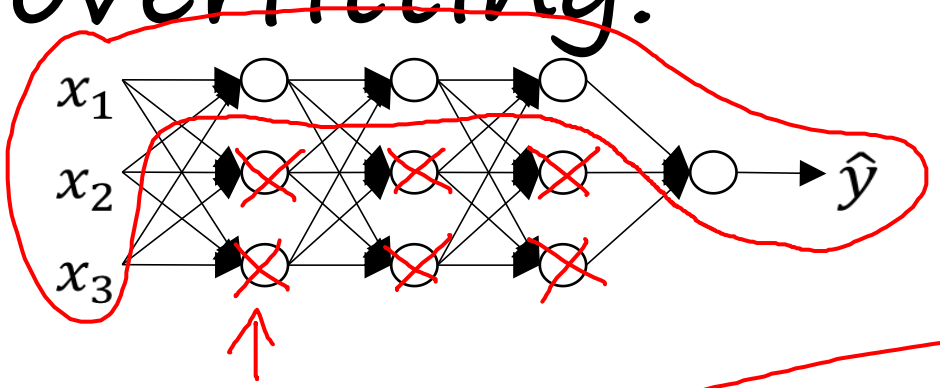
deeplearning.ai

# Regularizing your neural network

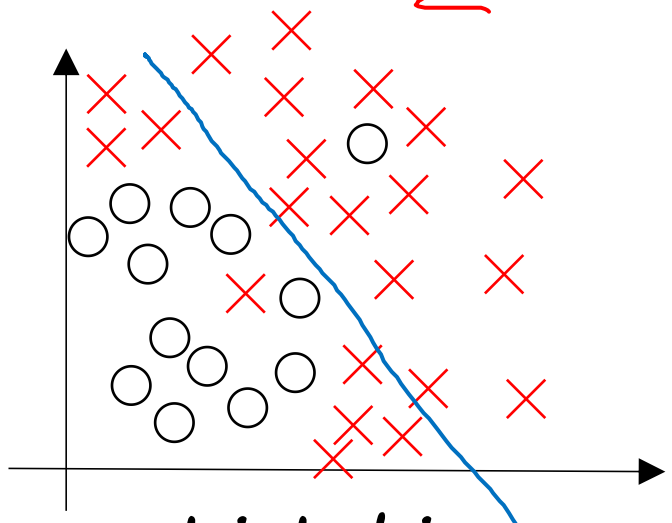
# Why regularization reduces overfitting

# How does regularization prevent overfitting?

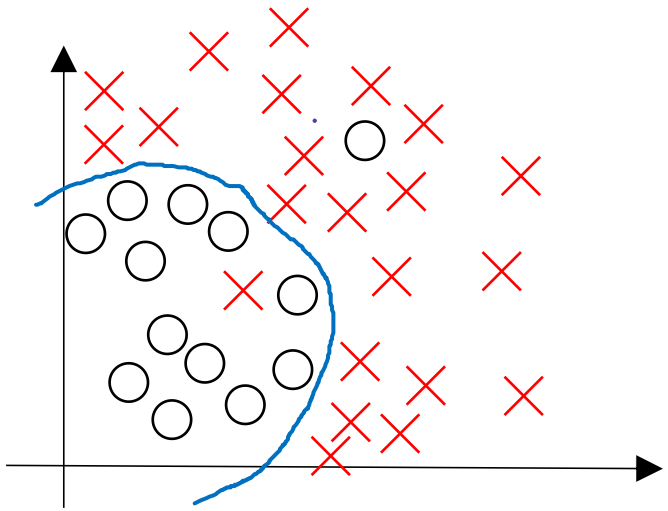



$$J(w^{[l]}, b^{[l]}) = \frac{1}{m}\sum_{i=1}^{m} \mathcal{L}(\hat{y}^{(i)}, y^{(i)}) + \frac{\lambda}{2m}\sum_{l=1}^{L} \|w^{[l]}\|_F^2$$

