EE 225A Spring 2005

Second Midterm Exam: Solutions

1. Which of the following properties apply to an *arbitrary* $N \otimes N$ circulant matrix? (Check all that apply; there may be more than one.)

   ☺ Toeplitz

   ☐ Positive definite

   ☐ Symmetric

   ☐ Hermitian

   ☐ Eigenvalues are non-negative real-valued

   ☺ Eigenvalues equal the discrete-Fourier transform of any row

   ☐ None of the above

2. Which of the following statements apply to the convergence of a deterministic *steepest descent algorithm* for solving the N -th order Weiner filtering design problem as covered in class and in the textbook? (Check all that apply; more than one may be correct.)

   ☐ For the optimum step size μ_{opt}, all modes of convergence are always at the same rate. (**Comment**: The modes corresponding to the maximum and minimum eigenvalues converge at the same rate, but all other modes converge at a faster rate.)

   ☺ If the initial condition of the tap weight vector **w** is optimum, then the tap weight vector will always be optimum even as we iterate the steepest descent algorithm.

   ☺ As the order N increases, and for the optimum step size μ_{opt}, the slowest converging mode(s) converge at a slower and slower rate. (**Comment**: This is a consequence of the Bordering Theorem on p. 48.)

   ☐ As the order N increases, the optimum step size μ_{opt} always remains the same. (**Comment**: In general the minimum eigenvalue will decrease, the maximum eigenvalue increase. Thus, the optimum step size will not vary much, but there is no result that says it will stay the same.)

☺ As the order N increases, the maximum step size permissible for convergence decreases. (**Comment**: the maximum eigenvalue increases, so the aximum step-size decreases.)

☺ As the order N increases, the eigenvalues of the autocorrelation matrix approach uniformly spaced samples of the corresponding power spectrum.

☐ None of the above.

3. Which of the following statements apply to the convergence of an LMS adaptive filter, where the input signals are assumed to be jointly wide-sense stationary? (Check all that apply; there may be more than one.)

☐ If the step size μ is chosen appropriately, the tap weight vector converges to the optimum Weiner filter solution with probability one. (**Comment**: It only converges in the mean, and then only if we make an independence assumption.)

☐ The step size μ that maximizes the rate of convergence of the mean-value of the tap weight vector (in the sense of fastest convergence of the slowest-converging mode) is the same as the step size μ that maximizes the rate of decrease of the mean-square estimation error. (**Comment**: Check this in Hayes—it is presumably different.)

☺ Choosing the step size μ involves a tradeoff between rate of convergence and the excess mean-square estimation error after convergence.

☐ If the initial condition of the tap weight vector $\mathbf{w}$ is optimum, then the tap weight vector will always be optimum. (**Comment**: Stochastic updates will be changing the tap weight vector at every iteration.)

☺ Characterizing the convergence properties usually requires making an assumption about the statistical independence of the tap weight vector and the input signals.

☐ None of the above.

4. Given a real-valued function $P(e^{j\omega}) > 0$, which of the following statements are correct? (Check all that apply; more than one may be correct.)

☺ There exists a wide-sense stationary random process for which this function is its power spectral density.

☐ If there exists a wide-sense stationary random process for which this function is the power spectral density, then the autocorrelation function of this random process is the inverse discrete-Fourier transform of a set of uniformly spaced samples $P(W_N^n), 0 \le n \le N-1$. (**Comment**: The DFT is a periodic version of the autocorrelation formed by taking a superposition of time-shifted autocorrelation functions.)

- ☐ There exists a complex-valued *minimum-phase* and *monic* function $S(z)$ such that $P(e^{j\omega}) = \left|S(e^{j\omega})\right|^2$. (**Comment**: There is a missing constant, necessary because of the monic assumption.)
- ☺ The geometric mean of $P(e^{j\omega})$ is always less than or equal to the arithmetic mean of $P(e^{j\omega})$.
- ☺ $P(e^{j\omega})$ is periodic in $\omega = 2\pi$.
- ☐ None of these statements is correct.

5. Regarding different methods of accommodating non-stationary signals, which of the following statements are correct? (Check all that apply; more than one may be correct.)

