36-720: Graphical Models

Brian Junker

September 19, 2007

- Collapsibility and Simpson's Paradox
- Decomposability
- Examples

One nice feature of graphical models is that they lead immediately to two imporant ideas in simplifying log-linear models. They are

- Decomposable models
- Collapsible models

Decomposable models are graphical models for which closed-form MLE's exist.

Collapsibility is closely related to (the non-occurrence of) Simpson's paradox like phenomena.

Collapsibility and Simpson's Paradox

Consider a hierarchical, graphical log-linear model $\mathcal{M}$ and partition the factors in $\mathcal{M}$ into three disjoint groups, A, B, and C. *Examples:*

- $\mathcal{M} = [12][23]$, $A = \{1\}$, $B = \{3\}$, $C = \{2\}$
- $\mathcal{M} = [123][145]$, $A = \{2, 3\}$, $B = \{4, 5\}$, $C = \{1\}$
- $\mathcal{M} = [126][234][46][56]$, $A = \{1, 2, 6\}$, $B = \{3, 5\}$, $C = \{4\}$

Often we are interested in the relationship between A and C

- Conditioning on B, vs.
- Summing (collapsing) over B.

Simpson's paradox happens when

$$P(A|C, B) \neq P(A|C)$$

[e.g. if A and C are single factors, $OR(A, C) \neq OR(A, C|B)$.]

We will say "the conditional relationship $A|C$ is collapsible over B" if

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

Clearly this can happen if and only if

$$B \perp\!\!\!\perp A \mid C$$

Since $\mathcal{M}$ is a graphical model, the *global Markov property* tells us this can happen iff

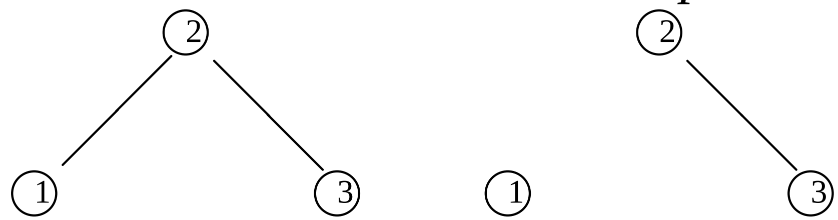
C separates $\mathcal{M}$ into disjoint subgraphs containing A and B.

and when we look at the graph we see this happens iff

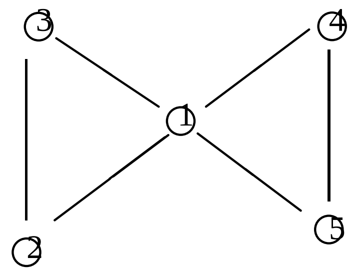
Every path from a variable in A to a variable
in B goes through at least one variable in C.

Thus we have a simple criterion for the collapsibility of relationships between groups of variables, over a "stratification" variable. *Examples:*

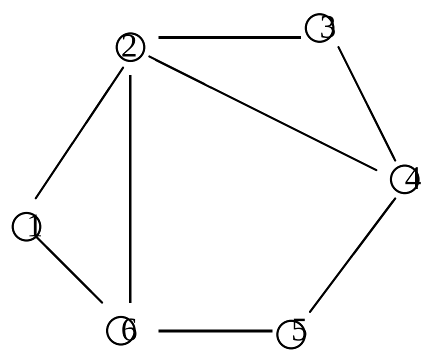
$\mathcal{M} = [12][23]$, $A = \{1\}$, $B = \{3\}$, $C = \{2\}$



$\mathcal{M} = [123][145]$, $A = \{2, 3\}$, $B = \{4, 5\}$, $C = \{1\}$



$\mathcal{M} = [126][234][45][56]$, $A = \{1, 2, 6\}$, $B = \{3, 5\}$, $C = \{4\}$



An Equivalent Form of Collapsibility

Sometimes we are not so much interested in the relationship between variables in A and those in C, but rather in *all* the relationships in A, afer collapsing over B.

Let A be a subset of variables in $\mathcal{M}$ and let $B = \mathcal{M} \setminus A$.

