High-Resolution Analysis of Quantizer Distortion

For fixed-rate, memoryless VQ, there are two principal results of high-resolution analysis:

Bennett's Integral

A formula for the mean-squared error distortion of a "high-resolution" VQ in terms of its "gross" characteristics.

A "high-resolution" VQ is one with "small" cells, so it has "small" distortion and, usually, "many" cells, and "large" rate. Later we'll see roughly how "small", how "many" and how "large" are adequate.

Question: What "gross" characteristics" distinguish different high-resolution quantizers?

Zador's formula

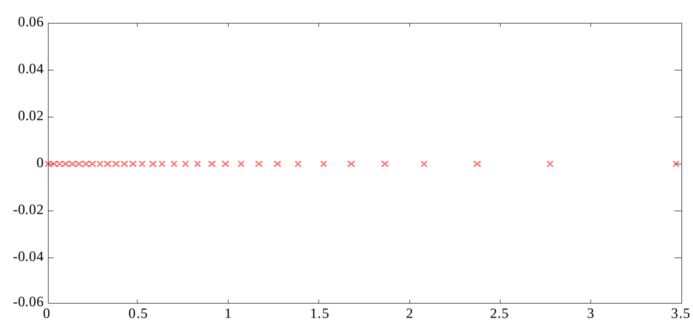
An approximation to the OPTA function $\delta(k,R)$. It applies when R is large.

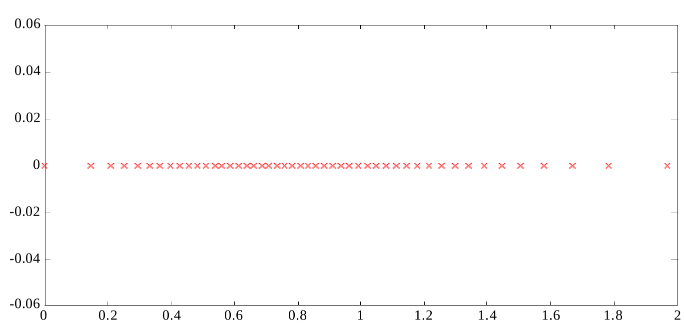
Question: We'd much rather have small rate, than large rate. Why should we be interested in high-resolution formulas, which apply when rate is larger?

Answer: We'll see that these formulas are accurate when $R \geq 3$, and accurate enough even when $R \cong 2$ to provide excellent insight.

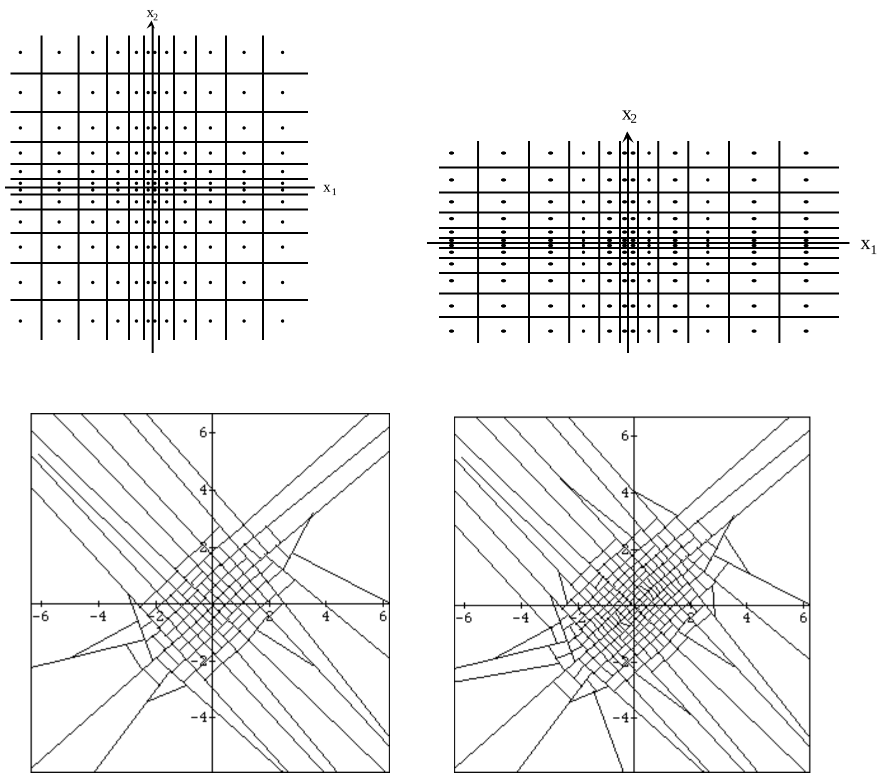
Examples of High-Resolution Quantizers:

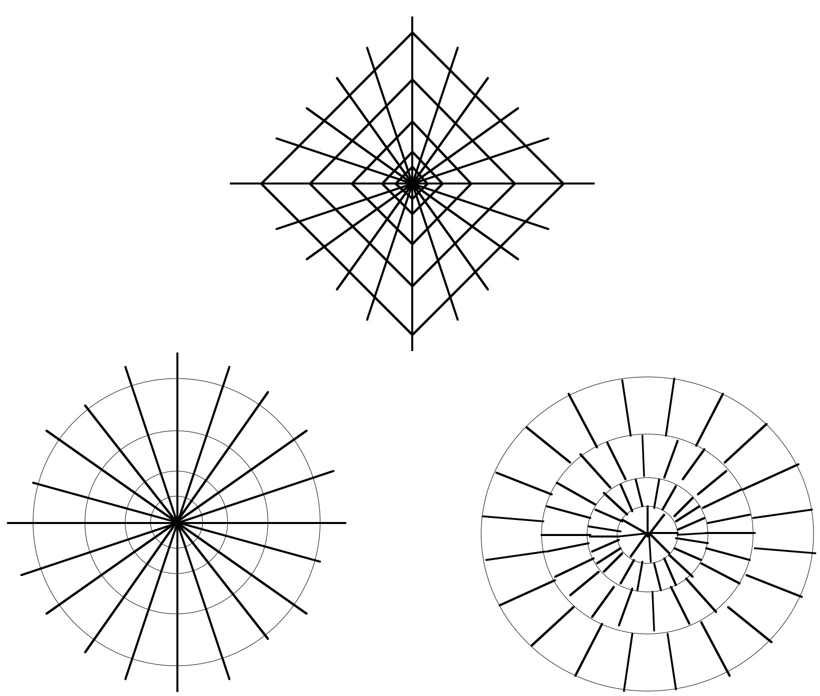
k=1 (scalar quantizers)

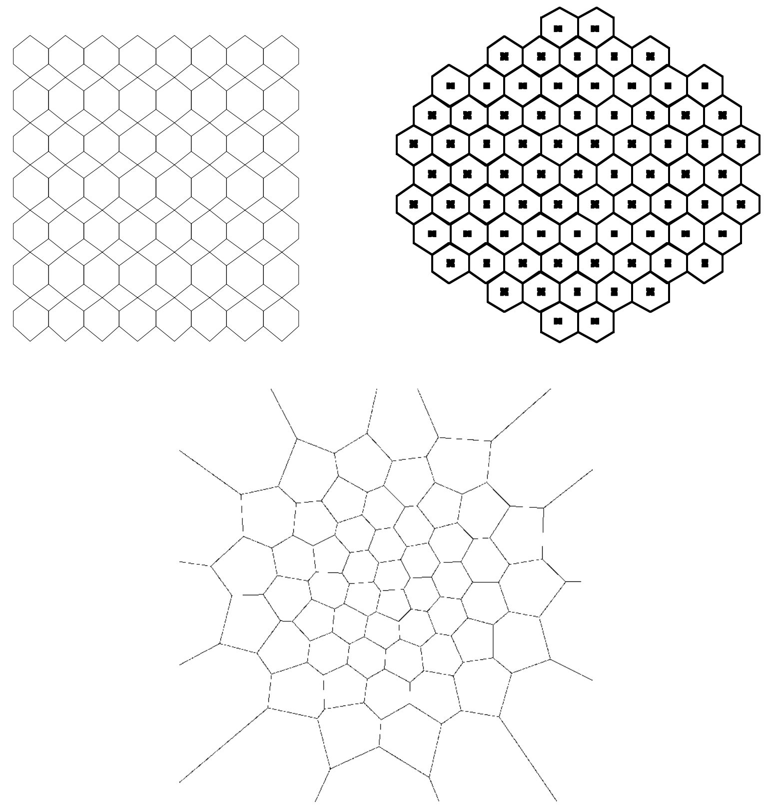




k=2







The Key Gross Characteristics of a High-Resolution Quantizer:

- Dimension k
- Size M
- The distribution or density of points/cells over $\mathcal{R}^k$
- Some way of characterizing the shapes of the cells as a function of location in $\mathcal{R}^k$.

