Prof. Dr. Rudolf Mathar, Dr.-Ing. Gholamreza Alirezaei, Emilio Balda, Vimal Radhakrishnan

Tutorial 8
- Proposed Solution -

Monday, January 28, 2019

Solution of Problem 1

One bit quantization of a single Gaussian random variable

Let $X \sim \mathcal{N}(0, \sigma^2)$ and let the distortion measure be squared error. Here we do not allow block descriptions. Show that the optimum reproduction points for 1 bit quantization are $\pm\sqrt{\frac{2}{\pi}}\sigma$, and that the expected distortion for 1 bit quantization is $\frac{\pi-2}{\pi}\sigma^2$. Compare this distortion with the rate distortion bound $D = \sigma^2 2^{-2R}$ for $R = 1$.

Solution:

Let $X \sim \mathcal{N}(0, \sigma^2)$ and let the distortion measure be squared error. With one bit quantization, the obvious reconstruction regions are the positive and negative real axes. The reconstruction point is the centroid of each region. For example, for the positive real line, the centroid a is

$$\begin{aligned}
a &= \int_0^\infty x \frac{2}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} dx \\
&= \int_0^\infty \sigma \sqrt{\frac{2}{\pi}} e^{-y} dy \\
&= \sigma \sqrt{\frac{2}{\pi}},
\end{aligned}$$

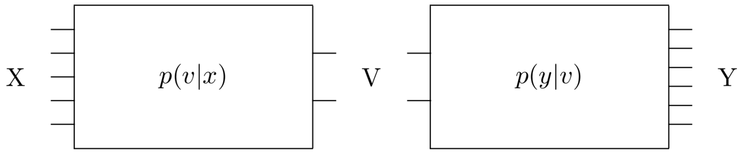
using the substitution $y = x^2/2\sigma^2$. The expected distortion for one bit quantization is

$$\begin{aligned}
D &= \int_{-\infty}^0 \left(x + \sigma\sqrt{\frac{2}{\pi}}\right)^2 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} dx \\
&\quad + \int_0^{-\infty} \left(x - \sigma\sqrt{\frac{2}{\pi}}\right)^2 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} dx \\
&= 2\int_{-\infty}^\infty \left(x + \sigma^2\frac{2}{\pi}\right)^2 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} dx \\
&\quad - 2\int_0^\infty \left(-2x\sigma\sqrt{\frac{2}{\pi}}\right) \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} dx \\
&= \sigma^2 + \frac{2}{\pi}\sigma^4 - 4\frac{1}{\sqrt{2\pi}}\sigma^2\sqrt{\frac{2}{\pi}} \\
&= \sigma^2 \frac{\pi - 2}{\pi}.
\end{aligned}$$

Solution of Problem 2

Bottleneck Channel

Suppose a signal $X \in \mathcal{X} = 1, 2, \ldots, m$ goes through an intervening transition $X \to V \to Y$:



where $x = \{1, 2, \ldots, m\}, y = \{1, 2, \ldots, m\}$, and $v = \{1, 2, \ldots, k\}$. Here $p(v|x)$ and $p(y|v)$ are arbitrary and the channel has transition probability $p(y|x) = \sum_v p(v|x)p(y|v)$. Show that $C \leq \log k$

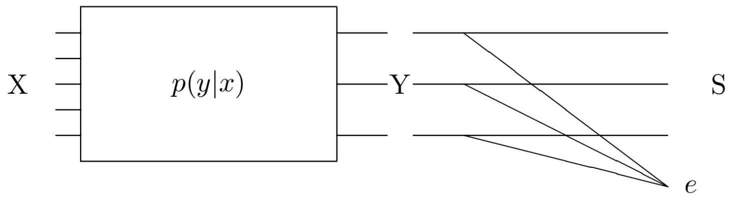
Solution:

The capacity of the cascade of channels is $C = \max_{p(x)} I(X;Y)$. By the data processing inequality, $I(X;Y) \leq I(V;Y) = H(V) - H(V|Y) \leq H(V) \leq logk$.

Solution of Problem 3

Erasure Channel

Let $\{\mathcal{X}, p(y|x), \mathcal{Y}\}$ be a discrete memoryless channel with capacity C. Suppose this channel is immediately cascaded with an erasure channel $\{\mathcal{Y}, p(s|y), \mathcal{S}\}$ that erases an α number of symbols.



Specifically, $\mathcal{S} = \{y_1, \ldots, y_m, e\}$ and

$$P(S = y|X = x) = (1 - \alpha)p(y|x), \quad y \in \mathcal{Y}$$
$$P(S = e|X = x) = \alpha .$$

Determine the capacity of this channel.

