ECE 231A HW 5

Lawrence Liu

December 1, 2022

Problem 1

(a)

First let us prove that

$$\min_i \frac{a_i}{b_i} \leq \frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n b_i} \leq \max_i \frac{a_i}{b_i}$$

We can prove this through induction, we already have that the base case when $n = 2$ is true. Now we consider the case the case of $n+1$, let us arrange the a_i and b_i in such a way such that $\frac{a_{n+1}}{b_{n+1}}$ be the minimum of the $n+1$ $\frac{a_i}{b_i}$'s. Then we have that

$$\frac{a_{n+1}}{b_{n+1}} \leq \min_{1\leq i\leq n} \frac{a_i}{b_i} \leq \frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n b_i}$$

Therefore we have that

$$\min_i \frac{a_i}{b_i} \leq \frac{\sum_{i=1}^{n+1} a_i}{\sum_{i=1}^{n+1} b_i} \leq \frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n b_i} \leq \max_i \frac{a_i}{b_i}$$

Thus we have proven that $\frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n b_i}$ is bounded by the minimum and maximum of the $\frac{a_i}{b_i}$'s. Using this property we get that

$$\frac{\sum_{x\in\mathcal{X}} \tilde{P}_X(x) D(W_{Y|X}||P_Y)}{\sum_{x\in\mathcal{X}} \tilde{P}_X(x)c(x)} \leq \max_{x\in\mathcal{X}} \frac{\tilde{P}_X(x) D(W_{Y|X}||P_Y)}{\tilde{P}_X(x)c(x)} = \max_{x\in\mathcal{X}} \frac{D(W_{Y|X}||P_Y)}{c(x)}$$

(b)

We have that

Problem 2

(a)

We have that in order to maximize the differential entropy of Z we want to maximize the variance of Z, this is given by

$$\mathrm{Var}(Z) = E[(Z_1 + Z_2)^2] - (E[Z_1 + Z_2])^2$$

Since Z_1 is independent of Z_2 we have that $E[Z_1 Z_2] = E[Z_1]E[Z_2]$, thus we get that

$$\mathrm{Var}(Z) = E[Z_1^2] + E[Z_2^2] - E^2[Z_1] - E^2[Z_2] \leq 2\sigma^2 - E^2[Z_1] - E^2[Z_2]$$

Therefore in order for this to be maximized, we want $E[Z_1] = E[Z_2] = 0$. Thus we get that both Z_1 and Z_2 are Gaussian with mean 0. And that the maximal differential entropy of Z is given by $\frac{1}{2}\log(4\pi e\sigma^2)$.

(b)

Once again we want to maximize the differential entropy of Z, and thus we want to maximize the variance of Z, this is given by

$$\text{Var}(Z) = E[(\sum_{i=1}^{n} Z_i)^2] - (E[\sum_{i=1}^{n} Z_i])^2$$

We have that

$$\left(\sum_{i=1}^{n} Z_i\right)^2 = \sum_{i=1}^{n} Z_i^2 + \sum_{i \neq j} Z_i Z_j$$

And

$$E^2[\sum_{i=1}^{n} Z_i] = \sum_{i=1}^{n} E^2[Z_i] + \sum_{i \neq j} E[Z_i]E[Z_j]$$

Since Z_i's are independent we have that

$$E[Z_i Z_j] - E[Z_i]E[Z_j] = 0$$

For all $i \neq j$, thus we get that

$$\text{Var}(Z) = \sum_{i=1}^{n} E[Z_i^2] - \sum_{i=1}^{n} E^2[Z_i] = n\sigma^2 - \sum_{i=1}^{n} E^2[Z_i]$$

Therefore we must have that $E[Z_i] = 0$ for all i, thus we get that the maximum differential entropy of Z is given by $\boxed{\frac{1}{2}\log(2\pi e n\sigma^2)}$.

