LINEAR AND NON-LINEAR DIGITAL PROCESSING

G. C. RANIERI and J. J. UHRAN, Jr.
Department of Electrical Engineering
University of Notre Dame
Notre Dame, Indiana.

Summary Three basic discrete estimators, each of increasing complexity, are presented for extracting a random signal sequence from one which has been distorted by an additive white noise sequence. Closed form solutions are obtained by minimizing the mean square error, convergence properties are indicated and comparisons made. Several examples are given to indicate what typically can be expected in each case.

Introduction Statistical communication and control theory have been researched a great deal over the last three decades; a large portion of the work being devoted to an expansion and application of the Wiener-Kolmogoroff [1] and the Kalman-Bucy [2] theories to the problems of estimation. While initial efforts were mainly directed towards continuous estimation problems [3-4], recently emphasis has shifted to systems which are basically discrete in nature [5-6]. The work which follows is concerned with solving the optimal estimator problem using both linear and non-linear solution techniques when the inputs are random sequences. Though this problem frequently encompasses prediction, filtering and smoothing, our concern will be mainly that of filtering. The theory can immediately be applied to the other two without any difficulty.

General Problem Statement The basic problem of interest here is modelled in Figure 1. A random signal sequence, $\{s_k\}$, is generated and distorted by a random, additive noise sequence, $\{n_k\}$, It is desired to operate on this new k sequence, $\{x_k\}$, in such a way as to extract an estimate, $\{\hat{s}_k\}$, of the signal sequence which is close to the original signal sequence in some statistical sense. The measure of goodness used is the mean-square difference, Eq. (1), between the signal and its estimate.

$$e_{ms} = E \left[(s_k - \hat{s}_k)^2 \right] \tag{1}$$

$E\{\cdot\}$ is defined to be the expectation operator. The problem is to minimize this measure over the class of estimators being used.

This work will propose a simple linear and non-linear estimating procedure and compare it to the well known optimum, discrete Wiener-Hopf linear estimator. Several examples are worked to give some insight into the development throughout the discussion. It will be assumed that the signal and noise sequences are statistically independent, when necessary for computations.

Discrete Wiener-Hopf Estimator For the optimum, discrete, linear estimator (Wiener-Hopf Estimator), the output, $\hat{s}_k$, Eq. (1), can be computed as a discrete convolution of x_k with a_k , that is,

$$\hat{s}_k = \sum_{j=-\infty}^{\infty} a_j \, x_{k-j} \tag{2}$$

where a_k is a sequence in ℓ_1^+ to be determined.

By substituting Eq. (2) into Eq. (1) and expanding the result, the mean-square error is found as in Eq. (3).

$$e_{ms} = \phi_s(o) - 2\sum_{j=-\infty}^{\infty} a_j \, \phi_{sx}(j) + \sum_{m=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} a_m a_j \phi_x(j-m) \tag{3}$$

where $\phi_s(\cdot)$ is the autocorrelation sequence for the actual signal, $\phi_{sx}(\cdot)$ is the cross correlation sequence for the actual sequence for the actual signal, and $\phi_x(\cdot)$ is the autocorrelation sequence for the distorted signal.

By any one of several methods, the mean-square error in Eq. (3) can be minimized over all possible sequences $\{a_k)$to obtain the equations

$$\sum_{j=-\infty}^{\infty} a_j \; {}_x(j-m) = \phi_{sx}(m) \tag{4}$$

which is the well known Wiener-Hopf Equation [5].

If the z-transform of $\phi_{sx}(\cdot)$ and $\phi_x(\cdot)$ are restricted to be rational functions of z, then the z-transform of the optimum realizable discrete WienerHopf estimator can be represented as

$$A(z) = \frac{1}{\Phi_x^+(z)} \left\{ \frac{\Phi_{sx}(z)}{\Phi_x^-(z)} \right\}_+ \tag{5}$$

where

$$\Phi_x(z) = \Phi_x^-(z) \; \Phi^+(z) \tag{6}$$

and the brackets {·} denote taking the realizable part of the function which the contain. The inverse transform of Eq. (5) yields the desired sequence, a_k}, of the estimator.

