HW02

Group 22

March 27, 2021

Theorem For every $x_1 \leq x_2 \leq ... \leq x_N \in \mathbb{R}$ there exists :

$$\widehat{z} = \{z : z = x_{\frac{N+1}{2}} \ for \ N = 2k+1 \ and \ z \in \left(x_{\frac{N}{2}}, x_{\frac{N}{2}+1}\right) \ for N = 2k, \ k \in \mathbb{N}\}$$
$$= median\{x_1, x_2, ..., x_n\}$$

such that minimizes the Mean Absolute Error J of z with respect to $x_1, x_2, ..., x_N$, meaning:

$$\widehat{J} = J(\widehat{z} \ ; x_1, x_2, ..., x_N) = min_{z \in \mathbb{R}} J(z \ ; x_1, x_2, ..., x_N)$$

Proof We know that the following equality is true:

$$\frac{dJ}{dz} = \frac{1}{N} \sum_{n=1}^{N} (|x_n - z|)' = \frac{1}{N} \sum_{n=1}^{N} \text{sgn}(z - x_n)$$

where:

$$\text{sgn}(z) = \begin{cases} -1 & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ 1 & \text{if } x > 0 \end{cases} \tag{1}$$

So when z is somewhere in $[x_1, x_N]$, the sum gives us +1 for every $x_n < z$ and -1 for every $x_n > z$

If $N = 2k+1$ $k \in \mathbb{N}$ we have:

$$\frac{dJ}{dz} = \frac{1}{N}\sum_{n=1}^{N} \operatorname{sgn}(z - x_n) \begin{cases} < 0 & \text{if } z < x_{\lfloor \frac{N+1}{2} \rfloor} \\ > 0 & \text{if } z > x_{\lceil \frac{N+1}{2} \rceil} \\ = 0 & \text{if } z \in \left[x_{\lfloor \frac{N+1}{2} \rfloor}, x_{\lceil \frac{N+1}{2} \rceil} \right] \end{cases}$$

and $z \in \left[x_{\lfloor \frac{N+1}{2} \rfloor}, x_{\lceil \frac{N+1}{2} \rceil} \right] = [x_{k+1}, x_{k+1}] = \{x_{k+1}\}$

So in this case the minimizing z is $\widehat{z} = x_{k+1}$

If $N = 2k$ $k \in \mathbb{N}$ we have:

$$\frac{dJ}{dz} = \frac{1}{N}\sum_{n=1}^{N} \operatorname{sgn}(z - x_n) \begin{cases} < 0 & \text{if } z < x_{\lfloor \frac{N+1}{2} \rfloor} \\ > 0 & \text{if } z > x_{\lceil \frac{N+1}{2} \rceil} \\ = 0 & \text{if } z \in \left(x_{\lfloor \frac{N+1}{2} \rfloor}, x_{\lceil \frac{N+1}{2} \rceil} \right) \end{cases}$$

And in this case the minimizing z is $\widehat{z} \in \left(x_{\lfloor \frac{N+1}{2} \rfloor}, x_{\lceil \frac{N+1}{2} \rceil} \right)$

So for every $k \in \mathbb{N}$, $\widehat{z} = median\{x_1, x_2, ..., x_n\}$