Lecture 13 - UGM

Tuesday, October 24, 2023 2:27 PM

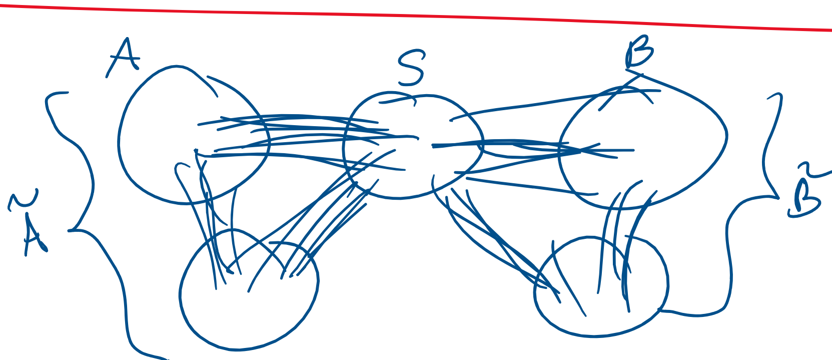
today: UGM (cont.)
DGM vs UGM

cond. independence for UGM:

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

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

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



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

proof: WLOG, we assume that $A \cup B \cup S = V$ & A separated from B by S

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

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

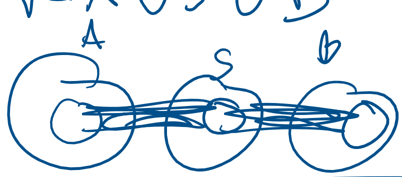
$A \subseteq \tilde{A}$
$B \subseteq \tilde{B}$

then if have $X_{\tilde{A}} \perp\!\!\!\perp X_{\tilde{B}} \mid X_S$

by decomposition $\Rightarrow X_A \perp\!\!\!\perp X_B \mid X_S$

[exercise: show that if $b \in \tilde{B}$ $\Rightarrow b$ is separated from $\tilde{A}$ 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 $p(x) = \overbrace{\frac{1}{Z} \prod_{\substack{C \in \mathcal{C} \\ C \subseteq A \cup S}} \psi_C(x_C)}^{} \overbrace{\prod_{\substack{C' \in \mathcal{C} \\ C' \not\subseteq A \cup S}} \psi_{C'}(x_{C'})}^{} = f(x_{A\cup S})\, g(x_{B\cup S})$

$\underbrace{C' \not\subseteq A \cup S}_{} \Rightarrow C' \subseteq B \cup S$

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

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

$$\Rightarrow p(x_A | x_S) = \frac{f(x_{A\cup S})}{\sum_{x'_A} f(x'_A, x_S)}$$

similarly, $p(x_B | x_S) = \dfrac{g(x_{B\cup S})}{\sum_{x'_B} g(x'_B, x_S)}$

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

i.e. $X_A \perp\!\!\!\perp X_B \mid X_S$

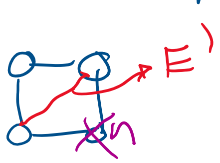
thm: (Hammersley–Clifford) if $\underline{p(x_V) > 0}$ $\forall x_V$

then $p \in \mathcal{L}(G) \Longleftrightarrow p$ satisfies the global Markov property for G

proof: see ch. 16 of Mike's book; use "Möbius inversion formula" (as exclusion-inclusion principle in prob.)

more properties of UGM:

* closure with respect to marginalization



let $V' = V \setminus \{n\}$

E' = edges in $G \setminus \{n\}$
+ connect all neighbors of n in G ~~together~~
(new clique)

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

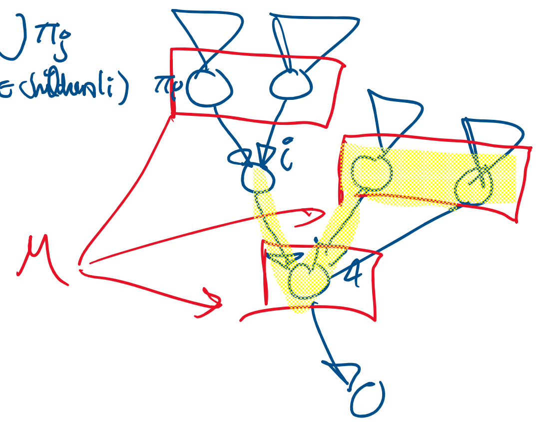
$\{$marginal on $x_{1:n-1}$ for some $p \in \mathcal{L}(G)\} = \mathcal{L}(G')$

DGM vs. UGM

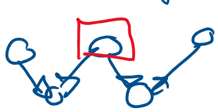
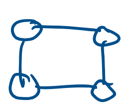
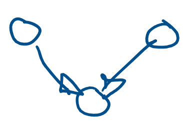
def: Markov blanket for i (for graph G) is the smallest set of nodes M

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

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



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	closed
	e.g. 	(connect all neighbors of removed node)
	but is fine for leaves	
cannot exactly capture some families		

Moralization:

Let G = DAG ...

