Announcements

If you are taking this class for credit, and this is your first class with me, please make an appointment (office hours preferable) via email to chat.

Graphical Models

Reference for much of today's material

Chapter 8 of Bishop

Available on-line

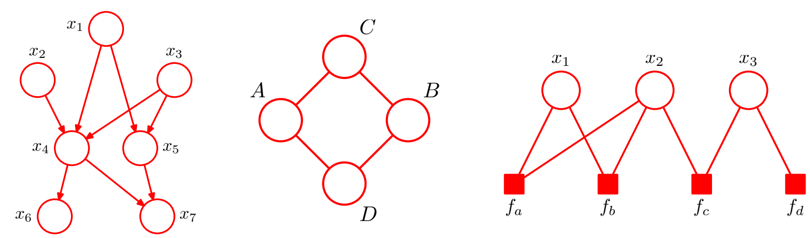
http://research.microsoft.com/~cmbishop/PRML

Graphical Models

- Graphical representation of statistical models
- Nodes
  - Random variables (or groups of them)
- Edges
  - Probabilistic relationships between nodes

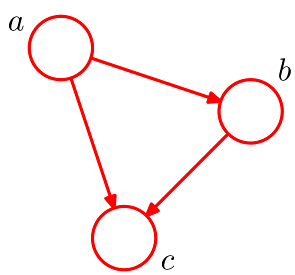
Graphical Models

- Various kinds
  - Directed (Bayesian networks)
  - Undirected (e.g., Markov random field)
  - Factor graphs (different representation)



Directed Graphical Models

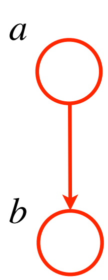
- Nodes represent random variables
- Edges between nodes have directed links
- No cycles



Directed Graphical Models

- Nodes represent random variables
- Edges between nodes have directed links
- No cycles
- The graph represents a factorization of the joint probability of all the random variables represented by the nodes.
- An arrow from one node (a) to another one (b) means that the second node (b) is conditioned on the first (a).

Directed Graphical Models



Arrow points from r.v. "a" to r.v. "b"

So, r.v. "b" is conditioned on "a" in the joint distribution.

In general:

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

In this example:

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

Ancestral sampling story:

To sample from $p(a, b)$

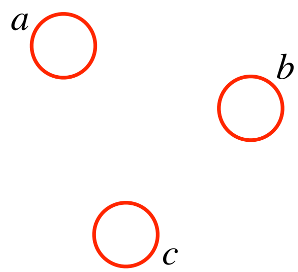
First sample $\tilde{a}$ from $p(a)$

Then sample $\tilde{b}$ from $p(b \mid \overline{a})$

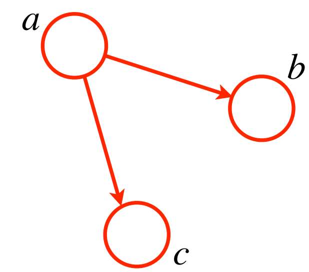
Directed Graphical Models

- A story of three random variables .... a, b, and c.
- General model is p(a,b,c)   (understand this!)
- What are possible relationships of a, b, and c?
  - Independence: p(a,b,c)=p(a)p(b)p(c)
  - Some structure: e.g., p(a,b,c)=p(a)p(b|a)p(c|a)
  - Arbitrary relationship

p(a,b,c) = p(a)p(b)p(c)



p(a,b,c) = p(a)p(b|a)p(c|a)



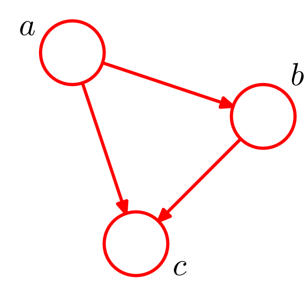
p(a,b,c) with no identified independence

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

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

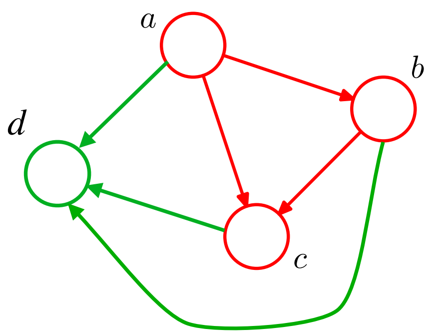
• • •

p(a,b,c) = p(a)p(b|a)p(c|a,b)



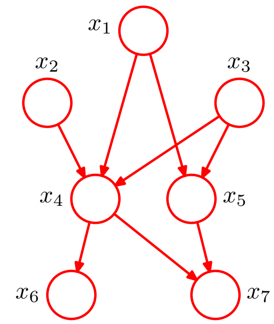
Note that the graph is fully connected

p(a,b,c,d) = p(d | a,b,c) p(a,b,c)



