6.854
ADVANCED ALGORITHMS
PROBLEM SET 1

JOHN WANG

1. Introduction - Inference

Phenomenon → (x latent variables) → Measurement Mechanism → (y observed variables) → Inference Engine → Inferences about x

- x summarizes aspecdts of interest in phenomenon.
- Can be discrete or continuous.
- Scalar or vector.
- Random or deterministic.

- y observations, available data.
- Discrete or continuous.
- Scalar or vector.
- Generally random.

1.1. Examples.

- Medical diagnosis. Someone has the flu or doesn't, vectors are the results of tests on a patient. Inference decides whether or not a person has the flu.
- Search for extraterrestrials. Make measurements of outerspace and try to infer whether life exists.
- Hedge funds. Infer whether there is value in a particular investment. Use previous history as observations.
- Amazon, Netflix. Predict what customers would like to buy things based on past purchases. What are you going to buy next, so they can put it in front of you.
- Siri. By observing samples of voice waveforms, Siri has to infer what it means and what you're asking for.
- Automatic face recognition. Given images, find faces of people by observing sample pictures.
- Unmanned autonomous vehicle navigation. Predict where the vehicle will be so that you can adjust its position.
- Bioinformatics. Many segments of DNA, must put together the segments to make a complete sequence.

2. Inter-related Issues

- Modelling phenomenon of interest
- Modelling observation mechanism
- Designing the inference engine
- Analyzing performance

Blend of art and science, and often the most straightforward way of modelling the problem is not the best way. Inference is some blending of statistics, information, and computation.

3. Bayesian Inference

- This is the case where x is random. Thus, we can associate with it some prior $p_x(.)$.
- Complete characterization of knowledge of x based on observing $y = y$ is posteriror $p_{x|y}(.|y)$. Posterior is what we know after the observation.

3.1. Bayesian Inference Engine. Bayes Rule:

$$p_{x|y}(x|y) = \frac{p_{y|x}(y|x)p_x(x)}{\sum_x p_{y|x}(y|x')p_x(x')} \tag{1}$$

Sometimes computing probabilities is not enough. Sometimes, you need to make a decision. There are two categories: soft and hard inference. Need to guess (educated guessing).

Landscape is such that Bayesian inference gives us soft decisions (all the possibilities with some weightings). Bayesian decision theory gives hard decisions.

3.2. Example. Big urn full of balls, all of the same size, but of different weights w. Pdf for w is such that mean is 2, median is 3, and mode is 4.

I pull two balls from the urn, and I let you weigh one of them. You have to then pick one of the balls, if you pick the heavier ball, you win $1,000,000$ otherwise you get nothing. You can pick a threshold above which you say its the other ball.

Classic Bayesian theoretic problem. There is some measure of goodness of guess and you need to choose a decision rule that matches the cost criterion. If w ¡ 3, then you should switch, otherwise you should keep the ball.

4. Systematic Approach

We'll develop a systematic approach with a heirarchy.

(1) 2 possible values of x is a detection problem.
(2) m possible values of x is a classification problem.
(3) ∞ possible values of x is an estimation problem.