- ☺ When using block processing, the covariance method is generally more accurate than the autocorrelation method, because it is less susceptible to artificial block end effects.
- ☐ When using block processing to find an all-pole model of the signal, the covariance method has an advantage over the autocorrelation method in that the model is always stable. (**Comment**: Other way around; autocorrelation method results in a stable model.)
- ☐ In an adaptive predictor based on the LMS algorithm, a lattice filter structure will always result in a minimum-phase model but the direct-form filter structure may not. (**Comment**: the Burg algorithm will result in a minimum-phase filter, but other adaptation criteria may not.)
- ☺ The recursive-least squares (RLS) method has the advantage that it is based on a clear and objective criterion, whereas the least-squares (LMS) method is not.
- ☺ The RLS method works better in some applications, but the LMS method works better in others.
- ☐ The LS method requires greater processing resources than the RLS method. (**Comment**: Other way around.)
- ☐ None of these statements is correct.

6. Which of the following statements regarding specific applications of adaptive filtering is correct? (Check all that apply; more than one may be correct.)

   ☺ The usual way of adaptively canceling noise added to a desired signal depends on a second input that is correlated with the noise but not the signal.

   ☺ Both differential pulse-code modulation (DPCM) and decision-feedback equalization (DFE) exploit the correlation of samples to reduce the variance of a signal.

   ☐ An adaptive LMS channel-inverse filter for digital communications does not require a training sequence of data symbols, unless there is a decision-feedback filter (because the error rate may initially be very high). (**Comment**: It does require training, because it requires knowledge of the data symbols for adaptation, regardless of whether there is a feedback filter or not.)

   ☺ A DFE may perform very poorly during the tracking mode if the error rate is high due to error-multiplication effects.

   ☐ DPCM requires a training mode to find a good initial approximation to the signal statistics before the actual signal to be coded is applied. (**Comment**: No, it is simply an adaptive predictor.)

   ☺ An adaptive LMS channel-inverse filter for digital communications will converge more rapidly (assuming an appropriate choice of step-size) if the data symbols are uncorrelated than if they are strongly correlated. (**Comment**: Yes, because the power spectrum of the input signal—data symbols plus noise—will be closer to white.)

   ☐ None of the above.

7. Which of the following statements regarding an all-pole lattice filter are correct? (Check all that apply; more than one may be correct.

   ☐ The filter is always causal but not necessarily monic. (**Comment**: It is monic, as the inverse is a prediction-error filter.)

   ☺ The filter is stable (all poles are interior to the unit circle) if and only if the magnitude of all the reflection coefficients are less than unity.

   ☐ The filter is minimum-phase (all poles are interior to the unit circle) if and only if the magnitudes of all the reflection coefficients are less than or equal to unity. (**Comment**: their magnitude must be strictly less than unity).

   ☺ The filter may have poles on the unit circle, but only if one or more reflection coefficients are unit-magnitude.

☐ The filter yields a minimum-phase allpass filter for free (no additional computation). (**Comment**: An allpass filter yes, but an allpass filter is never minimum-phase.)

☐ None of the above.

8. Which of the following statements regarding a bank of bandpass filters (such as derived in class, consisting of a demodulator, prototype lowpass filter which is the same design for each channel, and modulator) are correct? (Check all that apply; more than one may be correct.)

   ☺ When N (the number of filters) is larger, this results in improved frequency resolution and lower time resolution.

   ☐ The prototype lowpass filter must be bandlimited to less than $2\pi / N$ in radian frequency in order for recovery of the input signal to be possible. (**Comment**: A counter example would be $h_k = \delta(k)$.)

   ☐ The input signal can always be recovered by simply summing the filter outputs as long as the highest frequency passed by the prototype lowpass filter (in radian frequency) is greater than π / N. (**Comment**: This is neither a necessary nor sufficient condition.)

   ☐ If the highest frequency passed by the prototype lowpass filter (in radian frequency) is less than $2\pi / N$, the output of each lowpass filter can safely be decimated by a factor of N. (**Comment**: Decimation by $N/2$ would be safe.)

   ☺ Even if the input signal is real-valued, the signals at the bandpass filter outputs will generally be complex-valued.

   ☐ None of the above.

9. Which of the following statements about the eigenvalues and eigenvectors of a square Hermitian matrix is correct? (Check all that apply; more than one may be correct.)

   ☺ The eigenvalues are always real-valued.

   ☐ If one or more eigenvalues is complex-valued, the corresponding eigenvector must be zero.

   ☐ The eigenvalues are always real-valued and non-negative. (**Comment**: They could be negative.)

   ☺ If the eigenvalues are all non-zero, the matrix is non-singular.