Case 1: Bennett's Integral for Quantizers with Congruent Cells

Theorem

The MSE of a k-dimensional VQ with size M and "small" congruent cells, or at least "most" cells must be approximately congruent, is approximately

$$D \cong \frac{1}{M^{2/k}}\, m \int \frac{1}{\lambda^{2/k}(\underline{x})}\, f_{\underline{X}}(\underline{x})\, d\underline{x}$$

where

m = quantity, to be defined later, that depends only on the shape of the cells, but not on their size, nor on the source density

$\lambda(\underline{x})$ = "quantization density function". This is a function that characterizes, approximately, the density of cells/points in the vicinity of $\underline{x}$. It is also called the "point density" or "cell density" function.

The formula above is called "Bennett's integral". It was derived by Bennett[1] for scalar quantizers, and then extended to VQ's with congruent cells by Gersho[2].

Our derivation comes later. Examples of VQ's with congruent cells can be found on pages 4 and 5. Later we drop the congruent cell restriction.

"mostly" means: the contribution to distortion due to cells that are not approximately congruent is small. Among other things, this requires the overload region to have small probability.

---

[1] W. Bennett, "Spectra of quantized signals," *Bell Syst. Tech. J.*, vol. 27, pp. 446-472, July 1948.
[2] A. Gersho, "Asymptotically optimal block quantizers," *IEEE Trans. Inform. Thy.*, vol. 25, pp. 373-380, 1979.

Quantization Density $\lambda(\underline{x})$

The quantization density is a function $\lambda(\underline{x})$ that characterizes the distribution of quantization cells and points, by having the following properties:

1. $\int_A \lambda(\underline{x})\, d\underline{x} \cong$ fraction of codevectors (or cells) in region A
2. $\lambda(\underline{x}) \ge 0, \quad \int \lambda(\underline{x})\, d\underline{x} = 1$
3. if A is small, then $\lambda(\underline{x})\, |A| \cong \frac{\#\text{ cells/points in } A}{M}$, $\quad (|A| \triangleq \text{vol of } A)$
   (assuming A is much larger than cells in the vicinity of $\underline{x}$)
4. $\lambda(\underline{x})$ is a smooth or piecewise smooth function.
5. $\lambda(\underline{x}) \cong \frac{1}{M|S_i|}$ when $\underline{x} \in S_i$ or equivalently, $|S_i| \cong \frac{1}{M\lambda(\underline{x})}$

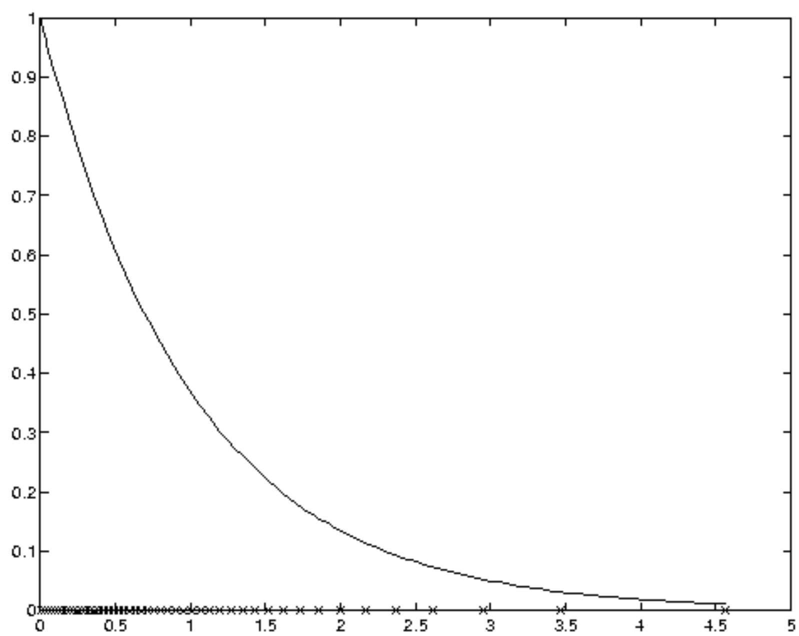
Why 5? In a small region A containing $\underline{x}$, one expects most cells to have approximately the same volume. Therefore,

$$\#\text{ cells in } A \cong \frac{|A|}{\text{cell vol}}, \quad \text{and by 3,} \quad \lambda(\underline{x})\, |A| \cong \frac{\#\text{ cells in } A}{M} = \frac{|A|/(\text{cell vol})}{M}$$

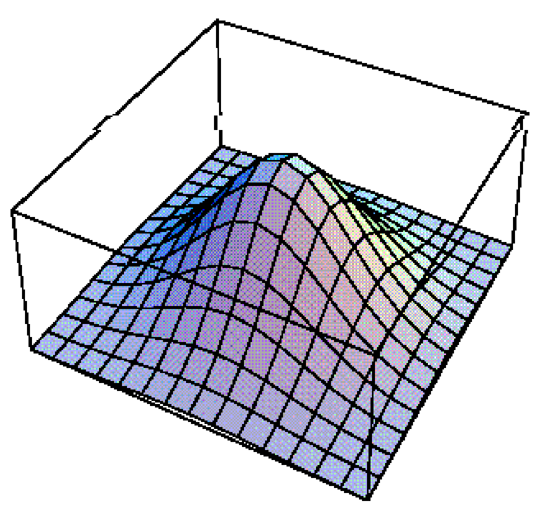
$$\Rightarrow \quad \lambda(\underline{x}) \cong \frac{1}{M|S_i|}$$

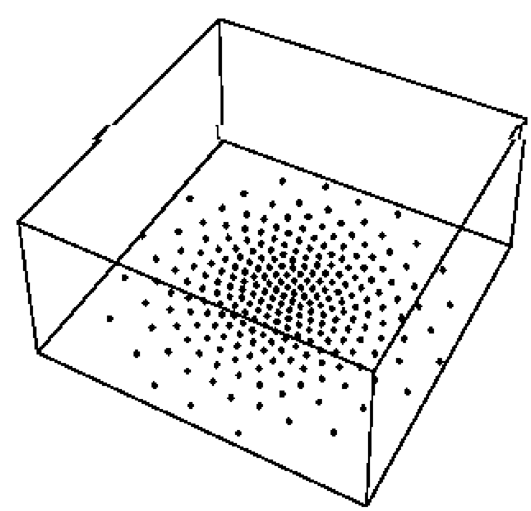
Examples of Point Densities

k=1



Optimal Quantizer for a 2-dimensional IID Gaussian random vector and its point density





Derivation of Bennett's Integral

$$D = \frac{1}{k} E\,||\underline{X}-Q(\underline{X})||^2 = \frac{1}{k}\sum_{i=1}^{M}\int_{S_i} ||\underline{x}-\underline{w}_i||^2 f_{\underline{X}}(\underline{x})\,d\underline{x}$$

$$\cong \frac{1}{k}\sum_{i=1}^{M} f_{\underline{X}}(\underline{w}_i)\int_{S_i} ||\underline{x}-\underline{w}_i||^2 d\underline{x} \quad \text{because } f_{\underline{X}}(\underline{x}) \cong f_{\underline{X}}(\underline{w}_i) \text{ when } \underline{x}\in S_i$$

$f_{\underline{X}}(\underline{x}) \cong f_{\underline{X}}(\underline{w}_i)$ when $\underline{x}\in S_i$ is a valid approximation because most S_i's are small, and because $f_{\underline{X}}(\underline{x})$ is ordinarily either continuous, in which case it changes little or over S_i and can be approximated by its value at $\underline{w}_i$, or piecewise continuous, in which case it is continuous on most cells and the previous argument applies.

$$= \sum_{i=1}^{M} f_{\underline{X}}(\underline{w}_i)\,\frac{1}{k}\,\mathcal{M}(S_i,\underline{w}_i)$$

where $\mathcal{M}(S,\underline{w}) \triangleq \int_S ||\underline{x}-\underline{w}||^2 d\underline{x}$ = "moment of inertia" (mi) of S about $\underline{w}$.

