Homework 7 Solutions **EECS 651** **March 26, 2001**

1. *A variable-rate k-dimensional transform code is optimized* ***for rate R*** *and for a zero-mean, stationary Gaussian source.* ***Assume R is large.***

(a) *Show that its point density is the optimal point density for variable-rate k-dimensional vector quantizers with n=1.*

We first find the point density for an optimal transform code. When a transform code is optimized for a Gaussian random vector $\underline{X}$ and rate R, the transform T is the KLT, and the ith quantizer is uniform with distortion

$$D_i \cong \delta_{transform}(k,R) \cong \frac{1}{12} |K_{\underline{X}}|^{1/k} \eta_{G,1} \, 2^{-2R} \quad \text{(it's the same for all i)}$$

We also know $D_i \cong \frac{\Delta_i^2}{12}$. Equating the two expressions for D_i shows that the ith quantizer has step size

$$\Delta_i = |K_{\underline{X}}|^{1/2k} (\eta_{G,1})^{1/2} \, 2^{-R} \quad \text{(same for all i)}$$

$$= |K_{\underline{X}}|^{1/2k} (2\pi e)^{1/2} \, 2^{-R} \quad \text{since } \eta_{G,1} = 2\pi e \text{ (see Prop. 4, p. 27 of lectures on VQ with EC}$$

and its point density (unnormalized, since variable-rate coding is used) is

$$\Lambda_i(u) = \frac{1}{\Delta_i} = |K_{\underline{X}}|^{-1/2k} (2\pi e)^{-1/2} \, 2^{R} \quad \text{(same for all, and constant with u)}$$

If we consider how the coefficient vector $\underline{U} = T\underline{X}$ is quantized into $\underline{V}$ we see that $\underline{U}$ is quantized with a product quantizer, whose point density is

$$\Lambda_{\underline{U}}(\underline{u}) = \Lambda_1(u_1)\, \Lambda_2(u_2) \ldots \Lambda_k(u_k) \quad \text{(product quantizers have product point densities.)}$$

$$= |K_{\underline{X}}|^{-1/2} (2\pi e)^{-k/2} \, 2^{kR} \quad \text{(notice that it is constant with } \underline{u}\text{)}$$

Finally, the point density of the transform code is

$$\Lambda(\underline{x}) = \Lambda_{\underline{U}}(T\underline{x}) \quad \text{(see Problem 2f, HW Set 3)}$$

$$= |K_{\underline{X}}|^{-1/2} (2\pi e)^{-k/2} \, 2^{kR}$$

In comparison the point density of a quantizer optimized for variable-rate coding with n = 1, is

$$\Lambda(\underline{x}) = 2^{k(R-h_k)} \quad \text{(p. 12 of the notes on VQ-EC)}$$

$$= |K_{\underline{X}}|^{-1/2} (2\pi e)^{-k/2} \, 2^{kR} \quad \text{since} \quad h_k = \frac{1}{2} \log_2 2\pi e |K|^{1/k}$$

which is the same as the point density of the optimized transform code.

(b) *Find its inertial profile.*

The product quantizer that effectively quantizes $\underline{U}$ has cubic cells, because each uniform scalar quantizer has the same step size. Therefore, its inertial profile is

$$m_{\underline{U}}(u) = \frac{1}{12}$$

From Problem 2f, HW Set 3,

$$\mathbf{m}(\underline{\mathbf{x}}) = m_U(T\underline{u}) = \mathbf{\frac{1}{12}}$$

2. (a) *Derive the OPTA function* **for large rate and** *k-dimensional transform coding in which the scalar quantizers are replaced by two-dimensional vector quantizers. That is, one vq is applied to* (U_1,U_2)*, another is applied to* (U_3,U_4)*, and so on. Assume* k *is even, and assume the source is zero mean, stationary and Gaussian. Your answer should include both the fixed-rate and variable-rate cases. You may parallel the derivation given in class, and you may use without rederivation any fact that was proven in class which is useful here. (Just remember to state the fact that you are using.)*

