Stochastic Processes

Fall 2016

Activity Descriptions
Intended Learning Outcomes
Reading Material
Exercises

Literature:
[KAY] Steven M. Kay, "Intuitive Probability and Random Processes using MATLAB"
[LN] Lecture Notes

This document will be gradually updated as the course progresses!

Session 1 – Recap Lecture: Random Variables, Random Vectors, Expectation and Conditional Expectation

Short Description

In this first lecture, we shortly review basic concepts of probability theory and random variables. We will recap the concepts of continuous and discrete random variables and explain how they are characterized by their probability density function (pdf) and probability mass function (pmf), or their cumulative distribution function (CDF). We will see how multiple, jointly distributed random variables can be described by their joint, conditional and marginal distributions, and introduce the concepts of independent random variables and covariance. We will then extend these ideas to the case of vector random variables, also called random vectors, and introduce covariance matrices. Finally, we will review how to compute the expectation and conditional expectation of a function of random variables.

All the concepts discussed during the lecture are fundamental to build and analyze stochastic processes, and their understanding is a prerequisite for the course. Therefore, this first lecture should be thought of as a way of checking that you have all the pieces of knowledge necessary to easily follow the course. If you can solve the proposed exercises without hesitation, it should be easy for you to start working with stochastic processes right away. Should this not be the case, however, do not panic! Luckily the proposed book for the course contains an extensive and intuitive explanation of these ideas, along with tons of exercises to practice them. Just check the additional background exercises for this lecture –and the associated reading material– and review those parts which you feel you need more help with.

Intended Learning Outcomes

After having attended this lecture and solved the exercises you should be able to:

- Explain the difference between a discrete and a continuous random variable to a fellow student.
- Relate the concepts of joint, conditional and marginal pdfs or pmfs.
- Write up and interpret the covariance matrix of a random vector.
- Compute the expectation of simple functions of random variables, such as a sum of random variables or the product of independent random variables.

Reading Material

The lecture will focus around the contents in:
[Kay] Sections 13.3, 13.4, 13.6, 14.3 and 14.5.
If you have problems understanding these sections, it may be useful for you to refresh as well Chapter 12 in [Kay].

Exercises

Since this is a recap lecture, we have included a large number of exercises, in case you need extra practice. If, after solving some exercises for a given topic you feel that you have a good understanding of the concepts, simply jump to the next group of exercises. We have also included additional reading material to help with the exercises, if needed.

Topics	**Exercises**	**Reading [Kay]**
Bivariate random variables	12.2, 12.5, 12.11, 12.15, 12.38, 12.56	12.3–.12.7, 12.11
Conditional probability densities and expectation	13.2, 13.4, 13.5, 13.8, 13.16	13.3–13.7
N-dimensional random vectors	14.1, 14.6, 14.11, 14.12, 14.19	14.3–14.5, 14.7, 14.9

Extra Recap Exercises

If you still have problems solving the exercises for this lecture, we have selected a set of more basic exercises and reading material to review the fundamental concepts of probability theory and random variables. Select those concepts you find difficulties with, use the [Kay] book for reading and try some of the exercises. **This extra recap material does not only apply to the first lecture: if you find yourself lost with any of the concepts used at any point during the course, make sure to go back to this table!**

Topics	**Exercises**	**Reading [Kay]**
Discrete random variables, probability mass functions and cumulative distribution functions	5.4, 5.31	Chapter 5
Continuous random variables, probability density functions and cumulative distribution functions	10.2, 10.4, 10.12, 10.21, 10.22	Chapter 10
Transformation of random variables	5.20, 10.40, 10.41, 10.46	Sections 5.7, 5.9, 10.7, 10.9
Expected value, variance	6.2, 6.13, 6.17, 9.10, 11.10, 11.23, 11.24	Cahpters 6 and 11
Joint distributions, convariance, independent random variables	7.10, 7.19, 7.34, 7.27, 7.33, 12.2, 12.5, 12.11, 12.15, 12.38, 12.56	Chapters 7 and 12
Conditional probability mass functions and expectation	8.2, 8.6, 8.13, 8.21	Chapter 8
M-dimensional (discrete) random vectors	9.12, 9.13, 9.17, 9.19, 9.20, 9.29, 9.30, 9.34, 9.35, 9.36	Chapter 9

Session 2 – Group Work: Monte Carlo Simulation of Random Vectors, Expectation and Conditional Expectation

Short Description

In this group work session, we will focus on computer simulation of random vectors and use this to inspect properties of random vectors. The type of computer simulation we consider it is used so frequently by engineers and scientists, that a name has been invented for it: Monte Carlo simulation. Monte Carlo simulation is a very effective tools to build up intuition about random variables and processes. But to do this effectively, there are several things you need to know about — this is what we will look at during this exercise session.

