Derivation of the
Fast Pole-Zero (ARMA) Recursive Least Squares
Algorithm
Using Geometric Projections

by

Sasan H. Ardalan

Center for Communications and Signal Processing
Department of Electrical and Computer Engineering
North Carolina State University
Box 7914
Raleigh, NC 27695-7914

February 1985

CCSP-TR-85/3

ABSTRACT

The Fast Pole-Zero (ARMA) Recursive Least Squares Algorithm is derived using geometric projections. The new algorithm has a reduced number of computations compared to the algorithm obtained using matrix methods. In particular two inner products are eliminated. This is particularly important in multiprocessor implementations of the algorithm. The geometric approach provides insight and useful interpretations of the various filters and residuals that form the algorithm. Such insights are absent in the matrix derivation. Simulations verifying the performance of the algorithm are presented where the poles and zeroes of an IIR filter are estimated. The algorithm finds application in echo cancellation and system identification.

1. Introduction

The Recursive Least Squares algorithm has been applied to a wide range of adaptive filtering problems. In many of these applications the system impulse response is estimated. In these cases the RLS algorithm has the structure of transversal filters and fast estimation schemes have been derived for efficient solution of these problems. These include the work by [1] and [2]. In the latter work the RLS problem is formulated geometrically and using geometric projections efficient solutions are derived. The resulting algorithms have fewer computations than the fast Kalman algorithm derived in [3] and used in [1]. Furthermore, the geometric approach provides insight into the interpretation and significance of the various filters and quantities originally encountered in [3].

Recently the RLS algorithm was applied to echo cancellation using pole-zero modeling [4,5]. In this case the fast Kalman algorithm derived in 3 is formulated to include the echo in addition to the echo generation signal to estimate the parameter vector containing the numerator and denominator coefficients of the echo path transfer function. In other words the echo path is modeled as an ARMA (pole-zero) process. The excellent performance of this algorithm in achieving very rapid and extremely high echo cancellation [6,4] has provided motivation for deriving a more efficient and robust algorithm.

In this paper, the geometric approach used in [7,8] and [2] is extended to the derivation of a fast pole-zero (ARMA) Recursive Least Squares algorithm. The work is also an extension of the formulation presented in [3] where a fast estimation scheme is developed using permutation matrices. The pole-zero algorithm differs from the multi-channel Fast Transversal Filter of [2] in that the order of poles and zeroes can be independently and arbitrarily specified. However, the same efficiency in terms of the elimination of some vector inner products and other operations in comparison with the fast Kalman algorithm are achieved as in the case of the Fast Transversal Filter.

Also in this paper, an alternative derivation of the projection update formula is presented using matrix methods.

2. Pole Zero Modeling

In the pole zero modeling of the linear system S, the transfer function H(z) is written as,

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum\limits_{i=0}^{i=M} b_i z^{-i}}{1 + \sum\limits_{i=1}^{i=L} a_i z^{-i}} \qquad (2.1)$$

Hence we can write in the discrete time domain,

$$y(n) = \sum_{i=0}^{i=M} b_i x(n-i) - \sum_{i=1}^{i=L} a_i y(n-i) \qquad (2.2)$$

Define the vector (N=M+L), note that "'" denotes the transpose of a vector,

$$\underline{z}(n) = [x(n)x(n-1)...x(n-M)y(n-1)...y(n-L)]' \quad Nx1 \qquad (2.3)$$

and

$$W_N^* = [a_0 a_1...a_M -b_1...-b_L]' \qquad (2.4)$$

Then we can write y(n) in vector notation as an inner product,

$$y(n) = \underline{z}'(n)W_N^* \qquad (2.5)$$

3. Geometric Projection Formulation

The solution to the RLS adaptive-filtering problem is stated as any vector $W_{N,T}$ for which

$$\xi_N(T) = \sum_{t=0}^{T} [d(t)+W_{N,T}\underline{z}(t)]^2 \qquad (3.1)$$

is minimized. In the case of pole/zero modeling the desired signal d(t)=y(t) the current system output sample.

Define the following vectors,

$$x_T = [x(T)x(T-1)...x(0)0...0]' \quad Jx1 \qquad (3.2)$$

$$y_T = [y(T-1)y(T-2)...y(0)0...0]' \quad Jx1 \qquad (3.3)$$

The above vectors group all measurements of x(n) and y(n-1) up to time T into vectors. Define,

$$X_{M,T} = [x_T x_{T-1}...x_{T-M+1}] \quad JxM \qquad (3.4)$$

$$Y_{L,T} = [y_T y_{T-1}...y_{T-L+1}] \quad JxL \qquad (3.5)$$

$$Z_{N,T} = [x_T x_{T-1} \cdots x_{T-M+1} y_{T-1} \cdots y_{T-L}] \quad Jx(M+L); \quad JxN \qquad (3.6)$$

$$Z_{N,T} = [X_{M,T} \quad Y_{L,T}] \quad JxN \qquad (3.7)$$

Note that N=M+L. Furthermore define the vector,

$$z_T = [x_T \quad y_T] \quad Jx2 \qquad (3.8)$$

which collects all measurements into a single vector. Also define the vector of desired response,

$$d_T = [d(T)d(T-1)\ldots d(0)0\ldots 0]^T \qquad (3.9)$$

Hence the prediction error becomes,

$$\varepsilon_{N,T} = d_T + Z_{N,T} W'_{N,T} \qquad (3.10)$$

From which,

$$\xi_N(T) = \varepsilon'_{N,T} \varepsilon_{N,T} \qquad (3.10)$$

In a vector space interpretation the minimizing solution is the $W_{N,T}$ that forms a linear combination of the input vectors

$$z_T, z_{T-1}, \ldots, z_{T-N-1} \qquad (3.11)$$

which has a minimum distance from the desired response vector, d_T. In the vector space $\varepsilon_{N,T}$ is orthogonal to the subspace that spans all the linear combinations of the N most recent input vectors. That is, $\varepsilon_{N,T}$ is orthogonal to the column space of $Z_{N,T}$ and thus satisfies,

$$z'_{T-i} \varepsilon_{N,T} = [0 \quad 0] \quad \text{for } i=0,\ldots,N-1. \qquad (3.12)$$

In matrix notation this becomes,

$$Z'_{N,T}\varepsilon_{N,T} = Z'_{N,T}(d_T + Z_{N,T}W'_{N,T}) = 0_N. \tag{3.13}$$

The solution is,

$$W_{N,T} = -d'_T Z_{N,T}(Z'_{N,T}Z_{N,T})^{-1} \tag{3.14}$$

As pointed out in [2] if the sample autocorrelation matrix is singular, a situation that always occurs when the discrete time index T is less that N-1, then any generalized inverse yields a solution to the RLS problem. We can write,

$$\varepsilon_{N,T} = [I - Z_{N,T}(Z'_{N,T}Z_{N,T})^{-1}Z'_{N,T}]d_T \tag{3.15}$$

Define the projection operators,

$$P_Z = Z(Z'Z)^{-1}Z' \tag{3.16}$$

and,

$$P^| = I - P_Z \tag{3.17}$$

Hence,

$$\varepsilon_{N,T} = P^|_{N,T}d_T \tag{3.18}$$

The matrix $P_{N,T}$ projects vectors onto the subspace spanned by the columns of Z, i.e. the past observations. The projection operator $P^|_{N,T}$ is its orthogonal complement and projects vectors onto a subspace that is orthogonal to the subspace associated with $P_{N,T}$. In other words [2], the prediction error $\varepsilon_T = P^|_{N,T}d_T = d_T - \hat{d}_T$ is orthogonal to the predicted estimate $\hat{d}_T = P_{N,T}d_T$.

