Session - 9, From ISI, Bangalore on 25 Feb 25

Marks breakup -

- Quiz - 5%
- Assignments - 20%
- Midterm - 25%
- End term - 50%

Note - Assignment
- Theory 10% (avg over 4 assignments work book)
- Group project - application 10% (single)

H.W. $(X_1, \cdots, X_n)$ r.v. $\sim N(\mu, \sigma^2)$.

Sample mean: $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

Sample max: $X_{max} = \max_{1 \le i \le n}(X_1, \cdots, X_n)$

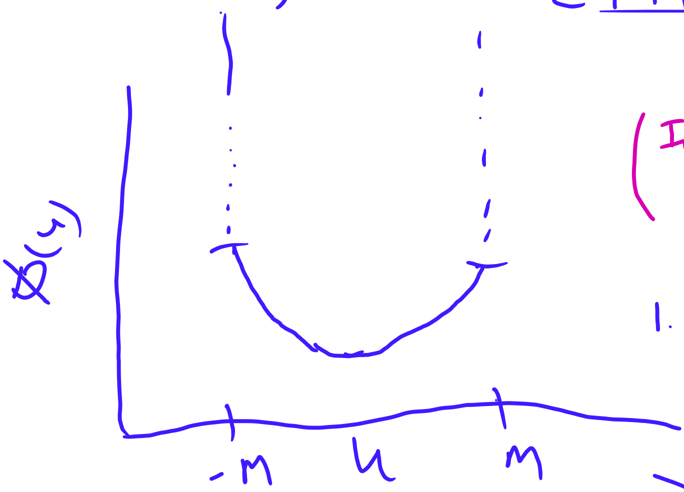
Derive the variances of $V(\bar{X})$, $V(X_{max})$ and compare.

(a) Robust least square (Huber penalty function)

$$\underset{x}{\text{minimise}} \sum_{i=1}^{n} \phi(a_i^T x - b_i)$$

where $\phi : R \to R$ defined as

$$\phi(u) = \begin{cases} u^2 & \text{if } |u| \le M \\ M(2|u| - M) & \text{if } |u| > M \end{cases}$$



(Interpretation - section 6.1 page - 299)

1. the penalty function agrees with LS when the residual is less than M.
2. It put fixed weights for residual > M.
3. Residual >M are ignored.

Weighted penalized LS

(b)
$$\underset{w, x}{\text{minimise}} \sum_{i=1}^{n} \frac{(a_i^T x - b_i)^2}{(w_i + 1)} + M^2 \mathbf{1}^T w$$

$w = (w_1 \ldots w_n)^T \succeq 0$

Second term in the objective function penalizes large value of w.

Question: Show that the two problems are equivalent.

Challeng: Objective function (a) is not convex. Hard combinatorial optimization problem.

Take problem (b):

$$\min_{x \in \mathbb{R}^n, w \in \mathbb{R}^n} \sum_{i=1}^{n} \frac{(a_i^T x - b_i)^2}{w_i + 1} + \frac{M^2}{2} 1^T w$$

it penalizes large weights

subject to $w \geq 0$

For given x, optimize i^{th} residual as a function of w_i

$$f(w_i) = \frac{(a_i^T x - b_i)^2}{w_i + 1} + \frac{M^2}{2} w_i$$

<u>Equate:</u> $\frac{\partial}{\partial w_i} f(w_i) = 0$

$$\Rightarrow w_i = \sqrt{\frac{2}{M}}\, |a_i^T x - b_i| - 1 .$$

weight to the i^{th} residual

residual

So, w_i is function of M and $\underline{|a_i^T x - b_i|}$

$\Rightarrow$ w_i depends on $|a_i^T x - b_i|$ when M is given

$\Rightarrow$ If residual is large $\Rightarrow$ w_i large.

$\Rightarrow$ weights will be small.

$\Rightarrow$ The contribution of the corresponding residual will reduce

Converse, small residual $\Rightarrow$ w_i small

$\Rightarrow$ more least square penalty.

The is same behavior as in the Robust least square function in problem (a). Also know as Huber penalty function.

Here the trick is to minimise the second function over w keeping x fixed.

Best linear unbiased estimator (BLUE) as a convex optimization problem.

(Ref: Section 4.7 page 176)

Suppose, $y = A\underline{x} + v$

where $v \in \mathbb{R}^m$ is the measurement noise

$y \in \mathbb{R}^m$ is the vector measurement

$x \in \mathbb{R}^n$ is unknown vector to be estimated when y is given

$A^{m \times n}$; fixed matrix, given

Target estimate x.

Assumption: (i) $E(v) = 0$ → centered

$E(vv^T) = I$ → uncorrelated.

(ii) A has rank 'n'

(H.W.) Derive the BLUE of x and show that the problem can be written as convex optimization problem.

Find the optimal solution and optimum value.

(Ref: Section 4.7.2, page 176)