HOMEWORK ASSIGNMENT #5

Due Tues. Oct. 26, 2005 (in class)

1. In this problem we will investigate *adaptive forward linear prediction.* Assume the configuration in Fig. 1, where $W^*(z) = \sum_{k=0}^{M-1} w_k^* z^{-k}$ and $\{d(n)\}$ is an AR-P process with feedback coefficients $\{a_1, a_2, \ldots, a_P\}$ driven by zero-mean white input noise $\{v(n)\}$ with variance σ_v^2. We refer to $e(n)$ as the "forward prediction error".

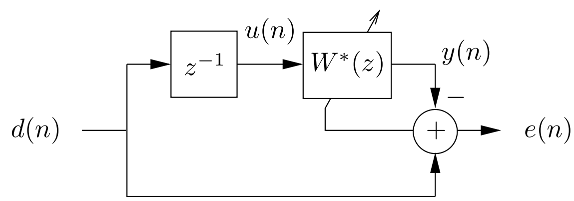

Figure 1: Adaptive Forward Linear Predictor

   (a) Derive expressions for the $\mathrm{E}\{|e(n)|^2\}$-minimizing coefficient vector $\boldsymbol{w}_\star$ and the corresponding minimum error variance $\sigma_e^2\big|_{\min}$ in terms of the autocorrelation sequence of $\{d(n)\}$.

   (b) Assuming that $M \geq P$, rewrite $\boldsymbol{w}_\star$ and $\sigma_e^2\big|_{\min}$ in terms of $\boldsymbol{a} = [a_1, a_2, \ldots, a_P]^t$ and σ_v^2.

   (c) What can we say about the ability of a M^{th}-order predictor to whiten an AR-P process when $M \geq P$? When $M < P$? (Hint: Examine the transer function from $v(n)$ to $e(n)$.)

   (d) From here on assume that $M = 2$ and that $\{d(n)\}$ is a real-valued AR-2 process generated from feedback taps $[a_1, a_2] = [0.6, -0.1]$ and driving noise variance $\sigma_v^2 = 0.2$. We would now like to simulate LMS adaptation of $\boldsymbol{w}$ and compare to GD-MSE optimization of $\boldsymbol{w}$. To do this, you will generate four plots, each showing MSE contours superimposed with principle axes, MMSE solution, GD-MSE trajectories, and ten LMS parameter trajectories. For each LMS trajectory, you will re-generate the random input $\{d(n)\}$. Use trajectory lengths of $N = 5000$. The four plots will correspond to the different combinations of

$$\boldsymbol{w}(0) = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \end{pmatrix} \right\} \qquad \text{and} \qquad \mu = \{.02, .2\}.$$

   (Note that most of your code can be reused from previous homework assignments.) An example of one plot is given in Fig. 2.

   (e) Finally, we would like to evaluate the excess MSE of LMS. To do this, empirically estimate the mean-square prediction error (i.e., $\widehat{\sigma_e^2} = \frac{1}{N-1}\sum_{n=1}^{N} |e(n)|^2$) using $\mu = 0.02$, $N = 10^6$, and $\boldsymbol{w}(0) = \boldsymbol{w}_\star$. Compare your answer to the expression derived in class:

$$J_{\text{emse}} \;\approx\; \frac{\mu}{2}\,\mathrm{tr}(\boldsymbol{R})J_{\min}.$$

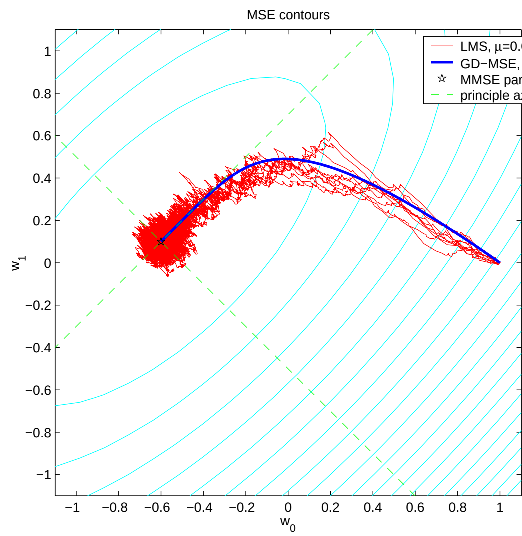

Figure 2: Example LMS and GD-MSE trajectories.

