

1. Fill in the blanks below with “an input feature”, “a parameter value”, “layer outputs”, “loss value”:

A **gradient** is the slope of the tangent line on a plot where the x axis is _____ and the y axis is _____. A **weight** in a linear layer is the slope of the tangent line on a plot where the x axis is _____ and the y axis is _____.

Sketch an example of both of these plots.

2. Let's use an MLP with one hidden layer to **classify** handwritten digits as 0, 1, ..., 9. We'll give the model a 28x28-pixel image of a digit, flattened into a 784-dimensional vector. Fill in the blanks below with reasonable values.

```
W1 = random_array(size=(____))
b1 = random_array(size=(____))
W2 = random_array(size=(____))
b2 = random_array(size=(____))

for x_batch, y_batch in training_data:
    # forward pass, starting with linear layer
    linear_out_1 = _____
    activations_1 = _____
    logits = _____
    probs = _____
    loss = cross_entropy_loss(probs, y_batch)

    # backward pass
    loss.backward() # grads now stored in .grad
    for param in [W1, b1, W2, b2]:
        param += _____
```

Fill in the shapes below:

```
x_batch.shape = (N, 784); y_batch.shape = (N,)
linear_out_1.shape = _____
activations_1.shape = _____
logits.shape = _____
probs.shape = _____
loss.shape = _____

W1.grad.shape = _____
b1.grad.shape = _____
W2.grad.shape = _____
b2.grad.shape = _____
```

1. Fill in the blanks below with “an input feature”, “a parameter value”, “layer outputs”, “loss value”:

A **gradient** is the slope of the tangent line on a plot where the x axis is _____ and the y axis is _____. A **weight** in a linear layer is the slope of the tangent line on a plot where the x axis is _____ and the y axis is _____.

Sketch an example of both of these plots.

2. Let's use an MLP with one hidden layer to **classify** handwritten digits as 0, 1, ..., 9. We'll give the model a 28x28-pixel image of a digit, flattened into a 784-dimensional vector. Fill in the blanks below with reasonable values.

```
W1 = random_array(size=(____))
b1 = random_array(size=(____))
W2 = random_array(size=(____))
b2 = random_array(size=(____))

for x_batch, y_batch in training_data:
    # forward pass, starting with linear layer
    linear_out_1 = _____
    activations_1 = _____
    logits = _____
    probs = _____
    loss = cross_entropy_loss(probs, y_batch)

    # backward pass
    loss.backward() # grads now stored in .grad
    for param in [W1, b1, W2, b2]:
        param += _____
```

Fill in the shapes below:

```
x_batch.shape = (N, 784); y_batch.shape = (N,)
linear_out_1.shape = _____
activations_1.shape = _____
logits.shape = _____
probs.shape = _____
loss.shape = _____

W1.grad.shape = _____
b1.grad.shape = _____
W2.grad.shape = _____
b2.grad.shape = _____
```

Before you leave, pick a couple of these questions to react to:

1. What was the most important concept from today for you?
2. What was the muddiest concept today?
3. How does what we did today connect with what you've learned before?
4. What would you like to review or clarify next time we meet?
5. What are you curious, hopeful, or excited about?

Before you leave, pick a couple of these questions to react to:

1. What was the most important concept from today for you?
2. What was the muddiest concept today?
3. How does what we did today connect with what you've learned before?
4. What would you like to review or clarify next time we meet?
5. What are you curious, hopeful, or excited about?