MLDL Homework-2

1) Assuming ~~input data~~ training data is:

$(x_1 = 2, y_1 = 8), (x_2 = 5, y_2 = 25)$
$(x_3 = 10, y_3 = 40)$

Let us consider that we use the least square to fit the data.

1) Linear Regression

a) overall Loss function:

$$y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} \quad X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ 1 & x_4 \end{bmatrix}$$

$$y = X\beta + \epsilon$$

$$e(\beta) = y - X\beta$$

loss function:

$$= e^T e$$

$$= (y - X\beta)^T (y - X\beta)$$

$$= (y^T - \beta^T X^T)(y - X\beta)$$

$$= y^T y - 2\beta^T X^T y + \beta^T X^T X \beta$$

b) Closed form normal equation:

$$\Rightarrow x^T x\beta - x^T y = 0$$

$$\therefore \hat{\beta} = (x^T x)^{-1} x^T y$$

here $y = \begin{bmatrix} 8 \\ 25 \\ 9 \\ 40 \end{bmatrix}$; $x = \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 3 \\ 1 & 10 \end{bmatrix}$

c) Closed form equation:

This closed form equation is used to obtain the ϕ value so that value can be used in the prediction for the implementation of linear regression

it is connected to project operation

ie: $\hat{y} = x\hat{\theta}$

$$\text{proj}_x(y) = x(x^T x)^{-1} x^T y$$

Now, $x(x^T x)^{-1} x^T = UU^T$

$$x(x^T x^{-1}) x^T y = (UU^T) y$$

$$= u_1 u_1^T y + \cdots + u_d u_d^T y.$$

2) PCA) b)

"X' be a data matrix of size $n \times d$

covariance matrix C

$$C = X^T X / (n-1)$$

It is a symmetric matrix It can be diagonalized

$C = ULU^T$ ("L" is a diagonal matrix with eigen values λ)

Let $X = USV^T$

U = Left singular vector matrix

V = Right singular vector matrix

S = diagonal matrix (or) singular values S_i

$$C = (VSU^T)(USV^T)/(n-1)$$

$$= V \cdot \frac{S^2}{(n-1)} \cdot V^T$$

Hence eigen values are given as $\lambda_i = \frac{s_i^2}{(n-1)}$

Scanned with CamScanner

3) a) for multivariable gausion

a) multivariable gaussion Distribution

$$p(x; \mu, \Sigma)$$

$$= \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp\left(-\tfrac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right)$$

b) Consider

$Y = ax + b$ when $X \sim N(\mu, \Sigma)$

$$p(Y) = \mu(a\mu + b, b\Sigma b^T)$$

c) For n independent samples $x_1, \dots x_n$

where $p(x_n) = \frac{1}{\sqrt{2\pi r^2}} \exp\left(-\frac{1}{2\sigma^2}(x_i - \mu)^2\right)$

$i(x_1, x_2 \dots x_n)$

$$= P(x_1) \cdot P(x_2) \cdots P(x_n)$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left(\frac{-1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2\right)$$

MLE estimation is given as

$$\mu_{MLE} = \max\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left[\frac{-1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2\right]$$

$$= \max_{\mu} = \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2$$

$$= \min \sum_{i=1}^{n} (x_i - \mu)^2$$

Differentiating w.r.t μ and equating to 0

we get

$$\sum_{i=1}^{n} 2(\mu_i - \mu) = 0$$

$$= \mu = \frac{1}{n} \sum_{i=1}^{n} x_i$$

Scanned with CamScanner