Discipline: Data Science, Optimization, Machine Learning / Production and Logistics, Supply Chain Management, Finance, Marketing, Management Science, Operations, and Information Systems

1. Language

English

2. Title

Approximate Dynamic Programming for Stochastic and Dynamic Decision Making

3. Lecturer

Dirk Mattfeld, Marlin Ulmer

4. Date and Location

25. – 28.08.2025

Technische Universität Braunschweig, Rebenring 58, Room RR58.4, 38106 Braunschweig

5. Course Description

5.1 Abstract and Learning Objectives

The 4-day course deals with anticipatory methods for dynamic decision making. It will address the following questions:

1. What are the components of dynamic decision processes and how do they interact?
2. How can dynamic decision processes be modeled mathematically?
3. What methods exist in approximate dynamic programming?
4. How can they be applied to different types of problems?

5.2 Content

Companies operate in an increasingly volatile environment. Within their respective business models, decisions are often made under incomplete information. These decisions are further adapted based on newly revealed information over time, for example, when customers place new orders, when the production process gets disturbed, when congestion delays parcel delivery or when stock exchange rates promise future opportunities. The challenges of incomplete information and real-time adaptions of plans affect many business fields, for example, logistics, production, sales, or finance. Fortunately, companies gain access to vast amounts of data. To provide effective support for subsequent decision making, advanced mathematical models and methods are needed allowing for integration of data into real-time decision support systems.

The resulting problems are stochastic and dynamic decision processes and can be modeled as a Markov decision process (MDP). An MDP models a problem as a sequence of states connected by decisions and stochastic transitions. A solution of an MDP is a decision policy assigning a decision to every potential decision state. An optimal solution maximizes the expected reward. Small MDPs can be solved to optimality, for example, by means of dynamic programming. Dynamic programming uses recursion and operates on the Bellman Equation maximizing the sum of immediate and expected future reward in every state. However, the MDPs for the business models under consideration are generally complex, the accessible calculation time is limited, and an optimal solution is not possible due to the infamous Curses of Dimensionality.

Instead, methods of approximate dynamic programming (ADP) are applied. Simple ADP-methods straightforwardly mimic practical rules-of-thumb. Advanced ADP-methods approximate the second term of the Bellman Equation, the expected future reward, by means of simulation. Dependent on the type of method, the approximated reward is either discarded after the decision is made or stored for future use. In the latter case, the algorithm learns subsequently letting the results guide future simulations for a more accurate approximation. However, to store the results, aggregation is necessary.

In this course, we describe the process of approaching complex stochastic and dynamic decision problems with ADP-methods. We present the required steps from business problem over MDP to the ADP-solutions in detail and give an overview over the most prominent ADP-methods. We especially focus on offline learning methods known as value function approximation. The theoretical content of this course is accompanied by many illustrative examples from the field of logistics and by a serious gaming application.

This course spreads over 4 days. On the first day, complex business models are analyzed with respect to dynamism and stochasticity. Stochastic and dynamic decision making is introduced both conceptually and based on examples. On the second day, the MDP-modeling framework is defined and solving MDPs by means of dynamic programming is discussed, particularly, in context of the curses of dimensionality. Furthermore, modeling MDPs is practiced in a hands-on part. On the third day, a selection of advanced ADP-methods is presented. First, solution methods are introduced and classified. Then, the course focuses on simulation methods, namely, lookahead models and value function approximation as well as generic combinations of different ADP-methods. On the last day, the results are summarized, and the participants conceptually traverse the steps from problem to method in a case-study.

The course combines the fields of process modeling, optimization, simulation, and machine learning. It is aimed at PhD-students from various disciplines, including operations management, production and logistics, supply chain management, finance, marketing, and management science. The course fits well for PhD-students working in the field of data science. The elements of the course range from lectures on theory, to the discussion of best practices and case studies all accompanied by regular presentations of work done by the students.

5.3 Schedule (including start and end time)

	Day 1	Day 2	Day 3	Day 4
Morning (9am-12am)	Arrival Check In (starting at 10:30am) Preliminary discussions (optional)	Presentation Modeling • Dynamism • Stochasticity • Markov Decision Processes • Solutions	Presentation Methodology • Classification • Rolling Horizon • Policy Function Approximation • Lookahead Models	Presentation Combined Approaches Serious Gaming Summary
Afternoon (1pm-6pm)	Dynamic Decision Process • Overview • Examples • Case Study	Problems • Preparation Serious Gaming • Modeling Discussion • Case Study	Value Function Approximation (VFA) • Concept • Discussion • Case Study	Individual discussions (optional)

5.4 Course format

Lecture

The first day is devoted to a system perspective of dynamic decision making. The field is illustrated by an example. For this example, the interrelation of information, decision and fulfilment processes are depicted. Dynamic decision making can be embedded in the context of distributed decision making. This paradigm is sketched to define anticipation as a building block of dynamic decision making. Finally, students are requested to depict their individual research projects in terms of processes in a hands-on group work.

