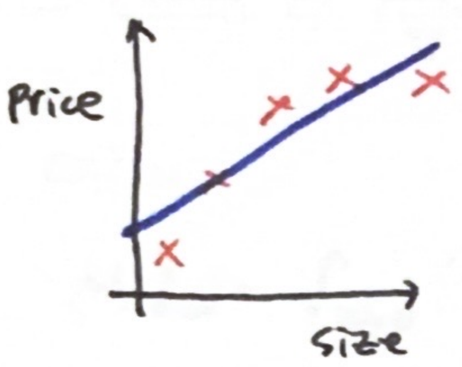
<Diagnosing bias and variance>

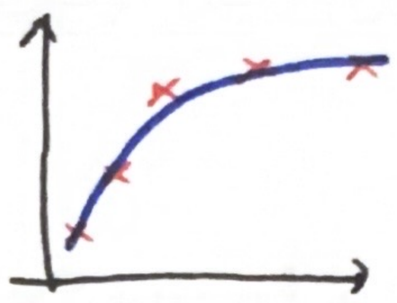


$f_{\vec{w},b}(x) = w_1 x + b$ d=1

"High bias"
(underfit)

⇓

d=1
(Jtrain is high
Jcv is high)



$f_{\vec{w},b}(x) = w_1 x + w_2 x^2 + b$ d=2

"Just right"

⇓

d=2
(Jtrain is low
Jcv is low)

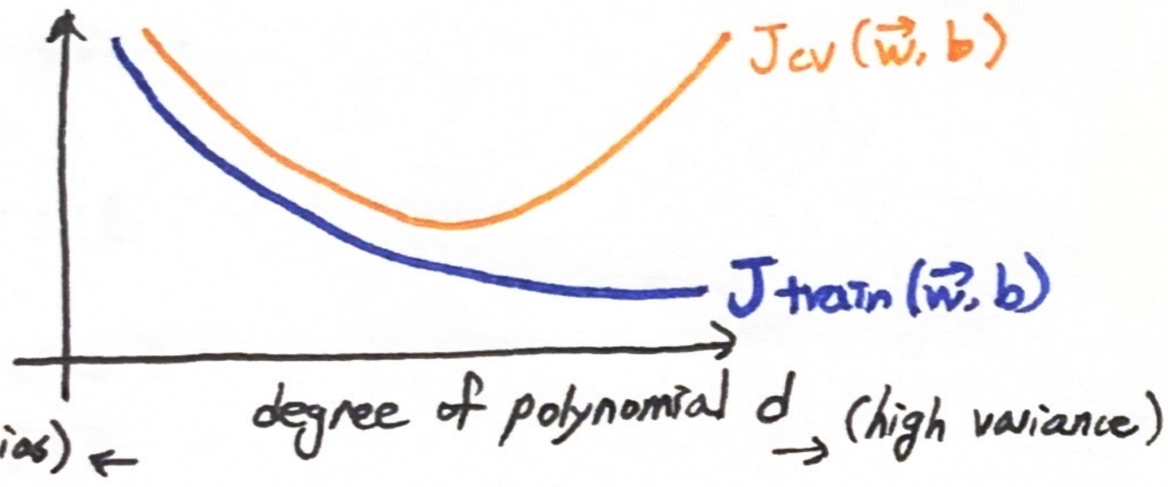
$f_{\vec{w},b}(x) = w_1 x_1 + w_2 x^2 + w_3 x^3 + w_4 x^4 + b$ d=4

"High variance"
(overfit)

⇓

d=4
(Jtrain is low
Jcv is high)

* Polynomial order d에 대한 Jcv 와 Jtrain의 상관관계

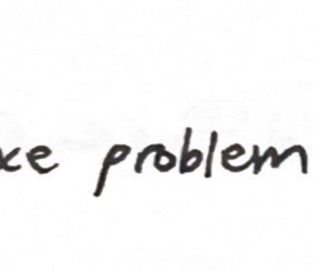


⇓

"Diagnosing bias or variance problem"

① High bias (underfit) : Jtrain will be high (Jtrain ≒ Jcv)

② High variance (overfit) : Jcv ≫ Jtrain (Jtrain may be low)

③ High bias and High variance : Jtrain will be high ⊕ Jcv ≫ Jtrain
(simultaneous)

↳ ex)