Machine Learning

HW 3 - Gradient-Based Learning
Spring 2025
Ali Ural

1. For the function $J = (x_1 w_1 - 5x_2 w_2 - 2)^2$, where $w = [w_1, w_2]$ are our weights to learn:

   (a) What are the partial gradients, $\frac{\partial J}{\partial w_1}$ and $\frac{\partial J}{\partial w_2}$? Show work to support your answer (6pts).

   $$f(\mathbf{w}) = x_1 w_1 - 5x_2 w_2 - 2, \qquad \text{so} \quad J = f^2.$$

   By the chain rule,

   $$\frac{\partial J}{\partial w_1} = 2f \frac{\partial f}{\partial w_1} = 2f\, x_1,$$

   $$\frac{\partial J}{\partial w_2} = 2f \frac{\partial f}{\partial w_2} = 2f\,(-5x_2) = -10x_2 f.$$

   So

   $$\frac{\partial J}{\partial w_1} = 2x_1\big(x_1 w_1 - 5x_2 w_2 - 2\big), \qquad \frac{\partial J}{\partial w_2} = -10x_2\big(x_1 w_1 - 5x_2 w_2 - 2\big).$$

   (b) What are the values of the partial gradients, given current values of $w = [0, 0]$, $x = [1, 1]$ (4pts)?

   $$f(0,0) = 1 \cdot 0 - 5 \cdot 1 \cdot 0 - 2 = -2.$$

   Substitute into the gradient formulas:

   $$\left.\frac{\partial J}{\partial w_1}\right|_{(0,0)} = 2(1)(-2) = -4,$$

   $$\left.\frac{\partial J}{\partial w_2}\right|_{(0,0)} = -10(1)(-2) = 20.$$

   So the gradient vector at the specified point is

   $$\nabla J|_{(0,0)} = (-4,\ 20).$$