Moralization:

let G a DAG; when can we get an 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

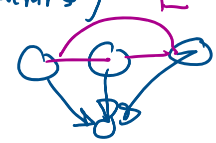
and $\bar{E} = \{ \{i,j\} : (i,j) \in E \}$ } undirected version of E

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

connect all the parents of i with i in big clique

only needed if $|\pi_i| > 1$ (i.e. v-structure)

$\bar{E}$



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

prop: for a DAG with no v-structure [forest] then $\mathcal{L}(G) = \mathcal{L}(\bar{G})$ (proof: see hwk 3?)

but in general, can only $\underset{\text{DAG}}{\mathcal{L}(G)} \subseteq \mathcal{L}(\bar{G})$

[note that $\bar{G}$ is the minimal undirected graph G' s.t. $\mathcal{L}(G) = \mathcal{L}(G')$]

$$p(x) = \prod_i \underbrace{p(x_i \mid x_{\pi_i})}_{\psi_C(x_C)} \quad \text{where } C = \{i\} \cup \pi_i$$

$\to \bar{G}$

15h39

general themes in this class:

A) representation → DGM
↳ UGM } model high dim.

A) representation $\longrightarrow$ DGM
$\quad\quad\quad\quad\quad\quad\quad\quad\;\; \hookrightarrow$ UGM

parameterization $\rightsquigarrow$ exponential family

} model high dim. distribution

B) inference $p(x_Q \mid x_E)$ $\longrightarrow$ next class: elimination algorithm
"query" "evidence" $\hookrightarrow$ next: sum-product / belief propagation (for trees)

later: approximate inference eg MCMC, variational

C) statistical estimation / learning $\longrightarrow$ MLE, maximum entropy, method of moments

Inference:

want to compute:

a) marginal: $p(x_F)$ for some $F \subseteq V$

b) conditional: $p(x_F \mid x_E)$
"query" "evidence"

c) for UGM: partition function $Z = \sum_{x_V} \left(\prod_C \psi_C(x_C) \right)$

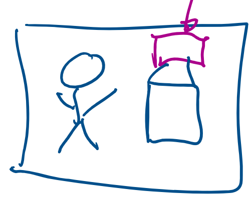
exponential sum

why?:

- missing data $p(x_{\text{unobs}} \mid x_{\text{obs}})$

example:

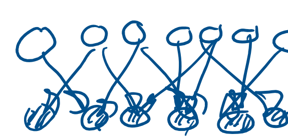
missing pixels



"infilling task"

- prediction $p(x_{\text{future}} \mid x_{\text{past}})$

- "latent cause" $p(x_{\text{cause}} \mid x_{\text{obs}})$



diseases (causes)

symptom

* also related inference:

* also related inference:

$$\underset{x_F}{\operatorname{argmax}}\; p(\underset{\substack{\text{could}\\ \text{be big}}}{x_F} \mid x_E) \qquad (\text{Viterbi alg.}) \qquad (\text{speech recognition})$$

* inference is also needed during <u>estimation</u> (parameter fitting MLE)

[e.g. during EM, E-step $p(z|x)$]

* we present inference alg. for UGM [for simplicity and more generality]

make DGM $\longrightarrow \subseteq$ UGM via moralization

i.e. $p \in$ DGM; $p(x) = \prod_i p(x_i | x_{\pi_i})$ moralize: $C_i \triangleq \{i\} \cup \pi_i$

$$= \frac{1}{Z} \prod_i \psi_{C_i}(x_{C_i}) \qquad \psi_{C_i}(x_{C_i}) \triangleq p(x_i | x_{\pi_i})$$

$Z = 1$

$$p(x_1) = \sum_{x_{2:n}} p(x) = \sum_{x_{2:n}} \prod_i \frac{1}{Z} \psi_{C_i}(x_{C_i})$$