Lecture 12 - scribbles

Friday, October 13, 2017

13:10

today: • finish DGM
• UGM
• graph eliminate

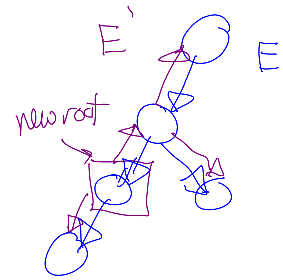
properties of DGM:

inclusion: $E \subseteq E' \Rightarrow \mathcal{L}(G) \subseteq \mathcal{L}(G')$

(adding edges increase # of dist.)

reversal: if G is a directed tree (or forest)

↳ ($|\pi_i| \leq 1$; i.e. there is no v-structure)



then let E' be another directed tree by choosing a different root (reverse some edges while keeping a directed tree)

$\Rightarrow \mathcal{L}(G) = \mathcal{L}(G')$
↳ (V, E')

⊛ this is why direction of edges are not causal

[e.g. $p(x,y) = p(x|y)p(y) = p(y|x)p(x)$

$p(x_V) = \prod_i p(x_i | x_{\pi_i}) = \prod_i p(x_i | x_{\pi_i'})$]

rephrasing: all directed trees from a undirected tree give same DGM

marginalizing:

• marginalizing a leaf n gives us a smaller DGM $= \sum_{x_n} p(x_{1:n})$
$\forall x \ldots$

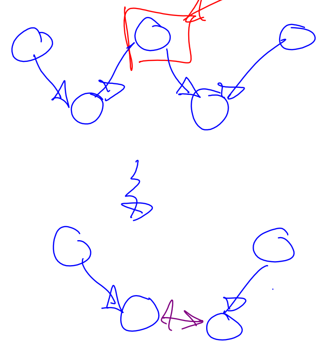
i.e. let $S = \{q$ is dist on $x_{1:n-1}$ s.t. $q(x_{1:n-1}) = p(x_{1:n-1})$ for $p \in \mathcal{L}(G)\}$

then $S = \mathcal{L}(G')$ where G' is G with leaf n removed

- not true for all marginalization:

e.g.

marginalizing out this node



get a set of dist. $\neq$ to any $\mathcal{L}(G')$ for some G'

i.e. marginalization is not a "closed operation" on DGM

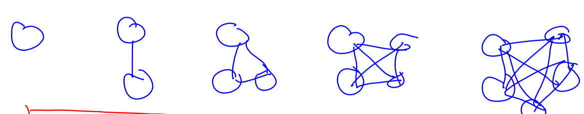
Undirected GM (UGM) (aka Markov random field or Markov network)

let $G = (V, E)$ be an undirected graph

let $\mathcal{C}$ be the set of cliques of G

clique = fully connected set of nodes

i.e. $C \in \mathcal{C} \iff \forall i, j \in C,\ i \neq j,\ \{i, j\} \in E$



then the UGM associated with G

is $\mathcal{F}(G) \triangleq \left\{ p : p(x_V) = \frac{1}{Z} \prod_{C \in \mathcal{C}} \psi_C(x_C) \right\}$

where $\psi_C(x_C) \geq 0 \ \forall x_C$ "potentials"

and $Z \triangleq \sum_x \left(\prod_{C \in \mathcal{C}} \psi_C(x_C) \right)$ "partition function"

notes: • unlike in a B.N. (where could think of C to be (i, π_i), $\psi_C(x_C) = p(x_i | x_{\pi_i})$)

$\psi_C(x_C)$ is <u>not directly</u> related to $p(x_C)$

• can multiply any $\psi_C(\cdot)$ by a constant without changing p

• it is sufficient to consider only $\mathcal{C}_{max}$, the set of <u>maximal cliques</u>

e.g. $C' \subseteq C$

redefine: $\psi_C^{new}(x_C) = \psi_C^{old}(x_C)\,\psi_{C'}(x_{C'})$

(we'll see later e.g. "overparametrization" of tree 
using both $\psi_i(x_i)$ and $\psi_{ij}(x_i, x_j)$)

<u>properties</u>:

as before, $E \subseteq E' \Rightarrow \mathcal{F}(G) \subseteq \mathcal{F}(G')$

$E = \emptyset \Rightarrow$ fully factorized dist.

