Transform coding, example

Problem

Suppose the signal is a stationary Gaussian source X_n with zero mean and auto correlation function $R_{XX}(k) = E\{X_n X_{n+k}\} = 0.9^{|k|}$. We want to transform code the source using zonal coding, such that the average rate is 2 bits/sample and the distortion is minimized. The quantizers are Lloyd-Max quantizers and the source coder is just a fixed length coder.

No transform

For reference, we first see what result we get when not using any transform. In the formula collection we can find the resulting distortion for a 2 bit (ie 4 level) Lloyd-Max quantizer given a Gaussian input signal:

$$D \approx 0.1175 \cdot \sigma_X^2 = 0.1175 \cdot R_{XX}(0) = 0.1175$$

Expressed as a signal-to-noise ration, we get

$$\mathrm{SNR} = 10 \cdot \log_{10} \frac{\sigma_X^2}{D} \approx 10 \cdot \log_{10} \frac{1}{0.1175} \approx 9.30 \text{ [dB]}$$

2-point DCT/DWHT

The transform matrix for a 2-point DCT or DWHT (they are the same) looks like

$$\mathbf{A} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

and it operates on vectors $\bar{\mathbf{x}}$ giving transform vectors $\bar{\theta}$:

$$\bar{\mathbf{x}} = \begin{pmatrix} X_0 \\ X_1 \end{pmatrix} \quad ; \quad \bar{\theta} = \begin{pmatrix} \theta_0 \\ \theta_1 \end{pmatrix} = \mathbf{A}\bar{\mathbf{x}}$$

Variances of the two transform components:

$$\begin{aligned}
\sigma_0^2 &= E\{\theta_0^2\} = \frac{1}{2}E\{(X_0 + X_1)^2\} = \frac{1}{2}E\{X_0^2 + X_1^2 + 2X_0X_1\} = \\
&= R_{XX}(0) + R_{XX}(1) = 1.9 \\
\sigma_1^2 &= E\{\theta_1^2\} = \frac{1}{2}E\{(X_0 - X_1)^2\} = \frac{1}{2}E\{X_0^2 + X_1^2 - 2X_0X_1\} = \\
&= R_{XX}(0) - R_{XX}(1) = 0.1
\end{aligned}$$

An alternative way of finding the variances is by using the correlation matrix for the input vector

$$\mathbf{R}_X = E\{\bar{\mathbf{x}}\bar{\mathbf{x}}^T\} = \begin{pmatrix} R_{XX}(0) & R_{XX}(1) \\ R_{XX}(1) & R_{XX}(0) \end{pmatrix} = \begin{pmatrix} 1 & 0.9 \\ 0.9 & 1 \end{pmatrix}$$

and calculate the correlation matrix for the transform vector

$$\mathbf{R}_\theta = E\{\bar{\theta}\bar{\theta}^T\} = \mathbf{A}\mathbf{R}_X\mathbf{A}^T = \begin{pmatrix} 1.9 & 0 \\ 0 & 0.1 \end{pmatrix}$$

The variances are found as the diagonal elements of $\mathbf{R}_\theta$.

We now want allocate bits to the two transform components such that the average distortion is minimized and the average rate is 2 bits/sample. We can do this by starting with rate 0 on each component and then iteratively giving a bit to the component with the currently highest distortion and then calculating the new distortion. Since the input is Gaussian, the transform components are also Gaussian. Distortion values are found in the formula collection.

R_i	D_i	θ_0	θ_1
0	σ_i^2	1.9	0.1
1	$0.3634 \cdot \sigma_i^2$	0.69046	0.03634
2	$0.1175 \cdot \sigma_i^2$	0.22325	
3	$0.03454 \cdot \sigma_i^2$	0.065626	
4	$0.009497 \cdot \sigma_i^2$		

The first bit is given to θ_0, the second to θ_0, the third to θ_0 and the fourth to θ_1. We now have the rates $R_0 = 3$ and $R_1 = 1$ and the distortions $D_0 \approx 0.065626$ and $D_1 \approx 0.03634$ which means we have the average rate

$$R = \frac{R_0 + R_1}{2} = 2$$

and the average distortion

$$D = \frac{D_0 + D_1}{2} \approx 0.050983$$

which gives us the signal-to-noise ratio

$$\text{SNR} = 10 \cdot \log_{10} \frac{\sigma_X^2}{D} \approx 10 \cdot \log_{10} \frac{1}{0.050983} \approx 12.93 \text{ [dB]}$$

4-point DWHT

The transform matrix for a 4-point DWHT looks like

$$\mathbf{A} = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{pmatrix}$$

and it operates on vectors $\bar{\mathbf{x}}$ giving transform vectors $\bar{\theta}$:

$$\bar{\mathbf{x}} = \begin{pmatrix} X_0 \\ X_1 \\ X_2 \\ X_3 \end{pmatrix} \quad ; \quad \bar{\theta} = \begin{pmatrix} \theta_0 \\ \theta_1 \\ \theta_3 \\ \theta_3 \end{pmatrix} = \mathbf{A}\bar{\mathbf{x}}$$