An important property of the projection operator is the idempotent property:

$$P^2_{N,T} = P_{N,T} \tag{3.19}$$

We also define the quantity,

$$K_Z = Z(Z'Z)^{-1} \tag{3.20}$$

From which using (3.14) we can write the optimum weights as,

$$W_{N,T} = -d_T' K_{N,T} \tag{3.21}$$

Therefore, $U'K_Z$ can be interpreted [2] as the linear transformation which acts on the subspace Z to yield the closest approximation to U in the least-squares or minimum distance sense. In other words $U'K_Z$ is a linear filter that acts on the input vectors to estimate U. In our case $U = d_T$. Also, K_Z can be interpreted as a generalized right inverse for Z:

$$Z'K_Z = I \tag{3.22}$$

Furthermore,

$$P_Z K_Z = K_Z \tag{3.23}$$

As in [7,9] and [2] we define the pinning vector,

$$\sigma = [1 \ 0 \ 0 \ ...0]' \quad Jx1 \ . \tag{3.24}$$

The pinning vector has the important property that its inner product with the time vectors $d_T, \varepsilon_{N,T},$ and z_T yields the most recent sample:

$$\sigma' d_T = d(T) \tag{3.25}$$

$$\sigma' z_T = [x(n)y(n-1)]' \tag{3.26}$$

$$\sigma' \varepsilon_{N,T} = \varepsilon(T) \qquad\qquad (3.27)$$

4. Filters for RLS Algorithm

We introduce the following filters for the pole/zero RLS algorithm:

$$A_{N,T} = [I_2 \ -z_T'K_{N,T-1}] \qquad 2\times(N+2) \qquad (4.1)$$

$$B_{N,T} = [-z_{T-N}'K_{N,T} \ I_2] \qquad 2\times(N+2) \qquad (4.2)$$

$$C_{N,T} = -\sigma' K_{N,T} \qquad 1\times N \qquad (4.3)$$

$$W_{N,T} = -d_T'K_{N,T} \qquad 1\times N \qquad (4.4)$$

Corresponding to these filters are the residuals obtained by appropriately exciting these filters. First define

$$\bar{z}(T) = [x(T) \ x(T-1)...x(T-M+1) \ y(T-1)...y(T-L+1)]' \qquad (4.5)$$

$$\varphi_F\bar{z}(T) = [x(T) \ y(T-1) \ \underline{z}'(T-1)]' \qquad (4.6)$$

$$\varphi_B\bar{z}(T) = [\ \underline{z}'(T) \ x(T-M+1) \ y(T-L+1)]' \qquad (4.7)$$

The permutation matrices φ_B and φ_F were introduced in [3] and have the important property that

$$\varphi \ \varphi' = I$$

Now,

$$e_N(T) = A_{N,T}\varphi_F\bar{z}'(T) = \underline{z}'(T)P_{N,T-1}^|\sigma \qquad 2\times1 \qquad (4.8)$$

$$r_N(T) = B_{N,T-1}\varphi_B\bar{z}'(T) = \underline{z}'(T-N)P_{N,T}^|\sigma \qquad 2\times1 \qquad (4.9)$$

$$\gamma_N(T) = 1 + C_{N,T}\underline{z}(T) \tag{4.10}$$

$$\gamma_N(T) = \sigma' P\bigg|_{N,T}\sigma \tag{4.11}$$

$$\varepsilon_N(T) = d_T' P\bigg|_{N,T}\sigma \tag{4.12}$$

The above quantities can be interpreted as the filtered residuals. In particular (4.10) provides an interesting interpretation for $C_{N,T}$ as the filter which when excited forms the best estimate of the quantity one. The residual being $\gamma_N(T)$. In [7] the term $\gamma_N(T)$ is seen to be the cosine squared of the angle between $\underline{z}(T)$ and its projection on the space $Z_{N,T-1}$. Hence as $\gamma_N(T)$ approaches one the gap between the spaces closes. In other words it is a measure of how much new information is provided.

We define the prediction residuals as,

$$e_N^p(T) = A_{N,T-1}\varphi_B\underline{\bar{z}}'(T) \qquad 2\times1 \tag{4.13}$$

$$r_N^p(T) = B_{N,T-1}\varphi_B\underline{\bar{z}}'(T) \qquad 2\times1 \tag{4.14}$$

$$\varepsilon_N^p = d(T) + W_{N,T-1}\underline{z}'(T) \tag{4.15}$$

Define the residual powers,

$$\alpha_N(T) = A_{N,T}\varphi_F Z_{N,T}Z_{N,T}'\varphi_F A_{N,T}' \qquad 2\times2 \tag{4.16}$$

$$\beta_N(T) = B_{N,T}\varphi_B Z_{N,T}Z_{N,T}'\varphi_B B_{N,T}' \qquad 2\times2 \tag{4.17}$$

Or, using the idempotent property of the projection operators,

$$\alpha_N(T) = z_T' P_{N,T-1}^{\mid} z_T \tag{4.18}$$

$$\beta_N(T) = z_{T-N}' P_{N,T}^{\mid} z_{T-N} \tag{4.19}$$

Furthermore,

$$\gamma_N(T) = ||\sigma + Z_{N,T} c_{N,T}'||^2 = \sigma' P_{N,T}^{\mid} \sigma \tag{4.20}$$

$$\xi_N(T) = ||d_T + Z_{N,T} w_{N,T}'||^2 = d_T' P_{N,T}^{\mid} d_T \tag{4.21}$$

5. Projection Operator and Filter Updating

There are two types of projection operator updating which for the case of the pole/zero RLS have to be distinguished. One involves order updating and the other involves time updating. In order updating, the order of the projection operator is increased by augmenting a new column space as follows:

$$P_{Z,W} = [Z \ W] \varphi_B' \left\{ \varphi_B \begin{bmatrix} Z \\ W \end{bmatrix} [Z \ W]' \varphi_B' \right\}^{\#} \varphi_B \begin{bmatrix} Z' \\ W' \end{bmatrix} \varphi_B' \tag{5.1}$$

