

**Homework #3**

Due: Monday 14 May 2007

This problem set has been expanded on 6 May with the addition of two new problems.

In compensation, this will be the *last* problem set. Out of class time following the midterm can be devoted fully to the project.

1. *Scaling and shifting densities*

Suppose that  $f_X$  is a well behaved pdf and that  $q^* = (\mathcal{E}^*, \mathcal{D}^*, \ell^*)$  is an optimal quantizer for  $f$  in the sense that for given  $\lambda > 0, \eta \in [0, 1]$

$$\rho(f_X, \lambda, \eta, q^*) = \rho(f_X, \lambda, \eta) = \inf_q \rho(f_X, \lambda, \eta, q)$$

where

$$\rho(f_X, \lambda, \eta, q) = D_{f_X}(q) + \lambda[(1 - \eta)Hf_X(q) + \eta \ln N]$$

where the distortion is squared error and the infimum is over all quantizers (including admissible length functions  $\ell$ ). The subscript denotes the pdf with respect to which the expectation is taken.

For a positive scalar constant  $a$  and a fixed vector  $b$ , define the random vector  $Y = aX + b$ , i.e.,  $Y$  is a scaled and shifted version of  $X$ .

(a) Consider a quantizer  $q = (\mathcal{E}, \mathcal{D}, \ell)$  defined in terms of  $q^*$  by

$$\begin{aligned} \mathcal{E}(y) &= \mathcal{E}^*((y - b)/a) \\ \mathcal{D}(i) &= a\mathcal{D}^*(i) + b \end{aligned}$$

Show that

$$\frac{\rho(f_Y, \lambda, \eta, q)}{\lambda} = \frac{\rho(f_X, a^{-2}\lambda, \eta, q^*)}{a^{-2}\lambda}$$

and show that  $q$  is optimal for  $Y$  with Lagrange multiplier  $a^{-2}\lambda$  if  $q^*$  is optimal for  $X$  with Lagrange multiplier  $\lambda$ . Describe the Zador/Gersho limiting behavior as  $\lambda \rightarrow 0$  for the special cases  $\eta = 0, 1$ .

(b) Derive a similar result for the operational distortion-rate function for the fixed and variable-rate case; that is, find the distortion-rate formulation rather than the Lagrangian formulation just derived.

(c) As a special case of the previous result, show how the operational DRF for a scalar random variable  $X$  relates to that for  $Y = aX + b$  where  $a$  and  $b$  are chosen so that  $Y$  has zero mean and unit variance.

- (d) As a further special case of the previous result, show in particular that in the case of a Gaussian random variable  $X$  with zero mean and variance  $\sigma_X^2$ , the average distortion at high rate for a fixed rate code is given approximately by

$$\delta_1(N) \approx \sigma_X^2 \frac{\sqrt{3}\pi}{2} N^{-2}$$

Where

$$\frac{\sqrt{3}\pi}{2} = \frac{1}{12} \left( \int_{-\infty}^{\infty} \left[ \frac{e^{-x^2/2}}{\sqrt{2\pi}} \right]^{1/3} dx \right)^3.$$

- (e) Derive a similar result to the previous part for the variable-rate case  $\delta_0(R)$ , that is, derive a formula for the general scalar Gaussian case in terms of the solution for a zero mean unit variance Gaussian.

2. *Transform coding* Recall the transform code of the first homework formed by applying a unitary matrix  $U$  to the input vector  $X$  to obtain a vector  $Y = UX$  which then has each component individually quantized to form an approximation  $\hat{Y} = (q_0(Y_0), \dots, q_{k-1}(Y_{k-1}))^t$ . The final approximation  $\hat{X}$  to  $X$  is formed by  $\hat{X} = U^{-1}\hat{Y}$ . In homework 1 it was shown that the mean squared error in both domains are the same, that is,  $\|X - \hat{X}\|^2 = \|Y - \hat{Y}\|^2$ .

Suppose now that each of the component scalar quantizers  $q_i$  has  $N_i$  codewords and that  $X = (X_0, \dots, X_{k-1})$  are samples from a zero mean stationary Gaussian random process with correlation  $R_X(j) = E(X_n X_{n-j})$ . Let  $R$  denote the corresponding  $k \times k$  correlation matrix  $E(XX^t)$ . Let  $r_i = \ln N_i$  denote the quantity of nats per component of the vector.

- (a) Use the high rate approximations to show that the average distortion for the overall quantizer is given approximately by

$$D(q) \approx \gamma \sum_{i=0}^{k-1} \sigma_{Y_i}^2 N_i^{-2}$$

where  $\gamma$  is a constant you must find.

- (b) Use the arithmetic/geometric mean inequality to show that by

$$D(q) \geq k\gamma \left( \prod_{i=0}^{k-1} \sigma_{Y_i}^2 \right)^{1/k} e^{-\frac{2}{k} \sum_{i=0}^{k-1} r_i}$$

and that equality holds if and only if

$$\frac{\sigma_{Y_i}^2}{N_i^2} = c, \quad i = 0, \dots, k-1$$

for a constant  $c$ .

If this is the case, note that each quantizer contributes the same partial distortion to the overall average. It is also the source of a famous rule of thumb that in the high rate regime the number of quantization levels devoted to each transform coefficient is proportional to the standard deviation of the coefficient being quantized.

