Title
How does uncertainty propagate through a transportation network under equilibrium?

Author
Liu, Ning

Publication Date
2021

Peer reviewed|Thesis/dissertation

How does uncertainty propagate through a transportation network under equilibrium?

By

NING LIU
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

APPLIED MATHEMATICS

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Yueyue Fan , Chair

Matthias Koeppe

Qinglan Xia

Committee in Charge

2021

To my grandma and uncle in heaven
-who I promised to dedicate this dissertation before they left this world

Contents

Abstract iv

Acknowledgments v

Chapter 1. Introduction 1

1.1. Introduction 1

1.2. Notation Introduction 5

1.3. Network Problem Setup 6

Chapter 2. A Minimum Norm Solution Mapping (MNSM) 11

2.1. Motivation 11

2.2. Minimum Norm Solution Mapping 11

2.3. Continuity of Minimum Norm Solution Mapping 14

2.4. Connectivity of the Partitions in Demand Space 21

Chapter 3. MNSM for Convex Quadratic Optimization (CQO) 23

3.1. The MNSM for a given region 23

3.2. Determine the Partition Region 32

3.3. Projection-based Solution Mapping 40

3.4. Discussion 41

Chapter 4. Approximated MNSM for Convex Optimization (CO) 44

4.1. Epigraph and Epi-convergence 44

4.2. Epi-splines 47

4.3. The MNSM for convex optimization 49

4.4. Quadratic Determination 54

Chapter 5. Measure push-forward by the MNSM 61

5.1. Preliminaries of probability space and measure 61

5.2. Measure Push-forward by Minimum Norm Solution Mapping 63
5.3. The MNSM of Convex optimization preserves Measure 65

Chapter 6. Numerical Methods 67
6.1. Find the MNSM in a given partition region 67
6.2. Preliminaries for Finding Partitions 69
6.3. Search for Partition Regions 72
6.4. Identify An Initial Region 75
6.5. Numerical Example and Discussion 78

Chapter 7. Applications 84
7.1. Projected Spaces and Uncertainty Prolongation 84
7.2. Identify Critical Network 93
7.3. Suboptimal assignment 97
7.4. Toll Policy Design 101

Chapter 8. Future Work 109
8.1. Relation between Different Partition Regions 109
8.2. Computational Robustness 109
8.3. Larger Network 110
8.4. Learning With Small Data 110

Appendix A. Details of 3×3 grid network 112

Bibliography 114

How does uncertainty propagate through a transportation network under equilibrium?

Abstract

Understanding the relationship between demand and traffic network flows in an uncertain setting is gaining more and more attention in the transportation science community. In this dissertation, we establish new theorems and methods for understanding the input-output relation of traffic network equilibrium problems under uncertainty. Approaching the problem from a fresh geometrical perspective, we provide new understanding of the uncertainty propagation process in the problem of traffic equilibrium. We first introduce a minimum norm solution mapping (MNS-M) between travel demand and network flows and explore its mathematical properties in terms of well-defined, continuity, induced partition and connectivity. Under the linearity assumption of the link cost function, we provide a stable analytic formula of the MNSM together with the criterion of partition region determination. We then extend those results to more general cost functions by using epi-splines to incorporate more realistic situation of nonlinear link cost under congestion. The new results associated with nonlinear cost functions can also maintain good geometric inherent characteristics. After completing these fundamental analyses, we demonstrate how the MNSM can help understand the uncertainty propagation process through the push-forward measure induced by the MNSM and we prove that the approximation process maintains strong convergence of the push-forward measure. We also provide an effective algorithm to compute the MNSM which avoids massive enumeration. Several important application examples are provided in the end to demonstrate how the new analytical and numerical methods established in the dissertation can be used to provide engineering and policy insights for transportation network planning and management.

Acknowledgments

Foremost, I would like to express my sincere gratitude to my advisor Prof. Yueyue Fan for the continuous support of my Ph.D. study and research, for her patience, motivation, enthusiasm, and immense knowledge. Her tolerance allowed me to turn my wild ideas into a complete study. Her guidance helped me in all the time of research and writing of this dissertation. I could not have imagined having a better advisor and mentor for my Ph.D study.

Besides my advisor, I would like to thank the rest of my dissertation committee: Prof. Matthias Koeppe and Prof. Qinglan Xia, for their encouragement, tolerance, insightful comments, and hard questions.

My sincere thanks also goes to Prof. Roger J-B Wets, Prof. Johannes O. Royset, Prof. Hao Chen, Prof. Xin Liu, and Prof. Jesus De Loera for offering me many valuable suggestions and discussions.

I would like to acknowledge and thank Becca Thomases, Robert Guy, Tina Denena and Sarah Driver for saving my life from the darkest moment. Without their help and support, I cannot live in the world and complete this work.

I thank my fellows in GGAM for the simulating discussions, for the brainstorming about the world, for the sleepless nights we were working together before deadlines, , for all the fun we have had in the past years and for all the moments we shared with each other. I didnt list your names for fear that I might miss one.

Last but not the least, I would like to thank my family: my parents Xiuqing Zhang and Zhonghai Liu, for giving birth to me at the first place and supporting me spiritually throughout my life.

CHAPTER 1

Introduction

1.1. Introduction

In transportation science, how to map travel demand to network flows is generally categorized as traffic assignment problems. The network abstraction naturally arises as transportation services are operated over spatially distributed but connected infrastructure systems. In a congested network, the travel time/cost of traversing a network segment depends on traffic condition of that segment. As individuals deciding how to go through the network to fulfill their travel needs, their collective actions influence the overall network condition, which in turn then affects their individual travel decisions. Therefore, in a traffic assignment problem, travelers decisions, though are made by individuals, are typically modeled simultaneously at a network level to capture the interactions among users.

The most widely adopted traffic assignment models are based on Wardrop's user equilibrium principle, with an assumption that all users choose their best perceived routes [War52]. A network is said to be under an equilibrium condition if no one can achieve a better result by unilaterally changing one's decision. Based on this principle, [BMW55] first formulated the Wardrop equilibrium for a general network as a mathematical programming problem, often referred as the BMW model in transportation and regional science [BMN05]. This classic network equilibrium model has been extensively studied over the last several decades. Just to name a few, efforts include modeling extensions to incorporate elastic demand [Daf82, ND02, Boy80] and temporal dynamics [WFT90, HA05, LS02, BPLM12], theoretical analyses on mathematical properties of the equilibrium solutions in terms stability, existence, and uniqueness [Smi79, Daf80, MN78, Car86], computational efforts in pursuing more efficient solution algorithms [Aka01, PY84, Nag86, PHRU12], and developments of computer simulation tools [FMT08, MFT08, Mah01]. Readers may refer to books by [She85, MP07, Pat15] for comprehensive reviews of the very rich literature on traffic network equilibrium models. Besides its foundational roles in transportation science problems, network equilibrium models can have equivalent presence in many other domain problems that involve non-cooperative games/competition, such as logistics and supply chain [ZDN+03, Nag06], power systems [NM07, GF17], and communication networks [BGH92, ORS93].

As stated by [BMN05], "BMW laid the intellectual and economic science foundation for transportation systems analysis planning and evaluation for the rest of the 20th century and beyond." In this dissertation, we revisit this classic traffic network equilibrium model and aim for a better understanding of the relation between travel demand and the network flows in the model. Following existing methods, one may immediately think of three possible ways of approaching this question. A straightforward way is to directly pinpoint the input-output relation by solving the network equilibrium model, i.e. to find a corresponding network flow solution or solution set for a given travel demand. To understand how small changes of travel demand may impact the network flow patterns, sensitivity analysis can be used. In a sensitivity analysis, the directions and rate of change of the network flows and travel costs are evaluated under perturbation of model parameters. Sensitivity analyses were carried out for the network equilibrium problem presented in the form of either convex optimization [TF88] or variational inequalities [QM89, Pat04, JP07]. Despite of the broad applications of sensitivity analyses of network equilibrium in network design, control, and estimation problems [Yan97, YY05, JP07], an apparent limitation of using sensitivity analysis to understand the global relation between model input and output is that the analysis results are only valid within a very small perturbed neighborhood. Alternatively, one may approach the problem using parametric optimization [GVJ90]. For example, [ESvBK18] used parametric optimization techniques to analyze the properties of the solution to a boundedly rational user equilibrium model under perturbation of model parameters. Parametric optimization was developed in parallel with sensitivity analysis [GG12], and has been applied to understanding the input-output relation in general mathematical programming and optimal control problems [CA13]. All approaches reported in literature for solving multiparametric programming problems involve two basic steps: (1) determination of the optimal solution as a parameter-dependent function, valid over a certain region in the parameter space; and (2) exploration of the remaining parameter space. Besides for limited special cases, an analytical solution can be extremely hard to obtain. Another challenge is that when the problem contains multiple optimal solutions, the corresponding parametric optimization problem cannot be solved because the inverse function cannot be constructed. This is why parametric optimization could not be directly applied to traffic assignment problems because solution uniqueness is typically not guaranteed. For example, in the study by [ESvBK18], the analyses were focused on the best and worst cases to avoid the solution non-uniqueness issue.

Deviating from existing schools of thought, we take a fresh geometric perspective combined with variational analysis, to understand how the travel demand space uncertainty transforms to the network flow space through physical and behavior rules imposed on the network. Our approach has some unique features compared to existing methods, which may offer advantages for downstream applications. First, one challenge arising often in decision making under uncertainty is the infinite problem dimension. Our approach can be used to overcome this challenge by focusing on the finite partitions of the input space instead of considering infinite possible realizations of the input parameters. Following our method, the partition of the input space could be determined directly and within each partition region only equality constraints exist for the decision making problem. Furthermore, under the assumption of linearity of the link cost function, analytical relation between the input space and the output space is easily presented within each partition region. For the problem without linearity assumption, we provide an approximation process by using epi-splines, which could extend the results from the linear link cost function to general convex link cost function. With the above results, the spread of uncertainty from the input to the output space can be clearly outlined.

This dissertation is organized as follows. In the remaining of chapter 1, we formulate the network problem. In chapter 2, we define the minimum norm solution mapping (MNSM) and we show that the MNSM is well-defined and it is a homeomorphism between corresponding spaces. With the homeomorphism property, we show that it preserves the connectivity and could form a natural partition of the input space. In chapter 3, we derive the analytical formula of the MNSM for convex quadratic optimization problem, including an analytic solution for a given partition region and the criterion to determine the partition region. The similar results for projected-based solution mapping are discussed as well. In chapter 4, we extend the analytical formula obtained in chapter 3 to a family of convex optimization problem by using epi-splines. In chapter 5, we focus on the measure push-forward by the MNSM and verify some basic properties of the measure. At the end of this chapter, we provide a convergence result of the push-forward measure under the process discussed in chapter 5. In chapter 6, the numerical methods to obtain the MNSM for convex quadratic problem are developed and explained. In chapter 7, we give four applications of the MNSM, projected spaces, identification of critical network, suboptimal assignment and toll policy design. In the last chapter, we discuss some unfinished thoughts of this dissertation research.

1.2. Notation Introduction

General Math (I)

$\mathbb{N}$	Natural number $\{1, 2, 3, \dots\}$
$\mathbb{R}$	Real number
$\mathbb{R}_+$	Nonnegative real number
$\mathbb{R}_+^n$	Nonnegative n-dimensional real number
$a \in A$	a is member of set A
$A \subseteq B$	A is a subset of B
$A \subset B$	A is a proper subset of B
$\emptyset$	Empty set
$A \cup B$	The union of A and B
$A \cap B$	The intersection of A and B
$A \setminus B$	Set difference
A^c	Complements of A
$A \Delta B = (A \setminus B) \cup (B \setminus A)$	Symmetric difference
$A \times B$	Cartesian product of A and B
$\|A\|$	Cardinality of the Cartesian product
$\text{Int}(A)$	Interior of set A
(x_1, x_2, x_3, x_4)	row vector with entries of x_1, x_2, x_3, x_4
$(x_1, x_2, x_3, x_4)^T$	column vector with entries of x_1, x_2, x_3, x_4
A^T	Transpose of matrix A
$A^\dagger$	Moore-Penrose inverse of matrix A
$\\| \cdot \\|$	2-norm of vector
$\\| \cdot \\|_F$	Frobenius norm of matrix
$\\| \cdot \\|_*$	Nuclear norm of matrix
$min f(x)$	Minimum value of function $f(x)$
$\arg\min f(x)$	Arguments of the minima of function $f(x)$
$\tilde{f}$	Extended function of function $f(x)$
$\partial f(x)$	Subgradient of function $f(x)$
$\text{epi} f(x)$	Epigraph of function $f(x)$

General Math (II)

$1:m$	The set $\{1,2,3,\ldots,m\}$
$\mathbb{P}(A)$	The power set of set A
C_b	Feasible set with parameter b
S_b	Solution set with parameter b
$\mathcal{A}_x$	Active set $\{i \mid x_i = 0\}$
$\mathcal{A}_x(b)$	Active set with parameter b $\{i \mid x(b)_i = 0\}$
$\mathcal{R}_U$	Partition set $\{b \mid \mathcal{A}_x(b) = U\}$
$\mathcal{D}^\dagger$	Image of the minimum norm solution mapping of the set $\mathcal{D}$

Transportation Assignment Problem

$\mathcal{L}$	Link flow space
x	Link flow
$\mathcal{P}$	Path flow space
f	path flow
$\mathcal{D}$	Demand space
q	Demand
B	Demand-path incidence matrix
F	Link-path incidence matrix
$c_l(x)$	Cost function on link l
$\{x_1, x_2, x_3, x_4\}$	Directed path connect link x_1, x_2, x_3, x_4 in turn

1.3. Network Problem Setup

In this section, we introduce some basic concepts and notion of the problem we studied in this dissertation. The core issue revolves around the two principles of Wardrop. According to

Beckmann, the two principles of Wardrop could be formulated as following:

$$
\begin{array}{r|l}
\textbf{UE:} & \textbf{SO:} \\
\min\limits_x \sum\limits_l \int_0^{x_l} t_l(x_l)dx & \min\limits_x \sum\limits_l x_l t_l(x_l)dx \\
\text{subject to} \sum\limits_k f_k^{rs} = q_{rs} : \forall r,s & \text{subject to} \sum\limits_k f_k^{rs} = q_{rs} : \forall r,s \\
x_l = \sum\limits_r \sum\limits_s \sum\limits_k \delta_{l,k}^{rs} f_k^{rs} : \forall l & x_l = \sum\limits_r \sum\limits_s \sum\limits_k \delta_{l,k}^{rs} f_k^{rs} : \forall l \\
f_k^{rs} \geq 0 : \forall k,r,s & f_k^{rs} \geq 0 : \forall k,r,s \\
x_l \geq 0 : \forall l \in \mathcal{L} & x_l \geq 0 : \forall l \in \mathcal{L}
\end{array}
\tag{1.1}
$$

where x_l is the equilibrium flow on link l, t_l is the travel time on link l, f_k^{rs} is the path flow on the path k connecting O-D pair $r - s$, q_{rs} is the demand between r and s. To make the formulation more compact, we give the O-D pairs an order so that we could use a vector $q \in \mathbb{R}_+^{n_d}$ to represent all the OD pairs in the system where n_d is the number of different OD pairs. Then we could rewrite the UE/SO as following:

$$
\begin{array}{r|l}
\textbf{UE:} & \textbf{SO:} \\
\min\limits_x \quad T_{ue}(x) & \min\limits_x \quad T_{so}(x) \\
\text{subject to } Bf = q & \text{subject to } Bf = q \\
x = Ff & x = Ff \\
f \geq 0 & f \geq 0
\end{array}
\tag{1.2}
$$

where $x \in \mathbb{R}_+^{n_l}$,$f \in \mathbb{R}_+^{n_p}$,n_l is the number of different links, n_p is the number of different paths, B is the incidence matrix between paths and demands,$B_{ij} = 1$ if the jth path connects the ith OD pair,$B_{ij} = 0$ otherwise; F is the incidence matrix between paths and links,$F_{jk} = 1$ if the jth path takes the kth link, $F_{jk} = 0$ otherwise; $T_{ue}(x) = \sum\limits_l \int_0^{x_l} t_l(x_l)dx$, and $T_{so}(x) = \sum\limits_l x_l t_l(x_l)dx$. If we closely look at the constraints in the above problems, it shows a connection among three spaces, the link flow space $\mathcal{L} \subset \mathbb{R}_+^{n_l}$ (the space of all the possible link flow), the path flow space $\mathcal{P} \subset \mathbb{R}_+^{n_p}$ (the space of all the possible path flow) and the demand space $\mathcal{D} \subset \mathbb{R}_+^{n_d}$ (the space of all the possible traffic demand).

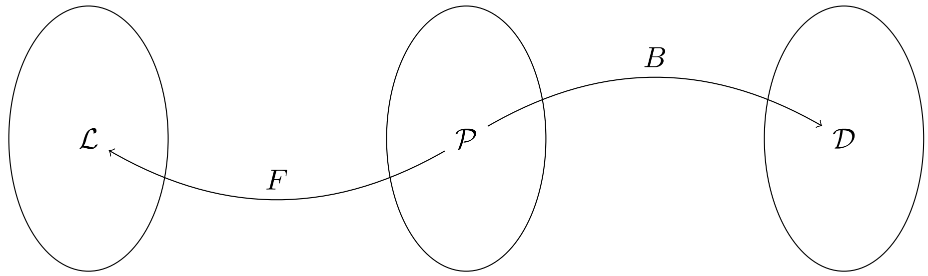


FIGURE 1.1. Relation among the link space, the path space and the demand space

To make it clear, we provide a four-link network as an example.

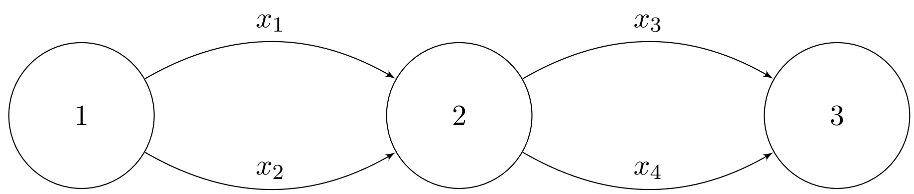


FIGURE 1.2. Four link network

EXAMPLE 1. In the four-link network, $1, 2, 3$ are nodes, x_1, x_2, x_3, x_4 are link 1, link 2, link 3, link 4 respectively. Then the link space $\mathcal{L}$ belongs to $\mathbb{R}^4_+$. The entries of $x = (x_1, x_2, x_3, x_4)^T \in \mathcal{L}$ represent the number of traffic on the corresponding link. For example, $x = (1, 2, 3, 4)^T$ means that there are 1 unit on the link 1, 2 units on the link 2, 3 units on the link 3, 4 units on the link 4.

There are three O-D pairs, $q_1 : 1 \to 2$,$q_2 : 1 \to 3$, and $q_3 : 2 \to 3$. Then the demand space $\mathcal{D}$ belongs to $\mathbb{R}^3_+$. The entries of $q \in \mathcal{D}$ represent the number of traffic demands. For example, $q = (2, 3, 5)$ means that there are 2 units starting from node 1 to node 2, 3 units starting from node 1 to node 3, 5 units starting from node 2 to node 3.

Moreover, there are eight paths, $f_1 : x_1$, $f_2 : x_2$, $f_3 : x_1 \to x_3$, $f_4 : x_1 \to x_4$, $f_5 : x_2 \to x_3$, $f_6 : x_2 \to x_4$, $f_7 : x_3$, $f_8 : x_4$. Then the path space $\mathcal{P}$ is a subset of $\mathbb{R}^8_+$. The entries of $f \in \mathcal{P}$ represent the number of traffic through the corresponding path.The incidence matrix between link flow space and path flow space F is given by,

$$F = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix},$$

and the incidence matrix between demand space and path flow space B is given by,

$$B = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

Then, $x = Ff$ and $q = Bf$.

Instead of considering the link-demand relation directly, we begin with the path-demand relation. Since the link-path relation is known, once we understand the relation between the path space and the demand space, we could easily push forward to the link space.Hence, we could further reduce the UE/SO as following:

$$\begin{array}{r|r} \textbf{UE:} & \textbf{SO:} \\ \min\limits_{f} \quad T_{ue}(Ff) & \min\limits_{f} \quad T_{so}(Ff) \\ \text{subject to } Bf = q & \text{subject to } Bf = q \\ f \geq 0 & f \geq 0 \end{array} \tag{1.3}$$

Let $T(f) = T_{ue}(Ff)$ if we consider UE problem or $T(f) = T_{so}(Ff)$ if we consider SO problem.Hence, the problem we are facing becomes

$$\min_{f} \quad T(f) \text{ subject to } Bf = q, f \geq 0 \tag{1.4}$$

Since we need the travel time on each link is a increasing function depending on the flow, the above problem, in fact, is a stochastic convex optimization problem for the demand q. Then, the solution mapping of above problem is

$$f(q) := \arg\min\{T(f) | Bf = q, f \geq 0\} \tag{1.5}$$

and the correponding assignment for link flow is

$$x(q) =: F(f(q)) = F(\arg\min\{T(f) | Bf = q, f \geq 0\}) \tag{1.6}$$

Clearly, without further conditions, those two mappings are set-valued. To handle set-valued mappings, we need the tools created by pioneers such as R. T. Rockafellar and Roger Wets, Jon Borwein and Adrian Lewis, and Boris Mordukhovich known as Variational Analysis.

On the other hand, since the solution mapping is set-valued, it means that for each realization of the traffic demand q, there might be finitely or infinitely many corresponding assignments which will dilute the corresponding probability. Hence, to understand the uncertainty prolongation, we have to give up the idea of studying the whole solution set and find a systematic way to track the uncertainty prolongation.

Due to these two observations, we come up with the idea of the minimum norm solution mapping (MNSM) that maps demand space to the path flow space as shown in Figure 1.3, which we will discuss in detail in Chapter 2.

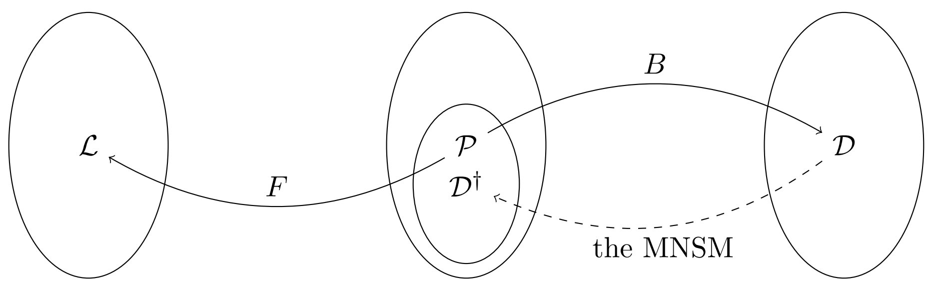


FIGURE 1.3. Relation among the link space, the path space and the demand space with the MNSM

CHAPTER 2

A Minimum Norm Solution Mapping (MNSM)

2.1. Motivation

Consider the convex optimization problem,

$$\min_{x} \quad g(x) \quad s.t. \quad Ax = b, x \geq 0 \tag{2.1}$$

where $g(x)$ is a convex function, A is a full row rank matrix. Our goal is to understand how the uncertainty of b prolongates to the corresponding solution $x(b)$, i.e.

$$S(b) := argmin\{g(x)|Ax = b, x \geq 0\} \tag{2.2}$$

In general, $S(b)$ may have multiple values for given b, i.e. $S(b)$ is set-valued mapping. This is not good news for studying the uncertainty prolongation because the probability of each given b will be corresponding to the whole set of the corresponding solutions $S(b)$, i.e. for any $x \in S(b)$

$$P(x|b) = \frac{P(b)}{|S(b)|} \tag{2.3}$$

where $|S(b)|$ is the number of elements in the set $S(b)$. It seems not a problem when $|S(b)|$ is finite but in the problem we are interested, the set of $S(b)$ always has infinity many elements so that the probability of each solution in $S(b)$ is 0s which does not help to understand how the uncertainty propagation. Hence, instead of considering the whole set $S(b)$, we want to build a systematic connection from the realization of b to some representative of $S(b)$, which leads to the Minimum Norm Solution Mapping discussed in the following section.

2.2. Minimum Norm Solution Mapping

Definition 2.2.1. (Minimum Norm Solution Mapping) Consider the problem (2.1) and let $C_b := \{x|Ax = b, x \geq 0\}$ and $S(b) := \arg\min_{x \in C_b} g(x)$. Then if g is coercive and lower-semicontinous,the minimum norm solution mapping $x^\dagger(b)$ is defined by

$$x^\dagger(b) := P_{S(b)}(0)$$

where $P_{S(b)}$ is the projection operator on $S(b)$, and a function g is coercive if and only if $g(x) \to +\infty$ as $\|x\| \to +\infty$. In fact $x^\dagger(b)$ is the solution of the following problem:

$$\begin{array}{ll} \min\limits_{y} & \|y\| \\ & y \in S(b) \end{array}$$

Now we want to show that the Minimum Norm Solution Mapping is well-defined, that is if C_b is nonempty, $x(b)$ is nonempty and single valued. To see this, recall the Weierstrass theorem.

PROPOSITION 2.2.1. *(Weierstrass' Theorem) Consider a closed proper function $g : \mathbb{R}^n \to (-\infty, \infty]$, and assume that any one of the following three conditions holds:*

(1) *dom(g) is bounded,*

(2) *there exists a scalar $\bar{\gamma}$ such that the level set*

$$\{x | g(x) \leq \bar{\gamma}\}$$

is nonempty and bounded,

(3) *g is coercive.*

Then the set of minima of g over $\mathbb{R}^n$ is nonempty and compact.

EXAMPLE 2.

$$\begin{array}{ll} \min & \frac{1}{x} \\ & x \geq 0 \end{array}$$

In this case, none of the above three is satisfied, and $g(x)$ is decreasing to 0 as $x \to \infty$.

Now applying the Weierstrass' Theorem to the extended real-valued function

$$\tilde{g}(x) = \begin{cases} g(x) & \text{if } x \in C_b \\ \infty & \text{otherwise} \end{cases}$$

we see that the set of minima of $g(x)$ over C_b is nonempty and compact if $g(x)$ is lower semicontinuous at each $x \in C_b$ and one of the following conditions holds:

(1) for some $\bar{\gamma}$, the set $\{x \in C_b | f(x) \leq \bar{\gamma}\}$ is nonempty and bounded

(2) $\tilde{g}$ is coercive

Since according to the definition that $g(x)$ is coercive, $\tilde{g}(x)$ is also coercive. Hence, the set of minima of $g(x)$ over C_b is nonempty and compact.

THEOREM 2.2.2. *If $g(x)$ is convex over C_b and x_1, x_2 are two distinct minima, then $\lambda x_1 + (1 - \lambda)x_2$, $\lambda \in [0, 1]$ is also a minimum which means that the set of minima of $g(x)$ over C_b is convex.*

Proof: Since x_1, x_2 are two distinct minima, $x_1 \neq x_2$ and $g(x_1) = g(x_2) = \inf_{x \in C_b} g(x)$. Since $g(x)$ is convex and C_b is a convex set, we obtain that

$$\inf_{x \in C_b} g(x) \leq g(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda g(x_1) + (1 - \lambda)g(x_2) = \inf_{x \in C_b} g(x)$$

Hence,

$$g(\lambda x_1 + (1 - \lambda)x_2) = \inf_{x \in C_b} g(x)$$

Therefore, $\lambda x_1 + (1 - \lambda)x_2$ is in the set of minima of $g(x)$ over C_b and the set of minima of $g(x)$ over C_b is convex. **Q.E.D.**

By now, we obtain that the set of minima of $g(x)$ over C_b is nonempty, convex and compact.

THEOREM 2.2.3. *Every closed convex subset of $\mathbb{R}^n$ has a unique element with minimum norm.*

THEOREM 2.2.4. *The Minimum Norm Solution Mapping is well-defined.*

Proof: According to previous discussion and definition 2.2.1, we obtain that the set of minima of $f(x)$ over C_b,$S(b)$ is a nonempty, convex and compact subset of $\mathbb{R}^n$. By theorem 2.2.3, for each given b or each realization of b such that C_b is nonempty,$x(b)$ is uniquely defined. **Q.E.D.**

Denote X to be the range of x, then $x(b)$ is the mapping from $\mathbb{R}^n_+$ to X.

COROLLARY 2.2.1. *The Minimum Norm Solution Mapping is bijective between $\mathbb{R}^n_+$ and X.*

Proof: By definition of X, the Minimum Norm Solution Mapping is surjective. According to theorem 2.2.4, it is injective. Therefore, it is bijective. **Q.E.D.**

Example 3. Consider the following problem:

$$\begin{aligned} \min_{x\in\mathbb{R}^3} \quad & x_1^2 \\ s.t. \quad & x_1+x_2+x_3=b \\ & x_1,x_2,x_3\geq 0 \end{aligned}$$

Then, it is easy to get that $S(b)=\{x\in\mathbb{R}^3|x_1=0,x_2=s,x_3=b-s,0\leq s\leq b\}$ while the optimal value is 0. Hence,

$$x(b)=P_{S(b)}(0)=(0,\frac{b}{2},\frac{b}{2})$$

where $P_{\mathcal{C}}(x)=\inf_{y\in\mathcal{C}}\|x-y\|$.

2.3. Continuity of Minimum Norm Solution Mapping

In this section, we want to show that the Minimum Norm Solution Mapping defined in the previous section is continuous, i.e. $x^\dagger(b):\mathbb{R}^n_+\to\mathbb{R}^m_+$ is continuous.

In order to explain the result clearly, we will borrow some notation introduced by R. Tyrrell Rockafellar and Roger J-B Wets [RW09]. Let $\mathbb{N}$ be the natural numbers,define the following collections of subsets of $\mathbb{N}$:

$$\begin{aligned} \mathcal{N}_\infty &:= \{N\subset\mathbb{N}|\mathbb{N}\backslash N \text{ finite }\} \\ &= \{\text{subsequences of } \mathbb{N} \text{ containing all } i \text{ beyond some } \bar{i}\} \\ \mathcal{N}_\infty^\# &= \{N\subset\mathbb{N}|\mathbb{N} \text{ infinite}\}=\{\text{all subsequences of } \mathbb{N}\} \end{aligned}$$

Definition 2.3.1 (inner and outer limits). For a sequence $\{C^i\}_{i\in\mathbb{N}}$ of subsets of $\mathbb{R}^n$, the *outer limit* is the set

$$\begin{aligned} \limsup_{i\to\infty} C^i &:= \{x|\exists N\in\mathcal{N}_\infty^\#, \quad \exists x^i\in C^i \quad (i\in N) \text{ with } x^i \underset{N}{\longrightarrow} x\} \\ &= \{x|\forall V\in\mathcal{N}(x),\exists N\in\mathcal{N}_\infty^\#,\forall i\in N: C^i\cap V\neq\emptyset\} \end{aligned}$$

while the *inner limit* is the set

$$\begin{aligned} \liminf_{i\to\infty} C^i &:= \{x|\exists N\in\mathcal{N}_\infty, \quad \exists x^i\in C^i \quad (i\in N) \text{ with } x^i \underset{N}{\longrightarrow} x\} \\ &= \{x|\forall V\in\mathcal{N}(x),\exists N\in\mathcal{N}_\infty,\forall i\in N: C^i\cap V\neq\emptyset\} \end{aligned}$$

The *limit* of the sequence exists if the outer and inner limit sets are equal:

$$\lim_{i\to\infty} C^i := \limsup_{i\to\infty} C^i = \liminf_{i\to\infty} C^i.$$

The next theorem provides the major criteria for checking set convergence which we will apply later.

THEOREM 2.3.2 (hit-and-miss criteria). *[RW09] For $C^i, C \subset \mathbb{R}^n$ with C closed, one has $C \subset \liminf_{i\to\infty} C^i$ if and only if for every open set $O \subset \mathbb{R}^n$ with $C \cap O \neq \emptyset$ there exists $N \in \mathcal{N}_\infty$ such that $C^i \cap O \neq \emptyset$ for all $i \in N$.*

Proof:[RW09] '⇒'holds by definition.
To show that '⇐' holds, consider any $x \in C$ and rational $\epsilon > 0$. There is a rational point $x' \in intB(x, \epsilon/2)$. For such a point x', we have $C \cap intB(x', \epsilon/2) \neq \emptyset$, so by assumption there exists $N \in \mathcal{N}_\infty$ with $C^i \cap intB(x', \epsilon/2) \neq \emptyset$ for all $i \in N$. Then, $x' \in C^i + (\epsilon/2)B$, so that $x \in C^n u + (\epsilon/2)B + (\epsilon/2)B = C^i + \epsilon B$ for all $i \in N$. Thus x satisfies the defining condition of $\liminf_i C^i$. **Q.E.D.**

Continuity properties of set-valued mappings $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ can be developed in terms of outer and inner limits as the limits:

$$\begin{aligned}
\limsup_{x\to\bar{x}} S(x) &:= \bigcup_{x^i\to\bar{x}} \limsup_{i\to\infty} S(x^i) \\
&= \{u | \exists x^i \to \bar{x}, \exists u^i \to u \text{ with } u^i \in S(x^i)\} \\
\liminf_{x\to\bar{x}} S(x) &:= \bigcap_{x^i\to\bar{x}} \liminf_{i\to\infty} S(x^i) \\
&= \{u | \forall x^i \to \bar{x}, \exists N \in \mathcal{N}_\infty, u^i \underset{N}{\to} u \text{ with } u^i \in S(x^i)\}
\end{aligned}$$

DEFINITION 2.3.3 (continuity and semicontinuity). [RW09] A set-valued mapping $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is *outer semicontinuous* (osc) at $\bar{x}$ if

$$\limsup_{x\to\bar{x}} S(x) \subset S(\bar{x}),$$

or equivalently

$$\limsup_{x\to\bar{x}} S(x) = S(\bar{x}),$$

but *inner semicontinuous* (isc) at $\bar{x}$ if

$$\liminf_{x \to \bar{x}} S(x) \supset S(\bar{x}),$$

or equivalently when S is closed-valued,

$$\liminf_{x \to \bar{x}} S(x) = S(\bar{x}).$$

It is called *continuous* at $\bar{x}$ if both conditions hold, i.e., if $S(x) \to S(\bar{x})$ as $x \to \bar{x}$.

The next theorem provides the convergence of solution to convex system which will lead to our continuous result of $C(b) := \{x | Ax = b, x \geq 0\}$

THEOREM 2.3.4 (convergence of solutions to convex system). *[RW09] Let*

$$C^i = \{x \in X^i | L^i(x) \in D^i\}, C = \{x \in X | L(x) \in D\},$$

for linear mappings $L^i, L : \mathbb{R}^n \to \mathbb{R}^m$ and convex sets $X^i, X \subset \mathbb{R}^n$ and $D^i, D \subset \mathbb{R}^m$, such that $L(x)$ cannot be separated from D. If $L^i \to L$, $\liminf_i X^i \supset X$ and $\liminf_i D^i \supset D$, then $\liminf_i C^i \supset C$. Indeed,

$$L^i \to L, X^i \to X, D^i \to D \Rightarrow C^i \to C.$$

Proof. The proof is in [RW09] page 130.