We say $\mathcal{M}$ *is collapsible* **onto** A iff the models

- $P(A)$
- $P(A|B)$

have the same graph [namely, the graph of $\mathcal{M}$ with the variables in B deleted]. I.e., $P(A)$ has exactly the same conditional independence relationships as $P(A|B)$.

The two forms of collapsibility are equivalent:

- Let $A = A' \,\dot\cup\, C'$ [disjoint union]. Clearly, if $\mathcal{M}$ is collapsible onto A, then $P(A'|C') = P(A'|C', B)$ since these are determined by $P(A)$ and $P(A|B)$.
- Conversely if $A'|C'$ is collapsible over B, then the table can be collapsed onto $A = A' \cup B'$. This follows from Theorem 2.3 and Corollary 2.5 of Asmussen & Edwards (1983, *Bmka*), which also establishes the criterion on the next slide.

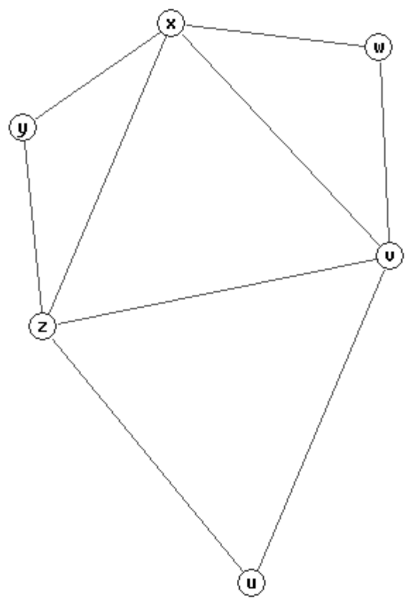
A criterion for collapsibility of $\mathcal{M}$ onto A

Let

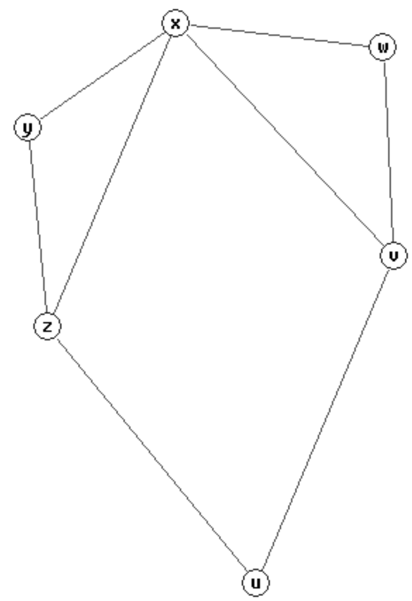
- $A \subseteq \mathcal{M}$, $B = \mathcal{M} \setminus A$
- $B_1, \ldots, B_k$ be the connected components of B
- $E_1, \ldots, E_k$ be their boundaries

Then $\mathcal{M}$ is collapsible onto A iff each E_j is complete.

Example



Model $\mathcal{M}$, $A = \{X, Y, Z, V\}$



$\mathcal{M}$, $A = \{X, Y, Z, V\}$

Decomposability

Definition (Lauritzen, 1996, pp 7–8) The graphical model $\mathcal{M}$ is *decomposable* iff

- $\mathcal{M}$ is complete; or
- $\mathcal{M}$ can be partitioned into disjoint node subsets A, B and C such that
  - $A \perp\!\!\!\perp B | C$
  - C is complete
  - $A \oplus C$ and $B \oplus C$ are decomposable

Note (abuse of notation):

- A can either mean the node set A, or the subgraph obtaind by retaining only the edges in $\mathcal{M}$ that begin and end with nodes in the node set A; similarly for B and C;
- $A \oplus B$ is the graph consisting of all the nodes in A and B and all the edges in $\mathcal{M}$ that begin and end in $A \cup B$.

