minimum value when the appropriate equivalent process parameter equals the model parameter, if convergence of the model parameters to the equivalent process parameters is to be guaranteed. An equivalent condition is that the first partial derivative of the cost function with respect to each of the parameter values may be equal to zero once only, or that the second partial derivative of the cost function with respect to each of the parameter values must always be greater than zero (with the first partial derivative of the cost function with respect to each of the parameter values being equal to zero at appropriate parameter values). These first and second partial derivative functions may be calculated analytically in a straightforward manner.

4. Model Structure Selection

One approach to determine the model order is to calculate the slope of the process magnitude versus frequency curve at high frequencies, though experimentally obtained frequency response data are seldom accurate enough to exhibit a slope more negative than -40 dB/decade [3]. Alternatively, the parameters of an arbitrary order model could be estimated; the most appropriate model order could be determined by calculating where the cost function, formed from the optimum parameters estimated (using the gradient method) as the model order is increased, levels out. This procedure is computationally intensive. A variation of the above strategy that is less computationally intensive would be to calculate the cost function based on the initial model parameter estimates (calculated using an analytical approach). A repeated pole model would simplify the calculations further.

Simulation results have revealed that convergence of the parameters to their optimum values using gradient methods is not always facilitated for higher order delayed process models (such as third order delayed models), due to cost function non-unimodality. Therefore, a strategy for the estimation of the parameters of an appropriate delayed model, which ensures cost function unimodality, is summarised in Figure 1.

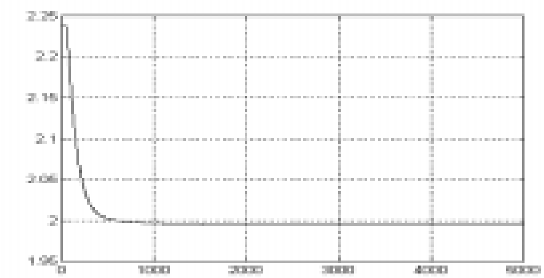
Fig. 1: Flowchart summarising the algorithm for model order and parameter estimation



One representative simulation result is included to demonstrate the applicability of the method. The results are determined when either +10% or -10% was added to the magnitude and phase values of the process frequency response (to simulate the effect of uncertain data); the parameter estimates are plotted against sample number. Ten values of the process frequency response, spaced equally between phase lags of 0^0 to 270^0 were used in the simulation; in the analytical stage, average values of the parameters are calculated over a number of points of the frequency response (to improve the robustness of the estimates).

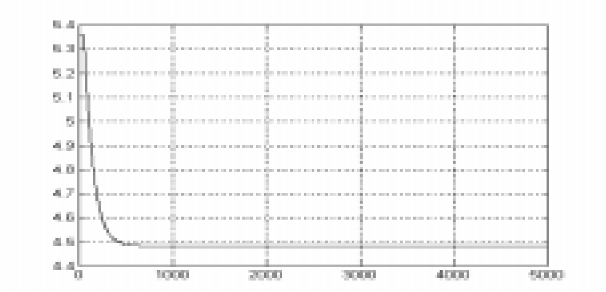
Simulation result: $G_p = 2e^{-1.0s}/1 + 4.5s + 4.5s^2$. Model order estimate: 2. The model parameters determined analytically are $K_m = 2.24$, $a_{1m} = 5.36$, $a_{2m} = 4.98$ and $\tau_m = 1.04$. These parameters are subsequently updated using the gradient algorithm (Figures 2 to 5). Time domain and frequency domain fitting is shown in Figures 6 and 7.

Fig. 2: $K_m - \mu = 0.1$



Sample Number

Fig. 3: $a_{1m} - \mu = 0.1$



Sample Number

Fig. 4: $a_{2m} - \mu = 0.1$

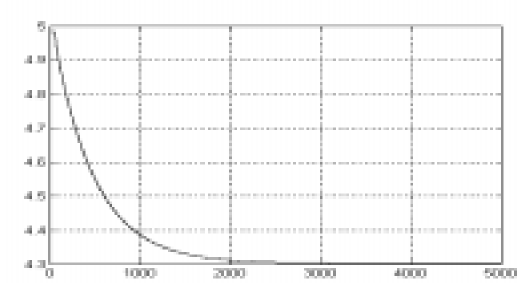
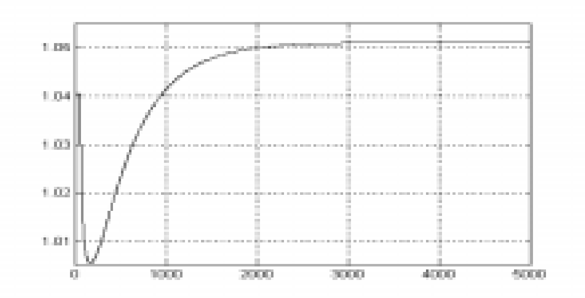


Fig. 5: $\tau_m - \mu = 0.01$



Sample Number

Fig. 6: Unit step response of the process and model

Fig. 7: Polar plot of the process and model

The full panorama of simulation results reveals that the estimation strategy summarised in the flowchart is a reasonable guide to the choice of an appropriate, sometimes reduced order, delayed model.

5 Other Issues

1. A recursive scheme to estimate model parameters as each data point became available has also been developed. For example, if the

parameters of a FOLPD model are to be estimated, a minimum of two data points are required to estimate the parameters analytically; the gradient technique appropriately updates the parameter estimates, as more data points became available. A flowchart for the recursive estimation algorithm and appropriate simulation results will be presented at the conference.

