Applied Probabilistic Programming & Bayesian Machine Learning

Syllabus

MIT IAP 2017 — Max W. Shen

1 About

This class, titled "Applied Probabilistic Programming & Bayesian Machine Learning", presents topics in Bayesian machine learning for the applied practitioner.

Our goal is for students to come away with a useful tool, probabilistic programming, that they can directly apply to datasets and problems of interest.

Towards this goal, we present material through a different lens than nearly any textbook, motivated first and foremost by the tool of probabilistic programming. We present topics in Bayesian statistics through the lens of "things you can do" with this tool of probabilistic programming, and focus on "how to do it".

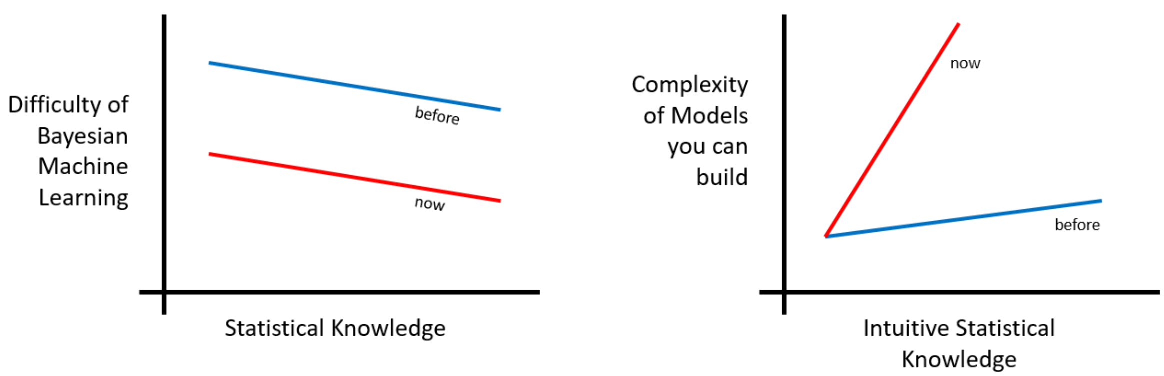

Figure 1: Assumptions

The reasoning for this approach is motivated by a couple core assumptions, including the ones depicted in 1.

Bayesian machine learning has traditionally required a deep foundation of statistics and math. With the advent of probabilistic programming languages, for the first time ever, it is now possible to teach Bayesian machine learning with applications in mind before theory. Is this a good idea? Well, in the grand spirit of IAP, this is our humble experiment,

2 The Experiment

This is a rough, first-draft version of the entire course. The scope is ambitious, and we are a few young grad students who sometimes feel like we've gotten way in over our heads. We ask your patience with mistakes and errors which are bound to exist, and also appreciate your assistance if you catch something. You can find our contact information at the end of this syllabus.

3 Logistics

We meet at 4-237 on Tuesdays and Thursdays during IAP 2017 at MIT, from 5pm - 6:30pm. To find building 4, use: whereis.mit.edu - room 273 is on the 2nd floor.

Our choice of language is Stan, with Python or R used for data pre-processing and simulations.

All of the class material is located at:

`https://github.com/maxwshen/iap-appbml`

This includes coding exercises, pdf's, and datasets in csv format. Coding exercises are performed in any environment you find comfortable with the intention of making it as easy as possible to apply class material to your own datasets and problems of interest.

Python, R, pystan, and RStan are used in the class - see the github repository for assistance on installation.

4 What is Probabilistic Programming?

For the purposes of this course, we choose to discuss probabilistic programming as representing the myriad of recent developments in automatic, model-agnostic approximate inference methods. Examples include languages such as Stan and Edward.

In plain English, recent developments in computational methods for Bayesian inference have made the power of fully Bayesian machine learning accessible beyond expert statisticians.

> ”Probabilistic programming could do for Bayesian ML what Theano has done for neural networks.”

5 Personalized Advice

- If you aren't confident with statistics or programming ...

  Read the glossary for definitions on terms we use in the lecture notes. Attend all the classes, talk to the instructors one-on-one to receive individualized help, and do all the exercises!

- If you are already comfortable with Bayesian statistics ...

You can probably skip all the lectures. You're probably just looking to learn Stan, so skim through the LaTeX material until you hit spots that specifically discuss Stan. We do recommend going in order and doing all the exercises to get a good grasp on Stan. If you have questions or would like some assistance, feel free to come towards the end of lectures for one-on-one help.

- If you aren't comfortable with coding in Python or R, but you have some programming experience ...

  I'm of the mindset that coding syntax can largely be solved by breaking the exercise down into tiny subtasks and googling "how to do this very specific subtask in this language".

- If you have a dataset in mind for your research ...

  The instructors are happy to consider your goals and discuss how probabilistic programming can, or cannot, help you!

- If you wish to revisit material presented in this class at a slower pace from a different perspective, focused on applications...

  Follow this class with Statistical Rethinking by McElreath (2015).

- If you wish to extend material presented in this class and dive deeper into theory...

  Follow this class with Bayesian Data Analysis 3 by Andrew Gelman.

6 Schedule & Topics

Two textbooks in particular have served as useful references in developing this class, and much of our material represents a "remixing" of content originally from Statistical Rethinking by McElreath (2015) and Bayesian Data Analysis 3 by Gelman et al. (2013).

Section 1: Max Shen

- A high-level introduction - what is probabilistic programming?
- Introduction to Stan: Hierarchical Models and Regression
- Identifiability
- Choosing Priors: Regularization, Sparsity, Conjugacy

Section 2: Max Shen

- Inference Methods: MCMC, HMC, NUTS (Alvin Shi)
- Model Checking: Sensitivity Analysis, Posterior Predictive Checks
- Continuous Model Expansion
- Information Criteria: AIC, DIC, WAIC, Cross-Validation
- Effective Parameters

- Robust Models

Section 3: Max Shen

- Exchangeability
- Generalized Linear Models
- Identifiability Part 2: Correlations
- Normal Approximations to the Posterior

Section 4: Carles Boix / Alvin Shi

* Subject to change

- Mixture Models
- Nonparametric Bayes: Gaussian Processes, Dirichlet Processes

Section 5: Max Shen

* Subject to change

- Bayesian Variable Selection
- Non-Linear Models
- Missing Data Imputation

Section 6: Carles Boix / Alvin Shi

* Subject to change

- Variational Inference
- Automatic Differentiation
- Bayesian Deep-Learning
- Edward

7 Contact

Max Shen

maxwshen@mit.edu