Proceedings of the Irish DSP and Control Colloquium, Dublin City University, Dublin, Ireland, July 1994, pp. 211-218

Model Parameter And Time Delay Estimation Using Gradient Methods

A. O' Dwyer and J.V. Ringwood***

** Department of Control Systems and Electrical Engineering,*
Dublin Institute of Technology,
Kevin St., Dublin 8.

*** School of Electronic Engineering,*
Dublin City University,
Glasnevin, Dublin 9.

Abstract

A number of approaches have been proposed for parameter and time delay estimation of process models in single input, single output (SISO) control systems using gradient descent algorithms; some of these approaches involve the selection of a rational polynomial that is used to approximate time delay variations. This paper takes a generalised approach to the investigation of the most appropriate choice of the rational polynomial, and the gradient descent algorithm, to be used.

1. INTRODUCTION:

The identification and control of many processes are frequently complicated by the time varying nature of both the process model parameters and the time delay. One strategy developed for the on line estimation of the parameters (by Kurz and Goedecke [1], Wong and Bayoumi [2] and De Keyser [3], amongst others) is to overparameterise the model in the discrete time domain. An identification algorithm (e.g. recursive least squares) may then be used to estimate the model parameters, and the time delay may be estimated based on the leading coefficients of the numerator polynomial converging to zero. However, the number of parameters that need to be estimated increases as either the time delay or the sampling frequency increases, which slows convergence of the estimates.

An alternative strategy was proposed by Durbin [4], in which the process is assumed to be modelled by a first order lag plus time delay (FOLPD) model. The time delay variation from a starting value is approximated by a rational polynomial, and a Gauss-Newton gradient descent algorithm is used to estimate the model parameters (including the time delay). The treatment in this paper is based on the above proposal, with the focus being on the most appropriate rational polynomial, and gradient descent algorithm, to be used.

Gradient methods of parameter estimation are based on updating the parameter vector by a vector that depends on information on the cost function to be minimised. The cost function is normally equal to the sum of the square of the error between the process and model outputs. In the more general algorithms, the update vector depends on the cost function, the partial derivative of the cost function with respect to the parameter vector and the second partial derivative of the cost function with respect to the parameter vector; an example is the Newton-Raphson algorithm. Other algorithms use an estimate of the second partial derivative of the cost function with respect to the parameter vector (e.g. the Gauss-Newton algorithm and the Levenberg-Marquardt algorithm). A popular algorithm for estimation purposes is a special case of the Gauss-Newton technique known as the steepest descent method; this algorithm is used for time delay estimation by Elnagger, Dumont and Elshafei [5]. Another, simpler, estimation algorithm is the least mean squares (LMS) algorithm in which the update is a function of the first partial derivative of the error with respect to the parameter vector.

The gradient algorithms are implemented by finding the partial derivatives of the error between the process output and the discretized model output, with respect to the gain, time constant and time delay index value (time delay index equals time delay divided by sample period). The partial derivatives are subsequently used to update the model parameters. However, prior to calculating the partial derivative of the error with respect to the time delay index value, the time delay index variation (from a starting time delay index value) is approximated by a rational polynomial. The most appropriate rational polynomial to use may be determined by finding the relationship of the mean squared error (MSE) function between the process and model outputs, to the time delay index parameter; this relationship, which is determined both graphically and analytically, must be unimodal for successful application of gradient descent algorithms, as must the relationship of the MSE function to the gain and time constant values, respectively.

2. THE CHOICE OF A RATIONAL POLYNOMIAL:

The time delay variation, r, may be approximated by two separate first order rational polynomials, namely the first order Taylor's series approximation and the first order Pade approximation (see Appendix A). The MSE function between the process and model outputs was calculated analytically, when the time delay variation was approximated by each of these approximations in turn; it was discovered that the MSE performance surface was unimodal with respect to all parameters when the first order Taylor's series approximation was used (see Appendix B). This result was confirmed by calculating the mean square value of the error over a large number of samples, for these conditions.

The use of the first order Pade approximation produced a non-quadratic MSE performance surface, which in general is non-unimodal. The model in the z domain, using this approximation, has two poles; one pole is at $e^{-T/TC}$, which is always stable (TC = time constant, T = sample time). The other pole is at $e^{-T/(0.5 \cdot r)}$; however, this is unstable when r < 0, where simulations that calculate the mean squared value of the error over a large number of samples confirm the non-unimodality to be located. In addition, even if r > 0, unimodality is only possible theoretically if $r < 2 \cdot TC$. Thus, the first order Taylor's series approximation for the time delay variation is the most appropriate one to use.

