Advanced probabilistic methods

Recap, Exam information

Pekka Marttinen

Aalto University

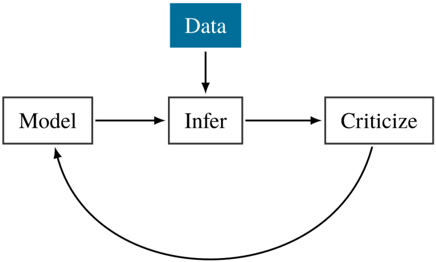
March, 2019

Lecture 10 overview

- Summary of the course
- Exam Information
- Recap, including last year's exam

Course contents

- Ingredients of probabilistic modeling
  - **Models**: Bayesian networks, Sparse Bayesian linear regression, Gaussian mixture models, latent linear models
  - **Methods for inference**: maximum likelihood, maximum a posteriori (MAP), Laplace approximation, expectation maximization (EM), Variational Bayes (VB), Stochastic variational inference (SVI)
  - **Ways to select between models:** Bayesian model selection, AIC, BIC, ...



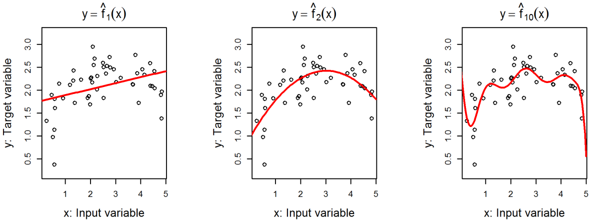
Box's loop (Blei, 2014)

Which model to use?

- Incremental approach
  - start with a simple model
  - assess fit (inspect residuals, compare model prediction with real data,...)
  - improve model if needed
  - See *Philosophy and practice of Bayesian statistics* (2012) by Gelman and Shalizi
- "Let the data speak"
  - start with a flexible model that can adapt to different kinds of data
  - often used in machine learning
- Construct alternative models based on understanding of the problem
  - Continue with model comparison (checking the fit still needed)

Which inference to use? (1/2)

- If the model is good, the results are often ok with any sensible model fitting technique.
- Bayesian approaches avoids some pitfalls of maximum likelihood (e.g. singularities in GMM) and yield uncertainty estimates and regularization
- Probabilistic programming (e.g. Stan for MCMC, Edward for VI) may be considered as the first thing to try, if analytical integration is not possible.
  - Easy and fast if they work
  - Difficult to debug if they don't



Which inference to use? (2/2)

- **MCMC**:
  - pros: asymptotically you get correct probabilities
  - cons: assessing convergence, computation time
  - Recommeneded usage[1]: any 'conventional' statistical analysis where it's important to get the probabilities correct and computation time is not an issue
- **Variational inference**:
  - pros: faster, may be more stable than MCMC
  - cons: underestimates uncertainty
  - Recommeneded usage: 'machine learning' type models, where the goal is more on prediction and exact probabilities are not needed (e.g. deep learning models)

[1] Based on the present understanding of the lecturer. Other views may also be justified, therefore: think what is suitable in your case and try different options.

Exam

- April 9th, 13:00-17:00, R037/AS2 (confirm from oodi!)
- In the exam, you are allowed to have with you
  - A laptop, ipad, or similar, with which you can read PDFs -> **must be disconnected** for the duration of the whole exam, i.e., turn off wifi, Blutooth etc.
  - Any documents you can find under 'Materials' and 'Assignments' in myCourses, **excluding 'Additional reading'**. **Download these before the exam**. Alternatively, these can be printed if you don't have the device to view the materials otherwise.
  - Calculator with memory erased.
  - Calculator with 'symbolic calculation' (i.e. ability to simplify formulas, integrals, etc.) is not allowed.
  - other conventional equipment: pencil, eraser,...
- Make sure you have a valid registration!

Exam contents

- 5 questions, 6 points each
- At least 1 question about Bayesian networks
- At least 2 questions about deriving (some part of) an inference algorithm for some *simple* model (EM algorithm, variational inference, Laplace approximation, Stochastic variational inference)
- 1-2 questions on other topics (for example, model selection, multivariate distributions, usage of the models, interpretation of model outputs, Edward)
- Questions will be similar to the exercises & examples on lectures

Reminders

- Please do give feedback, this will be used when developing the course further next year.

Questions?

- Good luck for the exam!
- Any questions?