Solution:

The capacity of the channel is

$$C = \max_{p(x)} I(X;S) .$$

Define a new random variable Z, a function of S, where $Z = 1$ if $S = e$ and $Z = 0$ otherwise. Note that $p(Z = 1) = \alpha$ independent of X. Expanding the mutual information we get

$$
\begin{aligned}
I(X;S) &= H(S) - H(S|X) \\
&= H(S,Z) - H(S,Z|X) \\
&\quad + H(Z) + H(S|Z) - H(Z|X) - H(S|X,Z) \\
&= I(X;Z) + I(S;X|Z) \\
&= 0 + \alpha I(X;S|Z=1) + (1-\alpha) I(X;S|Z=0)\,.
\end{aligned}
$$

When $Z = 1$, $S = e$ and $H(S|Z = 1) = H(S|X, Z = 1) = 0$. When $Z = 0, S = Y$, and $I(X; S|Z = 0) = I(X; Y)$. Thus

$$
I(X;S) = (1-\alpha) I(X;Y)
$$

and therefore the capacity of the cascade of a channel with an erasure channel is $(1-\alpha)$ times the capacity of the original channel.

Solution of Problem 4

Multiplier Channel

a) Consider the channel $Y = XZ$ where X and Z are independent binary random variables that take on values 0 and 1. Z is a Bernoulli(α), i.e. $P(Z = 1) = \alpha$. Find the capacity of this channel and the maximizing distribution on X .

Solution:

Let $P(X = 1) = p$. Then $P(Y = 1) = P(X = 1)P(Z = 1) = \alpha p$.

$$
\begin{aligned}
I(X;Y) &= H(Y) - H(Y|X) \\
&= H(Y) - P(X=1)H(Z) \\
&= H(\alpha p) - pH(\alpha)\,.
\end{aligned}
$$

Setting $\frac{\partial I(X;Y)}{\partial p} \overset{!}{=} 0$ allows us to find the p that maximizes $I(X; Y)$, that is

$$
p^* = \frac{1}{\alpha(2^{\frac{H(\alpha)}{\alpha}} + 1)\,.}
$$

Therefore, the capacity is

$$
C = H(\alpha p^*) - p^* H(\alpha) = \log(2^{\frac{H(\alpha)}{\alpha}} + 1) - \frac{H(\alpha)}{\alpha}\,.
$$

b) Now suppose the receiver can observe Z as well as Y. What is the capacity?

Solution:

Let $P(X = 1) = p$. Then

$$
\begin{aligned}
I(X;Y,Z) &= I(X;Z) + I(X;Y|Z) \\
&= H(Y|Z) - H(Y|X,Z) \\
&= H(Y|Z) \\
&= \alpha H(p).
\end{aligned}
$$

The expression is maximized for $p = 1/2$, resulting in $C = \alpha$. Intuitively, we can only get X through when Z is 1, which happens α of the time.

Solution of Problem 5

A channel with two independent looks at Y

Let Y_1 and Y_2 be conditionally indeopendent and conditionally identically distributed given X.

a) Show that $I(X;Y_1,Y_2) = 2I(X;Y_1) - I(Y_1;Y_2)$.

Solution:

$$\begin{aligned} I(X;Y_1,Y_2) &= H(Y_1,Y_2) - H(Y_1,Y_2|X) \\ &= H(Y_1) + H(Y_2) - I(Y_1;Y_2) - H(Y_1|X) - H(Y_2|X) \\ &\quad \text{(since } Y_1 \text{ and } Y_2 \text{ are conditionally independent given } X\text{)} \\ &= I(X;Y_1) + I(X;Y_2) - I(Y_1;Y_2) \\ &= 2I(X;Y_1) - I(Y_1;Y_2) \\ &\quad \text{(since } Y_1 \text{ and } Y_2 \text{ are conditionally i.i.d.)} \end{aligned}$$

b) Conclude that the capacity of the channel

$$X \to \boxed{\text{Channel}} \to (Y_1, Y_2)$$

is less than twice the capacity of the channel

$$X \to \boxed{\text{Channel}} \to Y_1\,.$$

Solution:

The capacity of the single look channel $X \to Y_1$ is

$$C_1 = \max_{p(x)} I(X;Y_1)\,.$$

The capacity of the double look channel $X \to (Y_1, Y_2)$ is

$$\begin{aligned} C_2 &= \max_{p(x)} I(X;Y_1,Y_2) \\ &= \max_{p(x)} 2I(X;Y_1) - I(Y_1,Y_2) \\ &\leq \max_{p(x)} 2I(X;Y_1) \\ &= 2C_1 \end{aligned}$$

Hence, two independent looks cannot be more than twice as good as one look.

