Study of Generalized Condition Number – Project Proposal

Yanyao Shen

1 Motivation

Condition number κ of a function describes how much the output value of the function can change for a small change in the input argument. For example, the condition number associated with the linear equation $Ax = b$, it is not hard to see that $\kappa = \|A^{-1}\| \cdot \|A\|$, which is the ratio of maximum and minimum eigenvalue of matrix A. The condition number has good properties: it describes how robust a linear system is, and also describes the convergence rate for optimizing some quadratic functions using first-order methods such as gradient descent.

For complex systems such as neural networks, people rarely use the concept of condition number, since it is easy to come up with a worst case example that is ill-conditioned. However, different deep architectures empirically behave quite differently in terms of convergence. Due to this gap, we decide to explore tools that are able to analyze the convergence performance given an architecture. Since there are always bad cases, we try to get some analysis based on mean field theory.

2 Technical Merit

The challenge of the problem is that analyzing the expected condition number is not trivial, even in the linear case. We need to calculate the expected condition number of a Wishart matrix ([1] may be a good reference for us). In the nonlinear case, we need to find a way to calculate the derivative due to non-linearity. Classical neural network theory literatures may be good references for us.

3 Broader Impact

The impact of studying this problem may help us design better architectures, as well as better analyzing an existing model. While there are works that try to explain neural network's convergence, the idea of calculating based on mean field theory and condition number is (to my knowledge) novel.

4 Project Plan

- Learn theory related to random matrices and statistical properties of eigenvalues.
- Learn theory related to the nonlinear transformation function.
- Calculate the expected condition number.
- Do experiments.

References

[1] T. Tao and V. Vu, "Random matrices: The distribution of the smallest singular values," *Geometric And Functional Analysis*, vol. 20, no. 1, pp. 260–297, 2010. 2