



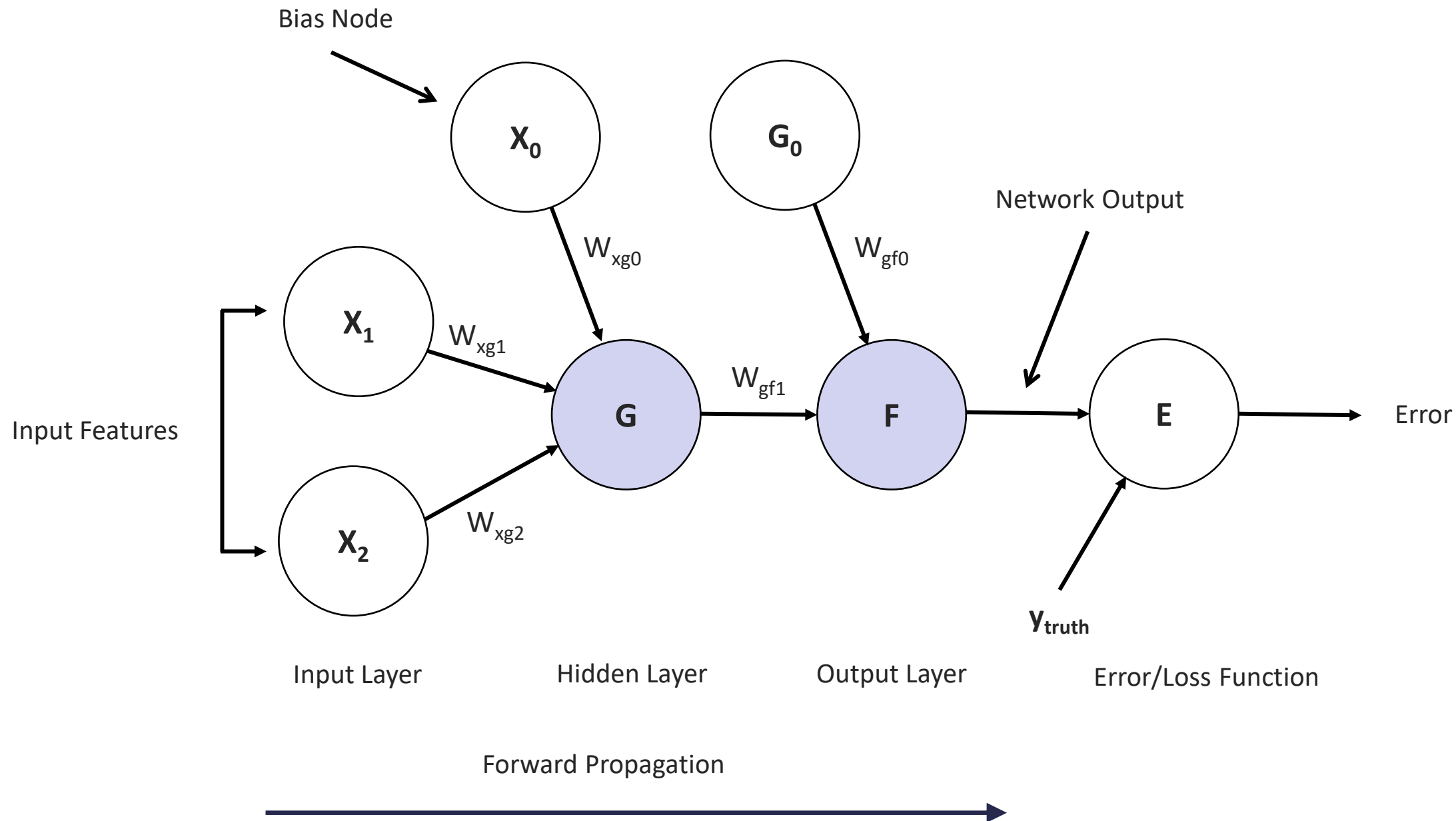
Gradient Descent

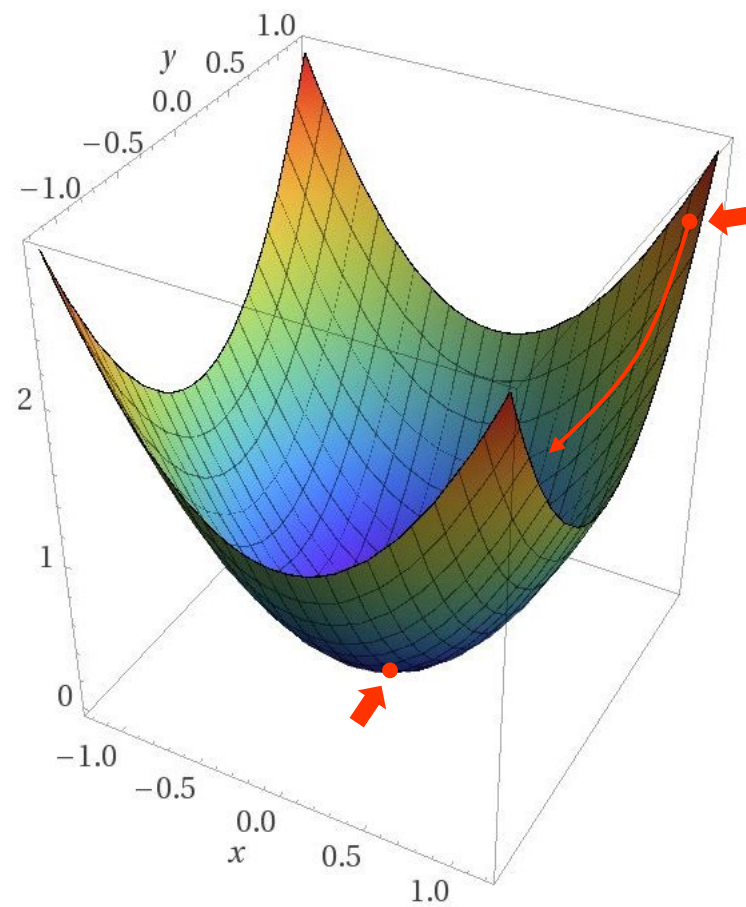


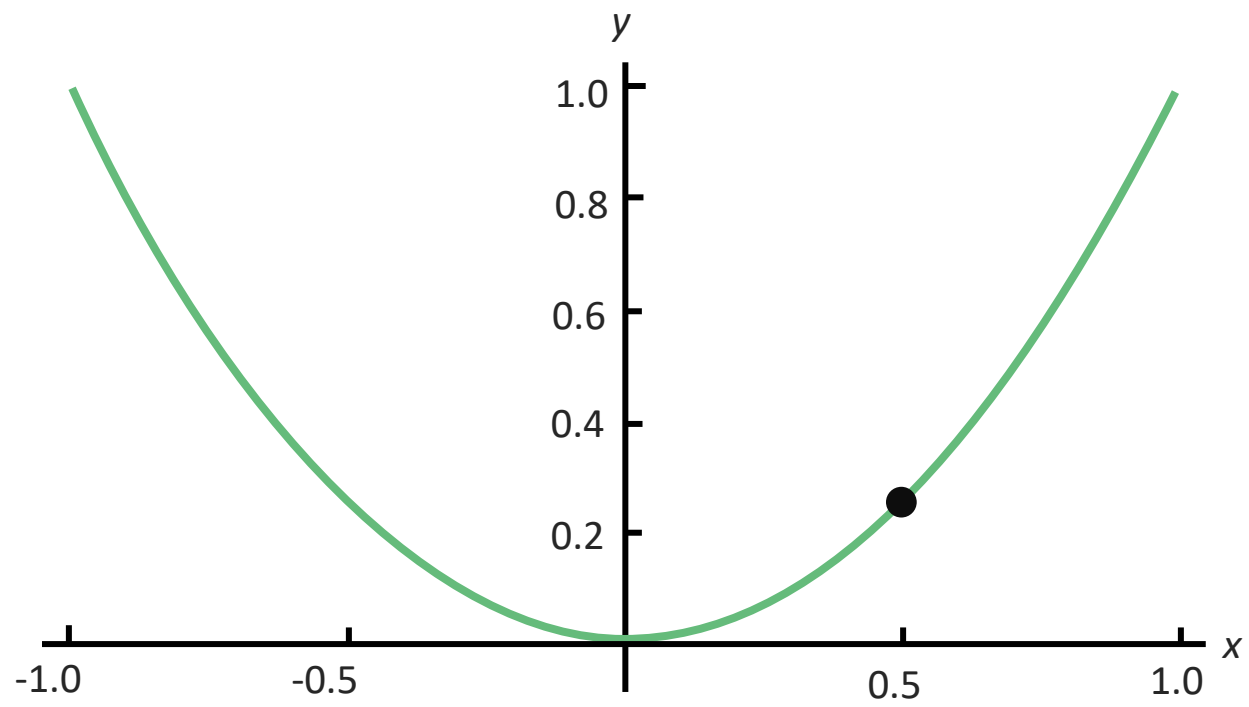
