**HOMEWORK SOLUTIONS #5**

1. (a) From Wiener theory, we know

$$\begin{aligned} \boldsymbol{w}_\star &= \boldsymbol{R}^{-1}\boldsymbol{p} \quad \text{for} \\ \boldsymbol{R} &= \mathrm{E}\{\boldsymbol{u}(n)\boldsymbol{u}^H(n)\} = \mathrm{E}\{\boldsymbol{d}(n)\boldsymbol{d}^H(n)\} \\ \boldsymbol{p} &= \mathrm{E}\{\boldsymbol{u}(n)d^*(n)\} = \mathrm{E}\{\boldsymbol{d}(n)d^*(n)\} \end{aligned}$$

and so

$$\begin{aligned} \boldsymbol{R} &= \begin{pmatrix} r(0) & r(1) & \dots & r(M-1) \\ r^*(1) & r(0) & & \\ \vdots & & \ddots & \vdots \\ r^*(M-1) & & \dots & r(0) \end{pmatrix} \\ \boldsymbol{p} &= \begin{pmatrix} r(1) & r(2) & \dots & r(M) \end{pmatrix}^H \end{aligned}$$

where $r(k) = \mathrm{E}\{d(n)d^*(n-k)\}$. We also know

$$\sigma_e^2\big|_{\min} = \sigma_d^2 - \boldsymbol{p}^H\boldsymbol{R}^{-1}\boldsymbol{p} = r(0) - \boldsymbol{p}^H\boldsymbol{R}^{-1}\boldsymbol{p}.$$

(b) From the Yule-Walker equations for an AR-$P$ model,

$$-\begin{pmatrix} r(0) & r(1) & \dots & r(P-1) \\ r^*(1) & r(0) & & \\ \vdots & & \ddots & \vdots \\ r^*(P-1) & & \dots & r(0) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_P \end{pmatrix} = \begin{pmatrix} r^*(1) \\ r^*(2) \\ \vdots \\ r^*(P) \end{pmatrix} \quad \text{and} \quad \sigma_v^2 = \begin{pmatrix} r^*(0) \\ r^*(1) \\ \vdots \\ r^*(P) \end{pmatrix}^H \begin{pmatrix} 1 \\ a_1 \\ \vdots \\ a_P \end{pmatrix}$$

so augmenting the coefficient vector with $M-P$ zeros (since $P \le M$),

$$-\begin{pmatrix} r(0) & r(1) & \dots & r(M-1) \\ r^*(1) & r(0) & & \\ \vdots & & \ddots & \vdots \\ r^*(M-1) & & \dots & r(0) \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_P \\ \boldsymbol{0} \end{pmatrix} = \begin{pmatrix} r^*(1) \\ r^*(2) \\ \vdots \\ r^*(M) \end{pmatrix} \quad \text{and} \quad \sigma_v^2 = \begin{pmatrix} r^*(0) \\ r^*(1) \\ \vdots \\ r^*(M) \end{pmatrix}^H \begin{pmatrix} 1 \\ a_1 \\ \vdots \\ a_P \\ \boldsymbol{0} \end{pmatrix}.$$

The last two equations can be rewritten using $\boldsymbol{a} = [a_1, \dots, a_P]^t$ and the quantities $\boldsymbol{R}$ and $\boldsymbol{p}$ defined earlier.

$$-\boldsymbol{R}\begin{pmatrix} \boldsymbol{a} \\ \boldsymbol{0} \end{pmatrix} = \boldsymbol{p}^H \quad \text{and} \quad \sigma_v^2 = r(0) + \boldsymbol{p}^H \begin{pmatrix} \boldsymbol{a} \\ \boldsymbol{0} \end{pmatrix}.$$

