



deeplearning.ai

## Regularizing your neural network

---

### Why regularization reduces overfitting

↳

■■■: 00:00:03.000



deeplearning.ai



■ 2 ■ / ■ 26 ■

■■■: keyframe\_00-00-03-000\_0001.jpg



deeplearning.ai

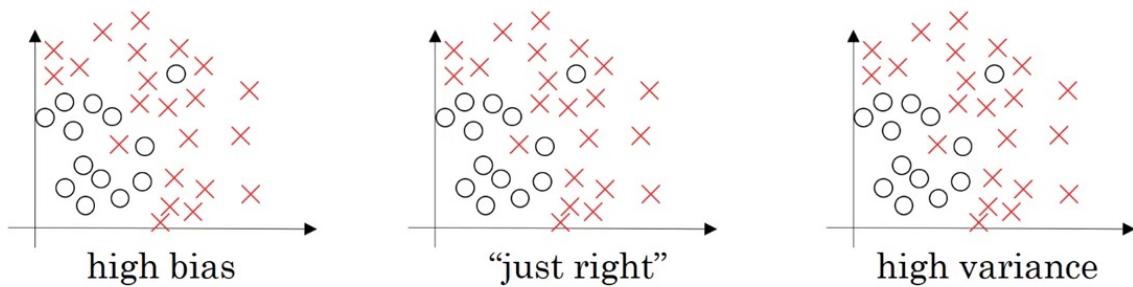
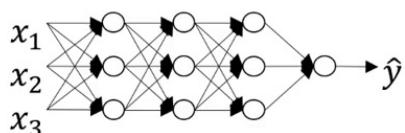
Regularizing your  
neural network

---

Why regularization  
reduces overfitting

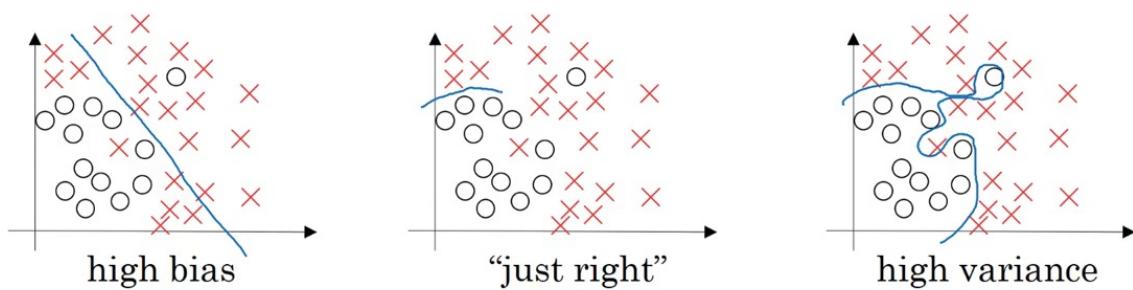
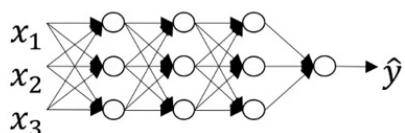
↳

## How does regularization prevent overfitting?



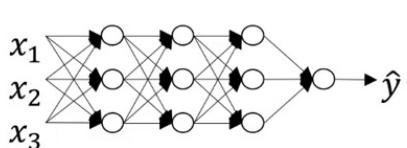
Andrew Ng

## How does regularization prevent overfitting?

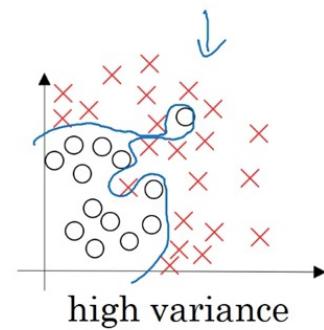
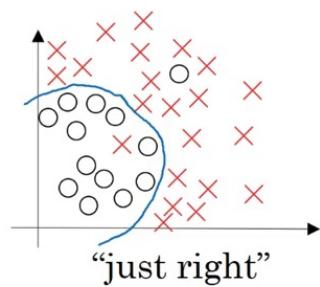
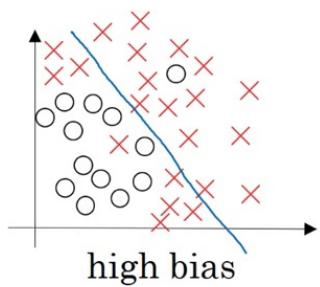


