

# Backpropagation

Vladimir Feinberg

July 11, 2017

Neural networks are real-valued circuits constructed from simple functions. This compositionality enables complex expressiveness of the function family through hidden units, while permitting efficient function evaluation and gradient evaluation. This trick is so essential to NN training that it is worthwhile to apply the trick to specific NN instances to understand how the flow of gradients travels.

[TODO]. Backprop derivations for MLPs, CNNs (with various forms of convolution/dilation/striding), forked FFNNs in general. Follow 6.5 in DL book. Resources probably exist online; prefer to link them instead of putting content here.

[TODO]. Batch Normalization backprop. See [notes here](#), but they include bias terms, which is redundant with mean.

[TODO]. Various RNNs.