☺ If the eigenvalues are all real-valued and positive, the matrix is non-singular and positive-definite.

☐ If one eigenvalue is zero, the corresponding eigenvector must also be zero. (**Comment**: *Any* vector can be a corresponding eigenvector.)

☺ There always exists an orthonormal set of eigenvectors.

☐ None of the above.

10. Which of the following statements regarding time-dependent Fourier transform representations of signals is correct? (Check all that apply; more than one may be correct.)

☺ Time resolution is a property of a signal analysis technique that characterizes how quickly changes in signal statistics can be accurately represented. Good time resolution means we can track more rapid changes in statistics.

☺ Frequency resolution is a property of a signal analysis technique that characterizes how well sinusoidal components at different frequencies can be resolved. Good frequency resolution means we can resolve frequency components with smaller frequency offsets.

☐ Estimating the frequency of a sinusoid more accurately requires better time resolution. (**Comment**: Worse time resolution, or more time to average the estimate.)

☐ Improving the time resolution has the benefit that frequency resolution is better as well. (**Comment**: Lower frequency resolution is the inevitable result.)

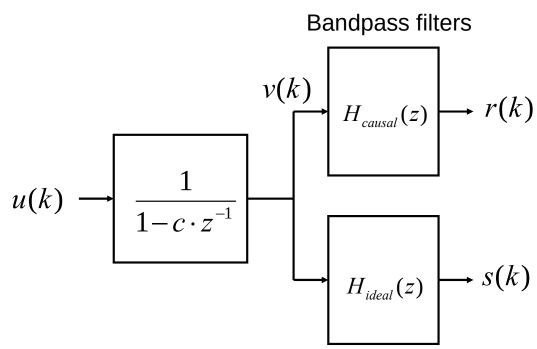
☺ Estimating the power spectrum more accurately (in terms of measures like variance of the estimate) requires worse time resolution or worse frequency resolution or both.

☐ None of the above.

Part II

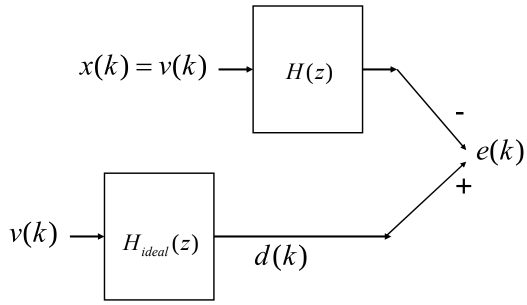
Problem 1. As shown in the figure below, in the course of designing a filterbank for time-frequency analysis, for implementation purposes we want a causal filter design $H_{causal}(z) \leftrightarrow h_c(k)$, which is not necessarily a rational transfer function. We design $H_{causal}(z)$ to be a good approximation to an ideal bandpass filter $H_{ideal}(z) \leftrightarrow h_i(k)$ (which is non-causal). The criterion for design of $H_{causal}(z)$ is to match a typical input signal $v(k)$, as shown in the figure below, where

- $u(k)$ is WSS white noise with mean zero and variance σ^2
- $|c|<1$
- $H_{causal}(z)$ is chosen to minimize $E|r(k)-s(k)|^2$



a. Find $H_{causal}(z)$ in terms of $H_{ideal}(z)$, σ^2, and c.

Solution: We can recast this as a Wiener filtering problem as shown below:



We identify:

$$P_x(z)=\sigma^2 Q(z)Q^*\left(\frac{1}{z^*}\right) \text{ where } Q(z)=\frac{1}{1-cz^{-1}}$$
$$P_{dx}(z)=P_x(z)H_{ideal}(z)$$

The Wiener solution is:

$$H_{causal}(z)=\frac{1}{\sigma Q(z)}\left[\frac{P(z)H_{ideal}(z)}{\sigma Q^*(1/z^*)}\right]_+ = \frac{1}{Q(z)}\left[Q(z)H_{ideal}(z)\right]_+$$

$$H_{causal}(z)=(1-cz^{-1})\cdot\left[\frac{H_{ideal}(z)}{1-cz^{-1}}\right]_+$$

b. Find $h_c(k)$ in terms of $h_i(k)$, σ^2, and c.

