

This is a note for the paper: Dropout as a Low-Rank Regularizer for Matrix Factorization (2018)

It is an interesting paper discussing about the relationship between dropout and nuclear norm penalty in the context of matrix approximation and matrix factorization. It reveals that in such context, dropout is equivalent to the squared nuclear norm regularizer which induces low rank of approximated matrix (low rank means sparsity of singular values) and acts similarly as data dependent adaptive PCA. Importantly, the author tells us how to choose the proper dropout rate when the size of matrices varies.