Mathematics for Deep Learning

Partial Derivatives for Backpropagation

Inputs and Output

Let x denote an input vector. For the moment we'll leave its dimension and how we index its components unspecified.

Similarly, let y denote the output vector corresponding to the input vector, x. Its dimension and how we index its components will for a moment also remain unspecified, but note that the dimension of x and y are generally completely different.

Neural Nets and Layers

We will conceive of a neural net as consisting of L layers. Each layer takes an input vector and produces an output vector. Then

$$y = x^{(L)} \leftarrow x^{(L-1)} \leftarrow x^{(L-2)} \leftarrow \ldots \leftarrow x^{(2)} \leftarrow x^{(1)} \leftarrow x^{(0)} = x$$

The parenthesized superscripts denote layer numbers, not exponents. A net with one layer ($L = 1$) has one input vector, one output vector, and no vectors in between those. An $L = 4$ neural net has five vectors $x^{(0)}$ to $x^{(4)}$.

We still haven't said what the dimensions of the vectors are or how we index their components, but again note that the dimensions of the various $x^{(l)}$, $l = 0, \ldots, l$ are in general unrelated.

Functions to Represent Layer Operations