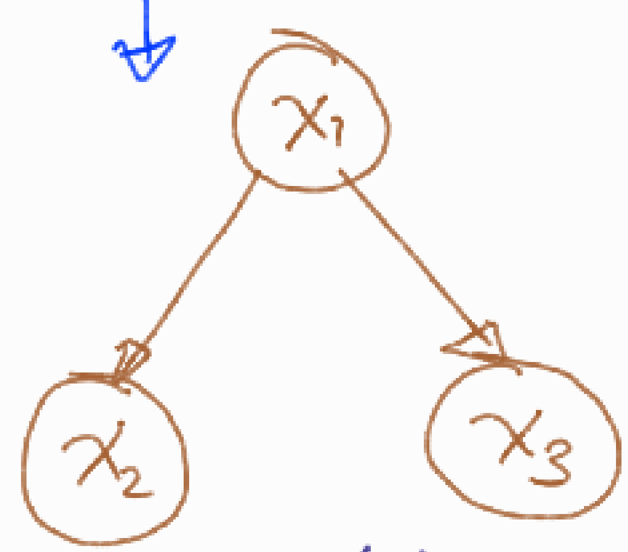
- Unsupervised Learning
- Expectation Maximization Algorithm
- Case of missing data / hidden data (Bayesian Network)

The story so far ......

x_1	x_2	x_3
1	1	0
0	1	1
1	0	0
...	...	...

Learn θ from data X $\Longrightarrow$

Design Network



parameters

$$\left.\begin{array}{l} P(x_1) \\ P(x_2/x_1) \\ P(x_3/x_1) \end{array}\right\} \theta$$

$$P(x_1, x_2, x_3) = P(x_2/x_1) \cdot P(x_3/x_1) \cdot P(x_1)$$

Generate X from θ

x_1	x_2	x_3
1	0	1
1	1	0

What if the data is not complete?

Case - I

x_1	x_2	x_3
1	?	1
1	0	1
0	1	1
0	?	0

x_2 has missing values

Case - II

x_1 & x_3 both have missing values

x_1	x_2	x_3
?	1	?
0	0	?
?	0	1
1	1	0

Missing Data

Case - III

	x_1	x_2	x_3	
D_1	1	0	?	(1)
D_2	0	1	?	(0)
D_3	0	0	?	(1)

class label?

Unsupervised Learning

(Hidden / Latent variable)

* Parameter Estimation from incomplete data

θ ? $\quad P_\theta(X, H)$

Hidden / Latent

observed

* Predict value for missing data / hidden variable

$P(H / X, \theta)$?

Find out H given X & θ

Maximum Likelihood Estimation

$$\theta_{MLE} \approx \underset{\theta}{\operatorname{argmax}}\ P_\theta(X, H)$$

Iterate over θ & "Find out θ that best describes distr. (X, H)"

Expectation Maximization Algorithm (EM Algo.)

1 - Initialize θ_0 (at random)

- For $t = 0, 1, 2 \ldots$ (until convergence)

2 - Calculate "Expected" value of hidden variables.

E-step: $P(H / X, \theta_i)$ ← Calculate H from assumed θ_i

3 - Re-estimate value of θ

M-step: $\underset{\theta}{\operatorname{argmax}}\ P_\theta(X, H)$

- Recalculate/update θ using H values calculated in E-step.

Example:

$$X_1 \to X_2 \to X_3$$

$$P(X_1, X_2, X_3) = \underbrace{P(X_3 / X_2) \cdot P(X_2 / X_1) \cdot P(X_1)}_{\theta}$$

	X_1	X_2	X_3 (H)
e_1	1	1	~~?~~ 1
e_2	0	1	~~?~~ 0
e_3	1	0	~~?~~ 0
e_4	0	0	~~?~~ 1

Assumed (Finish Computing & find exact)

Step: 1 Initialize θ_0

X_1	
1	.8
0	.2

$X_1 \backslash X_2$	0	1
0	0.4	0.6
1	0.7	0.3

$X_2 \backslash X_3$	0	1
0	0.6	0.4
1	0.2	0.8

CPT

Step-2 E-step : Calculate Exepected value of X_3

$$P(X_3 / X_1, X_2, \theta_0) = \frac{P(X_1, X_2, X_3)}{P(X_1, X_2)}$$

$$\boxed{e_1 \quad x_1 = 1, \quad x_2 = 1, \quad x_3 = ?}$$

$$P(x_3=1 / x_1=1, x_2=1) = \frac{P(x_1=1, x_2=1, x_3=1)}{P(x_1=1, x_2=1, x_3=1) + P(x_1=1, x_2=1, x_3=0)} \quad A$$

$$P(x_3=0 / x_1=1, x_2=1) = \frac{P(x_1=1, x_2=1, x_3=0)}{P(x_1=1, x_2=1, x_3=1) + P(x_1=1, x_2=1, x_3=0)} \quad B$$

$$A = P(x_3=1/x_2=1) \cdot P(x_2=1/x_1=1) \cdot P(x_1=1)$$
$$= 0.8 \times 0.3 \times 0.8 = 0.192$$

$$B = P(x_3=0/x_2=1) \cdot P(x_2=1/x_1=1) \cdot P(x_1)$$
$$= 0.048$$

$$P(x_3=0/x_1=1, x_2=1) = \frac{0.048}{0.192+0.048} = 0.2$$

$$P(x_3=1/x_1=1, x_2=1) = \frac{0.192}{0.192+0.048} = 0.8$$

$$e_2 = \quad x_1 = 0, \quad x_2 = 1, \quad x_3 = ?$$

$$P(x_3=1/x_2=1) \cdot P(x_2=1/x_1=0) \cdot P(x_1=0)$$
$$? \; (0.8)(0.6) = 0.096$$

$$P(x_3=0/x_2=1) \cdot P(x_2=1/x_1=0) \cdot P(x_1=0)$$
$$? \quad =$$

$$e_3 : \quad x_1 = 1, \quad x_2 = 0, \quad x_3 = ?$$

$$P(x_3=1/x_2=0) \cdot P(x_2=0/x_1=1) \cdot P(x_1=1)$$

$$P(x_3=0/x_2=0) \cdot P(x_2=0/x_1=1) \cdot P(x_1=1)$$

e_4 : $x_1 = 0, \; x_2 = 0, \; x_3 = ?$

$$P(x_3 = 1 / x_2 = 0) \cdot P(x_2 = 0 / x_1 = 0) \cdot P(x_1 = 0)$$

$$P(x_3 = 0 / x_2 = 0) \cdot P(x_2 = 0 / x_1 = 0) \cdot P(x_1 = 0)$$

Step 3: Maximization: Recalculate $\underline{\underline{\theta_1}}$

$$P(x_1) = 1/2$$

$$P(x_2 / x_1) = P(x_2 = 1 / x_1 = 1) = 1/2$$

$$P(x_3 / x_2) = P(x_3 = 1 / x_2 = 1) = 1/2$$

Repeat 2 & 3,