Optimizing Networked Systems and Inverse Optimal Control

Optimizing Networked Systems and Inverse Optimal Control

HAN ZHANG

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

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

TRITA-SCI-FOU 2019:04
ISBN 978-91-7873-085-8

Akademisk avhandling som med tillstånd av Kungliga Tekniska Högskolan framlägges till offentlig granskning för avläggande av teknologie doktorsexamen, måndagen den 18 februari 2019 klockan 10.00 i Kollegiesalen, Brinellvägen 8, Kungliga Tekniska Högskolan, Stockholm.

Print: Universitetsservice US AB, Stockholm, 2019

To my family

Abstract

This thesis is concerned with the problems of optimizing networked systems, including designing a distributed energy optimal consensus controller for homogeneous networked linear systems, maximizing the algebraic connectivity of a network by projected saddle point dynamics. In addition, the inverse optimal control problems for discrete-time finite time-horizon Linear Quadratic Regulators (LQRs) are considered. The goal is to infer the Q matrix in the quadratic cost function using the observations (possibly noisy) either on the optimal state trajectories, optimal control input or the system output.

In Paper A, an optimal energy cost controller design for identical networked linear systems asymptotic consensus is considered. It is assumed that the topology of the network is given and the controller can only depend on relative information of the agents. Since finding the control gain for such a controller is hard, we focus on finding an optimal controller among a classical family of controllers which is based on the Algebraic Riccati Equation (ARE) and guarantees asymptotic consensus. We find that the energy cost is bounded by an interval and hence we minimize the upper bound. Further, the minimization for the upper bound boils down to optimizing the control gain and the edge weights of the graph separately. A suboptimal control gain is obtained by choosing $Q = 0$ in the ARE. Negative edge weights are allowed, meaning that "competitions" between the agents are allowed. The edge weight optimization problem is formulated as a Semi-Definite Programming (SDP) problem. We show that the lowest control energy cost is reached when the graph is complete and with equal edge weights. Furthermore, two sufficient conditions for the existence of negative optimal edge weights realization are given. In addition, we provide a distributed way of solving the SDP problem when the graph topology is regular.

In Paper B, a projected primal-dual gradient flow of augmented Lagrangian is presented to solve convex optimization problems that are not necessarily strictly convex. The optimization variables are restricted by a convex set with computable projection operation on its tangent cone as well as equality constraints. We show that the projected dynamical system converges to one of the saddle points and hence finding an optimal solution. Moreover, the problem of distributedly maximizing the algebraic connectivity of an undirected network by optimizing the "port gains" of each nodes is considered. The original SDP problem is relaxed into a nonlinear programming (NP) problem that will be solved by the aforementioned projected dynamical system. Numerical examples show the convergence of the aforementioned algorithm to one of the optimal solutions. The effect of the relaxation is illustrated empirically with numerical examples. A methodology is presented so that the number of iterations needed to converge is reduced. Complexity per iteration of the algorithm is illustrated with numerical examples.

In Paper C and D, the inverse optimal control problems over finite-time horizon for discrete-time LQRs are considered. The well-posedness of the inverse optimal control problem is first justified. In the noiseless case, when these observations of the optimal state trajectories or the optimal control input are exact, we analyze the identifiability of the problem and provide sufficient conditions for uniqueness of the solution. In the noisy case, when the observations are corrupted by additive zero-mean noise, we formulate the problem as an optimization problem and prove that the solution to this problem is statistically consistent. The following two scenarios are further considered: 1) the distributions of the initial state and the observation noise are unknown, yet the exact observations on the initial states and the noisy observations on the system output are available; 2) the exact observations on the initial states are not available, yet the observation noises are known to be white Gaussian and the distribution of the initial state is also Gaussian (with unknown mean and covariance). For the first scenario, we show statistical consistency for the estimation. For the second scenario, we fit the problem into the framework of maximum-likelihood and Expectation Maximization (EM) algorithm is used to solve this problem. The performance of the proposed method is illustrated through numerical examples.

Keywords: Networked systems, energy optimal consensus control, semi-definite programming, distributed optimization, inverse optimal control

Sammanfattning

Denna avhandling handlar om problem inom optimering av nätverkssystem, inklusive utformning av den energioptimala konsensusregulatorn för identiska nätverksbaserade linjära system, vilket maximerar den algebraisk konnektiviteten av ett nätverk på ett distribuerat sätt genom att använda projicerad sadelpunktdynamik. Dessutom beaktas de inversa optimala kontrollproblemen för tidsdiskreta system med ändlig tidshorisont och linjär kvadratisk regulator (LQR). Målet är att uppskatta Q -matrisen i den kvadratiska kostnadsfunktionen med hjälp av observationerna av antingen optimala tillståndsbanor, optimal styrsignal, eller utsignalen från systemet.

I papper A beaktas en optimal regulator design, optimal i energi-mening, för att uppnå konsensus hos identisk nätverksbaserad linjära system. Det antas att nätverkets graftopologi är given och att regulatorn enbart kan använda relativ information mellan agenterna. Eftersom det är svårt att hitta regulatorförstärkningen för en sådan regulator fokuserar vi på att hitta en optimal regulator i en klassisk grupp av regulatorer, nämligen den grupp som bygger på den algebraisk Riccati Equation (ARE) och som garanterar asymptotisk konsensus. Vi finner att energikostnaden är begränsad till att ligga inom ett intervall, varefter vi minimerar den övre gränsen av detta intervall. Minimeringen av denna övre gränsen resulterar i att optimera regulatorförstärkningen och grafens kantvikter separat. En suboptimal regulatorförstärkningen erhålls genom att välja $Q = 0$ i ARE. Negativa kantvikter är tillåtna, vilket betyder att "konkurrens" mellan agenterna är tillåtna. Kantsviktoptimeringsproblemet formuleras som ett semidefinit optimeringsproblem (SDP). Vi visar att den lägsta kontrollenergikostnaden uppnås när grafen är komplett och med lika stora vikter på alla kanter. Vidare ges två tillräckliga villkor för förekomsten av optimala kantvikter som är negativa. Dessutom tillhandahåller vi ett distribuerat sätt att lösa SDP problemet när graftopologin är regelbunden.

I papper B presenteras en algoritm för att lösa konvexa optimeringsproblem som inte nödvändigtvis är strikt konvexa. Denna bygger på en projicerat primal-dual-gradientflödesalgoritm på "the augmented Lagrangian". Optimeringsvariablerna är begränsade till en konvex mängd på vilken projektionsoperationer på tangentkonen kan beräknas. De är även begränsade till att uppfylla vissa likhetsbivillkor. Vi visar att det projicerade dynamiska systemet konvergerar till en av sadelpunkterna och därigenom hittar en optimal lösning. Dessutom beaktas problemet med distribuerad maximering av den algebraiska konnektiviteten av ett oriktat nätverk genom optimering av nodernas "regulatorförstärkning". Det ursprungliga SDP-problemet relaxeras till ett icke-linjärt programmeringsproblem (NP) som kommer att lösas med hjälp av det ovan nämnda projicerade dynamiska systemet. Numeriska exempel visar konvergensen av den ovannämnda algoritmen till en av de optimala lösningarna. Effekten av relaxeringen illustreras empiriskt med numeriska exempel. Ett tillvägagångssätt presenteras, som reducerar antalet iterationer som behövs för att nå konvergens. Komplexiteten per iteration av algoritmen illustreras med numeriska exempel.

I papper C och D beaktas de inversa optimala kontrollproblemen för tidsdiskreta system med ändlig tidshorisont och LQR-regulator. Först motiveras att problemet är välställt. I fallet då data inte innehåller brus, det vill säga när observationerna av den optimala tillståndsbanan eller den optimala styrsignalen är exakta, analyserar vi problemets identifierbarhet och ger tillräckliga villkor för att lösningen ska vara unik. I fallet då data innehåller brus, det vill säga när observationerna är korrumperade med additivt brus, formulerar vi problemet som ett optimeringsproblem och bevisar lösningen till detta problem är statistiskt konsistent. Följande två scenarier behandlas mer ingående: 1) då fördelningen av initialtillståndet och bruset på observationerna är okänt, men exakta observationerna av initialtillståndet samt brusiga observationerna är tillgängliga; 2) Exakta observationerna av initialtillståndet är inte tillgängliga, men bruset på observationerna är vitt Gaussiskt brus och fördelningen av initialtillståndet är också Gaussik men med okänd medelvärde och kovarians. För det första scenariot visar vi att uppskattningen är statistiskt konsistent. För det andra scenariot omformulerar vi problemet till ett Maximum likelihood problem, och Expectation maximization-algoritmen används för att lösa det senare. Prestandan av den föreslagna metoden illustreras genom numeriska exempel.

