

Alignment and convergence of kernels in deep learning

Matus Telgarsky
University of Illinois
myemail@myemail.com

Despite nonconvexity and even nonsmoothness of deep network training, both the gradient descent trajectory and corresponding sequence of kernels are convergent under minimal assumptions. This talk will explore the basic proof technique, namely an alignment property of gradient descent that appears more fundamental than previously discovered implicit bias phenomena, and discuss consequences on sample complexity and comparison to the standard initial kernel (what is typically called the neural tangent kernel).