COLLOQUIUM

Sungwoo Jeong

Cornell University

Convergence of two kernel algorithms: Continuous analogues of the SVD and the Cholesky decomposition.

MONDAY, February 2nd at 11:30 AM
Jones 303, 5747 S. Ellis Ave. Chicago, IL 60637

ABSTRACT

Kernels have been shown to be effective in numerous applications. We discuss some misconceptions and facts about why kernels are powerful in practice. To investigate this, we consider two expansions that encode fundamental kernel structures: the kernel analogue of the SVD and the Cholesky decomposition.

First, the convergence of the kernel SVD (SVE) is equivalent to the existence of a corresponding functional space such as the reproducing kernel Hilbert space (RKHS). For general kernels, such as self-attention in neural networks, it is still unclear whether the SVE converges. We prove a surprising result showing that kernel continuity alone is not enough to guarantee this convergence. At the same time, we provide a new sufficient condition for convergence that helps explain why kernels work well in practice.

The kernel Cholesky algorithm is another fundamental tool used in applications such as Gaussian process regression (Bayesian inference on functions). While it is empirically observed that the Cholesky algorithm converges for smooth kernels, no rigorous result exists for kernels with weaker regularity than C^2. We prove a new convergence result for Lipschitz continuous kernels, together with an explicit convergence rate that sharply agrees with what is observed in practice.