The OPTA Function for Stationary Sources

For a stationary sources, we can use VQ's of any dimension and we seek the least possible distortion of any VQ with rate $R$ or less and any dimension. That is, we seek the "ultimate" VQ OPTA function:

$$\delta(R) = \inf_{\text{VQ's with rate } R \text{ or less}} D(VQ)$$

$$= \inf_k \delta(k,R)$$

Equivalently,

$$S(R) = \sup_{\text{VQ's with rate } R \text{ or less}} \text{SNR}(VQ)$$

$$= \sup_k S(k,R)$$

There is no VQ that achieves $\delta(R)$ exactly and no value of $k$ such that $\delta(k,R) = \delta(R)$. However, by definition of an "inf", for any $R$ and any small tolerance $\varepsilon > 0$, there is a VQ with rate $R$ or less and distortion $D \leq \delta(R) + \varepsilon$, and there is a value of $k$ such that $\delta(k,R) \leq \delta(R) + \varepsilon$. That is, one can come arbitrarily close to achieving the inf.

An "inf" is like a "min" except it works in cases where "min" does not.

Example,

$$\min_{x \in (0,1)} x^2 \text{ does not exist}$$

Defn: \[ \inf_{x \in G} f(x) = \text{largest number } y \text{ such that } y \leq f(x) \text{ for all } x \in G \]

Example,

$$\inf_{x \in (0,1)} x^2 = 0$$

because $0 \leq x^2$ for all $x \in (0,1)$ and there is no larger number $y$ such that $y \leq x^2$ for all $x \in (0,1)$

Defn: \[ \sup_{x \in G} f(x) = \text{smallest number } y \text{ such that } y \geq f(x) \text{ for all } x \in G \]

- "inf" and "sup" are short for "infimum" and "supremum"
- covered in Math 451
Properties of the OPTA functions of Stationary Sources

The OPTA's have a decreasing trend as \( k \) increases. However, it is not known if \( \delta(k,R) \geq \delta(k+1,R) \) for all \( k \). All that is known is:

**Fact:** The OPTA function \( \delta(k,R) \) is subadditive in \( k \); i.e.
\[
\delta(k+l,R) \leq \frac{k}{k+l} \delta(k,R) + \frac{l}{k+l} \delta(l,R)
\]
for any \( k,l \)

**Proof:** Let \( Q = \) product of \( Q_k \) and \( Q_l \), which are \( k \) and \( l \) dim'l VQ's with rate \( R \) or less, and with \( D(Q_k) \equiv \delta(k,R) \) and \( D(Q_l) \equiv \delta(l,R) \). As in Problem 8, HW 2.

\[
R(Q) = \frac{k}{k+l} R(Q_k) + \frac{l}{k+l} R(Q_l) \leq R
\]
\[
D(Q) = \frac{k}{k+l} D(Q_k) + \frac{l}{k+l} D(Q_l) \equiv \frac{k}{k+l} \delta(k,R) + \frac{l}{k+l} \delta(l,R)
\]

Therefore,
\[
\delta(k,R) \leq D(Q) \equiv \frac{k}{k+l} \delta(k,R) + \frac{l}{k+l} \delta(l,R)
\]

**Fact:** Subadditivity implies
\[
\delta(1,R) \geq \delta(k,R) \geq \delta(m,k,R) \text{ for any } m,k
\]

\[
\delta(R) = \inf_k \delta(k,R) = \lim_{k \to \infty} \delta(k,R)
\]

**Proof:** For any subadditive sequence, \( \lim = \inf \) (cf. Gallager, p. 112, 113)

---

Other Properties of the OPTA Function

For a stationary source with continuous random variables and fixed-rate VQ:

- \( \delta(0) = \sigma^2 \) (recall that \( \delta_{\text{vq,k}}(0) = \sigma^2 \))

- \( \delta(R) \) decreases monotonically toward zero as \( R \to \infty \).

- \( \delta(R) \) is a continuous, convex cup function of \( R \).

This is not like \( \delta(k,R) \), which has a stair-step form.
Sketch of proof of convexity: Given target rates $R_1$, $R_2$ and $\alpha$, $0 < \alpha < 1$, we must show

$$\delta_{vq}(\alpha R_2 + (1-\alpha) R_2) \leq \alpha \delta_{vq}(R_1) + (1-\alpha) \delta_{vq}(R_2).$$

First, suppose $\alpha = 1/2$. Let $Q_1$, $Q_2$ be VQ’s with large dimension $k$, with rates at most $R_1$ and $R_2$ and distortions $D_1 \equiv \delta(R_1)$ and $D_2 \equiv \delta(R_2)$, respectively. ($Q_1$, $Q_2$ exist by the defn of the OPTA function.)