For the session you will need the following things:

- Your laptops with Matlab installed. We suggest to use 1 computer per 2 students. You may team up in advance. See how to install Matlab with our university licence at `https://it-wiki.es.aau.dk/wiki/Matlab`
- Kay's book.
- Paper and pencil.

Intended Learning Outcomes

After this session you should be able to:

- Compute means and covariances of linear operations on random vectors.
- Simulate random vectors and conditional random vectors.
- Use the law of large numbers to compute and conditional expectations using Monte Carlo simulation.
- Plot simulation results in a form readable by peers. Discuss and interprete the plots and relate this to theory.

Preparation

[Kay] Chapter 2 (This is background material and a "light read")
[Kay] 13.3–13.7 (Recap of Conditional pdfs)
[Kay] 15.3 & 15.4 (Law of large numbers)

Exercises

Will be handed out at the beginning of the session.

Session 3 – Group Work / Video Lecture: Definition of Stochastic Processes and Strict-Sense Stationarity

Short Description

Many scientific and engineering problems are well-posed to be described using stochastic processes. Therefore, this is an important tool to engineers. In this lecture we define the “heart” of the course which is a *stochastic process*. In fact, we give two different, but equivalent, definitions. We also treat the concept of a *strict-sense stationary* (SSS) stochastic process as well as the mean function, variance function and covariance function.

Note that this will be a **video session**. Before the session, prepare by carefully going through the **reading material** and **watching the videos** specified below. The time in class will be spent by doing exercises and solving doubts. Please bring with you writing material, Kay’s book and your laptop – preferably with some headphones, so that you can watch parts of the videos again if needed.

Intended Learning Outcomes

After reading the texts, watching the videos, and solving the exercises, you should be able to:

- Explain the two definitions of a stochastic process to a fellow student. Relate these definitions to each other and to the definition of a random variable.
- Give a real-world example of a entity that can be modeled by a stochastic process from your field of study. For this example explain what the outcomes, events, and realizations are.
- Know and use the definition of an IID process.
- Compute the mean and covariance functions for a stochastic process using their definitions.
- State two examples of random processes that are SSS, and two examples of processes which are not SSS.

Study Material

[Kay] Section 5.3 (Recap of the definition of a random variable)
[Kay] Chapter 16 (Definition of stochastic processes and associated concepts)
Videos: SP 3.1 — SP 3.4. Links are provided in Moodle.

Exercises

16.1 (Also state an example from your own field of study)
16.2
16.4
16.7
16.14
16.15
16.16
16.24

16.26
16.30
16.31
Examine the different random processes described in all the exercises above. Which ones are SSS? Which ones are not SSS?

Go through the intended learning outcomes above with your fellow students. Did you achieve all of them? Did you learn something beyond the intended learning outcomes?

It is possible that some of you have already done a few of the above problems. If so, we suggest that you review the problems and consider if you can now solve them in a different and perhaps easier way.

Session 4 – Group Work: Simulation of Stochastic Processes

Short Description

In this group work session we will focus on computer simulation of stochastic processes. We will learn how to simulate realizations of i.i.d. random processes with a given distribution, as well as how to use such i.i.d. processes to construct other –more complex– types of processes. In addition to this, we will shortly review how to compute the pdf of a variable resulting of a transformation of one or several random variables, and how to use the transformation theorem to generate realizations of random variables with a desired distribution.

For the session you will need the following things:

- Your laptops with Matlab installed. We suggest to use 1 computer per 2 students. You may team up in advance. See how to install Matlab with our university licence at `https://it-wiki.es.aau.dk/wiki/Matlab`
- Kay's book.
- Paper and pencil.

