Lecture 7: Graphical Models
Machine Learning

Andrew Rosenberg

March 5, 2010

Last Time

- Logistic Regression

Today

- Graphical Models

Recap

Models we've looked at so far.

- Linear Regression
- Logistic Regression

Both make use of probabilistic models.

Graphical models are a way to structure and visualize probability models.

(Joint) Probability Tables.

We represent **multinomial** joint probabilities between K variables as K-dimensional tables.

$$p(x) = p(flu?, achiness?, headache?, \ldots, temperature?)$$

Assume D binary ("true/false") variables.

How big is this table? 2^D

Exponential Increase in size of the probability table.

- Related to the curse of dimensionality.

What if rather than a Bernouli (binary) variables, we had multinomials with M choices?

Probability Models

What if the variables are independent?

$$p(x) = p(\mathit{flu?}, \mathit{achiness?}, \mathit{headache?}, \ldots, \mathit{temperature?})$$

Recall, if x and y are independent:

$$p(x, y) = p(x)p(y)$$

The original probability distribution then **factorizes**.

$$p(x) = p(\mathit{flu?})p(\mathit{achiness?})p(\mathit{headache?}) \ldots p(\mathit{temperature?})$$

How big is this table (if each variable is binary)?

$p(\mathit{flu?}) =$

.2	.8

$p(\mathit{headache?}) =$

.6	.4

etc.

Total size $= 2 * D$

- Independence assumptions are convenient (Naive Bayes), but rarely true.
- More often some groups of variables are dependent, but others are independent.
- Moreover others are **conditionally independent**.

Conditional Independence

If two variables are **conditionally independent**, then:

$$p(x, z|y) = p(x|y)p(z|y)$$

but

$$p(x, z) \neq p(x)p(z)$$

e.g. $y = flu?$, $x = achiness?$, $z = headache?$.
Written as:

$$x \perp\!\!\!\perp z|y$$

Factorization of the joint

Assume

$$x \perp\!\!\!\perp z | y$$

How do you factorize $p(x, y, z)$?

$$\begin{aligned} p(x, y, z) &= p(x, z|y)p(y) \\ &= p(x|y)p(z|y)p(y) \end{aligned}$$

Factorization of the joint

Assume

$$x \perp\!\!\!\perp z | y$$

How do you factorize $p(x, y, z)$?

$$\begin{aligned} p(x, y, z) &= p(x, z|y)p(y) \\ &= p(x|y)p(z|y)p(y) \end{aligned}$$

What if x and z not conditionally independent?

$$\begin{aligned} p(x, y, z) &= p(x, z|y)p(y) \\ &= p(x|y, z)p(z|y)p(y) \end{aligned}$$

Structure of Graphical Models

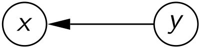
Graphical models allow us to represent dependence relationships between variables visually.

- Graphical models are **Graphs**
- Nodes: random variables
- Edges: Dependence relationship
- No Edge: Independent variables
- Direction of the edge: indicates a parent child relationship (like causality, but not exactly)
- Child: Destination of the edge – Response
- Parent: Source of the edge – Trigger
- Graphical models are **always** Directed Acyclic Graphs (DAG).

Some example models

Independence: $p(x, y) = p(x)p(y)$

Dependence: $p(x, y) = p(x|y)p(y)$



Parents of a node i denoted π_i or pa_i.

Factorization of the joint in a Graphical model

$$p(x_0, \ldots, x_{n-1}) = \prod_{i=0}^{n-1} p(x_i|pa_i) = \prod_{i=0}^{n-1} p(x_i|\pi_i)$$

Basic Graphical Models

Independent variables.

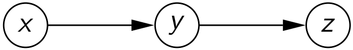
Observations.

- When we **observe** a variable – know it's value from data – we color the variable corresponding to that node grey.
- Observing a variable allows us to condition on it.
  E.g. p(x,z—y)
- Given an observation of any variable we can generate generate pdfs for the other variables.

Example Graphical Models

Basic Graphical models.

Markov Chain

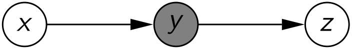


$$p(x, y, z) = \prod_{n \in \{x,y,z\}} (n|\pi_n) = p(x)p(y|x)p(z|y)$$

$x = cloudy?$
$y = raining?$
$z = wetground?$

Basic Graphical models.

