Dynamic Optimization for Agent-Based Systems and Inverse Optimal Control

Dynamic Optimization for Agent-Based Systems and Inverse Optimal Control

YIBEI LI

Licentiate Thesis
KTH Royal Institute of Technology
School of Engineering Sciences
Department of Mathematics
Division of Optimization and Systems Theory
Stockholm, Sweden 2019

Department of Mathematics
KTH Royal Institute of Technology
100 44 Stockholm, Sweden

TRITA-SCI-FOU 2019:36
ISBN 978-91-7873-248-7

Akademisk avhandling som med tillstånd av Kungliga Tekniska Högskolan framlägges till offentlig granskning för avläggande av teknologie doktorsexamen, måndagen den 30 Augusti 2019 klockan 10.00 i sal F11, Lindstedtsvägen 22, Kungliga Tekniska Högskolan, Stockholm.

Print: Universitetsservice US AB, Stockholm, 2019

To my family

Abstract

This dissertation is concerned with three problems within the field of optimization for agent–based systems. Firstly, the inverse optimal control problem is investigated for the single-agent system. Given a dynamic process, the goal is to recover the quadratic cost function from the observation of optimal control sequences. Such estimation could then help us develop a better understanding of the physical system and reproduce a similar optimal controller in other applications. Next, problems of optimization over networked systems are considered. A novel differential game approach is proposed for the optimal intrinsic formation control of multi-agent systems. As for the credit scoring problem, an optimal filtering framework is utilized to recursively improve the scoring accuracy based on dynamic network information.

In paper A, the problem of finite horizon inverse optimal control problem is investigated, where the linear quadratic (LQ) cost function is required to be estimated from the optimal feedback controller. Although the infinite-horizon inverse LQ problem is well-studied with numerous results, the finite-horizon case is still an open problem. To the best of our knowledge, we propose the first complete result of the necessary and sufficient condition for the existence of corresponding LQ cost functions. Under feasible cases, the analytic expression of the whole solution space is derived and the equivalence of weighting matrices is discussed. For infeasible problems, an infinite dimensional convex problem is formulated to obtain a best-fit approximate solution with minimal control residual, where the optimality condition is solved under a static quadratic programming framework to facilitate the computation.

In paper B, the optimal formation control problem of a multi-agent system is studied. The foraging behavior of N agents is modeled as a finite-horizon non-cooperative differential game under local information, and its Nash equilibrium is studied. The collaborative swarming behaviour derived from non-cooperative individual actions also sheds new light on understanding such phenomenon in the nature. The proposed framework has a tutorial meaning since a systematic approach for formation control is proposed, where the desired formation can be obtained by only intrinsically adjusting individual costs and network topology. In contrast to most of the existing methodologies based on regulating formation errors to the pre-defined pattern, the proposed method does not need to involve any information of the desired pattern beforehand. We refer to this type of formation control as intrinsic formation control. Patterns of regular polygons, antipodal formations and Platonic solids can be achieved as Nash equilibria of the game while inter-agent collisions are naturally avoided.

Paper C considers the credit scoring problem by incorporating dynamic network information, where the advantages of such incorporation are investigated in two scenarios. Firstly, when the scoring publishment is merely individual–dependent, an optimal Bayesian filter is designed for risk prediction, where network observations are utilized to provide a reference for the bank on future financial decisions. Furthermore, a recursive Bayes estimator is proposed to improve the accuracy of score publishment by incorporating the dynamic network topology as well. It is shown that under the proposed evolution framework, the designed estimator has a higher precision than all the efficient estimators, and the mean square errors are strictly smaller than the Cramér–Rao lower bound for clients within a certain range of scores.

Keywords: Inverse optimal control; formation control; differential game; credit scoring.

Sammanfattning

I denna avhandling behandlas tre problem inom optimering för agentbaserade system. Inledningsvis undersöks problemet rörande invers optimal styrning för ett system med en agent. Målet är att, givet en dynamisk process, återskapa den kvadratiska kostnadsfunktionen från observationer av sekvenser av optimal styrning. En sådan uppskattning kan ge ökad förståelse av det underliggande fysikaliska systemet, samt vara behjälplig vid konstruktion av en liknande optimal regulator för andra tillämpningar. Vidare betraktas problem rörande optimering över nätverkssystem. Ett nytt angreppssätt, baserat på differentialspel, föreslås för optimal intrinsisk formationsstyrning av system med fler agenter. För kreditutvärderingsproblemet utnyttjas ett filtreringsramverk för att rekursivt förbättra kreditvärderingens noggrannhet baserat på dynamisk nätverksinformation

I artikel A undersöks problemet med invers optimal styrning med ändlig tidshorisont, där den linjärkvadratiska (LQ) kostnadsfunktionen måste uppskattas från den optimala återkopplingsregulatorn. Trots att det inversa LQ-problemet med oändlig tidshorisont är välstuderat och med flertalet resultat, är fallet med ändlig tidshorisont fortfarande ett öppet problem. Så vitt vi vet presenterar vi det första kompletta resultatet med både tillräckliga och nödvändiga villkor för existens av en motsvarande LQ-kostnadsfunktion. I fallet med lösbara problem härleds ett analytiskt uttryck för hela lösningsrummet och frågan om ekvivalens med viktmatriser behandlas. För de olösbara problemen formuleras ett oändligtdimensionellt konvext optimeringsproblem för att hitta den bästa approximativa lösningen med den minsta styrresidualen. För att underlätta beräkningarna löses optimalitetsvillkoren i ett ramverk för statisk kvadratisk programmering.

I artikel B studeras problemet rörande optimal formationsstyrning av ett multiagentsystem. Agenternas svärmbeteende modelleras som ett icke-kooperativt differentialspel med ändlig tidshorisont och enbart lokal information. Vi studerar detta spels Nashjämvikt. Att, ur icke-kooperativa individuella handlingar, härleda ett kollaborativt svärmbeteende kastar nytt ljus på vår förståelse av sådana, i naturen förekommande, fenomen. Det föreslagna ramverket är vägledande i den meningen att det är ett systematiskt tillvägagångssätt för formationsstyrning, där den önskade formeringen kan erhållas genom att endast inbördes justera individuella kostnader samt nätverkstopologin. I motstat till de flesta befintliga metoder, vilka baseras på att reglera felet i formeringen relativt det fördefinierade mönstret, så behöver den föreslagna metoden inte på förhand ta hänsyn till det önskade mönstret. Vi kallar denna typ av formationsstyrning för intrinsisk formationsstyrning. Mönster så som regelbundna polygoner, antipodala formeringar och Platonska kroppar kan uppnås som Nashjämvikter i spelet, samtidigt som kollisioner mellan agenter undviks på ett naturligt sätt.

