Q2. One SGD Update Step

A sample has input vector x=[1,2], true label y=0, current weights w=[0.1,−0.2], bias b=0.05, and learning rate η=0.1.

Tasks:

- (a) Compute $z = w \cdot x + b$ and $\hat{y} = \sigma(z)$

  z=w·x+b=(0.1)(1)+(−0.2)(2)+0.05=0.1−0.4+0.05=−0.25

  y^ =σ(z)=1/1+e^-z = 1/1+e0.25 ≈ 0.4378

  z=−0.25, y^ =0.4378

- (b) Compute the gradients of the loss w.r.t. weights and bias

  ∂L/∂w = (ŷ − y)·x
  ∂L/∂b = (ŷ − y)

  ŷ − y = 0.4378 – 0 = 0.4378

  ∂L/∂w = [0.4378×1, 0.4378×2] = [0.4378, 0.8756]
  ∂L/∂b = 0.4378

  ∂L/∂w = [0.4378, 0.8756]
  ∂L/∂b = 0.4378

- (c) Update weights and bias using gradient descent

  w_new = w − η·∂L/∂w
  b_new = b − η·∂L/∂b

  w_new = [0.1, –0.2] – 0.1×[0.4378, 0.8756]
  = [0.1 – 0.04378, –0.2 – 0.08756]
  = [0.05622, –0.28756]

  b_new = 0.05 – 0.1×0.4378 = 0.00622

  Updated parameters:
  w_new = [0.0562, –0.2876]
  b_new = 0.0062