1 ML Week 5

Overfitting and Regularization

- If one neural network overfits the training set, one reasonable step is to increase the regularization parameter λ.
- For computational efficiency, after we performed gradient checking to verify that our back-propagation code is correct, we usually disable gradient checkking before using back-propagation to train the network

Exercise

Let $J(\theta) = 3\theta^2 + 2$

Let $\theta = 1$ and $\epsilon = 0.01$

Use the formula to numerically compute an approximation to the derivative of θ at $theta = 1$

$$\frac{J(\theta + \epsilon) - J(\theta + \epsilon)}{2\epsilon}$$

$$= \frac{(3(1.01)^2 + 2) - (3(0.99)^2 + 2)}{0.002} = 9.003$$