THEOREM 2.3.5. *Let $C : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be set-valued mapping defined by $C(b) := \{x | Ax = b, x \geq 0\}$ where A is full row rank matrix in $\mathbb{R}^n \times \mathbb{R}^m$ and $b \in \mathbb{R}^n$. Then it is continuous.*

Proof: Set $D^i = \{b^i\}$,$L^i(x) = L(x) = Ax$. Applying theorem 2.3.4,we obtain

$$\{x | Ax = b^i\} \to \{x | Ax = b\}, \text{ as } b^i \to b.$$

Since $C(b^i) = \{x | Ax = b^i\} \cap \mathbb{R}^n_+$,

$$C(b^i) \to \{x | Ax = b\} \cap \mathbb{R}^n_+ = C(b)$$

Therefore, $C(b)$ is continuous. **Q.E.D**.

Furthermore, we want $C(b)$ to be not only continuous but also continuous in Pompeiu-Hausdorff sense. To show this is also correct, recall the characterization of Pompeiu-Hausdorff continuity [DR09].

THEOREM 2.3.6 (characterization of PompeiuHausdorff continuity). *[DR09]A set-valued mapping $C : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is Pompeiu-Hausdorff at $\bar{b}$ if $C(\bar{b})$ is closed and both of the following condition hold:*

a. *For every open set $O \subset \mathbb{R}^n$ with $C(\bar{b}) \cap O \neq \emptyset$ there exists a neighborhood V of $\bar{b}$ such that $C(b) \cap O \neq \emptyset$ for all $b \in V$;*
b. *For every open set $O \subset \mathbb{R}^n$ with $C(\bar{b}) \subset O$ there exists a neighborhood V of $\bar{b}$ such that $C(b) \subset O$ for all $b \in V$.*

Proof: The proof is in page 150 of [DR09]

THEOREM 2.3.7. *$C(b)$ defined in Theorem 2.3.5 is Pomeiu-Hausdorff continuous.*

Proof: Clearly $C(b)$ is closed and nonempty (continuous of linear mapping). Condition (a) in Theorem 2.3.6 is satisfied according to Theorem 2.3.5 and Theorem 2.3.2. For every open set $O \subset \mathbb{R}^n$ with $C(\bar{b}) \subset O$, according to theorem 2.3.5, there exists a neighborhood V of $\bar{b}$ such that $C(b) \subset O$ for all $b \in V$. Therefore, $C(b)$ is Pomeiu-Hausdorff continuous.**Q.E.D.**

REMARK 2.3.1. *Another way to obtain similar result is applying the maximum theorem first proved by Claude Berge. The theorem is primarily used in mathematical economics and optimal control.*

THEOREM 2.3.8. *Let X and Θ be topological spaces, $f : X \times \Theta \to \mathbb{R}$ be a continuous function on the product $X \times \Theta$, and $C : \Theta \rightrightarrows X$ be a compact-valued correspondence such that $C(\theta) \neq \emptyset$ for all $\theta \in \Theta$. Define the marginal function (or value function) $f^* : \Theta \to \mathbb{R}$ by*

$$f^* := \sup\{f(x,\theta) : x \in C(\theta)\}$$

and the set of maximizers $C^* : \Theta \rightrightarrows X$ by

$$C^*(\Theta) = \arg\max\{f(x,\theta) : x \in C(\theta)\} = \{x \in C(\theta) : f(x,\theta) = f^*(\theta)\}$$

If C is continuous at θ, then f^* is continuous and C^* is upper semicontinuous with nonempty and compact values.

In fact, Pomeiu-Hausdorff continuous implies continuous. With the condition that the $C(b)$ is bounded, the reverse is also correct. Unfortunately, in our case, $C(b)$ is bounded for each given b but since $b \in \mathbb{R}^n_+$,

Now we continue to show that $S(b) := \underset{x \in C(b)}{\arg\min} f(x)$ is osc at $\bar{b}$ relative to Q where $Q \subset \mathbb{R}^n_+$ is some neighborhood of b.

THEOREM 2.3.9 (basic continuity properties of solution mappings in optimization). *[DR09] Let $\bar{b} \in Q \subset \mathbb{R}^n_+$ be fixed with $C(\bar{b})$ nonempty and bounded,and suppose that:*

a. *The mapping C is Pompeiu-Hausdorff continuous at $\bar{b}$ relative to Q*

b. *The function $\tilde{g}$ is continuous relative to $\mathbb{R}^m \times Q$ at $(\bar{x}, \bar{b})$ for every $\bar{x} \in C(\bar{b})$.*

Then the value mapping $P(b) := \min_{x \in C(b)} g(x)$ is continuous at $\bar{b}$ relative to Q,whereas $S(b)$ is osc at $\bar{b}$ relative Q.

In fact, according to theorem 7.41 in [RW09], we could relax the continuous condition in part b to lsc and the result would still hold. Moreover, since $g(x)$ is convex in our problem, Q could be as large as $domS = domP$.

THEOREM 2.3.10. *$S(b)$ in our problem is osc at $\bar{b}$ relative to Q.*

Proof: According to Theorem 2.3.7, condition a in Theorem 2.3.9 is satisfied. By definition and Theorem 2.3.4, condition b in Theorem 2.3.9 is satisfied. Therefore, we obtain $S(b)$ is osc at $\bar{b}$ relative Q. **Q.E.D**

To complete the proof of the continuity of $x^\dagger(b)$, the following proposition solves the last puzzle.

PROPOSITION 2.3.1 (continuous of distance). *[RW09] For a closed-valued mapping $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and a point $\bar{b}$ in a set $Q \subset \mathbb{R}^n$,*

(1) *S is osc at $\bar{b}$ relative to Q if and only if $\forall u \in \mathbb{R}^m$ the function $b \mapsto d(u, S(b))$ is lsc at $\bar{b}$ relative to Q;*

(2) *S is lsc at $\bar{b}$ relative to Q if and only if $\forall u \in \mathbb{R}^m$ the function $u \mapsto d(u, S(b))$ is usc at $\bar{b}$ relative to Q;*

(3) *S is continuous at $\bar{b}$ relative to Q if and only if $\forall u \in \mathbb{R}^m$ the function $b \mapsto d(u, S(b))$ is continuous at $\bar{b}$ relative to Q*

Proof: Proof is on page 157 [RW09].

Now we could obtain our main result of this section.

THEOREM 2.3.11. ***The Minimum Norm Solution Mapping** $x^\dagger(b)$ **is continuous in the** domS.*

Proof: First we want to show that the function $u \mapsto d(u, S(b))$ is usc at $\bar{b}$ relative to Q. Define $h(u) := d(u, S(b))$. Recall that $h(u)$ is usc at u for $\bar{b}$ relative to Q if and only if $\limsup_{u^i \to u} h(u^i) \leq h(u)$. Since $u^i \to u$, there exists a V such that when i is large enough, $u^i \in V$ as well. By the triangle inequality,for any $z \in S(b)$,

$$d(u^i, z) \leq d(u, z) + d(u, u^i)$$

Since $S(b)$ is nonempty and compact,take the minimum over $z \in S(b)$ in both sides of the above inequility,

$$d(u^i, S(b)) \leq d(u, S(b)) + d(u^i, u) \implies h(u^i) \leq h(u) + d(u^i, u).$$

Hence, the function $u \mapsto d(u, S(b))$ is usc at $\bar{b}$ relative to Q so that S is lsc at $\bar{b}$ relative to Q according to Proposition 2.3.1 part (2). Together with Theorem 2.3.10, S is continuous at $\bar{b}$ relative to Q. Hence, $\forall u \in \mathbb{R}^m$ the function $b \mapsto d(u, S(b)$ is continuous at $\bar{b}$ relative to Q. In fact, since $g(x)$ is convex, $Q = int(domS)$. Since $x^\dagger(b) := P_{S(b)}(0) = \min_{x \in S(b)} \|x\| = d(0, S(b))$, $x^\dagger(b)$ is continuous by setting $u = 0$. **Q.E.D.**

COROLLARY 2.3.1. *$x^\dagger(b)$ is a homeomorphism between $\mathbb{R}^n$ and $\mathcal{D}^\dagger$.*

Proof: According to Corollary 2.2.1, $x^\dagger(b)$ is a bijection between $\mathbb{R}^n$ and $\mathbb{R}^m$. According to Theorem 2.3.11, $x^\dagger(b)$ is continuous. Since the inverse of $x^\dagger(b)$ is Ax which is a linear map, it is continuous. Therefore, $x^\dagger(b)$ is a homeomorphism between $\mathbb{R}^n$ and $\mathcal{D}^\dagger$. **Q.E.D.**

In fact, according to the proof of theorem 2.3.11, we could define projection-based solution mapping for arbitrary u not just when $u = 0$.

DEFINITION 2.3.12 (projection-based solution mapping). Consider the problem (2.1) and let $C_b := \{Ax = b, x \geq 0\}$ and $S(b) := \arg\min_{x \in C_b} g(x)$. Then if $g(x)$ is coercive and lower-semicontinuous, the projection-based solution mapping $x_u^\dagger(b)$ is defined by

$$x_u^\dagger(b) := P_{S(b)}(u)$$

where $P_{S(b)}$ is the projection operator on $S(b)$.
In fact $x_u^\dagger(b)$ is the solution of the following problem:

$$\begin{aligned} \min_y \quad & \|y - u\| \\ & y \in S(b) \end{aligned}$$

THEOREM 2.3.13 (Projection theorem on Hilbert Space). *For every x in a Hilbert space H and every nonempty closed convex set $C \subseteq H$, there exists a unique vector $y \in C$ for which $\|x - z\|$ is minimized over the vectors $z \in C$.*

Based on the Projection theorem and above discussion, we could easily obtain the following corollary.

COROLLARY 2.3.2. *The projection-based solution mapping $x_u^\dagger(b)$ is well-defined. If $\mathcal{D}_u^\dagger$ is the image of the projection-based solution mapping $x_u^\dagger(b)$, then $x_u^\dagger(b)$ is a homeomorphism between $\mathbb{R}^n$ and $\mathcal{D}_u^\dagger$.*

Let $u(\theta)$ be a continuous function, then we could end up with a more general solution mapping by composition.

COROLLARY 2.3.3. *The projection-based solution mapping $x_{u(\theta)}^\dagger(b)$ is well-defined when $u(\theta)$ is a single-valued function of θ. If $\mathcal{D}_{u(\theta)}^\dagger$ is the image of the projection-based solution mapping $x_{u(\theta)}^\dagger(b)$, then $x_{u(\theta)}^\dagger(b)$ is a homeomorphism between $\mathbb{R}^n$ and $\mathcal{D}_{u(\theta)}^\dagger$ for each θ.*

Remark: Comparing the definition of the MNSM and projection-based solution mapping, it is easy to notice that those solution mappings could be connected by a simple translation. In this dissertation, we do not focus much on projection-based solution mapping due to this reason.

2.4. Connectivity of the Partitions in Demand Space

In this section, we plan to explain the key observation of the MNSM. One difficulty associated with stochastic optimization is that the problem is infinite dimensional if the uncertain parameter is not discrete. When the parameter is discrete, we could find a deterministic equivalent optimization problem which could be solved using decomposition methods such as L-shaped Algorithm and Bender's decomposition [Bnn62] or using scenario decomposition such as Progressive hedging method [RWW$^+$13]. Yet, when the parameter is not discrete but continuous, there is no such equivalence. To conquer this difficulty, we choose to study the finite partitions of the parameter space instead of infinite possible realizations. It makes this possible to decompose the original problem by discussing the different combination of activated inequality constraints. This allows us to consider finite many subproblems of the stochastic optimization with continuous parameter or infinitely many scenarios.

In the following, we will show how the Minimum Norm Solution Mapping could help us form such partition. Let $1:m := \{i|\forall i \in \mathbb{N}, i \leq m, m \in \mathbb{N}\}$,$\mathcal{P}(1:m)$ be the power set of $1:m$, i.e. the collection of all the subsets of $1:m$, and $\mathcal{A}_x$ be the active set, i.e. $\mathcal{A}_x := \{i|x_i = 0\}$. Now treat $\mathcal{A}_x$ as a set-valued mapping which maps $\mathbb{R}^n_+$ to $\mathcal{P}(1:m)$ such that $\mathcal{A}_x(b) := \{i|x^\dagger(b)_i = 0\}$. Then for any $U \in \mathcal{P}(1:m)$, define $\mathcal{R}_U := \{b|\mathcal{A}_x(b) = U\}$.

THEOREM 2.4.1. *The collection of $\mathcal{R}_U$ for all the $U \in \mathcal{P}(1:m)$ forms a partition of X.*

Proof: To show a collection forms a partition, we need to show that the union of this collection is the whole space while the intersection between different elememts is empty. Since $x^\dagger(b)$ is surjective, it turns out that

$$\bigcup_{U \in \mathcal{P}(1:m)} \mathcal{R}_U = X$$

Let $U, V \in \mathcal{P}(1:m)$ with $U \neq V$. Assume that $\mathcal{R}_U \cap \mathcal{R}_V \neq \emptyset$,then there exists a b' such that $\mathcal{A}_x(b') = U$ and $\mathcal{A}_x(b') = V$ which contradicts the fact that $U \neq V$. Hence, $\mathcal{R}_U \cap \mathcal{R}_V = \emptyset$.Therefore, it forms a partition. **Q.E.D.**

It is easy to see that the number of nonempty $\mathcal{R}_U$ for all $U \in \mathcal{P}(1:m)$ is at most 2^m. However, since we want $C(b)$ to be feasible for any b, the upper bound could reduce to $\begin{pmatrix} m \\ n \end{pmatrix} 2^{m-n}$. As we will show in numerical examples later, in practice due to some properties of $x^\dagger(b)$ this number is much smaller.

Corollary 2.4.1. *Let $U^i, U^j \in \mathcal{P}(1:m)$ be two different sequences with $U^i \to U$ and $U^j \to U$. Then*

$$\lim \mathcal{R}_{U^i} = \mathcal{R}_U = \lim \mathcal{R}_{U^j} \tag{2.4}$$

Proof: This is directly from the definition and theorem 2.3.4.

Corollary 2.4.2. *$\mathcal{R}_U$ is connected set for any $U \in \mathcal{P}(1:m)$.*

Proof: This is directly from the result that $x^\dagger(b)$ is a homeomorphism.

These corollaries allow us to create an algorithm to find $\mathcal{R}_U$ which does not depend on the choice of the order of activating the inequality constraints.

Therefore, we could decompose the problem (2.1) into finitely many subproblems systematically through the MNSM. In the next chapter, we will show that when (2.1) is a convex quadratic optimization problem, we could find an analytic formulation of the MNSM.

CHAPTER 3

MNSM for Convex Quadratic Optimization (CQO)

The MNSM provides a systematic way to find a representative of the solution set of equilibrium problems with good properties. It must be pointed out that there exist infinitely many ways to find the representative of the solution set of equilibrium problems. The advantages of the MNSM are not only the homeomorphism properties but also we could find an analytic formula when the objective function of the equilibrium is convex and quadratic. In this chapter, we are going to figure out the analytic formula of the MNSM for convex quadratic programming problem.

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & \tfrac{1}{2}x^TQx + c^Tx \\ \text{subject to} \quad & Ax = b \\ & x \geq 0 \end{aligned} \tag{3.1}$$

where $Q \in \mathbb{R}^{n\times n}$ is symmetric positive semi-definite matrix, $A \in \mathbb{R}^{m\times n}$ is full-row rank matrix, $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$.

3.1. The MNSM for a given region

If we already know that the parameter b in problem 3.1 belongs to a given region where certain inequalities are active, the parametric optimization problem with both equality and inequality constraints is equivalent to the one with only equality constraints. Hence, with the help of the well known KKT conditions for convex optimization problem we are able to derive the analytic formula of the MNSM for convex quadratic optimization problem. Now let's recall the KKT conditions.

Theorem 3.1.1 (Karush-Kuhn-Tucker (KKT) conditions). *[BV04] Given the convex optimization problem*

$$\min_{x \in \mathbb{R}^n} \quad f(x) \tag{3.2}$$

$$\textit{subject to} \quad g(x) \leq 0 \tag{3.3}$$

$$h(x) = 0 \tag{3.4}$$

where $f(x), g(x)$ *is convex and* $h(x)$ *is affine. If there exists a point* x *such that* $h(x) = 0$ *and* $g(x) < 0$*, then* x^* *is a minimum solution if and only if there exists a* λ *and* μ *such that*

$$0 \in \partial f(x^*) + \lambda^T \partial g(x^*) + \mu^T \partial h(x^*) \qquad (\textit{stationary condition}) \tag{3.5}$$

$$g(x^*) \leq 0, h(x^*) = 0 \qquad (\textit{primal feasible}) \tag{3.6}$$

$$\lambda \geq 0 \qquad (\textit{dual feasible}) \tag{3.7}$$

$$\lambda_i g_i(x^*) = 0, \forall i = 1, \ldots, m \quad (\textit{complementary slackness}) \tag{3.8}$$

Apply the KKT conditions to problem 3.1, we could obtain

$$Qx + A^T \mu + \lambda = -c \tag{3.9}$$

$$Ax = b, x \geq 0 \tag{3.10}$$

$$\lambda \geq 0 \tag{3.11}$$

$$\lambda_i x_i = 0, \forall i = 1, \ldots, n \tag{3.12}$$

where $Q \in \mathbb{R}^{n \times n}$ is symmetric positive semi-definite matrix, $A \in \mathbb{R}^{m \times n}$ is full-row rank matrix, $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$.

If we choose $b \in \mathcal{R}_\emptyset = \{b \mid x^\dagger(b) > 0\}$ such that no inequality constraints is activated, i.e. $x_i > 0$ for every $i = 1, \ldots, n$, then the KKT conditions reduce to

$$\begin{pmatrix} Q & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} x \\ \mu \end{pmatrix} = \begin{pmatrix} -c \\ b \end{pmatrix} \tag{3.13}$$

i.e. x is in the solution set of convex quadratic problem for $b \in \mathcal{R}_\emptyset$ if and only if there exists a μ such that the vector $(x, \mu)^T$ solves (3.13).

THEOREM 3.1.2. *If $b \in \mathcal{R}_\emptyset$ such that the feasible set of CQP is non-empty and no inequality constraint is activate, then finding the corresponding MNSM is equivalent to solving the following problem*

$$x^\dagger(b) := argmin \left\{ \|x\|_2 \middle| \begin{pmatrix} Q & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} x \\ \mu \end{pmatrix} = \begin{pmatrix} -c \\ b \end{pmatrix}, x \geq 0, \mu \in \mathbb{R}^m \right\} \tag{3.14}$$

Proof: Since the solution set of linear system is affine, the set of all the $(x, \mu)^T$ is a convex subset in $\mathbb{R}^n_+ \times \mathbb{R}^m$. Then the image of natural projection : $(x, \mu) \mapsto x$ of this set is still convex. According to the Projection theorem on Hilbert space, there exists a unique element having the minimum norm which is $x^\dagger(b)$ following by the definition of MNSM. **Q.E.D**

To derive a more explicit formula of the MNSM for CQP, we need the following lemmas.

LEMMA 3.1.3. *If A is full row rank, then*

$$\left\{ x \middle| \begin{pmatrix} Q & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} x \\ \mu \end{pmatrix} = \begin{pmatrix} -c \\ b \end{pmatrix}, x \geq 0, \mu \in \mathbb{R}^m \right\} = \left\{ x \middle| \begin{pmatrix} P^\perp_{A^T} Q \\ A \end{pmatrix} x = \begin{pmatrix} -P^\perp_{A^T} c \\ b \end{pmatrix}, x \geq 0 \right\}$$

Proof: Since A is full row rank, $\mathcal{N}(A^T) = \{0\}$. Hence

$$\exists \mu \text{ s.t. } Qx + A^T\mu = -c \tag{3.15}$$

$$\iff \quad Qx + c \in \mathcal{R}(A^T) \tag{3.16}$$

$$\iff \quad P^\perp_{A^T}(Qx + c) = 0 \tag{3.17}$$

$$\iff \quad P^\perp_{A^T} Qx = -P^\perp_{A^T} c \tag{3.18}$$

where $P^\perp_{A^T} = I - A^T(AA^T)^{-1}A$. Therefore, these two sets are identical. **Q.E.D.**

REMARK 3.1.1. *In fact, the condition in above lemma is not necessary if we redefine the projection $P^\perp_{A^T}$ as $I - A^T A^\dagger$ which is the orthogonal projector onto the kernel of A^T.*

LEMMA 3.1.4 (The least norm solution). *Let $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$ and suppose that $AA^\dagger b = b$. Then any vector of the form*

$$x = A^\dagger b + (I - A^\dagger A)y, \qquad \text{where } y \in \mathbb{R}^n \text{ is arbitrary} \tag{3.19}$$

is a solution of

$$Ax = b$$

Furthermore, all the solution of $Ax = b$ are of this form and the least norm solution is given by $x = A^\dagger b$, i.e. $x = A^\dagger b$ is the solution of

$$\min_{x} \quad \|x\|_2 \tag{3.20}$$

$$\textit{subject to} \quad Ax = b \tag{3.21}$$

The following theorem gives the analytic formula of the MNSM for CQP with no inequality activated.

THEOREM 3.1.5. *If $b \in \mathcal{R}_\emptyset$, then the MNSM of CQP 3.1 is*

$$x^\dagger(b) = M_\emptyset b + N_\emptyset \tag{3.22}$$

where $K_\emptyset = \begin{bmatrix} P_{A^T}^\perp Q \\ A \end{bmatrix}$, $M_\emptyset = K_\emptyset^\dagger(:, n + (1 : m))$, *and* $N_\emptyset = -K_\emptyset^\dagger(:, 1 : n)P_{A^T}^\perp c$.

REMARK 3.1.2. *Here, $K_\emptyset^\dagger(:, n + (1 : m))$ is MATLAB notation which means the columns of $K_\emptyset^\dagger$ from $(n + 1)^{th} to (n + m)^{th}$.*

Proof: Since $b \in \mathcal{R}_\emptyset$, no inequality constraints are activated. Then the CQP is equivalent to

$$\min_{x} \quad \frac{1}{2}x^T Qx + c^T x \tag{3.23}$$

$$\text{subject to} \quad Ax = b \tag{3.24}$$

Applying the above theorem 3.1.2 and lemma 3.1.3, 3.1.4, we obtain the result.**Q.E.D**
Follow the same idea, we could obtain the analytic formula of the MNSM for CQP with certain inequality activated.

THEOREM 3.1.6. *If $b \in \mathcal{R}_U$, that is $x_j = 0$ for all $j \in U$, the MNSM of CQP is*

$$x^\dagger(b) = M_U b + N_U \tag{3.25}$$

where $E_U = I_n(U,:)$,$L = \begin{bmatrix} A \\ E_U \end{bmatrix}$,$K_U = \begin{bmatrix} P^{\perp}_{L^T} Q \\ L \end{bmatrix}$,$M_U = K_U^{\dagger}(:, m + (1:n))$, *and* $N_U = -K_U^{\dagger}(:, 1:m) P^{\perp}_{L^T} c$.

Proof: Since $b \in \mathcal{R}_U$, $x_j = 0$ for all $j \in U$. The CQP is equivalent to

$$\min_x \quad \frac{1}{2} x^T Q x + c^T x \tag{3.26}$$

$$\text{subject to} \quad Ax = b \tag{3.27}$$

$$x_j = 0 \quad \forall j \in U \tag{3.28}$$

Now define $E_U = I_n(U,:)$ then

$$\min_x \quad \frac{1}{2} x^T Q x + c^T x \tag{3.29}$$

$$\text{subject to} \quad Ax = b \tag{3.30}$$

$$E_U x = 0 \tag{3.31}$$

Let $L = \begin{bmatrix} A \\ E_U \end{bmatrix}$ and $\tilde{b} = \begin{bmatrix} b \\ 0 \end{bmatrix}$,

$$\min_x \quad \frac{1}{2} x^T Q x + c^T x \tag{3.32}$$

$$\text{subject to} \quad Lx = \tilde{b} \tag{3.33}$$

Applying the above theorem 3.1.2 and lemma 3.1.3, 3.1.4, we obtain the result. **Q.E.D**

The following corollaries are motivated by the physical meaning of the MNSM. In this case, there is only one path for a given OD-pair so that the only possible choice for traveler is to choose the remain path.

Corollary 3.1.1. *Let A be the OD-path incidence matrix. If* $L = \begin{bmatrix} A \\ E_U \end{bmatrix}$ *is non-singular, then*

$$x^{\dagger}(b) = E^T_{U^c} b \tag{3.34}$$

where $E_U = I_n(U,:)$ *and* $U^c = 1:n - U$

Proof: Since L is non-singular, the feasible set $\{Ax = b, E_U x = 0, x \geq 0\}$ has only one element which is the solution to the minimum problem as well as the minimum norm solution. Solving the equation

$$\begin{matrix} Ax = b \\ E_U x = 0 \end{matrix} \implies x^\dagger(b) = E_{U^c}^T b \tag{3.35}$$

where $U^c = 1 : n - U$. **Q.E.D**

As a consequence of the above corollary, we obtain the following result which provides a quick result for the MNSM when the partition region corresponds to the case that there is only one path for a given OD-pair.

Corollary 3.1.2. *Let A be the OD-path incidence matrix. If $L = \begin{bmatrix} A \\ E_U \end{bmatrix}$ is non-singular, then*

$$L^\dagger(:, 1 : n) = E_{U^c}^T \tag{3.36}$$

Proof: Since L is non-singular, then $P_{L^T}^\perp$ maps every vector into $Ker(L) = \{0\}$

Then according to theorem 3.1.6 and corollary 3.1.1

$$K_U = \begin{bmatrix} P_{L^T}^\perp Q \\ L \end{bmatrix} = \begin{bmatrix} 0 \\ L \end{bmatrix} \implies K_U^\dagger = \begin{bmatrix} 0 & L^\dagger \end{bmatrix} \implies L^\dagger = E_{U^c}^T \tag{3.37}$$

Q.E.D

Now consider the convex quadratic optimization problem (CQP) with box constraints to prepare for further extension of the MNSM.

$$\min_{x \in \mathbb{R}^n} \quad \tfrac{1}{2} x^T Q x + c^T x \tag{3.38}$$

$$\text{subject to} \quad Ax = b \tag{3.39}$$

$$l \leq Bx \leq u, x \geq 0 \tag{3.40}$$

where Q is positive semi-definite, both A and B are full row rank matrix.

As with the definition of $\mathcal{R}_U$, define $\mathcal{R}_{U,V,g}$ as follow :

$$\mathcal{R}_{U,V,g} := \{b| \quad x^\dagger(b)_i = 0 \text{ if } i \in U \text{ and } x^\dagger(b)_j > 0 \text{ if } j \notin U \tag{3.41}$$

$$[Bx^\dagger(b)]_i = g_i \text{ if } i \in V \text{ and } l_i < [Bx^\dagger(b)]_i < u_i \text{ if } i \notin V\} \tag{3.42}$$

where $g_i = l_i$ if $[Bx]_i = l_i$ is activated and $g_i = u_i$ if $[Bx]_i = u_i$ is activated.

THEOREM 3.1.7 (The MNSM for CQP with general box constraints). *If $b \in \mathcal{R}_{U,V,g}$, the MNSM of CQP with general box constraints (3.38) is*

$$x^{\dagger}(b) = M_{U,V}b - N_{U,V}c + H_{U,V}g \tag{3.43}$$

where $E_U = I_N(U,:)$, $B_V = B(V,:)$, $L = \begin{bmatrix} A \\ E_U \\ B_V \end{bmatrix}$,$K_{U,V} = \begin{bmatrix} P_{L^T}^{\perp}Q \\ L \end{bmatrix}$, $M_{U,V} = K_{U,V}^{\dagger}(:, m + (1:n))$, $N_{U,V} = K_{U,V}^{\dagger}(:, 1:m)P_{L^T}^{\perp}$ *and* $H_{U,V} = K_{U,V}^{\dagger}(:, (m+n+1):end))$.

Proof: Since $b \in \mathcal{R}_{U,V,g}$, the CQP with general box constraints (3.38) is equivalent to

$$\min_{x \in \mathbb{R}^n} \quad \tfrac{1}{2}x^TQx + c^Tx \tag{3.44}$$

$$\text{subject to} \quad Ax = b \tag{3.45}$$

$$E_Ux = 0 \tag{3.46}$$

$$B_Vx = g \tag{3.47}$$

If $\begin{bmatrix} b \\ 0 \\ g \end{bmatrix} \notin R\left(\begin{bmatrix} A \\ E_U \\ B_V \end{bmatrix}\right)$, then the feasible region is empty so that there is no solution.

Now, we only consider that $\begin{bmatrix} b \\ 0 \\ g \end{bmatrix} \in R\left(\begin{bmatrix} A \\ E_U \\ B_V \end{bmatrix}\right)$. Applying theorem 3.1.2 and lemmas 3.1.3, 3.1.4, we obtain the result. **Q.E.D**

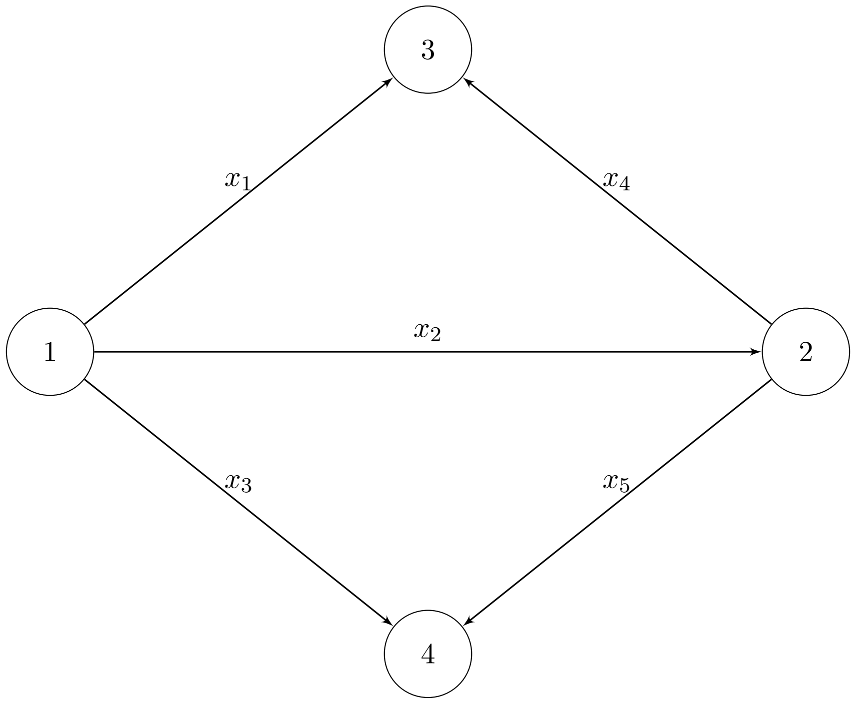


FIGURE 3.1. Network Topology

EXAMPLE 4. Consider the network shown in figure 3.1. The link costs are :

$$
\begin{aligned}
x_1 \quad &: \quad t_1(s) = s/7 + 82 \\
x_2 \quad &: \quad t_2(s) = s/5 + 32 \\
x_3 \quad &: \quad t_3(s) = s/4 + 82 \\
x_4 \quad &: \quad t_4(s) = 0 \\
x_5 \quad &: \quad t_5(s) = 0
\end{aligned}
$$

where the coefficients are randomly generated for the first three links and the last two links are free links.

In this network, we consider two O-D pairs, $q_1 : 1 \to 3$ and $q_2 : 1 \to 4$ and the corresponding paths, $f_1 : \{x_1\}$, $f_2 : \{x_2, x_4\}$, $f_3\{x_2, x_5\}$ and $f_4 : \{x_3\}$. Hence, the incidence matrix between O-D pairs and paths is given by

$$B = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

and the incidence matrix between links and paths is given by

$$F = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

So the following conservation relations are satisfied :

$$x = Ff, q = Bf$$

Then, the UE for this network could be formulated as follows:

$$\min_{x} \quad \frac{1}{2} f^T Q f + c^T f \tag{3.48}$$

$$\text{subject to} \quad Af = q \tag{3.49}$$

$$f \geq 0 \tag{3.50}$$

where $Q = F^T \begin{bmatrix} \frac{1}{7} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{5} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} F$ and $c = F^T \begin{bmatrix} 82 \\ 32 \\ 82 \\ 0 \\ 0 \end{bmatrix}$.

Since in the situation of UE, all the demand have to be assigned.

All the possible situations are: $\{f_1, f_3\}, \{f_1, f_4\}, \{f_1, f_3, f_4\}, \{f_2, f_3\}, \{f_2, f_4\}, \{f_2, f_3, f_4\}$, $\{f_3, f_1, f_2\}, \{f_4, f_1, f_2\}$, and $\{f_1, f_2, f_3, f_4\}$.

Therefore, $U \in \{\{2, 4\}, \{2, 3\}, \{2\}, \{1, 4\}, \{1, 3\}, \{1\}, \{4\}, \{3\}, \{\emptyset\}\}$.

If $U = \emptyset$, then according to theorem 3.1.5

$$f^{\dagger}(b) = M_{\emptyset} q + N_{\emptyset} = \begin{bmatrix} 7/16 & 7/16 \\ 9/16 & -7/16 \\ -1/4 & 3/4 \\ 1/4 & 1/4 \end{bmatrix} q + \begin{bmatrix} -875/8 \\ 875/8 \\ 125/2 \\ -125/2 \end{bmatrix} \tag{3.51}$$

If $U = \{2\}$, then according to theorem 3.1.6

$$f^{\dagger}(b) = M_U q + N_U = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 5/9 \\ 0 & 4/9 \end{bmatrix} q + \begin{bmatrix} 0 \\ 0 \\ 1000/9 \\ -1000/9 \end{bmatrix} \tag{3.52}$$

If $U = \{1, 3\}$, then according to corollary 3.1.1

$$f^{\dagger}(b) = M_U q + N_U = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} q \tag{3.53}$$

3.2. Determine the Partition Region

In this section, we establish the criteria that could determine the partition region directly. To derive those criteria, we need to find the boundary of each given partition region such that the MNSM satisfies the current active set and meanwhile the MNSM would not trigger other iequalities.

THEOREM 3.2.1. *The partition region $\mathcal{R}_U = \{b \mid x^{\dagger}(b)_i = 0 \text{ if } i \in U\}$ for problem 3.1 is given by if A and E_U are linearly independent,*

$$-\left[(L^T)^{\dagger}\left(QM_U b + QN_U + c\right)\right]_j \geq 0 \quad \text{for } j = 1 + n, \ldots, |U| + n. \tag{3.54}$$

$$[M_U b + N_U]_i \geq 0 \qquad \forall i \notin U \tag{3.55}$$

$$\tag{3.56}$$

if A and E_U are linearly dependent,

$$-\left[(\tilde{L}^T)^{\dagger}\left(QM_Ub + QN_U + c\right)\right]_j \geq 0 \quad \text{for } j = 1+n, \ldots, |U|+n. \tag{3.57}$$

$$[M_Ub + N_U]_i \geq 0 \qquad\qquad \forall i \notin U \tag{3.58}$$

where $L = \begin{bmatrix} A \\ E_U \end{bmatrix}$ *and* $\tilde{L} = \begin{bmatrix} \tilde{A} \\ E_U \end{bmatrix}$ *where $\tilde{A}$ is consisted by those rows of A which are linearly independent with the rows of E_U.*

Proof: To determine the partition region $\mathcal{R}_U$, recall the KKT conditions for CQP,

$$Qx + A^T\mu + \lambda = -c \tag{3.59}$$

$$Ax = b, x \geq 0 \tag{3.60}$$

$$\lambda \geq 0 \tag{3.61}$$

$$\lambda_i x_i = 0, \forall i = 1, \ldots, n \tag{3.62}$$

where $b \in \mathcal{R}_U$, $\mathcal{A}_x(b) = \{i | x^{\dagger}(b)_i = 0\} = U$.

