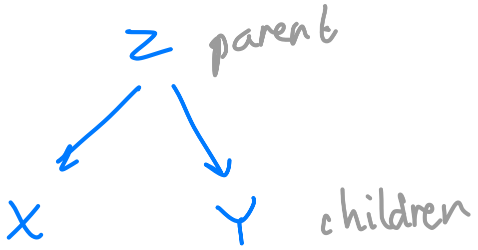
11/8/22

Conditional Independence

$X \perp Y | Z$

(1) Shared cause



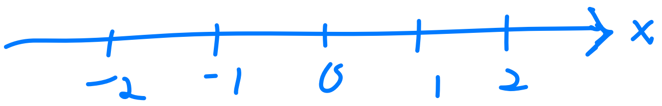
(2) Markovian

$$X \to Z \to Y$$

past / remote cause — present / immediate cause — future

markov chain/random walk

$$X_0 \to X_1 \to X_2 \to \dots \to X_t \to X_{t+1} \to \dots$$



Markovian property:

$$[X_{t+1} | X_t, X_{t-1}, \dots, X_0] \sim [X_{t+1} | X_t]$$

└ X_{t+1} only depends on previous step

$$X_{t+1} = F(X_t, \varepsilon_t)$$

$$\varepsilon_t \perp (X_0, \dots X_t)$$

joint density :

$$P(X_0, X_1, \dots, X_T) = p(x_0) p(x_1 | x_0) p(x_2 | x_1) \dots$$

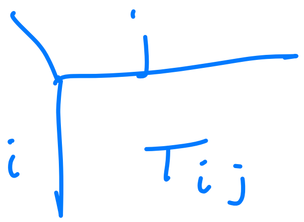
$$= p(x_0) \prod_{t=1}^{T} p(x_t | x_{t-1})$$ notation 1

Transition probability :

$$T(\underset{\substack{\downarrow \\ i}}{x}, \underset{\substack{\downarrow \\ j}}{y}) = P(X_{t+1} = \underset{\substack{\downarrow \\ j}}{y} | X_t = \underset{\substack{\downarrow \\ i}}{x})$$

notation 2

Notation 3



Recursive computation

notation 3

$$P_t(i) = P(X_t = i)$$

population migration interpretation

$P_t(i)$: # of people @ state i @ time t

T_{ij} : fraction of people in state i who will go to state j

notation 1

$$P(X_{t+1}) = \sum_{X_t} P(X_t, X_{t+1}) \longrightarrow \text{marginalization}$$

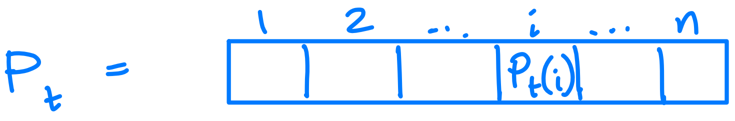
$$= \sum_{X_t} P(X_t) P(X_{t+1} | X_t) \longrightarrow \text{factorization}$$

