$$\hat{\Theta} = (\Phi^T \Phi)^{-1} \Phi^T y \qquad y = \begin{bmatrix} Y(1) \\ \vdots \\ Y(N) \end{bmatrix}_{N\times 1} \quad \Phi = \begin{bmatrix} \varphi^T(1) \\ \varphi^T(2) \\ \vdots \\ \varphi^T(N) \end{bmatrix}_{N\times 2}$$

$$R_N = \Phi^T \Phi \qquad f_N = \Phi^T y$$

R_N is 2×2, f_N is 2×1

$$\varphi(i) = \begin{bmatrix} 1 \\ x(i) \end{bmatrix}_{2\times 1}$$

2: number of parameters
N: data size

$$R_N = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ x(1) & x(2) & x(3) & & x(N) \end{bmatrix}_{2\times N} \begin{bmatrix} 1 & x(1) \\ 1 & x(2) \\ & \vdots \\ 1 & x(N) \end{bmatrix}_{N\times 2}$$

$$R_N = \begin{bmatrix} N & x(1)+x(2)+\cdots x(N) \\ x(1)+\cdots x(N) & x^2(1)+x^2(2)+\cdots x^2(N) \end{bmatrix} = \frac{1}{N}\begin{bmatrix} N & \sum_{i=1}^{N} x(i) \\ \sum_{i=1}^{N} x(i) & \sum_{i=1}^{N} x^2(i) \end{bmatrix}$$

$$f_N = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x(1) & x(2) & & x(N) \end{bmatrix}_{2\times N} \begin{bmatrix} Y(1) \\ Y(2) \\ \vdots \\ Y(N) \end{bmatrix}$$

$$f_N = \begin{bmatrix} Y(1)+Y(2)+\cdots+Y(N) \\ x(1)Y(1)+x(2)Y(2)+\cdots+x(N)Y(N) \end{bmatrix} = \frac{1}{N}\begin{bmatrix} \sum_{i=1}^{N} Y(i) \\ \sum_{i=1}^{N} x(i)Y(i) \end{bmatrix}$$

This was the solution to fit a line with 2 parameters.
What would change if we wanted to fit a polynomial of $a+bx+cx^2$ or a polynomial of degree K?

$a+bx+cx^2+dx^3+\cdots \omega x^K$