

# Jun 22 - Jun 29 Seminar Review

Jiaxin Hu

June 26, 2020

## 1 MAD 6.24

**Title:** The generalization error of overparametrized models: Insights from exact asymptotics

**Representer:** Andrea Montanari

In supervised learning, we want to learn a function that predicts the label given the new feature vectors. Classical learning theories address the empirical risk minimization problem and argue the uniform convergence of the estimation. The optimization is also discussed in the convex realm. However, modern deep learning challenges the above three viewpoints. The optimization problem is highly non-convex, training error is significantly smaller than testing error, and the minimization problem becomes GD or SGD. Dr. Montanari takes Ridge Regression with random feature maps and Max Margin as examples to understand this phenomenon in detail. He discusses the models in the high-dimensional regime, where the number of parameters is comparable with sample size. He also generalizes the discussion to kernel ridge regression and the soft-margin classification.

### Problem:

- What's the problem the seminar wants to solve?  
**Possible answer:** I guess Dr. Montanari aims to use ridge regression and max margin to address the non-convex optimization and non-uniform convergence problem in over parametrized model, such as deep learning.
- Dr. Montanari concludes the double descent phenomenon in model performance against sample size and good generalization with no/minimal regularization. Does this conclusion mean that we should not worry about the overparameter in the modern deep learning models and we should not use the "large" regularization?