

## ***Policy Evaluation, Policy Gradient, and Actor-Critic Learning in Continuous Time and Space: Theory and Algorithms***

ABSTRACT - We propose a unified framework to study policy evaluation (PE) and policy gradient (PG) for reinforcement learning (RL) in continuous time and space. Mathematically, PE is to devise a data-driven Feynman-Kac formula without knowing any coefficients of a partial differential equation. We show that this problem is equivalent to maintaining the martingale condition of a process. On the other hand, we represent the gradient of the value function with respect to a given parametric stochastic policy as the expected integration of an auxiliary running reward function that can be evaluated using samples and the current value function. This effectively turns PG into a PE problem, enabling us to apply the martingale approach. Based on this analysis, we propose various actor-critic algorithms for RL, where we learn and update value functions and policies simultaneously and alternately. We demonstrate the algorithms by simulations in concrete examples. The talk is based on two joint papers with Yanwei Jia.



**Dr. Xunyu Zhou**

Liu Family Professor of Financial Engineering; Director, Nie Center for Intelligent Asset Management, Dept. of Industrial Engineering and Operations Research, Columbia University

**SPEAKER BIO** – Xunyu Zhou is the Liu Family Professor of Financial Engineering and the Director of the Nie Center for Intelligent Asset Management at Columbia University. He was the Nomura Professor of Mathematical Finance at University of Oxford before joining Columbia in 2016.

His research covers stochastic control, dynamic portfolio selection, asset pricing, behavioral finance, and time inconsistency. Currently his research focuses on continuous-time reinforcement learning and applications to optimization broadly and to wealth management specifically. He is a recipient of the Wolfson Research Award from The Royal Society, the Outstanding Paper Prize from SIAM, the Alexander von Humboldt Research Fellowship, and the Croucher Senior Research Fellowship. He was an invited speaker at the 2010 International Congress of Mathematicians, a Humboldt Distinguished Lecturer at Humboldt University and an Archimedes Lecturer at Columbia. He is both an IEEE Fellow and a SIAM Fellow.

Xunyu Zhou received his PhD in Operations Research and Control Theory from Fudan University in 1989.