Journal of Machine Learning **ISSN**: 2790-2048(e), 2790-203X(p)

Enhancing Accuracy in Deep Learning Using Random Matrix Theory

Leonid Berlyand[1], Etienne Sandier[2], Yitzchak Shmalo[1], and Lei Zhang * [3]

[1]Department of Mathematics, Pennsylvania State University, University Park, PA 16802, USA.
[2]LAMA-CNRS UMR 8050, Université Paris-Est Créteil, Créteil 94010, France.
[3]Institute of Natural Sciences, Shanghai Jiao Tong University, Shanghai 200240, P.R. China.

Abstract. We explore the applications of random matrix theory (RMT) in the training of deep neural networks (DNNs), focusing on layer pruning that reduces the number of DNN parameters (weights). Our numerical results show that this pruning leads to a drastic reduction of parameters while not reducing the accuracy of DNNs and convolutional neural network (CNNs). Moreover, pruning the fully connected DNNs actually increases the accuracy and decreases the variance for random initializations. Our numerics indicate that this enhancement in accuracy is due to the simplification of the loss landscape. We next provide rigorous mathematical underpinning of these numerical results by proving the RMT-based Pruning Theorem. Our results offer valuable insights into the practical application of RMT for the creation of more efficient and accurate deep-learning models.

Keywords:
Deep learning,
Marchenko-Pastur distribution,
Random matrix theory,
Increasing accuracy,
Pruning.

Article Info.:
Volume: 3
Number: 4
Pages: 347 - 412
Date: December/2024
doi.org/10.4208/jml.231220

Article History:
Received: 20/12/2023
Accepted: 29/08/2024

Communicated by:
Zhi-Qin John Xu

Contents

*Corresponding author. lzhang2012@sjtu.edu.cn