Chapter 10

The Gaussian Channel

1. A mutual information game. Consider the following channel:

\[
\begin{aligned}
Z \\
X + \rightarrow Y
\end{aligned}
\]

Throughout this problem we shall constrain the signal power

\[
E X = u, \quad E X^2 = P,
\]

and the noise power

\[
E Z = 0, \quad E Z^2 = N,
\]

and assume that \( X \) and \( Z \) are independent. The channel capacity is given by \( I(X; X + Z) \).

Now for the game. The noise player chooses a distribution on \( Z \) to minimize \( I(X; X + Z) \), while the signal player chooses a distribution on \( X \) to maximize \( I(X; X + Z) \).

Letting \( X^* \sim \mathcal{N}(0, P), Z^* \sim \mathcal{N}(0, N) \), show that \( X^* \) and \( Z^* \) satisfy the saddlepoint conditions

\[
I(X; X + Z^*) \leq I(X^*; X^* + Z^*) \leq I(X^*; X^* + Z).
\]

Thus

\[
\min_Z \max_X I(X; X + Z) = \max_X \min_Z I(X; X + Z) = \frac{1}{2} \log \left( 1 + \frac{P}{N} \right),
\]

131
and the game has a value. In particular, a deviation from normal for either player worsens the mutual information from that player's standpoint. Can you discuss the implications of this?

*Note:* Part of the proof hinges on the entropy power inequality from Section 10.7, which states that if $X$ and $Y$ are independent random $n$-vectors with densities, then

$$e^{\frac{1}{n}h(X+Y)} \geq e^{\frac{1}{n}h(X)} + e^{\frac{1}{n}h(Y)}.$$  \hspace{1cm} (10.6)

**Solution: A mutual information game.**

Let $X$ and $Z$ be random variables with $EX = 0$, $EX^2 = P$, $EZ = 0$ and $EZ^2 = N$. Let $X^* \sim N(0,P)$ and $Z^* \sim N(0,N)$. Then as proved in class,

$$I(X; X + Z^*) = h(X + Z^*) - h(X + Z^* | X)$$

$$= h(X + Z^*) - h(Z^*)$$

$$\leq h(X^* + Z^*) - h(Z^*)$$

$$= I(X^*; X^* + Z^*),$$  \hspace{1cm} (10.7)

where the inequality follows from the fact that given the variance, the entropy is maximized by the normal.

To prove the other inequality, we use the entropy power inequality,

$$2^{2h(X+Z)} \leq 2^{2h(X)} + 2^{2h(Z)}.$$  \hspace{1cm} (10.8)

Let

$$g(Z) = \frac{2^{2h(Z)}}{2\pi e}$$  \hspace{1cm} (10.9)

Then

$$I(X^*; X^* + Z) = h(X^* + Z) - h(X^* + Z | X^*)$$

$$- h(Z)$$

$$\geq \frac{1}{2} \log \left(2^{2h(X^*)} + 2^{2h(Z)} \right) - h(Z)$$

$$= \frac{1}{2} \log \left(2^{2h(X^*)} + 2^{2h(Z)} \right) - 2 \log(2\pi e) g(Z)$$

$$= \frac{1}{2} \log \left(1 + \frac{P}{g(Z)} \right),$$  \hspace{1cm} (10.10)

where the inequality follows from the entropy power inequality. Now $1 + \frac{P}{g(Z)}$ is a decreasing function of $g(Z)$, it is minimized when $g(Z)$ is maximum, which occurs when $h(Z)$ is maximized, i.e., when $Z$ is normal. In this case, $g(Z^*) = N$ and we have the following inequality,

$$I(X^*; X^* + Z) \geq I(X^*; X^* + Z^*).$$  \hspace{1cm} (10.11)
The Gaussian Channel

Combining the two inequalities, we have

\[ I(X; X + Z^*) \leq I(X^*; X^* + Z^*) \leq I(X^*; X^* + Z). \]  \hspace{1cm} (10.19)

Hence, using these inequalities, it follows directly that

\[ \min_Z \max_X I(X; X + Z) \leq \max_X I(X; X + Z^*) \] \hspace{1cm} (10.20)
\[ = I(X^*; X^* + Z^*) \] \hspace{1cm} (10.21)
\[ = \min_Z I(X^*; X^* + Z) \] \hspace{1cm} (10.22)
\[ \leq \max_X \min_Z I(X^*; X^* + Z). \] \hspace{1cm} (10.23)

We have shown an inequality relationship in one direction between \( \min_Z \max_X I(X; X + Z) \) and \( \max_X \min_Z I(X, X + Z) \). We will now prove the inequality in the other direction is a general result for all functions of two variables.

