ECE1511,

In this problem we will compare LMS and RLS for adaptive linear prediction. [redacted] let $x(n)$ be a process that is generated according to the difference equation

$$x(n) = 1.2728x(n-1) - 0.81x(n-2) + v(n)$$

where $v(n)$ is unit variance white Gaussian noise. The adaptive linear predictor will be of the form

$$\hat{x}(n) = w_n(1)x(n-1) + w_n(2)x(n-2)$$

(a) Implement an RLS adaptive predictor with $\lambda = 1$ (growing window RLS) and plot $w_n(k)$ versus n for $k = 1, 2$. Compare the convergence of the coefficients $w_n(k)$ to those that are obtained using the LMS algorithm for several different values of the step size μ.

(b) Make a plot of the learning curve for RLS and compare it to the LMS learning curve [redacted] Example 9.2.2 on how to plot learning curves). Comment on the excess mean-square error for RLS and discuss how it compares to that for LMS.

(c) Repeat part (b) for exponential weighting factors of $\lambda = 0.99, 0.95, 0.92, 0.90$ and discuss the trade-offs involved in the choice of λ.

(d) Modify the m-file for the RLS algorithm to implement a sliding window RLS algorithm.

(e) Let $x(n)$ be generated by the time-varying difference equation

$$x(n) = a_n(1)x(n-1) - 0.81x(n-2) + v(n)$$

where $v(n)$ is unit variance white noise and $a_n(1)$ is a time-varying coefficient given by

$$a_n(1) = \begin{cases} 1.2728 & ; \quad 0 \le n < 50 \\ 0 & ; \quad 50 \le n < 100 \\ -1.2728 & ; \quad 100 \le n \le 200 \end{cases}$$

Compare the effectiveness of the LMS, growing window RLS, exponentially weighted RLS, and sliding window RLS algorithms for adaptive linear prediction. What approach would you propose to use for linear prediction when the process has step changes in its parameters?

Example 9.2.2 *LMS Misadjustment*

In this example, we look at the learning curves for the adaptive linear predictor considered in Example 9.2.1, and evaluate the excess mean-square error and the misadjustment for different step sizes. Since the learning curve is a plot of $\xi(n) = E\{|e(n)|^2\}$ versus n, we may approximate the learning curve by averaging plots of $|e(n)|^2$ that are obtained by repeatedly implementing the adaptive predictor. For example, implementing the adaptive predictor K times, and denoting the squared error at time n on the kth trial by $|e_k(n)|^2$, we have

$$\hat{\xi}(n) = \hat{E}\{|e(n)|^2\} = \frac{1}{K}\sum_{k=1}^{K}|e_k(n)|^2$$

With $K = 200$, an initial weight vector of zero, and step sizes of $\mu = 0.02$ and $\mu = 0.004$, these estimates of the learning curves are shown in Fig. 9.11. One property of the LMS algorithm that we are able to observe from these plots is that, when the step size is decreased, the convergence of the adaptive filter to its steady-state value is slower, but the average steady-state squared error is smaller.

We may estimate the steady-state mean-square error from these plots by averaging $\hat{\xi}(n)$ over n after the LMS algorithm has reached steady-state. For example, with

$$\hat{\xi}(\infty) = \frac{1}{100}\sum_{n=901}^{1000}\hat{E}\{|e(n)|^2\}$$

we find

$$\hat{\xi}(\infty) = \begin{cases} 1.1942 & ; \quad \text{for } \mu = 0.02 \\ 1.0155 & ; \quad \text{for } \mu = 0.004 \end{cases}$$

We may compare these results to the theoretical steady-state mean-square error using Eq. (9.43). With $\xi_{\min} = 1$, and eigenvalues $\lambda_1 = 9.7924$ and $\lambda_2 = 1.7073$ (see Example 9.2.1), it follows that

$$\xi(\infty) = \begin{cases} 1.1441 & ; \quad \text{for } \mu = 0.02 \\ 1.0240 & ; \quad \text{for } \mu = 0.004 \end{cases}$$

which is in fairly close agreement with the estimated values given above.

