$$y_i = \Sigma_{j=1}^{m} w_j h_j (x_i) + \epsilon_i$$, h - basis f^{o}

$$w = (H^T H)^{-1} H^T y \qquad (H - n \times m) \qquad (H_{ij} = h_j(x_i))$$

$$\hat{y} = h(x)^T w = h(x)^T (H^T H)^{-1} H^T y$$

$$= H (H^T H)^{-1} h(x))^T y$$

$$l(x) = H (H^T H)^{-1} h(x)$$

$$\hat{y} = l(x)^T . Y \rightarrow$$ Linear Smoother

2 a) $K(x_i, x) = \exp \frac{\|x_i - x\|_2^2}{\sigma}$ $\hat{y} = ?$

$$\hat{y} = \frac{\Sigma_{i=1}^{n} w_i y_i}{\sum_{i=1}^{n} w_i} \qquad \left(w_i = \exp\left(-\frac{D(x_i, x)^2}{K_w^2}\right)\right.$$

$$l_i(x) = \frac{w_i}{\Sigma_{i=1}^{n} w_i}$$

we expanded this eqn is form of $l(x)^T y$

$$\boxed{\hat{y} = l(x)^T y}$$

So the above linear Regression is Kernel ~~f~~ smoother.

Ans 2(b) sum of residual ~~error~~ square $\|Hw - Y\|_2$

$\|Hw - Y\|_1$

There is no way we can convert w that will ~~be~~ minimize $\underline{\Sigma \text{ abs(error)}}$

optimal w has same # of +ve & -ve error

$\rightarrow x_i = 1$ (or any constant) ($x_i \rightarrow$ input)

for different y

$w = \text{median}(y)$ in any of $(y_i \cdots y_n)$

w is not linear

(Median α Rank of ~~mat~~ matrix

2C

$$\hat{y} = \frac{1}{|B_k|} \Sigma_{q\, x_i \in B_k} y_i \quad \text{for } x \in B_k$$

$$l_j(x) = \frac{I(x_j \in B_i)}{|B_i|}$$

$I \rightarrow$ indicator Random Variable

since I is indicator Random variable

we can say regrossogram is linear smo