Consider another VQ, denoted $Q$, with dimension $2k$ created by using $Q_1$ followed by $Q_2$. ($Q$ time shares between $Q_1$ and $Q_2$.) Then

$$R(Q) = \frac{1}{2} (R_1 + R_2)$$
$$D(Q) = \frac{1}{2} (D_1 + D_2).$$

Since $\delta_{vq}(\frac{1}{2} R_1 + \frac{1}{2} R_2) =$ least dist’n of any quant, with rate $\frac{1}{2} R_1 + \frac{1}{2} R_2$,

$$\delta_{vq}(\frac{1}{2} R_1 + \frac{1}{2} R_2) \leq D(Q) = \frac{1}{2} (D_1 + D_2) \equiv \frac{1}{2} \left( \delta(R_1) + \delta(R_2) \right).$$

A somewhat sharper argument can demonstrate the above without the "$\equiv$". It could also use a time sharing that applies to any value of $\alpha$, thereby establishing convexity.

Proof of Continuity: Convexity implies continuity, except possibly at $R = 0$. But it can be shown that $\delta(R)$ is continuous at $R = 0$, too.

---

High Resolution Analysis of the OPTA’s of Stationary Sources

Recall that for large $R$

$$\delta(k,R) \equiv \mathcal{Z}(k,R) \overset{\Delta}{=} m_k^* \beta_k \sigma^2 2^{-2R}$$

What happens as $k$ increases?

Recall: $m_k^*$’s have a decreasing trend, they are subadditive.

How about the $\beta_k$’s? They, too, have a decreasing trend. It is not known if $\beta_{k+1} \leq \beta_k$ for all $k$ and all sources, however:

Fact: For a stationary source, $\beta_k$ is submultiplicative; i.e.

$$\beta_{k+l} \leq \left( \beta_k \beta_l \right)^{1/(k+l)} \text{ for any } k,l$$

equivalently $\log \beta_k$ is subbadditive:

$$\log \beta_{k+l} \leq \frac{k}{k+l} \log \beta_k + \frac{l}{k+l} \log \beta_l \text{ for any } k,l$$

From submultiplicativity it follows that

$$\log \beta_1 \geq \log \beta_k \geq \log \beta_{mk} \text{ for every } m, k \geq 1$$

and also that $\lim_{k \to \infty} \log \beta_k = \inf_k \log \beta_k$

and consequently that

$$\beta_{\infty} \overset{\Delta}{=} \lim_{k \to \infty} \beta_k = \inf_k \beta_k$$
Sketch of proof of Fact: Let \( f_k(x) \) denote the k-dimensional density of \( X_1, \ldots, X_k \). The proof uses the

**Triple Holder inequality:**

If \( f, g \) and \( h \) are nonnegative functions, and \( p, q, r \) are nonnegative numbers such that \( \frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1 \), then

\[
\int f(x) g(x) h(x) \, dx \leq \left( \int f^p(x) \, dx \right)^{1/p} \left( \int g^q(x) \, dx \right)^{1/q} \left( \int h^r(x) \, dx \right)^{1/r}
\]

We apply this inequality to \( \sigma_{\beta_{k+1}}^{2(k+1)/(k+l+2)} = \int f_{k+1}^{(k+l)/(k+l+2)}(x,y) \, dx \, dy \) with the integrand factored as \( f_{k+1}^{(k+l)/(k+l+2)}(x,y) = f(x,y) \, g(x,y) \, h(x,y) \), where

\[
\begin{align*}
f(x,y) &= \left( f_{k+1}(x,y) \right)^{k/(k+l+2)} \, f_k^{2/(k+2)}(x), \\
g(x,y) &= \left( f_{k+1}(x,y) \right)^{l/(k+l+2)} \, f_l^{2/(l+2)}(y), \\
h(x,y) &= \left( f_k^{2/(k+2)}(x) \right)^{2/(k+l+2)} \, f_k^{l/(l+2)}(y).
\end{align*}
\]

We let \( p = (k+l+2)/k \), \( q = (k+l+2)/l \) and \( r = (k+l+2)/2 \).

After simplifying, we find

\[
\sigma_{\beta_{k+1}}^{2(k+1)/(k+l+2)} \leq \left( \sigma_{\beta_k}^{2/(k+l+2)} \right)^{k/(k+l+2)} \left( \sigma_{\beta_l}^{2/(k+l+2)} \right)^{l/(k+l+2)}
\]

which implies the desired result.

