Problem 1.

$X \to$ have disease / not

$Y \to$ test positive / not

$P(X=1) = 1/10000$

$P(X=1 \mid Y=1) = P(Y=1 \mid X=1) \cdot P(X=1) / P(Y=1)$

$= 0.99 \times 1e^{-4} \Big/ \sum_x P(x, Y=1)$

$= 0.99 \times 1e^{-4} \Big/ 0.99 \times 1e^{-4} + 0.01 \times (1 - 1e^{-4})$

$= 0.99 \Big/ 0.99 + 99.99 \approx 1e^{-4}$

$\frac{1}{102}$

$P(Y=1 \mid X=1) \to 0.99$ 0.98

$P(Y=1 \mid X=0) = 0.1$

Problem 2.

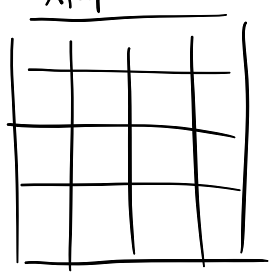
$$\max_{x_1, x_2, \dots x_n \in S^n} P(x_1, x_2, \dots, x_n)$$

$x_1 \to x_2 \to x_3 \to x_4 \cdots \to x_n$

$S = \{v_1, \dots v_m\}$

x_{i-1}

x_i



table

$m \times m$

Design $O(m^2 n)$ algorithm

$P(x_1 = v)$ for each $v \in S$

dp [i] [j]

$P(x_1, \cdots x_{i-1}) \cdot P(x_i | x_{i-1}) = P(x_1, \cdots x_i)$

$= dp[i]$

for a given col ?

x_{i-1} is given

$$\max\left(dp[i-1][j-1] \cdot T(j-1, k),\ dp[i][j]\right)$$

for 1 to n: ← i

for 1 to m: ← j

for 1 to m: ← k

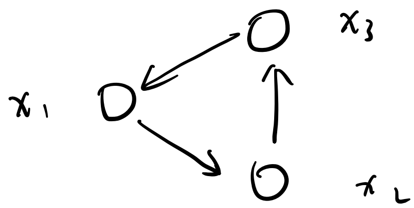
$O(m^2 n)$

Problem 3.

relax the defination of Bayesian network by removing the acyclic assumption

$$f(x_1, \cdots, x_n) = \prod_{v \in V} f_v(x_v | x_{pa(v)})$$

If G has a directed cycle



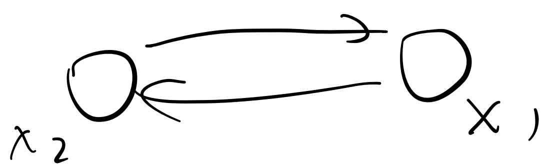
$$f(x_1, x_2, x_3) = f(x_1 \mid x_3) \cdot f(x_2 \mid x_1) \cdot f(x_3 \mid x_1)$$

$$\sum_{x_3} \sum_{x_2} \sum_{x_1} f(x_1, x_2, x_3) = 1 \quad \text{in theory}$$

$$\sum_{x_v \in Val(x_v)} f_v(x_v \mid x_{pa(v)}) = 1$$

$$\sum_{x_1} f(x_1 \mid x_3)\, f(x_2 \mid x_1) \cdot f(x_3 \mid x_1)$$

$x_3 \sim$
$x_1 \sim \{0, 1\}$
$x_2 \sim$



$$P(x_1, x_2) = P(x_1 \mid x_2) \cdot P(x_2 \mid x_1)$$

$$\begin{array}{cc|cc} & & x_1 & \\ & & 0 & 1 \\ \hline x_2 & 0 & p_1 & p_2 \\ & 1 & p_3 & p_4 \end{array} \qquad \begin{array}{cc|cc} & & x_2 & \\ & & 0 & 1 \\ \hline x_1 & 0 & q_1 & q_2 \\ & 1 & q_3 & q_4 \end{array}$$

$p_1 + p_2 = 1$, $p_3 + p_4 = 1$, $q_1 + q_2 = 1$
$q_3 + q_4 = 1$

$$\sum \sum p(x_1, x_2)$$

$$= (1,1) + (1,0) + (0,1) + (0,0)$$

$$= p_4 q_4 + p_2 q_3 + p_3 \cdot q_2 + p_1 q_1$$

$$= p_4 q_4 + (1-p_1) q_3 + p_3 (1-q_1) + p_1 q_1$$

$$= p_4 q_4 + q_3 - p_1 q_3 + p_3 - q_1 p_3 + p_1 q_1$$

$$\frac{0.05 + 0.05}{0.25 + 0.25}$$ 边缘

$p(x_1 | x_2)$ $p(x_2 | x_1)$

P_1 P_2

no constraints ?

$$\frac{p(x_1, x_2)}{p(x_2)} \qquad \frac{p(x_1, x_2)}{p(x_1)}$$

Problem 4

4.1

$$P(\alpha \mid \beta, \gamma) = \frac{P(\alpha, \beta, \gamma)}{P(\beta, \gamma)}$$

For set 1.

