Adaptive Filter for Noise Canceling

Adaptive signal processing systems have been used in several biomedical applications, such as removal of 60 Hz power line interference, and cancellation of maternal ECG from the ECG of the unborn baby (Widrow et al 1975). These systems are digital filters with coefficients, or "weights" (Steams, 1988). The driving force for adjusting the weights is the reduction of the mean square error between the desired and the actual response. A special case is the performance feedbac$_k$, adaptive interference canceling system (fig. 1.23) (Widrow et al., 1975). Consider d_k and x_k to be generated by, a pressure pulse associated reference (noise) sensor. The subscript $_k$, following Widrow's notation, denotes time[1] and can range from zero to infinity. The adaptive system see$_k$s to reduce the difference between the pulse signal d_k, and its own response Y_k, by updating weights, at each sample time $_k$, so that the mean square error MSE = $E[E_k^2) = (1/N)\ L\ (E_k^2)$ approaches its minimum value. When s_k and n_k are uncorrelated, the minimal error is obtained when $Y_k = n_k$. Output pulse points, calculated nonrecursively, are given by the error E_k. Although the use of multiple noise inputs are also possible (Steams, 1988), only a single noise input is considered here.

The adaptive process, by see$_k$ing the minimal mean squared error, amounts to a search for the lowest point on a "performance surface", or "error surface" in multidimensional space. Equations describing the surface can be developed by first defining the correlation functions for the signals in fig. 1:

cross correlation $r_{dx}(n) = E\,[d_k, x_{k+n}] = (1/N)\ \Sigma\ (d_k - x_{k+n})$ (17)

auto correlation $r_{xx}(n) = E\,[x_k, x_{k+n}] = (1/N)\ \Sigma\ (x_k - x_{k+n})$ (18)

where n represents a difference in starting points for correlation of the signals. By definition:

$r_{xd}\,(n) = r_{dx}\,(-n)$ (19)

The MSE can be estimated by assuming fixed filter weights so that y_k is stationary:

$\text{MSE} = \xi = E\,[\varepsilon_k^2\,] = E\,[(d_k - y_k)^2]$ (20)

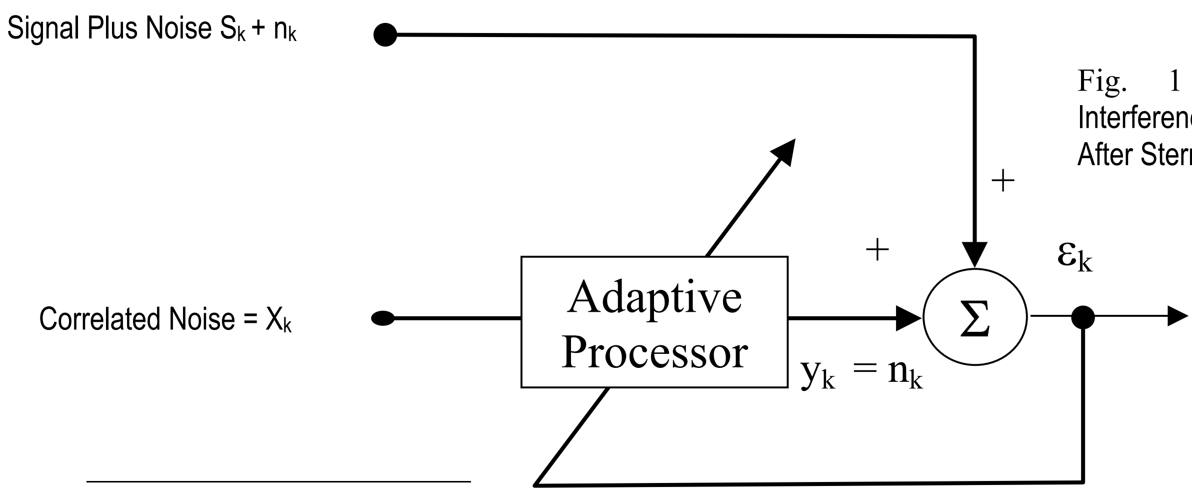

Fig. 1 Method of Adaptive Interference Canceling.
After Sterns, 1988.

[1]For instance, $d_k = d(k) = d(t)$

$$\xi = E [d_k^2] + E [y_k^2] - 2 E[d_k * y_k] \tag{21}$$

$$= r_{dd}(o) + re_{yy} (0) - 2 * r_{dy}(0) \tag{22}$$

It would be useful to $_k$now the MSE in terms of the correlation functions which pertain to the input signal x, rather than processed signal y. The power spectral relationships of the correlation functions can be used. The x - transform of any correlation function is its corresponding power spectrum (Oppenheim and Schafer, 1975), e.g.:

$$G_{uv}(z) = \Sigma\, r_{uv} (n) * z^{-n} \tag{23}$$

and the inverse z transform relationship is:

$$r_{uv}(n) = (1/2\pi j) \int G_{uv} (z) * z^n * (dz/z) \tag{24}$$

Using Eq. (1.24) with n = 0 in Eq. (1.22), the MSE is:

$$MSE = r_{dd} (0) + (1/2\pi j) \int [G_{yy} (z) - G_{dy} (z)] (dz/z) \tag{25}$$

The power spectral relationships for fig. 1.23 are (Widrow and Steams, 1985):

$$G_{yy} (z) = W(z) * W(z^{-1}) * G_{xx} (z) \tag{26}$$

$$G_{dy} (z) = W(z) * W(z^{-1}) * G_{dx} (z) \tag{27}$$

where W(z) represents the z-transform of the weight vector. Incorporating Eqs. (1.26) and (1.27) in Eq. (1.25), the expression for MSE is:

$$MSE = r_{dd}(0) + (1/2\pi\, j) \int [W(z^{-1}) * G_{xx}(z) - 2 * G_{dx}(z)] * W(z)] \ (dz/z) \tag{28}$$