Acknowledgments

First of all, I would like to extend my sincere gratitude towards my supervisor, Professor Xiaoming Hu for giving me the opportunity to do my Ph.D. study here in KTH. I really appreciate the mathematical knowledge as well as the living principle you taught me. Thank you for not only stimulating my interest towards mathematics, but also turning me, an engineer who even failed in single variable calculus during bachelor study, into a qualified Ph.D. student in the Department of Mathematics. Without your strict and patient supervision, I would have not make it.

Secondly, I wish to express my gratitude towards my co-supervisor, Professor Elias Jarlebring, for the knowledge of matrix analysis he taught me. I am very grateful to your patience and generosity for my drop-in discussions.

In addition, I would like to thank my co-authors, Jieqiang Wei, Peng Yi and Jack Umenberger. It has always been a great pleasure and honor to work with you guys. I enjoyed every discussions with you and I really appreciate your help, ideas and contributions to this thesis work. Thanks a lot for your support.

Further, I am also grateful to all faculty members of our division. I thank Professor Krister Svanberg for giving me a solid understanding of optimization as well as the discussions that contributed a lot to this thesis work. I will never forget the good Friday afternoons when we played table tennis together. I also would like to thank Professor Anders Forsgren, Professor Johan Karlsson and Professor Per Enqvist for the help and discussions during my Ph.D. study.

To my colleagues, Silun Zhang, Yibei Li, Axel Ringh, Emil Ringh, Alexander Aurell, David Ek, Michelle Böck, Lovisa Engberg, Isabel Haasler, Doctor Yuecheng Yang, Doctor Johan Thunberg, Doctor Ximei Wang, I would like to thank you all for the help and fantastic atmosphere offered during my Ph.D. study. It has always been a good time during the Friday afternoon fika. I would like to also thank Axel Ringh especially for helping me translating the abstract of this thesis into Swedish.

Moreover, I would like to express my gratefulness to my friends in Stockholm, they are: Thomas Ohlson Timoudas, Marvin Lübcke, Jaakko Rekola, Victoria Rivera, Changil Kim, Cynthia Ho, Sebastian Östlund, Qian Wang. You are my family here in Stockholm and helped me overcome the depression and anxiety during my Ph.D. study.

Last but the most, I would like to thank my family for all the support you gave me. Thank you Jingfang for being a perfect girlfriend and support me during my study. Thank you for tolerating the long-distance relationship. Thanks, Jingfang, dad, mom, grandma, I love you all. And grandpa, I wish you would have seen that.

Stockholm, January 2019

Han Zhang

Table of Contents

“Yesterday is history, tomorrow is a mystery, but today is a gift. That is why it is called present. ”

— Master Oogway

Part I: Introduction and Preliminaries

1. Introduction

In this thesis, the problems of optimizing networked systems as well as inverse optimal control are studied. More precisely, this thesis aims mainly to answer the following three questions:

1. What is the most energy efficient controller design for a networked identical linear systems so that they can reach asymptotic consensus?
2. How can we distributedly maximize the robustness of a network under some budget constraints?
3. Given a series of state trajectories, control input or system output (all of them are possibly noisy) of a Linear Quadratic Regulator (LQR), what is its corresponding cost function?

The mathematical tools of systems theory, optimization theory, optimal control theory, graph theory as well as some knowledge of statistics are used for the problems analysis and algorithm development. In this chapter, a brief introduction for the background and motivation of the problems is included. In addition, the contributions and the contents of all the appended papers are previewed. The mathematical tools that are used will be introduced in the next chapter.

1.1 Background and Motivation

Optimal Energy Control for Linear Systems Consensus

Cooperative control for multi-agent systems has been intensively studied in the past decade. The idea of cooperative control is inspired greatly by the collective and self-emerging behavior in the nature, such as bird, fish flocks, ant colony, etc. The cooperative control methods usually imply a distributed manner, which makes the systems enjoy many advantages such as being robust and economic and have found their applications in the field of aerospace [7], robots [13], etc.

Among the research topics of cooperative control for multi-agent, consensus control has been intensively studied in the literature. The goal for consensus control is to let the states or the outputs of all agents become the same eventually by designing the feedback control laws that only depend on the information of the agent's and its neighbours'. In this thesis, a series of identical linear systems

$$\dot{x}_i = Ax_i + Bu_i, \quad i = 1, \cdots, N$$

are considered, where $x_i \in \mathbb{R}^n, u_i \in \mathbb{R}^m$. We would like to design a distributed feedback controller to let the linear systems reach asymptotic consensus, i.e., $x_i(t) = x_j(t), \forall i \neq j$ as $t \to \infty$. As mentioned above, in order to let the control be distributed, only the information of the agent's and its neighbours can be used. Moreover, in some scenarios, for instance, the agents are robots that move on a 2-dimensional plane, it is difficult to get its own state x_i (pose in the world frame) while it is relatively easy to get relative state (pose) $x_j - x_i$, $i \neq j$ by observing other agents. Hence, we would like to design a distributed feedback consensus controller that only uses relative information $x_j - x_i$, $i \neq j$. More precisely, we are interested in designing the controller that has the form

$$u_i = K \sum_{j \in \mathcal{N}_i} w_{ij}(x_i - x_j),$$

where $\mathcal{N}_i$ denotes the neighbour set of agent i and w_{ij} is the gain added individually between i and j.

The method of designing such distributed asymptotic consensus controller for linear systems is well-known [6]. However, a question that is worth being asked is that what is the most energy efficient way of doing that? Hence we look into an optimal control problem whose cost function reads $\int_0^\infty U^T U dt$, where $U = [u_1^T, \cdots, u_N^T]^T$ and try to minimize the cost function by choosing appropriate control gains K and w_{ij}'s.

Distributed Network Algebraic Connectivity Maximization

A network can be typically seen as a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$. The edge connectivity $e(\mathcal{G})$ is usually used to indicate the robustness of a graph, namely, the smallest number of edges that can be removed such that the graph becomes disconnected. Algebraic connectivity is first introduced by [3]. It is defined as the second smallest eigenvalue of the Laplacian matrix, denoted as $\lambda_2(L)$, and it holds that [4]

$$e(\mathcal{G}) \geq \lambda_2(L) \geq 2\left(1 - \cos\frac{\pi}{|\mathcal{V}|}\right) e(\mathcal{G}),$$

where $|\cdot|$ denotes the cardinality of a set. Hence algebraic connectivity $\lambda_2(L)$ also indicates the robustness of a graph.

Moreover, algebraic connectivity plays a crucial role in the networked systems. It indicates the convergence rate of the variables consensus in multi-agent systems [11]. Hence maximizing the algebraic connectivity of a graph is an important and fundamental problem.

In this thesis, we consider a communication network modeled as an undirected edge-weighted graph $\mathcal{G}(\mathcal{V}, \mathcal{E}, \mathcal{W})$ whose nodes $i \in \mathcal{V} = \{1, 2, \cdots, N\}$ are homogeneous "base stations" and can control their "port gains" $w_k^{(i)} \in \mathcal{W}$ as illustrated in Fig. 1.1. The "gain" on each link $k \in \mathcal{E}$ is the sum of the "port gains" $w_k^{(i)}$ and $w_k^{(j)}$, $(i, j) = k \in \mathcal{E}$ contributed by the two end nodes connected by the edge. It is assumed that each agent can only get access to the information of its neighbours as well as the information of itself's. The goal is to develop a method so that each base station can adjust its own "port gains" only according to its neighbours' information, the number of nodes N and the information belonging to itself, so that the algebraic connectivity of the total network is maximized. Since the graph considered here has edge weights, therefore the algebraic connectivity concerned here is the second smallest eigenvalue of the *weighted* Laplacian matrix. A similar formulation of algebraic connectivity maximization can be found in [5], yet their aim is not solving the problem distributedly.

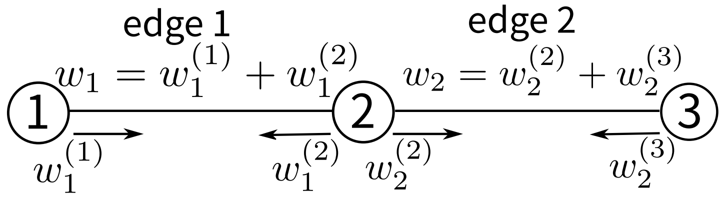

Figure 1.1: The edge-weighted graph.