Intended Learning Outcomes

After having attended this session you should be able to:

- Simulate iid. processes with Gaussian, uniform and exponentially distributed entries.
- Simulate random processes with uncorrelated Gaussian entries having arbitrary means and variances.
- Use the transformation theorem to simulate Gaussian random vectors with a certain prescribed covariance structure.
- Draw realizations of an MA(1) process.

Preparation

[Kay] Sections 10.7, 12.6, 14.4 (Transformation theorem)
[Kay] Sections 10.9, 12.11, 14.9 (Application of transformation theorem to computer simulation)

Exercises

Will be handed out at the beginning of the session.

Session 5 – Lecture: Wide-Sense Stationary Stochastic Processes

Short Description

While the concept of a strict-sense stationary (SSS) stochastic process is easy to define, it is much harder to verify that a stochastic process is indeed strictly stationary. The definition is simply "too strict" and most processes encountered in practice are not SSS. Therefore we introduce a less strict, but more (practically) useful concept, namely, *wide-sense stationary* (WSS) stochastic processes. The definition of such processes relies on the so-called *autocorrelation function* which we also define and interpret.

Intended Learning Outcomes

After having attended this lecture and solved the exercises you should be able to:

- Explain the meaning of an ACF to a fellow student.
- Compute the ACF for simple stochastic processes using its definition.
- Use the theoretical properties of any ACF as sanity checks of your derivations.
- Determine whether or not a stochastic process is WSS.
- Relate the definitions of SSS and WSS.
- Know (without hesitation and computation) the ACF of an iid. process.

Reading Material

[Kay] Sections 17.1-17.5

Exercises

Do exercise 17.1 and simulate 1000 realizations of this process to estimate the mean sequence. Try different values for p and compare with your analytically derived expression.

Do exercise 17.5 and plot $R = 100$ realizations for the time instances $n = 1, 2, \ldots, 10$. (See fig 16.15 for an example). Discuss the results in terms of WSS and/or SSS properties of this process. Afterwards, increase R to 1000 and estimate the mean sequence and the variance sequence.

Do exercises 17.8, 17.9 and 17.12

Let $W(n)$ be an iid. process with mean 1 and variance 4. Determine the mean sequence, variance sequence and autocorrelation sequence of the following three processes:

$$i) : \ W(n) \qquad ii) : \ X(n) = 3W(n) + W(n-1) \qquad iii) : \ Y(n) = nW(n)$$

Which of the three processes are SSS and/or WSS and why? Simulate outcomes of each of the three processes (Hint: you'll need to assume a pdf for the samples of $W(n)$. Try e.g. a normal or a uniform pdf. Furthermore, you'll need to select an appropriate range for n).

Go through the intended learning outcomes above. Did you achieve all of them? Did you learn something beyond the intended learning outcomes? Discuss this with your fellow group members.

Session 6 – Group Work: WSS Processes, the Autocorrelation Function and its Estimation

Short Description

In this group work session, we will mainly focus on the simulation and analysis of wide-sense stationary (WSS) processes. We will learn how to generate realizations of different types of WSS processes with specific first- and second-order statistical properties, i.e. specific mean and autocorrelation function (ACF), and how to estimate the ACF of a given process from some realization(s) of it. We will use the ACF to predict the outcome of future samples of a given process based on the observation of the current sample, and shortly explore the implications of a random process being *ergodic in the mean.*

For the session you will need the following things:

- Your laptops with Matlab installed. We suggest to use 1 computer per 2 students. You may team up in advance. See how to install Matlab with our university licence at `https://it-wiki.es.aau.dk/wiki/Matlab`
- Kay's book.
- Paper and pencil.

Intended Learning Outcomes

After having attended this session you should be able to:

- Generate realizations of moving average (MA) and autoregressive (AR) random processes and estimate their respective ACFs from them.
- Explain the difference between *temporal averaging* and *ensemble averaging*, and identify when they can be used interchangeably to estimate the mean of a WSS process.

Preparation

[Kay] Sections 7.9, 12.9 (Optimal Linear Prediction)
[Kay] Section 17.7 (Estimation of the ACF and the PSD – skip the PSD part for now!)

Recap of material from Lecture 3, with special attention on:
[Kay] Section 17.5 (Ergodicity and temporal averaging) and Example 17.5 (Autoregressive Random Process).

Exercises

Will be handed out at the beginning of the session.

