EE 372 | 4 April 2007
Quantization | Spring 2006-2007

Homework #1

Due: 18 April

1. Let $\{X_n\}$ be an iid sequence of random variables with a uniform probability density function on $[0,1]$. Suppose we define an N-level uniform quantizer $q(r)$ on [0,1] as follows: $q(r) = (k+1/2)/N$ if $k/N \leq r < (k+1)/N$. In other words, we divide [0,1] up into N intervals of equal size. If the input falls in a particular bin, the output is the midpoint of the bin. Application of the quantizer to the process $\{X_n\}$ yields the process $\{q(X_n)\}$. Define the quantizer error process $\{\epsilon_n\}$ by $\epsilon_n = q(X_n) - X_n$.

   (a) Find $E(\epsilon_n)$, $E(\epsilon_n^2)$, and $E(\epsilon_n q(X_n))$.

   (b) Find the (marginal) cumulative distribution function and the pdf for ϵ_n.

   (c) Find the covariance and power spectral density of ϵ_n.

   (d) Find the correlations $E(\epsilon_n X_k)$ and $E(\epsilon_n q(X_k))$.

2. *Transform coding* A k-dimensional real random vector X is coded as follows: First it is multiplied by a unitary matrix U to form a new vector $Y = UX$. (By unitary it is meant that $U^* = U^{-1}$, where U^* is the conjugate transpose of U). Each component Y_i, $i = 0, 1, \cdots, k-1$ is separately quantized by a quantizer q_i to form a reproduction $\hat{Y}_i = q_i(Y_i)$. Let $\hat{Y}$ denote the resulting vector. This vector is then used to produce a reproduction $\hat{X}$ of X by the formula $\hat{X} = U^{-1}\hat{Y}$. This code is called a *transform code*.

   (a) Suppose that we measure the distortion between input and output by the mean squared error
   $$(X - \hat{X})^*(X - \hat{X}) = \sum_{i=0}^{k-1} |X_i - \hat{X}_i|^2,$$
   where the asterisk denotes complex conjugate transpose. Show that the distortion is the same in the original and transform domain, that is,
   $$(X - \hat{X})^*(X - \hat{X}) = (Y - \hat{Y})^*(Y - \hat{Y}).$$

   (b) Find a matrix U such that Y has uncorrelated components, that is, its covariance matrix is diagonal.

3. Feedback/predictive quantizationConsider the coding scheme of Figure 1 in which the encoder is a simple quantizer inside a feedback loop.

   H and G are causal linear filters. Show that if the filters are chosen so that
   $$\hat{X}_n = \tilde{X}_n + q(\epsilon_n),$$
   then
   $$|X_n - \hat{X}_n| = |\epsilon_n - q(\epsilon_n)|$$

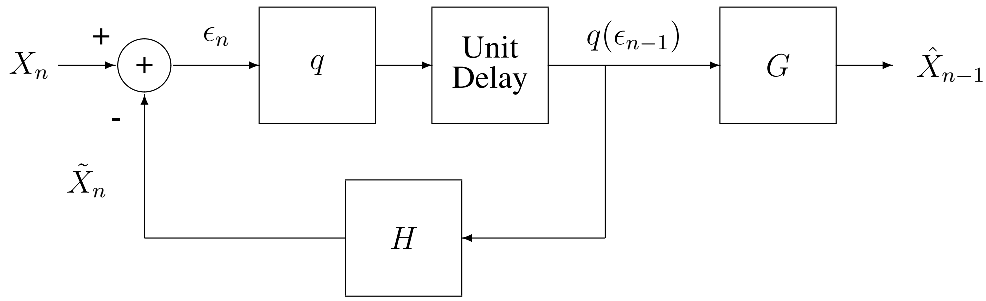


Figure 1: Coding Scheme

and hence

$$E(|X_n - \hat{X}_n|^2) = E(|\epsilon_n - q(\epsilon_n)|^2),$$

that is, the overall error is the same as the quantizer error. Suppose that G is fixed. What must H be in order to satisfy the given condition? What constraint is there on G so that H will be causal? (For the purpose of this problem, the "quantizer" can be any nonlinear memoryless mapping.)

4. Suppose that X is a random variable described by an exponential pdf

$$f_X(\alpha) = \lambda e^{-\lambda\alpha};\ \alpha \geq 0.$$

Define a function q which maps real numbers into integers by $q(x) =$ the largest integer less than or equal to x. In other words

$$q(x) = k \text{ if } k \leq x < k+1.$$

(This function is denoted by $q(x) = \lfloor x \rfloor$.) The function q is a form of quantizer, it rounds all real numbers downward to the nearest integer below the input real number. Define the following two random variables: the quantizer output

$$Y = q(X)$$

and the quantizer error

$$\epsilon = X - q(X).$$

Note: By construction ϵ can only take on values in $[0, 1)$.

(a) Find the pmf $p_Y(k)$ for Y.

(b) Find the expectations $E(X)$ and $E(Y)$.

(c) Derive the probability density function for ϵ.

