NORMALIZED INITIALIZATION

OF NEURAL NET PARAMETERS

Uniform distribution

$$W_{i,j} \sim U\left(-\sqrt{\frac{6}{m+n}}, \sqrt{\frac{6}{m+n}}\right)$$

Parameter in a fully-connected layer.

Number of inputs

Number of outputs

Chris Albon