Stochastic Processes, Sessions 16 and 17 — Group Work
Minimum Mean Squared Error (MMSE) Estimation

Troels Pedersen and Carles Navarro, Aalborg University

October, 2016

Sessions 16 and 17

Go through the exercises below. Allow yourself the time to reflect over your results and discuss them with other students! Use the book and lecture notes for inspiration and for further information.

Exercises from [Kay].

Solve the following problems from [Kay]:

- Problem 7.43.
- Problem 14.1.
- Problem 14.20.
- Problem 14.21.

LMMSE estimator of the outcome of a Gaussian random variable

Let $\boldsymbol{X} = [X_1, X_2, X_3, X_4]^{\mathrm{T}}$ be a Gaussian random vector with mean $[1, -3, 0, 2]^{\mathrm{T}}$ and covariance matrix

$$\boldsymbol{C}_{\boldsymbol{X}} = \begin{bmatrix} 1 & -1 & 0.5 & -1 \\ -1 & 5 & 2.5 & 3 \\ 0.5 & 2.5 & 6.5 & 2 \\ -1 & 3 & 2 & 2.5 \end{bmatrix}.$$

Assume that we can observe the outcome of the random variables X_1, X_2, X_3, but not the outcome of X_4, which we would like to estimate.

- Write the expression of the LMMSE estimator of X_4 given the observation of X_1, X_2 and X_3.
- Compute the coefficients of the LMMSE estimator of X_4 and the MSE of the estimate $\hat{X}_4$.

- Next, draw realizations of the random vector $\boldsymbol{X}$ and apply the estimator computed above to estimate X_4 from X_1, X_2, X_3. (*Hint: Recap [Kay, Sections 12.11 and 14.9] on how to draw correlated Gaussian vectors*).
- Average the error and squared error of the estimates over the multiple realizations and compare the result with the theoretical MSE computed above.
- Modify your program to estimate by Monte Carlo simulations the correlation between the estimation error and each of the observations, i.e. estimate $\mathbb{E}[(X_4 - \hat{X}_4)X_i]$ for $i = 1, 2, 3$. Discuss with your group mates whether you could have predicted the result beforehand, and why.

MMSE and LMMSE estimation of scalar random variable

In this exercise we compare the MMSE with the linear MMSE for the same estimation problem. This is in general a difficult task, since the MMSE is often hard to compute. Therefore we consider a specific case, where both estimators can be computed by simple hand derivations.

We observe the random variable Z which is defined as the sum of two independent uniformly distributed random variables:

$$Z = \Theta + W, \qquad \Theta \sim \mathcal{U}(0, 3), \text{ and } W \sim \mathcal{U}(0, 1).$$

This model could for example describe a measurement Z of a physical entity Θ with a noisy measurement apparatus adding the noise W. We want to know the value of Θ but we have only access to one observation of Z. However this interpretation is not essential for the exercise. The joint pdf of Z, Θ reads

$$p_{Z\Theta}(z, \theta) = \begin{cases} \frac{1}{3} & \text{for } 0 \leq \theta \leq 3 \text{ and } \theta \leq z \leq \theta + 1 \\ 0 & \text{otherwise.} \end{cases}$$

- Draw a sketch on paper of the above pdf in the z, θ-plane (it suffices to do a sketch in 2D indicating the support of the above pdf). Generate multiple joint realizations of Z, Θ and scatterplot these. Compare the scatterplot to your drawing of the pdf.
- Now we observe a realization of Z and use this to estimate a value of Θ using the (non-linear) MMSE. All you need to do is to calculate the conditional expectation $\hat{\Theta} = g(Z) = E[\Theta|Z]$.
  Hint: Do this without integration relying solely on graphical argumentation using your sketch of $p_{Z\Theta}(z, \theta)$ directly plotting the function $E[\Theta|z]$.
- Implement the MMSE in Matlab. Compute the bias and MSE via Monte Carlo simulations.
- Derive the LMMSE for estimating Θ given Z. Start by calculating $E[\Theta], E[Z], \mathrm{Var}(Z)$, and $\mathrm{Cov}(Z, \Theta)$. Plot a line representing the linear MMSE in the z, θ plane and compare to that of the non-linear MMSE.
- Implement the LMMSE in Matlab and compare its performance to the theoretical MSE from the lecture note and to the performance of the (non-linear) MMSE.

Filtering of a noisy signal

A realization of a random process $Y(n)$ is recorded with an instrument which introduces some measurement error. The samples recorded by the measurement device can be modeled as

$$X(n) = Y(n) + W(n)$$

with $W(n)$ being a white Gaussian noise process with variance σ_W^2.

From previous measurements of the process, we know that $Y(n)$ is well modeled by the following AR(1) process with feedback coefficient a:

$$Y(n) = aY(n-1) + U(n) \tag{1}$$

where $U(n)$ is white Gaussian noise with variance $\sigma_U^2 = 1 - a^2$.

- Compute the mean and auto-correlation functions (ACF) of the process $X(n)$. *Hint: recall that the ACF of the AR process in* (1) *is given by* $R_Y(k) = \sigma_U^2 \frac{a^{|k|}}{1-a^2}$.
- Compute the cross-correlation of the processes $X(n)$ and $Y(n)$ (i.e., derive the value of $\mathbb{E}[X(n)Y(m)]$).

Next, we want to filter in realtime the measurements provided by the recording device in order to reduce the distortion introduced by the noise process $W(n)$. To do so, we decide to apply LMMSE estimation of the sample $Y(n)$ based on the last $L+1$ observed measurements $\boldsymbol{X}_{\mathrm{RT}}(n) = [X(n), X(n-1), \ldots, X(n-L)]$.

- Compute the coefficients of the above LMMSE estimator for arbitrary values of L. Find as well the expression of the MSE of the estimates $\hat{Y}(n)$.
- Investigate how the theoretical MSE derived above varies with the parameters: L, a and σ_W^2 (for instance, plotting the MSE vs each of the parameters). Intuitively, how do you explain the behavior of the MSE with each of the parameters? (*Recall that for the AR process in* (1) *to be stable,* $|a| < 1$.)
- Generate 1000 samples of the processes $X(n)$ and $Y(n)$ with $a = 0.98$ and $\sigma_W^2 = 0.5$.
- Apply successively the LMMSE estimator with values $L = \{1, 5, 10\}$ to estimate the samples $Y(L+1), Y(L+2), \ldots, Y(1000)$. Plot your results and compare the processes $Y(n)$ and $X(n)$ with the estimates obtained.
- Average the squared error of the estimates over all samples and compare the result with the MSE derived above.

Assume now that the processing of the recorded samples of $X(n)$ can be done offline (for instance, all data is stored in a memory device and post-processed later on with MATLAB). If the processing need not be realtime, we can now use later samples of the measurement –which have been previously stored– i.e. $X(n+1), X(n+2), X(n+3)\ldots$, to estimate each sample of $Y(n)$.

- Modify the program to now perform LMMSE estimation of $Y(n)$ based on the $2L+1$ observed samples $\boldsymbol{X}_{\mathrm{OL}}(n) = [X(n-L), \ldots, X(n), \ldots, X(n+L)]$.
- Assuming that both estimators use the same total number of observed samples, which estimator is more accurate: the realtime or the offline one?