Problem 1

Properties of $R(D)$

Fix an allowed distortion $D$, and let $q^*(\hat{x}|x)$ be a conditional distribution that achieves the minimum in

$$\min_{q(\hat{x}|x)} \mathbb{E}[d(X,\hat{X})] \leq D.$$ 

Then,

$$D = \mathbb{E}_{Qq^*}[d(X,\hat{X})] = \sum_x \sum_{\hat{x}} Q(x)q^*(\hat{x}|x)d(x,\hat{x})$$

$$= \sum_x \sum_{\hat{x}} Q(x)q^*(\hat{x}|x)(d'(x,\hat{x}) + w(x))$$

$$= \sum_x \sum_{\hat{x}} Q(x)q^*(\hat{x}|x)d'(x,\hat{x}) + \sum_x Q(x)q^*(\hat{x}|x)w(x)$$

$$= \sum_x \sum_{\hat{x}} Q(x)q^*(\hat{x}|x)d'(x,\hat{x}) + \sum_x Q(x)w(x) \sum_{\hat{x}} q^*(\hat{x}|x)$$

$$= \sum_x \sum_{\hat{x}} Q(x)q^*(\hat{x}|x)d'(x,\hat{x}) + \sum_x Q(x)w(x)$$

$$= \mathbb{E}_{Qq^*}[d'(X,\hat{X})] + \bar{w},$$

where the last equality holds by definition of $\bar{w}$. Since $q^*(\hat{x}|x)$ achieves the minimum of $I(X;\hat{X})$ under the constraint that $\mathbb{E}[d(X,\hat{X})] = D$, it thus follows that it also achieves the minimum of $I(X;\hat{X})$ under the constraint that $\mathbb{E}[d'(X,\hat{X})] = D - \bar{w}$. Hence, $R'(D - \bar{w}) = R(D)$, or $R'(D) = R(D + \bar{w})$.

Therefore, if $\min_x d(x,\hat{x}) > 0$ for at least some $x$, we can define a new distortion measure

$$d'(x,\hat{x}) \triangleq d(x,\hat{x}) - \min_b d(x,b).$$

This new measure will result in a slightly smaller distortion, and hence we also need to adjust our requirement for the maximal distortion:

$$D' \triangleq D - \mathbb{E}_{b}\left[\min_b d(X,b)\right].$$

The resulting optimal rates are identical for both schemes and are achieved by the same distribution.

© Amos Lapidoth, 2015/2016
Problem 2

Erasure Distortion

a) Let’s start with the computation of the expected distortion:

\[
\mathbb{E}[d(X, \hat{X})] = \sum_{x \in \{0,1\}} \sum_{\hat{x} \in \{0,1,?\}} Q(x)q(\hat{x}|x)d(x, \hat{x})
\]

\[
= \frac{1}{2} q(1|1) \cdot 0 + \frac{1}{2} q(?|1) \cdot 1 + \frac{1}{2} q(0|1) \cdot \infty
\]

\[
+ \frac{1}{2} q(1|0) \cdot \infty + \frac{1}{2} q(?|0) \cdot 1 + \frac{1}{2} q(0|0) \cdot 0
\]

\[
= \frac{1}{2} q(?|1) + \frac{1}{2} q(?|0)
\]

\[
= \Pr[\hat{X} = ?],
\]

where we have to set \(q(0|1) = q(1|0) = 0\) since otherwise the expected distortion is always infinite. Hence, to find the rate distortion function we have to minimize \(I(X; \hat{X})\) subject to the constraint

\[
\Pr[\hat{X} = ?] \leq D. \tag{1}
\]

Since \(q(0|1) = q(1|0) = 0\) we know that if \(\hat{X} = 0\), then \(X = 0\) with probability one, and if \(\hat{X} = 1\), then \(X = 1\) with probability one. Thus,

\[
I(X; \hat{X}) = H(X) - H(X|\hat{X})
\]

\[
= \log 2 \left( \Pr[\hat{X} = 0] H(X|\hat{X} = 0) + \Pr[\hat{X} = 1] H(X|\hat{X} = 1) \right)
\]

\[
+ \Pr[\hat{X} =?] \cdot H(X|\hat{X} = ?)
\]

\[
\geq \log 2 - D \cdot \log 2,
\]

where the last inequality is met with equality if

\[
q(\hat{x}|x) = \begin{cases} 
1 - D & \text{if } \hat{x} = x, \\
D & \text{if } \hat{x} = ?.
\end{cases}
\]

Hence, the rate distortion function is given as follows:

\[
R(D) = \begin{cases} 
(1 - D) \log 2 & \text{if } 0 \leq D \leq 1, \\
0 & \text{if } D > 1.
\end{cases}
\]

The rate distortion region is depicted in Figure 1.

b) We use the following scheme: for every source sequence of length \(n\), choose a codeword that consists of the first \(n(1 - D)\) source symbols followed by \(nD\) question marks. This code needs \(2^{n(1-D)}\) codewords and has therefore a rate of \((1 - D)\) bits. The distortion achieved by this code is as follows: the first \(n(1 - D)\) digits have zero distortion, the rest has distortion 1. Hence, on average we get a distortion of \(D\).

