Engineering Tripos Part IIA, 3F8: Inference, 2017-18

Leader

Dr Jose Miguel Hernandez-Lobato

Lecturer

Dr Jose Miguel Hernandez-Lobato

Lab Leader

Dr Jose Miguel Hernandez-Lobato

Timing and Structure

Lent Term.

Prerequisites

3F3 Statistical Signal Processing

Aims

The aims of the course are to:

- Provide a thorough introduction into the topic of statistical inference including maximum-likelihood and Bayesian approaches
- Introduce inference algorithms for regression, classification, clustering and sequence modelling
- Introduce basic concepts in optimisation, dynamic programming and Monte Carlo methods

Objectives

As specific objectives, by the end of the course students should be able to:

- Understand the use of maximum-likelihood and Bayesian inference and the strengths and weaknesses of both approaches.
- Implement methods to solve simple regression, classification, clustering and sequence modelling problems.
- Implement simple optimisation methods (gradient and coordinate descent, stochastic gradient descent), dynamic programming (Kalman filter or Viterbi decoding) and simple Monte Carlo methods (importance sampling, rejection sampling, ancestral sampling).

Content

Introduction to inference (2L)

- [List of topics]

- [Additional notes or links]
Regression (2L)
- What are regression problems
- Revision of properties of Gaussian probability density
- Maximum likelihood and Bayesian fitting of Gaussians
- Linear regression and non-linear regression

Classification (2L)
- Classification problems
- Logistic regression probabilistic model
- Training logistic regression using optimisation
- Stochastic optimisation methods
- Non-linear feature expansions for logistic regression

Dimensionality Reduction (2L)
- What is dimensionality reduction
- Principal component analysis as minimising reconstruction cost
- Principal component analysis as inference

Clustering (2L)
- What is clustering
- The k-means algorithm
- Gaussian Mixture Models
- The Expectation Maximisation (EM) Algorithm

Sequence models (3L)
- Sequence modelling problems
- Markov Models and Hidden Markov models
- Inference in Hidden Markov Models using dynamic programming

Basic Monte Carlo (2L)
- The need for approximate inference methods
- Simple Monte Carlo
- Exact sampling
- Rejection sampling
- Importance sampling
- Ancestral sampling

Further notes
Lecture allocations above are approximate.

Coursework
To implement an algorithm for performing classification, called logistic regression, using gradient descent optimisation.

[Coursework Title]

Learning objectives:

Practical information:
- Sessions will take place in [Location], during week(s) [xxx].
- This activity [involves/doesn't involve] preliminary work ([estimated duration]).

Full Technical Report:
Students [will/won't] have the option to submit a Full Technical Report.

**Booklists**

There is no required textbook. However, the material covered is treated excellent recent text books:


David Barber Bayesian Reasoning and Machine Learning [3], Cambridge University Press (2012), available freely on the web.


**Examination Guidelines**

Please refer to Form & conduct of the examinations [6].

Last modified: 03/08/2017 15:42