國 立 交 通 大 學
運 輸 與 物 流 管 理 學 系

博 士 論 文

路網均衡流量之高階敏感度分析

High-Order Sensitivity Analysis of Equilibrium Network Flows

研 究 生：黃 友 恆

指導教授：卓 訓 榮 教授
方 述 誠 教授

中 華 民 國 一 〇 三 年 七 月

路網均衡流量之高階敏感度分析

High-Order Sensitivity Analysis of Equilibrium Network Flows

研　究　生：黃友恆　　　　Student：You-Heng Huang

指導教授：卓訓榮　　　　Advisors：Dr. Hsun-Jung Cho

　　　　　方述誠　　　　　　　　　Dr. Shu-Cherng Fang

國 立 交 通 大 學

運 輸 與 物 流 管 理 學 系

博 士 論 文

A Dissertation

Submitted to Department of Transportation and Logistics Management

College of Management

National Chiao Tung University

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

in

Traffic and Transportation

July 2014

Hsinchu, Taiwan, Republic of China

中　華　民　國 一 〇 三 年 七 月

High-Order Sensitivity Analysis of Equilibrium Network Flows

Student：You-Heng Huang　　　　　Advisors：Dr. Hsun-Jung Cho
Dr. Shu-Cherng Fang

Department of Transportation and Logistics Management
National Chiao Tung University

ABSTRACT

Sensitivity analysis of equilibrium network flows is useful in various fields, such as bilevel network design problems, road pricing and origin-destination matrix estimation problems. The problems mentioned above can be formulated as a Stackelberg game where the upper level problem aims to find the optimal strategy which maximizes the system performance, and the lower level problem aims to solve the user equilibrium problem, respectively. The reaction function of the lower level problem is the key to solving the Stackelberg game. Due to the characteristics of user equilibria, the lower level problem does not have an explicit reaction function. Usually, the reaction function is approximated by the sensitivity information of equilibrium network flows. By performing such sensitivity analysis, one can predict the directions of variation in the equilibrium patterns when the parameters of cost and demand functions are changed. With this information, the linear approximation of the reaction function can be obtained and applied to solve Stackelberg the game using a sensitivity analysis-based algorithm.

The models involved usually exhibit a user equilibrium constraint to form a difficult nonlinear, nonconvex optimization problem. Due to the computational difficulties, a nonlinear approximation of the reaction function is incorporated for solving the problem more efficiently. This research tries to establish the theory of higher-order sensitivity analysis of network equilibrium flows in order to solve the problem with a nonlinear approximation of the reaction function.

This research is also going to extend the applicability of directional derivative-based sensitivity analysis method. To generalize the directional derivative-based sensitivity analysis, the continuous differentiability assumption on the cost function is relaxed to be piecewise linear functions. Building on the original directional derivative-based method, an extended model will be studied for providing the required sensitivity information using piecewise linear cost functions.

Acknowledgements

I would like to express my gratitude to my advisors, Dr. Hsun-Jung Cho and Dr. Shu-Cherng Fang, for their guidance, patience, and encouragement during my doctoral research. Besides my advisors, I would like to thank my committee members, Dr. Hsin-Li Chang, Dr. Yu-Chiun Chiou, Dr. Chien-Hung Wei and Dr. Suh-Wen Chiou, for their valuable comments that improved the presentation and contents of this dissertation.

I would also like to thank my friends and labmates for their support and encouragement. Finally, I am truly grateful to my parents and family. This dissertation would not have been possible without their warm love, continued patience, and endless support.

Table of Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

1.1 Research Motivation

Sensitivity analysis has been one of the most important aspects of applied mathematical programming, particularly in linear programming applications. Sensitivity analysis of network equilibrium models is very different from linear programming sensitivity analysis, yet it is no less important. Because of the essential nonlinearity of network equilibrium problems, any parameters perturbations will generally result in a change in the equilibrium solution. The problem of sensitivity analysis then becomes the problem of numerically approximating a new equilibrium solution resulting from any of a variety of parameter perturbations which may occur simultaneously. Therefore, the sensitivity analysis is particularly useful in control and pricing applications because if we can anticipate the effects of a change in the traffic infrastructure on the behavior of travelers, then we can utilize this knowledge to optimize these changes according to some goal fulfillments, such as a reduction in flows, higher revenue from congestion tolls, etc.

Methods of sensitivity analysis for nonlinear programming problems [29] and for variational inequality problems [21, 40, 53, 59] have been applied to spatial price equilibrium problems [10, 22, 60]. However, direct application of these methods to the variational inequality formulation of the equilibrium network flow problem is not feasible since its solutions do not typically satisfy the required local uniqueness conditions. This is primarily due to the presence of path variables in the problem formulation. As a consequence, computational procedures which have thus far been proposed to find the gradients of arc-flow variables with respect to parameter perturbations require the determination of an unperturbed equilibrium path-flow vector with a restricted number of active paths (as for example in [24] and [61]). However, since traditional algorithms (such as the Frank-Wolfe feasible direction algorithm) usually terminate with approximate solutions which do not satisfy these restrictions, one must employ auxiliary search procedures to find such path-flow vectors (as in the linear-programming approach of [61]). Thus, two kinds of methods, gradient-based method and directional derivative-based method, are proposed to overcome this issue. Based on these two kinds of methods, the sensitivity analysis of network equilibrium problems and its widespread applications have received large attention recently. It can be applied to solve bilevel network design problems, network signal control problems, toll pricing and origin-destination (OD) matrix estimation problems [18, 19, 30, 35, 65, 69, 71, 72]. Furthermore, the sensitivity analysis for combined distribution assignment model [57], combined modal split assignment model [72], and combined travel demand model [68] were also proposed.