This implies that

$$\boldsymbol{w}_\star = \boldsymbol{R}^{-1}\boldsymbol{p} = -\begin{pmatrix} \boldsymbol{a} \\ \boldsymbol{0} \end{pmatrix}$$

and that

$$\sigma_e^2\big|_{\min} = r(0) - \boldsymbol{p}^H\boldsymbol{R}^{-1}\boldsymbol{p} = r(0) + \boldsymbol{p}^H \begin{pmatrix} \boldsymbol{a} \\ \boldsymbol{0} \end{pmatrix} = \sigma_v^2.$$

(c) For an AR-$P$ process $\{d(n)\}$ generated from white noise $\{v(n)\}$, we can say that $D(z) = \sum_n d(n)z^{-n}$ and $V(z) = \sum_n v(n)z^{-n}$ are related by

$$\begin{aligned} D(z) &= (1 + a_1^* z^{-1} + a_2^* z^{-2} + \cdots + a_P^* z^{-P})^{-1} V(z) \\ &= (1 + z^{-1} A^*(z))^{-1} V(z) \end{aligned}$$

where $A^*(z) = a_1^* + a_2^* z^{-1} + \cdots + a_P^* z^{-P+1}$. The forward linear prediction error $\{e(n)\}$ can be written in terms of $E(z) = \sum_n e(n)z^{-n}$, which in turn can be expressed as

$$E(z) = (1 - z^{-1} W^*(z)) D(z)$$

for $W^*(z) = w_0^* + w_1^* z^{-1} + \cdots + w_{M-1}^* z^{-M+1}$. Thus the transfer function from $V(z)$ to $E(z)$ is

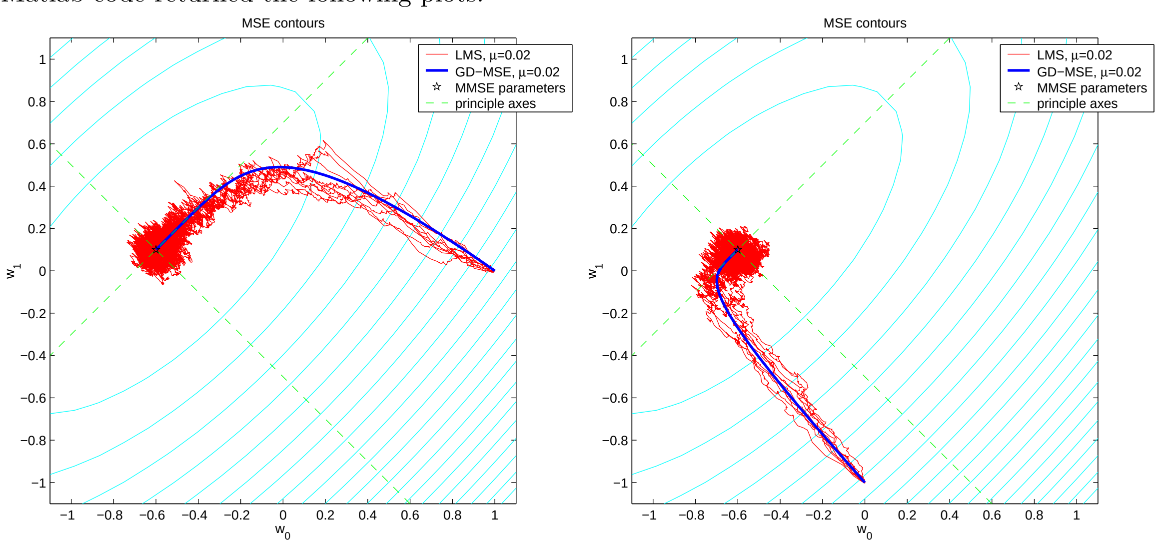
$$\frac{E(z)}{V(z)} = \frac{1 - z^{-1} W^*(z)}{1 + z^{-1} A^*(z)}$$