The use of higher order approximations for the time delay variation is conditioned on the use of a higher order model for the process than a FOLPD model (for example, a second order lag plus time delay model); some second order approximations that may be used are the Taylor's series approximation, the Pade approximation and the Laguerre [6], Product [6], Direct Frequency Response [7] and Marshall [8] approximations (see Appendix A). It was discovered that when the mean squared error was calculated for varying time delay index values over a large number of samples for the second order approximations above (and using a second order lag plus time delay model), unimodality was achieved only when a second order Taylor's series approximation was used. As before, unimodality is related to the location of the poles of the model (in the z domain) for each approximation taken; the poles are always within the unit circle when the Taylor's series approximation is used, are on the unit circle when the approximation defined by Marshall is used, and are outside the unit circle for r < 0 (and depending on the approximation used, other ranges of r values) when all of the other approximations are used. The conclusion is that if the process is modelled by a second order lag plus time delay model, then either a first or second order Taylor's series approximation for the time delay variation is appropriate.

3. CONVERGENCE OF THE MODEL PARAMETER AND TIME DELAY ESTIMATES:

It has been shown above that if the process is modelled as a FOLPD model, the most appropriate approximation to use for the time delay variation is a first order Taylor's series approximation. The following theorems show that the correct process delay (to the nearest integer multiple of the sample period) and the correct process parameters may be identified, under these conditions.

Theorem 1: For a first order discrete stable system of known parameters, when the time delay variation is approximated by a first order Taylor's series approximation, and if the measurement noise is uncorrelated with the system input, then the MSE performance surface is minimised when the model time delay index equals the process time delay index. The resolution on the process time delay is assumed to be equal to one sample period.

Proof: The process difference equation is

$$y(n) = e^{-T/TC_p} \cdot y(n-1) + K_p \cdot (1 - e^{-T/TC_p}) \cdot u(n - g_p - 1) + w(n) \quad (1)$$

with TC_p = process time constant, K_p = process gain, g_p = process time delay index and $w(n)$ = noise term.

The model difference equation is

$$y_m(n) = e^{-T/TC_m} \cdot y(n-1) - \frac{K_m \cdot (g_m - g_p) \cdot T}{TC_m} \cdot u(n - g_m) - K_m \cdot (e^{-T/TC_m} - 1 - \frac{(g_m - g_p) \cdot T}{TC_m}) \cdot u(n - g_m - 1) \quad (2)$$

with TC_m = model time constant, K_m = model gain and g_m = model time delay index.

If the parameters are known, then $TC_p = TC_m = TC$ and $K_p = K_m = K$ and therefore

$$e(n) = K \cdot (1 - e^{-T/TC}) \cdot u(n - g_p - 1) + \frac{K \cdot (g_m - g_p) \cdot T}{TC} \cdot u(n - g_m) + K \cdot (e^{-T/TC} - 1 - \frac{(g_m - g_p) \cdot T}{TC} \cdot u(n - g_m - 1) + w(n) \quad (3)$$

The MSE performance surface, $E[e^2(n)]$, may then be calculated to be equal to

$$2\cdot K^2\cdot(1-e^{-T/TC})^2\cdot[r_u(0)-r_u(g_p-g_m)]+\frac{2\cdot K^2\cdot(g_m-g_p)^2\cdot T^2}{TC^2}\cdot[r_u(0)-r_u(1)]$$
$$+\frac{2\cdot K^2\cdot(1-e^{-T/TC})\cdot T}{TC}\cdot(g_m-g_p)\cdot[r_u(0)-r_u(1)+r_u(g_p-g_m+1)-r_u(g_p-g_m)]+r_w(0) \quad (4)$$

where $r_u(n)$ and $r_w(n)$ are the autocorrelation functions of u and w, respectively. Therefore, $E[e^2(n)]=r_w(0)$ for $g_m=g_p$. Now $r_u(0)\geq r_u(n)\forall n$ and for $g_m>g_p$, $E[e^2(n)]>r_w(0)$ since $r_u(g_p-g_m+1)>r_u(g_p-g_m)$. For $g_m<g_p$, it may be shown by comparing the size of the individual terms that $E[e^2(n)]>r_w(0)$ for all values of g_m and g_p. Thus, the minimum value of the MSE function occurs at $g_m=g_p$ and the coloured noise has no effect on the estimated delay value. The only situation that arises for which $E[e^2(n)]=r_w(0)$ for $g_m\neq g_p$ is when the input has a flat autocorrelation function, which corresponds to a constant level input. Thus, any input change is sufficient for correct delay estimation.