Artikel C behandlar kreditutvärderingsproblemet genom att lägga till dynamisk nätverksinformation. Fördelarna med en sådan integrering undersöks i två scenarier. Då kreditvärdigheten enbart är individberoende utformas ett optimalt Bayesiskt filter för riskvärdering, där observationer från nätverket används för att tillhandahålla en referens för banken på framtida finansiella beslut. Vidare föreslås en rekursiv Bayesisk estimator (stickprovsvariabel) för att förbättra noggrannheten på den skattade kreditvärdigheten genom att integrera även den dynamiska nätverkstopologin. Inom den föreslagna ramverket för tidsutveckling kan vi visa att, för kunder inom ett visst intervall av värderingar, har den utformade estimatorn högre precision än alla effektiva estimatorer och medelkvadrafelet är strikt mindre än den nedre gränsen från Cramér–Raos olikhet.

Acknowledgments

First and foremost, I would like to express my sincere gratitude to my supervisor, Professor Xiaoming Hu. Thank you for stimulating my interest in Mathematics during my exchange study in 2013 and also providing such an opportunity for my doctoral study at KTH. Your patience, guidance, encouragements and positive attitude have been of immense support during this journey. I indeed learn a lot from your methodologies in academic research and your inspiring ideas never ceased to amaze me. Your sound knowledge, enthusiasm, and the sense of responsibility make you a great supervisor.

Secondly, I would like to express my gratitude to my advisor Professor Yu Yao as well as Professor Fenghua He at my home university, Harbin Institute of Technology. Thanks for leading me to the field of academia and every piece of knowledge you have shared with me. It is your strict and professional supervision that has endowed me with the basic abilities as an independent researcher.

Furthermore, I wish to thank the faculty at the Division of Optimization and Systems Theory. Thank you Professor Anders Forsgren and Professor Krister Svanberg for opening the door of optimization for me. It is your impressive lectures and patient discussions that help me establish a basic framework in this field. I am also grateful to Associate Professor Per Enqvist and Associate Professor Johan Karlsson, for improving my understanding in control theory and the interesting talks during go-outside lunch.

In addition, it is really fantastic to work with all my colleagues at the division. They are: Doctor Fei Wang, Doctor Ximei Wang, Doctor Axel Ringh, Doctor Han Zhang, Doctor Silun Zhang, Doctor Lovisa Engberg Sundström, Isabel Haasler, Alexander Aurell, Michelle Böck, Emil Ringh, David Ek, Sara Boldly, Martin Ryner and Göran Svensson. Thanks for your generous support and creating such a friendly and fantastic atmosphere during my study here. Especially, I would like to thank Emil Ringh for translating the abstract of this thesis into Swedish.

Last but most, a very special gratitude goes out to all my family members, especially my parents and my sister, for your unconditional support during my stay in Sweden. Thank you Yi for your accompany in all the ups and downs along this journey. I am also grateful to all my friends in Sweden for the memorable hours outside work of playing badminton, card games and pleasant dinners, which shed light on my life in the Swedish winter darkness.

Stockholm, April 2019

Yibei Li

Table of Contents

Part II: Research Papers

If I come up with 300 ideas in a year,
and only one of them is useful,
I am content.

– Alfred Nobel

Part I:
Introduction and Preliminaries

Chapter: 1

Introduction

In this thesis, the problems of inverse optimal control and dynamic optimization for agent-based systems are investigated. In particular, three subproblems are studies:

1. **Inverse linear quadratic optimal control problem:**
   Given the optimal controller of a dynamic system, recover the quadratic cost function based on which the optimal controller is generated.

2. **Optimal intrinsic formation control problem:**
   Design a systematic method for multi-agent systems to realize different formations optimally and intrinsically, where only local information from neighbors are available.

3. **Credit scoring problem incorporated by network information:**
   In addition to individual financial attributes, is it possible to use network-based measurements to further improve the credit score accuracy?

This chapter contains a brief introduction of the background and motivation of the problems studied in the thesis. A preview of all the appended papers is also included. Preliminary knowledge and mathematical tools used in this thesis can be found in the next chapter.

1.1 Background and Motivations

Inverse Optimal Control

Throughout the evolution of nature, the optimality principle has been investigated as an important tool to analyze natural phenomena, such as Fermat's law in optics and principle of least action in mechanics. In the field of biology, it is also a general hypothesis that the behavior of living systems are generated based on some optimal criteria, which leads to a promising topic of inverse optimization.

In recent years, the problem of inverse optimization has regained increasing popularity in the fields of robotics, economics, and bionics [22, 9, 5, 6]. The basic question is that given a dynamic system, when we observe the optimal policy of a specific task, how can we recover the optimization criterion based on which the optimal policy is generated? Such estimation could then help us develop a better understanding of the physical system and reproduce a similar optimal controller in other applications. For example, inverse optimal control is a promising tool to investigate the mechanisms underlying the human locomotion and to implement them in the design of humanoid robots [21].

As for the general case, the problem of reconstructing cost functions has been investigated intensively. Among the existing literatures, one well-studied direction is to treat it as a parameter identification problem, where numerous numerical results have been developed [22, 5, 12, 20, 15, 30, 29]. Under this situation the cost function is usually assumed to be a linear combination of certain basic functions, with the weights remaining to be identified. Then the unknown coefficients can be obtained whether by machine learning methods or residual optimization based on optimality conditions.

Among the various forms of the cost function, one important direction falls under the field of deterministic linear quadratic problems, which are not only well-defined but also popular for practical purposes. Some analytic results can also be obtained due to its special form. The inverse LQ problem is first proposed by Kalman in 1964 [17]. Compared to general cost estimation problems, the difficulty of the inverse LQ problem lies in the fact that the weighting matrix has to be estimated in the positive semi-definite cone. Although for infnite-horizon case there exist a number of results [1, 10, 8, 31], the finite-horizon problem is still an open problem, which is investigated in this thesis by the following steps:

1. Existence: determine the necessary and sufficient conditions on system matrices and the observed feedback control $u^*(t)$, such that $u^*(t)$ is an optimal control law for some linear quadratic cost function.

2. Solution: determine all pairs of weighting matrices in the cost function corresponding to the same optimal controller.
    - If the existence problem is feasible, give the analytic expression of the whole solution space.
    - For infeasible case, derive a best approximate cost function with minimal control residual.

Multi-Agent Systems and Formation Control

Multi-agent coordination is a fast-emerging field in control community, which has gained increasing attention in the past decades. Many research topics within this field are inspired by biological modeling of the collective phenomena in the nature, such as flocking of birds and fishes [11]. Multi-agent system is popular due to its advantages of better robustness as well as lower communication and computation burden. The theory has been widely applied to motion planning of multi-robots system, where the agents control their own dynamics to achieve a cooperative task by exchanging information with neighbors. Following significant results on consensus problem, realization of various formation patterns has attracted more attention in the past several years [27].