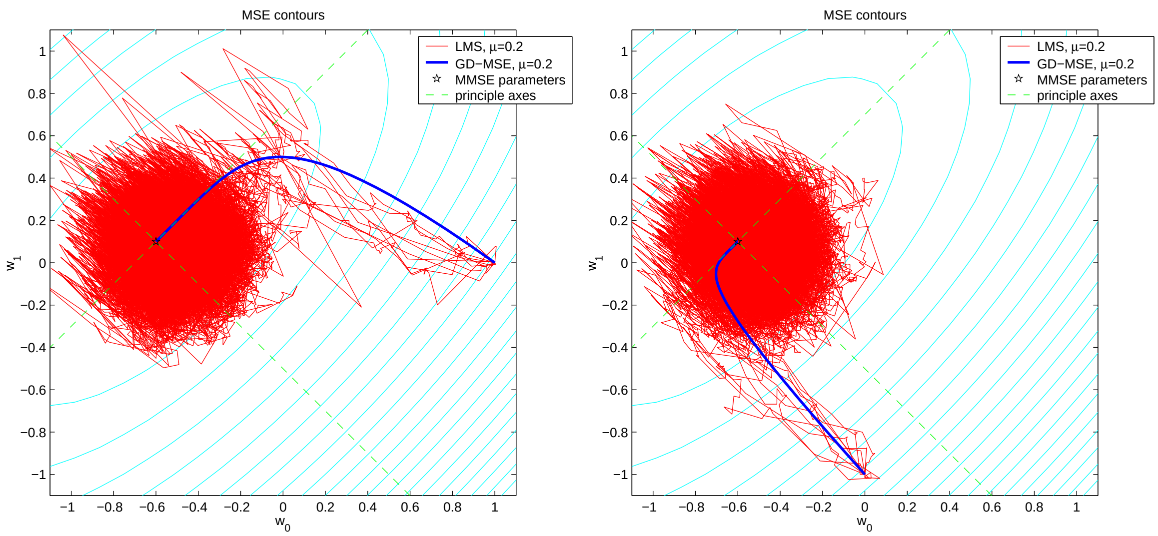
For $M \geq P$, we found that the Wiener solution yields $W_\star(z) = -A(z)$, in which case

$$\frac{E_\star(z)}{V(z)} = \frac{1 - z^{-1} W_\star(z)}{1 + z^{-1} A(z)} = \frac{1 + z^{-1} A^*(z)}{1 + z^{-1} A^*(z)} = 1.$$

Thus, since $\{v(n)\}$ is white, the Wiener error must be white. When $M < P$, however, we know that $W(z)$ does not have enough parameters to equate the numerator and denominator of $\frac{E(z)}{V(z)}$, and so $\{e(n)\}$ will be a filtered version of $\{v(n)\}$ and thus will not be white.

(d) Matlab code returned the following plots:






(e) After 5 million iterations, Matlab code returned simulated $J_{\text{emse}} = 0.00148$ and theoretical $J_{\text{emse}} = 0.00145$.

2. (a) From the block diagram,

$$\begin{aligned} e(n) &= \big(\boldsymbol{h}^H(n)\boldsymbol{x}(n) + z(n)\big) - \boldsymbol{w}^H(n)\boldsymbol{x}(n) \\ &= \big(\boldsymbol{h}(n) - \boldsymbol{w}(n)\big)^H \boldsymbol{x}(n) + z(n) \\ \mathrm{E}\{|e(n)|^2 | \boldsymbol{w}(n), \boldsymbol{h}(n)\} &= \big(\boldsymbol{h}(n) - \boldsymbol{w}(n)\big)^H \boldsymbol{R}_x \big(\boldsymbol{h}(n) - \boldsymbol{w}(n)\big) + \sigma_z^2 \qquad \text{since } \boldsymbol{x}(n), z(n) \text{ uncorrelated} \qquad (1) \end{aligned}$$

