A Real-Time QR implementation of a Constrained ARMA model for Line Enhancement

R.Geere and W.J.Fitzgerald
Department of Engineering
University of Cambridge
Cambridge U.K.

Abstract

In this paper, a new class of models is presented for the estimation of model parameters for narrow band signals in additive white noise. The new class of models are constrained ARMA models and these have been implemented using the QR algorithm using both off-line and real-time processing. The performance of this class of model will be described and applications for adaptive line enhancement will be given.

1 Introduction

Much effort has been devoted to the enhancement of narrow band signals corrupted by additive noise. This problem clearly has many applications ranging from Sonar to Biomedical signal processing, where signal spectra can be highly non-stationary. In general, the number and frequencies of the narrow band components may not be known, and indeed, these may change with time. The use of fixed filters is clearly inappropriate, and adaptive filters which can respond to the time series and optimise their transfer functions are necessary.

An appropriate parametric model for narrow band signals is the autoregressive (AR) model, and much work has taken place considering the spectral estimation of such signals that can be modelled by such a process, leading to AR spectral estimation and Maximum entropy methods. However, for real signals buried in noise, this simple model is not correct. A sum of L sinusoids can be written as;

$$s_n = \sum_{j=1}^{j=2L} f_j s_{n-j}$$

which defines an autoregressive process of order 2L. Adding white noise, u_n, results in the process

$$x_n = u_n + s_n = u_n + \sum_{j=1}^{j=2L} f_j s_{n-j} = \sum_{j=1}^{j=2L} f_j x_{n-j} - \sum_{j=1}^{j=2L} f_j u_{n-j} + u_n$$

Thus the appropriate model for narrow bandwidth signals in white Gaussian noise is an Auto-Regressive Moving Average (ARMA) model where the filter coefficients of the AR and MA parts are the same. The power spectral density $P_{xx}(z)$ of the process x_n can be written

$$P_{xx}(z) = \frac{[1 - F(z)][1 - F^*(1/z^*)]}{[1 - F(z)][1 - F^*(1/z^*)]}\rho_u$$

where $F(z)$ is the z-transform of the filter coefficients, and ρ_u is the power of the white noise process. Although accurate for pure tonals in white noise, the pole-zero degeneracy on the unit circle prevents the power spectral density being defined at all frequencies and creates numerical difficulties in estimating the model parameters. One way of extending the model to take account of finite bandwidth signals, is to separate the poles and zeroes in certain ways. The constraint introduced by Nehorai (1985), allows the poles and zeroes to move independently along a radius of the unit circle. This constraint was also discussed by Thompson (1979).

In this paper, a new constraint is introduced which still preserves the features of a constrained ARMA (CARMA) process, and has the form

$$x_n = \sum_{j=1}^{j=2L} (1+\lambda) f_j x_{n-j} - \sum_{j=1}^{j=2L} \lambda f_j u_{n-j} + u_{n-j}$$

Poles and zeroes still occur in pairs, but the problems resulting from pole-zero cancellation are avoided, with poles occuring at the solutions of the equation $(1+\lambda)F(z) = 1$ and zeroes at $\lambda F(z) = 1$.

In this class of model, λ defines the resolution of the spectral estimator, with the filter bandwidth narrowing as λ increases. In the limit of infinite λ, the degenerate case is reached with infinitely narrow resonances centered on the tonals, with the poles and zeroes coinciding. Decreasing λ moves the poles closer to the origin, with the limit as λ tends to zero corresponding to an autoregressive model.

As can be seen from the form of the transfer function, this model has interesting properties. By changing the value of the parameter λ one can modify the form of the model that one is trying to fit the data with. Since signal parameters, such as frequency and amplitude, of real signals, can be time dependent, an obvious method for the solution of the filter coefficients given above, is to use adaptive signal processing methods. The more commonly used adaptive algorithms assume that the signal to be detected can be modelled as an AR process, and hence the FIR adaptive filter is the appropriate 'matched filter' for this model. In the case of our constrained (CARMA) model, an adaptive filter of the form shown in Fig 1. is used. It should be noted that the form of constraint used by Nehorai (1985) is much more computationally intensive than our constraint.

