37458

Advanced Bayesian Methods

Professor Matt Wand

University of Technology Sydney

Major Goal for this Subject

By Week 6, tool you up to be able to do

ANY*

data analysis, no matter how complex,

– and understand the underlying maths.

* To be qualified later.

Path to Achieving Goal

- Draft book chapters of GRAPH THEORY AND STATISTICS by M.P. Wand.
- The R computing environment.
  Starting today!
- The BUGS inference engine and its interfaces with R, known as BRugs (for Windows) and rjags (for Mac OSX).
  Starting today! (later will a newer engine named Stan).
- IMPORTANT: no assumed knowledge of any of the above!

MARGINALISATION

and

CONDITIONAL MARGINALISATION

Connection to Assignment 2

Remember the Bob the DAG questions.

- Question 4 was on MARGINALISATION
- Question 5 was on CONDITIONAL MARGINALISATION

QUESTION: Why is (conditional) marginalisation important?

SOME ANSWERS:

- It is THE mathematical problem that has to be solved for Bayesian statistical inference, which is TOPIC 2 for this subject.
- Many modern Machine Learning algorithms (e.g. speech recognition, Internet searching, robot vision) require marginalisation over big probabilistic graphs.

Beyond Bob

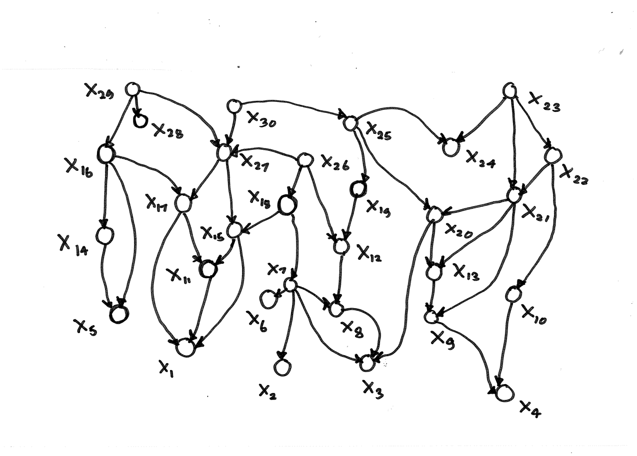
Bob the DAG questions had:

- A two-node DAGs.
- Binary random variables $\Longrightarrow$ simple Bernoulli distributions.

"Real" Statistics problems have DAGs with

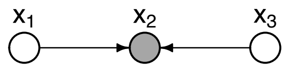
- nodes numbering in dozens, hundreds, thousands....
- more complicated (continuous) distributions.

Real Statistics (and Real Machine Learning)



Project and look at marginalisation section of notes.

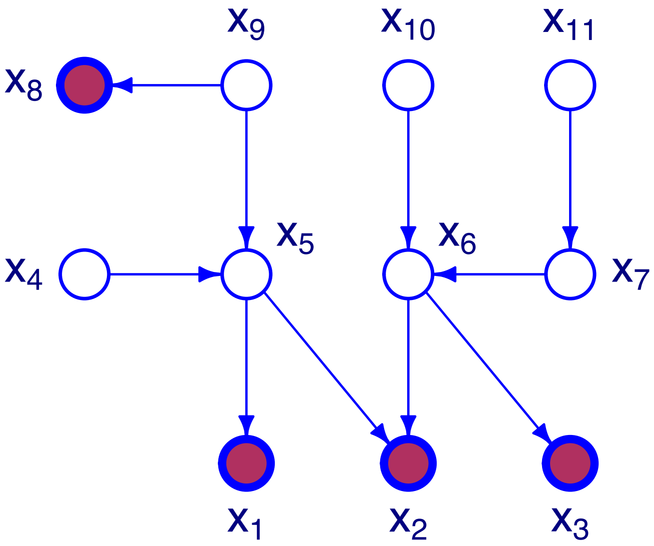
Innocent-Looking Example



$$x_2|x_1, x_3 \sim N(x_1, 1/x_3), \quad x_1 \sim N(0,1), \quad x_3 \sim \text{Gamma}(1,1).$$

BUT EXACT MARGINALISATION NOT POSSIBLE!

Look at mathematics of innocent-looking example in notes.



Example Conditional Marginalisation Problem

$$p(x_6|x_1 = \mathring{x}_1, x_2 = \mathring{x}_2, x_3 = \mathring{x}_3, x_8 = \mathring{x}_8)$$

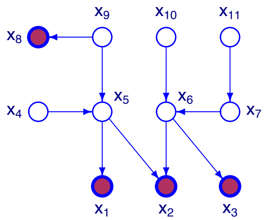
The required maths is:

$$p(x_6|x_1, x_2, x_3, x_8) = \frac{p(x_1, x_2, x_3, x_6, x_8)}{p(x_1, x_2, x_3, x_8)}$$

$$= \frac{\sum_{x_4=0}^{1} \int_0^\infty \sum_{x_7=0}^{1} \int_0^\infty \int_{-\infty}^\infty \int_0^1 p(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11})\, dx_5\, dx_9\, dx_{10}\, dx_{11}}{\sum_{x_4=0}^{1} \int_0^\infty \int_{-\infty}^\infty \sum_{x_7=0}^{1} \int_0^\infty \int_{-\infty}^\infty \int_0^1 p(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11})\, dx_5\, dx_6\, dx_9\, dx_{10}\, dx_{11}}$$

$=$ lots of difficult integration (and some integrals may not be tractable).