Or,

$$P_{Z,W} = [Z \ W] \left\{ \begin{bmatrix} Z \\ W \end{bmatrix} [Z \ W]' \right\}^{\#} \begin{bmatrix} Z' \\ W' \end{bmatrix} \tag{5.2}$$

In case of order updating $Z = Z_{N,T}$ and $W = z_{T-N}$. For time updating $W = \sigma$ as will be shown. The projection update can be shown to be [8] (also see appendix for a matrix derivation),

$$P_{Z,W} = P_Z + P_Z^{\mid} W (W' P_Z^{\mid} W)^{\#} W' P_Z^{\mid} \tag{5.3}$$

Define,

$$V = P_Z^\perp W \tag{5.4}$$

Then,

$$P_{Z,W} = P_Z + P_V \tag{5.5}$$

The update for the orthogonal projection becomes,

$$P_{Z,W}^\perp = (P_Z^\perp - P_Z^\perp W(W' \, P_Z^\perp W)^\# W' \, P_Z^\perp) \tag{5.6}$$

For filter updating we note that,

$$\varphi_B \begin{bmatrix} Z' \\ W' \end{bmatrix} K_{Z,W} = I_{Z,W} \tag{5.7}$$

Also,

$$P_{Z,W} K_{Z,W} = K_{Z,W} \tag{5.8}$$

Define,

$$\hat{K}_{Z,W} = K_{Z,W} \varphi_B \tag{5.9}$$

and,

$$\tilde{K}_{W,Z} = K_{W,Z} \varphi_F \tag{5.10}$$

Then substituting for $P_{Z,W}$ we obtain

$$\hat{K}_{Z,W} = P_Z \hat{K}_{Z,W} + P_Z^\perp W(W' \, P_Z^\perp W)^\# W' \, (I - P_Z) \hat{K}_{Z,W} \tag{5.11}$$

Or using (5.7),

$$\hat{K}_{Z,W} = (K_Z \; 0 \; 0) + P_Z^\perp W(W' \, P_Z^\perp W)^\# [-W' \, K_Z \; I] \tag{5.12}$$

Similarly,

$$\tilde{K}_{W,Z} = [0 \; 0 \; K_Z] + P_Z^\perp W(W' \, P_Z^\perp W)^\# [I \; -W' \, K_Z]. \tag{5.13}$$

Summary of Projection and Filter Updating

$$P_{Z,W} = \left[P_Z + P_Z^{\perp} W (W' P_Z^{\perp} W)^{\#} W' P_Z^{\perp} \right] \tag{5.1'}$$

$$U' P_{Z,W}^{\perp} V = U' P_Z^{\perp} V - (U' P_Z^{\perp} W)(W' P_Z^{\perp} W)^{\#}(W' P_Z^{\perp} V) \tag{5.2'}$$

$$U' \hat{K}_{Z,W} = [U' K_Z \ 0 \ 0] + (U' P_Z^{\perp} W)(W' P_Z^{\perp} W)^{\#}[-W' K_Z \ I] \tag{5.3'}$$

$$U' \tilde{K}_{W,Z} = [0 \ 0 \ U' K_Z] + (U' P_Z^{\perp} W)(W' P_Z^{\perp} W)^{\#}[I \quad -W' K_Z] \tag{5.4'}$$

$$\hat{K}_{Z,W} = K_{Z,W}^{\varphi} B \tag{5.5'}$$

$$\tilde{K}_{W,Z} = K_{W,Z}^{\varphi} F \tag{5.6'}$$

$$\hat{K}_{W,Z} = \tilde{K}_{W,Z}^{\varphi} F^{\varphi'} B \tag{5.7'}$$

6. Derivation of Algorithm

In the following derivation frequent use is made of the definitions of the filters and residuals of section 4. The derivation follows [2] closely.

First we derive the time updates for the forward and backward prediction filters: (use 5.3' with $W = \sigma, U = z_T, Z = Z_{N,T-1}$)

$$A_{N,T-1} = A_{N,T} - \frac{e_N(T)}{\gamma_N(T-1)}[0 \ 0 \ C_{N,T-1}] \tag{6.1}$$

(use 5.4' with $U = z_{T-N}, W = \sigma, Z = Z_{N,T}$)

$$B_{N,T-1} = B_{N,T} - \frac{r_N(T)}{\gamma_N(T)}[C_{N,T} \ 0 \ 0] \tag{6.2}$$

Order update $C_{N,T-1}$ by using (5.4') with $U = \sigma, W = z_T, Z = Z_{N,T-1}$.

$$C_{N+1,T} = [0 \ 0 \ C_{N,T-1}] - e_N(T)\alpha_N^{-1}(T)A_{N,T} \tag{6.3}$$

Also using (5.3') and $U=\sigma, W=z_{T-N}, Z=Z_{N,T}$,

$$\hat{C}_{N+1,T} = [C_{N,T} \ 0 \ 0] + r_N'(T)\beta_N^{-1}(T)B_{N,T} \tag{6.4}$$

Close examination of (6.4) reveals that the last two elements of $\hat{C}_{N+1,T}$ are equal to,

$$\mu(T) = r_N'(T)\beta_N^{-1}(T)$$

So we can write

$$\hat{C}_{N+1,T} = [C_{N,T} \ 0 \ 0] + \mu(T)B_{N,T} \tag{6.4'}$$

From (5.2') using $U=V=z_T, W=\sigma, Z=Z_{N,T-1}$,

$$\alpha_N(T) = \alpha_N(T-1) + e_N(T)\gamma_N^{-1}(T-1)e_N'(T) \tag{6.5}$$

Update $\beta_N(T)$ by using (5.2') with $U=V=z_{T-N}, W=\sigma, Z=Z_{N,T}$,

$$\beta_N(T) = \beta_N(T-1) + r_N(T)\gamma_N^{-1}(T)r_N'(T) \tag{6.6}$$

To obtain the prediction errors post multiply (6.1) by $\varphi_F \underline{z}'(T)$ and (6.2) by $\varphi_B \underline{z}'(T)$.

