Pontifícia Universidade Católica de Minas Gerais
ICEI - Curso de Ciência da Computação
Disciplina: Tópicos em Computação: Deep Learning em Visão Computacional
Prof. Alexei Machado

Lista No 1 - Respostas

A- Fundamentos Matemáticos

1. V; F; V; V se a=b; F se a!=1; F se a!=0 ou b!=0; F se x!=0; F se x!=0 ou y!=0; V.
2.
   b) 3/(3x+1) e 3exp(3x)
   c) V; F; F; V; V; V; F.
3.
   b) z=0; 0; 0.
   c) terceiro
4. Derivar a função L, considerando y constante. Para a 2ª derivada parcial, aplicar a regra da cadeia: dL/dz= dL/da x da/dz, onde da/dz foi calculado no exercício anterior.

B- MLP

1.
   a) (4,1), (4,2), (1,4) e (1,1)
   b) ambos são (4,m)

2. 23
3. max(0,s(s(x1+x2)+tanh(x1+x2))+tanh(s(x1+x2)_tanh(x1+x2)))

4- Seja $a_{h1} = max(0, b^{h1} + W_1^{h1}x_1 + W_2^{h1}x_2)$ a saída do nó h_1, então os pesos devem ser colocados de forma que somente a saída para entrada $x_1 = 1, x_2 = 0$ seja positiva, por exemplo:

$$b^{h1} = 0, W_1^{h1} = 1, W_2^{h1} = -1$$

Seja $a_{h2} = max(0, b^{h2} + W_1^{h2}x_1 + W_2^{h2}x_2)$ a saída do nó h_2, então os pesos devem ser colocados de forma que somente a saída para entrada $x_1 = 0, x_2 = 1$ seja positiva, por exemplo:

$$b^{h2} = 0, W_1^{h2} = -1, W_2^{h2} = 1$$

(Ou vice-versa)

Seja $a_{out} = max(0, b^{out} + W_1^{out}a_{h1} + W_2^{out}a_{h2})$ a saída da rede (do nó output). Os pesos devem ser distribuídos de forma que a saída seja positiva somente quando a_{h1} ou a_{h2} seja positivo, ou seja, quando $XOR(x_1, x_2)$ é verdadeiro. Por exemplo:

$$b^{h2} = 0, W_1^{h2} = 1, W_2^{h2} = 1$$

5- O nó de output na rede de duas camadas receberá a seguinte função:

$$b^{[2]} + W^{[2]}(b^{[1]} + W^{[1]}X)$$

$$= (b^{[2]} + W^{[2]}b^{[1]}) + (W^{[2]}W^{[1]})X$$

Seja $b^{[1]*} = (b^{[2]} + W^{[2]}b^{[1]})$ e $W^{[1]*} = (W^{[2]}W^{[1]})$, então se uma rede de uma camada estimar $b^{[1]*}$ e $W^{[1]*}$ como seus parâmetros, então ela gerará o mesmo resultado da rede de duas camadas acima, logo elas são equivalentes.

Não pois o XOR não é linearmente separável e a saída da rede será uma reta.

6. nH(1+d+c)+c

C- Grafos Computacionais

1-
a)

$$q = x + y \qquad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$
$$f = qz \qquad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$

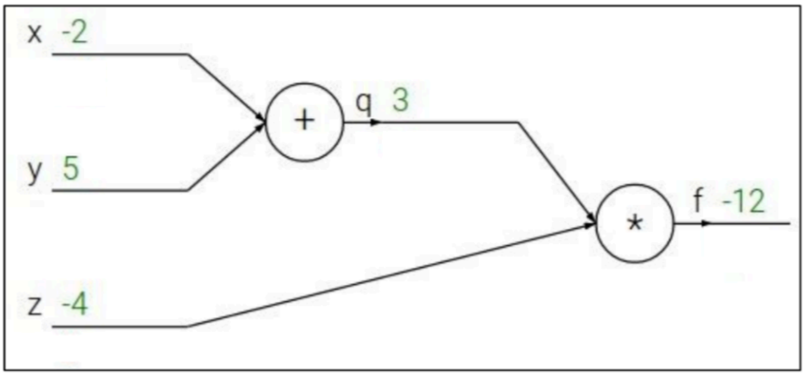
Para o backpropagation, precisamos calcular as derivadas parciais $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$