Session 7 – Lecture: The Power Spectral Density and its Estimation

Short Description

For WSS random processes, the *power spectral density* (or just the *power spectrum*) is defined as the Fourier transform of the autocorrelation function. While the autocorrelation function yields information about the temporal correlation structure of a WSS process, the power spectrum reveals information about the periodicities (or the "frequency content") of a random process. The power spectrum also plays an important role for the input-output relationships of linear time-invariant systems when fed with random WSS signals (the topic of Lecture 5). In this lecture we define the power spectrum and discuss its properties. Finally, we elaborate on how to estimate the power spectrum via observations of a WSS random process.

Intended Learning Outcomes

After having attended this lecture and solved the exercises you should be able to:

- Explain the definition and the meaning of a PSD to a fellow student.
- Compute the PSD given a particular ACF.
- Use the theoretical properties of any PSD as sanity checks of your derivations.
- Know (without hesitation and computation) the PSD of a white process.
- Know the definition of bias and MSE of an estimator and explain their meaning.

Reading Material

[Kay] Sections 17.6-17.9

Exercises

17.30
17.34
17.36
17.37
17.38
17.40

Go through the intended learning outcomes above. Did you achieve all of them? Did you learn something beyond the intended learning outcomes? Discuss this with your fellow group members.

Session 8 – Group Work: Estimation of ACF and PSD: Sample Autocorrelation Function and Periodogram

Short Description

The second-order characteristics of a WSS random process are described by its autocorrelation function (ACF) or, equivalently, by its power spectral density (PSD). The knowledge of either of these functions enables a set of statistical methods to process realizations of WSS processes. For instance, linear prediction, interpolation, or filtering of noisy samples of a WSS process can be optimally performed when the second-order properties of the process are known, as we will see later in the course. The Kalman Filter, which is widely applied in many engineering areas, is another example of such tools.

A natural question that arises is then how to estimate the ACF or PSD of a given WSS process when these are not known. In this group work session we will explore a few non-parametric (or model-free) approaches for the estimation of the ACF and the PSD of a WSS process from the observation of some of its samples. In particular, we will focus on the sample autocorrelation function (biased and unbiased), the periodogram and some slight modifications of those.

For the session you will need the following things:

- Your laptops with Matlab installed. We suggest to use 1 computer per 2 students. You may team up in advance. See how to install Matlab with our university licence at `https://it-wiki.es.aau.dk/wiki/Matlab`
- Kay's book.
- Paper and pencil.

Intended Learning Outcomes

After having attended this session you should be able to:

- Infer the behavior that realizations of a WSS process will have from analyzing its ACF or its PSD.
- Estimate the ACF and PSD of a process from its samples using the sample autocorrelation function and the periodogram.
- Explain the differences, advantages and disadvantages between the biased and unbiased sample autocorrelation functions as estimators of the ACF of a WSS process.
- Use the periodogram to identify periodicities in a given WSS process from the observation of one of its realizations.

Preparation

[Kay] Section 17.7 (Recap)
[LN] Section 6.1 (Very detailed read of this part is recommended).

Exercises

Will be handed out at the beginning of the session.

Session 9 – Lecture: Response of Linear Time-Invariant Systems to Random Inputs

Short Description

Many systems encountered by the practitioning engineer are well-approximated by *linear* and *time-invariant* (LTIV) systems. This is a great advantage since the input-output relationships of such systems are particularly simple to compute. In this lecture we consider the input-output relationships of LTIV systems for WSS random processes.

Intended Learning Outcomes

After having attended this lecture and solved the exercises you should be able to:

- Use the input-output relationships to compute the first- and second-order properties of the output when a WSS process is fed as input to an LTIV system.
- Explain and demonstrate how white Gaussian processes can be used for simulation of zero mean Gaussian processes with relatively simple autocorrelation structures.

Reading Material

[Kay] Sections 18.1-18.4 and 18.6
[LN] Chapter 1

Exercises

18.1
18.2

18.3 (But notice that the problem is ill-posed as given in the book. Can you identify what's wrong?) Instead, solve the problem assuming that the impulse response of the system is $h(k) = a^{|k|}$ with $|a| < 1$. (Hint: Use the geometric series from [LN] page 1-10 bottom). Plot the power spectrum for different values of a. What happens if you set $a = 1$?