For future reference, we note that the key approximation made here is

$$\int_{S_i} ||\underline{x}-\underline{w}_i||^2 f_{\underline{X}}(\underline{x})\,d\underline{x} \cong f_{\underline{X}}(\underline{w}_i)\int_{S_i} ||\underline{x}-\underline{w}_i||^2 d\underline{x}$$

Normalized Moment of Inertia (nmi)

Let us now separate the effect on $\mathcal{M}(S,\underline{w})$ of the shape of S from its size.

$$\frac{1}{k}\mathcal{M}(S,\underline{w}) = \frac{\int_S ||\underline{x}-\underline{w}||^2 d\,\underline{x}}{k\,|S|^{1+2/k}} \times |S|^{1+2/k}, \qquad \text{where } |S| = \text{vol of } S = \int_S 1\,d\underline{x}$$

$$= m(S,\underline{w}) \times |S|^{1+2/k}$$

where

$$m(S,\underline{w}) = \frac{\int_S ||\underline{x}-\underline{w}||^2 d\,\underline{x}}{k\,|S|^{1+2/k}}$$

= "normalized moment of inertia" (nmi) of S about $\underline{w}$.

Fact: $m(S,\underline{w})$ is not affected by scaling nor translation. Thus, it is determined only by its shape, but not its size or position.

Proof: First, consider scaling by a factor $a>0$: $S \to aS = \{\underline{z} = a\underline{x} : \underline{x} \in S\}$; $\underline{w} \to a\underline{w}$

$$m(aS,a\underline{w}) = \frac{\int_{aS} ||\underline{x}-a\underline{w}||^2 \, d\underline{x}}{k|aS|^{1+2/k}} = \frac{\int_S ||a\underline{z}-a\underline{w}||^2 a^k d\underline{z}}{k|aS|^{1+2/k}} \quad \text{where } a\underline{z} = \underline{x},\ a^k d\underline{z} = d\underline{x}$$

$$= \frac{\int_S a^2 ||\underline{z}-\underline{w}||^2 a^k d\underline{z}}{k|S|^{1+2/k}(a^k)^{1+2/k}} \quad \text{because } |aS| = a^k |S|$$

$$= \frac{\int_S ||\underline{z}-\underline{w}||^2 \, d\underline{z}}{k|S|^{1+2/k}} = m(S,\underline{w})$$

Next, consider translating by a vector $\underline{v}$: $S \to S+\underline{v} = \{\underline{z} = \underline{x}+\underline{v} : \underline{x} \in S\}$; $\underline{w} \to \underline{w}+\underline{v}$

$$m(S+\underline{v},\underline{w}+\underline{v}) = \frac{\int_{S+\underline{v}} ||\underline{x}-\underline{w}-\underline{v}||^2 \, d\underline{x}}{k|S+\underline{v}|^{1+2/k}} = \frac{\int_S ||\underline{z}-\underline{w}||^2 d\underline{z}}{k|S|^{1+2/k}} \quad \text{where } \underline{z} = \underline{x}-\underline{v},\ d\underline{z} = d\underline{x}$$

$$= m(S,\underline{w})$$

Completion of the Derivation of Bennett's Integral

$$D \cong \sum_{i=1}^{M} f_{\underline{X}}(\underline{w}_i) \frac{1}{k} \mathcal{M}(S_i,\underline{w}_i) \qquad \text{derived earlier}$$

$$= \sum_{i=1}^{M} f_{\underline{X}}(\underline{w}_i)\, m(S_i,\underline{w}_i)\, |S_i|^{1+2/k} \qquad \text{by the definition of nmi}$$

$$= m(S_o,\underline{w}_o) \sum_{i=1}^{M} f_{\underline{X}}(\underline{w}_i)\, |S_i|^{1+2/k} \qquad \text{since all } S_i\text{'s are congruent to } S_o$$

Note: For the last equality to hold, each $\underline{w}_i$ must be in the same relative position within S_i that $\underline{w}_o$ is in S_o. We define "congruence" so as to imply this.

Note: We can now see how cell sizes and shape separately affect distortion.

Now recall: $|S_i| \cong \frac{1}{M\lambda(\underline{w}_i)}$, which implies $|S_i|^{1+2/k} = \frac{1}{M^{2/k}} \frac{1}{\lambda^{2/k}(\underline{w}_i)} |S_i|$

Therefore,

$$D \cong \frac{1}{M^{2/k}} m(S_o,\underline{w}_o) \sum_{i=1}^{M} f_{\underline{X}}(\underline{w}_i) \frac{1}{\lambda^{2/k}(\underline{w}_i)} |S_i|$$

$$\cong \frac{1}{M^{2/k}} m(S_o,\underline{w}_o) \int f_{\underline{X}}(\underline{x}) \frac{1}{\lambda^{2/k}(\underline{x})} d\underline{x} \quad = \text{ Bennett's integral}$$

where the last "$\cong$" is by the definition of an integral and the fact that most $|S_i|$'s are small.

Special case: Scalar quantization (k=1)

Cells are intervals.

If codepoints (levels) are in the centers of the cells, then it is easy to show that

$$m(\text{interval}) = \frac{1}{12}$$

and

$$D \cong \frac{1}{12M^2} \int_{-\infty}^{\infty} \frac{f_X(x)}{\lambda^2(x)}\, dx$$

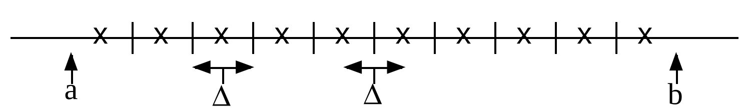
This is what Bennett originally derived.

Special case: Uniform scalar quantizer.

Consider a quantizer that is *uniform* over the interval [a,b], in the sense that

(a) the partition divides [a,b] into M cells of width $\Delta = \frac{b-a}{M}$

(c) the codepoints (levels) are in the centers of the cells (i.e. they are uniformly spaced, Δ apart)



Δ is called the "quantizer stepsize".

Suppose Δ is small. Then

$$\lambda(x) \cong \frac{1}{M\Delta} = \frac{1}{b-a}$$

Suppose also that $\Pr(a \le X \le b) \cong 1$. Then

$$D \cong \frac{1}{12M^2} \int_a^b \frac{f_X(x)}{1/(b-a)^2}\, dx \cong \frac{(b-a)^2}{12M^2} \int_a^b f_X(x)\, dx \cong \frac{\Delta^2}{12}$$

This is a formula worth remembering.

Bennett's Integral for Vector Quantizers -- General Case

Theorem[3]: Under the *high-resolution conditions* stated below, the MSE distortion of a k-dimensional VQ with size M applied to random vector $\underline{X}$ with pdf $f_{\underline{X}}(\underline{x})$ can be approximated by

$$D \cong \frac{1}{M^{2/k}} \int \frac{m(\underline{x})}{\lambda^{2/k}(\underline{x})} f_{\underline{X}}(\underline{x})\, d\underline{x}$$

High-resolution conditions:

+ Most cells are small enough that the prob. density can be approximated as being constant on each. (Union of cells for which prob. density cannot be so approximated has very small probability. The overload distortion is negligible.)
+ Neighboring cells have similar sizes and shapes, i.e. cell size & shape change slowly, if at all, with $\underline{x}$. ("similar shape" also refers to placement of codevectors within cells.)
+ The quantization density is approximately $\lambda(\underline{x})$.
+ The *inertial profile* is approximately $m(\underline{x})$, where the inertial profile of a VQ is a function such that

$$m(\underline{x}) \cong \text{NMI of cell containing } \underline{x} = m(S_i,\underline{w}_i) \text{ if } \underline{x} \in S_i$$

We also require $m(\underline{x}) \geq 0$, all $\underline{x}$.

[3] S. Na and D. Neuhoff, "Bennett's integral for vector quantizers," *IEEE Trans. Inform. Thy.*, vol. 41, pp. 886-900, July 1995.

