Advanced Statistical Inference
Introduction

Motonobu Kanagawa
motonobu.kanagawa@eurecom.fr

Data Science Department
EURECOM

February 21, 2024

Contents of the Course

- Introduction to Bayesian statistics/machine learning.
  - Statistics: Estimating a population parameter (e.g., mean) from finite data.
  - Machine learning: Learning a predictive model (=AI) from finite training data.

Contents of the Course

- Introduction to Bayesian statistics/machine learning.
  - Statistics: Estimating a population parameter (e.g., mean) from finite data.
  - Machine learning: Learning a predictive model (=AI) from finite training data.

- Data being finite causes uncertainties in parameter estimation and prediction.
  - For example, for autonomous driving, training data cannot cover all possible situations.

Contents of the Course

- Introduction to Bayesian statistics/machine learning.
  - Statistics: Estimating a population parameter (e.g., mean) from finite data.
  - Machine learning: Learning a predictive model (=AI) from finite training data.

- Data being finite causes uncertainties in parameter estimation and prediction.
  - For example, for autonomous driving, training data cannot cover all possible situations.

- Bayesian approaches enable statistics and machine learning quantifying uncertainties in a principled manner.

Schedule (tentative)

- Feb 21: Lecture (Intro, linear algebra, probability - recaps).
- Mar 06: Lecture (Bayesian inference 1)
- Mar 13: Lecture (Bayesian inference 2 + Bayes linear regression 1)
- Mar 20: Lecture (Bayes linear regression 2)
- Mar 27: Lab (Bayesian linear regression, Coin toss experiment)
- Apr 03: Lecture (Gaussian process regression)
- Apr 10 Lab (Gaussian process regression)
- Apr 17:Lecture (Logistic regression with MCMC 1)
- Apr 24 Lecture (Logistic regression with MCMC 2)
- May 15 Lab (Logistic regression with MCMC)
- May 22 Lecture: (ML performance evaluation/Variational Inference 1)
- May 29 Lecture: (Variational Inference 2)
- Jun 05: Lab (Variational inference)
- Jun 12: Lecture TBD

Pre-requisites

- We represent objects as vectors/matrices (arrays of numbers), so we have to do maths.
- Being familiar with calculus (integration, differentiation ...)
- Good understanding of probabilities
- Good understanding of linear algebra
- Today we will do a brief recap of these math subjects.

Grade Evaluation

- Total of 20 points
  - Labs ($1 \times 4 = 4$ points)
  - Attendance of lectures and labs.
    - During each lecture, I will uniformly randomly choose students to ask several questions. If you are chosen but not at the lecture, I will subtract 2 points from your grade. (The same student can be selected multiple times during the course, of course.)
    - If you cannot come to the lectures for a valid reason (e.g., being sick, strike, etc), please write me an email.
  - Written exam (16 points)

Lecturers

- **Lectures**: Motonobu Kanagawa
  - Assistant Professor in the Data Science Department
  - Statistics, machine learning, and simulation
  - motonobu.kanagawa@eurecom.fr
- **Labs**: Nugzar Gognadze
  - PhD student in the Data Science Department
  - Statistical learning, application in geophysics
  - nugzar.gognadze@eurecom.fr

Disclaimer

- Until the last year, the ASI course was taught by Prof Maurizio Filippone.
- However, Prof Filippone left EURECOM last December ...
- That's why I am teaching this course.
- Therefore, most of the teaching materials are based on Prof Filippone's materials.

Suggested readings

A First Course in Machine Learning

S. Rogers and M. Girolami

Pattern Recognition and Machine Learning

C. Bishop

Information Theory, Inference, and Learning Algorithms

D. MacKay

Machine Learning: A Probabilistic Perspective

K. P. Murphy

Suggested readings

Bayesian Data Analysis

Andrew Gelman

Bayesian Reasoning and Machine Learning

David Barber

Machine Learning

Peter Flach