



復旦大學
FUDAN UNIVERSITY

Applied Math
Ph.D. Seminar

**ϕ -Update: A Class of Policy Update Methods with
Policy Convergence Guarantee**

Speaker: Wenye Li (Fudan University)

Time: 2025-04-17, 16:10 to 17:00

Location: Rm 1801, Guanghai East Tower

Advisor: Ke Wei (Fudan University)

Abstract: Policy optimization refers to a family of effective algorithms which search in the policy space based on policy parameterization to solve reinforcement learning problems. Inspired by the similar update pattern of softmax natural policy gradient and Hadamard policy gradient, we propose to study a general policy update rule called ϕ -update, where ϕ refers to a scaling function on advantage functions. Under very mild conditions on ϕ , the global asymptotic state value convergence of ϕ -update is firstly established. Then we show that the policy produced by ϕ -update indeed converges, even when there are multiple optimal policies. This is in stark contrast to existing results where explicit regularizations are required to guarantee the convergence of the policy. The exact asymptotic convergence rate of state values is further established based on the policy convergence. Lastly, we establish the global linear convergence of ϕ -update.