This theorem is similar to one developed by Elnagger, Dumont and Elshafei [5], for a FOLPD system in which the time delay is not approximated.

<u>Theorem 2</u>: For a first order discrete stable system of unknown parameters, when the time delay variation is approximated by a first order Taylor's series approximation, and if measurement noise is assumed to be absent, then the MSE performance surface is minimised when the model time delay index equals the process time delay index, provided certain conditions are observed on the model parameters.
<u>Proof</u>: From Appendix B, it may be shown that $(g_m-g_p)\cdot T$ is minimised when

$$(g_m-g_p)=\frac{TC_m}{2\cdot T}\cdot\left[\frac{K_p}{K_m}\cdot\frac{(1-e^{-T/TC_p})\cdot(1-e^{-2\cdot T/TC_m})}{(1-e^{-T/TC_m-T/TC_p})}-(1-e^{-T/TC_m})\right] \quad (5)$$

If the time delay is calibrated as integer multiples of the sample period, then $g_m=g_p$ when

$$\left|\frac{TC_m}{2\cdot T}\cdot\left[\frac{K_p}{K_m}\cdot\frac{(1-e^{-T/TC_p})\cdot(1-e^{-2\cdot T/TC_m})}{(1-e^{-T/TC_m-T/TC_p})}-(1-e^{-T/TC_m})\right]\right|<1 \quad (6)$$

If $K_m=a\cdot K_p$, $TC_m=b\cdot TC_p$ and $T=0.1\cdot TC_m$ then it may be shown that a and b must satisfy the equation

$$1.6309+(1/a-1.4756)\cdot e^{-0.1\cdot b}-(1/a)>0. \quad (7)$$

This gives a lower limit of a = 0.6, with b = 0.1. An upper limit of a and b is undefined, though an upper limit of b = 3.0 will adequately fulfil the Nyquist criterion. Thus, convergence of the model time delay index to the process time delay index is possible when $0.1\cdot TC_m\leq TC_p\leq 3\cdot TC_m$ and $0.6\cdot K_m\leq K_p$, where TC_m and K_m are the values of the model time constant and gain, respectively, at the start of the identification process.

The presence of measurement noise may result in a narrowing of the allowed range of the above parameters. If a more conservative bound of 0.5 is placed on $|g_m-g_p|$, then $0.1\cdot TC_m\leq TC_p\leq 3\cdot TC_m$ and $0.8\cdot K_m\leq K_p$.

It may also be shown that the identification method proposed will only facilitate identification of the time delay to a resolution equal to the sample period. A generalisation of Theorem 1 demonstrates that exact identification of the time delay, when coloured measurement noise is present, is possible either when the other model parameters equal the corresponding process parameters or if the model time constant equals infinity with the process being excited by a white noise signal. It is interesting that Elnagger, Dumont and Elshafei [9] prove that unbiased time delay estimation is possible in the presence of coloured noise when the model time constant equals infinity, for a FOLPD system in which the time delay is not approximated (when the excitation signal is white noise).

<u>Theorem 3</u>: For a first order discrete stable system of unknown parameters, when the time delay variation is approximated by a first order Taylor's series approximation, and if measurement noise is assumed to be absent, then the MSE performance surface is minimised when the model gain equals the process gain and the model time constant equals the process time constant, provided the model time delay index equals the process time delay index.
<u>Proof</u>: From Appendix B, it may be shown that when $g_m=g_p$, the MSE function is minimised with respect to K_m when

$$K_m = \frac{K_p \cdot (1-e^{-T/TC_p}) \cdot (1-e^{-2\cdot T/TC_m})}{(1-e^{-T/TC_m}) \cdot (1-e^{-T/TC_m - T/TC_p})} \tag{8}$$

Similarly, it may be shown that under the same conditions, the MSE function is minimised with respect to TC_m when

$$K_m = \frac{K_p \cdot (1-e^{-T/TC_p})^2 \cdot (1-e^{-2\cdot T/TC_m})^2}{(1-e^{-T/TC_m})^2 \cdot (1-e^{-T/TC_p - T/TC_m})^2} \tag{9}$$

It is straightforward to demonstrate that the MSE function is minimised with respect to both parameters when $TC_p = TC_m$ and $K_p = K_m$.

