

# Weekly Learning Plan — MATH38161

Korbinian Strimmer

29 September 2020

## Overview

### General plan

This module is taught based on the concept of “flipped learning”:

- 1) First, you study the asynchronous material by yourself:
  - i) read the [lecture notes](#),
  - ii) watch the [lecture videos](#) and
  - iii) try to solve the questions and problems given in the [worksheets](#) and study their solutions.

For each of the semester weeks 1–11 the following pages provide a detailed list which material you will need to study in a specific week.

- 2) Subsequently, you attend the live synchronous sessions:
  - i) the review lecture and
  - ii) the tutorial session.

The review lecture recapitulates the key content from the lecture notes and videos each week. The tutorial session discuss the questions and problems in an interactive fashion.

### Links

You can find the course material both on [Blackboard](#) and on the [MATH38161 course homepage](#). There can also find the scheduled times and locations for the two weekly live sessions.

### Workload

For this course you will need to spend about 6-7 hours study each week:

- 2-3h watching videos, reading and understanding lecture notes
- 2h answering questions and solving and understanding problems in worksheets
- 2h attending live sessions

The in-semester assessment will require about 10h of additional work.

## Week 1 (from 5 October 2020)

**LN:** lecture notes — [view online](#)

**Video:** lecture videos — [view online](#)

**Worksheets:** [download worksheets](#)

*Topics:*

In this week first an overview over the course content and course organisation is given. The essentials in matrix theory are reviewed and an introduction to multivariate random variables is given. The tutorial provides information about installing R and R Studio as well as the R packages used later in the course. It also introduces R Markdown for document preparation mixing text, formulas, R code and R output.

*Study:*

---

**Video 1** Introduction to the course

**LN Preface**

**LN Section 1.1**

---

**Video 2** Refresher on matrices and matrix calculus

**LN Appendix A**

---

**Video 3** Multivariate random variables

**LN Section 1.2**

---

*Questions and problem solving:*

---

**Worksheet 1**

---

## Week 2 (from 12 October 2020)

**LN:** lecture notes — [view online](#)

**Video:** lecture videos — [view online](#)

**Worksheets:** [download worksheets](#)

*Topics:*

In this week we study the multivariate normal distribution and how to estimate its parameters. In addition, a number of further continuous and discrete multivariate distributions are discussed as generalisation of their univariate counterparts.

*Study:*

---

**Video 4** Multivariate normal distribution  
**LN Section 1.3**

---

**Video 5** Estimation of multivariate mean and covariance  
**LN Section 1.4**

---

**Video 6** Further multivariate distributions  
**LN Sections 1.5 and 1.6**

---

*Questions and problem solving:*

---

**Worksheet 2**

---

## Week 3 (from 19 October 2020)

**LN:** lecture notes — [view online](#)

**Video:** lecture videos — [view online](#)

**Worksheets:** [download worksheets](#)

*Topics:*

The focus of this week are linear and nonlinear transformations for random vectors and their properties. As special class we consider general whitening transformations that form the basis of many multivariate statistical models and procedures.

*Study:*

---

**Video 7** Linear transformations

**LN Section 2.1**

---

**Video 8** Nonlinear transformations

**LN Section 2.2**

---

**Video 9** General whitening transformation

**LN Section 2.3**

---

*Questions and problem solving:*

---

**Worksheet 3**

---

## Week 4 (from 26 October 2020)

**LN:** lecture notes — [view online](#)

**Video:** lecture videos — [view online](#)

**Worksheets:** [download worksheets](#)

*Topics:*

In this week we study properties of five commonly used natural whitening procedures and discuss related methods such as principal component analysis and canonical correlation analysis.

*Study:*

---

**Video 10** Natural whitening transformations

**LN Section 2.4**

---

**Video 11** Principal component analysis

**LN Sections 2.5 and 2.6**

---

**Video 12** Canonical correlation analysis

**LN Section 2.7**

---

*Questions and problem solving:*

---

**Worksheet 4**

---

## Week 5 (from 2 November 2020)

**LN:** lecture notes — [view online](#)

**Video:** lecture videos — [view online](#)

**Worksheets:** [download worksheets](#)

*Topics:*

In this week we discuss the problem of unsupervised learning and focus on algorithmic methods for clustering such as hierarchical clustering based on trees and the  $K$ -means algorithm.