5. The transform/characteristic function method

Suppose that q is a uniform quantizer with bin width Δ and having the form $q(x) = k\Delta + \Delta/2$ for $x \in [k\Delta, (k+1)\Delta)$ for all integer k. (We allow the quantizer to have an infinite number of levels.)

Define the resulting quantization error

$$\epsilon = \epsilon(x) = q(x) - x.$$

This quantizer will be applied to a stationary random process X_n with 0 mean, characteristic function

$$\Phi_X(\nu) = E(e^{j\nu X_n})$$

and joint characteristic function

$$\Phi_{X_n,X_m}(\nu_n, \nu_m) = E(e^{j(\nu_n X_n + \nu_m X_m)})$$

.

(a) Prove that the quantization error can be expressed as a function of the input as

$$\epsilon(x) = \Delta(\frac{1}{2} - < \frac{x}{\Delta} >), \quad (1)$$

where $< r >$ denotes the fractional part of r or $r \bmod 1$, that is, $< r >= r - \lfloor r \rfloor$ where $\lfloor r \rfloor$ is the largest integer less than or equal to r. Provide a labeled sketch of this quantizer error function.

(b) The error function ϵ is periodic in Δ, so it can be expressed as a Fourier series of the form

$$\epsilon(x) = \sum_i a_i e^{2\pi j i x/\Delta}.$$

Find the a_i.

(c) Use the Fourier series representation to find formulas for the mean $E(\epsilon_n)$ and autocorrelation function $R_\epsilon(n, m) = E(\epsilon_n \epsilon_m)$ of the noise process $\epsilon_n = q(X_n) - X_n$ in terms of the characteristic function and joint characteristic function of X_n. This method is referred to variously as the *transform method* or the *characteristic function method* and is often used in the analysis of nonlinear systems.

(d) Assume that X_n is a Gaussian random process and that $R_X(n, m) = \sigma^2 r^{|n-m|}$ for $|r| < 1$. Simplify the answers to the previous part as much as possible.

6. Do a Web search for "centroidal Voronoi" and comment (briefly, with a supporting URL) on their applications to or relations to the topics listed below.

- Delauney triangulation
- Mapping population densities
- Art
- Three and higher dimensions
- Climate modeling
- Natural occurrences
- Image compression

- Computational geometry
- Two other distinct subjects

7. Suppose that you have a k-dimensional random vector X with a probability density function f_X that has the following form:

$$f_X(x) = \begin{cases} a & x \in R_0 \\ 2a & x \in R_1 \\ 0 & otherwise \end{cases}$$

where R_0 and R_1 are two k-dimensional spheres having equal radius r and volume V. Assume the spheres have Euclidean centers of gravity (centroids) c_0 and c_1. (To be precise, $R_l = \{x : ||x - c_l|| \leq r\}$, $l = 0, 1$.) Assume that the spheres are separated by more than twice the radius, i.e., $||c_0 - c_1|| > 2r$. The distortion measure is squared error.

   (a) Find a in terms of V. (The remaining final answers should depend only on the known volume and not on a.) What is the optimal zero bit ($N = 1$) vector quantizer?

   (b) Find a one bit vector quantizer ($N = 2$) which satisfies the Lloyd conditions.

   (c) Next assume that N is quite large. Assuming that Gersho's conjecture is true, find the optimal quantizer point density and the resulting average squared error.

   (d) What fraction of the total number of quantizer levels will lie in R_0?

8. Scalar and 2-dimensional vector quantizers for iid Gaussian sources.

   Suppose that $\{X_n\}$ is an iid sequence of Gaussian random variables with mean 0 and variance $\sigma_X^2 = 1$.

   (a) Analytically find a one bit (2 levels) scalar quantizer (dimension $k = 1$) that satisfies the Lloyd conditions and evaluate the resulting average squared error. State the codewords to an accuracy of four decimal places.

   (b) Next generate a set S of 10,000 iid Gaussian sample values using a random number generator (say which one). Use the first 9,000 of these as a training set and use the Lloyd algorithm I to design a one-bit scalar quantizer. Use a convergence threshold of $\epsilon = 0.01$. Test your quantizer on the remaining 1000 samples. Report the test average squared error and the final codewords, along with the number of iterations it took to converge. Also report the SNR (with respect to the optimal 0 rate code). Try this design with the following three initial codebooks and comment on the differences if any. In particular, do any guesses provide faster convergence?

   - $\{-1, +1\}$
   - $\{$ centroid , centroid $+ (\epsilon, \epsilon)^t\}$, where centroid is the Euclidean centroid of your training set, and $\epsilon = 0.01$.
   - $\{$ centroid , worst$\}$, where worst = the vector in the training set farthest from the centroid, that is, the vector with the biggest distortion with respect to the centroid.

How close are the test distortions to the training distortion? Compare the test distortion to the high-rate and Shannon bounds.

(c) Now consider the following variations. Break the original set of 10,000 samples into 10 groups of 1000 consecutive samples, say $T_i = \{X_{1000i}, X_{1000i+1}, \ldots, X_{1000(i+1)-1}\}$. For $i = 0, 1, 2, \ldots 8$, use the 9000 samples formed by removing T_i from S as a training set and then test the result on T_i. Note we have already done this for $i = 9$. Use the same initial codebook (your choice) throughout. Finally, average the average squared error achieved for all 10 of the designs, i.e., for all 10 possible training sets and test sequences. Compare this number with the theoretical result of the first part. This technique to estimating the performance of a design based on a training set is called *cross validation* in statistics.