In this thesis, the formation control problem is investigated in a differential game framework, where the novelty are motivated from the following aspects.

Intrinsic control: As for the non-consensus formation control problem, most existing methodologies are based on the pre-defined formation pattern, where the formation error is regulated to zero [18, 19]. Inspired by [38, 34], in this thesis the formation is achieved in an "intrinsic" way in the sense that it is only attributed to the inter-agent interaction and geometric properties of the network, where the desired formations are not designated beforehand. The reduced attitude on the compact manifold $\mathcal{S}^2$ is considered in [38, 34], and symmetric formations can be obtained by designing inter-agent repulsion. However, those results cannot be applied to the Euclidean space directly since it is unbounded.

Distributed framework: Implementing the system in a distributed framework is mainly motivated by two aspects. Firstly, the distributed manner facilitates more robustness than centralized structures, where the collapse of a single node might lead to the breakdown of the whole system. Secondly, Modeling and computation tasks are becoming much more complex as the size of the system continues to increase. It is communicationally heavy or sometimes impossible for each agent to know the absolute positions of all other agents. Relying merely on local information of neighbors can significantly decrease the burden and facilitate the controller design.

Non-cooperative differential game framework: In recent years, the game theoretic approach has shed new light on the formation control problems. Game theory, in particular evolutional game theory, has been applied to multi-agent systems such as [23, 26]. Although there are numerous results on situations in which agents cooperative to achieve a common task, there are more practical scenarios where agents have individual and partial conflicting goals, thus leading to a non-cooperative setting. Differential games focus on multi-player decision making problems over a given time interval, where each agent aims to optimize its own, individual cost subject to the common state dynamics [2]. Furthermore, game theoretical approach has the advantage of realizing desired formations from an optimization perspective. The derived controller is then not only optimal, but also possesses better performance like robustness. However, among the existing results of formation control problem based on differential game theory, most papers focus on the *synchronization* problem [32, 3, 33]. To the best of our knowledge, only a few papers in this field have considered the case outside the consensus framework, such as [28, 37].

Network-Based Credit Scoring

When it comes to the purchase of financial assets such as deposits, loans and securities, one major issue with granting loans is whether the clients could fulfill their obligation or not, which is characterized in the form of credit scores. The credit scoring problem has been one of the key topics in financial risk management analysis for individual customers, giant commercial loans, and even governments [4]. The purpose of credit scoring is to assess the ability and willingness of individuals to meet their financial obligations on time. The scorings by the credit bureau (e.g., financial institutions or governments) play an important role in the investors' decisions.

The overall objective of credit scoring prediction is to build models that can extract knowledge of credit risk evaluation from past observations [13]. There exist numerous promising results on credit prediction based on different choices of statistical models, such as the Credit Risk+ model by Credit Suisse Financial Products (CSFP) [35] and the KMV model developed by KMV company [7]. In some methods, credit rating is also formulated as a classification problem and various tools of machine learning can be utilized [24, 16].

Up until recently, assessing consumers' creditworthiness relies merely on their own financial attributes. In the above existing methods, individual data like salary, debt value and length of credit history has been considered as main attributes in credit rating models. However, in recent years the credit scoring industry has witnessed a dramatic change in data sources [36]. Network information such as users' social networking profiles or the trading network information provided by the bank has attracted increasing attention.

Motivated by the growing use of network information, in this thesis the network-based credit scoring problem is investigated. In particular, the following two questions are addressed:

1. How to model the relationship between network information and individual credit scores?

2. Is there any advantage of network-based credit scoring over methods only based on individual data? If the answer is affirmative, how to use network information to derive a more accurate scoring?

1.2 Summary of the Appended Papers

Paper A Continuous-Time Inverse Quadratic Optimal Control Problem

This paper is co-authored with Yu Yao and Xiaoming Hu, and is submitted to *Automatica, 2018.*

Summary: In paper A, the problem of finite horizon inverse optimal control is investigated. The goal is to recover the quadratic cost function of a dynamic process based on the observation of optimal control sequences. As for the inverse Linear Quadratic (LQ) problem, although the infinite-horizon case is well-studied with numerous results, the finite-horizon problem is still an open problem. In contrast to existing works like [25] and [14], here we focus on the standard form without cross terms, which is more advantageous in its practical meaning. To the best of our knowledge, we propose the first complete result of the necessary and sufficient condition for the existence of corresponding LQ cost functions. Under feasible cases, the analytic expression of the whole solution space is derived and the equivalence of weighting matrices in LQ problems is discussed. For infeasible problems, an infinite dimensional convex problem is formulated to obtain a best-fit approximate solution with minimal control residual. And the optimality condition is solved under a static quadratic programming framework to facilitate the computation. The performance of the proposed methods are also illustrated by numerical examples and simulations.

Contribution: The main idea is initiated by the co-author Xiaoming Hu. The contribution by the author of the thesis is the main part of theoretical results and other authors primarily provide suggestions to improve the quality of the paper.

Paper B A Differential Game Approach to Optimal Intrinsic Formation Control

This paper is co-authored with Xiaoming Hu, and is submitted to *IEEE Annual Conference on Decision and Control (CDC), 2019.*

Summary: Paper B concerns the optimal formation control problem of a multi-agent system. The foraging behavior of N agents is modeled as a finite-horizon non-cooperative differential game under local information, and its Nash equilibrium is studied. In contrast to most of the existing methodologies based on regulating formation errors to the pre-defined pattern, in this paper formations are achieved in an intrinsic manner that does not need to involve any information of the desired pattern beforehand. The proposed framework has a tutorial meaning since a novel systematic approach for formation control is proposed, where the desired formation can be obtained by only intrinsically adjusting individual costs and network topology. Patterns of regular polygons, antipodal formations and Platonic solids can be achieved as Nash equilibria of the game while inter-agent collisions are naturally avoided. Furthermore, the results not only lead to a better understanding of the natural phenomenon where a collaborative swarming behaviour can result from non-cooperative individual actions, but also bring new inspiration in the construction of other formations. Numerical simulations are also provided in both two-dimensional and three-dimensional Euclidean space to demonstrate the effectiveness and feasibility of the proposed methods.

Contribution: The main idea is initiated by the co-author Xiaoming Hu. The theoretical results are developed by the author of this thesis under the supervision of Xiaoming Hu.

Paper C Credit Scoring by Incorporating Dynamic Network Information

This paper is co-authored with Ximei Wang, Boualem Djehiche and Xiaoming Hu, and is submitted to *Journal of Economic Dynamics and Control, 2019*.