The network design problem is, in its most general form, "to choose facilities to add to a transportation network or to determine capacity enhancements of existing facilities of a transportation network which are, in some sense, optimal" [33]. Usually, this kind of problem can be casted as a Stackelberg game or a bilevel problem where the upper level problem aims to find the optimal signal setting or capacity enhancement of arcs which maximizes system performance and the lower level problem aims to solve the user equilibrium flows, respectively. In this research we will be concerned with one particular manifestation of this problem. For our purposes it is the problem of determining the link additions or link capacity enhancements (within an existing transportation network) that will "best" improve the traffic situation.

Network design models can be classified in the following ways [25]: whether the investment decision variables are discrete (i.e., network design models are formulated in terms of discrete links to be considered for addition to an existing network) or continuous (i.e., network design models are formulated in terms of increase in the divisible capacity of each link); whether the traffic assignment is a user equilibrium (i.e., each user tries to minimize his/her own journey time) or system equilibrium (i.e., the total journey time is minimized). In any case, the network might include any combination of rail, highway, or mass transit modes.

For highway networks, which are the focus of this research, the user equilibrium model is much more realistic than the system equilibrium model; the system traffic equilibrium is an idealized target which will not be observed in practice unless cooperation among individual users is introduced, or, what is equivalent, a central authority with power to modify individual route choices is established. In other words, we recognize that individual drivers are free to make decisions, and that the transportation flow pattern which actually occurs will be derived from these decisions. This greatly complicates the solution of the network design problem. Clearly the mathematical program employed to determine optimal design or capacity enhancement program for such a network must include wither implicit or explicit constraints to ensure that the flow pattern is in accord with the expected behavior of users and providers of transportation. However, when we include such constraints, which create a nonconvex feasible set, we leave open the possibility that the new equilibrium flow redistribution actually degrades the optimization criterion employed. This phenomenon was firstly pointed out by Braess [8] for scalar network design problems employing user optimization as the behavioral description (when congestion externalities exist). Yet, in spite of these difficulties (i.e., the nonconvexities) some progress in the development of solution algorithms has been made.

LeBlanc [41] formulated the network design model as a mixed integer programming problem, the integer variables being the set of arcs considered for addition to the network. This version of the network design problem is one of the largest of all spatial combinational problems. In the terminology of combinational problems, this problem is NP-hard. The number of alternative solutions is 2^L, where L is the number of arcs. The feasible solution set which has to be search is enormous, and, despite much progress, the techniques for searching it are still quite crude.

Abdulaal and LeBlanc [2] formulated an alternative version of the network design problem in which continuous improvement variables are used instead of integer addition variables. Hence, their formulation is not a mixed programming problem, but it does include user equilibrium assignment constraints. They proposed two methods to solve the problem, one based on the work of Powell and the other using the Hook and Jeeves (H-J) algorithm. Morcotte [46], using the fact that the network assignment constraints can be formulated as a variational inequality problem, proposed the Constraint Accumulation Algorithm (CAA). However, he admits that this approach is still very difficult to apply even when few constraints are active. Friesz and Harker [34] solved the same problem using, what they called, the Iterative-Optimization Assignment Algorithm. Also, Marcotte [47] compared four heuristic algorithms on a small example problem of this type, and, most recently, Suwansirikul et al. [56] suggested an alternative heuristic called Equilibrium Decomposed Optimization (EOD). Many algorithms have been proposed for the bilevel programming problem, in which both levels of the problem are mathematical programs. Kolstad [39] suggests that these algorithms can be fit into following typology: (i) extreme point search methods, (ii) Kuhn-Tucker-Karush methods, and (iii) descent methods. Finally, and most importantly for our purposes, De Silva [26] presents a method which employs Fiacco and McCormick's [30] and Fiacco's [28] nonlinear programming sensitivity analysis results to calculate the implicit reaction functions for the second level in the bilevel problem. Chiou developed a series of solution methods such as gradient-based method, generalized bundle subgradient projection method, conjugate subgradient projection method to solve the continuous network design problems [12-14].

In regard to the network signal control problem (NSCP), the problem is to find the optimal signal setting which improves the performance of existing facilities in a transportation network. Conventional methods for optimizing signal settings can be divided into two types: stage-based and group-based approaches [4, 9, 11, 54, 58, 63]. The stage-based approach divided the signal cycle into separate stages and solved the optimal signal settings for each group of compatible traffic movements in stages. This approach is regarded as superior in the concern for safety and loss of capacity with phase switching [58]. The group-based approach considered each group of traffic streams having right-of-way in the time domain directly. Compared with the stage-based approach, the group-based approach has a higher degree of flexibility in signal timing arrangement [63, 64]. However, the most optimization models proposed so far usually converged to a local optimal solution and without taking traffic rerouting effects into account when solving NSCP [9]. The equilibrium network signal control problem (ENSCP) is used to find an optimal network signal design when the network flow pattern is constrained to be equilibrium. Friesz [33] points out that this is a problem of interest because of Braess' paradox [8]. This paradox shows that the congestion of the network may be severer when adding capacity to a congested network without taking the reaction of network users into consideration. Hence, in practice, the equilibrium network signal design problem must be solved by constraining the network flow pattern to meet user equilibrium. The user equilibrium network design with fixed transportation demand has been studied in both discrete [41] and continuous [2] versions. To help solve the signal control problem, Allsop [3] pointed out that the route choices of road users should be considered as the impacts of signal settings changing. Gartner et al. [36] and Fisk [31] described the signal control problem is a Stackelberg or leader-follower game between road users and the administration. The Stackelberg game can be represented as a bilevel problem where the upper level problem aims to find the optimal signal setting or link capacity expansions which maximizes system performance, and the lower level problem aims to solve the user equilibrium (UE) flows, respectively [5, 45].

