

数学与系统科学研究院
计算数学所网络学术报告

报告人: **Dr. Bin Shi**

(*University of California, Berkeley, U.S.*)

报告题目:

On gradient- based optimization algorithms

邀请人: 刘歆 研究员

报告时间: 2021 年 4 月 20 日 (周二)

上午 9:00-10:00

报告工具: 腾讯会议 (ID: 707 294 582)

入会密码: 311311

Abstract:

In these talks, I will present the understandings about gradient-based optimization algorithms from the continuous perspective.

(1) Nesterov's accelerated gradient descent is one of the most beautiful but mysterious algorithms. It is an important question to understand the acceleration phenomena behind Nesterov's "Estimate Sequence". Here, I will show the understanding from the ODE techniques. The key point is the introduction of first-order approximation, which leads to the discovery of gradient correction. Also, I will present the phase-space representation is the core technique connecting the continuous ODE and the discrete algorithms. Meanwhile, I will present the dimensional analysis from physics is used to discriminate the inequality intuitively.

(2) Stochastic gradient descent succeeds in the nonconvex optimization, especially in deep learning. Here, I will, from the continuous perspective, show the key characteristics of SGD --- learning-rate dependent noise, contrasting to GD and SGLD. The evolutive behavior of the expectation is different from the single particles. The noise plays the essential role to make the particles jump out of local minima through the lowest saddle. Meanwhile, the learning rate leads to the convergence rate can be quantified. Finally, I will show the learning rate will sharply change the convergence rate.

欢迎大家参加！