So by complementary slackness condition and the MNSM for CQP, we have for $i \in U$, $x^{\dagger}(b)_i = 0 \implies \lambda_i \geq 0$ and for $i \notin U$, $x^{\dagger}(b)_i \geq 0 \implies \lambda_i = 0$.

Hence,

(1) $i \notin U$. According to theorem 3.1.6, when $b \in \mathcal{R}_U$,

$$[x^{\dagger}(b)]_i = [M_Ub + N_U]_i \geq 0 \tag{3.63}$$

(2) $i \in U$. Let $L = \begin{bmatrix} A \\ E_U \end{bmatrix}$. Then equation 3.59 becomes

$$Qx + L^T \begin{bmatrix} \mu \\ \lambda \end{bmatrix} = -c \tag{3.64}$$

If L is full row rank, LL^T is invertable.

Multiplying L on both sides, we obtain

$$Qx + L^T \begin{bmatrix} \mu \\ \lambda \end{bmatrix} = -c \implies LL^T \begin{bmatrix} \mu \\ \lambda \end{bmatrix} = -L\left(Qx + c\right) \implies \begin{bmatrix} \mu \\ \lambda \end{bmatrix} = -(LL^T)^{-1}L\left(Qx + c\right) \tag{3.65}$$

Since $(L^T)^\dagger = (LL^T)^{-1}L$ and $x^\dagger(b) = M_U b + N_U$ for $b \in \mathcal{R}_U$, we have got

$$\begin{bmatrix} \mu \\ \lambda \end{bmatrix} = -(L^T)^\dagger \left(QM_U b + QN_U + c\right) \tag{3.66}$$

Since $x^\dagger(b)_i = 0$, we have

$$\lambda_i \geq 0 \implies -\left[(L^T)^\dagger \left(QM_U b + QN_U + c\right)\right]_j \geq 0 \quad \text{for } j = 1+n, \ldots, |U| + n. \tag{3.67}$$

On the other hand, if L is not full row rank, the corresponding feasible region is given by

$$Ax = b, E_U x = 0, x_i > 0 \quad \text{if } i \notin U \tag{3.68}$$

Since L is not full row rank matrix, the feasible region is empty if $\begin{bmatrix} b \\ 0 \end{bmatrix} \notin R(L)$. If $\begin{bmatrix} b \\ 0 \end{bmatrix} \in R(L)$ and L is not full row rank matrix, there must be some rows of A such that these rows are linearly independent with the rows of E_U since both A and E_U are full row rank matrix.

Hence those rows of A who are linearly dependence with the rows of E_U are redundant constraints for the feasible region. Let $\tilde{A}$ consist of those rows who are linearly independent with the rows of E_U. Then in this case $\tilde{L} = \begin{bmatrix} \tilde{A} \\ E_U \end{bmatrix}$ is full row rank matrix. Repeating the above process we have

$$\lambda_i \geq 0 \implies -\left[(\tilde{L}^T)^\dagger \left(QM_U b + QN_U + c\right)\right]_j \geq 0 \quad \text{for } j = 1+n, \ldots, |U| + n. \tag{3.69}$$

Therefore, putting 3.83, 3.88 and 3.89 together, we obtain the criteria to determine the partition region $\mathcal{R}_U$.

If A and E_U are linearly independent,

$$-\left[(L^T)^\dagger (QM_U b + QN_U + c)\right]_j \geq 0 \quad \text{for } j = 1+n, \ldots, |U|+n. \tag{3.70}$$

$$[M_U b + N_U]_i \geq 0 \qquad \forall i \notin U \tag{3.71}$$

$$\tag{3.72}$$

If A and E_U are linearly dependent,

$$-\left[(\tilde{L}^T)^\dagger (QM_U b + QN_U + c)\right]_j \geq 0 \quad \text{for } j = 1+n, \ldots, |U|+n. \tag{3.73}$$

$$[M_U b + N_U]_i \geq 0 \qquad \forall i \notin U \tag{3.74}$$

Q.E.D

EXAMPLE 5. Consider the network showing in figure 3.1. Now applying the methods we provide, we have the demand space of the UE separated into 4 parts shown in figure **??**

$$\begin{aligned}
\mathcal{R}_\emptyset &:= \{q|q_1 + q_2 - 250 \geq 0, 9q_1 - 7q_2 + 1750 \geq 0, q_1 - 3q_2 - 250 \geq 0\} \\
\mathcal{R}_{\{2\}} &:= \{q|9q_1 - 7q_2 + 1750 < 0, q_1 \geq 0\} \\
\mathcal{R}_{\{3\}} &:= \{q|q_1 - 3q_2 - 250 < 0, q_1 \geq 250\} \\
\mathcal{R}_{\{1,4\}} &:= \{q|q_1 + q_2 - 250 < 0, q_1 \geq 0, q_2 \geq 0\}
\end{aligned}$$

Similarly, we could consider a more complicated case

$$\begin{aligned}
\min_x \quad & \frac{1}{2}x^T Q x + c^T x \\
s.t. \quad & Ax = b, Bx \leq u, x \geq 0
\end{aligned} \tag{3.75}$$

where the feasible set is nonempty, both A and B are full row rank matrix.

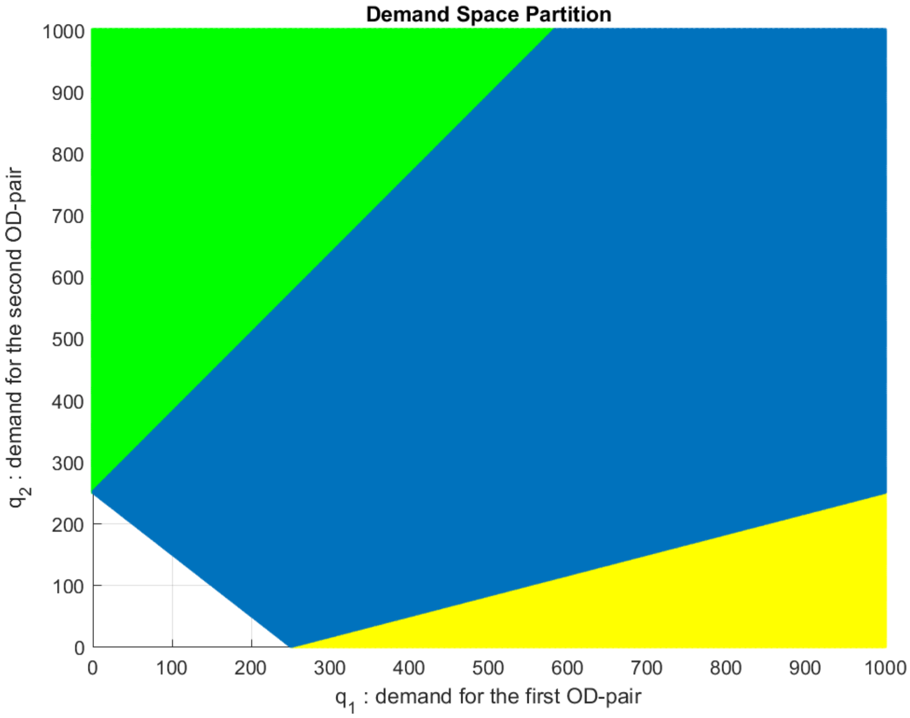


FIGURE 3.2. Demand space partition: $\mathcal{R}_{\emptyset}$ is blue region, $\mathcal{R}_{\{2\}}$ is green region, $\mathcal{R}_{\{3\}}$ is yellow region, $\mathcal{R}_{\{1,4\}}$ is white region

THEOREM 3.2.2. *The partition region* $\mathcal{R}_{U,V} = \{b \mid x^{\dagger}(b)_i = 0,\ \text{if } i \in U \quad Bx^{\dagger}(b)_j = u_j,\ \text{if } j \in V\}$ *for problem 3.75 is given by if* $L = \begin{bmatrix} A \\ E_U \\ B_V \end{bmatrix}$ *is full row rank,*

$$\left[M_{U,V}b + N_{U,V} + H_{U,V}\right]_i \geq 0 \qquad i \notin U \tag{3.76}$$

$$\left[B(M_{U,V}b + N_{U,V} + H_{U,V}) - u\right]_j \leq 0 \qquad j \notin V$$

$$-\left[(L^T)^{\dagger}\left(QM_{U,V}b + QN_{U,V} + QH_{U,V} + c\right)\right]_i \geq 0 \qquad \text{for } i = 1+n, \ldots, |U|+n$$

$$-\left[(L^T)^{\dagger}\left(QM_{U,V}b + QN_{U,V} + QH_{U,V} + c\right)\right]_j \leq 0 \qquad \text{for } j = 1+n+|U|, \ldots, |U|+|V|+n$$

Or

if $L = \begin{bmatrix} A \\ E_U \\ B_V \end{bmatrix}$ *is not full row rank*

$$[M_{U,V}b + N_{U,V} + H_{U,V}]_i \geq 0 \qquad i \notin U \tag{3.77}$$

$$[B(M_{U,V}b + N_{U,V} + H_{U,V}) - u]_j \leq 0 \qquad j \notin V$$

$$-\left[(\tilde{L}^T)^\dagger \left(QM_{\tilde{U},\tilde{V}}b + QN_{\tilde{U},\tilde{V}} + QH_{\tilde{U},\tilde{V}} + c\right)\right]_i \geq 0 \qquad \text{for } i = 1+n,\ldots,|\tilde{U}|+n$$

$$-\left[(\tilde{L}^T)^\dagger \left(QM_{\tilde{U},\tilde{V}}b + QN_{\tilde{U},\tilde{V}} + QH_{\tilde{U},\tilde{V}} + c\right)\right]_j \leq 0 \qquad \text{for } j = 1+n+|\tilde{U}|,\ldots,|\tilde{U}|+|\tilde{V}|+n$$

where $\tilde{L}$ *consist of the rows of* L *who are linear independent.*

Proof: To determine the partition region $\mathcal{R}_{U,V}$, recall the KKT conditions for CQP,

$$Qx + A^T\mu + \lambda + B^T\kappa = -c \tag{3.78}$$

$$Ax = b, x \geq 0 \tag{3.79}$$

$$\kappa \geq 0, \lambda \leq 0 \tag{3.80}$$

$$\lambda_i x_i = 0, \forall i = 1,\ldots,n \tag{3.81}$$

$$\kappa_j(B_U x - u)_j = 0, \forall j = 1,\ldots,n \tag{3.82}$$

where $b \in \mathcal{R}_{U,V}$

So by complementary slackness condition and the MNSM for CQP, we have for $i \in U$, $x^\dagger(b)_i = 0 \implies \lambda_i \leq 0$ and for $i \notin U$, $x^\dagger(b)_i \neq 0 \implies \lambda_i = 0$ and we have for $j \in V$, $[Bx - u]_j = 0 \implies \kappa_j \geq 0$ and for $j \notin V$, $[Bx^\dagger(b) - u]_j \neq 0 \implies \kappa_j = 0$

Hence,

(1)$i \notin U$. According to theorem 3.1.6, when $b \in \mathcal{R}_{U,V}$,

$$[x^\dagger(b)]_i = [M_{U,V}b + N_{U,V} + H_{U,V}]_i \geq 0 \tag{3.83}$$

(2)$j \notin V$. According to theorem 3.1.6, when $b \in \mathcal{R}_{U,V}$,

$$[Bx^\dagger(b) - u]_j = [B(M_{U,V}b + N_{U,V} + H_{U,V}) - u]_j \leq 0 \tag{3.84}$$

(3)$i \in U, j \in V$. Let $L = \begin{bmatrix} A \\ E_U \\ B_V \end{bmatrix}$. Then 3.78 becomes

$$Qx + L^T \begin{bmatrix} \mu \\ \lambda \\ \kappa \end{bmatrix} = -c \tag{3.85}$$

If L is full row rank, LL^T is invertible.

Multiplying L on both sides, we obtain

$$Qx + L^T \begin{bmatrix} \mu \\ \lambda \\ \kappa \end{bmatrix} = -c \implies LL^T \begin{bmatrix} \mu \\ \lambda \\ \kappa \end{bmatrix} = -L\,(Qx + c) \implies \begin{bmatrix} \mu \\ \lambda \\ \kappa \end{bmatrix} = -(LL^T)^{-1}L\,(Qx + c) \tag{3.86}$$

Since $(L^T)^\dagger = (LL^T)^{-1}L$ and $x^\dagger(b) = M_{U,V}b + N_{U,V} + H_{U,V}$ for $b \in \mathcal{R}_U$, we have got

$$\begin{bmatrix} \mu \\ \lambda \\ \kappa \end{bmatrix} = -(L^T)^\dagger\,(QM_{U,V}b + QN_{U,V} + QH_{U,V} + c) \tag{3.87}$$

Since $\mu \in \mathbb{R}^n, i \in U \implies \lambda_i \geq 0$, and $j \in V \implies \kappa \leq 0$, we have

$$\lambda_i \geq 0 \implies -\Big[(L^T)^\dagger\,(QM_{U,V}b + QN_{U,V} + QH_{U,V} + c)\Big]_i \geq 0 \quad \text{for } i = 1+n, \ldots, |U|+n. \tag{3.88}$$

$$\kappa_j \leq 0 \implies -\Big[(L^T)^\dagger\,(QM_{U,V}b + QN_{U,V} + QH_{U,V} + c)\Big]_j \leq 0 \quad \text{for } j = 1+n+|U|, \ldots, |U|+|V|+n \tag{3.89}$$

where $|\cdot|$ is the total number of elements of the set.

If L is not full row rank, the corresponding feasible region is given by

$$Ax = b, E_U x = 0, B_V x = u. \tag{3.90}$$

If $\begin{bmatrix} b \\ 0 \\ u \end{bmatrix} \notin R\left(\begin{bmatrix} A \\ E_U \\ B_V \end{bmatrix}\right)$, then the feasible region is empty so that there is no such region.

If $\begin{bmatrix} b \\ 0 \\ u \end{bmatrix} \in R\left(\begin{bmatrix} A \\ E_U \\ B_V \end{bmatrix}\right)$, we know some equality constraints are redundant so that we could eliminate those equality constraint without changing the feasible region. Then, we first eliminate those equality constrains in $Ax = b$ since we are more interested in those inequality constraints which could be inactivated. After removing redundant, we obtain that $\tilde{L} = \begin{bmatrix} \tilde{A} \\ E_{\tilde{U}} \\ B_{\tilde{V}} \end{bmatrix}$ is full row rank.

Hence we could repeat above process to obtain that

$$-\left[(\tilde{L}^T)^\dagger \left(QM_{\tilde{U},\tilde{V}}b + QN_{\tilde{U},\tilde{V}} + QH_{\tilde{U},\tilde{V}} + c\right)\right]_i \geq 0 \quad \text{for } i = 1+n, \ldots, |\tilde{U}| + n \tag{3.91}$$

$$-\left[(\tilde{L}^T)^\dagger \left(QM_{\tilde{U},\tilde{V}}b + QN_{\tilde{U},\tilde{V}} + QH_{\tilde{U},\tilde{V}} + c\right)\right]_j \leq 0 \quad \text{for } j = 1+n+|\tilde{U}|, \ldots, |\tilde{U}| + |\tilde{V}| + n \tag{3.92}$$

There, the criteria that determine the partition region $\mathcal{R}_{U,V}$ is given by the following :

$$\begin{aligned}
[M_{U,V}b + N_{U,V} + H_{U,V}]_i &\geq 0 && i \notin U \\
[B(M_{U,V}b + N_{U,V} + H_{U,V}) - u]_j &\leq 0 && j \notin V \\
-\left[(L^T)^\dagger (QM_{U,V}b + QN_{U,V} + QH_{U,V} + c)\right]_i &\geq 0 && \text{for } i = 1+n, \ldots, |U| + n \\
-\left[(L^T)^\dagger (QM_{U,V}b + QN_{U,V} + QH_{U,V} + c)\right]_j &\leq 0 && \text{for } j = 1+n+|U|, \ldots, |U| + |V| + n
\end{aligned}$$

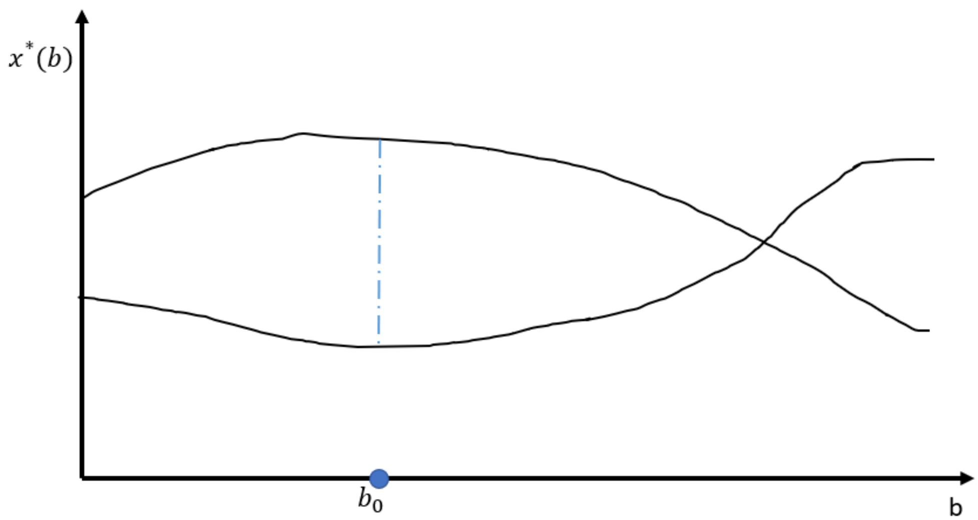
Or

$$\begin{aligned}
[M_{U,V}b + N_{U,V} + H_{U,V}]_i &\geq 0 && i \notin U \\
[B(M_{U,V}b + N_{U,V} + H_{U,V}) - u]_j &\leq 0 && j \notin V \\
-\left[(\tilde{L}^T)^\dagger \left(QM_{\tilde{U},\tilde{V}}b + QN_{\tilde{U},\tilde{V}} + QH_{\tilde{U},\tilde{V}} + c\right)\right]_i &\geq 0 && \text{for } i = 1+n, \ldots, |\tilde{U}| + n \\
-\left[(\tilde{L}^T)^\dagger \left(QM_{\tilde{U},\tilde{V}}b + QN_{\tilde{U},\tilde{V}} + QH_{\tilde{U},\tilde{V}} + c\right)\right]_j &\leq 0 && \text{for } j = 1+n+|\tilde{U}|, \ldots, |\tilde{U}| + |\tilde{V}| + n
\end{aligned}$$

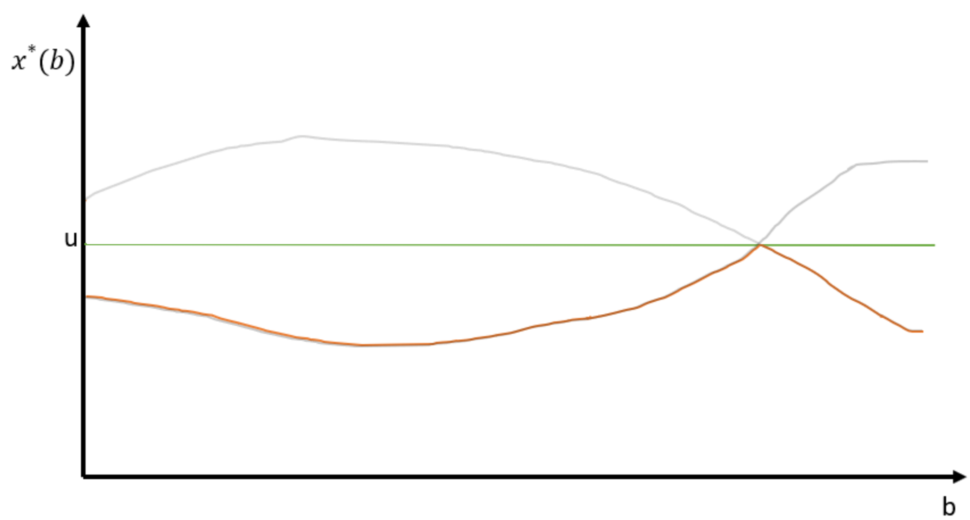
Q.E.D

3.3. Projection-based Solution Mapping

In this section, we would extend the results from previous sections to the projection-based solution mapping with respect to arbitrary point $u \in \mathbb{R}^m$. In fact, when u walks through all the possible points, we could potentially handle almost all solution sets that can arise in the traffic equilibrium problems.



(a) The optimal solution set for all input parameters b. The blue dash line shows the optimal solution set for b_0.



(b) The MNSM vs The PBSM. The orange curve represents the MNSM. The green line represents the PBSM when the basis is u

FIGURE 3.3. Geometry of the Solution Set

Again, there are two steps to obtain the analytic formula, getting the analytic formula for a given region and determining the partition region. According to the definition 2.3.12, the projection-based solution mapping $x_u^\dagger(b)$ with respect to u is given as following:

$$\begin{aligned} \min_x \quad & \|x-u\| \\ s.t. \quad & x \in S(b) \end{aligned} \tag{3.93}$$

where $S(b)$ is the solution set of the problem 3.1.

THEOREM 3.3.1. *If $b \in \mathcal{R}_{u,\emptyset}$, then the projection-based solution mapping with respect to u for the problem 3.1 is*

$$x_u^\dagger(b) = M_\emptyset b + N_\emptyset + (I - K_\emptyset^\dagger K_\emptyset)u \tag{3.94}$$

where $K_\emptyset = \begin{bmatrix} P_{A^T}^\perp Q \\ A \end{bmatrix}$, $M_\emptyset = K_\emptyset^\dagger(:, n+(1:m))$, *and* $N_\emptyset = -K_\emptyset^\dagger(:, 1:n)P_{A^T}^\perp c$.

3.4. Discussion

The solution to equation 3.13 has been studied extensively. Whenever KKT conditions are used to solve the constrained quadratic optimization problem, equation 3.13 would appear. Mainly there are two kinds of methods to solve this problem: Direct methods and Iteration methods. The ideas behind direct methods are usually applying some techniques such as factorization and null space approaching so that we could solve equation 3.13 by solving some smaller size equations. Iteration methods such as precondition conjugate gradient and projected conjugate gradient are sequentially tracking the information given by the initial condition to satisfy the terminal condition. More details could be found in [BV04]. However, neither of those methods meets the requirement of building a numerical stable solution mapping without knowing the initial data or input. In the case of direct method, if we follow the steps of those techniques, one cannot avoid matrix multiplications involving inverse of matrices. It can enhance numerical instability. In the case of iteration method, one needs the information of initial data or input to start the process. With these limitations, we need another approach to get the pre-computed optimal solution. This is a question studied in the area of multiparametric quadratic optimization such as [FJO13]. However, existing approaches are limited to the case of strictly convex quadratic optimization, which has an invertible hessian, and again the analytic formula of the solution mapping involves the inverse of some matrices.

A naive choice to build the solution mapping is applying the pseudoinverse of the left hand side on the equation 3.13. For a problem with unique solution, it turns out that the results of using pseudoinverse may be the same as using the regular inverse. If we limit our framework on the quadratic convex optimization, both choices give a representative of the solution set. However, it is not a good choice by a deep geometric reason. In figure 3.4, the black line is the solution set in

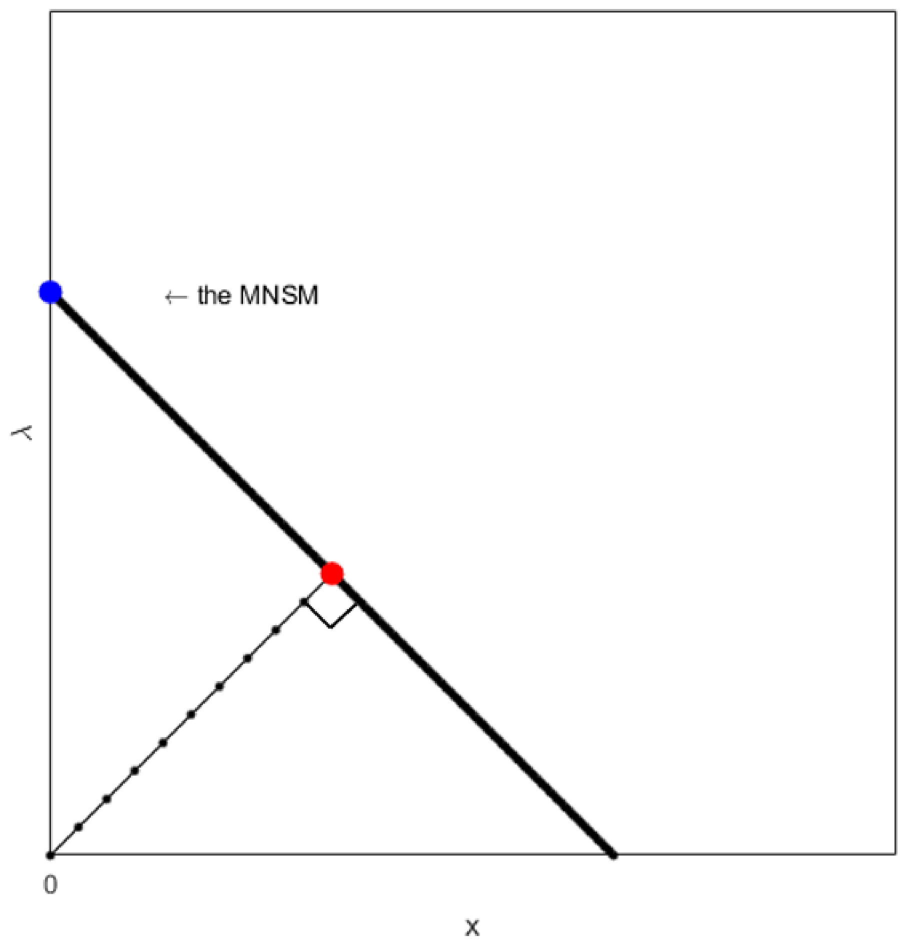


FIGURE 3.4. Difference between the MNSM and pseudoinverse of the KKT matrix on the equation 3.13

the space containing both primal variable and dual variable. If we choose to use the pseudoinverse, we will end up with the red point which has the minimum two norm in the space. But the blue point represents the choice of the MNSM. The difference becomes significant when one tries to extend the result from convex quadratic optimization problems to general convex optimization problems. A natural approach is to use a sequence of quadratic problems to approximate the convex problem. Using the pseudoinverse of the KKT matrix directly implies considering the convergence process in the space both involving primal and dual variables with the topology induced by epi-hypo convergence, which is not separable according to [AW83]. On the other hand, our MNSM approach considers the primal variables only. Hence the approximation process is in the space with only primal variables whose topology is induced by the so-called epi-convergence. This gives great advantages of extending our results from the convex quadratic case to the general convex case, which will be discussed in details in the next Chapter.

CHAPTER 4

Approximated MNSM for Convex Optimization (CO)

In this section, we want to generalize the results obtained from CQP to convex optimization rising from traffic equilibrium problem with general cost functions under congestion. In fact, even though we have an identical definition of the MNSM, we could barely derive an analytic solution for the general convex optimization problem. Hence, we explore the MNSM for a convex optimization problem by using a sequence of the MNSM for a suitable choice of sequence of CQP. Before we go further in the technical methods, we need to introduce concepts of epi-convergence and epi-splines. Those preliminary materials will be covered in the first two sections. In section 3, we first define the so-called approximate MNSM (aMNSM) and then we provide two situations for which the aMNSM matches the MNSM we defined before.

4.1. Epigraph and Epi-convergence

A central question with a sequence of approximated optimization problems is whether they provide a good approximation to the solution of the original optimization problem. That is, ideally we would like for $V_n \to V$, where $V_n = min f_n(x)$ and $V = min f(x)$, and we would also like that for any minimizer sequence $x_n \in \arg\min f_n(x)$, we could have $x \in \arg\min f(x)$ if $x_n \to x$. Unfortunately, the two well-known convergence processes, point-wise convergence and uniform convergence are not suitable for this purpose. At its surface, point-wise convergence seems like an easy exercise but certain pathology could occur as evidenced by the following example : Consider a sequence of optimization problem

$$\begin{aligned} \min \quad & \min\{1-x, 1, 2n|x+1/n|-1\} \\ s.t. \quad & x \in [-1, 1] \end{aligned} \tag{4.1}$$

The objective functions $f_n(x) = \min\{1-x, 1, 2n|x+1/n|-1\}$ converge pointwise to $f(x) = \min\{1-x, 1\}$. It is not hard to figure out that the minimizer of $f_n(x)$ is $x_n = -1/n$ and it converge to zero as $n \to \infty$. However, the minimizer of the optimization problem that we are approximating by this sequence is $x = 1$. The minimizer of the sequence does not converge to the minimizer of the limiting optimization problem. The reason that uniform convergence is not suitable is more straight forward since the requirement is way too strong to achieve in most cases. Hence, epi-convergence comes into the picture. The materials about epigraph and epi-convergence reported in this section are from the related chapter of 'Variational Analysis' [RW09] for the convenience of readers.

Definition 4.1.1 (Epigraph). For $f : \mathbb{R}^n \to \bar{\mathbb{R}}$, the epigraph of f is the set

$$\text{epi}f := \{(x, \alpha) \in \mathbb{R}^n \times \mathbb{R} \mid f(x) \leq \alpha\} \tag{4.2}$$

The epigraph thus consists of all the points of $\mathbb{R}^{n+1}$ lying on or above the graph of f.

Definition 4.1.2 (Epigraph for vector-valued function). For $f : \bar{\mathbb{R}}^n \to \bar{\mathbb{R}}^m$, the epigraph of f is the set

$$\text{epi}f := \left\{(x, \alpha) \in \mathbb{R}^n \times \mathbb{R}^m \mid f(x) \leq_{\mathbb{R}^m_+} \alpha\right\} \tag{4.3}$$

Proposition 4.1.1. *$f : \bar{\mathbb{R}}^n \to \bar{\mathbb{R}}^m$ is lower-semi continuous function if and only if epif is a closed subset of $\mathbb{R}^n \times \mathbb{R}^m$. f is a convex function if and only if epif is a convex set.*

Definition 4.1.3 (lower and upper epi-limits). For any sequence $\{f^k\}_{k\in\mathbb{N}}$ of functions on $\mathbb{R}^n$, the lower epi-limit $\text{e-}\liminf_k f^k$ is the function having as its epigraph the outer limit of the sequence of sets $\text{epi}f^k$:

$$\text{epi}(\text{e-}\liminf_k f^k) := \limsup_k(\text{epi}f^k). \tag{4.4}$$

The upper epi-limit $\text{e-}\limsup_k f^k$ is the function having as its epigraph the inner limit of the sets $\text{epi}f^k$:

$$\text{epi}(\text{e-}\limsup_k f^k) := \liminf_k(\text{epi}f^k). \tag{4.5}$$

Thus $\text{e-}\liminf_k f^k \leq \text{e-}\limsup_k f^k$ in general. When these two functions coincide, the epi-limit function $\text{e-}\lim_k f^k$ is said to exist : $\text{e-}\lim_k f^k :=\text{e-}\liminf_k f^k = \text{e-}\limsup_k f^k$. In this event the functions f^k are said to epi-converge to f. Thus,

$$f^k \xrightarrow{\text{e}} f \Leftrightarrow \text{epi}f^k \to \text{epi}f \tag{4.6}$$

Proposition 4.1.2 (characterization of epi-limits). *Let $\{f^k\}_{k\in\mathbb{N}}$ be any sequence of functions on $\mathbb{R}^n$, and let x be any point of $\mathbb{R}^n$. Then*

$$(e\text{-}\liminf_k f^k)(x) = \min\{\alpha \in \bar{\mathbb{R}} | \exists x^k \to x \text{ with } \liminf_k f^k(x^k) = \alpha\} \tag{4.7}$$

$$(e\text{-}\limsup_k f^k)(x) = \min\{\alpha \in \bar{\mathbb{R}} | \exists x^k \to x \text{ with } \limsup_k f^k(x^k) = \alpha\} \tag{4.8}$$

Thus, $f^k \xrightarrow{e} f$ *if and only if at each point x one has*

$$\begin{cases} \liminf_k f^k(x^k) \geq f(x) & \text{for every sequence } x^k \to x \\ \limsup_k f^k(x^k) \leq f(x) & \text{for some sequence } x^k \to x \end{cases} \tag{4.9}$$

Proposition 4.1.3 (properties of epi-limits). *The following properties hold for any sequence $\{f^k\}_{k\in\mathbb{N}}$ of functions on $\mathbb{R}^n$*

(1) *The functions* e-$\liminf_k f^k$ *and* e-$\limsup_k f^k$ *are lower semi-continuous, and* e-$\lim_k f^k$ *is also lower semi-continuous when it exists.*

(2) *The functions* e-$\liminf_k f^k$ *and* e-$\limsup_k f^k$ *depend only on the sequence* $\{cl f^k\}_{k\in\mathbb{N}}$

(3) *If the sequence $\{f^k\}$ is non-increasing ($f^k \geq f^{k+1}$), then* e-$\lim_k f^k$ *exists and equals* $cl[\inf_k f^k]$*;*

(4) *If the sequence $\{f^k\}$ is non-decreasing ($f^k \leq f^{k+1}$), then* e-$\lim_k f^k$ *exists and equals* $\sup_k cl[f^k]$ *(rather than $cl[\sup_k f^k]$).*

(5) *For subsets C^k and C of $\mathbb{R}^n$, one has $C = \liminf_k C^k$ if and only if $I_C = e\text{-}\limsup_k I_{C^k}$, while one has $C = \limsup_k C^k$ if and only if $I_C = e\text{-}\liminf_k I_{C^k}$*

(6) *If $f_1^k \leq f^k \leq f_2^k$ with $f_1^k \xrightarrow{e} f$ and $f_2^k \xrightarrow{e} f$, then $f^k \xrightarrow{e} f$*

(7) *If $f^k \xrightarrow{e} f$, or just $f =$e-$\liminf_k f^k$, then $dom f \subset \limsup_k[dom f^k]$*

Proposition 4.1.4 (characterization of epi-convergence via minimization). *For functions f^k and f on $\mathbb{R}^n$ with f lsc, one has*

(1) e-$\liminf_k f^k \geq f$ *if and only if* $\liminf_k(\inf_B f^k) \geq \inf_B f$ *for every compact set* $B \subset \mathbb{R}^n$

(2) e-$\limsup_k f^k \leq f$ *if and only if* $\limsup_k(\inf_O f^k) \leq \inf_O f$ *for every open set* $O \subset \mathbb{R}^n$

Thus, e-$\lim_k f^k = f$ *if and only if both conditions hold.*

Let ϵ-$\arg\min f := \{x \mid f(x) \leq \inf f + \epsilon\}$

PROPOSITION 4.1.5 (epigraphical nesting). *If* $e\text{-}\limsup_k f^k \leq f$, *then*

$$\limsup_k (\inf_k f^k) \leq \inf f \tag{4.10}$$

Furthermore, the inclusion

$$\limsup_k (\epsilon^k\text{-}\arg\min f^k) \subset \arg\min f \tag{4.11}$$

holds for any sequence $\epsilon^k \searrow 0$ *such that whenever* $N \in \mathcal{N}_\infty^\#$ *and* $x^k \underset{N}{\to} x$ *with* $x^k \in \epsilon^k\text{-}\arg\min f^k$, *then* $f^k(x^k) \underset{N}{\to} f(x)$

THEOREM 4.1.4 (inf and argmin). *Suppose* $f^k \xrightarrow{e} f$ *with* $-\infty < \inf f < \infty$.

