FoDA L14

Linear Regression

Review & Misc.

Polynomial Regression

<u>Input</u> (X, y)

$X \in \mathbb{R}^{n \times 1}$ $\quad y \in \mathbb{R}^n$

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

$x_i \in \mathbb{R}$

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

$v_i = (1, x_i, x_i^2, x_i^3, \dots, x_i^p) \in \mathbb{R}^{p+1}$

$$\tilde{X}_p = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{R}^{n \times p}$$

$$\alpha^* = \underset{\alpha}{\text{minimize}} \sum_{i=1}^{n} \left(M_{\alpha, p}(x_i) - y_i\right)^2$$

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

$$M_{\alpha, p}(x_i) = \langle \alpha, v_i \rangle$$

Cross-Validation

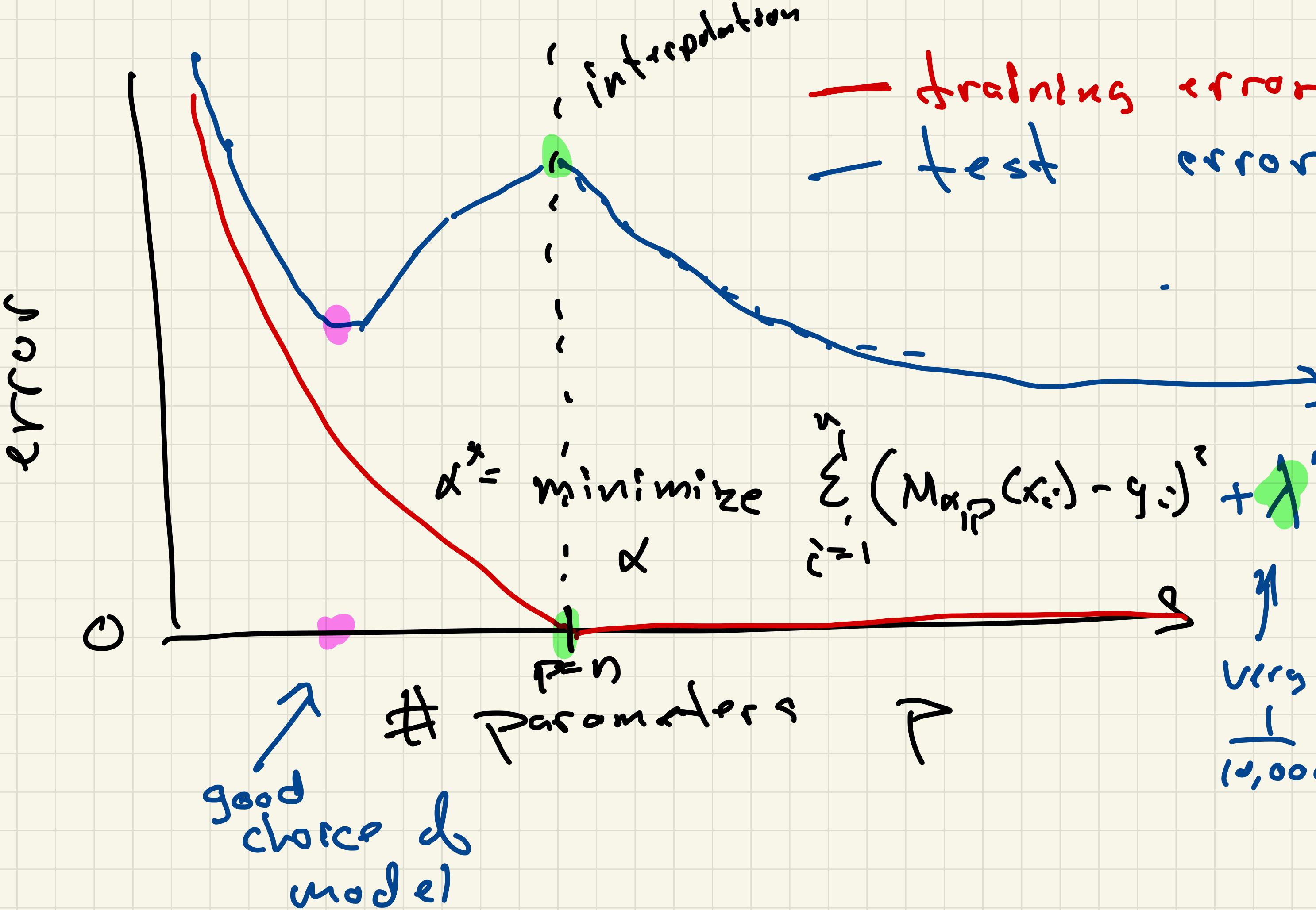
Split $(X, y) \rightarrow$ training set (X_R, y_R)

test set (X_E, y_E)

