



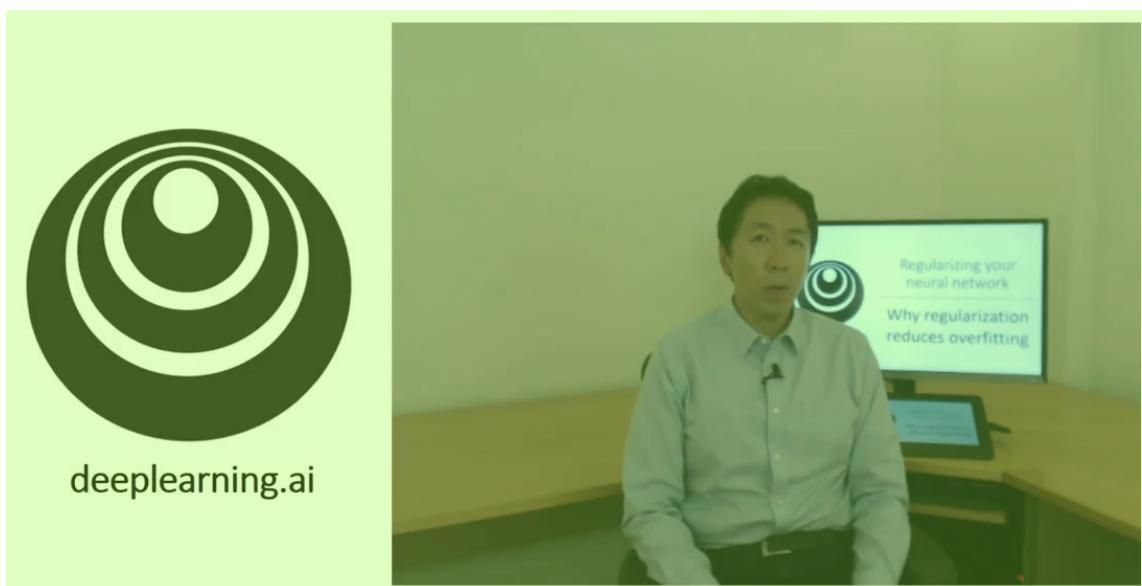
3

Regularizing your²
neural network

Why regularization⁴
reduces overfitting

↳

[■■■] 00:00:03.000 1



[■ 2 ■ / ■ 26 ■]
[■■]: keyframe_00-00-03-000_0001.jpg



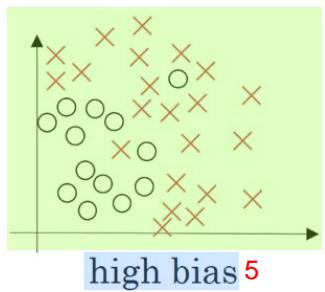
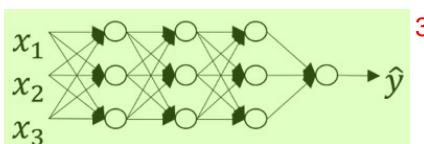
3

Regularizing your²
neural network

Why regularization⁴
reduces overfitting

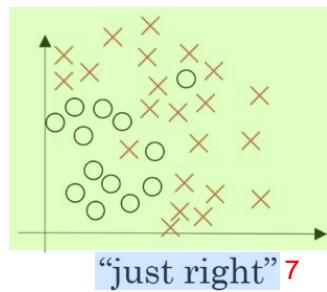
↳

How does regularization prevent overfitting?²



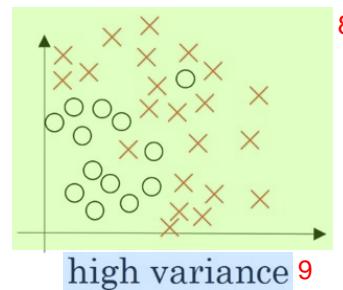
4

high bias 5



6

“just right” 7

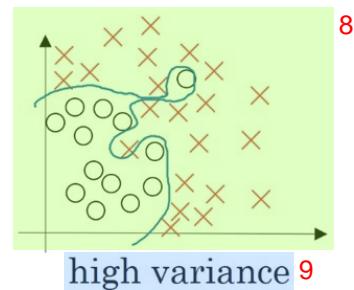
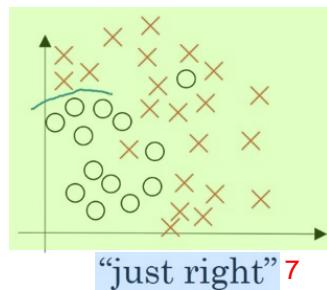
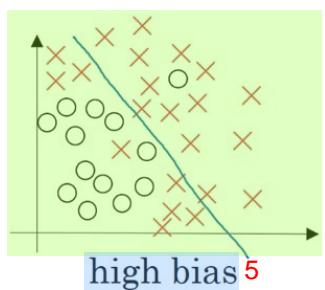
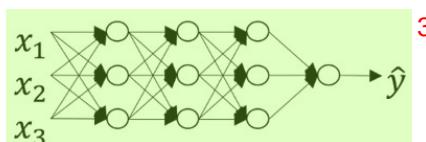


