The Hong Kong University of Science and Technology

Department of Mathematics

<u>Seminar on Data Science and Machine Learning</u>

Global Convergence of Block Coordinate Descent in Deep Learning

by

Mr. Tim Tsz-Kit Lau
Department of Statistics,
Northwestern University, Illinois, USA

<u>Abstract</u>

Deep learning has aroused extensive attention due to its great empirical success. The efficiency of the block coordinate descent (BCD) methods has been recently demonstrated in deep neural network (DNN) training. However, theoretical studies on their convergence properties are limited due to the highly nonconvex nature of DNN training. In this paper, we aim at providing a general methodology for provable convergence guarantees for this type of methods. In particular, for most of the commonly used DNN training models involving both two- and three-splitting schemes, we establish the global convergence to a critical point at a rate of $O(1/k)$, where k is the number of iterations. The results extend to general loss functions which have Lipschitz continuous gradients and deep residual networks (ResNets). Our key development adds several new elements to the Kurdyka-Lojasiewicz inequality framework that enables us to carry out the global convergence analysis of BCD in the general scenario of deep learning.

Date: Friday, 12 July 2019
Time: 3:00p.m. - 4:00p.m.
Venue: Room 2463, Academic Building
(near Lifts 25 - 26), HKUST

All are welcome!