Markov Chain

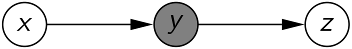


$$p(x, y, z) = \prod_{n \in \{x,y,z\}} (n|\pi_n) = p(x)p(y|x)p(z|y)$$

Is $x \perp\!\!\!\perp z|y$? That is... Does $p(x, z|y) = p(x|y)p(z|y)$?

Basic Graphical models.

Markov Chain



$$p(x, y, z) = \prod_{n \in \{x,y,z\}} (n|\pi_n) = p(x)p(y|x)p(z|y)$$

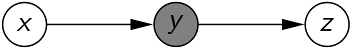
Is $x \perp\!\!\!\perp z|y$? That is... Does $p(x, z|y) = p(x|y)p(z|y)$?

$$p(x, z|y) = p(x|z, y)p(z|y)$$

Example Graphical Models

Basic Graphical models.

Markov Chain



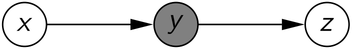
$$p(x,y,z) = \prod_{n \in \{x,y,z\}} (n|\pi_n) = p(x)p(y|x)p(z|y)$$

Is $x \perp\!\!\!\perp z|y$? That is... Does $p(x,z|y) = p(x|y)p(z|y)$?

$$\begin{aligned} p(x,z|y) &= p(x|z,y)p(z|y) \\ p(x|z,y) &= \frac{p(x,y,z)}{p(y,z)} = \frac{p(x)p(y|x)p(z|y)}{p(y)p(z|y)} \end{aligned}$$

Basic Graphical models.

Markov Chain



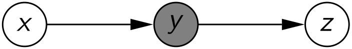
$$p(x,y,z) = \prod_{n \in \{x,y,z\}} (n|\pi_n) = p(x)p(y|x)p(z|y)$$

Is $x \perp\!\!\!\perp z|y$? That is... Does $p(x,z|y) = p(x|y)p(z|y)$?

$$\begin{aligned} p(x,z|y) &= p(x|z,y)p(z|y) \\ p(x|z,y) &= \frac{p(x,y,z)}{p(y,z)} = \frac{p(x)p(y|x)p(z|y)}{p(y)p(z|y)} \\ &= \frac{p(x)p(y|x)}{p(y)} = \frac{p(x,y)}{p(y)} = p(x|y) \end{aligned}$$

Basic Graphical models.

Markov Chain



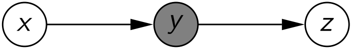
$$p(x,y,z) = \prod_{n \in \{x,y,z\}} (n|\pi_n) = p(x)p(y|x)p(z|y)$$

Is $x \perp\!\!\!\perp z|y$? That is... Does $p(x,z|y) = p(x|y)p(z|y)$?

$$\begin{aligned}
p(x,z|y) &= p(x|z,y)p(z|y) \\
p(x|z,y) &= \frac{p(x,y,z)}{p(y,z)} = \frac{p(x)p(y|x)p(z|y)}{p(y)p(z|y)} \\
&= \frac{p(x)p(y|x)}{p(y)} = \frac{p(x,y)}{p(y)} = p(x|y) \\
p(x,z|y) &= p(x|y)p(z|y)
\end{aligned}$$

Basic Graphical models.

Markov Chain

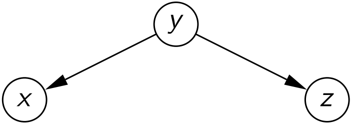


$$p(x,y,z) = \prod_{n \in \{x,y,z\}} (n|\pi_n) = p(x)p(y|x)p(z|y)$$

Is $x \perp\!\!\!\perp z|y$? That is... Does $p(x,z|y) = p(x|y)p(z|y)$?

$$\begin{aligned}
p(x,z|y) &= p(x|z,y)p(z|y) \\
p(x|z,y) &= \frac{p(x,y,z)}{p(y,z)} = \frac{p(x)p(y|x)p(z|y)}{p(y)p(z|y)} \\
&= \frac{p(x)p(y|x)}{p(y)} = \frac{p(x,y)}{p(y)} = p(x|y) \\
p(x,z|y) &= p(x|y)p(z|y) \\
x &\perp\!\!\!\perp z|y
\end{aligned}$$

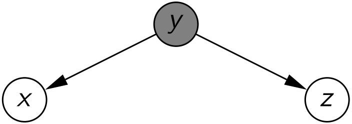
One cause two effects



$$p(x, y, z) = \prod_{n \in \{x,y,z\}} (n|\pi_n) = p(x|y)p(y)p(z|y)$$

$x = achiness?$
$y = flu?$
$z = fever?$

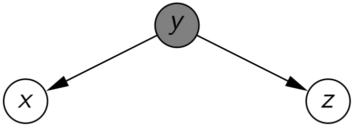
One cause two effects



$$p(x, y, z) = \prod_{n \in \{x,y,z\}} (n|\pi_n) = p(x|y)p(y)p(z|y)$$

Is $x \perp\!\!\!\perp z|y$? That is... Does $p(x, z|y) = p(x|y)p(z|y)$?