For any function \( f(a, b) \) of two variables, for all \( b \), for any \( a_0 \),

\[ f(a_0, b) \geq \min_a f(a, b). \] \hspace{1cm} (10.24)

Hence

\[ \max_b f(a_0, b) \geq \max_a \min_b f(a, b). \] \hspace{1cm} (10.25)

Taking the minimum over \( a_0 \), we have

\[ \min_{a_0} \max_b f(a_0, b) \geq \min_a \max_b \min_{a_0} f(a, b). \] \hspace{1cm} (10.26)

or

\[ \min_a \max_b f(a, b) \geq \max_{a_0} \min_b f(a, b). \] \hspace{1cm} (10.27)

From this result,

\[ \min_Z \max_X I(X; X + Z) \geq \max_X \min_Z I(X; X + Z). \] \hspace{1cm} (10.28)

From (10.23) and (10.28), we have

\[ \min_Z \max_X I(X; X + Z) = \max_X \min_Z I(X, X + Z) \] \hspace{1cm} (10.29)
\[ = \frac{1}{2} \log \left( 1 + \frac{P}{N} \right). \] \hspace{1cm} (10.30)

This inequality implies that we have a saddlepoint in the game, which is the value of the game. If signal player chooses \( X^* \), the noise player cannot do any better than choosing \( Z^* \). Similarly, any deviation by the signal player from \( X^* \) will make him do worse, if the noise player has chosen \( Z^* \). Any deviation by either player will make him do worse.

Another implication of this result is that not only is the normal the best possible signal distribution, it is the worst possible noise distribution.
2. A channel with two independent looks at Y. Let \( Y_1 \) and \( Y_2 \) be conditionally independent and conditionally identically distributed given \( X \).

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

(b) Conclude that the capacity of the channel

\[
X \rightarrow (Y_1, Y_2)
\]

is less than twice the capacity of the channel

\[
X \rightarrow Y
\]

Solution: A channel with two independent looks at Y. Channel with two independent looks at Y.

(a)

\[
I(X; Y_1, Y_2) = H(Y_1, Y_2) - H(Y_1, Y_2|X) \tag{10.31}
\]

\[
- H(Y_1) + H(Y_2) - I(Y_1; Y_2) - H(Y_1|X) - H(Y_2|X) \tag{10.32}
\]

(since \( Y_1 \) and \( Y_2 \) are conditionally independent given \( X \))

\[
= I(X; Y_1) + I(X; Y_2) - I(Y_1; Y_2) \tag{10.33}
\]

\[
- 2I(X; Y_1) - I(Y_1; Y_2) \tag{since \( Y_1 \) and \( Y_2 \) are conditionally identically distributed}
\]

(b) The capacity of the single look channel \( X \rightarrow Y_1 \) is

\[
C_1 = \max_{p(x)} I(X; Y_1). \tag{10.34}
\]

The capacity of the channel \( X \rightarrow (Y_1, Y_2) \) is

\[
C_2 = \max_{p(x)} I(X; Y_1, Y_2) \tag{10.35}
\]

\[
= \max_{p(x)} 2I(X; Y_1) - I(Y_1; Y_2) \tag{10.36}
\]

\[
= \max_{p(x)} 2I(X; Y_1) \tag{10.37}
\]

\[
- 2C_1 \tag{10.38}
\]

