Exercise 8.9 (Johnson Jonaris GadElkarim)

Gaussian mutual information. Suppose that \((X, Y, Z)\) are jointly Gaussian and that \(X \rightarrow Y \rightarrow Z\) forms a Markov chain. Let \(X\) and \(Y\) have correlation coefficient \(\rho_1\) and let \(Y\) and \(Z\) have correlation coefficient \(\rho_2\). Find \(I(X; Z)\).

Solution

\[ I(X; Z) = h(X) + h(Z) - h(X, Z) \]

Since \(X, Y, Z\) are jointly Gaussian, hence \(X\) and \(Z\) are jointly Gaussian, their covariance matrix will be:

\[
K = \begin{bmatrix}
\sigma_x^2 & \sigma_x \sigma_z \rho_{xz} \\
\sigma_x \sigma_z \rho_{xz} & \sigma_z^2
\end{bmatrix}
\]

Hence

\[
I(X; Z) = 0.5 \log(2\pi e \sigma_x^2) + 0.5 \log(2\pi e \sigma_z^2) - 0.5 \log(2\pi e |K|)
\]

\[
|K| = \sigma_x^2 \sigma_z^2 (1 - \rho_{xz}^2)
\]

\[
I(X; Y) = -0.5 \log(1 - \rho_{xz}^2)
\]

Now we need to compute \(\rho_{xz}\), using markovity \(p(x, z | y) = p(x | y) p(z | y)\) we can get

\[
\rho_{xz} = \frac{E(xz)}{\sigma_x \sigma_z} = \frac{E((xz | y))}{\sigma_x \sigma_z} = \frac{E(E(x | y) E(z | y))}{\sigma_x \sigma_z}
\]

Since \(X, Y\) and \(Z\) are jointly Gaussian: \(E(x | y) = \frac{\sigma_x \rho_{xy}}{\sigma_y} Y\), we can do the same for \(E(z | y)\)

\[
\rho_{xz} = \rho_{xy} \rho_{zy}
\]

\[
I(X; Z) = -0.5 \log(1 - (\rho_{xz})^2)
\]

Exercise 9.2 (Johnson Jonaris GadElkarim)

Two-look Gaussian channel. Consider the ordinary Gaussian channel with two correlated looks at \(X\), that is, \(Y = (Y_1, Y_2)\), where

\[
Y_1 = X + Z_1 \\
Y_2 = X + Z_2
\]
with a power constraint $P$ on $X$, and $(Z_1, Z_2) \sim \mathcal{N}_2(0, K)$, where

$$K = \begin{bmatrix} N & N\rho \\ N\rho & N \end{bmatrix}.$$ 

Find the capacity $C$ for

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

**Solution** The capacity will be

$$C = \max I(X; Y_1, Y_2)$$

$$I(X; Y_1, Y_2) = h(Y_1, Y_2) - h(Y_1, Y_2|X) = h(Y_1, Y_2) - h(Z_1, Z_2)$$

$$h(Z_1, Z_2) = 0.5 \log(2\pi e)^2|k| = 0.5 \log(2\pi e)^2 N^2(1 - \rho^2)$$

The mutual information will be maximized when $Y_1, Y_2$ are jointly Gaussian with covariance matrix $K_y = P.I_{2 \times 2} + K_z$ where $I_{2 \times 2}$ is an identity matrix of dimension 2.

$$|K_y| = N^2(1 - \rho^2) + 2PN(1 - \rho)$$

Hence the capacity will be: $C = 0.5 \log(1 + 2P/N(1 + \rho))$

(a) $\rho = 1$, $C = 0.5 \log(1 + P/N)$
(b) $\rho = 0$, $C = 0.5 \log(1 + 2P/N)$
(c) $\rho = -1$, $C = 0.5 \log(1 + \infty) = \infty$

**Exercise 9.4 (Shu Wang)**

*Exponential noise channel. $Y_i = X_i + Z_i$, where $Z_i$ is i.i.d. exponentially distributed noise with mean $\mu$. Assume that we have a mean constraint on the signal (i.e. $E[X_i] \leq \lambda$). Show that the capacity of such a channel is $C = \log \left(1 + \frac{\lambda}{\mu}\right)$.*

**Solution**

From the textbook, maximize the entropy $h(f)$ over all probability densities $f$ satisfying