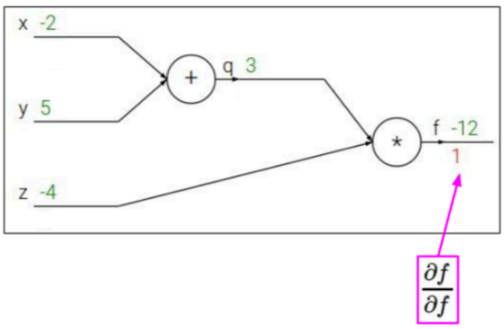
$$\frac{\partial f}{\partial f} = 1$$

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial q}\frac{\partial q}{\partial x} = z$$

$$\frac{\partial f}{\partial y} = \frac{\partial f}{\partial q}\frac{\partial q}{\partial y} = z$$

b)
De forma gráfica $f(x, y, z) = (x + y)z$, onde $x = -2, y = 5, z = -4$:





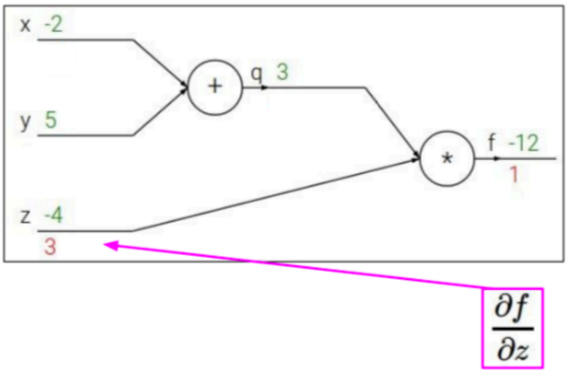


Figure 2: $\dfrac{\partial f}{\partial z} = q = 3$

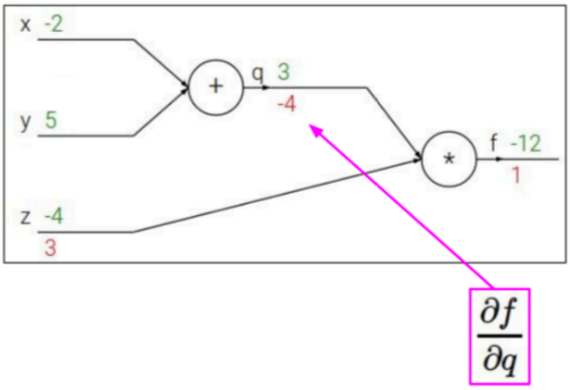
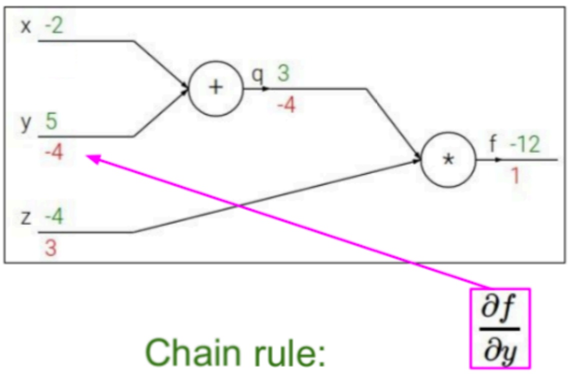


Figure 3: $\dfrac{\partial f}{\partial q} = z = -4$



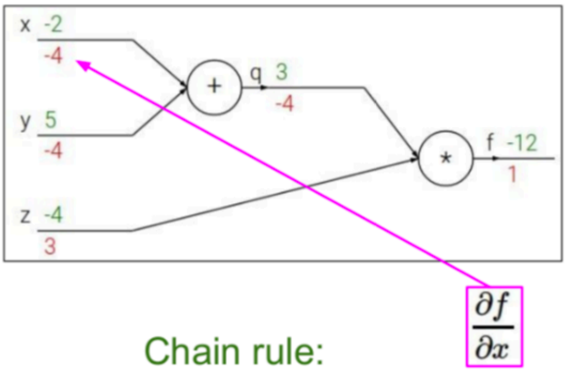


Figure 5: $\frac{\partial f}{\partial x} = \frac{\partial f}{\partial q}\frac{\partial q}{\partial x} = z * 1 = z = -4$

2-

De forma gráfica $f(w,x) = \frac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$, onde $w_0 = 2.0, x_0 = -1.0, w_1 = -3.0, x_1 = -2.0, w_2 = -3$:

$$
\begin{aligned}
f(x) &= e^x && \rightarrow && \frac{\partial f}{\partial x} = e^x \\
f_a(x) &= ax && \rightarrow && \frac{\partial f}{\partial x} = a \\
f(x) &= \tfrac{1}{x} && \rightarrow && \frac{\partial f}{\partial x} = -1/x^2 \\
f_c(x) &= c + x && \rightarrow && \frac{\partial f}{\partial x} = 1
\end{aligned}
$$