Derivation of Bennett's Integral -- General Case

$$D = \frac{1}{k} E\, ||\underline{X}-Q(\underline{X})||^2$$

$$= \frac{1}{k} \sum_{i=1}^{M} \int_{S_i} ||\underline{x}-\underline{w}_i||^2 f_{\underline{X}}(\underline{x})\, d\underline{x}$$

$$\cong \frac{1}{k} \sum_{i=1}^{M} f_{\underline{X}}(\underline{w}_i) \int_{S_i} ||\underline{x}-\underline{w}_i||^2 d\underline{x}$$ by same argument as in congruent case

$$= \sum_{i=1}^{M} f_{\underline{X}}(\underline{w}_i) \frac{1}{k} \mathcal{M}(S_i,\underline{w}_i)$$ $\mathcal{M}(S_i,\underline{w}_i)$ is the MI of S_i about $\underline{w}_i$

$$= \sum_{i=1}^{M} f_{\underline{X}}(\underline{w}_i)\, m(S_i,\underline{w}_i)\, |S_i|^{1+2/k}$$ recall: $\mathcal{M}(S_i,\underline{w}_i) = m(S_i,\underline{w}_i)\, k\, |S_i|^{1+2/k}$

$$\cong \frac{1}{M^{2/k}} \sum_{i=1}^{M} f_{\underline{X}}(\underline{w}_i) \frac{m(\underline{w}_i)}{\lambda^{2/k}(\underline{w}_i)} |S_i|$$ recall: $|S_i| \cong \frac{1}{M\lambda(\underline{w}_i)}$, $m(\underline{w}_i) \cong m(S_i,\underline{w}_i)$

$$\cong \frac{1}{M^{2/k}} \int \frac{m(\underline{x})}{\lambda^{2/k}(\underline{x})} f_{\underline{X}}(\underline{x})\, d\underline{x}$$ by the definition of an integral

Notes on Bennett's Integral

- Bennett's integral identifies point density and inertial profile as key characteristics of VQ's, in addition to k and M.
- When M is large, both left and righthand sides of the Bennett integral relation are approximately zero; so when one does a careful derivation, what really needs to be show (and what has been shown) is

$$\frac{D}{\frac{1}{M^{2/k}} \int \frac{m(\underline{x})}{\lambda^{2/k}(\underline{x})} f_{\underline{X}}(\underline{x})\, d\underline{x}} \cong 1$$

- Bennett's integral shows that distortion decreases as $\frac{1}{M^{2/k}}$, assuming point density and inertial profile stay the same.

  To see that this makes sense, consider what happens when M doubles, while maintaining the same point density function and the same inertial profile. Doubling M cuts the volumes of cells in a given region in half. Assuming the cell shapes stay the same, this decreases the linear dimensions of the cells in any region by the factor $1/2^{1/k}$, and causes the average squared distance between points $\underline{x}$ and the codevector of the cell in which $\underline{x}$ lies to decrease on the average by $1/2^{2/k}$. This suggests that mean-squared error decreases as $1/M^{2/k}$.

- Equivalently, SNR increases 6 dB for each one bit increase of rate.

$$\frac{1}{M^{2/k}} = 2^{-2R} \quad \Rightarrow \quad D \cong 2^{-2R} \int \frac{m(\underline{x})}{\lambda^{2/k}(\underline{x})} f_{\underline{X}}(\underline{x})\, d\underline{x}$$

$$\Rightarrow \text{SNR} = 10 \log_{10} \frac{\sigma^2}{D} = 10 \log_{10} 2^{2R} + 10 \log_{10} \frac{\sigma^2}{\int m(\underline{x})\, \lambda^{-2/k}(\underline{x})\, f_{\underline{X}}(\underline{x})\, d\underline{x}}$$

$$= 6.02\, R + 10 \log_{10} \frac{\sigma^2}{\int m(\underline{x})\, \lambda^{-2/k}(\underline{x})\, f_{\underline{X}}(\underline{x})\, d\underline{x}}$$

This is called the "6 dB per bit rule".

- Usually, we don't employ a quantization density or inertial profile to describe a VQ unless most cells are small, where "small" means that the probability density changes little across the cell and "most" means that the probability of the cells that are small is large.
- Usually, λ and m are fairly smooth functions that do not convey the detailed locations of codevectors and cells. VQ's with the same quantization density can differ in the number of points, in the exact placement of codepoints and in the shapes of the cells. VQ's with the same inertial profile can differ in the number and placement of codepoints
- Quantization density and inertial profile are, generally, idealizations or models. Often we pick a target quantization density and/or target inertial profile and try to make our quantizer approximate it.

  Later we'll find what quantization densities and inertial profiles are desirable.

- We don't use the following as definitions because if we did, a quantizer would almost never have a specified quantization density or inertial profile:

$$\lambda(\underline{x}) = \frac{1}{M|S_i|} \text{ when } \underline{x} \in S_i \;, \quad m(\underline{x}) = m(S_i,\underline{w}_i) \text{ if } \underline{x} \in S_i$$

- Sketch of why Property 5 on p. 8 implies Property `

$$\int_A \lambda(\underline{x})\, d\underline{x} = \sum_{i=1}^{M} \int_{S_i \cap A} \lambda(\underline{x})\, d\underline{x} \cong \sum_{i=1}^{M} \int_{S_i \cap A} \frac{1}{M\, vol(S_i)}\, d\underline{x}$$

$$= \frac{1}{M} \sum_{i=1}^{M} \frac{vol(S_i \cap A)}{vol(S_i)} \cong \frac{\#\,\text{cells in } A}{M}$$

because $\frac{vol(S_i \cap A)}{vol(S_i)} = \begin{cases} 1 \;, & \text{if } S_i \subset A \\ 0 \;, & \text{if } S_i \cap A = \phi \\ <1 \;, & \text{otherwise} \end{cases}$, and because most S_i's are small

- When, as usual, $\lambda(x)$ is smooth, Prop. 5, p. 8 implies neighboring cells have similar sizes; e.g. it rules out quantizers with alternating large and small cells.
- When, as usual, $m(x)$ is smooth, the defining property of m implies neighboring cells mostly have similar NMI; e.g. it rules out quantizers whose cell shapes change rapidly.
- We see from Bennett's integral that to make D small, we want larger quantization density where $f_{\underline{X}}(\underline{x})$ is larger. Small inertial profile is desired, everywhere.

Properties and Examples of Normalized Moment of Inertia (NMI)

- **Definition:** NMI of S about point $\underline{w}$ is

$$m(S) = m(S,\underline{w}) = \frac{1}{k}\,\frac{\int_S \|\underline{x}-\underline{w}\|^2 \, d\underline{x}}{\mathrm{vol}^{1+2/k}(S)}$$

- NMI of a k-dimensional cube is 1/12, for all k.

This is why the definition of $m(S)$ includes $1/k$.

Defn: k-dimen'l cube = $\{\underline{x} : a \le x_i \le b, i = 1,\ldots,k\}$ for some $a < b$

Proof: Let $S = \{\underline{x} : -1/2 \le x_i \le 1/2, i = 1,\ldots,k\}$ and $\underline{w} = (0,\ldots,0)$. Then $\mathrm{vol}(S) = 1$ and

$$m(S) = \frac{1}{k}\,\frac{1}{\mathrm{vol}(S)^{1+2/k}} \int_{-1/2}^{1/2}\cdots\int_{-1/2}^{1/2} \sum_{i=1}^{k} x_i^2 \, dx_1 \ldots dx_k$$

$$= \frac{1}{k} \int_{-1/2}^{1/2}\cdots\int_{-1/2}^{1/2} \left(\sum_{i=2}^{k} x_i^2 + \frac{x_1^3}{3}\Big|_{-1/2}^{1/2} \right) dx_2 \ldots dx_k$$

$$= \frac{1}{k} \int_{-1/2}^{1/2}\cdots\int_{-1/2}^{1/2} \left(\sum_{i=2}^{k} x_i^2 + \frac{1}{12} \right) dx_2 \ldots dx_k$$

$$= \frac{1}{k} \int_{-1/2}^{1/2}\cdots\int_{-1/2}^{1/2} \left(\sum_{i=3}^{k} x_i^2 + \frac{x_2^3}{3}\Big|_{-1/2}^{1/2} + \frac{1}{12} \right) dx_3 \ldots dx_k$$