$$e_N^p(T) = e_N(T) - e_N(T)\gamma_N^{-1}(T-1)(\gamma_N(T-1)-1) = \frac{e_N(T)}{\gamma_N(T-1)} \tag{6.7}$$

$$r_N^p(T) = \frac{r_N(T)}{\gamma_N(T)} \tag{6.8}$$

To obtain the order update for $\gamma_{N+1}(T)$ use (5.2') with $W=z_{T-N}$, $Z=Z_{N,T-1}$, $U=V=\sigma$,

$$\gamma_{N+1}(T) = \gamma_N(T) - r_N'(T)\beta_N^{-1}(T)r_N(T) \tag{6.9}$$

Again in (5.2') let $W=z_T$, $Z=Z_{N,T-1}$, $U=V=\sigma$,

$$\gamma_{N+1}(T) = \gamma_N(T-1) - e_N'(T)\alpha_N^{-1}(T)e_N(T) \tag{6.10}$$

Define,

$$\tilde{C}_{N,T} = \gamma_N^{-1}(T)C_{N,T} \tag{6.11}$$

$$\hat{C}_{N+1,T} = \gamma_{N+1}^{-1}(T)\varphi_B C_{N+1,T} \tag{6.12}$$

$$\tilde{C}_{N+1,T} = \gamma_{N+1}^{-1}(T)\varphi_F C_{N+1,T} \tag{6.13}$$

Then from (6.3) substituting (6.1) and simplifying using (6.11) we obtain,

$$C_{N+1,T} = [0 \ 0 \ C_{N,T-1}] - e_N(T)\alpha_N^{-1}(T)A_{N,T-1} - $$

$$e_N(T)\alpha_N^{-1}(T)e_N(T)[0 \ 0 \ C_{N,T-1}]$$

$$C_{N+1,T} = [0 \ 0 \ C_{N,T-1}] - $$

$$e_N(T)\alpha_N^{-1}(T)A_{N,T-1}[\gamma_{N+1}(T)-\gamma_N(T-1)]\gamma_N^{-1}(T-1)[\ 0 \ 0 \ C_{N,T-1}].$$

$$C_{N+1,T} = \gamma_{N+1}(T)\gamma_N^{-1}(T-1)[0 \ 0 \ C_{N,T-1}] - e_N(T)\alpha_N^{-1}(T)\gamma_{N+1}^{-1}(T)A_{N,T-1}$$

Using the definition (6.11) we obtain,

$$\hat{C}_{N+1,T} = [0 \ 0 \ \tilde{C}_{N,T-1}] - e_N(T)\alpha_N^{-1}(T)\gamma_{N+1}^{-1}(T)A_{N,T-1} \tag{6.14}$$

Now from (6.11) through (6.13),

$$\tilde{C}_{N+1,T} = \hat{C}_{N+1,T}\varphi_F\varphi_B \tag{6.15}$$

Substituting (6.2) into (6.4) we get,

$$\hat{C}_{N+1,T} = [C_{N,T} \; 0 \; 0 \;] + \mu(T)B_{N,T-1} + \mu(T)\frac{r_N(T)}{\gamma_N(T)}[C_{N,T} \; 0 \; 0] \tag{6.16}$$

But $\mu(T) = r'_N(T)\beta_N^{-1}(T)$. Hence, using (6.9) we get,

$$C_{N+1,T} = [C_{N,T} \; 0 \; 0] + \mu(T)B_{N,T-1} - [\gamma_{N+1}(T) - \gamma_N(T)]\frac{1}{\gamma_N(T)}[C_{N,T} \; 0 \; 0]$$

$$\tilde{C}_{N+1,T} = [\tilde{C}_{N,T} \; 0 \; 0] + \mu(T)B_{N,T-1} \tag{6.17}$$

Simultaneous solution of (6.4) and (6.2) yields

$$[C_{N,T} \; 0 \; 0] = [1+\mu(T)r_N^p(T)]^{-1}[C_{N+1,T} - \mu(T)B_{N,T-1}] \tag{6.18}$$

Excite by $\varphi_B\bar{z}_T$,

$$\gamma_{N,T}^{-1}(T) = [1+\mu(T)r_N^p(T)]^{-1}[\gamma_{N+1,T}^{-1}(T) - r'_N(T)\beta_N^{-1}(T)r_N^p(T)]$$

Substituting from (6.9) and simplifying,

$$\gamma_N(T) = [1+\mu(T)r_N^p(T)]^{-1}\gamma_{N+1}(T) \tag{6.19}$$

Once $\tilde{C}_{N,T}$ has been updated then the joint process updates can be obtained:

$$\varepsilon_N^p(T) = d(T) + W_{N,T-1}\underline{z}(T) \tag{6.20}$$

$$\varepsilon_N(T) = \varepsilon_N^p(T)\gamma_N(T) \tag{6.21}$$

$$W_{N,T} = W_{N,T-1} + \varepsilon_N(T)\tilde{C}_{N,T} \tag{6.22}$$

Summary Fast Pole/Zero Adaptive Filter

$$e_N^p(T) = A_{N,T-1} \; \varphi_F \underline{\bar{z}}(T)$$

$$e_N(T) = e_N^p(T)\gamma_N(T-1)$$

$$\alpha_N(T) = \alpha_N(T-1) + e_N^p(T)e_N'(T)$$

$$\gamma_{N+1}(T) = \gamma_N(T-1) - e_N'(T)\alpha_N^{-1}(T)e_N(T)$$

$$\hat{C}_{N+1,T} = [0 \; 0 \; \tilde{C}_{N,T-1}] - e_N'(T)\alpha_N^{-1}(T) \; \gamma_{N+1}^{-1}(T)A_{N,T-1}$$

$$\tilde{C}_{N+1,T} = \hat{C}_{N+1,T} \; \varphi_F \varphi_B$$

$$A_{N,T} = A_{N,T-1} + e_N(T)[0 \; 0 \; \tilde{C}_{N,T-1}]$$

$$r_N^{p'}(T) = -\tilde{\mu}(T)\beta_N(T-1)$$

Note that $\tilde{\mu}(T) = \tilde{C}_{N+1}^N$ is the last two elements of $\tilde{C}_{N+1,T}$.

$$\gamma_N(T) = [1+\tilde{\mu}(T)r_N^p(T)\gamma_{N+1}(T)]^{-1}\gamma_{N+1}(T)$$

$$r_N(T) = r_N^p(T)\gamma_N(T)$$

$$\beta_N = \beta_N(T-1) + r_N^p(T)r_N'(T)$$

$$[\tilde{C}_{N,T} \; 0 \; 0] = \tilde{C}_{N,T} - \tilde{\mu}(T)\tilde{B}_{N,T-1}$$

$$B_{N,T} = B_{N,T-1} + r_N(T)[\tilde{C}_{N,T} \; 0 \; 0]$$

Joint-Process Extension.

$$\varepsilon_N^p(T) = d(T) + W_{N,T-1}\underline{z}(T)$$

$$\varepsilon_N(T) = \varepsilon_N^p(T)\gamma_N(T)$$

$$W_{N,T} = W_{N,T-1} + \varepsilon_N(T)\tilde{C}_{N,T}$$

7. Simulation Results

In this section the results of simulating the Fast
Pole-Zero Recursive Least Squares algorithm, derived and
summarized in the previous section, are presented. The
algorithm was programmed in C. To test the algorithm the
poles and zeroes of a 10th order elliptic bandpass digital
filter were estimated. The pole-zero distribution of the
filter is shown in figure 1. In figure 2 the estimated
poles and zeroes using the algorithm when a white Gaussian
noise sequence is used to excite the filter is shown. Note
the close estimation of the pole and zero locations. Figure
3 shows the prediction error plotted in dB's against the
number of iterations. Note the rapid convergence of the
algorithm. Also the prediction error is more than 100 dB's
below the desired signal, an indication of the high degree
of accuracy achievable by pole-zero modeling in comparison
to impulse response modeling. In an echo cancellation
application this corresponds to a 100 dB echo return loss
enhancement. Where appropriate and as a measure of the
algorithm's performance the ratio of the power of the
desired signal over the power of the prediction error is
given in dB's. Ideally, in the absence of noise, this ratio
is very large. In all cases the first N iterations are dis-
carded.