Inverse Optimal Control for Linear Quadratic Regulators

In nature, after millions of years of evolution, it is regarded that animals tend to behave optimally. For instance, their gaits [1], the timing to hatch queens in wasp hives [2]. On the other hand, in the stock market, each bond trader tends to sell and buy stocks in an optimal way. All of these "optimal behaviors" can be seen as solutions to some specific optimal control problems. Although we can observe the solutions to the optimal control problem, for example in the prementioned cases, the gaits, the timing to morph queens and the amount of the trade that has been made by bond traders, the cost function of the corresponding optimal control problems remains unknown. The cost function is of great interest because it can help one to understand how do the individuals make decisions and take actions. Moreover, once the cost functions of the corresponding optimal control problems were known, one would be able to predict the behaviors of the individuals.

Speaking more mathematically, the problem of inverse optimal control is first proposed by [8] and aims to "reverse engineer" the cost function, given observations of optimal trajectories or control inputs, for known system dynamics in constrast to classical "forward" optimal control problems. In this thesis, the problem of inverse optimal control for the discrete-time Linear Quadratic Regulators (LQRs) over finite-time horizons is concerned. We aim to find the parameters in the quadratic objective funtion given the discrete-time linear system dynamics and (possibly noisy) observations of the optimal trajectory or control input. We first justify the well-posedness for the problem and consider the problem under four different scenarios:

1. The observations on the optimal trajectories of the full-state variable $x_{1:N}$ or the optimal control input $u_{1:N-1}$ and the initial value x_1 are available and exact;

2. The observations on inital value x_1 are exact while the observations on the optimal trajectories of the full-state variable $x_{2:N}$ or the optimal control input $u_{1:N-1}$ are available but contaminated by some zero-mean, finite variance noises;

3. The observations on inital value x_1 are exact while the observations of the system output $y_t = Cx_t, t = 2 : N$ are available but contaminated by some zero-mean, finite variance noises;

4. The observations of the system output $y_t = Cx_t, t = 1 : N$ are available but contaminated by white Gaussian noise, while the intial value x_1 is distributed according to some unknown Gaussian distribution.

1.2 Summary and Contribution of the Appended Papers

Paper A: Consensus Control for Linear Systems with Optimal Energy Cost

Summary: In this paper, an optimal energy controller that depends only on the relative information between the agents is constructed. Finding the optimal solution for the problem is hard and we focus on finding an optimal controller among a classical family of controller designs that is based on the Algebraic Riccati Equation (ARE) and guarantees asymptotic consensus. Such controller design mainly involves determining the Q matrix in ARE and the edge weights of the graph. But finding the optimal solution among this family of controller is still hard. Nevertheless, we find that the energy cost is bounded by an interval and hence we minimize the upper bound. A suboptimal control gain is obtained by choosing $Q = 0$ in the ARE; the edge weights of the graph is optimized by minimizing the so-called "synchronizability". The controller that we designed enjoys several favourable properties:

1. The controller coincides with the optimal control in [12] when the graph is complete. It has been pointed out in [12] that any other distributed control laws constructed by Laplacian matrices that do not correspond to complete graphs with equal edge weights are suboptimal.

2. When optimizing the edge weights, "competitions" are allowed between the connected agents. By doing so, the feasible region of the optimization problem is enlarged, and hence a smaller control energy cost might be obtained. We offer two sufficient conditions for when "competitions" will happen between agents. These two conditions help to determine whether the two agents will compete if we add a link between them based on the old optimal solution.

3. When the graph topology is regular, namely, every node has the same number of neighbours, the controller can be calculated in a distributed manner.

Contribution: The problem was initiated by the co-author Xiaoming Hu. The theoretical results are developed by the first author under the supervision of the co-author Xiaoming Hu.

This paper is referred to *Zhang H, Hu X. Consensus control for linear systems with optimal energy cost[J]. Automatica, 2018, 93: 83-91.*

Paper B: Projected Primal-Dual Gradient Flow of Augmented Lagrangian with Application to Distributed Maximization of the Algebraic Connectivity of a Network

Summary: In this paper, an algorithm that is based on projected saddle point dynamics is presented to solve a convex optimization problem that is not necessarily strictly convex. As a supplement to [10] and its conference version [9], we propose a novel analysis line regarding the convergence of the dynamical system to reach comparable results. Moreover, the problem of distributedly maximizing the algebraic connectivity of an undirected network by adjusting the "port gains" of each nodes is considered. The problem is formulated as a Semi-Definite Programming (SDP) Problem. In order to solve it distributedly, we first re-formulate the original problem into another SDP problem and prove the equivalence. Then the problem is relaxed into a nonlinear programming problem. We also show how well the result would be after the relaxation compared to the original problem. Meanwhile, the nonlinear programming problem is not strictly convex, hence the projected saddle point dynamics method mentioned in the first part of the paper is applied. Numerical experiments are used to illustrate the performance as well as the effect of different choice of variables in the algorithm.

Contribution: This work is an extension from the first paper and utilizes the idea of "consensus" when developing the distributed algorithm. The theoretical results are mainly joint work with the co-author Jieqiang Wei and the modifications of the paper during the reviewing process are done under the discussions with the co-authors Jieqiang Wei, Peng Yi and the supervision of Xiaoming Hu. The numerical simulations are coded by the author of this thesis.

This paper is referred to *Zhang H, Wei J, Yi P, et al. Projected primal–dual gradient flow of augmented Lagrangian with application to distributed maximization of the algebraic connectivity of a network[J]. Automatica, 2018, 98: 34-41.*

Paper C: Inverse Quadratic Optimal Control for Discrete-Time Linear Systems

Summary: In this paper, the inverse optimal control problem for discrete-time LQRs over finite-time horizons is considered. The goal is to infer the parameters that define the quadratic cost function by using the observations of the optimal trajectories or the optimal control inputs of a linear time-invariant system. The well-posedness of the inverse optimal control problem is first justified. In the noiseless case, when these observations are exact, we analyze the identifiability of the problem and provide sufficient conditions for uniqueness of the solution. In the noisy case, when the observations are corrupted by additive zero-mean noise, we formulate the problem as an optimization problem and prove the statistical consistency of the problem later. The performance of the proposed method is illustrated through numerical examples.

Contribution: The idea of this work is initiated by the co-author Xiaoming Hu. The theoretical results are worked out through discussions with the co-author Jack Umenberger as well as under the supervision of Xiaoming Hu. The numerical simulations in this paper utilize partially some code of the previous work by the co-author Jack Umenberger.

This paper has been submitted to *Automatica*.

Paper D: Inverse Optimal Control for Finite-Horizon Discrete-Time Linear Quadratic Regulator Under Noisy Output

Summary: In this paper, the problem of inverse optimal control for finite-horizon discrete-time LQRs is considered. The goal in this paper is similar to that of Paper C, nevertheless, in this paper, the problem of inverse optimal control is considered in two scenarios:

1. the distributions of the initial state and the observation noise are unknown, yet the exact observations on the initial states and the noisy observations on system output $y_t = Cx_t$ are available (while in Paper C, the full-state can be observed);
2. the exact observations on the initial states are not available, yet the observation noises are known to be white Gaussian and the distribution of the initial state is also Gaussian (with unknown mean and covariance). Only the noisy observations on the system output $y_t = Cx_t$ are available (while in Paper C, the full-state can be observed).

For the first scenario, we show the statistical consistency for the estimation. For the second scenario, we fit the problem into the framework of maximum-likelihood and Expectation Maximization (EM) algorithm is used to solve this problem. The performance for the estimations is shown by numerical examples.

Contribution: This work is a natural extension of the previous paper. We would like to extend the problem into a more general setting, namely, we can only get noisy observations of the system output. The theoretical result is worked out under the supervision of the co-author Xiaoming Hu.

This paper has been submitted to *IEEE 58th Conference on Decision and Control (CDC).*

References

[1] R McN Alexander. The gaits of bipedal and quadrupedal animals. *The International Journal of Robotics Research*, 3(2):49–59, 1984.

[2] R McNeill Alexander. *Optima for animals.* Princeton University Press, 1996.