(b)

$$\begin{aligned} \nabla_{\boldsymbol{w}(n)} \mathrm{E}\{|e(n)|^2 | \boldsymbol{w}(n), \boldsymbol{h}(n)\} &= 2\boldsymbol{R}_x\big(\boldsymbol{h}(n) - \boldsymbol{w}(n)\big) \\ &= 0 \\ &\Leftrightarrow \boldsymbol{w}_\star(n) = \boldsymbol{h}(n) \qquad \text{for full rank } \boldsymbol{R}_x. \qquad (2) \\ J_{\min} &= \sigma_z^2 \qquad \text{after plugging (2) into (1).} \end{aligned}$$

(c) We could use the Cholesky decomposition:

$$\boldsymbol{Q} = \boldsymbol{A}^H \boldsymbol{A}$$

where $\boldsymbol{A}$ is an upper triangular matrix, since

$$\boldsymbol{q}(n) = \boldsymbol{A}^H \boldsymbol{\nu}(n) \quad \Rightarrow \quad \mathrm{E}\{\boldsymbol{q}(n)\boldsymbol{q}^H(n)\} = \boldsymbol{A}^H \mathrm{E}\{\boldsymbol{\nu}(n)\boldsymbol{\nu}^H(n)\}\boldsymbol{A} = \boldsymbol{A}^H\boldsymbol{A} = \boldsymbol{Q}$$

Or we could use the eigenvalue decomposition

$$\boldsymbol{Q} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^H = \boldsymbol{U}\boldsymbol{\Lambda}^{1/2}\boldsymbol{\Lambda}^{1/2}\boldsymbol{U}^H$$

where $\boldsymbol{U}$ is unitary, $\boldsymbol{\Lambda}$ is diagonal with positive elements, and $\boldsymbol{\Lambda}^{1/2}$ is diagonal with non-zero elements equal to the square-roots of the corresponding elements in $\boldsymbol{\Lambda}$. In this case,

$$\boldsymbol{q}(n) = \boldsymbol{U}\boldsymbol{\Lambda}^{1/2}\boldsymbol{\nu}(n) \quad \Rightarrow \quad \mathrm{E}\{\boldsymbol{q}(n)\boldsymbol{q}^H(n)\} = \boldsymbol{U}\boldsymbol{\Lambda}^{1/2}\mathrm{E}\{\boldsymbol{\nu}(n)\boldsymbol{\nu}^H(n)\}\boldsymbol{\Lambda}^{1/2}\boldsymbol{U}^H = \boldsymbol{Q}$$

The Cholesky approach can be implemented more efficiently since $\boldsymbol{A}$ has many zeros.

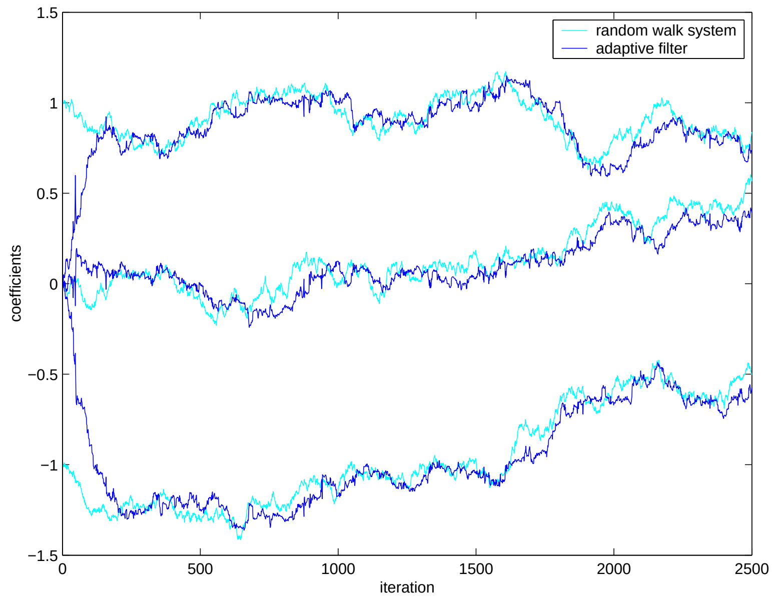



Figure 1: LMS: parameter trajectories.

