9.1 Motivation: Quantization of Random Variables

- A continuous source contains an infinite amount of information and cannot be represented exactly using a finite number of bits.

- In lossy source coding, we seek instead a representation that is close to the source (with respect to some fidelity criterion) and can be represented using a finite number of bits.

- **Quantization:**
  - Let $X$ be a random variable
  - For every value $X = x$, we would like to find a representation $\hat{x}(x)$ where $\hat{x}$ can take on only $2^R$ different values for a given rate $R$ (measured in bits).

- **Example:** Quantizing a Gaussian random variable with squared error distortion:
  - Let $X \sim N(0, \sigma^2)$
  - Consider mean squared error distortion
    $$\mathbb{E}[(X - \hat{x}(X))^2]$$
  - If we use 1 bit to represent $X$ (i.e., we can chose only two different reconstruction symbols), then we should use the bit to indicate whether $X$ is positive or negative.
  - To minimize the square error distortion, the reconstruction symbols should be the conditional mean given the sign
    $$\hat{x}(x) = \begin{cases} 
    \mathbb{E}[X|X > 0], & x > 0 \\
    \mathbb{E}[X|X < 0], & x < 0 
    \end{cases} = \begin{cases} 
    \sqrt{\frac{2}{\pi}} \sigma^2, & x > 0 \\
    -\sqrt{\frac{2}{\pi}} \sigma^2, & x < 0 
    \end{cases}$$
The average distortion will then be
\[
E[(X - \hat{x}(X))^2] = \frac{1}{2}E[(X - \hat{x}(X))^2|X > 0] + \frac{1}{2}E[(X - \hat{x}(X))^2|X \leq 0] = \frac{1}{2}E[(X - \hat{x}(X))^2|X > 0] = \left(1 - \frac{2}{\pi}\right) \sigma^2
\]

If we are given 2 bits to represent $X$, we should divide the real line into 4 regions and use the conditional means within each region as the reconstruction points. However, it is not obvious how we should choose these regions.

- In general, a quantization scheme is characterized by a partition $\{V_i\}$ of the space and the corresponding reconstruction points $\{\hat{x}_i\}$.

\[
x \in V_i \implies \hat{x}(x) = \hat{x}_i
\]

The regions and reconstruction points should satisfy:

- Given a set of reconstruction points, the regions should be chosen to minimize the distortion. Under the squared-error distortion, this occurs if the regions are the Voronoi cells

\[
V_i = \{x : \|x - x_i\| < \|x - x_j\| \text{ for all } j \neq i\}
\]

- Given a set of regions, the reconstruction points should be chosen to minimize the distortion. Under squared error distortion, this is given by the conditional mean

\[
\hat{x}_i = E[X|X \in V_i]
\]

- **Lloyd’s Algorithm** is an iterative algorithm for constructing a quantization function. Starting with an initial set of reconstruction points, the algorithm repeats the following two steps:

  1. Given reconstruction points, find optimal set of regions
  2. Given regions, find optimal set of reconstruction points.

This algorithm will converge to a local optimum (but not necessarily the global optimum).

- **Vector Quantization**

  - Let $X^n = [X_1, \ldots, X_n]$ be an length-$n$ random vector with iid entries
  - For every realization $X^n = x^n$ we would like to find a representation $\hat{x}^n(x^n)$ where $\hat{x}^n$ can take on only $2^{nR}$ different values for a given rate $R$ (measured in bits).
  - One option is to use the rate $R$ scalar quantization strategy outlined above.
  - It turns out that quantizing jointly can be much better than quantizing separately.

- Recap:

  - scalar quantization: Approximate $X$ with $R$ bits
  - vector quantization: Approximate $X^n$ with $nR$ bits
9.2 Lossy Source Coding Definitions

- Intuition: Using more bits reduces quantization error. How can we quantify this tradeoff?
- Illustration of a $(2^n R, n)$ lossy source code

\[
\begin{array}{c}
\text{source} \\
X^n \\
\end{array} \xrightarrow{f_n} \begin{array}{c}
\text{encoder} \\
W \in \{1, \ldots, 2^n R\} \\
\end{array} \xrightarrow{g_n} \begin{array}{c}
\text{decoder} \\
\hat{X}^n \\
\end{array}
\]

- The source produces a sequence $X_1, X_2, \cdots$ of iid random variable with distribution $p(x)$ supported on a finite alphabet $\mathcal{X}$.
- The encoder is a mapping $f_n : \mathcal{X}^n \rightarrow \{1, 2, \cdots, 2^n R\}$ that describes every source sequence by an index $w$. The rate is given by

\[
R = \frac{\log \# \text{ of indices}}{n} \quad \text{bits per symbol}
\]

- The decoder $g_n : \{1, 2, \cdots, 2^n R\} \rightarrow \hat{X}$ maps each index to an estimate $\hat{X} \in \hat{X}^n$ where $\hat{X}$ is the reconstruction alphabet.

