FoDA L12 : Multiple Linear Regression & Polynomial Regression

Input data

$(X, y) = \{(x_1, y_1), (x_2, y_2), \dots (x_n, y_n)\}$

$x_i \in \mathbb{R}^d \quad y_i \in \mathbb{R}$

Ex $x_2 = (3, 2, 5, 6) \in \mathbb{R}^4 \quad y_2 = -3$

linear model

$x_i \in \mathbb{R}^d \quad x_i = (x_{i1}, x_{i2}, \dots, x_{id})$

dependent var

b a

$$\hat{y}_i = M_\alpha(x_i) = \alpha_0 + x_{i1}\alpha_1 + x_{i2}\alpha_2 + \dots + x_{id}\alpha_d$$

$$= \alpha_0 1 + \sum_{j=1}^{d} x_{ij}\alpha_j$$

$$\alpha = (\alpha_0, \alpha_1, \alpha_2 \dots \alpha_d) \in \mathbb{R}^{d+1}$$

$$= \langle \alpha, (1, x_{i1}, x_{i2}, \dots x_{id}) \rangle = \langle \alpha, (1, x_i) \rangle$$

$\in \mathbb{R}^{d+1}$

Data (X, y) Model M_α $\alpha \in \mathbb{R}^{d+1}$

$X \subset \mathbb{R}^d$ $y \in \mathbb{R}$

Error $SSE((X,y), M_\alpha) = \sum_{i=1}^{n} \underbrace{r_i}_{\text{residual}}^2 = \sum_{i=1}^{n} (y_i - \overbrace{M_\alpha(x_i)}^{\hat{y}_i})^2$

$X \in \mathbb{R}^{n \times d}$

$\hookrightarrow \tilde{X} \in \mathbb{R}^{n \times (d+1)}$

$$X = \begin{bmatrix} 2 & 4 & 3 \\ 1 & 7 & 6 \end{bmatrix} \in \mathbb{R}^{2\times 3}$$

$$\tilde{X} = \begin{bmatrix} 1 & 2 & 4 & 3 \\ 1 & 1 & 7 & 6 \end{bmatrix} \in \mathbb{R}^{2\times 4}$$

$\boxed{d=1 \quad \alpha = \frac{\langle \tilde{x}, \tilde{y} \rangle}{\|x\|^2}}$

all 1s column

$$\tilde{X} = [\mathbb{1}; X_1, X_2, \ldots X_d]$$

$$X_j = \begin{bmatrix} x_{1j} \\ x_{2j} \\ \vdots \\ x_{nj} \end{bmatrix}$$

jth column

$\boxed{\text{optimal SSE soln} \quad \alpha^* = (\tilde{X}^T \tilde{X})^{-1} \tilde{X}^T y}$

$\hat{X}$ is M_α (linear model)

?

input to linear model

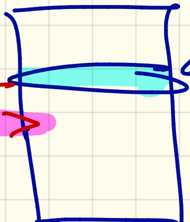
to make a prediction

Models

α is one row $x_i \in \mathbb{R}^d$

$$\hat{y}_i = M_\alpha(x_i) = \alpha_0 + \alpha_1 \bigcirc + \alpha_2 \bigcirc + \dots + \alpha_d \bigcirc \qquad X \in \mathbb{R}^{n \times d}$$

x_{i1} $\quad$ x_{i2} $\quad$ x_{id}



$x_i = (x_{i1}, x_{i2}, \dots x_{id})$

$\hat{y}_i \leftarrow \langle \alpha, \rangle$

website tracking customers

$n = 11$ customers

track 3 explanatory variables

$X \in \mathbb{R}^{n \times d} = \mathbb{R}^{11 \times 3}$

1. time on site (sec)
2. jiggle (cm)
3. scroll (cm)

track 1 dependent var

sales in cents

$\tilde{X} =$

x_3

$x_1 =$

time: X_1	jiggle: X_2	scroll: X_3	sales: y
232	33	402	2201
10	22	160	0
6437	343	231	7650
512	101	17	5599
441	212	55	8900
453	53	99	1742
2	2	10	0
332	79	154	1215
182	20	89	699
123	223	12	2101
424	32	15	8789

customer

$= y \in \mathbb{R}^n$

$$\alpha^* = (\tilde{X}^T \tilde{X})^{-1} \tilde{X}^T y$$

$\alpha_0 = 2626$ ← come on site

$\alpha_1 = 0.42$ (time on site)

