Shengjie Li
RUID: 188008047

Midterm Exam

Problem 1: Let x_0 be deterministic and $x_1, \ldots, x_N$ denote random variables satisfying (an autoregressive model of order 1)

$$x_n = \alpha x_{n-1} + w_n, \quad n = 1, \ldots, N,$$

where $w_1, \ldots, w_N$ are independent and identically distributed Gaussian random variables with mean 0 and variance 1 while α denotes an unknown parameter.

a) Find the joint density of $x_1, \ldots, x_N$ given α (remember x_0 is deterministic).

The joint density function of $x_1, \ldots, x_N$ is:

$$\begin{aligned}
&f_{X_1,\ldots,X_N}(x_1, x_2, \ldots, x_N) \\
&= f_{X_N|X_1,\ldots,X_{N-1}}(x_N|x_1, \ldots, x_{N-1}) \\
&\times f_{X_{N-1}|X_1,\ldots,X_{N-2}}(x_{N-1}|x_1, \ldots, x_{N-2}) \\
&\ldots \\
&\times f_{X_1}(x_1).
\end{aligned}$$

Then we can get the CDFs and PDFs:

$F_{X_1}(X_1) = P(x_1 < X_1) = P(\alpha x_0 + \omega_1 < X_1) = P(\omega_1 < X_1 - \alpha x_0) = F_w(X_1 - \alpha x_0)$

$f_{X_1}(X_1) = f_w(X_1 - \alpha x_0)$

$\vdots$

$F_{X_N}(X_N) = P(x_N < X_N) = P(\alpha x_{N-1} + \omega_N < X_N) = P(\omega_N < X_N - \alpha x_{N-1}) = F_w(X_N - \alpha x_{N-1})$

$f_{X_N}(X_N) = f_w(X_N - \alpha x_{N-1})$

Then the joint density function of $x_1, \ldots, x_N$ will be:

$$f_{X_1,\ldots,X_N}(x_1, x_2, \ldots, x_N) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x_i - \alpha x_{i-1})^2}$$

b) Compute the maximum likelihood estimate of α when you are given x_0 and a realization of $x_1, \ldots, x_N$.

From a) we know that

$$\mathcal{L}(\alpha) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x_i - \alpha x_{i-1})^2}.$$

Let $l(\alpha) = \ln(\mathcal{L}(\alpha))$,

$$l(\alpha) = -\frac{N}{2}\ln(2\pi) - \sum_{i=1}^{N}\frac{1}{2}(x_i - \alpha x_{i-1})^2$$

$$l'(\alpha) = -\sum_{i=1}^{N}(x_i - \alpha x_{i-1})(-x_{i-1})$$

$$= \sum_{i=1}^{N}(x_i x_{i-1}) - \sum_{i=1}^{N}(\alpha x_{i-1}^2)$$

$$= \sum_{i=1}^{N}(x_i x_{i-1}) - \alpha\sum_{i=1}^{N}(x_{i-1}^2)$$

$$l''(\alpha) = -\sum_{i=1}^{N}(x_{i-1}^2)$$

Because $l''(\alpha) = -\sum_{i=1}^{N}(x_{i-1}^2 < 0$, $l(\alpha)$ is concave.
Let $l'(\alpha) = 0$, then we can get the maximum likelihood estimate of α:

$$\alpha = \frac{\sum_{i=1}^{N}(x_i x_{i-1})}{\sum_{i=1}^{N}(x_{i-1}^2)}.$$

Problem 2: Let $x_n, n = 1, \ldots, N$ be random variables and consider the two scenarios:

$\mathsf{H}_0 : x_n = -s\alpha_n + w_n$,
$\mathsf{H}_1 : x_n = s\alpha_n + w_n$,

where w_n are independent and identically distributed Gaussian random variables with mean 0 and variance σ_2 where σ_2 is unknown, $\alpha_1, \ldots, \alpha_N$ are deterministic and known and, finally $s > 0$ is a deterministic and unknown parameter. If the prior probabilities are $P(\mathsf{H}_0) = P(\mathsf{H}_1) = 0.5$

a) Find the optimum decision mechanism that decides between the two scenarios and minimizes the probability of making an error. Start by assuming that all unknown parameters are magically known.

Let D be the actual outcome, C be the cost, then we have:

$$\{D_0, H_0\} \text{ with cost } C_{00},$$
$$\{D_0, H_1\} \text{ with cost } C_{01},$$
$$\{D_1, H_0\} \text{ with cost } C_{10},$$
$$\{D_1, H_1\} \text{ with cost } C_{11}.$$

Then the average error(cost) will be:

$$C(\delta_0, \delta_1) = C_{00}\mathbb{P}(D_0\&_0) + C_{01}\mathbb{P}(D_0\&_1) + C_{10}\mathbb{P}(D_1\&_0) + C_{11}\mathbb{P}(D_1\&_1).$$