2. In this problem we study fixed-stepsize LMS for time-varying system identification. In Fig. 3, $\boldsymbol{h}(n)$ represents the time-n impulse response of an unknown linear system, $\boldsymbol{w}(n)$ is the time-n impulse response of an adaptive filter that is used to identify the system, $\{z(n)\}$ represents measurement noise, and $\{x(n)\}$ the input process. As usual, our goal is minimization of $\mathrm{E}\{|e(n)|^2\}$.

   Assume that $\{x(n)\}$ is a zero-mean WSS random process with full rank autocorrelation matrix $\boldsymbol{R}_x$, that $\{z(n)\}$ is zero-mean white noise with variance σ_z^2 uncorrelated with $\{x(n)\}$, and that $\{\boldsymbol{h}(n)\}$ is generated by the random walk model $\boldsymbol{h}(n+1) = \boldsymbol{h}(n) + \boldsymbol{q}(n)$ with zero-mean i.i.d. vector process $\{\boldsymbol{q}(n)\}$ having autocorrelation $\boldsymbol{Q} = \mathrm{E}\{\boldsymbol{q}(n)\boldsymbol{q}^H(n)\}$.

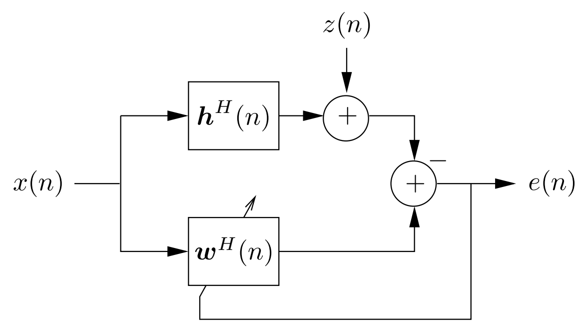

Figure 3: Adaptive Time-Varying System Identification

   (a) For $\boldsymbol{w}$ and $\boldsymbol{h}$ of the same length, derive an expression for $\mathrm{E}\{|e(n)|^2|\boldsymbol{w}(n), \boldsymbol{h}(n)\}$. In other words, fix $\boldsymbol{w}(n)$ and $\boldsymbol{h}(n)$, then express your answer in terms of the statistics of $\{x(n)\}$ and $\{z(n)\}$.

   (b) Under the same assumptions as (a), derive expressions for the MMSE identifier

$$\boldsymbol{w}_\star(n) = \arg\min_{\boldsymbol{w}(n)} \mathrm{E}\{|e(n)|^2|\boldsymbol{w}(n), \boldsymbol{h}(n)\}$$

   and the corresponding MMSE error $J_{\min}(n) = \min_{\boldsymbol{w}(n)} \mathrm{E}\{|e(n)|^2|\boldsymbol{w}(n), \boldsymbol{h}(n)\}$.

(c) Say that we want to generate $\{\boldsymbol{q}(n)\}$ for a particular $\boldsymbol{Q} = \mathrm{E}\{\boldsymbol{q}(n)\boldsymbol{q}^H(n)\}$ using $\boldsymbol{q}(n) = \boldsymbol{B}\boldsymbol{\nu}(n)$ where $\{\boldsymbol{\nu}(n)\}$ has uncorrelated elements (i.e., $\mathrm{E}\{\boldsymbol{\nu}(n)\boldsymbol{\nu}^H(n)\} = \boldsymbol{I}$). How might we choose $\boldsymbol{B}$?

