EE 121: Introduction to Digital Communication Systems

Midterm Solutions

1. Consider the following discrete-time communication system. There are two equallly likely messages to be transmitted, and they are encoded into an input process $\{X_n\}$ such that if message 0 is transmitted, $X_n = -1$ for all $n \geq 0$ and if message 1 is transmitted, $X_n = 1$ for all $n \geq 0$. The process is then passed through an AWGN channel such that the output process $\{Y_n\}$ is given by

$$Y_n = X_n + W_n$$

where $\{W_n\}$ is a sequence of i.i.d. $N(0, \sigma^2)$ random variables. The received sequence then goes through a threshold receiver whose output is the process $\{Z_n\}$, such that $Z_n = \mathrm{sgn}(Y_n)$.

[4] a) Is the process $\{X_n\}$ stationary? Explain.

[4] b) Is the process $\{Z_n\}$ stationary? Explain.

[4] c) Compute $P(X_k = 1|Z_k = 1)$. (Express your answer in terms of the Q-function.)

[4] d) Compute $P(Z_{k+1} = 1|Z_k = 1)$. Are the Z_n's independent?

Solution:

a) Yes. For all $k, n_1, n_2, \ldots, n_k$,

$$P(X_{n_1} = 1, \ldots, X_{n_k} = 1) = \frac{1}{2}, \qquad P(X_{n_1} = -1, \ldots, X_{n_k} = -1) = \frac{1}{2}.$$

Thus, the joint distribution does not depend on the times $n_1, \ldots, n_k$ and the process is (strict-sense) stationary.

Some students argue that the process is not stationary because for negative times, the process is not defined. I accept that although it is taken for granted that by stationarity here, one means only over positive times.

Some students only verify wide-sense stationarity. I accept that too.

b) Yes. $\{Z_n\}$ is a time-invariant transformation of two stationary processes $\{X_n\}$ and $\{W_n\}$ and hence is stationary. (This is similar to one of the homework questions.)

c)

$$P(X_k = 1|Z_k = 1) = \frac{P(Z_k = 1|X_k = 1)P(X_k = 1)}{P(Z_k = 1)}$$

By symmetry, $P(Z_k = 1) = \frac{1}{2}$, $P(X_k = 1) = \frac{1}{2}$. Also,

$$P(Z_k = 1|X_k = 1) = P(W_k > -1) = Q(\frac{-1}{\sigma})$$

and hence

$$P(X_k = 1|Z_k = 1) = Q(\frac{-1}{\sigma}).$$

d)

$$
\begin{aligned}
& P(Z_{k+1}=1|Z_k=1) \\
= \quad & P(Z_{k+1}=1|Z_k=1,X_k=1)P(X_k=1|Z_k=1)+P(Z_{k+1}=1|Z_k=1,X_k=-1)P(X_k=-1|Z_k=1) \\
= \quad & P(Z_{k+1}=1|X_k=1)P(X_k=1|Z_k=1)+P(Z_{k+1}=1|X_k=-1)P(X_k=-1|Z_k=1) \\
= \quad & P(Z_{k+1}=1|X_{k+1}=1)P(X_k=1|Z_k=1)+P(Z_{k+1}=1|X_{k+1}=-1)P(X_k=-1|Z_k=1) \\
= \quad & Q(\frac{-1}{\sigma}Q(\frac{-1}{\sigma})+Q(\frac{1}{\sigma}(1-Q(\frac{-1}{\sigma})) \\
= \quad & Q^2(\frac{-1}{\sigma})+Q^2(\frac{1}{\sigma})
\end{aligned}
$$

2. Suppose a discrete-time source is modelled by a stationary zero-mean Gaussian process $\{X_n\}$ with autocorrelation function R_X.

[5] a) Is the variance of $X_n - X_{n-1}$ always less than that of X_n? Explain. If not, when does it hold? Does DPCM coding where the difference between two consecutive samples are quantized always improve the performance over PCM? Explain.

[5] b) Find the optimal linear predictor of X_n given X_{n-1}, and calculate the variance of the resulting error. How do you compare the performance of a DPCM coder using this predictor with that in (a)?