Andrew Ng

## How does regularization prevent overfitting?

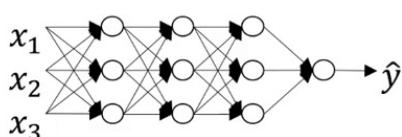


$\mathcal{J}(\omega^{(1)}, b)$

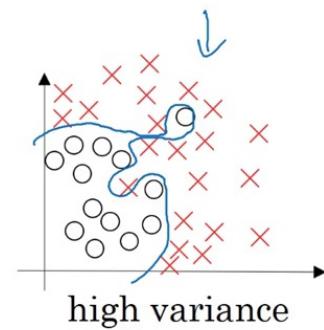
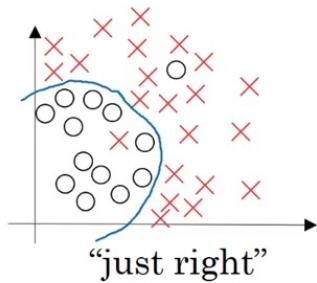
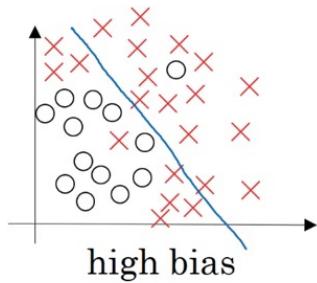


Andrew Ng

## How does regularization prevent overfitting?

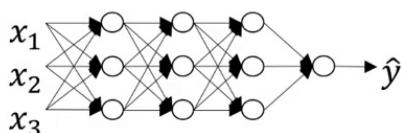


$$\mathcal{J}(\boldsymbol{\omega}^{(m)}, b^{(m)}) = \frac{1}{m} \sum_{i=1}^m L(y^{(i)}, \hat{y}^{(i)}) + \frac{\lambda}{2m} \sum_{l=1}^L \|\boldsymbol{\omega}^{(l)}\|$$



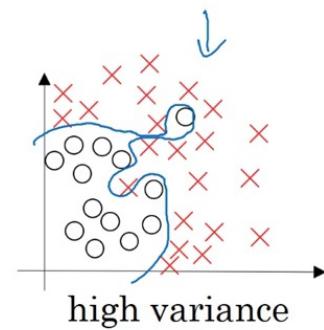
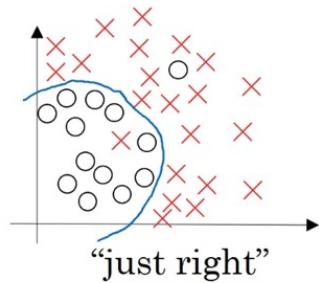
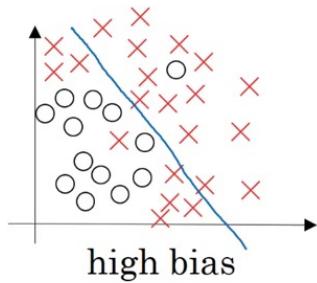
Andrew Ng

## How does regularization prevent overfitting?



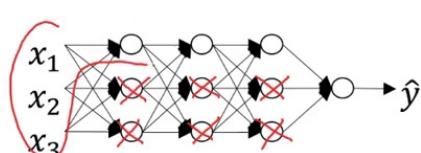
$$J(\omega^{(0)}, b^{(0)}) = \frac{1}{m} \sum_{i=1}^m L(y^{(i)}, \hat{y}^{(i)}) + \frac{\lambda}{2m} \sum_{l=1}^L \|\omega^{(l)}\|_F^2$$

