6.438 Fall 2020 Exam 1

Rylan Schaeffer

Massachusetts Institute of Technology
Department of Electrical Engineering and Computer Science
6.438 Algorithms for Inference
Fall 2020

QUIZ 1 ANSWER BOOKLET

Thursday, October 15, 2020
9:00 am – 11:00 am / 7:00 pm – 9:00 pm

NAME: Rylan Schaeffer

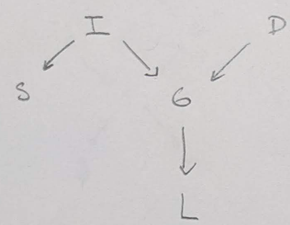
- Don't forget to put your name on all sheets.
- Remember that only this answer booklet will be considered in the grading of your exam.
- Be sure to **show all relevant work and reasoning.**
- Please be neat! You may want to first work things through on scratch paper and then neatly transfer to this answer booklet the work you would like us to look at.

Problem	Score
1	
2	
3	
Total	

NAME: Rylan Schaeffer

Problem 1

(a)(i) Draw the DAG here:



(a)(ii) Determine if the CI relations are true or false:

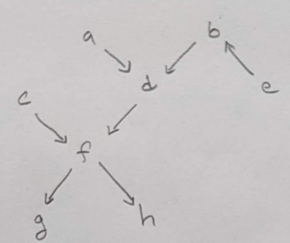
- $D \perp\!\!\!\perp I$: True / False
- $D \perp\!\!\!\perp S \mid L$: True / False
- $S \perp\!\!\!\perp L \mid I$: True / False

Reasoning/Work to be looked at for Problem 1(a):

$D \perp\!\!\!\perp I$ is True b/c G is not conditioned, meaning $I \to G \leftarrow D$ is unblocked

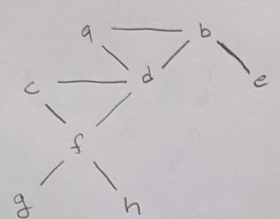
$D \perp\!\!\!\perp S \mid L$ is False b/c we condition on L, child of G, unblocking path $D \to G \leftarrow I \to S$

$S \perp\!\!\!\perp L \mid I$ is True b/c $S \leftarrow I \to G \to L$ is blocked by conditioning on I (b)(i) Draw the I-equivalent DAG here:



I flipped b → e. Any other flip adds/removes immoralities

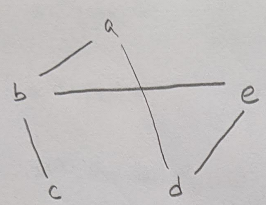
(b)(ii) Draw the minimal undirected I-map here:



Reasoning/Work to be looked at for Problem 1(b):

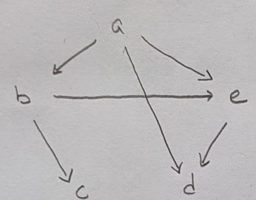
NAME: Rylan Schaeffer

(c)(i) Draw the undirected graph here:



I hope I didn't make a mistake, but this graph satisfies all CI. & I don't think I can remove more edges

(c)(ii) Draw the DAG here:



This is a valid DAG. I add a directed edge $a \to e$ to make DAG chordal. By theorem in class, this directed graph is minimal I-map of (c)(i)

Reasoning/Work to be looked at for Problem 1(c):

(c)(i) Start fully connected. Remove edges.

e.g. $a \perp\!\!\!\perp c \mid b \Rightarrow$ no edge $a - c$

$a \perp\!\!\!\perp c \mid b, d \Rightarrow$ no edge $a - e - c$

$c \perp\!\!\!\perp d \mid b \Rightarrow$ no edge $c - d$

$c \perp\!\!\!\perp e \mid b \Rightarrow$ no edge $c - e$

$a \perp\!\!\!\perp e \mid b, d \Rightarrow$ no edge $a - e$

$b \perp\!\!\!\perp d \mid a, e \Rightarrow$ no edge $b - d$

~~$a \perp\!\!\!\perp d \mid a, e \Rightarrow$ at least 1 edge e-b-e~~

NAME: Rylan Schaeffer

Problem 2

(a) $f(\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(N)}; \beta_{12}) = \dfrac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(x_2 - \beta_{12} x_1)^2\right\}$

$g(\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(N)}; \beta_{13}) = \dfrac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(x_3 - \beta_{13} x_1)^2\right\}$

$h(\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(N)}; \beta_{24}, \beta_{34}) = \dfrac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(x_4 - \beta_{24} x_2 - \beta_{34} x_3)^2\right\}$

Reasoning/Work to be looked at for Problem 2(a):

