EPSTEIN INSTITUTE SEMINAR ▪ ISE 651

Asymptotic Analysis of Deep Residual Networks

ABSTRACT - Residual networks (ResNets) have displayed impressive results in pattern recognition and, recently, have garnered considerable theoretical interest due to a perceived link with neural ordinary differential equations (neural ODEs). This link relies on the convergence of network weights to a smooth function as the number of layers increases. We investigate the properties of weights trained by stochastic gradient descent and their scaling with network depth through detailed numerical experiments. We observe the existence of scaling regimes markedly different from those assumed in neural ODE literature. Depending on certain features of the network architecture, such as the smoothness of the activation function, one may obtain an alternative ODE limit, a stochastic differential equation or neither of these. These findings cast doubts on the validity of the neural ODE model as an adequate asymptotic description of deep ResNets and point to an alternative class of differential equations as a better description of the deep network limit.

Further, we show in each case the convergence of the classical backpropagation algorithm to a backward differential equation, suggesting ways to efficiently train the deep network limits.

Alain Rossier
3rd Year PhD Student,
Oxford University

SPEAKER BIO – Alain Rossier is a 3rd year PhD student at the University of Oxford under the supervision o Prof. Rama Cont. He is interested in the theory of deep learning, in particular in the asymptotic behavior anc the dynamics of gradient descent of deep networks. He is visiting Prof. Renyuan Xu at USC to pursue ongoing collaborations.

USC Viterbi
School of Engineering
Daniel J. Epstein Department of Industrial and Systems Engineering

TUESDAY, MARCH 22, 2022
3:30 PM – 4:50 PM
ZOOM/ONLINE *PLEASE EMAIL OWH@USC.EDU FOR PASSWORD*