(1) $\inf f^k \to \inf f$ *if and only if there exists for every* $\epsilon > 0$ *a compact set* $B \subset \mathbb{R}^n$ *along with an index set* $N \in \mathcal{N}_\infty$ *such that*

$$\inf_B f^k \leq \inf f^k + \epsilon \text{ for all } k \in N \tag{4.12}$$

(2) $\limsup_k(\epsilon\text{-}\arg\min f^k) \subset \epsilon\text{-}\arg\min f$ *for every* $\epsilon \geq 0$ *and consequently*

$$\limsup_k (\epsilon\text{-}\arg\min f^k) \subset \arg\min f \text{ whenever } \epsilon \searrow 0 \tag{4.13}$$

(3) *Under the assumption that* $\inf f^k \to \inf f$, *there exists a sequence* $\epsilon^k \searrow 0$ *such that* ϵ^k-$\arg\min f^k \to \arg\min f$. *Conversely, if such a sequence exists, and if* $\arg\min f \neq \emptyset$, *then* $\inf f^k \to \inf f$.

Theorem 4.1.4 demonstrated that epi-convergence is the suitable choice of convergence process. The next step is to find a suitable way to generate a sequence of functions which epi-converges to the problem we are interested in. It could be solved by using the epi-splines we introduce in the next section.

4.2. Epi-splines

We review some pertinent facts regarding the space of lower semi-continuous functions before introducing the epi-spline technique based on the series papers by Johannes O. Royset and Roger J-B Wets [RW16, RW12, RW14].

Let $lsc\text{-}fcns(\mathbb{R}^n)$ be the space of lower-semi continuous functions on $\mathbb{R}^n$ with the epi-distance d, which for any $f, g : \mathbb{R}^n \to \bar{\mathbb{R}}$, not identically equal to ∞, is defined as

$$d\!l(f,g) := \int_0^\infty d\!l_\rho(f,g)e^{-\rho}d\rho, \tag{4.14}$$

where ρ-epi-distance, $\rho \geq 0$ is given by

$$d\!l_\rho(f,g) := \max_{\|\bar{x}\|} |d(\bar{x}, \text{epi} f) - d(\bar{x}, \text{epi} g)|, \tag{4.15}$$

the standard distance between a point $\bar{x} = (x, \alpha) \in \mathbb{R}^{n+1}$ with $x \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$ and a set $S \subset \mathbb{R}^{n+1}$ is given by

$$d(\bar{x}, S) := \inf_{\bar{y} \in S} \|\bar{x} - \bar{y}\| \tag{4.16}$$

Notice that the space $lsc\text{-}fcns(\mathbb{R}^n)$ is not a vector space but a convex cone. The epi-distance is a metric on $lsc\text{-}fcns(\mathbb{R}^n)$ and induces the Attouch-Wets topology. With this topology, we have $f^k : \mathbb{R}^n \to \bar{\mathbb{R}}$ epi-converge to $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ if $d\!l(f^k, f) \to 0$. Even better the space $lsc\text{-}fcns(\mathbb{R}^n)$ with Attouch-Wets topology has been shown to be a Polish (complete separable metric) space which forms a foundation of the epi-splines.

Definition 4.2.1 (partition). A finite collection $R_1, R_2, \ldots R_n$ of open subsets of $\mathbb{R}^n$ is a partition of a closed set $B \subseteq \mathbb{R}^n$ if $\cup_{k=1}^N clR_k = B$ and $R_k \cap R_l = \emptyset$ for all $k \neq l$.

A polynomial in n dimensions is of total degree p if it is expressed as a finite sum of polynomial terms each having the sum of powers of the variables being no larger than p. The set of all such polynomials is denoted by $poly^p(\mathbb{R}^n)$ and the total number of terms in such a polynomial is at most

$$n_p := (n+p)!/(n!p!)$$

Definition 4.2.2 (lsc epi-splines). A (lsc) epi-spline $s : \mathbb{R}^n \to \overline{\mathbb{R}}$ of order $p \in \mathbb{N}_0$, with partition $\mathcal{R} = \{R_k\}_{k=1}^N$ of a closed set $B \subseteq \mathbb{R}^n$, is a function that on each $R_k, k = 1, \ldots, N$ is polynomial of total degree p, has $s(x) = \infty$ for $x \notin B$ and for every $x \in \mathbb{R}^n$, has $s(x) = \liminf_{x' \to x} s(x')$. The family of all such epi-splines is denoted by $e\text{-}spl_n^p(\mathcal{R})$.

Proposition 4.2.1. *For any partition $\mathcal{R}$ of a closed set $B \subseteq \mathbb{R}^n$, $p \in \mathbb{N}_0$, and $n \in \mathbb{N}_0$, $e\text{-}spl_n^p(\mathcal{R}) \subset lsc\text{-}fcns(B) \subseteq lsc\text{-}\ fcns(\mathbb{R}^n)$.*

DEFINITION 4.2.3 (infinite refinement). A sequence $\{\mathcal{R}^i\}_{i=1}^{\infty}$ of partitions of a closed set $B \subseteq \mathbb{R}^n$, with $\mathcal{R}^i = \{R_k^i\}_{k=1}^{N^i}$ is an infinite refinement if for every $x \in B$ and $\epsilon > 0$, there exists $\bar{i} \in \mathbb{N}$ such that $R_k^i \subset \mathbb{B}(x, \epsilon)$ for every $i \geq \bar{i}$ and k satisfying $x \in cl R_k^i$.

THEOREM 4.2.4 (dense approximation). *For any $p \in \mathbb{N}_0$ and $\{\mathcal{R}^i\}_{i=1}^{\infty}$, an infinite refinement of a closed set $B \subset \mathbb{R}^n$,*

$$\bigcup_{i=1}^{\infty} e\text{-}spl_n^p(\mathcal{R}^i) \text{ is dense in } lsc\text{-}fcns(B). \tag{4.17}$$

THEOREM 4.2.5 (decomposition). *For every $s \in e\text{-}spl_n^p(\mathcal{R})$ with $n > p$ and $\mathcal{R} = \{R_k\}_{k=1}^{N}$, there exists $q_k \in poly^p(\mathbb{R}^n)$ and $q_{k,j} \in poly^p(\mathbb{R}^p), j = 1, 2 \ldots, \binom{n}{p}$, such that*

$$s(x) = q_k(x) = \sum_{j=1}^{\binom{n}{p}} q_{k,j}(x_{[j]}), \text{ for all } x \in R_k \tag{4.18}$$

COROLLARY 4.2.1. *For every $s \in e\text{-}spl_n^p(\mathcal{R})$, with $n > p$ and $\mathcal{R} = \{R_k\}_{k=1}^{N}$ there exist $q_{k,i} \in poly^p(\mathbb{R}^{n-1})$, $i = 1, 2, \ldots, n$ such that*

$$s(x) = \sum_{i=1}^{n} q_{k,i}(x_{-i}), \text{ for all } x \in R_k \tag{4.19}$$

4.3. The MNSM for convex optimization

Now, we are ready to construct the MNSM for the following convex optimization problem.

$$\min_x \quad f(x) \tag{4.20}$$

$$\text{subject to} \quad Ax = b \tag{4.21}$$

$$x \geq 0 \tag{4.22}$$

where $f(x)$ is a convex function arising from the traffic equilibrium problem with general cost function. Let $\mathcal{R}^k$ be a sequence of partitions of a closed set $B \subseteq \mathbb{R}^n$ with $\mathcal{R}^k = \{R_i^k\}_{i=1}^{N^k}$ such that $\mathcal{R}^{k+1} \subseteq \mathcal{R}^k$. According to theorem 4.2.4, there exists a sequence of epi-splines $s^k(x)$ of total degree 2 such that for any infinite refinement $\mathcal{R}^k$

$$s^k(x) \xrightarrow{e} f(x) \quad as \quad k \to \infty \tag{4.23}$$

where $s^k(x) \in$ e-$spl_n^2(\mathcal{R}^k)$. Then $\limsup_k \arg\min s^k(x) \subseteq \arg\min f(x)$ in general, i.e. if $x^k \in \arg\min s^k(x)$ and $x^k \to x$, then $x \in \arg\min f(x)$. Hence, we obtain the following definition of the approximated MNSM.

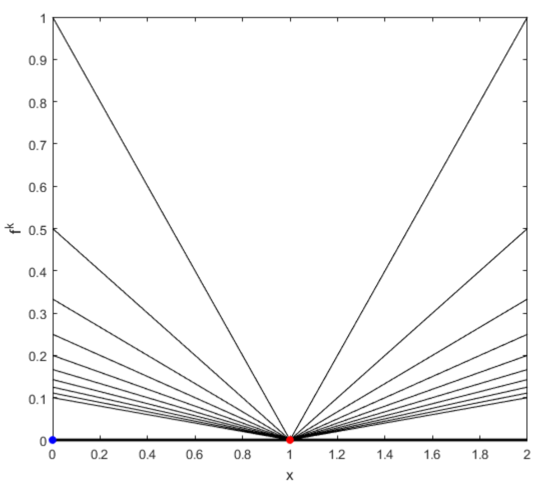
DEFINITION 4.3.1. If $\arg\min f(x)$ is not empty and $\arg\min s^k(x) \ni (x^k)^\dagger(b) \to x^*(b)$, then the approximated MNSM is defined by

$$\tilde{x}^\dagger(b) := x^*(b) \tag{4.24}$$

In general, the approximated MNSM $\tilde{x}^\dagger(b)$ of a convex problem is not necessarily equal to the MNSM of the convex problem $x^\dagger(b)$. The following example shows that even when a sequence of functions converges uniformly, the approximated MNSM does not necessarily match the MNSM we defined.

EXAMPLE 6. Let $f^k(x) = \frac{1}{k}|x-1|$, then $f^k(x) \to f \equiv 0$ uniformly on [0,2]. The MNSM for each $f^k(x)$ is $x = 1$ so that the limit of this sequence is 1. But the MNSM of $f = 0$ is $x = 0 \neq 1$. See figure 4.1.

FIGURE 4.1. the approximated MNSM does not match the MNSM



Fortunately, we have established that the approximated MNSM matches the MNSM under the following two situations.

THEOREM 4.3.2. *The approximated MNSM coincides with the MNSM if one of the following two conditions holds:*

(1) $\arg\min f(x)$ *is singleton.*

(2) *There exists a strictly convex funtion g and a full row-rank matrix L such that $f = g \circ L$ and there exists a sequence $g^k \xrightarrow{e} g$ with $f^k = g^k \circ L$*

To prove this theorem, we first recall the following proposition from [RW09]

PROPOSITION 4.3.1 (set convergence through projections). *For nonempty, closed sets C^k and C in $\mathbb{R}^n$, one has $C^k \to C$ if and only if $\limsup_k d(0, C^k) < \infty$ and the projection mappings P_{C^k} have the property that*

$$\limsup_k P_{C^k}(x) \subset P_C(x) \text{ for all } x. \tag{4.25}$$

When the sets C^k and C are also convex, one simply has that $C^k \to C$ if and only if $P_{C^k}(x) \to P_C(x)$ for all x.

Proof: Proof of this proposition is in [RW09] chapter 4 page 114.

Proof of theorem 4.3.2: (1) is trivial. (2) Since g is strictly convex, then g has a unique minimizer for every suitable $b \in \mathbb{R}^n_+$ denoted by $y(b)$. Hence for k in some index set $N \in \mathcal{N}_\infty$, $y^k(b) \to y(b)$ where $y^k(b) \in \arg\min g^k$ since $g^k \xrightarrow{e} g$. According to exercise 7.47 of [RW09], $g^k \circ L \xrightarrow{e} g \circ L$, i.e., $f^k \xrightarrow{e} f$. Meanwhile, $\arg\min f(b) = \{x | Lx = y(b)\}$ and $\{x^k | L^k x^k = y^k(b)\} \subseteq \arg\min f^k(b)$. Hence, according to theorem 2.3.4, we obtain $\arg\min f^k(b) \to \arg\min f(b)$. According to theorem 2.2.2 tells that both $\arg\min f^k(b)$ and $\arg\min f(b)$ are convex. Therefore according to proposition 4.3.1 and the definition of the MNSM,

$$x^{k\dagger}(b) \to x^\dagger(b) \text{ for every } b \in \mathbb{R}^n_+ \tag{4.26}$$

Q.E.D.

The second condition is, in general, not easy to achieve but in the traffic equilibrium problem where the full row rank matrix L could be simply chosen by using the link path incidence matrix. Also, the objective function of the equilibrium problem is $\sum_i \int_0^{x_i} t_i(s)ds$ which allows us to decompose the objective function into each links and the epi-spline could be considered as single-variable case.

EXAMPLE 7. Consider the network showing in figure 3.1 again. The link costs are given by :

$$
\begin{aligned}
x_1 &\quad : \quad t_1(s) = 82 + \frac{s}{7} + (\frac{s}{7})^4 \\
x_2 &\quad : \quad t_2(s) = 32 + \frac{s}{5} + (\frac{s}{5})^4 \\
x_3 &\quad : \quad t_3(s) = 82 + \frac{s}{4} + (\frac{s}{4})^4 \\
x_4 &\quad : \quad t_4(s) = 0 \\
x_5 &\quad : \quad t_5(s) = 0
\end{aligned}
$$

So the objective function for UE is given by

$$UE(x_1, x_2, x_3, x_4, x_5) = \sum_{i=1}^{5} \int_0^{x_i} t_i(s)ds \tag{4.27}$$

Even though, this is a function with five variables but it is a summation of five function with only one variable so that we could choose the epi-spline as single-variable case for each links. Choosing $t_3(s)$ as an example, we have

$$f(x_3) = \int_0^{x_3} t_3(s)ds = 82x_3 + \frac{x_3^2}{8} + \frac{x_3^5}{5 \cdot 4^4}, x_3 \in \mathbb{R}_+ \tag{4.28}$$

Then one piece second order epi-spline for above equation could choose as

$$f^1(x_3) = 82x_3 + \frac{x_3^2}{8}, x_3 \in \mathbb{R}_+ \tag{4.29}$$

Hence, it is easy to verify that the UE in example 4 is same as the result of using one piece second order epi-spline to approximate the UE in this example.

Furthermore, we could construct a sequence of function $f^k(x_3)$ such that $f^k(x_3) \xrightarrow{e} f(x_3)$ in the following way:

(1) $f^k(0) = f(0) = 0$

(2) Let $\mathcal{R}^k = \bigcup_{i=1}^{n} \mathcal{R}_i^k = \bigcup_{i=1}^{n-1} (a_i^k, a_{i+1}^k] \cup (a_n^k, \infty)$ be a finite partition of $(0, \infty)$ with $a_1 = 0$. For any $x \in \mathcal{R}_i^k$,

$$f^k(x) = f(a_i^k) + f'(a_i^k)(x - a_i^k) + \frac{f''(a_i^k)}{2}(x - a_i^k)^2 \tag{4.30}$$

(3) $\mathcal{R}^{k+1}$ is a refinement of $\mathcal{R}^k$ i.e.

$$\mathcal{R}^{k+1} \subseteq \mathcal{R}^k \tag{4.31}$$

It is easy to verify that $f^k(x_3) \xrightarrow{\text{e}} f(x_3)$. By construction f^k is lower semi-continuous function and non-decreasing. Also f^k converge pointwise to f. Hence, according to proposition 4.1.3,

$$f^k(x_3) \xrightarrow{\text{e}} f(x_3). \tag{4.32}$$

Remark: Here we only provide one way to construct the epi-spline which is not unique. But in the following of the dissertation, we only consider the epi-spline constructed according to above example.

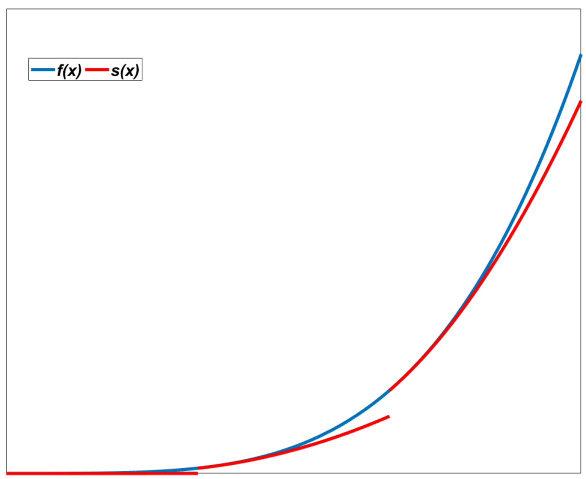


FIGURE 4.2. Second order epi-spline example

Now we turn our attention to how the corresponding partition regions change based on the approximated MNSM.

COROLLARY 4.3.1. *Let* $\mathcal{R}_U^k := \{b \mid \mathcal{A}_x^k(b) = U\}$ *where* $\mathcal{A}_x^k(b) := \{i \mid x^{k\dagger}(b)_i = 0\}$. *If the second condition in theorem 4.3.2 holds, then for every* $U \in \mathcal{P}(1:m)$

$$\mathcal{R}_U^k \to \mathcal{R}_U \tag{4.33}$$

Proof: From the proof of theorem 4.3.2, we obtain $\arg\min f^k(b) \to \arg\min f(b)$. Denote $\Omega := \{x \mid x_i = 0 \text{ for } i \in U\}$. Then $\arg\min f^k(b) \cap \Omega \to \arg\min f(b) \cap \Omega$. According to proposition 4.3.1 and the definition of the MNSM,

$$\mathcal{R}_U^k \to \mathcal{R}_U \tag{4.34}$$

Q.E.D

In the above corollary, we have shown for each U, $\mathcal{R}_U^k$ converges to $\mathcal{R}_U$. Also, in chapter 2, we discussed that both $\mathcal{R}^k := \bigcup_{U \in \mathcal{P}(1:m)} \mathcal{R}_U^k$ and $\mathcal{R} := \bigcup_{U \in \mathcal{P}(1:m)} \mathcal{R}_U$ form a partition of $\mathbb{R}_+^m$. Following the process defined by the second condition in theorem 4.3.2, the boundary of the partition generated by this process will converge to the origin one in some sense eventually. However, through our numerical simulation, we observed that it seems all the boundaries are the same which means the sequence of the partition regions is constant. More details will be discussed in next section.

4.4. Quadratic Determination

In this section, we share an important observation when we follow the method given by theorem 4.3.2 condition 2 and provide a possible explanation for this observation.

Example 8. Consider the network showing in figure 3.1 again. The general link costs are :

$$\begin{aligned}
x_1 \quad &: \quad t_1(s) = 82 + \frac{s}{7} + (\frac{s}{7})^4 \\
x_2 \quad &: \quad t_2(s) = 32 + \frac{s}{5} + (\frac{s}{5})^4 \\
x_3 \quad &: \quad t_3(s) = 82 + \frac{s}{4} + (\frac{s}{4})^4 \\
x_4 \quad &: \quad t_4(s) = 0 \\
x_5 \quad &: \quad t_5(s) = 0
\end{aligned}$$

From example 7, we know that the UE in example 4 is the same as the result of using one piece second order epi-spline to approximate the UE in this example.

As in example 5, we consider that demands q are in the box $[0, 1000] \times [0, 1000]$. Then, $0 \leq f_i \leq 1000$, $i = 1, \dots, 4$ and $0 \leq x_i \leq 1000$, $i = 1, \dots, 5$ by flow conservation. Also we apply the way shown in example 7 to construct two pieces epi-spline $\mathcal{R}^2 = \{[0, 500], (500, 1000]\}^5$.

So the UE :

$$\begin{array}{llcll} \min\limits_{f} & T_{UE}(Ff) & \rightarrow & \min\limits_{f} & T_{UE}^2(Ff) \\ s.t. & Bf = q, f \geq 0 & & s.t. & Bf = q, f \geq 0 \end{array} \tag{4.35}$$

where $T_{UE}(x) = \sum\limits_{i=1}^{5} \int_0^{x_i} t_i(s)ds$, $T_{UE}^2(x) = \sum\limits_{i=1}^{5} h_i^2(x_i)$, and for $i = 1, \dots 5$, $h_i(x_i) = \int\limits_0^{x_i} t_i(s)ds$

$$h_i^2(x_i) = \begin{cases} h_i(0) + h_i'(0)x_i + \frac{h_i''(0)}{2}x_i^2 & 0 \leq x_i \leq 500 \\ h_i(500) + h_i'(500)(x_i - 500) + \frac{h_i''(500)}{2}(x_i^2 - 500)^2 & 500 < x_i \leq 1000 \end{cases} \tag{4.36}$$

Hence, the UE with general link cost function decomposes into 2^5 subproblems and each subproblem is a quadratic optimization problem with general box constraints.

For example, when $x_1, x_3, x_4 \in [0, 500]$ and $x_2, x_5 \in (500, 1000]$, the corresponding subproblem is given by

$$\begin{array}{llcll} \min\limits_{f} & T_{UE}^2(Ff) & = & \min\limits_{f} & \frac{1}{2}f^T F^T Q F f + f^T F^T c \\ s.t. & Bf = q, l \leq Ff \leq u, f \geq 0 & & s.t. & Bf = q, l \leq Ff \leq u, f \geq 0 \end{array} \tag{4.37}$$

where $Q = \begin{bmatrix} h_1''(0) & 0 & 0 \\ 0 & h_2''(500) & 0 \\ 0 & 0 & h_3''(0) \end{bmatrix}$, $c = \begin{bmatrix} h_1'(0) \\ h_2'(500) - 500h_2''(500) \\ h_3'(0) \end{bmatrix}$, $l = \begin{bmatrix} 0 \\ 500 \\ 0 \\ 0 \\ 500 \end{bmatrix}$, and $u = \begin{bmatrix} 500 \\ 1000 \\ 500 \\ 500 \\ 1000 \end{bmatrix}$.

Remark 1: Since x_2, $x_5 > 500$, so that $\mathcal{R}_{U,2,5,[0,500,0,0,500]}$ could never happen.

Remark 2: Even though there are 2^5 subproblems, due to the topology of the network, some of the subproblems are not feasible such as the subproblem with $x_2 \in [0, 500]$ and $x_4, x_5 \in (500, 1000]$.

Now according to theorem 3.2.2, we could find the corresponding MNSM partition regions in the demand space shown in figure 4.3. Similarly, we could also find the corresponding MNSM partition regions in the demand space with epi-spline defined on

$\mathcal{R}^4 = \{\{0\}, (0, 250], (250, 500], (500, 750], [750, 1000]\}^5$ shown in figure 4.4.

Comparing with the MNSM partition region in figure 3.2, we observe that those main boundaries appear in figure 4.3 and 4.4 as well. Here a main boundary means where the corresponding inequality constraint $f_i \geq 0$ is activate.

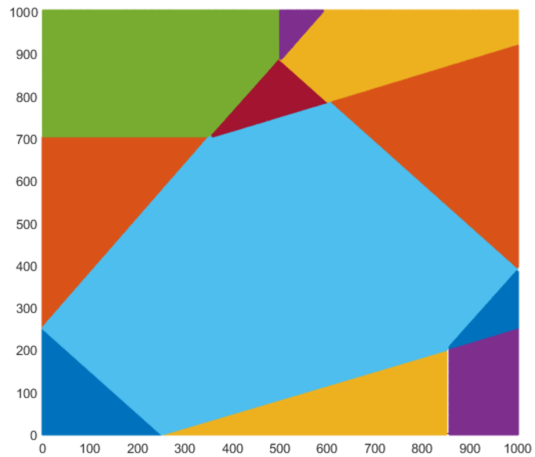


FIGURE 4.3. the MNSM partition region corresponding to epi-splines approximation with $\mathcal{R}^2$

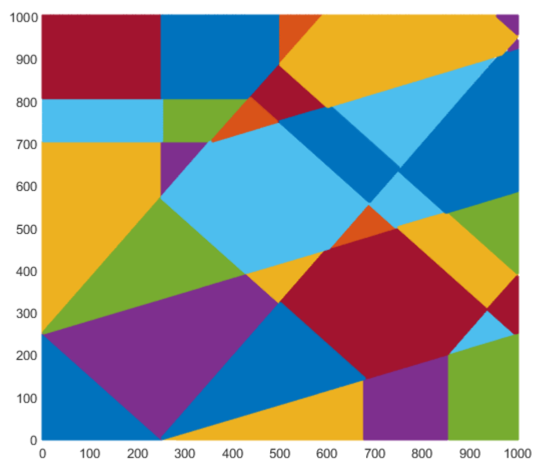


FIGURE 4.4. the MNSM partition region corresponding to epi-splines approximation with $\mathcal{R}^4$

To make it clear, let's focus on the main boundary corresponding to $f_2 \geq 0$. As in example 5, we know that the boundary of $\mathcal{R}_{\{2\}}$ is corresponding to $9b_1 - 7b_2 = 1750$ (red line in figure 4.5) which represents the boundary for $f_2 = 0$. Now choose three different regions from above, $\mathcal{R}^2_{\{2\},\{3\}}$, $\mathcal{R}^4_{\emptyset,\{1,3\}}$, and $\mathcal{R}^4_{\{2\},\{3\}}$ shown in figure 4.5. Also, we know the analytic form for those regions as following:

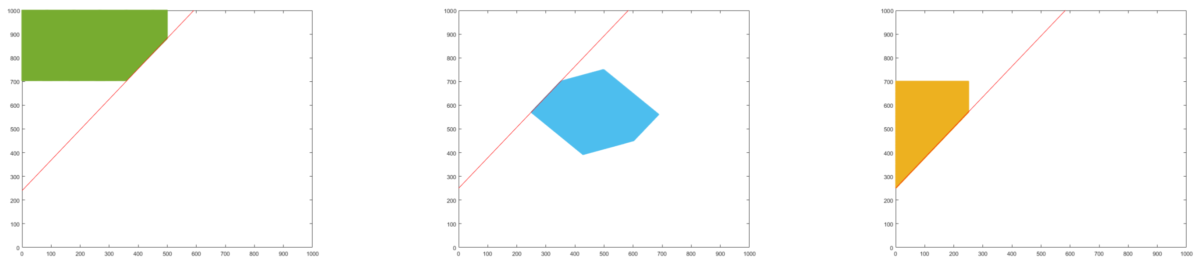
$$\mathcal{R}^2_{\{2\},\{3\}} = \{b \in \mathbb{R}^2 \mid \begin{bmatrix} 1.0000 & 0.0000 & -0.0000 \\ 0.0000 & 0.5435 & -382.6087 \\ -0.0000 & 0.4565 & -117.3913 \\ -1.0000 & -0.0000 & 500.0000 \\ -0.0000 & -0.5435 & 882.6087 \\ 0.0000 & -0.4565 & 617.3913 \\ -0.1429 & 0.1141 & -29.3478 \end{bmatrix} \begin{bmatrix} b \\ 1 \end{bmatrix} \geq 0\} \tag{4.38}$$

$$\mathcal{R}^4_{\emptyset,\{1,3\}} = \{b \in \mathbb{R}^2 \mid \begin{bmatrix} 0.4377 & 0.4377 & -359.2720 \\ 0.5623 & -0.4377 & 109.2720 \\ -0.2503 & 0.7497 & -187.4151 \\ 0.2503 & 0.2503 & -62.5849 \\ -0.4377 & -0.4377 & 609.2720 \\ -0.5623 & 0.4377 & 140.7280 \\ 0.2503 & -0.7497 & 437.4151 \\ -0.2503 & -0.2503 & 312.5849 \end{bmatrix} \begin{bmatrix} b \\ 1 \end{bmatrix} \geq 0\} \tag{4.39}$$

$$\mathcal{R}^4_{\{2\},\{3\}} = \{b \in \mathbb{R}^2 \mid \begin{bmatrix} 1.0000 & 0.0000 & -0.0000 \\ 0.0000 & 0.5435 & -382.6087 \\ -0.0000 & 0.4565 & -117.3913 \\ -1.0000 & -0.0000 & 500.0000 \\ -0.0000 & -0.5435 & 882.6087 \\ 0.0000 & -0.4565 & 617.3913 \\ -0.1429 & 0.1141 & -29.3478 \end{bmatrix} \begin{bmatrix} b \\ 1 \end{bmatrix} \geq 0\} \tag{4.40}$$

It is easy to check that the 7th row of $\mathcal{R}^2_{\{2\},\{3\}}$ and $\mathcal{R}^4_{\{2\},\{3\}}$ are parallel to $9b_1 - 7b_2 = 1750$ as well as the second row of $\mathcal{R}^4_{\emptyset,\{1,3\}}$. Hence, when b is approaching to those boundaries, we have that the corresponding value of f_2 is approaching to 0 as well. Also, the rest f_i are not matched.

Based on the observation shown in example 8, we propose the following conjecture:



(a) $\mathcal{R}^2_{\{2\},\{3\}}$ (b) $\mathcal{R}^4_{\emptyset,\{1,3\}}$ (c) $\mathcal{R}^4_{\{2\},\{3\}}$

FIGURE 4.5. Three different partition regions

The MNSM partition region of the UE with any general link cost function is determined by its quadratic approximation constructed as illustrate by example 7.

In fact, we have one step left to make above hypothesis into a theorem. The missing piece of the puzzle is that under certain condition, the MNSM defined in chapter 2 is a convex function, i.e. $x^\dagger(b)$ is vector-valued convex function in "some sense". The certain condition to guarantee $x^\dagger(b)$ convex should be positive non-decreasing lower semicontinuous objective function. The reason that we say in "some sense" is because in the space with dimension higher than 1, there are infinitely many ways to define a partial order. Actually, any non-empty pointed convex cone could induce a partial order. A natural guess is using so called Lorentz cones $L^n = \{(x,t) \in \mathbb{R}^m \times \mathbb{R} | \|x\| \leq t\}$ since the only geometric property of the MNSM that we could rely on is the minimum norm property. Unfortunately, we could not prove it at this time.

In the rest of this section, we assume that the MNSM is convex. We want to show that under this assumption, the hypothesis we claimed above would be true.

The proof has three steps. The first step is to show that for each given U the corresponding partition region is a convex set. The second step is to show that a finite collection of convex sets forms a partition of a Euclidean space only if each convex set is a polyhedral. The last step is to show that two partitions are the same.

Recall the problem we are dealing with

$$\min_x \quad f(x) \tag{4.41}$$

$$\text{subject to} \quad Ax = b \tag{4.42}$$

$$x \geq 0 \tag{4.43}$$

where $f(x)$ is non-decreasing convex function.

Let $b \in \mathcal{R}_U$ and $x^\dagger(b)$ be the image of MNSM, then the KKT conditions are given by

$$\nabla f(x^\dagger(b)) + L^T\lambda = 0 \tag{4.44}$$

$$Ax^\dagger(b) = b, x^\dagger(b) \geq 0 \tag{4.45}$$

$$\lambda \geq 0 \tag{4.46}$$

$$x^\dagger(b)_i\lambda_i = 0 \quad \forall i \tag{4.47}$$

Since $x^\dagger(b)$ is convex, the epigraph of $x^\dagger(b)$, epi$x^\dagger(b)$, is a convex set.

Since $f(x)$ is convex, then $\nabla f(x)$ is a monotone operator so that $\nabla f(\text{epi}x^\dagger(b))$ is still a convex set according to [RB16]

Since L is full column rank, according to (4.22), $\lambda = -(L^T)^\dagger \nabla f(x^\dagger(b))$. Together with (4.24) and (4.23), we obtain that $\mathcal{R}_U$ is a convex set for any U.

Theorem 4.4.1. *A finite collection of convex sets forms a partition of a Euclidean space only if each convex set is a polyhedral*

Proof: Assume Ω_1 in this collection is nonempty and not a polyhedral, and Ω_2 shares part of the boundary with Ω_1. i.e. $\partial\Omega_1 \cap \partial\Omega_2 \neq \emptyset$ and $\partial\Omega_1 \cap \partial\Omega_2$ is not singleton. Let $x, y \in \partial\Omega_1 \cap \partial\Omega_2$, then since Ω_1 is convex, the line segment between x, y lies in Ω_1. Also since Ω_2 is convex as well, the line segment between x, y lies in Ω_2 which contradicts $\Omega_1 \cap \Omega_2 = \emptyset$. Hence, the line segment between x, y must lie in $\partial\Omega_1 \cap \partial\Omega_2$. So the boundary of each set in this collection must be linear which means each convex set is a polyhedral. **Q.E.D.**

According to this theorem, we obtain that $\mathcal{R}_U$ is not only a convex set for each given U but also a polyhedral. Let $\mathcal{R}$ be the MNSM partition of problem (4.19) and $\hat{\mathcal{R}}$ be the MNSM partition of its quadratic approximation problem. For arbitrary set $U \in \mathcal{P}(1:m)$, we claim that $\hat{\mathcal{R}}_U \supset \mathcal{R}_U$.

This could prove by contradiction. Since $\mathcal{R}_U$ is not a subset of $\hat{\mathcal{R}}_U$, there exist $x_0 \in \mathcal{R}_U \setminus \hat{\mathcal{R}}_U$. Then according to theorem 18.1 [NW06], there exists a neighborhood of x_0 such that x_0 is the minimizer of the quadratic approximation problem with the same inequality be activated. This is a contradiction. Also, according to theorem 2.4.1, $\hat{\mathcal{R}}_U \cap \hat{\mathcal{R}}_V = \emptyset$ when $U \neq V$. Since both $\mathcal{R}$ and $\hat{\mathcal{R}}$ are partitions, we obtain that $\mathcal{R} = \hat{\mathcal{R}}$.

Back to the beginning of this chapter, a major barrier that may discourage this method from being implemented in wider applications is the number of partition regions, as it will increase the computational burden tremendously. If the conjecture is true, we would not need to use a very fine epi-spline to approximate the original problem; rather we can use its quadratic approximation. Since each partition region is independent from each other, we could use the result obtained from the quadratic problem as a baseline, then use the epi-spline approximation to build a parallel algorithm to compute the aMNSM.

CHAPTER 5

Measure push-forward by the MNSM

In this chapter, we will show how the uncertainty of input propagates through problem 2.1 to the uncertainty of the output under the Minimum Norm Solution Mapping. We first introduce the preliminaries of the probability space and measure. We then show that the measure push-forward from the input space to the output space forms a probability measure on the output space. At the end, we will show how the measure behaves under the process discussed in chapter 4.

5.1. Preliminaries of probability space and measure

We use Ω to denote an abstract space whose elements are called points. These points are denoted by ω.