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

3. The two-look Gaussian channel.

\[
X \rightarrow (Y_1, Y_2)
\]
The Gaussian Channel

Consider the ordinary Shannon Gaussian channel with two correlated looks at $X$, i.e., $Y = (Y_1, Y_2)$, where

$$
Y_1 = X + Z_1 \quad (10.41)
$$
$$
Y_2 = X + Z_2 \quad (10.42)
$$

with a power constraint $P$ on $X$, and $(Z_1, Z_2) \sim \mathcal{N}_2(0, K)$, where

$$
K = \begin{bmatrix}
N & N\rho \\
N\rho & N
\end{bmatrix}. \quad (10.43)
$$

Find the capacity $C$ for

(a) $\rho = 1$
(b) $\rho = 0$
(c) $\rho = -1$

**Solution:** The two look Gaussian channel

It is clear that the input distribution that maximizes the capacity is $X \sim \mathcal{N}(0, P)$. Evaluating the mutual information for this distribution,

$$
C_2 = \max I(X; Y_1, Y_2) \quad (10.44)
$$

$$
= h(Y_1, Y_2) - h(Y_1, Y_2 | X) \quad (10.45)
$$

$$
= h(Y_1, Y_2) - h(Z_1, Z_2 | X) \quad (10.46)
$$

$$
= h(Y_1, Y_2) - h(Z_1, Z_2) \quad (10.47)
$$

Now since

$$(Z_1, Z_2) \sim \mathcal{N} \left( 0, \begin{bmatrix}
N & N\rho \\
N\rho & N
\end{bmatrix} \right), \quad (10.48)$$

we have

$$
h(Z_1, Z_2) = \frac{1}{2} \log(2\pi e)^2 |K_2| = \frac{1}{2} \log(2\pi e)^2 N^2 (1 - \rho^2). \quad (10.49)
$$

Since $Y_1 = X + Z_1$, and $Y_2 = X + Z_2$, we have

$$(Y_1, Y_2) \sim \mathcal{N} \left( 0, \begin{bmatrix}
P + N & P + \rho N \\
P + \rho N & P + N
\end{bmatrix} \right), \quad (10.50)$$

and

$$
h(Y_1, Y_2) = \frac{1}{2} \log(2\pi e)^2 |K_Y| = \frac{1}{2} \log(2\pi e)^2 (N^2 (1 - \rho^2) + 2PN(1 - \rho)). \quad (10.51)
$$

Hence the capacity is

$$
C_2 = h(Y_1, Y_2) - h(Z_1, Z_2) \quad (10.52)
$$

$$
= \frac{1}{2} \log \left( 1 + \frac{2P}{N(1 + \rho)} \right). \quad (10.53)
$$
(a) $\rho = 1$. In this case, $C = \frac{1}{2} \log (1 + \frac{P^2}{N})$, which is the capacity of a single look channel. This is not surprising, since in this case $Y_1 = Y_2$.

(b) $\rho = 0$. In this case,

$$C = \frac{1}{2} \log \left( 1 + \frac{2P}{N} \right),$$

(10.54)

which corresponds to using twice the power in a single look. The capacity is the same as the capacity of the channel $X \to (Y_1 + Y_2)$.

(c) $\rho = -1$. In this case, $C = \infty$, which is not surprising since if we add $Y_1$ and $Y_2$, we can recover $X$ exactly.

Note that the capacity of the above channel in all cases is the same as the capacity of the channel $X \to Y_1 + Y_2$.

4. Parallel channels and waterfilling. Consider a pair of parallel Gaussian channels, i.e.,

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} + \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix},$$

(10.55)

where

$$\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \sim \mathcal{N} \left( 0, \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} \right),$$

(10.56)

and there is a power constraint $E(X_1^2 + X_2^2) \leq 2P$. Assume that $\sigma_1^2 > \sigma_2^2$. At what power does the channel stop behaving like a single channel with noise variance $\sigma_2^2$, and begin behaving like a pair of channels?

Solution: Parallel channels and waterfilling. By the result of Section 10.4, it follows that we will put all the signal power into the channel with less noise until the total power of noise + signal in that channel equals the noise power in the other channel. After that, we will split any additional power evenly between the two channels.

Thus the combined channel begins to behave like a pair of parallel channels when the signal power is equal to the difference of the two noise powers, i.e., when $2P = \sigma_1^2 - \sigma_2^2$. 