For Favour of Posting

Department of Mathematics
The Chinese University of Hong Kong

數學系
香港中文大學

Phone: (852) 3943 7988 • Fax: (852) 2603 5154 • Email: dept@math.cuhk.edu.hk (Math. Dept.)
Room 220, Lady Shaw Building, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong

SIAM Student Chapter Seminar Series

On the Convergence and Generalization of Neural networks: Applications of Statistical Mechanics, Random Matrix Theory and Langevin Dynamics

Mr. Xinjiang Wang
SenseTime Corporation

Abstract: Analyzing learning from examples in feedforward neural networks is significant for both understanding and designing neural networks. The learning process using stochastic gradient descent has deep analogy to the Langevin dynamics in diffusion process. The stochastic training leads to a Gibbs distribution of networks characterized by a temperature related with the learning rate and batch size during training. When a network converges to an equilibrium point, an analysis on the generalization ability follows. Exact treatment of the quenched disorder of the sample entails the use of replica theory in describing generalization. Of special interest is the generalization curve, namely, the curve describing the relation between generalization error and the number of training examples. Different learning tasks including activation functions and regularization effect are analyzed within this framework. Lastly, the training dynamics and generalization ability of a normalization trick called batch normalization (BN) are explored. It is found that BN enables higher learning rates and prevents overfitting when training deep neural networks (DNNs). First, we show an analytical expression that represents the regularization effect in BN, which can be decoupled into two parts, a weight normalization and a penalty on its scale parameters. Its properties can be well explained by studying this penalty. Second, convergence behavior of BN is analyzed by exploring its dynamic of online training, showing that it enables
large maximum and effective learning rates.

Date:	April 27, 2018 (Friday)
Venue:	Room 222, Lady Shaw Building, The Chinese University of Hong Kong, Shatin
Time:	3:00pm – 4:30pm

All are Welcome