THEORY ASSIGNMENT

Ans 4.

For uniform distributed random variables

$$P(x_i|\theta) = \begin{cases} \frac{1}{\pi\theta^2} & |x| \leq \theta \\ 0 & \text{otherwise} \end{cases}$$

for uniform random variable

likelihood function

$$L(\theta) = \prod_{1}^{n} P(x_i|\theta) = \prod_{i=1}^{n} \frac{1}{\pi\theta^2} = \left(\frac{1}{\pi}\right)^n \left(\frac{1}{\theta^2}\right)^n$$

$$= \pi^{-n}\,\theta^{-2n}.$$

log likelihood function

$$\ln(L(\theta)) = -n\log\pi - 2n\log\theta.$$

$$\frac{d(\ln(L\theta))}{d\theta} = \frac{-2n}{\theta}$$

which is <0 for $\theta>0$.

Hense $L(\theta)$ is a decreasing function and maximised at $\underline{\theta = X_n}$.

Ans 1

In case of simple linear regression model, finding minima ~~a~~ will also fulfil the purpose that is done by gradient descent, in most cases ~~better~~ giving better results than gradient descent.

But for large dataset, calculation of minima often becomes computationally challenging, ~~thus~~ and gradient descent is relatively easy to implement, and as pin point accuracy is not required, a more general approach is followed by approximating minima using Gradient descent

Ans 2.

Function Approximation on a given dataset tends to build function to best accomodate given set and ~~not to predict new~~ results ~~is~~ a function giving whose output main purpose is to match given dataset, whereas in Machine Learning the aim is to genralise come up with a genralise function to estimate data from unknown ~~d~~ data & result. If function approximation is used in ML, there will be high tendency that the hypothesis with overfit the given data.

If we are given all possible data available before hand then function approximation will ~~in~~ ~~mos~~ give better results than Machine learning Techniques. As ML Techniques tend to genralise given data to predict unknown data results