Lesson Overview

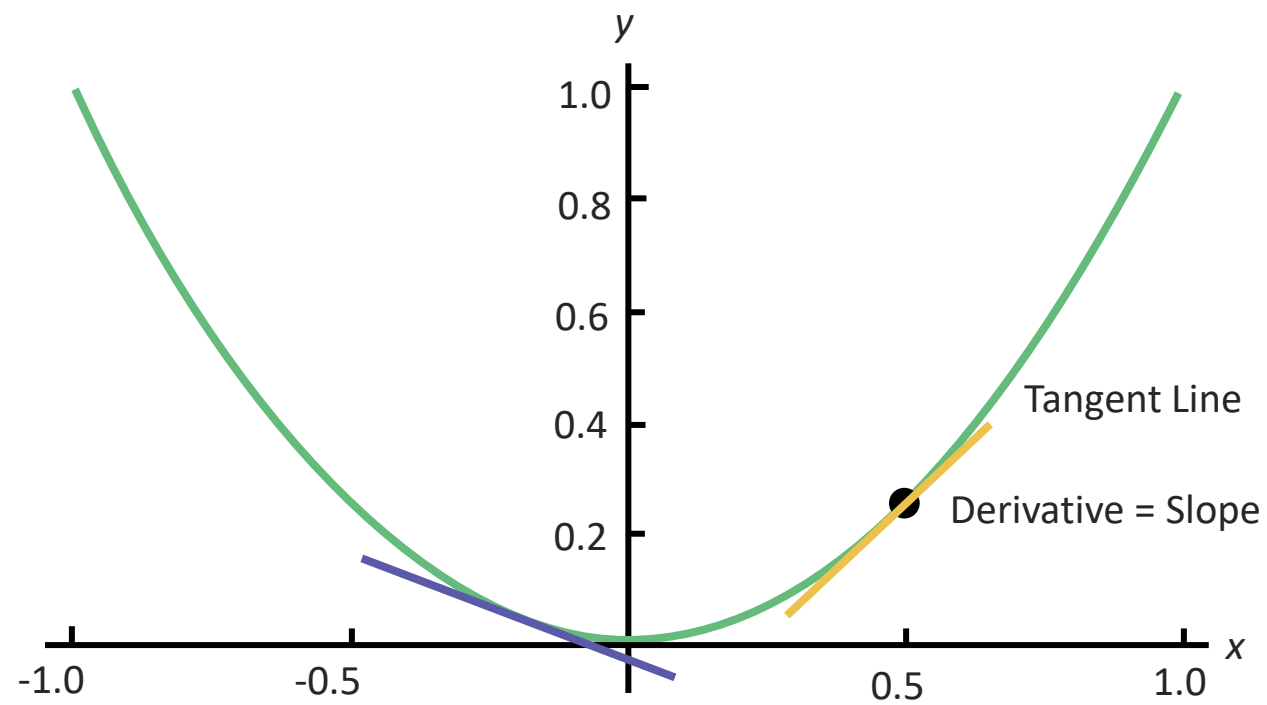
1. The Big Picture
2. The Loss Function
3. A Race to the Bottom
4. Optimizers