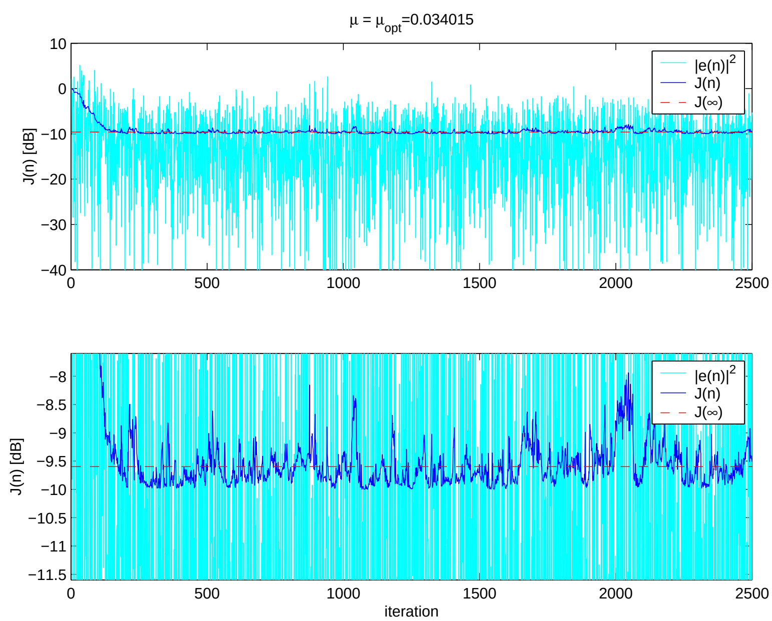



Figure 2: LMS: MSE trajectory.

(d) Matlab returns $\mu_{\text{opt}} = \texttt{0.0340}$ and $J(n) \to \texttt{0.1097 = -9.60 dB}$.

(e) See Fig. 1 and Fig. 2.

(f) Using 50000 samples, we obtain Fig. 3. Note that the EMSE values for $\mu = 0.1$ are far from theoretical; recall that at various points in our EMSE derivation we assumed that $\mu$ is very small, which is not true in this case. The other values are very close to theoretical.

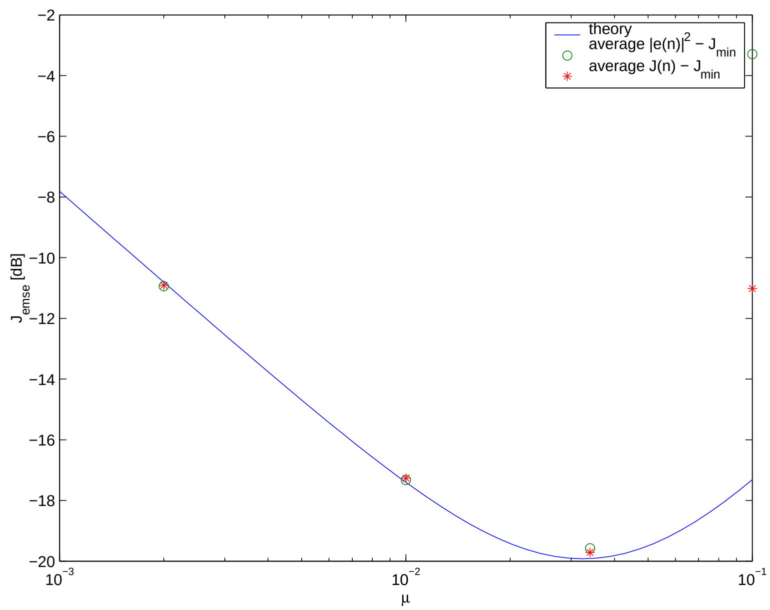



Figure 3: LMS: excess-MSE versus stepsize.