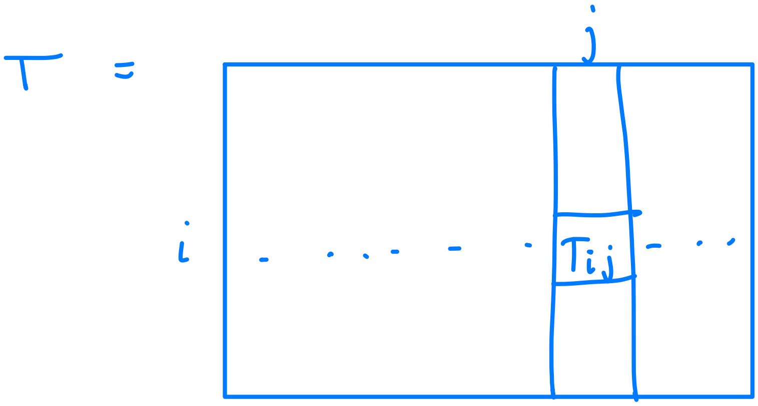
notation 2

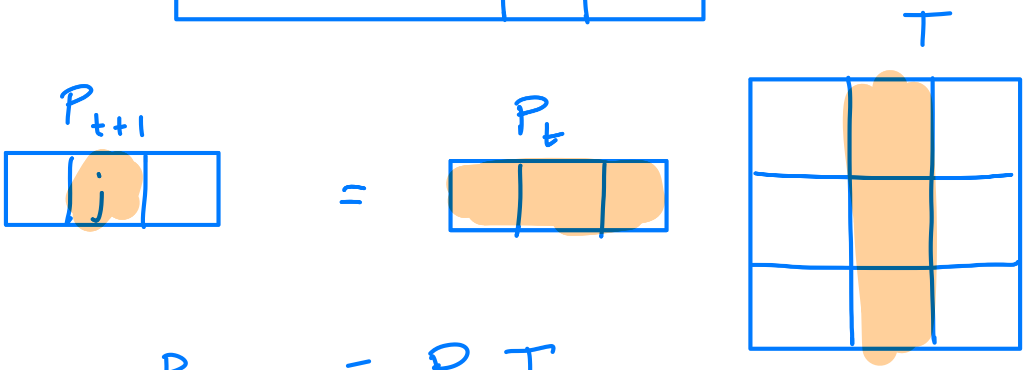
$$P_{t+1}(y) = \sum_x P_t(x)\, T(x,y) \quad \text{or} \quad p^{(t+1)}(y) = \sum_x p^{(t)}(x)\, T(x,y)$$

notation 3

$$P_{t+1}(j) = \sum_i P_t(i) \cdot T_{ij} \quad \text{or} \quad p^{(t+1)}_j = \sum_i p^{(t)}_i T_{ij}$$



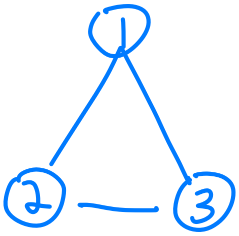




markov matrix

Each row sum = 1

$$P_{t+1} = P_t T$$

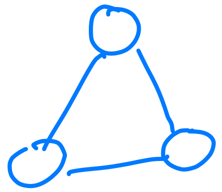


pi
stationary distribution

$$P_t = P_0 T^t \longrightarrow \pi$$

ex.) if after diagonalization $\begin{bmatrix} 1 & & \\ & .99 & \\ & & -.8 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & & \\ & 0 & \\ & & 0 \end{bmatrix}$

ex.)



with prob $\frac{1}{2}$ stay

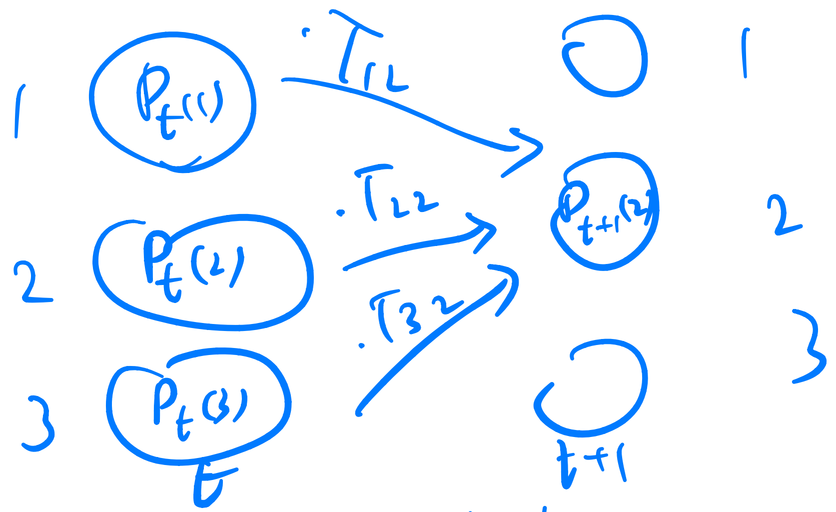
with prob $\frac{1}{4}$ go to one of the remaining 2 states,

$$T = \begin{array}{c|ccc} & 1 & 2 & 3 \\ \hline 1 & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 2 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 3 & \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{array}$$