Let T be an orthogonal transform, let $\underline{U} = (U_1,...,U_k) = T\underline{X}$ be the vector of coefficients. We also write $\underline{U} = (\underline{U}_1, \underline{U}_2, ..., \underline{U}_{k/2})$ where $\underline{U}_i = (U_{i,1},U_{i,2})$ is the ith pair of coefficients. That is, $\underline{U}_1 = (U_1,U_2)$, $\underline{U}_2 = (U_3,U_4)$, ..., $\underline{U}_{k/2} = (U_{k-1},U_k)$. Let Q_i denote the 2-dimensional vq applied to $\underline{U}_i$. Then, the distortion and rate of a transform code with two-dimensional vector quantizers applied to successive pairs of coefficients are

$$D = \frac{2}{k}\sum_{i=1}^{k/2} D_i(Q_i) \quad \text{and} \quad R = \frac{2}{k}\sum_{i=1}^{k/2} R_i(Q_i)$$

where $D_i(Q_i)$ and $R_i(Q_i)$ are the distortion and rate, respectively of Q_i applied to $\underline{U}_i$.
Given some set of R_i's to be determined later, the vq's should have rates at most R_i and distortion equal to the opta funcion $\delta_i(2,R_i)$ for 2-dimensional vq (either fixed-rate or variable-rate) applied to $\underline{U}_i$, where

$$\delta_i(2,R_i) = m_2^* \tilde{\sigma}_i^2 \alpha_i 2^{-2R_i}, \quad m_2^* = .0802, \quad \bar{\sigma}_i^2 = \frac{1}{2}\Big(\mathrm{var}(U_{i,1}) + \mathrm{var}(U_{i,2})\Big)$$

$$\alpha_i = \frac{1}{\bar{\sigma}_i^2}\alpha_G |K_i|^{1/2}, \quad \alpha_G = \begin{cases} 2\pi\, 3^{3/2} = 32.6, & \text{fixed rate coding} \\ 2\pi e = 17.1, & \text{variable-rate coding}\end{cases}$$

$K_{\underline{U}_i}$ is the 2×2 covariance matrix of $\underline{U}_i$

$$= m_2^* |K_{\underline{U}_i}|^{1/2} \alpha_G 2^{-2R_i}$$

It follows that when the 2-dimensional vector quantizers are optimized for a given transform T and a given rate R, then

$$D = \min_{R_1,...,R_k:\ \frac{2}{k}\sum_{i=1}^{k/2} R_i \le R} \frac{2}{k}\sum_{i=1}^{k/2} \delta_i(2,R_i) = \min_{R_1,...,R_k:\ \frac{2}{k}\sum_{i=1}^{k/2} R \le R} \frac{2}{k}\sum_{i=1}^{k/2} m_2^* \alpha_G |K_{\underline{U}_i}|^{1/2} 2^{-2R_i}$$

In class we minimized: $\frac{1}{k}\sum_{i=1}^{k} \frac{1}{12}\sigma_i^2\alpha_i 2^{-2R_i}$ subject to $\frac{1}{k}\sum_{i=1}^{k} R_i = R$,

and found $R_i = R + \frac{1}{2}\log_2 \frac{\sigma_i^2\alpha_i}{\left(\prod_{j=1}^{k}\sigma_j^2\alpha_j\right)^{1/k}}$ and $D \cong \frac{1}{12}\left(\prod_{j=1}^{k}\sigma_j^2\alpha_j\right)^{1/k} 2^{-2R}$

We have the same sort of thing here, except k is replaced by $k/2$ and $\frac{1}{12}\sigma_i^2\alpha_i$ is replaced by $m_2^*\alpha_G |K_i|^{1/2}$. The result is

$$R_i = R + \frac{1}{2}\log_2 \frac{|K_{\underline{U}_i}|^{1/2}}{\left(\prod_{j=1}^{k/2}|K_{\underline{U}_j}|^{1/2}\right)^{2/k}}$$

and

$$D \cong m_2^* \alpha_G \left(\prod_{j=1}^{k/2} |K_{\underline{U}_j}|^{1/2}\right)^{2/k} 2^{-2R} = m_2^* \alpha_G \left(\prod_{j=1}^{k/2} |K_{\underline{U}_j}|\right)^{1/k} 2^{-2R} \qquad (**)$$

It remains only to choose the transform to minimize $\prod_{j=1}^{k/2} |K_{\underline{U}_j}|$. I will show that the KLT does this.

