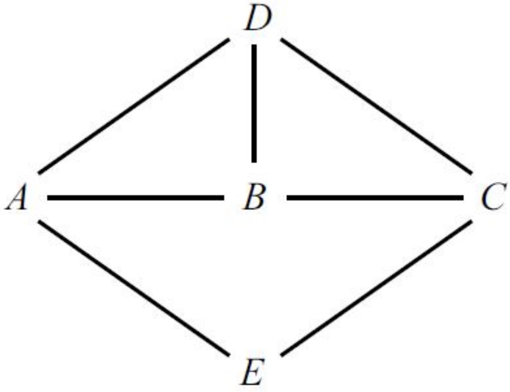
CRG 17/18 Meeting 2:
Chapter 2: Preliminary things
Spirtes Glymour & Scheines (2000)

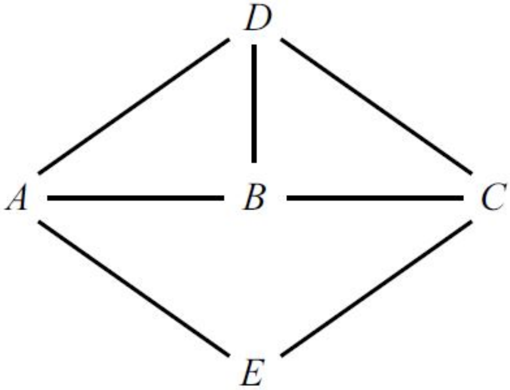
Discussant: Oisín Ryan

October 27, 2017

Name that graph!

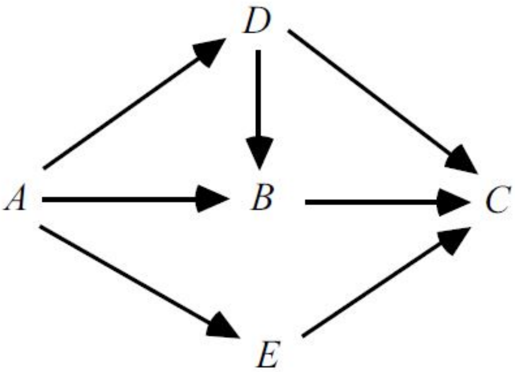


Name that graph!

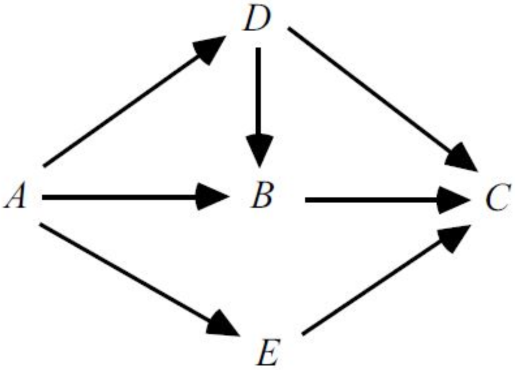


Undirected Graph

Name that graph!

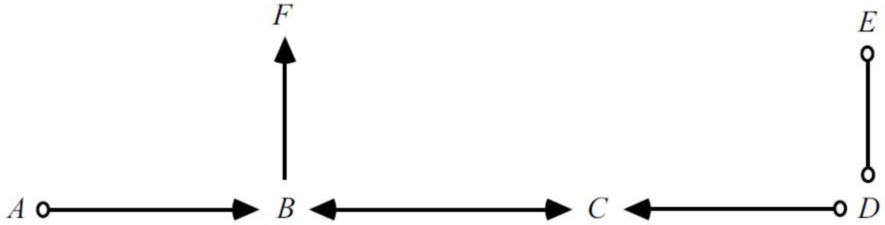


Name that graph!

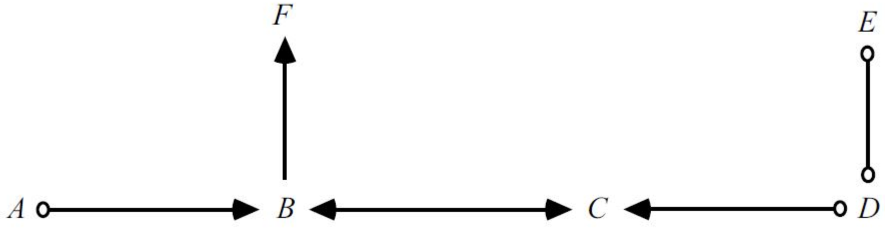


Directed Graph

Name that graph!