18.4
18.5
18.6 (Notice that the process $X[n]$ is a special case of the process from exercise 18.5)
18.7
18.10

Go through the intended learning outcomes above. Did you achieve all of them? Did you learn something beyond the intended learning outcomes? Discuss this with your fellow group members.

Session 10 – Group Work: AR, MA and ARMA Processes

Short Description

In this group work session, we will study a useful class of discrete-time stochastic processes called *autoregressive moving average* (ARMA) processes. We have already in previous sessions of the course seen special cases of ARMA processes, namely the *autoregressive* (AR) and *moving average* (MA) processes. As the name indicates, ARMA processes are generated by recursive difference equations. Alternatively, an ARMA process may be viewed as a white Gaussian noise passed through a cascade of two filters — one with infinite impulse response (IIR) and one with finite impulse response (FIR). This simple makes it possible to obtain *analytical* expressions for both autocorrelation function and power spectral density.

ARMA processes are particularly useful since they are easy to fit to data and for this reason they are widely used in real-world data analysis. Specifically, any WSS process can be approximated by an ARMA process of "appropriate" order.

For the session you will need the following things:

- Your laptops with Matlab installed. We suggest to use 1 computer per 2 students. You may team up in advance. See how to install Matlab with our university licence at `https://it-wiki.es.aau.dk/wiki/Matlab`
- Kay's book.
- Paper and pencil.

Intended Learning Outcomes

The knowledge and experience gained throughout this session should enable you to:

- Identify whether a random process is of AR, MA or ARMA type and determine its order and coefficients.
- Identify the parameters of a given ARMA process, calculate its autocorrelation function and power spectral density.
- Use the Yule-Walker equations to set the parameters of an AR process to yield a process which approximates a given target autocorrelation function.
- Simulate any prescribed ARMA(p, q) process.

Reading Material

[LN] Chapter 2
[Kay] Examples 16.7, 16.10, 17.3, 17.5, 17.6, 18.3

Exercises

Will be handed out at the beginning of the session.

Session 11 – Group Work: Parametric Estimation of Random Processes

Short Description

In parametric estimation of the ACF and PSD of random processes, one typically assumes that the data used for estimation is a realization of a WSS process specified by a finite set of parameters. In our case, we will model the observations as realizations of ARMA processes of "appropriate" order. By finding the values of the parameters that best fit the available data, one obtains analytical expressions of the estimated PSD and ACF. This is in contrast to non-parametric methods, as the sample autocorrelation function or the periodogram. The estimated model can then be used for multiple purposes: among others, it can be used as input for the derivation of signal processing algorithms – e.g. estimators, like in the Kalman filter – or it can be used to generate synthetic – i.e. computer generated – realizations of the process that keep the statistical properties of the original data.
In this group work session, we will explore parametric estimation of random processes by applying one of the popular methods available: the Box-Jenkins method.
For the session you will need the following things:

- Your laptops with Matlab installed. We suggest to use 1 computer per 2 students. You may team up in advance. See how to install Matlab with our university licence at `https://it-wiki.es.aau.dk/wiki/Matlab`
- Kay's book.
- Paper and pencil.

Intended Learning Outcomes

After having attended this session you should be able to:

- Preprocess measurement data to transform it into a sequence that can be assumed to be the realization of a WSS process.
- Fit an "appropriate" ARMA model from measurement data by use of the associated Yule-Walker equations (and similar preprogrammed routines).
- With the fitted parameters, obtain analytical estimates of the ACF and PSD of the process.

Preparation

[LN] Recap Section 6.1: Model-Free Estimation of Random Processes
[LN] Section 6.2: Parametric (Model-Based) Estimation of Random Processes

Exercises

Will be handed out at the beginning of the exercise session.

Session 12 – Lecture: Point Processes in 2D (binomial and Poisson)

Short Description

Random point patterns show up in numerous practical situations. Observed point patterns together with stochastic models for such patterns can be used to infer about underlying mechanisms governing where points occur. However, stochastic models for point patterns are also very important in their own right. In particular, they can be used as building blocks for generating ordinary random processes. Such constructions are useful for instance in queuing theory where stochastic models of queues are widely used for analyzing the behavior of time-shared computer and communication systems. In this lecture we introduce the basics of the theory of point processes and we take a closer look at the two simplest classes of such point processes.