Solution of Problem 6

Bounds on the rate distortion function for squared error distortion.

For the case of a continuous random variable X with mean zero and variance σ^2 and squared error distortion, show that

a) $h(X) - \frac{1}{2}\log(2\pi e D) \leq R(D)$,

Solution:

We assume that X has zero mean and variance σ^2 . To prove the lower bound, we use the same techniques as used for the Gaussian rate distortion function. Let $(X, \hat{X})$ be random variables such that $E(X - \hat{X})^2 \leq D$. Then

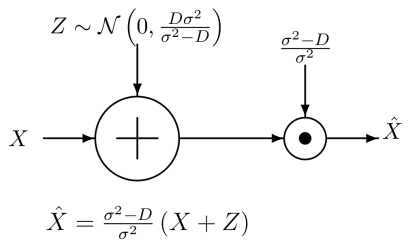
$$\begin{aligned} I(X;\hat{X}) &= h(X) - h(X|\hat{X}) \\ &= h(X) - h(X - \hat{X}|\hat{X}) \\ &\geq h(X) - h(X - \hat{X}) \\ &\geq h(X) - h(\mathcal{N}(0, E(X - \hat{X})^2)) \\ &= h(X) - \frac{1}{2}\log(2\pi e)E(X - \hat{X})^2 \\ &\geq h(X) - \frac{1}{2}\log(2\pi e)D\,. \end{aligned}$$

which combined with the definition of the rate distortion function gives us the required lower bound.

b) $R(D) \leq \frac{1}{2}\log\frac{\sigma^2}{D}$.

Solution:

To prove this upper bound, we consider the following joint distribution



and calculate the distortion and the mutual information between X and $\hat{X}$. Since

$$\hat{X} = \frac{\sigma^2 - D}{\sigma^2}(X + Z)\,,$$

we have

$$\begin{aligned} E(X - \hat{X})^2 &= E(\frac{D}{\sigma^2}X - \frac{\sigma^2 - D}{\sigma^2}Z)^2 \\ &= (\frac{D}{\sigma^2})\sigma^2 + (\frac{\sigma^2 - D}{\sigma^2})^2\frac{D\sigma^2}{\sigma^2 - D} \\ &= D\,, \end{aligned}$$

since X and Z are independent and zero mean. Also the mutual information is

$$I(X;\hat{X}) = h(\hat{X}) - h(\hat{X}|X) = h(\hat{X}) - h(\frac{\sigma^2 - D}{\sigma^2} Z)\,.$$

Now

$$\begin{aligned}
E\hat{X}^2 &= \left(\frac{\sigma^2 - D}{\sigma^2}\right) E(X+Z)^2 \\
&= \left(\frac{\sigma^2 - D}{\sigma^2}\right) (E(X)^2 + E(Z)^2) \\
&= \left(\frac{\sigma^2 - D}{\sigma^2}\right) (\sigma^2 + \frac{D\sigma^2}{\sigma^2 - D}) \\
&= \sigma^2 - D\,.
\end{aligned}$$

Hence, we have

$$\begin{aligned}
I(X;\hat{X}) &= h(\hat{X}) - h(\hat{X}|X) \\
&= h(\hat{X}) - h(\frac{\sigma^2 - D}{\sigma^2} Z) \\
&= h(\hat{X}) - h(Z) - \log\frac{\sigma^2 - D}{\sigma^2} \\
&\le h(\mathcal{N}(0, \sigma^2 - D)) - \frac{1}{2}\log(2\pi e)\frac{D\sigma^2}{\sigma^2 - D} - \log\frac{\sigma^2 - D}{\sigma^2} \\
&= \frac{1}{2}\log(2\pi e)(\sigma^2 - D) - \frac{1}{2}\log(2\pi e)\frac{D\sigma^2}{\sigma^2 - D} - \frac{1}{2}\log\left(\frac{\sigma^2 - D}{\sigma^2}\right)^2 \\
&= \frac{1}{2}\log\frac{\sigma^2}{D}\,,
\end{aligned}$$

which combined with the definition of the rate distortion function gives us the required upper bound.

Remark: For a Gaussian random variable, $h(X) = \frac{1}{2}\log(2\pi e)\sigma^2$ and the lower bound is equal to the upper bound. For any other random variable, the lower bound is strictly less than the upper bound and hence non-Gaussian random variables cannot require more bits to describe to the same accuracy than the corresponding Gaussian random variables. This is not surprising, since the Gaussian random variable has the maximum entropy and we would expect that it would be the most difficult to describe.