Welcome to ISTA 410/510

Inaugural offering of Bayesian Modeling and Statistics

Today

- Course mechanics, syllabus, etc.
- Brief course outline
- Introduce the topic

Course mechanics

Course page is at: http://www.sista.arizona.edu/classes/ista410/spring11
(Linked from instructor's home page http://kobus.ca

Lectures and assignments will require either connecting from a UA machine, OR a login id ("me") and password ("bayes4fun").

Significant communication for the course will happen using the class mail list (ista410@listserve.arizona.edu).

Office hours: Request an appointment during open times as described at:
http://kobus.ca/calendar_info.html.

Monday: 4:00 - 5:00.
Wednesday: 12:00 - 1:00.
Wednesday: 4:00 - 5:00.
Friday: 4:15 - 6:00.

Course mechanics (2)

TA: Kyle Simek

Office hours will be on at least one of Tuesday/Thursday (poll)

Office hours will be in GS 927-C

Lecture notes

PDFs for lectures will be posted on the class web page after class.

Videos for lectures will also be linked.

Course mechanics (3)

This course requires a SISTA/CS account. Apparently (new for 2010), if you have a UA email and are registered, an account is automatically created for you.

The 9th floor lab will be available for this course.

Assignments will be posted on the web page

The course will have both written and programming assignments

Programming language will be Matlab (at least to start)

Assignment zero will be posted soon!

Course mechanics (4)

Books and materials
- No required text (all material will be lecture notes and assignments)
- Good reference (too extensive to be our text) is Koller and Friedman
- Another good reference is Bishop

Co-convened course
- Grad students will have longer assignments
- Grad students will be expected to do more/better on exams

Grade distribution
- Assignments: 60% (there will be 6-10 assignments)
- Midterms: 20% (there will be two midterms).
- Final Exam: 20%

Additional policies and procedures available in syllabus linked from class page

Course outline

Blurb: To develop a solid fundamental understanding of Bayesian methods and how to apply them to diverse problems. Skills developed will include: 1) creating graphical models for data; 2) specifying distributions for parameters of model components that link the model to data; 3) applying inference methods to estimate model parameters; 4) setting up learning model structure from data; and 5) applying Bayesian methods to decision making processes.

Topics: Probabilistic foundations
Introduction to the Bayesian methodology and introductory examples
Representing models using graphs
Inference for graphical models
Learning model structure
Actions and decisions

What is this course about?

Recommended reading

http://mitpress.mit.edu/books/chapters/0262013193chap1.pdf

Introductory Questions

What is a model?

What is a statistical model?

What is a parameterized model?

Why do we like probabilistic / statistical models?

What is a Bayesian model?

What do we mean by inference?

Why does modeling work at all?

Example one:

Video feed from a camera watching the world

Summary points:

There is structure in the world

Real world high dimensional data isn't!

Brute force representation of a high dimensional distribution is a bad idea for two reasons

It is completely impractical

It misses the forest for the trees

More summary

Our models should capture what is important

1) Mechanistically
2) Statistically (simplify the joint density)

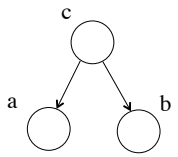
Clustering in high dimensions

$$p(x_i, x_2, \ \ldots \ , x_n) = \sum_c p(c) \prod_i p(x_i \mid c)$$

Generative models

Informally, tells a story about how data comes to be

Illustrated using ancestral sampling



$$p(a,b,c) = p(c)p(a \mid c)p(b \mid c)$$