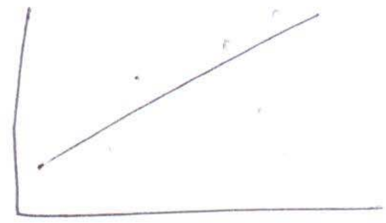
Non-Linear Regression

02/25

$x \in \mathbb{R}$



x	y
4	13
6	16
9	29
2	7

Polynomial Expansion

x^0 1	x	x^2	y
1	4	16	13
1	6	36	16
1	9	81	29
1	2	4	7
w_0	w_1	w_2	

$\begin{bmatrix} \vdots \end{bmatrix} W$

New test instance $\quad x^* = 5$

$\bar{x}^* \rightarrow [1 \;\; 5 \;\; 25]$

$$\boxed{y = W^T x^*}$$

$\sin(x) \qquad \cos(x) \qquad \exp(-2x^2)$

$x \in R^2$

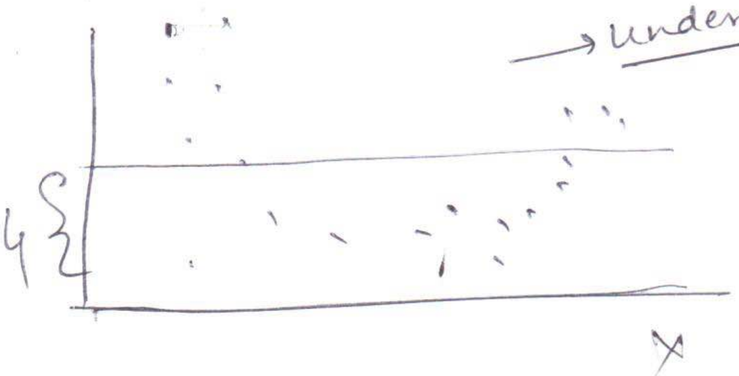
x_1	x_2	y
3	4	3
2	9	1
7	-4	2
8	-6	4
⋮		

$$\boxed{1 \quad x_1 \quad x_1^2 \quad x_2 \quad x_2^2}$$

$1, \ (x_1+x_2)^2 \quad (x_1+x_2)^3 \ \ldots$

$x^* = [4 \quad 3]$

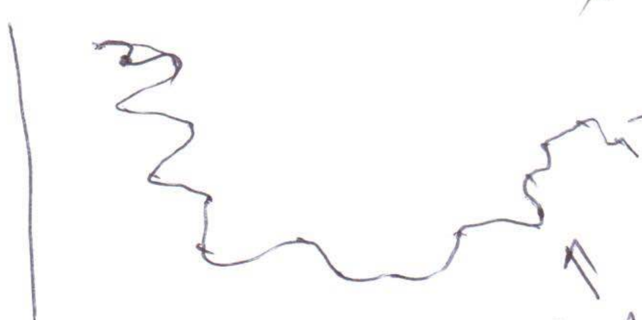
Bias – Variance Tradeoff



→ underfit.

$\boxed{4 + \overset{20}{0} \cdot x}$

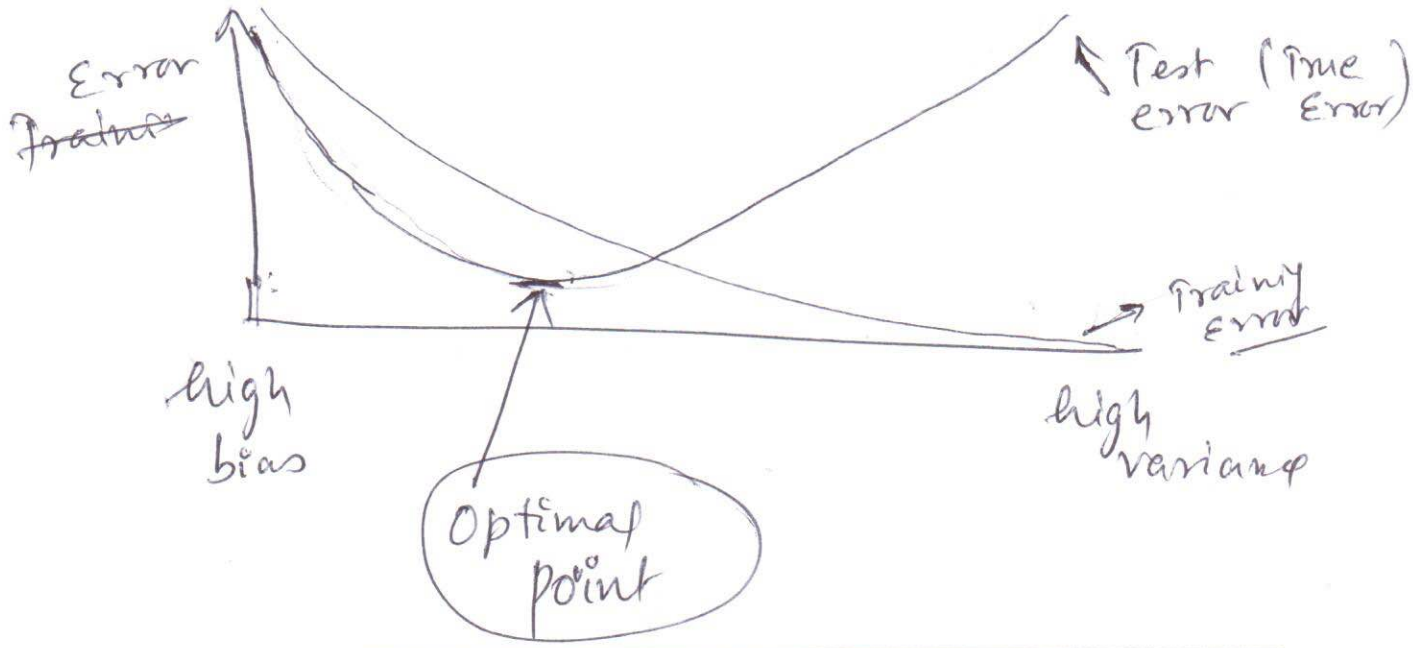
↑ high bias.