Summary: Paper C considers the credit scoring problem for a group of clients. In contrast to most of the existing methodologies where only individual financial attributes are utilized, this paper also investigates the advantages of incorporating network information. The client network is modeled according to homogeneous preference based on others' credit assessments reported by the bank. It is shown that such correlation can be developed to further improve the scoring precision in two scenarios. Firstly, a Bayesian optimal filter is proposed to provide a risk prediction for banks assuming that published credit scores are estimated merely from structured individual data. Such prediction can then serve as a reference for the bank on future financial decisions. We further propose a recursive Bayes estimator to improve the accuracy of score publishment by incorporating the dynamic interaction topology of clients as well. It is shown that under the proposed evolution framework, the estimation variance can be significantly reduced, which is strictly smaller than the Cramér–Rao lower bound. In addition, as for the uniformly distributed scores, the mean square error of the proposed estimator is strictly smaller than that of all efficient estimators for clients in the middle class.

Contribution: The main ideas emerged from the discussion among all the authors. The theoretical results are mainly developed by the author of this thesis, and the numerical simulations are conducted in cooperation with Ximei Wang.

References

[1] J. Moore B. Anderson. *Optimal Control: Linear Quadratic Methods*. Prentice-Hall International Inc., 1989.

[2] Tamer. Basar and Geert Jan. Olsder. *Dynamic noncooperative game theory*. SIAM, 1999.

[3] D. Bauso, L. Giarré, and R. Pesenti. Non-linear protocols for optimal distributed consensus in networks of dynamic agents. *Systems & Control Letters*, 55(11):918–928, nov 2006.

[4] William H Beaver. Financial ratios as predictors of failure. *Journal of accounting research*, pages 71–111, 1966.

[5] Bastien Berret, Enrico Chiovetto, Francesco Nori, and Thierry Pozzo. Evidence for composite cost functions in arm movement planning: an inverse optimal control approach. *PLoS computational biology*, 7(10):e1002183, 2011.

[6] Bastien Berret and Frédéric Jean. Why don't we move slower? the value of time in the neural control of action. *Journal of neuroscience*, 36(4):1056–1070, 2016.

[7] Jeff Bohn and Peter Crosbie. Modeling default risk. *KMV Corporation*, 2003.

[8] Stephen Boyd, Laurent El Ghaoui, Eric Feron, and Venkataramanan Balakrishnan. *Linear matrix inequalities in system and control theory*, volume 15. Siam, 1994.

[9] Chelsea Finn, Sergey Levine, and Pieter Abbeel. Guided cost learning: Deep inverse optimal control via policy optimization. In *International Conference on Machine Learning*, pages 49–58, 2016.

[10] Takao Fujii. A new approach to the lq design from the viewpoint of the inverse regulator problem. *IEEE Transactions on Automatic Control*, 32(11):995–1004, 1987.

[11] Musad A Haque, Amir R Rahmani, and Magnus B Egerstedt. Biologically inspired confinement of multi-robot systems. *International Journal of Bio-Inspired Computation*, 3(4):213–224, 2011.

[12] Kathrin Hatz, Johannes P Schloder, and Hans Georg Bock. Estimating parameters in optimal control problems. *SIAM Journal on Scientific Computing*, 34(3):A1707–A1728, 2012.

[13] Zan Huang, Hsinchun Chen, Chia-Jung Hsu, Wun-Hwa Chen, and Soushan Wu. Credit rating analysis with support vector machines and neural networks: a market comparative study. *Decision support systems*, 37(4):543–558, 2004.

[14] Frédéric Jean and Sofya Maslovskaya. Inverse Optimal Control Problem: The Linear-Quadratic Case. working paper or preprint, March 2018.

[15] Miles Johnson, Navid Aghasadeghi, and Timothy Bretl. Inverse optimal control for deterministic continuous-time nonlinear systems. In *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on*, pages 2906–2913. IEEE, 2013.

[16] Stewart Jones, David Johnstone, and Roy Wilson. Predicting corporate bankruptcy: an evaluation of alternative statistical frameworks. *Journal of Business Finance & Accounting*, 44(1-2):3–34, 2017.

[17] Rudolf Emil Kalman. When is a linear control system optimal? *Journal of Basic Engineering*, 86(1):51–60, 1964.

[18] Ali Karimoddini, Hai Lin, Ben M. Chen, and Tong Heng Lee. Hybrid three-dimensional formation control for unmanned helicopters. *Automatica*, 49(2):424–433, feb 2013.

[19] Hiroaki Kawashima and Magnus Egerstedt. Manipulability of leader–follower networks with the rigid-link approximation. *Automatica*, 50(3):695–706, mar 2014.

[20] Arezou Keshavarz, Yang Wang, and Stephen Boyd. Imputing a convex objective function. In *Intelligent Control (ISIC), 2011 IEEE International Symposium on*, pages 613–619. IEEE, 2011.

[21] Jim Mainprice, Rafi Hayne, and Dmitry Berenson. Goal set inverse optimal control and iterative replanning for predicting human reaching motions in shared workspaces. *IEEE Trans. Robotics*, 32(4):897–908, 2016.

[22] Katja Mombaur, Anh Truong, and Jean-Paul Laumond. From human to humanoid locomotion-an inverse optimal control approach. *Autonomous robots*, 28(3):369–383, 2010.

[23] Thulasi Mylvaganam and Alessandro Astolfi. Approximate optimal monitoring: preliminary results. In *2012 American Control Conference (ACC)*, pages 4745–4750. IEEE, 2012.

[24] Natalia Nehrebecka. Predicting the default risk of companies. comparison of credit scoring models: Logit vs support vector machines. *Econometrics*, 22(2):54–73, 2018.

[25] Francesco Nori and Ruggero Frezza. Linear optimal control problems and quadratic cost functions estimation. In *Proceedings of the 12th IEEE Mediterranean Conference on Control and Automation (MED'04)*, pages 6–9, 2004.

[26] Mojtaba Nourian, Peter E. Caines, Roland P. Malhame, and Minyi Huang. Nash, Social and Centralized Solutions to Consensus Problems via Mean Field Control Theory. *IEEE Transactions on Automatic Control*, 58(3):639–653, mar 2013.

[27] Kwang-Kyo Oh, Myoung-Chul Park, and Hyo-Sung Ahn. A survey of multi-agent formation control. *Automatica*, 53:424–440, mar 2015.

[28] Arif Bülent Özgüler and Aykut Yıldız. Foraging swarms as nash equilibria of dynamic games. *IEEE transactions on cybernetics*, 44(6):979–987, 2014.