$P(\beta, \gamma)$, $P(\alpha)$, $P(\alpha, \beta)$, $P(\gamma, \alpha)$

Nop

For set 2.

$P(\beta, \gamma)$, $P(\alpha)$, $P(\beta, \gamma, \alpha)$

Yep

For set 3.

$P(\beta \mid \alpha)$ $P(\gamma \mid \alpha)$ $P(\alpha)$

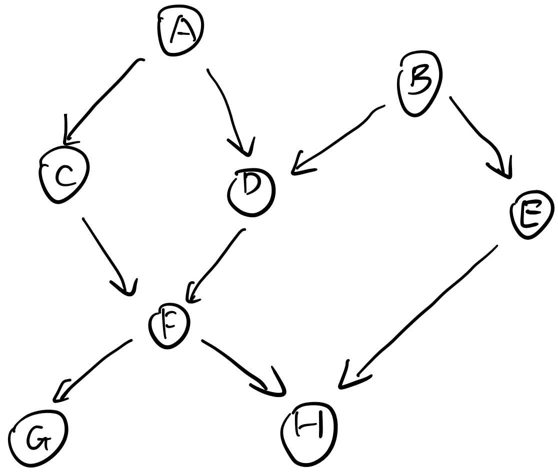
Nop

4.2

$$P(\beta \mid \alpha) \cdot P(\gamma \mid \alpha) = P(\beta, \gamma \mid \alpha)$$

Set 1 yep Set 2 yep Set 3 yep

Problem 5.



B

We have $P(A)$, $P(B)$, $P(E|B)$, $P(D|A,B)$ and we can get the independent relationship from the graph B.

$$P(A=0, B=0) = P(A=0) \cdot P(B=0) = 0.8 \times 0.3 = 0.24$$

$$P(E=1 | A=1) = P(E=1, A=1) / P(A=1)$$

$$P(A, B, E) = P(A) \cdot P(B) \cdot P(E|B)$$

$$P(E=1, A=1) = \sum_{B} P(A=1, B, E=1)$$

$$= 0.2 \times 0.7 \times 0.1 + 0.2 \times 0.3 \times 0.9$$

$$= 0.014 + 0.054 = 0.068$$

$$P(E=1|A=1) = 0.34$$

2. $\text{d-sep}_B(A ; E \mid \{B, H\})$ No

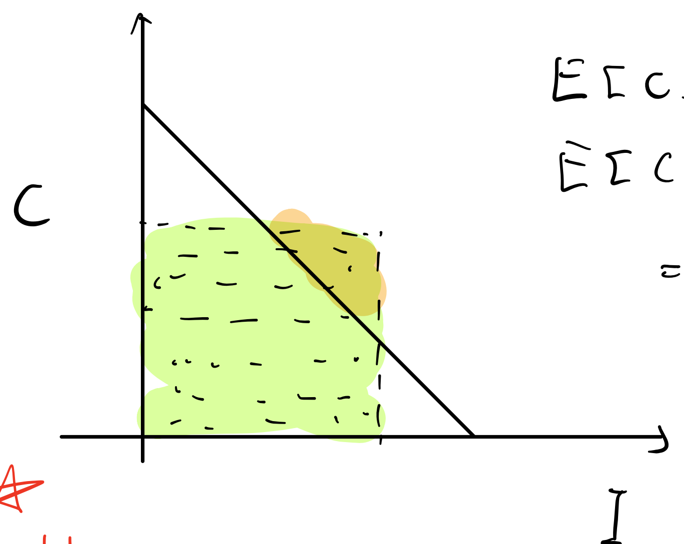
$\text{d-sep}_B(G ; E \mid D)$ No

$\text{d-sep}_B(\{A, B\} ; \{G, H\} \mid F)$ No

Problem 6

admitted

$A \sim \{0, 1\}$



$E[C] = 0.5$

$E[C \mid A=1] = \int_{0.5}^{1} P(C \mid A=1)\, c\, dc$

$= \int_{0.5}^{1} \frac{P(C, A=1)}{P(A=1)}\, c\, dc$

$= 8 \int_{0.5}^{1} c \cdot P(C, I+C \geq 1.5)\, dc$

another way:

$f_{C \mid I+C \geq 1.5} = \frac{f_{I+C \geq 1.5 \mid c} \cdot f_c}{f_{I+C \geq 1.5}}$