The effect of zero overspecification is presented in
figures 4 and 5. In this case the number of zeroes was

overspecified by 10 (M=20,L=10). The zeroes are distributed around the origin and two of them cancel out two of the poles. The rapid convergence and high degree of prediction accuracy is again observed in figure 4 with a slight degradation in comparison with no zero overspecification.

The interesting case of both pole and zero overspecification is presented in figure 6 (M=20, L=20). Notice that the extra poles and zeroes cancel each other out.

An important property of the algorithm is its ability to predict the desired signal in the presence of noise. In figure 7, the prediction error is plotted against the number of iterations for the case where white Gaussian noise with an rms value of 0.1 was added to the desired signal. Note that the prediction error is uncorrelated in contrast to the filtered and correlated desired signal. In other words the desired signal has been eliminated from the prediction error. In an echo cancellation application this corresponds to the elimination of the echo even in the presence of noise. The pole-zero distribution for this case is shown in figure 8. Note that noise does effect the location of the poles and zeroes. In fact, since the noise contaminates $y(n)$ which is used in the prediction of the desired signal a bias is introduced in the estimation of the weight vector $W_{N,T}$.
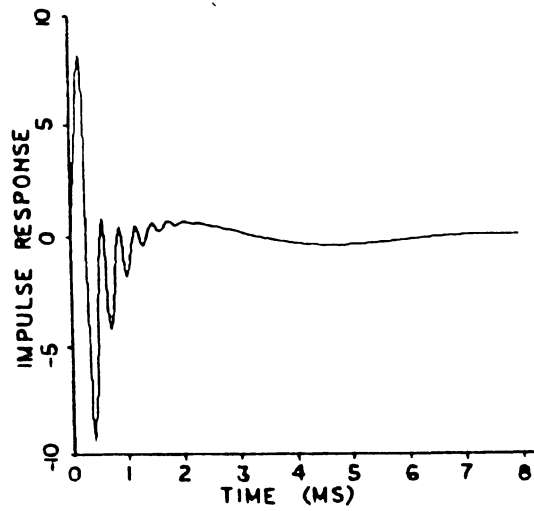
Digital IIR Filter Impulse Response

Sampling Rate 32 kHz
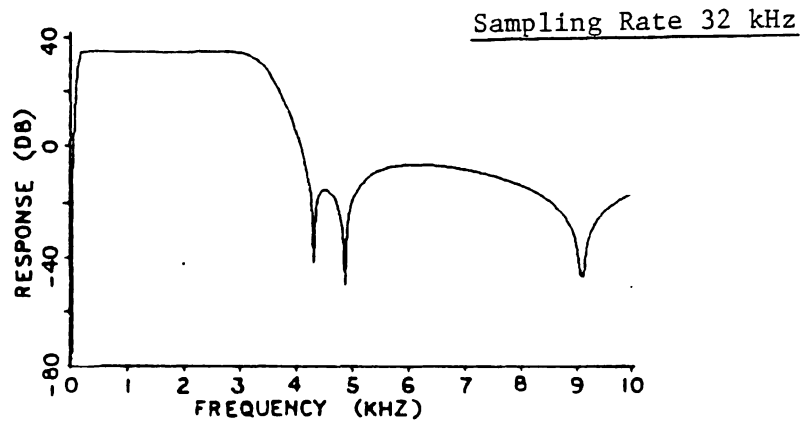
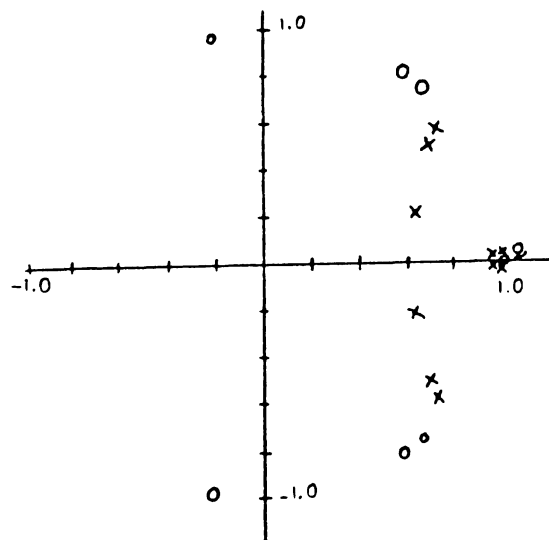Digital IIR Filter Frequency Response

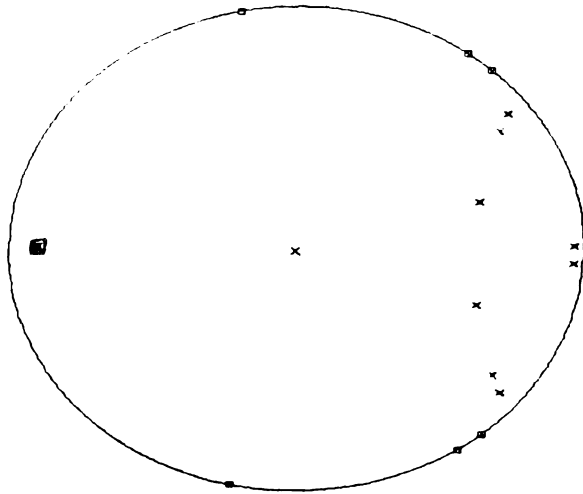Fig. 1 Pole-Zero Distribution of IIR Digital Filter