The decreasing trends for \( m_k^* \) and \( \beta_k \) indicate that one can't do better than to choose \( k \) large. Indeed, we need \( k \) to be large in order to approach the best possible performance. The following theorem summarizes.

**Theorem:** For a stationary source and large values of \( R \),

\[
\delta(R) \equiv Z(R) = \frac{1}{2\pi e} \sigma^2 \sigma^2 \beta_{\infty} 2^{-2R}
\]

Equivalently,

\[
S(R) \equiv 10 \log_{10} \frac{Z(R)}{\sigma^2} = 6.02 R - 10 \log_{10} \frac{1}{2\pi e} \beta_{\infty}
\]

We again see the 6 dB gain per bit.

**Relationship of kth-order OPTA to overall OPTA**

\[
S(k,R) = 6.02 R - 10 \log_{10} m_k^* \beta_k
\]

\[
= 6.02 R - 10 \log_{10} m_{\infty} \beta_{\infty} - 10 \log_{10} \frac{m_k^*}{m_{\infty}} - 10 \log_{10} \frac{\beta_k}{\beta_{\infty}}
\]

\[
= S(R) - 10 \log_{10} \frac{m_k^*}{m_{\infty}} - 10 \log_{10} \frac{\beta_k}{\beta_{\infty}}
\]

From this we see explicitly how \( S_{vq,k}(R) \) improves with \( k \) through decreases in \( m_k^* \) and \( \beta_k \).
Gauss-Markov Source, $\rho = .9$

$$S(k,R) = S(R) - 10 \log_{10} \frac{m_k}{m_\infty} - 10 \log_{10} \frac{\beta_k}{\beta_\infty}, \quad R = 3$$

IID Gaussian source

$$S(k,R) = S(R) - 10 \log_{10} \frac{m_k^*}{m_\infty^*} - 10 \log_{10} \frac{\beta_k}{\beta_\infty}, \quad R = 3$$
Question:

Why does $\beta_k$ get better as $k$ increases, even for an IID source?

We'll consider this question after discussing some properties of $\beta_\infty$.

Properties of $\beta_\infty$

(1) For a Gaussian source, recall

$$\beta_k = 2\pi \left( \frac{k+2}{k} \right)^{(k+2)/2} \frac{|K^{(k)}|^{1/k}}{\sigma^2}$$

where $K^{(k)}$ is the $k \times k$ by covariance matrix of $(X_1, \ldots, X_k)$. Recall that

$$2\pi \left( \frac{k+2}{k} \right)^{(k+2)/2} \downarrow 2\pi e \text{ as } k \to \infty$$

For a stationary Gaussian source

$$|K^{(k+1)}|^{1/(k+1)} \leq |K^{(k)}|^{1/k} \quad \text{(this will be shown later in the course)}$$

$$|K^{(k)}|^{1/k} \downarrow \frac{Q}{\sigma^2} \quad \text{(will also be shown later)}$$

where

$$Q \equiv \exp \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln S_X(\omega) \, d\omega \right\}$$

= MSE of optimum linear predictor for $X_i$ based on $X_{i-1}, X_{i-2}, \ldots$

= "one-step prediction error"

and $S_X(\omega)$ = power spectral density of $X$.

It follows that

$$\beta_k \downarrow \beta_\infty = \frac{2\pi e Q}{\sigma^2} \text{ as } k \to \infty$$
(2) Upper bound for an arbitrary stationary source

\[ \beta_\infty \leq \frac{2\pi e Q}{\sigma^2} \]

with equality iff \( X \) is Gaussian.

This shows that Gaussian sources have the largest \( Z(R) \) among sources with a given power spectral density or autocorrelation function. In other words, Gaussian sources are the hardest to quantize.

Note, for example, that even though an IID Laplacian source has a larger \( \beta_1 \) it has a smaller \( \beta_\infty \).

Proof: Postponed to variable-rate VQ discussion, for reasons that will be clear then.

(3) \( \beta_k \) tends to be smaller and to decrease more with \( k \) for sources with memory than for memoryless sources. (This is an admittedly rough rule of thumb.)

