Deep Neural Networks

Assignment 11

Assignment due by: 22.01.2020, Discussions on: 29.01.2020

Question 1 Backpropagation in a Deep Learning Framework (10 points)

We have made a small deep learning 'framework' to demonstrate some of the principles that drive the big frameworks like *TensorFlow*, *PyTorch*, etc. We have omitted a few pieces of code for you to complete and to demonstrate that this simple framework can be used to successfully learn network weights.

The *autograd* framework is mostly inspired by PyTorch and only capable of first-order derivatives. It uses the `Tensor` as its central object but removes the need to deal with an explicit graph and session. There is no graph/runtime optimization, each forward pass will build a new graph on the fly. The different operations, variables and inputs are all subclasses of `Tensor` and each of them is responsible for implementing its behavior during the forward and backward pass.

(a) Complete the backward method for the base class Tensor. (2 points)

(b) Assign the gradients for all inputs of ReduceMean, Add, Mul, MatMul, ReLU, and Sigmoid. Neg and Sqrt are already completed as an example. (6 points)

(c) Implement the mean squared error (mse) function. (1 point)

(d) Run the provided linear regression script and attach the resulting plot. Comment on the result. (1 point)

Question 2 Backpropagation through a convolution (6 points)

The formula to calculate a one-channel 2D convolution is:

$$\mathsf{Z}_{ij} = \sum_{k,l} \mathsf{V}_{i+k,j+l}\mathsf{K}_{k,l}$$

where $\mathsf{V}_{i,j}$ is the input image and $\mathsf{K}_{k,l}$ is the kernel. Assuming that this convolution appears somewhere in a neural network and given the partial derivative $\frac{\partial \mathcal{L}}{\partial \mathsf{Z}_{ij}}$ calculate the partial derivatives:

$$\frac{\partial \mathcal{L}}{\partial \mathsf{K}_{mn}} \text{ and } \frac{\partial \mathcal{L}}{\partial \mathsf{V}_{mn}}$$

Show that the resulting formulae are again convolutions.

Hint: You don't need to worry about edge effects (padding etc.) for this question. As a first step you should calculate the partial derivatives $\frac{\partial \mathsf{Z}_{ij}}{\partial \mathsf{K}_{mn}}$ and $\frac{\partial \mathsf{Z}_{ij}}{\partial \mathsf{V}_{mn}}$. Also, for the Kronecker delta: $\delta_{a,b} = \delta_{a-b,0}$

Question 3 Convolution by Hand (4 points)

Convolve the following two kernels by hand:

$$k_1 = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ -2 & 0 & -1 \end{bmatrix},\ k_2 = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$$

use a stride of 1 for k_1 and a stride of 2 for k_2,

over the following (pre-padded) image:

$$I = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 2 & 3 & 5 & 6 & 0 \\ 0 & 0 & 1 & 5 & 4 & 2 & 0 & 0 \\ 0 & 3 & 4 & 8 & 0 & 1 & 2 & 0 \\ 0 & 3 & 2 & 3 & 0 & 1 & 3 & 0 \\ 0 & 0 & 4 & 6 & 3 & 1 & 5 & 0 \\ 0 & 1 & 2 & 5 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$