

$$\underline{w} = \begin{pmatrix} w_0 \\ \vdots \\ w_k \end{pmatrix} \quad \underline{t} = \begin{pmatrix} t_1 \\ \vdots \\ t_N \end{pmatrix}$$

$$\underline{X} = \begin{pmatrix} 1 & x_1 & x_1^2 & x_1^3 \\ 1 & x_2 & x_2^2 & x_2^3 \\ 1 & x_3 & x_3^2 & x_3^3 \\ 1 & x_4 & x_4^2 & x_4^3 \end{pmatrix} : \downarrow 1 \text{ row per observation}$$

$\rightarrow 1 \text{ col per input / feature}$

$$\underline{w} = \underbrace{(\underline{X}^\top \underline{X})^{-1}}_{=} \underline{X}^\top \underline{t}$$

$$\text{np.linalg}(\underline{X}^\top \underline{X})^{-1} \underline{X}^\top \underline{t}$$

$$x_{\text{test}} \rightarrow \begin{pmatrix} 1 & x_{\text{test}} & x_{\text{test}}^2 & \dots \end{pmatrix}$$

$$t_{\text{test}} = \underline{w}^\top x_{\text{test}}$$

$$\downarrow \text{np.linalg}(\underline{X}^\top \underline{X}, \underline{X}^\top \underline{t})$$