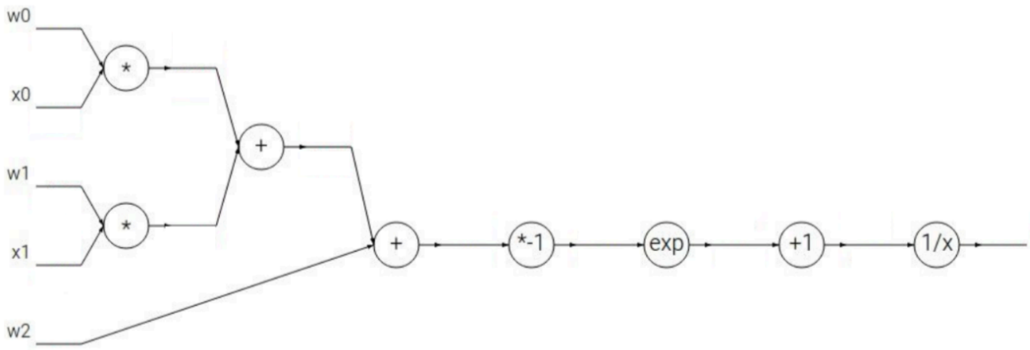
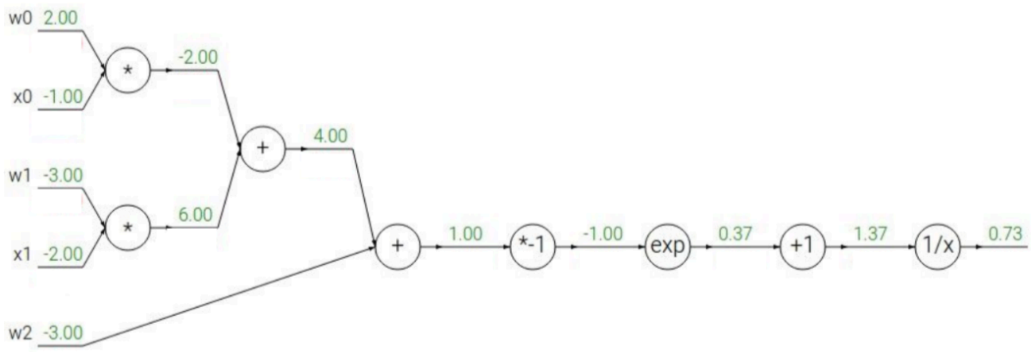


Figure 6: Diagrama da função f



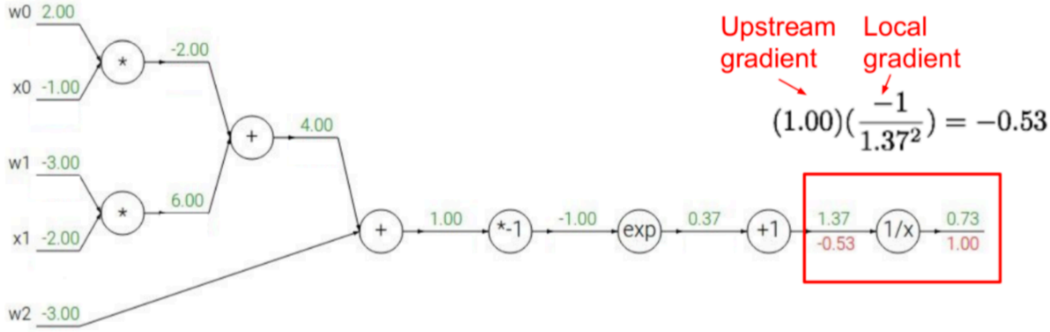


Figure 8: $f(x) = \frac{1}{x} \rightarrow \quad \frac{\partial f}{\partial x} = -1/x^2$