Intended Learning Outcomes

After having attended this lecture and solved the exercises you should be able to:

- Give examples of practical occurrences of random point patterns from your own field of study.
- Explain intuitively what a region count is and discuss its main properties.
- Relate the interpretation of an intensity function to the interpretation of a probability density function (pdf). Discuss similarities and distinctions.
- Simulate realizations of 2D binomial point processes.
- Briefly outline a particular way of characterizing a 1D point process which is neither natural nor possible in 2D.

Reading Material

Read the 9 page study note "2D_point_processes_S12.pdf" which can be found in Moodle (general course material). Think about how to solve the built-in exercises while reading.

Exercises

Go through the exercises which are built into the study note (increasing level of difficulty). Make sure to cover a broad selection of the exercises. If you happen to get stuck somewhere then try to go on anyway.

Go through the intended learning outcomes above. Did you achieve all of them? Did you learn something beyond the intended learning outcomes? Discuss this with your fellow group members.

Session 13 – Group Work: Point Processes in 2D (binomial and Poisson)

Short Description

In this group work session, we will work with one- and two-dimensional point processes, as well as shot noise processes. We will learn how to simulate and draw realizations of them. We will construct counting and queueing processes using Poisson processes and, in addition, we will apply Campbell's theorem to compute their expected value.
For the session you will need the following things:

- Your laptops with Matlab installed. We suggest to use 1 computer per 2 students. You may team up in advance. See how to install Matlab with our university licence at `https://it-wiki.es.aau.dk/wiki/Matlab`
- Kay's book and the lecture notes on 2D Point Processes.
- Paper and pencil.

Intended Learning Outcomes

After having attended this session you should be able to:

- Draw sequences of Poisson and binomially distributed random numbers with given parameters.
- Simulate realizations of 2D binomial point processes.
- Simulate realizations of 2D Poisson point processes (both the homogeneous and the inhomogeneous case).
- Use point processes to simulate and analyze the behavior of counting and queuing processes.
- Apply Campbell's theorem to calculate the expected value of functions dependent on point processes.

Reading Material

Lecture notes "2D_point_processes_S12.pdf" (recap).

Exercises

Will be handed out at the beginning of the session.

Session 14 – Lecture: Hypothesis Testing and Detection Theory

Short Description

In this lecture we introduce the basic concepts of *detection theory* and apply them to concrete engineering problems. Detection theory is an application of *hypothesis testing*, which is a branch of statistics. It is useful in numerous engineering fields, e.g. radar and wireless communication. We introduce the maximum-a-posteriori (MAP) decision rule which minimizes the probability of making a decision error and we also introduce Bayes decision rule that minimizes a given average cost.

Intended Learning Outcomes

After having attended this lecture and solved the exercises you should be able to:

- Explain the premises of hypothesis testing and discuss how this framework differs from a parameter estimation framework.
- Apply the following decision rules to signal detection problems: ML, MAP, and Bayes.
- Compute the associated decision error probabilities and expected costs.
- Compare and relate the ML, MAP, and Bayes decision rules in terms of their optimality criteria.

Reading Material

Main:
[KAY] Section 14.10 (real-world example)
Lecture notes in file: "decisionRules_S14.pdf", (see Moodle)

Additional Reading:
[LN] Chapter 3

Exercises

Exercises are built in within the notes "decisionRules_S14.pdf".

Session 15 – Group Work: Hypothesis Testing and Detection Theory

Short Description

Detection theory is an application of hypothesis testing, which is a branch of statistics useful in numerous engineering fields. In this group work session we will develop intuition and comprehension related to hypothesis testing, the different decision rules that can be applied, and their respective probabilities of error. This will be done through a combination of calculation and simulation tasks.
For the session you will need the following things:

- Your laptops with Matlab installed. We suggest to use 1 computer per 2 students. You may team up in advance. See how to install Matlab with our university licence at `https://it-wiki.es.aau.dk/wiki/Matlab`
- Kay's book and the lecture notes on Decision Rules.
- Paper and pencil.

Intended Learning Outcomes

After having attended this session you should be able to:

- Derive and implement the ML, MAP and Bayes' decision rules for a given hypothesis testing problem.
- Derive the error probability of a decision rule and compute it numerically using Monte Carlo simulations.
- Apply Bayes' costs to obtain decision rules that fulfill a given design criterion.