The second day starts with a presentation of the hands-on work by the students. The remainder of the day focuses on mathematical modeling. The concept of Markov decision processes is introduced and connected to the dynamic decision process introduced on the first day. The serious gaming application is introduced and embedded in the modeling framework. Finally, the student groups will mathematically model the dynamic decision processes derived on the first day.

The third day starts with the students presenting their modeling-results. Then, a classification of approximate dynamic programming methods is given. The individual methods are defined and illustrated with examples from the literature. Then, the concept of value function approximation is discussed in detail. Finally, the student groups derive suitable steps of value function approximation tailored to their problems.

On the fourth day, the students summarize their work. Then, the serious gaming application is used to illustrate the impact of different value function approximation choices and to evaluate the students learning process. Finally, the content of the course is summarized.

6. Preparation and Literature

6.1 Prerequisites

The course requires basic skills in process modeling, mathematics, statistics, and classical optimization methods on a master's level. Computer programming experience is not necessary but can be helpful.

6.2 Essential Reading Material

Participants are required to read some overview literature as part of their preparation for the course.

1. Powell, W. B. (2009). What you should know about approximate dynamic programming. Naval Research Logistics (NRL), 56(3), 239-249.
2. Soeffker, Ni, Ulmer M. W., and Mattfeld, D.C (2022). Stochastic dynamic vehicle routing in the light of prescriptive analytics: A review." European Journal of Operational Research, 3(1), 801-820.
3. Ulmer, M. W., Goodson, J. C., Mattfeld, D. C., & Thomas, B. W. (2020). On modeling stochastic dynamic vehicle routing problems. *EURO Journal on Transportation and Logistics*, *9*(2), 100008

6.3 Additional Reading Material

The potential reading material is vast. In the following, we give a few examples. We will provide additional reading material during the course. We are also happy to provide additional reading material for a specific field. Please just request the material per email.

Books:

- Powell, W. B. (2022). Reinforcement learning and stochastic optimization: A unified framework for sequential decisions. Hoboken: John Wiley & Sons, Inc.
- Ulmer, M. W. (2017). Approximate Dynamic Programming for Dynamic Vehicle Routing, Springer.

Tutorials:

- Mes, M. R., & Rivera, A. E. P. (2017). Approximate Dynamic Programming by Practical Examples. Markov Decision Processes in Practice, 63-101. Springer.
- Powell, W. B., & Meisel, S. (2016). Tutorial on stochastic optimization in energy—Part I: Modeling and policies. IEEE Transactions on Power Systems, 31(2), 1459-1467.

Miscellaneous:

- Meisel, S., & Mattfeld, D. (2010). Synergies of operations research and data mining. European Journal of Operational Research, 206(1), 1-10.
- Mes, M.R., & van Heeswijk, Wouter (2020). Comparison of Manual and Automated Decision-Making with a Logistics Serious Game, E. Lalla-Ruiz et al. (Eds.): ICCL 2020, LNCS 12433, pp. 698–714
- Mortenson, M. J., Doherty, N. F., & Robinson, S. (2015). Operational research from Taylorism to Terabytes: A research agenda for the analytics age. European Journal of Operational Research, 241(3), 583-595.

6.4 To prepare

Students are required to read the mandatory literature.

3. In the beginning of the course, each participant will briefly present their research topic. Because of the large group, the presentations will be in PechaKucha-style. Please prepare 12 slides that motivate and describe your problem, highlight the stochasticity and dynamism, and if applicable, discuss the idea of the method and results. Slides should not have any text. The presentation should transition to the next slide automatically after 20 seconds; thus, the entire presentation will be 4 minutes long.

7. Administration

7.1 Max. number of participants

20

7.2 Assignments

The students will work in groups on selected problems at the end of every day. They will present their results at the beginning of the following day.

Every student will write a succinct summary of the developed models and methods for the individual case studies (about 8 pages).

7.3 Exam

The final grade will be based on class participation during the lectures and case studies (50%) and the quality of the summary to be submitted after the course (50%).

7.4 Credits

The course corresponds to a scope of 6 LP/ECTS

8. Working Hours

Working Hours	**Stunden**
Preparation I: Literature	30h
Preparation II: Research Definition	20h
Preparation II: Presentation	10h
Active Participation:	40h
Assignments:	10h
Writing Summary:	30h
Comprehensive recap of course content:	40h
SUMME	**180 h**