Chen Tang's Knowledge Database

ChenTang@link.cuhk.edu.cn

Last updated on August 31, 2023

Preface

The following is a compendium of my academic notes spanning various domains. I present these notes publicly to share my methodological framework for managing and structuring an individual's knowledge networks.

The inevitability of encountering occasional errors is acknowledged.

This notebook will undergo continuous updates.

Contents

Chapter 1

Mathematics and Optimization

1.1 Calculus and Linear Algebra

1.2 Analysis and Algebra

1.3 Probability Theory

1.4 Stochastic Process

1.5 Linear Programming

1.6 Convex Optimization

1.7 Non-Convex Optimization

Chapter 2

Statistics and Econometrics

2.1 Statistics Theory

2.2 Causal Model

2.2.1 Rubin's Causal Model

2.3 Reduced-Form Identification

2.4 Advanced Econometrics

2.5 Machine Learning Interface

Chapter 3

Economics

3.1 Microeconomics

3.2 Macroeconomics

3.3 Game Theory

3.4 Development Economics

3.5 Data Science Interface

Chapter 4

Computer Science and Data Science

4.1 Machine Learning

4.2 Deep Learning

4.3 Reinforcement Learning

Chapter 5

Information Systems and Operations Management

5.1 Empirical Operations Management

5.2 Revenue Management

Revenue management is a data-driven system to price perishable assets tactically at the micro-market level to maximize expected revenue or profit. Some important reviews before 2009 can be found in the book [Gallego and Topaloglu, 2019]. There are some extra summary papers like [Strauss et al., 2018], [Klein et al., 2020].

5.2.1 Traditional RM

Keywords 5.1

- Protection Level, Booking Limit, Littlewood's Rule

Assumption 5.2.1.1: What does **"traditional"** means in RM?

1. The traditional RM system doesn't consider the choice model, in particular, it assumes the demands are independent random variables;
2. Further assumption: consumer will leave without purchasing if preferred fare class is unavailable (holds when gaps in fares are large enough);
3. The capacity is fixed, the capacity's marginal profit is zero(can be relaxed);
4. All booked consumers would arrive (another circumstance see 5.2.2).

Assumption 5.2.1.2: Single Resource RM

1. The units of capacity is c, pricing at multiple different level $p_n < \cdots < p_1$;
2. Low-before-high fare class arrival order: D_2 before D_1 for example (this is the worst case for revenue);
3. **Protection level** for customer j: leave $y \in \{0, 1, \cdots, c\}$ for $D_{j-1}, , D_1$; $c - y$ is the **booking limit** which serves D_j;

😎 So the problem is to solve the optimal protection level given the current consumer level j.

Let $V_j(x)$ be the optimal revenue given D_j coming in, x units remained. $V_0(x) = 0$ by design. Let y be the protection level for $D_{j-1}, \cdots, D_1$: sales at $p_j = \min\{x - y, D_j\}$. The remaining capacity for $D_{j-1}, \cdots, D_1$ is $x - \min\{x - y, D_j\} = max\{y, x - D_j\}$. Now let $W_j(y, x)$ be the optimal solution. We have:

$$W_j(y,x) = p_j\mathbb{E}\{\min\{x-y, D_j\}\} + \mathbb{E}\{V_{j-1}(\max\{y, x-D_j\})\} \tag{5.2.1.1}$$

$$V_j(x) = \max_{y\in\{0,\dots,x\}} W_j(y,x) = \max_{y\in\{0,\dots,x\}} \left\{p_j\mathbb{E}\{\min\{x-y, D_j\}\} + \mathbb{E}\{V_{j-1}(\max\{y, x-D_j\})\}\right\} \tag{5.2.1.2}$$

Proposition 5.2.1.1: Structure of the Optimal Policy

$$y^*_{j-1} = \max\{y \in \mathbb{N}_+ : \Delta V_{j-1}(y) > p_j\}. \tag{5.2.1.3}$$

The maximizer of $W_j(y,x)$ is given by $y^*_j, , y^*_1$

Remark 5.2.1

The optimal solution for $y_j$is independent of the distribution of D_j;

Corollary 5.2.1.1: When $j = 2$:

Theorem 5.2.1.1: Littlewood's rule

$$y^*_1 = \max\{y \in \mathbb{N}_+ : \mathbb{P}\{D_1 \geq y\} > r\} \tag{5.2.1.4}$$

;

Remark 5.2.2

The Littlewoood's Rule:

1. The solution depends on the **fare ratio**: $r := p_2/p_1$;
2. When the distribution of D_2 is continuous: $F_1(y) = \mathbb{P}\{D_1 \leq y\}$. The optimal protection level is $y^*_1 = F_1^{-1}(1-r) = \mu_1 + \sigma_1 \notin^{-1} (1-r)$:
   (a) if $r > \frac{1}{2}$, $y^*_1 < \mu_1$ and y^*_1 decreases with σ_1;
   (b) if $r < \frac{1}{2}$, $y^*_1 < \mu_1$ and y^*_1 increases with σ_1;
   (c) if $r = \frac{1}{2}$, $y^*_1 = \mu_1$;
3. Using Littlewood's rule would result some D_1 served by competitors (high spill rates). Solution: add penalty to save more seats for the high fare consumers:

$$y^*_1 = \max\left\{y \in \mathbb{N}_+ : \mathbb{R}\{D_1 \geq y\} > \frac{p_2}{p_1+\rho}\right\} \tag{5.2.1.5}$$

5.2.2 Overbooking

5.2.3 Traditional Consumer Choice Model

5.2.4 Current Cosumer Choice Model

5.3 Platform Operations Management

5.3.1 Consumer Polarization

Consumer polarization is a topic in consumer research. **Group-Polarization Hypothesis** suggests that group discussion generally produces attitudes that are more extreme in the direction of the average of prediscussion attitudes in a variety of situations. Works like [Rao and Steckel, 1991] provides a mathematical presentation for this phenomenon (in the domain of preference):

$$U_{\mathrm{s}} = \sum_{i=1}^{m} \lambda_i u_i + \phi(\bar{u} - K) 0 \leq \lambda_i \leq 1, \sum_{i=1}^{m} \lambda_i = 1, \phi \geq 0 \tag{5.3.1.1}$$

In this model, the $\bar{u}$ is the algebraic mean of all consumers' utility, and K is the **Pivot Point**. Rewrite this formula:

$$\begin{aligned} U_{\mathrm{g}} &= \sum_{i=1}^{m} \left(\lambda_i + \frac{\phi}{m}\right) u_i - \phi K \\ &= w_0 + w_1 u_1 + w_2 u_2 + \cdots + w_m u_m \end{aligned} \qquad \begin{aligned} & w_0 \leq 0; \\ & \sum_{i=1}^{m} w_i \geq 1; \\ & 0 \leq \frac{w_0}{1 - \sum w_i} \leq 1. \end{aligned} \tag{5.3.1.2}$$

Remark 5.3.1

- [Zhao et al., 2023] use experiment results to suggest that eWOM (electronic word of mouth) polarization (the degree of eWOM to which positive and negative sentiments are simultaneously strong) would decrease the consumers' intention to purchase, mediating by the enhance of attitude ambivalence. (Ambivalence is a psychological state where a person endorses both positive and negative attitudinal positions)
- [Iyer and Yoganarasimhan, 2021] use game theory framework, get the conclusion that sequential decision making could reduce the polarization.

5.3.2 Network Effect

Network Effect and **Network Externality**:

[Narayan et al., 2011] verifies that the peer influence affects attribute preferences via a bayesian updating mechanism. In their model, the utility is given as follow:

$$U_{ijp}^{R} = X_{jp} \beta_i^R + \lambda_i \varepsilon_{ijp}^{R} \tag{5.3.2.1}$$

Where U_{ijp} is the utility of consumer i for product j given choice set p, X_{jp} is the attribute of product j in choice set p, β_i^R is the customers' weights. The bayesian updating process is given below:

$$\beta_{ik}^{R} = \rho_{ik}\beta_{ik}^{l} + (1-\rho_{ik})\frac{\sum_{i=1,i\neq i}^{N} w_{ii}\beta_{ik}^{l}}{\max\left[\left(\sum_{i=1,i\neq i}^{N} w_{ii}\right), 1\right]}, \tag{5.3.2.2}$$

$$\text{where } 0 \leq \rho_{ik} \leq 1.$$

Other research on peer-influence:

Remark 5.3.2

The consumers' interaction and social connections have a proposition proposed in [Zhang et al., 2017] for their goal attainment and spending: a positive linear term plus a negative squared term;

5.3.3 Online Gaming

Many industrial news about online gaming can be found in [Chen et al., 2017]. In [Lei, 2022], the dissertation fully discussed loot box pricing, matchmaking and price discrimination with fairness constraints.