8

high variance 9

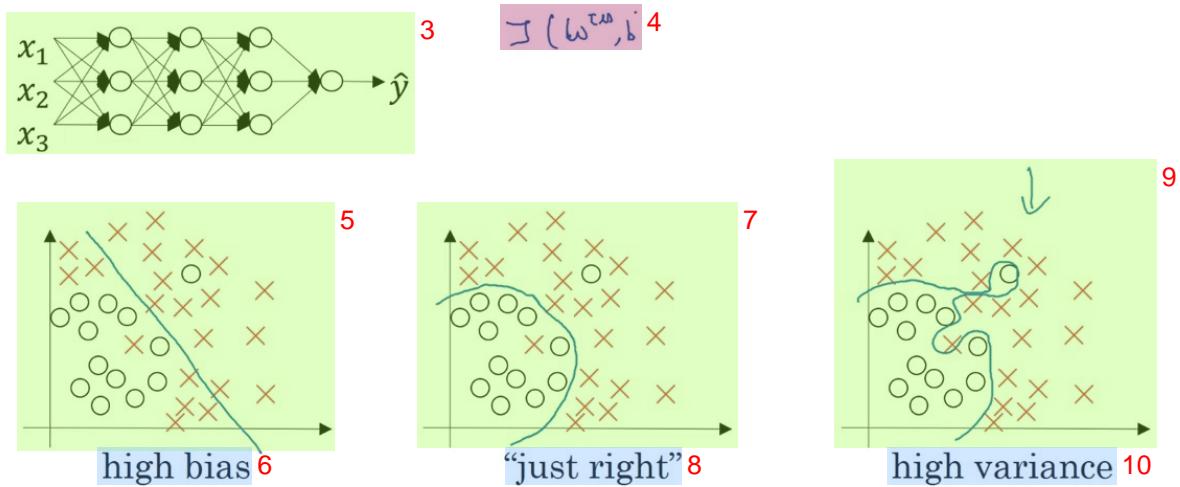
Andrew Ng 10

How does regularization prevent overfitting?²



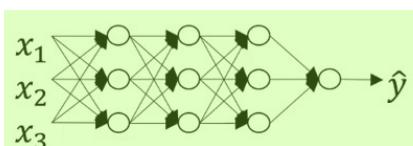
Andrew Ng 10

How does regularization prevent overfitting?²



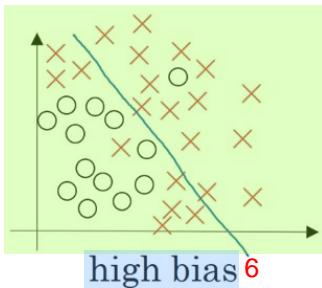
Andrew Ng 11

How does regularization prevent overfitting?²

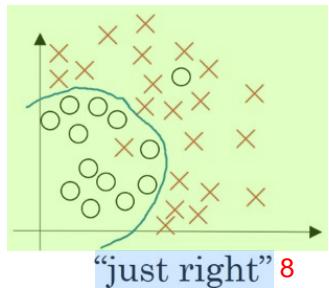


3

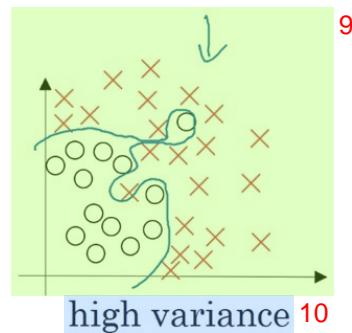
$$J(\omega^{(0)}, b^{(0)}) = \frac{1}{m} \sum_{i=1}^m L(y^{(i)}, \hat{y}^{(i)}) + \frac{\lambda}{2m} \sum_{l=1}^L \|\omega^{(l)}\|^2$$



