Homework #5

1. **Vector Gaussian noise channel.** The output of a vector digital communication channel with Additive Gaussian Noise is given by:

\[ Y = S + Z, \]

where the signal vector is

\[ S = \begin{cases} \begin{bmatrix} 2 & 1 & \frac{1}{4} & \cdots & \frac{1}{2^{n-2}} \end{bmatrix}^T, & \text{with probability } \frac{1}{2}, \\ -[2, 1, \frac{1}{4}, \cdots, \frac{1}{2^{n-2}}]^T, & \text{with probability } \frac{1}{2}, \end{cases} \]

and the noise vector \( Z = [Z_1, Z_2, \ldots, Z_n]^T \) is

\[ Z_1 = W_1, \quad \text{and} \quad Z_i = \frac{1}{2}Z_{i-1} + W_i, \quad 2 \leq i \leq n, \]

where \( W \sim \mathcal{N}(0, I) \) is independent of \( S \). Determine the optimal estimate \( \hat{\Theta}(Y) \) for deciding which of the two signals was sent and the corresponding probability of error. (We are expecting numerical answers).

(Hint: By inspection, you can express \( Z = AW \), where \( W \sim \mathcal{N}(0, I) \), and the inverse of \( A \) can also be found by inspection).

2. **Prediction.** Let \(|\alpha| < 1\) and let \( X = [X_1, X_2, \ldots, X_n]^T \) be a zero-mean random vector with covariance matrix

\[ \Sigma_X = \begin{bmatrix} 1 & \alpha & \alpha^2 & \cdots & \alpha^{n-2} & \alpha^{n-1} \\ \alpha & 1 & \alpha & \cdots & \alpha^{n-3} & \alpha^{n-2} \\ \alpha^2 & \alpha & 1 & \cdots & \alpha^{n-4} & \alpha^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \alpha^{n-2} & \alpha^{n-3} & \alpha^{n-4} & \cdots & 1 & \alpha \\ \alpha^{n-1} & \alpha^{n-2} & \alpha^{n-3} & \cdots & \alpha & 1 \end{bmatrix}. \]

Find the best linear MSE estimate of \( X_n \) given \( X_1, X_2, \ldots, X_{n-1} \) and its MSE.

3. **Noise cancellation.** A classic problem in statistical signal processing involves estimating a weak signal (e.g., the heart beat of a fetus) in the presence of a strong interference (the heart beat of its mother) by making two observations—one with the weak signal present and one without (by placing one microphone on the mother’s belly and another close to her heart). The observations can then be combined to estimate the weak signal by “canceling out” the interference. The following is a simple version of this application.

Let the weak signal \( X \) be a random variable with mean \( \mu \) and variance \( P \). Let the observations be \( Y_1 = X + Z_1 \) and \( Y_2 = Z_1 + Z_2 \), where \( Z_1 \) is the strong interference and \( Z_2 \) is measurement noise. Assume that \( Z_1 \) and \( Z_2 \) are zero mean with variances \( N_1 \) and \( N_2 \), respectively. Further assume that \( X, Z_1, \) and \( Z_2 \) are uncorrelated. Find the best linear MSE estimate of \( X \) given \( Y_1 \) and \( Y_2 \) and the corresponding MSE. Interpret the results.
4. **Reconstruction on the tree.** In this problem, you will explore reconstruction on the tree (see lecture notes pages 4-3 to 4-7). Our goal is to reconstruct the value of the root random variable given the observations at the leaf nodes of the tree.

As the depth of the tree is varied, two opposing effects are at work: On the one hand, a larger-depth tree has a larger number of leaf nodes, and the detection of the root node is thus based on more observed data. On the other hand, with growing tree depth, each layer introduces additional noise to the observations, and thus the random variable at each leaf node becomes less correlated with the root random variable. Thus a larger tree depth corresponds to a larger amount of less accurate data.

Download the posted Matlab file `treeReconstruction.m`. It contains skeleton code to generate tree realizations and evaluate error probabilities, but it is missing the actual code for the detection rules.

a. Write the function `MAP` that computes the likelihoods $L_{1,0}$ and $L_{1,1}$ of the root random variable from the likelihoods $L_{i,0}$ and $L_{i,1}$ of the leaf nodes $i$. Hand in your code.

   *Hint:* In order to avoid numerical instabilities in the MAP iterations, you can scale the likelihoods $L_{i,0}$ and $L_{i,1}$ on each layer such that their maximum across the layer is always 1.

b. An alternative detection rule is **majority voting**, where we declare the root random variable to be 1 if the majority of leaf random variables is 1. Write the function `MajorityVote` to implement this rule and hand in your code.

c. Plot the empirical probability of error as a function of $\epsilon$ for both MAP and majority voting. Generate trees with 4, 8, and 12 layers, and vary $\epsilon$ in 15 steps in the interval $[0, 0.5]$. For plotting the empirical probability of error, use as many realizations as you reasonably can (at least 10000 per value of $\epsilon$).

d. How does the MAP rule and majority voting compare in terms of computational complexity and probability of detection error?

e. Does increasing the tree size improve or worsen the error probability under the MAP rule?

5. **Gaussian MMSE estimator is linear.** It can be shown that if $X$ and $Y$ are jointly Gaussian, then $X \mid \{Y = y\}$ is Gaussian and the conditional mean $E(X \mid Y = y)$ is linear in $y$. Thus the best MSE estimate of $X$ given $Y$ is linear. In this problem you will prove this fact using orthogonality and properties of jointly Gaussian random variables discussed in the lecture notes. Assume that $X$ and $Y$ are jointly Gaussian. Let $\hat{X}_1$ be the best MSE estimate of $X$ given $Y$, and let $\hat{X}_2$ be the best MSE linear estimate of $X$ given $Y$. Justify each of the following steps of the proof.

   a. $E((X - \hat{X}_2)(Y - E(Y))) = 0$

   b. $(X - \hat{X}_2)$ and $(Y - E(Y))$ are jointly Gaussian

   c. Parts (a) and (b) imply that $E(X - \hat{X}_2 \mid Y) = 0$

   d. Part (c) implies that $\hat{X}_2 = E(X \mid Y) = \hat{X}_1$

6. **Transformation of Gaussians (II).** Let

$$\mathbf{X} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\right).$$

Find $g_1(\mathbf{X})$ and $g_2(\mathbf{X})$ such that

$$\mathbf{Y} = \begin{bmatrix} g_1(\mathbf{X}) \\ g_2(\mathbf{X}) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\right).$$
7. Additive nonwhite Gaussian noise channel. Let \( Y_i = X + Z_i \) for \( i = 1, 2, \ldots, n \) be \( n \) observations of a signal \( X \sim \mathcal{N}(0, P) \). The additive noise random variables \( Z_1, Z_2, \ldots, Z_n \) are zero mean jointly Gaussian random variables that are independent of \( X \) and have correlation \( \mathbb{E}(Z_i Z_j) = N \cdot 2^{-|i-j|} \) for \( 1 \leq i, j \leq n \). Find the best MSE estimate of \( X \) given \( Y_1, Y_2, \ldots, Y_n \), and its MSE. Hint: the coefficients for the best estimate are of the form \( h^T = [a \ b \ b \ \cdots \ b \ b \ a] \).