Intepret $P_{t+1}(j) = \sum_i P_t(i)\, T_{ij}$:

$P_t(i)$ ↓ # people in state i

T_{ij} ↓ fraction that go to j



Multivariate Normal / Gaussian

$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim N\left(0,\ \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}\right)$$

symmetric positive definite

Σ_{21}: $cov(x_2, x_1)$

Σ_{22}: $cov(x_2, x_2) = Var(x_2)$

$$[x_2|x_1] \sim N\Big(\underset{E(x_2|x_1)}{\Sigma_{21}\Sigma_{11}^{-1}x_1},\ \underset{Var(x_2|x_1)}{\Sigma_{22}-\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}}\Big)$$

proof :

$$\varepsilon = x_2 - A x_1$$

$$x_2 = \underset{\overset{\downarrow}{\text{regression coefficient}}}{A}\, x_1 + \underset{\overset{\downarrow}{\text{regression error}}}{\varepsilon} \longrightarrow \text{like regression}$$

we want x_1 to be independent of ε

$$\operatorname{cov}(\varepsilon, x_1) = 0$$

$$\operatorname{cov}(x_2 - Ax_1, x_1) = \operatorname{cov}(x_2, x_1) - \operatorname{cov}(Ax_1, x_1)$$

$$= \Sigma_{21} - A\Sigma_{11} = 0$$

$$A = \Sigma_{21}\Sigma_{11}^{-1}$$

$$\operatorname{var}(\varepsilon) = \operatorname{cov}(\varepsilon, \varepsilon) = \operatorname{cov}(x_2 - Ax_1, \varepsilon)$$

$$= \operatorname{cov}(x_2, \varepsilon) - \underset{0}{\cancel{\operatorname{cov}(AX_1, \varepsilon)}}$$

$$= \operatorname{cov}(x_2, x_2 - Ax_1)$$

$$= \Sigma_{22} - \operatorname{cov}(x_2, Ax_1)$$

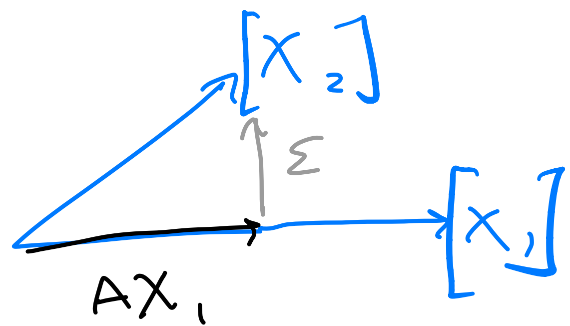
$$= \Sigma_{22} - \operatorname{cov}(x_2, x_1)A^T$$

$$= \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$$

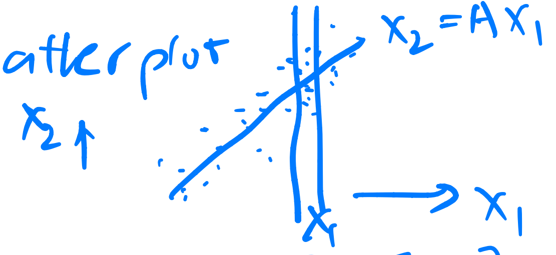
data perspective

	$[X_1]$	$[X_2]$
1		
⋮	x_{i1}	x_{12}
n		

vector perspective



Scatterplot



$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \to \begin{bmatrix} X_1 \\ \varepsilon \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 - AX_1 \end{bmatrix} = \begin{bmatrix} I & 0 \\ -A & I \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$

$$\mathrm{Var}\begin{bmatrix} X_1 \\ \varepsilon \end{bmatrix} = \begin{bmatrix} I & 0 \\ -A & I \end{bmatrix} \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \begin{bmatrix} I & -AI \\ 0 & I \end{bmatrix}$$

$$= \begin{bmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \end{bmatrix}$$

HW* Still need to prove that transformation $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ is still normal.

