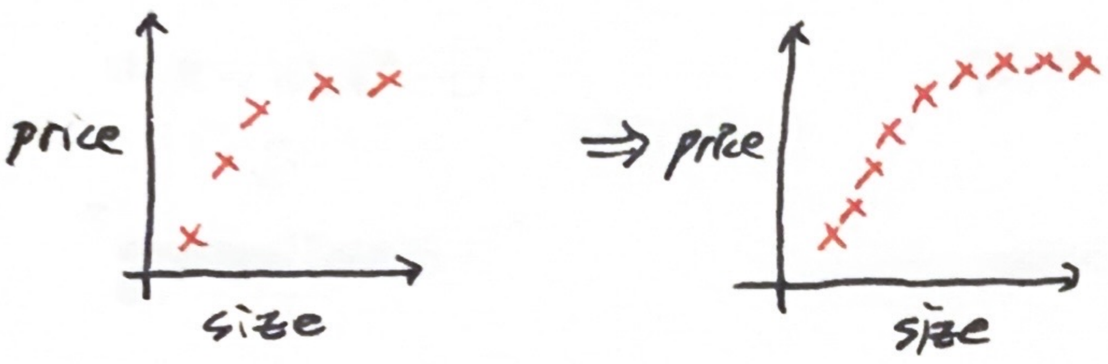
* Addressing Overfitting

- overfitting을 제한할 수 있는 방법

① "Collect more training data"



⇒ with larger training set,
the learning algorithm will learn to fit
a function that is less wiggly

② "Select features to include/exclude"

= "Select fewer features"

ex) training set

size x_1	bedrooms x_2	floors x_3	age x_4	avg income x_5	...	distance to coffee shop x_{100}	price y
					...		

a lot of features ⊕ insufficient data

⇓

occurs overfitting

⇒ select fewer features for training

= "feature selection"

* disadvantage of feature selection

- 모든 feature가 조금씩은 target 예측에 도움이 되는 경우도 있음

feature 개수를 유지

⊕

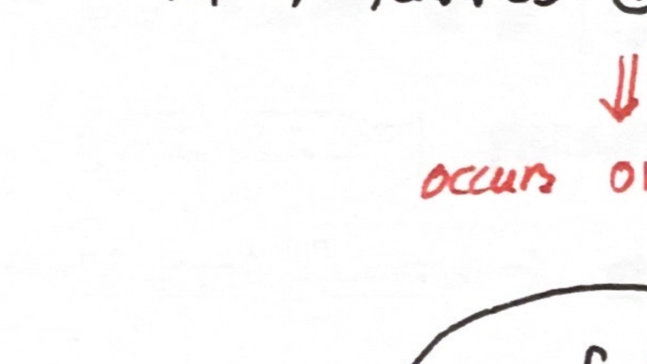
overfitting을 피하는 방법

③ "Regularization"

- 많은 feature가 target (y) 예측에 조금씩 기여하는 경우 유용

- make learning algorithm to shrink the values of parameters without eliminating features

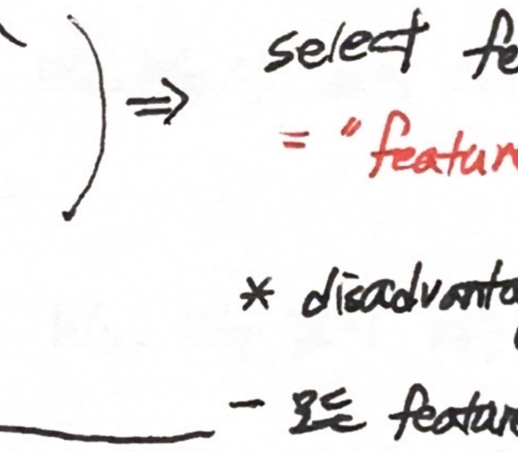
ex)



$f(x) = 28x - 385x^2 + 39x^3$

$+ 39x^3 - 1174x^4 + 400$

regularization ⟹



$f(x) = 13x - 0.23x^2 + 0.00001 4x^3$

$- 0.00001x^4 + 10$

small values for W_j parameters