Data-Driven Decision-Making
In Enterprise Applications

Introduction

Rainer Schlosser, Martin Boissier, Matthias Uflacker

Hasso Plattner Institute (EPIC)

April 18, 2019

The World is Full of Decision Problems

What Constitutes a Decision Problem?

Decisions

Objectives

Constraints

How to Approach Decision Problems?

Decisions x	When can I do what?	Identify.
Objective $F(x)$	What do I want to optimize?	Define.
Constraints $C(x)$	What has to be satisfied?	Determine.

How to Approach Decision Problems?

Decisions x	When can I do what?	Identify.
Impact of x	What happens if a certain decision is made?	Estimate.
Objective $F(x)$	What do I want to optimize?	Define.
Constraints $C(x)$	What has to be satisfied?	Determine.
Optimization	Max $F(x)$ over x such that $C(x)$ is satisfied.	Solve!

Agenda

- Introduction ✓
- **Personal Background**
- Goals of the Course & Grading
- Examples: Decision Problems in Data-Driven Applications

Personal Background

- Ph.D. Operations Research (2014), Humboldt-University of Berlin
- Hasso Plattner Institute, EPIC, since 2015
- Field of Research
  – Data-driven decision support
  – Focus on stochastic dynamic models
- Current Areas of Applications
  – Operations management (e.g., dynamic pricing, ordering, advertising)
  – Database configuration (e.g., data placement problems, index selection)

Agenda

- Introduction ✓
- Personal background ✓
- **Goals of the Course & Grading**
- Examples: Decision Problems in Data-Driven Applications

Technical Information

- Credits? 4 SWS (V/Ü), 6 ECTS (graded)
- When? Monday 13.30 - 15.00 / Thursday 11.00 – 12.30
  Start: April 18, 2019, End: July 11, 2019
- Where? Room D-E. 9/10
- Who? Rainer Schlosser, rainer.schlosser@hpi.de
  Martin Boissier, martin.boissier@hpi.de
- Slides? EPIC, Teaching, Summer 2019

Structure of the Course

- April/May: Lectures on „Optimization Techniques“:
  - (i) Linear Programming
  - (ii) Integer Linear Programming
  - (iii) Linear/Logistic Regression
  - (iv) Dynamic Programming
  - (v) Robust Optimization
- June/July: Choose Projects, Apply/Extend Suitable Techniques, Work in Teams, Input/Support, Presentations
- Aug/Sep: Documentation of Projects Results

Overview

2	April 25	Linear Programming
3	April 29	Integer Linear Programming
4	May 2	Linear/Logistic Regression
5	May 6	Exercise Implementations
6	May 16	Dynamic Programming I
7	May 20	Dynamic Programming II
8	May 23	Response Strategies / Game Theory
9	May 27	Project Assignments
10	June 3	Robust Optimization
11	June 13	Workshop / Group Meetings
12	June 20	Presentations (First Results)
13/14	June 24/27	Workshop / Group Meetings
15/16	July 1/4	Workshop / Group Meetings
17	July 11	Presentations (Final Results), Feedback, Documentation (Aug 31)

Goals of the Course & Grading

- Goal: Develop models to compute optimized decisions for data-driven applications
- Learn: Optimization techniques
- Do: Apply & extend different optimization approaches
- Grading:
  - 10% Regular attendance / Personal engagement
  - 20% Results / Homework
  - 30% Presentations
  - 40% Documentation / Paper (End of semester)