[3] Miroslav Fiedler. Algebraic connectivity of graphs. *Czechoslovak mathematical journal*, 23(2):298–305, 1973.

[4] Miroslav Fiedler. Laplacian of graphs and algebraic connectivity. *Banach Center Publications*, 25(1):57–70, 1989.

[5] Frank Göring, Christoph Helmberg, and Markus Wappler. Embedded in the shadow of the separator. *SIAM Journal on Optimization*, 19(1):472–501, 2008.

[6] Xiaoming Hu and Anders Lindquist. Geometric control theory, lecture notes. KTH, 2012.

[7] Steven Hughes. Formation design and sensitivity analysis for the magnetospheric multiscale mission (mms). In *AIAA/AAS Astrodynamics Specialist Conference and Exhibit*, page 7357, 2012.

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

[9] Simon K Niederländer, Frank Allgöwer, and Jorge Cortés. Exponentially fast distributed coordination for nonsmooth convex optimization. In *2016 IEEE 55th Conference on Decision and Control (CDC)*, pages 1036–1041. IEEE, 2016.

[10] Simon K Niederländer and Jorge Cortés. Distributed coordination for nonsmooth convex optimization via saddle-point dynamics. *arXiv preprint arXiv:1606.09298*, 2016.

[11] Reza Olfati-Saber, J Alex Fax, and Richard M Murray. Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE*, 95(1):215–233, 2007.

[12] Johan Thunberg and Xiaoming Hu. Optimal output consensus for linear systems: a topology free approach. *Automatica*, 68:352–356, 2016.

[13] Matthew Turpin, Nathan Michael, and Vijay Kumar. Decentralized formation control with variable shapes for aerial robots. In *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, pages 23–30. IEEE, 2012.

2. Preliminaries

In this chapter, an overview of some concepts, theories as well as some algorithms that will be used in this thesis will be provided. The purpose of this chapter is to offer a brief introduction of the mathematical tools and knowledge to the readers who might not be familar with the topics and field.

This chapter is organized as follows: in Section 2.1, we will briefly introduce the model for dynamical systems, its solutions as well as its extension. In Section 2.2, some knowledge about convex optimization including saddle points, optimality conditions and Semi-Definite Programming (SDP) problem is included. In Section 2.3, a brief review of the optimality conditions for discrete-time optimal control problem is included, especially for the discrete LQ optimal control problem. In Section 2.4, we briefly introduce some knowledge of graph theory that will be used in the remainder of the thesis. Some knowledge about statistics, including the definition of convergence of random variables, some useful theorems as well as introductions of Expectation Maximization (EM) algorithm and fixed point smoothing are included in Section 2.5.

2.1 Dynamical Systems Model

In the field of systems theory, it is conventional to consider continuous-time and discrete-time input-output dynamical systems:

$$\begin{aligned} \dot{x} &= f(t, x, u), \\ y &= h(t, x, u), \end{aligned} \tag{2.1.1}$$

where $x \in \mathbb{R}^n$ is the state variable, $u \in \mathbb{R}^m$ is the control input, while $y \in \mathbb{R}^p$ is the system output; and

$$\begin{aligned} x_{t+1} &= f(t, x_t, u_t), \\ y_t &= h(t, x_t, u_t), \end{aligned} \tag{2.1.2}$$

where here $t \in \mathbb{Z}$ and $x_t \in \mathbb{R}^n$, $u_t \in \mathbb{R}^m$, $y_t \in \mathbb{R}^p$, $\forall t$ is referred as the same name as (2.1.1).

The continuous-time system (2.1.1) is defined through Ordinary Differential Equations (ODEs), of which the solution may not exist globally. Here we include some classical ODE theory.

Definition 2.1.1. Consider the following ODE

$$\dot{x} = f(t, x), \tag{2.1.3}$$

where $f : D \subset \mathbb{R} \times \mathbb{R}^n \mapsto \mathbb{R}^n$ and D is open. A point $(t_0, x_0) \in D$ is called initial value point. A *(classical)* solution to an initial value problem at (t_0, x_0) is a differentiable function $x(t)$ that is a solution to (2.1.3) and satisfies $x(t_0) = x_0$.

Theorem 2.1.2. *Consider the initial value problem*

$$\dot{x} = f(t, x), \; x(t_0) = x_0. \tag{2.1.4}$$

If f is continuous on $D = \{(t, x) | t_0 \leq t \leq T, -\infty < x < \infty\}$ and f is Lipschitz continuous with respect to x, then the initial value problem (2.1.4) *has a unique solution.*

Sometimes, the condition of Lipschitz continuity does not hold and the function f can be even discontinuous. For such systems, more general solutions need to be defined.

Definition 2.1.3. A solution $x(t)$ to the initial value problem (2.1.4) is called a *Carathéodory* solution if it is continuous, $x(t_0) = x_0$ and satisfies the ODE in (2.1.4) for almost all t.

Roughly speaking, Carathéodory solutions are absolutely continuous curves that satisfy

$$x(t) = x_0 + \int_{t_0}^{t} f(\tau, x(\tau)) d\tau, \; t > t_0,$$

where the integral is the Lebesgue integral [2].

Now, let us consider a class of discrete-time input-output system

$$\begin{aligned} x_{t+1} &= f(x_t) + B(x_t) u_t, \\ y_t &= g(x_t), \end{aligned} \tag{2.1.5}$$

where $x_t \in \mathbb{R}^n$, $u_t \in \mathbb{R}^m$, $y_t \in \mathbb{R}^p$. The functions $f : \mathbb{R}^n \mapsto \mathbb{R}^n$, $B : \mathbb{R}^n \mapsto \mathbb{R}^{n \times m}$, $g : \mathbb{R}^n \mapsto \mathbb{R}^p$ are smooth. Denote the composite function of f and g as $g \circ f$. Denote

$$f^i(x) = f^{i-1} \circ (f(x)), \quad f^0(x) = x,$$

then the relative degree of (2.1.5) can be defined as [8]:

Definition 2.1.4. The input-output system has relative degree $(r_1, \cdots, r_p)$ if

$$\frac{\partial}{\partial u} g_j \circ f^i \left(f(x) + B(x) u \right) = 0, \quad 0 \leq i \leq r_p - 2, j = 1, \cdots, m,$$

and the $p \times m$ matrix

$$\begin{bmatrix} \frac{\partial}{\partial u} g_1 \circ f^{r_1 - 1} \left(f(x) + B(x) u \right) \\ \vdots \\ \frac{\partial}{\partial u} g_p \circ f^{r_p - 1} \left(f(x) + B(x) u \right) \end{bmatrix}$$

has full column rank for all $x \in \mathbb{R}^n$.

In addition, perhaps one of the most well-studied systems is linear invariant systems, namely,

$$\dot{x} = Ax + Bu, \quad y = Cx_t, \tag{2.1.6}$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $y \in \mathbb{R}^p$ and its discrete version

$$x_{t+1} = \hat{A} x_t + \hat{B} u_t, \quad y_t = \hat{C} x_t, \tag{2.1.7}$$

where x_t, u_t and y_t have the same dimension as the continuous time case (2.1.6). The solution to (2.1.6) given the initial value $x(0) = x_0$ is

$$x(t) = e^{At}x_0 + \int_0^t e^{A(t-\tau)}Bu(\tau)d\tau. \tag{2.1.8}$$

Based on (2.1.8), if (2.1.7) is a discretization of (2.1.6), then it holds that $\hat{A} = e^{A\Delta t}$, $\hat{B} = \int_0^{\Delta t} e^{A\tau}Bd\tau$, $\hat{C} = C$, where Δt is the sampling period.

Two fundamental properties of control system models are controllability and observability. Roughly speaking, controllability indicates that whether there exists a control input that can drive the system from an arbitrary initial state x_0 to x_T in the time interval $[0, T]$ (or within T time-steps in the discrete-time case); observability indicates that if it is possible to distinguish between two different initial values by observing the output y and u within the time interval $[0, T]$ (or T time steps in the discrete-time case). For both continuous-time (2.1.6) and discrete-time (2.1.7) linear invariant systems, their controllability and observability can be determined by the following theorem.

Theorem 2.1.5. *For both continuous-time* (2.1.6) *and discrete-time linear system* (2.1.7)*, the systems are controllable if and only if*

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

has full row rank; the systems are observable if and only if

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

has full column rank.