2. Appropriate estimates of the learning rate, μ, (equation (22)) have been found in simulation. The best setting of this value to allow rapid convergence of the parameter estimates appears to be related to the process order and to whether the process is underdamped or overdamped. Unfortunately, it is very possible for the model parameters to converge to non-optimum values if the value of the learning rate is too large. A trial and error procedure to choose the learning rate was the only satisfactory method developed; further work will consider the development of an adaptive learning rate.
3. The normalising used in the cost function (equation (19)) has the effect of weighting the cost function more equally over a wide range of frequencies. This facilitates the convergence of the model parameters to their optimum values, using the gradient method, over a wider range of initial model parameters than if no cost function weighting is used.

6 Conclusions

1. The two-stage method defined has successfully allowed the estimation of the parameters of SISO delayed process models from an appropriate number of arbitrarily specified points on the process frequency response in a wide variety of simulations.
2. It has been shown in simulation that convergence of the initial model parameter estimates, calculated using the analytical approach, to the optimum model parameter estimates calculated using the gradient approach, is possible if the model parameters are sufficiently close to the optimum parameters. An analytical proof of the convergence properties is the subject of future work. If the second partial derivative of the cost function with respect to the denominator parameter value(s) is less than zero, then a simple strategy that involves the commencement of iteration at different values of the parameter estimates could be employed to increase the probability that the parameters estimated using the gradient approach will correspond to the global minimum of the cost function. Similar methods based on this multiple model estimation technique have been well explored in the time domain.
3. An arbitrary order delayed model may be fitted to the data. Alternatively, a lower order model, such as a FOLPD model, may be estimated. The trade-off of relatively poor fitting of the original process (by estimating a FOLPD model) may be balanced by an increase in the parameter space for which the cost function is unimodal, faster convergence of the parameter estimates to their final values and a smaller computational burden. Of course, the acceptability of the fitting depends on the use to which the process model is applied. It appears reasonable that, for many applications, the phase lag range taken of 0^o to 270^o will be the maximum range over which good fitting will be required. This is true for many compensation strategies (e.g. PID controller design); in addition, most processes, being low pass in nature, will have a small magnitude at larger phase lags, making the measurement problem greater. These and the other considerations mentioned provide a cogent argument for estimating the parameters of a low order process model (such as a FOLPD model).

References:

[1] Unbehauen, H. and Rao, G.P., *Identification of continuous systems*, Volume 10, North Holland System and Control Series, Elsevier Science Publishers B.V., 1987.

[2] Deshpande, P.B. and Ash, R.H., *Elements of Computer Process Control with Advanced Control Applications*, Instrument Society of America, Prentice-Hall Inc., 1983.

[3] Seborg, D.E., Edgar, T.F. and Mellichamp, D.A., *Process dynamics and control*, John Wiley and Sons, 1989.

[4] Isermann, R., Baur, U., Bamberger, W., Kneppo, P. and Seibert, H., Comparison of six on line identification and parameter estimation methods, *Automatica*, Vol. 10, 1974, pp. 81-103.

[5] Sundaresan, K.R. and Krishnaswamy, P.R., Estimation of time delay, time constant parameters in time, frequency and Laplace domains, *The Canadian Journal of Chemical Engineering*, Vol. 56, 1978, pp. 257-262.

[6] Koganezawa, K., On-line parameter identification of non-stationary continuous system with time-variant delay, Proceedings of the *1991 International Conference on Industrial Electronics*, Kobe, Japan, Vol. 3, 1991, pp. 1990-1993.

[7] Dos Santos, P.L. and De Carvalho, J.L.M., Automatic transfer function synthesis from a Bode plot, Proceedings of the *29th Conference on Decision and Control*, Honolulu, Hawaii, U.S.A., 1990, pp. 1093-1098.

[8] Lilja, M., Least squares fitting to a rational transfer function with time delay, *IEE Control Conference*, 1988, pp. 143-146.

[9] Palmor, Z.J. and Blau, M., An auto-tuner for Smith dead time compensator, *International Journal of Control*, 1994**,** Vol. 60, pp. 117-135.

Appendix 1: The matrix in equation (13) is

$$A = \begin{bmatrix}
1 & \omega_1^2 & . \; . & \omega_1^{2u} & -\omega_1^2 \left|G_p(j\omega_1)^2\right| & . \; . & -\omega_1^{2v}\left|G_p(j\omega_1)\right|^2 \\
1 & \omega_2^2 & . \; . & \omega_2^{2u} & -\omega_2^2 \left|G_p(j\omega_2)\right|^2 & . \; . & -\omega_2^{2v}\left|G_p(j\omega_2)\right|^2 \\
. & . & . \; . & . & . & . \; . & . \\
. & . & . \; . & . & . & . \; . & . \\
.1 & \omega_u^2 & . \; . & \omega_u^{2u} & -\omega_u^2 \left|G_p(j\omega_u)\right|^2 & . \; . & -\omega_u^{2v}\left|G_p(j\omega_u)\right|^2 \\
1 & \omega_{u+1}^2 & . \; . & \omega_{u+1}^{2u} & -\omega_{u+1}^2 \left|G_p(j\omega_{u+1})\right|^2 & . \; . & -\omega_{u+1}^{2v}\left|G_p(j\omega_{u+1})\right|^2 \\
. & . & . \; . & . & . & . \; . & . \\
. & . & . \; . & . & . & . \; . & . \\
1 & \omega_{u+v+1}^2 & . \; . & \omega_{u+v+1}^{2u} & -\omega_{u+v+1}^2 \left|G(j\omega_{u+v+1})\right|^2 & . \; . & -\omega_{u+v+1}^{2v}\left|G_p(j\omega_{u+v+1})\right|^2
\end{bmatrix}$$