$\omega$



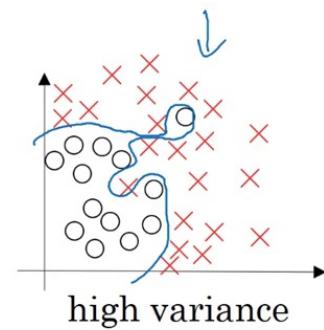
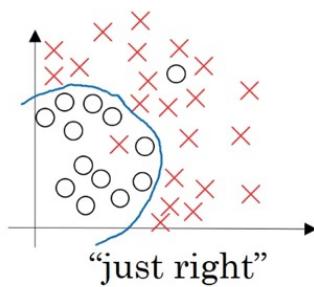
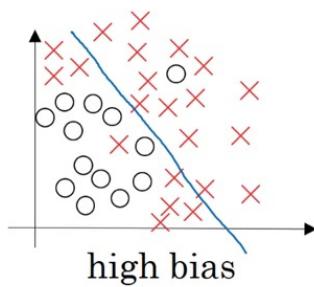
Andrew Ng

## How does regularization prevent overfitting?



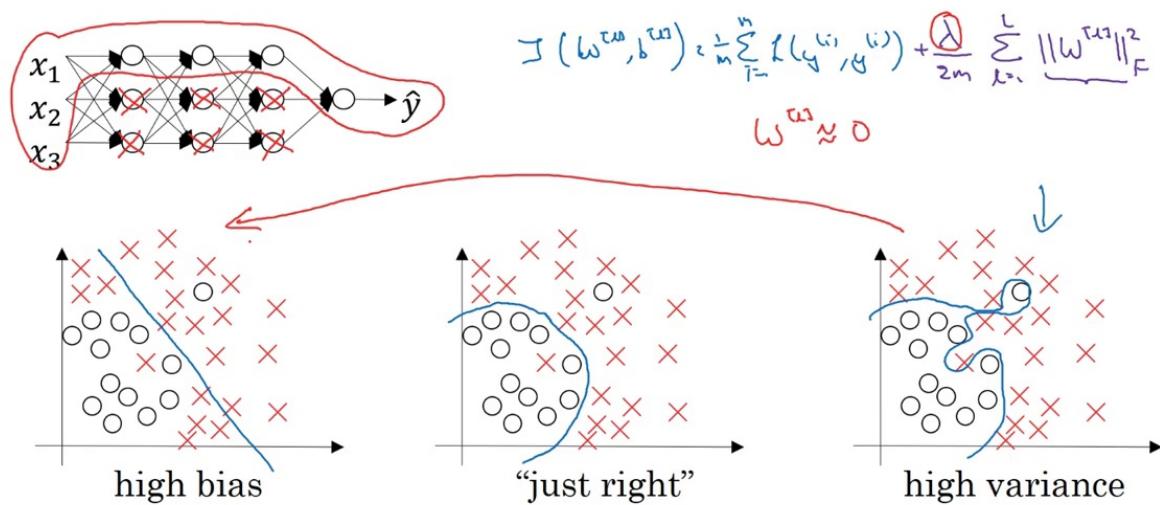
$$J(\omega^{(m)}, b^{(m)}) = \frac{1}{m} \sum_{i=1}^m L(y_i^{(i)}, \hat{y}_i^{(i)}) + \frac{\lambda}{2m} \sum_{l=1}^L \|\omega^{(l)}\|_F^2$$