Definition 5.1.1. Let $\mathcal{F}$ be a collection of subsets of Ω. $\mathcal{F}$ is called a field (algebra) if $\Omega \in \mathcal{F}$ and $\mathcal{F}$ is closed under complementation and finite union. That is,

(1) $\Omega \in \mathcal{F}$

(2) $A \in \mathcal{F} \implies A^c \in \mathcal{F}$

(3) $A_1, A_2, \ldots, A_n \in \mathcal{F} \implies \cup_{j=1}^{n} A_j \in \mathcal{F}$

If in addition, (3) can be replaced by countable unions, that is

(iv) $A_1, \ldots, A_n \ldots \in \mathcal{F} \implies \cup_{j=1}^{\infty} A_j \in \mathcal{F}$,

then $\mathcal{F}$ is called a σ-algebra or often also a σ-field.

Here are three simple examples of σ-algebras

(1) $\mathcal{F} = \{\emptyset, \Omega\}$

(2) $\mathcal{F} = \{\text{all subsets of } \Omega\}$

(3) $\mathcal{F} = \{\emptyset, \Omega, A, A^c\}$, if $A \subset \Omega$.

Definition 5.1.2. Given any collection $\mathcal{A}$ of subsets of Ω, let $\sigma(\mathcal{A})$ be the smallest σ-algebra containing $\mathcal{A}$. That is if $\mathcal{F}$ is another σ-algebra and $\mathcal{A} \subset \mathcal{F}$, then $\sigma(\mathcal{A}) \subset \mathcal{F}$.

Definition 5.1.3. Let $(\Omega, \mathcal{F})$ be a measurable space. By a measure on this space we mean a function $\mu : \mathcal{F} \to [0, \infty]$ with the properties:

(1) $\mu(\emptyset) = 0$

(2) if $A_j \in \mathcal{F}$ are disjoint then

$$\mu\left(\cup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mu\left(A_j\right). \tag{5.1}$$

Remark 5.1.1. *We will refer to the triple $(\Omega, \mathcal{F}, \mu)$ as a measure space. If $\mu(\Omega) = 1$, we refer to it as a probability space.*

Definition 5.1.4. A Lebesgue-Stieltjes measure on $\mathbb{R}$ is a measure on $\mathcal{B} = \sigma(\mathcal{B}_0)$ such that $\mu(I) < \infty$ for each bounded interval I. By an extended distribution function on $\mathbb{R}$ we shall mean a map $F : \mathbb{R} \to \mathbb{R}$ that is increasing, $F(a) \leq F(b)$ if $a < b$, and right continuous, $\lim_{x \to x_0^+} F(x) = F(x_0)$. If in addition the function F is nonnegative satisfying $\lim_{x \to \infty} F(x) = 1$ and $\lim_{x \to -\infty} F(x) = 0$, we shall simply call it a distribution function.

Proposition 5.1.1. *Let μ be a Lebesgue-Stieltjes measure on $\mathbb{R}$. Define $F : \mathbb{R} \to \mathbb{R}$ up to additive constants, by $F(b) - F(a) = \mu(a, b]$. For example, fix $F(0)$ arbitrary and set $F(x) - F(0) = \mu(0, x], x \geq 0, F(0) - F(x) = \mu(x, 0], x < 0$. Then F is an extended distribution.*

Theorem 5.1.5. *Suppose F is a distribution function on $\mathbb{R}$. There is a unique measure μ on $\mathcal{B}(\mathbb{R})$ such that $\mu(a, b] = F(b) - F(a)$.*

Definition 5.1.6. Suppose $\mathcal{A}$ is an algebra. μ is a measure on $\mathcal{A}$ if $\mu : \mathcal{A} \to [0, \infty]$,$\mu(\emptyset) = 0$ and if $A_1, A_2, \ldots$ are disjoint with $A = \cup_j^{\infty} A_j \in \mathcal{A}$, then $\mu(A) = \sum_{j=1}^{\infty} \mu(A_j)$. The measure is σ-finite if the space $\Omega = \cup_{j=1}^{\infty} \Omega_j$ where the $\Omega_j \in \mathcal{A}$ are disjoint and $\mu(\Omega_j) < \infty$.

Theorem 5.1.7 (Caratheodory's Extension Theorem). *.Suppose μ is σ-finite on an algebra $\mathcal{A}$. Then μ has a unique extension to $\sigma(\mathcal{A})$.*

Definition 5.1.8 (Product σ-Field). let $(\Omega_1, \mathcal{F}_1), (\Omega_2, \mathcal{F}_2)$ be two measurable space. The product σ-field $\mathcal{F}_1 \bigotimes \mathcal{F}_2$ on $\Omega_1 \times \Omega_2$ is defined as the σ-field generated by the collection of all sets of the form $\{A_1 \times A_2 \quad | \quad A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2\}$. The sets in this collection are called measurable rectangles.

REMARK 5.1.2. *$\mathcal{F}_1 \bigotimes \mathcal{F}_2 \neq \mathcal{F}_1 \times \mathcal{F}_2$ because $\mathcal{F}_1 \times \mathcal{F}_2$ may not be closed on A^c or $A_1 \cup A_2$*

THEOREM 5.1.9 (Product Measure). *Let $(\Omega_1, \mathcal{F}_1, \mu_1)$ and $(\Omega_2, \mathcal{F}_2, \mu_2)$ be two measurable spaces where μ_1 and μ_2 are σ-finite measures. There exists a unique measure μ on $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \bigotimes \mathcal{F}_2)$ that satisfies $\mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2)$ for all $A_1 \in \mathcal{F}_1$ and $A_2 \in \mathcal{F}_2$. This measure is called the product measure, written as $\mu = \mu_1 \times \mu_2$*

THEOREM 5.1.10 (Tonelli/Fubini theorem). *Let $(\Omega_1, \mathcal{F}_1, \mu_1)$ and $(\Omega_2, \mathcal{F}_2, \mu_2)$ be two measurable spaces where μ_1 and μ_2 are σ-finite measures. Let $\mu = \mu_1 \times \mu_2$ be the product measure on $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \bigotimes \mathcal{F}_2)$. Let $f : \Omega_1 \times \Omega_2 \to \mathbb{R}$ be a nonnegative measurable function. Then the following holds:*

$$\int f d\mu = \int \left[\int f(\omega_1, \omega_2) d\mu_1(\omega_1)\right] d\mu_2(\omega_2) = \int \left[\int f(\omega_1, \omega_2) d\mu_2(\omega_2)\right] d\mu_1(\omega_1) \tag{5.2}$$

Remark: the condition of f be a nonnegative measurable function can be extended to integrable functions with respect to the product measure μ,i.e. $\int |f| d\mu < \infty$.

5.2. Measure Push-forward by Minimum Norm Solution Mapping

The way we provided to understand how uncertainty propagates from the input space to output space is to understand the measure push-forward by the MNSM.

THEOREM 5.2.1. *Let $(\mathbb{R}^n_+, \mathcal{F}, \mu)$ be a probability space with the Lebesgue σ-algebra $\mathcal{F}$ and $x^\dagger : \mathbb{R}^n_+ \to \mathcal{D}^\dagger$ be the MNSM defined in previous section. Then the collection*

$$\mathcal{X} := \{B \subseteq \mathcal{D}^\dagger | (x^\dagger)^{-1}(B) \in \mathcal{F}\} = \{B \subseteq \mathcal{D}^\dagger | A(B) \in \mathcal{F}\} \tag{5.3}$$

is a σ-algebra on $\mathcal{D}^\dagger$

Proof: One have to show that the three conditions of definition 5.1.1 are satisfied. Since $x^\dagger$ is a homeomorphism,

(1)
$$(x^\dagger)^{-1}(\mathcal{D}^\dagger) = A(\mathcal{D}^\dagger) = \mathbb{R}^n_+ \in \mathcal{F} \implies \mathcal{D}^\dagger \in \mathcal{X} \tag{5.4}$$

(2) Let $B \subseteq \mathcal{X}$. Since $B \subseteq \mathcal{X}$,

$$A(B) \in \mathcal{F} \tag{5.5}$$

Since $\mathcal{F}$ is σ-algebra,

$$A(B)^c \in \mathcal{F} \tag{5.6}$$

By preimage of set difference under mapping and $A(\mathcal{D}^\dagger) = \mathbb{R}^n_+$,

$$A(\mathcal{D}^\dagger \setminus B) = A(\mathcal{D}^\dagger) \setminus A(B) = \mathbb{R}^n_+ \setminus A(B) = A(B)^c \tag{5.7}$$

Hence

$$B^c \subseteq \mathcal{X} \tag{5.8}$$

(3) Let $B_i \subseteq \mathcal{X}$ for $i \in \mathbb{N}$. Then,

$$B_i \subseteq \mathcal{X} \implies A(B_i) \in \mathcal{F} \text{ for } i \in \mathbb{N} \tag{5.9}$$

Since $\mathcal{F}$ is σ-algebra,

$$\bigcup_{i \in \mathbb{N}} A(B_i) \in \mathcal{F} \tag{5.10}$$

By Preimage of Union under Mapping

$$A\left(\bigcup_{i \in \mathbb{N}} B_i\right) = \bigcup_{i \in \mathbb{N}} A(B_i) \tag{5.11}$$

Hence

$$\bigcup_{i \in \mathbb{N}} B_i \in \mathcal{X} \tag{5.12}$$

Therefore, $\mathcal{X}$ is σ-algebra on $\mathcal{D}^\dagger$.**Q.E.D.**

DEFINITION 5.2.2. Let (X, Σ) and (X', Σ') be measurable spaces. Let μ be a measure on (X, Σ). $f : X \to X'$ is a measurable mapping. Then the pushforward measure of μ under f is the mapping $f_*\mu : \Sigma' \to [0, \infty]$ defined by

$$\forall E' \in \Sigma' : f_*\mu(E') := \mu(f^{-1}(E')) \tag{5.13}$$

THEOREM 5.2.3. *Let $(\mathbb{R}^n_+, \mathcal{F}, \mu)$ be a probability space with the Lebesgue σ-algebra $\mathcal{F}$ and probability measure μ. Let $x^\dagger : \mathbb{R}^n_+ \to \mathcal{D}^\dagger$ be the MNSM defined in previous section and $\mathcal{X}$ be pushforward σ-algebra in theorem 5.2.1. Then the pushforward measure $x^\dagger_*\mu : \mathcal{D}^\dagger \to [0, \infty]$ is a probability measure.*

Proof: To show $x_*^\dagger\mu$ is a probability measure, one need to check the conditions of definition 5.1.3.

(1)

$$x_*^\dagger\mu(\emptyset) = \mu((x^\dagger)^{-1}(\emptyset)) = \mu(A(\emptyset)) = \mu(\emptyset) = 0 \tag{5.14}$$

(2)

$$x_*^\dagger\mu(\mathcal{D}^\dagger) = \mu((x^\dagger)^{-1}(\mathcal{D}^\dagger)) = \mu(A(\mathcal{D}^\dagger)) = \mu(\mathbb{R}_+^n) = 1 \tag{5.15}$$

(3) Let $B_i, i \in \mathbb{N}$ be pairwise disjoint sets in $\mathcal{X}$.

$$\begin{aligned} x_*^\dagger\mu\left(\bigcup_{i\in\mathbb{N}} B_i\right) &= \mu\left((x_*^\dagger)^{-1}\left(\bigcup_{i\in\mathbb{N}} B_i\right)\right) \\ &= \mu\left(A\left(\bigcup_{i\in\mathbb{N}} B_i\right)\right) \\ &= \mu\left(\bigcup_{i\in\mathbb{N}} A\left(B_i\right)\right) \\ &= \sum_{i\in\mathbb{N}} \mu\left(A(B_i)\right) \\ &= \sum_{i\in\mathbb{N}} x_*^\dagger\mu\left(B_i\right) \end{aligned}$$

Also, for $i \neq j$,

$$\left(x_*^\dagger\right)^{-1}(B_i) \cap \left(x_*^\dagger\right)^{-1}(B_j) = \left(x_*^\dagger\right)^{-1}(B_i \cap B_j) = \left(x_*^\dagger\right)^{-1}(\emptyset) = \emptyset \tag{5.16}$$

Therefore, $x_*^\dagger\mu$ is a probability measure on $(\mathcal{D}, \mathcal{X})$.

Q.E.D.

5.3. The MNSM of Convex optimization preserves Measure

In this section, we want to show that the sequence of the push-forward measures induced by the MNSM in chapter 4 converges strongly.

THEOREM 5.3.1. *Let μ be the probability measure on the demand space $(\mathbb{R}_+^n, \mathcal{F})$, ν_k be the induced probability measure on the path flow space by the MNSM $x^{k\dagger}(b)$ of the second order approximation of the convex problem, and ν be the induced probability measure on the path flow space by the MNSM $x^\dagger(b)$ of the convex problem. Then for every $\Omega \in \mathcal{F}$, $\nu_k(x^{k\dagger}(\Omega)) \to \nu(x^\dagger(\Omega))$*

Proof: According to theorem 4.3.2, $x^{k\dagger}(b) \to x^{\dagger}(b)$ for every $b \in \mathbb{R}^m_+$. Hence, for every $\Omega \in \mathcal{F}$, $x^{k\dagger}(\Omega) \to x^{\dagger}(\Omega)$. Since both $x^{k\dagger}$ $x^{\dagger}$ have continuous inverse mapping A, $\nu^k = \mu \circ (x^{k\dagger})^{-1} = \mu \circ A$ and $\nu = \mu \circ (x^{\dagger})^{-1} = \nu \circ A$. Therefore, for every $\Omega \in \mathcal{F}$,

$$\nu_k(x^{k\dagger}(\Omega)) \to \nu(x^{\dagger}(\Omega)) \tag{5.17}$$

Q.E.D.

From the proof of theorem 5.3.1, we could see that the sequence of the MNSMs generates a sequence of subsets in $\mathbb{R}^m_+$ which eventually deforms into the subset generated by the MNSM of the origin problem. Meanwhile, the corresponding measures are, in fact, unchanged in some sense. Geometricaly speaking, those image of $x^{k\dagger}$ and the corresponding measure are the same essentially but with varying degrees of deformation which could not be captured by the natural topology in $\mathbb{R}^m_+$. Moreover, as the sequence of measures converges strongly, those measures converge weakly as well. So according to portmanteau theorem [Hos80] for all bounded continuous functions f

$$\lim E_k[f] \to E[f] \tag{5.18}$$

, and for every lower semi-continuous function bounded from below,

$$\liminf E_k[f] \geq E[f] \tag{5.19}$$

where E_k is the expectation with respect to ν^k and E is the expectation with respect to ν.

In conclusion, under the MNSM framework, it needs to pullback the information from the output space to understand the uncertainty. And what matters is those pullback image of the input space. This provides a systematic way to observe and study the uncertainty of the output space. Examples will be provided in Chapter 7 to demonstrate how the theories established in this chapter can be used to understand uncertainty propagation from the demand space to the network flow space in a traffic equilibrium network.

CHAPTER 6

Numerical Methods

In this chapter, we are going to explore suitable numerical methods to obtain the MNSM of the traffic equilibrium problem. According to the discussion in chapter 3 and chapter 4, the MNSM of the traffic equilibrium problem with linear cost function is a piecewise affine function with respect to the input and the MNSM of the traffic equilibrium problem with general cost function could be approximated by second order epi-splines which is also a piecewise affine function with more partition regions. In each partition region, the MNSM is given by theorem 3.1.6 and theorem 3.1.7. Hence, there are three steps to obtain the MNSM of the traffic equilibrium problem. First, one needs to calculate the pseudoinverse of the corresponding matrix in the theorem 3.1.6 or theorem 3.1.7. Second, one needs to identify the region of the corresponding partition based on theorem 3.2.1. Lastly, one needs to find all the regions or obtain the regions which people are interesting.

6.1. Find the MNSM in a given partition region

When one identifies a demand belonging to a known region, the MNSM is given by theorem 3.1.6 or theorem 3.1.7. For each case, we need to calculate the pseudoinverse of a matrix [BIG03] to obtain the MNSM. There are two types of algorithms to calculate the pseudoinverse. One category is based on certain decomposition of the given matrix such as Rank decomposition, QR decomposition and Singular Value Decomposition. Another way is based on rank-one-update such as Greville's method and Kishi's method. In general, the best way to obtain the pseudoinverse is using the methods based on decomposition with the complexity $O(min(n^2m, m^2n))$. However, in our case, computing the pseudoinverse is not an isolated problem. The MNSM of adjacent partition regions has internal relation – the corresponding matrices could be obtained by appending or deleting rows (or columns) from one to another. In this chapter, we introduce the so-called direct method based on decomposition technique. In chapter 8, we will discuss a recursive method specifically suitable for our case and discuss its importance as well as challenges.

The so-called direct method is to obtain the pseudoinverse only based on the given matrix itself without any other addition information. A common algorithm is based on the singular value decomposition, which has very stable implementation in Matlab and Julia as pinv function. Here we recall the theorem for singular value decomposition.

THEOREM 6.1.1 (Singular Value Decomposition). *[BIG03]. Let $A \in \mathbb{R}^{n\times m}$ be a matrix of rank r with singular values $\sigma_1, \dots, \sigma_r$. Then there exists an orthogonal matrix $U \in \mathbb{R}^{n\times n}$ and an orthogonal matrix $V \in \mathbb{R}^{m\times m}$ such that*

$$A = U\Sigma V^T, \tag{6.1}$$

where $\Sigma = diag\{\sigma_1, \dots, \sigma_r\, 0, \dots, 0\} \in \mathbb{R}^{n\times m}$.

Define $\Sigma^\dagger = diag\{1/\sigma_1, \dots, 1/\sigma_r, 0, \dots, 0\}$, then the pseudoinverse of A is given by

$$A^\dagger = V\Sigma^\dagger U^T \tag{6.2}$$

As easily noticed, there are some rows of U and V corresponding to the 0 singular values. Those rows will not affect the results of multiplication. Hence, we could have a compact singular value decomposition without considering those rows.

COROLLARY 6.1.1 (Compact SVD). *Let $A \in \mathbb{R}^{n\times m}$ be a matrix of rank r, $r \geq 1$, and let $u_1, \dots, u_n$ and $v_1, \dots, v_m$ denote the columns of the orthogonal matrices $U \in \mathbb{R}^{n\times n}$ and $V \in \mathbb{R}^{m\times m}$ that appear in the theorem 6.1.1 of A. Let $U_c = [u_1, \cdots, u_r]$, $V_c = [v_1, \cdots, v_r]$,$\Sigma_c = diag\{\sigma_1, \dots, \sigma_r\} \in \mathbb{R}^{r\times r}$. Then*

$$A = U_c\Sigma_c V_c^T = \sum_{i=1}^{r} \sigma_i u_i v_i^T \text{ and } A^T = V_c\Sigma_c U_c^T = \sum_{i=1}^{r} \sigma_i v_i u_i^T \tag{6.3}$$

Applying the corollary 6.1.1 to theorem 3.1.6, we obtain the MNSM of a given region of partition.

THEOREM 6.1.2. *Let $b \in \mathcal{R}_U$ and $L = \begin{bmatrix} A \\ E_U \end{bmatrix}$. $K_U = \begin{bmatrix} P_{L^T}^{\perp} Q \\ L \end{bmatrix}$ has the compact singular value decomposition*

$$K_U = U_c\Sigma_c V_c^T \implies K_U^\dagger = V_c\Sigma_c^\dagger U_c^T \tag{6.4}$$

Then the MNSM

$$x(b) = M_U b + N_U c = V_c \Sigma_c^+ U_c^T(:, (n+1):(n+m))b - V_c \Sigma_c^+ U_c^T(:, 1:n) P_{A^T}^{\perp} c \tag{6.5}$$

Proof: The proof is applying theorem 6.1.2 to theorem 3.1.6.**Q.E.D.**

The theorem 6.1.2 not only gives a formula of the MNSM for given region $\mathcal{R}_U$, but also potentially provides a suboptimal assignment. For example, one may approach an idea of the suboptimal assignment by ignoring some of the small singular values and the corresponding vectors: $\tilde{x}(b) = \tilde{V}_c \tilde{\Sigma}_c \tilde{U}_c^T(:, (n+1):(n+m))b - \tilde{V}_c \tilde{\Sigma}_c^+ \tilde{U}_c^T(:, 1:n) P_{A^T}^{\perp} c$. This idea will be discussed in more details in chapter 7.

6.2. Preliminaries for Finding Partitions

In the previous section, we assume that one knows which given partition region the demand belongs to. In this section, we present numerical methods for exploring the partition regions. Naively speaking, the number of partition regions could be 2^n, where n is the total number of paths in the network, since each path could be either occupied or idle. In fact, due to the connectivity of the partition region and the low-rank of the MNSM in a given region, the number of partition regions should be much smaller than the worst case. Furthermore, in the real world, the demands tend to have a concentrated distribution, i.e. some part of the demand space would occupy the majority of the probability of the demands. Hence, the number of partition regions which need to be considered can be further reduced. In this dissertation, we have developed two types of algorithms based on different problem requirements. One is to obtain a comprehensive set of all partition regions, the other one is to obtain the partition regions focusing on neighborhoods of interested demand values. Approaching with either method, there involve two important sub-algorithms: redundancy removal and feasibility check as preliminaries.

6.2.1. H-Redundancy Removal. A redundant constraint is a constraint that can be removed from a system of constraints without changing the feasible region. As we discussed above, the MNSM of CQP is a piecewise affine function defined on the input space. In this specific problem, the input domain corresponds to the demand space $\mathbb{R}_+^m$. In this section, we are interested in how to remove the redundant constraints of the pre-image of $\{f | f_i = 0, \text{ if } i \in U \text{ and } f_i > 0 \text{ if } i \notin U\}$, where $U \in \mathcal{P}(1:m)$ follows the same definition as before. Applying the theorem 3.2.1, the closure of the pre-image set is given by

$$
\begin{aligned}
-\left[(L^T)^\dagger \left(QM_U b + QN_U + c\right)\right]_j &\geq 0 \quad \text{for } j = 1+n, \ldots, |U|+n. \\
[M_U b + N_U]_i &\geq 0 \qquad\qquad\quad \forall i \notin U
\end{aligned}
\tag{6.6}
$$

Hence, we want to remove the redundant constraints from above set and identify the non-redundant constraints. In our case, only the non-redundant constraints are corresponding to the paths having potential to be eliminated as one goes from the current region to its neighboring region. The main difference between our case and general linear programming problem (or polyhedral computation problem [Fuk04]) is those redundant constraints which parallel to some non redundant constraints are still meaningful. The parallel constraints mean that the corresponding paths should be eliminated or added to the system simultaneously. Topologically speaking, those constraints are identical under the framework of the MNSM. Identifying redundant constraints and parallel constraints helps to reduce the number of searching options which we will discuss in next section.

It is easy to see that (6.6) is a convex polytope since it is the intersection of half-spaces. Now define the polytope in (6.6) as $Ab \leq c$ for convenience. Define $H = [A \quad c] = [h_{ij}]$. Hence it is easy to find the parallel constraints by considering $\tilde{H}$, where $\tilde{H}$ is given by dividing the first non-zero entry of each row of H. If $\|\tilde{H}(i,:) - \tilde{H}(j,:)\| \leq tol$, we say that ith row and jth row are parallel with tolerance tol where $\tilde{H}(i,:)$ means ith row of $\tilde{H}$. Meanwhile, if two rows are parallel, we should group them together. Next step is to remove the redundant constraints and identify the non-redundancy constraints of (6.6). It could be done by considering the following linear programming problem:

$$
P := \max_b \quad 0 \quad \text{s.t. } Ab \leq c \quad b \in \mathbb{R}_+^m \tag{6.7}
$$

Let P_i denote the sub-problem of problem P which has the objective function $z_i = \max_b \sum_j^m A_{ij} b_j$ and constraints $\sum_j^m A_{kj} b_j \leq c_k$ for $k = 1, 2, \ldots, i-1, i+1, \ldots, n$. Then, we compare z_i and c_i. If $z_i \leq c_i$ then, we say ith constraint is redundant otherwise it is non-redundant.

input : A, c
output: the index of non-redundant constraints: I^{nr}
$I^{nr} = \emptyset$;
for i ***in*** U **do**
    $z_i \leftarrow$ Find the optimal objective value of Problem P_i;
    if $z_i \leq c_i$ **then**
        The i-th constraint is redundant;
        continue
    else
        The i-th constraint is non-redundant;
        $I^{nr} \leftarrow I^{nr} \cup \{i\}$;
    end
end

Algorithm 1: H-Redundancy Removal

6.2.2. Feasible Partition Region. A set $\{x|Ax \leq b\}$ is feasible if it is not empty. In this section, we summarize well known duality theorems that can be used to check the feasibility based on Linear Programming theory.

LEMMA 6.2.1 (Farkas's lemma). *The set $\{x|Ax \leq b\}$ is non-empty if and only if the $\{y|y^T A = 0, y^T b < 0, y \geq 0\}$ is empty.*

Now we could form a LP as following :

$$(\mathcal{P}): \quad \max_x \quad 0 \tag{6.8}$$

$$\text{s.t.} \quad Ax \leq b \tag{6.9}$$

Then the corresponding Dual problem is given by:

$$(\mathcal{D}): \quad \min_y \quad b^T y \tag{6.10}$$

$$\text{s.t.} \quad A^T y = 0 \tag{6.11}$$

$$y \geq 0 \tag{6.12}$$

THEOREM 6.2.2 (Weak duality theorem). *Let $\mathcal{P} = max(c^T x|Ax \leq b)$ and $\mathcal{D} = min(b^T y|A^T y = c, y \geq 0)$ be its dual. If x is a feasible solution for $\mathcal{P}$ and y is a feasible solution for $\mathcal{D}$, then $c^T x \leq b^T y$.*

Theorem 6.2.3 (Duality Theorem for LPs). *If $\mathcal{P}$ and $\mathcal{D}$ are a primal-dual pair of LPs, then one of these four cases occurs:*

(1) *Both are infeasible;*

(2) *$\mathcal{P}$ is unbounded and $\mathcal{D}$ is infeasible;*

(3) *$\mathcal{D}$ is unbounded and $\mathcal{P}$ is infeasible;*

(4) *Both are feasible and there exist optimal solutions x, y to $\mathcal{P}$ and $\mathcal{D}$ such that $c^T x = b^T y$.*

Hence, to check the feasibility of $\mathcal{P}$, we could check whether $\mathcal{D}$ is unbounded or not. This can be done easily by applying the well-known Simplex method [NM65].

6.3. Search for Partition Regions

One of the most difficult part of multi-parametric quadratic optimization problem is to identify the partition regions of the input space. In [BMDP02],[Bao02],[TJB03a],[TJB03b],[SKJ$^+$06],[SGDD03], and[OD04], researchers created searching algorithms based on geometric properties of the partition regions especially based on reversing recursively the facet-defining hyperplanes of all previously identified regions with or without the assumption that for each facet of a parition region, there exists only one neighboring partition region that is adjacent to this facet. In [GBN11] and [FJO13], the authors presented a combinatorial multi-parametric approach that is based on an implicit enumeration of all possible constraint combinations in the form of candidate active sets. In our cases, thanks to theorem 3.2.1, we could determine the partition region when the index of activated inequality constraints are known. Meanwhile, we could find a way such that we could avoid enumeration of all possible constraint combinations by locating an initial region.

In this section, we provide a method to find the partition regions based on the Depth First Search (DFS) and wall-crossing technique [KZ16]. Depth-first search (DFS) is an algorithm for traversing or searching tree or graph data structures. The algorithm starts at the root node (selecting some arbitrary node as the root node in the case of a graph) and explores as far as possible along each branch before backtracking, as shown in algorithm 2. Meanwhile, wall-crossing is to flip the corresponding inequality constraints of the boundary for the certain partition boundary in our consideration. It provides a way to generate next node on the searching graph since one would not have the whole graph at the beginning.

```
input  : A graph G and a vertex v of G
output: All vertices reachable from v labeled as discovered
procedure DFS(G, v);
label v as discovered;
for all directed edges from v to w that are in G.adjacentEdges(v) do
    if vertex w is not labeled as discovered then
        recursively call DFS(G, w)
    end
end
```

Algorithm 2: Pseudocode of DFS

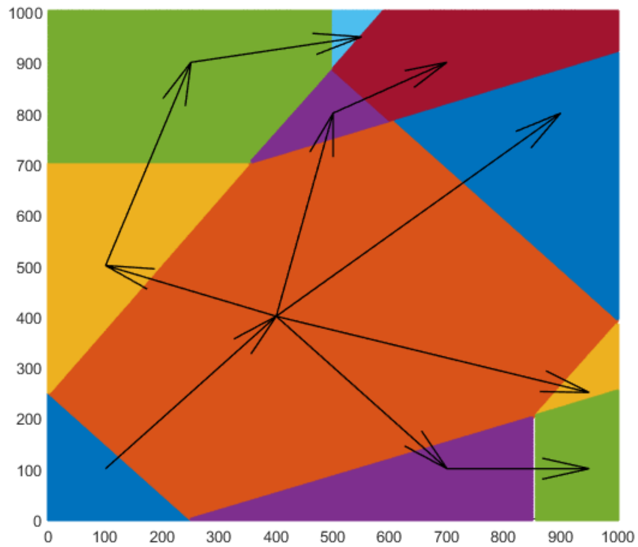
Assume that the initial region is known, which means the index of activated inequality constraints is known, then according to theorem 3.1.6 (or theorem 3.1.7) and theorem 3.2.1 we could obtain the analytic representation of the initial region, which could be treated as our root node. Then according to the method discussed in section 6.2, we could identify those non-redundant inequality constraints. If the non-redundant inequality constraints are in the top half of (6.6), then eliminate the corresponding indexes together with those parallel ones from the index set of the initial region which forms a new node. If the non-redundant inequality constraints are in the bottom half, then add the corresponding indexes together with those parallel ones to the index set of the initial region which forms a new node. Each of the non-redundant inequality constraints could be treated as the edge connecting to the next node.There are two situations which could help to prune the searching tree. One is pruning by infeasibility. Another one is that there only exists visited node. Notice that some boundaries of partition regions are based on domain of the input space, i.e. we could have the partition regions intersecting with the domain of the input which is a trivial non-redundant constraint and could not do wall-crossing. Also, notice that due to numerical error of the analytic formula of the MNSM, we need to add the reversed inequality constraints to the new region since otherwise it is not guaranteed that the same boundary computed by the top half of (6.6) and the bottom half of (6.6) are perfectly attached. Now we want to show one example to illustrate the process of the DFS.

Example 9. Consider the problem in example 7 with the general cost function case. As in example 7, we still consider that the demands q are in the box $[0, 1000] \times [0, 1000]$ with the second order epi-spline given by $R^2 = \{[0, 500], [500, 1000]\}^2$. Assume that we start at the region $\mathcal{R}_{1,4} = \{(q_1, q_2) \mid q_1 \geq 0, q_2 \geq 0, q_1 \leq 500, q_2 \leq 500, q_1 + q_2 \leq 250, q_1 + q_2 \leq 250\}$ given by theorem

3.2.2, then it is easy to see that $q_1 \geq 0, q_2 \geq 0, q_1 + q_2 \leq 250$ are non-redundant constraints, $q_1 \leq 500, q_2 \leq 500$ are redundant constraints. Among the non-redundant constraints, $q_1 \geq 0, q_2 \geq 0$ are corresponding to the domain of the demand space which could not apply wall-crossing technique. $q_1 + q_2 \leq 250$ has one parallel inequality constraint so that we need to group the corresponding indices which are $\{1, 4\}$ according to theorem 3.2.2. So $q_1 + q_2 \leq 250$ is the only non-redundant constraint which could apply wall-crossing technique. Also, since $q_1 + q_2 \leq 250$ is in the top half of (6.6), we need to eliminate the corresponding indices together with those parallel ones to the index set $\{1, 4\}$. Hence, the only candidate region for next step is

$$\begin{aligned}
\mathcal{R}_\emptyset \quad = \quad & \{(q_1, q_2) \mid 7/16q_1 + 7/16q_2 \geq 875/8, && (6.13)\\
& 9/16q_1 - 7/16q_2 + 875/8 \geq 0, && (6.14)\\
& -1/4q_1 + 3/4q_2 + 125/2 \geq 0, 1/4q_1 + 1/4q_2 - 125/2 \geq 0, && (6.15)\\
& -7/16q_1 - 7/16q_2 + 4875/8 \geq 0, -9/16q_1 + 7/16q_2 + 3125/8 \geq 0, && (6.16)\\
& 1/4q_1 - 3/4q_2 + 875/2 \geq 0, -1/4q_1 - 1/4q_2 + 1125/2 \geq 0\} && (6.17)
\end{aligned}$$

given by theorem 3.2.2. Since we apply wall-crossing to the inequality $q_1 + q_2 \leq 250$, the region that we should deal with now is defined by $\mathcal{R}_\emptyset \cap \{(q_1, q_2) \mid q_1 + q_2 \geq 250\}$. Through redundancy check, we obtain that $1/4q_1 + 1/4q_2 \leq 1125/2$ is the redundant constraint and the rest are non-redundant. Among those non-redundant constraints, $7/16q_1 + 7/16q_2 \geq 875/8$ and $1/4q_1 + 1/4q_2 \geq 125/2$ are parallel to $q_1 + q_2 \geq 250$ which could not form a new candidate. Hence there are five new candidate regions for the next steps. Repeating this process until no new candidate region generated, we obtain the partition regions and searching tree shown in figure 6.1.



(a) Partition Regions

$$
\begin{tikzcd}
 & & \mathcal{R}_{\{1,4\}} \arrow[d] & & \\
 & & \mathcal{R}_{\emptyset} \arrow[dll] \arrow[dl] \arrow[d] \arrow[dr] \arrow[drr] & & \\
\mathcal{R}_{\{2\}} \arrow[d] & \mathcal{R}_{\{3\}} \arrow[d] & \mathcal{R}_{\emptyset,\{1\}} & \mathcal{R}_{\emptyset,\{2\}} & \mathcal{R}_{\emptyset,\{3\}} \arrow[d] \\
\mathcal{R}_{\{2\},\{3\}} \arrow[d] & \mathcal{R}_{\{3\},\{2\}} & & & \mathcal{R}_{\emptyset,\{1,3\}} \\
\mathcal{R}_{\{2\},\{1,3\}} & & & &
\end{tikzcd}
$$

(b) Searching Tree

FIGURE 6.1. Partition Regions and Searching Tree

6.4. Identify An Initial Region

In this section, we provide two ways to obtain the initial region for the searching process of the partition region.

6.4.1. Minimal Demand. When the demand for each OD pairs is small enough, the cost function could be approximated by its free flow time on each link.

$$t(x) \approx t(0) \tag{6.18}$$

Hence, the equilibrium problem reduced to the shortest path problem so that we could use algorithm such as Dijkstras algorithm and Bellman-Ford Algorithm to get the solutions. Notice that since there might be multiple paths having minimum flow cost at the same time, in this case the MNSM would be the average of those paths. Then the index of those zero components of the solution forms the index set U. Then $\mathcal{R}_U$ would be the first known partition region in the demand space.