Example Distributions on This DAG



$x_1|x_5 \sim$ Poisson$(3\,x_5)$, $x_2|x_5, x_6 \sim N(2x_6 + 5, 9/x_5)$,

$x_3|x_6 \sim N(x_6, 36)$, $x_4 \sim$ Bernoulli(0.37),

$x_5|x_4, x_9 \sim$ Gamma$(x_4 + 3, 4x_9)$, $x_6|x_7, x_{10} \sim N(x_7 x_{10}, 16)$

$x_7|x_{11} \sim$ Bernoulli(x_{11}), $x_8|x_9 \sim$ Poisson(x_9),

$x_9 \sim$ Gamma$(4, 13)$, $x_{10} \sim N(0, 1)$

and $x_{11} \sim$ Beta$(9, 3)$.

MARKOV CHAIN MONTE CARLO

to the rescue!!! ...

Markov Chain Monte Carlo (MCMC) Alternative

This requires that we draw samples from the full conditionals:

$p(x_4|\text{rest of graph})$

$p(x_5|\text{rest of graph})$

$p(x_6|\text{rest of graph})$

$p(x_7|\text{rest of graph})$

$p(x_9|\text{rest of graph})$

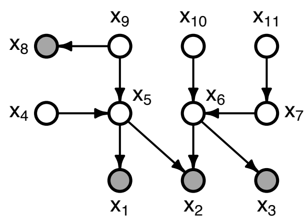
$p(x_{10}|\text{rest of graph})$

$p(x_{11}|\text{rest of graph})$

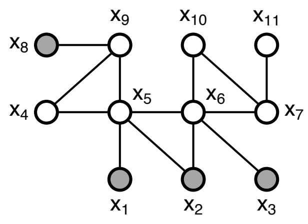
GRAPH THEORY HELPS US

TO SIMPLIFY THESE!

a probabilistic DAG



corresponding moral graph



full conditional density functions

functions of unshaded nodes

$p(x_4 \mid \text{rest}) = p(x_4 \mid x_5, x_9)$

$p(x_5 \mid \text{rest}) = p(x_5 \mid x_1, x_2, x_4, x_6, x_9)$

$p(x_6 \mid \text{rest}) = p(x_6 \mid x_2, x_3, x_5, x_7, x_{10})$

$p(x_7 \mid \text{rest}) = p(x_7 \mid x_6, x_{11})$

$p(x_9 \mid \text{rest}) = p(x_9 \mid x_4, x_5)$

$p(x_{10} \mid \text{rest}) = p(x_{10} \mid x_6)$

$p(x_{11} \mid \text{rest}) = p(x_{11} \mid x_7)$

Run demoDAGandMCMC.Rs ...

The BRugs and rjags Computer Packages

BRugs (Windows) or rjags (Mac OSX)

Allows running of MCMC within the R computing environment.

All we have to do is specify the distributions that make up the DAG.

You get to learn some BRugs/rjags before your next meal! (Laboratory 1 – starting at 11 o'clock)

One last sub-topic...

CONDITIONAL INDEPENDENCE THEOREMS

The Stan Computing Environment

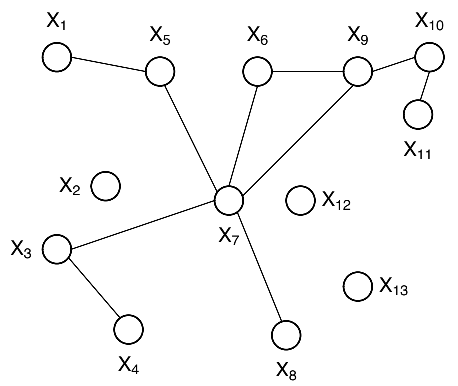
R version: rstan

Since late 2013 also allows running of MCMC within the R computing environment.

Has starting to replace BRugs/rjags as main package of type.

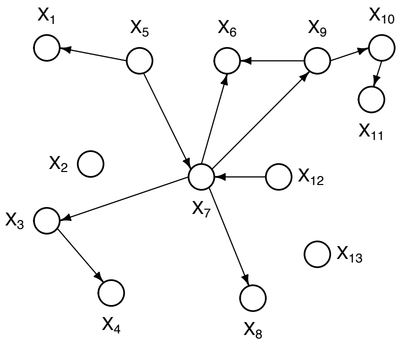
We will use rstan a lot in the second half of this subject.

Class Exercise



Write down two conditional independence statements for this undirected graph.

The DAG Case



PROBLEM: Is $x_3 \perp\!\!\!\perp x_{10} | \{x_7, x_9\}$?

Laboratory 1

Time: 11:00–12:00 Place: Room 011, Level 1, Building 5C.

- Best to use Mozilla Firefox and not Microsoft Explorer.
- For Windows users: safe choice editor to use is WordPad. Don't use NotePad.
- Upper-case versus lower-case.
- Must save files with exactly same as appears in lab.
- We can use any extension name we want on a filename – not just those that Microsoft Inc. chooses! Lecturer likes filename.Rs for "R script".