$\omega^{(l)} \approx 0$



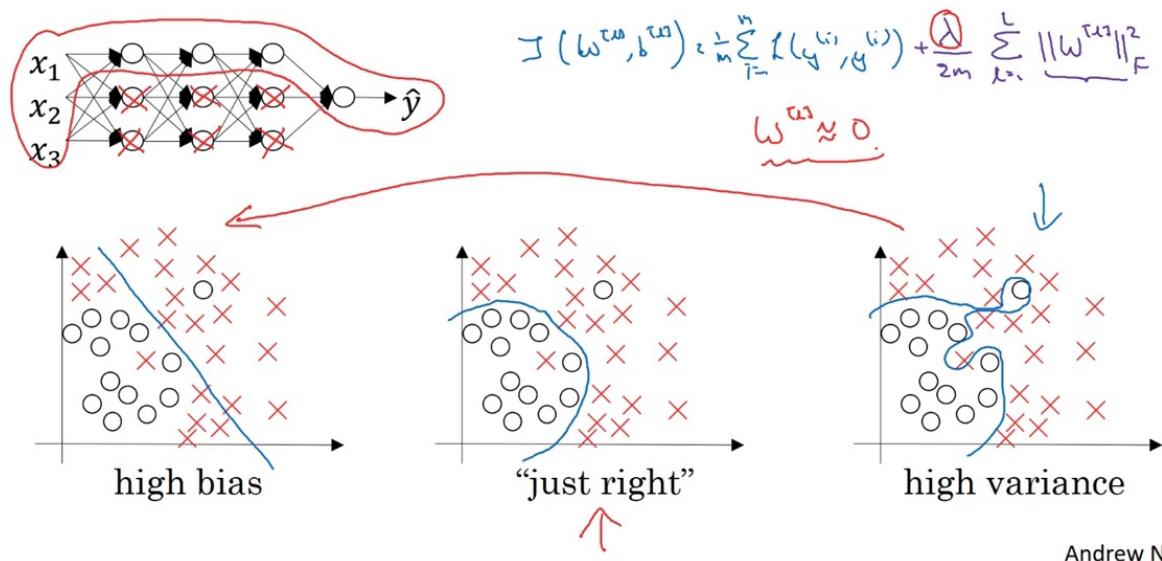
Andrew Ng

## How does regularization prevent overfitting?

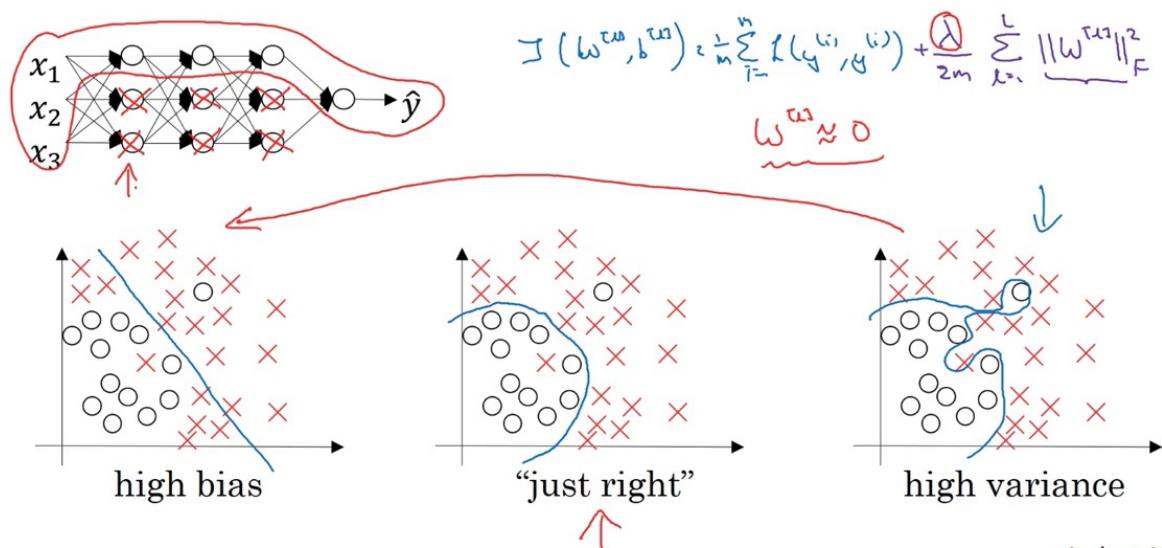


Andrew Ng

## How does regularization prevent overfitting?

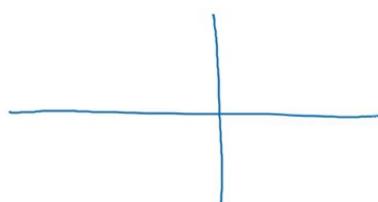


## How does regularization prevent overfitting?



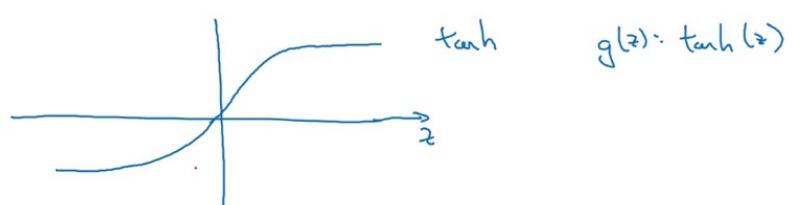
■■■: 00:03:00.000

How does regularization prevent overfitting?