*Study:*

---

**Video 13** Unsupervised learning and clustering

**LN Section 3.1**

---

**Video 14** Hierarchical clustering

**LN Section 3.2**

---

**Video 15**  $K$ -means algorithm

**LN Section 3.3**

---

*Questions and problem solving:*

---

**Worksheet 5**

---

## Week 6 (from 9 November 2020)

**LN:** lecture notes — [view online](#)

**Video:** lecture videos — [view online](#)

**Worksheets:** [download worksheets](#)

*Topics:*

In this week we continue with supervised learning, focusing on probabilistic models in the form of (Gaussian) mixture models, and show how to fit these models to observations using the EM algorithm.

*Study:*

---

**Video 16** Mixture models

**LN Section 3.4**

---

**Video 17** Fitting mixture models to data

**LN Section 3.5**

---

*Questions and problem solving:*

---

**Worksheet 6**

---

## Week 7 (from 16 November 2020)

**LN:** lecture notes — [view online](#)

**Video:** lecture videos — [view online](#)

**Worksheets:** [download worksheets](#)

*Topics:*

In this week we consider supervised learning problems and introduce the general Bayes classifier. Subsequently, three widely used classification methods (QDA, LDA and DDA) are developed and demonstrated how they can be applied to data.

*Study:*

---

**Video 18** Supervised learning and the Bayes classifier  
**LN Sections 4.1 and 4.2**

---

**Video 19** Normal Bayes classifiers (QDA, LDA, DDA)  
**LN Section 4.3**

---

**Video 20** Training QDA and LDA classifiers from data  
**LN Section 4.4**

---

*Questions and problem solving:*

---

**Worksheet 7**

---



## Week 8 (from 23 November 2020)

**LN:** lecture notes — [view online](#)

**Video:** lecture videos — [view online](#)

**Worksheets:** [download worksheets](#)

*Topics:*

In this week we focus how to select features suitable for prediction in classification and how to evaluate prediction accuracy in supervised learning settings by cross-validation.

*Study:*

---

**Video 21** Variable ranking in classification

**LN Section 4.5**

---

**Video 22** Variable selection and cross-validation

**LN Section 4.6** + further reading

---

Note: in LN Section 4.6.3 you find specific directions for further reading.

*Questions and problem solving:*

---

**Worksheet 8**

---

*In-semester assessment:*

Instructions for the project are available on [Blackboard](#) on **Monday 23 November 2020, 12 noon**.

## Week 9 (from 30 November 2020)

**LN:** lecture notes — [view online](#)

**Video:** lecture videos — [view online](#)

**Worksheets:** [download worksheets](#)

*Topics:*

In this week we study the vector correlation to measure linear association between two random vectors, discuss mutual information to measure nonlinear association and introduce the essentials of graphical models to describe the dependency patterns among random variables.

*Study:*

---

**Video 23** Vector correlation and RV coefficient

**LN Section 5.1**

---

**Video 24** Mutual information as generalisation of correlation

**LN Section 5.2**

---

**Video 25** Graphical models

**LN Section 5.3**

---

*Questions and problem solving:*

---

**Worksheet 9**

---

*In-semester assessment:*

Work on project! Submission deadline is next week.

## Week 10 (from 7 December 2020)

**LN:** lecture notes — [view online](#)

**Video:** lecture videos — [view online](#)

**Worksheets:** [download worksheets](#)

*Topics:*

In this week we move away from linear models to nonlinear models and introduce random forests, a nonparametric prediction model.

*Study:*

---

**Video 26** Limits of linear models and correlation  
**LN Section 6.1**

---

**Video 27** Random forests  
**LN Section 6.2** + further reading

---

Note: in LN Sections 6.2 and 6.3 you find specific directions for further reading.

*Questions and problem solving:*

---

**Worksheet 10**

---

*In-semester assessment:*

Project submission deadline is on **Monday 7 December 2020, 12 noon**.

## Week 11 (from 14 December 2020)

**LN:** lecture notes — [view online](#)

**Video:** lecture videos — [view online](#)

**Worksheets:** [download worksheets](#)

*Topics:*

In this week we provide a broad overview of Gaussian processes and neural networks, two commonly used machine learning approaches for nonlinear modeling and prediction.

*Study:*

---

**Video 28** Gaussian processes  
**LN Section 6.3** + further reading

---

**Video 29** Neural networks  
**LN Section 6.4** + further reading

---

Note: in LN Sections 6.3 and 6.4 you find specific directions for further reading.

*Questions and problem solving:*

---

**No worksheet this week**

---

The tutorial this week is an exam Q & A.