(4) For a stationary, Gaussian first-order autoregressive source with correlation coefficient \( \rho \), \( Q = 1 - \rho^2 \) and

\[ \beta_k = 2\pi \left( \frac{k+2}{k} \right)^{(k+2)/2} (1-\rho^2)^{(k-1)/k} \rightarrow \beta_\infty = 2\pi e (1-\rho^2) \]

---

**Why Fixed-Rate VQ Outperforms Fixed-Rate SQ**

We want to understand what specific characteristics of vector quantization improve with dimension, and by how much.

We will first compare a fixed-rate k-dimensional VQ, denoted \( Q_k \), to a fixed-rate scalar quantizer \( Q_1 \), both having rate \( R \).

To make it fair, we compare characteristics (point density and inertial profile) of \( Q_k \) to those of the k-dimen'l "product" VQ, denoted \( Q_{pr,k} \), formed by using \( Q_1 \) \( k \) times in succession.

The "product" codebook contains all k-tuples formed from scalar quant. levels.

In the "product" partition, each cell is the Cartesian product of the scalar cells corresponding to the components of its codevector.

The "product" quantization rule is

\[ Q_{pr,k}(x) = ( Q_1(X_1), Q_2(X_2), \ldots, Q_k(X_k) ) \]

It is easy to see that

\[ R(Q_{pr,k}) = R(Q_1) = R \]
\[ D(Q_{pr,k}) = D(Q_1) = \delta(1,R) \]
Assume the scalar quantizer has point density $\lambda_1$.

The cell $S_x$ of the product quantizer containing $x$ is a rectangle:

$$S_x = \frac{1}{2^k \lambda_1(x_1)} \times \frac{1}{2^k \lambda_1(x_2)} \times \frac{1}{2^k \lambda_1(x_3)} \times \ldots \times \frac{1}{2^k \lambda_1(x_k)}$$

with volume

$$|S_x| = \frac{1}{2^{kR}} \times \frac{1}{\lambda_1(x_1) \lambda_1(x_2) \ldots \lambda_1(x_k)}$$

The point density of the product quantizer is

$$\lambda_{pr,k}(x) = \frac{1}{2^{kR} |S_x|} = \frac{1}{\lambda_1(x_1) \lambda_1(x_2) \ldots \lambda_1(x_k)}$$

(It's a product!)

The inertial profile of the product quantizer is

$$m_{pr,k}(x) = \frac{1}{12} \sum_{i=1}^{k} \frac{1}{\lambda_1(x_i)^2} = m^*_i \frac{1}{\left(\prod_{i=1}^{k} \frac{1}{\lambda_1(x_i)^2}\right)^{1/k}}$$

Consider the "loss" of the scalar quantizer (i.e. the product quantizer) relative to the VQ:

$$L = \frac{D(Q_1)}{D(Q_k)} = \frac{D(Q_{pr,k})}{D(Q_k)}$$

Apply Bennett's integral to both terms. Let

$$B(k, m, \lambda, f) = \int \frac{m(x)}{\lambda^{2/k}(x)} f(x) \, dx = \text{Bennett's integral}$$

Then

$$L = \frac{D(Q_1)}{D(Q_k)} = \frac{D(Q_{pr,k})}{D(Q_k)} \equiv \frac{B(k, m_{pr,k}, \lambda_{pr,k}, f)}{B(k, m_k, \lambda_k, f)}$$

$$= \frac{B(k, m_{pr,k}, \lambda_{pr,k}, f)}{B(k, m_k, \lambda_{pr,k}, f)} \times \frac{B(k, m_k, \lambda_{pr,k}, f)}{B(k, m_k, \lambda_k, f)}$$

$$= \text{cell shape loss} \times \text{point density loss}$$

$$= L_{\text{ce}} \times L_{\text{pt}}$$

Now assume $k$ is large and $Q_k$ is optimal, so that

$$m_k(x) \equiv m^*_k = \frac{1}{2\pi e}$$

and

$$\lambda_k(x) \equiv \lambda^*_k(x) = cf^{k/(k+2)}(x) \equiv f(x)$$
Then the loss of the scalar quantizer relative to the optimal high dim'l VQ is

\[
L = \frac{m_1}{m_\infty} \times \frac{\int \frac{1}{k} \sum_{i=1}^{k} \frac{1}{\lambda_1(x_i)} f(x) \, dx}{\int \left(\prod_{i=1}^{k} \lambda_1^2(x_i)\right)^{1/k} f(x) \, dx} \times \frac{B(k, m_1^*, \lambda_{pr,k}, f)}{B(k, m_1^*, \lambda^*_k, f)}
\]

\[
= \text{cubic loss} \times \text{oblongitis loss} \times \text{point density loss}
\]

\[
= L_{cu} \times L_{ob} \times L_{pt}
\]

where we have factored the cell shape loss \(L_{ce}\) into the product of a "cubic loss" \(L_{cu}\) and an "oblongitis loss".