$= \frac{(0.5-c) \cdot 1}{\frac{1}{8}}$

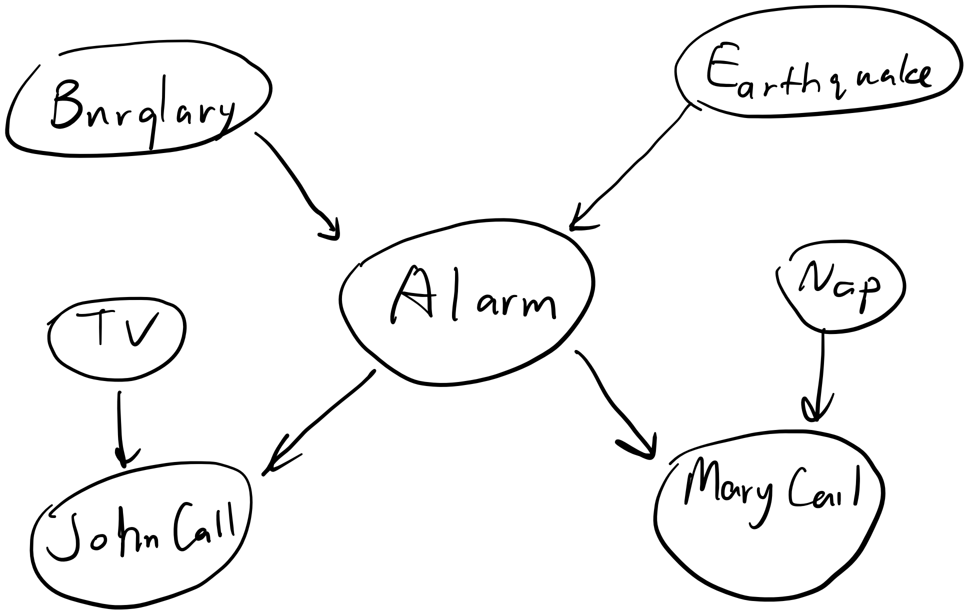
$= \iint_{I+C \geq 1.5} P(C \mid A=1)\, dI\, dc$

$= 8 \int_{0.5}^{1} \int_{1.5-c}^{1} c \cdot dc$

$E[C \mid I=0.95] = 0.5$

$E[C \mid I=0.95, A=1] = E[c \mid c \geq 0.55] = 0.775$

Problem 7



用 $x_j \to x_i$

引入 independence

$$(X_i \perp \underbrace{x_j}_{} \mid \{Pa(x_i) - x_j\})$$

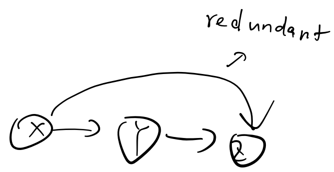
in fact is a set $\{x_j, x_k \dots\}$

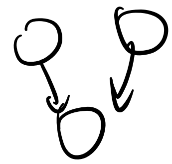
check if this is hold in I

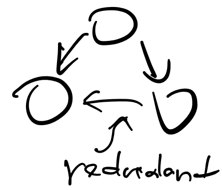
BN 定义, given parents then independent of non decent nodes

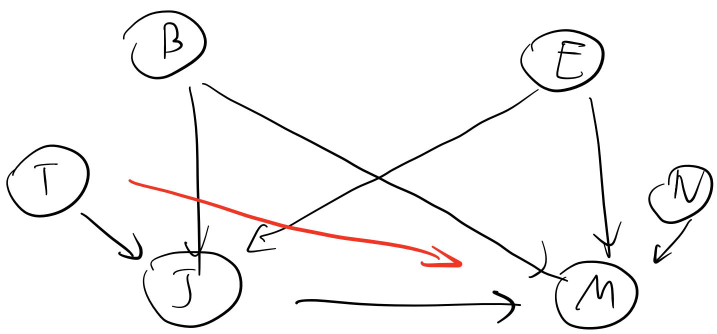
这里其实就是通过原图，看有没有dependence
要是有就加边就行了

Alarm 就是 unobserved









T → M ?