Example Let us assume that the signal has an autocorrelation sequence

$$\phi_s(k) = \nu \exp \{-\alpha |k|\} \tag{7}$$

and the additive noise has an autocorrelation sequence

$$\phi_n(k) = N_o \delta_{o,k} \tag{8}$$

where $\delta_{0,k}$ is a Kronecker delta. The optimum, realizable filter thus has the weighting sequence (the realizable solution to Eq. (4))

$$a_k = \begin{cases} \dfrac{2\nu \operatorname{Sinh} \alpha}{N_o(z_1^* - e^{-\alpha})} \cdot (z_1)^k , & k \geq o \\ o , & k < o \end{cases} \tag{9}$$

where

$$z_1 = \beta - \sqrt{\beta^2 - 1} = 1/z_1^* \tag{10}$$

and

$$\beta = \operatorname{Cosh} \alpha + \frac{\nu}{N_o} \operatorname{Sinh} \alpha \tag{11}$$

The effect that the signal to noise ratio,

$$\mathrm{SNR} = \frac{\text{Signal Power}}{\text{Noise Density}} = \frac{\nu}{N_o} \tag{12}$$

has on the mean-square error of this estimator was computed for a wide range of α, the signal correlation coefficient, and SNR, and is shown in Figure 2. From these curves, it is evident that as the correlation between elements in the signal sequence increases - that is, as $\propto$ decreases - the overall meansquare error decreases. It should also be noted that as a signal to noise ratio increases - that is, as N_o decreases - the mean-square error decreases to 0.

Discrete Linear Nonrecursive Estimator While the Wiener-Hopf Estimator of the previous section has potentially infinite memory (recursive), Figure 3 shows a model of a linear nonrecursive, finite memory estimator (transversal equalizer). In this case the estimated signal is a finite convolution of the input and the delay line weighting sequence, Eq.(13).

$$\hat{s}_k = \sum_{j=o}^{n} A_{(j)} x_{k-j} \tag{13}$$

By substituting Eq. (13) into Eq. (1), a finite version of Eq. (3) is obtained, that is

$$e_{ms} = \phi_s(o) - 2\sum_{j=o}^{n} A_{(j)} \phi_{sx}(j) + \sum_{m=o}^{n} A_{(m)} \sum_{j=o}^{n} A_{(j)}\phi_x(j-m) \tag{14}$$

The mean-square error will be a minimum if [7]

$$\sum_{m=o}^{n} A_{(m)} \phi_x(m-j) = \phi_{sx}(j), \quad j = o,1,2, \ldots n \tag{15}$$

or in vector notation

$$\Phi_x \underline{A} = \underline{\phi}_{sx}$$

where

$$\underline{A} = \begin{bmatrix} A_{(o)} \\ A_{(1)} \\ \vdots \\ A_{(n)} \end{bmatrix}, \quad \underline{\phi}_{sx} = \begin{bmatrix} \phi_{sx}(o) \\ \phi_{sx}(1) \\ \vdots \\ \phi_{sx}(n) \end{bmatrix} \tag{16}$$

and

$$\Phi_x = \begin{bmatrix} \phi_x(o) & \phi_x(1) \cdots & \phi_x(n) \\ \phi_x(1) & \phi_x(o) \cdots & \phi_x(n-1) \\ \cdot & \cdot & \cdot \\ \phi_x(n) & \phi_x(n-1) \cdots & \phi_x(o) \end{bmatrix} \tag{17}$$

Note that $\phi_x(-j) = \phi_x(j)$ in Eq. (7).

It is sufficient that $\phi_x(\cdot)$ be a real positive definite sequence to insure the invertibility of Φ_x. Thus, one finds that

$$\underline{A} = \Phi_x^{-1} \underline{\phi}_{sx} \tag{18}$$

which is the optimum solution to the discrete, linear, nonrecursive estimator equation.

It can be shown [7] that the mean-square error for this estimator decreases monotonically as the length of the delay line increases and will converge to the realizable Wiener-Hopf mean-square error as the delay line length goes to infinity.