One cause two effects

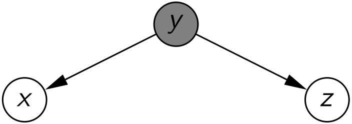


$$p(x,y,z) = \prod_{n \in \{x,y,z\}} (n|\pi_n) = p(x|y)p(y)p(z|y)$$

Is $x \perp\!\!\!\perp z|y$? That is... Does $p(x,z|y) = p(x|y)p(z|y)$?

$$p(x,z|y) \;\; = \;\; p(x|z,y)p(z|y)$$

One cause two effects

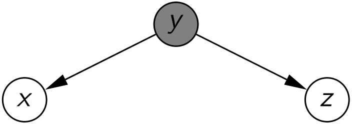


$$p(x, y, z) = \prod_{n \in \{x,y,z\}} (n|\pi_n) = p(x|y)p(y)p(z|y)$$

Is $x \perp\!\!\!\perp z|y$? That is... Does $p(x, z|y) = p(x|y)p(z|y)$?

$$\begin{aligned} p(x, z|y) &= p(x|z, y)p(z|y) \\ p(x|z, y) &= \frac{p(x, y, z)}{p(y, z)} = \frac{p(x|y)p(y)p(z|y)}{p(y)p(z|y)} \end{aligned}$$

One cause two effects

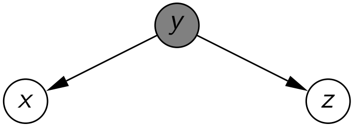


$$p(x,y,z) = \prod_{n\in\{x,y,z\}} (n|\pi_n) = p(x|y)p(y)p(z|y)$$

Is $x \perp\!\!\!\perp z|y$? That is... Does $p(x,z|y) = p(x|y)p(z|y)$?

$$\begin{aligned} p(x,z|y) &= p(x|z,y)p(z|y) \\ p(x|z,y) &= \frac{p(x,y,z)}{p(y,z)} = \frac{p(x|y)p(y)p(z|y)}{p(y)p(z|y)} \\ &= p(x|y) \end{aligned}$$

One cause two effects

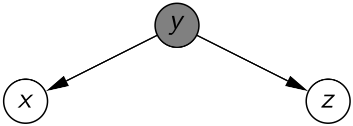


$$p(x, y, z) = \prod_{n \in \{x,y,z\}} (n|\pi_n) = p(x|y)p(y)p(z|y)$$

Is $x \perp\!\!\!\perp z|y$? That is... Does $p(x, z|y) = p(x|y)p(z|y)$?

$$\begin{aligned} p(x, z|y) &= p(x|z, y)p(z|y) \\ p(x|z, y) &= \frac{p(x, y, z)}{p(y, z)} = \frac{p(x|y)p(y)p(z|y)}{p(y)p(z|y)} \\ &= p(x|y) \\ p(x, z|y) &= p(x|y)p(z|y) \end{aligned}$$

One cause two effects



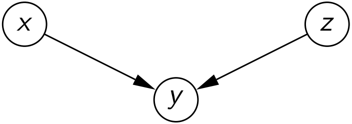
$$p(x, y, z) = \prod_{n \in \{x,y,z\}} (n|\pi_n) = p(x|y)p(y)p(z|y)$$

Is $x \perp\!\!\!\perp z|y$? That is... Does $p(x, z|y) = p(x|y)p(z|y)$?

$$\begin{aligned}
p(x, z|y) &= p(x|z, y)p(z|y) \\
p(x|z, y) &= \frac{p(x, y, z)}{p(y, z)} = \frac{p(x|y)p(y)p(z|y)}{p(y)p(z|y)} \\
&= p(x|y) \\
p(x, z|y) &= p(x|y)p(z|y)
\end{aligned}$$