1.2 Research Objectives and Scopes

This research is going to extend the applicability of these two methods respectively. For gradient-based method, we will introduce Kronecker product to calculate the second-order sensitivity information of network equilibrium flow. With the second-order sensitivity information, the response of the travelers with respect to the change in the traffic infrastructure or control policy can be approximated more precisely by a nonlinear reaction function. Based on the nonlinear reaction function, we can solve the network design problems and signal control problems with fewer iterations. For directional derivative-based method, we will relax the continuously differentiable assumption of arc cost functions in the network equilibrium problem by introducing piecewise linear arc cost functions instead. We will analyze the properties of the equilibrium solution in this generalized problem. Based on the original directional derivative-based method, an extended model will be proposed to solve the sensitivity information of the network equilibrium problem with piecewise linear arc cost functions.

To develop sensitivity results, we begin in the Section 2.1 with a brief review of the basic network equilibrium problem, and in particular, the formulation of that problem as a variational inequality. This is followed in Section 2.2 with a formulation of an associated perturbed version of the equilibrium problem in which the perturbation parameters are associated with the arc-flow cost functions and/or travel demands. Two major types of the sensitivity analysis of static network equilibrium problems are reviewed in Section 2.3 and 2.4, one is gradient-based method and the other is directional derivative-based method, respectively. Section 2.5 introduces applications of sensitivity analysis of network equilibrium problems.

There are some issues There are some issues about TFM has been proposedby Patriksson (2004) [51], Josefsson and Patriksson (2007) [52], Marcotte and Patriksson (2007) [48] and Yang and Bell (2007) [67]. Section 3 has been written to clarify the regularity conditions of the Tobin-Friesz method (TFM) for user equilibrium sensitivity analysis presented in Tobin and Friesz [61] and rederived in Cho et al. [16]. We will discuss and demonstrate some numerical examples appearing in the literature on user equilibrium sensitivity analysis about the issues in this section.

In Section 4, we will introduce Kronecker product to calculate the second-order sensitivity information of network equilibrium flow based on row reduction gradient-based method. For directional derivative-based method, we will relax the continuously differentiable assumption of arc cost functions in the network equilibrium problem by introducing piecewise linear arc cost functions instead in Section 5. Finally, we provide numerical examples of calculation of high-order sensitivity information and directional derivatives for the extensions of gradient-based method and directional derivative-based method, respectively, and make a brief conclusion about the computational results and research findings so far. The notation used throughout the study is listed in Table 1.

Table 1: The notation used in this study

N	=	the set of nodes of the network
$i, j \in N$	=	specific nodes in the network
A	=	the set of arcs of the network
$a \in A$	=	an arc in the network; $a = (i, j)$
W	=	the set of origin-destination pairs
$w \in W$	=	an origin destination pair; $w = (i, j)$
P_w	=	the set of paths between origin-destination pair w
$p \in P_w$	=	a path between origin-destination pair w
$\Delta = [\Delta_{ap}]$	=	the arc/path incidence matrix, where $\Delta_{ap} = 1$ if arc a is in path p, 0 otherwise
$\Lambda = [\Delta_{wp}]$	=	the origin-destination/path incidence matrix, where $\Lambda_{wp} = 1$ if path $p \in P_w$, 0 otherwise

T_w = the number of trips between origin destination pair *w*

$T = [T_w]$ = the vector of all trips

f_a = the flow on arc *p*

$h = [h_p]$ = the vector of all path-flows

f_a = the flow on arc *a*

$f = [f_a]$ = the vector of all arc-flows; note that $f = \Delta h$

$y = [y_a]$ = the vector of all capacity variables that are desired to be increased

$t_a(f)$ = the cost on arc *a* as a function of all path-flows

$t(f) = [t_a(f)]$ = the vector of arc cost functions

$c_p(h)$ = the cost on path *p* as a function of all path-flows

$c(h) = [c_p(h)]$ = the vector of path cost functions, note that $c(h) = \Delta^T c(f)$

CHAPTER 2

Literature Review

2.1 Formulations of Network Equilibrium Problem

Most sensitivity analysis methods are developed with the fundamental theory of for nonlinear programming problems or variational inequality problems. With different problem formulations, various sensitivity analysis methods can be applied to calculate the sensitivity information respectively. Therefore, the essential of the sensitivity analysis of network equilibrium problems is the formulations of the network equilibrium problems. In the following sections, we summarize some important results from Tobin and Friesz [61] and Cho et al. [16].

Consider a transportation network $G(N, A)$ with a finite set of nodes $i \in N$, and a finite set of links $a \in A$, together with a nonempty set of origin-destination (OD) pairs $w \in W$. Each $w \in W$ is joined by a nonempty finite set of paths, $p \in P_w$ and the set P is the union of path set P_w for all OD pairs w. Let real numbers, nonnegative reals, and positive reals are denoted respectively by R; R_+ and R_{++}, and the cardinalities of A, W and P are denoted respectively by $\alpha = |A|$, $\omega = |W|$, and $\rho = |P|$. Each positive column vector, $T = (T_w: w \in W) \in R_{++}^{\omega}$, is designated as a possible travel demand vector. Each nonnegative column vector, $h = (h_p: p \in P) \in R_{+}^{\rho}$, is designated as a path-flow vector. Each nonnegative column vector, $f = (f_a: a \in A) \in R_{+}^{\alpha}$, is designated as a arc-flow vector. The relationship between arc-flows, path-flows, and travel demand are given by

$$f = \Delta h, \tag{2.1}$$

$$T = \Lambda h, \tag{2.2}$$

where Δ is an $\alpha \times \rho$ matrix, with $\Delta_{ap} = 1$ if arc a belongs to path p and $\Delta_{ap} = 0$ otherwise; Λ is an $\omega \times \rho$ matrix, with $\Lambda_{wp} = 1$ if OD-pair w belongs to path p and $\Lambda_{wp} = 0$ otherwise. Generally, Δ and Λ are the link/path and OD/path matrices associated with equilibrium paths respectively.