(c)

We have that

$$\text{Var}(Z) = E[(\sum_{i=1}^{n} Z_i)^2] - (E[\sum_{i=1}^{n} Z_i])^2$$

We can rewrite this as

$$\begin{aligned}\text{Var}(Z) &= E[\sum_{i=1}^{n} Z_i^2] - E[\sum_{i=1}^{n} Z_i]^2 + \sum_{i\neq j} E[Z_iZ_j] - E[Z_i]E[Z_j] \\ &= \sum_{i=1}^{n} Var(Z_i) + \sum_{i\neq j} Cov(Z_i, Z_j)\end{aligned}$$

We have that $Cov(Z_i, Z_j) \leq \sqrt{Var(Z_i)Var(Z_j)}$, and since $Var(Z_i) \leq \sigma^2$ we get that

$$Var(z) \leq n^2\sigma^2$$

With equality only happening when $E[Z_i] = 0$, therefore we get that the joint distribution of Z_i's as expressed by a vector $\mathbf{Z} = (Z_1, Z_2, \ldots, Z_n)$ is distrubuted as a multivariate Gaussian with mean $\mu = [0, \ldots, 0]$ and covariance matrix Σ where $\Sigma_{ij} = \sigma^2$ for all $1 \leq i, j \leq n$. And with this the maximum differential entropy of Z is given by $\boxed{\frac{1}{2}\log_2(2\pi e n^2\sigma^2)}$.

Problem 3

(a)

We have:

$$\begin{aligned}I(Y_1, Y_2; X) &= H(Y_1, Y_2) - H(Y_1, Y_2|X) \\ &= H(Y_1, Y_2) - H(Z_1, Z_2) \\ &= H(Y_1, Y_2) - H(Z_1) - H(Z_2) \\ &= H(Y_{1,2}) - \log_2(2\pi e\sigma^2) \\ &\leq H(Y_1) + H(Y_2) - \log_2(2\pi e\sigma^2)\end{aligned}$$

Since $E[Y_1^2] = E[Y_2^2] = \sigma^2 + P$ we get that $H(Y_1) \leq \frac{1}{2}\log_2(2\pi e(P + \sigma^2))$ we have:

$$\begin{aligned}I(Y_1, Y_2; X) &\leq \log_2(2\pi e(P + \sigma^2)) - \log_2(2\pi e\sigma^2) \\ &= \log_2\left(\frac{P}{\sigma^2} + 1\right)\end{aligned}$$

Therefore we get that the channel capacity is $\boxed{\log_2\left(\frac{P}{\sigma^2}+1\right)}$ and that we can achieve this capacity by having an X distributed as a Gaussian with mean 0 and variance σ^2.

(b)

We have that we want to maximize

$$\begin{aligned}
I(Y_1, Y_2, Y_3, Y_4; X) &= H(Y_1, Y_2, Y_3, Y_4) - H(Y_1, Y_2, Y_3, Y_4|X) \\
&= H(Y_1, Y_2, Y_3, Y_4) - H(Z_1, Z_2, Z_3, Z_4) \\
&\leq H(Y_1) + H(Y_2) + H(Y_3) + H(Y_4) - \log_2(2\pi e\sigma_1^2) - \log_2(2\pi e\sigma_2^2) \\
&\leq \log_2\left(\left(\frac{P_1}{\sigma_1^2}+1\right)\right) + \log_2\left(\left(\frac{P_2}{\sigma_2^2}\right)\right)
\end{aligned}$$

Therefore we have the optimization problem for the channel capacity is just maximizing $\log_2\left(\frac{P_1}{\sigma_1^2}+1\right)+\log_2\left(\frac{P_2}{\sigma_2^2}\right)$ subject to the constrain that $P_1+P_2 = P$.

(c)

Using a langrangian multiplier we get that to find the optimal solution we must find P_1 and P_2 that satisfy

$$\nabla f(P_1, P_2) = 0$$

where

$$f(P_1, P_2) = \log_2\left(\frac{P_1}{\sigma_1^2}+1\right) + \log_2\left(\frac{P_2}{\sigma_2^2}+1\right) + \lambda(P_1 + P_2 - P)$$

We have that

$$\frac{\partial f}{\partial P_1} = \frac{1}{P_1 + \sigma_1^2} + \lambda = 0$$

and

$$\frac{\partial f}{\partial P_2} = \frac{1}{P_2 + \sigma_2^2} + \lambda = 0$$

Therefore we get that

$$P_2 = -\frac{1}{\lambda} - \sigma_2^2$$

and

$$P_1 = -\frac{1}{\lambda} - \sigma_1^2$$

Thus we get that

$$\lambda = -\frac{2}{\sigma_1^2 + \sigma_2^2 + P}$$

and