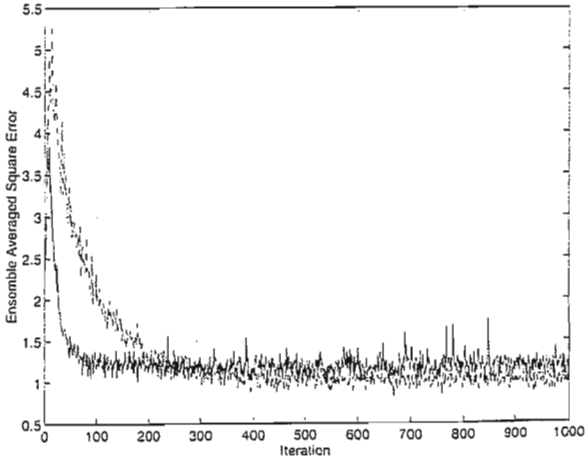

Figure 9.11 *Approximations to the learning curves for a second-order LMS adaptive linear predictor using step sizes of $\mu = 0.02$ (solid line) and $\mu = 0.004$ (dotted line).*

Example 9.2.1 *Adaptive Linear Prediction Using the LMS Algorithm*

Let $x(n)$ be a second-order autoregressive process that is generated according to the difference equation

$$x(n) = 1.2728x(n-1) - 0.81x(n-2) + v(n) \tag{9.42}$$

where $v(n)$ is unit variance white noise. As we saw in Section 7.3.4, the optimum causal linear predictor for $x(n)$ is

$$\hat{x}(n) = 1.2728x(n-1) - 0.81x(n-2)$$

However, in order to design this predictor (i.e., to know that the optimum predictor coefficients are 1.2728 and -0.81) it is necessary to know the autocorrelation sequence of $x(n)$. Therefore, suppose we consider an adaptive linear predictor of the form:

$$\hat{x}(n) = w_n(1)x(n-1) + w_n(2)x(n-2)$$

as shown in Fig. 9.9. With the LMS algorithm, the predictor coefficients $w_n(k)$ are updated as follows:

$$w_{n+1}(k) = w_n(k) + \mu e(n)x^*(n-k)$$

If the step size μ is sufficiently small, then the coefficients $w_n(1)$ and $w_n(2)$ will converge in the mean to their optimum values, $w(1) = 1.2728$ and $w(2) = -0.81$, respectively. Note that the prediction error is

$$e(n) = x(n) - \hat{x}(n) = \big[1.2728 - w_n(1)\big]x(n-1) + \big[-0.81 - w_n(1)\big]x(n-2) + v(n)$$

Therefore, when $w_n(1) = 1.2728$ and $w_n(2) = -0.81$, the error becomes $e(n) = v(n)$, and the minimum mean-square error is[3]

$$\xi_{\min} = \sigma_v^2 = 1$$

Although we might expect the mean-square error $E\{|e(n)|^2\}$ to converge to $\xi_{\min}$ as $\mathbf{w}_n$ converges to $\mathbf{w}$, as we will soon discover, this is not the case.

To see how this adaptive linear predictor behaves in practice, suppose that the weight vector is initialized to zero, $\mathbf{w}_0 = \mathbf{0}$, and that the step size is $\mu = 0.02$. Shown in Fig. 9.10*a*

[3] This may also be shown by evaluating the expression for the minimum mean-square error given in Eq. (9.26).

