

Lecture 0. Syllabus

Sim, Min Kyu, mksim@seoultech.ac.kr



서울과학기술대학교 데이터사이언스학과

1 I. Fundamental Information

2 II. Background

3 III. Topics

4 IV. Evaluation and others

I. Fundamental Information

Are we in the right place?

- 확률모형 및 시뮬레이션 (9490030-101)
- Monday 10am - 1pm
- Frontier Rm. 108 → **subject to change** ♥
- Zoom classroom
 - URL (No password required): ♥
<https://seoultech.zoom.us/j/6490221641?pwd=c0RHaXlkMXRLU1h0WkZ0TWp6RVNqZz09>
 - URL (Password required): <https://seoultech.zoom.us/j/6490221641> ♥
 - Password: 355785 ♥

2시간 - 비공개 ♥
1시간 - 공개 ♥

Professor

- Sim, Min Kyu. mksim@seoultech.ac.kr ✓
- Frontier Rm. 915
- In-person availability
 - Feel free to email me to setup a online/offline meeting. ✓
 - Generally available on Monday/Wednesday afternoon
- 주요 연구 분야
 - ① 확률과정론 (Stochastic Processes) ✓
 - ② 강화학습 응용 (Reinforcement learning application) ✓
 - ③ 신재생에너지 운영 (Renewable energy operation) ✓
 - ④ 금융데이터 모델링 및 기타 (Quantitative finance and others) ✓

[Stochastic]

This course will emphasize...

- 수리적 모형의 principle
- 수리적 모형의 application
- 강화학습의 수강의 토대

< 강화학습! >

II. Background

About

What is 확률과정 (stochastic process)?

- Stochastic이라는 말은 Time과 Random의 의미를 포함하고 있다.
 - 시간이 가면서 변화하는 것들에 관한... ♥
 - 시간이 가면서 확정적이지 않은 사건들에 관한.. ♥
 - 즉, 시간이 가면서 불확정적으로 변화하는 시스템을 다룬다.
- 예시
 - 날씨, 태양광 발전량, 풍력 발전량
 - 혈압, 체온
 - 공장의 생산라인의 모습
 -

What is simulation?

- ① 컴퓨터를 이용해 현실의 어떠한 현상을 가상으로 구현하고,
- ② 가상 환경에 대한 분석으로 insight을 얻고,
- ③ 이를 현실에 적용하는 것을 목표로 함
 - 본 강좌에서는 Stochastic process에 대한 확률적 시뮬레이션을 주로 다룸

인공지능과 강화학습

- 인공지능?
 - 기계가 시연하는 지능
 - 환경을 인식하고 목표를 성공적으로 달성 할 수 있는 기회를 극대화 하는 주체인 ‘지능형 에이전트’에 대한 연구
- 강화학습?
 - ① 환경 및 목표로 구성된 시스템을 설정하고
 - ② 지능형 에이전트가 시스템 내에서
 - ③ 기회를 최대화 하는 법을 학습하게 함
- 강화학습은 인공지능의 본질이라고 할 수 있다. 그 이유는
 - 다른 기계 학습의 영역에 비해
 - 인간이 지능을 사용하여 세상을 배워가는 방식과 닮아있다. ♥

Brief motivation for 강화학습

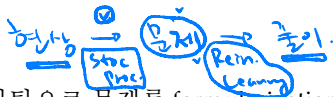
학습의 과정 - 우리는 어떻게 배우는가?

“지식과 편견이 부족한 어린 아이의 행동은 거침이 없습니다. 위험한지 모른 채 높은 곳에 올라가서 넘어지고, 아무 것이나 입에 넣어 씹니다. 그 결과, 때로는 다치기도 하고 엄마한테 혼나기도 하면서 세상에 대한 경험을 쌓아가면서 올바른 행동 방식을 찾아갑니다.”

“그렇게 성장한 아이가 성인이 되면 어떻게 되나요? 이제는 위험하지 않게 안전하고 효율적인 방식으로 행동합니다. 본인이 이해할 수 있는 세상속에서 가장 바람직한 행동을 선택하여 행동합니다. 새로운 시도를 통해서 배우는 빈도는 점점 더 줄어듭니다. 즉, 사람은 나이가 들면서 기존의 경험과 질서내에 적응하고 보수적으로 생각하고 살아가게 됩니다.”