Prerequisites

- Programming

  Parameters, Data Preparation
  Loops, Recursions, Simulations

- Basic Mathematical Background

  Sets, Vectors
  Probabilities, Random Variables, Expected Values

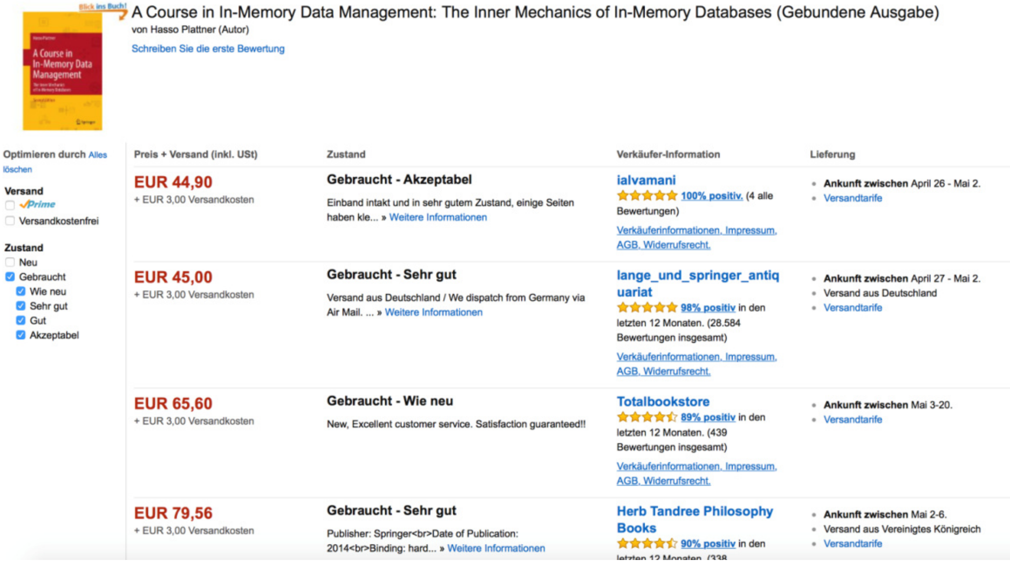
- More does not harm

  Regression Analysis
  Experience with Solvers
  Game Theory

Agenda

- Introduction ✓
- Personal Background ✓
- Goals of the Course & Grading ✓
- **Examples: Decision Problems in Data-Driven Applications**

Problem Example 1 – Dynamic Pricing



A Course in In-Memory Data Management: The Inner Mechanics of In-Memory Databases (Gebundene Ausgabe)
von Hasso Plattner (Autor)
Schreiben Sie die erste Bewertung

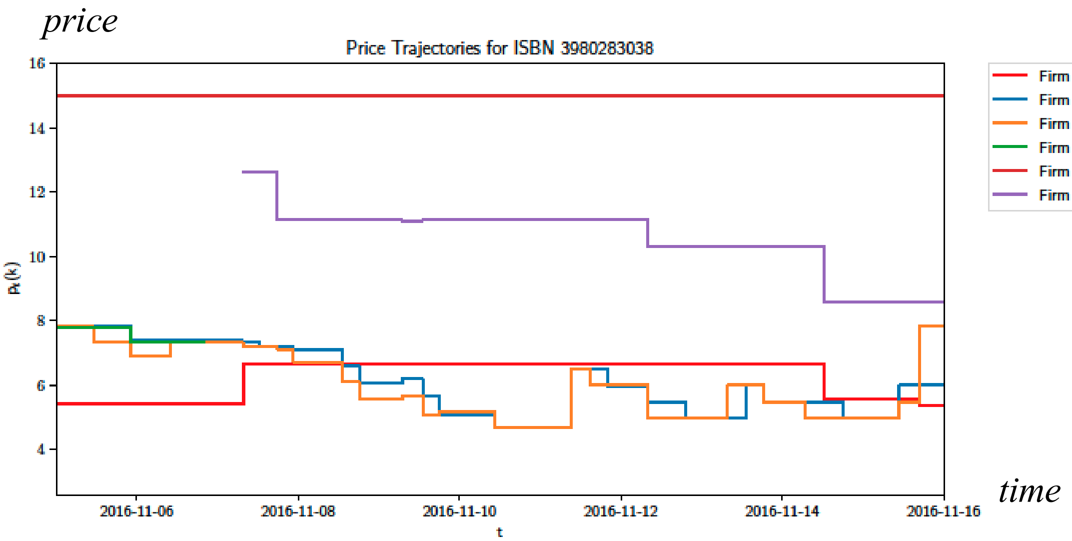
Optimieren durch Alles löschen
Versand
Prime
Versandkostenfrei
Zustand
Neu
Gebraucht
Wie neu
Sehr gut
Gut
Akzeptabel

Preis + Versand (inkl. USt)	Zustand	Verkäufer-Information	Lieferung
EUR 44,90 + EUR 3,00 Versandkosten	Gebraucht - Akzeptabel Einband intakt und in sehr gutem Zustand, einige Seiten haben kle... » Weitere Informationen	ialvamani 100% positiv. (4 alle Bewertungen) Verkäuferinformationen, Impressum, AGB, Widerrufsrecht.	Ankunft zwischen April 26 - Mai 2. Versandtarife
EUR 45,00 + EUR 3,00 Versandkosten	Gebraucht - Sehr gut Versand aus Deutschland / We dispatch from Germany via Air Mail. ... » Weitere Informationen	lange_und_springer_antiquariat 98% positiv in den letzten 12 Monaten. (28.584 Bewertungen insgesamt) Verkäuferinformationen, Impressum, AGB, Widerrufsrecht.	Ankunft zwischen April 27 - Mai 2. Versand aus Deutschland Versandtarife
EUR 65,60 + EUR 3,00 Versandkosten	Gebraucht - Wie neu New, Excellent customer service. Satisfaction guaranteed!!	Totalbookstore 89% positiv in den letzten 12 Monaten. (439 Bewertungen insgesamt) Verkäuferinformationen, Impressum, AGB, Widerrufsrecht.	Ankunft zwischen Mai 3-20. Versandtarife
EUR 79,56 + EUR 3,00 Versandkosten	Gebraucht - Sehr gut Publisher: Springer Date of Publication: 2014 Binding: hard... » Weitere Informationen	Herb Tandree Philosophy Books 90% positiv in den	Ankunft zwischen Mai 2-6. Versand aus Vereinigtes Königreich Versandtarife