Example For the previous example, described in Eqs. (7) and (8), the optimum weighting sequence, Eq.(18), was found by the Gauss-Seidel method of solving simultaneous equations and the mean-square error of Eq.(14) computed. Figure 4 displays the mean-square error versus SNR for a length 3 and length 15 filter. The comments concerning this estimator are borne out by the example.

Discrete Nonrecursive Polynomial-Type Estimator The discrete, nonrecursive, polynomial-type estimator is a simple, zero memory nonlinearity. The purpose of this design, shown in Figure 5, is to try to improve performance by using information contained in the higher order moments of the input sequence rather than only the correlation of the sequence elements as done in the previous cases. In this situation, the estimated sequence, $\hat{s}_k$, is found as

$$\hat{s}_k = \sum_{j=1}^{p} A_{(j)} x_k^j \qquad (19)$$

where each $A_{(j)}$ is a constant to be determined. Substitution of Eq.(19) into Eq.(1) leads to the mean-square error

$$e_{ms} = \phi_s(0) - 2\sum_{j=1}^{p} A_{(j)} E\{s_k x_k^j\} + \sum_{m=1}^{p} A_{(m)} \sum_{j=1}^{p} A_{(j)} E\{x_k^{m+j}\} \qquad (20)$$

It is evident that both the marginal and joint moments of the signal and noise sequences are necessary to solve Eq. (20).

By a straight forward application of the calculus of variations, the mean-square error can be minized to give the optimum weighting sequence as

$$\underline{A} = \Lambda_x^{-1} \underline{\lambda}_{sx} \qquad (21)$$

where

$$\underline{A} = \begin{bmatrix} A_{(1)} \\ A_{(2)} \\ \cdot \\ \cdot \\ A_{(p)} \end{bmatrix} , \qquad \underline{\lambda}_{sx} = E\begin{bmatrix} s_k x_k \\ s_k x_k^2 \\ \cdot \\ \cdot \\ s_k x_k^p \end{bmatrix} \qquad (22)$$

and

$$\Lambda_x = E\begin{bmatrix} x_k^2 & x_k^3 & \cdot & \cdot & \cdot & x_k^{p+1} \\ x_k^3 & x_k^4 & \cdot & \cdot & \cdot & x_k^{p+2} \\ \vdots & \vdots & & & & \vdots \\ x_k^{p+1} & x_k^{p+2} & \cdot & \cdot & \cdot & x_k^{2p} \end{bmatrix} \tag{23}$$

The matrix in Eq.(23) is always invertible [7], which always guarantees a solution to Eq. (21).

This particular model allows us in principal to relate the amount of information contained in the higher order moments as compared to the information in the second moments which are used in the linear estimators of the previous cases. There is no advantage to be obtained if the signals are gaussian, however one might hope to achieve some additional improvement when the distributions are non-gaussian in nature. This in fact is the case in general [7], though the relative improvements are not yet known.

Example Let the signal sequence be defined by the probability density function of Eq.(24),

$$f(s) = e^{-(s+1)}u_{-1}(s+1) \tag{24}$$

and let the noise sequence be defined by the probability density function of Eq.(25),

$$f(n) = \rho e^{-(\rho n+1)}u_{-1}(\rho n+1) \tag{25}$$

where u_{-1} is the unit step function. The unusual shape of these density functions was chosen so that the mean value of each of the sequences is zero. Assuming that both the signal and noise sequences are stationary and statistically independent, the mean-square error of Eq.(20) was computed for several values of p. The mean-square error is tabulated in Figure 6 for two different values of ρ for polynomials of order up to 5. For these cases, the mean-square error is a monotone non-increasing function, indicating that the higher order moments did indeed contain some additional information.

Conclusions In the case of the linear estimators, the Wiener-Hopf estimator is optimum. However, a small number of delays might only be necessary if the correlation of the signal sequence is rather low. As the correlation between adjacent values in the sequence increases, the optimum filter gives a better performance than the nonrecursive filter for a fixed number of delays. As the signal to noise ratio improves, the mean-square error performance approaches zero in either case.