Fig. 2 Estimated Pole-Zero Distribution (M=10, L=10)
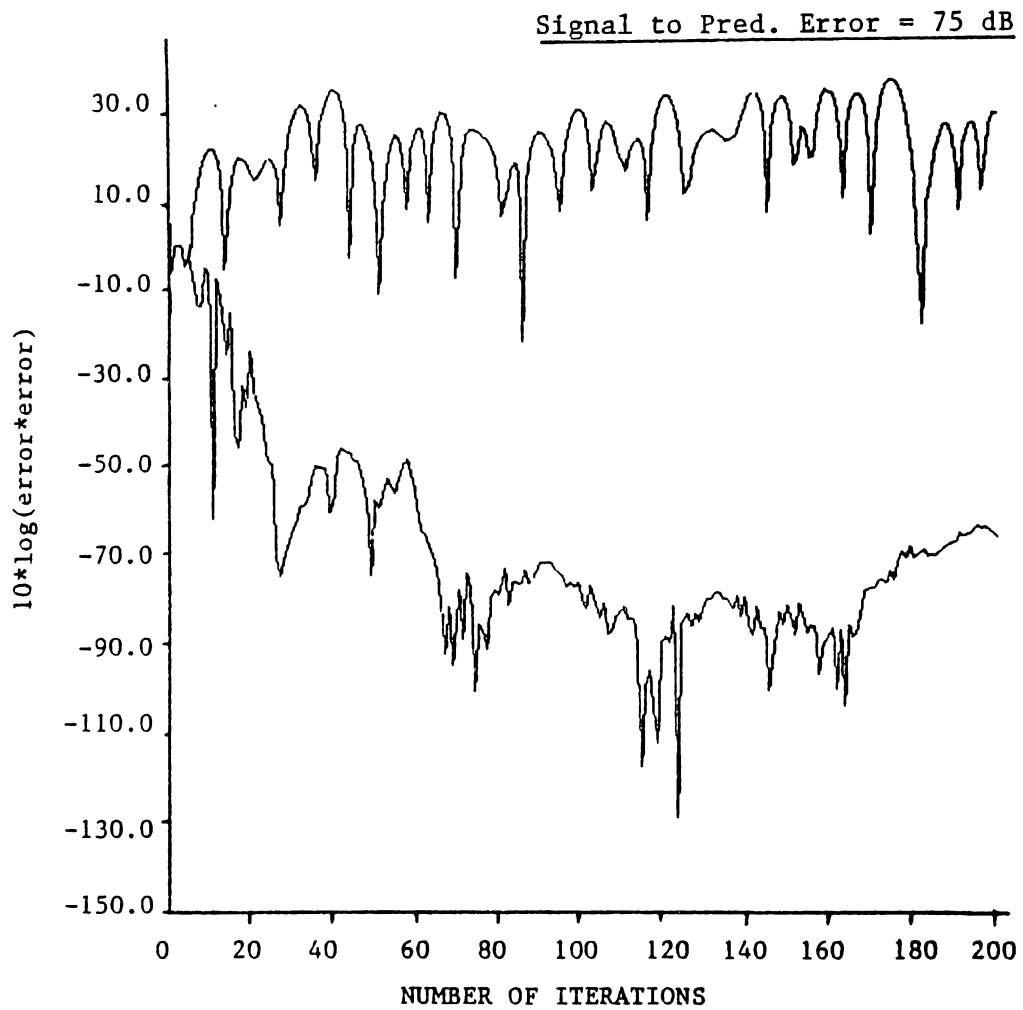

Fig. 3 Convergence of Fast Pole-Zero Algorithm (M=10, L=10)
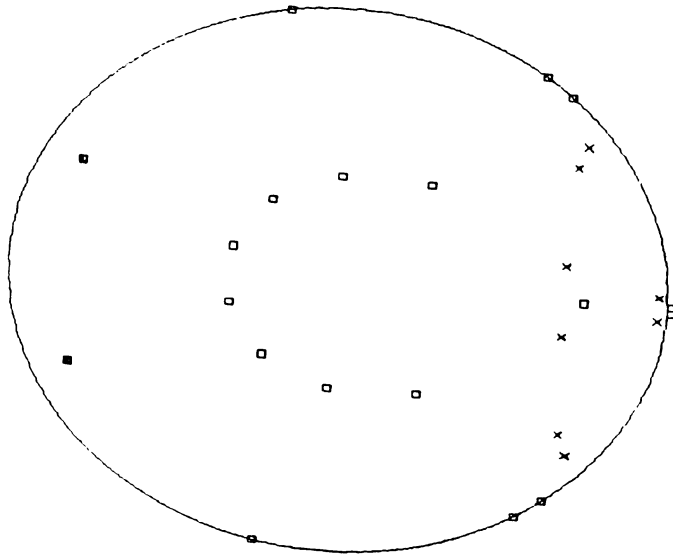
Fig. 4 Pole-Zero Distribution Zero Overspecification
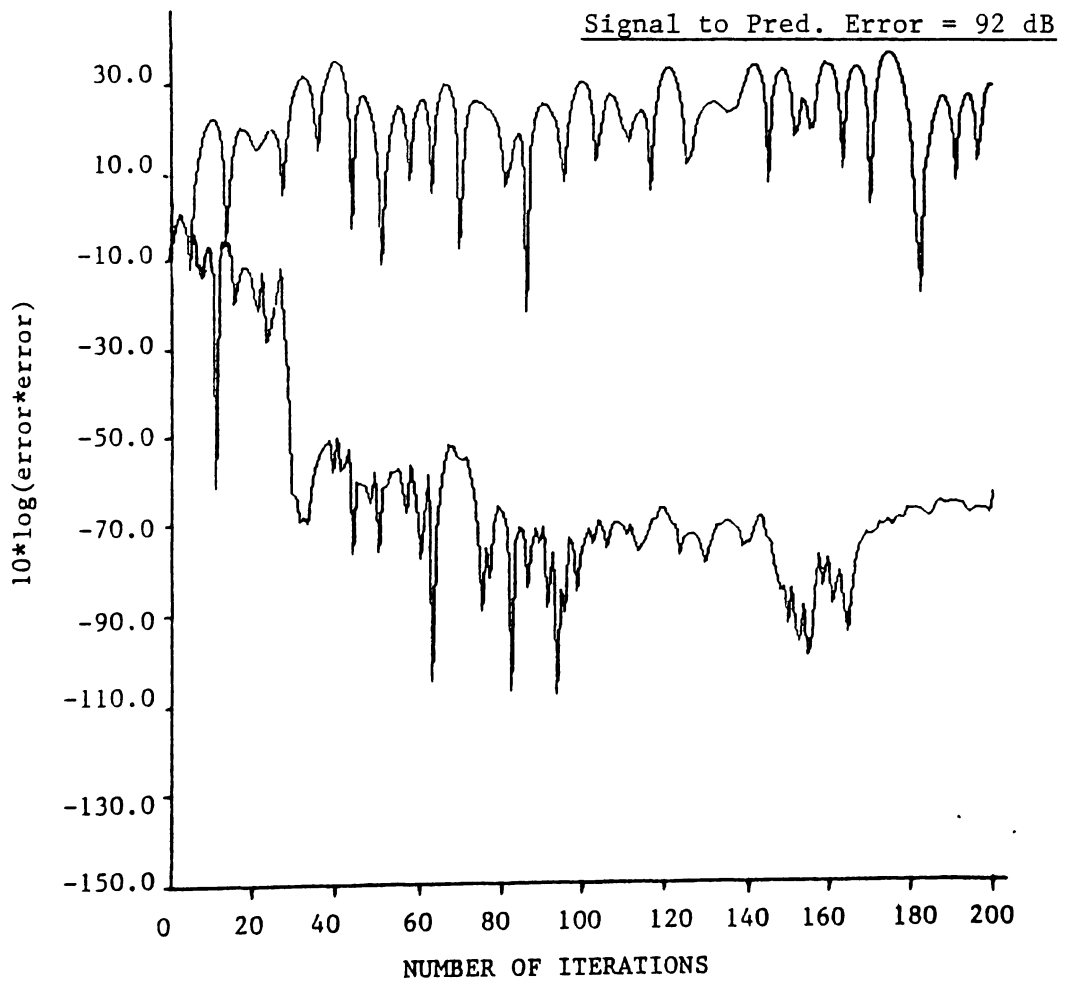(M=20, L=10)