First, consider this product for the KLT. When T is the KLT, the components of $\underline{U}$ are uncorrelated, so the components of each $\underline{U}_j$ are uncorrelated, so $K_{\underline{U}_i}$ is a diagonal 2×2 matrix with

$$|K_{\underline{U}_j}| = K_{\underline{U}_j}(1,1)\, K_{\underline{U}_j}(2,2) = \mathrm{var}(U_{j,1})\mathrm{var}(U_{j,2}) .$$

It follows that

$$\prod_{j=1}^{k/2} |K_{\underline{U}_j}| = \prod_{j=1}^{k} \mathrm{var}(U_j) = |K|.$$

Now suppose T is any arbitrary orthogonal transform. We will create another orthonormal transform $\tilde{T}$ whose product is the same as T and at least as large as that of the KLT, which will demonstrate that T is not better than the KLT. Accordingly, given some T, each 2-dimensional vector of coefficients $\underline{U}_i$ may have correlated components, i.e. $U_{i,1}$ and $U_{i,2}$ may be correlated. However, after computing $\underline{U} = T\underline{X}$, one can apply an orthogonal 2×2 transform S_1 to $\underline{U}_1$ such that $\tilde{\underline{U}}_1 = S_1 \underline{U}_1$ has uncorrelated components. One can similarly apply an orthogonal 2×2 transform S_2 to $\underline{U}_2$ to obtain $\tilde{\underline{U}}_2 = S_2 \underline{U}_2$ with uncorrelatd components, and so on. Overall we create a new transform $\tilde{T}$ that produces

$$\tilde{\underline{U}} = (\tilde{U}_1,\ldots,\tilde{U}_k) = (\tilde{\underline{U}}_1,\ldots,\tilde{\underline{U}}_{k/2}) = \tilde{T}\, \underline{X},$$

where each $\tilde{\underline{U}}_i$ has uncorrelated components. It is easy to see that $||\tilde{\underline{U}}|| = ||\underline{U}|| = ||\underline{X}||$. Therefore, $\tilde{T}$ is orthogonal. Let $K_{\tilde{\underline{U}}_i}$ denote the covariance matrix of $\tilde{\underline{U}}_i$. As discussed in class, the facts that $\tilde{\underline{U}}_i = S_i\, \underline{U}$ and S_i is orthogonal imply

$$|K_{\underline{U}_j}| = |K_{\tilde{\underline{U}}_i}| = K_{\tilde{\underline{U}}_i}(1,1)\;\; K_{\tilde{\underline{U}}_i}(2,2)$$ where the last equality is because $K_{\tilde{\underline{U}}_i}$ is diagonal.

It now follows that

$$\prod_{j=1}^{k/2} |K_{\underline{U}_j}| = \prod_{j=1}^{k} K_{\tilde{\underline{U}}}(j,j) \geq |K_{\tilde{\underline{U}}}| = |K|.$$

where the inequality is from Fact 6, p. 26 of the transform coding notes, and the last equality is from Fact 8, p. 26. We now see that the KLT minimizes $\prod_{j=1}^{k/2} |K_{\underline{U}_j}|$.

Substituting $\prod_{j=1}^{k/2} |K_{\underline{U}_j}| = |K|$ into (**), gives the opta function for transform coding with two-dimensional vq's

$$\boldsymbol{\delta_{\mathrm{transf}}(R) = m_2^* \alpha_G \;\; |K| \;\; 2^{-2R}} \; .$$

Notice that this only differs from the opta function for transform coding with scalar quantization in that $m_1^* = 1/12$ has been replaced by m_2^* .