Let $t: R_+^\alpha \to R_+^\alpha$ be an arc cost function and $t(f)$ = ($t_a(f)$: $a \in A$) be the vector of arc cost functions on each arc, $a \in A$, for a given arc-flow f. Hence, the path cost of each path, $p \in P$ is given by

$$c_p(h) = \Delta^T t(f). \tag{2.3}$$

For each $f \in R_+^a$ and $w \in W$, the minimum path cost is denoted by

$$\mu_w = \min\{c_p(h) : p \in P_w\}. \tag{2.4}$$

Based on the preliminary above, Wardrop stated two principles that tend to model the nature of traffic behavior [62]. The first principle states the nature from a user's point of view, and the second principle describes the desirable behavior from the system designer's standpoint.

Wardrop's first principle:

The travel time on all the routes actually used are equal, and less than those which would be experienced by a single vehicle on any unused route.

Wardrop's second principle:

The average journey time is a minimum. This amounts to minimizing $t(f)^T f$ over feasible aggregate flows f.

Then, we may now make the following definition:

Definition 2.1 (User equilibrium)

A flow pattern (f, μ) satisfying the following conditions is a user equilibrium:

$$h_p[c_p(h) - \mu_w] = 0, \forall p \in P_w, \tag{2.5}$$

$$c_p(h) - \mu_w \geq 0, \forall p \in P_w, \tag{2.6}$$

$$\sum_{p \in P_w} h_p - T_w = 0, \forall w, \tag{2.7}$$

$$f = \Delta h, \tag{2.8}$$

$$h \geq 0, \tag{2.9}$$

$$\mu \geq 0. \tag{2.10}$$

Expressions (2.5) and (2.6) are recognized as equivalent to Wardrop's first principle [62]. They require that for utilized paths between a given OD pair, path cost equals the minimum path cost; paths whose costs exceed that minimum are not utilized. A user equilibrium flow pattern will ensure when no individual user has an incentive for deviating from the current chosen path. Expression (2.7) is a statement of flow conservation; (2.8) is definitional; (2.9) and (2.10) are nonnegativity conditions. Based on the definition of user equilibrium, the common formulations of network equilibrium problems are introduced as follows.

2.1.1 Nonlinear Complementarity Formulation

It is well known that the conditions defining user equilibrium can be formulated as a nonlinear complementarity problem (NCP).

A NCP is as follows: find x such that

$$F(x)x = 0, F(x) \geq 0, x \geq 0. \tag{2.11}$$

Aashitiani and Magnanti [1] showed that the user equilibrium conditions can be placed in this form when

$$x = (h, \mu), \tag{2.12}$$

$$F(x) = \left[c_p(h) - \mu_w, \forall p \in P_w \; ; \; \sum_{p \in P_w} h_p - T_w = 0, \forall w \right], \tag{2.13}$$

provided arc costs are positive.

By transforming the NCP to a fixed-point problem and applying Brouwer's theorem, Aashitiani and Magnanti [1] were able to establish a quite general existence theorem for user equilibrium. They found that a user equilibrium will exist when the arc cost functions are positive and continuous, and the travel demands are nonnegative, continuous and bounded from above. They also employed results from complementarity theory to establish that a user equilibrium is unique if arc cost functions and negative demand functions are strictly monotone increasing.

2.1.2 Variational Inequality Formulation

It is also well known that the user equilibrium conditions can be formulated as a variational inequality problem (VIP).

A VIP is as follows: find $x^* \in X$ such that

$$F(x^*)(y - x^*) \geq 0, \forall y \in X . \tag{2.14}$$

It can be shown that x^* is a solution of NCP (2.11) if, and only if, it solves the VIP of finding $x^* \in R_+$ [38], such that

$$F(x^*)(y - x^*) \geq 0, \forall y \in R_+ . \tag{2.15}$$

Variational inequality formulations of network equilibrium derived from NCPs have been studied by Fisk and Boyce [32] and Pang [50]. It is well known that if $F(x)$ is continuous and X is compact and convex, Eq. (2.14) has a solution. It is also well known that if $F(x)$ is strictly monotone on X, any solution of Eq. (2.14) is unique. In particular, Dafermos [20] showed that the user equilibrium conditions are completely equivalent to the VIP of finding $(x^*, T^*) \in \Omega$, such that

$$t(f^*)(f - f^*) - \xi(T^*)(T - T^*) \geq 0, \forall (f, T) \in \Omega . \tag{2.16}$$

where $\xi(T)$ is the inverse demand function for OD pairs, and

$$\Omega = \left\{ (f, T) : \sum_{p \in P_w} h_p - T_w = 0, \forall w; f = \Delta h; h \geq 0; T \geq 0 \right\} . \tag{2.17}$$

Smith also derived a similar result for the fixed demand case [55]. Dafermos observed that such formulations have powerful theorem for VIPs may be employed to establish the qualitative properties if a user equilibrium [20].

2.1.3 Mathematical Programming Formulation

It is also well known that the VIP is a necessary condition for x^* to be a local minimum of the mathematical program:

$$\min \oint_x F(y)dy, \text{ s.t. } x \in X, \tag{2.18}$$

provided the Jacobian $\nabla F(x)$ is a symmetric matrix. The VIP (2.14) and the mathematical program (2.18) are completely equivalent and possess a unique solution when X is a convex set and $\nabla F(x)$ is symmetric and positive definite.