Defining $\frac{H_{ideal}(z)}{1-cz^{-1}} \leftrightarrow g(k)$ and $\left[\frac{H_{ideal}(z)}{1-cz^{-1}}\right]_+ \leftrightarrow g_+(k)$ we note that

$$h_c(k)=g_+(k)-cg_+(k-1)=\begin{cases} 0, & k<0 \\ g_+(0), & k=0 \\ g_+(k)-cg_+(k-1), & k>0 \end{cases}$$

$$\text{or } h_c(k)=\begin{cases} 0, & k<0 \\ g(0), & k=0 \\ g(k)-cg(k-1), & k>0 \end{cases}$$

$$\text{or } h_c(k) = \begin{cases} 0, & k < 0 \\ g(0), & k = 0 \\ h_i(k), & k > 0 \end{cases}$$

Finally, it is left to determine $g(0)$,

$$g(k) = \sum_{m=0}^{\infty} c^m h_i(k-m) \text{ or } g(0) = \sum_{m=0}^{\infty} c^m h_i(-m) .$$

Problem 2. Assume a received signal is the sum of an amplitude- and phase-modulated complex exponential plus noise,

$$d(k) = r(k) \cdot \exp\{j(\omega_0 k + \phi(k))\} + u(k)$$

and $u(k)$ is complex-valued zero-mean white noise with variance σ^2.

a. Assuming that the amplitude $r(k)$ and phase $\phi(k)$ vary sufficiently slowly and that the frequency ω_0 is known exactly, find an LMS adaptive filter that forms an estimate of the time-varying amplitude and also the time-varying phase.

Define the error as $e(k) = d(k) - a \cdot \exp\{j\omega_0 k\}$ with a complex-valued coefficient a. Then

$$\frac{\partial}{\partial a^*} |e(k)|^2 = -e(k) \cdot \exp\{-j\omega_0 k\}$$

and thus an LMS algorithm that attempts to minimize this error is

$$a(k+1) = a(k) + \mu \cdot e(k) \cdot \exp\{-j\omega_0 k\} = (1-\mu) \cdot a(k) + \mu \cdot d(k) \cdot \exp\{-j\omega_0 k\} .$$

The amplitude and phase estimates are given by

$$a(k) = r(k) \cdot \exp\{j\phi(k)\} .$$

b. Assuming that the amplitude and phase are constant, characterize the convergence of the mean value of the amplitude/phase estimate. Is it asymptotically unbiased, and if so under what conditions? How does the convergence depend on the step size? State any assumptions you have to make.

Assuming constant amplitude and phase,

$$a(k+1) = (1-\mu) \cdot a(k) + \mu \cdot r \cdot \exp\{j\phi\} + \mu \cdot v(k)$$

where $v(k) = u(k) \cdot \exp\{-j\omega_0 k\}$ is also white noise. Defining a tap-misalignment signal $c(k) = a(k) - r \cdot \exp\{j\phi\}$, this becomes

$$c(k+1) = (1-\mu) \cdot c(k) + \mu \cdot v(k) \, .$$

Then we get as the mean value of the misalignment

$$E[c(k+1)] = (1-\mu) \cdot E[c(k)] + \mu \cdot E[v(k)] \Rightarrow E[c(k)] = (1-\mu)^k \cdot c(0) \, .$$

No assumptions were necessary. Yes, it is an asymptotically unbiased estimate if $0 < \mu < 2$.

c. After convergence of the mean-value of the estimate, what is the asymptotic variance of the estimate? How does it depend on the step size? State any assumptions you have to make.

Since know the mean of $c(k)$ converges to zero, the variance is asymptotically equal to the second moment,

$$E\,|\,c(k+1)\,|^2 = (1-\mu)^2 \cdot E\,|\,c(k)\,|^2 + \mu^2 \cdot E\,|\,v(k)\,|^2 \rightarrow \frac{\mu \cdot \sigma^2}{2-\mu}$$

To set the cross-term equal to zero, we had to assume that $c(k)$ and $v(k)$ are independent. Thus, we can make the variance of the estimate as small as we like by choosing a small step size, but of course at the expense of speed of convergence.

Problem 3. An observation $\mathbf{d}$ is given by

$$\mathbf{d} = \mathbf{s} + \mathbf{u}$$

where $\mathbf{d}$, $\mathbf{s}$, and $\mathbf{u}$ are all random N-dimensional vectors. Assume that $\mathbf{s}$ and $\mathbf{u}$ are independent, with first and second-order statistics

$$E[\mathbf{s}] = \bar{\mathbf{s}} \neq \mathbf{0} \text{ and } E[\mathbf{u}] = \mathbf{0}$$

$$\mathbf{E[ss^H]} = \mathbf{R_s} \text{ and } \mathbf{E[uu^H]} = \sigma^2 \cdot \mathbf{I} \, .$$

a. We desire to form $\hat{\mathbf{s}}$, a linear estimate of $\mathbf{s}$ based on $\mathbf{d}$, of the form $\hat{\mathbf{s}} = \mathbf{Ad}$. Prove that, if we use a minimum mean-square error (MMSE) criterion, there always exists an $\bar{\mathbf{s}}$ such that the estimate is biased.

