Hertie School **Master of Data Science for Public Policy**
Spring Semester 2022

Course Syllabus, Version 8.12.2021

GRAD-C23: Mathematics for Data Science

Lynn Kaack and Slava Jankin

1. General information

Class time	Tuesday, 12:00-14:00 in the Forum.
Course Format	Both the lectures and the accompanying labs take place onsite.
Instructor	Lynn Kaack Slava Jankin
Instructor's office	3.22 (Lynn Kaack)
Instructor's e-mail	jankin@hertie-school.org kaack@hertie-school.org
Assistant	Name: Alex Karras Email: karras@hertie-school.org Phone: +49 30 259 219 156 Room: 3.45
Instructor's Office Hours	Email assistant

Link to MIA and MPP **Module Handbooks**
Link to **Study, Examination and Admission Rules**

Instructor Information:
Lynn Kaack is Assistant Professor of Computer Science and Public Policy at the Hertie School. Her research and teaching focus on methods from statistics and machine learning to inform climate mitigation policy across the energy sector, and she also has an interest in climate-related AI policy. She is a co-founder and chair of the organisation Climate Change AI, and was a member of the Austrian Council on Robotics and Artificial Intelligence. Previously she was Postdoctoral Researcher and Lecturer in the Energy and Technology Policy Group at ETH Zürich. She obtained a PhD in Engineering and Public Policy and a Master's in Machine Learning from Carnegie Mellon University, as well as a MS and BS in Physics from the Free University of Berlin.

Slava Jankin is Professor of Data Science and Public Policy at the Hertie School. He is the Director of the Hertie School Data Science Lab. His research and teaching are primarily in the field of natural language processing and machine learning. Before joining the Hertie School faculty, he was a Professor of Public Policy and Data Science at University of Essex, holding a joint appointment in the Institute for Analytics and Data Science and Department of Government. At Essex, Slava served as a Chief Scientific Adviser to Essex County Council, focusing on artificial intelligence and data science in public services. He previously worked at University College London and London School of Economics. Slava holds a PhD in Political Science from Trinity College Dublin.

2. Course Contents and Learning Objectives

Course contents:

This course aims to deliver a compact and tailored introduction to the core mathematical concepts of data science. Linear algebra, probability theory, statistics, and optimisation are mathematical pillars underlying the practice of data science. The course covers foundational mathematical concepts such as statistical estimation, norms, matrix algebra, Lagrange Multipliers and many more in theory and practice. Upon completing the course, students will have a broad knowledge of linear algebra, probability theory, statistics, and optimisation necessary to understand the theoretical underpinnings of modern statistics and machine learning methods. **Note: "Mathematics for Data Science" accompanies the "Machine Learning" course, it covers fundamental concepts that are applied in "Machine Learning."**

Main learning objectives:

This course covers the foundational mathematical concepts underlying statistical and data science techniques. Students will be introduced to many different mathematical areas and will find that much of what is covered in this course will be directly applicable in other courses they are taking, most importantly "Machine learning" as well as "Causal inference"

Teaching style:

This course is taught with lectures that cover the main content of the course and small-group support lab session taught weekly by a teaching assistant. Lectures will be accompanied by regular homework assignments and hands-on tutorials for students to develop deeper understanding of the material.

Prerequisites:

You should be familiar with the most common approaches in calculus, optimization, and linear algebra, e.g., you should know how to multiply matrices. The course will recap those basics but move quickly to more advanced concepts.

Diversity Statement:

This course covers many different concepts, and students may have had different previous exposure to these concepts. We encourage the class participants to acknowledge the diversity in experience and backgrounds and create an environment where everyone feels comfortable to ask clarifying questions and voice their struggle. Mutually supporting each other will be at the heart of success in this class, and we hope that more advanced students will help their peers. Students are encouraged to bring up any hardships that they encounter to the lecturers, and we will make an effort to resolve them.

3. Grading and Assignments

Composition of Final Grade:

Assignment 1: Problem sets	Deadline: As indicated on each problem set	Submit via Moodle or as otherwise indicated	30%

Assignment 2: Midterm exam	Deadline: Mid-term exam week	In-class exam	35%
Assignment 3: Final exam	Deadline: Final exam week	In-class exam	35%

Assignment Details

Assignment 1: Problem sets
This class will have up to five regular homework assignments in the form of problem sets. The assignments consist of different mathematical problems that need to be solved by hand and potentially by writing and running code. By completing the assignments students gain a deeper understanding of the topics covered in class, and experience in solving those problems.