How can we assist an e-commerce merchant in optimizing his/her prices?

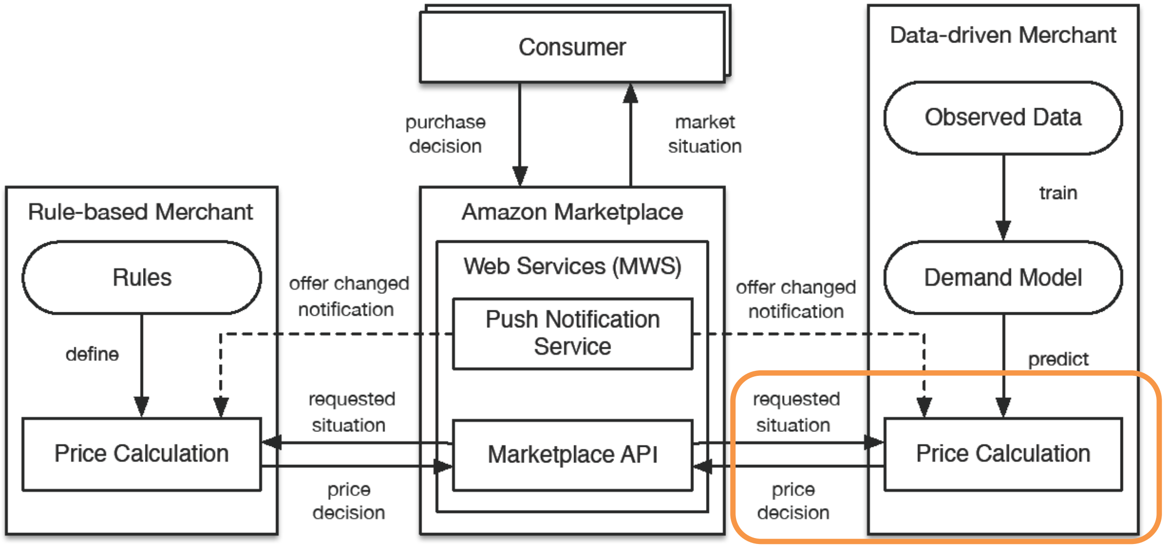
Impact of Price Decisions and Changing Markets

Characteristics:
- Exits & entries of competitors
- Active and passive competitors
- Price cycles



Pricing Options: Price Updates on Amazon

- Price update process on Amazon: (i) request a market situation (ii) optimize price based on demand model, (iii) send price update



Estimation of Price Impacts and Optimization

(1) Estimation of Sales Probabilities

- ca. 10 market situations/day/item with 1-20 firms (100 Mio obs.)
- ca. 2 000 sales/month (1 year of data)
- Predict sales probabilities (for time intervals and market situations)

(2) Price *Optimization*

Estimation of Price Impacts and Optimization

(1) Estimation of Sales Probabilities

- ca. 10 market situations/day/item with 1-20 firms (100 Mio obs.)
- ca. 2 000 sales/month (1 year of data)
- Predict sales probabilities (for time intervals and market situations)

(2) Price *Optimization*

- Maximize expected discounted long-term profit
- **Dynamic programming**

Estimation of Price Impacts and Optimization

(1) Estimation of Sales Probabilities

- ca. 10 market situations/day/item with 1-20 firms (100 Mio obs.)
- ca. 2 000 sales/month (1 year of data)
- Predict sales probabilities (for time intervals and market situations)

