

強化学習を使った最適化手法をゼロから理解する！

Python による問題解決シリーズ第 3 巻

最適化のための 強化学習

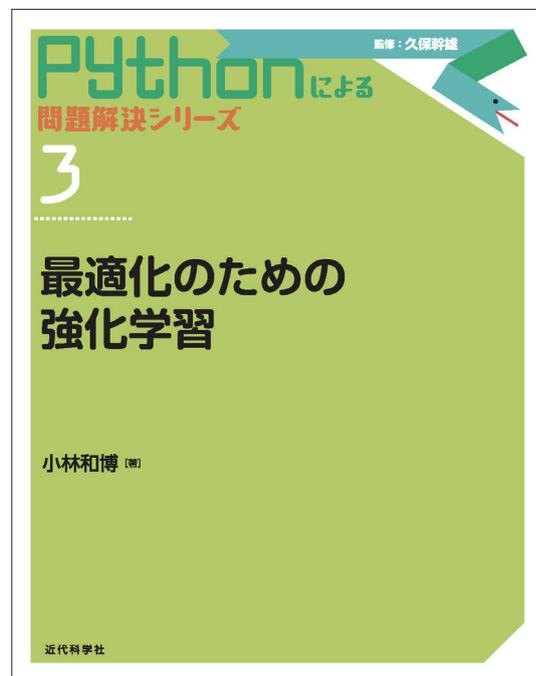
著者：小林 和博

仕様：B5 変形判・並製・204 頁

定価：3,200 円（税抜）

ISBN：978-4-7649-0710-2 C3304

発売：近代科学社



内容紹介

本書では、強化学習における基本的な考え方や計算手法を紹介し、実際に Python を用いて実現する方法をまとめている。具体的にはマルコフ決定過程、価値関数、方策評価、方策反復、価値反復、モンテカルロ評価、SARSA、Q 学習を扱う。ほとんどの内容は簡単な数学の知識があれば問題なく理解できるように記述し、読者の学びやすさを優先して繰り返しの説明や既出の数式を再掲するなど工夫している。Python の基礎から強化学習の利用までを詳細に解説した充実の一冊。

全国の書店・ネット書店にてお求めいただけます。お取り扱い店は以下のウェブページをご覧ください。

https://www.kindaikagaku.co.jp/book_list/detail/9784764907102/



お問い合わせ先

株式会社近代科学社

〒101-0051 東京都千代田区神田神保町 1-105

神保町三井ビルディング

電子メール：contact@kindaikagaku.co.jp

目次

第1章 Pythonで強化学習を行うための環境構築

- 1.1 オンラインサービスを利用する方法
- 1.2 手元のコンピュータに実行環境を整える方法
- 1.3 パッケージのインストール
- 1.4 実行環境

第2章 Pythonの基礎

- 2.1 データ構造
- 2.2 科学技術計算パッケージ NumPy
- 2.3 条件分岐
- 2.4 繰り返し処理
- 2.5 擬似乱数生成パッケージ random
- 2.6 可視化ライブラリ Matplotlib
- 2.7 関数
- 2.8 内包表記

第3章 強化学習の概要

第4章 マルコフ決定過程

- 4.1 マルコフ性
- 4.2 推移確率行列
- 4.3 マルコフ過程
- 4.4 マルコフ報酬過程
- 4.5 リターン
- 4.6 価値関数
- 4.7 方策
- 4.8 マルコフ決定過程

第5章 動的計画

- 5.1 例 1: 整数の和
- 5.2 例 2: 最短路問題
- 5.3 動的計画による価値関数の評価
- 5.4 方策評価
- 5.5 方策改善
- 5.6 方策反復
- 5.7 価値反復

第6章 モンテカルロ学習

- 6.1 全幅探索とサンプル探索
- 6.2 モンテカルロ方策評価
- 6.3 First-visit モンテカルロ方策評価
- 6.4 Every-visit モンテカルロ方策評価
- 6.5 平均の増分計算

第7章 Temporal Difference 学習

- 7.1 TD(0) 学習
- 7.2 オンポリシー学習とオフポリシー学習
- 7.3 オンポリシーモンテカルロ学習
- 7.4 オンポリシー TD 学習 — SARSA
- 7.5 オフポリシー TD 学習 — Q 学習

著者紹介

小林 和博 (こばやし かずひろ)

2003年3月 慶應義塾大学大学院文学研究科前期博士課程修了
1998年 東京大学工学部計数工学科卒業
2000年 東京大学大学院工学系研究科計数工学専攻修士課程修了, 修士(工学)
2009年 博士(理学)
現在青山学院大学理工学部准教授

主要著書

『サプライチェーンリスク管理と人道支援ロジスティクス』(共著), 近代科学社 (2015)
『航海応用力学の基礎』(共著), 成山堂書店 (2015)
『Python 言語によるビジネスアナリティクス—実務家のための最適化・統計解析・機械学習—』(共著), 近代科学社 (2016)
『最適化問題入門』(Python による問題解決シリーズ 2), 近代科学社 (2020)