Note that once \(D \geq 1\) we will only use question marks, i.e., we only use one codeword \(\hat{x} = (?,\ldots,?)\). This code has rate zero and it achieves a distortion of 1.
Problem 3  

**Rate Distortion Function with Infinite Distortion**

Let’s start by computing the average distortion for a certain choice of \( q(\cdot|\cdot) \):

\[
E[d(X, \hat{X})] = Q_X(0)q(0|0)d(0, 0) + Q_X(0)q(1|0)d(0, 1) + Q_X(1)q(0|1)d(1, 0) + Q_X(1)q(1|1)d(1, 1)
\]

\[
= \frac{1}{2}q(1|0) \cdot \infty + \frac{1}{2}q(0|1) \leq D.
\]

Hence, we see that \( q(1|0) \equiv 0 \). Let \( q(0|1) = p \) for some parameter \( p, 0 \leq p \leq 1 \). Then, the above computation implies that

\[
0 \leq p \leq 2D.
\]

The joint distribution of \((X, \hat{X})\) is given by

\[
Q_{X,\hat{X}}(x, \hat{x}) = \begin{cases} \frac{1}{2} & \text{if } x = 0, \hat{x} = 0, \\ 0 & \text{if } x = 0, \hat{x} = 1, \\ \frac{1}{2}p & \text{if } x = 1, \hat{x} = 0, \\ \frac{1}{2}(1-p) & \text{if } x = 1, \hat{x} = 1, \end{cases}
\]

and therefore

\[
Q_\hat{X}(\hat{x}) = \begin{cases} \frac{1}{2} + \frac{p}{2} & \text{if } \hat{x} = 0, \\ \frac{1}{2} - \frac{p}{2} & \text{if } \hat{x} = 1. \end{cases}
\]

We next compute the mutual information \( I(X; \hat{X}) \) for this joint distribution:

\[
I(X; \hat{X}) = \sum_{x, \hat{x}} Q_{X,\hat{X}}(x, \hat{x}) \log \frac{Q_{X,\hat{X}}(x, \hat{x})}{Q_X(x)Q_\hat{X}(\hat{x})}
\]

\[
= \frac{1}{2} \log \frac{1}{2(\frac{1}{2} + \frac{p}{2})} + \frac{p}{2} \log \frac{\frac{p}{2}}{\frac{1}{2} + \frac{p}{2}} + \left(\frac{1}{2} - \frac{p}{2}\right) \log \frac{\frac{1}{2} - \frac{p}{2}}{\frac{1}{2}(\frac{1}{2} - \frac{p}{2})}
\]

\[
= \frac{1}{2} \log 2 - \frac{1}{2} \log(1 + p) + \frac{p}{2} \log 2 + \frac{p}{2} \log p - \frac{p}{2} \log(1 + p) + \left(\frac{1}{2} - \frac{p}{2}\right) \log 2
\]

\[
= \log 2 - \frac{1}{2}(1 + p) \log(1 + p) + \frac{p}{2} \log p \triangleq f(p).
\]

To find the optimal \( p \), let’s compute the derivative:

\[
\frac{\partial f(p)}{\partial p} = \frac{1}{2} \log \frac{p}{1 + p} < 0.
\]
Since the derivative is negative, the mutual information decreases as $p$ increases, and we thus want to choose $p$ as large as possible:

$$p^* = \begin{cases} 2D & \text{if } 2D \leq 1, \\ 1 & \text{if } 2D > 1. \end{cases}$$

This leads to the following rate distortion function:

$$R(D) = \begin{cases} \log 2 - \left( \frac{1}{2} + D \right) \log(1 + 2D) + D \log 2D & \text{if } 0 \leq D \leq \frac{1}{2}, \\ 0 & \text{if } D > \frac{1}{2}. \end{cases}$$

Note that for $D > \frac{1}{2}$ we take $\hat{X} = 0$.

