$$\{X = [\vec{1} \mid \vec{x}_1 \mid \dots \mid \vec{x}_p], \vec{y}\} = \mathbb{D}$$

$$\mathcal{X}_{\mathbb{D}} = [\min(\vec{x}_1), \max(\vec{x}_1)] \times [\min(\vec{x}_2), \max(\vec{x}_2)] \times \dots \times [\min(\vec{x}_p), \max(\vec{x}_p)]$$

$\vec{x}_* \in \mathcal{X}_{\mathbb{D}}$ $\quad g(\vec{x}_*)$ is called interpolation

$\vec{x}_* \notin \mathcal{X}_{\mathbb{D}}$ $\quad g(\vec{x}_*)$ " " " " extrapolation

---

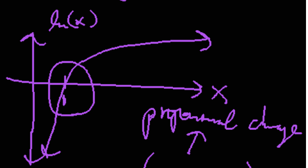
Non-linear models fit with OLS. Log transformations.

$$\hat{y} = b_0 + b_1 \ln(x) \qquad X = \begin{bmatrix} 1 & x_{11} \\ 1 & x_{12} \\ 1 & x_{13} \\ \vdots & \vdots \\ 1 & x_{1n} \end{bmatrix} \rightarrow X = \begin{bmatrix} 1 & \ln(x_{11}) \\ 1 & \ln(x_{12}) \\ \vdots & \vdots \\ 1 & \ln(x_{1n}) \end{bmatrix}$$

$$\ln(x+1) = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots + \dots$$

for x small ≈ 0 $\quad \approx x$

$$\ln(x) = \ln((x+1)-1) \approx x - 1 \quad \text{for } x \text{ near } 1$$



$$b_1 \Delta \ln(x) = b_1(\ln(x_f) - \ln(x_0)) = b_1 \ln\left(\frac{x_f}{x_0}\right) \approx b_1\left(\frac{x_f}{x_0} - 1\right) = b_1\left(\frac{x_f - x_0}{x_0}\right)$$

if $\frac{x_f}{x_0} \approx 1$

if x increases by 25% $\Rightarrow$ $\hat{y}$ increases by $0.25\, b_1$

$$\hat{y}' = \ln(\hat{y}) = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_p x_p \Rightarrow \hat{y} = e^{\boxed{b_0 + b_1 x_1 + \dots + b_p x_p}}$$

log-linear model

$$= e^{b_0} e^{b_1 x_1} \cdot \dots \cdot e^{b_p x_p}$$

$$= \underbrace{e^{b_0}}_{m_0} \underbrace{(e^{b_1})}_{m_1}^{x_1} \cdot \dots \cdot \underbrace{(e^{b_p})}_{m_p}^{x_p}$$

$$\hat{y} = m_0 m_1^{x_1} \cdot \dots \cdot m_p^{x_p}$$

$$\ln(\hat{y}) = b_0 + b_1 \ln(x)$$

$$X = \begin{bmatrix} 1 & x_{11} & x_{21} \\ 1 & x_{12} & x_{22} \\ \vdots & \vdots & \vdots \\ 1 & x_{1n} & x_{2n} \end{bmatrix} \xrightarrow{\text{first-order interaction}} X = \begin{bmatrix} 1 & x_{11} & x_{21} & x_{11}x_{21} \\ 1 & x_{12} & x_{22} & x_{12}x_{22} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{1n} & x_{2n} & x_{1n}x_{2n} \end{bmatrix}$$

derived feature

$p_{raw} = 2 \Rightarrow p = 3$

$$g(\vec{x}) = \hat{y} = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_1 x_2$$

$$= b_0 + (b_1 + b_3 x_2) x_1 + b_2 x_2$$

$$= b_0 + b_1 x_1 + (b_2 + b_3 x_1) x_2$$

$$f: \mathbb{R}^2 \rightarrow \mathbb{R} = \mathcal{Y}$$

non-linear linear model fit with OLS.

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_4 x_4 + b_5 x_1 x_2 + b_6 x_1 x_3 + b_7 x_1 x_4$$