Reading Material

Main:
[KAY] Section 14.10 (real-world example)
Lecture notes in file: "decisionRules_S14.pdf", (see Moodle)

Additional Reading:
[LN] Chapter 3

Exercises

Will be handed out at the beginning of the session.

Session 16 – Group Work / Video Lecture: (Linear and Non-linear) MMSE Estimation and Vector LMMSE Estimation

Short Description

The problem of estimating the value of a certain unobservable quantity Y based on an available observation X is quite common. For this purpose we use an estimator, which is a function $\hat{Y} = g(X)$. There are many possible choices for g.

In this lecture we consider the *linear* estimator which *minimizes* the *mean square error* $\mathbb{E}[(Y - \hat{Y})^2]$. This is the so-called LMMSE estimator. Why is this particular estimator worth considering? Firstly, because requiring the mean square error to be as small as possible is in many cases sensible, secondly, since by restricting the function g to linear, the estimate $\hat{Y}$ is simple to compute. In this lecture we discuss the LMMSE estimator (scalar version) and the *orthogonality principle* which occurs when minimizing the mean square error of a linear estimator.

The LMMSE estimator forms the basis for the later derivation of so-called Kalman filters (the topic of Lecture 8 and Exercise Sessions 8 and 9).

In the previous lecture we have seen that the coefficient of the LMMSE estimator can be easily computed from the first-order moments (means) and second-order central moments (variances, covariances) of the unknown random variable Y and the observed random variables X. This is precisely what makes the LMMSE estimator so attractive, i.e. its simplicity (it is linear) and the fact that only these moments needs to be known. We do not need any further information on the statistical properties of X and Y.

Nevertheless, it is interesting to ask the question of an MSE-optimum estimator $\hat{Y} = g(X)$, when the linearity constraint on g is dropped. We show that this estimator exists and, in fact, is equal to the conditional expectation of Y given X, i.e. $\hat{Y}_{\text{opt}} = \mathbb{E}[Y|X]$. In this lecture we also recap the derivation of the vector LMMSE estimator and its mean square error.

Intended Learning Outcomes

After having attended this lecture and solved the exercises you should be able to:

- Recognize an estimation problem in an engineering context and state it rigorously in mathematical terms.
- Explain the optimality criterion and the main steps in the derivation of the orthogonality principle and the scalar LMMSE estimator.
- Derive and apply a scalar LMMSE estimator given the necessary covariance and cross-covariance structures.
- Compute and interpret the mean square error for the scalar LMMSE estimator.
- Explain the differences between the linear and non-linear MMSE estimators.
- Derive and apply a vector LMMSE estimator given the necessary covariance and cross-covariance structures.

- Apply scalar and vector LMMSE estimators to filtering, interpolation and prediction problems.

Reading Material

[Kay] Sections 7.9, 12.9, 14.3 and 14.8
Session 16 Notes: `MMSE_LMMSEnotes_S16.pdf`
Videos: SP 16.1 — SP 16.4. Links are provided in Moodle.

Extra Material:
[LN] Sections 4.1 and 4.2
Videos: SP 16.5 — SP 16.7.

Exercises

Will be handed out at the beginning of the session.

Session 17 – Group Work: Implementation, Simulation and Assessment of (L)MMSE estimators

Short Description

In this group work session, we focus on computing, implementing and simulating scalar LMMSE estimators in MATLAB.

Starting from the simple problem of linear prediction of the outcome of a random variable from the observation of another, we explore how the correlation between variables impacts the accuracy of the estimates. We then extend the problem to that of estimating the outcome of a random variable from the outcome of several observed variables.
To finalize, we approach the problem of linear filtering of a noisy random signal, which is commonly referred to as Wiener filtering/smoothing.
For the session you will need the following things:

- Your laptops with Matlab installed. We suggest to use 1 computer per 2 students. You may team up in advance. See how to install Matlab with our university licence at `https://it-wiki.es.aau.dk/wiki/Matlab`
- Kay's book and Lecture Notes.
- Paper and pencil.

