1. **(Optimal linear prediction)** Given the random variables \(X_{n-m}, X_{n-m+1}, \ldots, X_{n-1}, X_n\) having finite variance, for any vector \(a = (a_1, \ldots, a_m)^T \in \mathbb{R}^m\) of (linear) prediction coefficients define

\[
g(a) = E\left[ \left( X_n - \sum_{i=1}^{m} a_i X_{n-i} \right)^2 \right].
\]

Let \(R\) be the autocorrelation matrix \(R = \{r_{jk}\} = \{E(X_{n-j}X_{n-k})\}_{1 \leq j,k \leq m}\) and \(v = (r_{01}, \ldots, r_{0m})^T\). Show that if \(\hat{a} = (\hat{a}_1, \ldots, \hat{a}_m)^T\) satisfies \(R\hat{a} = v\) then \(\hat{a}\) is an optimal prediction coefficient vector, i.e.,

\[
g(\hat{a}) = \min_{a \in \mathbb{R}^m} g(a).
\]

**Hint:** Show that \(g(a) - g(\hat{a}) = (a - \hat{a})^T R (a - \hat{a})\).

2. **(Predictive quantization)**

(a) Use the orthogonality principle to find the optimal linear predictor coefficients for a 2nd order predictor for the process \(X_n = W_n + W_{n-1} + W_{n-2}\) where the process \(\{W_n\}\) is zero mean unit variance and white (uncorrelated, i.e., \(E(W_iW_j) = 0\) if \(i \neq j\)).

(b) The predictor coefficients obtained in part (a) are used in a DPCM coder for encoding \(\{X_n\}\). It is assumed that \(\{W_n\}\) is a Gaussian process (in addition to being unit variance and uncorrelated). An 8-bit quantizer is used and the quantizer design assumes that the closed-loop prediction error has approximately the same distribution as the open-loop prediction error (i.e., Gaussian). Using the high-rate approximation, calculate the overall SNR of the DPCM system. Explain your reasoning.

**Hint:** As part of this problem, you’ll have to show that \(\frac{1}{12} \|\phi\|_{1/3} = \frac{\sqrt{3\pi}}{2}\) for the standard normal pdf \(\phi\).

(c) What is the approximate gain (in SNR) of the system in part (b) compared with a system that directly encodes each sample \(X_n\) using an optimal 8-bit quantizer?

3. **(Transform coding)** Let \(X_n = W_n + W_{n-1}\), where the sequence \(\{W_n\}\) is zero mean, unit variance, and white (uncorrelated).

(a) Find the Karhunen-Loeve transform for \(X = (X_1, X_2, X_3)^T\).

(b) Assume \(\{W_n\}\) is a Gaussian process. Method I is a transform coder for \(X\) using the Karhunen-Loeve transform, optimal quantization of the transform coefficients, and optimal bit allocation with a total of \(B\) bits. Method II separately quantizes the components of \(X\) with optimal scalar quantizers having \(b = B/3\) bits. Using high-resolution approximations, calculate the performance gain (in decibels) of Method I over Method II.
4. *For MATH 877 students:* The discrete Fourier transform (DFT) of the \( k \)-vector \( \mathbf{x} = (x_0, \ldots, x_{k-1})' \) is defined as

\[
\mathbf{y} = \mathbf{W} \mathbf{x}
\]

where \( \mathbf{W} \) is the \( k \times k \) matrix with (complex) entries

\[
w_{mn} = \frac{1}{\sqrt{k}} e^{i2\pi \frac{mn}{k}}; \quad m, n = 0, \ldots, k - 1.
\]

Show that \( \mathbf{W} \) is a *unitary* matrix, i.e., \( \mathbf{WW}^* = \mathbf{I} \), where \( \mathbf{W}^* \) is the conjugate transpose of \( \mathbf{W} \).