Assignment 2: Midterm exam
This class will have a midterm exam. The exam will cover the most important topics discussed in class or covered in core readings until that date. The exam may include mathematical exercises, multiple choice questions and questions that require written answers.

Assignment 3: Final exam
This class will have a final exam. The exam will cover the most important topics discussed in class or covered in core readings. The exam may include mathematical exercises, multiple choice questions and questions that require written answers.

Late submission of assignments: For each day the assignment is turned in late, the grade will be reduced by 10% (e.g. submission two days after the deadline would result in 20% grade deduction).

Attendance: Students are expected to be present and prepared for every class session. Active participation during lectures and seminar discussions is essential. If unavoidable circumstances arise which prevent attendance or preparation, the instructor should be advised by email with as much advance notice as possible. Please note that students cannot miss more than two out of 12 course sessions. For further information please consult the Examination Rules §10.

Academic Integrity: The Hertie School is committed to the standards of good academic and ethical conduct. Any violation of these standards shall be subject to disciplinary action. Plagiarism, deceitful actions as well as free-riding in group work are not tolerated. See Examination Rules §16 and the Hertie Plagiarism Policy.

Compensation for Disadvantages: If a student furnishes evidence that he or she is not able to take an examination as required in whole or in part due to disability or permanent illness, the Examination Committee may upon written request approve learning accommodation(s). In this respect, the submission of adequate certificates may be required. See Examination Rules §14.

Extenuating circumstances: An extension can be granted due to extenuating circumstances (i.e., for reasons like illness, personal loss or hardship, or caring duties). In such cases, please contact the course instructors and the Examination Office *in advance* of the deadline.

4. General Readings

- On linear algebra: Strang, G., 2019. *Linear algebra and learning from data*. Cambridge: Wellesley-Cambridge Press. [We will designate it as **Strang** throughout]

- On statistics and probability theory: Wasserman, Larry. All of statistics: a concise course in statistical inference. Vol. 26. New York: Springer, 2004. [We will designate it as **Wasserman** throughout]
- On optimization: Vandenberghe, Lieven, and Stephen Boyd. Convex optimization. Vol. 1. Cambridge: Cambridge University Press, 2004. https://stanford.edu/~boyd/cvxbook/ [We will designate it as **Vandenberghe and Boyd** throughout]
- On statistical learning: Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. (2009). *The Elements of Statistical Learning: Data mining, inference, and prediction*, 2nd edition, available at https://web.stanford.edu/~hastie/ElemStatLearn/. [We will designate it as **HTF** throughout]

Other general references:

- James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. An introduction to statistical learning. Vol. 112. New York: Springer, 2013. [We will designate it as **ISL** throughout]
- Bishop, Christopher M. "Pattern recognition." Machine learning128, no. 9 (2006). Chapter 1.6 [We will designate it as **Bishop** throughout.]
- Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. Deep learning. MIT press, 2016. Chapter 3.1-11 https://www.deeplearningbook.org [We will designate it as **GBC** throughout.]
- Kolter, Zico “Linear Algebra Review and Reference” https://www.cs.cmu.edu/~zkolter/course/15-884/linalg-review.pdf

5. Session Overview

Session	Session Date	Session Title
1	08.02.2022	Introduction and Probability Theory
2	15.02.2022	Calculus
3	22.02.2022	Linear Algebra (1)
4	01.03.2022	Statistical Learning (1)
5	15.03.2022	Statistical Learning (2): Linear Regression and Regularization
Mid-term Exam Week: 21 – 25.03.2022 – no class		
6	29.03.2022	Statistical Learning (3): Model Assessment and Selection
7	05.04.2022	Information Theory
8	12.04.2022	Optimization and Search (2)
9	19.04.2022	Linear Algebra (2) and Related Methods
10	26.04.2022	Linear Algebra (3) and Optimization (2): Neural Networks
11	03.05.2022	Sample Complexity and Generalization
12	10.05.2022	Data Science and Ethics
Final Exam Week: 16 – 20.05.2022 – no class		

6. Course Sessions and Readings

All readings will be accessible on the Moodle course site before semester start. In the case that there is a change in readings, students will be notified by email.

All readings are intended to supplement the class session and facilitate your own learning. Core readings are more fundamental to the course content, please skim them before class to ensure you are familiar with the concepts they build on. Optional readings are intended to either provide more material on fundamental concepts or broaden your knowledge in the respective area, and it is highly recommended to at least skim them.

