X, Y

$K \times N$ $\quad 1 \times N$

$Y = \Theta^T X$

.predict()

$\hat{Y} = \Theta^T X$

Θ

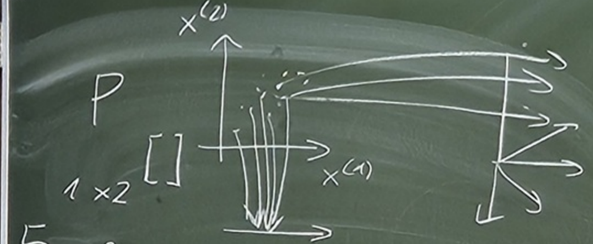
$K \times 1$ $\quad B$

$\hat{\Theta} = (XX^T)^{-1} X Y^T$

.fit()

$\hat{\Theta} = BY^T$?

$Av = \lambda v$



P

1×2 []

5×2

$y = ax \quad y = ax \quad a = \frac{y}{x}$

$Y = [y_1 \cdots y_N]$

$\hat{Y} = [\hat{y}_1 \cdots \hat{y}_N]$

$E = Y - \hat{Y} = [e_1 \cdots e_N]$

$1 \times N$

$Q(\Theta) = \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{N} e_i^2 = EE^T = (Y - \hat{Y})(Y - \hat{Y})^T = (Y - \Theta^T X)(Y - \Theta^T X)^T$

$\Theta^T x_i$

I $\nabla_{\tilde{\Theta}} YY^T = O_K$

II $\nabla_{\tilde{\Theta}} \underbrace{YX^T}_{1\times K} \tilde{\Theta} = (YX^T)^T$

III $\nabla_{\tilde{\Theta}} \tilde{\Theta}^T \underbrace{XY^T}_{} = XY^T$

IV $\nabla_{\tilde{\Theta}} \tilde{\Theta}^T XX^T \tilde{\Theta} = 2XX^T \tilde{\Theta}$

$\nabla_{\tilde{\Theta}} =$ w domu $= O_K \xrightarrow{?} \hat{\tilde{\Theta}}$

$$\nabla_{\tilde{\Theta}} Q = \begin{bmatrix} \frac{\partial Q}{\partial \tilde{\Theta}^{(1)}} \\ \vdots \\ \frac{\partial Q}{\partial \tilde{\Theta}^{(K)}} \end{bmatrix}$$

$$Q(\tilde{\Theta}) = (Y - \tilde{\Theta}^T X)(Y - \tilde{\Theta}^T X)^T$$

$$Q(\tilde{\Theta}) = (Y - \tilde{\Theta}^T X)(Y^T - X^T \tilde{\Theta})$$

$$Q(\tilde{\Theta}) = \underbrace{YY^T}_{I} - \underbrace{Y X^T \tilde{\Theta}}_{II} - \tilde{\Theta}^T X Y^T + \tilde{\Theta}^T X X^T \tilde{\Theta}$$