Deep Learning PhD course: Pre-course assignment

Niklas Wahlström
February 3, 2019

This is a pre-course assignment which you need to complete as a part of you course application. The functions you implement here you also need later for your first hand-in assignment in the course. To complete this assignment you may use any software you want, but we recommend to use the programming language Python. You do not need any knowledge about Python to use the language to complete the assignments in the course, but you do need some general programming experience and be willing to learn it on your own. If you choose to do this assignment in Python, a notebook with some hints is available from the course homepage.

Note: Try to avoid using for-loops to complete this assignment.

1 Sigmoid function

In deep learning we will often use the sigmoid function. It is defined as

\[\sigma(x) = \frac{1}{1 + e^{-x}}\]  \(1\)

and is displayed below for scalar input and output.

![Sigmoid Function Graph]

**Exercise 1.1** Implement a function `sigmoid()` which takes a scalar, vector or matrix \(x\) as input and elementwise computes the sigmoid function (1) as output. Evaluate the function on the input \(x = [1, 2, 3]\).

For example if the input is a vector then \(x = [x_1, x_2, \ldots, x_n] \in \mathbb{R}^n\) then the output should be of a vector of the same size

\[\sigma(x) = \begin{bmatrix} \sigma(x_1) \\ \vdots \\ \sigma(x_n) \end{bmatrix} = \begin{bmatrix} \frac{1}{1 + e^{-x_1}} \\ \vdots \\ \frac{1}{1 + e^{-x_n}} \end{bmatrix}\]

2 Derivative of the sigmoid function

In deep learning we also need the derivatives of all our functions. The derivative of the sigmoid function is

\[\sigma'(x) = \frac{e^{-x}}{(1 + e^{-x})^2} = \sigma(x)(1 - \sigma(x))\]  \(2\)

**Exercise 2.1** Implement a function `sigmoid_diff()` which takes a scalar, vector or matrix \(x\) as input and element-wise computes (2) as output. Evaluate the function on the input \(x = [1, 2, 3]\). You might find your function `sigmoid()` useful.

Also this function should act elementwise, meaning that if \(x = [x_1, x_2, \ldots, x_n] \in \mathbb{R}^n\) then

\[\sigma'(x) = \begin{bmatrix} \sigma'(x_1) \\ \vdots \\ \sigma'(x_n) \end{bmatrix} = \begin{bmatrix} \sigma(x_1)(1 - \sigma(x_1)) \\ \vdots \\ \sigma(x_n)(1 - \sigma(x_n)) \end{bmatrix}\]
3 Reshaping

In deep learning we often need to reshape data. For example, a gray-scaled image is represented by a 2D matrix of shape \((\text{length}, \text{height})\). When you read this into the algorithms we learn in the beginning of the course we need to reshape these 2D matrices into 1D vectors.

$$\begin{bmatrix}
0.0 & 0.0 & 0.8 & 0.9 & 0.6 & 0.0 \\
0.0 & 0.9 & 0.6 & 0.0 & 0.8 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.9 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.9 & 0.6 & 0.0 \\
0.0 & 0.0 & 0.9 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.8 & 0.9 & 0.9 & 0.9 & 0.9 \\
\end{bmatrix} \Rightarrow \text{Reshape} \begin{bmatrix}
0.0 & 0.0 & 0.8 & 0.9 & 0.6 & 0.0 \\
\cdots \\
0.0 & 0.8 & 0.9 & 0.9 & 0.9 \\
\end{bmatrix}$$

**Exercise 3.1** Implement a function `image2vec()` that takes an input of shape \((\text{length}, \text{height})\) and returns a reshaped vector of shape \((\text{length} \times \text{height})\). Evaluate the function on the matrix

$$X = \begin{bmatrix}
0.1 & 0.3 \\
0.2 & 0.0 \\
0.9 & 0.4 \\
\end{bmatrix}$$

Do not hardcode the dimensions of the matrix to reshape. Instead, look up the quantities you need, for example with the `shape` command in Python or `size` command in Matlab.

4 Normalization

In deep learning we often need to normalize our data. If we want to normalize the columns in a matrix of shape \((n \times p)\), we change each column by dividing with its norm

$$X = \begin{bmatrix} x_1 & \cdots & x_m \end{bmatrix} \Rightarrow X_{\text{norm}} = \begin{bmatrix} x_1 \|x_1\| & \cdots & x_p \|x_p\| \end{bmatrix}.$$ 

For example if we use the 2-norm, we would for the following matrix get

$$X = \begin{bmatrix} 2 & 1 & 1 \\
0 & -4 & 3 \end{bmatrix} \Rightarrow X_{\text{norm}} = \begin{bmatrix} \frac{2}{\sqrt{2^2+0^2}} & 0 & \frac{1}{\sqrt{1^2+(-4)^2}} & \frac{1}{\sqrt{1^2+3^2}} \\
\frac{1}{\sqrt{1^2+0^2}} & -4 & \frac{1}{\sqrt{1^2+(-4)^2}} & \frac{1}{\sqrt{1^2+3^2}} \end{bmatrix} = \begin{bmatrix} 0.71 & 0 & 0.2 & 0.3 \\
0.2 & -0.57 & 0.2 & 0.3 \end{bmatrix}. \quad (3)$$

**Exercise 4.1** Implement `normalizeColumns()` which takes a matrix \(X\) as the input and computes an output of the same shape where the columns have been normalized using the 2-norm. Evaluate it on the matrix above.