$\alpha_2 = 17.72$ (jiggle cm)

$\alpha_3 = -6.50$ (scroll cm)

Polynomial Regression

Input (X, y) $X \in \mathbb{R}^{n \times d}$ (assume $d=1$) $y \in \mathbb{R}^n$

poly-degree-2 (quadratic) model

$$\hat{y} = M_\alpha^{(2)}(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2$$

$M_\alpha^{(2)} : \mathbb{R} \to \mathbb{R}$

$$\hat{y} = M_\alpha^{(p)}(x) = \alpha_0 + \alpha_1 X + \dots + \alpha_p x^p$$

$$= \alpha_0 + \sum_{j=1}^{p} \alpha_j x^j = \sum_{j=0}^{p} \alpha_j x^j \qquad \alpha \in \mathbb{R}^{p+1}$$

$$= \langle \alpha, (1, x, x^2, \dots, x^p) \rangle \qquad (1, x, x^2, \dots x^p) \in \mathbb{R}^{p+1}$$

Residual for (x_i, y_i) as

$$r_i = \hat{y}_i - y_i = M_\alpha^{(p)}(x_i) - y_i$$

$$SSE\left((x,y), M_\alpha^{(p)}\right) = \sum_{i=1}^{n} r_i^2 = \sum_{i=1}^{n} \left(M_\alpha^{(p)}(x_i) - y_i\right)^2$$

$$= \sum_{i=1}^{n} \left(\left\langle \alpha, (1, x, x^2, \dots, x^p)\right\rangle - y_i\right)^2$$

initial input

$$x \in \mathbb{R}^1 \Longrightarrow \tilde{x} = (1, x, x^2, \dots, x^p) \in \mathbb{R}^{p+1}$$

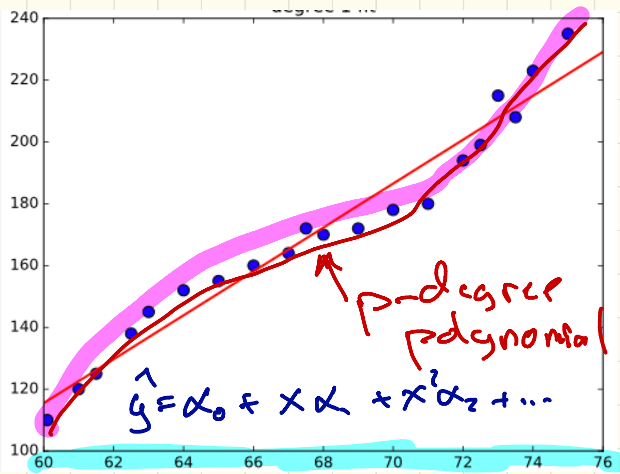
$$\tilde{X}_p = \begin{bmatrix} 1 & x_1 & x_1^2 & & x_1^p \\ 1 & x_2 & x_2^2 & \cdots & x_2^p \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & & x_n^p \end{bmatrix} \qquad \alpha^*_{(p)} = \left(\tilde{X}_p^T \tilde{X}_p\right)^{-1} \tilde{X}_p^T y$$

x y

height (in)	weight (lbs)
66	160
68	170
60	110
70	178
65	155
61	120
74	223
73	215
75	235
67	164

x y

height (in)	weight (lbs)
61.5	125
73.5	208
62.5	138
63	145
64	152
71	180
69	172
72.5	199
72	194
67.5	172



$$x = \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix} \in \mathbb{R}^{n=3} \qquad y = \begin{bmatrix} 1 \\ 6 \\ 5 \end{bmatrix}$$

$$p = 5$$

$$\left\{ \begin{array}{c} \tilde{X}_p = \begin{bmatrix} 1 & 2 & 4 & 8 & 16 & 32 \\ 1 & 4 & 16 & 64 & 256 & 1024 \\ 1 & 3 & 9 & 27 & 81 & 243 \end{bmatrix} \\ \begin{array}{cccccc} \alpha_0 & \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_5 \end{array} \end{array} \right\} \Rightarrow \hat{y}$$