4. GRADIENT DESCENT ALGORITHMS:

The following gradient descent algorithms were used for the estimation of the parameters, $\theta^T(n) = [K_m(n) \;\; 1/TC_m(n) \;\; g_m(n)]$, of a FOLPD model.

4.1. Levenberg-Marquardt algorithm / Gauss-Newton algorithm [10]:

$$\theta(n+1) = \theta(n) + \alpha(n)\cdot e(n) \tag{10}$$

with

$$\alpha(n) = \frac{-H^{-1}(\theta, n-1)\cdot \dfrac{\partial e(n)}{\partial \theta(n)} \cdot \mu}{\lambda(n) + \left[\dfrac{\partial e(n)}{\partial \theta(n)}\right]^T \cdot H^{-1}(\theta, n-1) \cdot \left[\dfrac{\partial e(n)}{\partial \theta(n)}\right]} \tag{11}$$

and

$$\lambda(n) = 0.995\cdot\lambda(n-1) + 0.005, \tag{12}$$

and

$$H^{-1}(\theta, n) = \frac{1}{\lambda(n)} \cdot \left[H^{-1}(\theta, n-1) + \alpha(n) \cdot \left[\frac{\delta e(n)}{\delta \theta(n)}\right] \cdot H^{-1}(\theta, n-1) \right] + \delta \cdot I \tag{13}$$

with $\theta, \alpha \in \Re^n$ and $H, I \in \Re^{nxn}$. In addition, when K_m and TC_m are being updated,

$$e(n) = y(n) - e^{-T/TC_m} \cdot y(n-1) - K_m \cdot (1 - e^{-T/TC_m}) \cdot u(n - g_m - 1) \tag{14}$$

and when g_m is being updated

$$e(n) = y(n) - e^{-T/TC_m} \cdot y(n-1) + \frac{K_m \cdot (g_m - g_p)}{TC_m} \cdot u(n - g_m) + K_m \cdot (e^{-T/TC_m} - 1 - \frac{(g_m - g_p)}{TC_m}) \cdot u(n - g_m - 1) \tag{15}$$

Also, $\mu = 4.0$ (when K_m is being updated), $\mu = 2.5$ (when TC_m is being updated) and $\mu = 0.4$ (when g_m is being updated). In addition, $H^{-1}(\theta,0) = 25\cdot I$, $\delta = 0.001$, $\lambda(0) = 0.90$ and $\theta(0)$ = known starting values. The values of μ, δ, $H^{-1}(\theta,0)$, $\lambda(0)$ and the equation for $\lambda(n)$ quoted were found from simulation results to be appropriate to the application.

The Gauss-Newton algorithm omits the addition of the $\delta \cdot I$ term.

4.2. Steepest descent algorithm [10]:

$$\theta(n+1) = \theta(n) + \alpha(n)\cdot e(n) \tag{16}$$

with

$$\alpha(n) = \frac{-H^{-1} \cdot \dfrac{\partial e(n)}{\partial \theta(n)} \cdot \mu}{\lambda(n) + \left[\dfrac{\partial e(n)}{\partial \theta(n)}\right]^T \cdot H^{-1} \cdot \left[\dfrac{\partial e(n)}{\partial \theta(n)}\right]} \tag{17}$$

and with $\theta, \alpha \in \Re^n, H \in \Re^{nxn}$. In addition, $\mu = 4.0$ (when K_m is being updated), $\mu = 2.5$ (when TC_m is being updated) and $\mu = 0.03$ (when g_m is being updated). Also, $\lambda(n) = 0.90$, $H^{-1} = I$ and $\theta(0)$ = known starting values. As before, the value of $\lambda(n)$ and μ quoted were found from simulation.