Consider just 1 sample for simplicity

$$p(x; \beta_{12}, \beta_{13}, \beta_{24}, \beta_{34}) = p(x_1) \underbrace{p(x_2 | x_1, \beta_{12})}_{f(x^1 \ldots x^N; \beta_{12})} \underbrace{p(x_3 | x_1, \beta_{13})}_{g(x^1 \ldots x^N; \beta_{13})} \underbrace{p(x_4 | x_2, x_3, \beta_{24}, \beta_{34})}_{p(x^1 \ldots x^N; \beta_{24}, \beta_{34})}$$

$$f(x^1 \ldots x^N; \beta_{12}) = p(x_2^1 \ldots x_2^N | x_1^1 \ldots x_1^N, \beta_{12})$$

$$\cancel{\mathbb{E}[x_2] = \mathbb{E}[\beta_{12} x_1 + \varepsilon_1] = \beta_{12} \mathbb{E}[x_1] + \mathbb{E}[\varepsilon_1] = 0 + 0}$$

$$\cancel{\mathbb{V}[x_2] = \mathbb{V}[\beta_{12} x_1] + \mathbb{V}[\varepsilon_2]}$$

$$= \frac{1}{\sqrt{2\pi (1)^2}} \exp\left\{-\frac{1}{2} \frac{(x_2 - \beta_{12} x_1)^2}{(1)^2}\right\}$$

Similarly:

$$g(x^1 \ldots x^N; \beta_{13}) = p(x_3^1 \ldots x_3^N | x_1^1 \ldots x_1^N, \beta_{13}) = N(\beta_{13} x_1, \sigma^2 = 1)$$

$$h(x^1 \ldots x^N; \beta_{24}, \beta_{34}) = N(\beta_{24} x_2 + \beta_{34} x_3, \sigma^2 = 1)$$

NAME: Rylan Schaeffer

(b) $\hat{\beta}_{12} = \left(\sum_{n=1}^{N} x_2^{(n)} x_1^{(n)}\right) \Big/ \left(\sum_{n=1}^{N} (x_1^{(n)})^2\right)$

$\hat{\beta}_{13} = \left(\sum_{n=1}^{N} x_3^{(n)} x_1^{(n)}\right) \Big/ \left(\sum_{n=1}^{N} (x_1^{(n)})^2\right)$

$\hat{\beta}_{24} = \left[\left(\sum_{n=1}^{N} \begin{bmatrix} x_2^n \\ x_3^n \end{bmatrix} \begin{bmatrix} x_2^n & x_3^n \end{bmatrix}\right)^{-1} \left(\sum_{n=1}^{N} \begin{bmatrix} x_2^n \\ x_3^n \end{bmatrix} x_4^n\right)\right]_1 \longleftarrow$ i.e. first element of this 2×1 vector

$\hat{\beta}_{34} = \left[\left(\sum_{n=1}^{N} \begin{bmatrix} x_2^n \\ x_3^n \end{bmatrix} \begin{bmatrix} x_2^n & x_3^n \end{bmatrix}\right)^{-1} \left(\sum_{n=1}^{N} \begin{bmatrix} x_2^n \\ x_3^n \end{bmatrix} x_4^n\right)\right]_2 \longleftarrow$ i.e. second element of this 2×1 vector

---

Reasoning/Work to be looked at for Problem 2(b):

Define $l_1(\beta_{12}) = \log \prod_{n=1}^{N} p(x_2^n \mid x_1^n, \beta_{12})$

$$\propto \sum_{n=1}^{N} \log \exp\left\{-\frac{1}{2}(x_2^n - \beta_{12} x_1^n)^2\right\}$$

Take derivative wrt parameter, set equal to 0, Solve for $\hat{\beta}_{12}$

$$\partial_{\beta_{12}} l(\beta_{12}) = 0 = \sum_{n=1}^{N} (x_2^n - \hat{\beta}_{12} x_1)\, x_1^n$$

$$\Rightarrow \hat{\beta}_{12} = \left(\sum_{n=1}^{N} x_2^n x_1^n\right) \Big/ \left(\sum_{n=1}^{N} (x_1^n)^2\right)$$

Similarly define $l(\beta_{13}) = \log \prod_{n=1}^{N} p(x_3^n \mid x_1^n, \beta_{13})$ & take same approa

$$\Rightarrow \hat{\beta}_{13} = \left(\sum_{n=1}^{N} x_3^n x_1^n\right) \Big/ \left(\sum_{n=1}^{N} (x_1^n)^2\right)$$

Then define $l(\beta_{24}, \beta_{34}) = \log \prod_{n=1}^{N} p(x_4^n \mid x_2^n, x_3^n, \beta_{24}, \beta_{34})$

$$\partial_{\beta_{24}} l(\beta_{24}, \beta_{34}) = 0 = \sum (x_4^n - \hat{\beta}_{24} x_2 - \hat{\beta}_{34} x_3)\, x_2$$

$$\Rightarrow \hat{\beta}_{24} =$$

Similarly for $\partial_{\beta_{34}} l(\beta_{24}, \beta_{34})$ by symmetry

(c) **Sufficient Statistics and Natural Parameters:**

Natural parameters: $\{J_{ij}\} \cup \{h_i\}$ for $i,j \in \{1,2,3,4\}$

Sufficient statistics: $\mathbb{E}_{p(x_1 x_2 x_3 x_4)}[x_i x_j]$ and $\mathbb{E}_{p(x_1 x_2 x_3 x_4)}[x_i]$