(2) Price *Optimization*

$$\max E(G_t \,|\, X_t = n, \vec{S}_t = \vec{s}_t), \qquad G_t := \sum_{s=t}^{T-1} \delta^{s-t} \cdot \Big(\big(a(X_s, \vec{S}_s) - c \big) \cdot \big(X_s - X_{s+1} \big) - l \cdot X_s \Big) \qquad (1)$$

$$a(n, \vec{s}) = \underset{a \in A}{\arg\max} \left\{ \sum_{i \geq 0} \tilde{P}(i, a \,|\, \vec{s}) \cdot \Big((a-c) \cdot \min(n,i) - n \cdot l + \delta \cdot V\big((n-i)^+, \vec{s} \big) \Big) \right\} \qquad (2)$$

$$V(n, \vec{s}) = \max_{a \in A} \left\{ \sum_{i > 0} \tilde{P}(i, a \,|\, \vec{s}) \cdot \begin{pmatrix} (a-c) \cdot \min(n,i) - n \cdot l \\ -z \cdot \delta \cdot V\big((n-i)^+, \vec{s} \big) \end{pmatrix} \Big/ \Big(1 - \tilde{P}(0, a \,|\, \vec{s}) \cdot z \cdot \delta \Big) \right\} \qquad (3)$$

Comparison of Performance Results

Comparison: Our *data-driven* strategy vs. the seller's *rule-based* strategy

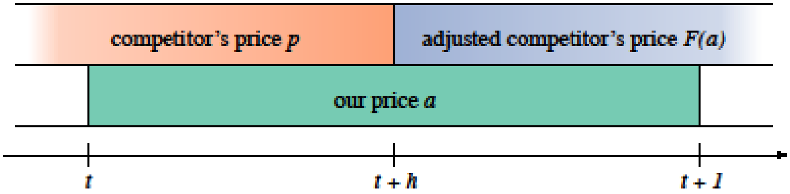
Strategy	#Books	% Sold (3 months)		Profit per sale (EUR)		Acc. profit
Rule-Based	5,534	42 %	100.0 %	2.56 €	100.0 %	100.0 %
HPI1 (high prices)	5,206	29 %	–30 %	3.58 €	**+40 %**	–1.5 %
HPI2	5,407	37 %	–12 %	3.03 €	+19 %	+4.3 %
HPI3	**5,241**	**44 %**	**+7 %**	**2.94 €**	**+15 %**	**+23.1 %**
HPI4 (low prices)	5,200	45 %	**+8 %**	2.52 €	–1 %	+6.4 %

Publications: Computers and Operations Research (2018)

KDD 2018

Optimal Response Strategies in Duopoly Settings

Question: How do *optimal* price adjustment strategies look like?



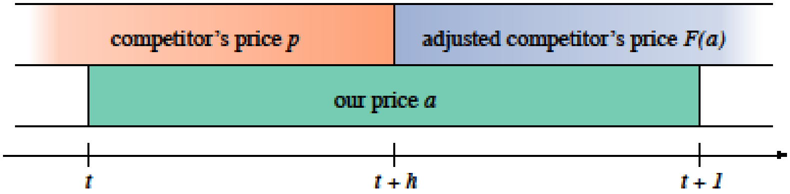
Setting: Infinite horizon, competitor's response strategy *is known*

Results:

against $F(a) := \max(a-1, 1)$

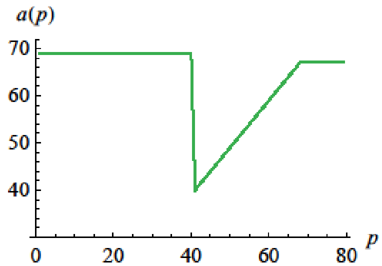
Optimal Response Strategies in Duopoly Settings

Question: How do *optimal* price adjustment strategies look like?



Setting: Infinite horizon, competitor's response strategy *is known*

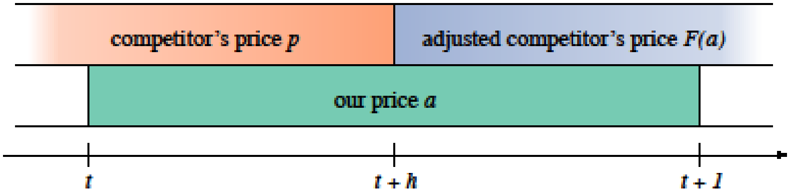
Results:



against $F(a) := \max(a-1,1)$ *mutual optimal* (*equilibrium*)

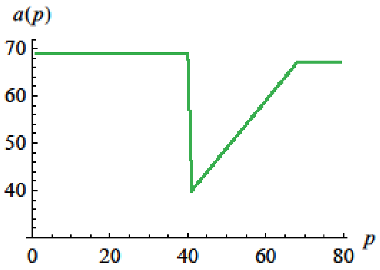
Optimal Response Strategies in Duopoly Settings

Question: How do *optimal* price adjustment strategies look like?