Session 1: Introduction and Probability Theory	
Learning Objective	We will cover course logistics, and dive into core concepts of probability theory such as random variables, common distributions, and Bayes' rule.
Core Readings	GBC Chapter 3.1-11
Optional Readings	Strang Part V, Wasserman Chapter 1-3

Session 2: Calculus	
Learning Objective	We recap calculus in this session and familiarize ourselves with multivariate calculus. We also cover the basics of optimization.
Core Readings	Zhang, Lipton, Li, and Smola, *Dive into Deep Learning:* • 2.4 Calculus http://d2l.ai/chapter_preliminaries/calculus.html • 18.3 and 18.4 Appendix: Mathematics for Deep Learning https://d2l.ai/chapter_appendix-mathematics-for-deep-learning/index.html
Optional Readings	
Session 3: Linear Algebra (1)	
Learning Objective	We recap foundational concepts of linear algebra, including eigenvalues, the scalar product, matrix calculus etc. We then introduce norms and other linear algebra concepts useful for statistical learning.
Core Readings	Strang Part I 1-7, 11
Optional Readings	GBC Chapter 2 Z. Kolter, Linear Algebra Review and Reference https://www.cs.cmu.edu/~zkolter/course/15-884/linalg-review.pdf

Session 4: Statistical Learning (1)	
Learning Objective	This class provides a reminder about some of the most important concepts from statistics. We will take a closer look at statistical estimation of parameters from data, and cover Maximum Likelihood Estimation, Maximum a Posteriori Estimation, and cross validation, among others.

Core Readings	HTF: Chapter 2
Optional Readings	ISL: Chapter 2

Session 5: Statistical Learning (2): Linear Regression and Regularization	
Learning Objective	We look at a familiar model through a different lens: linear regression as the least squares estimator. We then cover regularized methods such as Lasso and Ridge Regression.
Core Readings	HTF Chapter 3.2.1-2
Optional Readings	Remainder of HTF Chapter 3 HTF: Chapter 3 Bishop Chapter 3 through 3.2

Session 6: Statistical Learning (3): Model Assessment and Selection	
Learning Objective	We discuss bias, variance, and model complexity. We focus on the bias-variance tradeoff. We cover approaches to assess model performance. We discuss cross-validation and bootstrap methods.
Core Readings	HTF: Chapter 7
Optional Readings	ISL: Chapter 5

Mid-term Exam Week: 21 – 25.03.2022 – no class

Session 7: InformationTheory	
Learning Objective	This class provides an introduction to information theory, which provides the basis for decision tree learning. We cover concepts such as the information gain and entropy.
Core Readings	Bishop Chapter 1.6
Optional Readings	GBC Chapter 3.13

Session 8: Optimization and Search (1)	
Learning Objective	After a brief recap of the foundational concepts needed for optimization, we build the basis for computational approaches to find the optimal hypothesis to fit the data. We also cover search algorithms.
Core Readings	Strang Part VI
Optional Readings	Vandenberghe and Boyd: Introduction

Session 9: Linear Algebra (2) and Related Methods

Learning Objective	We discuss the curse of dimensionality and applications in high-dimensional space. We also cover core dimensionality reduction techniques, including principle component analysis.
Core Readings	HTF: Chapters 14.5, 14.7-14.9
Optional Readings	Remainder of HTF: Chapter 14 HTF: Chapter 18 Strang Part I remainder

Session 10: Linear Algebra (3) and Optimization (2): Neural Networks	
Learning Objective	This class will cover more concepts from matrix algebra and optimization. We will take a closer look at gradient descent and the backpropagation algorithm.
Required Readings	HTF Chapter 11
Optional Readings	Bishop Chapter 5.2 ISL Chapter 10.7 Strang VII.1, 3

Session 11: Sample Complexity and Generalization	
Learning Objective	We take a closer look at how learning algorithms generalize. Specifically, we cover concepts around the question of how many training examples are needed for a particular algorithm to learn the target function and cover the "No free lunch theorem."
Required Readings	HTF Chapter 7.9
Optional Readings	

Session 12: Data Science and Ethics	
Learning Objective	We will discuss common concepts related to responsible AI, such bias and fairness.
Required Readings	Barocas, Hardt, and Narayanan, *Fairness and Machine Learning*, https://fairmlbook.org/index.html
Optional Readings	Kleinberg, Mullainathan, Raghavan, *Inherent Trade-Offs in the Fair Determination of Risk Scores* https://arxiv.org/pdf/1609.05807.pdf

Final Exam Week: 16 – 20.05.2022 – no class