5



7

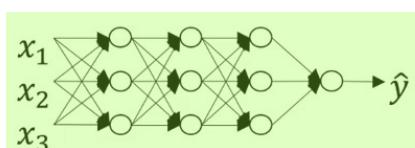


9

high variance 10

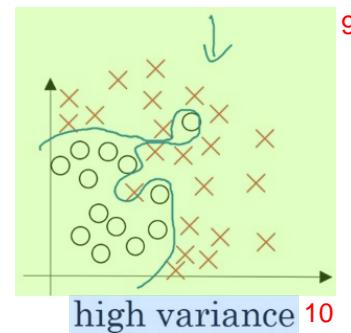
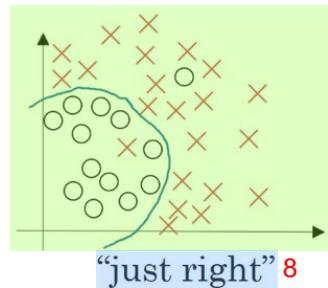
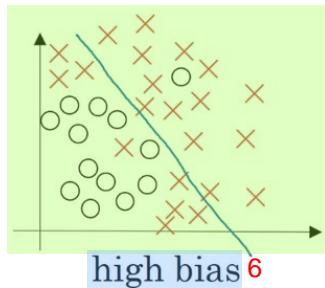
Andrew Ng 11

How does regularization prevent overfitting? ²



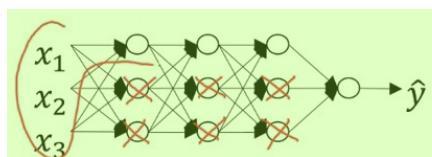
$$J(\omega^{(m)}, b^{(m)}) = \frac{1}{m} \sum_{i=1}^m L(y_i^{(i)}, \hat{y}_i^{(i)}) + \frac{\lambda}{2m} \sum_{l=1}^L \|\omega^{(l)}\|_F^2$$

4



Andrew Ng 11

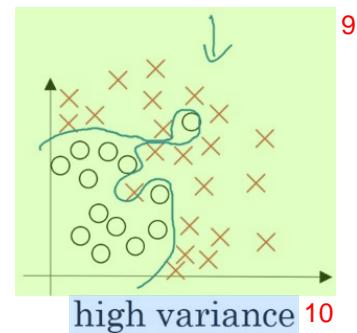
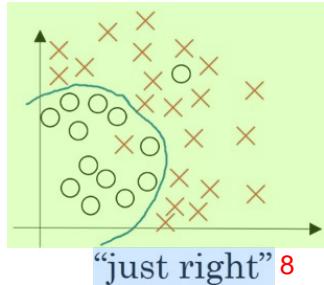
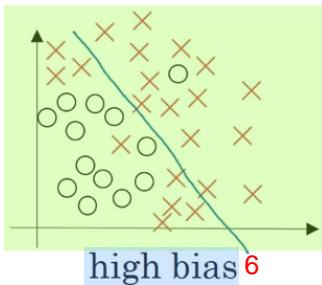
How does regularization prevent overfitting? ²