Lyapunov equation

$$A^TP + PA = -Q \tag{2.1.9}$$

plays an important role in linear systems control, where Q is positive semidefinite. If A is Hurwitz, then the unique solution to (2.1.9) has the explicit expression as $P = \int_0^\infty e^{A^Tt}Qe^{At}dt$.

In addition, later in this thesis, a so-called "projected dynamical system" will be considered. We first define the projection operator. Assume $K \subset \mathbb{R}^n$ is a closed and convex set, the projection of a point x to the set K is defined as $P_K(x) = \arg\min_{y \in K} \|x - y\|$. For $x \in K$, $v \in \mathbb{R}^n$, the projection of the vector v at x with respect to K is defined as [5],[1]:

$$\Pi_K(x, v) = \lim_{\delta \to 0} \frac{P_K(x + \delta v) - x}{\delta} = P_{T_K(x)}(v)$$

where $T_K(x)$ denotes the tangent cone of K at x. We denote the boundary and the interior of K as ∂K and $int(K)$, respectively. The set of inward normals of K at x is defined as

$$n(x) = \{\gamma \mid \|\gamma\| = 1, \langle \gamma, x - y \rangle \le 0, \forall y \in K\},$$

where $\|\cdot\|$ denotes l_2-norm and $\langle \cdot, \cdot \rangle$ denotes the inner-product and the projection $\Pi_K(x, v)$ fulfills the following theorem [5]:

Theorem 2.1.6. *If $x \in int(K)$, then $\Pi_K(x,v) = v$; if $x \in \partial K$, then $\Pi_K(x,v) = v + \beta(x)n^*(x)$, where $n^*(x) = \arg\max_{n \in n(x)} \langle v, -n \rangle$ and $\beta(x) = \max\{0, \langle v, -n^*(x) \rangle\}$.*

Further, suppose $F : K \mapsto \mathbb{R}^n$ be a vector field, the projected dynamical system is given by

$$\dot{x} = \Pi_K(x, F(x)). \tag{2.1.10}$$

Note that the right hand side of (2.1.10) can be discontinuous on ∂K. Hence given an initial value $x_0 \in K$, the system does not necessarily have a classical solution. However, under certain conditions, it is possible to get an "*almost classical*" solution to the projected dynamical system (2.1.10).

Theorem 2.1.7. *[5] Consider the projected dynamical system* (2.1.10). *If $F(x)$ is Lipschitz continuous, then it has a unique Carathéodory solution that continuously depends on the initial value.*

2.2 Convex Optmization

In this section, we will present some brief knowledge about optimization that is useful in the remainder of the thesis.

Suppose $D \subset \mathbb{R}^n$ is a convex set.

Definition 2.2.1. A function $f : D \mapsto \mathbb{R}$ is said to be convex if

$$f(\alpha x_1 + (1-\alpha)x_2) \leq \alpha f(x_1) + (1-\alpha) f(x_2), \quad \forall x_1, x_2 \in D, \forall \alpha \in [0,1].$$

f is said to be *strictly* convex if

$$f(\alpha x_1 + (1-\alpha)x_2) < \alpha f(x_1) + (1-\alpha) f(x_2), \quad \forall x_1 \neq x_2 \in D, \forall \alpha \in (0,1).$$

f is said to be concave if $-f$ is convex.

Proposition 2.2.2. *If $f : D \mapsto \mathbb{R}$ is differentiable, then f is convex if and only if*

$$f(y) \geq f(x) + \nabla f(x)^T (y - x)$$

holds for all $x, y \in D$. It is concave if and only if

$$f(y) \leq f(x) + \nabla f(x)^T (y - x)$$

holds for all $x, y \in D$.

Definition 2.2.3. A differentiable function $f : D \mapsto \mathbb{R}$ is called strongly convex with parameter $m > 0$ if

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq m \|x - y\|^2, \quad \forall x, y \in D,$$

where $\langle \cdot, \cdot \rangle$ is the inner-product and $\| \cdot \|$ is the natural norm associated with the inner-product.

Convex optimization is one of the most well-studied problems in the field of optimization. A convex problem can be expressed as

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f_0(x) \\ \text{s.t.} \quad & f_i(x) \leq 0, \quad i = 1, \cdots, m \\ & Ax - b = 0, \end{aligned} \tag{2.2.1}$$

where $f_i : D_i \mapsto \mathbb{R}$ are convex for all i and $A \in \mathbb{R}^{p \times n}$. Denote $D = \cap_{i=0}^m D_i$. Suppose f_i's are all smooth for all i and D is closed.

For problem (2.2.1), it is conventional to consider its *Lagrangian*. The Lagrangian $\mathscr{L} : D \times \mathbb{R}^m \times \mathbb{R}^p \mapsto \mathbb{R}$ associated with (2.2.1) is defined as

$$\mathscr{L}(x, \lambda, v) = f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + v^T (Ax - b). \tag{2.2.2}$$

Note that $\mathscr{L}(x, \lambda, v)$ is convex with respect to x and concave with respect to λ and v, hence we can define the saddle points for the Lagrangian.

Definition 2.2.4. If it holds for every $x \in D$, $\lambda \in \mathbb{R}_+^m$ and $v \in \mathbb{R}^p$ that

$$\mathscr{L}(x^*, \lambda, v) \leq \mathscr{L}(x^*, \lambda^*, v^*) \leq \mathscr{L}(x, \lambda^*, v^*), \tag{2.2.3}$$

then (x^*, λ^*, v^*) is a *saddle point* of the Lagrangian $\mathscr{L}(x, \lambda, v)$.

Once the Lagrangian is obtained, sometimes it is convenient to work with its dual problem, namely,

$$\sup_{\lambda \geq 0, v} \quad g(\lambda, v) := \inf_{x \in D} \mathscr{L}(x, \lambda, v). \tag{2.2.4}$$

Suppose x^* is optimal to the primal problem (2.2.1) and the optimum of the primal problem is attained, then it holds that $\sup_{\lambda \geq 0, v} g(\lambda, v) \leq f_0(x^*)$.

Before introducing the optimality conditions to (2.2.1), we would like to introduce the *Slater's condition*. It provides a sufficient condition for strong duality, namely $g(\lambda^*, v^*) = f_0(x^*)$.

Theorem 2.2.5. *For the convex optimization problem* (2.2.1)*, if there exists an* $x \in$ *relint*(D) *such that* $f_i(x) < 0$, $i = 1, \cdots, m$, $Ax = b$, *where relint*(D) *denotes the relative interior of* D*, then the optimum of the dual problem* (2.2.4) *is attained and strong duality holds.*

Now we introduce some optimality conditions to (2.2.1).

Theorem 2.2.6. (x^*, λ^*, v^*) *is a primal-dual optimal solution pair if and only if it is a saddle point of the Lagrangian* $\mathscr{L}(x, \lambda, v)$.

Theorem 2.2.7. *For convex optimization problem* (2.2.1)*,* (x^*, λ^*, v^*) *is a primal-dual optimal solution pair if and only if*

$$\begin{aligned} & \nabla f_0(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) + v^{*T} (Ax^* - b) = 0 \\ & f_i(x^*) \leq 0, \ Ax^* - b = 0, \ \lambda_i^* \geq 0, \ \lambda_i^* \nabla f_i(x^*) = 0. \end{aligned} \tag{2.2.5}$$

(2.2.5) is referred as Karush-Kuhn-Tucker (KKT) conditions. Note that although Theorem 2.2.6 and Theorem 2.2.7 provide sufficient and necessary conditions for primal-dual optimal solution pair to (2.2.1), it does not mean that there exists such (x^*, λ^*, v^*) that satisfies (2.2.3) and (2.2.5). If Slater's conditions hold, then the existence of such (x^*, λ^*, v^*) is guaranteed.

Proposition 2.2.8. *Suppose the optimal value f_0^* of* (2.2.1) *can be attained, then the set of optimal solutions to* (2.2.1) *is closed.*

To see that, denote $\Omega = \{x \in D \mid f_i(x) \leq 0,\ i = 1, \cdots, m,\ Ax - b = 0\}$ and the set of optimal solutions can be expressed as $\Omega \cap \{x \in D \mid f_0(x) \leq f_0^*\}$. Therefore, the set of optimal solutions is closed since the intersection of closed sets is closed.