Variances of the two transform components:

$$\begin{aligned}
\sigma_0^2 &= E\{\theta_0^2\} = \frac{1}{4}E\{(X_0+X_1+X_2+X_3)^2\} = \\
&= \frac{1}{4}E\{X_0^2+X_1^2+X_2^2+X_3^2+2X_0X_1+2X_0X_2+2X_0X_3+2X_1X_2+2X_1X_3+2X_2X_3\} = \\
&= \frac{1}{4}(4\cdot R_{XX}(0)+6\cdot R_{XX}(1)+4\cdot R_{XX}(2)+2\cdot R_{XX}(3)) = 3.5245 \\
\sigma_1^2 &= E\{\theta_1^2\} = \frac{1}{4}E\{(X_0+X_1-X_2-X_3)^2\} = \\
&= \frac{1}{4}E\{X_0^2+X_1^2+X_2^2+X_3^2+2X_0X_1-2X_0X_2-2X_0X_3-2X_1X_2-2X_1X_3+2X_2X_3\} = \\
&= \frac{1}{4}(4\cdot R_{XX}(0)+2\cdot R_{XX}(1)-4\cdot R_{XX}(2)-2\cdot R_{XX}(3)) = 0.2755 \\
\sigma_2^2 &= E\{\theta_2^2\} = \frac{1}{4}E\{(X_0-X_1-X_2+X_3)^2\} = \\
&= \frac{1}{4}E\{X_0^2+X_1^2+X_2^2+X_3^2-2X_0X_1-2X_0X_2+2X_0X_3+2X_1X_2-2X_1X_3-2X_2X_3\} = \\
&= \frac{1}{4}(4\cdot R_{XX}(0)-2\cdot R_{XX}(1)-4\cdot R_{XX}(2)+2\cdot R_{XX}(3)) = 0.1045 \\
\sigma_3^2 &= E\{\theta_3^2\} = \frac{1}{4}E\{(X_0-X_1+X_2-X_3)^2\} = \\
&= \frac{1}{4}E\{X_0^2+X_1^2+X_2^2+X_3^2-2X_0X_1+2X_0X_2-2X_0X_3-2X_1X_2+2X_1X_3-2X_2X_3\} = \\
&= \frac{1}{4}(4\cdot R_{XX}(0)-6\cdot R_{XX}(1)+4\cdot R_{XX}(2)-2\cdot R_{XX}(3)) = 0.0955
\end{aligned}$$

An alternative way of finding the variances is by using the correlation matrix for the input vector

$$\begin{aligned}
\mathbf{R}_X &= E\{\bar{\mathbf{x}}\bar{\mathbf{x}}^T\} = \begin{pmatrix} R_{XX}(0) & R_{XX}(1) & R_{XX}(2) & R_{XX}(3) \\ R_{XX}(1) & R_{XX}(0) & R_{XX}(1) & R_{XX}(2) \\ R_{XX}(2) & R_{XX}(1) & R_{XX}(0) & R_{XX}(1) \\ R_{XX}(3) & R_{XX}(2) & R_{XX}(1) & R_{XX}(0) \end{pmatrix} = \\
&= \begin{pmatrix} 1 & 0.9 & 0.81 & 0.729 \\ 0.9 & 1 & 0.9 & 0.81 \\ 0.81 & 0.9 & 1 & 0.9 \\ 0.729 & 0.81 & 0.9 & 1 \end{pmatrix}
\end{aligned}$$

and calculate the correlation matrix for the transform vector

$$\mathbf{R}_\theta = E\{\bar{\theta}\bar{\theta}^T\} = \mathbf{A}\mathbf{R}_X\mathbf{A}^T = \begin{pmatrix} 3.5245 & 0 & -0.0855 & 0 \\ 0 & 0.2755 & 0 & 0.0855 \\ -0.0855 & 0 & 0.1045 & 0 \\ 0 & 0.0855 & 0 & 0.0955 \end{pmatrix}$$

The variances are found as the diagonal elements of $\mathbf{R}_\theta$.

We now want allocate bits to the four transform components such that the average distortion is minimized and the average rate is 2 bits/sample. We can do this by starting with rate 0 on each component and then iteratively giving a bit to the component with the currently highest distortion and then calculating the new distortion. Since the input is Gaussian, the transform components are also Gaussian. Distortion values are found in the formula collection.

R_i	D_i	θ_0	θ_1	θ_2	θ_3
0	σ_i^2	3.5245	0.2755	0.1045	0.0955
1	$0.3634 \cdot \sigma_i^2$	1.2808	0.1001	0.03798	0.03470
2	$0.1175 \cdot \sigma_i^2$	0.4141	0.03237		
3	$0.03454 \cdot \sigma_i^2$	0.1217			
4	$0.009497 \cdot \sigma_i^2$	0.03347			

The first bit is given to θ_0, the second to θ_0, the third to θ_0, the fourth to θ_1, the fifth to θ_0, the sixth to θ_2, the seventh to θ_1 and the eighth to θ_3. We now have the rates $R_0 = 4$, $R_1 = 2$, $R_2 = 1$ and $R_3 = 1$ and the distortions $D_0 \approx 0.03347$, $D_1 \approx 0.03237$, $D_2 \approx 0.03798$ and $D_3 \approx 0.03470$ which means we have the average rate

$$R = \frac{R_0 + R_1 + R_2 + R_3}{4} = 2$$

and the average distortion

$$D = \frac{D_0 + D_1 + D_2 + D_3}{4} \approx 0.03463$$

which gives us the signal-to-noise ratio

$$\text{SNR} = 10 \cdot \log_{10} \frac{\sigma_X^2}{D} \approx 10 \cdot \log_{10} \frac{1}{0.03463} \approx 14.61 \text{ [dB]}$$