

Exercícios Deep Learning - Lista Teórica 02

Elaboração de Carolina Coimbra, Guilherme Borges e Edemir Andrade Jr.

August 23, 2019

1 Redes Neurais

1- Considere a rede com uma camada escondida mostrada abaixo:



Assumindo que cada neurônio de uma camada $[l]$ aplica uma função $f(Z^{[l]}) = A^{[l]}$ onde Z é calculado como $W^{[l]}A^{[l-1]} + b^{[l]}$ e que a entrada pode ser escrita como a camada zero: $A[0] = \mathbf{x}$.

a) Quais são as dimensões de $W^{[1]}$, $b^{[1]}$, $W^{[2]}$ e $b^{[2]}$?

b) Quais são as dimensões de $Z^{[1]}$ e $A^{[1]}$?

2- Além da sigmóide, a tangente hiperbólica e ReLU são funções de ativação comumente utilizadas em redes neurais.

a) Encontre a derivada da tangente hiperbólica $\tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ com respeito a x . Em qual valor de x a derivada atinge seu valor máximo?

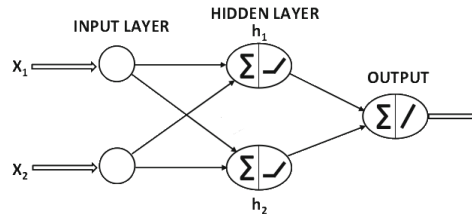
b) Encontre a derivada da função $ReLU = \max(0, x)$ para $x < 0$ e $x > 0$.

c) Sejam a, b constantes positivas. Expresse a derivada da função $f(x) = a * \tanh(bx)$ (sigmoide anti-simétrica) em termos da própria função.

3- Considere uma rede neural com dois inputs X_1 e X_2 e com duas camadas escondidas, cada uma com dois nós. Assuma que os pesos estão distribuídos de forma que os nós em cima na camada aplicam a sigmoide na soma dos seus inputs e que os nós em baixo aplicam a função tanh em seus inputs. O nó de saída aplica a ReLU na soma dos dois inputs. Desenhe esta rede. Escreva a saída dessa rede neural como uma função de x_1 e x_2 em forma fechada.

4- Suponha que você tenha uma sequência de dados com duas features binárias, isto é, $X = \{(x_1^{(0)}, x_2^{(0)}), \dots, (x_1^{(n)}, x_2^{(n)})\}$, onde $x_1^{(i)}, x_2^{(i)} \in \{0, 1\}$. Imagine uma rede neural com 3 nós, como na figura abaixo, todos utilizam a função de ativação $ReLU(x)$, o último nó (output) deve retornar 0 se o resultado é Falso e um valor maior que zero se o resultado é Verdadeiro. Mostre como você pode usar esta rede para calcular $XOR(x_1, x_2)$. Ou seja, mostre os pesos dos parâmetros dos nós h_1, h_2 e output para calcular o XOR.

Lembre-se que, o $XOR(x_1, x_2)$, também chamado de *ou exclusivo*, é Verdadeiro quando $x_1 = 1$ OU $x_2 = 1$, porém é Falso quando $x_1 = 1$ E $x_2 = 1$ e quando $x_1 = 0$ E $x_2 = 0$.

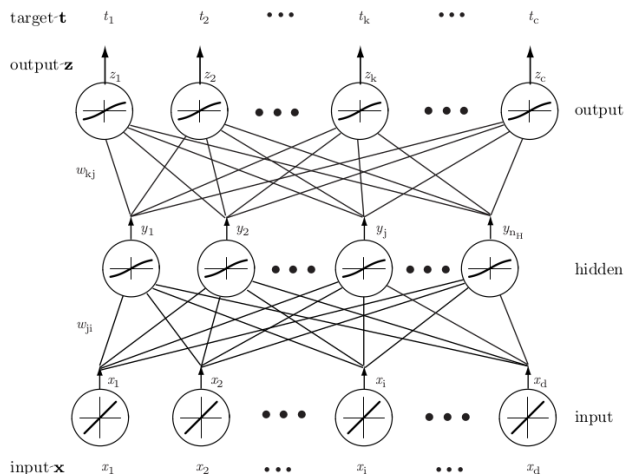


5- Suponha que todos os nós de uma rede neural utilizam uma função de ativação linear, isto é, $g(x) = x$. Mostre que, quando utilizada essa função, qualquer rede neural com 2 camadas é equivalente a uma rede de 1 camada. Verifique intuitivamente que esse resultado é válido independentemente do número de camadas (não precisa fazer contas).

Esse tipo de rede consegue classificar corretamente inputs de acordo com o XOR (problema 4)?

6- Considere uma rede padrão de 3 camadas cuja entrada \mathbf{x} possui dimensão $d \times 1$, a primeira camada da rede possui d unidades de entrada e possui somente uma ativação linear do tipo $f(\mathbf{x}) = \mathbf{x}$, a camada escondida possui n_H unidades

escondidas e a camada final possui c unidades de saída e o bias.

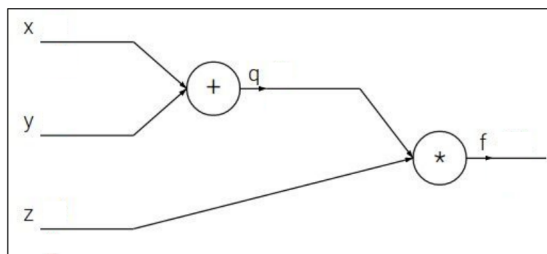


a) Qual o número total de pesos que existem na rede?

b) Mostre que se o sinal de cada peso da rede for trocado, a função da rede permanece inalterada, caso a função de ativação usada for uma função par. Lembre-se que, uma função é par se e somente se $f(-x) = f(x)$.

2 Backpropagation

7- Seja a função $f(x, y, z) = (x + y)z$. dada pelo diagrama da figura abaixo.



a) Calcule as seguintes derivadas parciais:

$$\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}, \frac{\partial f}{\partial q}, \frac{\partial f}{\partial z}, \frac{\partial q}{\partial x}, \frac{\partial q}{\partial y},$$

b) Mostre passo a passo o calculo do backpropagation, considerando as entradas $x = -2, y = 5, z = -4$

2- Seja a função $f(w, x) = \frac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$, onde $w_0 = 2.0, x_0 = -1.0, w_1 = -3.0, x_1 = -2.0, w_2 = -3$. Construa o diagrama assim como na questão (1) e refaça os passos das letras (a) e (b).