The error vector is

$$\mathbf{e} = \mathbf{s} - \hat{\mathbf{s}} = \mathbf{s} - \mathbf{A(s+u)} = \mathbf{(I-A)s} - \mathbf{Au}$$

Then

$$E[\mathbf{e}] = \mathbf{(I-A)\bar{s}} \text{ and } E[\mathbf{ee^H}] = \mathbf{(I-A)R_s(I-A)^H} + \sigma^2 \cdot \mathbf{AA^H}$$

The key observation is that $E\|\mathbf{e}\|^2 = Tr\{E[\mathbf{ee^H}]\}$.

Unless $\mathbf{A} = \mathbf{I}$, $\mathbf{(A-I)}$ must be at least rank one, and there exists an $\mathbf{\bar{s}}$ such that $E[\mathbf{e}] \neq 0$. So we must show that $\mathbf{A} = \mathbf{I}$ is *not* the MMSE estimator. To show that there exists an estimator with a smaller error, choose $\mathbf{A} = \alpha \cdot \mathbf{I}$, in which case

$$E[\mathbf{ee^H}] = (1-\alpha)^2 \cdot \mathbf{R_s} + \alpha^2\sigma^2 \cdot \mathbf{I} \text{ and } E\|\mathbf{e}\|^2 = (1-\alpha)^2 \cdot Tr\{\mathbf{R_s}\} + \alpha^2\sigma^2 N$$

The minimum of $E\|\mathbf{e}\|^2$ does not occur at $\alpha = 1$ unless $\sigma = 0$.

b. Assuming that we further *constrain* the estimate in a. to be unbiased for all $\mathbf{\bar{s}}$, find the optimum MMSE unbiased estimator $\mathbf{A}$ and the resulting MSE.

As shown before we must have $\mathbf{A} = \mathbf{I}$ ($\alpha = 1$) and $E\|\mathbf{e}\|^2 = \sigma^2 N$.

c. We decide to simultaneously estimate $s(k)$ and estimate its frequency-domain content by modeling its components as a sum of complex exponentials

$$\hat{s}(k) = \sum_{n=1}^{N} a_n \cdot W_N^{nk} \text{ for } W_N = e^{j2\pi/N},\ 1 \le k \le N,$$

We assume a linear estimate of the complex exponential amplitudes $\mathbf{a}$ of the form $\mathbf{\hat{a}} = \mathbf{Bd}$. Repeat a. by showing that, here again, the MMSE linear estimator is biased.

Define the unitary matrix

$$\mathbf{W} = \begin{bmatrix} W_N & W_N^2 & \cdots & W_N^N \\ W_N^2 & W_N^4 & \cdots & W_N^{2N} \\ & & & \\ W_N^N & W_N^{2N} & & W_N^{N^2} \end{bmatrix} \text{ where } \mathbf{W^{-1}} = \mathbf{W^H}$$

Then

$$\mathbf{\hat{s}} = \mathbf{Wa}$$

$$\mathbf{e} = \mathbf{s} - \mathbf{\hat{s}} = \mathbf{s} - \mathbf{Wa} = \mathbf{s} - \mathbf{WB(s+u)} = \mathbf{(I-WB)s} - \mathbf{WBu}$$

This is the same as before, where we associate $\mathbf{A} = \mathbf{WB}$, and thus, the same conclusion as before holds.

d. Repeat b. by finding the MMSE *unbiased* estimator $\mathbf{B}$ and the resulting MSE.

To get an unbiased estimator, let $\mathbf{WB} = \mathbf{I}$ or $\mathbf{B} = \mathbf{W^H}$. Then $\mathbf{e} = -\mathbf{WBu}$ and

$$E[\mathbf{ee^H}] = \sigma^2 \cdot \mathbf{WBB^H W^H} = \sigma^2 \cdot \mathbf{I} \text{ and } E\|\mathbf{e}\|^2 = \sigma^2 N.$$