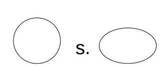
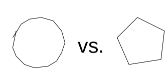
$$= \frac{1}{k} \int_{-1/2}^{1/2}\cdots\int_{-1/2}^{1/2} \left(\sum_{i=4}^{k} x_i^2 + \frac{x_3^3}{3}\Big|_{-1/2}^{1/2} + \frac{1}{12} + \frac{1}{12} \right) dx_4 \ldots dx_k$$

$$= \frac{1}{k} \sum_{i=1}^{k} \frac{1}{12} = \frac{1}{12}$$

- The NMI of various cell shapes

cell shape	dimension	NMI
1×2 rectangle	2	.104
cube	any	$.0833 = \frac{1}{12}$
hexagon	2	$.0802 = \frac{5\sqrt{3}}{108}$
circle	2	$.0796 = \frac{1}{4\pi}$
sphere	3	$.0770 = \frac{(4\pi/3)^{-2/3}}{5}$
sphere	k	$\frac{1}{(k+2)(V_k)^{2/k}}$
sphere	∞	$.0585 = \frac{1}{2\pi e} \cong \frac{1}{17}$
$s_1 \times s_2 \times \ldots \times s_k$ rectangle	k	$\frac{1}{12} \frac{\frac{1}{k}\sum_{i=1}^{k}(s_i)^2}{\left(\prod_{i=1}^{k}(s_i)^2\right)^{1/k}} = \frac{1}{12} \frac{\text{arith mean of sides}^2}{\text{geom mean of sides}^2}$

where V_k = volume of k-dimensional sphere with radius 1

- Shapes that tend to make NMI smaller
  - + Spheroidal rather than oblong 
  - + More finely faceted (many sides rather than few) 
  - + Higher rather than lower dimension
- A sphere has the lowest NMI of any cell of a given dimension.
- NMI of a sphere decreases with dimension to the limit $1/2\pi e = .0585$

Volume of k-dimensional sphere

(from Wozencraft & Jacobs, p. 357)

$$V_k = \begin{matrix}\text{vol. of k-dimensional} \\ \text{sphere with radius 1}\end{matrix} = \begin{cases} \dfrac{\pi^{k/2}}{(k/2)!}, & k \text{ even} \\ \dfrac{2^k \pi^{(k-1)/2} (\frac{k-1}{2})!}{k!}, & k \text{ odd} \end{cases}$$

From Stirling's approximation,

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{\varepsilon_N} \cong \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \quad \text{where } 0 < \varepsilon_n < \frac{1}{12n},$$

which somewhat underestimates n!, one can show

$$V_k \cong \frac{N}{\sqrt{\pi k}}\left(\frac{2\pi e}{k}\right)^{k/2}.$$

From the above,

$$V_k \cong \exp\left\{\frac{k}{2}\left(\ln(2\pi e) - \ln(k)\right)\right\}$$

$$\to 0 \text{ as } k \to \infty \quad \text{because } -k \ln k \to -\infty$$

It follows that

$$m(\text{k-dim'l sphere}) = \frac{1}{(k+2)(V_k)^{2/k}} \cong \frac{1}{k+2}\left(\frac{1}{\sqrt{\pi k}}\left(\frac{2\pi e}{k}\right)^{k/2}\right)^{-2/k}$$

$$= (\pi k)^{1/k}\frac{k}{k+2}\frac{1}{2\pi e} = \exp\{\frac{1}{k}\ln(\pi k)\}\frac{k}{k+2}\frac{1}{2\pi e}$$

$$\to \frac{1}{2\pi e} \text{ as } k \to \infty.$$

Formal Statement of the Validity of Bennett's Integral

Theorem[4]: $\lim_{M\to\infty} M^{2/k} D(Q_M) = \int \frac{m(\underline{x})}{\lambda^{2/k}(\underline{x})} f_{\underline{X}}(\underline{x})\, d\underline{x}$

Assuming

+ $Q_1, Q_2, \ldots$ is sequence of k-dim'l VQ's, with Q_M having size M, partition S_M
+ $\lambda_M(\underline{x}) \to \lambda(\underline{x})$ in probability as $M \to \infty$

  where $\lambda_M(\underline{x}) \stackrel{\Delta}{=} \frac{1}{M\ \text{vol(cell containing } \underline{x})}$ = *specific point denstiy* of Q_M
+ $m_M(\underline{x}) \to m(\underline{x})$ as $M \to \infty$ in prob.

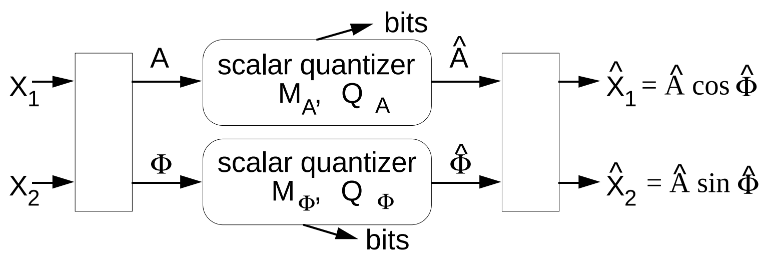
  where $m_M(\underline{x}) \stackrel{\Delta}{=} m(S_{\underline{x}}, \underline{y}_{\underline{x}})$ = *specific inertial profile* of Q_M
+ $\{M^{2/k}||\underline{x}-Q_M(\underline{x})||^2\}$ has *uniformly absolutely continous integrals*
+ diam(cell of S_M containing $\underline{X}$) $\to 0$ in prob.
+ $f_{\underline{X}}(\underline{x})$ is piecewise continuous
+ Bennett's integral is finite

"Sequence approach" first used by Bucklew & Wise (1982), for scalar quantizers.

[4]S. Na and D. Neuhoff, "Bennett's integral for vector quantizers," *IEEE Trans. Inform. Thy.*, vol. 41, pp. 886-900, July 1995.

Example: Polar Quantization

Quantize $\underline{X} = (X_1, X_2)$ by independently scalar quantizing its magnitude and phase:



Magnitude:

$$A = ||\underline{X}|| = \sqrt{X_1^2 + X_2^2} = \text{magnitude/amplitude}, \quad A \geq 0$$

Phase:

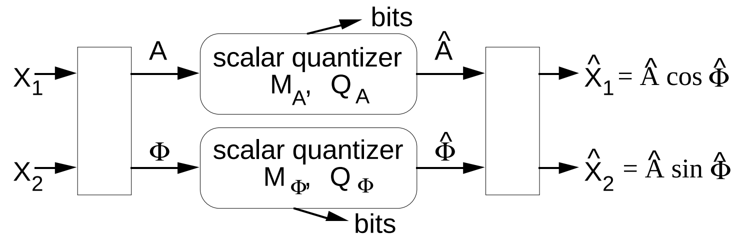
$$\Phi = \angle \underline{X} = \text{angle or phase of } \underline{X} = (X_1, X_2), \quad 0 \leq \Phi < 2\pi$$

$$= \begin{cases} \tan^{-1}\left(\frac{x_2}{x_1}\right), & x_2 \geq 0 \\ \tan^{-1}\left(\frac{x_2}{x_1}\right) + \pi, & x_2 < 0 \end{cases}, \qquad (\text{assuming } 0 \leq \tan^{-1} z < \pi)$$

Magnitude quantizer: $\hat{A} = Q_A(A)$, $C_A = \{w_1, \ldots, w_{M_A}\}$, $S_A = \{S_1, \ldots, S_{M_A}\}$

Phase quantizer: $\hat{\Phi} = Q_\Phi(\Phi)$, $C_\Phi = \{v_1, \ldots, v_{M_\Delta}\}$, $T_\Phi = \{T_1, \ldots, T_{M_\Phi}\}$

Polar Quantization is a kind of 2-dimensional VQ.