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

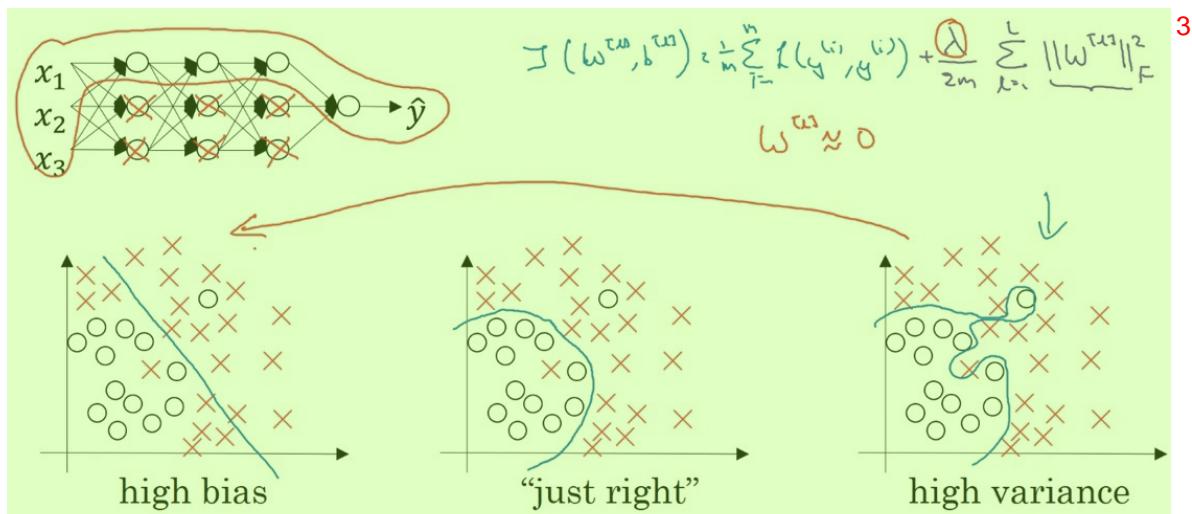
$\omega^{(l)} \approx 0$

4



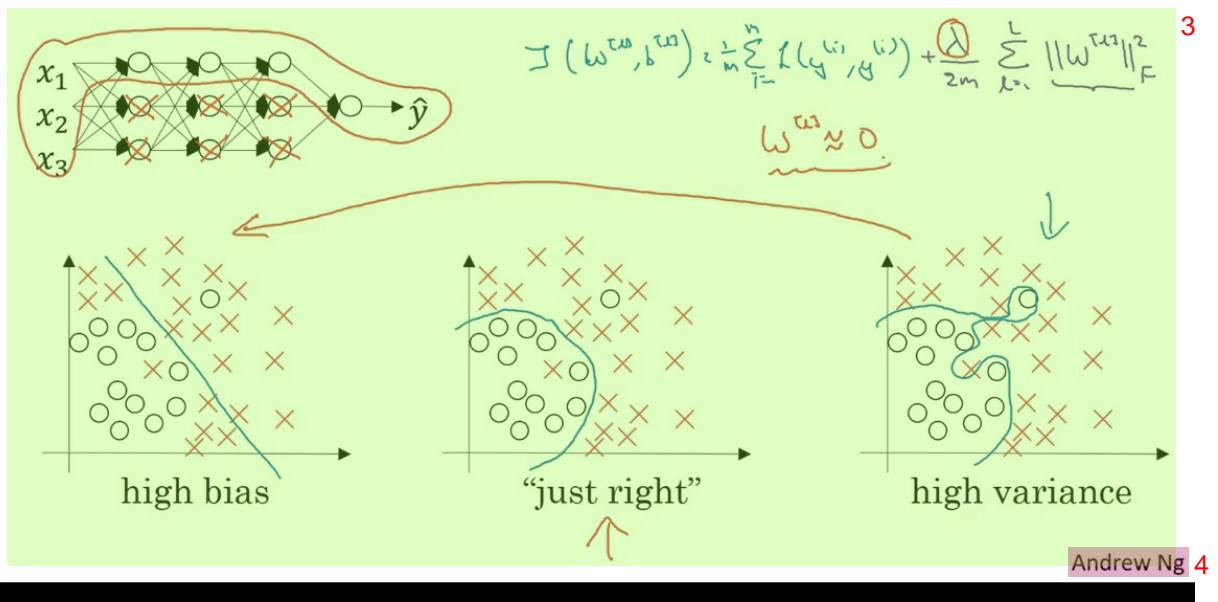
Andrew Ng 11

How does regularization prevent overfitting?²

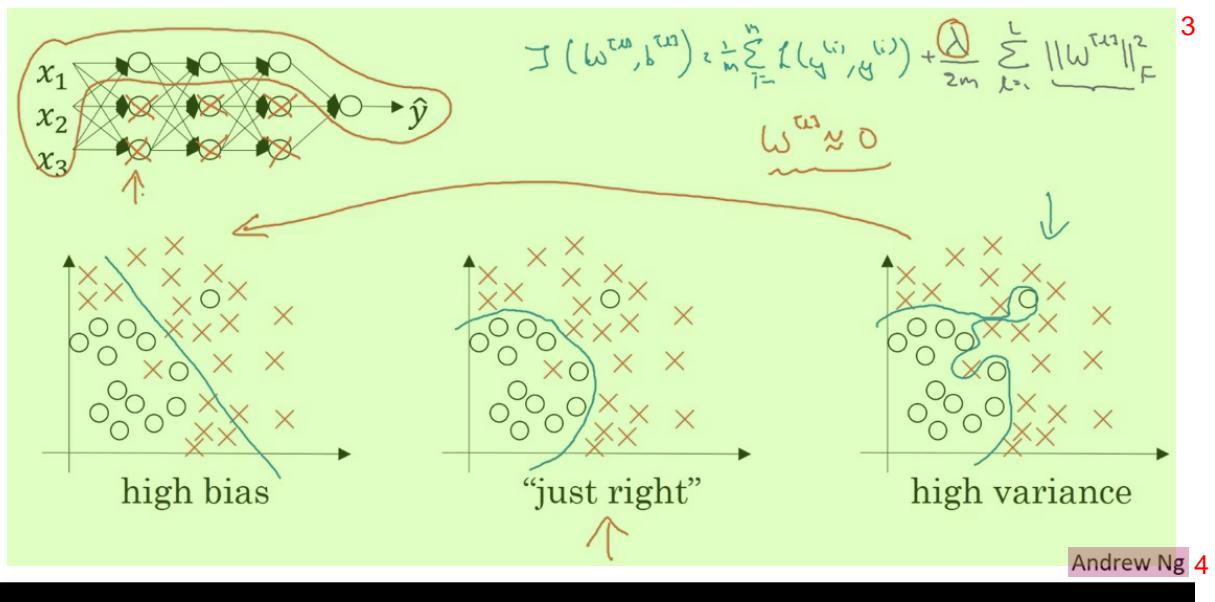


Andrew Ng 4

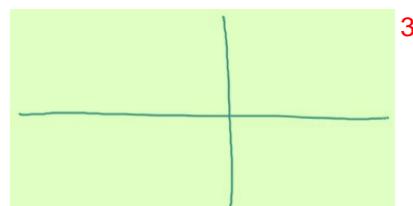
How does regularization prevent overfitting? ²