[29] Edouard Pauwels, Didier Henrion, and Jean-Bernard Lasserre. Linear conic optimization for inverse optimal control. *SIAM Journal on Control and Optimization*, 54(3):1798–1825, 2016.

[30] Edouard Pauwels, Didier Henrion, and Jean-Bernard Bernard Lasserre. Inverse optimal control with polynomial optimization. In *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*, pages 5581–5586. IEEE, 2014.

[31] M Cody Priess, Richard Conway, Jongeun Choi, John M Popovich, and Clark Radcliffe. Solutions to the inverse lqr problem with application to biological systems analysis. *IEEE Transactions on Control Systems Technology*, 23(2):770–777, 2015.

[32] E. Semsar and K. Khorasani. Optimal Control and Game Theoretic Approaches to Cooperative Control of a Team of Multi-Vehicle Unmanned Systems. In *2007 IEEE International Conference on Networking, Sensing and Control*, pages 628–633. IEEE, 2007.

[33] F. Shamsi, F. Abdollahi, and K.Y. Nikravesh. Time varying Formation Control Using Differential Game Approach. *IFAC Proceedings Volumes*, 44(1):1114–1119, jan 2011.

[34] Wenjun Song, Johan Markdahl, Silun Zhang, Xiaoming Hu, and Yiguang Hong. Intrinsic reduced attitude formation with ring inter-agent graph. *Automatica*, 85:193–201, nov 2017.

[35] Credit Suisse. Creditrisk+: A credit risk management framework. *Credit Suisse Financial Products*, pages 18–53, 1997.

[36] Yanhao Wei, Pinar Yildirim, Christophe Van den Bulte, and Chrysanthos Dellarocas. Credit scoring with social network data. *Marketing Science*, 35(2):234–258, 2015.

[37] Aykut Yildiz and A. Bülent Özgüler. Foraging motion of swarms with leaders as Nash equilibria. *Automatica*, 73:163–168, nov 2016.

[38] Silun Zhang, Wenjun Song, Fenghua He, Yiguang Hong, and Xiaoming Hu. Intrinsic tetrahedron formation of reduced attitude. *Automatica*, 87:375–382, jan 2018.

Chapter: 2

Preliminaries

This chapter provides an overview of some well-known concepts and theories that will be used in this thesis. Mathematical tools of control theory, optimal control, differential games, convex optimization and estimation theory are covered. Since all the theory in this chapter has been well-established, the proof of the theorems in this chapter will be omitted. Further details can be found in the books listed on the reference list at the end of this chapter.

2.1 Control Systems

In this part, some fundamental concepts in control theory are introduced. In general, a continuous-time dynamic system can be modeled in its *state-space form* by a set of ordinary differential equations (ODE)

$$\begin{aligned} \dot{x}(t) &= f(t, x(t), u(t)), \\ y(t) &= h(t, x(t), u(t)), \end{aligned} \tag{2.1}$$

where $x(t) \in \mathbb{R}^n$ is the *state variable*, $u(t) \in \mathbb{R}^m$ is the *control input* and $y(t) \in \mathbb{R}^p$ is the *output*.

In order for (2.1) to be a well-posed model, some fundamental properties of the ODE has to be investigated.

Theorem 2.1.1 (Local Existence and Uniqueness)**.** *Consider the initial value problem*

$$\dot{x}(t) = f(t, x(t)), \quad x(t_0) = x_0. \tag{2.2}$$

Let $f(t, x)$ be piecewise continuous in t and satisfy the Lipschitz condition on a neighborhood $x \in B(x_0)$ and $t \in [t_0, t_1]$. Then there exists some $\delta > 0$ such that (2.2) has a unique solution over $[t_0, t_0 + \delta]$.

Roughly speaking, the *Linear time-invariant (LTI)* system serves as the fundamental model of control theory, which also provides the theoretic basis for many advanced nonlinear systems. In general, a LTI system is modeled by

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t) + Du(t), \end{aligned} \tag{2.3}$$

where A, B, C and D are constant matrices with proper dimensions.

The solution of state trajectory $x(t)$ to (2.3) and be determined by its state transition matrix

$$\Phi(t, s) = e^{A(t-s)}, \tag{2.4}$$

and thus

$$x(t) = e^{A(t-t_0)} x_0 + \int_{t_0}^{t} e^{A(t-s)} B(s)u(s)ds. \tag{2.5}$$

Controllability and observability are two fundamental properties of the control system. For LTI systems, they can be determined by the following theorem.

Theorem 2.1.2. *The LTI system (2.3) is controllable if and only if the controllability matrix*

$$\Gamma = \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix},$$

has full row rank; the system is observable if and only if the observability matrix

$$\Omega = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix},$$

has full column rank.

2.2 Optimal Control

Optimal control theory is an important component of modern control theory, which provides a systematic tool for a large variety of control design problems. Optimal control theory is closely related in its origins to the theory of calculus of variations, and is further developed by giants as Bernoulli, Euler, Lagrange, and so on. In particular, Pontryagin minimum principle (PMP) and Bellman's dynamic programming are two powerful tools for solving most optimal control problems. The linear quadratic regulator (LQR) of particular importance

A general optimal control problem can be formulated as:

$$\begin{aligned} \min_{u} \ & \phi(x(t_f)) + \int_{t_0}^{t_f} f_0\left(t, x\left(t\right), u\left(t\right)\right) dt \\ s.t. \ & \dot{x}\left(t\right) = f\left(t, x\left(t\right), u\left(t\right)\right) \\ & x\left(t_0\right) = x_0 \\ & x\left(t_f\right) \in \mathcal{S}_f \\ & u(t) \in U(t, x(t)). \end{aligned} \tag{2.6}$$

Pontryagin minimum principle

Define the *Hamiltonian* function

$$H(t, x, u, \lambda) = f_0(t, x, u) + \lambda^T f(t, x, u), \tag{2.7}$$

where $\lambda(t)$ is called the *adjoint function*.

Then the Pontryagin minimum principle (PMP) provides a necessary condition for a control input $u^*(t)$ and the corresponding state trajectory $x^*(t)$ to be optimal.

Theorem 2.2.1 (PMP). *Suppose $(x^*(t), u^*(t))$ is an optimal solution of (2.6). Then there exists a nonzero adjoint variable $\lambda(t)$ such that*

(i) $\lambda(t)$ satisfies the adjoint equation

$$\dot{\lambda}(t) = -\frac{\partial H}{\partial x}(t, x^*(t), u^*(t), \lambda(t)). \tag{2.8}$$

(ii) $u^*(t)$ is the pointwise minimizer of the Hamiltonian

$$u^*(t) = \underset{u \in U(t, x^*)}{\arg\min} H(t, x^*(t), u, \lambda(t)). \tag{2.9}$$

(iii) $\lambda(t_f)$ *satisfies the boundary condition*

$$\lambda(t_f) - \bigtriangledown\phi(x^*(t_f)) \perp \mathcal{S}_f. \tag{2.10}$$

Then with PMP, candidates for optimality can be found by solving a two-point boundary value problem.