It is then immediate from Eq. (2.16) that a user equilibrium may be found when $\nabla c(f)$ and $\nabla\theta(T)$ are symmetric matrices, by solving

$$\min \sum_a \oint_f t_a(y)dy_a - \sum_i \sum_j \oint_T \xi_{ij}(z)dz_{ij}, \text{ s.t. } (f, T) \in \Omega, \tag{2.19}$$

which is ca convex mathematical program with an unique global minimum when $\nabla t(f)$ and $-\nabla\xi(T)$ are positive definite. The mathematical program (2.19) is essentially Beckmann's equivalent optimization problem for user equilibrium [6]. Actually Beckmann's original formulation dealt only with separable functions so that the symmetric restrictions necessary for writing down Eq. (2.19) are satisfied trivially. Originally Beckmann derived Eq. (2.19) by first postulating its validity and then showing that the associated Karush-Kuhn-Tucker (KKT) conditions are identical to the equilibrium conditions.

2.2 Perturbation System of Network Equilibrium Problem

To formulate this perturbation problem, suppose that the flow-cost function and travel demand vector are influenced by some finite-dimensional vector of perturbations, $\Theta \in R^k$. In particular, suppose that a given function, θ_0, and that it is meaningful to consider changes in c_0 and T_0 corresponding to parameter values in some neighborhood, Θ, of θ_0 in the parameter space R^k. Then, if for each $\theta \in \Theta$ we define the corresponding perturbation vector, $\varepsilon = \theta - \theta_0$, we may reparameterize these functions in term of the associated set of perturbation vectors, $D = \{\varepsilon \in R^k : \varepsilon + \theta_0 = \theta\}$. Of special interest is the zero perturbation vector, $0 \in D$, which corresponds to the initial (unperturbed) parameter vector, θ_0. In particular, we shall be primarily concerned with small perturbations in this initial vector θ_0, and hence assume for convenience that all sufficiently small perturbations are possible. To be more precise, if for each $x \in R^n$ and positive scalar, $\delta > 0$, we designate the set $B(x) = \{y \in R^n : \| x - y \| < \delta\}$, as an x-neighborhood in R^n, then we now assume that D is bounded (i.e., is contained some 0-neighborhood). Finally, to study the continuity properties of perturbations, it is convenient to assume that D is a closed set (and hence that D is compact in R^k). In summary then, we now say that:

Definition 2.1

Each compact set, $D \subseteq R^k$, containing a 0-neighborhood is designated as an admissible perturbation domain.

Given any perturbation domain, D, it is postulated that for each $\varepsilon \in D$ we may associate a unique arc cost function, $t(\cdot,\varepsilon)$, and positive travel demand vector $T(\varepsilon)$ (where by definition $t_0 = t(\cdot,0)$ and $T_0 = T(0)$). Hence if for each $\varepsilon \in D$ and $f \in R^{\alpha}$ we now let

$$H(f,\varepsilon) = \{h \in R_+^{\rho} : f = \Delta h \text{ and } \Lambda h = T(\varepsilon)\}, \tag{2.20}$$

and define the feasible arc-flow set corresponding to (2.20) by

$$\Omega(\varepsilon) = \{f \in R^{\alpha} : H(f,\varepsilon) \neq \phi\} \subseteq R_+^{\alpha}, \tag{2.21}$$

then the set if equilibrium arc-flow vectors for the equilibrium problem defined by $t(\cdot,\varepsilon)$ and $T(\varepsilon)$ is now given by the solution set for the associated VIP, i.e. by

$$VI[t(\cdot,\varepsilon), \Omega(\varepsilon)] = \{f \in \Omega(\varepsilon) : t(f,\varepsilon)^T (g - f) \geq 0, \forall g \in \Omega(\varepsilon)\}, \tag{2.22}$$

Our primary concern in the present study is with those perturbation problems for which these equilibrium arc-flow vectors are at least locally unique. Hence we now define a general class of perturbation systems with this property as follows. If the closure of a set $X \subseteq R^n$, is designed by $cl(X)$, then we now say that:

Definition 2.2

For any perturbation domain, D, continuous functions, $t : R_+^\alpha \times D \to R_+^\alpha$, $T : D \to R_{++}^\alpha$, and open set, $F \subseteq R^\alpha$, the ordered collection (D, F, T, t) is designated as a perturbation system iff the following local uniqueness condition is satisfied:

Condition 2.1 (Local uniqueness)

For all perturbation vectors, $\varepsilon \in D$,

$\left| VI[t(\cdot,\varepsilon), \Omega(\varepsilon)] \cap F \right| = 1 = \left| VI[t(\cdot,\varepsilon), \Omega(\varepsilon)] \cap cl(F) \right|$.

Condition 2.1 asserts that for each perturbation vector, $\varepsilon \in D$, the VIP in (2.22) has exactly one solution, $f(\varepsilon) \in \Omega(\varepsilon) \cap F$, and that there exist no other solutions in $\Omega(\varepsilon) \cap cl(F)$. If $VI[t(\cdot,\varepsilon), \Omega(\varepsilon)] - F \neq \phi$, then $f(\varepsilon)$ is locally unique with respect to F, and if $VI[t(\cdot,\varepsilon), \Omega(\varepsilon)] - F = \phi$, then $f(\varepsilon)$ is globally unique in $\Omega(\varepsilon)$. In all cases, the solution vectors, $f(\varepsilon)$, define a unique equilibrium arc-flow function, $f : D \to F$, with respect to the region $F \subseteq R^\alpha$. In these terms, our primary objective is to study the properties of such functions.

Theorem 2.1 (Continuity of arc-flows)

For each perturbation system, (D, F, T, t), the associated equilibrium arc-flow function, $f : D \to F$, is continuous.

Notice also that the continuity of perturbed equilibrium arc-flows depends only on the continuity of t and T. In particular, this continuity property is independent of any monotonicity properties of c, as employed for example in the continuity theorems of Fang

[27], Dafermos and Nagurney [23] and Dafermos [21].

But while this result may be said to provide a satisfactory conceptual framework for the analysis of small perturbations in equilibrium arc-flows, it fails to yield any operational procedures for doing so. Hence, our main objective is to impose stronger structural conditions on perturbation system, (D, F, T, t), which will yield a procedure for approximating equilibrium arc-flow function, $f : D \rightarrow F$ in some small neighborhood of the unperturbed state, $\varepsilon=0$.