$x \perp\!\!\!\perp z|y$

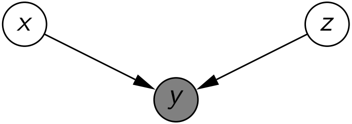
Two causes One effect



$$p(x, y, z) = \prod_{n \in \{x,y,z\}} (n|\pi_n) = p(x)p(y|x, z)p(z)$$

$x = rain?$
$y = wetsidewalk?$
$z = spilledcoffee?$

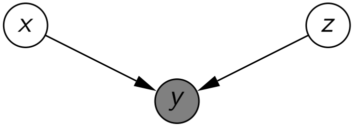
Two causes One effect



$$p(x, y, z) = \prod_{n \in \{x,y,z\}} (n|\pi_n) = p(x)p(y|x, z)p(z)$$

Is $x \perp\!\!\!\perp z|y$? That is... Does $p(x, z|y) = p(x|y)p(z|y)$?

Two causes One effect

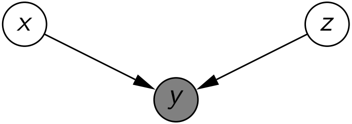


$$p(x,y,z) = \prod_{n \in \{x,y,z\}} (n|\pi_n) = p(x)p(y|x,z)p(z)$$

Is $x \perp\!\!\!\perp z|y$? That is... Does $p(x,z|y) = p(x|y)p(z|y)$?

$$p(x,z|y) \quad = \quad p(x|z,y)p(z|y)$$

Two causes One effect

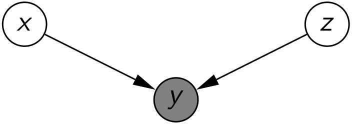


$$p(x, y, z) = \prod_{n \in \{x,y,z\}} (n|\pi_n) = p(x)p(y|x, z)p(z)$$

Is $x \perp\!\!\!\perp z|y$? That is... Does $p(x, z|y) = p(x|y)p(z|y)$?

$$\begin{aligned} p(x, z|y) &= p(x|z, y)p(z|y) \\ p(x|z, y) &= \frac{p(x, y, z)}{p(y, z)} = \frac{p(x)p(y|x, z)p(z)}{p(y|z)p(z)} \end{aligned}$$

Two causes One effect

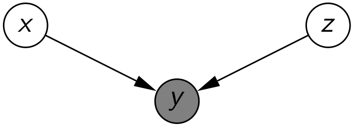


$$p(x,y,z) = \prod_{n\in\{x,y,z\}} (n|\pi_n) = p(x)p(y|x,z)p(z)$$

Is $x \perp\!\!\!\perp z|y$? That is... Does $p(x,z|y) = p(x|y)p(z|y)$?

$$\begin{aligned} p(x,z|y) &= p(x|z,y)p(z|y) \\ p(x|z,y) &= \frac{p(x,y,z)}{p(y,z)} = \frac{p(x)p(y|x,z)p(z)}{p(y|z)p(z)} \\ &= = \frac{p(x)p(y|x,z)}{p(y|z)} \end{aligned}$$

Two causes One effect



$$p(x,y,z) = \prod_{n \in \{x,y,z\}} (n|\pi_n) = p(x)p(y|x,z)p(z)$$

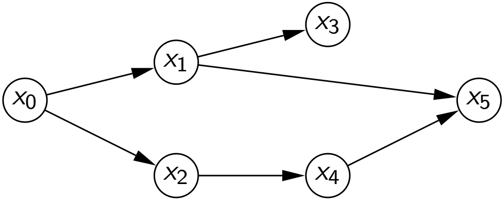
Is $x \perp\!\!\!\perp z|y$? That is... Does $p(x,z|y) = p(x|y)p(z|y)$?

$$\begin{aligned} p(x,z|y) &= p(x|z,y)p(z|y) \\ p(x|z,y) &= \frac{p(x,y,z)}{p(y,z)} = \frac{p(x)p(y|x,z)p(z)}{p(y|z)p(z)} \\ &= = \frac{p(x)p(y|x,z)}{p(y|z)} \end{aligned}$$