Assuming W(Z) is a causal filter with n weights (Stearns, 1988):

$$W(Z) = \Sigma\, w_i * z^{-I} \qquad (i = 0, L-1) \tag{29}$$

Then:

$$MSE = r_{dd} (0) + \Sigma\ \Sigma\, w_i\, w_m\, r_{xx} (i-m) -2\ \Sigma\ w_i\ r_{xd}(i) \qquad \text{for } (i,m = 0, L-1) \tag{30}$$

This is the general expression for the performance surface of a causal FIR adaptive filter with L weights. The MSE is a quadratic surface because the weights appear only to the first and second degrees in Eq. (1.30). Using the following definitions, which assume a single (noise) input, Eq. (1.30) can be simplified:

single input signal: $\mathbf{X}_k = [x_k\ x_{k-1} \ldots x_{k-L+1}]^T$ (31)

auto correlation matrix $\mathbf{R} = E[\mathbf{X}_k\ \mathbf{X}_k]$ (32)

cross conelation vector: $\mathbf{P} = E[d_k \mathbf{X}_k]$ (33)

$= [r_{xd}(0)\ r_{xd}(1) \ldots x_{xd}(L\text{-}1)]^T$

weight vector: $\mathbf{W} = [w_o\ w_1 \ldots w_{L-1}]^T$ (34)

Then Eq. 1.30 can be written equivalently, as:

$$MSE = r_{dd}(0) + \mathbf{W}^T \mathbf{R}\, \mathbf{W} - 2\, \mathbf{P}^T \mathbf{W} \quad (35)$$

Since the MSE is always positive, the performance surface, being quadratic, must be parabolic, and "concave up" (Stearns, 1988). For stationary signals, the performance surface is thus a parabolic "bowl" in L+ l dimensional space (fig. 2). In typical applications with nonstationary , slowly varying signal statistics, the bowl drifts in space as the signal properties change slowly with time. Adaptation is the process of searching for and continually tracking the bottom of the bowl.

One method to seek the bottom of the weight bowl is to use a gradient (Widrow et al., 1975). Differentiating Eq. 1.35 with respect to the weight vector, the gradient is:

$$\nabla = \partial (MSE) / \partial \mathbf{W} = \partial (MSE) / \partial w_o \quad \partial (MSE) / \partial w_1 \ldots \partial (MSE) / \partial w_{L-1} \quad (36)$$

$$= 2\, \mathbf{R}\, \mathbf{W} - 2\, \mathbf{P}$$

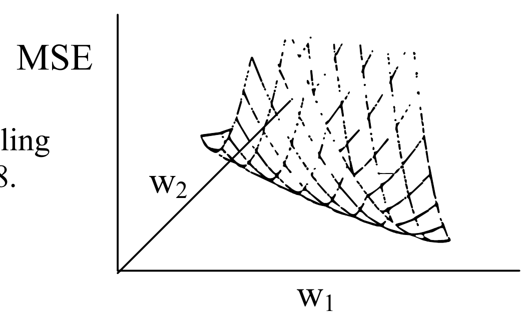

Fig. 2. Example of an interference canceling performance surface. After Stearns, 1988.

Since the bowl is quadratic, the global minimum MSE is obtained when $\nabla = 0$. The optimum weight vector, $\mathbf{W}^*$, is then obtained:

$$\mathbf{W}^* = \mathbf{R}^{-1} \mathbf{P} = \mathbf{W} - (1/2) * \mathbf{R}^{-1} * \nabla \quad (37)$$

The least mean squares (LMS) algorithm of Widrow et al. (1975) is based on Making a local estimate of the gradient ∇, and moving incrementally downward toward the bottom of the bowl. Therefore one adaptive cycle consists of determining ∇ at time k and making an appropriate adjustment in the weight vector from $\mathbf{W}_k$ to $\mathbf{W}_{k+l}$ in order to move toward $\mathbf{W}^*$ at the bottom of the bowl, which can be written as (Steams, 1988):

$$\mathbf{W}_{k+l} = \mathbf{W}_k - \mu * \mathbf{R}^{-1} * \nabla_k \quad (38)$$

Where u is a constant and determines the time for convergence. In general the **R** matrix is not known, but it is related to the signal power, and Eq. (1.38) can be approximated as (Stearns, 1988):

$$\mathbf{W}_{k+l} = \mathbf{W}_k + 2\,\mu * \varepsilon_k * \mathbf{X}_k * (1/L\,\sigma^2) \qquad (39)$$

for 0 < u < 1. This is the LMS algorithm. The input signal power sig2 can be estimated, for instance, by squaring the noise signal and averaging it with an exponential low pass filter and suitably long time constant. Employing one weight, Eq. (1.39) becomes:

$$w_k{+}l = w_k + 2\mu * (d_k - w_k\, x_k + [LPF[x_k] - LPF\,[d_k]) * x_k / LPF[x_k^2] \qquad (40)$$

where signal offset LPF $[x_k]$ -LPF $[d_k]$ is subtracted to provide best signal overlap of pulse signals to noise signals, and the signal power μ is estimated by LPF $[x_k^2]$. Lin and Chang (1988) implemented the LMS algorithm for processing of the pulmonary artery pressure signal. Low frequency artifact, added to simulate periodic respiration, was removed adaptively to obtain a stable baseline. Addition of more weights was found to enhance accuracy but also prolonged the convergence time.