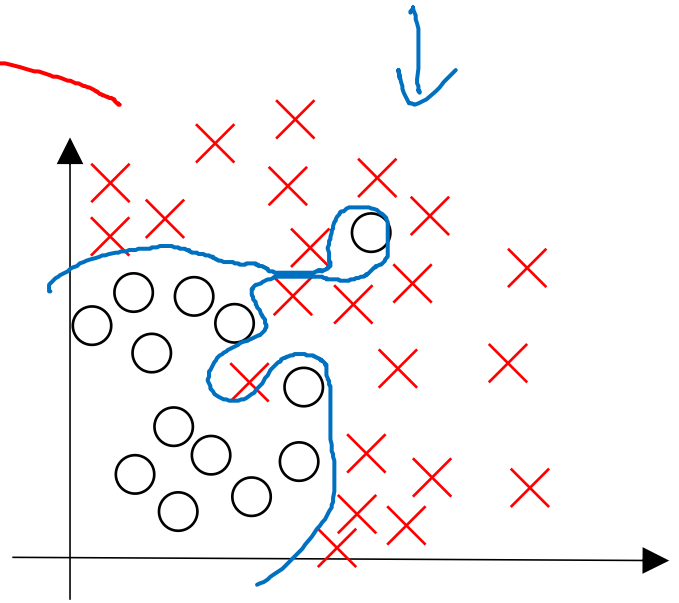
$w^{[l]} \approx 0$



high bias



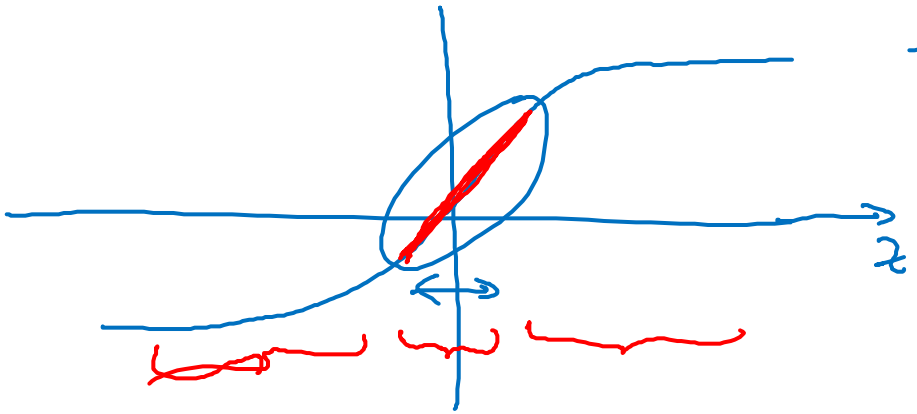
"just right"



high variance

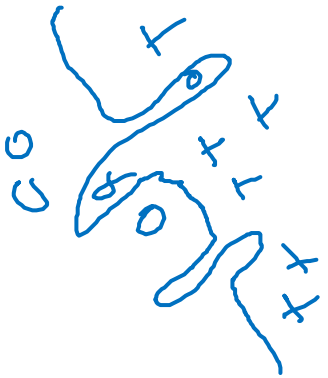
Andrew Ng

# How does regularization prevent overfitting?



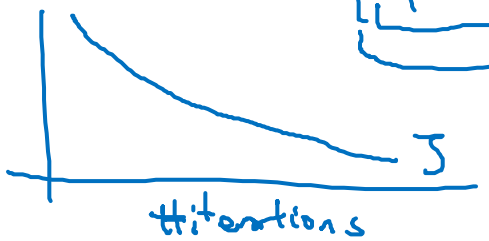

$g(z) = \tanh(z)$

$\lambda \uparrow \quad W^{[l]} \downarrow \quad z^{[l]} = W^{[l]} a^{[l-1]} + b^{[l]}$



Every layer $\approx$ linear

$$J(\ldots) = \sum_{i} \mathcal{L}(\hat{y}^{(i)}, y^{(i)}) + \frac{\lambda}{2m} \sum_{l} \| W^{[l]} \|_F^2$$




Andrew Ng