Let $C_{00} = C_{11} = 0$, $C_{01} = C_{10} = 1$, then

$$\begin{aligned} C(\delta_0, \delta_1) &= \mathbb{P}(D_0 \& H_1) + \mathbb{P}(D_1 \& H_0) \\ &= \int \delta_0(X) f_1(X) dX \cdot \mathbb{P}(H_1) + \int \delta_1(X) f_0(X) dX \cdot \mathbb{P}(H_0). \end{aligned}$$

By definition, we can get:

$$F_0(X) = \prod_{i=1}^{N} P(x_i < X_i) = \prod_{i=1}^{N} P(-s\alpha_i + \omega_i < X_i) = \prod_{i=1}^{N} P(\omega_i < X_i + s\alpha_i) = \prod_{i=1}^{N} F_{w_i}(X_i + s\alpha_i),$$

$$f_0(X) = \prod_{i=1}^{N} f_{w_i}(X_i + s\alpha_i).$$

$$F_1(X) = \prod_{i=1}^{N} P(x_i < X_i) = \prod_{i=1}^{N} P(-s\alpha_i - \omega_i < X_i) = \prod_{i=1}^{N} P(\omega_i < X_i - s\alpha_i) = \prod_{i=1}^{N} F_{w_i}(X_i - s\alpha_i),$$

$$f_1(X) = \prod_{i=1}^{N} f_{w_i}(X_i - s\alpha_i).$$

We want to minimize the cost,

$$\arg\min_{\delta_0, \delta_1} C(\delta_0, \delta_1) = \arg\min_{\delta_0, \delta_1} \int (\delta_0(x) f_1(x) dx \cdot \mathbb{P}(H_1) + \delta_1(x) f_0(x) dx \cdot \mathbb{P}(H_0)).$$

If $\frac{f_1(x)}{f_0(x)} \geq \frac{\mathbb{P}(H_0)}{\mathbb{P}(H_1)}$, we choose H_1; If $\frac{f_1(x)}{f_0(x)} < \frac{\mathbb{P}(H_0)}{\mathbb{P}(H_1)}$, we choose H_0.

$$\frac{f_1(x)}{f_0(x)} = \frac{\prod_{i=1}^{N} f_w(x_i - s\alpha_i)}{\prod_{i=1}^{N} f_w(x_i + s\alpha_i)} = \prod_{i=1}^{N} \frac{\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x_i - s\alpha_i)^2}}{\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x_i + s\alpha_i)^2}} = \prod_{i=1}^{N} e^{\frac{4x_i s\alpha_i}{2\sigma^2}}$$

The optimum decision mechanism is, if $\prod_{i=1}^{N} e^{\frac{4x_i s\alpha_i}{2\sigma^2}} \geq 1$, then we choose H_1, otherwise we choose H_0.

b) The decision mechanism you found in a) depends on the unknown parameters s and σ_2 . Apply suitable transformations to find an equivalent mechanism (by taking for example the logarithm and removing unnecessary terms) which does not depend on these two unknown parameters.

From what we got in a), we know that the optimum decision mechanism is, if $\prod_{i=1}^{N} e^{\frac{4x_i s\alpha_i}{2\sigma^2}} \geq 1$, then we choose H_1, otherwise we choose H_0.

If we do logarithm on both sides, the results would not change.

$$ln(\prod_{i=1}^{N} e^{\frac{4x_i s\alpha_i}{2\sigma^2}}) \geq ln(1)$$

$$\implies \sum_{i=1}^{N} \frac{4x_i s\alpha_i}{2\sigma^2} \geq 0$$

$$\frac{4s}{2\sigma^2} \sum_{i=1}^{N} x_i\alpha_i \geq 0$$

Because $s > 0$, $\sigma^2 \geq 0$,

$$\implies \sum_{i=1}^{N} x_i\alpha_i \geq 0$$

The equivalent mechanism is: if $\sum_{i=1}^{N} x_i\alpha_i \geq 0$, we choose H_1, otherwise we choose H_0.

c) Explain what are the optimality properties of the mechanism you ended up with.

(a) By using this mechanism, we will have least probability of making an error.

(b) Each decision is independent on previous decisions.

Problem 3: Consider a random vector X for which we have three possible scenarios

$\mathsf{H}_0 : X \sim f_0(X)$,
$\mathsf{H}_1 : X \sim f_1(X)$,
$\mathsf{H}_2 : X \sim f_2(X)$,

with all the prior probabilities assumed equal. Find the optimum decision mechanism that minimizes the probability of making an error. Consider now the two likelihood ratios $\mathsf{L}_1 = \frac{f_1(X)}{f_0(X)}$ and $\mathsf{L}_2 = \frac{f_2(X)}{f_0(X)}$. For every realization X you can compute the two likelihood ratios which are in fact all you need to make your decision.

a) In the 2D space with axes $\mathsf{L}_1, \mathsf{L}_2$ identify the regions for which you decide in favor of each of the three scenarios $\mathsf{H}_0, \mathsf{H}_1, \mathsf{H}_2$.