Dynamic programming

Based on the *principle of optimality* proposed by Richard Bellman, the dynamic programming equation can be derived for solving the optimal control problem. And the continuous time dynamic programming is closely related to the Hamilton-Jacobi-Bellman equation (HJBE), which provides a sufficient condition for optimality.

Theorem 2.2.2. *Suppose there exists a continuous function* $V : [t_0, t_f] \times \mathbb{R}^n \to \mathbb{R}$ *such that:*

(i) $V(t,x)$ *is* C^1 *(in both arguments) and solves the HJBE*

$$\begin{aligned} -\tfrac{\partial V}{\partial t}(t,x) &= \min_{u \in U(t,x)} \left\{ f_0(t,x,u) + \tfrac{\partial V}{\partial x}(t,x)^T f(t,x,u) \right\}, \\ V(t_f, x) &= \phi(x). \end{aligned} \tag{2.11}$$

(ii) $\mu(t,x) = \underset{u \in U(t,x)}{\arg\min} \left\{ f_0(t,x,u) + \frac{\partial V}{\partial x}(t,x)^T f(t,x,u) \right\}$ *is a piecewise continuous function for any closed loop solution.*

Then

(a) $V(t_0, x_0)$ *is the optimal cost.*

(b) $\mu(t,x)$ *is the optimal feedback control law, i.e.* $u^*(t) = \mu(t,x)$.

Linear quadratic regulator

In general, there is no systematic way to solve the HJBE (2.11). However, the special case of linear quadratic regulator (LQR) is well-studied, where the necessary and sufficient condition for optimality can be obtained. And a unique feedback optimal control can be computed by solving the Riccati equation.

Consider the continuous time LQR problem:

$$\begin{aligned}\min_{u}\ & x^{T}\left(t_{f}\right)Fx\left(t_{f}\right)+\int_{0}^{t_{f}}\left(x^{T}\left(t\right)Qx\left(t\right)+u^{T}\left(t\right)Ru\left(t\right)\right)dt\\ s.t.\ & \dot{x}\left(t\right)=Ax\left(t\right)+Bu\left(t\right)\\ & x\left(t_{0}\right)=x_{0},\end{aligned} \tag{2.12}$$

where $Q = Q^T \succeq 0$, $F = F^T \succeq 0$, and $R = R^T \succ 0$.

Theorem 2.2.3. *The necessary and sufficient conditions for $u^*(t)$ to be the unique optimal controller of (2.12) is:*

$$u^*(t) = -R^{-1}B^T P\left(t\right)x\left(t\right), \tag{2.13}$$

where $P(t) = P(t)^T \succeq 0$ is the unique nonnegative semi-definite solution to the differential Riccati equation (DRE):

$$-\dot{P}(t) = P(t)A + A^T P(t) - P(t)BR^{-1}B^T P(t) + Q, \tag{2.14}$$

with boundary condition $P(t_f) = F$.

Furthermore, the optimal performance is given by

$$J^*(x_0) = \frac{1}{2}x_0^T P(t_0)x_0. \tag{2.15}$$

For the infinite-horizon case where $t_f = +\infty$, the problem ends up with the algebraic Riccati equation (ARE)

$$PA + A^T P - PBR^{-1}B^T + Q = 0, \tag{2.16}$$

and the unique optimal feedback controller is stabilizing under assumptions of controllability and observability.

2.3 Differential Games

In a nutshell, game theory involves a multi-player decision making process. In this thesis, noncooperative games are studied, where each player pursues his own interests which are partially conflicting with each others'.

In particular, we consider the class of differential games, where the evolution of states is described by a differential equation and the players act throughout a time interval. As for a system of N players, let $x_i(t) \in \mathbb{R}^n$ denote the state of player i and $x(t) = [x_1(t); ...; x_N(t)]$ be the system state. Then a noncooperative differential game for N players can be defined as following.

Definition 2.3.1. An N-player differential game of prespecified fixed duration involves the following:

(i) An index set $N = \{1, ..., N\}$ called the players' set.

(ii) A time interval $[0, T]$ which is specified and denotes the duration of the evolution of the game.

(iii) An infinite set S_0 with some topological structure, called the trajectory space of the game. Its elements are denoted as $\{x(t), 0 < t < T\}$ and constitute the permissible state trajectories of the game. Furthermore, for each fixed $t \in [0, T], x(t) \in S_0$, where S_0 is a subset of a finite dimensional vector space, say $\mathbb{R}^{Nn}$.

(iv) An infinite set U_i with some topological structure, defined for each $i \in N$ and which is called the control (action) space of game P_i, whose elements $\{u_i(t), 0 < t < T\}$ are the control functions or simply the controls of P_i.

(v) A differential equation

$$\dot{x}(t) = f(t, x(t), u_1(t), ..., u_N(t)), \quad x(0) = x_0, \tag{2.17}$$

whose solution describes the state trajectory of the game corresponding to the N-tuple of control functions $\{u_i(t), 0 < t < T\}$ $(i = 1, ..., N)$ and the given initial state x_0.

(vi) Each player i has his own individual cost to minimize, which is given by a well-defined functional

$$J_i(x(0), u_1, ..., u_N) = \int_0^T l_i(t, x, u_1, ..., u_N)dt + \phi_i(x(T)). \tag{2.18}$$

Then the dynamic non-cooperative game played by player i is given by (P_i):

$$\begin{aligned} \min_{u_i} \;\; & J_i(x(0), u_i, u_{-i}) \\ s.t. \;\; & \dot{x}(t) = f(t, x(t), u_1(t), ..., u_N(t)), \\ & x(0) = x_0. \end{aligned} \tag{2.19}$$

where u_{-i} is the strategy profile of all players except for player i.

The solution of a non-cooperative game is usually characterized by its *Nash Equilibrium*, where no player can benefit by changing strategies while the other players keep theirs unchanged.

Definition 2.3.2 (Nash Equilibrium Strategy)**.** The set of admissible control functions $(u_1^*, u_2^*, ..., u_N^*)$ is called *Nash Equilibrium Strategy* if

$$J_i(x(0), u_i^*, u_{-i}^*) \leq J_i(x(0), u_i, u_{-i}^*), \tag{2.20}$$

for all $u_i \neq u_i^*$, $i = 1, 2, ..., N$.

For continuous-time differential games, necessary conditions for the existence of a Nash equilibrium is given by the following theorems.

