


复旦大学数学科学学院

数学综合报告会

报告题目：Policy Gradient Flow for Stochastic Control

报告人： 任振杰 教授（University Evry Paris-Saclay）

时间：2026-08-25 星期二 10:10-11:10

地点：光华楼东主楼1801

报告摘要:

Reinforcement learning has achieved remarkable success in landmark applications such as AlphaGo and AlphaFold, and is increasingly expected to play a central role in emerging domains such as autonomous driving. A substantial body of work has established theoretical foundations for classical reinforcement learning algorithms, including temporal-difference learning, policy iteration, and Q-learning. By contrast, policy gradient descent, despite its widespread use, remains less well understood from a theoretical perspective, largely because of its non-convex structure and the complexity of the underlying functional space. Existing convergence analyses typically require uniform regularity assumptions on the policy, viewed as a feedback control function, along the gradient flow ; moreover, the resulting convergence rates depend on these regularity bounds. In this work, we significantly strengthen the existing convergence theory. Our key insight is to relate policy gradient descent to a mirror flow on the space of probability measures over controlled trajectories, where the Bregman divergence is induced by the convex control cost. We rigorously prove a JKO-type convergence result showing that the discrete-time mirror flow converges to its continuous-time counterpart, and we identify the limiting dynamics as a preconditioned policy gradient descent flow on the space of control processes. Leveraging this mirror-flow perspective, we establish an exponential convergence rate that is independent of policy regularity. The convergence holds both in Bregman divergence and in the value of the associated control problem.

非线性数学模型与方法教育部重点实验室
中法应用数学国际联合实验室
上海市现代应用数学重点实验室
复旦大学数学研究所