**Definition:** A *per-letter distortion measure* is a mapping

\[
d : \mathcal{X} \times \hat{\mathcal{X}} \rightarrow \mathbb{R}^+
\]

from the set of source alphabet-reconstruction pairs to the nonnegative real numbers. In most cases, the reconstruction alphabet is equal to the source alphabet. The distortion measure is *bounded* if the maximum value of the distortion is finite

\[
d \text{ is bounded} \iff \max_{x \in \mathcal{X}, \hat{x} \in \hat{\mathcal{X}}} d(x, \hat{x}) < \infty
\]

- The Hamming distortion is defined as

\[
d(x, \hat{x}) = \begin{cases} 
0, & \text{if } x = \hat{x} \\
1, & \text{if } x \neq \hat{x}
\end{cases}
\]

- The squared-error distortion is defined as

\[
d(x, \hat{x}) = (x - \hat{x})^2
\]

- The distortion between two sequences $x^n$ and $\hat{x}^n$ is given by the average per-letter distortion

\[
d(x^n, \hat{x}^n) = \frac{1}{n} \sum_{i=1}^{n} d(x_i, \hat{x}_i)
\]

**Definition:** A $(2^n R, n)$ rate distortion coding scheme consists of

- a source alphabet $\mathcal{X}$ and reconstruction alphabet $\hat{\mathcal{X}}$.
- encoding function $f_n : \mathcal{X}^n \rightarrow \{1, 2, \cdots, 2^n R\}$
- decoding function $g_n : \{1, 2, \cdots, 2^n R\} \rightarrow \hat{\mathcal{X}}^n$. 
• a distortion measure $d$

• The distortion associate with the $(2^{nR}, n)$ coding scheme is given by the expected distortion

$$\mathbb{E}[d(X^n, \hat{X}^n)] = \sum_{x^n} p(x^n) d(x^n, g_n(f_n(x^n)))$$

• **Definition:** A rate distortion pair $(R, D)$ is said to be *achievable* for a source $p(x)$ and distortion measure $d(\cdot, \cdot)$ if there exists a sequence of $(2^{nR}, n)$ rate-distortion coding schemes with

$$\lim_{n \to \infty} \mathbb{E}[d(X^n, g_n(f_n(X^n)))] < D$$

• **Definition:** the *rate distortion region* for a source is the closure of the set of achievable distortion pairs $(R, D)$.

• **Definition:** The *rate distortion function* $R(D)$ is the infimum of rates $R$ such that $(R, D)$ is in the rate distortion region. Likewise, the *distortion rate function* $D(R)$ is the infimum of distortions such that $(R, D)$ is in the rate distortion region.

![](image)

### 9.3 The Rate Distortion Coding Theorem

• **Definition:** The *information rate distortion function* $R^{(I)}(D)$ for a source $p(x)$ with distortion measure $d(x, \hat{x})$ is defined as

$$R^{(I)}(D) = \min_{p(\hat{x}|x)} I(X; \hat{X})$$

where the minimum is over all distributions $p(\hat{x}|x)$ such that the pair $(X, \hat{X})$ satisfy the distortion constraint

$$\mathbb{E}[d(X, \hat{X})] \leq D, \quad (X, \hat{X}) \sim p(x)p(\hat{x}|x)$$

• **Theorem:** The rate distortion function for an i.i.d. source $p(x)$ and bounded distortion measure $d(\cdot, \cdot)$ is equal to the associated information rate distortion function, i.e.

$$R(D) = R^{(I)}(D)$$

  - Achievability: $R > R^{(I)}(D) \implies$ the pair $(R, D)$ is achievable

  - Converse: the pair $(R, D)$ is achievable $\implies R \geq R^{(I)}(D)$
9.3.1 Example: Binary Source

- **Theorem**: The rate distortion function for an iid Bernoulli($p$) source with Hamming distortion is given by

$$R(D) = \begin{cases} 
H(p) - H(D), & 0 \leq D \leq \min\{p, 1-p\} \\
0, & D > \min\{p, 1-p\}
\end{cases}$$