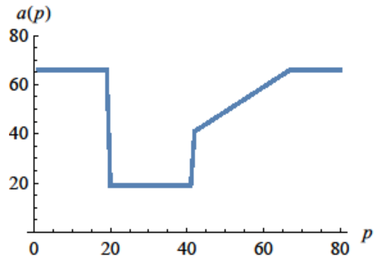


Setting: Infinite horizon, competitor's response strategy *is known*

Results:



$against\ F(a) := \max(a-1, 1)$

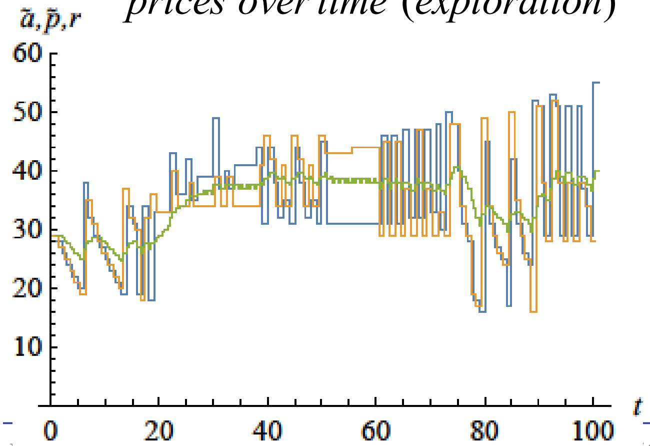


mutual optimal (*equilibrium*)

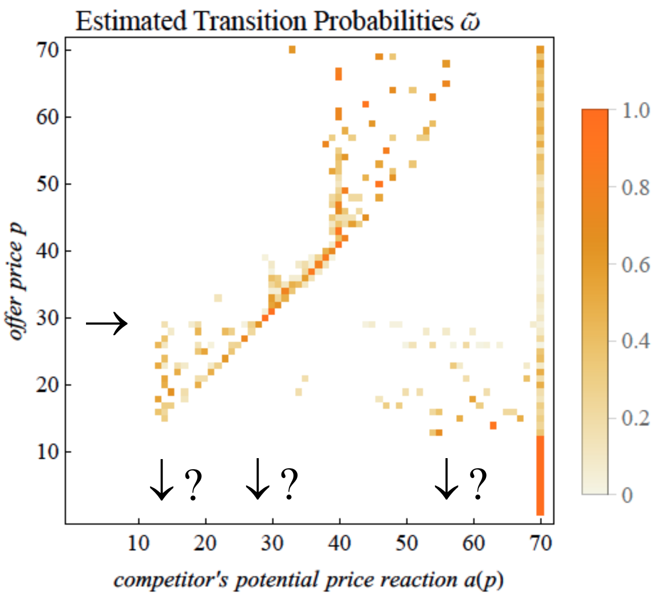
Interaction of Self-Adapting Strategies (Short-Term)

- Now, price responses *have to be learned*!
- Both players update their strategies
- Do equilibria exist?

prices over time (exploration)



anticipated price reactions



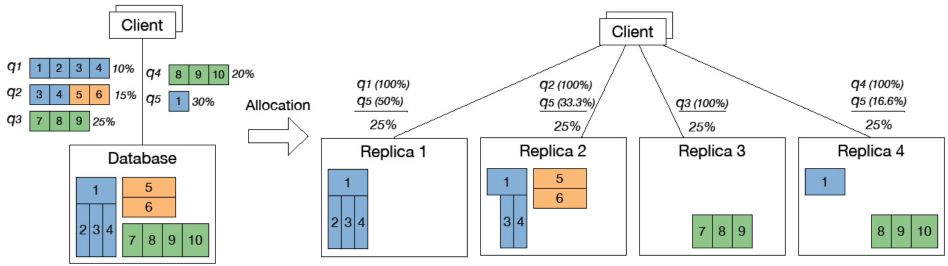
Further Decision Problems

Revenue Management (Dynamic Programming)

- Inventory Management
- Advertising

Database Configuration (Linear Programming)

- Database Replication
- Data Tiering
- . . .



Overview

2	**April 25**	**Linear Programming**
3	April 29	Integer Linear Programming
4	May 2	Linear/Logistic Regression
5	May 6	Exercise Implementations
6	May 16	Dynamic Programming I
7	May 20	Dynamic Programming II
8	May 23	Response Strategies / Game Theory
9	May 27	Project Assignments
10	June 3	Robust Optimization
11	June 13	Workshop / Group Meetings
12	June 20	Presentations (First Results)
13/14	June 24/27	Workshop / Group Meetings
15/16	July 1/4	Workshop / Group Meetings
17	July 11	Presentations (Final Results), Feedback, Documentation (Aug 31)