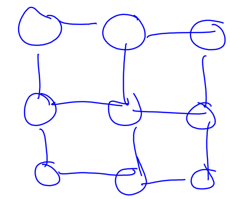
$E =$ all pairs (i.e. one big clique) $\Rightarrow$ all distributions

"exponential family" (we'll see later)

- for $\psi_c(x_c) > 0 \ \forall x_c$ $\quad \langle \theta_c, T_c(x_c) \rangle$

can write $p(x_V) = \exp\Big(\underbrace{\sum_{C \in \mathcal{C}} \log \psi_c(x_c)}_{\text{negative energy function}} - \log Z\Big)$

e.g. Ising model in physics : $x_i \in \{0,1\}$



node potentials → E_i
edge " → E_{ij}

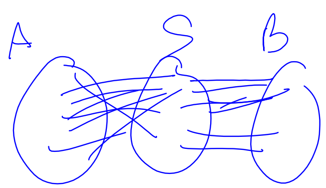
think social network modeling

Conditional independence for UGM

def: we say that p satisfies global Markov property (with respect to undirected graph G)

iff $\forall A, B, S \subseteq V$ disjoint s.t. S separates A from B in G

then $X_A \perp\!\!\!\perp X_B \mid X_S$



S separates A & B
if all paths from A to B pass through S

prop: $p \in \mathcal{F}(G) \Rightarrow p$ satisfies the global Markov property

proof: WLOG, we assume $A \cup B \cup S = V$

and disjoint

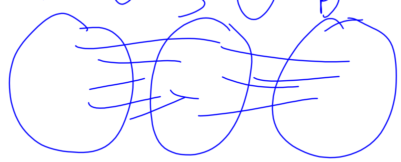
(why? if not, let $\tilde{A} \triangleq A \cup \{a \in V : a \text{ and } A \text{ are not separated by } S\}$

$\tilde{B} \triangleq V \setminus (S \cup \tilde{A})$)

then if have $X_{\tilde{A}} \perp\!\!\!\perp X_{\tilde{B}} \mid X_S$, we have $X_A \perp\!\!\!\perp X_B \mid X_S$ by decomposition

[exercise: show that $b \in \tilde{B} \Rightarrow b$ separated from A by S]

$V = A \cup S \cup B$



let $C \in \mathcal{C}$

cannot have $C \cap A \neq \emptyset$ and $C \cap B \neq \emptyset$

thus $p(x) = \frac{1}{Z} \prod_{\substack{C \in \mathcal{C} \\ C \subseteq A \cup S}} \psi_C(x_C) \underbrace{\prod_{\substack{C' \in \mathcal{C} \\ C' \not\subseteq A \cup S}}}_{\Rightarrow\, C' \subseteq B \cup S \text{ and } C' \not\subseteq S} \psi_{C'}(x_{C'}) = f(x_{A \cup S})\, g(x_{B \cup S})$

$$p(x_A \mid x_S) \propto p(x_{A \cup S}) = \sum_{x_B} p(x_V) = \sum_{x_B} f(x_{A \cup S})\, g(x_{B \cup S})$$

$$= f(x_{A \cup S}) \underbrace{\sum_{x_B} g(x_{B \cup S})}_{\text{constant w.r.t. } x_A}$$

$$\Rightarrow p(x_A \mid x_S) = \frac{f(x_A, x_S)}{\sum_{x'_A} f(x'_A, x_S)}$$

similarly, $p(x_B \mid x_S) = \dfrac{g(x_B, x_S)}{\sum_{x'_B} g(x'_B, x_S)}$

$$p(x_A|x_S)\,p(x_B|x_S) = \frac{f(x_{A\cup S})\,g(x_{B\cup S})}{\sum_{x_B'}\sum_{x_A'} f(x_A', x_S)\,g(x_B'|x_S)} = \frac{p(x_V)}{p(x_S)} = p(x_A, x_B | x_S)$$

ie. $X_A \perp\!\!\!\perp X_B \mid X_S$ //

thm: (Hammersley-Clifford) if $p(x_V) > 0 \;\forall x_V$

then $p \in \mathcal{F}(G) \iff p$ satisfies global Markov property

$\xLeftarrow{(HC)}$

proof: see ch. 16 of Mike's book; use "Möbius inversion formula"