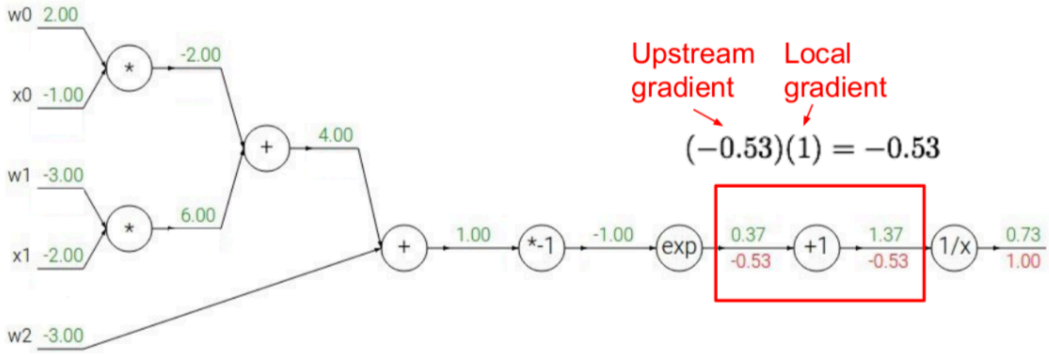
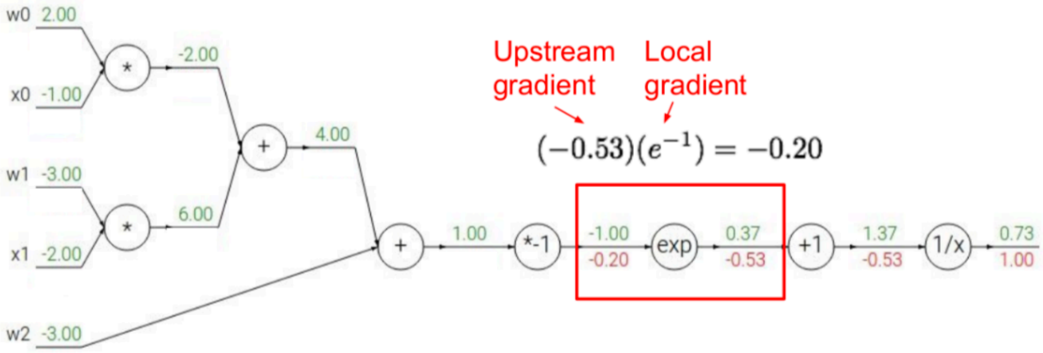


Figure 9: $f_c(x) = c + x \rightarrow \quad \frac{\partial f}{\partial x} = 1$



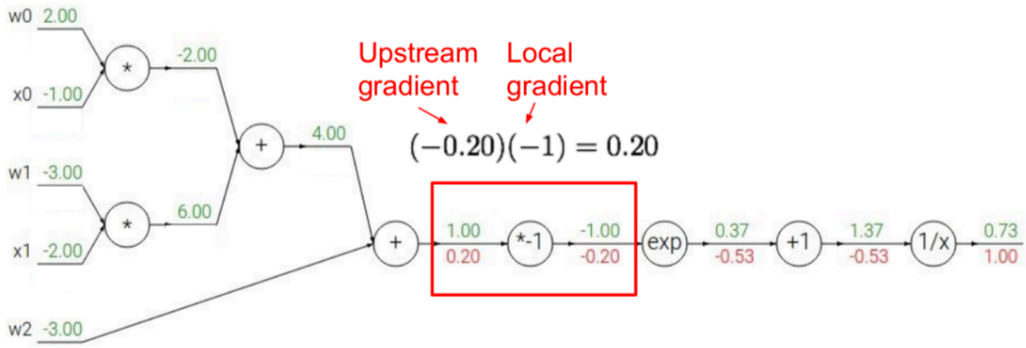


Figure 11: $f_a(x) = ax \rightarrow \quad \frac{\partial f}{\partial x} = a$

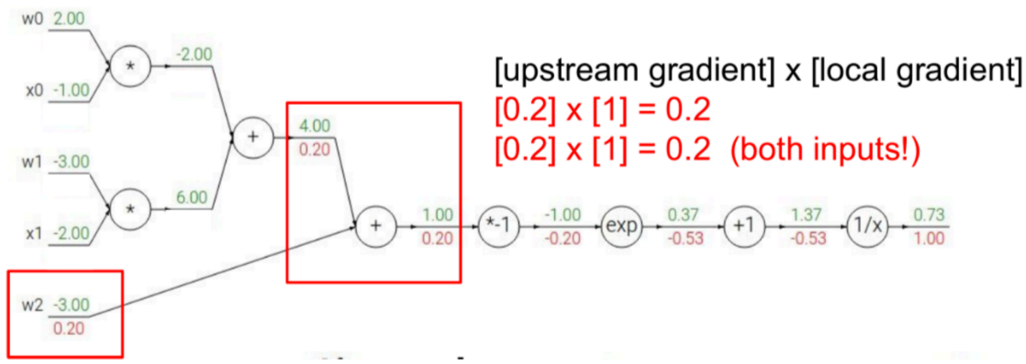


Figure 12: $\frac{\partial f}{\partial x} = 1 \frac{\partial f}{\partial w_2} = 1$