Name that graph!



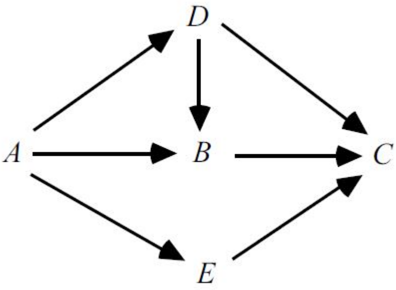
Partially oriented inducing path graph

Graphs

Graphs (G) are made up of

- Vertices (**V**)
  - Variables/nodes
- Marks (**M**)
  - Types of edges
- Edges (**E**)
  - Which vertices are linked and how

2.1: Graphs



- $\mathbf{V} = \{A, B, C, D, E\}$
- $\mathbf{M} = \{EM, >\}$
- $\mathbf{E} = [A, EM], [B, >]...$
- Parents(B) $= \{A, D\}$
- Children(B) $= \{C\}$
- A is an **ancestor** of C
- Undirected path from B to E $=$ $B \rightarrow C \leftarrow E$
- C is a collider on that path

2.1: Graphs

Graph

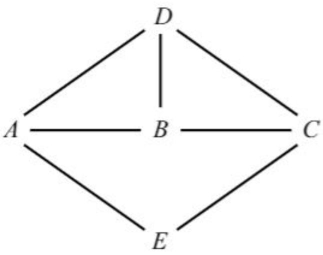


Figure 2.1

Sub-graph

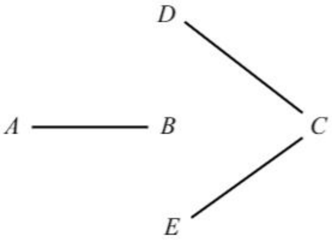


Figure 2.7

2.2: Probability

- Independence
- Joint independence - all pairs in a set are independent
- Conditional independence
- Marginal probability

2:3 Graphs and Probability Distributions

Markov Condition: A directed acyclic graph G over $\mathbf{V}$ and a probability distribution $P(\mathbf{V})$ satisfy the Markov condition if and only if for every W in $\mathbf{V}$, W is independent of $\mathbf{V}\backslash(\mathbf{Descendants}(W) \cup \mathbf{Parents}(W))$ given $\mathbf{Parents}(W)$.

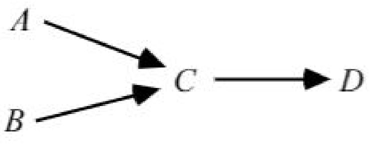


Figure 2.8

- **Descendants**(D)= D
- **Parents**(D) = C
- **V** $\backslash$**Descendants**(D) U **Parents**(D) = $\{A, B\}$
- D is independent of A and B, conditional on C

2:3 Graphs and Probability Distributions

For all values of **v** of **V** for which $f(\mathbf{v}) \neq 0$, the joint density function $f(\mathbf{V})$ satisfying the Markov Condition is given by

$$f(\mathbf{V}) = \prod_{V \in \mathbf{V}} f(V|\mathbf{Parents}(V))$$

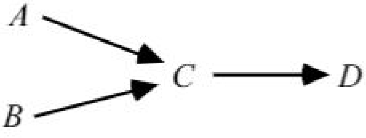


Figure 2.8

$$P(A, B, C, D) =$$
$$P(A)P(B)P(C|A, B)P(D|C)$$

Minimality

- If G is a directed acyclic graph over $\mathbf{V}$ and P a probability distribution over $\mathbf{V}$, $< G, P >$ satisfies the Minimality Condition if and only if for every proper subgraph H of G with vertex set $\mathbf{V}$, $< H, P >$ does not satisfy the Markov condition
- **Discussant interpretation:** Minimality **and** the Markov condition are met if all of the edges in G are necessary to fully describe the dependency structure in the joint density. I cannot build a simplified version of the graph which also satisfies the markov condition

Progress Meeting 2

Material discussed:

- Chapter 2 p. 5-12 (sections 2.1 to 2.3.1)

Readings of this week to discuss next week

- Chapter 2, sections 2.3.2 to 2.6 (inclusive)

Readings for next week

- Chapter 3, up to and including section 3.3 (p.19-29)

