

Backpropagation

Vladimir Feinberg

July 4, 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 6.5 in DL book] -> give derivations for all DNNs, conv, etc in a new doc.

[TODO Batch Normalization backprop], see [notes here](#), but they include bias...