

Generalization Ideas in Deep Learning

Marcel Bruckner

Seminar: Optimization and Generalization in Deep Learning

Abstract

Write a brief abstract here (around 150 words).
Generalization is nice.

1 What is generalization and why do we want it

2 What is a matrix norm and which do we use

2.1 What is a matrix norm

2.2 Why do we use matrix norms as measures for capacity bounds

2.3 Which norms do we use

- l_2 norm
- l_1 -path norm
- l_2 -path norm
- spectral norm
- spectral norm
- spectral norm
- spectral norm
- spectral norm
- spectral norm

3 What is sharpness

4 A few remarks

Each report should include an introduction describing the problem, the motivations and a brief outline. The main approach should then be described and discussed in separate sections, followed by experimental results (when applicable) and conclusions.

- Please use citations when appropriate. Again, you are not expected to read through all the references appearing in your assigned paper. Add your citations in bibtex format into the file `egbib.bib`. An example is [3].
- You can use the theorem environment to write theorems. An example:

Theorem 1. *Let p be a prime number. Then, for any $a \in \mathbb{N}$, $a^p - a$ is evenly divisible by p . More formally,*

$$a^p \equiv a \pmod{p}. \quad (1)$$

- Please keep all your formulas numbered.
- The report should be 4 to 6 pages long (not including citations).
- Reports must be in English.
- Please do not change the layout (*e.g.*, do not change page margins, font size, etc.).

References

- [1] Peter L. Bartlett and Shahar Mendelson. Rademacher and gaussian complexities: Risk bounds and structural results. In David Helmbold and Bob Williamson, editors, *Computational Learning Theory*, pages 224–240, Berlin, Heidelberg, 2001. Springer Berlin Heidelberg.
- [2] Kenji Kawaguchi, Leslie Pack Kaelbling, and Yoshua Bengio. Generalization in deep learning, 2017.
- [3] Behnam Neyshabur, Srinadh Bhojanapalli, David McAllester, and Nati Srebro. Exploring generalization in deep learning. In *Advances in Neural Information Processing Systems*, pages 5947–5956, 2017.
- [4] Behnam Neyshabur, Ruslan Salakhutdinov, and Nathan Srebro. Path-sgd: Path-normalized optimization in deep neural networks. *CoRR*, abs/1506.02617, 2015.
- [5] Behnam Neyshabur, Ryota Tomioka, and Nathan Srebro. Norm-based capacity control in neural networks. *CoRR*, abs/1503.00036, 2015.
- [6] Chiyuan Zhang, Samy Bengio, Moritz Hardt, Benjamin Recht, and Oriol Vinyals. Understanding deep learning requires rethinking generalization. *CoRR*, abs/1611.03530, 2016.