왜 강화학습은 어렵다고 하는가?

- 주어진 문제를 푸는 과정 자체가 어렵다기 보다는
- 강화학습의 이론적인 배경이 되는 아래의 문제들의 내용 자체가 어려움
 - ✓ Sequential Decision Making
 - ✓ Stochastic Dynamic Programming
 - ✓ Markov Decision Process
- 주어진 문제를 푸는게 아니라 현상을 바탕으로 문제를 formalization 하는 부분이 어려움. 그 이유는
 - 현상의 특성과 에이전트의 목적에 맞게 Markov Decision Process로 잘 정의해야 함. (충분히 정확하게 모델링해야함)
 - 강화학습 에이전트가 풀 수 있는 문제의 규모로 formalization 해야함 (충분히 simple해야 문제가 풀림)
- 이러한 이슈 때문에 게임과 같은 상대적으로 closed system이 아닌 다양한 분야에 적용시키는데에는 많은 투자와 연구가 필요함
- 다양한 문제의 특성 만큼이나 다양한 모델링 기법등이 요구됨



왜 강화학습은 어렵다고 하는가?

- 수학적 내용 외에도 통계학, Computer Simulation, Deep Learning, Experiment Design 등에 대한 넓은 배경지식이 때때로 필요함
- 요구되는 선수과목 및 배경 지식의 수준
 - 수학적 엄밀함
 - 혼동없이 superscript, subscript를 사용하면서 의사소통 할 수 있어야 함
 - 행렬 연산
 - System of linear equation의 답을 찾는 법
 - 역행렬과 determinant의 의미
 - 확률통계학
 - 기초 통계학 수준의 추정에 대한 의미 (모평균등)
 - 분포의 의미와 기대값, 분산에 대한 의미
 - 답리닝
 - 반드시 필요한 것은 아니며, 기초적인 내용은 다루게 됨
 - R/Python
 - 두 개다 사용 경험이 있고, 하나는 편안하게 사용할 수 있는 수준이 필요함
- 어렵다고 느끼면 두 명 정도가 같이 복습하고 숙제하는 것을 권장함



$$Ax = b$$
$$x = A^{-1}b$$

개요

- Python, R, SQL 등의 프로그래밍 언어를 interactive 한 tutorial 형식으로 배울수 있는 사이트입니다.
- 아래의 invitation link로 들어가서 @seoultech.ac.kr 계정을 사용하면 무료로 가입이 됩니다.
- Invitation Link: https://www.datacamp.com/groups/shared_links/16bd199efae66e1aaf73c390fddeaed78e453d52f21f52480f8888e84d084a5
- 매년 3월과 9월에 계정이 만료됩니다. 그 때는 다시 가입을 하면 됩니다.
- 다음의 링크에 항상 초대장을 게시합니다:
<https://github.com/aceMKSIm/teaching/> ✓

추천 강좌

● Statistics

- [Statistical Thinking in Python (Part 1)]
<https://learn.datacamp.com/courses/statistical-thinking-in-python-part-1>
- [Statistical Thinking in Python (Part 2)]
<https://learn.datacamp.com/courses/statistical-thinking-in-python-part-2>

● R

- [Intro to R] <https://learn.datacamp.com/courses/free-introduction-to-r>
- [Intermediate R] <https://learn.datacamp.com/courses/intermediate-r>
- [Intro to tidyverse]
<https://learn.datacamp.com/courses/introduction-to-the-tidyverse>

● Python

- [Intro to Python]
<https://learn.datacamp.com/courses/intro-to-python-for-data-science>
- [Intermediate Python] <https://learn.datacamp.com/courses/intermediate-python>

✓ ● Deep Learning

- [Intro to Deep Learning]
<https://learn.datacamp.com/courses/introduction-to-deep-learning-with-keras>

강화학습을 어떻게 배울 것인가?

- 본 강좌는 강화학습 알고리즘을 이해하는 것을 넘어서
- ✓ ● 실제로 문제를 정의하고 알고리즘을 적용하여 문제를 풀 수 있는 능력을 갖추는 것을 목표로 함
- 따라서 강화학습의 이론적인 배경이 되는 Markov Decision Process (MDP) ✓
에 대해서 첫 번째 학기에서 다룸
- 두 번째 학기에서는 MDP에 대한 배경 지식을 바탕으로 본격적으로 강화학습 방법론을 적용함

강화학습은 왜 배우는가?