Fig. 5 Convergence With Zero Over Specification
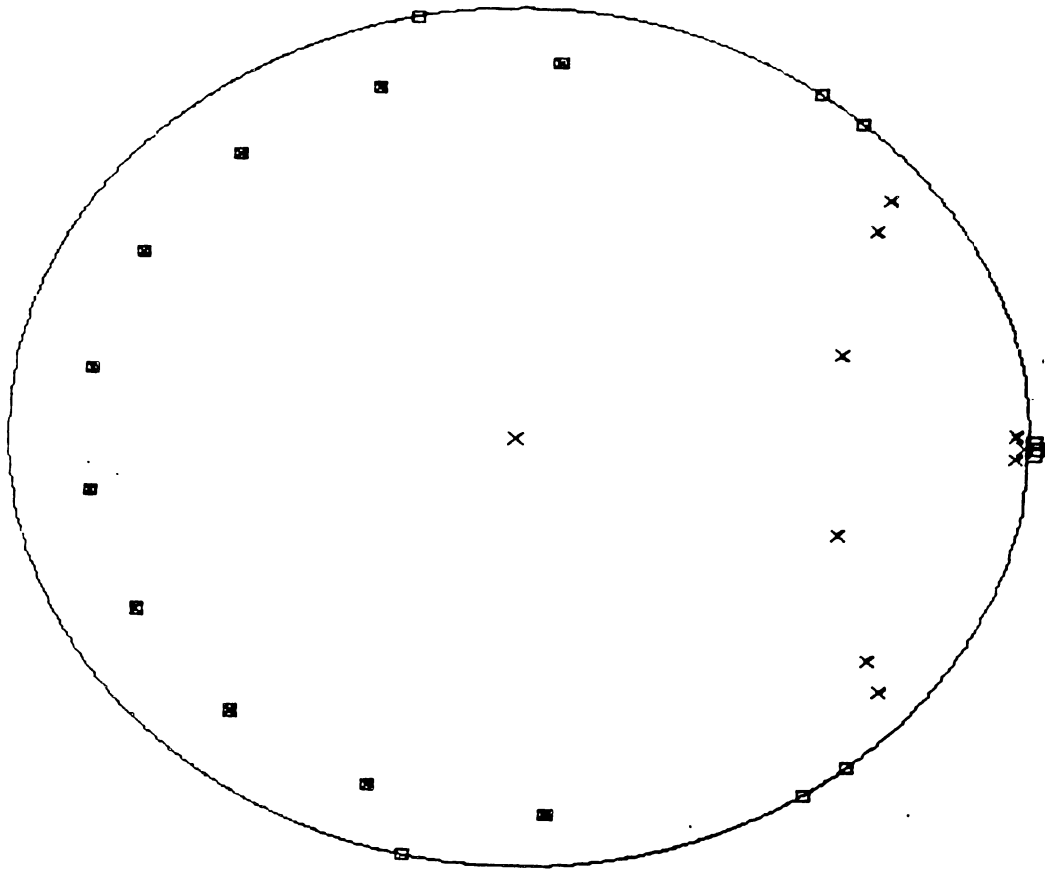(M=20, L= 10)

Fig. 6 Pole-Zero Distribution With Pole and Zero
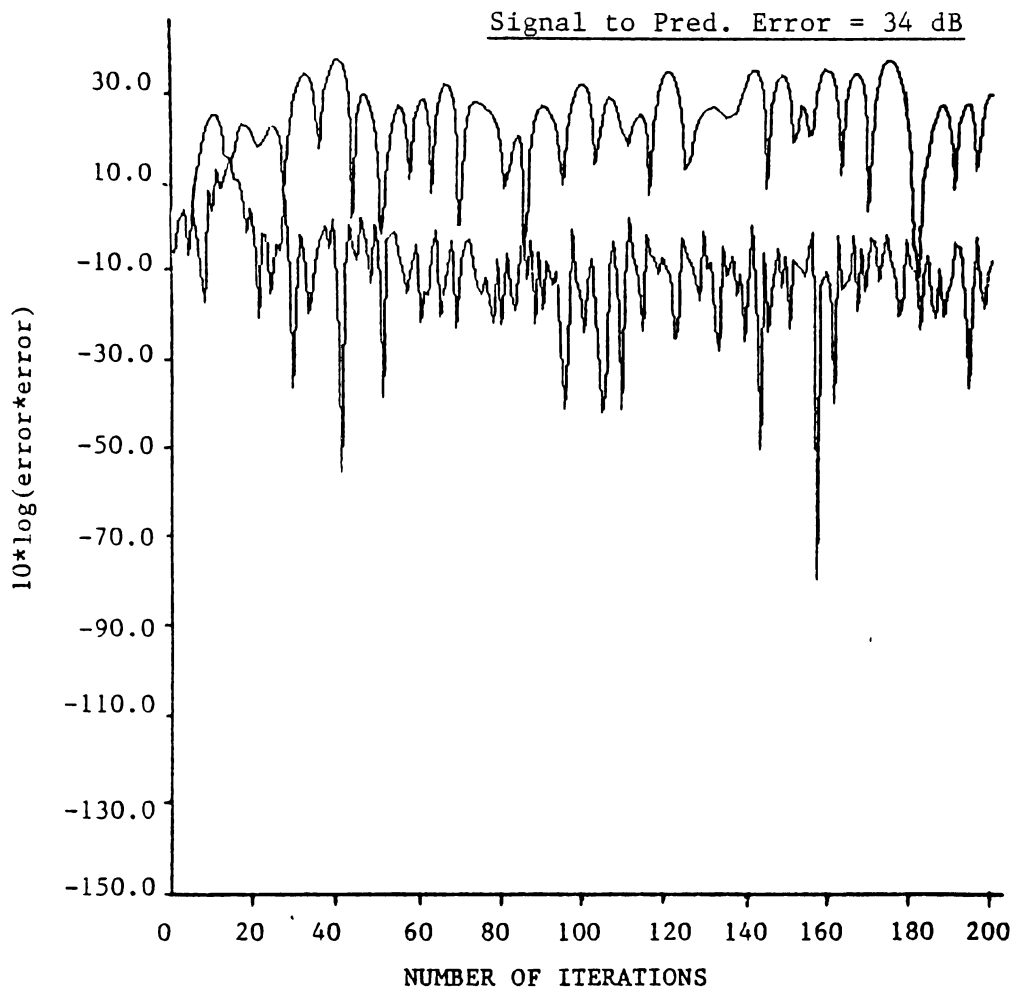Overspecification
        (M=20, L=20)

