

Some notes on Reinforcement Learning for stochastic control in finance

Samy Mekkaoui

April 1, 2025

These notes are complementary to the course *Apprentissage automatique et contrôle stochastique* taught by Huy  n Pham at the M2 Probabilit  s et Finance. They aim to provide a deeper understanding of the *Reinforcement Learning for stochastic control problems* section.

Contents

1	Basis of Reinforcement Learning	2
1.1	An overview of MDP	2
1.2	From MDP to Learning	2
1.3	Value-based methods	2
1.4	Policy-based methods	2
2	RL in Continuous time	3
2.1	The continuous-time framework	3
2.2	Policy gradient methods in continuous time	3
2.2.1	Policy gradient representation	3
2.2.2	Actor-critic algorithms	3
2.3	Q-learning in continuous time	3
2.3.1	Q-learning and Hamiltonian function	3
2.3.2	Q-learning and actor critic algorithms	3
3	Some applications in finance	4
3.1	Systemic risk	4
3.2	Portfolio Optimization	4
3.3	Optimal execution	4

1 Basis of Reinforcement Learning

1.1 An overview of MDP

1.2 From MDP to Learning

1.3 Value-based methods

1.4 Policy-based methods

2 RL in Continuous time

2.1 The continuous-time framework

2.2 Policy gradient methods in continuous time

2.2.1 Policy gradient representation

2.2.2 Actor-critic algorithms

2.3 Q-learning in continuous time

2.3.1 Q-learning and Hamiltonian function

2.3.2 Q-learning and actor critic algorithms

3 Some applications in finance

3.1 Systemic risk

3.2 Portfolio Optimization

3.3 Optimal execution