- **Proof of $\geq$**:
  - Need to show that for any distribution $p(\hat{x}|x)$,
    $$\mathbb{P}[X \neq \hat{X}] \leq D \implies I(X; \hat{X}) \geq R(D)$$
  - Without loss of generality, assume $p < 1/2$.
  - Note that $X \oplus \hat{X} = 1 \iff \hat{X} \neq X$
  - For any $p(\hat{x}|x)$ obeying $\mathbb{P}[X \neq \hat{X}] \leq D$, we have
    $$I(X; \hat{X}) = H(X) - H(X|\hat{X})$$
    $$= H(p) - H(X \oplus \hat{X}|\hat{X})$$
    $$\geq H(p) - H(X \oplus \hat{X})$$
    $$= H(p) - H_b(\mathbb{P}[X \neq \hat{X}])$$
    $$\geq H(p) - H_b(D)$$
    since $H(D) > H(D')$ for all $D' < D$
  - Thus, for $D < p$,
    $$R(D) \geq H(p) - H(D)$$

- **Proof of $\leq$**:
  - The goal is to show that there exists $p(\hat{x}|x)$ such that $\mathbb{P}[X \neq \hat{X}] \leq D$ and $I(X; \hat{X}) \leq R(D)$
  - Without loss of generality, assume $p < 1/2$.
  - If $D \geq p$ then distortion is achieved automatically and $R = 0$ suffices.
  - Henceforth, we assume $D < p$.
  - The trick is to consider a channel with input $\hat{X}$ and output $X$ defined by $p(x|\hat{x})$. Note that given $p(x)$, there is a one-to-one mapping between $p(\hat{x}|x)$ and $p(x|\hat{x})$
○ For this problem, it is useful to consider the binary symmetric channel with cross over probability $D$ shown below:

![Binary Symmetric Channel Diagram]

○ We want choose the distribution on $\hat{X}$ so that $X$ is Bernoulli($p$). This requires that

$$P[X = 1] = P[\hat{X} = 1](1 - D) + P[\hat{X} = 0]D = p$$

or equivalently,

$$P[\hat{X} = 1] = \frac{p - D}{1 - 2D}, \quad P[\hat{X} = 0] = \frac{1 - p - D}{1 - 2D}$$

○ With this marginal distribution $\hat{X}$, it then follows that

$$I(X; \hat{X}) = H(X) - H(X|\hat{X}) = H(p) - H(D)$$

and the expected distortion is $P[X \neq \hat{X}] = D$.

9.3.2 Example: Gaussian Source

The rate distortion coding theorem stated above assumes discrete sources and bounded distortion measures. It turns out that the theorem can also be proved for well-behaved continuous sources, such as for the case of Gaussian source with squared error distortion.

- **Theorem:** The rate distortion function for an $\mathcal{N}(0, \sigma^2)$ source with squared-error distortion is

$$R(D) = \begin{cases} 
\frac{1}{2} \log \left( \frac{\sigma^2}{D} \right), & 0 \leq D \leq \sigma^2 \\
0, & D > \sigma^2 \end{cases}$$

- **Proof of $\geq$:**

○ Let $(X, \hat{X})$ be distributed such that $X \sim \mathcal{N}(0, \sigma^2)$ and $\mathbb{E}[(X - \hat{X})^2] \leq D$. 

![Rate Distortion Function Graph]
• The mutual information obeys the following inequalities:

\[
I(X; \hat{X}) = h(X) - h(X|\hat{X}) \\
= \frac{1}{2} \log(2\pi e\sigma^2) - h(X - \hat{X} | \hat{X}) \\
\geq \frac{1}{2} \log(2\pi e\sigma^2) - h(X - \hat{X}) \\
\geq \frac{1}{2} \log(2\pi e\sigma^2) - \frac{1}{2} \log(2\pi e \mathbb{E}[(X - \hat{X})^2]) \\
\geq \frac{1}{2} \log(2\pi e\sigma^2) - \frac{1}{2} \log(2\pi e D) \\
= \frac{1}{2} \log \frac{\sigma^2}{D}
\]

• Hence, if \( D < \sigma^2 \),

\[R(D) \geq \frac{1}{2} \log \frac{\sigma^2}{D}\]

• **Proof of \( \leq \):**

  • Consider the inverse channel with input \( \hat{X} \) and output \( X \).
  • In this case, we choose the Gaussian channel:

  \[X = \hat{X} + Z, \quad \hat{X} \sim \mathcal{N}(0, \sigma^2 - D), \quad Z \sim \mathcal{N}(0, D)\]

  where \( \hat{X} \) and \( Z \) are independent.