Note that the graph is fully connected

Another example (8.2 in Bishop)



$$p(x_1)p(x_2)p(x_3)p(x_4 \mid x_1,x_2,x_3)p(x_5 \mid x_1,x_3)p(x_6 \mid x_4)p(x_7 \mid x_4,x_5)$$

Univariate Gaussian with known variance

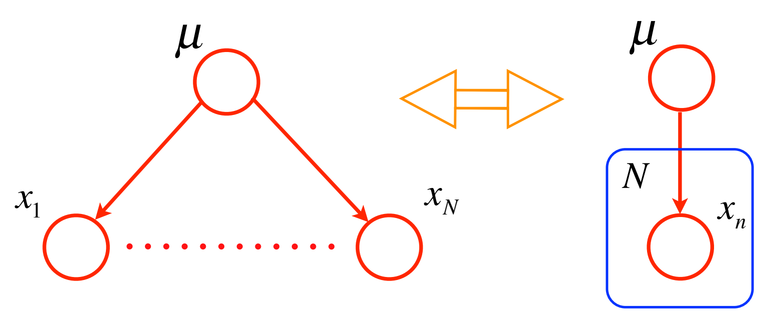
$$D = \{x_1, x_2, x_3, \ldots, x_N\}$$

$$p(D, \mu) = p(\mu)\prod_{n=1}^{N} p(x_n \mid \mu)$$

where

$$p(x_n \mid \mu) = \mathbb{N}(x_n \mid \mu; \sigma^2)$$

Univariate Gaussian with known variance

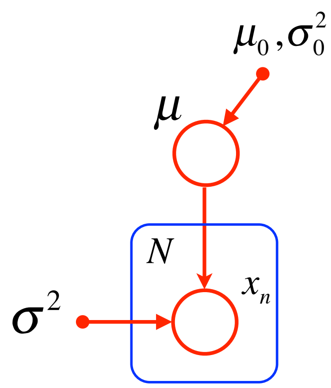


More compact notation
(plate representation)

Deterministic parameters

Our univariate Gaussian has some known parameters: the variance and the prior on the mean.

If we wish to illustrate them, we use a small filled in circle.



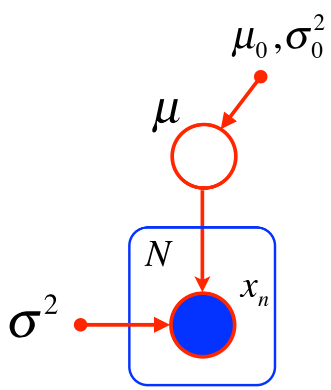
Observed variables

We indicate observed variables by shading them

Alternatively, this indicates conditioning

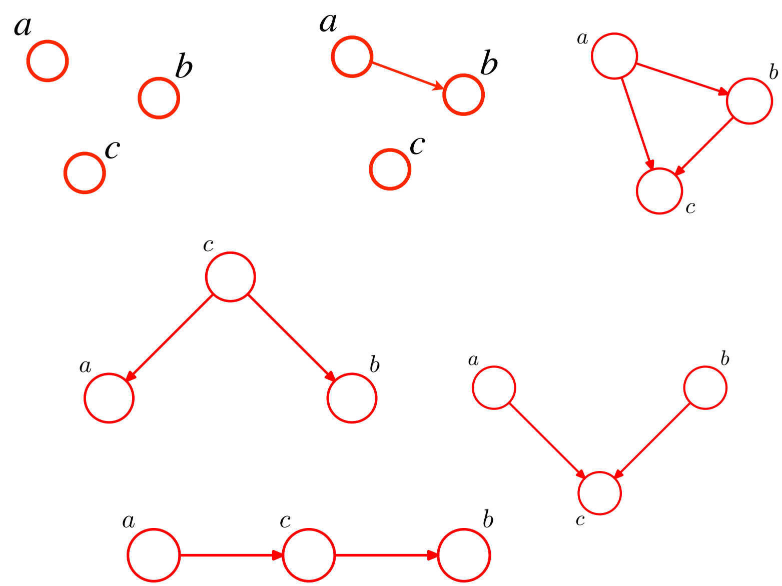
Observed variables

Example: Inferring the mean of the univariate

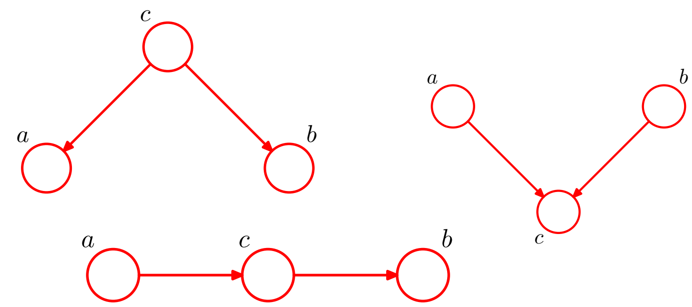


Back to three variables

What are the possible Bayes nets with three variables?



