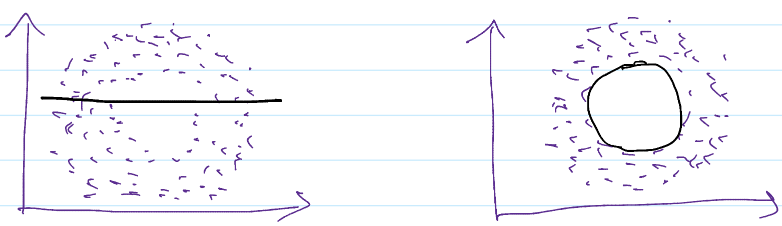
HWO (EE19MTECH01008)

Friday, 20 September 2019 3:42 PM

Q2) Find examples where PCA fails?

Ans:- • Main job of PCA is to represent data in lower dimension by removing redundant features. It achieves that through finding orthogonal principal components.
This is not possible if the joint distribution of data follows other distribution instead of multivariate Gaussian.
We also use covariance matrix to find principal components.
The only distribution to allow us to represent whole data in compact form is Gaussian distribution.
So, PCA makes implicit assumption that data should follow Gaussian distribution.

• Standard PCA finds linear principal components to represent the data in lower dimension.



PCA Kernel PCA

Standard PCA can not good representative direction
Kernal PCA can overcome standard PCA limitations

Q3) Find maximum likelihood of the following distributions.

a) Binomial,

x is total number of success & x_i is single trial

$$\prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} = p^{\sum_i^n x_i}(1-p)^{\sum_i^n 1-x_i} = p^{x}(1-p)^{n-x}$$

We find the log-likelihood,

$$\ln({}^{n}C_{x}\, p^{x}(1-p)^{n-x}) = \ln({}^{n}C_{x}) + x\ln(p) + (n-x)\ln(1-p)$$

Take derivative & put it to 0,

$$\sum_{i=1}^{n} \frac{d}{dp} \ln({}^{n}C_{x}) + x\ln(p) + (n-x)\ln(1-p) = \frac{x}{p} - \frac{n-x}{1-p} = 0$$

$$\Rightarrow \sum_{i=1}^{n} \frac{n}{x} = \frac{1}{p}$$

$$\Rightarrow \boxed{p_{MLE} = \sum_{i=1}^{n} \frac{x}{n} = \hat{X}}$$

b) Poisson,

$$f(x;\lambda) = \frac{\lambda^{x} e^{-\lambda}}{x!}$$

$$L(x,\theta) = \prod_{i=1}^{N} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}$$

$$\log(L(x;\theta)) = \sum_{i=1}^{n} [x_i \log\lambda - \lambda - \log(x_i!)]$$

$$\frac{d}{d\lambda} \log(L(x;\theta)) = 0$$

$$\Rightarrow \frac{1}{\lambda} \sum_{i=1}^{N} x_i - N = 0$$

$$\Rightarrow \boxed{\lambda_{MLE} = \frac{1}{N} \sum_{i=1}^{N} x_i}$$

c) Exponential,

$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x_i}, & x_i \geq 0 \end{cases}$$

$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x}, & x_i \geq 0 \\ 0, & \text{else} \end{cases}$$

N observations are positive

$$\Rightarrow \mathcal{L}(x,\lambda) = \prod_{i=1}^{N} \lambda e^{-\lambda x_i}$$

$$\ln[\mathcal{L}(x,\lambda)] = \sum_{i=1}^{N} [\ln\lambda - \lambda x_i]$$

$$\ln[\mathcal{L}(x,\lambda)] = N\ln\lambda - \lambda\sum_{i=1}^{N} x_i$$

$$\lambda_{MLE}, \frac{\partial}{\partial\lambda}\left(\ln[\mathcal{L}(x,\lambda)]\right) = 0$$

$$\Rightarrow \frac{N}{\lambda_{MLE}} - \sum_{i=1}^{N} x_i = 0$$

$$\Rightarrow \boxed{\lambda_{MLE} = \frac{N}{\sum_{i=1}^{N} x_i}}$$

d) Gaussian,

Let $x_i \sim \mathcal{N}(\mu, \sigma)$

$$\mathcal{L}(\vec{x},\theta) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(x_i-\mu)^2}{2\sigma^2}\right)$$

$$\log(\mathcal{L}(\vec{x},\mu,\sigma)) = \frac{-N}{2}\log(2\pi\sigma^2) + \sum_{i=1}^{N} \frac{-(x_i-\mu)^2}{2\sigma^2}$$