Intended Learning Outcomes

After having attended this session you should be able to:

- Compute and interpret the correlation coefficient of two random variables.
- Derive and apply a scalar LMMSE estimator given the necessary covariance and cross-covariance structures.
- Implement LMMSE estimators in MATLAB and evaluate their accuracy by means of numerical simulations.
- Double-check the correctness of the implementation by comparing the numerical results obtained with those predicted by theory.

Reading Material

Session 16 Notes: `MMSE_LMMSEnotes_S16.pdf`
Videos: SP 16.1 — SP 16.4. Links are provided in Moodle.
[Kay] Section 7.9 (prediction, MSE and correlation)
[Kay] Section 12.9 (linear vs. non-linear prediction)
[Kay] Section 13.6 (conditional expectation)
[Kay] Section 14.8 (linear prediction)
[Kay] Section 18.5 (Wiener Filtering)

Extra Material:
[LN] Section 4.1
Videos: SP 16.5 — SP 16.7.

Exercises

Will be handed out at the beginning of the session.

Session 18 – Lecture: Kalman Filters (scalar and vector version)

Short Description

The *Kalman filter* is a recursive implementation of the LMMSE estimator exploiting the linear structure of the *system model* generating the unobservable sequence $\{Y(m)\}$ and of the *observation model* describing how the observable sequence $\{X(m)\}$ depends on $\{Y(m)\}$. The derivation of the equations of the Kalman filter can be easily done based on one of its important inherent property: the computation of the estimate of $Y(m+1)$ can be decomposed into two steps: (1) a prediction step computing the LMMSE estimators of $Y(m+1)$ and $X(m+1)$ based on the currently available observations $X(m), X(m-1), \ldots, X(0)$, (2) followed by an update step that updates the LMMSE prediction of $Y(m+1)$ based on the new observation $X(m+1)$. We show how the orthogonality principle can be extensively used to easily obtain the equations of the Kalman filter.

For the sake of simplicity we perform the derivation for the case where the system and observation models have scalar parameters. Then, we formulate the Kalman filter for the case where these models contain vector-valued parameters. This lecture concludes, if time permits, with a steady-state analysis when the system and observation models are time-invariant.

Intended Learning Outcomes

After having attended this lecture and solved the exercises you should be able to:

- Explain the relationship between the system model and the induced structure of the Kalman filter. State one or two main features of the Kalman filter.
- Derive the scalar Kalman filter equations for given system and observation models.
- Explain and compare how the filter equations would change for the vector case.
- Calculate the mean square error(s) of the scalar and vector Kalman filters.

Reading Material

[LN] Chapter 5 (Kalman Filters)

Exercises

Available in the file `ExercisesS18.pdf` in Moodle (“Lecture Notes” folder).

Session 19 – Group Work: Kalman Filtering

Short Description

In the last two group work sessions in the course, we study the *Kalman filter*. A typical engineering task in which the Kalman filter is employed usually consists of three steps:
(1) formulating the linear dynamic system model and the linear observation model
(2) deriving the Kalman filter for these models
(3) implementing the filter and assessing its performance.

Sessions 19 and 20 give you the opportunity to address all three steps.

Intended Learning Outcomes

The knowledge and experience gained throughout this session should enable you to:

- Formulate system models and observation models for practical problems.
- Implement and test a Kalman filter using computer simulation and subsequently apply it on measurement data.
- Identify and discuss if a given practical engineering problem is suitable for being handled/solved using a Kalman filter approach.

Reading Material

[LN] Chapter 5

Exercises

Will be handed out at the beginning of the session.

Session 20 – Group Work: Kalman Filtering (cont'd)

Short Description

In the last two group work sessions in the course, we study the *Kalman filter*. A typical engineering task in which the Kalman filter is employed usually consists of three steps:
(1) formulating the linear dynamic system model and the linear observation model
(2) deriving the Kalman filter for these models
(3) implementing the filter and assessing its performance.

Sessions 19 and 20 give you the opportunity to address all three steps.

Intended Learning Outcomes

The knowledge and experience gained throughout this session should enable you to:

- Formulate system models and observation models for practical problems.
- Implement and test a Kalman filter using computer simulation and subsequently apply it on real-world measurement data.
- Identify and discuss if a given practical engineering problem is suitable for being handled/solved using a Kalman filter approach.

Reading Material

[LN] Chapter 5

Exercises

Will be handed out at the beginning of the session.