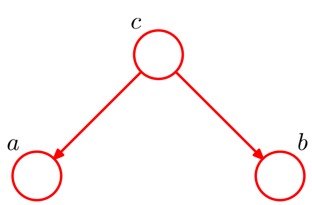
Three interesting cases



For each case, consider two questions:

1) Is $a \perp b$?
2) Is $a \perp b \mid c$? (i.e. c is observed)

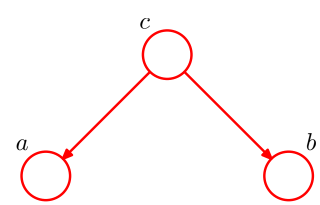
Case one (already used as an example)



$$a \not\perp b$$

If you know a, that informs you about c (by Bayes) which informs you about b.

Case one (already used as an example)

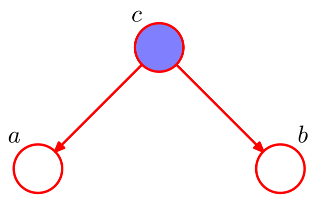


$$a \not\perp b$$

Can prove this using a counter example noting that

if $a \perp b$ then $p(a,b) = p(a)p(b)$ and $p(a|b) = p(a)$

Case one where c is observed



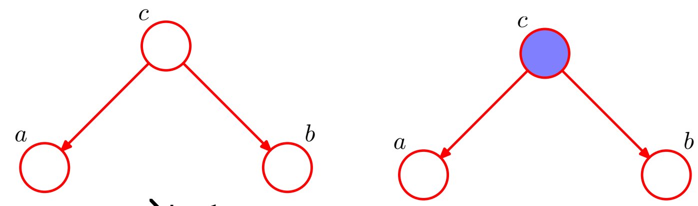
$$a \perp b \mid c$$

$p(a,b,c) = p(a|c)\,p(b|c)$ (what the graph represents)

This is the definition of $a \perp b | c$

Or notice that if c is constant, then $p(a|c) = p(a|\tilde{c}) = p_{\tilde{c}}(a)$

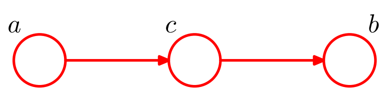
Case one (tail-to-tail) summary



$a \not\perp b$ $\qquad$ $a \perp b \mid c$

Tail-to-tail case
- With no conditioning, no independence
- With conditioning, we have independence

Case two

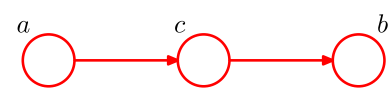


$$a \not\perp b$$

The graph represents $p(a,b,c) = p(a)\,p(c|a)\,p(b|c)$

If you know a, that informs you about c which informs you about b.

Case two



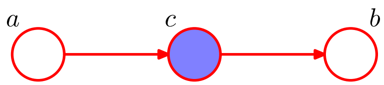
$$a \not\perp b$$

The graph represents $p(a,b,c) = p(a)\,p(c|a)\,p(b|c)$

Algebraically,

$$p(a,b) = \sum_c p(a,b,c) = p(a) \sum_c p(c|a)\,p(b|c) = p(a)\,p(b|a)$$

Case two where c is observed



$$a \perp b \mid c$$

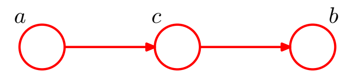
$$p(a,b \mid c) = \frac{p(a,b,c)}{p(c)} \qquad \text{(definition)}$$

$$= \frac{p(a)\,p(c|a)\,p(b|c)}{p(c)} \qquad \text{(from graph)}$$

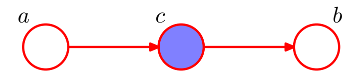
$$= \frac{p(a)\,p(a|c)\,p(c)\,p(b|c)}{p(a)\,p(c)} \qquad \text{(Bayes on } p(c|a)\text{)}$$

$$= p(a|c)\,p(b|c) \qquad \text{(canceling factors)}$$

Case two (head-to-tail) summary



$$a \not\perp b$$



$$a \perp b \mid c$$

Head-to-tail case
- With no conditioning, no independence
- With conditioning, we have independence