4.3. LMS algorithm [11]:

$$\theta(n+1) = \theta(n) + \alpha(n) \cdot e(n) \quad \text{with} \quad \alpha(n) = -2 \cdot \frac{\partial e(n)}{\partial \theta(n)} . \mu \tag{18}$$

and with $\theta, \alpha \in \Re^n$. In addition, $\mu = 0.25$ (when K_m and TC_m are being updated) and $\mu = 0.005$ (when g_m is being updated). These values were found from simulation.

5. RESULTS:

The simulation results reported below were performed in MATLAB. The starting parameters of the process and model were defined to be equal, with $K_p = K_m = 1.0$, $TC_p = TC_m = 1.0$ and $g_p = g_m = 0$. A step change is made in the process parameters to $K_p = 2.0$, $TC_p = 0.7$ and $g_p = 100$. Noise free conditions were assumed. The time delay index variation (which equals 100 in this case) was approximated by a first order Taylor's series approximation and the four gradient algorithms were used to track the changing model parameters. Limits were placed on the allowed variation of the model parameters according to the results of Theorem 2.

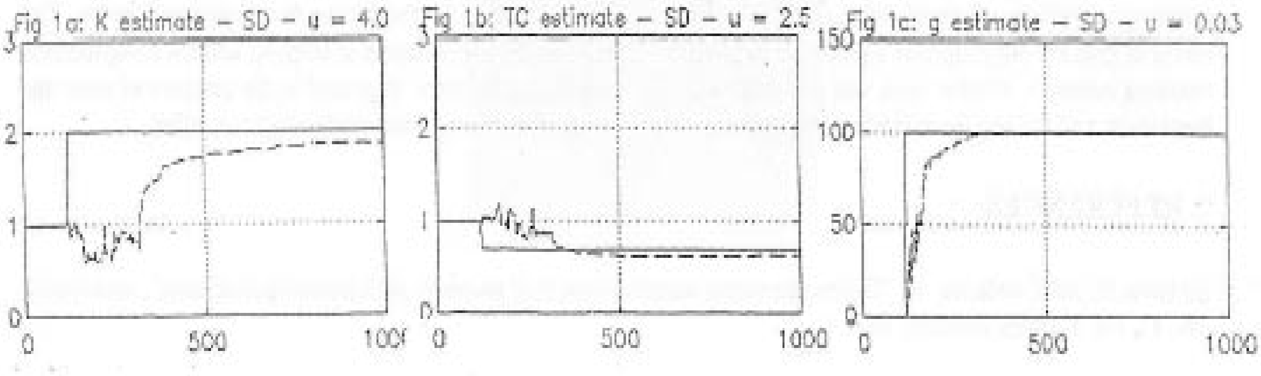
The time delay index was allowed to vary by up to ten sample periods per sample (which is a form of filtering on the time delay index). The values of the time constant and gain estimates were also filtered, by passing the parameters through a first order filter given by

$$\Phi^{k+1} = \alpha \cdot \Phi^k + (1 - \alpha) \cdot \Phi \tag{19}$$

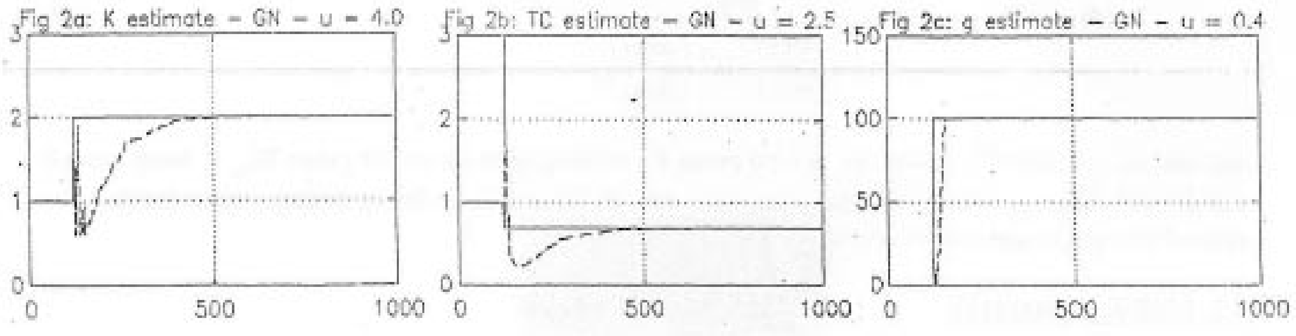
with Φ^k = vector of previous parameter estimates and Φ = updated vector of parameter estimates. α is a filter factor with a value between 0.0 and 1.0. Vogel and Edgar [12] suggest that filtering of the parameter estimates is appropriate to prevent relatively large fluctuations of the parameter values that may result in large, sudden changes in corresponding adaptive controller parameters. Durbin [4] implements a similar scheme to that given above. The choice of the filter factor was defined equal to 0.5 in the simulations.