Moreover, in order to preserve the uniqueness of the arc-flow equilibria, $f(\varepsilon)$, we require a stronger condition on arc cost functions $c(\cdot, \varepsilon)$. In particular, a function $\rho : R_{+}^{\alpha} \rightarrow R_{+}^{\alpha}$, will be said to be strictly monotone on $S \subseteq R_{+}^{\alpha}$ iff $[\rho(x) - \rho(x)]^T(x - y) > 0$ for the distinct $x, y \in S$, and we now require that (D, F, T, t) satisfy the following local strict monotonicity condition:

Condition 2.2 (Local strict monotonicity)

There exists some 0-neighborhood, $B_s \subseteq D$ such that $t(\cdot, \varepsilon)$ is strictly monotone on $\Omega(\varepsilon) \cap cl(F)$ for all $\varepsilon \in B_s$.

In the following gradient-based sensitivity analysis method, they all restrict the network to those arcs with positive flows and only consider path variables h_p which are positive in h^*. To guarantee positivity of $h(\varepsilon, h_0)$ for all ε sufficiently close to zero, it suffices to require that in the perturbed network equilibrium problem, there exists at least one equilibrium path-flow vector in which all minimum-cost paths are used. Hence, if we let

$$H_{+}(\varepsilon) = H(f(\varepsilon), \varepsilon) \cap R_{++}^{\rho_0},$$

denote the set of positive flow vector in $H(f(\varepsilon), \varepsilon)$, then the desired local positivity condition can be stated as follows:

Condition 2.3 (Local positivity)

$H_{+}(0) \neq \phi$.

Based on Condition 2.2 and 2.3, we can obtain the local existence condition of the perturbation system as following:

Definition 2.3

For any perturbation domain, D, continuous functions, $c: R_{+}^{\alpha} \times D \to R_{+}^{\alpha}$, $T: D \to R_{++}^{\alpha}$, and open set $F \subseteq R^{\alpha}$, the ordered collection (D, F, T, t) is designated as a locally regular perturbation system iff (D, F, T, t) satisfies Condition 2.3 and 2.4 together with the following local existence condition:

Condition 2.4 (Local existence)

For all perturbation vectors, $\varepsilon \in D$, $VI[t(\cdot, \varepsilon), \Omega(\varepsilon)] \cap F \neq \phi$.

Moreover, if we want to analyze the local sensitivity of equilibrium arc-flow functions, we may now define the relevant class of differentiable systems for our purposes as follows:

Definition 2.4

A locally regular perturbation system, (D, F, T, t), is said to be locally smooth iff there exists a 0-neighborhood, $B(0) \subseteq D$, and an $f(0)$-neighborhood, $F(0) \subseteq F(0) \cap R_{++}^{\varepsilon}$, such that the restricted functions, $t: F(0) \times B(0) \to R_{+}^{\alpha}$ and $T: B(0) \to R_{+}^{\omega}$, are continuously differentiable, and the following additional condition is satisfied:

Condition 2.5 (Local positive definiteness)

$\nabla_{f} t(f(0), 0)$ is positive definite.

Based on Definition 2.1~2.4, several gradient-based sensitivity analysis methods can be developed. The following section will summarize the major results of gradient-based sensitivity analysis methods.

2.3 Gradient-Based Sensitivity Analysis

The sensitivity analysis methods for nonlinear programming problems [19] and variational inequality problems [21, 40, 59] cannot evaluate the gradient-based sensitivity analysis formula for the traffic equilibrium problem directly due to the path-flow solution does not satisfy the local uniqueness condition. As a sequence, a restricted equilibrium problem is developed by Tobin and Friesz [61] and the gradient-based sensitivity analysis formula of deterministic traffic equilibrium can be applied. In Tobin and Friesz method (TFM) [61], they restricted the network to those with positive flow. According to the restricted network, assuming that a strict complementary slackness condition is satisfied and the extreme point is nondegenerate as follows:

Condition 2.6 (Strict complementary slackness)

If $h_p^*(0)=0$, then $\mu_w > 0$ for origin-destination pair $w \in W$.

Condition 2.7 (Nondegenerate extreme point)

There exists $h^*(0)$ that is a nondegenerate extreme point of $H(f^*,0)$ in the sense that $h^*(0)$ corresponds to a unique basis and the number of paths with positive flow is equal to the rank of $[\Delta^T, \Lambda^T]$ after restricting the network to arcs with positive flows.

Also, the arc cost functions are assumed to be continuously differentiable and strictly monotone increasing. Under these assumptions, the user equilibrium conditions can reduce to a perturbation system of equations when perturbation parameter is zero:

$$c_p\left(h^*,0\right)-\Lambda^T\mu=0\,, \tag{2.23}$$

$$\Lambda h^*-T(0)=0\,. \tag{2.24}$$

In this perturbation system, cost functions and OD demand are functions of perturbation parameter, ε. Moreover, the first equation in the system represents the equilibrium condition, and the second one represents the flow conservation. If taking the derivative of the system with respect to perturbation parameter and applying the implicit function theorem, one can calculate a unique arc-flow sensitivity that does not depend on which extreme point was chosen. The sensitivity formula of network equilibria at $\varepsilon = 0$ is

$$\begin{bmatrix}\nabla_\varepsilon h(0)\\ \nabla_\varepsilon \mu(0)\end{bmatrix}=\begin{bmatrix}B_{11} & B_{12}\\ B_{21} & B_{22}\end{bmatrix}\begin{bmatrix}-\nabla_\varepsilon c_p\left(h^*,0\right)\\ \nabla_\varepsilon T(0)\end{bmatrix}. \tag{2.25}$$