6.4.2. Particular Demand. The particular demand case means that for a given demand, we could allocate this demand into a region such that the original problem is equivalent to a subproblem with only equality constraints. As we have seen, the MNSM is a solution mapping for given demand which means the corresponding feasible direction should be zero according to the first order optimality condition. Since x is the solution to the problem

$$\min_x \quad \frac{1}{2}x^TQx + c^Tx \tag{6.19}$$

$$\text{subject to} \quad Ax = b \tag{6.20}$$

$$x \geq 0 \tag{6.21}$$

where $b \in \mathcal{R}_U$ for given U, the corresponding feasible direction should solve the following problem:

$$\begin{aligned} \min_q \quad & \frac{1}{2}p^TQp + g^Tp \\ \text{subject to} \quad & Ap = 0 \\ & p_i = 0, i \notin U \end{aligned}$$

where $g = Qx + c$. If the current region is not the correct one, then there are two situations. First, there exist some blocking constraints, i.e. we could not move from the current point along the feasible direction as many as we want.Hence, we should add the corresponding path into the system. Second, there are some inequalities that have negative dual variable, i.e. we should eliminate the corresponding path out of the system. Here are two theorems which guarantee the above two strategy are valid.

THEOREM 6.4.1. *Suppose that the point $\hat{x}$ satisfies first-order conditions for the equality constrained subproblem with $\hat{U}$. Suppose that the constraint gradients $a_i, i \notin \hat{U}$ are linearly independent and that there is an index $j \notin \hat{U}$ such that $\hat{g}_j < 0$. Let p be the solution obtained by dropping the constraint j and solving the following subproblem :*

$$\begin{aligned} \min_{q} \quad & \frac{1}{2}p^T Qp + (G\hat{x} + c)^T p \\ \textit{subject to} \quad & Ap = 0 \\ & p_i = 0, i \notin U \end{aligned}$$

Then p is a feasible direction for constraint i, that is $p_i \geq 0$. Moreover if p satisfies second-order sufficient conditions, then we have that $p_i > 0$, and p is a decent direction.

THEOREM 6.4.2. *Suppose that the solution p is nonzero and satisfies the second-order sufficient conditions for optimality for that problem. Then the function is strictly decreasing along the direction p.*

Both proofs could be found in chapter 16 in the book "Numerical Optimization" written by Jorge Nocedal and Stephen J. Wright . In fact, if people are not interested in finding the whole partition region but some particular regions, they could use algorithm 3 together with theorem 3.2.1 to obtain those interested regions.

```
input  : q
output: U, M_U, N_U
U = 1 : n ;
M_U, N_U ← Find the MNSM corresponding to U;
x ← M_U q + N_U;
g ← Qx + c;
while not all the g_i is positive do
    Solve p;
    if p ≠ 0 then
        α ← min(1, min_{i∉U, p_i<0} (−x_i/p_i));
        if α = 1 then
            U ← U ;
        else
            i ← arg min_{i∉U, p_i<0} (−x_i/p_i);
            U ← U ∪ {i};
        end
    else
        j ← arg min_{j∈U} g_j;
        U ← U\{j};
    end
    M_U, N_U ← Find the MNSM corresponding to U;
    x ← M_U q + N_U;
    g = Qx + c;
end
```

Algorithm 3: Demand allocation

6.5. Numerical Example and Discussion

As we discussed in previous sections, there are two concerns of the MNSM approaching. One is the uncontrollable growth of the number of partition regions, and another one is the computing time for predetermine of the partition regions. In this section, we provide two examples to demonstrate promising results regarding those two issues.

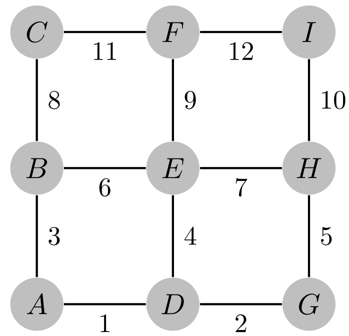


FIGURE 6.2. 3×3 Grid Network

The first example is a 3 by 3 grid network as shown in figure 6.2, with link cost functions given by :

$$t(s) = \begin{bmatrix} 0.0313s + 39.0000 \\ 0.0455s + 42.0000 \\ 0.0111s + 72.0000 \\ 0.0135s + 58.0000 \\ 0.0175s + 100.0000 \\ 0.0357s + 44.0000 \\ 0.0145s + 51.0000 \\ 0.0167s + 71.0000 \\ 0.0109s + 10.0000 \\ 0.0417s + 81.0000 \\ 0.0238s + 57.0000 \\ 0.0244s + 44.0000 \end{bmatrix} \tag{6.22}$$

In this example, we focus on two OD pairs with two demands in $[0, 10000] \times [0, 10000]$. One is from node A to node I and another one is from node B to node H. In our consideration, there are 16 paths corresponding to first OD pair and there are 9 paths corresponding to the second OD pair (details shown in appendix). Then, we have total 25 different paths and two OD pairs. Hence, the worst case analysis tells there would be 2^{23} different partition regions since each path could be used or unused in each partition region and by the flow conservation, each OD pair must have at least one path connected. However, there are in fact only 23 different regions as shown in figure

6.3 (details shown in appendix). Also, increasing the range of two demands, the number of the partition regions would not increase any more.

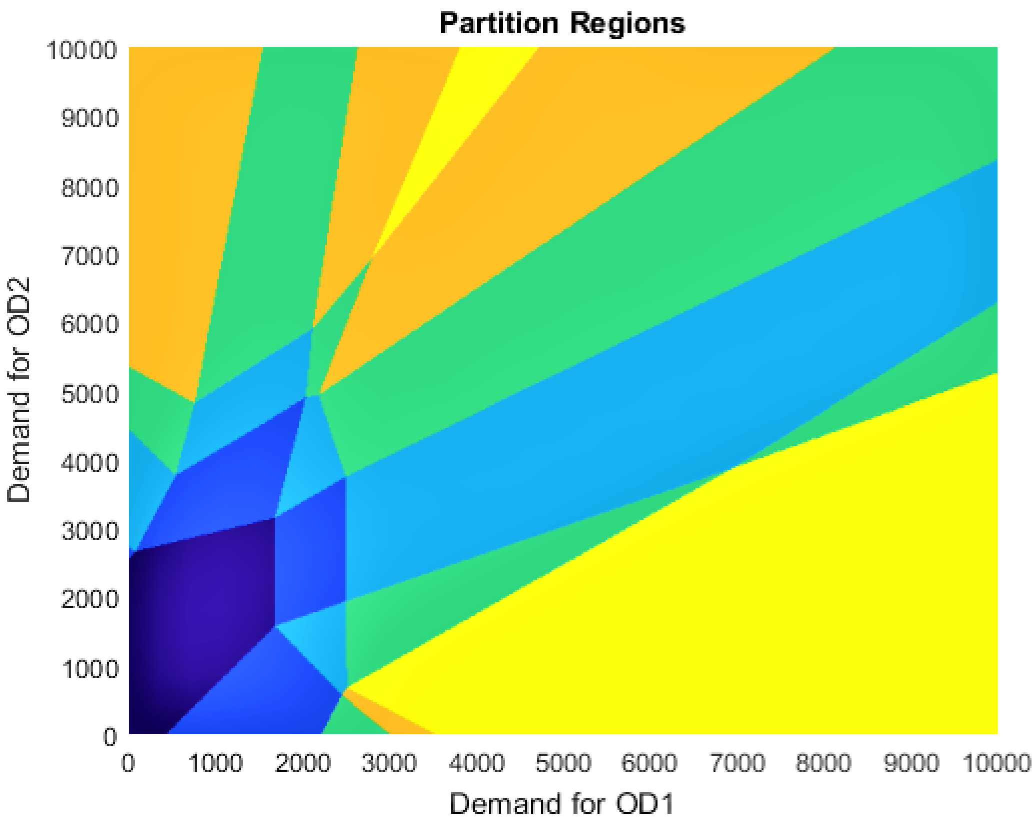


FIGURE 6.3. Partition Regions of the 3×3 grid network

Here we observe that the total number of the partition regions is much less than the worst case analysis. A possible explanation for this phenomenon is due to the low rankness of the MNSM according to theorem 3.1.6 and 3.2.1. As we have shown in 3×3 grid network, the MNSM for each given partition region $\mathcal{R}_U$ is

$$f = M_U q + N_U \tag{6.23}$$

where $M_U \in \mathbb{R}^{25\times 2}$. According to theorem 3.2.1, the partition region is determined by about 50 inequality constraints. Due to the low rankness, if we flip 4 inequality constraints, there would be a high chance to end up with an empty region.

The second example is based on Berlin Friedrichshain network, for which detail information is available at https://github.com/bstabler/TransportationNetworks/blob/master/Berlin-Friedrichshain/README.md

This network has 224 nodes and 523 links, as shown in figure 6.4. We use this example to demonstrate computational applicability of our method towards medium-large networks.

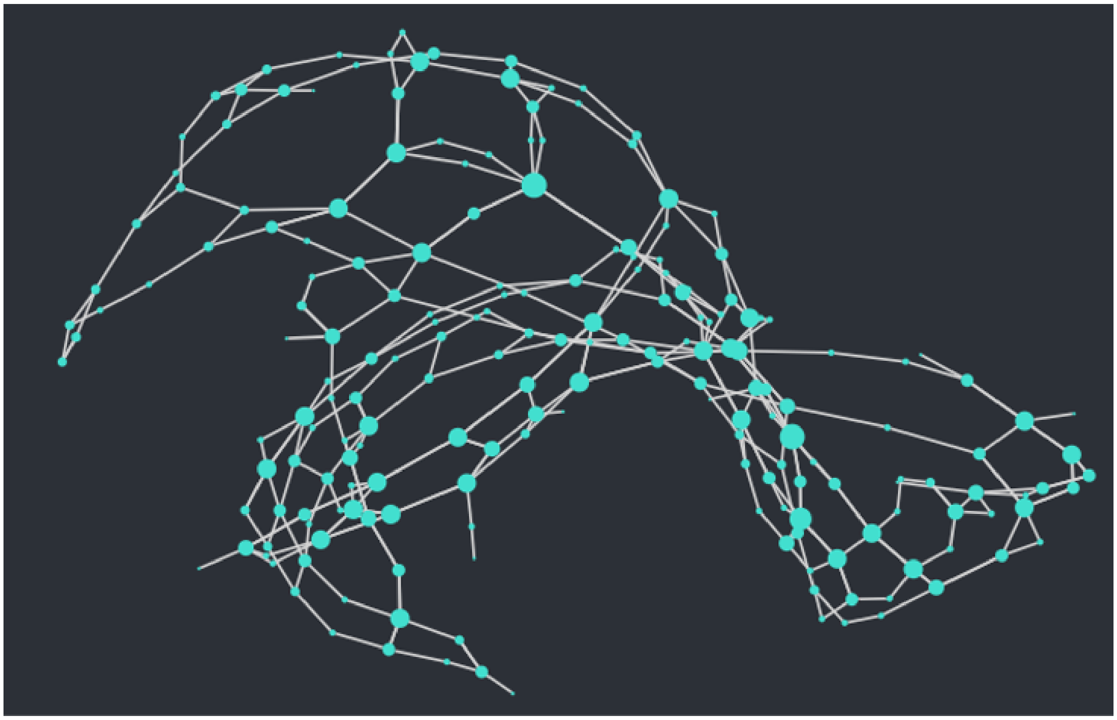


FIGURE 6.4. Berlin Friedrichshain Network Schematic diagram

The initial region is chosen by the minimal demand strategy. In each numerical experiment, we consider a certain number of OD pairs generated randomly. For each chosen OD pair, we take five paths into consideration which are calculated by Yen's k-shortest path algorithm. The programs are implemented in Julia version 1.0.2 with Intel(R) Core(TM) i5-8400 CPU@ 2.80GHz.

Figure 6.5 reports computing time associated with problems of different sizes by varying the number of OD pairs from 4 to 15. When the number of OD pairs is between 4 and 12, we repeat 1000 times and calculate the average running time for searching partition regions. When the number of OD pairs is between 12 and 15, we repeat 500 times due to the increasing running time and calculate the average running time for searching partition regions.

In figure 6.6, the blue curve represents the average running time, the orange one represents the exponential growth and the yellow one represents the fourth-order polynomial growth. As one may

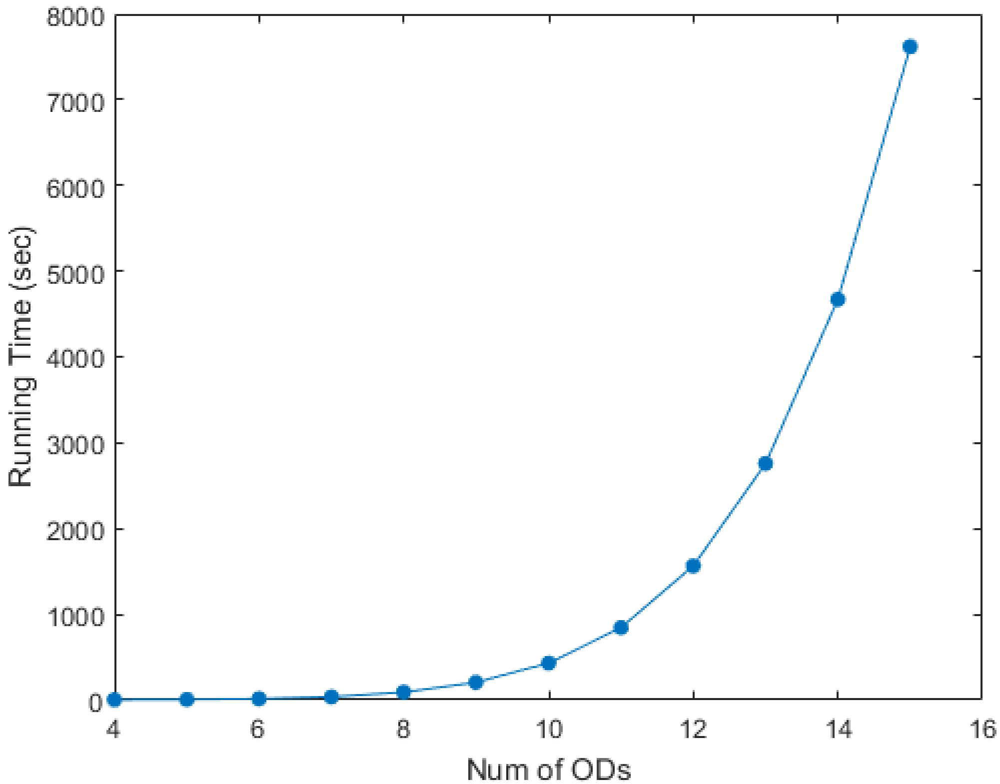


FIGURE 6.5. Average running time

notice, the average running time growth of our algorithm is faster than fourth-order polynomial, but significantly slower than exponential growth.

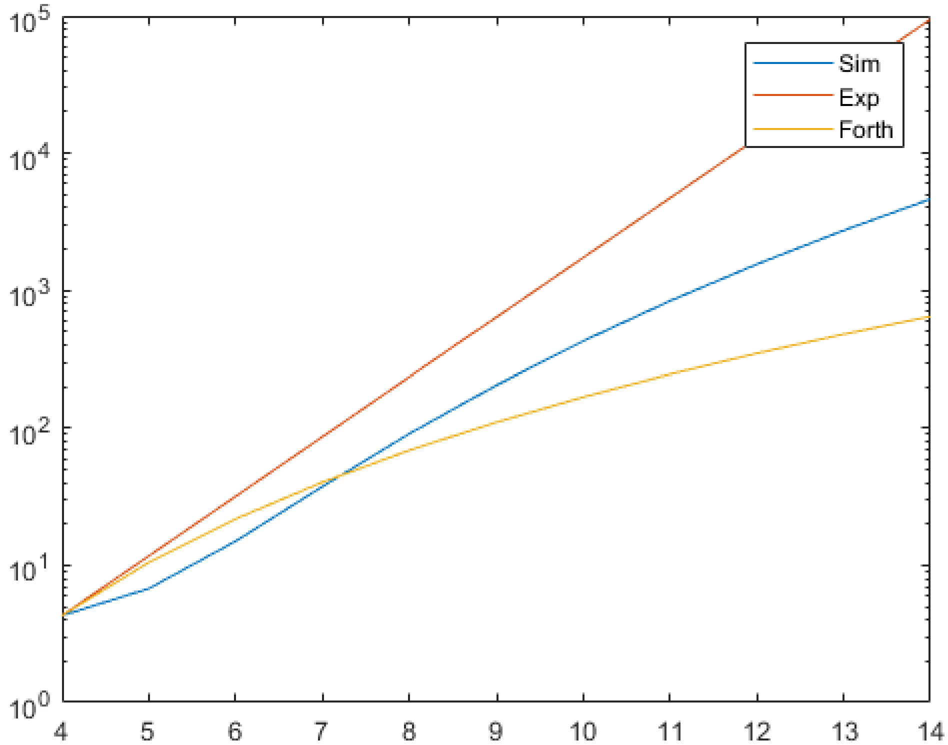


FIGURE 6.6. Comparison with exponential growth and fourth-order polynomial growth

CHAPTER 7

Applications

7.1. Projected Spaces and Uncertainty Prolongation

In this section, we discuss how the uncertainty propagates from the demand space to the path flow space in the traffic equilibrium problems. For $q \in \mathcal{R}_\Phi$, as we discussed in chapter 3, the corresponding path flow under the MNSM is given by

$$f_\Phi = M_\Phi q + N_\Phi \tag{7.1}$$

where $\Phi \in \mathcal{P}(1:m)$. Applying the SVD to M_Φ, we know that there exist two orthonormal matrices U_Φ and V_Φ such that $M_\Phi = U_\Phi \Sigma V_\Phi^T$. Substituting into equation (7.1), we obtain that

$$f_\Phi = U_\Phi \Sigma V_\Phi^T q + N_\Phi \implies U_\Phi^T f_\Phi = \Sigma V_\Phi^T q + U_\Phi^T N_\Phi \tag{7.2}$$

where $\Sigma = diag\{\sigma_1, \sigma_2, \ldots, \sigma_n, 0 \ldots, 0\}$. Let $\tilde{f}_\Phi = U_\Phi^T f_\Phi$ and $\tilde{q} = V_\Phi^T q$ be the projected path flow and the projected demand respectively. Denote $\tilde{N}_\Phi = U_\Phi^T N_\Phi$. Then, we end up with a neat relation between the projected path flow space and the demand space.

$$\tilde{f}_\Phi = \Sigma \tilde{q}_\Phi + \tilde{N}_\Phi \quad or \quad \tilde{f}_{\Phi i} = \sigma_i \tilde{q}_{\Phi i} + \tilde{N}_{\Phi i} \tag{7.3}$$

Furthermore since U_Φ is an orthonormal matrix $U_\Phi U_\Phi^T = U_\Phi^T U_\Phi = I$, we could recover the path flow by

$$f_\Phi = U_\Phi \tilde{f}_\Phi \tag{7.4}$$

From the above calculation, we could split the uncertainty propagation process into 4 steps:

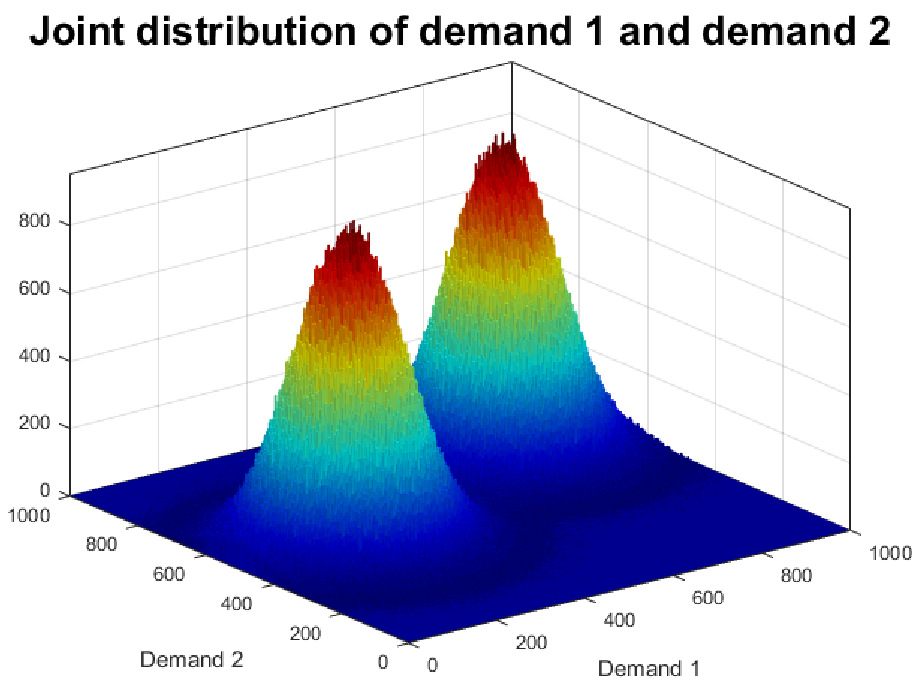
(1) Rotation in Demand space $\mathbb{R}_+^n : q \to V^T q = \tilde{q}$;
(2) Dilation : $\tilde{q} \to \Sigma \tilde{q}$;
(3) Translation: $\Sigma \tilde{q} \to \Sigma \tilde{q} + \tilde{N} = \tilde{f}$
(4) Rotation in path flow space $\mathbb{R}_+^m : \tilde{f} \to U \tilde{f} = f$

Now we use three different distributions of the demand to illustrate how projected spaces could clear the uncertainty propagation process.

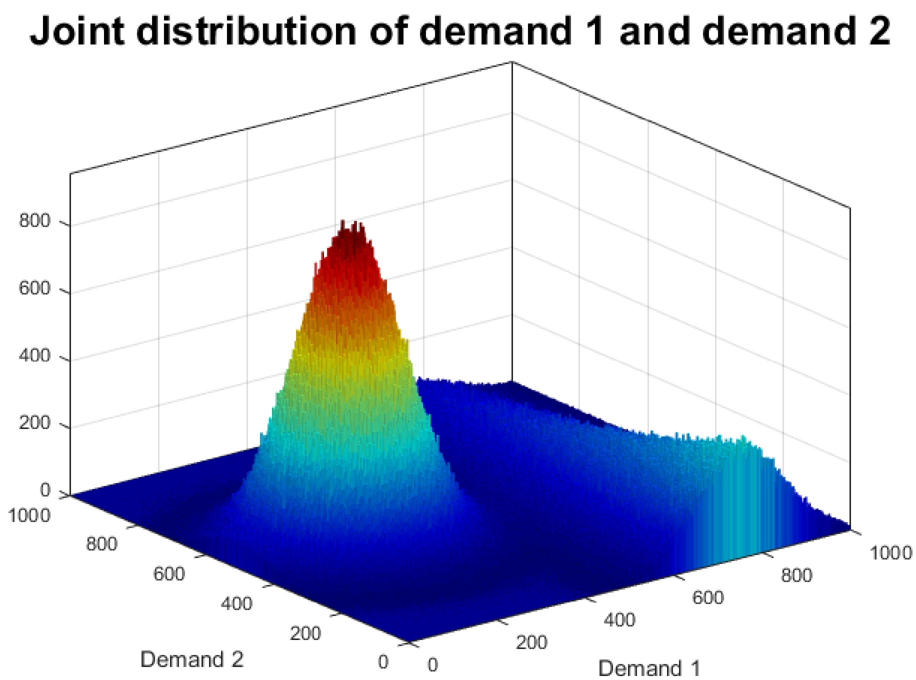
Example 10. we simulate three different distributions of the demand on the toy model network discussed in example 4.

(1) Uniform distribution where $(q_1, q_2) \in [0, 1000] \times [0, 1000]$.
(2) Two-peak distribution where $(q_1, q_2) \in [0, 1000] \times [0, 1000]$, one peak is at $(250, 500)$ with variation 100 and another peak is at $(750, 700)$ with variation 100 shown in figure 7.1(a).
(3) Mix distribution where $(q_1, q_2) \in [0, 1000] \times [0, 1000]$. This distribution is shown in figure 7.1(b)

In this simulation, we randomly sampled 10 millions demands $(q_1, q_2) \in [0, 1000] \times [0, 1000]$ obeying different distributions. In figure 7.2, the histogram of the path flows shows that no clear relation between the demand obeying uniform distribution and the corresponding path flows. Now we change the perspective to the projected spaces. As shown in figure 7.3, the projected demands result in some similar patterns of the corresponding projected paths but not exact the same when we look at the entire space. When we focus the comparisons within each partition region, the situation becomes more clear. In figure 7.4, 7.5, 7.6,7.7,7.8 and 7.9, the histograms of the projected demands exhibit similar patterns of the histograms of the corresponding projected path flows in each partition region except dilation and translation. These results indicate that the relation between projected demands and projected paths in each region is essential and independent of the input distribution. Furthermore, if we follow the traditional perspective, we could compare the relation between the demands and link flows. In these cases, even though we observe clear relation between the projected paths and projected demands, there is still no clear relation between demand and link flows. This provides a strong support to focus our attention of uncertainty propagation through exploring the demand-path instead of demand-link flow relation.

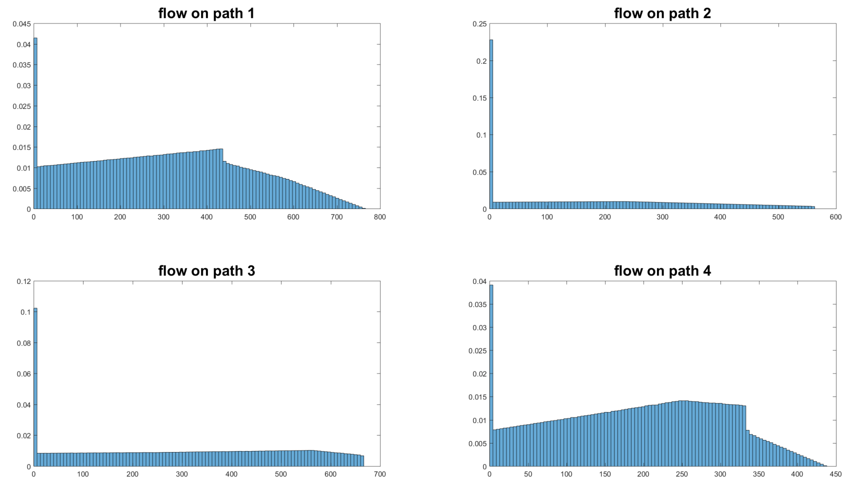


(a) Histogram of the joint distribution between demand q_1 and demand q_2 obeying two-peak distribution

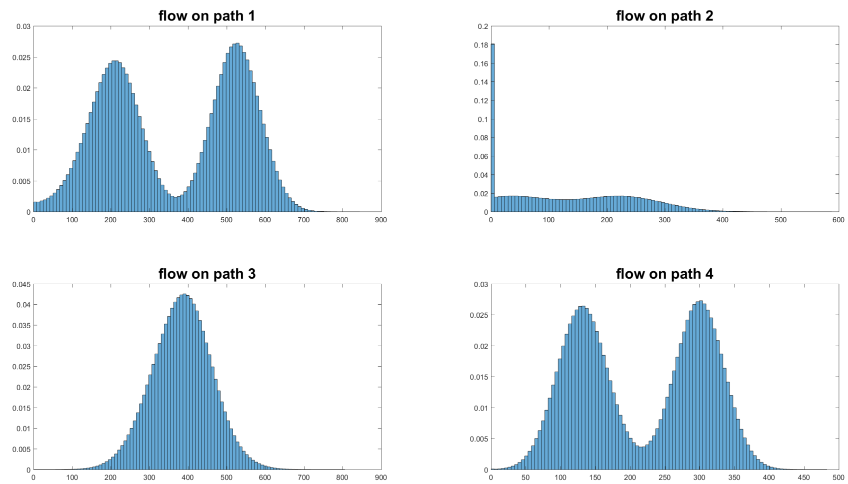


(b) Histogram of the joint distribution between demand q_1 and demand q_2 obeying mixed distribution

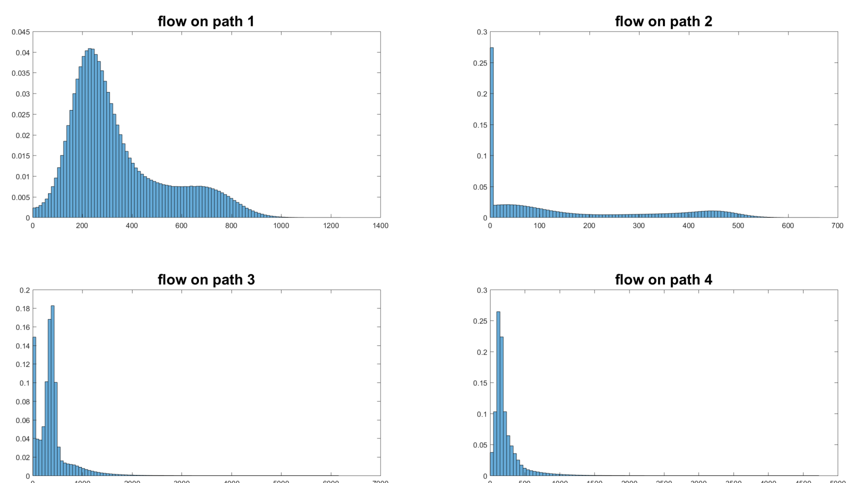
FIGURE 7.1. Distributions for simulation



(a) Demands obeying the uniform distribution

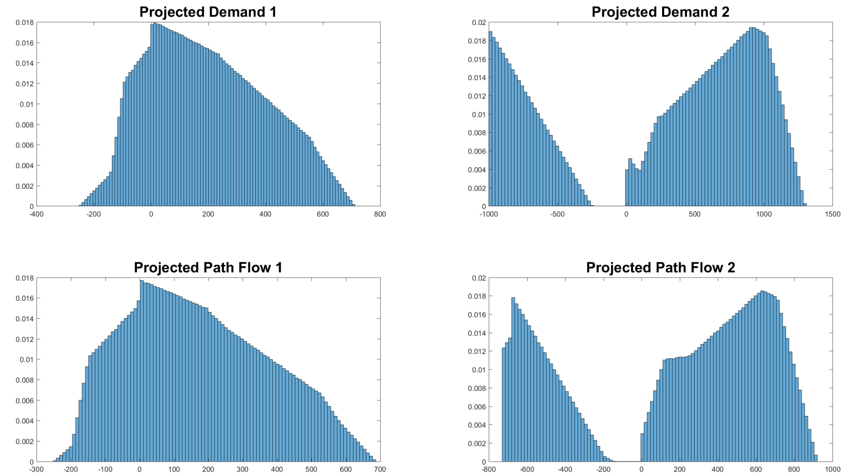


(b) Demands obeying the two-peak distribution

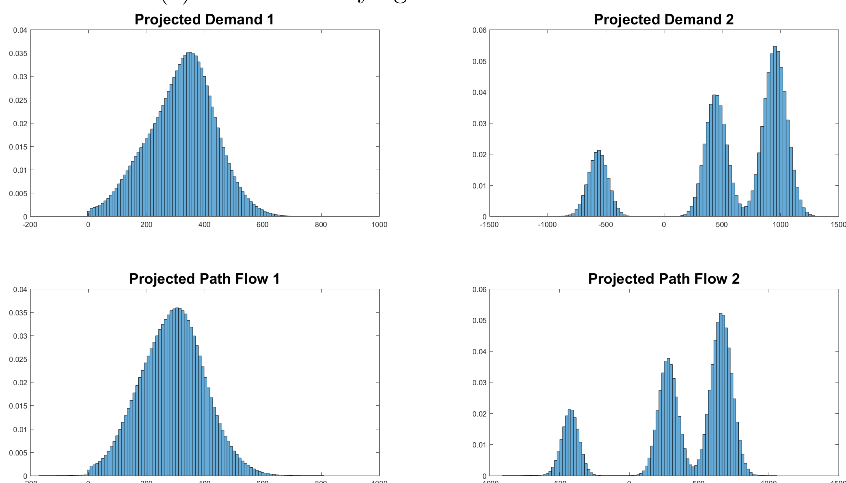


(c) Demands obeying the mixed distribution

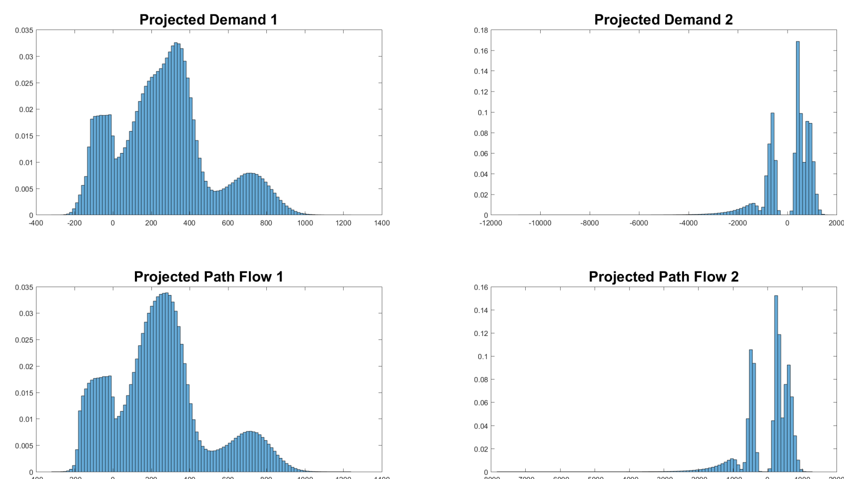
FIGURE 7.2. Histogram of the path flows



(a) Demands obeying the uniform distribution



(b) Demands obeying the two-peak distribution



(c) Demands obeying the mixed distribution

FIGURE 7.3. Histogram of the projected demands and the projected path flows for the entire space

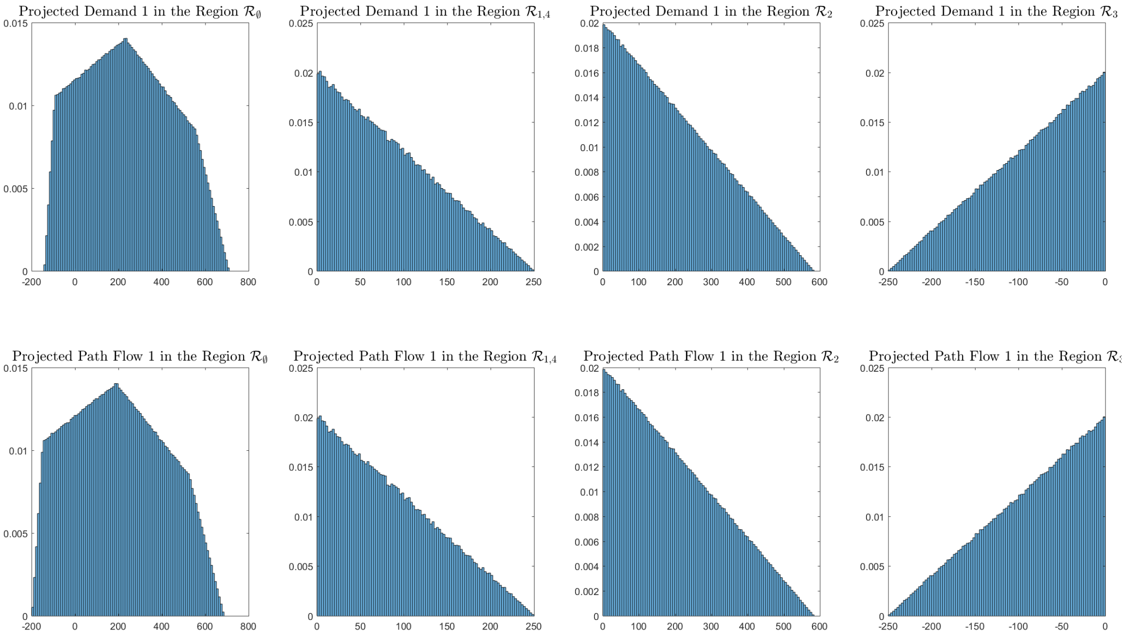


FIGURE 7.4. Comparison of the histogram of the projected demand 1 and the projected path flow 1 by regions