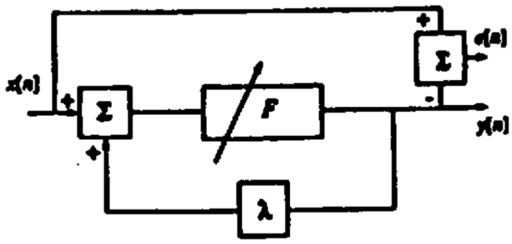

Fig 1. The line enhancement linear prediction filter for the CARMA process

However, for IIR filters the error surface is not, in general, unimodal, and much effort has been directed to the problem of finding efficient algorithms for the solution of the IIR filter coefficients. The adaptive line enhancer is basically a least squares predictor for the observed time series, and in its realisation as a non-recursive tapped delay line, it lends itself to simple adaptation procedures. Most adaptive IIR algorithms proposed for use with recursive filters have been derived from a direct form implementation of the filter coefficients. However, some disadvantages of implementation with respect to the direct recursive filter structure, such as finite precision effects and the complexity of stability monitoring, have led to the development of alternative realisations. In this paper, a real-time implementation of the constained (CARMA) filter given above is described. One method chosen for the solution of the filter coefficients was a method based on the QR decomposition, although other methods were also implemented, Fitzgerald and Pedley (1991). The next section will decribe the theory and implementation of the QR alogorithm for our form of transfer function, and the following section will describe the results obtained using this model.

2 Theory

Solutions to the constrained ARMA model can be obtained by minimising the power of the error process, from the spectral inverting filter, which is one minus the filter transfer function, and is given by

$$H_e = \frac{1-(\lambda+1)F}{1-\lambda F}$$

The error power $\zeta = E\{e[n]^2\}$ is defined by the inverse Wiener-Khintchine relation between the power spectral density and the autocorrelation sequence

$$\zeta = \frac{1}{2\pi i}\oint P_{xx}(z)H_e(z)H_e^*(1/z^*)dz/z$$

which from above can be written as

$$\zeta = \frac{1}{2\pi i}\oint P_{xx}(z)\frac{[1-(\lambda+1)F(z)][1-(\lambda+1)F^*(1/z^*)]}{[1-\lambda F(z)][1-\lambda F^*(1/z^*)]}$$

At the least squares configuration, the error power will be at a minimum defined by the set of M constraints given by

$$\frac{\partial\zeta}{\partial f_j} = \frac{-1}{\pi i}\oint P_{xx}(z)\frac{[1-(\lambda+1)F(z)]z^j}{[1-\lambda F(z)][1-\lambda F^*(1/z^*)]^2}$$

for j=1,M.

In the adaptive filtering problem, the 'data matrix' has a Toeplitz structure so that each row of the data matrix contains only one new datum when compared to the previous row. Various algorithms have been devised that take advantage of this structure to reduce the computational load , for a p th order filter from $O(p^2)$ to $O(p)$, but many of these algorithms are not well-conditioned, and can led to numerical instabilities. It is possible to solve the least squares minimisation problem using the QR decomposition, which has the advantage that it operates on the data matrix directly, rather than on the corresponding covariance matrix, and only involves orthogonal rotations, which are numerically stable.

To produce a QR decomposition, a series of rotations Q are applied to a matrix to produce an upper triangular form R. The solution of a series of linear equations is straightforward, and even when the system is singular the least squares solution is readily found.

Givens rotations are used, and each rotation can be represented by a 2-D rotation of two of the rows i and j, with all other rows unaltered.

$$\begin{pmatrix} r_i' \\ r_j' \end{pmatrix} = \begin{pmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{pmatrix}\begin{pmatrix} r_i \\ r_j \end{pmatrix}$$

Q is a product of such rotations such that

$$QX = R$$

where X is the original matrix. The rotations are numerically stable and Q^{-1} can readily be calculated if required.

Such a decomposition requires $O(n^3)$ operations. The method can be applied to non square matrices. If an extra row is added to X to give X' then a series of n extra Givens rotations Q' will be required to produce the new decomposition.

$$QX = R$$

$$Q'\begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} X \\ x \end{pmatrix} = Q'\begin{pmatrix} R \\ x \end{pmatrix} = \begin{pmatrix} R' \\ 0 \end{pmatrix}$$

where 0 represents zero vectors of the appropriate dimensions. The update of the decomposition requires $O(n^2)$ operations.

The filtering problem for a three element filter and a data series $x_0, x_1, ---, x_n$ is to choose $y_1 --- y_3$ so that the filter error vector e has minimum norm.

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ . \\ . \\ . \\ y_n \end{pmatrix} - \begin{pmatrix} x_2 & x_1 & x_0 \\ x_3 & x_2 & x_1 \\ x_4 & x_3 & x_2 \\ . & . & . \\ . & . & . \\ . & . & . \\ x_n & x_{n-1} & x_{n-2} \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ . \\ . \\ . \\ e_n \end{pmatrix}$$

where y is the desired filter output.