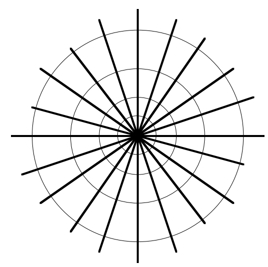
Dimension: $k = 2$

Size: $M = M_A \times M_\Phi$

Codebook: $C = \{ \underline{w}_{i,j} \}$, where $\underline{w}_{i,j} = (w_i \cos v_j, w_i \sin v_j)$, $i = 1,...,M_A$, $j = 1,...,M_\Phi$

Partition: $S = \{ S_{i,j} \}$, where $S_{i,j} = \{ \underline{x} : ||\underline{x}|| \in S_i,\ \angle\underline{X} \in T_j \}$, $i = 1,...,M_A$, $j = 1,...,M_\Phi \}$

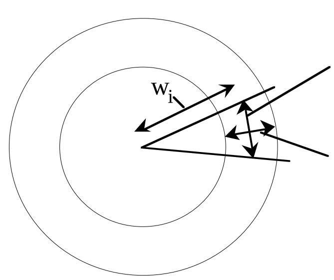
Quantization rule: $Q(\underline{x}) = \left(Q_A(||\underline{x}||\cos Q_\Phi(\angle\underline{x}),\ Q_A(||\underline{x}||\sin Q_\Phi(\angle\underline{x}) \right)$



Distortion analysis via Bennett's integral

Assumptions: M_A, M_Φ are large,
magnitude quantizer has point density $\lambda_A(a)$,
phase quantizer is uniform with step size $\Delta = \frac{2\pi}{M_\Phi}$

Consequence: cells of polar quant cells are, approximately, small rectangles:



$$W \cong \frac{2\pi w_i}{M_\Phi}$$

$$H \cong \frac{1}{M_A \lambda_A(w_i)}$$

Volume of cell near $\underline{x}$

$$HW \cong \frac{2\pi||\underline{x}||}{M_A M_\Phi \lambda_A(||\underline{x}||)}$$

Point density of the VQ: $\lambda(\underline{x})$ = \f(1,M H W) = $\frac{\lambda_A(||\underline{x}||)}{2\pi||\underline{x}||}$

Inertial profile of the VQ:

$$m(\underline{x}) = \frac{1}{12}\frac{\frac{1}{2}(H^2+W^2)}{\sqrt{H^2W^2}} = \frac{1}{24}\left(\frac{M_\Phi}{2\pi||\underline{x}||M_A\lambda_A(||\underline{x}||)} + \frac{2\pi||\underline{x}||M_A\lambda_A(||\underline{x}||)}{M_\Phi}\right)$$

Substituting: $\lambda(\underline{x})$ and $m(\underline{x})$ into Bennett's integral and simplifying gives:

$$D \cong \frac{1}{M}\int \frac{m(\underline{x})}{\lambda(\underline{x})} f_{\underline{X}}(\underline{x})\, d\underline{x}$$

$$= \frac{1}{M}\frac{1}{24}\left(\frac{\sqrt{M}}{M_A}\right)^2 \int_0^\infty \frac{1}{\lambda_A(a)^2} f_A(a)\,da + \frac{1}{M}\frac{\pi^2}{6}\left(\frac{M_A}{\sqrt{M}}\right)^2 \int_0^\infty a^2 f_A(a)\,da$$

$$= \frac{1}{M}\frac{1}{L^2}\int_0^\infty \frac{1}{\lambda_A(a)^2} f_A(a)\,da + \frac{1}{M}\frac{\pi^2}{6} L^2 \int_0^\infty a^2 f_A(a)\,da$$

where

$$L \triangleq \frac{M_A}{\sqrt{M}} = \sqrt{\frac{M_A}{M_\Phi}} = \text{magnitude level allocation}$$

(Recall: $M = M_A \times M_\Phi$)

Notice that D does not depend on the probability distribution of the angle. This is because the angle quantizer is uniform. So it does not favor some angles over others.

Optimizing Polar Quantization

Given M (large), choose L and λ_A (the key characteristics) to minimize distortion

$$D \cong \frac{1}{M}\frac{1}{L^2}\int_0^\infty \frac{1}{\lambda_A(a)^2} f_A(a)\,da + \frac{1}{M}\frac{\pi^2}{6} L^2 \int_0^\infty a^2 f_A(a)\,da$$

Approach 1: For given choice of λ_A, find best L by equating to zero the derivative of D with respect to L. Then find best λ_A by calculus of variations.

Approach 2: For given choice of L, find best λ_A by calculus of variations or Holder's inequality. Then find best L by equating to zero the derivative wrt L of the resulting expression for distortion.

The result:

$$\lambda_A(a) = c\, p_A(a)^{1/3} \qquad (c \text{ chosen to make } \lambda_A(a) \text{ integrate to one})$$

$$L^2 = \frac{M_A}{M_\Phi} = \frac{1}{2\pi}\left(EA^2 \int_0^\infty f_A^{1/3}(a)\,da\right)^{-1/2}$$

$$D \cong \frac{1}{12} 2\pi \sqrt{E\ A^2}\left(\int_0^\infty f_A^{1/3}(a)\,da\right)^{3/2}\frac{1}{M}$$

We'll review calculus of variations shortly.

Prime Example: X_1, X_2 IID Gaussian

$f_{\underline{X}}(\underline{x}) = \frac{1}{2\pi} e^{-\|\underline{x}\|^2/2}$ and $f_A(a) = a\, e^{-a^2/2}, \; a \ge 0$ (Rayleigh density)

$$EA = \sqrt{\pi/2}, \quad EA^2 = E(X_1^2+X_2^2) = 2\,EX^2 = 2, \quad \sigma_A^2 = 2 - \frac{\pi}{2} = .429$$

For optimized polar quantization

$$\lambda_A(a) = c\, p_A(a)^{1/3} = \frac{a}{\sqrt{3}} e^{-a^2/6}, \quad a \ge 0$$

$$L^2 = \frac{M_A}{M_\Phi} = \frac{1}{2\pi}\left(EA^2 \int_0^\infty f_A^{1/3}(a)\,da \right)^{-1/2} = .376$$

$$\frac{M_A}{M_\Phi} = .613 \quad \text{(more phase levels than amplitude levels)}$$

$$D \cong \frac{1}{12} \sigma_X^2 \, 29.7 \frac{1}{M}$$

Later we'll show that for optimal scalar quantization applies directly to X_1, X_2

$$D \cong \frac{\mathbf{1}}{\mathbf{12}} \sigma_X^2 \, 32.6 \frac{\mathbf{1}}{\mathbf{M}}$$

Gain of polar quantization over conventional optimal scalar quantization

$$10 \log_{10} \frac{\mathbf{32.6}}{\mathbf{29.7}} = .41 \text{ dB}$$

Gain of optimal two-dimensional VQ over scalar quantization = 1.30 dB

Comments on Polar Quantization

- Polar quantization would seem to be "especially well suited" to quantizing $\underline{X}$ when $f_{\underline{X}}(\underline{x})$ is circularly symmetric; i.e. when

  $f_{\underline{X}}(\underline{x})$ depends only on $\|\underline{x}\|$.

  equivalently,

  A and Φ are indep., and Φ is uniformly distributed between 0 and 2π.

  equivalently,

  $f_{\underline{X}}(\underline{x}) = \frac{1}{2\pi} \frac{1}{\|\underline{x}\|} f_A(\|\underline{x}\|)$, where $f_A(a)$ is pdf of amplitude a

  because

  $$f_A(\|\underline{x}\|)\,\Delta \cong \Pr(\|\underline{x}\| \le A \le \|\underline{x}\|+\Delta) = \Pr(\|\underline{x}\| \le \|\underline{X}\| \le \|\underline{x}\|+\Delta) \cong 2\pi \underline{x}\,\Delta\, f_{\underline{X}}(\underline{x})$$

- On the other hand, we found that distortion does not depend on the angle distribution.

  The explanation: The opta function for a noncircularly symmetric density is less than the opta function for a circularly symmetric density with the same magnitude density. Therefore, it's not that polar quantization attains less distortion for a density that is circularly symmetric than for a density that is not. Rather, for a circularly symmetric density, polar quantization is closer to being an optimal 2-dimensional VQ than for a noncircularly symmetric pdf.

Calculus of Variations

Fix M and L. Let $J(\lambda_A)$ be the functional defined by the formula for the distortion of a polar quantizer. Calculus of variations finds an equation that λ_A must solve.