x **not** $\perp\!\!\!\perp z|y$

Factorization

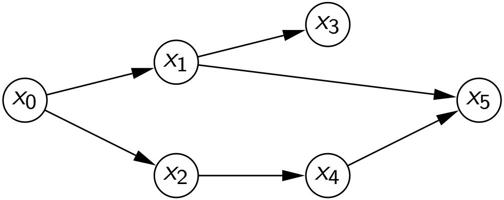
A more complicated factorization



$$p(x_0, x_1, x_2, x_3, x_4, x_5) \quad = \quad ?$$

Factorization

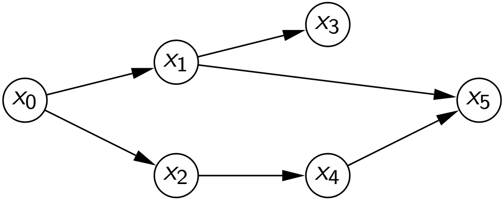
A more complicated factorization



$$
\begin{aligned}
p(x_0, x_1, x_2, x_3, x_4, x_5) &= ? \\
&= p(x_0) \ldots
\end{aligned}
$$

Factorization

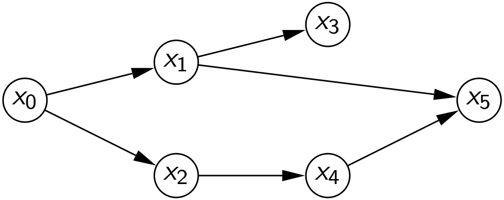
A more complicated factorization



$$
\begin{aligned}
p(x_0, x_1, x_2, x_3, x_4, x_5) &= ? \\
&= p(x_0)\ldots \\
&= p(x_0)p(x_1|x_0)\ldots
\end{aligned}
$$

Factorization

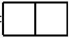
A more complicated factorization

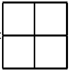


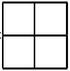
$$
\begin{aligned}
p(x_0, x_1, x_2, x_3, x_4, x_5) &= ? \\
&= p(x_0) \ldots \\
&= p(x_0)p(x_1|x_0) \ldots \\
&= p(x_0)p(x_1|x_0)p(x_2|x_0)p(x_3|x_1)p(x_4|x_2)p(x_5|x_1, x_4)
\end{aligned}
$$

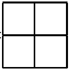
How big are the probability tables?

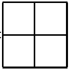
$$p(x_0, x_1, x_2, x_3, x_4, x_5) = p(x_0)p(x_1|x_0)p(x_2|x_0)p(x_3|x_1)p(x_4|x_2)p(x_5|x_1, x_4)$$

$p(x_0) =$ 

$p(x_1|x_0) =$ 

$p(x_2|x_0) =$ 

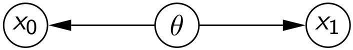
$p(x_3|x_1) =$ 

$p(x_4|x_2) =$ 

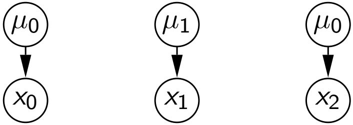
$p(x_5|x_1, x_4) =$ 

Model Parameters as nodes

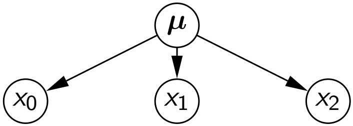
If we model the parameters, θ, as a random variable, we can include these in the graphical model.



Multivariate Bernouli

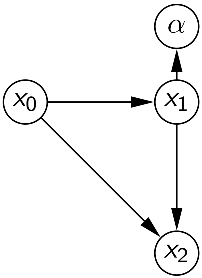


Multinomial



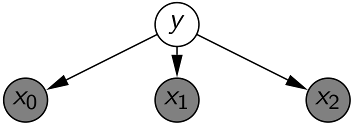
Continuous models

Graphical models can incorporate both discrete and continuous nodes.



Naive Bayes

Naive Bayes Classification.



- Observation variables, x_i are each independent given the class y.
- A distribution is optimized using maximum likelihood for each variable separately.
- Can easily combine multinomial, bernouli and continuous (e.g. Gaussian) distributions from the variables.

$$p(y|x_0 x_1, x_2) \propto p(x_0, x_1, x_2|y)p(y)$$
$$p(y|x_0 x_1, x_2) \propto p(x_0|y)p(x_1|y)p(x_2|y)p(y)$$

Graphical Models

- Graph representation of dependency relationship
- Directed Acyclic Graph (DAG)
- Nodes are random variables
- Edges define dependence relationships.

What can we do with Graphical models

- Learn Parameters – to fit data
- Understand the independence relationships between variables
- Perform inference (marginals and conditionals)
- Compute Likelihoods for classification

Bye

- Next
  - More fun with Graphical Models