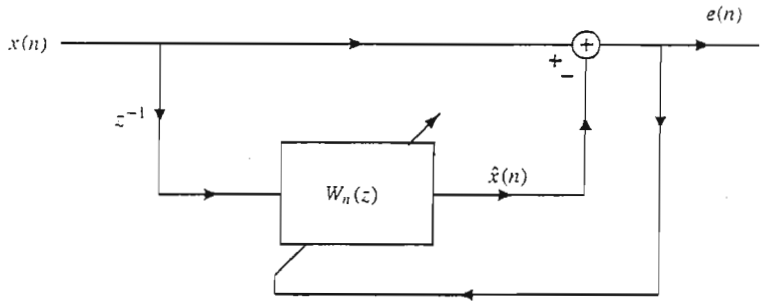

Figure 9.9 *An adaptive filter for linear prediction.*

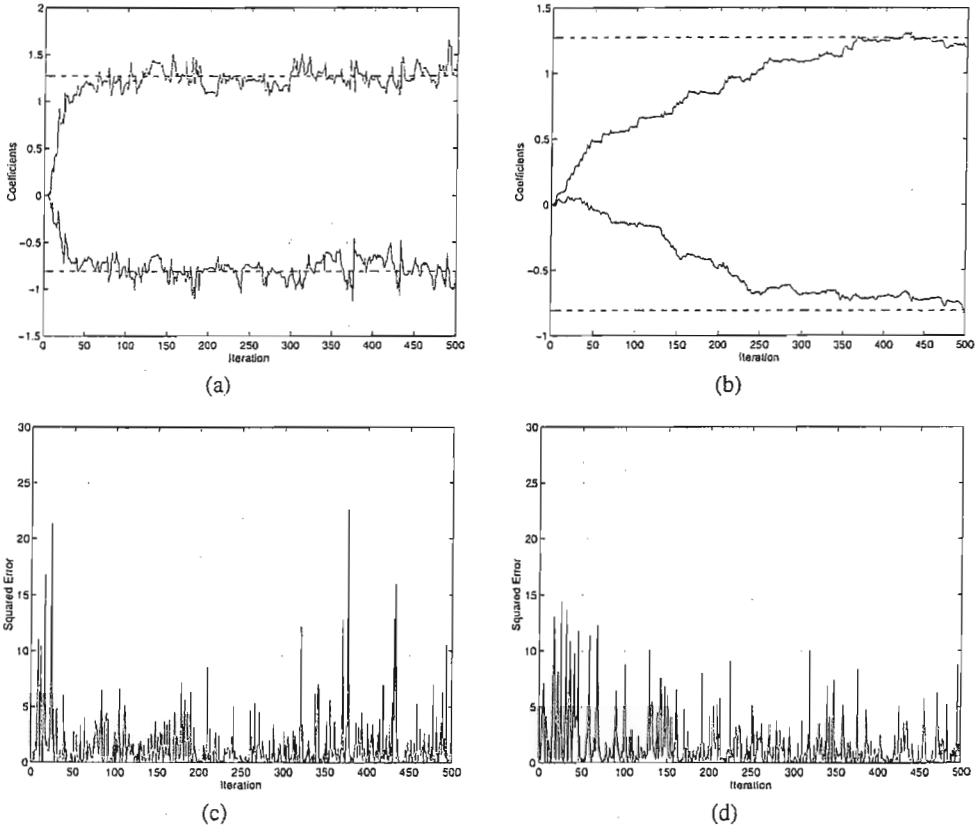

Figure 9.10 *Performance of a two-coefficient LMS adaptive linear predictor. The trajectories of the predictor coefficients are shown for step sizes of (a) $\mu = 0.02$ and (b) $\mu = 0.004$ with the correct values indicated by the dashed lines. A plot of the squared error, $e^2(n)$, is shown in (c) and (d) for step sizes of $\mu = 0.02$ and $\mu = 0.004$, respectively.*

are the trajectories of $w_n(1)$ and $w_n(2)$ versus n. As we see, there is a great deal of fluctuation in the weights, even after they have converged to within a neighborhood of their steady-state values. By contrast, shown in Fig. 9.10*b* are the trajectories of the predictor coefficients for a step size $\mu = 0.004$. Compared to a step size of $\mu = 0.02$, we see that, although the weights take longer to converge, the trajectories are much smoother, illustrating the basic trade-off