- (c) A further lower bound can be obtained by using the *covariance determinant bound* which states that for zero mean random variables  $Y_i$ ,

$$\left(\prod_{i=0}^{k-1} \sigma_{Y_i}^2\right) \geq \det(R_Y)$$

where  $R_Y = E(YY^t)$ . (See Gersho and Gray Section 8.6 for a proof.) Equality holds if and only if  $R_Y$  is diagonal. Show that for the decorrelating unitary transform (called the *Karhunen-Loeve transform*) of homework 1, these lower bounds hold with equality and express the final approximation in terms of the input covariance matrix  $R_X$ .

In summary, the Karhunen-Loeve transform is the optimal transform for transform coding Gaussian vectors in the high rate regime in the sense that it achieves an unbeatable lower bound over all transforms. Some more general results along this line have been obtained by Vivek Goyal.

### 3. Classifier-based distortion measure

A fixed-rate vector quantizer  $Q$  is described by its codebook of reproduction labels  $\mathcal{C} = \{y_l; l = 1, \dots, N\}$  and its encoder partition  $\mathcal{S} = \{S_l; l = 1, \dots, N\}$ . The quantizer is to be designed for a  $k$ -dimensional input vector  $X$ .

Suppose that the input space can be divided into two classes of vectors, say  $H_0$  and  $H_1$ , where both  $H_0$  and  $H_1$  are considered to be disjoint subsets of  $k$  dimensional Euclidean space. Suppose that  $\Pr(X \in H_0) = p = 1 - \Pr(X \in H_1)$ . Define the classifier function

$$\delta(x) = \begin{cases} 1 & x \in H_0 \\ 0 & x \in H_1 \end{cases}$$

The vectors in  $H_0$  are considered to be very important and those in  $H_1$  are considered to be not so important. This is reflected by defining a distortion measure  $d$  by

$$d(x, y) = C_0 \|x - y\|^2 \delta(x) + C_1 \|x - y\|^2 (1 - \delta(x))$$

$$\begin{cases} C_0 \|x - y\|^2 & x \in H_0 \\ C_1 \|x - y\|^2 & x \in H_1 \end{cases}$$

Where  $C_0 > C_1 \geq 0$  represent the relative costs of squared error for the two types of input vectors. In words: If the input vector is in class  $H_0$ , then its squared error counts  $C_0/C_1$  times more than squared error in class  $H_1$ . Counting distortion more in this set should force the code design to do a better job of such vectors.

Describe the Lloyd necessary conditions for a vector quantizer with codebook  $\mathcal{C}$  and encoder partition  $\mathcal{S}$  to be optimal with respect to this distortion measure.

4. *The “quantization theorem”* This problem considers classic results on uniform quantizers developed by Widrow (1960) and Sripad and Snyder (1977), although the basic technique dates back to Rice and Davenport and Root.

Revisit problem 5 of Homework 1 for a description of the characteristic function method for analyzing a uniform quantizer.

- (a) Prove that that a necessary and sufficient condition for the quantizer error pdf to be uniform is that  $\Phi_X(2\pi n/\Delta) = 0$  for all  $n \neq 0$ . Give at least two examples of pdf's  $f_X$  with the property. Can you describe in general pdf's for which this will be true?

*Hint:* Show that for  $u \in [-\Delta/2, \Delta/2]$ ,

$$f_\epsilon(u) = \frac{1}{\Delta} + \frac{1}{\Delta} \sum_{n \neq 0} \Phi_X\left(\frac{2\pi n}{\Delta}\right) e^{-2\pi n u / \Delta}$$

*Note:* A special case of the condition is Widrow's assumption that  $\Phi_X(\nu) = 0$  for all  $|\nu| \geq 2\pi/\Delta$

- (b) Prove that a necessary and sufficient condition for the joint pdf for  $\epsilon_n$  and  $\epsilon_m$  to have the form

$$f_{\epsilon_n, \epsilon_m}(u, v) = \frac{1}{\Delta^2} \text{ for } u, v \in [-\Delta/2, \Delta/2]$$

is that the joint input characteristic function satisfies

$$\Phi_{X_n, X_m}\left(\frac{2\pi i}{\Delta}, \frac{2\pi \ell}{\Delta}\right) = 0 \text{ all } i \neq 0, \ell \neq 0.$$

Thus in this case quantization errors will be uniform and independent!

*Hint:* Show that for  $u, v \in [-\Delta/2, \Delta/2]$ ,

$$f_{\epsilon_n, \epsilon_m}(u, v) = \frac{1}{\Delta^2} + \frac{1}{\Delta^2} \sum_{i \neq 0} \sum_{\ell \neq 0} \Phi_{X_n, X_m}\left(\frac{2\pi i}{\Delta}, \frac{2\pi \ell}{\Delta}\right) e^{-2\pi j(iu + \ell v) / \Delta}$$

These results characterize the sources for which the Bennet approximations are exact. The manipulations should be reminiscent of the sampling theorem, except here a Fourier series is found for a pdf (the "signal") defined only on a finite interval  $[-\Delta/2, \Delta/2]$  and the Fourier series only equals the signal on the interval. The samples of the transform are simply Fourier coefficients.

5. (a) Design a Shannon lossless code for the source with 8 symbols with probabilities:

$$0.04, 0.07, 0.09, 0.10, 0.10, 0.15, 0.2, 0.25$$

- (b) Design a Huffman code for the same source  
 (c) Compute the average word length for the above codes and compare with the entropy.