Welcome to ISTA 410/510

Bayesian Modeling and Statistics

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

Course mechanics

Instructor
- Kobus Barnard
- kobus@sista.arizona.edu (put ISTA 410 or 510 in the subject line!)
- http://kobus.ca
- GS 927A

TA
- Ernesto Brau
- GS 909B (office hours will be in GS 927C)

Course mechanics (II)

Prerequisites (ugrad)
- Some programming (e.g., 130)
- Basic probability and stats (e.g., ISTA 116 or 521)
- Linear algebra (implies calculus)

This course is not especially demanding in terms of programming (alone), or in terms of math (alone), but the combination leads to many drops

Alternatives to consider if you are doubting your preparation (ugrad)
- ISTA 371: Foundations of Information and Inference (will be a prerequisite for this class in the future)
- ISTA 370: Research Methods for the Information Age

Course mechanics (III)

We will make some use of D2L (details to be worked out)

Default course page is at: http://www.sista.arizona.edu/classes/ista410/spring13
(Linked from instructor's home page http://kobus.ca)

Lectures, videos, and assignments posted on the course page may require either connecting from a UA machine, OR a login id ("me") and password ("bayes4fun").

Course mechanics (IV)

For an individual appointment with the instructor, send email, with proposed availability (if possible) during likely open times as described at:
http://kobus.ca/calendar_info.html.

Current list of times

Monday / Wednesday / Friday: 10:00 - 11:00 (10:30 preferred)

Instructor office hour slots must be claimed 18 hours in advance

TA office hours will be Tuesday at 1pm and Thursday at 11am, in GS 927C.

Course mechanics (V)

- Lecture note previews will be posted sporadically for those who want to look at them
  - The longer in the future the material is, the less accurate it will be!
- Official PDFs for lectures will be posted after class.
- Videos for lectures will also be linked.

Course mechanics (VI)

- Assignments will be posted either on the web page or on D2L.
- Assignments will be handed in using D2L
- The course will have both written and programming assignments
  - Key deliverable will typically be a PDF with answers, results, etc.
  - If programming was involved, code needs to be submitted.
  - Recommended programming language is Matlab
  - C/C++ is also an option (library support is available)
  - Others languages can also be used (but I won't look at the code in detail)

Course mechanics (VII)

Books and materials
No required text (all material will be lecture notes and assignments)
Important reference is Bishop (key chapters will be put online)
Good reference (too extensive to be our text) is Koller and Friedman
Third reference is Murphy (very thorough for temporal models)

Co-convened course (roughly 2/1 grad/ugrad)
Grad students will have longer assignments
Grad students will be expected to do more/better on exams

Grade distribution
Assignments: 60% (there will be about 6 assignments)
Midterms: 20% (there will be two midterms, likely take home)
Final Exam: 20% (likely take home)

Participation in experiments
Extra credit, TBA (an alternative will be available)

Additional policies and procedures available in syllabus linked from class page

Course mechanics (VIII)

Tentative schedule of due dates posted on the class web page

First assignment will be posted very soon — due January 23

First assignment will introduce Matlab and “remind” everyone about material that may (or may not) be in prerequisite courses.

Course context

This topic is one of three main topics in modern machine learning

- Discriminative methods (much of ISTA 521, AKA Plan A)
- Generative models (this course, AKA Plan B)
- Reinforcement learning

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

Core Philosophical Content

Going from data to knowledge

Familiar terms that are tricky to define

- Theory
- Representation
- Computational Model

A computer vision example of

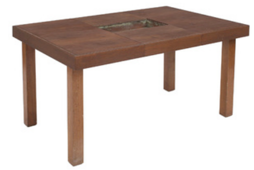
Data ⇨ Knowledge

Visual Representation ⇨ Semantic Representation

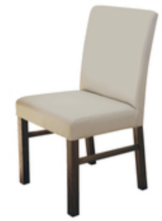
⇨ A tiger lying in the grass

How do you go from data to knowledge?

For example, let's build a system to recognize furniture.



table



chair

other

How do you go from data to knowledge?

Plan A (bottom up)

Study the **data** and the problem and figure out something that might be useful (e.g., extract edges, or color if you think it is salient ...)

Having found (and committed) to features, now aggregate the information (perhaps generically, e.g., stick them into a histogram)

Now give the histogram to a pattern classifier trained to separate the "table" regions of histogram space versus "chairs" and "other"

To find where things are, apply the detector to different parts of the image.

How do you go from data to knowledge?

Plan A (bottom up)

Extract features ⇨ Group features ⇨ Extract relational characterizations ⇨ Train a classifier

or ⇨ Estimate quantity

Note "pipeline" flow of information.

Each step "re-presents" information for the next step, and **restricts** the nature of the representation.

Each step commits to an answer and mistakes are hard to **undo**.

How do you go from data to knowledge?

Problems with plan A

The hard part is representation, but the notion of representation is weak (just labeling)

Every step of the process restricts the representation, typically biased by what is easy to do (existing tools).

Exploiting "high-level" knowledge about the world is hard

We do not know much about what we recognize. We know image location, but not where it is in 3D space or its geometry.

How do you go from data to knowledge?

Plan B (top down)

Think about **objects** and imaging processes

Create models about the world (prior)

Build models that tell you how these models become images (likelihood).

Invert the processes using Bayesian inference

Clean separation of modeling and inference!
Clean separation of modeling the world and how it links to data.

How do you go from data to knowledge?

Solution to Plan B problems

Challenging to implement

Take this course (as a start)

Inference can be computationally expensive

Good programming is important

Use fast computers

Going from data to knowledge (summary)

- Many problems are nicely handled by plan A
- However, complex problems often are **not** well handled by plan A.
- When problems are complex, **and** representation (theory) matters, go with plan B.
  - Your model will be more about the world, and less about measurement processes and data analysis tools.

But what is a "complex" problem

- Attributes that make problems fit into the previous mold (thus by circular reasoning are complex)
  - When hard to model "variance" is property of the world. Examples:
    - Tracking an object with varying illumination (shadows)
    - Tracking an object with occlusions
  - The model is likely to have a number of interacting components, each if which is complex
  - When the number of parameters is unknown and can vary a lot

Recommended reading for Plan B

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

(linked from class page)