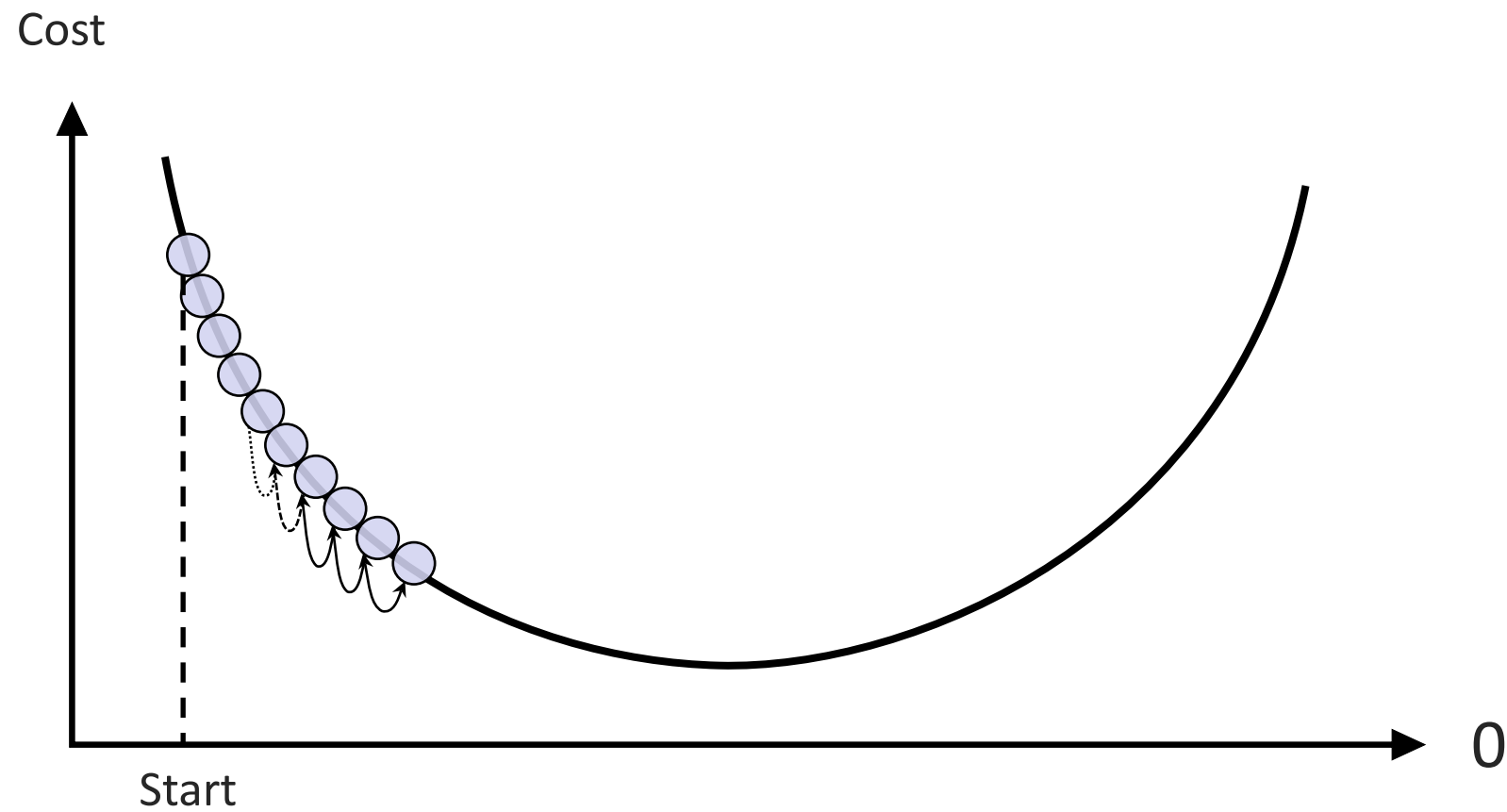






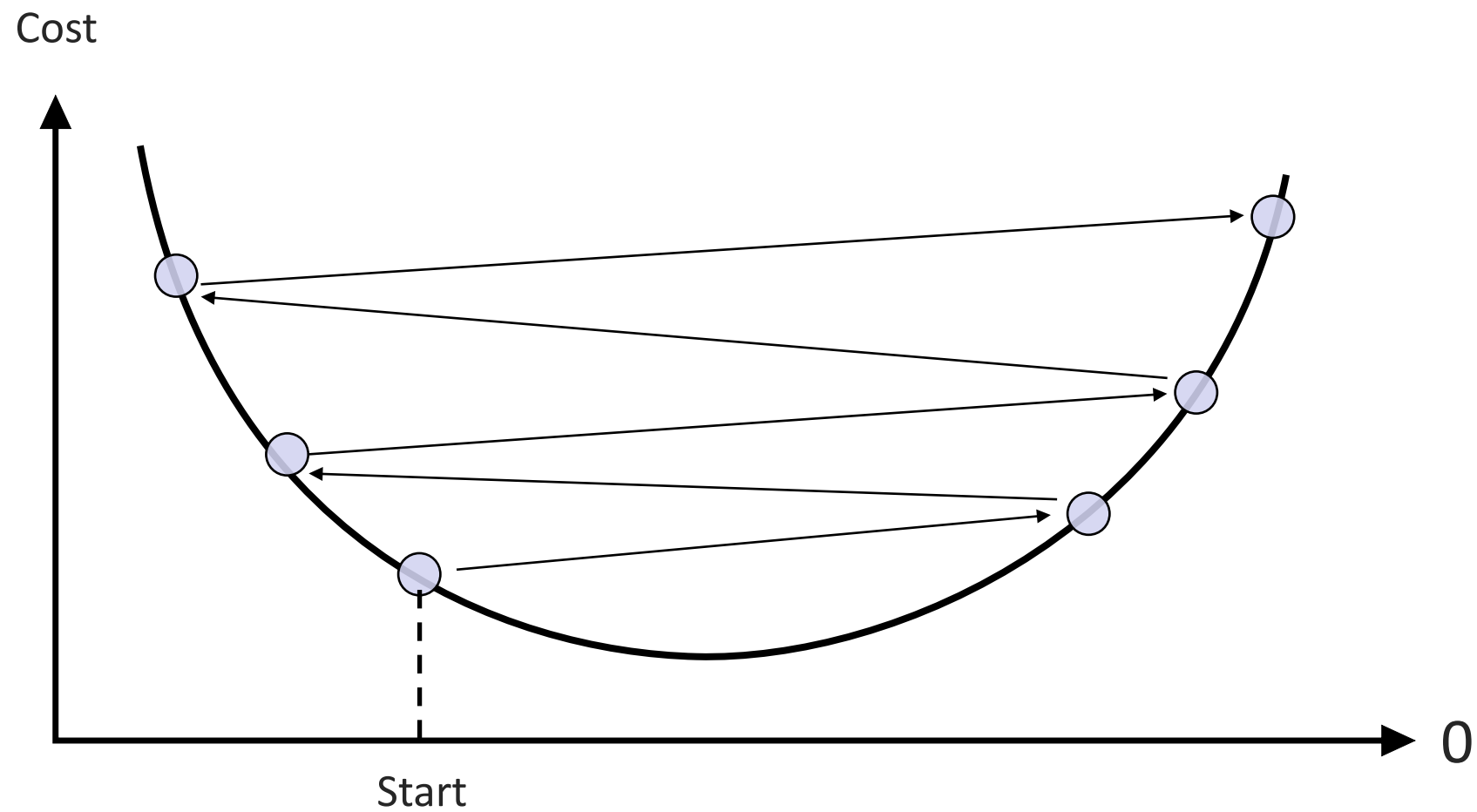






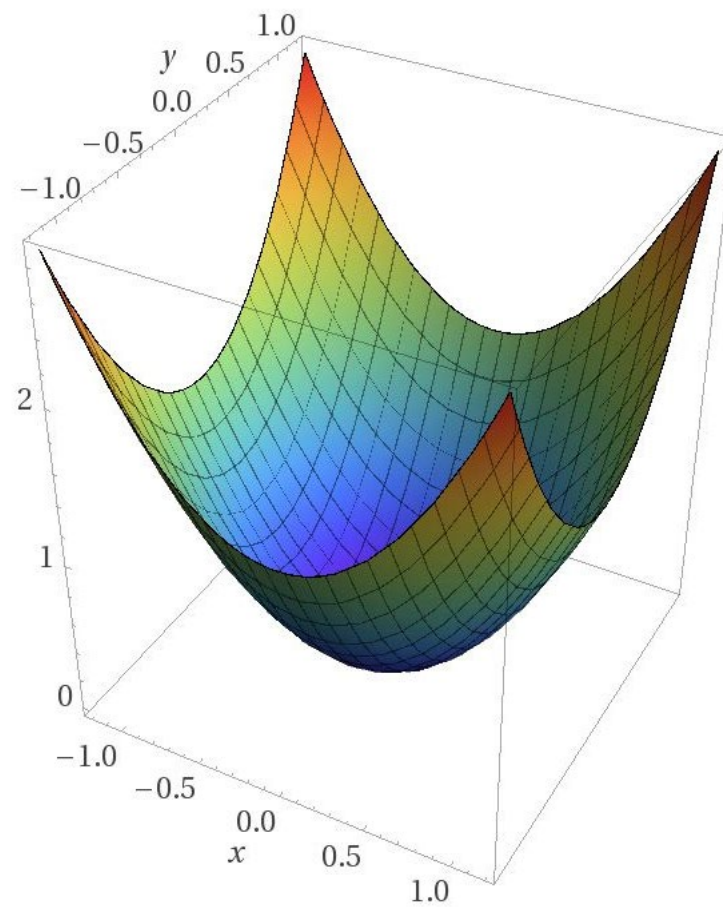