Theorem 2.3.3. *For an N–person differential game on time interval $[0, T]$, let*

(i) $f(t, \cdot, u_1(t), ..., u_N(t))$ *be continuously differentiable on* $\mathbb{R}^{Nn}$*,* $\forall t \in [0, T]$*,*

(ii) $l_i(t, \cdot, u_1, ..., u_N)dt$ *and* $\phi_i(\cdot)$ *be continuously differentiable on* $\mathbb{R}^{Nn}$*,* $\forall t \in [0, T]$ *and* $\forall i = 1, \cdots, N$.

If $\{\gamma_i^*(t, x_0) = u_i^*(t)\}_{i=1}^N$ *provides an open-loop Nash equilibrium solution on* $[0, T]$*, and* $x^*(t)$ *is the corresponding state trajectory, then there exist N costate functions* $p_i(\cdot) : [0, T] \to \mathbb{R}^{Nn}$ *, such that the following equations are satisfied:*

$$\begin{aligned}
&\dot{x}^*(t) = f(t, x(t), u_1^*(t), ..., u_N^*(t)), x^*(0) = x_0, \\
&\gamma_i^*(t, x_0) := u_i^*(t) = \arg\min_{u_i \in U_i} H(t, x^*(t), u_1^*(t), ..., u_{i-1}^*(t), u_i(t), \\
&\qquad\qquad\qquad\qquad\qquad\qquad u_{i+1}^*(t), ...u_N^*(t), p_i(t)), \\
&\dot{p}_i(t) = -\frac{\partial}{\partial x} H_i(t, x^*(t), u_1^*(t), ..., u_N^*(t), p_i(t)), \\
&p_i(T) = \frac{\partial}{\partial x} \phi_i(x^*(T)),
\end{aligned} \tag{2.21}$$

where $i = 1, ..., N$ *and*

$$H_i(t, x, u_1, ..., u_N, p_i) = l_i(t, x, u_1, ..., u_N)dt + p_i^T f(t, x, u_1, ..., u_N).$$

As for the closed-loop Nash equilibrium, the following sufficient conditions can also be derived.

Theorem 2.3.4. *For an N-person differential game on* $[0, T]$*, an N-tuple strategies* $\{\gamma_i^*\}$ *provides a feedback Nash equilibrium solution if there exist functions* V_i :

$[0,T] \times \mathbb{R}^{Nn} \to \mathbb{R}, i = 1, \cdots, N$, *such that*

$$\begin{aligned} -\frac{\partial V_i(t,x)}{\partial t} &= \min_{u_i \in U_i} \{ \frac{\partial V_i(t,x)}{\partial x} f(t,x,\{u_i,\gamma^*_{-i}\}) + l_i(t,x,\{u_i,\gamma^*_{-i}\}) \} \\ &= \frac{\partial V_i(t,x)}{\partial x} f(t,x,\gamma^*) + l_i(t,x,\gamma^*), \\ V_i(T,x) &= \phi_i(x). \end{aligned} \tag{2.22}$$

And the corresponding Nash equilibrium cost for (P_i) is $V_i(0, x_0)$.

2.4 Convex Optimization and Duality

Optimization is a useful tool in many fields. In this part, we give a brief introduction of the convex optimization theory, which is frequently used throughout the thesis.

Definition 2.4.1. A real-valued functional f defined on a convex subset C of a linear vector space is called a *convex functional* if

$$f(\alpha x_1 + (1-\alpha)x_2) \leq \alpha f(x_1) + (1-\alpha) f(x_2), \tag{2.23}$$

for all $x_1, x_2 \in C$ and $\alpha \in [0,1]$. If strict inequality holds whenever $x_1 \neq x_2$, f is then said to be *strictly convex*. And a *convex mapping* can be defined in a similar way.

By introducing a cone defining the positive vectors in a given space, it is then possible to consider inequality constraints for optimization problems.

Definition 2.4.2. Let P be a convex cone in a vector space X. For any $x, y \in X$, we write $x \geq_P y$ (i.e. with respect to P) if $x - y \in P$. The cone P defining this relation is called the *positive cone* in X. The cone $N = -P$ is called the *negative cone* in X and we write $y \leq_P x$ for $y - x \in N$.

Let X be a linear vector space, Z a normed space, Ω a convex subset of X, and P the positive cone in Z. Assume that P contains an interior point. In general, a convex optimization problem can be considered as:

$$\begin{aligned} \min_x \ & f(x) \\ s.t. \ & G(x) \leq_P 0 \\ & x \in \Omega, \end{aligned} \tag{2.24}$$

where f is a real-valued convex functional on Ω and G is a convex mapping from Ω into Z.

Duality theory is one of the dominating tools in solving optimization problems. Before considering dual optimization problems, firstly the definition of the dual space is given.

Definition 2.4.3. Let X be a normed linear vector space. The space of all bounded linear functionals on X is called the *dual* of X and is denoted by X^*.

Based on Lagrange duality, the Lagrange Multiplier Theorem can be derived for convex optimization problems.

Theorem 2.4.4. *Consider the convex optimization problem in (2.24). Assume that there exists of a point $x_1 \in \Omega$ for which $G(x_1) <_P 0$ (i.e. $G(x_1)$ is an interior point of $N = -P$).*

Let

$$\mu_0 = \inf f(x) \qquad s.t.\ x \in \Omega, G(x) \leq_P 0, \tag{2.25}$$

and assume μ_0 is finite. Then there is an element $z_0^* \geq 0$ (with respect to the corresponding positive cone) in Z^* such that

$$\mu_0 = \inf_{x \in \Omega} \{f(x) + \langle G(x), z_0^* \rangle\}. \tag{2.26}$$

Furthermore, if the infimum is achieved in (2.25) by an $x_0 \in \Omega$, $G(x_0) \leq_P 0$, it is then achieved by x_0 in (2.26) and

$$\langle G(x_0), z_0^* \rangle = 0. \tag{2.27}$$

Without assumptions of convexity, local optimality can still be studied and necessary conditions for a local minimum can be derived from Lagrange duality theory, such as the Karush–Kuhn–Tucker condition.

Theorem 2.4.5 (Generalized Kuhn-Tucker Theorem)**.** *Let X be a vector space and Z a normed space having positive cone P. Assume that P contains an interior point. Let f be a Gateaux differentiable real-valued functional on X and G a Gateaux differentiable mapping from X into Z. Assume that the Gateaux differentials are linear in their increments. Suppose x_0 minimizes f subject to $G(x) \leq_P 0$ and that x_0 is a regular point of the inequality $G(x) \leq_P 0$. Then there is a $z_0^* \geq 0$ in Z^* such that the Lagrangian*

$$f(x) + \langle G(x), z_0^* \rangle,$$

is stationary at x_0; furthermore, $\langle G(x_0), z_0^* \rangle = 0$.