How does regularization prevent overfitting? ²



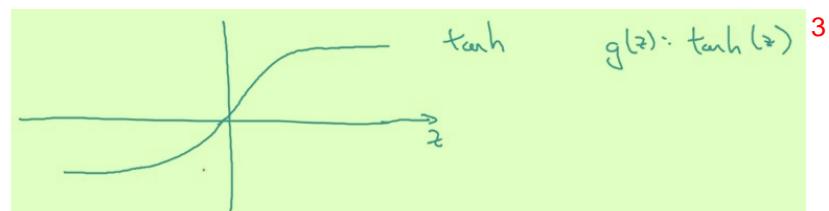
How does regularization prevent overfitting?²



3

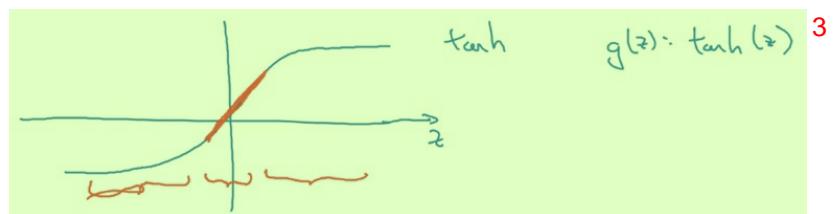
Andrew Ng

How does regularization prevent overfitting?²



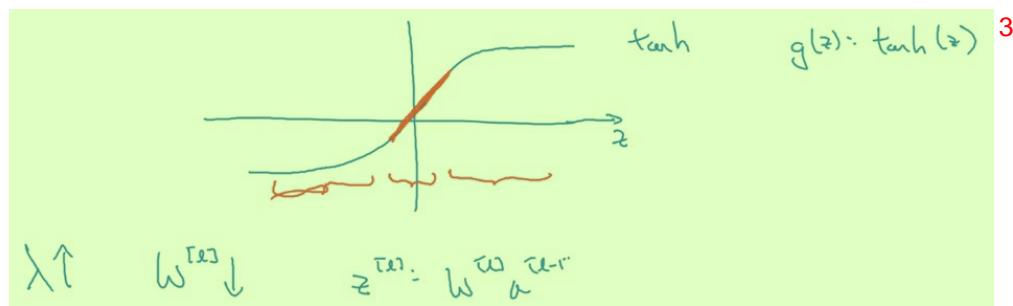
Andrew Ng

How does regularization prevent overfitting?²