where

$$B_{11}=\nabla c_p\left(h^*,0\right)^{-1}\left[I-\Lambda^T\left[\Lambda\nabla c_p\left(h^*,0\right)^{-1}\Lambda^T\right]^{-1}\Lambda\nabla c_p\left(h^*,0\right)^{-1}\right],$$

$$B_{12}=\nabla c_p\left(h^*,0\right)^{-1}\Lambda^T\left[\Lambda\nabla c_p\left(h^*,0\right)^{-1}\Lambda^T\right]^{-1},$$

$$B_{21}=-\left[\Lambda\nabla c_p\left(h^*,0\right)^{-1}\Lambda^T\right]^{-1}\Lambda\nabla c_p\left(h^*,0\right)^{-1},$$

$$B_{22}=\left[\Lambda\nabla c_p\left(h^*,0\right)^{-1}\Lambda^T\right]^{-1}.$$

The derivatives of path-flows with respect to ε at $\varepsilon = 0$ are

$$\nabla_\varepsilon h=-B_{11}\nabla_\varepsilon c_p\left(h^*,0\right)+B_{12}\nabla_\varepsilon T(0). \tag{2.26}$$

And the derivatives of arc-flows with respect to ε at $\varepsilon = 0$ for the restricted problem can be written as

$$\nabla_\varepsilon f(0)=\Delta\nabla_\varepsilon h(0). \tag{2.27}$$

Since Tobin and Friesz proposed their method, it has been the most popular tool for producing sensitivity information in network equilibrium problems. However, the assumptions of their method are too strong to find a nondegenerate extreme point in the large scale network topology. Consequently, two reduction methods are proposed to resolve this restriction.

2.3.1 Row Reduction Gradient-Based Method

In order to overcome the restriction of Tobin and Friesz method, Cho et al. [16] proposed a row reduction gradient-based method to calculate the sensitivity information of equilibrium network flows in the large scale network topology. The row reduction method assumed the arc cost functions are strictly monotone and the Jacobian matrix is positive definite. Also assuming that there exists a set of strictly positive flows on all the equilibrium paths when perturbation parameter is zero, then the solution of the perturbation system of equations exist and satisfy the local uniqueness condition.

In the row reduction method, a maximum number of linearly independent equations from the system (2.1) and (2.2) is chosen to express dependent arc-flows in terms of independent arc-flows and OD demand vectors. In other words, they select a maximal set of rows from Δ, says Δ_1, for which the combined matrix $[\Delta_1; \Lambda]$ is of full row rank. Hence, we can partition Δ as

$$\Delta = \begin{bmatrix} \Delta_1 \\ \Delta_2 \end{bmatrix}. \tag{2.28}$$

Therefore, there must exist matrices M_1 and M_2 such that

$$\Delta_2 = M_1 \Delta_1 + M_2 \Lambda . \tag{2.29}$$

Moreover, one can induce that the dependence arc-flow vector, f_2, can be expressed as

$$f_2 = \Delta_2 h = M_1 \Delta_1 h + M_2 \Lambda h = M_1 f_1 + M_2 T(\varepsilon). \tag{2.30}$$

Based on Δ_1, Λ and independence arc-flow vector, f_1, the arc-based reduction method employed a “minimum-distance” technique to select a unique equilibrium path-flow vector for each equilibrium arc-flow vector as follows:

$$h(\varepsilon, f_1, h_0) = h_0 + \begin{bmatrix} \Delta_1 \\ \Lambda \end{bmatrix}^T \begin{bmatrix} \Delta_1 \Delta_1{}^T & \Delta_1 \Lambda^T \\ \Lambda \Delta_1{}^T & \Lambda \Lambda^T \end{bmatrix}^{-1} \left(\begin{bmatrix} f_1 \\ T(\varepsilon) \end{bmatrix} - \begin{bmatrix} \Delta_1 \\ \Lambda \end{bmatrix} h_0 \right). \tag{2.31}$$

Moreover, let

$$\begin{bmatrix} \Delta_1 \Delta_1{}^T & \Delta_1 \Lambda^T \\ \Lambda \Delta_1{}^T & \Lambda \Lambda^T \end{bmatrix}^{-1} = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}, \tag{2.32}$$

Eq.(2.3-28) can be rewritten in compact form as:

$$h(\varepsilon, f_1, h_0) = N_0 h_0 + N_1 f_1 + N_2 T(\varepsilon), \tag{2.33}$$

where

$$N_0 = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} - \begin{bmatrix} \Delta_1 \\ \Lambda \end{bmatrix}^T \begin{bmatrix} \Delta_1 \Delta_1{}^T & \Delta_1 \Lambda^T \\ \Lambda \Delta_1{}^T & \Lambda \Lambda^T \end{bmatrix}^{-1} \begin{bmatrix} \Delta_1 \\ \Lambda \end{bmatrix}, \tag{2.34}$$

$$N_1 = \Delta_1{}^T M_{11} + \Lambda^T M_{21}, \tag{2.35}$$

$$N_2 = \Delta_1{}^T M_{12} + \Lambda^T M_{22}. \tag{2.36}$$

Moreover, from Eq. (2.29) and $\begin{bmatrix} \Delta_1 \\ \Lambda \end{bmatrix}$ is full row rank, one can induce that

$$\begin{bmatrix} M_1 & M_2 \end{bmatrix} = \Delta_2 \begin{bmatrix} \Delta_1 \\ \Lambda \end{bmatrix}^T \left(\begin{bmatrix} \Delta_1 \\ \Lambda \end{bmatrix} \begin{bmatrix} \Delta_1 \\ \Lambda \end{bmatrix}^T \right)^{-1} = \Delta_2 \begin{bmatrix} \Delta_1 \\ \Lambda \end{bmatrix}^T \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}. \tag{2.37}$$