Figures 1a to 4c below show the tracking of changing process parameters using the respective gradient algorithms. These results demonstrate that the time delay index estimate must converge before the other model parameter estimates may converge, which is in agreement with Theorem 3. As expected, each of the gradient algorithms facilitated correct tracking of the changing process parameters.

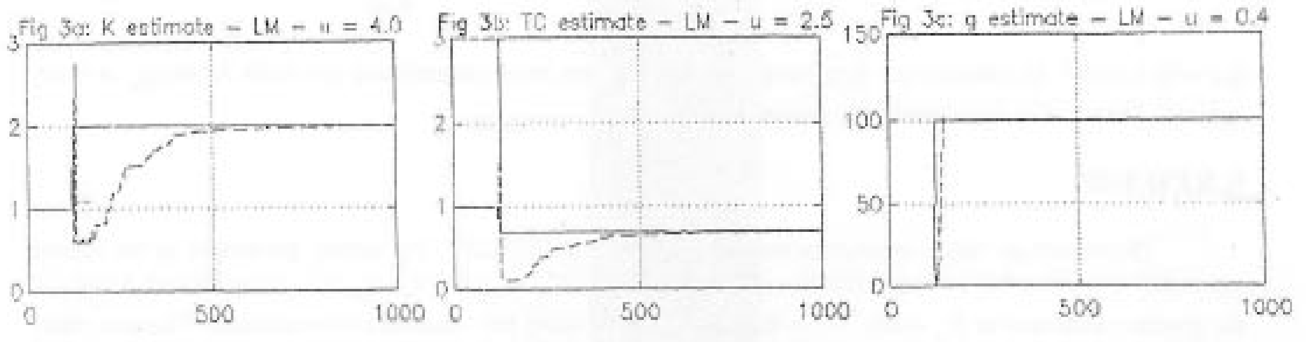
5.1 Steepest descent algorithm:



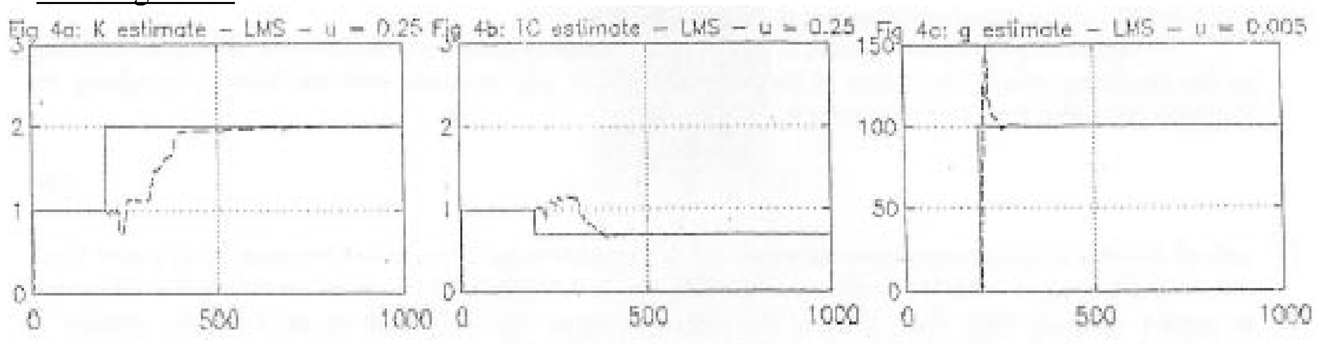
5.2 Gauss-Newton algorithm:



5.3 Levenberg-Marquardt algorithm:



5.4 LMS algorithm:



6. CONCLUSIONS:

This paper has shown that the most appropriate choices of first and second order rational polynomials to use to approximate the time delay variation of a process modelled by an appropriate model, if the model parameters and time delay are to estimated using a gradient descent algorithm, are the first and second order Taylor's series polynomials, respectively. Convergence of the time delay estimate is guaranteed in the presence of coloured measurement noise when the model gain and time constants equal the process gain and time constants, respectively. When the model and process parameters differ, convergence of the time delay estimate is guaranteed when the starting model gain and time constants are within a defined range about the corresponding parameters, in the absence of measurement noise. The choice of gradient algorithm for a particular application depends on the desired speed of tracking and the computational resources available. Further work will concentrate on the implementation of the algorithm in the presence of noise and bias inputs, and the implementation of the algorithm when the process structure and model structure differ.

7. REFERENCES:

[1] Kurz, H. and Goedecke, W. "Digital parameter-adaptive control of processes with unknown dead time", *Automatica*, Vol. 17, No. 1, pages 245-252, 1981.

[2] Wong, K.Y. and Bayoumi, M.M. "A self tuning control algorithm for systems with unknown time delay", *Proceedings of the IFAC Identification and System Parameter Estimation Conference*, pages 1193-1198, Washington, D.C., U.S.A., 1982.
[3] De Keyser, R.M.C. "Adaptive dead-time estimation*", Proceedings of the IFAC Adaptive Systems in Control and Signal Processing Conference*, pages 385-389, Lund, Sweden, 1986.
[4] Durbin, L.D. "Deadtime approximations with adaptive deadtime compensation", *Proceedings of the American Control Conference*, Vol. 3, pages 1707-1712, 1985.
[5] Elnagger, A., Dumont, G.A. and Elshafei, A.-L. "System identification and adaptive control based on a variable regression for systems having unknown delay", *Proceedings of the 29th IEEE Conference on Decision and Control*, Part 3 of 6, pages 1445-1450, Honolulu, Hawaii, U.S.A., December 5-7, 1990.
[6] Piche, R. "Low-order rational all pass approximations to e^{-s}",*Proceedings of the 11th IFAC World Congress on Automatic Control*, Vol. 2, pages 134-137, Tallinn, Estonia, 1990.
[7] Stahl, H. and Hippe, P. "Design of pole placing controllers for stable and unstable systems with pure time delay", *International Journal of Control*, Vol. 45, No. 6, pages 2173-2182, 1987.
[8] Marshall, J.E., *Control of time-delay systems*, IEE Control Engineering Series. Peter Peregrinus Ltd., 1979.
[9] Elnagger, A., Dumont, G.A. and Elshafei, A.-L. "Delay estimation using variable regression", *Proceedings of the American Control Conference*, pages 2812-2817, Mass., USA, June 1991.
[10] Ljung, L., *System identification: theory for the user*, Prentice-Hall Inc., 1987.
[11] Widrow, B. and Stearns, S.D., *Adaptive signal processing*, Prentice-Hall Inc., 1985.
[12] Vogel, E.F. and Edgar, T.F. "Application of an adaptive pole-zero placement controller to chemical processes with variable dead time", *Proceedings of the American Control Conference*, pages 536-544, 1982.

APPENDIX A: Some first and second order rational approximations for the time delay.

A1: First order approximations:

Taylor: $e^{-s\cdot r} \approx 1 - s\cdot r$

Pade: $e^{-s\cdot r} \approx \dfrac{1-0.5\cdot s\cdot r}{1+0.5\cdot s\cdot r}$

A2: Second order approximations:

Taylor: $e^{-s\cdot r} \approx 1 - s\cdot r + 0.5\cdot s^2\cdot r^2$

Pade: $e^{-s\cdot r} \approx \dfrac{1-0.5\cdot s\cdot r+0.0833\cdot s^2\cdot r^2}{1+0.5\cdot s\cdot r+0.0833\cdot s^2\cdot r^2}$

Laguerre [6]: $e^{-s\cdot r} \approx \dfrac{1-0.5\cdot s\cdot r+0.0625\cdot s^2\cdot r^2}{1+0.5\cdot s\cdot r+0.0625\cdot s^2\cdot r^2}$

Product [6]: $e^{-s\cdot r} \approx \dfrac{1-0.5\cdot s\cdot r+0.1013\cdot s^2\cdot r^2}{1+0.5\cdot s\cdot r+0.1013\cdot s^2\cdot r^2}$

Direct Freq. Response [7]: $e^{-s\cdot r} \approx \dfrac{1-0.49\cdot s\cdot r+0.0953\cdot s^2 . r^2}{1+0.49\cdot s\cdot r+0.0953\cdot s^2\cdot r^2}$

Marshall [8]: $e^{-s\cdot r} \approx \dfrac{1-0.0625\cdot s^2\cdot r^2}{1+0.0625\cdot s^2\cdot r^2}$

APPENDIX B: Analytical calculation of the performance function when the time delay variation is approximated by a first order Taylor's series.