**Problem 4**

**Rate Distortion for Uniform Source with Hamming Distortion**

First, we note that the expected distortion $E[d(X, \hat{X})]$ is equal to $\Pr[X \neq \hat{X}]$:

$$E[d(X, \hat{X})] = \sum_x p(x) \sum_{\hat{x}} p(\hat{x}|x) d(x, \hat{x}) = \sum_x p(x) \sum_{\hat{x} \neq x} p(\hat{x}|x) = \Pr[X \neq \hat{X}].$$

Hence, to find $R(D)$ we have to minimize $I(X; \hat{X})$ subject to the constraint that $\Pr[X \neq \hat{X}] \leq D$. First we note that $R(D) = 0$ for $D \geq (m-1)/m$, because if $\hat{X}$ is independent of $X$, then, irrespective of $P_\hat{X}$, the expected distortion equals $(m-1)/m$ and the mutual information between $X$ and $\hat{X}$ equals 0. In the following we thus consider the case where $D < (m-1)/m$. Introduce the auxiliary random variable

$$E = \begin{cases} 0 & \text{if } \hat{X} = X, \\ 1 & \text{if } \hat{X} \neq X, \end{cases}$$

and note that $\Pr[X \neq \hat{X}] \leq D$ is equivalent to $\Pr[E = 1] \leq D$. For $D < (m-1)/m$

$$I(X; \hat{X}) = H(X) - H(X|\hat{X}) = \log m - H(X,E|\hat{X}) \geq \log m - \Pr[E = 1]H(X|\hat{X}, E = 1) - H(E|\hat{X}) \geq \log m - \Pr[E = 1] \log(m-1) - H(E) \geq \log m - D \log(m-1) - H_b(D),$$

where $H_b(\cdot)$ is the binary entropy function: (a) holds because $X$ is uniform over $\{1, 2, \ldots, m\}$ and $E$ is a function of $(X, \hat{X})$; (b) holds because $E = 1$ implies that $X \in \{1, 2, \ldots, m\} \setminus \{\hat{X}\}$, because the uniform distribution maximizes entropy, and because conditioning cannot increase entropy; and (c) holds because $\Pr[X \neq \hat{X}] \leq D$ (which follows from the rate distortion constraint) and because the function

$$t \mapsto t \log(m-1) + H_b(t)$$

is nondecreasing in $t$ on the interval $[0, (m-1)/m]$. Next, we argue that both inequalities (b) and (c) can be met with equality. Indeed, if we let

$$P_{X|\hat{X}}(\hat{x}|x) = \begin{cases} 1-D & \text{if } \hat{x} = x, \\ \frac{D}{m-1} & \text{if } \hat{x} \neq x, \end{cases}$$
then by

\[ P_{\hat{X}}(\hat{x}) = \sum_x P_{X,\hat{X}}(x, \hat{x}) = \sum_x P_X(x)P_{\hat{X}|X}(\hat{x}|x) = \frac{1}{m}(1 - D) + \frac{m - 1}{m} \frac{D}{m - 1} = \frac{1}{m} \]

\( \hat{X} \) is uniform over \( \{1, 2, \ldots, m\} \), by

\[ P_{E|\hat{X}}(1|\hat{x}) = \sum_{x \neq \hat{x}} P_{X,\hat{X}}(x, \hat{x}) = (m - 1) \frac{D}{m - 1} = D \]

\( E \) is independent of \( \hat{X} \) and \( \Pr[E = 1] = D \), and since for all \( x, \hat{x} \in \{1, 2, \ldots, m\} \) for which \( x \neq \hat{x} \)

\[ P_{X|\hat{X},E}(x|\hat{x}, 1) = \frac{P_{X,\hat{X}}(x, \hat{x})}{P_{X,E}(\hat{x}, 1)} = \frac{P_X(x)P_{\hat{X}|X}(\hat{x}|x)}{\frac{1}{m} \frac{D}{m - 1}} = \frac{1}{\frac{1}{m} \frac{D}{m - 1}} = \frac{1}{m - 1} \]

\( X \) is uniform over \( \{1, 2, \ldots, m\} \setminus \{\hat{X}\} \) conditional on \( \hat{X} \) and \( E = 1 \). Consequently, for this choice of \( P_{\hat{X}|X} \) both inequalities (b) and (c) hold with equality. Therefore:

\[ R(D) = \begin{cases} \log m - D \log(m - 1) - H_b(D) & 0 \leq D \leq \frac{m - 1}{m}, \\ 0 & D > \frac{m - 1}{m}. \end{cases} \]