The learning rate is too small





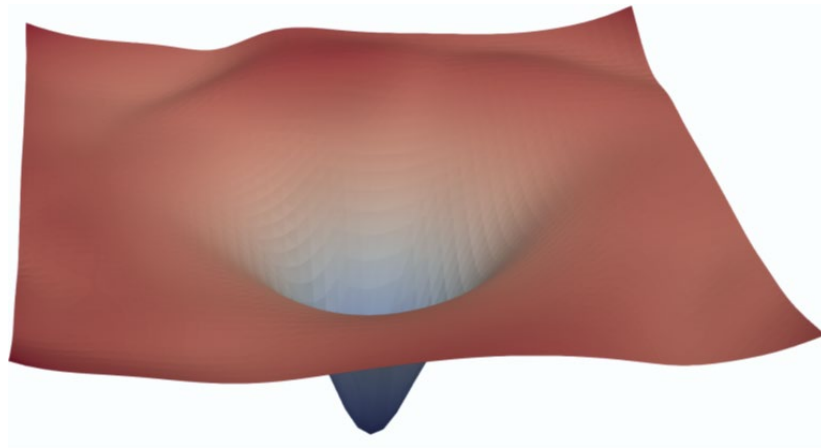
The learning rate is too large



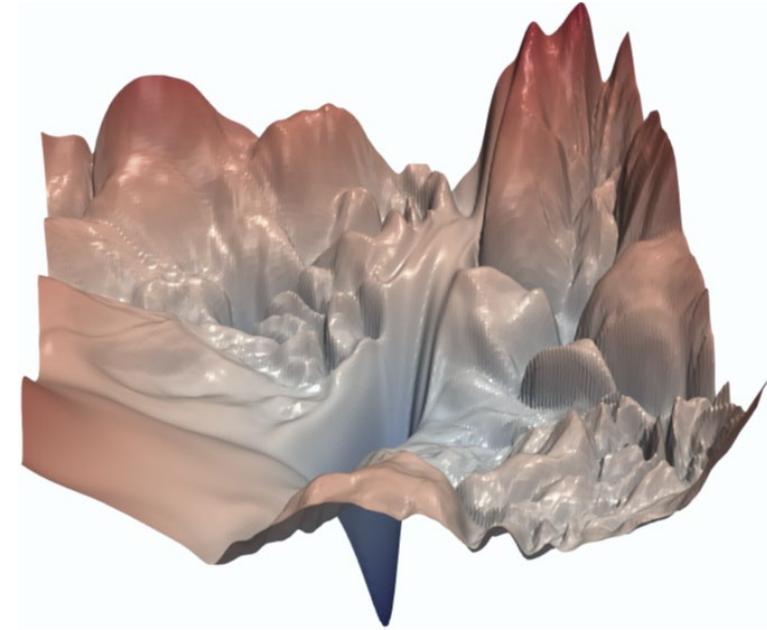