The process transfer function $G_p(s) = \dfrac{K_p\cdot e^{-s\cdot TD_p}}{1+s\cdot TC_p} = \dfrac{Y(s)}{U(s)}$ may be calculated in the z domain to be

$$G_p(z) = \frac{K_p\cdot(1-e^{-T/TC_p})\cdot z^{-g_p}}{z-e^{-T/TC_p}} \qquad \text{(B1)}$$

assuming $TD_p = g_p\cdot T$, T = sample time.

The model transfer function $G_m(s) = \dfrac{K_m\cdot e^{-s\cdot TD_m}}{1+s\cdot TC_m}$ may be calculated in the z domain to be

$$G_m(z) = \frac{-K_m\cdot[(e^{-T/TC_m}-1-\frac{r}{TC_m})+\frac{r}{TC_m}\cdot z]\cdot z^{-g_p}}{z-e^{-T/TC_m}} \qquad \text{(B2)}$$

with $TD_m = g_p\cdot T + r$ and $e^{-r\cdot s} = 1 - r\cdot s$.

The input to both processes is assumed to be a white noise signal. The procedure defined by Widrow and Stearns [11] is used to calculate the MSE performance surface i.e.

$$E[e^2(n)] = \phi_{yy}(0) + \frac{1}{2 \cdot \pi \cdot j} \oint [G_m(z^{-1}) \cdot \Psi_{uu}(z) \cdot G_m(z) - 2 \cdot \Psi_{yu}(z) \cdot G_m(z)] \cdot \frac{dz}{z} \qquad \text{(B3)}$$

with $\phi_{yy}(0) = E[y(n) \cdot y(n)]$, $\Phi_{uu}(z) = \sum_{n=-\infty}^{\infty} \phi_{uu}(n) \cdot z^{-n}$, $\Phi_{yu}(z) = \sum_{n=-\infty}^{\infty} \phi_{uy}(n) \cdot z^{-n}$

Using the residue theorem to calculate the closed curve integral, it may be determined that the MSE performance surface is quadratic in r and is thus unimodal in this variable i.e.

$$E[e^2(n)] = A \cdot \left(\frac{r}{TC_m}\right)^2 + B \cdot \left(\frac{r}{TC_m}\right) + C \qquad \text{(B4)}$$

with $$A = \frac{2 \cdot K_m{}^2 \cdot (1 - e^{-T/TC_m})}{1 - e^{-(2 \cdot T)/TC_m}} \qquad \text{(B5)}$$

$$B = \frac{2 \cdot K_m{}^2 \cdot (1 - e^{-T/TC_m})^2}{1 - e^{-(2 \cdot T)/TC_m}} - \frac{2 \cdot K_m \cdot K_p \cdot (1 - e^{-T/TC_m}) \cdot (1 - e^{-T/TC_p})}{1 - e^{-T/TC_m - T/TC_p}} \qquad \text{(B6)}$$

$$C = \frac{K_m{}^2 \cdot (1 - e^{-T/TC_m})^2}{1 - e^{-(2 \cdot T)/TC_m}} - \frac{2 \cdot K_m \cdot K_p \cdot (1 - e^{-T/TC_m}) \cdot (1 - e^{-T/TC_p})}{1 - e^{-T/TC_m - T/TC_p}} + \frac{K_p{}^2 \cdot (1 - e^{-T/TC_p})^2}{1 - e^{-(2 \cdot T)/TC_p}} \qquad \text{(B7)}$$

Similarly, it may be found that the performance surface is unimodal with respect to the model gain, K_m, and the model time constant, TC_m. It is not necessary to approximate the time delay variation for these calculations. Plots of the MSE performance surface versus each parameter are shown below.