Machine Learning - Homework Week 2

Nguyen Nhu Quynh

October 2021

1 Multivariate Gaussian distribution

For a D-dimensional vector x, the multivariate Gaussian distribution takes the form

$$p(x|\mu,\sigma^2) = \frac{1}{(2\pi)^2|\Sigma|^1/2}exp(-\frac{1}{2}(x-\mu)^T|\Sigma|^{-1}(x-\mu))$$

where μ is a D-dimensional mean of vector, Σ is a DxD covariance matrix, $|\Sigma|$ denotes the determinant of Σ

Set $\triangle^2 = (x-\mu)^T|\Sigma|^{-1}(x-\mu)$ (1)

The eigenvector equation for the covariance matrix $\Sigma u_i = \lambda_i u_i$ where i = 1,...D

Because Σ is symmetric $\Rightarrow$ Its eigenvalues is real and its eigenvectors have form of orthonormal set is $u_i^T u_j = I_{ij}$ where I_{ij} is the i,j element of identity matrix

The covariance matrix Σ can be expressed as an expansion in terms of its eigenvectors

$$\Sigma = \Sigma_{i=1}^D \lambda_i u_i u_i^T$$

$$\Rightarrow \Sigma^{-1} = \Sigma_{i=1}^D \frac{1}{\lambda_i} u_i u_i^T \ (2)$$

(1)(2) $\Rightarrow \triangle^2 = \Sigma_{i=1}^D \frac{y_i^2}{\lambda_i}$ with $y_i = u_i^T(x-\mu)$

Because the determinant of a matrix is equal to the product of its eigenvalues $\Rightarrow |\Sigma|^{1/2} = \Pi_{j=1}^D \lambda_j^{1/2}$

Then, the Multivariate Gaussian distribution can be written as:

$$p(y) = \Pi_{j=1}^D \frac{1}{(2\pi\lambda_j)^{1/2}} exp(\frac{y_i^2}{2\lambda_j})$$

$$\Rightarrow \int_{-\infty}^{\infty} p(y)dy = \Pi_{j=1}^D \frac{1}{(2\pi\lambda_j)^{1/2}} exp(\frac{y_i^2}{2\lambda_j})\, dy_j$$

The right equation has the form of Univariate Gaussian distribution and Univariate Gaussian distribution is normalized

$$\Rightarrow \int_{-\infty}^{\infty} p(y)dy = \Pi_{j=1}^D \frac{1}{(2\pi\lambda_j)^{1/2}} exp(\frac{y_i^2}{2\lambda_j})\, dy_j = 1$$

Hence, Multivariate Gaussian distribution is normalized.

2 Conditional Gaussian distribution

$\mu_{ab} = \mu_a + \Sigma_{ab}\Sigma_{bb}^{-1}(x_b - \mu_b)$

$\Sigma_{a|b} = \Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba}$

$\Rightarrow p(x_a|x_b) = N(x_{a|b}|\mu_{a|b}, \Sigma_{a|b})$

3 Marginal Gaussian distribution

$E[x_a] = \mu_a$

$cov[x_a] = \Sigma_{aa}$

$\Rightarrow p(x_a) = N(x_a|\mu_a, \Sigma_{aa})$