→ overfit.

↑ high variance

high Bias —————— high variance



Maximum Margin Methods (SVM) try to reach the optimal point using $\underline{\underline{C}}$

Regularization $\quad W$

$$W = [\,w_0 \mid w_1 \mid w_2 \mid w_3 \mid \cdots \mid w_d\,]$$

$$W = [w_0 \;\; w_1 \;\; w_2]^T$$

$$\|W\|_2 = (w_0^2 + w_1^2 + w_2^2)^{1/2} \rightarrow l2$$

$$\|W\|_3 = (w_0^3 + w_1^3 + w_3^3)^{1/3} \rightarrow l3$$

$$\|W\|_\infty = \max(w_0, \ldots, w_2) \rightarrow l_\infty$$

$$\|W\|_1 = |w_0| + |w_1| + |w_2| \rightarrow l_1$$

$$|W|$$

$$\boxed{\|W\|_2^2 = w_0^2 + w_1^2 + w_2^2}$$

Occam's Razor 02/27/19

$$\underset{w}{\operatorname{argmin}} \boxed{J(w) + \lambda \|w\|_2^2}$$

$$\frac{1}{2}\sum_{i=1}^{n}(y_i - w^T x_i)^2 + \lambda \|w\|_2^2$$

$$\frac{1}{2}(\underbrace{y}_{n\times 1} - \underbrace{X}_{n\times d}\underbrace{w}_{d\times 1})^T(y - Xw) + \lambda\frac{\|w\|^2}{2} = \frac{\lambda w^T w}{2}$$

$$J(w) \quad \underset{w}{\operatorname{argmin}} \quad \frac{1}{2}(y - Xw)^T(y - Xw) + \frac{\lambda}{2} w^T w$$

$$\nabla J = X^T X w - X^T y + \lambda w$$

$$\frac{d}{w}(w^T w) = 2w$$

$$\nabla J = 0$$

$$X^T X w - X^T y + \lambda w = 0$$

$$(X^T X + \lambda I_d) w = X^T y$$

$$\boxed{w = (X^T X + \lambda I_d)^{-1} X^T y}$$

Ridge regression solution

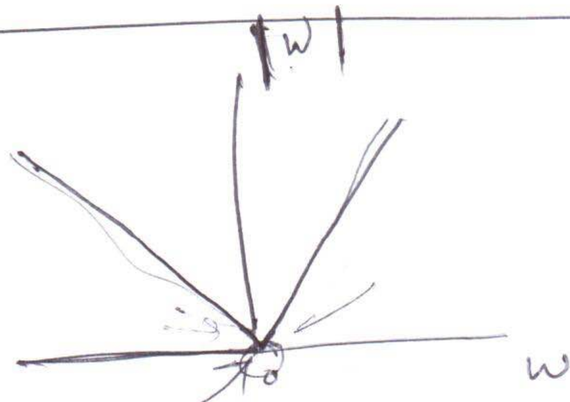
$$\boxed{\lambda > 0}$$

$$\boxed{I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad AI = A}$$

age X_1	age-spa X_2	age-children X_3	Y life expectancy
0·4	0·41	0·42	0·4 + 0·41 + 0·42
0·7	0·69	0·7	0·7 + 0·69 + 0·7
0·8	0·82	0·78	
10·4	1·35	1·41	

W	1	1	1
W	2	0	1
W	3	0	0
W	0	0	3
W	0	3	0

l1



Cross-validation

K-fold CV

10-fold CV