(b) *Would the OPTA function change if instead of applying 2-dimensional vq's to (U_1,U_2), (U_3,U_4), oen applied two-dimensional vq's to (U_1,U_k), (U_2,U_{k-1}), ... ?*

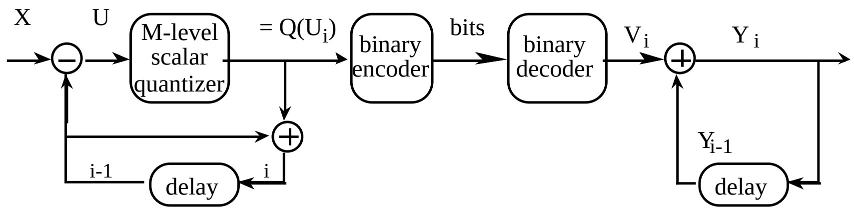
No. Since the minimal distortion again depends only on the product of the eigenvalues, it does not matter how coefficients are paired.

3. (a) *Design and describe a fixed-rate DPCM code with rate 3 and with a linear predictor of order two, for a zero-mean, stationary Gaussian source with autocorrelation function:*

$$R_X(k) = \frac{-128}{105}\,\frac{1}{4^{|k|}} + \frac{64}{21}\,\frac{1}{2^{|k|}} \;.$$

The distortion should be as small as you can make it. Hints: In designing the code, you may use the assumptions we used in our high resolution analysis, and you may use the scalar quantization table that was used in the previous homework.

We **describe** the DPCM code by giving a block diagram and by describing the predictor, quantizer and binary decoder. The block diagram is:



We will follow the high-resolution design procedure described in class.

We choose the predictor to be the optimal linear predictor for X_i from X_{i-1}, X_{i-2}. This is

$$X_i = a_1 X_{i-1} + a_2 X_{i-2}$$

where $\underline{a} = (a_1,a_2)^t = K^{-1}\,\underline{r}$, $K = \begin{bmatrix} R_X(0) & R_X(1) \\ R_X(1) & R_X(0) \end{bmatrix} = \frac{64}{35}\begin{bmatrix} 1 & 2/3 \\ 2/3 & 1 \end{bmatrix}$, $\underline{r} = \begin{bmatrix} R_X(1) \\ R_X(2) \end{bmatrix} = \frac{8}{105}\begin{bmatrix} 16 \\ 9 \end{bmatrix}$

Inverting K gives $K^{-1} = \frac{63}{64}\begin{bmatrix} 1 & -2/3 \\ -2/3 & 1 \end{bmatrix}$ and solving for the predictor coefficients gives

$$\underline{\mathbf{a}} = K^{-1}\underline{r} = \begin{bmatrix} \mathbf{3/4} \\ \mathbf{-1/8} \end{bmatrix}.$$

The resulting MSE is $M_2 = \sigma^2 - \underline{a}^t\,\underline{r} = 1$.

We now choose the quantizer to be the optimal rate 3 scalar quantizer for $\tilde{U}_i = a_1 X_{i-1} + a_2 X_{i-2}$ which is Gaussian with zero mean and variance equal to $M_2 = 1$.

From the optimal quantizer tables, we find that the quantizer should have:

levels: -2.152, -1.344, -0.756, -0.245, 0.245, 0.756, 1.344, 2.152

thresholds: -∞, -1.748, -1.050, -0.501, 0.0, 0.501, 1.050, 1.748, ∞

The binary encoder assigns binary sequences of length 3 to the quantization cells in some arbitrary order.

(b) *This part is not required and will not be graded. Show that a DPCM code with a higher order predictor would do no better.*

One can use the optimality property. One merely has to check that

$$E\left(X_i - \frac{3}{4}X_{i-1} + \frac{1}{8}X_{i-2}\right)X_{i-j} = 0 \quad \text{for } j = 3, 4, 5, \ldots$$

It turns out that this means that X is a second-order AR process, i.e. $X_i = \frac{3}{4}X_{i-1} - \frac{1}{8}X_{i-2} + Z_i$,

where Z_i's are IID Gaussian with mean zero and variance 1.

(c) *This part is not required and will not be graded. Find, approximately, the OPTA function assuming no restriction on the order of the linear predictor.*

Since higher-order predictors are no better, the OPTA function is the same as for part (a) namely,

$$\boldsymbol{\delta_{dpcm}(R) = \frac{1}{12} \times 32.6 \times 1 \times 2^{-2R}}\;.$$