Assuming multivariate sample,

For maximum likelihood, $\frac{\partial}{\partial \mu} \log(L)$ & $\frac{\partial}{\partial \sigma} \log(L) = 0$

$$\frac{\partial}{\partial \mu}\left(\log\left(L(\vec{x}, \mu, \sigma)\right)\right) = \cancel{2} \sum_{i=1}^{N} \left(\frac{\vec{x}^{(i)} - \mu}{\cancel{2}\sigma^2}\right) = 0$$

$$\Rightarrow \boxed{\mu_{MLE} = \frac{1}{N} \sum_{i=1}^{N} \vec{x}^{(i)}}$$

$$\frac{\partial}{\partial \sigma}\left(L(x, \mu, \sigma)\right) = \frac{\partial}{\partial \sigma}\left[-\frac{N}{2} \log(2\pi\sigma^2) - \sum_{i=1}^{N}\left(\frac{(x-\mu)^2}{2\sigma^2}\right)\right] = 0$$

$$\Rightarrow \frac{N}{2\pi} \sigma^2 = \sum_{i=1}^{N} (x-\mu)^2$$

$$\Rightarrow \sigma^2_{MLE} = \frac{1}{N} \sum_{i=1}^{N} (x-\mu)^2$$

e) Laplacian,

$$f(x; \mu, \beta) = \frac{1}{2\beta} e^{-\left\{\frac{|x-\mu|}{\beta}\right\}}$$

$$L(x; \mu, \beta) = \prod_{i=1}^{N} \frac{1}{2\beta} \exp\left\{-\frac{|x_i - \mu|}{\beta}\right\}$$

$$\ln\left(L(x; \mu, \beta)\right) = -N \ln 2\beta - \sum_{i=1}^{N} \frac{|x_i - \mu|}{\beta}$$

$$\frac{\partial}{\partial \beta} \ln(L) = 0$$

$$\Rightarrow \frac{\partial}{\partial \beta}\left[-N \ln 2\beta - \sum_{i=1}^{N} \frac{|x_i - \mu|}{\beta}\right] = 0$$

$$\Rightarrow \left[-\frac{N}{2\beta}(2) + \frac{1}{\beta^2} \sum_{i=1}^{N} |x_i - \mu|\right] = 0$$

$$\Rightarrow \left[-\frac{N}{2\beta}(2) + \frac{1}{\beta^2} \sum_{i=1}^{N} |x_i - \mu| \right] = 0$$

$$\Rightarrow N\beta = \sum_{i=1}^{N} |x_i - \mu|$$

$$\Rightarrow \boxed{\beta_{MLE} = \frac{\sum_{i=1}^{N} |x_i - \mu|}{N}}$$

$$\frac{\partial}{\partial \mu}(L) = 0$$

$$\Rightarrow \frac{\partial}{\partial \mu} \left[-N \ln(2\beta) - \sum_{i=1}^{N} \frac{|x_i - \mu|}{\beta} \right] = 0$$

$$\Rightarrow -\cancel{\frac{1}{\beta}} \sum_{i=1}^{N} \frac{|x_i - \mu|}{x_i - \mu} = 0$$

$$\Rightarrow \sum_{i=1}^{N} \operatorname{sgn}(x_i - \mu) = 0$$

$$\Rightarrow \sum \operatorname{sgn}(x) = 0$$

Equal samples of x below and above zero

μ is middle value among collection of data

$$\boxed{\mu = \text{median}(x_i)}$$