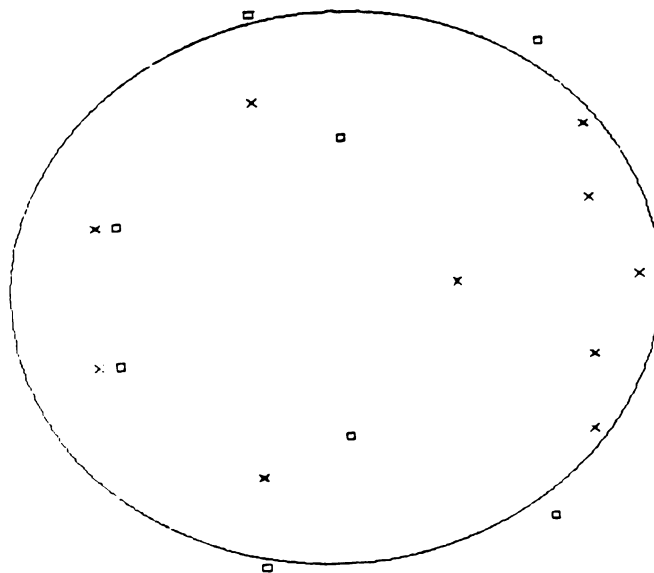Fig. 7 Prediction Error in the Presence of Noise
(rms value 0.1)

Fig. 8 Pole-Zero Distribution With Noise

8. Appendix

In this appendix the projection update is derived directly from matrix methods. The projection obtained by augmenting the column space W to Z is,

$$P_{Z,W} = (Z\ W) \left[\begin{pmatrix} Z' \\ W' \end{pmatrix} \begin{pmatrix} Z & W \end{pmatrix} \right]^{-1} \begin{pmatrix} Z' \\ W' \end{pmatrix}$$

We must find an inverse such that,

$$\begin{bmatrix} ZZ' & Z'W \\ W'Z & W'W \end{bmatrix} \begin{bmatrix} A_1 & A_2 \\ B_1 & B_2 \end{bmatrix} = I$$

Or,

$$\begin{bmatrix} ZZ'A_1 + Z'WB_1 & Z'ZA_2 + Z'WB_2 \\ W'ZA_1 + W'WB_1 & W'ZA_2 + W'WB_2 \end{bmatrix} = \begin{bmatrix} I_{NN} & 0_p \\ 0_N & I_{pp} \end{bmatrix}$$

This leads to four equations:

$$Z'ZA_1 + Z'WB_1 = I_{NN} \tag{1}$$

$$W'ZA_1 + W'WB_1 = 0_p \tag{2}$$

$$Z'ZA_2 + Z'WB_2 = 0_N \tag{3}$$

$$W'ZA_2 + W'WB_2 = I_{pp} \tag{4}$$

From (3),

$$A_2 = -(Z'Z)^{-1}Z'WB_2$$

Substituting into (4) and rearranging,

$$B_2 = [W'W - W'Z(Z'Z)^{-1}Z'W]^{-1}$$

Hence,

$$A_2 = -(Z'Z)^{-1}Z'W[WW' - W'Z(Z'Z)^{-1}Z'W]^{-1}$$

Similarly we obtain,

$$B_1 = [W'W - W'Z(Z'Z)^{-1}Z'W]^{-1}(W'Z)(Z'Z)^{-1}$$

$$A_1 = -(Z'Z)^{-1}Z'W[WW' - W'Z(Z'Z)^{-1}Z'W]^{-1}(W'Z)(Z'Z)^{-1}+(Z'Z)^{-1}$$

Substituting the inverse into the projection update and after some algebra we obtain,

$$P_{Z,W} = P_Z + P_Z^{\perp}W(W'P_Z^{\perp}W)^{-1}W'P_Z^{\perp} \tag{5}$$

Note that,

$$P_Z^{\perp} = I - P_Z$$

The above relationships could have been derived using the generalized inverses. It is interesting to note that if we define,

$$V = P_Z^{\perp}W \tag{6}$$

Then the projection update (5) can be written,

$$P_{Z,W} = P_Z + P_V \tag{7}$$

In fact an alternative derivation of (5) is suggested by (7). The vector V represents that part of W that cannot be estimated by the space spanned by the columns of Z. It is orthogonal to the sub space Z. Hence the projection operator that includes the subspace Z and W can be expressed as a direct sum of the projection operator of the subspace Z (P_Z) and the projection operator P_V of the subspace spanned by V which is orthogonal to Z. Thus,

$$P_{Z,W} = P_Z + P_V$$

REFERENCES

[1]   D.D. Falconer and L. Ljung, "Application of Fast Kal-
       man Estimation to Adaptive Equalization," IEEE Tran-
       sactions on Communications, vol. COM-26, pp. pp.
       1439-1446., October 1978.

[2]   John M. Cioffi and Thomas Kailath, "Fast, Recursive-
       Least-Squares Transversal Filters for Adaptive
       Filtering," IEEE Transactions on Acoustics,Speech,
       and Signal Processing, vol. ASSP-32, no. 2, pp. pp.
       304-337, April 1984.

[3]   L. Ljung, M. Morf, and D. D. Falconer, "Fast Calcula-
       tion of Gain Matrices for Recursive Estimation
       Schemes," Int. Journal of Control, vol. 27, no. 1,
       pp. pp. 1-19, 1978.

[4]   Sasan Ardalan and S. T. Alexander, "Rapid Echo Can-
       cellation Using Fast Kalman Pole-Zero Modeling," IEEE
       Transactions on Communications, Submitted.

[5]   Sasan Ardalan, "Pole/Zero Echo Path Modeling and
       Recursive Least Squares Echo Cancellation," IEEE
       Workshop On Digital Signal Processing, October 1984.

[6]   Sasan Ardalan, "Pole/Zero Fast Kalman Echo Cancella-
       tion and Application to Actual Measured Echo," Inter-
       national Conference on Communications, ICC-85, June
       1985.

[7]   Daniel T. L. Lee, Martin Morf, and Benjamin
       Friedlander, "Recursive Least Squares Ladder Estima-
       tion Algorithms," IEEE Transactions on Acoustics,
       Speech, and Signal Processing, vol. ASSP-29, no. 3,
       June 1981.

[8]   Boaz Porat, Benjamin Friedlander, and Martin
       Morf, "Square Root Covariance Ladder Algorithms,"
       IEEE Transactions on Automatic Control, vol. AC-27,
       no. 4, pp. pp. 813-825, August 1982.