19/9/19 EE5801: Representation Learning

- Quiz review
- Review
- Exp. Maximization

- Q1: $X = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 \end{bmatrix}$

Since the observations are not zero mean, subtract mean from each row (each row's corresponding mean)

$$\hat{X} = \begin{bmatrix} -2 & -1 & 0 & 1 & 2 \\ 2 & 1 & 0 & -1 & -2 \end{bmatrix}$$

$$C_{\hat{X}\hat{X}} = \frac{1}{5}\begin{bmatrix} 10 & -10 \\ -10 & 10 \end{bmatrix} = \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}$$

$P = E^T$ where $C_{\hat{X}\hat{X}} = E\,D\,.E^T$. (see previous notes).

2a. $$f_X(x; x_0, \theta) = \begin{cases} \theta \cdot x_0^{\theta} \, x^{-\theta-1} & x > x_0 ;\ \theta > 0 \\ 0 & , \quad \text{else} \end{cases}$$

$$L(x; x_0, \theta) = \prod_{i=1}^{N} f_X(x_i; x_0, \theta)$$

$$\log(L(x; x_0, \theta)) = \sum_{i=1}^{N} \log\Big[\underbrace{f_X(x; x_0, \theta)}\Big]$$

$$= \sum_{i=1}^{N} \log\left[\theta \cdot x_0^{\theta} \cdot x_i^{-\theta-1}\right] \qquad \left[\begin{array}{l}\text{For the case}\\ x_i > x_0\end{array}\right]$$

$$= \sum_{i=1}^{N} \left[\log\theta + \theta \cdot \log x_0 - (\theta+1) \cdot \log x_i\right]$$

$\theta_{ML} = \underset{\theta}{\operatorname{argmax}} \log\left[L(\underline{x}; x_0, \theta)\right]$. To find this

Set $\dfrac{\partial \log\left[L(x; x_0, \theta)\right]}{\partial\theta} = 0$

$$= \frac{N}{\theta} + N \cdot \log x_0 - \sum_{i=1}^{N} \log x_i$$

$$\Rightarrow \frac{N}{\theta_{ML}} = \sum_{i=1}^{N} \log x_i - N \log x_0$$

$$\boxed{\theta_{ML} = \frac{N}{\sum_{i=1}^{N} \log x_i - N \log x_0}}$$

- Review ML: Gaussian RV where $f_X(x;\theta) = \dfrac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left[-\dfrac{(x-\mu)^2}{2\sigma^2}\right]$

$$\log\left[L(X:\theta)\right] = \sum_{i=1}^{N} \log f_X(x;\theta)$$

$$= \sum_{i=1}^{N}\left[\log\frac{1}{\sqrt{2\pi}\sigma} - \frac{(x_i-\mu)^2}{2\sigma^2}\right]$$

- Mixture models: Gaussian Mixture Models

$$p(x) = \sum_{k=1}^{K} \pi_k \mathcal{N}(\underline{x}; \underline{\mu}_k, \Sigma_k)$$

π_k : mixture weights

$\underline{\mu}_k$: mean vector of the k^{th} Gaussian component

Σ_k : covariance matrix of the k^{th} Gaussian component

Ex of latent variables:

P1: Given 2 biased coins A, B. Find the bias on each of them empirically.

Soln: Set1: $[x_1 \dots x_{10}]_{z_1}$ $z \in \{A, B\}$

$[x_1 \dots x_{10}]_{z_2}$ $x \in \{H, T\}$

$\vdots$

Set5. $[x_1 \dots x_{10}]_{z_5}$

Let p_A be the bias on A, p_B be the bias on B

$$p_A = \frac{\#\text{ heads when A is flipped}}{\#\text{ times A is flipped}}$$

$$p_B = \frac{\#\text{ heads when B is flipped}}{\#\text{ time B is flipped}}$$

P2: Find p_A, p_B when z is hidden or unknown

Soln: Since p_A, p_B are unknown or cannot be empirically found, assume initial values $p_A^{(0)}, p_B^{(0)}$ and iterate until convergence