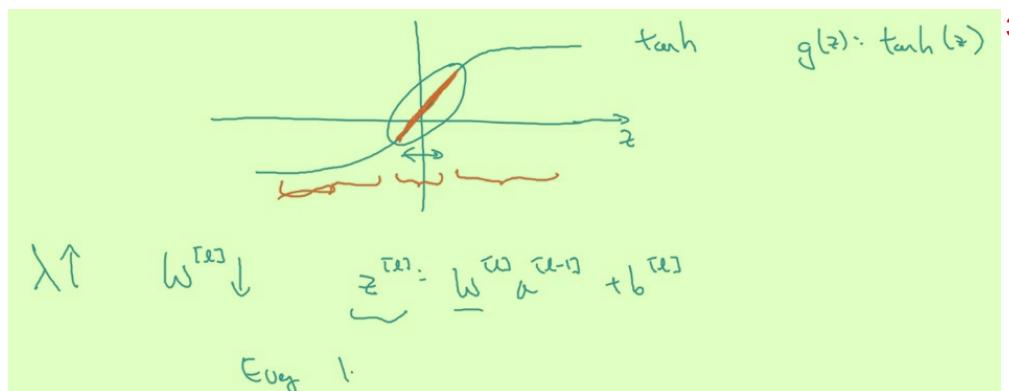
Andrew Ng

How does regularization prevent overfitting? ²



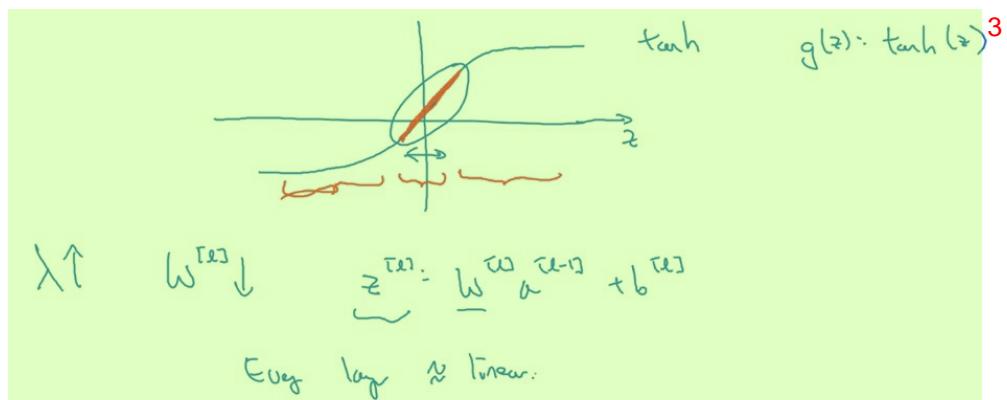
Andrew Ng

How does regularization prevent overfitting? ²



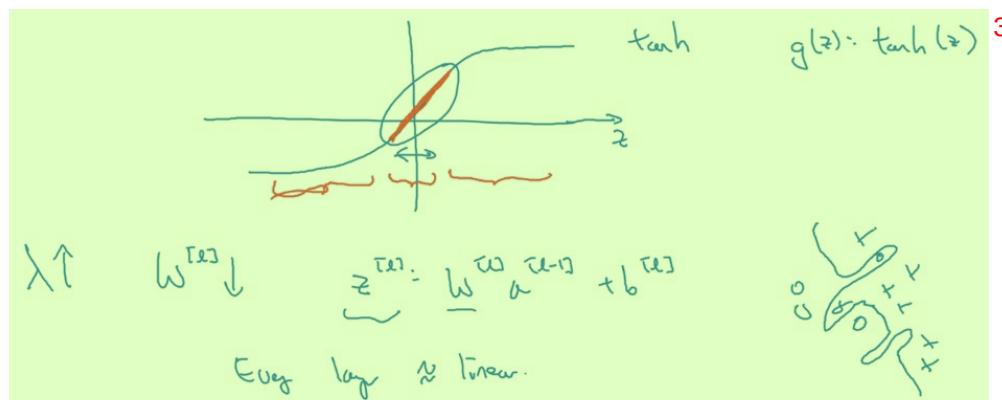
Andrew Ng

How does regularization prevent overfitting? ²



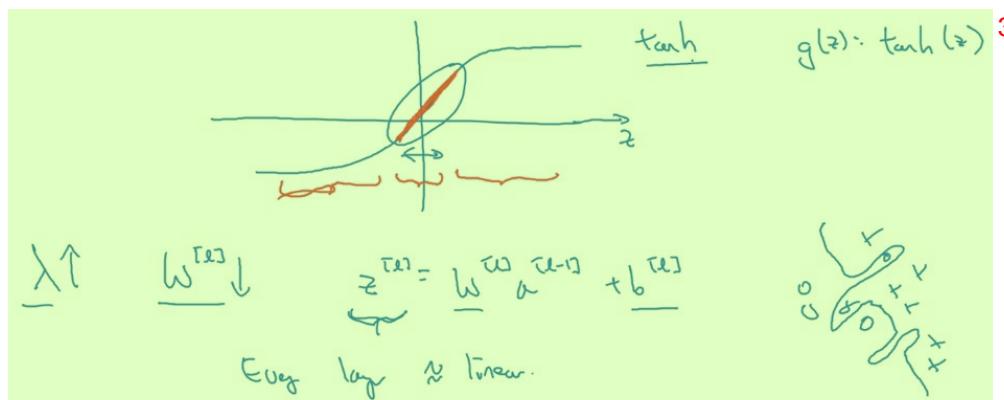
Andrew Ng