1. $f(x) \geq 0$ with equality outside the support set $S$
2. $\int_{S} f(x)dx = 1$
3. $\int_{S} f(x)dx = \alpha_i$ for $1 \leq i \leq m$
In this problem the support set is $[0, \infty)$. Over this support set, if $f(x)$ is Gaussian, $\int_S f(x)dx \neq 1$. So Gaussian cannot maximize the entropy. But if $f(x) = \lambda e^{-\lambda x}$, we can find that $\int_S f(x)dx = 1$. So if $f(x)$ is exponential distribution, $h(x)$ can be maximized. We can prove this:

$$f_e(x) = \lambda e^{-\lambda x}$$

$$D(f||f_e) = \int f(x) \log \frac{f(x)}{\lambda e^{-\lambda x}} dx$$

$$= \int f(x)(\log f(x) - \log(\lambda e^{-\lambda x})) dx$$

$$= -h(x) - \int f(x) \log(\lambda e^{-\lambda x}) dx$$

$$= -h(x) + h(x_e) \geq 0$$

So $h(x_e) \geq h(x)$. So when $f(x)$ is exponential distributed, $h(x)$ will be maximized.

Because $EX_i \leq \lambda$, $EY_i = E[X_i + Z_i] = E[X_i] + E[Z_i] \leq \mu + \lambda$. So

$$C = \max I(X_i; Y_i)$$

$$= \max [h(Y_i) - h(Y_i|X_i)]$$

$$= \max [h(Y_i) - h(X_i + Z_i|X_i)]$$

$$= \max [h(Y_i) - h(Z_i|X_i)]$$

$$= \max [h(Y_i) - h(Z_i)]$$

$$= h_e(Y_i) - h(Z_i)$$

Because $EZ_i = \mu$, $f(Z_i) = \frac{1}{\mu} e^{-\frac{z}{\mu}}$. Suppose $a = \frac{1}{\mu}$

$$h(Z_i) = -\int ae^{-az} \log ae^{-az} dz$$

$$= \log_2(e)$$

$$= \log_2(e\mu)$$

According to the analysis above, $h(Y_i) \leq \log_2(e(\mu + \lambda))$. So $C \leq \log_2(1 + \frac{\lambda}{\mu})$

If we want $C = \log_2(1 + \frac{\lambda}{\mu})$, that means we can find $X_i$ which makes $X_i + Z_i$ is an exponential distribution with mean $\mu + \lambda$. I will show how to find it.

The characteristic function of $Z$ with mean $\mu$ is
\[
\Phi_Z(\omega) = E[e^{j\omega z}]
= \int_0^\infty e^{j\omega z} e^{-\frac{z}{\mu}} dz
= \frac{1}{\mu} \frac{1}{\mu - j\omega}
\]

Also we can have \(\Phi_Y(\omega) = \frac{1}{\mu + \lambda} \frac{1}{\mu + \lambda - j\omega}\). Because \(\Phi_Y(\omega) = \Phi_X(\omega)\Phi_Z(\omega)\). We will get:

\[
\Phi_X(\omega) = \frac{\Phi_Y(\omega)}{\Phi_Z(\omega)}
= \frac{1 - j\mu\omega}{1 - j(\mu + \lambda)\omega}
= \frac{1}{1 - j(\mu + \lambda)\omega} - \frac{j\mu\omega}{1 - j(\mu + \lambda)\omega}
= \frac{1}{1 - j(\mu + \lambda)\omega} - \frac{\mu}{\mu + \lambda} \left(1 - \frac{1}{1 - j(\mu + \lambda)\omega}\right) + \frac{\mu}{\mu + \lambda}
= \frac{\lambda}{\mu + \lambda} \left(1 - \frac{1}{1 - j(\mu + \lambda)\omega}\right) + \frac{\mu}{\mu + \lambda}
\]

If we do the inverse, we will get \(f(x) = (\frac{\lambda}{\mu + \lambda}) (\frac{1}{\mu + \lambda} e^{-\frac{x}{\mu + \lambda}}) + \frac{\mu}{\mu + \lambda} \delta(x)\) and \(x \geq 0\)

So we can find a \(X\) which makes \(C = \log_2(1 + \frac{\lambda}{\mu})\)

**Exercise 9.8 (Johnson Jonaris GadElkarim)**

*Parallel Gaussian channels.* Consider the following parallel Gaussian channel: [see the figures in the book], where \(Z_1 \sim \mathcal{N}(0, N_1)\) and \(Z_2 \sim \mathcal{N}(0, N_2)\) are independent Gaussian random variables and \(Y_i = X_i + Z_i\). We wish to allocate power to the two parallel channels. Let \(\beta_1\) and \(\beta_2\) be fixed. Consider a total cost constraint \(\beta_1 P_1 + \beta_2 P_2 \leq \beta\), where \(P_i\) is the power allocated to the \(i\)th channel and \(\beta_i\) is the cost per unit power in the channel. Thus, \(P_1 \geq 0\) and \(P_2 \geq 0\) can be chosen subject to the cost constraint \(\beta\).

