FoDA L12

Multi-linear and Polynomial Regression

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

multiple factors

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

explanatory $\quad$ dependent

$M(x_i) \approx y_i$

linear

$$\hat{y}_i = M_\alpha(x_i) = M_\alpha(x_{i,1}, x_{i2}, \dots, x_{id}) = \alpha_0 + \sum_{j=1}^{d} \alpha_j x_{ij}$$

linear model

$$= \alpha_0 1 + \alpha_1 x_{i1} + \alpha_2 x_{i2} + \dots + \alpha_d x_{id}$$

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

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

line $\quad \hat{y}_i = a x_i + b\, 1$

$\tilde{X} \in \mathbb{R}^{11 \times 4}$

offset X_0

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

$\alpha_0 + \alpha_1 \cdot 424 + \alpha_2 \cdot 32 + \alpha_3 \cdot 15 \approx 8789 \quad y$

$X \in \mathbb{R}^{11 \times 3}$ $\quad y \in \mathbb{R}^{11}$

model M_α s.t.

$y \approx \tilde{X}\alpha$

for all $i \in [1 \dots n]$

$y_i \approx \langle \tilde{x}_i, \alpha \rangle$

$\alpha = (\alpha_0, \alpha_1, \alpha_2, \alpha_3)$

$\alpha_0 = 2626$

$\alpha_1 = 0.42$

$\alpha_2 = 12.72$

$\alpha_3 = -6.50$

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

$$M_\alpha(x_i) = \langle \alpha, (1, x_i) \rangle$$

$$SSE((x, y), M_\alpha) = \sum_{i=1}^{n} r_i^2 = \sum_{i=1}^{n} \big(\underbrace{y_i - \overbrace{M_\alpha(x_i)}^{\hat{y}_i}}\big)^2$$

goal $\quad \alpha^* = \underset{\alpha \in \mathbb{R}^{d+1}}{\arg\min} \; S(\alpha) = SSE((X, y), M_\alpha)$

$$\tilde{X} = [\mathbb{1}; X] \in \mathbb{R}^{n \times (d+1)}$$

$$a = \frac{\langle \bar{x}, \bar{y} \rangle}{\langle \bar{x}, \bar{x} \rangle}$$

$\mathbb{1} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$

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

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

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

Polynomial Regression

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

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

$y_i \approx \hat{y}_i = M_2(x_i)$

$$y \approx \hat{y} = M_2(x) = \overbrace{\alpha_0 + \alpha_1 x + \alpha_2 x^2}^{\text{polynomial model degree ?}} + \alpha_3 x^3$$

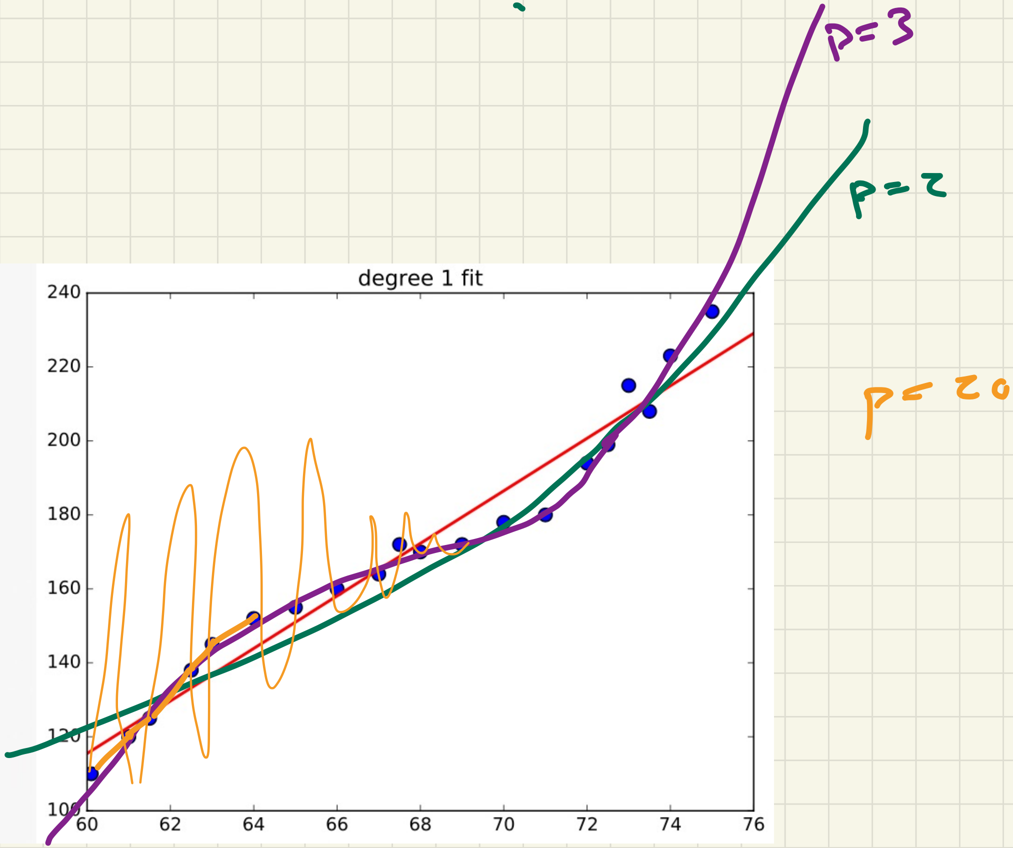
$$= M_p(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \dots + \alpha_p x^p$$

$$= \alpha_0 + \sum_{j=1}^{d} \alpha_j x^j = \sum_{j=0}^{d} \alpha_j x^j$$

$\alpha \in \mathbb{R}^{p+1}$

$$= \langle \alpha, (1, x, x^2, \dots, x^p) \rangle$$

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



Input $(x, y) \in \mathbb{R}^{n \times 1} \times \mathbb{R}^n$

$$M_{\alpha,p}(x_i) \approx y_i$$

$$M_{\alpha,p} = \langle \alpha, (1, x, \dots x^p) \rangle$$

$$\alpha^* = \underset{\alpha \in \mathbb{R}^{p+1}}{\arg\min} \; SSE((X, y), M_{\alpha,p}) = \sum_{i=1}^{n} (y_i - M_{\alpha,p}(x_i))^2$$

convert each $x_i \in X$ into $v_i \in \mathbb{R}^{p+1}$

$$v_i = (1, x, x^2, \dots, x^p)$$

$$\tilde{X}_p = \begin{bmatrix} - v_1 - \\ - v_2 - \\ \vdots \\ - v_n - \end{bmatrix} = \begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^p \\ 1 & x_2 & x_2^2 & \dots & x_2^p \\ & & & & \\ 1 & x_n & x_n^2 & \dots & x_n^p \end{bmatrix}$$

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

$n=3$

$$x = \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix} \qquad y = \begin{bmatrix} 1 \\ 6 \\ 5 \end{bmatrix}$$

$$\tilde{X}_p = \begin{bmatrix} 1 & 2 & 4 & 8 & 16 & 32 \\ 1 & 4 & 16 & 64 & 256 & 1024 \\ 1 & 3 & 9 & 27 & 81 & 243 \end{bmatrix}$$

$p=5$

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