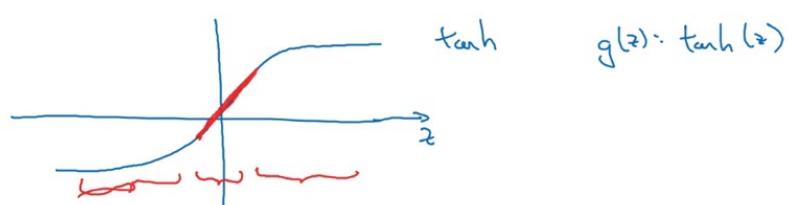
Andrew Ng

How does regularization prevent overfitting?



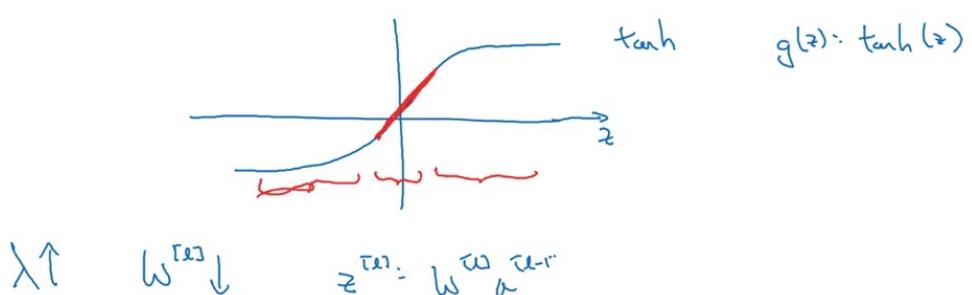
Andrew Ng

How does regularization prevent overfitting?



