

1 Minimax Estimation

Recall the minimax example in the lecture notes. Define a sequence of i.i.d. Bernoulli random variables $\{b_i\}_{i=1}^N$, each taking the value 1 with probability p and 0 with probability $1 - p$. Consider estimators of the form

$$\hat{p}_\beta = \frac{\beta + \sum_{i=1}^N b_i}{2\beta + N}, \quad (1)$$

which give an approximation of the parameter p . We wish to choose β in a way that minimizes the worst case mean squared error, that is to say

$$\beta_{\minimax} = \operatorname{argmin}_\beta \max_{p \in [0,1]} \mathbb{E} [(\hat{p}_\beta - p)^2 | p]. \quad (2)$$

This is a simple enough problem that we can compute the parameter directly.

(a) Notionally, p is decided as a function of β . Please identify the function

$$C(\beta) = \max_{p \in [0,1]} \mathbb{E} [(\hat{p}_\beta - p)^2 | p], \quad (3)$$

which represents our worst case cost for a chosen β .

Your Solution Here!

(b) Now, minimize our worst case error

$$\beta_{\minimax} = \operatorname{argmin}_\beta C(\beta) \quad (4)$$

Your Solution Here!

(c) For β chosen above, what is the MSE as a function of p ?

Hint: This is somewhat of a trick question

Your Solution Here!

2 Kalman Filter

This problem builds off of some of the notions in the previous homework about linear dynamical systems. Suppose we have a system

$$x_{n+1} = Fx_n + u_n \quad (5)$$

$$y_n = Hx_n + v_n, \quad (6)$$

where $x_0 = \mathbf{0}$ (this is a slightly different definition from the previous problem set due to an indexing issue). Let $u_n \stackrel{iid}{\sim} \mathcal{N}(0, Q)$, and $v_n \stackrel{iid}{\sim} \mathcal{N}(0, R)$.

In this problem, you will derive the MMSE estimator of the state $x_i \in \mathbb{R}^N$ given the set of observations $\{y_j \in \mathbb{R}^M\}_{j=1}^i$.

The Kalman filter is generally broken into two steps, a prediction step, and an update step. The orthogonality principle will be used throughout.

If you have trouble with the general cases, start with the first couple of timesteps before trying to generalize.

(a) What is the MMSE linear predictor $\hat{x}_{i+1|i}$ given the previous data? That is to say, the MMSE estimator of x_{i+1} given observations $\{y_1, \dots, y_i\}$. For the sake of this problem, assume that you have the MMSE linear estimator for timestep i , denoted $\hat{x}_{i|i}$.

Your Solution Here!

(b) In order to use the predictor in the previous part, we need the MMSE linear estimator of x_i given $\{y_1, \dots, y_i\}$. Using the fact that we have the MMSE linear predictor $\hat{x}_{i|i-1}$ (which is time-independent), what is the MMSE linear estimator $\hat{x}_{i|i}$?

Your Solution Here!

(c) We seem to have a bit of a chicken and egg problem here. To predict, we need to estimate. To estimate, we need to predict. We need a base case to resolve this problem. What is the best linear estimator of x_1 given the observation y_1 ?

Your Solution Here!

(d) For good measure, please verify that your solutions for parts (a) and (b) are indeed the MMSE estimators of x_2 . This completes the base case for our inductive estimator.

Your Solution Here!

3 Frequency Band Occupancy (Computational)

Consider designing a wireless communication system. Your communication link is not isolated in the environment, there are other similar devices trying to communicate in the same general frequency band. For this reason, you wish to infer which frequency bands are occupied in order to minimize inter-channel interference.

Fortunately, each of these devices only occupy a relatively narrow frequency band. Unfortunately, links can hop between frequency bands, start, or stop seemingly randomly.

You may consider the sampled received signal to be of the form

$$y[n] = w[n] + \sum_{i=1}^N x_i[n], \quad (7)$$

where the number of signals N is unknown and each $x_i \in BL(-B_i, B_i)$ is an unknown, but approximately bandlimited function of a potentially finite non-zero length. Assume $w[n] \stackrel{i.i.d.}{\sim} \mathcal{N}(b, \sigma^2)$ be a random process comprised of (possibly non-zero) Gaussian noise with variance σ^2 .

Please design an algorithm that, given a stream of data, attempts to determine the set of occupied frequency bands *using only the prior data*.

We have provided a data file (containing amplitude modulated speech). Please run your algorithm on the file and provide an image displaying the results. Additionally, please provide a statistical analysis and justification of your algorithm.

4 Linear Predictive Coding (Computational)

Implement a linear predictive coding system using 10 ms frames of data as described in class. Apply it with orders $p \in \{6, 12, 20, 50\}$ to the provided speech file or your own recording.