ResNet-56 Model

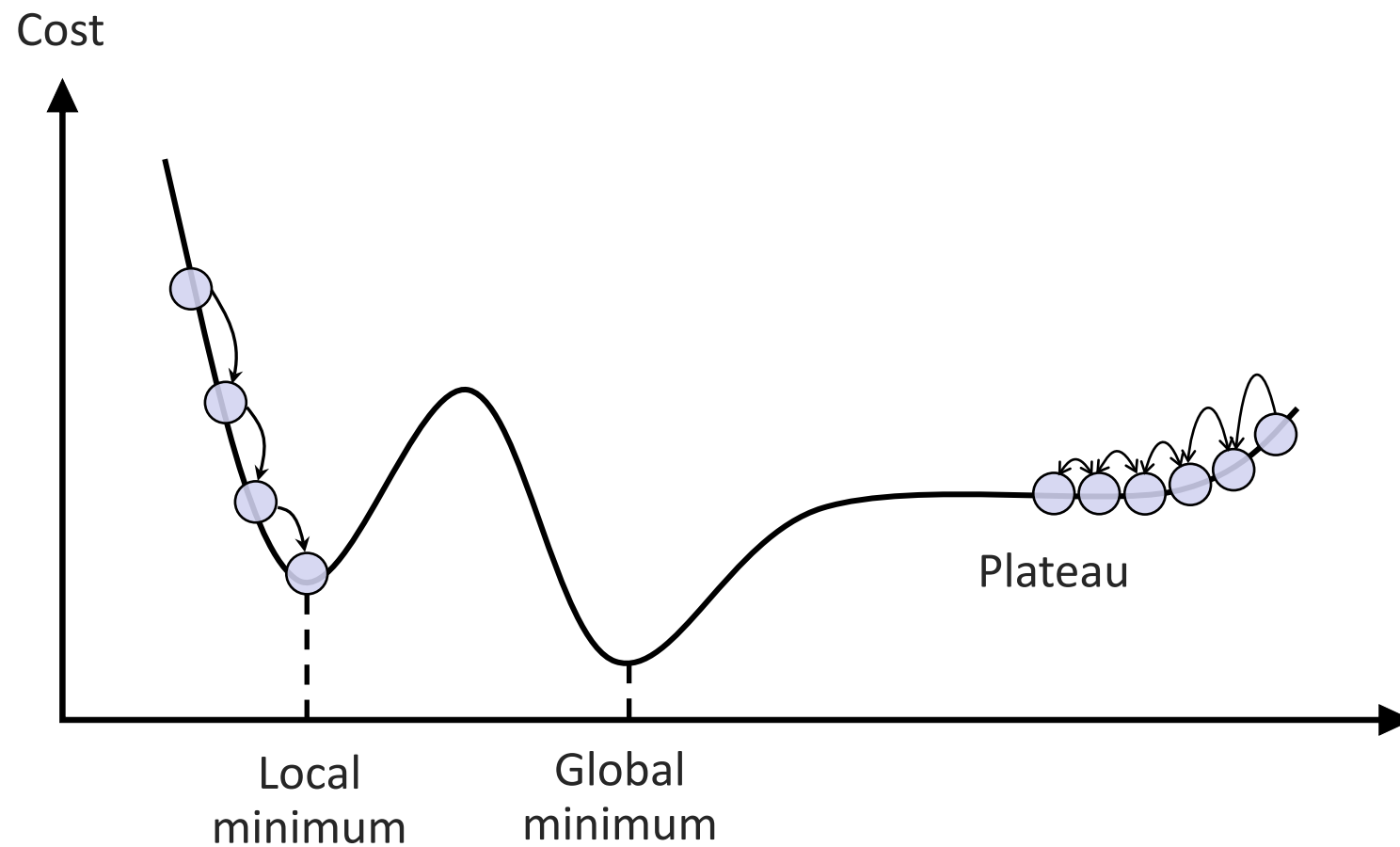


(a) with skip connections



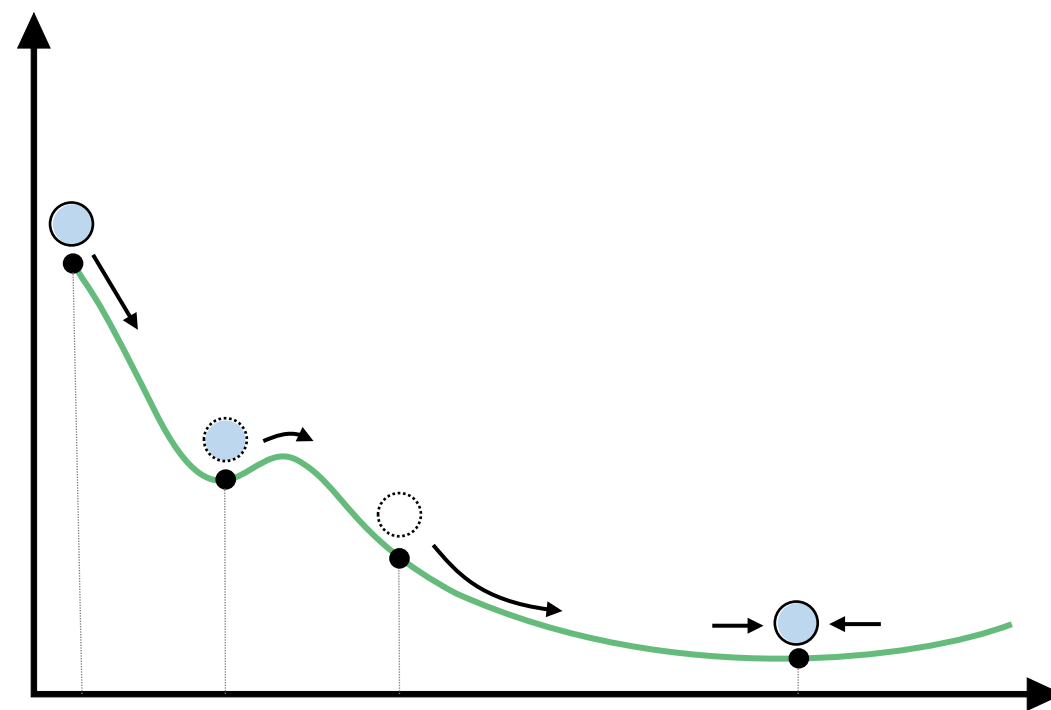