Reasoning/Work to be looked at for Problem 2(c):

We know x_1, x_2, x_3, x_4 are jointly Gaussian.

$$p(x) = p(x_1, x_2, x_3, x_4) \propto \exp\left(-\frac{1}{2} x^T J x + h^T x\right)$$

$$\propto \exp\left(-\frac{1}{2} \sum_{i,j} J_{ij} x_i x_j + \sum_i h_i x_i\right)$$

$\Rightarrow$ Natural Parameters are $\{J_{ij}\} \cup \{h_i\}$

Sufficient statistics are $\mathbb{E}_p[x_i x_j]$, $\mathbb{E}_p[x_i]$

(d) Reasoning/Work to be looked at for Problem 2(d):

Let x_1 be first node in topological ordering.

By construction, $x_1 = \varepsilon_1$ is Gaussian

B/c linear transform of Gaussian remains Gaussian,

$p(x_i \mid x_{\pi_i})$ is Gaussian

The joint distribution, given by the graph factorization, is Gaussian

$$p(x_V) = \prod_{i=1} p(x_i \mid x_{\pi_i})$$

$\Rightarrow$ sufficient statistics for Gaussian include $\mathbb{E}_p[x_i x_j]$

~~If i, j has no edge in undirected graph $\Rightarrow$ $J_{ij} = 0$~~

Suppose $\exists$ edge $(i,j) \in \mathcal{E}$ in undirected I-map. Then

$J_{ij} \neq 0 \Rightarrow \mathbb{E}_p[x_i x_j]$ is a sufficient statistic for natural parameter J_{ij}

Suppose $\mathbb{E}_p[x_i x_j]$ is a sufficient statistic for $J_{ij} \Rightarrow J_{ij} \neq 0$

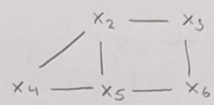
$\Rightarrow \exists$ edge $(i,j) \in \mathcal{E}$ in undirected I-map

NAME: Rylan Schaeffer

Problem 3

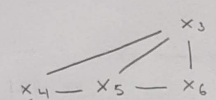
(a)(i) Draw the sequence of graphs:

Elim 1:



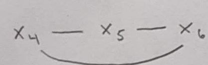
⇓

Elim 2:



⇓

Elim 3



Elim 4:

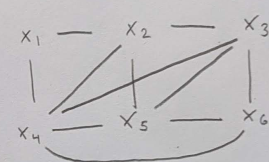
$x_5 - x_6$

Elim 5:

x_6

Elim 6:

Reconstituted Graph



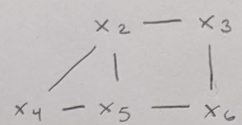
Draw the reconstituted graph:

See

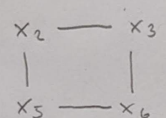
(a)(ii) Choose your elimination order: 1, 4, 6, 3, 5, 2

Draw the sequence of graphs:

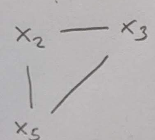
Elim x_1



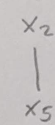
Elim x_4



Elim x_6



Elim x_3

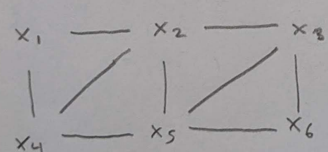


Elim x_5

x_2

Elim x_2

Draw the reconstituted graph:



max clique size: 3

NAME: Rylan Schaeffer

(b)(i) Draw the interventional graph:

$$\begin{tikzcd}
X_1 \arrow[r] & X_2 \arrow[r] \arrow[d] & X_3 \arrow[d] \\
X_4 \arrow[r] & X_5 \arrow[r] & X_6
\end{tikzcd}$$

(b)(ii) Write the distribution $p(x_1, x_2, x_3, x_5, x_6 \mid \text{do}(x_4 = x_4))$ in terms of factors:

$$p(x_1, x_2, x_3, x_5, x_6 \mid do(x_4 = x_4)) =$$

$$= \mathbb{1}_{x_4 = x_4} \; p(x_1)\, p(x_2 \mid x_1)\, p(x_3 \mid x_2)\, p(x_5 \mid x_2, x_4)\, p(x_6 \mid x_3, x_5)$$

NAME: Rylan Schaeffer

(b)(iii) **Reasoning/Work to be looked at for Problem 3(b)(iii):**

Causally operating to set $x_4 = X_4$ effectively removes edge

$X_1 \longrightarrow X_4$

First, eliminate $\underline{X_1}$ and then $\underline{X_4}$, each in

$O(|X|^2)$ time.

Then, proceed w/ elimination on remaining graph

(i.e. X_2, X_3, X_5, X_6) in $O(|X|^4)$ time

Total time: $2\,O(|X|^2) + O(|X|^4) = O(|X|^4)$