Selection Dynamics for Deep Neural Networks

Abstract: We present a partial differential equation framework for deep residual neural networks and for the associated learning problem. This is done by carrying out the continuum limits of neural networks with respect to width and depth. We study the wellposedness, the large time solution behavior, and the characterization of the steady states of the forward problem. Several useful time-uniform estimates and stability/instability conditions are presented. We state and prove optimality conditions for the inverse deep learning problem, using standard variational calculus, the Hamilton-Jacobi-Bellmann equation and the Pontryagin maximum principle. This serves to establish a mathematical foundation for investigating the algorithmic and theoretical connections between neural networks, PDE theory, variational analysis, optimal control, and deep learning.