(b) without skip connections





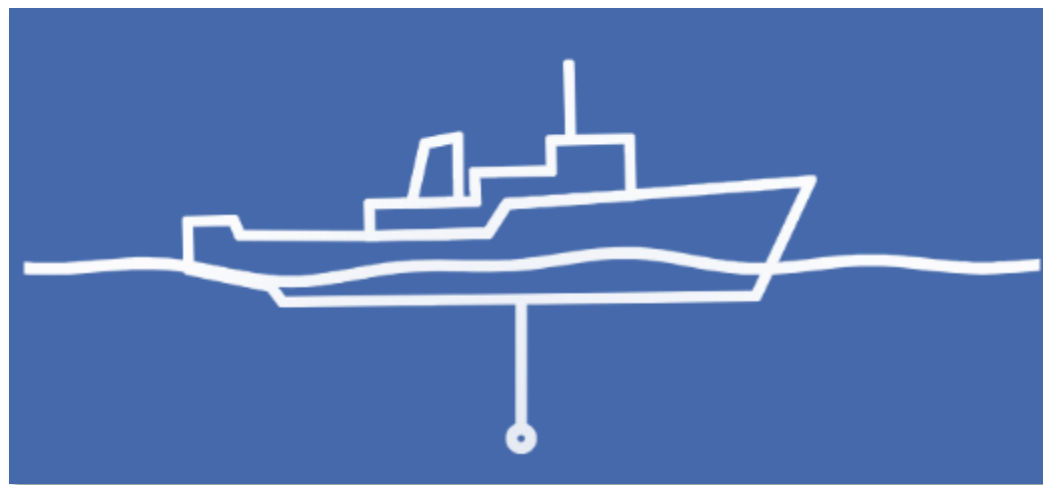
Gradient Descent pitfalls







Gradient Descent Game