↳ use change of variable & Jacobian

recall:

$\mathrm{var}(X) = \Sigma$

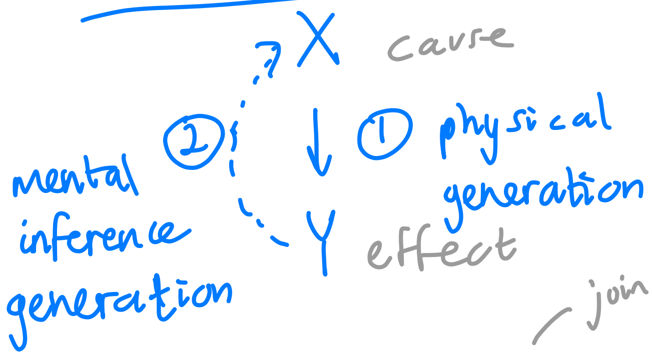
$\mathrm{var}(AX) = A\Sigma A^T$

$X \sim N(0, \Sigma)$

$Y = AX \sim N(0, A\Sigma A^T)$

$$① \xrightarrow[\text{use Jacobian}]{} ②$$

Bayes' Rule



prior belief

prior : $X \sim p(x)$

generation : $[Y \mid X = x] \sim p(y|x)$

posterior : $[X \mid Y = y] \sim p(x|y)$

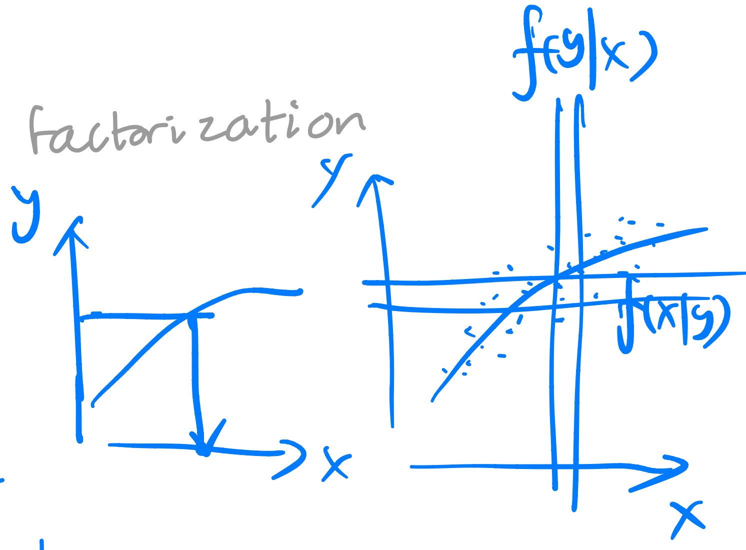
updated belief

joint

$$p(x|y) = \frac{p(x,y)}{p(y)} \longrightarrow \text{conditioning/normalization}$$

$$= \frac{p(x,y)}{\sum_x p(x,y)} \longrightarrow \text{marginalization}$$

$$= \frac{p(x)\,p(y|x)}{\sum_{x'} p(x')\,p(y|x')} \longrightarrow \text{factorization}$$



ex.) rare disease example

X : discrete $\qquad Y$: discrete

$X \in \{0, 1\}$

0 → no disease, 1 → have disease

x	0	1
p(x)	99%	1%

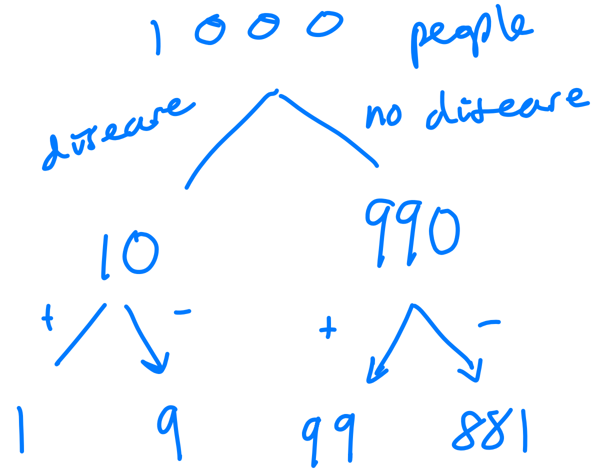
$Y \in \{+, -\}$ test for disease

y \ x	0	1
+	10%	90%
−	90%	10%

false positive (10%) false negative (10%)