Solution:

a)

$$Var(X_n - X_{n-1}) = Var(X_n) + Var(X_{n-1}) - 2Cov(X_n X_{n-1}) = 2R_X(0) - 2R_X(1)$$

Thus, the variance of $X_n - X_{n-1}$ is smaller than that of X_n if and only if

$$R_X(0) < 2R_X(1)$$

If this is not satisfied, the variance of the difference is actually greater than the variance of the individual sample and DPCM will in fact do worse.

b) By solving the Yule-Walker equation, one finds that

$$\hat{X}_n = \frac{R_X(1)}{R_X(0)} X_{n-1}$$

and the variance of the resulting error is

$$E[(X_n - \frac{R_X(1)}{R_X(0)} X_{n-1})^2] = R_X(0) - \frac{R_X(1)^2}{R_X(0)}.$$

Since this is always less than the variance of X_n, it always improves over PCM.

3. Consider the following modulation scheme for symbols of n bits long. The first k bits picks the amplitude A_i and the last $n-k$ bits picks the phase ϕ_j. The modulated waveform is

$$A_i cos(\frac{2\pi t}{T} + \phi_j)$$

on $[0, T]$.

[6] (a)Find an orthonormal basis for the waveforms.

[2] (b) What is the dimension of the space spanned by the waveforms?

[4] (c) What is the geometry of the signal constellation?

Solution

(a)

$$A_i cos(\frac{2\pi t}{T} + \phi_j) = A_i cos(\phi_j) cos(\frac{2\pi t}{T}) + A_i sin(\phi_j) sin(\frac{2\pi t}{T}).$$

Therefore, an orthonormal basis for the waveforms is given by

$$\{\sqrt{\frac{2}{T}} sin(\frac{2\pi t}{T}), \sqrt{\frac{2}{T}} cos(\frac{2\pi t}{T})\}$$

.

b) The dimension is 2.

c) Each signal can be represented as a point in 2D plane with (A_i, ϕ_j) as the polar coordinates.

4. Consider a source with an alphabet of N letters, probabilities $p_1, \ldots, p_N$, all non-zero.

[6] a) Argue that any prefix-conditioned binary code with codeword lengths $l_1, \ldots, l_N$ must satisfy:

$$\sum_{i=1}^{N} 2^{-l_i} \leq 1$$

(Hint: think in terms of the binary tree representation of the code.)

[6] b) Argue that equality must hold if the code minimizes the expected codeword lengths.

[6] c) (Bonus) Show that the expected codeword length of any prefix-conditioned binary code cannot be smaller than

$$-\sum_{i=1}^{N} p_i \log_2 p_i$$

(Hint: considering the constraint in (b) on the codeword lengths, how small can the expected codeword length be even if the lengths could be fractional?)

Solution:

a) A rigorous derivation is by induction. All prefix-condition code correponds to a binary tree where all codewords are at the leaves. If there are only two codewords, the assertion is certainly true since $l_1 = l_2 = 1$. Assume it is true for all prefix-conditioned code with n or less codewords. Consider now a code with $n+1$ codewords. Let L and R be the codeword leaves in the left and right subtrees respectively. A codeword of length l_i can be thought of as a codeword of length $l_i - 1$ in one of the subtrees. Applying our induction hypothesis on the left and right subtrees separately,

$$\sum_{i \in L} 2^{-l_i - 1} \leq 1, \qquad \sum_{i \in R} 2^{-l_i - 1} \leq 1.$$

This implies that

$$\sum_{i} 2^{-l_i} \leq 1.$$

Any informal argument which is along these lines is given full credit.

b) When the average codeword length is minimized, all interior nodes have two siblings. Using a similar argument as above, one can show that equality must hold.

c) Consider the problem of minimizing $\sum_i p_i l_i$ subject to the constraint of $\sum_i 2^{-l_i} = 1$, without requiring that l_i's are integer. By using the method of Lagrange multipler, one can show that the optimal solution is given by $l_i^* = -log_2 p_i$. Hence the average codeword length for any code must be lower bounded by $-\sum_i p_i \log_2 p_i$.