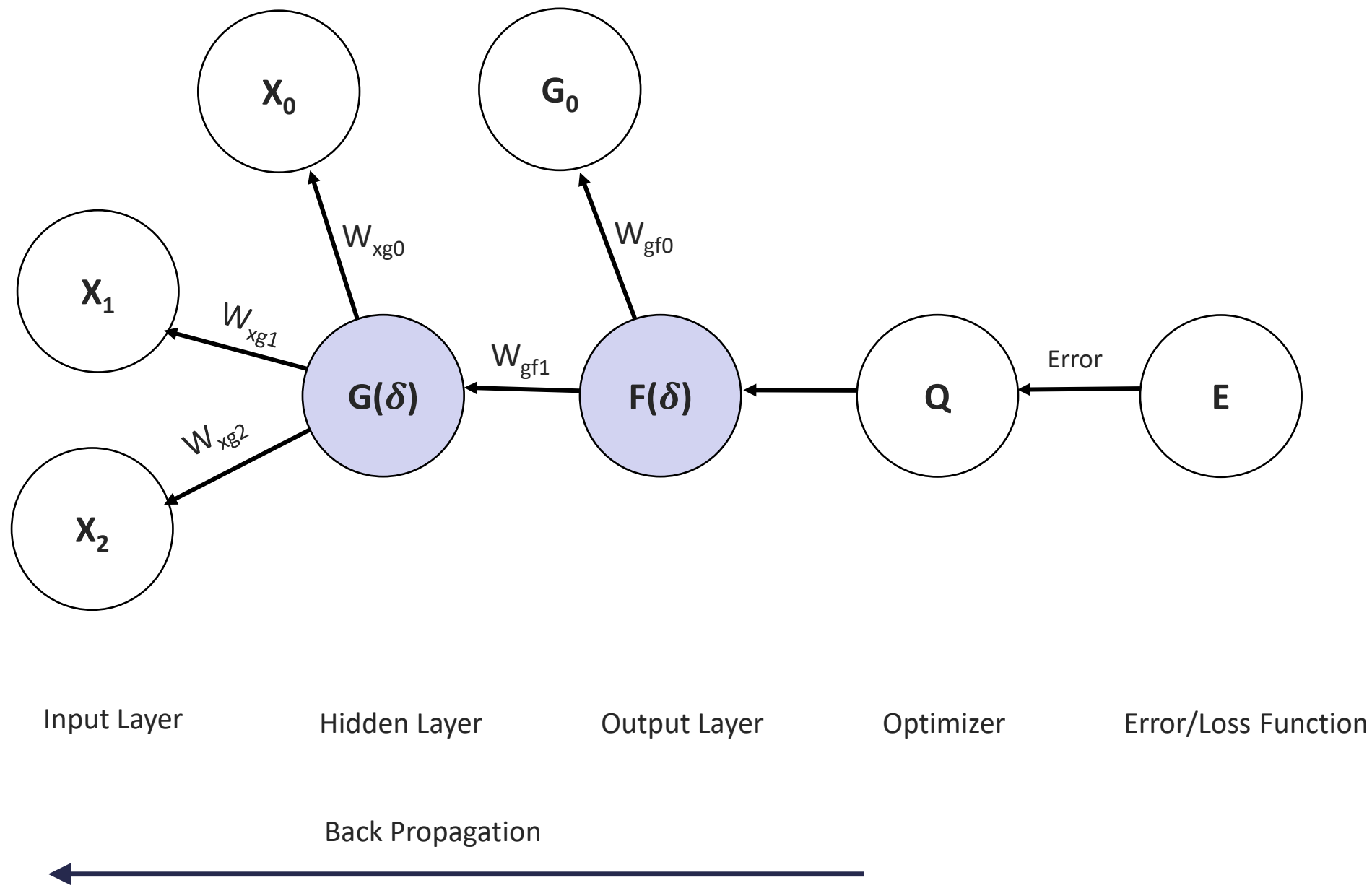
Exercise

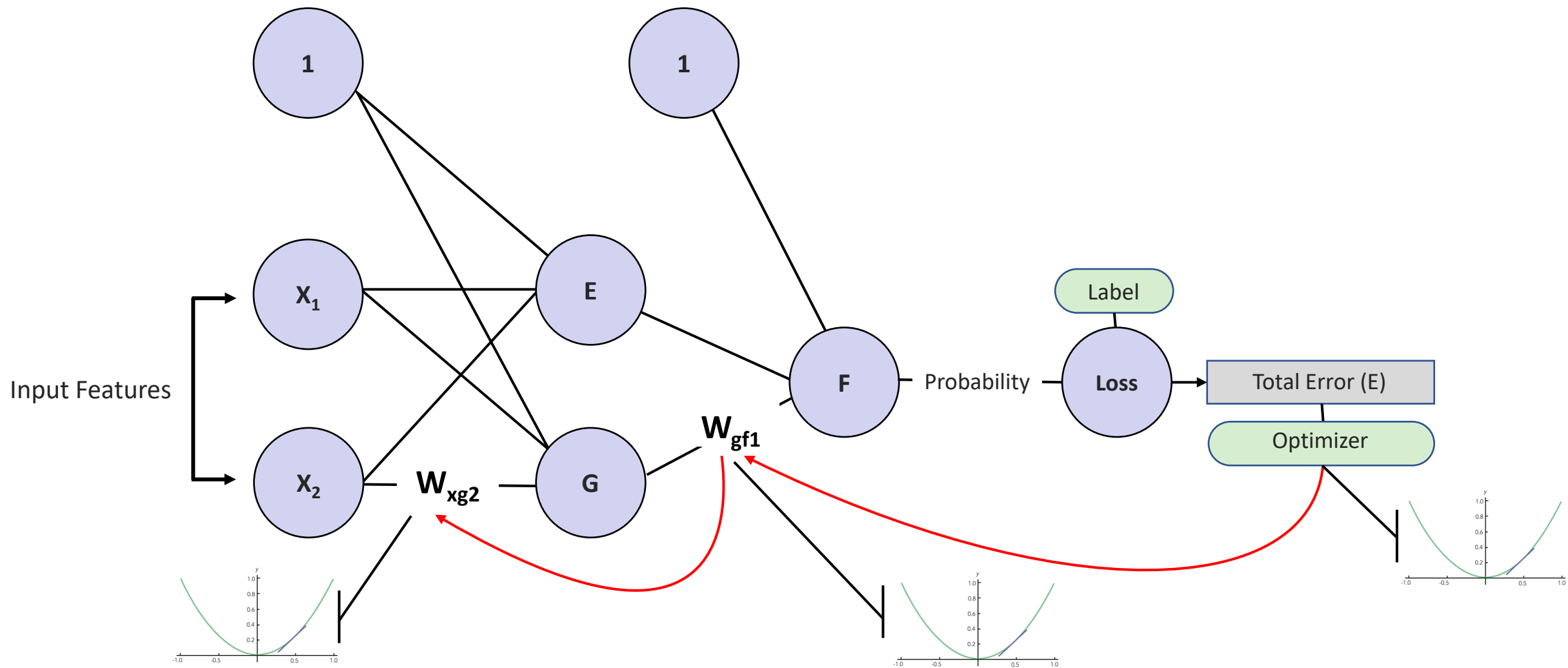
(A Keras Classifier)
03.1_keras_classifier.ipynb