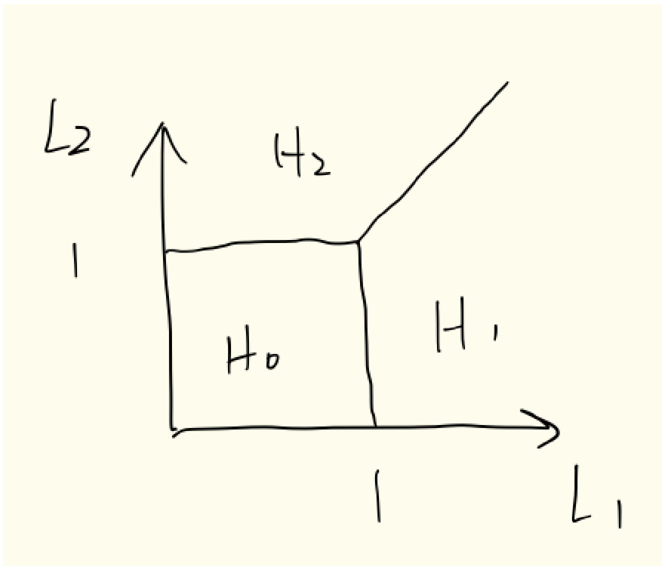
Denote the cost of outcome i and hypothesis j by C_{ij}. Let $\begin{cases} C_{ij} = 1, i \neq j \\ C_{ij} = 0, i = j \end{cases}$, then we have

$$C(\delta_0, \delta_1, \delta_2) = \int ((\delta_0(f_1(x)\mathbb{P}(H_1) + f_2(x)\mathbb{P}(H_2)) + (\delta_1(f_0(x)\mathbb{P}(H_0) + f_2(x)\mathbb{P}(H_2))$$
$$+ (\delta_2(f_0(x)\mathbb{P}(H_0) + f_1(x)\mathbb{P}(H_1)))dx.$$

Because $P(H_0) = P(H_1) = P(H_2)$,

$$\arg\min_{\delta_0,\delta_1,\delta_2} C(\delta_0,\delta_1,\delta_2) = \int ((\delta_0(f_1(x) + f_2(x)) + (\delta_1(f_0(x) + f_2(x)) + (\delta_2(f_0(x) + f_1(x)))dx.$$

Denote $f_1(x) + f_2(x)$ by C_0, $f_0(x) + f_2(x)$ by C_1, $f_0(x) + f_1(x)$ by C_2, then we can get the optimum decision mechanism: let $C_i = min(C_0, C_1, C_2)$, set δ_i to 1 and any other δ to 0. The region would look like this:



b) What happens at the boundaries between two regions? What happens at the single point which belongs to the boundary of all three regions?

We already knew that all prior probabilities are equal.
Therefore, for $L_1 = 1$, we let $C_i = min(C_0, C_1)$, set δ_i to 1 and any other δ to 0.
For $L_2 = 1$, we let $C_i = min(C_0, C_2)$, set δ_i to 1 and any other δ to 0.
For $L_1 = L_2$, we let $C_i = min(C_1, C_2)$, set δ_i to 1 and any other δ to 0.
For the boundary of all three regions, let $C_i = min(C_0, C_1, C_2)$, set δ_i to 1 and any other δ to 0.

Problem 4: As discussed in the class the space of all random variables constitutes a vector space. We can also define an inner product (also mentioned in class) between two random vectors x, y

$$< x, y >= E[xy].$$

Consider now the following random variables x, z, w. We are interested in linear combinations of the form $\hat{x} = az + bw$ where a, b are real deterministic quantities.

a) By using the orthogonality principle find the $\hat{x}_*$ (equivalently the optimum coefficients a_*, b_*) that is closest to x in the sense of the norm induced by the inner product.

Suppose the vector space is $V = \{y; y = az + bw\}$, W is a subspace of V.
We want to find the $\hat{x}_*$ that is closest to x, that is:

$\arg\min_{\hat{x}} ||x - \hat{x}||, < x - \hat{x}, y >= 0$ for all y in W.
Then we can get
$< x - \hat{x}, z >=< x - \hat{x}, w >= 0$
$< x - \hat{x}, z >=< x - (az + bw), z >= E[(x - (az + bw))z] = E[xz] - aE[z^2] - bE[wz] = 0$
$E[xz] = aE[z^2] - bE[wz]$
$< x - \hat{x}, w >=< x - (az + bw), w >= E[(x - (az + bw))w] = E[xw] - aE[z^2] - bE[wz] = 0$
$E[xw] = aE[zw] - bE[w^2]$
$a_* = \dfrac{E[xz]E[w^2] - E[xw]E[wz]}{E[z^2]E[w^2] - E[wz]^2}$
$b_* = \dfrac{E[xw]E[z^2] - E[xz]E[wz]}{E[z^2]E[w^2] - E[wz]^2}$

b) Compute the optimum (minimum) distance and its optimum approximation $\hat{x}_*$ in terms of $E[xz], E[xw], E[z^2], E[zw], E[w^2]$.

c) Explain what is the physical meaning of this approximation.