Group discussion: Unresolved questions

1. **Subgraphs** (p. 9-10)
   - A subgraph of $<$ **V, M, E** $>$ is any graph $<$ **V', M', E'** $>$ such that **V'** is included in **V**, **M'** is included in **M** and **E'** is included in **E**
   - Can we define an undirected version of a directed graph as a subgraph? I.e. is Fig 2.1 a subgraph of Fig2.2? Can we change only the set of marks **M** to obtain a subgraph?
   - Argument for: the definition of the subgraph includes that the set of marks in the subgraph must be a subset of the marks included in the (super)graph. We see examples of subgraphs in which only the vertices or edges are changed. Changing only the marks retains all of the *adjacencies* in the graph, even if the specific edges contain less information.
   - Argument against: **E** is defined as combinations of elements of **M** and **V**. For example take the graph G=$<$ **V, M, E** $>$, where $\mathbf{V} = \{A, B\}, \mathbf{M} = \{EM, >\}, \mathbf{E} = \{[A, EM], [B, >]\}$. We can also define an undirected version of the above graph, given as U=$<$ **V*, M*, E*** $>$, where $\mathbf{V*} = \{A, B\}, \mathbf{M*} = \{EM\}, \mathbf{E*} = \{[A, EM], [B, EM]\}$. Strictly speaking, **V*** is included in **V**, **M*** is included in **M**. However **E*** is not included in **E**. Therefor U is not a subgraph of G

Group discussion: Unresolved questions

2 **Minimality** (pg 12)

- The text suggests that it is possible to meet the Markov condition without meeting the minimality condition.
- We can see this if we add an edge between A and D in Figure 2.8. Then the markov condition is met - each variable is still independent of its non-descendants and non-parents, given its parents. (A is now a parent of D, and D is still independent of B given given A and C).
- Is it possible to satisfy the minimality condition, while violating the Markov condition?
- Trivially we can define a graph G with no edges, which does not meet the markov condition (assuming some dependencies between variables). No subgraph of this graph will satisfy the Marko condition either. Therefore, we have minimality without the markov condition.
- Discussant interpretation on minimality slide updated accordingly

Group discussion: Marginal probability

The definition of the marginal on pg 11 is a little difficult to penetrate.

Erik-Jan offered the following example in the case of **two binary variables**

First, recall the definition of the special summation sign (notations chapter, p. xx - xxi). The sign means we have to take the sum over the set of **values of the random variables which appear under the sign.**

$$\overset{->}{\sum_{\mathbf{V'}}} P(\mathbf{V'}) = P(V_1 = 0, V_2 = 0) + P(V_1 = 0, V_2 = 1) + P(V_1 = 1, V_2 = 0) + P(V_1 = 0, V_2 = 0)$$

	$V_2 = 0$	$V_2 = 1$	Marginal $P(V_1)$
$V_1 = 0$	.3	.4	.7
$V_1 = 1$	.2	.1	.3
Marginal $P(V_2)$	.5	.5	

Let $\mathbf{V'} = \{V_1, V_2\}$

The **joint probability distribution** $P(\mathbf{V'})$ is given by the table to the left

Group discussion: Marginal probability

	$V_2 = 0$	$V_2 = 1$	Marginal $P(V_1)$
$V_1 = 0$	.3	.4	.7
$V_1 = 1$	.2	.1	.3
Marginal $P(V_2)$	.5	.5	

Let $\mathbf{V'} = \{V_1, V_2\}$
Let $\mathbf{V} = \{V_1\}$

Pg. 11 defines the **marginal** of $P(\mathbf{V'})$ over $\mathbf{V}$ as $P(\mathbf{V})$, in the case where $\mathbf{V}$ is a subset of $\mathbf{V'}$, and where

$$P(\mathbf{V}) = \overset{->}{\sum_{\mathbf{V'} \setminus \mathbf{V}}} P(\mathbf{V'})$$

In our example, to take the **marginal** of $P(\mathbf{V'})$ over V_1, that is, to find the Marginal Probabilty $P(V_1)$, we take the sum over $P(\mathbf{V'})$ of the values of $\mathbf{V'} \setminus \mathbf{V} = \{V_2\}$.

$$P(V_1) = \overset{->}{\sum_{V_2}} P(\mathbf{V'}) = P(V_1, V_2 = 0) + P(V_1, V_2 = 1)$$

The marginal probabilities for values of V_1 can be read off the table above, $P(V_1 = 0) = .7$ and $P(V_1 = 1) = .3$ respectively