The adaptive filtering problem will normally weight the error vector in some appropriate way (e.g. exponentially) and adjusts the filter weights as successive data inputs, i.e. rows of X, are added. Each row of X is just the previous row shifted to the right and one new value, x_n, added. This form allows a fast algorithm to be developed, which only requires $O(n)$ operations for each new row added.

For our real-time implementation using the QR decomposition, a DSP32C processor installed in an IBM/PC was used together with a Tektronix graphics board for the real-time display of the filter transfer function. The QR algorithm produces the output of the filter, and every now and then the filter weights are frozen and the filter excited with a delta function to give the transfer function which is then displayed. For the current real-time system the sampling rate used is 512 Hz for a filter with 256 weights.

The QR algorithm implementation is described in this paper due to its much faster response compared to other algorithms. This fast reaction time together with many other features of the algorithm will prove to be important for many applications.

3 Results

The results obtained using the QR algorithm, for the solution of the constrained ARMA (CARMA) model, using off-line simulated data, are shown in Fig 2., where the filter transfer function is plotted against frequency for various values of λ. The individual curves correspond to a) $\lambda = 0$, b) $\lambda = 4$, c) $\lambda = 8$, d) $\lambda = 16$. It will be seen that the transfer function corresponding to $\lambda = 0$ has the familiar FIR shape, with a broad central lobe, and high side-lobes. The effect of increasing λ is to sharpen up the central line, and also to depress the side-lobe levels. The data used was a single sinusoid centered at 60 Hz buried in noise, and the signal-to-noise ratio was -10 dB. Fig 3. show the results obtained for a fixed value of λ and a varying number of samples, corresponding to a) 1,000 points, b) 10,000 points, and c) 100,000 points.

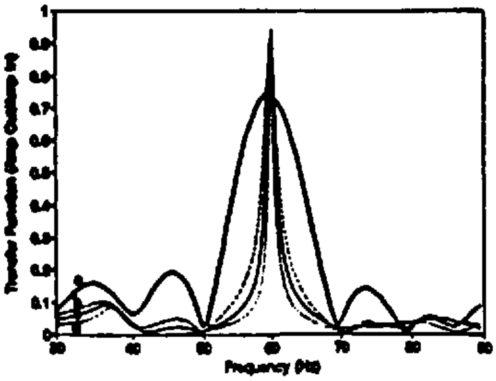

Fig 2.

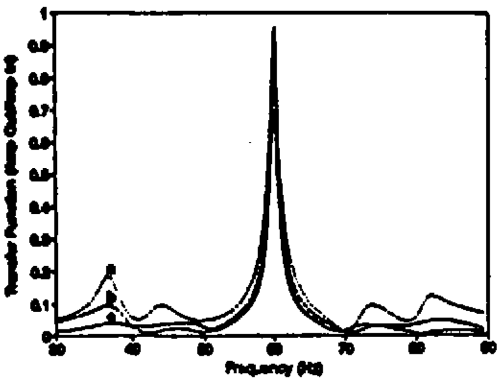

Fig 3.

The length of the filter was 52, and the sampling rate was 512 Hz. The algorithm has operated with many closely spaced tonals, and due to the extremely sharp central lobe in the transfer function, the filter can achieve super resolution, with also the added advantage of a reduction in false alarm rate since the side-lobes are very significantly reduced.

4 Conclusions

In this paper we have introduced a new constrained pole-zero (CARMA) model which is the appropriate model for finite bandwidth sinusoids in white Gaussian noise. The model has the feature that the bandwidth of the signal can be tuned using the parameter λ, and in the limit as λ tends towards zero, the standard FIR filter is obtained. The model has been implemented to real-time on a DSP32C, and the method choosen in this paper for the solution of the adaptive filter coefficients was based on the QR decomposition, although other implementations have also been used and will be reported separately.

The model has shown very good performance for data consisting of many closely spaced tonals in the presence of high noise levels and the model has many useful applications ranging from sonar to biomedical processing.

5 References

1. Nehorai, A., IEEE Trans. ASSP, 33, 983, 1985.

2. Thompson, P.A. Ph.D. Thesis, Dept. of Electrical Engineering, Stanford University, 1979.

3. Fitzgerald, W.J, and Pedley, M. Submitted to IEE Proceedings F, 1991.