PHD PROJECT: DEEP LEARNING AND HIGH-DIMENSIONAL STATISTICS

JOHANNES HEINY

Residual Networks (ResNet) unquestionably represent successful deep learning models, whose remarkable performance in image classification tasks has made them an essential tool for various applications such as object recognition, facial recognition, and medical image analysis [2]. However, theoretical analyses explaining their success have been limited in recent years.

This project focuses on the intersection of deep learning theory, random matrix theory and high-dimensional statistics. In deep learning, a notable phenomenon is the presence of a greater number of parameters compared to the available sample size. This scenario, known as the over-parameterized regime, presents intriguing challenges and opportunities. One particularly potent mathematical tool that can shed light on this situation is random matrix theory [1, 3, 4, 5].

Random matrix theory offers a powerful framework for exploring the theoretical properties of deep neural networks. This framework enables us to analyze the behavior of large parameter matrices that arise in the context of deep learning models. By leveraging concepts from random matrix theory, we aim to derive valuable insights into the intricate dynamics of over-parameterized neural networks.

Contemporary machine learning models, like ResNet [2] or some other models, tend to exhibit a characteristic of being significantly over-parameterized. This characteristic allows them to be finely tuned and adapt to the nuances of the training data. The prevalence of these over-parameterized models has propelled them to the forefront of cutting-edge performance across a spectrum of learning tasks. However, this situation contrasts with classical statistical learning theory, where models with excessive parameters are traditionally associated with a higher risk of overfitting, leading to poor generalization. The success of these over-parameterized models, despite the theoretical concerns, has spurred a recent surge in research efforts aimed at unraveling the underlying reasons behind their remarkable performance.

This project requires a **strong background in probability theory and statistics**, as well as self-motivation, enthusiasm and the willingness to carry out significant research within a lively area of modern mathematics. The supervisor and the PhD student will choose the PhD topic together to ensure that personal preferences and strengths are accounted for, with the student taking as much initiative as possible.

References

[1] Basrak, B., Cho, Y., Heiny, J., and Jung, P. Extreme eigenvalue statistics of m-dependent heavy-tailed matrices. *Ann. Inst. Henri Poincaré Probab. Stat. 57*, 4 (2021), 2100–2127.

[2] He, K., Zhang, X., Ren, S., and Sun, J. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2016), pp. 770–778.

Date: March 10, 2024.

[3] Heiny, J. Large sample correlation matrices: a comparison theorem and its applications. *Electron. J. Probab. 27* (2022), Paper No. 94, 20.

[4] Heiny, J., Mikosch, T., and Yslas, J. Point process convergence for the off-diagonal entries of sample covariance matrices. *Ann. Appl. Probab. 31*, 2 (2021), 538–560.

[5] Heiny, J., and Yao, J. Limiting distributions for eigenvalues of sample correlation matrices from heavy-tailed populations. *Ann. Statist. 50*, 6 (2022), 3249–3280.

More information about my research can be found at
https://www.su.se/english/profiles/johe3032-1.640812

Department of Mathematics, Stockholm University, Albano hus 1, 10691 Stockholm, Sweden
Email address: johannes.heiny@math.su.se