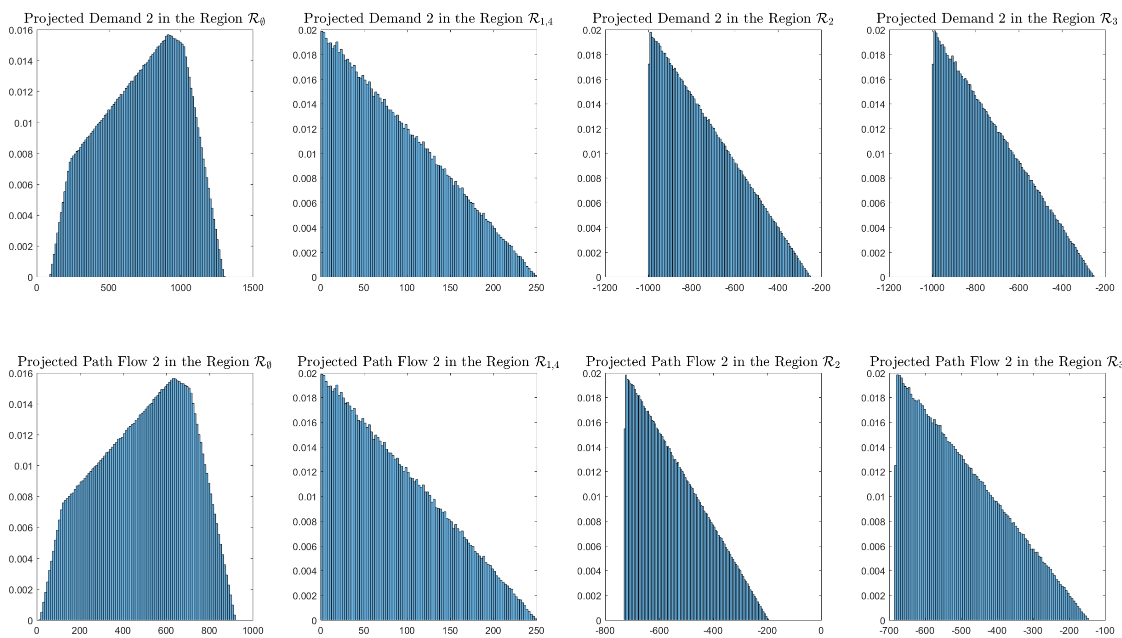


FIGURE 7.5. Comparison of the histogram of the projected demand 2 and the projected path flow 2 by regions

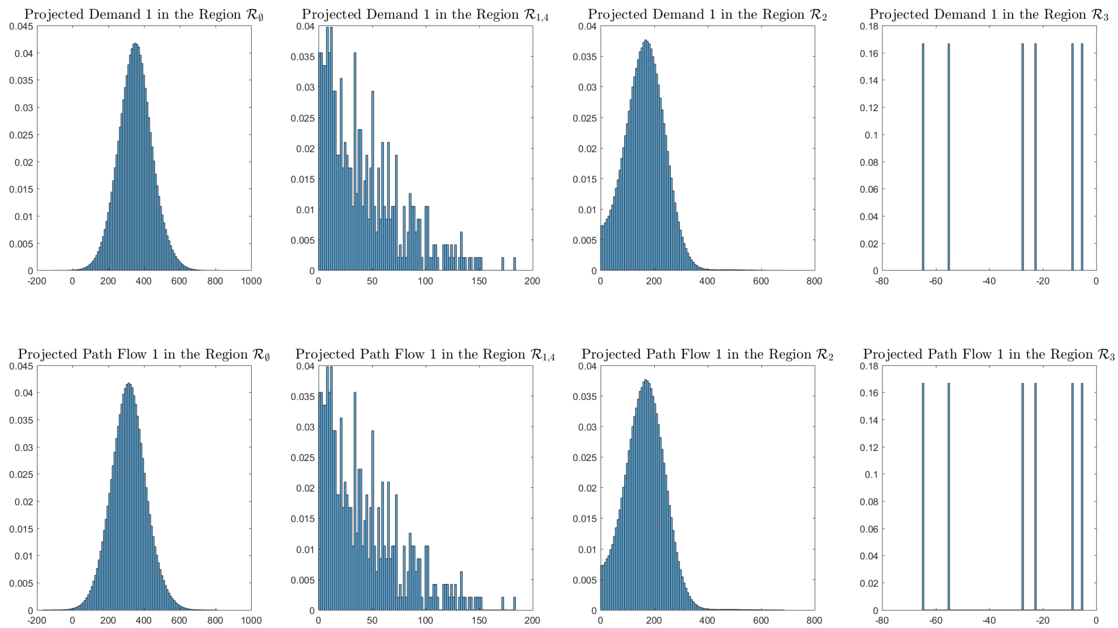


FIGURE 7.6. Comparison of the histogram of the projected demand 1 and the projected path flow 1 by regions

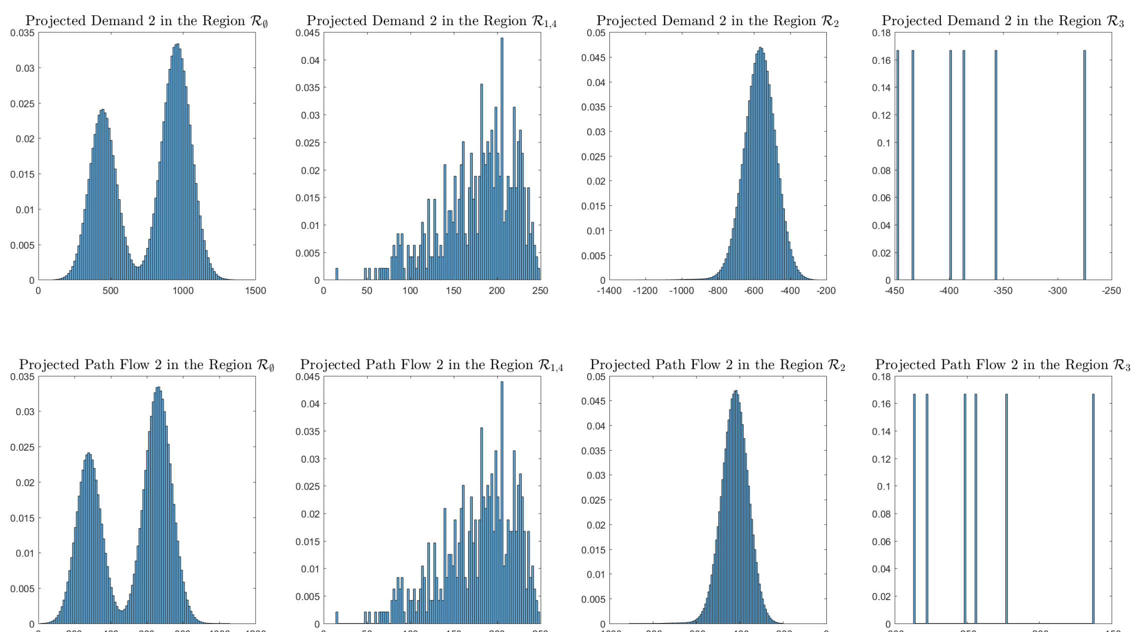


FIGURE 7.7. Comparison of the histogram of the projected demand 2 and the projected path flow 2 by regions

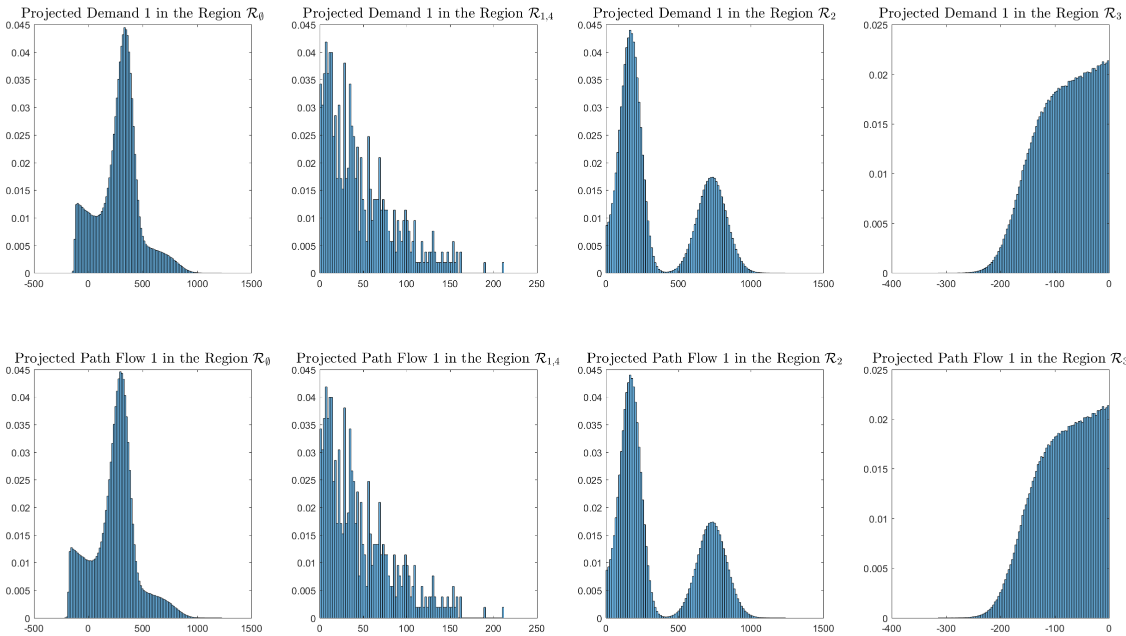


FIGURE 7.8. Comparison of the histogram of the projected demand 1 and the projected path flow 1 by regions

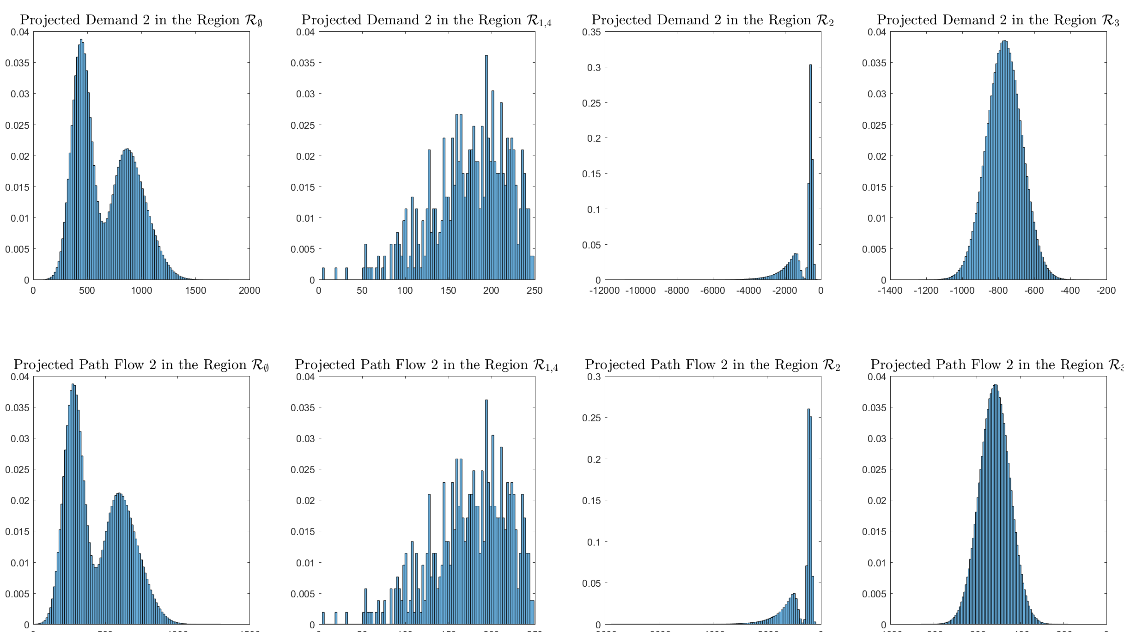
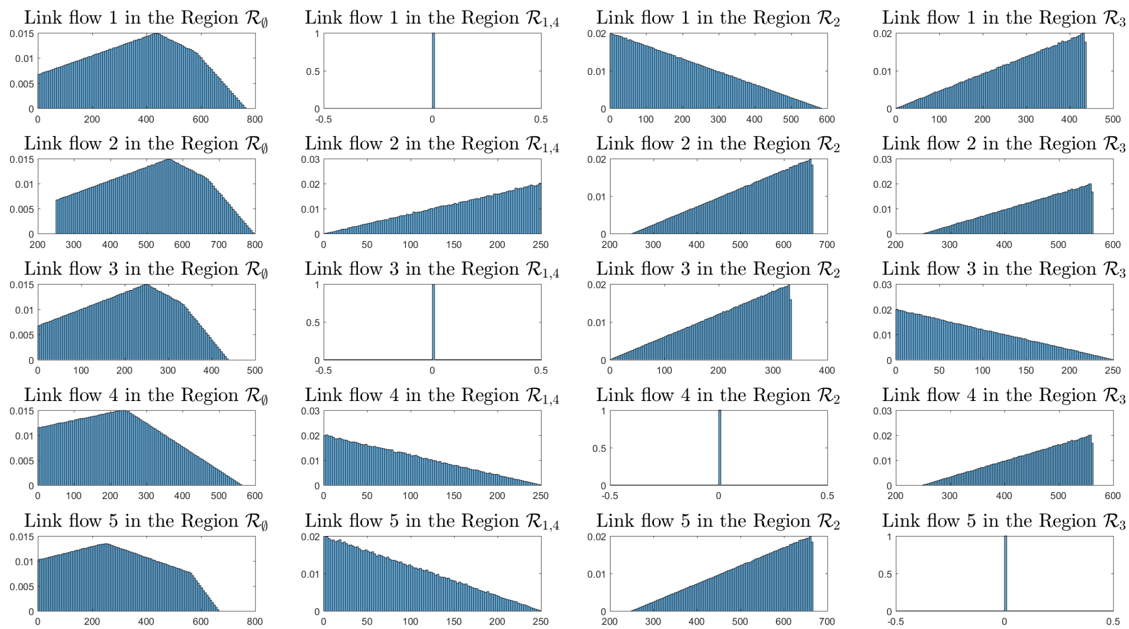


FIGURE 7.9. Comparison of the histogram of the projected demand 2 and the projected path flow 2 by regions


Link flow 1 in the Region $\mathcal{R}_\emptyset$
Link flow 1 in the Region $\mathcal{R}_{1,4}$
Link flow 1 in the Region $\mathcal{R}_2$
Link flow 1 in the Region $\mathcal{R}_3$
Link flow 2 in the Region $\mathcal{R}_\emptyset$
Link flow 2 in the Region $\mathcal{R}_{1,4}$
Link flow 2 in the Region $\mathcal{R}_2$
Link flow 2 in the Region $\mathcal{R}_3$
Link flow 3 in the Region $\mathcal{R}_\emptyset$
Link flow 3 in the Region $\mathcal{R}_{1,4}$
Link flow 3 in the Region $\mathcal{R}_2$
Link flow 3 in the Region $\mathcal{R}_3$
Link flow 4 in the Region $\mathcal{R}_\emptyset$
Link flow 4 in the Region $\mathcal{R}_{1,4}$
Link flow 4 in the Region $\mathcal{R}_2$
Link flow 4 in the Region $\mathcal{R}_3$
Link flow 5 in the Region $\mathcal{R}_\emptyset$
Link flow 5 in the Region $\mathcal{R}_{1,4}$
Link flow 5 in the Region $\mathcal{R}_2$
Link flow 5 in the Region $\mathcal{R}_3$

7.2. Identify Critical Network

In this section, we introduce an application to show how the MNSM could help identify the critical components of a network. In transportation assignment problems, there is a great interest in understanding which demand pairs might affect the network most significantly. In our setting, we separately consider the effects of a demand pair on the network locally (by considering a particular demand partitioning region) and globally (by considering the entire demand space).

Following the idea from the previous section, the uncertainty propagation along the network is determined by the relation between the projected demand and the projected path flow. For the given region $\mathcal{R}_U$, $f = M_U q + N_U$ and $\tilde{f} = \sigma_U \tilde{q} + \tilde{N_U}$. Hence, we have the following criteria to show which demand affects the network flow uncertainty more.

Definition 7.2.1. If $q \in \mathcal{R}_U$, then define the local criteria as

$$\eta_U := \mathbb{1}\Sigma|V_U^T| \tag{7.5}$$

where $\mathbb{1}$ is the row vector with all ones, $|V_U^T| = [|V_U^T|_{ij}]$ is element-wise absolute value. And define the global criteria

$$\eta := \frac{1}{N_{partition}} \sum_{\mathcal{R}_U \neq \emptyset} \eta_U \tag{7.6}$$

where $N_{partition}$ is the total number of nonempty partition regions.

Hence, the OD pair who has a larger η value implies a higher effect on the network flow uncertainty.

Example 11. Consider the network showing in figure 3.1 again.
If $U = \emptyset$, then

$$f^\dagger(b) = M_\emptyset q + N_\emptyset = \begin{bmatrix} 7/16 & 7/16 \\ 9/16 & -7/16 \\ -1/4 & 3/4 \\ 1/4 & 1/4 \end{bmatrix} q + \begin{bmatrix} -875/8 \\ 875/8 \\ 125/2 \\ -125/2 \end{bmatrix} \implies \tilde{f}^\dagger(b) = \begin{bmatrix} 1.0392 & 0 \\ 0 & 0.7487 \end{bmatrix} \tilde{q} - \begin{bmatrix} 51.2416 \\ 53.8786 \end{bmatrix} \tag{7.7}$$

So

$$\eta_\emptyset = \begin{bmatrix} 1.0822 & 1.2435 \end{bmatrix} \tag{7.8}$$

Since $\eta_{\emptyset}(2) > \eta_{\emptyset}(1)$, we conclude that if $q \in \mathcal{R}_{\emptyset}$, the second OD pair will affect the network flow uncertainty more than the first OD pair.

If $U = \{2\}$, then

$$f^{\dagger}(b) = M_U q + N_U = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 5/9 \\ 0 & 4/9 \end{bmatrix} q + \begin{bmatrix} 0 \\ 1000/9 \\ -1000/9 \end{bmatrix} \implies \tilde{f}^{\dagger}(b) = \begin{bmatrix} 1.0000 & 0 \\ 0 & 0.7115 \end{bmatrix} \tilde{q} - \begin{bmatrix} 0 \\ 17.3526 \end{bmatrix} \tag{7.9}$$

So

$$\eta_{\{2\}} = [1.00000.7115] \tag{7.10}$$

Since $\eta_{\emptyset}(1) > \eta_{\emptyset}(2)$, we conclude that if $q \in \mathcal{R}_{\{2\}}$, the first OD pair will affect the network flow uncertainty more than the second OD pair.

If $U = \{3\}$, then

$$f^{\dagger}(b) = M_U q + N_U = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 5/9 \\ 0 & 4/9 \end{bmatrix} q + \begin{bmatrix} 0 \\ 1000/9 \\ -1000/9 \end{bmatrix} \implies \tilde{f}^{\dagger}(b) = \begin{bmatrix} 1.0000 & 0 \\ 0 & 0.7169 \end{bmatrix} \tilde{q} + \begin{bmatrix} 0 \\ 33.9056 \end{bmatrix} \tag{7.11}$$

So

$$\eta_{\{3\}} = [0.71691.0000] \tag{7.12}$$

Since $\eta_{\emptyset}(2) > \eta_{\emptyset}(1)$, we conclude that if $q \in \mathcal{R}_{\{3\}}$, the second OD pair will affect the network flow uncertainty more than the first OD pair.

If $U = \{1, 4\}$, then

$$\eta_{\{1,4\}} = [1, 1]; \tag{7.13}$$

Since $\eta_{\emptyset}(2) = \eta_{\emptyset}(1)$, we conclude that if $q \in \mathcal{R}_{\{1,4\}}$, the second OD pair will affect the network flow uncertainty same as the first OD pair.

Therefore

$$\eta = \frac{1}{4}\left(\eta_{\emptyset} + \eta_{\{2\}} + \eta_{\{3\}} + \eta_{\{1,4\}}\right) = \begin{bmatrix} 0.9498 & 0.9887 \end{bmatrix} \tag{7.14}$$

which means over all the region the second OD pair has the most effect to the network flow uncertainty.

On the other hands, we could also obtain which path flow is more sensitive to the demands with the help of the MNSM.

DEFINITION 7.2.2. If $q \in \mathcal{R}_U$, then define the local criteria as

$$\tau_U := |M_U|\mathbb{1} \tag{7.15}$$

where $\mathbb{1}$ is the column vector with all ones, and $|M_U|$ is element-wise absolute value. And define the global criteria

$$\tau := \frac{1}{N_{partition}} \sum_{\mathcal{R}_U \neq \emptyset} \tau_U \tag{7.16}$$

Hence, the path flow who has larger τ value implies more sensitive to the OD demand.

EXAMPLE 12. Consider the same network as previous example.
If $U = \emptyset$, then

$$\tau_\emptyset = \begin{bmatrix} 0.8750 \\ 0.1250 \\ 0.5000 \\ 0.5000 \end{bmatrix} \tag{7.17}$$

Hence, the sensitivity of the path flow is in the order $f_1 > f_3, f_4 > f_2$.
If $U = \{2\}$, then

$$\tau_{\{2\}} = \begin{bmatrix} 1.0000 \\ 0 \\ 0.5556 \\ 0.4444 \end{bmatrix} \tag{7.18}$$

Hence, the sensitivity of the path flow is in the order $f_1 > f_3 > f_4 > f_2 = 0$.

If $U = \{3\}$, then

$$\tau_{\{3\}} = \begin{bmatrix} 0.5833 \\ 0.4167 \\ 0 \\ 1.0000 \end{bmatrix} \tag{7.19}$$

Hence, the sensitivity of the path flow is in the order $f_4 > f_1 > f_2 > f_3 = 0$.

If $U = \{1, 4\}$, then

$$\tau_{\{1,4\}} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \tag{7.20}$$

Hence, the sensitivity of the path flow is in the order $f_2 = f_3 > f_1 = f_4 = 0$.

Therefore,

$$\tau = \frac{1}{4}\left(\tau_{\{1,4\}} + \tau_{\{3\}} + \tau_{\{2\}} + \tau_{\emptyset}\right) = \begin{bmatrix} 0.6146 \\ 0.3854 \\ 0.5139 \\ 0.4861 \end{bmatrix} \tag{7.21}$$

which means overall the sensitivity of the path flow is in the order $f_1 > f_3 > f_4 > f_2$.

Furthermore, it is easy to extend the idea to link-level sensitivity analysis. Recall that the MNSM for a given region $\mathcal{R}_\Phi$ is $f_\Phi = M_\Phi q + N_\Phi$. Then according to the path-link incidence matrix, we have $x_\Phi = F f_\Phi = F M_\Phi q + F N_\Phi$. Hence, instead of applying SVD on M_Φ, we could find the SVD of $F M_\Phi$. Assume that the SVD of $F M_\Phi$ is given by

$$F M_\Phi = U'_\Phi \Sigma'_\Phi V'^{T}_\Phi \tag{7.22}$$

Then, we could define the criteria for link-level sensitivity analysis similar to definition 7.2.1 and 7.2.2.

7.3. Suboptimal assignment

In this section, we explore a suboptimal assignment scheme under the assumption of UE/SO with linear cost function. The main idea is that for a given demand q, we first ignore a small part of the demand q_ϵ and only assign the main part of the demand q_m, and then assign the ignored part of the demand. To gain benefit from this idea under the MNSM framework, we need the following three requirements:

(1) The demand q_ϵ is small enough to ignore;
(2) The partition region which contains the demand q_m is known;
(3) The assignment scheme for q_ϵ has less cost.

REMARK 7.3.1. *In fact, the demand q_ϵ small enough implies that $q_m, q \in \mathcal{R}_U$ which means q_ϵ is the direction such that $q - q_\epsilon$ do not cross the boundary of the region $\mathcal{R}_U$.*

Let f_ϵ and f_m be the corresponding path flow of q_ϵ and q_m respectively. Then, we have $Bf_\epsilon = q_\epsilon$ and $Bf_m = q_m$ with $q = q_\epsilon + q_m$. Furthermore, we want that f_m is assigned by the MNSM with q_m i.e. $f_m = M_U q_m + N_U$ when $q_m \in \mathcal{R}_U$. f_ϵ is assigned by some fast algorithm. Recall that when we have linear cost function on each link the corresponding UE/SO could be formulated as following:

$$\begin{aligned} \min_{f} \quad & \frac{1}{2} f^T A f + b^T f \\ s.t. \quad & Bf = q, f \geq 0 \end{aligned} \tag{7.23}$$

Define $H(f) = \frac{1}{2} f^T A f + b^T f$ for convenience. Then for $q \in \mathcal{R}_U$ the error of the suboptimal assignment is given by

$$err(q_\epsilon) = H(Bq_\epsilon + f_m) - H(f) \tag{7.24}$$

where $f = M_U q + N_U$ is the MNSM with demand q. Then we have

$$
\begin{aligned}
err(q_\epsilon) &= H(f_\epsilon + M_U q_m + N_U) - H(f) && (7.25)\\
&= H(f_\epsilon + M_U(q - q_\epsilon) + N_U) - H(f) && (7.26)\\
&= H(f_\epsilon + M_U q + N_U - M_U q_\epsilon) - H(f) && (7.27)\\
&= H(f_\epsilon - M_U q_\epsilon + f) - H(f) && (7.28)\\
&= (f_\epsilon - M_U q_\epsilon)^T \nabla H(f) + (f_\epsilon - M_U q_\epsilon)^T \nabla^2 H(f)(f_\epsilon - M_U q_\epsilon) && (7.29)\\
&= (f_\epsilon - M_U q_\epsilon)^T (A M_U q + A N_U + b) + (f_\epsilon - M_U q_\epsilon)^T A (f_\epsilon - M_U q_\epsilon) && (7.30)
\end{aligned}
$$

There are two ways to assign f_ϵ fast. One is the fixed assignment and another one is the random assignment. The fixed assignment is to assign the small part of demands to the certain paths which has been fixed. A special case of the fixed assignment is to assign the small part of demands to the path which has the maximum cost, i.e. f_ϵ is the solution of

$$
\begin{aligned}
\max_f \quad & \frac{1}{2} f^T A f + b^T f && (7.31)\\
\text{subject to} \quad & \mathcal{B} f = I && (7.32)\\
& f = \{0, 1\} && (7.33)
\end{aligned}
$$

THEOREM 7.3.1. *If the demand $q \in \mathcal{R}_U$ and the small part of the demands ignored first is q_ϵ, the error bound for the fixed assignment is given by*

$$|err(q_\epsilon)| \leq \|I - M_U B\|(\|A\|\|M_U\|\|q\| + \|A\|\|N_U\| + \|b\|)\|q_\epsilon\|_1 + \|A\|\|I - M_U B\|^2 \|q_\epsilon\|_1^2 \quad (7.34)$$

Proof: Since f is the MNSM with demand q,

$$H(f) \leq H(f'), \text{ where } f' \in \{f \mid Bf = q, f \geq 0\} \quad (7.35)$$

So the error bound $err(q_\epsilon) = H(Bq_\epsilon + f_m) - H(f) \geq 0$.

Applying the Cauchy-Schwarz inequality and the convexity of norm to (7.30), we have

$$\begin{aligned}
|(7.30)| &\leq |(f_\epsilon - M_U q_\epsilon)^T (AM_U q + AN_U + b)| + |(f_\epsilon - M_U q_\epsilon)^T A(f_\epsilon - M_U q| && (7.36)\\
&\leq \|f_\epsilon - M_U q_\epsilon\| \|AM_U q + AN_U + b\| + \|A\| \|f_\epsilon - M_U q_\epsilon\|^2 && (7.37)\\
&\leq \|f_\epsilon - M_U q_\epsilon\| (\|AM_U q\| + \|AN_U\| + \|b\|) + \|A\| \|f_\epsilon - M_U q_\epsilon\|^2 && (7.38)\\
&\leq \|f_\epsilon - M_U q_\epsilon\| (\|A\| \|M_U\| \|q\| + \|A\| \|N_U\| + \|b\|) + \|A\| \|f_\epsilon - M_U q_\epsilon\|^2 && (7.39)
\end{aligned}$$

Furthermore, we have

$$\begin{aligned}
\|f_\epsilon - M_U q_\epsilon\| &= \|f_\epsilon - M_U B f_\epsilon\| && (7.40)\\
&= \|(I - M_U B) f_\epsilon\| && (7.41)\\
&\leq \|(I - M_U B)\| \|f_\epsilon\| && (7.42)\\
&\leq \|(I - M_U B)\| \|f_\epsilon\|_1 && (7.43)
\end{aligned}$$

where the last inequality is based on the relation of 2-norm and 1-norm in vector space $\|x\| \leq \|x\|_1$ and I is identity matrix. Notice that $\|q_\epsilon\|_1 = \|Bf_\epsilon\|_1 = \|f_\epsilon\|_1$ since B is the incidence matrix between path and demand.

Therefore,

$$|err(q_\epsilon)| \leq \|I - M_U B\| (\|A\| \|M_U\| \|q\| + \|A\| \|N_U\| + \|b\|) \|q_\epsilon\|_1 + \|A\| \|I - M_U B\|^2 \|q_\epsilon\|_1^2 \quad (7.44)$$

Q.E.D.

COROLLARY 7.3.1. *If* $q_\epsilon \in Null(M_U)$, *then*

$$|err(q_\epsilon)| \leq (\|A\| \|M_U\| \|q\| + \|A\| \|N_U\| + \|b\|) \|q_\epsilon\|_1 + \|A\| \|q_\epsilon\|_1^2 \quad (7.45)$$

Proof: Since $q_\epsilon \in Null(M_U)$, $M_U q_\epsilon = 0$. Then following the proof of the above theorem, we could get

$$|err(q_\epsilon)| \leq (\|A\| \|M_U\| \|q\| + \|A\| \|N_U\| + \|b\|) \|q_\epsilon\|_1 + \|A\| \|q_\epsilon\|_1^2 \quad (7.46)$$

Q.E.D.

Another way is to assign the demand in a random way. In this case, f_ϵ is a multidimensional random variable based on a predefined random assignment scheme such that $Bf_\epsilon = q_\epsilon$, $E[f_\epsilon] = Pq_\epsilon$ and $Var[f_\epsilon] = \Sigma_f$ where $P = [p_{ij}]$ and $\sum_j p_{ij} = 1$ for all i. To show the error bound in this case, we need the following lemma.

LEMMA 7.3.2. *[MP92] Let ξ be a random variable with expected value μ and variance-covariance matrix Σ. Then*

$$E[\xi^T \Lambda \xi] = tr[\Lambda\Sigma] + \mu^T \Lambda \mu \tag{7.47}$$

THEOREM 7.3.3. *If the demand $q \in \mathcal{R}_U$ and the small part of the demands ignored first is q_ϵ The expected error bound for the random suboptimal assigment is given by*

$$E[|err(q_\epsilon)|] \le \|P - M_U\|(\|A\|\|M_U\|\|q\| + \|A\|\|N_U\| + \|b\|)\|q_\epsilon\| + \|A\|\|P - M_U\|^2\|q_\epsilon\|^2 \tag{7.48}$$

Proof: Take expected value on 7.30 and apply lemma 7.3.2, we have

$$\begin{aligned}
E[(7.30)] &= E\left[(f_\epsilon - M_U q_\epsilon)^T(AM_U q + AN_U + b) + (f_\epsilon - M_U q_\epsilon)^T A(f_\epsilon - M_U q_\epsilon)\right] \\
&= (E[f_\epsilon] - M_U q_\epsilon)^T(AM_U q + AN_U + b) + E\left[f_\epsilon^T A f_\epsilon\right] \\
&- E[f_\epsilon^T]AM_U q_\epsilon - (M_U q_\epsilon)^T A E[f_\epsilon] + (M_U q_\epsilon)^T A(M_U q_\epsilon) \\
&= (Pq_\epsilon - M_U q_\epsilon)^T(AM_U q + AN_U + b) + tr[A\Sigma_f] + (Pq_\epsilon)^T A(Pq_\epsilon) \\
&- (Pq_\epsilon)^T AM_U q_\epsilon - (M_U q_\epsilon)^T APq_\epsilon + (M_U q_\epsilon)^T AM_U q_\epsilon \\
&= (Pq_\epsilon - M_U q_\epsilon)^T(AM_U q + AN_U + b) + tr[A\Sigma_f] + (Pq_\epsilon - M_U q_\epsilon)^T A(Pq_\epsilon - M_U q_\epsilon)(\ *\ )
\end{aligned}$$

where Pq_ϵ is the expected value of the random assignment scheme and Σ_f is the variance-covariance matrix of the random assignment scheme. Applying the Cauchy-Schwarz inequality and the convexity of the norm to $(*)$, we obtain

$$\begin{aligned}
|(*)| &\le |(Pq_\epsilon - M_U q_\epsilon)^T(AM_U q + AN_U + b)| + |tr[A\Sigma_f]| + |(Pq_\epsilon - M_U q_\epsilon)^T A(Pq_\epsilon - M_U q_\epsilon)| \\
&\le \|(P - M_U)q_\epsilon\|\|AM_U q + AN_U + b\| + |tr[A\sigma]| + \|A\|\|(P - M_U)q_\epsilon\|^2 \\
&\le \|P - M_U\|(\|A\|\|M_U\|\|q\| + \|A\|\|N_U\| + \|b\|)\|q_\epsilon\| + \|A\|\|P - M_U\|^2\|q_\epsilon\|^2
\end{aligned}$$

Therefore,

$$E[|err(q_\epsilon)|] \leq \|P - M_U\|(\|A\|\|M_U\|\|q\| + \|A\|\|N_U\| + \|b\|)\|q_\epsilon\| + \|A\|\|P - M_U\|^2\|q_\epsilon\|^2 \quad (7.49)$$

Q.E.D.

Corollary 7.3.2. *If $q_\epsilon \in Null(M_U)$, then*

$$E[|err(q_\epsilon)|] \leq \|P\|(\|A\|\|M_U\|\|q\| + \|A\|\|N_U\| + \|b\|)\|q_\epsilon\| + \|A\|\|P\|^2\|q_\epsilon\|^2 \quad (7.50)$$

Proof: Repeat above proof, and notice $M_U q_\epsilon = 0$.**Q.E.D.**

In conclusion, when one chooses to follow the suboptimal assignment, the error or expected error is bounded by

$$C_1(\|q\| + \|q_\epsilon\|) + C_2\|q_\epsilon\|^2 \quad (7.51)$$

7.4. Toll Policy Design

In the transportation, adding a toll to a link is an effective way to change the behavior of each participant in a network system. A widely used way to design the toll policy relies on the assumption that adding a special designed toll based on the demand or link flow could transform the solution of a user equilibrium problem into the solution of the system optimal problems which means the selffish behavior of each participant coincides with the system optimal decision. In this section, we will show that with the help of the MNSM, we are able to design a global toll policy based on the information of the demand space instead of the work done by the predecessors which only have a toll policy for single demand or certain demand level. Meanwhile, we could also know the rate of change for the toll based on the demand and how a specially designed toll policy could affect the original network systems especially the link cost.