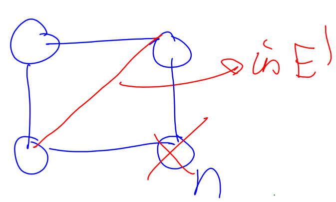
(as exclusion-inclusion principle in prob.)

properties:

* closure with respect to marginalization:

let $V' = V \setminus \{n\}$ $\quad E' =$ edges in $G \setminus \{n\}$

+ connect all neighbors of n in G together (new clique)



if $\{$marginal on $x_{1:n-1}$ for $p \in \mathcal{F}(G)\} = \mathcal{F}(G')$

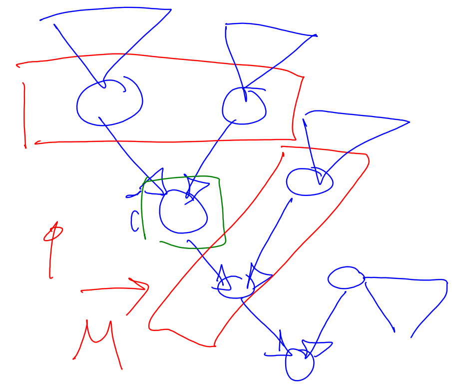
DGM vs UGM:

def: Markov blanket for i is the smallest set of nodes M

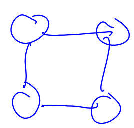
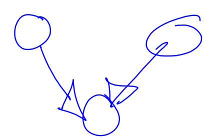
(for graph G)

s.t. $X_i \perp\!\!\!\perp X_{V\setminus ...} \mid X_M$ "rest"

- for UGM : $M = \{ j : \{i,j\} \in E \}$ = set of neighbors of i
- for DGM : $M = \pi_i \cup \text{children}(i) \cup \pi_j \big|_{j \in \text{children}(i)}$



recap:

	DGM	UGM
factorization:	$p(x) = \prod_i p(x_i \mid x_{\pi_i})$	$\frac{1}{Z} \prod_C \psi_C(x_C)$
cond indep.:	d-separation	separation
marginalization:	not closed in general  but fine for leaves	closed
cannot exactly capture:		 v-structure

Moralization

let G be a DAG; when can we transform to equivalent UGM?

def: for G a DAG, we call $\bar{G}$ the moralized graph of G

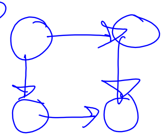
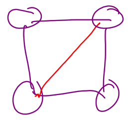
where $\bar{G}$ is an undirected graph with same V

and $\bar{E} = \{ \{i,j\} : (i,j) \in E \}$ } undirected version of E

$\cup \{ \{k,l\} : k \neq l \in \pi_i \text{ for some } i \}$ } "moralization" "marrying the parents" (???)

↑ connect all the parents of i with i in big clique

↳ only add edges when $|\pi_i| > 1$ i.e. v-structure

e.g. G  $\rightsquigarrow$ $\bar{G}$ 

prop: for a DAG G with no v-structure [forest]

then $\underset{\text{DGM}}{\mathcal{L}(G)} = \underset{\text{UGM}}{\mathcal{L}(\bar{G})}$

but in general, can only say that $\mathcal{L}(G) \subseteq \mathcal{L}(\bar{G})$

[note that $\bar{G}$ is the minimal undirected graph s.t. $\mathcal{L}(G) \subseteq \mathcal{L}(\bar{G})$]