# BP Network

Saturday, 31 March 2018 22:05

Back-propagation:

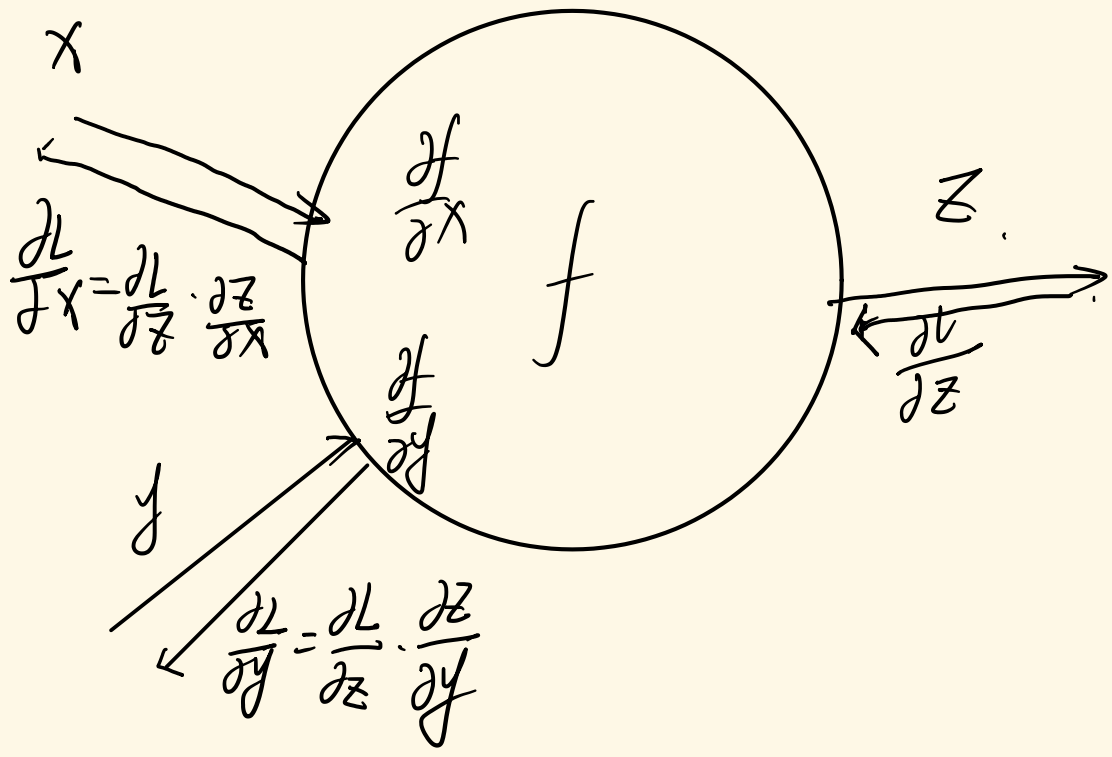
$f(x,y,z) = (x+y)z$.

$q = x+y \quad \frac{\partial q}{\partial x}=1, \frac{\partial q}{\partial y}=1$

$f = qz \quad \frac{\partial f}{\partial q}=z, \frac{\partial f}{\partial z}=q$

Want: $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$:

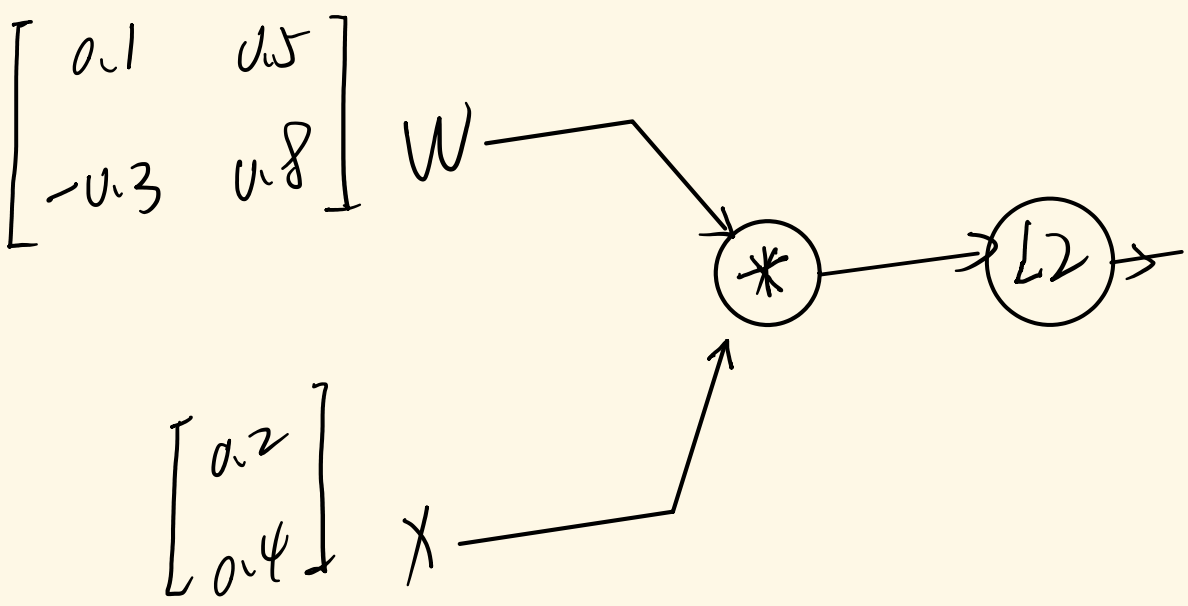
$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial q} \cdot \frac{\partial q}{\partial x}$: (Chain rule).




Vectorized Operations:

Jacobi matrix

$f(x,W) = \|W \cdot x\|^2 = \sum_{i=1}^{n} (W \cdot x)_i^2$




$q = W \cdot x = \begin{pmatrix} w_{1,1}x_1 + \cdots + w_{1,n}x_n \\ w_{n,1}x_1 + \cdots + w_{n,n}x_n \end{pmatrix}$

$f = \|x\| = x_1^2 + x_2^2 + \cdots + x_n^2$.

$\frac{\partial q_k}{\partial w_{i,j}} = 1_{k=i} x_j$

$\frac{\partial f}{\partial w_{i,j}} = \sum_k \frac{\partial f}{\partial q_k} \cdot \frac{\partial q_k}{\partial w_{i,j}}$

$\nabla_W f = 2q \cdot x^T$

$= \sum_k (2q_k)(1_{k=j} x_j)$

$= 2q_i x_j$. // diag.

Neural Network:

2-layer Neural Network:

$f = W_2 \underbrace{\max(0, W_1 x)}_{\text{non-linear function}}$.