  • Then,

  \[
  I(X; \hat{X}) = h(X) - h(X|\hat{X}) \\
  = \frac{1}{2} \log(2\pi e\sigma^2) - \frac{1}{2} \log(2\pi e D) \\
  = \frac{1}{2} \log \frac{\sigma^2}{D}
  \]

  and also

  \[\mathbb{E}[(X - \hat{X})^2] = \mathbb{E}[Z^2] = D\]

• The distortion rate function is given by

\[D(R) = \sigma^2 2^{-2R}\]

• Comparison of scalar quantization and rate-distortion at rate \( R = 1 \)

  • Scalar quantization: \( D = (1 - \frac{2}{\pi})\sigma^2 \approx 0.36\sigma^2 \)
  • Distortion rate function: \( D = \frac{1}{4}\sigma^2 \).

9.3.3 Properties of the Rate Distortion Functions

• **Theorem:** The rate distortion function \( R(D) \) is a non-increasing convex function of \( D \).

• **Proof:**
The rate distortion function is the minimum of the mutual information over increasingly larger sets as $D$ increases. Thus $R(D)$ is non-increasing.

Consider two rate distortion pairs $(R_1, D_1)$ and $(R_2, D_2)$ which lie on the rate distortion curve.

Let $p_1(x, \hat{x}) = p(x)p_1(\hat{x}|x)$ and $p_2(x, \hat{x}) = p(x)p_2(\hat{x}|x)$ be the distributions that achieve these pairs.

For $\lambda \in [0, 1]$ consider the distribution $p_\lambda(\hat{x}|x) = \lambda p_1(\hat{x}|x) + (1 - \lambda) p_2(\hat{x}|x)$.

By the linearity of expectation, the distortion associated with $p_\lambda$ is given by $D_\lambda = \lambda D_1 + (1 - \lambda) D_2$.

Since mutual information $I(X; \hat{X})$ is convex with respect to the conditional distribution $p(x|\hat{x})$, the rate distortion function at $D_\lambda$ obeys

\[
R(D_\lambda) \leq I(p(x), p_\lambda(\hat{x}|x)) \\
\leq \lambda I(p(x), p_1(\hat{x}|x)) + (1 - \lambda) I(p(x), p_2(\hat{x}|x)) \\
= \lambda R(D_1) + (1 - \lambda) R(D_2)
\]

### 9.4 Converse to the Rate Distortion Coding Theorem

- We now prove the converse to the rate distortion coding theorem. In particular, we show that for any $(2^{nR}, n)$ coding scheme,

\[
\mathbb{E}[d(X^n, g_n(f_n(X^n)))] \leq D \implies R \geq R(D)
\]

- Let $\hat{X}^n = g_n(f_n(X^n))$ and observe that:

\[
nR \geq H(f_n(X^n)) \quad \text{range of } f_n \text{ at most } 2^{nR} \\
\geq H(f_n(X^n)) - H(f_n(X^n)|X^n) \\
= I(X^n; f_n(X^n)) \\
\geq I(X^n; \hat{X}^n) \quad \text{data processing inequality} \\
= H(X^n) - H(X^n|\hat{X}^n) \\
= \sum_{i=1}^{n} H(X_i) - H(X^n|\hat{X}^n) \\
= \sum_{i=1}^{n} H(X_i) - \sum_{i=1}^{n} H(X_i|\hat{X}_i, X_{i-1}, \ldots, X_1) \quad \text{chain rule} \\
\geq \sum_{i=1}^{n} H(X_i) - \sum_{i=1}^{n} H(X_i|\hat{X}_i) \quad \text{conditioning cannot increase entropy} \\
= \sum_{i=1}^{n} I(X_i; \hat{X}_i) \\
\geq \sum_{i=1}^{n} R\left(\mathbb{E}[d(X_i, \hat{X}_i)]\right) \quad \text{definition of rate distortion function}
\]
\[
= n \left( \frac{1}{n} \sum_{i=1}^{n} R\left( \mathbb{E} \left[ d(X_i, \hat{X}_i) \right] \right) \right)
\]
\[
\geq n R \left( \frac{1}{n} \sum_{i=1}^{n} \mathbb{E} \left[ d(X_i, \hat{X}_i) \right] \right) \quad \text{convexity of } R(D) \text{ & Jensen’s inq.}
\]
\[
= n R \left( \mathbb{E} \left[ d(X^n, \hat{X}^n) \right] \right) \quad \text{extension of distortion function}
\]
\[
\geq n R(D) \quad \text{since } R(D) \text{ is nonincreasing}
\]