C19 : Unsupervised Machine Learning

Frank Wood

University of Oxford

January, 2017

Introduction

- Supervised machine learning
  - = regression/classification = conditional density/distrib. estimation
  - Requires training pairs; input/output
  - Sometimes "easier"
    - Lower dimensional parameter space (needn't model inputs)
- Unsupervised Machine Learning
  - = dimensionality reduction = joint density estimation
  - Requires data + model
  - Often harder
    - Must model inputs
    - Richer models
  - Many supervised models arrive from conditioning unsupervised models on observed data
    - Conditioning on observed data = fixing values of some variables in joint
  - Inference techniques for unsupervised models automatically work for supervised models

Learning Goals

- Understand connection between graphical models and joint distributions (Bishop, Ch. 8; Murphy Ch. 10, 19)
- Be able to "invent" graphical models for problems of interest
- Understand sampling techniques (Bishop, Ch. 11; Murphy Ch. 24)
  - Basic sampling techniques
  - Markov chain Monte Carlo (MCMC)
    - Metropolis Hasting (MH)
    - Gibbs
- Understand conjugacy and how to exploit it for analytic marginalization (Bishop pg. 117; Murphy pg. 74)
- Understand how to derive and implement MCMC samplers for arbitrary graphical models
- Understand Monte Carlo integration
- Understand how to formulate inference questions in terms of integrals

Resources

- Lecture Notes and Problem Sheet :

  `http://www.robots.ox.ac.uk/~fwood/teaching/C19_hilary_2016_2017/`

- Books :
  - Pattern Recognition and Machine Learning [Bishop, 2007]
  - Bayesian data analysis [Gelman et al., 1995]
  - Machine Learning : a probabilistic perspective [Murphy, 2012]
  - Sequence Monte Carlo Methods in Practice [Doucet et al., 2001]
  - (free online) Information Theory, Inference, and Learning Algorithms [MacKay, 2003]
- Tutorials :
  - Probabilistic Inference using MCMC methods [Neal, 1993]
  - A Tutorial on particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking [Arulampalam et al., 2002]

Bibliography I

M Sanjeev Arulampalam, Simon Maskell, Neil Gordon, and Tim Clapp. A Tutorial on Particle Filters for Online Nonlinear / Non-Gaussian Bayesian Tracking. *IEEE Transactions on Signal Processing*, 50(2): 174–188, 2002. ISSN 1053587X. doi: 10.1109/78.978374. URL `http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=978374`.

Christopher M Bishop. *Pattern Recognition and Machine Learning*. Springer, 2007.

A Doucet, N de Freitas, and N Gordon, editors. *Sequential Monte Carlo Methods in Practice*. Springer-Verlag, 2001.

A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin. *Bayesian data analysis*. Chapman & Hall, New York, 1995.

D. J. C. MacKay. *Information theory, inference, and learning algorithms*. Cambridge University Press, Cambridge, 2003.

Bibliography II

Kevin P Murphy. *Machine learning: a probabilistic perspective*. The MIT Press, 2012.

Radford M. Neal. Probabilistic inference using Markov chain Monte Carlo methods. *Technical Report*, 1993.