数学与系统科学研究院

计算数学所网络学术报告

报告人： 史斌 博士

（ *University of California, Berkeley* ）

报告题目：

Understanding the Acceleration Phenomenon via High-Resolution Differential Equations ---- The Effect of Gradient Correction

邀请人： 周爱辉 研究员

报告时间：2021 年 4 月 27 日（周二）

上午 **10:00-11:00**

报告工具：腾讯会议（ID：759 117 436）

Abstract：

During the past decade, artificial intelligence has obtained a great success. Meanwhile, the brand-new math tool --- the continuous differential equation -- is introduced to understand and analyze the discrete algorithms. In ICM 2018, Michael I. Jordan gives a plenary talk and proposes some open problems about this rising and hot topic. One of the most significant problems is why the simple first-order algorithm proposed by Yurii Nesterov in 1983 will lead to the essential acceleration, compared to the classical gradient descent. (Actually, Nesterov also mentioned it in his invited talk of ICM 2010.) In this talk, I will present a new high-resolution ODE framework, based on the first-order approximation for simplifying the differential equations in physics, to understand and analyze Nesterov's accelerated algorithms. The gradient correction, a small but key term, is discovered as an essential factor to bring about the acceleration. Meanwhile, I will show the numerical comparison between the ODEs and the algorithms to show the intuition. Accompanying with the new discovery, I will also show the two new math techniques used in Lyapunov analysis,

(1) phase-space representation: directly transfer the continuous Lyapunov functions to the discrete ones;

(2) dimensional analysis: as an intuition used to estimate the derivative of Lyapunov functions.

Moreover, I will present a new discovery about Nesterov's algorithm itself, which minimizes the squared gradient norm at an inverse cubic rate. Finally, I will show some new progresses and open questions for the high-damped case ($\alpha > 3$) based on the high-resolution ODE framework above.

欢迎大家参加！