Obtaining MLE's from a decomposable graph

Suppose $\mathcal{M} = A \cup B \cup C$ where

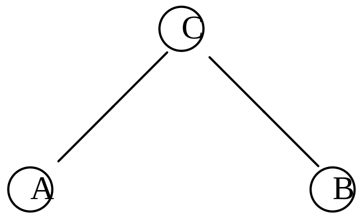
- A, B, C are disjoint
- $A \oplus C$, $B \oplus C$ and C are all complete
- $A \perp\!\!\!\perp B|C$

Then $\mathcal{M}$ is decomposable, and

- $P(\mathcal{M}) = P(A, B, C) = P(A|C)P(B|C)P(C) = P(A \oplus C)P(B \oplus C)/P(C)$
- Complete graph $\Rightarrow$ saturated model $\Rightarrow \hat{m}_{ij\ldots} = n_{ij\ldots}$, so we know how to write the MLE's for $P(A \oplus C)$, $P(B \oplus C)$ and $P(C)$.

If $A \oplus C$ or $B \oplus C$ are not complete, but are decomposable themselves, then *recurse*...

Example $\mathcal{M}$ = [AC][BC]



- [AC], [BC], and [C] are all complete, and $A \perp\!\!\!\perp B|C$ by construction, so the graph is decomposable.
- The probability calculation is

$$\begin{aligned} P(\mathcal{M}) = p_{ijk} &= P(A|C)P(B|C)P(C) \\ &= P(A \oplus C)P(B \oplus C)/P(C) \\ &= p_{i+k}p_{+jk}/p_{++k} \end{aligned}$$

- Multiplying by m_{+++} on both sides, and then multiplying and dividing by m_{+++} on the right, we have

$$\begin{aligned} m_{ijk} = p_{ijk}m_{+++} &= p_{i+k}m_{+++}p_{+jk}/p_{++k} \\ &= p_{i+k}m_{+++}p_{+jk}m_{+++}/(p_{++k}m_{+++}) \\ &= m_{i+k}m_{+jk}/m_{++k} \end{aligned}$$

- For MLE's, replace m's by n's on the right (obs suff stats = exp suff stats).

Interpretation

The idea of decomposability is to reduce the model formulae for $f \in \mathcal{M}$ to a sequence of saturated models. This allows:

- Closed-form MLE's, for log-linear models (and other cases);
- Simple interpretations;
- Suggestions of a causal or time-order sequence for the variables.

Criterion [Leimer, 1989; discussed in Frydenberg & Lauritzen, 1989, Sect. 5]

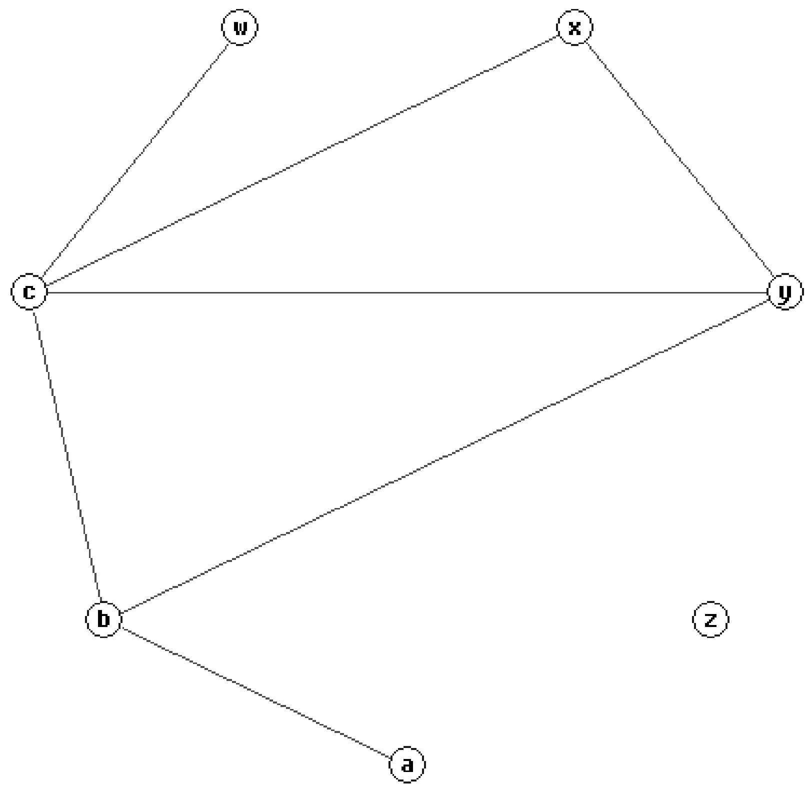
Let $\mathcal{M}$ be a graphical log-linear model.