In addition, SDP is a very important topic in the field of optimization. The optimization problem is defined in the space of symmetric matrices $\mathbb{S}^n$. Denote the cone of positive semi-definite matrix as $\mathbb{S}^n_+$. $G_1 \preceq G_2$ and $H_1 \succeq H_2$ means that $G_2 - G_1$ and $H_1 - H_2$ are positive semi-definite matrices. $tr(\cdot)$ denotes matrix trace. If $A, B \in \mathbb{R}^{n\times n}$, then $tr(A^T B) = \langle A, B\rangle$ is defined as the inner-product in the Hilbert space $\mathbb{R}^{n\times n}$.

A standard linear SDP takes the form of

$$\begin{aligned} \underset{X}{\text{minimize}} \quad & tr(CX) \\ \text{s.t.} \quad & tr(A_i X) = b_i, \quad i = 1, \cdots, m \\ & X \succeq 0, \end{aligned} \tag{2.2.6}$$

where X, C, A_i, $i = 1, \cdots, m$ are all symmetric matrices. The KKT condition of (2.2.6) reads

$$\begin{aligned} & C - \sum_{i=1}^m v_i A_i - S = 0, \quad tr(SX) = 0, \quad S \succeq 0, \\ & tr(A_i X) = b_i, \quad i = 1, \cdots, m, \quad X \succeq 0. \end{aligned}$$

And the dual problem of (2.2.6) reads

$$\begin{aligned} \underset{\{y_i\}_{i=1}^m, Z}{\text{maximize}} \quad & \sum_{i=1}^m b_i y_i \\ \text{s.t.} \quad & \sum_{i=1}^m y_i A_i + Z = C \\ & Z \succeq 0. \end{aligned} \tag{2.2.7}$$

Denote $\mathcal{M}_r = \{X \in \mathbb{S}^n \mid rank(X) = r\}$. The tangent space of $\mathcal{M}_r$ at X reads

$$\mathcal{T}_X = \left\{ Q \begin{bmatrix} U & V \\ V^T & 0 \end{bmatrix} Q^T \mid U \in \mathbb{S}^r, V \in \mathbb{R}^{r\times(n-r)} \right\},$$

where $X = Qdiag(\lambda_1, \cdots, \lambda_r)Q^T$, $Q^T Q = I$. Then the *nondegeneracy* of feasible solutions for the primal problem (2.2.6) can be defined as follows.

Definition 2.2.9. X is *primal nondegenerate* if it is primal feasible and $\mathcal{T}_X + \mathcal{N} = \mathbb{S}^n$, where $\mathcal{N} = \{Y \in \mathbb{S}^n \mid tr(A_i Y) = 0, i = 1, \cdots, m\}$ and the sum between $\mathcal{T}_X$ and $\mathcal{N}$ is in the sense of Minkowski.

Based on Definition 2.2.9, the following theorem provides a sufficient condition for the uniqueness of the solution for the dual problem (2.2.7).

Theorem 2.2.10. *Let X be primal nondegenerate and optimal, then there exists a unique optimal dual solution for* (2.2.7).

Note that $\mathcal{T}_X + \mathscr{N} = \mathbb{S}^n$ is equivalent to $\mathcal{T}_X^\perp \cap \mathscr{N}^\perp = \{0\}$, where $\mathcal{T}_X^\perp$ and $\mathscr{N}^\perp$ denote the orthogonal complements of $\mathcal{T}_X$ and $\mathscr{N}$ respectively, namely,

$$\mathcal{T}_x^\perp = \left\{ Q \begin{bmatrix} 0 & 0 \\ 0 & W \end{bmatrix} Q^T \mid W \in \mathbb{S}^{n-r} \right\}, \quad \mathscr{N}^\perp = span\{A_i\}.$$

On the other hand, SDPs are usually solved by the interior point method. Before introducing the interior point method, we would like to introduce two operators that would be useful in the method. For any symmetric matrix $G \in \mathbb{S}^n$, $svec(G)$ is defined as $svec(G) = [G_{11}, \sqrt{2}G_{21}, \cdots, \sqrt{2}G_{n1}, G_{22}, \sqrt{2}G_{32}, \cdots, \sqrt{2}G_{n1}, \cdots, G_{nn}]^T$. It follows from the above definition that $tr(DG) = svec(D)^T svec(G)$, $\forall D, G \in \mathbb{S}^n$. The *Symmetric Kronecker Product* between two matrices R_1 and R_2 is defined by the following identity

$$(R_1 \otimes_s R_2) svec(G) = \frac{1}{2} svec(R_2 G R_1^T + R_1 G R_2^T),$$

where $G \in \mathbb{S}^n$. Note that R_1 and R_2 are not required to be symmetric. $mat(\cdot)$ is the inverse operation of $svec(\cdot)$.

Now we are ready to introduce the interior point method for solving SDP problems. First, we introduce log-barrier function $-\rho \ln \det X$ to (2.2.6) and get penalized barrier problems as

$$\begin{aligned} \underset{X}{\text{minimize}} \quad & tr(CX) - \rho \ln \det X \\ \text{s.t.} \quad & tr(A_i X) = b_i, \quad i = 1, \cdots, m, \quad (X \succ 0). \end{aligned} \tag{2.2.8}$$

Therefore, the central path reads

$$\begin{bmatrix} \sum_{i=1}^m y_i A_i + Z - C \\ tr(A_1 X) - b_1 \\ \cdots \\ tr(A_m X) - b_m \\ XZ - \rho I \end{bmatrix} = 0.$$

Since XZ is in general not symmetric, to make the algorithm robust, we can rewrite the last equation in the central path as $XZ + ZX = 2\rho I$. Compute the Newton step on the central path and it follows that

$$\sum_{i=1}^m \Delta y_i A_i + \Delta Z = C - \sum_{i=1}^m y_i A_i - Z, \tag{2.2.9}$$

$$tr(A_i \Delta X) = b_i - tr(A_i X), \tag{2.2.10}$$

$$\Delta X Z + Z \Delta X + X \Delta Z + \Delta Z X = 2\rho I - (XZ + ZX) \tag{2.2.11}$$

Symmetric vectorize (2.2.9) and (2.2.11) on both hand sides and use the property of symmetric vectorization, we can write (2.2.9)-(2.2.11) into the following compact form

$$\begin{bmatrix} 0 & \mathscr{A}^T & I \\ \mathscr{A} & 0 & 0 \\ \mathscr{E} & 0 & \mathscr{F} \end{bmatrix} \begin{bmatrix} svec(\Delta X) \\ \Delta y \\ svec(\Delta Z) \end{bmatrix} = \begin{bmatrix} r_d \\ r_p \\ r_c \end{bmatrix} \tag{2.2.12}$$

where

$$\mathscr{A} = \begin{bmatrix} svec(A_1)^T \\ \vdots \\ svec(A_m)^T \end{bmatrix}, \quad \mathscr{E} = Z \otimes_s I, \quad \mathscr{F} = X \otimes_s I$$

and $r_p = b - \mathscr{A}x$, $r_d = C - Z - mat(\mathscr{A}^T y)$, $r_c = svec\left(\rho I - \frac{1}{2}(XZ + ZX)\right)$. Using Gauss elimination, (2.2.12) can further be simplified as

$$\begin{bmatrix} -\mathscr{F}^{-1}\mathscr{E} & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} svec(\Delta X) \\ \Delta y \end{bmatrix} = \begin{bmatrix} r_d - \mathscr{F}^{-1} r_c \\ r_p \end{bmatrix}$$

and we can solve ΔX, Δy and ΔZ respectively. Once the search direction $(\Delta X, \Delta y, \Delta Z)$ is obtained, the step-length α_k is computed by

$$\alpha_k = \max\{\alpha \in (0,1] \mid X_k + \alpha \Delta X \succ 0, Z_k + \alpha \Delta Z \succ 0\}$$

and update $X_{k+1} = X_k + \alpha_k \Delta X$, $y_{k+1} = y_k + \alpha_k \Delta y$ and $Z_{k+1} = Z_k + \alpha_k \Delta Z$. Shrink the parameter $\rho \leftarrow \sigma\rho$ when necessary (note $\sigma \in (0,1)$).