How does regularization prevent overfitting?²



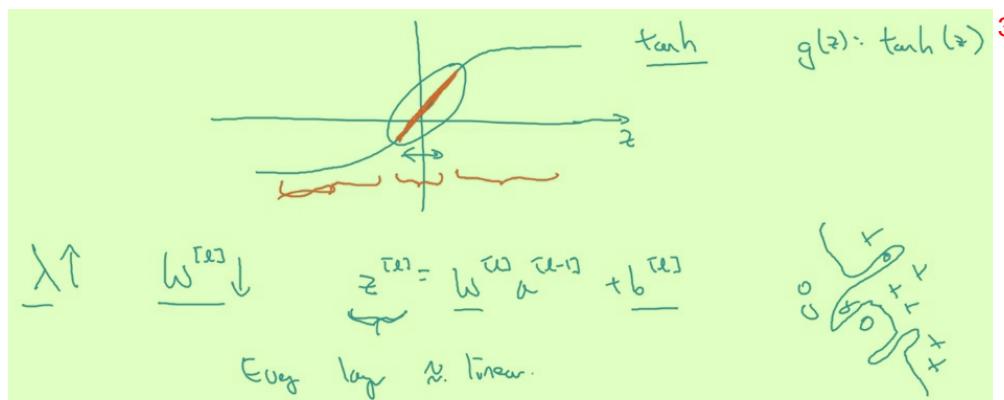
Andrew Ng

How does regularization prevent overfitting? ²



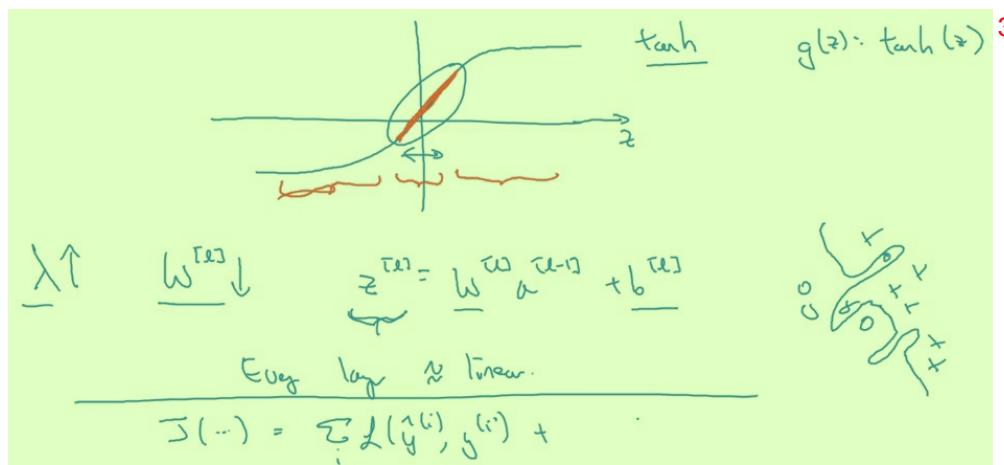
Andrew Ng

How does regularization prevent overfitting? ²



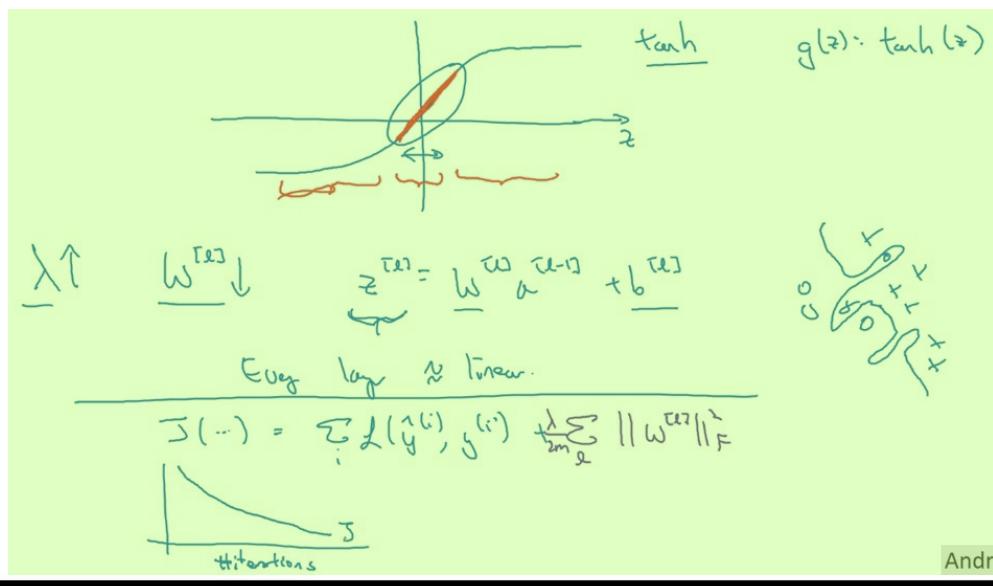
Andrew Ng

How does regularization prevent overfitting? ²

