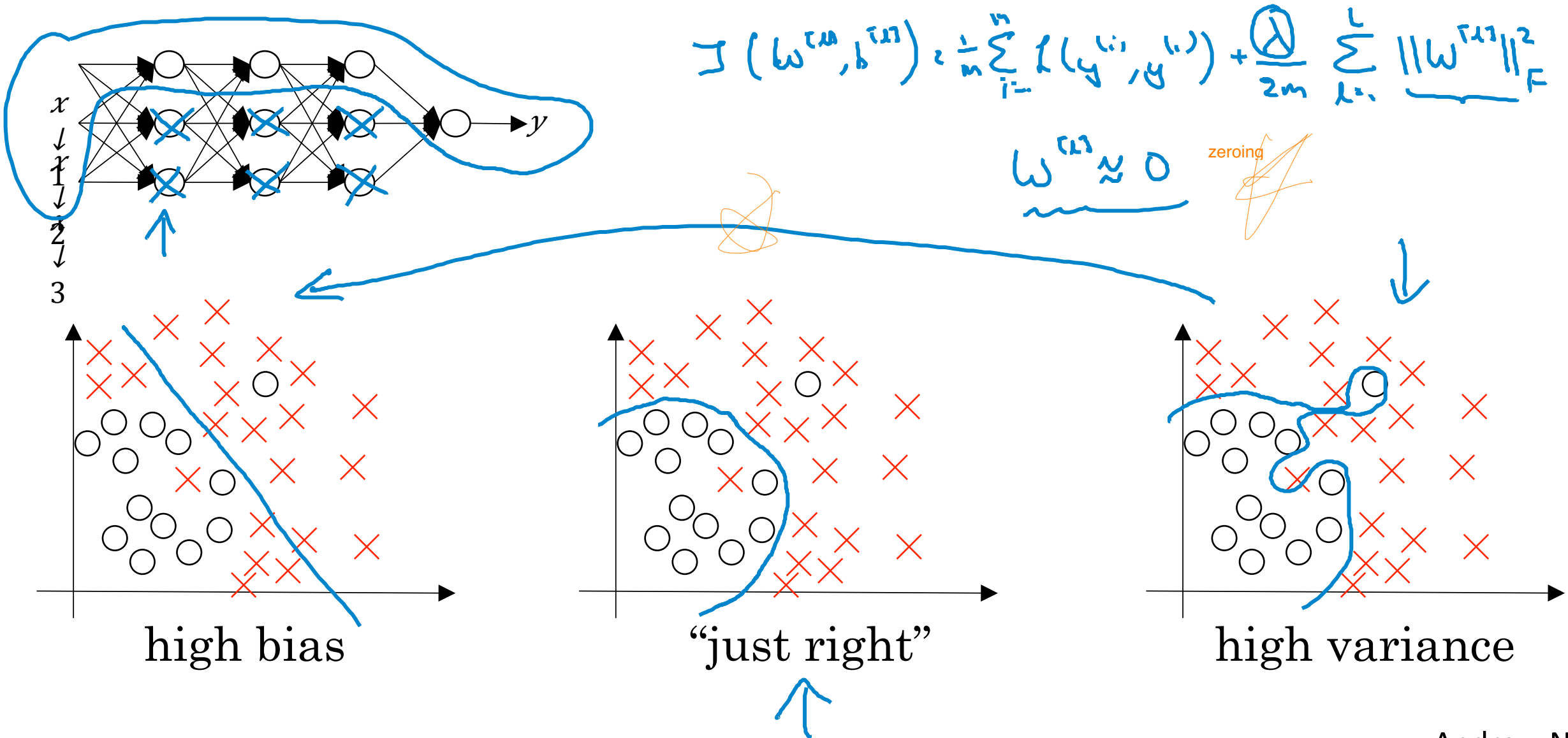Regularizing your
neural network

Why regularization
reduces overfitting

이전 강의 ppt 뒷부분에 동일한 내용 붙어있는데 거기에 정리했음.

deeplearning.ai

How does regularization prevent overfitting?

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

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

zeroing

high bias   "just right"   high variance

Andrew Ng

How does regularization prevent overfitting?

tanh

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

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

Every layer $\approx$ linear

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

Andrew Ng