If λ_A is the optimal point density, i.e the one that makes $J(\lambda_A)$ smallest among all nonnegative functions λ that integrate to one, then for any function[5] g such that $\int_0^\infty g(a)\,da = 0$ and any $\varepsilon > 0$, $\lambda_A(a)+\varepsilon g(a)$ cannot be better than $J(\lambda_A)$, i.e.

$$J(\lambda_A+\varepsilon g) \text{ has a local minimum with respect to } \varepsilon \text{ when } \varepsilon = 0$$

Therefore,

$$\frac{d}{d\varepsilon} J(\lambda_A+\varepsilon g)\Big|_{\varepsilon=0} = 0$$

and this must be true for every function g such that $\int_0^\infty g(a)\,da = 0$.

Therefore, we take the derivative of $J(\lambda_A+\varepsilon g)$ with respect to ε. We set $\varepsilon = 0$ and we equate the derivative at $\varepsilon=0$ to zero. This gives an equation that λ_A must satisfy for every function g such that $\int_0^\infty g(a)\,da = 0$. Solving this equation gives the answer[6]. Typically, the equation is something like

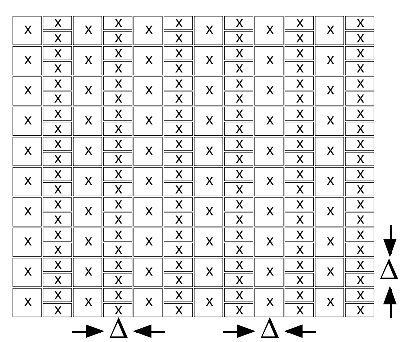
$$\int g(a)\,(\lambda_A(a)-h(a))\,da = 0 \quad \text{for all functions } g \text{ such that } \int g(a)\,da = 0$$

which implies that $\lambda_A(a) = h(a) + c$, where c is some constant.

[5] To be perfectly rigorous, we need to restrict attention to functions g such that $g(a) = 0$ wherever $\lambda_A(a) = 0$.

[6] One should check that $g(a) = 0$ wherever $\lambda_A(a) = 0$.

Generalizing Bennett's Integral to Quantizers Whose Neighboring Cells Do Not Have Similar Sizes or Shapes[7]



Can we apply Bennett's integral to the two-dimensional quantizer shown to the right in which cells come in identical groups of three?

For concreteness and tractability, assume $\underline{X}$ is uniformly distributed on the support of this quantizer.

As a first try, the point density would appear to be

$$\lambda(\underline{x}) = \frac{3}{2M\Delta^2}$$

because there are 3 cells with in each $\Delta \times 2\Delta$ rectangular group, and the inertial profile would appear to be

$$m(\underline{x}) = \frac{1}{2}\frac{1}{12} + \frac{1}{2} \times 0.104 = 0.0919$$

because half the area is contained in squares with nmi 1/12, and half is in 1×2 rectangles with nmi 0.104. Substituting these into Bennett's integral gives

$$D \cong \frac{1}{M^{2/k}} \int \frac{m(\underline{x})}{\lambda^{2/k}(\underline{x})} f_{\underline{X}}(\underline{x})\,d\underline{x} = \frac{1}{M^{2/2}} \int \frac{0.0919}{(3/(2M\Delta^2))^{2/2}} f_{\underline{X}}(\underline{x})\,d\underline{x} = 0.0613\,\Delta^2$$

[7] This section is optional reading.

As a check, let us do a direct computation of distortion.

When $\underline{X}$ lies in a square cell, the distortion is the same as when the quantizer consists entirely of square cells, namely, $\Delta^2/12$, which is the distortion of a uniform scalar quantizer with step size Δ.

Similarly, when $\underline{X}$ lies in a 1×2 rectangular cell, the distortion will be the same as a quantizer consisting entirely of such cells. Such a quantizer is equivalent to using a uniform scalar quantizer with step size Δ on X_1 and a uniform scalar quantizer with step size $\Delta/2$ on X_2. The distortion for X_1 is $\Delta^2/12$, the distortion for X_2 is $(\Delta/2)^2/12$, and these average to give distortion $(5/96)\Delta^2$ for $\underline{X}$ in a 1×2 cell.

Since $\underline{X}$ is in square or rectangular cells with probability 1/2, the overall distortion is

$$D = \frac{1}{2}\frac{\Delta^2}{12} + \frac{1}{2}\frac{5\Delta^2}{96} = \frac{13}{192}\Delta^2 = 0.0677\,\Delta^2$$

This differs considerably from the value $0.0613\,\Delta^2$ computed with Bennett's integral.

Why does Bennett's integral give the wrong answer? Notice the cells within each group do not all have the same, or even similar, sizes and shapes. This invalidates the high-resolution conditions, used in deriving Bennett's integral.

To see explicitly where these conditions were used, let us examine the three approximations in the derivation of Bennett's integral, which is repeated below:

$$D = \frac{1}{k}\sum_{i=1}^{M}\int_{S_i} ||\underline{x}-\underline{w}_i||^2 f_{\underline{X}}(\underline{x})\, d\underline{x}$$

$$\overset{(1)}{\cong} \sum_{i=1}^{M} f_{\underline{X}}(\underline{w}_i)\frac{1}{k}\mathcal{M}(S_i,\underline{w}_i)$$

this approximation is OK; it requires only that the cells be small

$$= \sum_{i=1}^{M} f_{\underline{X}}(\underline{w}_i)\, m(S_i,\underline{w}_i)\, |S_i|^{1+2/k}$$

$$\overset{(2)}{\cong} \sum_{i=1}^{M} f_{\underline{X}}(\underline{w}_i)\frac{m(\underline{w}_i)}{M^{2/k}\lambda^{2/k}(\underline{w}_i)}|S_i| \quad (***)$$

this approximation requires $|S_i| \cong \frac{1}{M\lambda(\underline{w}_i)}$ and $m(S_i,\underline{w}_i) \cong m(\underline{w}_i)$

\s((3),≅,) \f(1,M\s(2/k,)) \i(,)\d\ba4()\f(m(x),λ\s(2/k,)(x)) fX(x) dx \s(this approximation is OK; it requires only that the,Si's form a partition with small cells and wi∈Si)

The problem is that the approximations in (***) are not valid because $\lambda(\underline{x})$ and $m(\underline{x})$ are, generally, slowly varying smooth functions, but cell size and shape change abruptly from cell to cell.

To rectify this problem, we need to consider groups of three cells as one unit. Previously, we approximated distortion one cell at a time:

$$\frac{1}{k}\int_{S_i} ||\underline{x}-\underline{w}_i||^2 f_{\underline{X}}(\underline{x})\, d\underline{x} \;\cong\; f_{\underline{X}}(\underline{w}_i)\,\frac{1}{k}\,\mathcal{M}(S_i,\underline{w}_i) \;=\; f_{\underline{X}}(\underline{w}_i)\, m(S_i,\underline{w}_i)\, |S_i|^{1+2/k}$$

$$\cong\; f_{\underline{X}}(\underline{w}_i)\,\frac{m(\underline{w}_i)}{M^{2/k}\lambda^{2/k}(\underline{w}_i)}\,|S_i|$$

Now, instead, we compute the distortion of the three cells as group.

Let $S_{n,1}, S_{n,2}, S_{n,3}$ denote the cells in the nth group, let $\underline{w}_{n,1}$, $\underline{w}_{n,2}$, $\underline{w}_{n,3}$ denote the corresponding codevectors, let $\tilde{S}_n = S_{n,1} \cup S_{n,2} \cup S_{n,3}$, and let $\tilde{\underline{w}}_n$ be some point in $\tilde{S}_n$.

Point density:

With a macroscopic view, we take the point density to be a function such that

$$\lambda(\underline{x}) \;\cong\; \frac{3}{M\,|\tilde{S}_n|} = \frac{3}{2M\Delta^2} \quad \text{when } \underline{x} \in \tilde{S}_n$$

Inertial profile:

Let $\tilde{m}(S_{n,1}, \underline{w}_{n,1}, S_{n,2}, \underline{w}_{n,1}, S_{n,3}, \underline{w}_{n,1})$ denote the following average nmi of the group of three cells:

$$\tilde{m}(S_{n,1}, \underline{w}_{n,1}, S_{n,2}, \underline{w}_{n,1}, S_{n,3}, \underline{w}_{n,1}) \;=\; \frac{1}{k}\,\frac{\frac{1}{3}\left(\mathcal{M}(S_{n,1},\underline{w}_{n,1})+\mathcal{M}(S_{n,2},\underline{w}_{n,2})+\mathcal{M}(S_{n,3},\underline{w}_{n,3})\right)}{\left(\frac{1}{3}(|S_{n,1}|+|S_{n,2}|+|S_{n,3}|)\right)^{1+2/k}}$$

and let $\tilde{m}(\underline{x})$ denote the corresponding inertial profile:

$$\tilde{m}(\underline{x}) \;\cong\; \tilde{m}(S_{n,1}, \underline{w}_{n,1}, S_{n,2}, \underline{w}_{n,1}, S_{n,3}, \underline{w}_{n,1}) \quad \text{when } \underline{x} \in \tilde{S}_n$$

One can straightforwardly check that $\tilde{m}$ is invariant to the scale and position of the group of cells. Thus, it depends only the shapes and relative sizes of the cells in the group.