Now assume source is IID.

Consider the choice of \(\lambda_1(x)\) to minimize the loss, i.e. to minimize \(L_{ob} \times L_{pt}\).

- Choosing \(\lambda_1(x)\) to be a constant on the set where \(f_1(x)\) is large makes \(L_{ob} \approx 1\). However, \(L_{pt}\) becomes very large.

- On the other hand choosing \(\lambda_1(x) = f_1(x)\) causes

\[
\lambda_{pr,k}(x) = \prod_{i=1}^{k} f_1(x_i) = f(x) = c f_{k/(k+2)}(x) = \lambda^*_k(x) \quad (k \text{ large } \Rightarrow k/(k+2) \approx 1)
\]

so that \(L_{pt} = 1\).

In other words scalar quantization can produce the optimal point density!

This fact is often overlooked, because for IID Gaussian case, the product quantizer looks like it has a "cubical" point density, when it actually has a spherical one. Unfortunately, however, for this \(\lambda_1\), \(L_{ob} = \infty\).
The point density that minimizes $L_{ob} \times L_{pt}$ is the compromise

$$\lambda_1(x) = c f_1(x)^{1/3}$$

This makes $\lambda_1(x)$ "flatter" (more uniform) in regions where $f_1(x)$ is large than the previous choice of $\lambda_1(x) = f(x)$. Therefore, the rectangular cells are more nearly cubical in the important region where $f_1(x)$ is large, so there is less oblongitis loss.
In summary, for an IID source the shortcomings of scalar quantization relative to high-dimensional VQ are
(a) The cubic loss $L_{cu} = 1.53 \text{ dB}$, which is a measure of its inability to produce cells with smaller nmi than a cube.
(b) The lack of sufficient degrees of freedom to simultaneously attain good inertial profile and good point density.

The oblongitis and point density losses are larger for Laplacian than for the Gaussian density, because the Laplacian’s sharper peak at the origin and heavier tail means that a good scalar quantizer must be more nonuniform. This causes more oblongitis, which in turn causes more compromising of the optimal point density in order to reduce the oblongitis.

The total losses in the right hand column are the potential gains of high dimensional VQ over scalar quantization.
Shortcomings of Optimal k'-dimensional VQ

One can similarly decompose the loss of optimal k'-dimensional VQ (k' \geq 2) relative to high-dimensional VQ into point density, oblongitis, and space-filling losses by comparing the characteristics of an optimal k'-dimensional VQ to that of an optimal high-dimensional VQ. To make the comparison, one considers the point density and inertial profile of the product quantizer formed by using the k'-dimensional VQ many times. The point density and oblongitis losses are then defined in the same way as before. The space-filling loss, which is

\[ L_{sp} = \frac{m_{k'}}{m_{\infty}}, \]

generalizes the "cubic loss" we considered for k'=1. It is called the "space-filling" loss because it represents the loss due to the product quantizers inability to fill space with cells whose NMI is as good as those that induce \( m_{k'}. \) (For k=1 it's a "cubic loss", for k=2 it's a "hexagonal loss".

As k' increases, one finds:

(a) The "space filling" loss decreases to 1 (0 dB).

(b) There are more degrees of freedom, so less compromise is needed between the k'-dimensional point density that minimizes oblongitis and that which minimizes point density loss. Consequently, when optimized, the oblongitis & point density losses are smaller, and decrease to 1 (0 dB).

IID Gaussian Source

\[ L_{ob} = \left(\frac{k+2}{k}\right)^{k/2} e^{-k/(k+2)}, \quad L_{pt} = \frac{k+2}{k} e^{-2/(k+2)}, \quad \frac{L_{pt}}{L_{ob}} = \frac{L_{shape}}{L_{shape_\infty}} = \frac{\beta_k}{\beta_{\infty}} = \frac{1}{e} \left(\frac{k+2}{k}\right)^{(k+2)/2} \]

![Graph showing the losses L_{ob}, L_{pt}, L_{shape} vs dimension k]
Memory Loss

For sources with memory (e.g. autoregressive), scalar and low-dimensional quantization suffer an additional loss, namely the inability to exploit or fully exploit the memory, i.e. the dependence between source samples.

Both oblongitis and point density losses can be factored into two terms, one of which expresses the quantizers inability to fully exploit the source memory.