Information Theory
Homework #3

1. Find optimal quantizer $\phi : \mathbb{R} \to \{a, b\}$ for zero-mean Gaussian r.v. X with variance σ^2.

Claim: $a = -\sqrt{\frac{2}{\pi}}\,\sigma$, $b = \sqrt{\frac{2}{\pi}}\,\sigma$ minimizes MSE $\mathbb{E}[(X - \phi(X))^2]$.

Pf: We first prove $\mathbb{E}[\phi(X)] = \mathbb{E}[X - \phi(X)] = 0$.

If not, we would have

$$\mathbb{E}[(X - \phi(X))^2] > \mathbb{E}[(X - \phi(X) - \mathbb{E}[X - \phi(X)])^2]$$

but $\tilde{\phi}(X) = \phi(X) + \mathbb{E}[X - \phi(X)]$

has range $\operatorname{ran}(\tilde{\phi}) = \{a + \mathbb{E}[X - \phi(X)],\ b + \mathbb{E}[X - \phi(X)]\}$

so is an admissible quantizer, better than ϕ! Since ϕ was assumed optimal,

$\mathbb{E}[\phi(X)] = \mathbb{E}[X] = 0$.

Next we argue that ϕ must round to nearest of a, b. If not, re-assigning ϕ would lower MSE.

Thus,

$$0 = \mathbb{E}[\phi(X)] = a \int_{-\infty}^{c} \underbrace{\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}x^2}}_{p(x)} dx$$

$$+ b \int_{c}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}x^2} dx$$

and $c = \frac{a+b}{2}$. Writing $a = c - \delta$, $b = c + \delta$, $\delta > 0$, gives

$$0 = (c-\delta) \int_{-\infty}^{c} p(x)\,dx + (c+\delta) \int_{c}^{\infty} p(x)\,dx.$$

We note differentiating w.r.t. δ gives

$$0 = -\underbrace{\int_{-\infty}^{c} p(x)\,dx}_{\text{monotonic in } c} + \int_{c}^{\infty} p(x)\,dx$$

$\Rightarrow$ $c = 0$ is unique solution.

Finally, we claim

$$\delta = 2 \int_{0}^{\infty} x\, p(x)\,dx.$$

Otherwise, $\mathbb{E}[(X-\phi(X))^2 \mid X \geq 0]$ could be reduced by making $\mathbb{E}[X-\phi(X) \mid X \geq 0] = 0$.

We compute

$$\delta = 2\int_0^\infty x \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}x^2}\, dx$$

$$= \int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}u}\, du = \frac{2\sigma^2}{\sqrt{2\pi\sigma^2}} = \sqrt{\frac{2}{\pi}}\,\sigma$$

$$\Rightarrow \quad a = -\sqrt{\frac{2}{\pi}}\,\delta, \quad b = \sqrt{\frac{2}{\pi}}\,\delta.$$

2. Exp. dist. has density

$$q(x) = \mu e^{-\mu x} \quad \text{on } \mathbb{R}^+$$

and μ is expected value.

Note: For $\underline{\text{any}}$ density p on $\mathbb{R}^+$, $xp \in L^1(\mathbb{R}^+)$, with expected value μ,

$$-\int_{\mathbb{R}^+} p(x) \ln q(x)\, dx$$

$$= -\int_{\mathbb{R}^+} p(x) (\ln(\mu) - \mu x)\, dx$$

$$= -\ln\mu + \mu^2 = -\int_{\mathbb{R}^+} q(x) \ln q(x)\, dx$$

Consequently,

$$h(Y) - h(X) = -\int_{\mathbb{R}^+} q \ln q \, dx + \int_{\mathbb{R}^+} p \ln p \, dx$$

$$= + \int p (\ln p - \ln q) \, dx$$

$$= - \int p \ln \frac{q}{p} \, dx$$

$$= D(X \| Y) \geq 0$$

and $h(X) \leq h(Y)$, with equality iff

$q(x) = p(x)$ a.e.

3. a) Let $E = \begin{cases} 1, & X \neq Y \\ 0, & \text{else} \end{cases}$

We have $H(X|Y) = H(E|Y)$

because $\phi(X,Y) = (E,Y)$

is 1-1, and $\mathbb{P}(X = a \mid Y = b)$

$= \mathbb{P}(E = |a-b| \mid Y = b)$

for all $a, b \in \{0,1\}$.