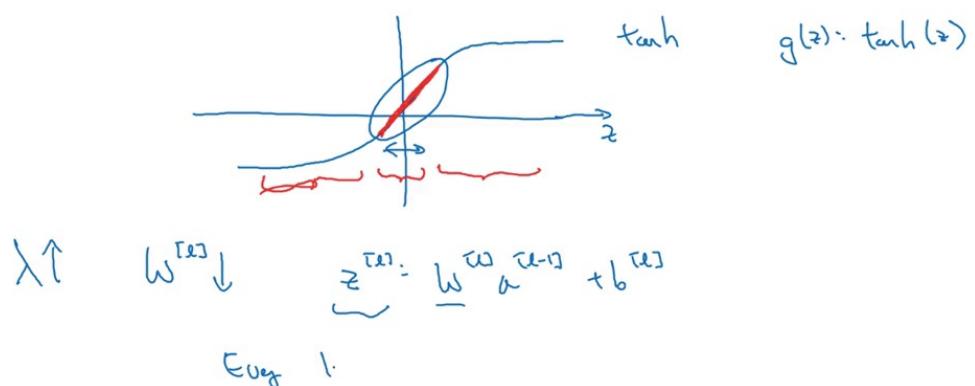
Andrew Ng

## How does regularization prevent overfitting?



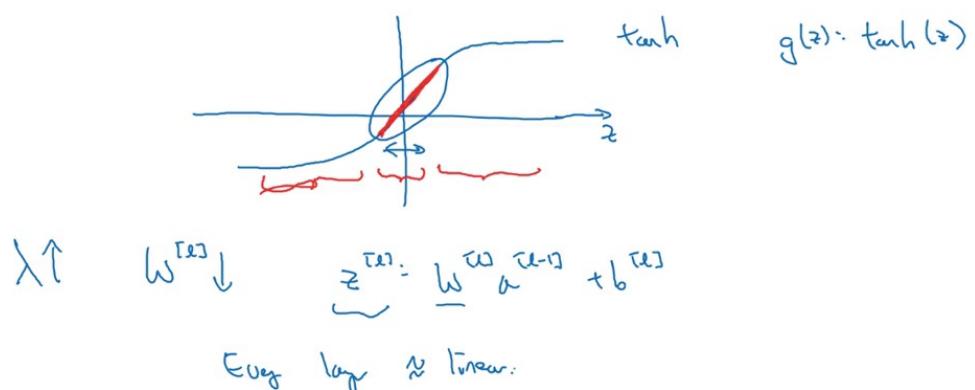
Andrew Ng

## How does regularization prevent overfitting?



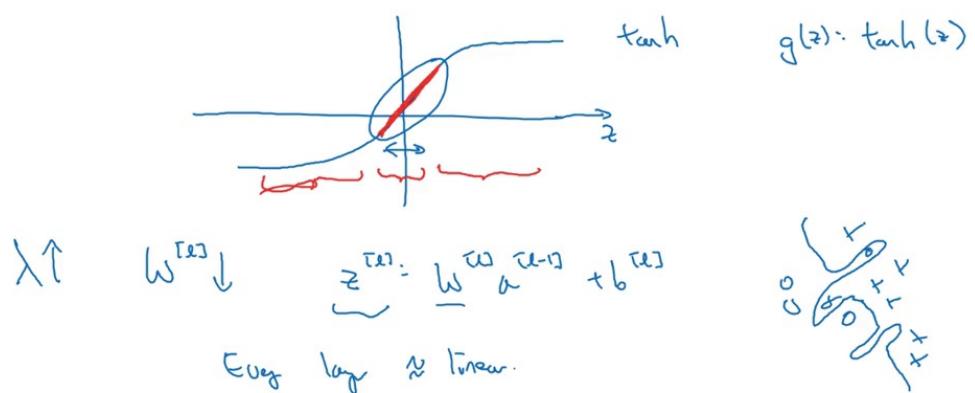
Andrew Ng

## How does regularization prevent overfitting?



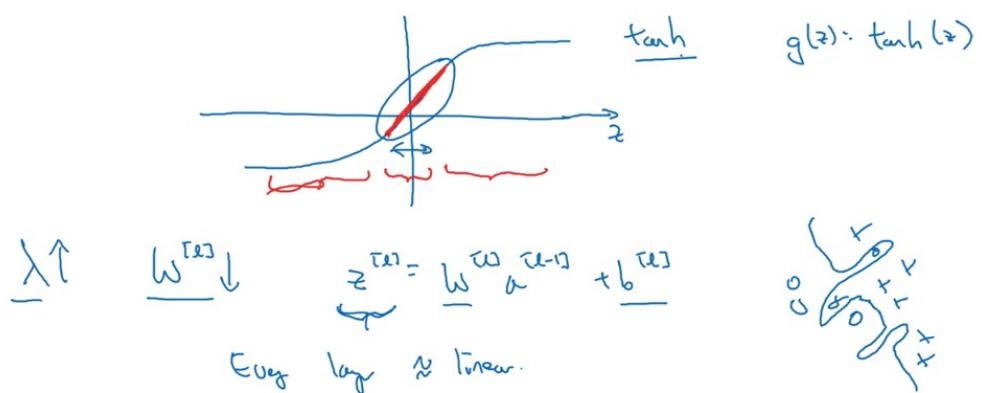
Andrew Ng

