Probabilistic Reasoning

2 main imp Rule

* Sum Rule ⇒ Sum of exclusive Prob = 1
* Product Rule ⇒ 5 dice Rolled simultaneously

↓

Conditional Prob

Prob $P(1, 2, 3, 4, 5)$

$$P(A \cap B) = P(A) \times P(A \mid B)$$

$$\frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} \times \frac{1}{6}$$

$$= \boxed{\frac{1}{7776}}$$

Graphical Networks

to represent probabilistic Distribution

* easy visualization
* properties like independence can be seen in the Graph
* Complex computation can be represented by Graphical manipulations

Node	⇒	~~Prob~~ Random Variable
Link	⇒	Probabilistic Relations

↓

Bayesian Network (OR) Directed Graphical models

Directed Graph → Bayesian Networks → usefull for Casual relationship btw R ~~B~~ Varia → ~~B~~ Random Variable

undirected Graph → Markov Random fields → Soft Constraints btw R variable

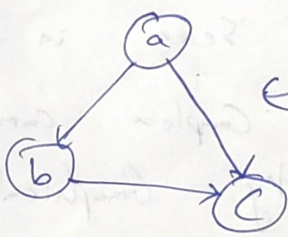
---

Factor Graph

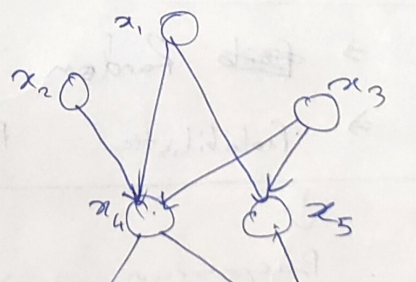
Probability distribution : $P(a,b,c) = P(c|a,b) \times P(a,b)$

Decomposition of Joint probability

$$= P(c|a,b) \cdot P(b|a) \cdot P(a)$$



Given a Graph : the decomposition of the joint probability is



$$P(x_4|x_1, x_2, x_3) \cdot P(x_5|x_1, x_3) \cdot P(x_7|x_4, x_5) \cdot P(x_6|x_4) \cdot P(x_1) P(x_2) P(x_3)$$

Child node $\rightarrow x_k$

Set of Parent nodes = Pa_k

Joint distribution $\rightarrow$ Product of all the nodes in the graph

thus it can be written as

K variables $\rightarrow$ K nodes

Joint distribution : $P(x_1, x_2, \ldots\ldots, x_K)$

$$\boxed{= P(x_K \mid x_1, \ldots\ldots x_{K-1}) \ldots P(x_2 \mid x_1) \cdot P(x_1)}$$

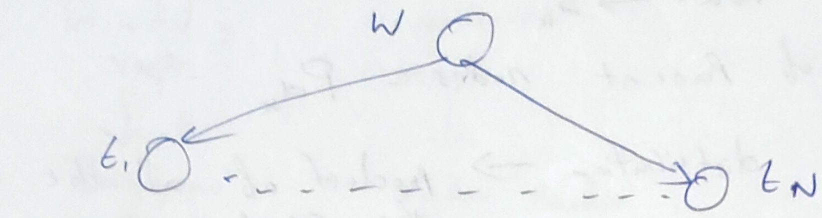
$$\Rightarrow \boxed{P(X) = \prod_{k=1}^{K} P(x_k \mid Pa_k)}$$

where

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_K \end{bmatrix}$$

↳ there are no directed cycles in said Bayesian Networks which makes them Directed Acyclic Graphs short for DAG

↳ implies that there exists some ordering of the nodes such that there are no links to go from any node to any lower numbered node



$\mathbf{t} = t_1 \ldots t_N \rightarrow$ N observed data

w - polynomial coefficients

input data $X = (x_1 \ldots x_N)$

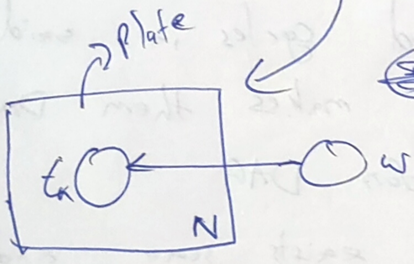
noise variance (σ^2)

