From Markov Decision Processes to Reinforcement Learning with Python

Saúl Díaz Infante, David González Sánchez

2024-05-30

Table of contents

Preface

This notes are based in the course from Berstekas for the MIT see all lectures and other resources for complete the understanding.

Outline

The textbook for chapter one is Bertsekas' book (Bertsekas 2005). Chapters 2 and 3 are adapted from Sutton's book (Ch. 3, Ch. 4, Sutton and Barto 2018). For application and broad connection with more machine learning applications, we refer to (Brunton and Kutz 2019). Also, we recommend a handbook of algorithms (Szepesvári 2022). For applications with implemented code, we follow the books (Bilgin 2020).

1 The Dynamic Programming Algorithm

1.1 Introduction

1.2 The Basic Problem

1.3 The Dynamic Programming Algorithm

1.4 State Augmentation and Other Reformulations

1.5 Some Mathematical Issues

1.6 Dynamic Programming and Minimax Control

1.7 Notes, Sources, and Exercises

2 Dynamic Programming

2.1 Policy Evaluation (Prediction)

2.2 Policy Improvement

2.3 Policy Iteration

2.4 Value Iteration

2.5 Asynchronous Dynamic Programming

2.6 Generalized Policy Iteration

2.7 Efficiency of Dynamic Programming

2.8 Summary

3 Applications

3.1 Recycling Robot

3.2 A robot with randomly moves in a grid world.

4 Finite Markov Decision Processes

4.1 The Agent–Environment Interface

4.2 Goals and Rewards

4.3 Returns and Episodes

4.4 Unified Notation for Episodic and Continuing Tasks

4.5 Policies and Value Functions

4.6 Optimal Policies and Optimal Value Functions

4.7 Optimality and Approximation

4.8 Summary

5 Finite Markov Decision Processes

5.1 The Agent–Environment Interface

5.2 Goals and Rewards

5.3 Returns and Episodes

5.4 Unified Notation for Episodic and Continuing Tasks

5.5 Policies and Value Functions

5.6 Optimal Policies and Optimal Value Functions

5.7 Optimality and Approximation

5.8 Summary

6 Dynamic Programming

6.1 Policy Evaluation (Prediction)

6.2 Policy Improvement

6.3 Policy Iteration

6.4 Value Iteration

6.5 Asynchronous Dynamic Programming

6.6 Generalized Policy Iteration

6.7 Efficiency of Dynamic Programming

6.8 Summary

References

Bertsekas, Dimitri P. 2005. *Dynamic Programming and Optimal Control. Vol. I.* Third. Athena Scientific, Belmont, MA.

Bilgin, E. 2020. *Mastering Reinforcement Learning with Python: Build Next-Generation, Self-Learning Models Using Reinforcement Learning Techniques and Best Practices*. Packt Publishing. https://books.google.com.mx/books?id=s0MQEAAAQBAJ.

Brunton, Steven L., and J. Nathan Kutz. 2019. *Data-Driven Science and Engineering*. Cambridge University Press, Cambridge. https://doi.org/10.1017/9781108380690.

Sutton, Richard S., and Andrew G. Barto. 2018. *Reinforcement Learning: An Introduction*. Second. Adaptive Computation and Machine Learning. MIT Press, Cambridge, MA.

Szepesvári, Csaba. 2022. *Algorithms for Reinforcement Learning*. Vol. 9. Synthesis Lectures on Artificial Intelligence and Machine Learning. Springer, Cham. https://doi.org/10.1007/978-3-031-01551-9.