- Let $c_1, \ldots, c_k$ be all the cliques in $\mathcal{M}$
- Let $b_2 = c_2 \cap c_1$, $b_3 = c_3 \cap (c_2 \cup c_1)$, $\ldots$, $b_k = c_k \cap (\cup_{t=1}^{k-1} c_t)$ be the "running intersections sets"

The clique ordering $c_1, \ldots, c_k$ is an *SD-clique ordering* if, for every $s = 2, 3, \ldots, k$, there is a $j < s$ such that $b_s \subseteq c_j$.

The graphical model $\mathcal{M}$ is *decomposable* iff there is an SD-clique ordering.

Example



The cliques are $c_1 = \{A, B\}$, $c_2 = \{B, C, Y\}$, $c_3 = \{C, W\}$, $c_4 = \{C, X, Y\}$, $c_5 = \{Z\}$.

The running intersection sets are $b_2 = \{B\}$, $b_3 = \{C\}$, $b_4 = \{C, Y\}$, $b_5 = \emptyset$.

Some properties of decomposible graphs

- Decomposible graphs are collapsible.
  - In fact, $\mathcal{M}$ is collapsible onto each $u_s = \cup_{t=1}^{s} c_t$, $s < k$;
  - Clearly, each u_s is also collapsible, onto each u_r, $r < s$;
- Using the relation $f = f_{b|a} f_a$ recursively we have that

$$f = f(c_1) f(c_2|u_1) f(c_3|u_2) \cdots f(c_k|u_{k-1})$$

  for all $f \in \mathcal{M}$.
- Thus to find the MLE $\hat{f}$ for a decomposible log-linear graphical model we should fit the *saturated model* $f(c_s|u_{s-1})$ to each clique, and then multiply the fitted models together.
- MLE's are not much harder to find for general decomposible models.
- It can be shown that log-linear graphical models are decomposible iff they are *triangulated*. (A stronger condition is needed if both discrete and continuous nodes are present). [Lauritzen, 1996, p. 11] See next slides.

- It turns out that decomposable graphical models are exactly those models that are "triangulated":
  - A *cycle* is a sequence of edges $(v_1, v_2), (v_2, v_3), \ldots, (v_{k-1}, v_k)$ in $\mathcal{M}$ such that $v_1, \ldots, v_{k-1}$ are distinct vertices, and $v_1 = v_k$. A *chord* is an edge (v_i, v_j) in $\mathcal{M}$ between non-adjacent vertices ($|i - j| > 1$) not already in the cycle.
  - A *triangulated graph* is one whose *chordless cycles* contain no more than three vertices.
- For a triangulated graphical model it is easy to read off the MLE's:
  - The cell counts will have the form

$$m_{ijklpqr(etc.)} = \frac{\prod(\textit{minimal sufficient margins})}{\prod(\textit{separator margins})}$$

    where
    - The *"minimal sufficient margins"* are the margins with fixed indices corresponding to the terms in the generator (cliques in the graph!); and
    - the *"separator margins"* are the margins with fixed indices corresponding to variables common to terms in the generator (these terms are *minimal* sets of vertices separating the graph into disconnected parts).

Examples

- All log-linear models for three-way tables are decomposable, except for the model of no three-way interaction (which is not even graphical).
- [12][26][235][345]
  - Graph is triangulated
  - The separator sets are [2] (with multiplicity* 2) and [35] (multiplicity 1)

$$m_{ijklrq} = \frac{m_{ij++++}m_{+j+++r}m_{+jk+q+}m_{++klq+}}{(m_{+j++++})^2 m_{++k+q+}}$$

- [12][26][235][345][245] not decomposable—not even graphical!
- [126][234][45][56] is graphical but not triangulated.
- [126][234][456][246] is graphical and triangulated. What are the MLE's $\hat{m}_{ijklrq}$?

**Multiplicity is one less than the number of disjoint subgraphs after removing the separator set.*