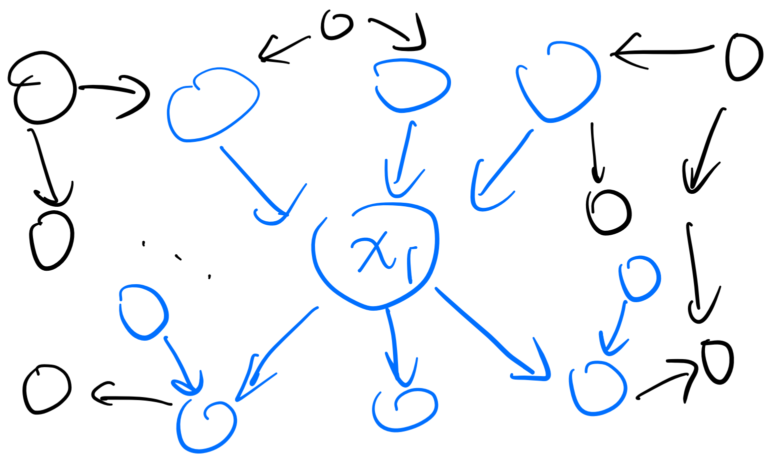
Generalize :

$$P_{BN}(x_1, \dots, \boxed{x_i}, x_{i+1}, \dots, x_n)$$

remove

Min im I-Map of BN

$$P_{BN}(x_1, \dots, \quad , x_{i+1}, \dots, x_n)$$



Markov Blanket''

其实就是只关注与 x_i 有联系的部分

Problem 8

8.1: $(X_1, X_2, \dots, X_n)$

Use Markov Blanket

Notation: $\pi(i) \to$ set of parent nodes of X_i,

$C_{(i)} \to$ set of children nodes of X_i

$\pi(C_i) \to$ set of parent of children nodes of X_i (excluding X_i)

$O(i)$: everything outside the Market blanket of X_I, formally defined as

$$O(i) = V \setminus [\pi(i) \cup C_i \cup \pi(C_i) \cup \{X_i\}]$$

$$P(x_i \mid x_1, \dots, x_n) = P(x_i \mid O(i), \pi(i), \pi(C_i), C_i)$$

$$= \frac{P(x_i, O(i), \pi(i), \pi(C_i), C_i)}{\sum_{x_i} P(x_i, O(i), \pi(i), \pi(C_i), C_i)}$$

$$= \frac{P(x_i \mid \pi(i)) \cdot P(C_i \mid \pi(C_i), x_i) P(\pi(i), O_i, \pi(C_i))}{\sum_{x_i} P(x_i \mid \pi(i)) \cdot P(C_i \mid \pi(C_i), x_i) P(\pi(i), O_i, \pi(C_i))}$$

$$= \frac{P(x_i \mid \pi(i)) P(C_i \mid \pi(C_i), x_i)}{\sum_{x_i} P(x_i \mid \pi(i)) P(C_i \mid \pi(C_i), x_i)}$$

8.2: Topologically sort X

for each X_i in X sample $X_i = x_i$ from $P(X_i \mid X_{pa(i)} = x_{pa(i)})$

Coding Part:

Q6

$$P(x_1,\dots x_{784} \mid z_1 z_2) = P(x_1 \mid z_1 z_2)\cdots P(x_{784} \mid z_1 z_2)$$

$$P(x_1,\dots x_{784}) = \sum_{z_1 z_2} P(z_1 z_2)\cdot P(x_1 \mid z_1 z_2)\cdots P(x_{784} \mid z_1 z_2)$$

$$\log P(x_1,\dots, x_{784}) = \log \sum_{z_1 z_2} \underbrace{P(z_1 z_2)\cdot P(x_1 \mid z_1 z_2)\cdots P(x_{784} \mid z_1 z_2)}_{q}$$

point: $\prod P(x_i \mid z_1 z_2)$ can be very small as to cause underflow, in the end, we are just summing 0. and $\log 0 \to -\inf$

The sum prevent us from transforming q to $\log q$ $\to$ a reasonable number for calculation

Consider $\log \sum_{z_1 z_2} e^{\log q}$

$= \log \sum e^{\log q - a}$

$- \log (\sum_{z_1 z_2} \quad)$

$= a + \log \sum_{z_1 z_2} e^{\log q - a}$

Calculate q using log → To reduce the computation cost

We have conditional prob vector
so just log and np. sum

For an image vector [1 x 784]
use the np.where function

image: [0, 1, 1, 0, 0, ...]

prob vect [P_{x_1}, P_{x_2}, ... $P_{x_{784}}$]

the prob equals 0

[log(map [Px_1, Px_2, ... Px_{784}])].sum

$\begin{cases} x_i = 0 & px_i \\ x_i = 1 & 1 - px_i \end{cases}$

using np. select( )

Q7

$$P\left(z_1, z_2 \mid X_{1:784} = I^k\right) \qquad P(z_1, z_2 \mid X_{1:784})$$

$$\sum_{z_2} \boxed{\frac{P(z_1, z_2, X_{1:784})}{P(X_{1:784})\,[Q_6]}} \qquad z_2 \to \text{mean } z_2$$

$$\sum_{z_1} \frac{P(z_1, z_2, X_{1:784})}{P(X_{1:784})\,[Q_6]} \qquad z_1 \to \text{mean } z_1$$