hyperparameter α

$$P(\mathbf{t}, w) = P(w) \prod_{n=1}^{N} P(t_n | w)$$

↳ it is hard representing N nodes

thus new form

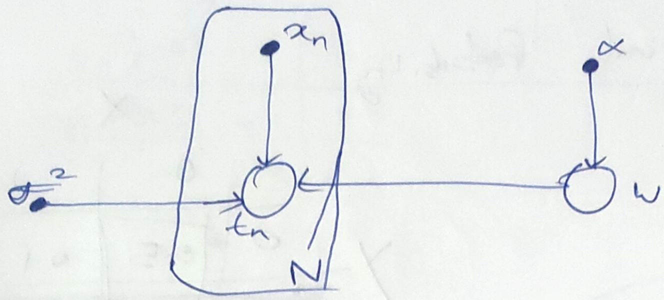


$$P(\mathbf{t}, w | X, \alpha, \sigma^2) = P(w|\alpha) \prod_{n=1}^{N} P(t_n | x_n, \sigma^2, w)$$

deterministic values

↳ X, σ^2, α not included in graph

thus new form



deterministic parameters

$\Rightarrow \bullet$

X, α, σ^2

Random variables

↳ ~~t~~ t, w

latent / hidden variable

↳ a variable which is not observed

latent hidden variable

~~Anc~~ Ancestral sampling

when $\prod P(x_i | pa\, x_i)$ a valid distribution, conditional independence

~~≡~~ $\perp\!\!\!\perp$

marginal distribution

markov probab...

Pg 365 till 378

X → has Disease

Y ↓ shows symptoms

Y \ X	0	1
0	0.5	0.1
1	0.1	0.3

Given Probs

Joint Probability

$P(X=0, Y=0) = 0.5$

Prob that Patient has no symptom & no disease

$P(X=1, Y=1) = 0.3$

$P(X=1, Y=0) = 0.1$

Marginal Probability

Y \ X	0	1	
0	0.5	0.1	0.6
1	0.1	0.3	0.4
	0.6	0.4	1

$P(Y=1) = 0.1 + 0.3$

$P(Y=1) = \underline{0.4}$

Marginal Prob is a subset of the original Prob dist, it loses some dimension to Produce new probabilities

$P(X=0) =$

$0.5 + 0.1 = 0.6$

$P(X=0) = 0.6$

Marginal prob

General

$$\boxed{P(X = x_k) = \sum_{i=1}^{k} P(x_1 = i, x_2 = i, \ldots, x_k = x_k)}$$

* Joint Dist is a Probability distribution

	X	0	1	
Y	0	0.5	0.1	0.6
	1	0.1	0.3	0.4
		0.6	0.4	1

~~P(x=0~~

$P(Y=0 \mid X=0) = \frac{0.5}{0.6} = 0.834$

0.6 → total of col

I know X is 0
but what is prob of $y=0$

$$\boxed{P(Y \mid X) = \frac{P(X, Y)}{P(X)}}$$

$P(X,Y)$ → Joint
$P(Y \mid X)$ → conditional
$P(X)$ → marginal

$$P(x \mid y) = \frac{P(x, y)}{P(y)}$$

$$P(x, y) = P(y|x)\,P(x)$$

Joint Distribution

$$P(a,b|c) = \frac{P(c, a,b)}{P(c)}$$

Conditional Independence → Pg 392 (pdf)
Pg 372 (book)

$$P(a,b,c) = P(c|a,b)\cdot P(b|a)\,P(a)$$

Joint distribution of a & b

= Product of marg.. of a
× Marg.. of b

$$P(a, b|c) = P(a|c)\,P(b|c)$$

$P(A \cap B) = P(A)\cdot P(B)$ ⇒ A & b are independant

Conditional independance

when a & b are said to be Condi.... Independan...

$$\boxed{a \perp\!\!\!\perp b \mid C}$$

given c

$$P(y|x) = \frac{P(x, y)}{P(x)}$$

$P(y|x)$ → Conditional

$P(x, y)$ → Joint

$P(x)$ → marginal

Bayes Rule

Lets say $P(X_1 \cdots\cdots X_n)$ → Joint

$$\boxed{P(x_1 \cdots\cdots x_n) = \prod_{i=1}^{N} P(x_i \,|\, P_a(x_i))}$$

→ Joint Distribution

Marginal Distribution

Lets say $P(x_1 \cdots\cdots x_n \neq x_j)$

means doesn't include x_j

$$\boxed{P(x_1 \cdots\cdots x_n \neq x_j) = \int P(x_1 \cdots\cdots x_n)\, dx_j}$$

the full term from x_1 to x_n including x_j (basically Joint Dist.)

basically integration dissolves x_j term, thus providing the marginal distribution

$$P(x_1 \ldots\ldots x_n \neq x_j) = \int P(x_1 \ldots\ldots x_n)\, dx_j$$

$$= \boxed{\int \prod_{i=1}^{N} P(x_i \mid Pa(x_i))\, dx_j}$$

~~Form~~ Shortened form