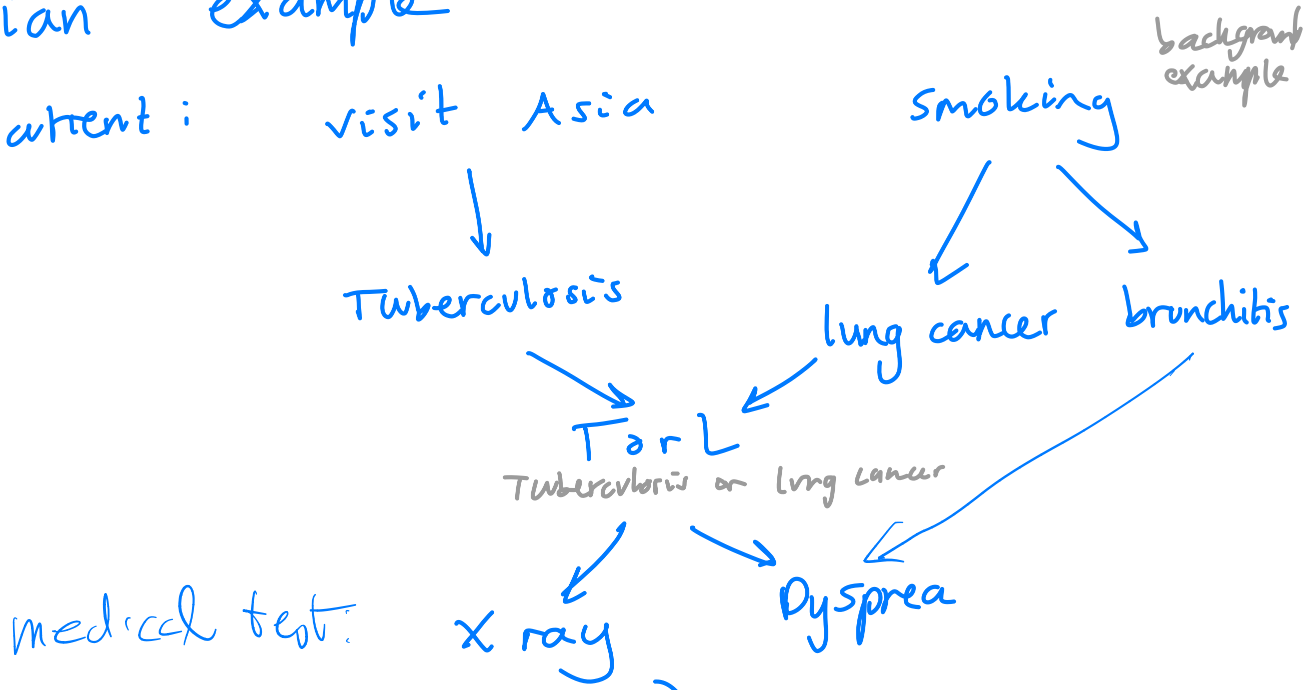
$$p(\overset{\overset{1}{\uparrow}}{x} \mid \overset{\overset{+}{\uparrow}}{y}) = \frac{1\% \; 90\%}{99\% \cdot 10\% + 1\% \cdot 90\%} = \frac{90}{990 + 90}$$

$$x' \in \{0, 1\} \qquad = \frac{1}{12}$$



ex.) Asian example

ask patient :



$$P(T=1 \mid X=1, D=0, V=1, S=0)$$

probability that patient has tuberculosis given that they have a positive X ray, no dysprea, visit asia, non smoking

$$P(v, s, t, l, b, o, x, d) = p(v)p(s)p(t|v)p(l|s)p(b|s)p(o|t,l)$$
$$\cdot p(x|o)p(d|o,b)$$

$\downarrow$ marginal

factorization

$$P(T=1 | X=1, D=0, V=1, S=0) = p(t|x,d,v,s)$$
$$= \frac{p(t,x,d,v,s)}{p(x,d,v,s)}$$

conditioning