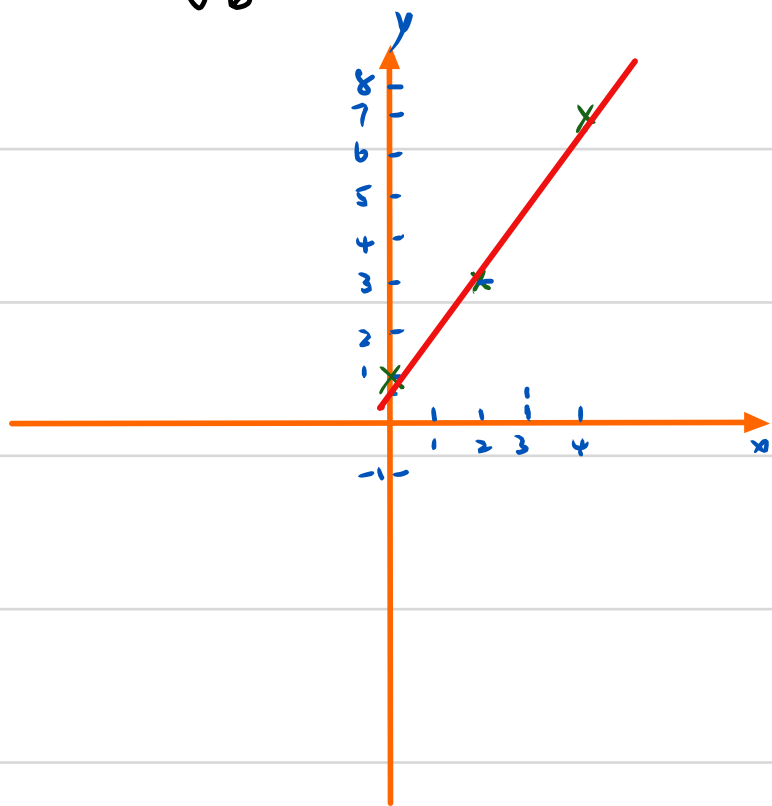


$$L = \sum_i (wx_i + b - y_i)^2 = (b-1)^2 + (2w+b-3)^2 + (4w+b-7)^2$$

$$\frac{\partial L}{\partial w} = 2(2w+b-3) \cdot 2 + 2(4w+b-7) \cdot 4 = 40w + 12b - 68$$

$$\frac{\partial L}{\partial b} = 2(b-1) + 2(2w+b-3) + 2(4w+b-7) = 6b + 12w - 22$$

$$\begin{cases} \frac{\partial L}{\partial w} = 0 \\ \frac{\partial L}{\partial b} = 0 \end{cases} \Rightarrow \begin{cases} w = \frac{3}{2} \\ b = \frac{2}{3} \end{cases}$$



$$\begin{aligned} 1) & \quad y_i = 1: p(y=1 | x=x_i) = \sigma(w^T x_i + b) \\ & \quad y_i = 0: p(y=0 | x=x_i) = 1 - \sigma(w^T x_i + b) = \sigma(-w^T x_i - b) \end{aligned}$$

$$2) \quad \max_{w, b} \sum_{i \in [n]} \log[p(y=y_i | x=x_i)]$$

$$= \max_{w, b} \sum_{i \in [n]} \log[p(y=1 | x=x_i)^{y_i} p(y=0 | x=x_i)^{1-y_i}]$$

$$= \max_{w, b} \sum_{i \in [n]} y_i \log[\sigma(w^T x_i + b)] + (1-y_i) \log[\sigma(-w^T x_i - b)]$$

$$= \max_{w, b} \left\{ - \sum_{i \in [n]} y_i \log[1 + e^{-w^T x_i - b}] - \sum_{i \in [n]} (1-y_i) \log[1 + e^{w^T x_i + b}] \right\}$$

三. D 为样本

1) $H(D) = - \sum_{k \in [K]} \frac{|D_k|}{|D|} \log_2 \frac{|D_k|}{|D|} = - \left(\frac{4}{9} \log_2 \frac{4}{9} + \frac{5}{9} \log_2 \frac{5}{9} \right) \doteq 0.991076$

$$H(D^{A=0}) = - \left(\frac{1}{4} \log_2 \frac{1}{4} + \frac{3}{4} \log_2 \frac{3}{4} \right) \doteq 0.811278$$

$$H(D^{A=1}) = - \left(\frac{3}{5} \log_2 \frac{3}{5} + \frac{2}{5} \log_2 \frac{2}{5} \right) \doteq 0.970951$$

$$H(A) = - \left(\frac{4}{9} \log_2 \frac{4}{9} + \frac{5}{9} \log_2 \frac{5}{9} \right) \doteq 0.991076$$