## How does regularization prevent overfitting?



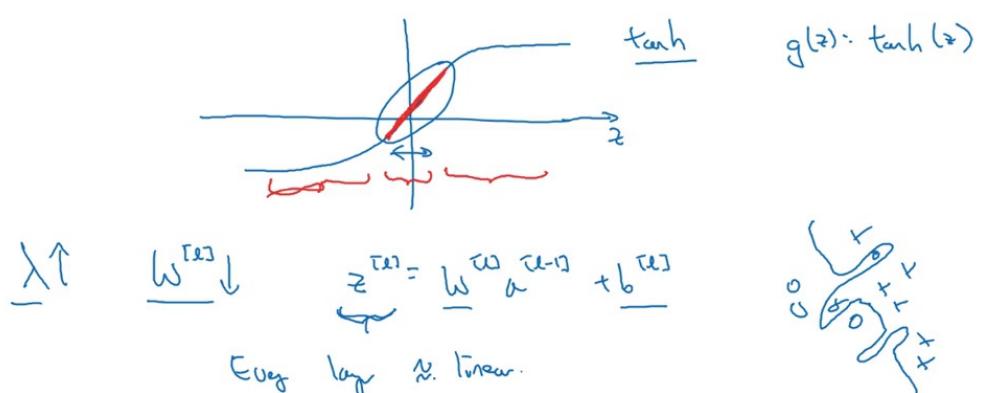
Andrew Ng

## How does regularization prevent overfitting?



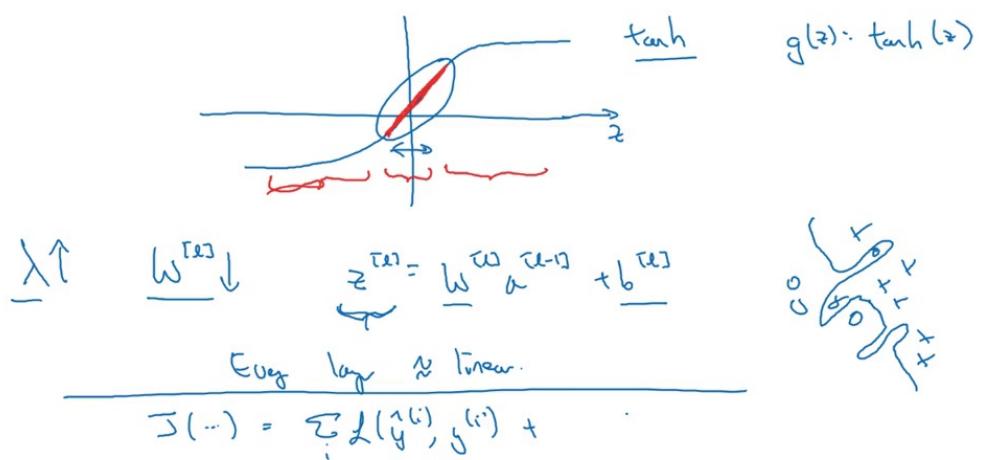
Andrew Ng

## How does regularization prevent overfitting?



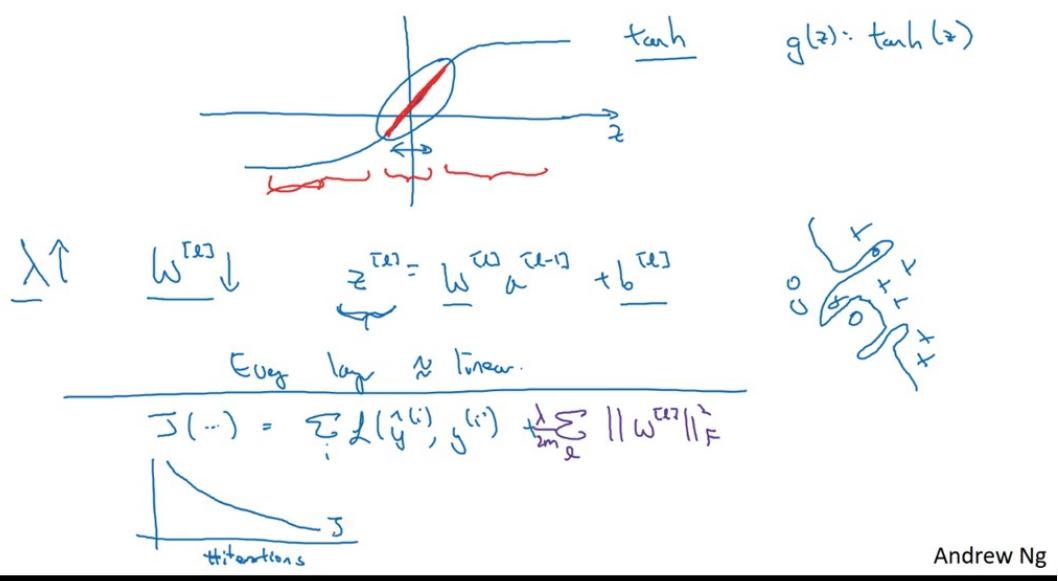
Andrew Ng

## How does regularization prevent overfitting?

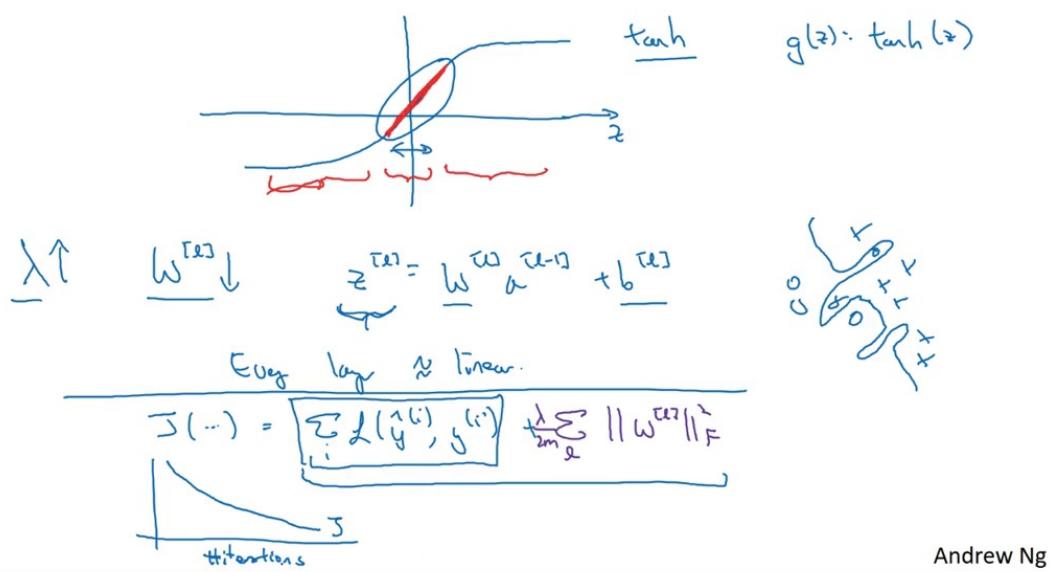


Andrew Ng