Since $\tilde{m}$ does not depend on Δ, we may assume $\Delta = 1$ and compute

$$\tilde{m} \;=\; \frac{1}{2}\,\frac{\frac{1}{3}\left(\frac{1}{6}+\frac{5}{96}+\frac{5}{96}\right)}{\left(\frac{1}{3}(1+\frac{1}{2}+\frac{1}{2})\right)^2} \;=\; \frac{13}{128} \;=\; 0.102$$

which lies between the nmi of a square (0.0833),,and that of a 1×2 rectangle (0.104))

Then the distortion due to the group of three cells is

$$\frac{1}{k}\sum_{i=1}^{3}\int_{S_{n,i}} ||\underline{x}-\underline{w}_{n,i}||^2 f_{\underline{X}}(\underline{x})\, d\underline{x} \;\cong\; f_{\underline{X}}(\underline{\tilde{w}}_n) \sum_{i=1}^{3}\frac{1}{k}\mathcal{M}(S_{n,},\underline{w}_{n,i}) \quad \text{since } \tilde{S}_n \text{ is also small}$$

$$= f_{\underline{X}}(\underline{\tilde{w}}_n)\, 3\,\tilde{m}(S_{n,1},\underline{w}_{n,1},S_{n,2},\underline{w}_{n,2},S_{n,3},\underline{w}_{n,3})\; \left(\tfrac{1}{3}\right)^{1+2/k} |\tilde{S}_n|^{1+2/k}$$

$$= f_{\underline{X}}(\underline{\tilde{w}}_n)\, \tilde{m}(S_{n,1},\underline{w}_{n,1},S_{n,2},\underline{w}_{n,2},S_{n,3},\underline{w}_{n,3})\; \left(\tfrac{1}{3}|\tilde{S}_n|\right)^{2/k} |\tilde{S}_n|$$

$$\cong f_{\underline{X}}(\underline{\tilde{w}}_n)\, \frac{\tilde{m}(\underline{\tilde{w}}_n)}{M^{2/k}\lambda^{2/k}(\underline{\tilde{w}}_n)}\, |\tilde{S}_n|$$

where we have used the approximation $\frac{1}{3}|\tilde{S}_n| \cong \dfrac{1}{M\lambda(\underline{\tilde{w}}_n)}$

Now, summing the approximate distortions of all groups gives

$$D = \sum_{n=1}^{M/3}\frac{1}{k}\sum_{i=1}^{3}\int_{S_{n,i}} ||\underline{x}-\underline{w}_{n,i}||^2 f_{\underline{X}}(\underline{x})\, d\underline{x} \;\cong\; \sum_{n=1}^{M/3} f_{\underline{X}}(\underline{\tilde{w}}_n)\frac{\tilde{m}(\underline{\tilde{w}}_n)}{M^{2/k}\lambda^{2/k}(\underline{\tilde{w}}_n)}\, |\tilde{S}_n|$$

$$\cong \frac{1}{M^{2/k}}\int \frac{\tilde{m}(\underline{x})}{\lambda^{2/k}(\underline{x})}\, f_{\underline{X}}(\underline{x})\, d\underline{x}$$

This is just like the original Bennett's integral, but with a definition of inertial profile that averages over cells in a group.

As a check, substituting for $\lambda(\underline{x})$ and $\tilde{m}(\underline{x})$ gives

$$D \cong \frac{1}{M}\int \frac{13/128}{3/(2M\Delta^2)}\, f_{\underline{X}}(\underline{x})\, d\underline{x} = \frac{13}{192}\Delta^2 = 0.0677\,\Delta^2,$$

which is the correct answer.

More generally if a partition comes in groups of M_o cells, then the Bennett integral approximation is:

$$D \cong \frac{1}{M^{2/k}} \int \frac{\tilde{m}(\underline{x})}{\lambda^{2/k}(\underline{x})} f_{\underline{X}}(\underline{x})\, d\underline{x}$$

where

$$\lambda(\underline{x}) \cong \frac{M_o}{M \sum_{i=1}^{M_o} |S_{n,i}|} \quad \text{when } \underline{x} \in S_{n,1} \cup \ldots \cup S_{n,M_o}$$

$$\tilde{m}(\underline{x}) \cong \tilde{m}(S_{n,1}, \underline{w}_{n,1}, \ldots, S_{n,M_o}, \underline{w}_{n,M_o}) \quad \text{when } \underline{x} \in S_{n,1} \cup \ldots \cup S_{n,M_o}$$

$$\tilde{m}(S_{n,1}, \underline{w}_{n,1}, \ldots, S_{n,M_o}, \underline{w}_{n,M_o}) = \frac{1}{k} \frac{\frac{1}{M_o}\sum_{i=1}^{M_o} \mathcal{M}(S_{n,i},\underline{w}_{n,i})}{\left(\frac{1}{M_o}(|S_{n,1}|+\ldots+|S_{n,M_o}|)\right)^{1+2/k}}$$

Distortion Density and Local Distortion[8]

Let G be some subset of $\Re^k$. If one repeats the derivation of Bennett's integral, but for the distortion computed only $\underline{X}$ in the set G, then one obtains

$$\int_G ||\underline{x}-Q(\underline{x})||^2 f_{\underline{X}}(\underline{x})\, d\underline{x} \cong \int_G \frac{m(\underline{x})}{M^{2/k}\lambda^{2/k}(\underline{x})} f_{\underline{X}}(\underline{x})\, d\underline{x} = \int_G \rho(\underline{x})\, d\underline{x}$$

where

$$\rho(\underline{x}) \triangleq \frac{m(\underline{x})}{M^{2/k}\lambda^{2/k}(\underline{x})} f_{\underline{X}}(\underline{x})$$

Since ρ is a function that when integrated over a region G yields the distortion in region G, it can be called the *distortion density* of the quantizer.

One may also rewrite the distortion in G as

$$\int_G ||\underline{x}-Q(\underline{x})||^2 f_{\underline{X}}(\underline{x})\, d\underline{x} \cong \int_G d(\underline{x})\, f(\underline{x})\, d\underline{x}$$

where

$$d(\underline{x}) \triangleq \frac{m(\underline{x})}{M^{2/k}\lambda^{2/k}(\underline{x})}$$

is considered the *local distortion* of the quantizer at $\underline{x}$.

[8]This section is optional reading.

Notice that the distortion in region G is not the same as the conditional distortion given region G, which is

$$E\left[\|\underline{X}-Q(\underline{X})\|^2 \mid \underline{X}\in G\right] = \int_G \|\underline{x}-Q(\underline{x})\|^2 \frac{f_{\underline{X}}(\underline{x})}{\Pr(\underline{X}\in G)}\, d\underline{x}$$

$$\cong \int_G \frac{\rho(\underline{x})}{\Pr(\underline{X}\in G)}\, d\underline{x}$$

$$= \int_G d(\underline{x}) \frac{f_{\underline{X}}(\underline{x})}{\Pr(\underline{X}\in G)}\, d\underline{x}$$

This shows that conditional distortion is the average of local distortion.

When G is small and $\underline{x} \in G$, approximating d and f as being constant on G yields

$$E\left[\|\underline{X}-Q(\underline{X})\|^2 \mid \underline{X}\in G\right] \cong \frac{\rho(\underline{x})}{\Pr(\underline{X}\in G)} |G| \cong \frac{\rho(\underline{x})}{f_{\underline{X}}(\underline{x})} = d(\underline{x})\,.$$

Thus for a small region, the conditional distortion approximately equals the local distortion in that region.