(a) For what value of \(\beta\) does the channel stop acting like a single channel and start acting like a pair of channels?

(b) Evaluate the capacity and find \(P_1\) and \(P_2\) that achieve the capacity for \(\beta_1 = 1\), \(\beta_2 = 2\), \(N_1 = 3\), \(N_2 = 2\) and \(\beta = 10\).
**Solution** a) We have power budget: $\beta_1 P_1 + \beta_2 P_2 \leq \beta$

The Lagrange function will be:

$$J(P_1, P_2) = \sum_{i=1}^{2} \log(1 + P_i/N_i) + \lambda(\beta_1 P_1 + \beta_2 P_2)$$

Differentiating with respect to $P_i$ and equating to zero and noting that the power must be positive we get:

$$\beta_i P_i = \left(-\frac{1}{2}\lambda - \beta_i N_i\right)^+$$

For the 2 channels to act as a pair of channels we need $\beta$ to fill the gap between them, i.e. $\beta \geq |\beta_1 N_1 - \beta_2 N_2|

---

### Acting as 2 Channels

<table>
<thead>
<tr>
<th>$\beta_1 N_1$</th>
<th>$\beta_2 N_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

b) $C = 0.5\left(\log(1 + P_1/N_1) + \log(1 + P_2/N_2)\right)$

$$P_1 + P_2 = 10$$

Since $\beta_1 N_1 = 3$, $\beta_2 N_2 = 4$, we will fill the gap first with 1 (filling the first channel), then the rest 9 will be divided equally onto the 2 channels.

Hence $P_1 = 5.5$, $P_2 = 4.5/2 = 2.25$

---

**Exercise 9.9 (Matteo Carminati)**

*Vector Gaussian channel.* Consider the vector Gaussian noise channel

$$Y = X + Z$$

where $X = (X_1, X_2, X_3)$, $Z = (Z_1, Z_2, Z_3)$, $Y = (Y_1, Y_2, Y_3)$, $E||X||^2 \leq P$, and

$$Z \sim \mathcal{N}\left(0, \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}\right).$$

Find the capacity. The answer may be surprising.

**Solution**

The channel presented in this exercise can be modelled by considering a group of three channels with colored Gaussian noise. In order to optimize the distribution of power among the channels,
we must apply the algorithm described in the book. First of all, the covariance matrix of the noise must be diagonalized; to do that its eigenvalues and its eigenvectors must be computed:

\[
p(x) = |K_Z - xI| = \begin{bmatrix} 1 - x & 0 & 1 \\ 0 & 1 - x & 1 \\ 1 & 1 & 2 - x \end{bmatrix} = (1 - x)((1 - x)(2 - x) - 1) - (1 - x)
= (1 - x)x(x - 3)
\]

Thus, the eigenvalues of this matrix are \(\lambda_1 = 1, \lambda_2 = 0\) and \(\lambda_3 = 3\). Since one of the eigenvectors is 0 and since the determinant of a squared matrix can be computed as the product of the eigenvectors, the determinant of \(K_Z\) is 0. This means that the columns (or rows since the matrix is symmetric) are linearly dependent: in particular we can see that the last column (row), can be computed as the summation of the first two columns (rows).

Since \(K_Z\) is a covariance matrix, this implies that \(Z_3\) can be rewritten as a function of \(Z_1\) and \(Z_2\): \(Z_3 = Z_1 + Z_2\). As seen in exercise 9.2 this fact can be exploited to nullify the effect of the noise. In particular if the same signal \(X = X_1 = X_2 = X_3\) is sent over the three channels, its value can be perfectly derived by subtracting \(Y_3\) to the summation of \(Y_1\) and \(Y_2\), in fact:

\[
Y_1 + Y_2 - Y_3 = X_1 + Z_1 + X_2 + Z_2 - X_3 - Z_3
= X + Z_1 + X + Z_2 - X - Z_1 - Z_2 = X
\]

Thus, as in exercise 9.2, the capacity of the channel can be considered to be infinite!