Introduction to Deep Learning - Week 2

<http://playground.tensorflow.org> for an online platform to compose MLPs

To train an MLP we need to calculate the hidden layer hyper-parameters automatically and fast.

The graph of derivatives is the MLP graph with reversed arrows. The chain rule is the sum of all the paths on this diagram. If you use the chain rule to expand all the derivatives then each term will relate to one backwards path through the MLP.

To use SGD we use a loss function on the sigmoid-ed output of the neuron.

Backpropagation uses the fact that the later derivatives in an MLP are reused as you go backwards in the network, meaning the value of each node derivative only has to be calculated once. First you make a forward-pass to determine how the node transforms its inputs and get the values at that point.

Matrix multiplication can be done fast in comparison to loops, and done on both CPU (e.g. BLAS) and GPU (e.g. cuBLAS).

Tensorflow

Whilst more recent versions of Tensorflow support eager execution, typically In Python you merely define the graph and create a session which executes the graph using C++.

Note that the session owns all of the resources and needs cleaning as it eats RAM.

Tensorflow API:

* Tf.InteractiveSession()
* tf.trainable\_variables() – list all the trainable variables
* tf.Print() – is not visible but stores the values synchronously so they can be viewed later (available in jupyter stdout or in Tensorboard)
* tf.placeholder()
* tf.reduce\_mean() – to define a loss function
* tf.train.GradientDescentOptimizer() – define an optimiser than can be called with opt.minimise(loss\_function)
* tf.Model checkpoints help when a model takes a long time and you want to be able to perform training in multiple sessions.

Summaries to view data

* tf.summary.scalar()
* tf.summary.FileWriter()
* tf.summary.merge\_all()
* writer.add\_summary()
* writer.flush() – this forces the summary writer to send any buffered data to storage. To view in web interface use $ tensorboard –logdir=./logs and is available at <http://localcost:6006>

Additional Reading

* <https://jacobbuckman.com/2018-06-25-tensorflow-the-confusing-parts-1/>
* <https://www.tensorflow.org/tutorials>