(d) For the remaining parts of this problem, assume that

- $\{x(n)\}$ is generated by an AR-4 model with feedback coefficients $[a_1, a_2, a_3, a_4] = [0.6, -0.1, 0.1, 0.1]$ and driving noise variance $\sigma_v^2 = 0.5$.
- $\sigma_z^2 = 0.1$.
- $\boldsymbol{Q} = 10^{-4} \times \begin{pmatrix} 1.1 & 0.1 & 0.1 \\ 0.1 & 1.1 & 0.1 \\ 0.1 & 0.1 & 1.1 \end{pmatrix}$
- $\boldsymbol{h}(0) = [-1, 0, 1]^t$.
- $\boldsymbol{h}$ and $\boldsymbol{w}$ have equal length.

Calculate, in Matlab, the optimal stepsize μ_{opt} and the resulting steady-state MSE, i.e., $\lim_{n\to\infty} J(n)$, using the non-stationary EMSE equations derived in the lecture.

(e) Simulate LMS adaptation of $\boldsymbol{w}(n)$ using length-2500 sequences, the optimal stepsize, and initial value $\boldsymbol{w}(n) = \boldsymbol{0}$. Generate a plot that superimposes the true parameters $\{\boldsymbol{h}(n)\}$ and adapted estimates $\{\boldsymbol{w}(n)\}$ versus n. Generate a second plot that superimposes $|e(n)|^2$, $\mathrm{E}\{|e(n)|^2 | \boldsymbol{w}(n), \boldsymbol{h}(n)\}$, and the value of $\lim_{n\to\infty} J(n)$ that you calculated in (d), *all on a dB scale*. The results should look something like Fig. 4.

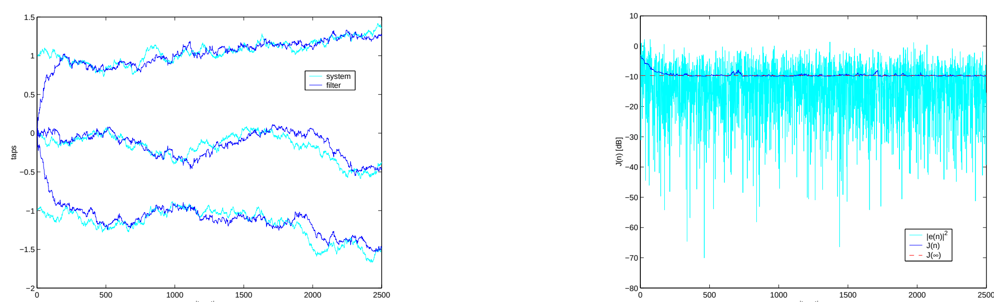

Figure 4: LMS time-varying system identification.

(f) Now we take a closer look at the non-stationary EMSE approximation. Plot theoretical EMSE as a function of μ over the range 10^{-3} to 10^{-1}. For each value $\mu \in \{0.1, \mu_{\text{opt}}, 0.01, 0.002\}$, run LMS for many iterations (≥ 50000 samples) intialized at $\boldsymbol{w}(0) = \boldsymbol{w}_\star(0)$ and empirically estimate the EMSE in two ways: using time-averages of $|e(n)|^2$ and of $\mathrm{E}\{|e(n)|^2 | \boldsymbol{w}(n), \boldsymbol{h}(n)\}$. Superimpose these two sets of EMSE estimates onto the theoretical EMSE plot. The result should look something like Fig. 5.

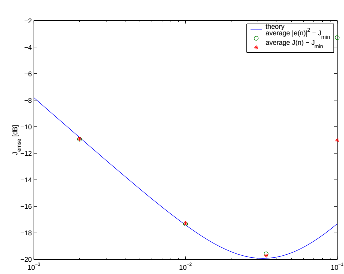

Figure 5: EMSE versus μ.