Play-Duration and Spending

[Zhang et al., 2017]'s work show that there is a nonlinear effect of social connections and interactions on consumers' goal attainment and spending: A positive linear terms and the negative squared term. Mechanism: functional in providing useful information or tips that can facilitate goal attainment, but would rise information overload problem.

Player engagement can be embodied by many specific metrics, such as time or money spent in the game, the number of matches played within a time window, or churn risk. [Chen et al., 2017] define churn risk as the proportion of total players stopping playing the game over a period of time.

Matchmaking

Matchmaking connects multiple players to participate in online PvP games. (PvP(Player-versus-Player) games, which cover many popular genres, such as multiplayer online battle arena (MOBA), first-person shooting (FPS), and e-Sports, have increased worldwide popularity in recent years.)

The past matchmaking strategy matches similar skilled players in the same round (SBMM), the current MM system focus on impoving the players' engagement and decreasing the churn rate. Foe example, in [Chen et al., 2017] EOMM (Engagement Optimization MatchMaking) is proposed to minimize the churn rate.

[Chen et al., 2021] propose an algorithm to maximize the cummulative active players.

Assumption 5.3.3.1: Chen 2021 MatchMaking

1. players can have heterogeneous skill levels: level 1 to level K;
2. the outcome of each match is a Bernoulli random varaible: $p_{kj} = 1 - p_{jk}, p_{kk} = 0.5$, $p_{kj} > 0.5$ if $k > j$;
3. player's skill level is fixed: *relative* level;
4. and their state depends on the win-loss outcomes of the past m matches: $g \in \mathcal{G}$ ($2^m + 1$ possible cardinality);
5. A geometric losing churn model: palyers churn with a fixed probability, starting from the second loss in a row;
6. $P^k_{win}, P^k_{lose} \in [0, 1]^{|\mathcal{G}|\times|\mathcal{G}|}$ is the transition matrix of level k player's engagement state;
7. $M_{kj} = p_{kj}P^k_{win} + (1 - p_{kj})P^k_{lose}$ is the aggregate transition matrix. ($\bar{G}$ is the reduced aggregate transition matrix);
8. using the fluid mathcing model and assume players are infinitely divisible;

The **Dynamic Programming** formulation: $f_{kg,jg'}$ is the amount of kg players matched with jg' players, s^t_{kg} is the number of kg players at time t.
FB *flow balance* constraints:

$$\begin{aligned}
\sum_{j=1}^{K}\sum_{g'\in\bar{\mathcal{G}}} f^t_{kg,jg'} &= s^t_{kg}, k = 1, \dots, K, \forall g \in \bar{\mathcal{G}}, \\
\sum_{j=1}^{K}\sum_{g'\in\bar{\mathcal{G}}} f^t_{jg',kg} &= s^t_{kg}, k = 1, \dots, K, \forall g \in \bar{\mathcal{G}}, \\
f^t_{kg,jg'} &= f^t_{jg',kg}, j = 1, \dots, K, k = 1, \dots, K, \forall g \in \bar{\mathcal{G}}, g' \in \bar{\mathcal{G}} \\
f^t_{kg,jg'} &\geq 0, j = 1, \dots, K, k = 1, \dots, K, \forall g \in \bar{\mathcal{G}}, g' \in \bar{\mathcal{G}}
\end{aligned} \tag{5.3.3.1}$$

ED *evolution of demographics*:

$$\mathbf{s}^{t+1}_k = \sum_{j=1,\dots,K} \left(\mathbf{f}^t_{kj}\mathbf{1}\right)^\top \left(\bar{M}_{kj} + N_k\right) k = 1, \dots, K \tag{5.3.3.2}$$

The value-to-go function is:

$$V^{\pi}(\mathbf{s}^t) = \sum_{k=1}^{K} \sum_{g \in \bar{\mathcal{G}}} s_{kg}^{t+1} + \gamma V^{\pi}(\mathbf{s}^{t+1}) \quad \text{subject to (FB), (ED).} \tag{5.3.3.3}$$

The above model can be formulated in a linear progarmming style:

Theorem 5.3.3.1: Chen 2021 MM LP Formulation

$$\begin{aligned}
V^*(\mathbf{s}^0) = \max & \sum_{t=1}^{\infty} \gamma^{t-1} \sum_{k} \sum_{g \in \bar{\mathcal{G}}} s_{kg}^t \\
\text{s.t.} & \sum_{j=1}^{K} \sum_{g' \in \bar{\mathcal{G}}} f_{kg,jg'}^t = s_{kg}^t, \forall k, \forall g \in \bar{\mathcal{G}}, t = 0, 1, \ldots \\
& \sum_{j=1}^{K} \sum_{g' \in \bar{\mathcal{G}}} f_{jg',kg}^t = s_{kg}^t, \forall k, \forall g \in \bar{\mathcal{G}}, t = 0, 1, \ldots \\
& f_{kg,jg'}^t = f_{jg',kg}^t, j = 1, \ldots, K, k = 1, \ldots, K, \forall g \in \bar{\mathcal{G}}, g' \in \bar{\mathcal{G}}, t = 0, 1, \ldots \\
& f_{kg,jg'}^t \geq 0, j = 1, \ldots, K, k = 1, \ldots, K, \forall g \in \bar{\mathcal{G}}, g' \in \bar{\mathcal{G}}, t = 0, 1, \ldots \\
& \mathbf{s}_k^{t+1} = \sum_{j=1,\ldots,K} \left(\mathbf{f}_{kj}^t \mathbf{1}\right)^\top \left(\bar{M}_{kj} + N_k\right), \forall k, t = 0, 1, \ldots
\end{aligned} \tag{5.3.3.4}$$

Remark 5.3.3

- Using an optimal matchmaking policy instead of SBMM may reduce the required bot ratio significantly while maintaining the same level of engagement.

5.4 Behavioral Operations Management

5.5 Data-Driven Operations Management

Chapter 6

Miscellaneous

6.1 Notes on Tools

6.1.1 LaTeX Shortcuts

There are 6x6 colors in the preset preamble:

aa	ab	ac	ad	ae	af
ba	bb	bc	bd	be	bf
ca	cb	cc	cd	ce	cf
da	db	dc	dd	de	df
ea	eb	ec	ed	ee	ef
fa	fb	fc	fd	fe	ff

Using `\href{URL}{text}` to refer a website.
Using `\eq` to write equation, `\tab` to get an unordered list, `\lis` to get an ordered list.

$$E = mc^2 \tag{6.1.1.1}$$

- item 1;
- item 2;
- item 3.

1. item 1;
2. item 2;

3. item 3.

Format of Words

`\hl{highlighted},`
`\ul{underlined},`
`\st{strikethrough}\\`
`\rt{red}, \yt{yellow},`
`\bt{blue}, \gt{green}`
	highlighted, underlined, ~~strikethrough~~
	red, yellow, blue, green

Shortcuts

`\RR, \NN, \ZZ, \QQ\\` `\bA, \bB, \bC, \bD`	$\mathbb{R}$, $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ $\mathbb{A}, \mathbb{B}, \mathbb{C}, \mathbb{D}$

Emoji

`\emogood, \emobad, \emocool,` `\emoheart, \emotree`	, , , ,

Use `\ass, \ax, \thm, \co, \pro, \defi, \re, \key, \ex, \proo` to use preset tcolorboxes template.

Assumption 6.1.1.1: Example

1. item 1;
2. item 2;
3. item 3.

Axiom 6.1.1.1: Exmaple

Test

Theorem 6.1.1.1: Exmaple

Test

Corollary 6.1.1.1: Exmaple

Test

Proposition 6.1.1.1: Exmaple

Test

Definition 6.1.1.1: Exmaple

Test

Remark 6.1.1

Expamle

Keywords 6.1

Example

Example 6.1.1.1

Problem example

Solution:

The solution is omitted.

Proof 6.1.1: proposition 6.1.1

The Formal Proof

Q.E.D.

Using `\label, \ref` to refer to chapters 6, sections 6.1, equations 6.1.1, and boxes 6.1.1.

Using `\cite` to cite the literature in apa style. For example: [Klein et al., 2020]

Using `\sep` to insert a horizontal line with words in the middle:

Compilation

Clearning all the auxiliary files: LaTeXmk → BibTeX → LaTeXmk → LaTeXmk. Or, zip the main files and upload to Overleaf.

Put photos in the *pic* file and use `\fig` to show it.

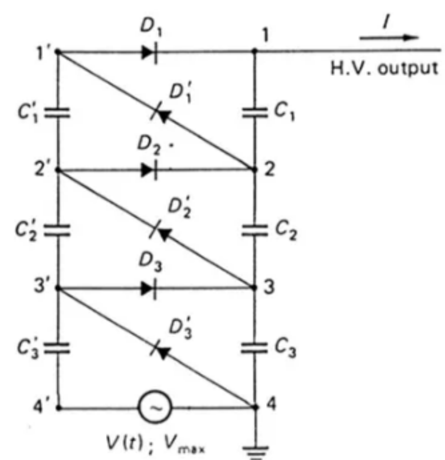

Figure 6.1: Example.png

6.2 Important Proofs

6.3 Beautiful Phrases

Introduction

To the limit of our knowledge, this study is among the first to unveil ...