Note that during the iterations of the interior point method, a system of linear equations need to be solved in order to get the Newton search direction. More generally speaking, suppose N agents would like to solve a system of linear equations $Ax = b$, where A and b can be row-segmented as $A = [A_1^T, \cdots, A_N^T]^T$ and $b = [b_1^T, \cdots, b_N^T]^T$. Further, suppose each agent i knows matrix A_i and vector b_i, then the system of linear equation can be solved distributedly using the following iteration [4]:

$$x_i(t+1) = x_i(t) - \frac{1}{m_i(t)} P_i \left(m_i(t) x_i(t) - \sum_{j \in \mathcal{N}_i(t)} x_j(t) \right),$$

where $m_i(t)$ is the number of neighbours of agent i at time step t, $\mathcal{N}_i(t)$ is the neighbour set of agent i at time step t and $P_i(\cdot)$ is the projection operator that projects on the kernel space of A_i. It has been proved in [4] that the algorithm converges exponentially to a solution of $Ax = b$ if the time-dependent graph is jointly strongly connected.

On the other hand, it is also possible to relax the SDP problem into a standard nonlinear optimization problem. For a symmetric matrix X being positive semidefinite, it is the same as saying $\lambda_{\max}(-X) \leq 0$, where $\lambda_{\max}(\cdot)$ denotes the largest eigenvalue. Proposed by [6], the largest eigenvalue of a symmetric matrix $X \in \mathbb{S}^n$ can be approximated by a function $f_\varepsilon : \mathbb{S}^n \mapsto \mathbb{R}$ and has the form

$$f_\varepsilon(X) = \varepsilon \ln tr(e^{X/\varepsilon}) = \varepsilon \ln \left[\sum_{i=1}^{N} e^{\lambda_i(X)/\varepsilon} \right]$$

and its derivative with respect to X reads

$$\nabla_X f_\varepsilon(X) = \left[\sum_{i=1}^{n} e^{\lambda_i(X)/\varepsilon} \right]^{-1} \left[\sum_{i=1}^{n} e^{\lambda_i(X)/\varepsilon} \nu_i \nu_i^T \right],$$

where λ_i denotes the i'th smallest eigenvalue of X and ν_i is the eigenvector corresponds to $\lambda_i(X)$ with $\|\nu_i\| = 1, \forall i$. It holds for f_ε that

$$\lambda_{\max}(X) \leq f_\varepsilon(X) \leq \lambda_{\max}(X) + \varepsilon \ln n.$$

Hence for ε sufficiently small, $f_\varepsilon(X) \approx \lambda_{\max}(X)$. Numerically, the derivative $\nabla_X f_\varepsilon(X)$ only depends on few largest eigenvalues and correpondant eigenvectors since the factors of the other terms in the sum decrease very rapidly. On the other hand, the Lipschitz constant of the gradient is inverse proportional to ε. Further, $f_\varepsilon(X)$ is convex, but not strictly convex. To see that, for all $0 \leq \alpha \leq 1$, it holds that

$$\begin{aligned}\alpha f_\varepsilon(I) + (1-\alpha) f_\varepsilon(2I) &= \alpha\varepsilon \ln(ne^{\frac{1}{\varepsilon}}) + (1-\alpha)\varepsilon \ln(ne^{\frac{2}{\varepsilon}}) = \varepsilon \ln n + 2 - \alpha \\ &= \varepsilon \ln(ne^{\frac{2-\alpha}{\varepsilon}}) = f_\varepsilon((2-\alpha)I) = f_\varepsilon(\alpha I + (1-\alpha)2I).\end{aligned}$$

2.3 Optimal Control Theory

Consider a general discrete-time optimal control problem

$$\begin{aligned}\underset{x_{0:N}, u_{0:N-1}}{\text{minimize}} \quad & \phi(X_N) + \sum_{k=0}^{N-1} f_0(k, x_k, u_k) \\ \text{s.t.} \quad & x_{k+1} = f(k, x_k, u_k) \\ & x_0 = \bar{x}, \quad G(x_N) = 0,\end{aligned} \tag{2.3.1}$$

where the initial value $\bar{x}$ is given, $G(x) = [g_1^T(x), \cdots, g_p^T(x)]^T$ and the gradients $\nabla g_k(x), k = 1, \cdots, p$ are linearly independent. Pontryagin Maximum Principle (PMP) provides necessary optimality conditions for (2.3.1) and it reads

Theorem 2.3.1. *Let $u^*_{0:N-1}$ be an optimal control for* (2.3.1) *and let $x^*_{0:N}$ be the corresponding trajectory. Then there exists an adjoint variable $\lambda_{1:N}$ such that*

1. *(Adjoint equation)*
$$\lambda_k = \frac{\partial H}{\partial x}(k, x_k^*, u_k^*, \lambda_{k+1}), \quad k = 1, \cdots, N-1.$$
2. *("Pointwise minimization")*
$$\frac{\partial H}{\partial u}(k, x_k^*, u_k^*, \lambda_{k+1}) = 0, \quad k = 0, 1, \cdots, N-1.$$
3. *(Boundary condition)*
$$\lambda_N = \frac{\partial \phi}{\partial x}(x_N^*) + \nabla_x G(x_N^*)^T \nu,$$
for some $\nu \in \mathbb{R}^p$.

And the Hamiltonian is

$$H(k, x, u, \lambda) = f_0(k, x, u) + \lambda^T f(k, x, u).$$

If (2.3.1) is linear quadratic, namely,

$$\begin{aligned}\underset{x_{0:N}, u_{0:N-1}}{\text{minimize}} \quad & x_N^T S x_N + \sum_{k=0}^{N-1} x_k^T Q x_k + u_k^T R u_k \\ \text{s.t.} \quad & x_{k+1} = Ax_k + Bu_k, \\ & x_0 = \bar{x},\end{aligned} \tag{2.3.2}$$

where Q, S are positive semi-definite and R is strictly positive semi-definite, then PMP is also sufficient for optimality since (2.3.2) has a unique solution.

Besides PMP, the solution to the discrete LQ optimal control problem can be also computed as follows,

$$u_k^* = -(B^T P_{k+1} B + R)^{-1} B^T P_{k+1} A x_k,$$

where $P_{1:N} \succeq 0$ is the solution to the discrete-time Riccati Equation (DRE)

$$\begin{aligned} &P_k = A^T P_{k+1} A + Q - A^T P_{k+1} B (B^T P_{k+1} B + R)^{-1} B^T P_{k+1} A, \quad k = 1 : N-1 \\ &P_N = S. \end{aligned}$$

2.4 Some Kowledge about Graph Theory

An edge-weighted undirected graph $\mathcal{G}(\mathcal{V}, \mathcal{E}, \mathcal{W})$ is composed of a node set $\mathcal{V} = \{1, 2, \cdots, N\}$, an edge set $(i, j) \in \mathcal{E},\ i, j \in \mathcal{V}$ which describes the connection topology between the nodes and the edge weight set $w_{ij} \in \mathcal{W}, i, j \in \mathcal{V}$ which includes all the weights of the corresponding edges. In the remainder of this thesis, in order to abbreviate the notation, we label the edges with numbers. For example, an edge with label l is denoted as $l \in \mathcal{E}$. On the other hand, seen from the nodes' perspective, the set of edges that is connected to node i is denoted as $\mathcal{E}(i)$, which can be interpreted as "communication channels" of node i. Note that $\mathcal{E} = \bigcup_{i \in \mathcal{V}} \mathcal{E}(i)$. Similarly, the set of the edge weights belonging to node i is denoted as $\mathcal{W}(i)$. $\mathcal{N}(i)$ denotes the neighbour vertices set of node i.

Note that in this thesis, only undirected graphs are considered, hence the edge-weighted Laplacian matrix L_w is symmetric and defined as

$$\begin{aligned} [L_w]_{ij} &= \begin{cases} \sum_l w_{il} & \text{if } i = j \text{ and } (i, l) \in \mathcal{E} \\ -w_{ij} & \text{if } i \neq j \text{ and } (i, j) \in \mathcal{E} \\ 0 & \text{otherwise} \end{cases} \\ \Leftrightarrow L_w &= \sum_{k \in \mathcal{E}} w_k E_k, \end{aligned}$$

where k is the label of the edges, $w_k \in \mathcal{W}$, $\forall k \in \mathcal{E}$ are the edge weights. If nodes i and j are connected via edge k, then $[E_k]_{ii} = [E_k]_{jj} = 1, [E_k]_{ij} = [E_k]_{ji} = -1$, and the other elements of E_k are zero. If $w_k = 1, \forall k$, then L_w becomes a "classic" Laplacian matrix, i.e., without edge-weights. L_w is symmetric and we denote $\lambda_i(L_w)$ as the i'th smallest eigenvalue of L_w.

2.5 Some Knowledge about Statistics

In this thesis, we will also look into some problems that involves random variables. Here we would like to introduce some definitions and theorems that are useful in the remainder of the thesis.

