Dr. Anjan Kumar Chakrabarty, IIT Guwahati, delivered lectures in the Guest lecture series for the course Introduction to Functional Analysis (MATH F341) on February 13, 2019.

Prof. B.L.S. Prakasa Rao, Emeritus Professor of Indian Statistical Institute, delivered a talk on 'Big Data and High Dimensional Data Analysis' on February 28, 2019

Abstract

Over the last twenty years, more and more corporations are adapting to data-driven approach to have targeted services, reduced risks and improved performance. They are implementing specialised data analytic programs to collect, store, manage and analyse large data sets or what is now called BIG DATA. Analysing large size of economic and financial data is challenging. BIG DATA has unique features that are not shared by the traditional data sets. BIG DATA sets are characterised by massive sample size and high-dimensionality. Massive sample size allows one to unravel hidden patterns associated with small sub populations. Modelling the intrinsic heterogeneity of BIG DATA requires better statistical methods. There are several phenomena associated with high-dimensionality such as noise accumulation, spurious correlation and incidental endogeneity. Traditional statistical methods are inappropriate to tackle such problems. There are also many types of event we can think of when we have a potentially large number of measurable parameters/covariates quantifying the event but a relatively few instances of that event. Example: few patients with a given genetic disease but a large number of genes which might cause this event. In statistical terms, the number of parameters p is large as compared to the number of observation. This type of data is termed as HIGH-DIMENSIONAL DATA. The basic methodology which was used in classical statistical methods is not applicable for analysing such data. He has discussed some problems that arises in analysis of BIG DATA and HIGH-DIMENSIONAL DATA.