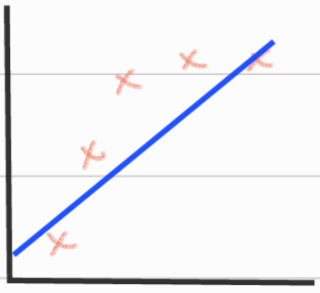
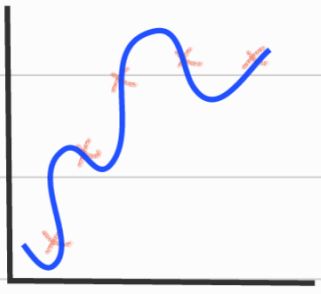
Diagnosing Bias vs. Variance

편향 분산

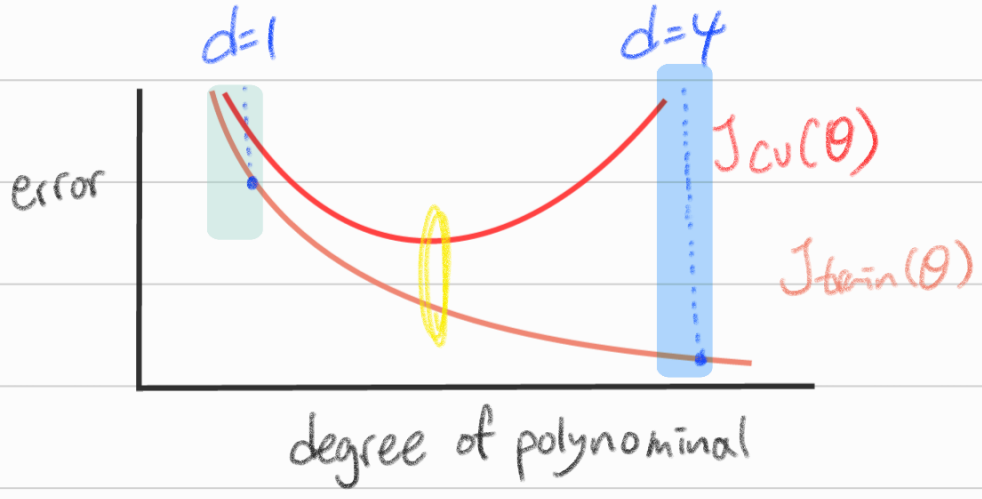


Underfit



Overfit

	Underfit	Overfit
Bias	High	Low
Variance	Low	High



Bias (underfit):

$J_{train}(\theta)$ & $J_{cv}(\theta)$ both high

Variance (overfit):

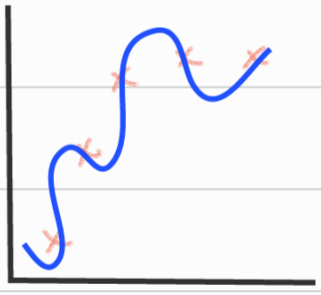
$J_{train}(\theta)$ low, but $J_{cv}(\theta)$ high

Regularization and Bias / Variance

Large λ

Underfit

Small λ



Overfit

Bias	High	Low
Variance	Low	High

Choosing the regularization parameter λ

1. Try $\lambda = 0$
2. Try $\lambda = 0.01$
3. Try $\lambda = 0.02$
4. $\vdots$

Iterate through the λs

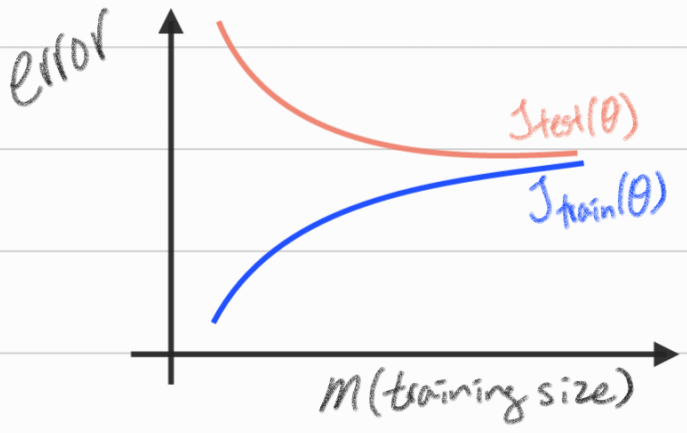
and go through all the models to learn some θ

Compute the cv error using the learned θ

Learning Curves

$$J_{train}(\theta) = \frac{1}{2m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)})^2$$

$$J_{cv}(\theta) = \frac{1}{2m_{cv}} \sum_{i=1}^{m_{cv}} (h_\theta(x_{cv}^{(i)}) - y_{cv}^{(i)})^2$$



If the training set size is small, then the training error is small.

(Very easy to fit)

cv & test error decreases as m increases

(more data, generalization↑)

If a learning algorithm is suffering from high bias, (편향)

getting more training data will not help much

If a learning algorithm is suffering from high variance (분산)

getting more training data is likely to help