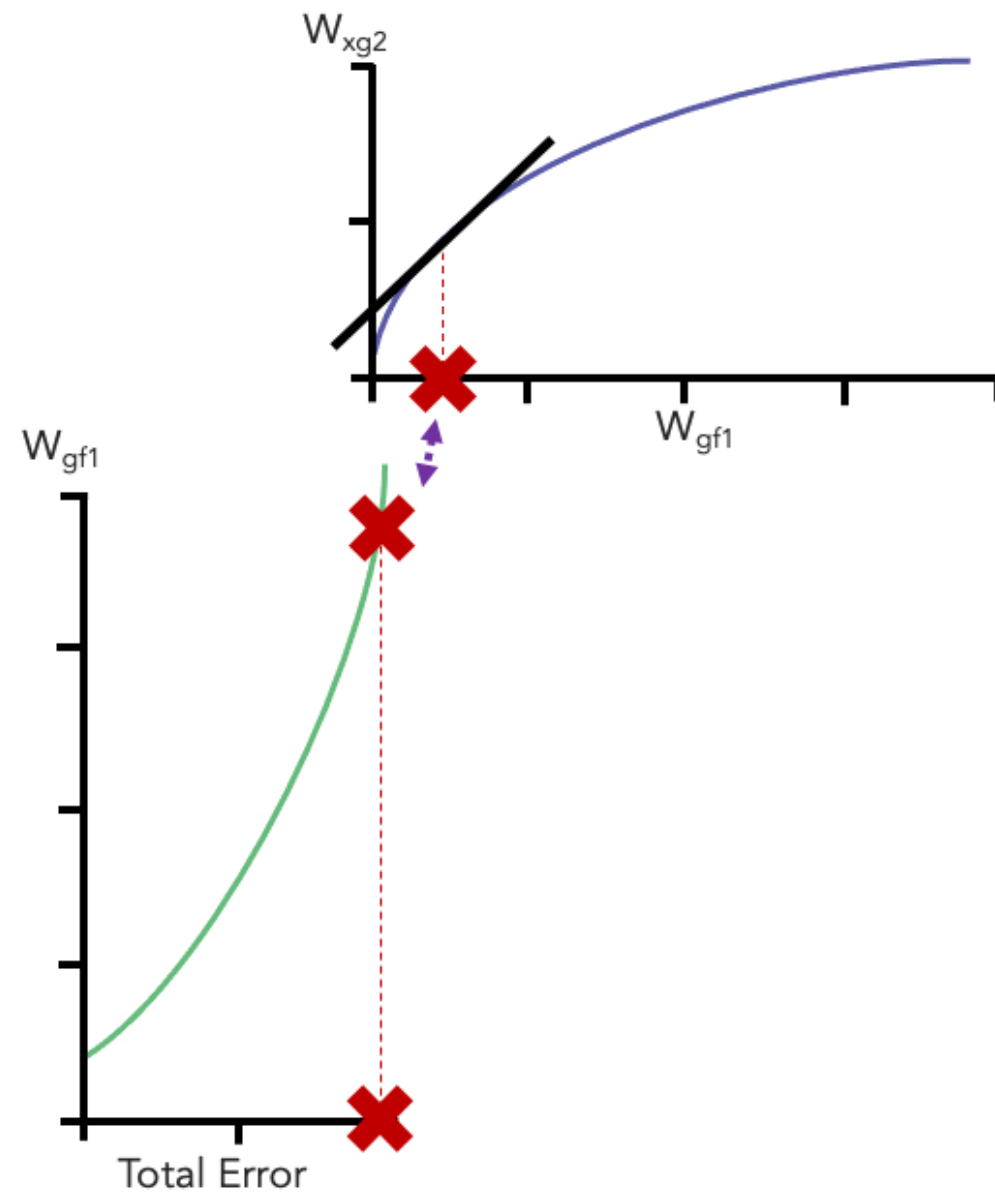


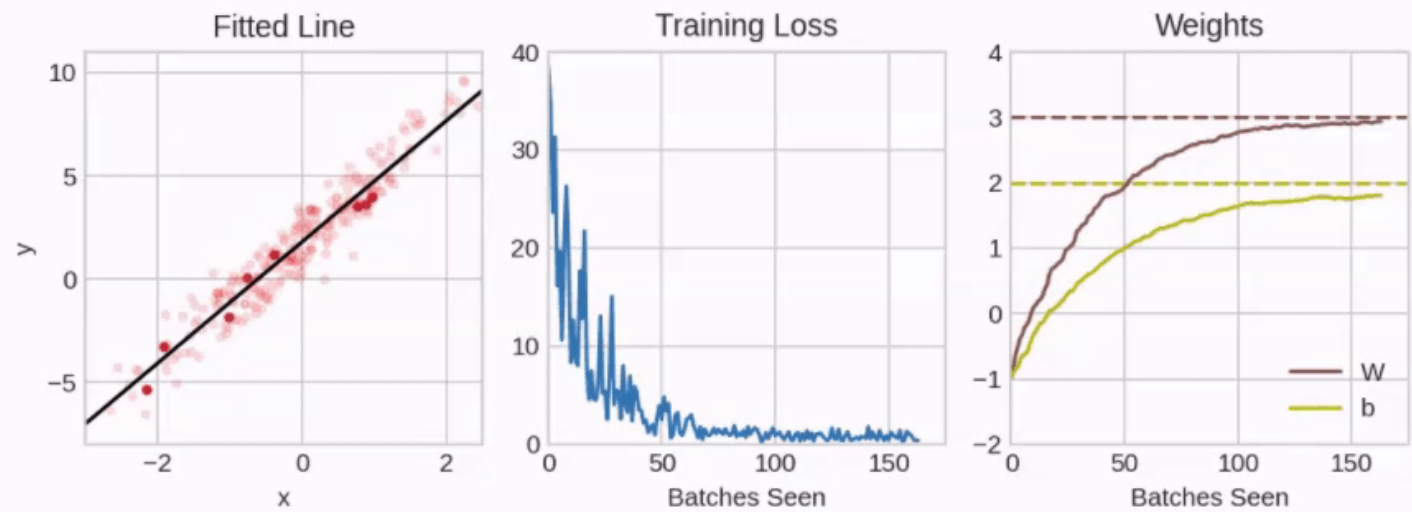


The Chain Rule









Training a neural network with Stochastic Gradient Descent.





Summary

- Error (Loss) Function
- Gradient Descent & Partial Derivatives
- Complex Gradients
- Optimizers
- Chain Rule





Exercise

(MNIST Exercise)

03.2_mnist_classifier.ipynb