1. Train α on X_R, y_R
2. Test error on X_E, y_E

- how well on new data
- choose model parameter e.g. p

$$\alpha^* = \underset{\alpha}{\text{minimize}} \sum_{i=1}^{n} \left(M_{\alpha, p}(x_i) - y_i \right)^2$$



$$\alpha^* = \underset{\alpha}{\text{minimize}} \sum_{i=1}^{n} \left(M_{\alpha,p}(x_i) - y_i \right)^2 + \lambda \|\alpha\|^2$$

λ very small $\frac{1}{10{,}000}$

Why $\alpha^* = (X^T X)^{-1} X^T y$ $\alpha \in \mathbb{R}^d$

$\alpha = (\alpha_1 \dots \alpha_d)$

multi-variable linear regression

Input X, y $X \in \mathbb{R}^{n \times d}$ $y \in \mathbb{R}$ $x_{i1} = 1$ for all i

$x_i \in \mathbb{R}^d$ $x_i = (1, x_{i2}, x_{i3} .. x_{id})$

$$S(\alpha) = SSE((X, y), M_\alpha) = \sum_{i=1}^{n} r_i^2 = \sum_{i=1}^{n} (y_i - \langle \alpha, x_i \rangle)^2$$

fixed

$$\frac{\partial S(\alpha)}{\partial \alpha_j} = 2 \sum_{i=1}^{n} r_i \frac{\partial r_i}{\partial \alpha_j} = 2 \sum_{i=1}^{n} r_i (-x_{ij}) = 2 \sum_{i=1}^{n} (y_i - \langle \alpha, x_i \rangle)(-x_{ij})$$

$$\frac{\partial S(\alpha)}{\partial \alpha_j} = \cancel{2} \sum_{i=1}^{n} \left(y_i - \langle \alpha, x_i \rangle\right)\left(\cancel{-} X_{ij}\right) = 0$$

all $j = 1 \ldots d$

$$0 = \sum_{i=1}^{n} X_{ij}\, y_i - X_{ij} \langle x_i, \alpha \rangle$$

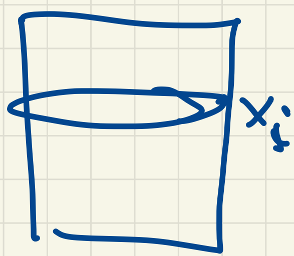
$$\sum_{i=1}^{n} x_{ij} \langle x_i, \alpha \rangle = \sum_{i=1}^{n} x_{ij}\, y_i$$

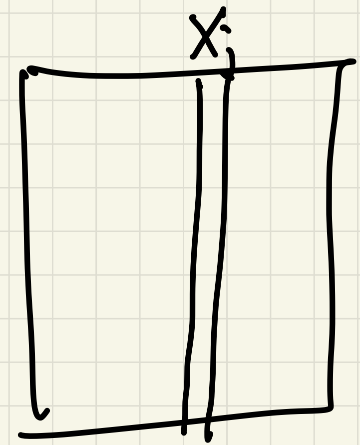
all $j = 1 \ldots d$

$$\left\langle \sum_{i=1}^{n} x_{ij}\, x_i, \alpha \right\rangle$$

$$\left(X^T X\right) \alpha = X^T y$$

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





$$(X^T X)^{-1}$$

$$(X^T X)\alpha = X^T y$$

$$X^T(X\alpha - y) = 0$$

$$(y - X\alpha)^T X = 0$$

$$\underline{0} = (y - X\alpha)^T X$$

pick any vector $v \in \mathbb{R}^n$

$$\underline{0} = (y - X\alpha)^T X v = \langle \underbrace{y - X\alpha}_{r}, Xv \rangle$$

$$v = \alpha$$

$$X\alpha = \hat{y} = \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \end{bmatrix} = M(x_i)$$

$$r = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix}$$

$$\underline{0} = \langle r, \hat{y} \rangle$$

$$\underline{0} = \langle r, \hat{y} \rangle$$