Now

$$I(X;Y) = H(X) - H(X|Y)$$

$$= h(p) - H(E|Y)$$

$$\geq h(p) - H(E)$$

$$\geq h(p) - h(\varepsilon)$$

where we assumed $\varepsilon \leq p < \frac{1}{2}$, and used that $h(\varepsilon)$ is increasing on $[0, \frac{1}{2})$.

b) We compute

$$\mathbb{P}(X=0) = \mathbb{P}(X=0 \mid Y=0)\, \mathbb{P}(Y=0) + \mathbb{P}(X=0 \mid Y=1)\, \mathbb{P}(Y=1)$$

$$= (1-\varepsilon)\frac{1-p-\varepsilon}{1-2\varepsilon} + \varepsilon \frac{p-\varepsilon}{1-2\varepsilon}$$

$$= \frac{1-p-\varepsilon - \varepsilon + \varepsilon(p+\varepsilon) + \varepsilon(p-\varepsilon)}{1-2\varepsilon} = \frac{1-(1-2\varepsilon)p - 2\varepsilon}{1-2\varepsilon} = 1-p$$

consequently, $\mathbb{P}(X=1) = p$

and since $H(E|Y) = H(E)$, following same as a)

E indep. of Y

$$I(X;Y) = h(p) - H(E)$$

$$= h(p) - h(\varepsilon).$$

c) By lower bound from a), and

$$\mathbb{E}\left[\frac{1}{n} d_n(X,Y)\right] = \mathbb{P}(E=1) = \varepsilon = D$$

for E in b), we have

$$R(D) = h(p) - h(D), \quad \text{if } D \le p.$$

Moreover, if $D > p$, set $Y_j = 1$ regardless of input, then X_j, Y_j are indep. and

$$I(X_1, \dots X_n ; Y_1, \dots Y_n) = 0$$

but

$$\mathbb{E}\left[\frac{1}{n} d_n(X,Y)\right] = \mathbb{P}(X_1 = 1) = 1-p < D$$

so this choise is allowed and thus

$$R(D) = 0, \quad \text{if } D > p.$$

4. Eigenvalues of $C_N = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \rho \\ 0 & \rho & 1 \end{pmatrix}$

are $1+\rho$, $1-\rho$, 1. So, change of sign for ρ amounts to reordering $\sigma_1^2 = 1$, $\sigma_2^2 = 1-\rho$, $\sigma_3^2 = 1+\rho$. Enough to consider $\rho \geq 0$.

By result from class,

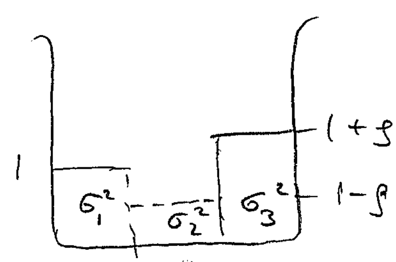
$$C(S) = \sum_{j=1}^{3} \frac{1}{2} \log\left(1 + \frac{S_j}{\sigma_j^2}\right)$$

where

$$S_j = (\theta - \sigma_j^2)$$

and θ is chosen s.th. $\sum_{j=1}^{3} S_j = S$.

Water filling sketch



If $\theta \leq 1$, then only $S_2 \geq 0$

and thus $\sum_{j=1}^{3} S_j = S_2 = S$,

and $S_2 + \sigma_2^2 \leq \sigma_1^2 \iff S \leq \sigma_1^2 - \sigma_2^2$

$= \rho$.

Conversely, if S_2 is only non-zero S_i, we know $\theta \le 1$. We summarize

Case I: If $S \le \rho$, then

$$C(S) = \frac{1}{2} \log \left(1 + \frac{S}{1-\rho}\right).$$

Now assume $1 \le \theta \le 1 + \rho$, then

$$S_1 = \theta - \sigma_1^2$$

$$S_2 = \theta - \sigma_2^2, \quad S_3 = 0$$

$$\Rightarrow S = S_1 + S_2 + S_3 = 2\theta - (1 + 1 - \rho)$$

$$\le 2(1+\rho) - (2-\rho)$$

$$= 3\rho.$$

By monotonicity of S in θ, if $\rho \le S \le 3\rho$, then $S_3 = 0$ and we have:

Case II: If $\rho \le S \le 3\rho$

$$(S = 2\theta - (1 + 1 - \rho) \Rightarrow \theta = \frac{S}{2} + 1 - \rho/2$$

$$\Rightarrow C(S) = \frac{1}{2} \log \left(1 + \frac{S-\rho}{2(1-\rho)}\right) + \frac{1}{2} \log \left(1 + \frac{S+\rho}{2(1-\rho)}\right)$$

Finally, the remaining case $S > 3\rho$ gives all $S_j \neq 0$ and

Case III: $S = S_1 + S_2 + S_3 = 3\theta - (1 + 1 - \cancel{\rho} + 1 + \cancel{\rho})$

$$\Rightarrow \quad \theta = \frac{S}{3} + 1$$

$$\Rightarrow \quad C(S) = \frac{1}{2}\log\left(1 + \frac{S}{3(1-\rho)}\right) + \frac{1}{2}\log\left(1 + \frac{S/3 + \rho}{1-\rho}\right) + \frac{1}{2}\log\left(1 + \frac{S/3 - \rho}{1-\rho}\right).$$