Fundamentos matemáticos del aprendizaje profundo

1° cuatrimestre 2025

Práctica 6: Redes neuronales

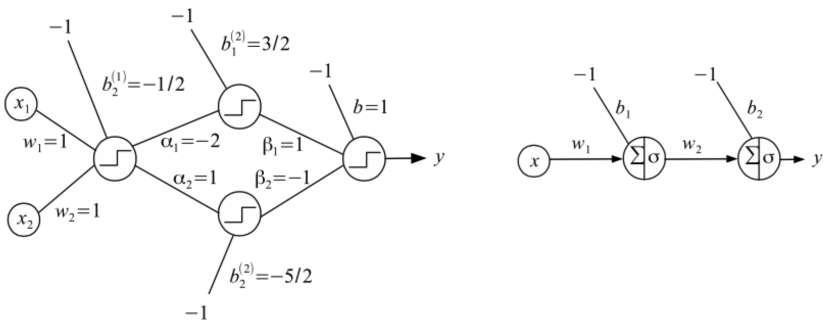
Ejercicio 1. Dibujar el esquema de la red dada por la Observación 6.1.2 de las notas de clase.

Ejercicio 2. (a) ¿Puede un único perceptrón aprender el siguiente mapa:

$$(0,0,0) \mapsto 1, \quad (1,0,0) \mapsto 0, \quad (0,1,0) \mapsto 0, \quad (0,0,1) \mapsto 0,$$
$$(0,1,1) \mapsto 1, \quad (1,1,0) \mapsto 0, \quad (1,0,1) \mapsto 0, \quad (1,1,1) \mapsto 1?$$

(b) Dibujar el esquema de una red neuronal con perceptrones que aprenda el mapa del ítem anterior. Dar los pesos y sesgos de todos los perceptrones de la red.

Ejercicio 3. Escribir una fórmula explícita para la salida de la red neuronal multi-perceptrón dada por la figura parte **a**.



a **b**

a. La red neuronal usada en el Ejercicio 3. **b.** La red neuronal usada en el Ejercicio 5.

Ejercicio 4. Considerar una neurona sigmoide con una entrada unidimensional x, peso w, sesgo b y salida $y = \sigma(wx+b)$. El objetivo es una variable unidimensional z. Consideremos la función de costo $C(w,b) = \frac{1}{2}(y-z)^2$.

(a) Hallar $\nabla C(w,b)$ y mostrar que $\|\nabla C(w,b)\| < \frac{1}{4}\sqrt{1+x^2}(1+|z|)$.

(b) Escribir la iteración de descenso de gradiente para la sucesión de pesos (w_n, b_n).

Ejercicio 5. Consideremos la red neuronal dada por la figura parte **b**, con la función de costo dada por $C(w,b) = \frac{1}{2}(y-z)^2$.

(a) Calcular las deltas $\delta^{(\ell)} = \frac{\partial C}{\partial s^{(\ell)}}$, $\ell = 1,2$.

(b) Usar el algoritmo de retropropagación para hallar el gradiente $\nabla C(w,b)$.

Ejercicio 6. Asumamos que la función de activación en la capa ℓ es lineal y que los pesos y entradas son variables aleatorias.

(a) Mostrar que $\delta_j^{(\ell)}$ y $W_{ij}^{(\ell)}$ son variables aleatorias independientes.

(b) Mostrar que $\delta_j^{(\ell)}$ y $X_i^{(\ell-1)}$ son variables aleatorias independientes.

Ejercicio 7. Considerar una neurona sigmoide con entrada aleatoria normalmente distribuida, $X \sim N(0,1)$ y con salida $Y = \sigma(wX+b)$. Mostrar que $\mathbb{V}[Y] \simeq w^2\sigma'(b)^2$. Notar que la varianza de la salida decrece para valores pequeños del peso w o grandes valores del sesgo b.

Ejercicio 8. Considerar una red neuronal con una capa escondida que posee neuronas sigmoides en esa capa. Suponer que las entradas están normalmente distribuidas, $X \sim N(0,1)$ y que la salida es $Y = \sum_{i=1}^N \alpha_i \sigma(w_i X + b_i)$. Mostrar que $\mathbb{V}[Y] \simeq \sum_j \sigma'(b_i)^2 \alpha_i^2 w_i^2$.