2.5 Estimation and Filtering

Pointwise Estimation

Given a *probability space* $(\Omega, \mathcal{F}, \mathcal{P})$, a *random variable* Y takes values in the sample space according to a probability density function (pdf) $p(y;\theta)$, where $\theta \in \mathbb{R}^n$ is an unknow parameter vector. The observed value y of Y constitutes the *data*. The task of an *estimation problem* is to estiamte parameter θ from the observed data. A real-valued function of Y is called an *estimator* of θ, which is denoted by $\hat{\theta}(Y)$.

The *mean squared error* (MSE) is a widely-used risk function to measure the quality of an estimator, which is defined by

$$\begin{aligned} MSE[\hat{\theta}(Y)] &:= \mathbb{E}[\|\hat{\theta}(Y) - \theta\|^2] \\ &= \mathbb{E}[\|\hat{\theta}(Y) - \mathbb{E}[\hat{\theta}(Y)]\|^2] + \|\mathbb{E}[\hat{\theta}(Y) - \theta]\|^2. \end{aligned} \tag{2.28}$$

The first term is called the *variance* error and the second term denotes the square of the *bias* error.

When we restrict our focus on unbiased estimators, the MSE is equal to the variance error. Then a lower bound on the variance of any unbiased estimator can be given by the inverse of the *Fisher Information Matrix* (FIM).

Definition 2.5.1. The *Fisher Information* that Y contains about the parameter θ is defiend by

$$I(\theta) = \mathbb{E}[(\frac{\partial}{\partial \theta} \log p(Y;\theta))(\frac{\partial}{\partial \theta} \log p(Y;\theta))^T]. \tag{2.29}$$

Theorem 2.5.2 (Information inequality)**.** *Suppose that $\hat{\theta}(Y)$ is an estimator of parameter θ and we have a family of density functions $\{p(y;\theta), \theta \in D_\theta\}$. Assume that*

(i) D_θ is an open interval.

(ii) Distributions $p(y;\theta)$ have common support, which is independent of θ.

(iii) $\frac{\partial}{\partial \theta} p(y;\theta)$ exists and is finite for all $\theta \in D_\theta$ and y in the common support.

Then the variance of any unbiased estimator $\hat{\theta}(Y)$ must satusfy

$$Var[\hat{\theta}(Y)] \geq I(\theta)^{-1}, \tag{2.30}$$

which is called the Cramér–Rao lower bound (CRLB). And an unbiased estimator whose variance reaches the CRLB is called an efficient estimator.

As is shown above, efficient estimators are uniformly optimal in the class of unbiased estimators. However, when we take into account all the estimators without any restrictions, it is impossible to find a uniformly optimal estimator that minimizes the risk at every value of θ. Then some weaker optimality properties can be considered. Denote $R(\theta, \hat{\theta}(Y))$ as some risk function, below are some common principles for deriving optimal biased estimators

(i) Minmax estimation: a *minimax estimator* is obtained by minimizing the worst-case risk

$$\sup_{\theta \in D_\theta} R(\theta, \hat{\theta}(Y)).$$

(ii) Average risk minimization: a *Bayes estimator* is computed by minimizing the average risk

$$r(\hat{\theta}, f_\Theta) = \int_{D_\theta} R(\theta, \hat{\theta}(Y)) f_\Theta(\theta) d\theta,$$

where $R(\theta, \hat{\theta}(Y))$ is some risk function and $f_\Theta(\theta)$ is a positive weighting function indicating how important it is to have a low risk for different values of θ.

Bayesian Theory and Bayesian Filtering

As for filtering problems, Bayesian inference has become an important branch in statistics inference, which has been applied successfully in statistical decision, detection and estimation, pattern recognition, and machine learning. In particular, one of the fundamental principles is the *Bayes rule*, where the posterior probability can be derived from a prior probability and a statistical model for the observed data.

Theorem 2.5.3 (Bayes rule)**.** *Given the prior $p(x)$ and likelihood $p(y|x)$, the posterior $p(x|y)$ is obtained by the product of prior and likelihood divided by a normalizing factor as*

$$p(x|y) = \frac{p(x)p(y|x)}{p(y)} = \frac{p(x)p(y|x)}{\int_X p(x)p(y|x)dx}. \tag{2.31}$$

Consider the general filtering problem

$$\begin{aligned} x_{n+1} &= f(x_n, u_n, w_n), \\ y_n &= g(x_n, u_n, v_n), \end{aligned} \tag{2.32}$$

where x_n is the state vector, y_n is the measurement vector, u_n is the system input, w_n and v_n represent the process noise and measurement noise respectively.

Assume that the states follow a first-order Markov process and the observations are independent of the given states. We denote Y_n as a sequence of observations, i.e. $Y_n = \{y_0, ..., y_n\}$. Based on the *Bayes rule*, the rule for *Bayesian Filtering* can be derived as following, where states are regarded as random variables instead of unknown constants.

$$p(x_n|Y_n) = \frac{p(x_n)p(Y_n|x_n)}{p(Y_n)} = \frac{p(y_n|x_n)p(x_n|Y_{n-1})}{p(y_n|Y_{n-1})}. \tag{2.33}$$

Bayesian optimal filtering is aimed to apply the Bayesian statistics and Bayes rule to sequential state estimation problem under some optimal sense. For instance, some potential criteria for measuring the optimality can be minimum mean-squared error (MMSE), maximum a posteriori (MAP), maximum likelihood (ML), min-max, and so on.

References

[1] Tamer. Basar and Geert Jan. Olsder. *Dynamic noncooperative game theory*. SIAM, 1999.

[2] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[3] Zhe Chen et al. Bayesian filtering: From kalman filters to particle filters, and beyond. *Statistics*, 182(1):1–69, 2003.

[4] Hassan K Khalil and Jessy W Grizzle. *Nonlinear systems*, volume 3. Prentice hall Upper Saddle River, NJ, 2002.

[5] Erich L Lehmann and George Casella. *Theory of point estimation*. Springer Science & Business Media, 2006.

[6] A. Lindquist and J. Sand. An introduction to mathematical system theory, lecture notes, 2002.

[7] David G Luenberger. *Optimization by vector space methods*. John Wiley & Sons, 1997.

[8] Harry L Trentelman, Anton A Stoorvogel, and Malo Hautus. *Control theory for linear systems*. Springer Science & Business Media, 2012.

To live is to choose. But to choose well,
you must know who you are and what you stand for,
where you want to go and why you want to get there.

– Kofi Annan

Part II:
Research Papers