How Neural Networks Learn?

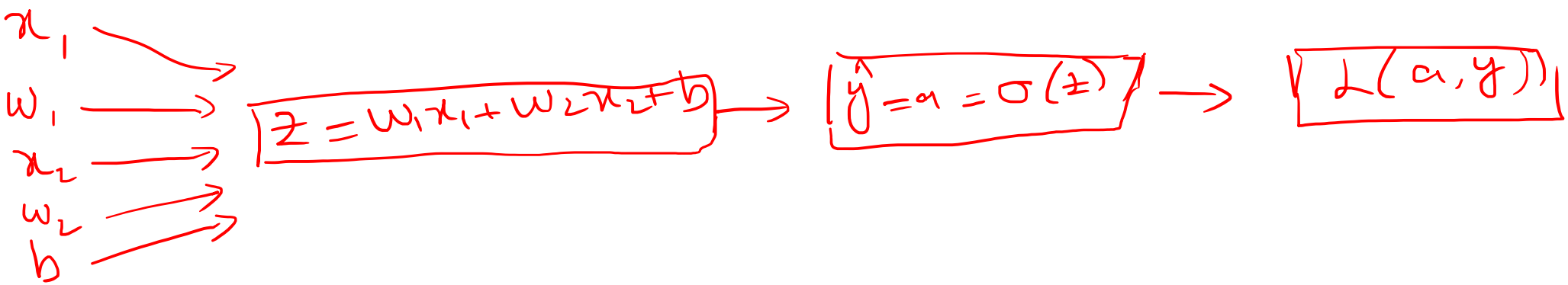
Dr. Muhammad Wasim

Logistic Regression as a Computation Graph

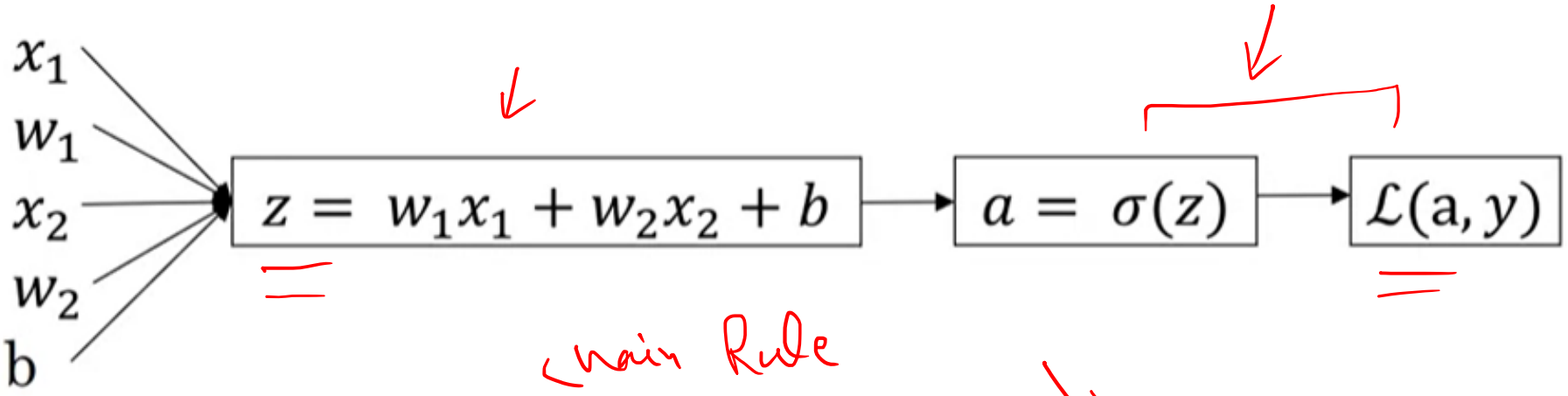
$$z = w^T\theta + b$$

$$\hat{y} = a = \sigma(z)$$

$$Loss(a, y) = -(ylog(a) + (1 - y)\log(1 - a))$$



Logistic Regression as a Computation Graph



Chain Rule

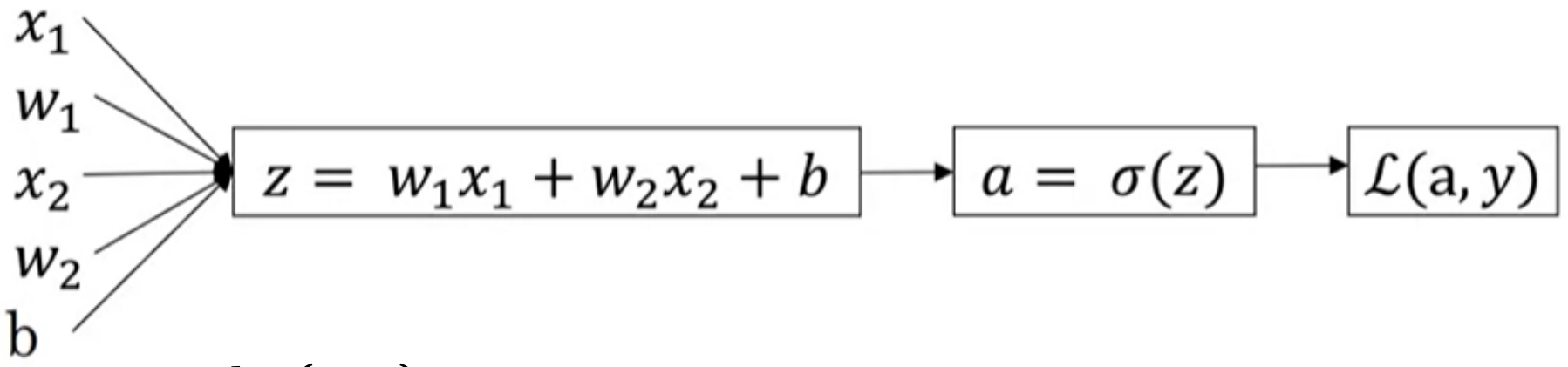
$$dz = \frac{d\,\mathcal{L}(a,y)}{dz} = \frac{d\,\mathcal{L}(a,y)}{da} \times \frac{da}{dz}$$

$$= a - y$$

$$da = \frac{d\,\mathcal{L}(a,y)}{da} = -\frac{y}{a} + \frac{1-y}{1-a}$$

$$\frac{da}{dz} = a\,(1-a)$$

Logistic Regression Derivatives



$$\frac{d\,\mathcal{L}(a,y)}{dw_1} = "dw_1" = x_1.dz \qquad dw_2 = x_2.dz \qquad db = dz$$

$$w_1 = w_1 - \alpha\, dw_1$$

$$w_2 = w_2 - \alpha\, dw_2$$

$$b = b - \alpha\, db$$

The Learning Process of a NN