For the polynomial-type estimator(non-linear), it has been demonstrated that an improvement in the mean-square error performance for a specific signal-to-noise ratio can be obtained by using only up to a third order polynomial estimate. Although the improvement is rather mild in this case, an efficient means for designing non-linear estimators is believed to have been accomplished. At the least, the computation is straight-forward and convenient in most cases. The combination of non-linear estimators and linear estimators should provide a significant improvement over either used alone. Applications in terms of adaptive estimators in cases where the statistics required are not known a-priori.

References

1. N.Wiener, "Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications",Wiley, New York, N.Y.;1948.

2. R.E.Kalman and R.S. Bucy,"New Results in Linear Filtering and and Prediction Theory," Trans. of ASME, Series D, Journal of Basic Engineering, pp. 95-108;March, 1961.

3. H.L. Van Trees,"Detection, Estimation and Modulation Theory, Part I," Wiley, New York, N.Y;1968.

4. A.H.Haddad and J.B.Thomas,"on Optimal and Suboptimal Nonlinear Filters For Discrete Inputs," IEEE Trans On Information Theory, vol. IT-14, pp. 16-21; January, 1968.

5. J.T.Tou,"Digital and Sampled-Data Control Systems," Mc-Graw Hilli New York, N.Y; 1959.

6. R.A. Roberts and J.R.Tooley, "Estimation and Finite Memory," IEEE Trans. on Information Theory," vol. IT-16, pp. 685-691; November, 1970.

7. G.C.Ranieri, "Linear and Nonlinear Digital Processing of Noisy Signals," M.S. Thesis, Dept. of Electrical Engineering, Univ. of Notre Dame;1971.

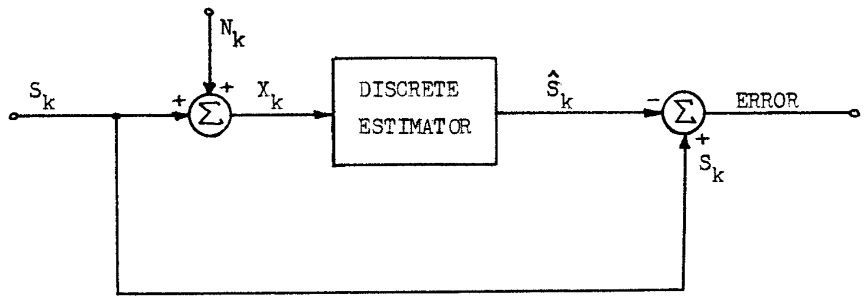

FIGURE 1. Discrete estimation system model.

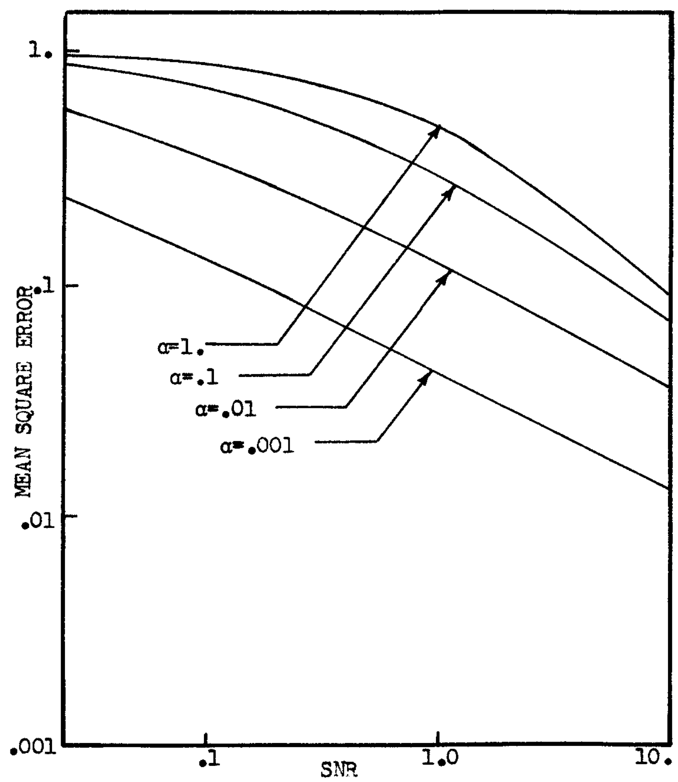

FIGURE 2. Mean square error versus SNP for the optimum realizable discrete Wiener-Hopf filter, where v=1.0.