- 미래에 대한 조망을 바탕으로 한 현명한 의사결정은 지적 능력의 정수
- 최적의 판단은 부가가치가 매우 높은 영역
- 물고기 → 코끼리 → 인간 → 인공지능
- 다양한 영역의 매우 어려운 문제를 해결할 수 있는 잠재력을 가지고 있음
 - 분배와 할당의 문제
 - 도박 및 투자에 관한 결정의 문제
 - 맞춤 광고 및 두 개의 party가 interaction하는 상황에서의 문제
 - 자율 주행, 인공지능 제어 등의 많은 공학적 문제

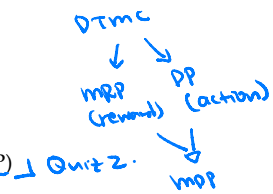
III. Topics

Topics to be covered

1st semester

- Part A. Review on quantitative methods
 - Math review
 - Prob/stat review
 - Simulation review
- Part B. Newsvendor model
 - One-period decision making under uncertainty
 - Optimization (always) involves trade-off
- Part C. Discrete-time Markov chain (DTMC)
 - Multi-period description on stochastic system
 - As a foundation of Markov decision process (MDP)
- Part D. Bridges between DTMC and MDP
 - Extending system description (DTMC) to decision making problem (MDP)
 - Dynamic programming (DP) and Markov reward process (MRP) to be covered
- Part E. MDP w/ model
 - Formulating MDP ✓
 - Approaches for finding an optimal policy ✓

↓ Quiz 1.



↓ Quiz 2.

2nd semester

- ✓ ● Part F. MDP w/o model
 - How to find an optimal policy when model is not present?
 - Core components of reinforcement learning
 - ✓ ● Part G. Functional approximation
 - Overcome curse of dimensionality by functional approximation
 - Deep neural network to be reviewed
 - Part H. Value based agent
 - Part I. Policy based agent
 - Part J. Recent advancements in DRL (Deep reinforcement learning)
- Deep Reinforcement Learning*

Material



① This Lecture Notes

- ② The following book is recommended for basic quantitative methods related to part A to part C in this course.

- Ross, S. M. (2014). Introduction to probability models. Academic press. ✓

- ③ The following book covers more on stochastic process up to MRP:

- Serfozo, R. (2009). Basics of applied stochastic processes. Springer Science & Business Media.

- ④ The following book is thin, but contains essence of dynamic programming:

- ✓ Bellman, R. (2003). Dynamic Programming (ISBN-13:978-0486428093)

- ⑤ The following book has a nice mathematical treatment with a lot of example on MDP:

- ✓ Ross, S. M. (1995). Introduction to Stochastic Dynamic Programming (ISBN-13:978-0125984218)

- ✓ ⑥ The following book shows good build-up from DTMC, MRP, and MRP. I highly recommend you to get the book.
- 노승은. (2020). 바닥부터 배우는 강화 학습.
- ⑦ Following book provides rather complete treatment on reinforcement learning. Maybe fairly readable after this course.
- ✓ ● Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction. MIT press.

IV. Evaluation and others

Evaluation

1 Quizzes

- Covers Part A. (10%) ✓
- Covers Part B & C. (20%) ✓

2 Final Exam (30%)

- Covers Part D & E. ✓

3 Homework (30%)

- Problem sets
- Case studies ✓

4 Discussion (10%)

- Class discussion
- Final case study presentation

당부의 말씀

- Youtube - 이론, Zoom, 오프라인 - 논의 및 복습의 형태로 구성되어 있습니다.
- 이 과목의 특성상 배우는 내용이 cumulative 하고, 난이도가 계속 올라갑니다.
- 숙제를 성실히 하지 않으면 내용을 따라가기가 어려워집니다.
- 이 과목은 소규모의 대학원 수업입니다. 적극적인 참여를 기대합니다.
- 끝까지 다 배우면 매우 재밌습니다!
- 강화학습은 교수의 현재 main research area이며, 저도 계속 배우고 있습니다. 본인 연구 주제와의 결합등의 연구 아이디어에 대해서 언제나 open되어 있습니다.

"A ship in harbor is safe, but that is not what ships are built for. - A.Einstein"

[1] "A ship in harbor is safe, but that is not what ships are built for. - A.Einstein"