Recall the formulation of UE and SO:

$$
\begin{array}{rl|rl}
\textbf{UE:} & & \textbf{SO:} & \\
\min\limits_{x} & T_{ue}(x) & \min\limits_{x} & T_{so}(x) \\
\text{subject to} & Bf = q & \text{subject to} & Bf = q \\
& x = Ff & & x = Ff \\
& f \geq 0 & & f \geq 0
\end{array}
\tag{7.52}
$$

where $T_{ue}(x) = \sum\limits_{l} \int_0^{x_l} t_l(s)ds$, and $T_{so}(x) = \sum\limits_{l} x_l t_l(x_l)$.

As we known, those two equilibrium problems are both convex so that the KKT conditions provide a sufficient and necessary condition for the solution. Comparing those two formulations, it is easy to figure out that the only difference of the KKT conditions of those two problems is the gradient of the objective function.

$$
\begin{array}{l|l}
\textbf{UE:} & \textbf{SO:} \\
\dfrac{\partial T_{ue}(x)}{\partial x_l} = t_l(x_l) & \dfrac{\partial T_{so}(x)}{\partial x_l} = t_l(x_l) + x_l \dfrac{dt_l(x_l)}{dx_l}
\end{array}
\tag{7.53}
$$

As we discussed, we want to add tolls on the link cost based on link flow.

$$
\tilde{t_l}(x_l) = t_l(x_l) + T_l(x_l) \implies T_l(x_l) = x_l \frac{dt_l(x_l)}{dx_l} \tag{7.54}
$$

where T_l is toll price on link l and.

Clearly, if we set the toll price on link l equal to $x_l \frac{dt_l(x_l)}{dx_l}$, the solutions of the modified UE will automatically become the solutions of the original SO. With the help of the MNSM, we could further extend the tolls function depending on the demands.

$$
T_l(q) = [Ff^{\dagger}(q)]_l \frac{dt_l}{dx_l} \left([Ff^{\dagger}(q)]_l \right) \tag{7.55}
$$

where $[\cdot]_l$ means the lth row. Hence, the total toll cost function $PT(q)$ for paths would be

$$
PT(q) = F^T T(q) \tag{7.56}
$$

where $T(q) = \begin{bmatrix} T_1(q) \\ T_2(q) \\ \vdots \\ T_n(q) \end{bmatrix}$ and F is the link-path incidence matrix

Example 13. Once more, let's consider the toy model in example 4.

Then, we have

$$\frac{dt_1}{dx_1} = \frac{1}{7} \tag{7.57}$$

$$\frac{dt_2}{dx_2} = \frac{1}{5} \tag{7.58}$$

$$\frac{dt_3}{dx_3} = \frac{1}{4} \tag{7.59}$$

$$\frac{dt_4}{dx_4} = 0 \tag{7.60}$$

$$\frac{dt_5}{dx_5} = 0 \tag{7.61}$$

If $q \in \mathcal{R}_\emptyset$, we obtain the toll function on each link given by

$$T_1(q) = [Ff^\dagger(q)]_1 \frac{dt_1}{dx_1}\left([Ff^\dagger(q)]_1\right) = \frac{q_1}{16} + \frac{q_2}{16} - \frac{125}{8} \tag{7.62}$$

$$T_2(q) = [Ff^\dagger(q)]_2 \frac{dt_2}{dx_2}\left([Ff^\dagger(q)]_2\right) = \frac{q_1}{16} + \frac{q_2}{16} + \frac{275}{8} \tag{7.63}$$

$$T_3(q) = [Ff^\dagger(q)]_3 \frac{dt_3}{dx_3}\left([Ff^\dagger(q)]_3\right) = \frac{q_1}{16} + \frac{q_2}{16} - \frac{125}{8} \tag{7.64}$$

$$T_4(q) = [Ff^\dagger(q)]_4 \frac{dt_4}{dx_4}\left([Ff^\dagger(q)]_4\right) = 0 \tag{7.65}$$

$$T_5(q) = [Ff^\dagger(q)]_5 \frac{dt_5}{dx_5}\left([Ff^\dagger(q)]_5\right) = 0 \tag{7.66}$$

The total toll cost for f_1 is $PT_1(q) = T_1(q) = \frac{q_1}{16} + \frac{q_2}{16} - \frac{125}{8}$.

The total toll cost for f_2 is $PT_2(q) = T_2(q) + T_4(q) = \frac{q_1}{16} + \frac{q_2}{16} + \frac{275}{8}$.

The total toll cost for f_3 is $PT_3(q) = T_2(q) + T_5(q) = \frac{q_1}{16} + \frac{q_2}{16} + \frac{275}{8}$.

The total toll cost for f_4 is $PT_4(q) = T_3(q) = \frac{q_1}{16} + \frac{q_2}{16} - \frac{125}{8}$.

If $q \in \mathcal{R}_2$, we obtain the toll function on each link given by

$$T_1(q) = [Ff^\dagger(q)]_1 \frac{dt_1}{dx_1}\left([Ff^\dagger(q)]_1\right) = \frac{q_1}{7} \tag{7.67}$$

$$T_2(q) = [Ff^\dagger(q)]_2 \frac{dt_2}{dx_2}\left([Ff^\dagger(q)]_2\right) = \frac{q_2}{9} + \frac{200}{9} \tag{7.68}$$

$$T_3(q) = [Ff^\dagger(q)]_3 \frac{dt_3}{dx_3}\left([Ff^\dagger(q)]_3\right) = \frac{q_2}{9} - \frac{250}{8} \tag{7.69}$$

$$T_4(q) = [Ff^\dagger(q)]_4 \frac{dt_4}{dx_4}\left([Ff^\dagger(q)]_4\right) = 0 \tag{7.70}$$

$$T_5(q) = [Ff^\dagger(q)]_5 \frac{dt_5}{dx_5}\left([Ff^\dagger(q)]_5\right) = 0 \tag{7.71}$$

The total toll cost for f_1 is $PT_1(q) = T_1(q) = \frac{q_1}{7}$.

The total toll cost for f_2 is $PT_2(q) = T_2(q) + T_4(q) = \frac{q_2}{9} + \frac{200}{9}$.

The total toll cost for f_3 is $PT_3(q) = T_2(q) + T_5(q) = \frac{q_2}{9} + \frac{200}{9}$.

The total toll cost for f_4 is $PT_4(q) = T_3(q) = \frac{q_2}{9} - \frac{250}{8}$.

If $q \in \mathcal{R}_3$, we obtain the toll function on each link given by

$$T_1(q) = [Ff^\dagger(q)]_1 \frac{dt_1}{dx_1}\left([Ff^\dagger(q)]_1\right) = \frac{q_1}{12} - \frac{125}{6} \tag{7.72}$$

$$T_2(q) = [Ff^\dagger(q)]_2 \frac{dt_2}{dx_2}\left([Ff^\dagger(q)]_2\right) = \frac{q_1}{12} + \frac{125}{6} \tag{7.73}$$

$$T_3(q) = [Ff^\dagger(q)]_3 \frac{dt_3}{dx_3}\left([Ff^\dagger(q)]_3\right) = \frac{q_2}{4} \tag{7.74}$$

$$T_4(q) = [Ff^\dagger(q)]_4 \frac{dt_4}{dx_4}\left([Ff^\dagger(q)]_4\right) = 0 \tag{7.75}$$

$$T_5(q) = [Ff^\dagger(q)]_5 \frac{dt_5}{dx_5}\left([Ff^\dagger(q)]_5\right) = 0 \tag{7.76}$$

The total toll cost for f_1 is $PT_1(q) = T_1(q) = \frac{q_1}{12} - \frac{125}{6}$.

The total toll cost for f_2 is $PT_2(q) = T_2(q) + T_4(q) = \frac{q_1}{12} + \frac{125}{6}$.

The total toll cost for f_3 is $PT_3(q) = T_2(q) + T_5(q) = \frac{q_1}{12} + \frac{125}{6}$.

The total toll cost for f_4 is $PT_4(q) = T_3(q) = \frac{q_2}{4}$.

If $q \in \mathcal{R}_{1,4}$, we obtain the toll function on link 2 given by

$$T_2(q) = [Ff^\dagger(q)]_2 \frac{dt_2}{dx_2}\left([Ff^\dagger(q)]_2\right) = \frac{q_1}{5} + \frac{q_2}{5} \tag{7.77}$$

and the rest are equal to 0.

The total toll cost for f_2, f_3 are $PT_2(q) = PT_3(q) = \frac{q_1}{5} + \frac{q_2}{5}$.

Besides, the rate of change of the toll function depending on the demand

$$\frac{dT}{dx} = \frac{dt}{dx} + x\frac{d^2t}{dx^2} = h(x) \tag{7.78}$$

Solving this equation, we obtain that the cost function would be

$$t(x) = \int_1^x h(s)(\ln(x) - \ln(s))ds \tag{7.79}$$

$$\frac{dT}{dq} = h(x)\frac{dx}{dq} = h(x)E_lF\nabla f^\dagger(q) \quad a.e. \tag{7.80}$$

Example 14. If the cost function on the link is given by

$$t(x) = t_0 + \frac{Cx^4}{16} \tag{7.81}$$

then

$$h(x) = Cx^3 \tag{7.82}$$

The rate of change the tolls function depending on the demand is given by

$$\frac{dT}{dq} = C\left(E_lFf^\dagger(q)\right)E_lF\nabla f^\dagger(q) \quad a.e. \tag{7.83}$$

Example 15. If the cost function on the link is given by

$$t(x) = t_0 + C_1x + \frac{C_2x^4}{16} \tag{7.84}$$

then

$$h(x) = C_1 + C_2x^3 \tag{7.85}$$

The rate of change the tolls function depending on the demand is given by

$$\frac{dT}{dq} = \left(C_2 + C_1\left(E_lFf^\dagger(q)\right)\right)E_lF\nabla f^\dagger(q) \quad a.e. \tag{7.86}$$

where the almost everywhere is because the MNSM has discontinuities To quantify the critical demand for the toll cost function in each region, we have the following definition

DEFINITION 7.4.1. If $q \in \mathcal{R}_U$, then define the local criteria as

$$[\zeta_U]_j := \sum_i \left[|\frac{dT}{dq}| \right]_{i,j} \tag{7.87}$$

where j is corresponding to jth demand. And define the global criterion

$$\zeta = \frac{1}{N_{partition}} \sum_{\mathcal{R}_U \neq \emptyset} \zeta_U \tag{7.88}$$

Hence, the demand that has larger ζ value implies more contributions to the toll cost change. Similarly, we could have the following definition to see the toll on which link is more sensitive to the demand than others.

DEFINITION 7.4.2. If $q \in \mathcal{R}_U$, then define the local criteria as

$$[\kappa_U]_i := \sum_j \left[|\frac{dT}{dq} \right|]_{i,j} \tag{7.89}$$

where i is corresponding to ith link. And define the global criterion

$$\kappa = \frac{1}{N_{partition}} \sum_{\mathcal{R}_U \neq \emptyset} \zeta_U \tag{7.90}$$

EXAMPLE 16. Consider the network in the example 4, we could easily get if $q \in \mathcal{R}_\emptyset$,

$$\frac{dT}{dq} = \begin{bmatrix} \frac{1}{16} & \frac{1}{16} \\ \frac{1}{16} & \frac{1}{16} \\ \frac{1}{16} & \frac{1}{16} \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \tag{7.91}$$

if $q \in \mathcal{R}_2$,

$$\frac{dT}{dq} = \begin{bmatrix} \frac{1}{7} & 0 \\ 0 & \frac{1}{9} \\ 0 & \frac{1}{9} \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \tag{7.92}$$

if $q \in \mathcal{R}_3$,

$$\frac{dT}{dq} = \begin{bmatrix} \frac{1}{12} & 0 \\ \frac{1}{12} & 0 \\ 0 & \frac{1}{4} \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \tag{7.93}$$

and if $q \in \mathcal{R}_{1,4}$,

$$\frac{dT}{dq} = \begin{bmatrix} 0 & 0 \\ \frac{1}{5} & \frac{1}{5} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \tag{7.94}$$

Then, we have

$$\zeta_\emptyset = \left[\frac{3}{16}, \frac{3}{16}\right], \zeta_2 = \left[\frac{1}{7}, \frac{2}{9}\right], \zeta_3 = \left[\frac{1}{6}, \frac{1}{4}\right], \zeta_{1,4} = \left[\frac{1}{5}, \frac{1}{5}\right] \tag{7.95}$$

Hence, we could conclude that in the region $\mathcal{R}_\emptyset$ and $\mathcal{R}_{1,4}$ demand q_1 and demand q_2 have the same impact to the change of the tolls. But in the region $\mathcal{R}_2$ and $\mathcal{R}_3$, demand q_2 has the larger impact. Also, we have

$$\kappa_\emptyset = \begin{bmatrix} \frac{1}{8} \\ \frac{1}{8} \\ \frac{1}{8} \\ 0 \\ 0 \end{bmatrix}, \kappa_2 = \begin{bmatrix} \frac{1}{7} \\ \frac{1}{9} \\ \frac{1}{9} \\ 0 \\ 0 \end{bmatrix}, \kappa_3 = \begin{bmatrix} \frac{1}{12} \\ \frac{1}{12} \\ \frac{1}{4} \\ 0 \\ 0 \end{bmatrix}, \kappa_{1,4} = \begin{bmatrix} 0 \\ \frac{2}{5} \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{7.96}$$

Then, we could get that in the region $\mathcal{R}_\emptyset$, the toll functions are equally behaved, in the region $\mathcal{R}_2$ the toll function on link 1 is more sensitive to the demand change, in the region $\mathcal{R}_3$ the toll function on link 3 is more sensitive to the demand change, and in the region $\mathcal{R}_{1,4}$, only the toll function on link 2 is affected by the demand change.

CHAPTER 8

Future Work

8.1. Relation between Different Partition Regions

In this work, we find the analytic formula of the MNSM for CQP as a foundation to see how the uncertainty propagates. When different inequality constraints are activated, one could reach different partition region of the input space (demand space for traffic equilibrium problem). Due to this reason, our MNSM is defined piecewisely. Moreover, in different partition regions $\mathcal{R}_U$ and $\mathcal{R}_V$, the essential problem is how to compute the corresponding $K_U^\dagger$ and $K_V^\dagger$ to obtain the MNSM. In chapter 6, we have shown that we could use matrix decomposition technique to find those pseudoinverses of K_U or K_V. But as we have known, the cost of computing pseudo-inverse is very expensive. Hence, using some existing information of $K_U^\dagger$ to find a efficient way to compute $K_V^\dagger$ would be an interesting question to ask when we have some information of U and V such as $U \cap V$ is singleton. In fact, through a simple calculation, we have observed that if $U = V \cup \{i\}$ for some i, then K_U could treat as a rank-two update of K_V.

From a geometric perspective, the MNSM in different partition region is the image of the minimum norm point of the solution set under different projection mapping in some sense. So, to find the relation between different partition regions is equivalent to find a mapping between different partition regions which preserves the property of the minimum norm point unchanged. Moreover, if such mapping exists, is it possible to represent it as composition of functions?

$$MNSM = \Phi_1(\Phi_2(\cdots(\Phi_n(b)))) \tag{8.1}$$

8.2. Computational Robustness

From the numerical experiments, we realized that the algorithm to find the MNSM highly relies on the numerical accuracy. Even though the pseudoinverse of a matrix is numerically more stable than the inverse of the matrix, the partition region of the MNSM depends on the accuracy of each row of the pseudoinverse. Sometimes the inaccuracy causes two different partition regions having overlap and sometimes the inaccuracy causes the theoretically adjacent regions having gaps. These facts could make us overestimate or underestimate the number of partition regions. Hence, a more robust algorithm is needed to conquer this problem.

8.3. Larger Network

In this dissertation, we mainly focus on the theoretical development in building up the basis of the MNSM approaching to understand the uncertainty. The examples we implemented are limited to small or middle size networks. As we have shown in chapter 3, the criteria of the partition region are defined separately for each region which provide us an opportunity to build up a parallel algorithm to search the partition region. The only thing we don't know is whether those conditions defined an empty set or not. We believe that the number of regions in the partition would be a super small number comparing with the combinatorial upper bound.

8.4. Learning With Small Data

In recent years, machine learning has received more and more attention. Many excellent results and applications have emerged. But it is unavoidable that this method can only accomplish "big data and small tasks." Inspired by Boris Hanin and David Rolnick's recent paper [HR19], and combined with our work in this article, we propose the following imagination model shown in figure 8.1.

In the paper [HR19], Boris and David have showed that the neural network with ReLU function could partition the corresponding space into linear regions or polytopes. Meanwhile, through our work in this dissertation, the MNSM or PBSM of the equilibrium problem with linear cost function could also partition the corresponding space into linear regions or polytopes. Then if the last question in section 8.1 is true, each MNSM could correspond to a neural network with ReLU function and each neural network with ReLU function could correspond to a MNSM of some equilibrium problem. Based on this observation, training a neural network with ReLU and quadratic objective function could translate into find the MNSM of a suitable equilibrium problem. The advantage of this idea is that finding the MNSM only needs four information Q, c, A, b which would be much smaller number of parameters than a neural network. If the objective function is a

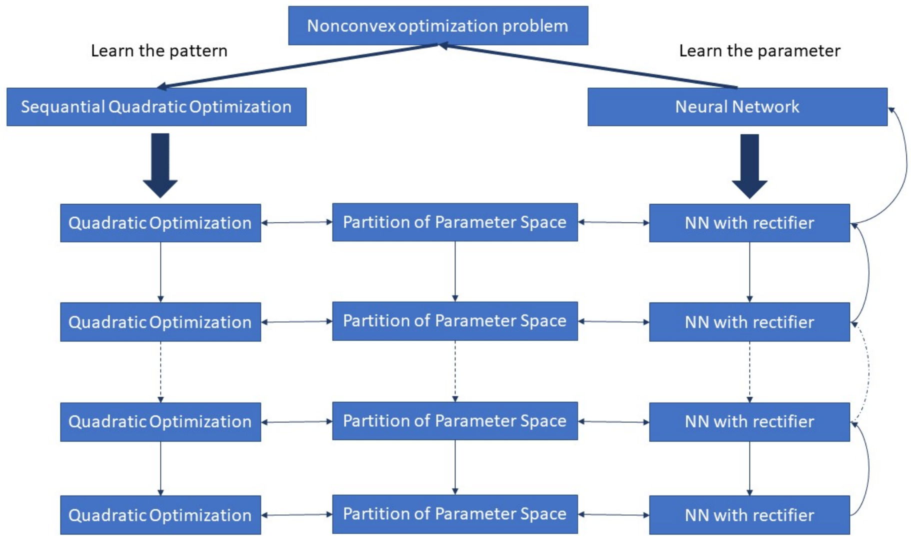


FIGURE 8.1. Framework of Learning with Small Data

general non-convex function, then following the idea of a very successful algorithm in non-convex optimization area – sequential quadratic programming, the strategy of training a ReLU network with non-convex objective function would be three steps. First, find a proper way to generate a sequence of optimization problem with quadratic objective function and only linear constraints. Second, for each subproblem, find the corresponding MNSM and its composition formula. At last, interpret the relevant results as a ReLU network.

Of course, there is no free lunch. Each step mentioned above may be very expensive. But I believe that the final answer to learn with small data would be finding a pattern or functional instead of point or function as we are doing recently.

APPENDIX A

Details of 3×3 grid network

The paths from node A to node I are following:

(1) $1 \rightarrow 2 \rightarrow 5 \rightarrow 10$

(2) $1 \rightarrow 2 \rightarrow 5 \rightarrow 7 \rightarrow 9 \rightarrow 12$

(3) $1 \rightarrow 2 \rightarrow 5 \rightarrow 7 \rightarrow 6 \rightarrow 8 \rightarrow 11 \rightarrow 12$

(4) $1 \rightarrow 4 \rightarrow 9 \rightarrow 12$

(5) $1 \rightarrow 4 \rightarrow 9 \rightarrow 11 \rightarrow 8 \rightarrow 6 \rightarrow 7 \rightarrow 10$

(6) $1 \rightarrow 4 \rightarrow 7 \rightarrow 10$

(7) $1 \rightarrow 4 \rightarrow 6 \rightarrow 8 \rightarrow 11 \rightarrow 12$

(8) $1 \rightarrow 4 \rightarrow 6 \rightarrow 8 \rightarrow 11 \rightarrow 9 \rightarrow 7 \rightarrow 10$

(9) $3 \rightarrow 8 \rightarrow 11 \rightarrow 12$

(10) $3 \rightarrow 8 \rightarrow 11 \rightarrow 9 \rightarrow 7 \rightarrow 10$

(11) $3 \rightarrow 8 \rightarrow 11 \rightarrow 9 \rightarrow 4 \rightarrow 2 \rightarrow 5 \rightarrow 10$

(12) $3 \rightarrow 6 \rightarrow 9 \rightarrow 12$

(13) $3 \rightarrow 6 \rightarrow 7 \rightarrow 10$

(14) $3 \rightarrow 6 \rightarrow 7 \rightarrow 5 \rightarrow 2 \rightarrow 4 \rightarrow 9 \rightarrow 12$

(15) $3 \rightarrow 6 \rightarrow 4 \rightarrow 2 \rightarrow 5 \rightarrow 10$

(16) $3 \rightarrow 6 \rightarrow 4 \rightarrow 2 \rightarrow 5 \rightarrow 7 \rightarrow 9 \rightarrow 12$

The paths from node B to node H are following:

(1) $6 \rightarrow 7$

(2) $6 \rightarrow 4 \rightarrow 2 \rightarrow 5$

(3) $6 \rightarrow 9 \rightarrow 12 \rightarrow 10$

(4) $3 \rightarrow 1 \rightarrow 2 \rightarrow 5$

(5) $3 \rightarrow 1 \rightarrow 4 \rightarrow 7$

(6) $3 \rightarrow 1 \rightarrow 4 \rightarrow 9 \rightarrow 12$

(7) $8 \to 11 \to 12 \to 10$

(8) $8 \to 11 \to 9 \to 7$

(9) $8 \to 11 \to 9 \to 4 \to 2 \to 5$

Region of Partitions:(23 different regions for demand in $[0, 10000] \times [0, 10000]$ v.s worst case analysis $\approx 2^{23}$)

(1) 1 4 6 9 12 13 17

(2) 1 4 6 9 12 17

(3) 1 4 6 9 17

(4) 1 4 9 12 17

(5) 1 4 9 17

(6) 1 4 9 17 20 23 24

(7) 1 4 9 17 20 24

(8) 1 4 9 17 24

(9) 4 6 9 12 13 17

(10) 4 6 9 12 17

(11) 4 6 12 13 17

(12) 4 6 12 17

(13) 4 9 12 17

(14) 4 9 17

(15) 4 9 17 20 23 24

(16) 4 9 17 20 24

(17) 4 9 17 24

(18) 4 12 17

(19) 4 17

(20) 4 17 20 23 24

(21) 4 17 20 24

(22) 4 17 23 24

(23) 4 17 24

Bibliography

[Aka01] T. Akamatsu, *An efficient algorithm for dynamic traffic equilibrium assignment with queues*, Transportation Science **35** (2001), no. 4, 389–404.

[AW83] H. Attouch and R. J.-B. Wets, *A convergence theory for saddle functions*, Transactions of the American Mathematical Society **280** (1983), no. 1, 1–41.

[Bao02] M. Baotić, *An efficient algorithm for multiparametric quadratic programming*, (2002).

[BGH92] D. P. Bertsekas, R. G. Gallager, and P. Humblet, *Data networks*, vol. 2, Prentice-Hall International New Jersey, 1992.

[BIG03] A. Ben-Israel and T. N. Greville, *Generalized inverses: theory and applications*, vol. 15, Springer Science & Business Media, 2003.

[BMDP02] A. Bemporad, M. Morari, V. Dua, and E. N. Pistikopoulos, *The explicit linear quadratic regulator for constrained systems*, Automatica **38** (2002), no. 1, 3–20.

[BMN05] D. E. Boyce, H. S. Mahmassani, and A. Nagurney, *A retrospective on Beckmann, Mcguire and Winsten's studies in the economics of transportation*, Papers in regional science **84** (2005), no. 1, 85–103.

[BMW55] M. J. Beckmann, C. B. McGuire, and C. B. Winsten, *Studies in the economics of transportation*, (1955).

[Bnn62] J. BnnoBRs, *Partitioning procedures for solving mixed-variables programming problems* , (1962).

[Boy80] D. E. Boyce, *A framework for constructing network equilibrium models of urban location*, Transportation Science **14** (1980), no. 1, 77–96.

[BPLM12] X. J. Ban, J.-S. Pang, H. X. Liu, and R. Ma, *Modeling and solving continuous-time instantaneous dynamic user equilibria: A differential complementarity systems approach*, Transportation Research Part B: Methodological **46** (2012), no. 3, 389–408.

[BV04] S. Boyd and L. Vandenberghe, *Convex optimization*, Cambridge university press, 2004.

[CA13] E. F. Camacho and C. B. Alba, *Model predictive control*, Springer Science & Business Media, 2013.

[Car86] M. Carey, *A constraint qualification for a dynamic traffic assignment model*, Transportation Science **20** (1986), no. 1, 55–58.

[Daf80] S. Dafermos, *Traffic equilibrium and variational inequalities*, Transportation science **14** (1980), no. 1, 42–54.

[Daf82] ______, *The general multimodal network equilibrium problem with elastic demand*, Networks **12** (1982), no. 1, 57–72.

[DR09] A. L. Dontchev and R. T. Rockafellar, *Implicit functions and solution mappings*, Springer Monographs in Mathematics. Springer **208** (2009).

[ESvBK18] O. A. Eikenbroek, G. J. Still, E. C. van Berkum, and W. Kern, *The boundedly rational user equilibrium: a parametric analysis with application to the network design problem*, Transportation Research Part B: Methodological **107** (2018), 1–17.

[FJO13] C. Feller, T. A. Johansen, and S. Olaru, *An improved algorithm for combinatorial multi-parametric quadratic programming*, Automatica **49** (2013), no. 5, 1370–1376.

[FMT08] M. Florian, M. Mahut, and N. Tremblay, *Application of a simulation-based dynamic traffic assignment model*, European Journal of Operational Research **189** (2008), no. 3, 1381–1392.

[Fuk04] K. Fukuda, *Lecture: polyhedral computation*, Tech. report, Research Report, Department of Mathematics, and Institute of Theoretical , 2004.

[GBN11] A. Gupta, S. Bhartiya, and P. Nataraj, *A novel approach to multiparametric quadratic programming*, Automatica **47** (2011), no. 9, 2112–2117.

[GF17] Z. Guo and Y. Fan, *A stochastic multi-agent optimization model for energy infrastructure planning under uncertainty in an oligopolistic market*, Networks and Spatial Economics **17** (2017), no. 2, 581–609.

[GG12] T. Gal and H. J. Greenberg, *Advances in sensitivity analysis and parametric programming*, vol. 6, Springer Science & Business Media, 2012.

[GVJ90] J. Guddat, F. G. Vazquez, and H. T. Jongen, *Parametric optimization: singularities, pathfollowing and jumps*, Springer, 1990.

[HA05] B. G. Heydecker and J. Addison, *Analysis of dynamic traffic equilibrium with departure time choice*, Transportation Science **39** (2005), no. 1, 39–57.

[Hos80] J. R. Hosking, *The multivariate portmanteau statistic*, Journal of the American Statistical Association **75** (1980), no. 371, 602–608.

[HR19] B. Hanin and D. Rolnick, *Complexity of linear regions in deep networks*, arXiv preprint arXiv:1901.09021 (2019).

[JP07] M. Josefsson and M. Patriksson, *Sensitivity analysis of separable traffic equilibrium equilibria with application to bilevel optimization in network design*, Transportation Research Part B: Methodological **41** (2007), no. 1, 4–31.

[KZ16] M. Köppe and Y. Zhou, *Toward computer-assisted discovery and automated proofs of cutting plane theorems*, International Symposium on Combinatorial Optimization, Springer, 2016, pp. 332–344.

[LS02] H. K. Lo and W. Y. Szeto, *A cell-based dynamic traffic assignment model: formulation and properties*, Mathematical and computer modelling **35** (2002), no. 7-8, 849–865.

[Mah01] H. S. Mahmassani, *Dynamic network traffic assignment and simulation methodology for advanced system management applications*, Networks and spatial economics **1** (2001), no. 3-4, 267–292.

[MFT08] M. Mahut, M. Florian, and N. Tremblay, *Comparison of assignment methods for simulation-based dynamic-equilibrium traffic assignment*, Proceeding of the Transportation Research Board 87th Annual Meeting (DVD), Washington, DC, 2008.

[MN78] D. K. Merchant and G. L. Nemhauser, *Optimality conditions for a dynamic traffic assignment model*, Transportation Science **12** (1978), no. 3, 200–207.

[MP92] A. M. Mathai and S. B. Provost, *Quadratic forms in random variables: theory and applications*, Dekker, 1992.

[MP07] P. Marcotte and M. Patriksson, *Traffic equilibrium*, Handbooks in Operations Research and Management Science **14** (2007), 623–713.

[Nag86] A. Nagurney, *Computational comparisons of algorithms for general asymmetric traffic equilibrium problems with fixed and elastic demands*, Transportation Research Part B: Methodological **20** (1986), no. 1, 78–84.

[Nag06] ———, *On the relationship between supply chain and transportation network equilibria: A supernetwork equivalence with computations*, Transportation Research Part E: Logistics and Transportation Review **42** (2006), 293–316.

[ND02] A. Nagurney and J. Dong, *A multiclass, multicriteria traffic network equilibrium model with elastic demand*, Transportation Research Part B: Methodological **36** (2002), no. 5, 445–469.

[NM65] J. A. Nelder and R. Mead, *A simplex method for function minimization*, The computer journal **7** (1965), no. 4, 308–313.

[NM07] A. Nagurney and D. Matsypura, *A supply chain network perspective for electric power generation, supply, transmission, and consumption*, Optimisation, Econometric and Financial Analysis (Berlin, Heidelberg) (E. J. Kontoghiorghes and C. Gatu, eds.), Springer Berlin Heidelberg, 2007, pp. 3–27.

[NW06] J. Nocedal and S. Wright, *Numerical optimization*, Springer Science & Business Media, 2006.

[OD04] S. B. Olaru and D. Dumur, *A parameterized polyhedra approach for explicit constrained predictive control*, 2004 43rd IEEE Conference on Decision and Control (CDC)(IEEE Cat. No. 04CH37601), vol. 2, IEEE, 2004, pp. 1580–1585.

[ORS93] A. Orda, R. Rom, and N. Shimkin, *Competitive routing in multiuser communication networks*, IEEE/ACM Transactions on Networking **1** (1993), no. 5, 510–521.

[Pat04] M. Patriksson, *Sensitivity analysis of traffic equilibria*, Transportation Science **38** (2004), no. 3, 258–281.

[Pat15] ———, *The traffic assignment problem: models and methods*, Courier Dover Publications, 2015.

[PHRU12] J.-S. Pang, L. Han, G. Ramadurai, and S. Ukkusuri, *A continuous-time linear complementarity system for dynamic user equilibria in single bottleneck traffic flows*, Mathematical programming **133** (2012), no. 1-2, 437–460.

[PY84] J.-S. Pang and C.-S. Yu, *Linearized simplicial decomposition methods for computing traffic equilibria on networks*, Networks **14** (1984), no. 3, 427–438.

[QM89] Y. Qiu and T. L. Magnanti, *Sensitivity analysis for variational inequalities defined on polyhedral sets*, Mathematics of Operations Research **14** (1989), no. 3, 410–432.

[RB16] E. K. Ryu and S. Boyd, *Primer on monotone operator methods*, Appl. Comput. Math **15** (2016), no. 1, 3–43.

[RW09] R. T. Rockafellar and R. J.-B. Wets, *Variational analysis*, vol. 317, Springer Science & Business Media, 2009.

[RW12] J. Royset and R. Wets, *Epi-splines and exponential epi-splines: Pliable approximation tools*, University of California Davis, Tech. Rep (2012).

[RW14] J. O. Royset and R. J. Wets, *From data to assessments and decisions: Epi-spline technology*, Bridging data and decisions, INFORMS, 2014, pp. 27–53.

[RW16] ———, *Multivariate epi-splines and evolving function identification problems*, Set-Valued and Variational Analysis **24** (2016), no. 4, 517–545.

[RWW+13] S. M. Ryan, R. J.-B. Wets, D. L. Woodruff, C. Silva-Monroy, and J.-P. Watson, *Toward scalable, parallel progressive hedging for stochastic unit commitment*, 2013 IEEE Power & Energy Society General Meeting, IEEE, 2013, pp. 1–5.

[SGDD03] M. M. Seron, G. C. Goodwin, and J. A. De Doná, *Characterisation of receding horizon control for constrained linear systems*, Asian Journal of Control **5** (2003), no. 2, 271–286.

[She85] Y. Sheffi, *Urban transportation networks*, vol. 6, Prentice-Hall, Englewood Cliffs, NJ, 1985.

[SKJ+06] J. SpjoTvold, E. C. Kerrigan, C. N. Jones, P. ToNdel, and T. A. Johansen, *On the facet-to-facet property of solutions to convex parametric quadratic programs*, Automatica **42** (2006), no. 12, 2209–2214.

[Smi79] M. J. Smith, *The existence, uniqueness and stability of traffic equilibria*, Transportation Research Part B: Methodological **13** (1979), no. 4, 295–304.

[TF88] R. L. Tobin and T. L. Friesz, *Sensitivity analysis for equilibrium network flow*, Transportation Science **22** (1988), no. 4, 242–250.

[TJB03a] P. ToNdel, T. A. Johansen, and A. Bemporad, *An algorithm for multi-parametric quadratic programming and explicit mpc solutions*, Automatica **39** (2003), no. 3, 489–497.

[TJB03b] P. Tondel, T. A. Johansen, and A. Bemporad, *Further results on multiparametric quadratic programming*, 42nd IEEE International Conference on Decision and Control

(IEEE Cat. No. 03CH37475), vol. 3, IEEE, 2003, pp. 3173–3178.

[War52] J. G. Wardrop, *Road paper. some theoretical aspects of road traffic research.*, Proceedings of the institution of civil engineers **1** (1952), no. 3, 325–362.

[WFT90] B.-W. Wie, T. L. Friesz, and R. L. Tobin, *Dynamic user optimal traffic assignment on congested multidestination networks*, Transportation Research Part B: Methodological **24** (1990), no. 6, 431–442.

[Yan97] H. Yang, *Sensitivity analysis for the elastic-demand network equilibrium problem with applications*, Transportation Research Part B: Methodological **31** (1997), no. 1, 55–70.

[YY05] J. Q. Ying and H. Yang, *Sensitivity analysis of stochastic user equilibrium flows in a bi-modal network with application to optimal pricing*, Transportation Research Part B: Methodological **39** (2005), no. 9, 769–795.

[ZDN+03] D. Zhang, J. Dong, A. Nagurney, et al., *Innovations in financial and economic networks*, (2003).