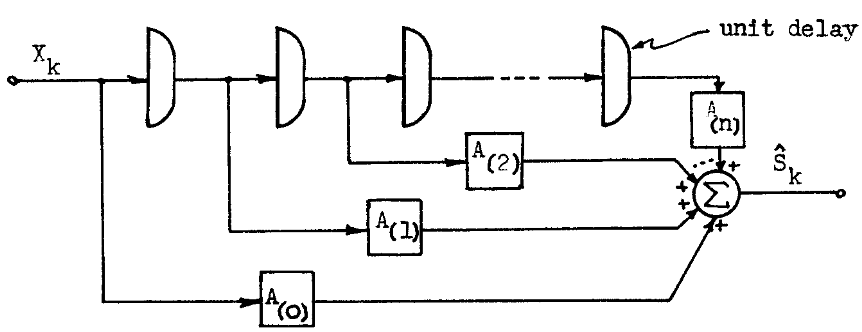

FIGURE 3. Model of the discrete linear nonrecursive estimator.

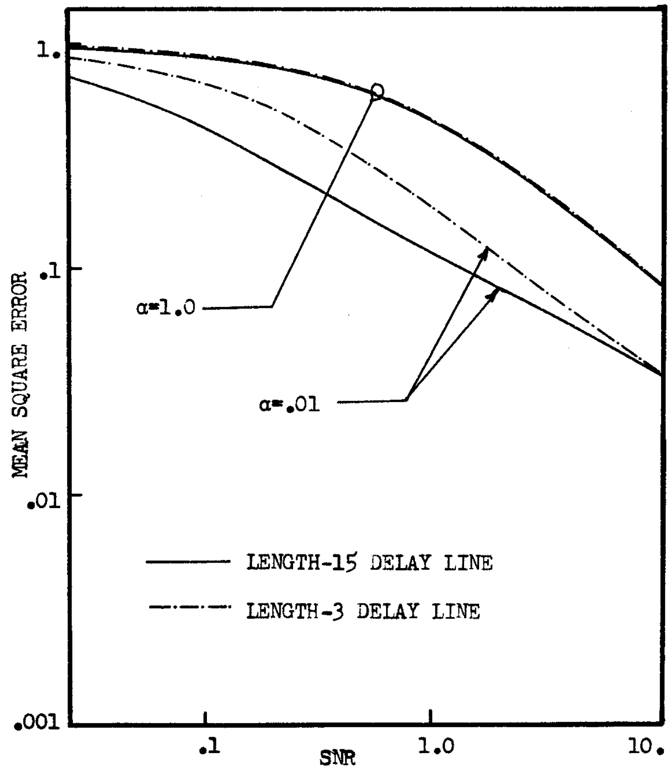

FIGURE 4. Mean square error versus SNR for the discrete linear non-recursive filters of lengths 3 and 15 with v=1.0.

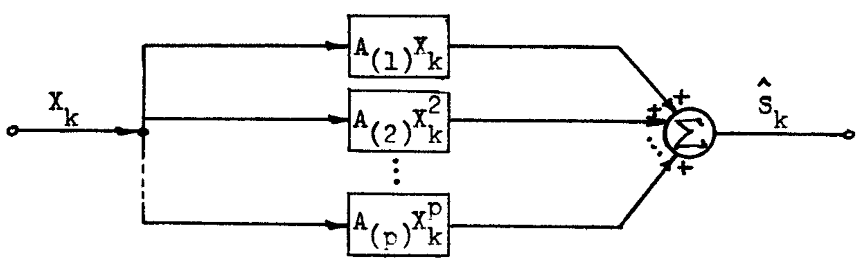

FIGURE 5. The model of the discrete nonrecursive Type estimator.

POLYNOMIAL DEGREE	MEAN SQUARE ERROR ρ=.5	ρ=2.
1	.800000	.200000
2	.777183	.194297
3	.777089	.194274
4	.776643	.194171
5	-------	.194166

FIGURE 6. Mean square error versus polynomial degree for the discrete nonrecursive Polynomial type filter, where p of the noise probability density is .5 and 2.0.