PUBG Paper

The bot can be designed so that it is competitive but still loses to the human player, which may result in the human breaking their losing streak and remaining in the system longer. On the other hand, due to the limitations of technology, AI-powered bots can be identified by experienced players. If a human player is frequently matched with bots, they may find out that their opponents are not human and perhaps be discouraged from playing the game.

6.4 Eureka Ideas

6.4.1 RNN and Causal Model

On **1st Aug 2023**.
The RNN structure and rubin's causal model under changing environment looks similar;
To-do:

1. Inspect the papers containing keywords "RNN" and "Causal";
2. Explorations of Causal Models in Bayesian Causal Framework, with a Focus Beyond Rubin's Work;
3. explore further connection between RNN and causal model.

6.4.2 Earthquakes on Immigration and Investment

On **19th Aug 2023**.
Since 2019, numerous enterprises have engaged in shale gas extraction across various cities and counties in southern Sichuan, leading to frequent yet non-hazardous seismic activities. The geographical locations subjected to shale gas extraction remain independent of the local economic conditions, and the introduction of these enterprises has shown no discernible positive impact on the fiscal health of the local governments or the regional employment scenario. Essentially, this situation represents a discontinuity, wherein the extraction of shale gas corresponds to a quasi-random selection of regions experiencing seismic events. This phenomenon can be effectively studied using the Differences-in-Differences (DiD) methodology to investigate the causal effects of this unstable geological activity on the migration rates of local residents and the influx of external capital.
To Do:

1. Reviewing Literature to Investigate the Independence of Shale Gas from Local Economic Levels;
2. Assessing the Accessibility of Relevant Data for Investigation.

List of Figures

Bibliography

[Chen et al., 2021] Chen, M., Elmachtoub, A. N., and Lei, X. (2021). Matchmaking strategies for maximizing player engagement in video games. *Available at SSRN 3928966*.

[Chen et al., 2017] Chen, Z., Xue, S., Kolen, J., Aghdaie, N., Zaman, K. A., Sun, Y., and Seif El-Nasr, M. (2017). EOMM: An Engagement Optimized Matchmaking Framework. In *Proceedings of the 26th International Conference on World Wide Web*, pages 1143–1150, Perth Australia. International World Wide Web Conferences Steering Committee.

[Gallego and Topaloglu, 2019] Gallego, G. and Topaloglu, H. (2019). *Revenue Management and Pricing Analytics*, volume 279 of *International Series in Operations Research & Management Science*. Springer New York, New York, NY.

[Iyer and Yoganarasimhan, 2021] Iyer, G. and Yoganarasimhan, H. (2021). Strategic Polarization in Group Interactions. *Journal of Marketing Research*, 58(4):782–800.

[Klein et al., 2020] Klein, R., Koch, S., Steinhardt, C., and Strauss, A. K. (2020). A review of revenue management: Recent generalizations and advances in industry applications. *European Journal of Operational Research*, 284(2):397–412.

[Lei, 2022] Lei, X. (2022). *Revenue Management in Video Games and with Fairness*. PhD thesis, Columbia University.

[Narayan et al., 2011] Narayan, V., Rao, V. R., and Saunders, C. (2011). How Peer Influence Affects Attribute Preferences: A Bayesian Updating Mechanism. *Marketing Science*, 30(2):368–384.

[Rao and Steckel, 1991] Rao, V. R. and Steckel, J. H. (1991). A polarization model for describing group preferences. *Journal of Consumer Research*, 18(1):108–118.

[Strauss et al., 2018] Strauss, A. K., Klein, R., and Steinhardt, C. (2018). A review of choice-based revenue management: Theory and methods. *European Journal of Operational Research*, 271(2):375–387.

[Zhang et al., 2017] Zhang, C., Phang, C. W., Wu, Q., and Luo, X. (2017). Nonlinear Effects of Social Connections and Interactions on Individual Goal Attainment and Spending: Evidences from Online Gaming Markets. *Journal of Marketing*, 81(6):132–155.

[Zhao et al., 2023] Zhao, P., Ma, Z., Gill, T., and Ranaweera, C. (2023). Social media sentiment polarization and its impact on product adoption. *Marketing Letters*.