Definition 2.5.1. A sequence $\{X_n\}$ of random variables is said to *converge in probability* towards the random variable X if

$$\lim_{n \to \infty} \mathbb{P}\left(|X_n - X| > \varepsilon\right) = 0, \quad \forall \varepsilon > 0,$$

and is denoted as $X_n \xrightarrow{p} X$.

Uniform law of large numbers is useful when proving statistical consistency of an estimator. It reads as follows.

Theorem 2.5.2. *If*

1. *Θ is compact,*
2. *$f(x,\theta)$ is continuous at each $\theta \in \Theta$ for almost all x's, and $f(x,\theta)$ is a measurable function of x at each θ,*
3. *there exists a dominating function $d(x)$ such that $\mathbb{E}[d(x)] < \infty$ and $\|f(x,\theta)\| \leq d(x)$, $\forall \theta \in \Theta$,*

then $\mathbb{E}[f(x,\theta)]$ is continuous in θ and $\sup_{\theta\in\Theta} \|\frac{1}{n}\sum_{i=1}^{n} f(X_i,\theta) - \mathbb{E}[f(X,\theta)]\| \xrightarrow{p} 0$.

To prove statistical consistency of an estimator, the following theorem [9] is also useful.

Theorem 2.5.3. *Let M_n be random functions and let M be a fixed function of θ such that for every $\varepsilon > 0$*

$$\sup_{\theta\in\Theta} |M_n(\theta) - M(\theta)| \xrightarrow{p} 0, \qquad \sup_{\theta:\|\theta-\theta_0\|\geq\varepsilon} M(\theta) \leq M(\theta_0),$$

then any sequence of estimators $\hat{\theta}_n$ with $M_n(\hat{\theta}_n) \geq M_n(\theta_0) - o_p(1)$ converges in probability to θ_0, where $o_p(1)$ denotes a sequence that converges in probability to zero.

Maximum likelihood is a powerful estimation method to infer parameters in probabilistic models from observations. However, due to the complicated probability density functions and latent variables, the log-likelihood often results in a complicated expression that is difficult to analyze and solve. Proposed by [3], the EM algorithm provides a systematic way of decomposing the maximum likelihood problem into two sub-problems and hopefully both of the problems are easier to solve than the original problem. The EM-algorithm is well-celebrated today and has found its applications in parameter inference.

The key idea of the EM algorithmis to consider the joint-likelihood function of the latent variable x and the measurements $y_{1:N}$, namely,

$$L_\theta(x, y_{1:N}) = \ln p_\theta(x, y_{1:N}),$$

where $p_\theta(x, y_{1:N})$ is the joint probability density function of x and $y_{1:N}$ that is parametrized by coefficient θ. By Bayes' rule, it follows that

$$\ln p_\theta(y_{1:N}) = \ln p_\theta(x, y_{1:N}) - \ln p_\theta(x|y_{1:N}). \tag{2.5.1}$$

Let θ_k is an estimate of θ, taking the expected value with respect to $p_{\theta_k}(x|y_{1:N})$ on both sides of (2.5.1), we have that

$$\begin{aligned}&\ln p_\theta(y_{1:N}) = \int \ln p_\theta(x, y_{1:N}) p_{\theta_k}(x|y_{1:N}) dx - \int \ln p_\theta(x|y_{1:N}) p_{\theta_k}(x|y_{1:N}) dx \\ &= \mathbb{E}_{\theta_k}\left[\ln p_\theta(x, y_{1:N})|y_{1:N}\right] - \mathbb{E}_{\theta_k}\left[\ln p_\theta(x|y_{1:N})|y_{1:N}\right] \\ &= \mathcal{Q}(\theta, \theta_k) - \mathcal{V}(\theta, \theta_k).\end{aligned}$$

Note that

$$L_\theta(y_{1:N}) - L_{\theta_k}(y_{1:N}) = \left(\mathcal{Q}(\theta,\theta_k) - \mathcal{Q}(\theta_k,\theta_k)\right) + \left(\mathcal{V}(\theta_k,\theta_k) - \mathcal{V}(\theta,\theta_k)\right),$$

where $(\mathcal{V}(\theta_k, \theta_k) - \mathcal{V}(\theta, \theta_k)) \geq 0$ is Kullback-Leibler information divergence. Hence if a new estimation θ_{k+1} is such that $\mathcal{Q}(\theta_{k+1}, \theta_k) \geq \mathcal{Q}(\theta_k, \theta_k)$, then there will be an increase in the likelihood function, i.e., $L_{\theta_{k+1}}(y_{1:N}) \geq L_{\theta_k}(y_{1:N})$ and hence result in an at-least-as-good, or even better estimation of θ. When using the EM-algorithm, one first calculates $\mathcal{Q}(\theta, \theta_k)$ by the current estimation θ_k and then let $\theta_{k+1} = \arg\max_\theta \mathcal{Q}(\theta, \theta_k)$ and repeat the procedure until $\mathcal{Q}(\theta, \theta_k)$ converges.

In this thesis, the EM-algorithm is used to infer the parameter in the objective functions of optimal control problems and as mentioned before, this involves calculating the expected value and it is done through fixed point smoothing [7]. The purpose of fixed point smoothing is to estimate $\hat{x}_{t_0|t}, t > t_0$, where t_0 is an arbitrary fixed time.

Consider a discrete-time stochastic system

$$x_{t+1} = Fx_t + v_t, y_t = Hx_t + e_t,$$

where x_t is the state vector and v_t, e_t are independent white Gaussian noises with covariances R_1 and R_2 respectively. By introducing $z_t = [x_t^T, x_t'^T]^T$, the above system can be re-written as

$$\begin{aligned} z_{t+1} &= \begin{bmatrix} F & 0 \\ 0 & I \end{bmatrix} z_t + \begin{bmatrix} v_t \\ 0 \end{bmatrix}, \quad z_{t_0} = \begin{bmatrix} x_{t_0} \\ x_{t_0} \end{bmatrix}, \\ y_t &= \begin{bmatrix} H & 0 \end{bmatrix} z_t + e_t. \end{aligned}$$

A standard Kalman filter can be applied to this system to obtain the optimal estimate $\hat{z}(t+1|t)$. The lower half of this vector is exactly $\hat{x}_{t_0|t}$.

References

[1] Bernard Brogliato, Aris Daniilidis, Claude Lemaréchal, and Vincent Acary. On the equivalence between complementarity systems, projected systems and differential inclusions. *Systems & Control Letters*, 55(1):45–51, 2006.

[2] Jorge Cortes. Discontinuous dynamical systems. *IEEE control Systems*, 28(3), 2008.

[3] Arthur P Dempster, Nan M Laird, and Donald B Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society. Series B (methodological)*, pages 1–38, 1977.

[4] Shaoshuai Mou, Ji Liu, and A Stephen Morse. A distributed algorithm for solving a linear algebraic equation. *IEEE Transactions on Automatic Control*, 60(11):2863–2878, 2015.

[5] Anna Nagurney and Ding Zhang. *Projected dynamical systems and variational inequalities with applications*, volume 2. Springer Science & Business Media, 2012.

[6] Yurii Nesterov. Smoothing technique and its applications in semidefinite optimization. *Mathematical Programming*, 110(2):245–259, 2007.

[7] Torsten Söderström. *Discrete-time stochastic systems: estimation and control*. Springer Science & Business Media, 2012.

[8] Mingxuan Sun and Danwei Wang. Analysis of nonlinear discrete-time systems with higher-order iterative learning control. *Dynamics and Control*, 11(1):81–96, 2001.

[9] Aad W Van der Vaart. *Asymptotic statistics*, volume 3. Cambridge university press, 2000.

"There's no such a thing as a painless lesson. They don't exist. Sacrifices are necessary. You can't gain anything without losing something first. Although, if you can endure that pain and walk away from it, you'll find that you now have a heart strong enough to overcome any obstacles — A heart made of Fullmetal."

— Edward Elric

Part II: Research Papers

Paper A

Consensus Control for Linear Systems with Optimal Energy Cost

Paper B

Projected Primal-Dual Gradient Flow of Augmented Lagrangian with Application to Distributed Maximization of the Algebraic Connectivity of a Network

Paper C

Inverse Quadratic Optimal Control for Discrete-Time Linear Systems

Paper D

Inverse Optimal Control for Finite-Horizon Discrete-Time Linear Quadratic Regulator Under Noisy Output