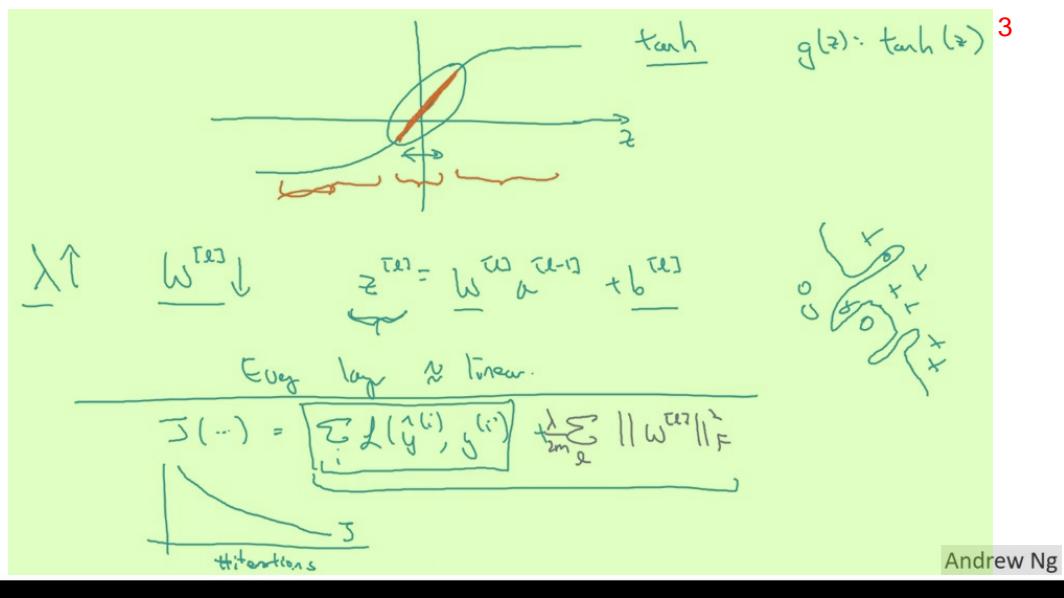


Andrew Ng

How does regularization prevent overfitting? ²



How does regularization prevent overfitting? ²



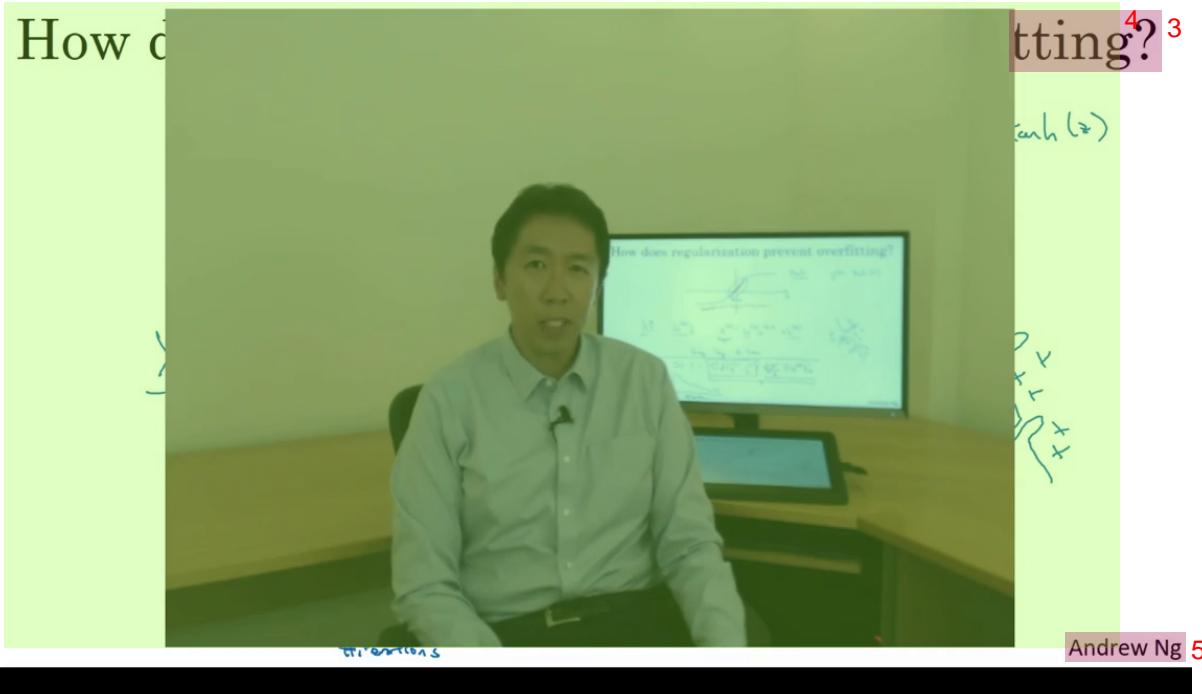
[■■■: 00:06:54.000] 1

How does regularization prevent overfitting? 2 3 4 5



Andrew Ng 5

[■■■: 00:07:00.000] 1



[■ 26 ■ / ■ 26 ■

[■■: keyframe_00-07-00-000_0025.jpg]