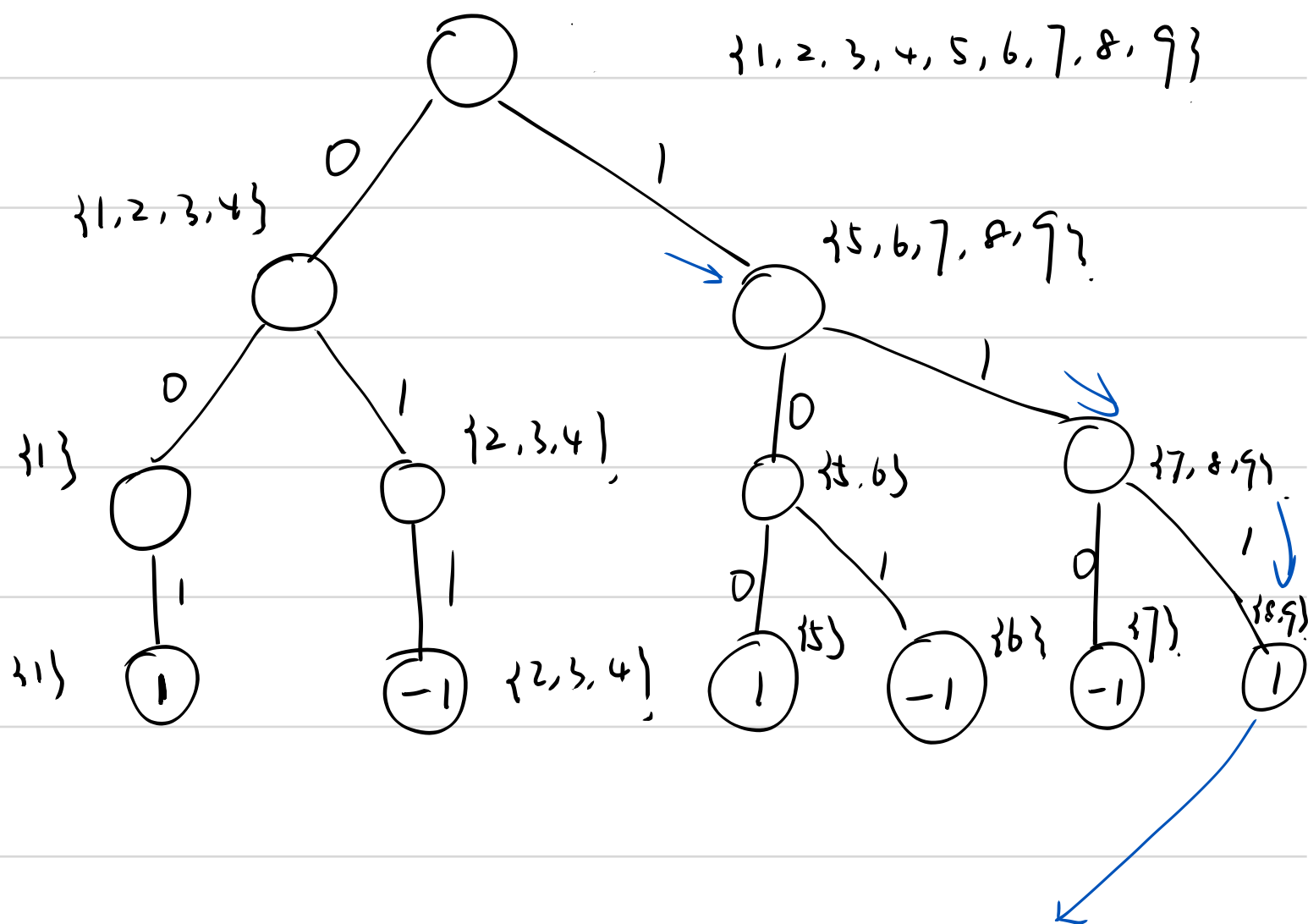
$$g_R(D, A) = \frac{g(D, A)}{H(A)} = \frac{H(D) - \frac{4}{9} H(D^{A=0}) - \frac{5}{9} H(D^{A=1})}{H(A)}$$

$$\doteq 0.091910 = 9.2\%$$

2) A

C

B



$\chi_* = [1, 1, 1]$ 标签为 1.

$$\text{IV. 1) } \frac{\partial a_j}{\partial z_i} = \frac{-e^{z_j} \cdot e^{z_i}}{(\sum_k e^{z_k})^2} = -a_j \cdot a_i \quad (j \neq i).$$

$$\frac{\partial a_i}{\partial z_i} = \frac{e^{z_i} (\sum_k e^{z_k}) - e^{z_i} e^{z_i}}{(\sum_k e^{z_k})^2} = \frac{e^{z_i}}{\sum_k e^{z_k}} \left(1 - \frac{e^{z_i}}{\sum_k e^{z_k}} \right) = a_i (1 - a_i)$$

$$\begin{aligned} \frac{\partial L}{\partial z_i} &= \sum_j \frac{\partial L}{\partial a_j} \frac{\partial a_j}{\partial z_i} = \left(- \sum_{j \neq i} \frac{\partial L}{\partial a_j} a_j \cdot a_i \right) + \frac{\partial L}{\partial a_i} a_i (1 - a_i) \\ &= (-a_i \sum_j \frac{\partial L}{\partial a_j} a_j) + a_i \frac{\partial L}{\partial a_i} \end{aligned}$$

$$2) \frac{\partial a_j}{\partial z_i} = \frac{(\sum_k e^{z_k})}{e^{z_j}} \cdot \frac{-e^{z_j} \cdot e^{z_i}}{(\sum_k e^{z_k})^2} = \frac{-e^{z_i}}{\sum_k e^{z_k}} = -a_i \quad (i \neq j)$$

$$\frac{\partial a_i}{\partial z_i} = \frac{(\sum_k e^{z_k})}{e^{z_i}} \cdot \frac{e^{z_i}}{\sum_k e^{z_k}} \left(1 - \frac{e^{z_i}}{\sum_k e^{z_k}} \right) = 1 - a_i$$

$$\frac{\partial L}{\partial z_i} = \sum_j \frac{\partial L}{\partial a_j} \frac{\partial a_j}{\partial z_i} = \left(-a_i \sum_{j \neq i} \frac{\partial L}{\partial a_j} \right) + (1 - a_i) \frac{\partial L}{\partial a_i}$$

$$= \left(-a_i \sum_j \frac{\partial L}{\partial a_j} \right) + \frac{\partial L}{\partial a_i}$$