(d) Hopefully all of your work to this point suggests that there is essentially only one scalar quantizer for this source satisfying the Lloyd conditions. You found it exactly in the first part and approximately in the remaining parts. If this is true, you have found the globally optimal solution. See if you can prove it is true or suggest how you might convincingly demonstrate it by simulation.

(e) Next consider two-dimensional VQs for the same source. Again consider codes with a rate of one bit per sample, so now the 2D codebook has 4 codewords. Describe at least three genuinely distinct codebooks that satisfy the fixed-rate Lloyd conditions. You are asked for genuinely distinct codebooks because a little thought should show that in this example any rotation of a Lloyd codebook around the origin yields another Lloyd codebook. State and sketch the codewords and the average distortion.

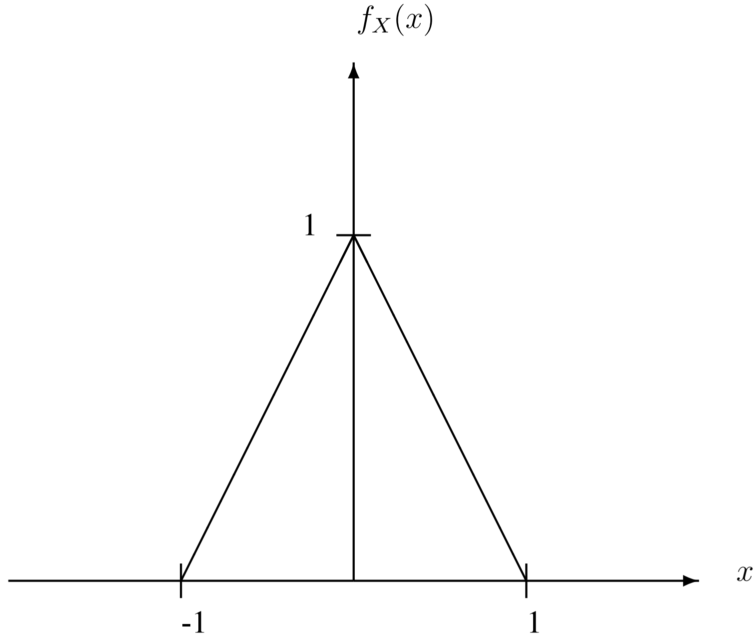
Hint: A little thought should assist you to find two of these codebooks analytically before you run Lloyd, and you should be able to guess the structure of the third. Lloyd will likely be needed, however, to find the distortion and exact codeword location of the third code. Run the Lloyd algorithm using the data already generated, using the first 4500 pairs to design a 2D VQ and test it on the remaining 500 pairs.

This simple example makes the point that locally optimal codebooks exist and can be quite poor. No other locally optimal codebooks for this case have been found, but no one has actually proved that they do not exist.

(f) How will the results of this problem change if you are told that the variance of X_n is 4 instead of 1?

Keep in mind you can obtain and use the C software at the University of Washington compression Web site, or write your own matlab or c code. There is also scattered Matlab and C/C++ on the Web.

9. Describe the Lloyd optimal fixed-rate quantizer for a random variable with pdf shown below and $N = 3$.



10. Approximating continuous random vectors by discrete random vactors

The *Wassershtein* distance or ρ-bar distance (or Ornstein distance) between two random vectors X and Y with distributions P_X and P_Y, respectively, is defined in terms of a distortion measure d such as squared error as

$$\overline{\rho}(P_X, P_Y) = \inf_{P_{X,Y}} Ed(X, Y)$$

where the infimum is over all joint distributions having P_X and P_Y as marginal distributions. In the squared error case this becomes

$$\overline{\rho}(P_X, P_Y) = \inf_{P_{X,Y}} E(\|X - Y\|^2).$$

A reasonable way to fit a discrete distribution to a continuous distribution would be to find the discrete distribution minimizing the $\overline{\rho}$ distance subject to some constraints, for example, letting A_Y denote the discrete alphabet of Y,

$$\overline{\rho}_X(R) = \begin{cases} \inf_{P_Y : H(Y) \leq R} \overline{\rho}(P_X, P_Y) & \text{entropy constrained} \\ \inf_{P_Y : \ln N(A_Y) \leq R} \overline{\rho}(P_X, P_Y) & \text{alphabet size constrained} \end{cases}$$

What can you say about the relationships between these optimization problems and the operational rate distortion functions for fixed and variable rate quantizers?

Another way to look at optimizing fits of discrete sources to continuous ones is to use a Lagrangian approach and define

$$\overline{\rho}_X(\lambda) = \inf_{P_Y} \left(\overline{\rho}(P_X, P_Y) + \lambda R(P_Y)\right)$$

where $R(P_Y) = \ln N(A_Y)$ or $H(Y)$. What can you say about relationships between this optimization and the Lagrangian operational DRF $\rho(\lambda)$ for the purely fixed and variable rate cases?