CDCI 567-Machine Learning Assignment-1

1 Question-1

1.1 MLE

1.1.1 Beta(α,1)

$x_1, x_2, ..., x_n$ are samples of $Beta(\alpha, 1)$ distribution.

In this case, pdf of $x_(i), i = 1, ..., n$ is

$$f(x_i|\alpha, 1)) = \frac{x^{\alpha-1}}{\beta(\alpha, 1)} = \frac{x^{\alpha-1}}{\int_0^1 t^{\alpha-1} dt} = \alpha x^{\alpha-1}$$

Likelihood function and log likelihood function:

$$L(\alpha, 1|x) = \prod_{i=1}^{n} \alpha x^{\alpha-1}$$

$$l(\alpha, 1|x) = log(\alpha^n) + log(\prod_{i=1}^{n} x_i^{\alpha-1}) = nlog(\alpha) + \sum_{i=1}^{n} (\alpha - 1) log(x_i)$$

Derivative of $l(\alpha, 1|x)$:

$$\frac{\partial l(\alpha, 1|x)}{\partial \alpha} = \frac{n}{\alpha} + \sum_{i=1}^{n} log(x_i) = 0$$

$$\frac{\partial^2 l(\alpha, 1|x)}{\partial \alpha} = -\frac{n}{\alpha^2} < 0$$

$$\hat{\alpha} = -\frac{n}{\sum_{i=1}^{n} log(x_i}$$

1.1.2 Normal(θ,diag(θ))

$x_1, x_2, ..., x_n$ are samples of $Normal(\theta, diag(\theta))$ distribution. In this case, pdf of $x_i, i = 1, ..., n$ is

$$f(x_i|\theta, diag(\theta)) = \frac{1}{\sqrt{2\pi diag(\theta)}} \exp(-\frac{(x_i - \theta)}{2\theta})$$

Likelihood function and log likelihood function:

$$L(\theta, diag(\theta)|x_i) = (2\pi)^{-\frac{d}{2}} |diag(\theta)|^{-\frac{1}{2}} \exp(-\frac{1}{2}(x - \theta)^T diag(\theta)^{-1}(x - \theta))$$

$$|diag(\theta)| = \theta^d$$

$$l(\theta, diag(\theta)|x_i) = -\frac{d}{2}log(2\pi\theta) - \sum_{i=1}^{n}\frac{1}{2}\frac{(x_i-\theta_i)^T(x_i-\theta_i)}{\theta_i}$$

Derivative of $l(\theta, diag(\theta)|x_i)$:

$$\frac{\partial l(\theta, diag(\theta)|x_i)}{\partial\theta} = -\frac{dn}{2\theta} - \frac{n}{2} + \frac{1}{2\theta^2}$$