## How does regularization prevent overfitting?

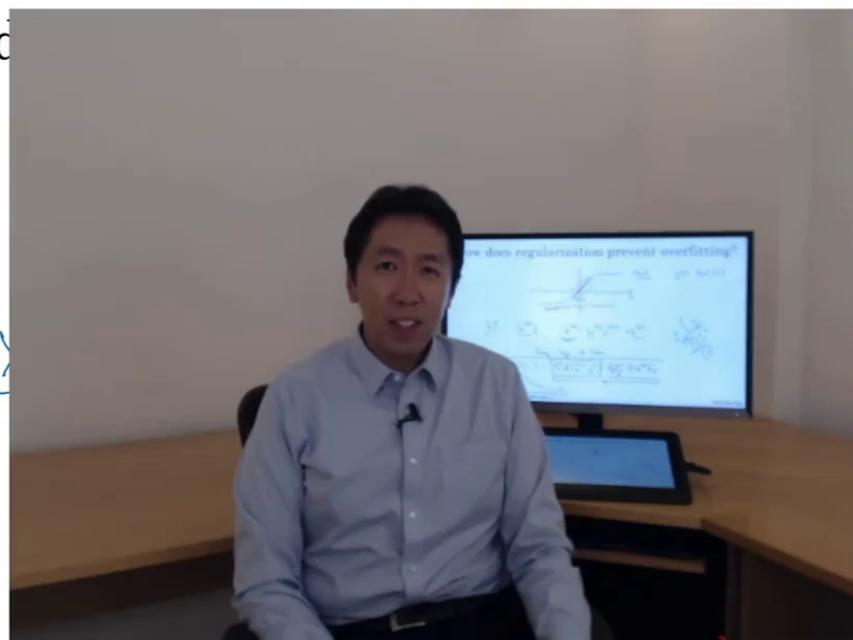


## How does regularization prevent overfitting?



■■■: 00:06:54.000

How does regularization prevent overfitting?



carh (≈)

Andrew Ng

■■■: 00:07:00.000

How do we prevent overfitting?

lambda (<math>\lambda</math>)



Andrew Ng

■ 26 ■ / ■ 26 ■

■■■: keyframe\_00-07-00-000\_0025.jpg