By Eq. (2.35) and (2.36), we can conclude that the flow-conservation conditions in Eq. (2.30) can be written as

$$\Delta_2 N_1 f_1 - f_2 + \Delta_2 N_2 T(\varepsilon) = 0, \tag{2.38}$$

$$\Rightarrow Mf + \Delta_2 N_2 T(\varepsilon) = 0, \tag{2.39}$$

where

$$M = \begin{bmatrix} \Delta_2 N_1 & -I \end{bmatrix}. \tag{2.40}$$

According to the perturbed network equilibrium problem, we can obtain the following system equations:

$$G(f,\mu,\varepsilon)=\begin{bmatrix} t(f,\varepsilon)-M^T\mu \\ Mf+\Delta_2^0 N_2 T(\varepsilon) \end{bmatrix}. \tag{2.41}$$

The above system equations are continuously differentiable on all feasible solution sets. After differentiating Eq. (2.41) and applying the implicit function theorem, the sensitivity formula of row reduction method can be obtained and the invertibility of the inverse matrix is guaranteed exists due to the matrix [Δ_1; Λ] is of full row rank.

$$\nabla_{(f,\mu)}G(f_0,\mu_0,\varepsilon_0)=\begin{bmatrix} \nabla_f t(f_0,\varepsilon_0) & -M^T \\ M & 0 \end{bmatrix}, \tag{2.42}$$

$$\nabla_\varepsilon G(f_0,\mu_0,\varepsilon_0)=\begin{bmatrix} \nabla_\varepsilon t(f_0,\varepsilon_0) \\ \Delta_2 N_2 \nabla_\varepsilon T(\varepsilon_0) \end{bmatrix}, \tag{2.43}$$

$$\begin{bmatrix} \nabla_f t(f_0,\varepsilon_0) & -M^T \\ M & 0 \end{bmatrix}\begin{bmatrix} \nabla_\varepsilon f(\varepsilon_0) \\ \nabla_\varepsilon \mu(\varepsilon_0) \end{bmatrix}=\begin{bmatrix} -\nabla_\varepsilon t(f_0,\varepsilon_0) \\ \Delta_2 N_2 T_\varepsilon(\varepsilon_0) \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} \nabla_\varepsilon f(\varepsilon_0) \\ \nabla_\varepsilon \mu(\varepsilon_0) \end{bmatrix}\Bigg|_{\varepsilon_0=0}=\begin{bmatrix} \nabla_f t(f_0,\varepsilon_0) & -M^T \\ M & 0 \end{bmatrix}^{-1}\begin{bmatrix} -\nabla_\varepsilon t(f_0,\varepsilon_0) \\ \Delta_2 N_2 T_\varepsilon(\varepsilon_0) \end{bmatrix}\Bigg|_{\varepsilon_0=0}. \tag{2.44}$$

2.3.2 Column Reduction Gradient-Based Method

The other gradient-based sensitivity analysis approach, column reduction method [67], was proposed to deal with the problem in the large scale network topology also. In the column reduction method, the link cost function $t(f, \varepsilon)$ is assumed to be positive and strictly monotone in f for $f \geq 0$, and once continuously differentiable in (f, ε); the travel demand function is assumed to be once continuously differentiable in ε. Furthermore, it is assumed that there exists a set of strictly positive flows on all the equilibrated paths when the perturbation parameter is zero.

To overcome the non-uniqueness of path-flows in the equilibrium network flow problem, column reduction approach choose a maximal set of equilibrated and linearly independent (ELI) paths or columns in $[\Delta; \Lambda]^T$. Denote the set of ELI paths as $\tilde{R}$ and the corresponding path-flow variables as $\tilde{h}$, and the further reduced arc/path and OD/path incidence matrices as $\tilde{\Delta}$ and $\tilde{\Lambda}$. Eq. (2.1) and (2.2) can be rewritten as following:

$$\begin{bmatrix} \tilde{\Delta} & \tilde{\Delta}^c \\ \tilde{\Lambda} & \tilde{\Lambda}^c \end{bmatrix} \begin{bmatrix} \tilde{h} \\ \tilde{h}^c \end{bmatrix} = \begin{bmatrix} f^*(0) \\ T(0) \end{bmatrix}. \tag{2.45}$$

where "c" denotes the corresponding "complementary" matrices and vector to the ELI paths.

For sufficiently small ε near zero, one can always fix the complementary or non-basic path-flow variables as $\tilde{f}^c = \tilde{f}^{c+}$ and solve the following linear system of equations for $\tilde{f}$ for any ε near zero:

$$\begin{bmatrix} \tilde{\Delta} \\ \tilde{\Lambda} \end{bmatrix} \tilde{h}(\varepsilon) = \begin{bmatrix} f^*(\varepsilon) \\ T(\varepsilon) \end{bmatrix} - \begin{bmatrix} \tilde{\Delta}^c \\ \tilde{\Lambda}^c \end{bmatrix} \tilde{h}^{c+}. \tag{2.46}$$

In this method, it is sufficient to consider the ELI working paths only. Hence, the equilibrium system, Eq. (2.23) and (2.24), can be reduced to:

$$\tilde{c}(h^*,0) - \tilde{\Lambda}^T \mu = 0, \tag{2.47}$$

$$\tilde{\Lambda}\tilde{h}^* - T(0) = 0, \tag{2.48}$$

where $\tilde{c}(h^*,0)$ represents the corresponding reduced cost vector. Based on the equilibrium system in Eq. (2.47) and (2.48), differentiating both sides of the system with respect to perturbations ε yields:

$$\begin{bmatrix} \nabla_{\tilde{h}} \tilde{c}(h^*,0) & -\tilde{\Lambda}^T \\ \tilde{\Lambda} & 0 \end{bmatrix} \begin{bmatrix} \nabla_\varepsilon \tilde{h} \\ \nabla_\varepsilon \mu \end{bmatrix} = \begin{bmatrix} -\nabla_\varepsilon \tilde{c}(h^*,0) \\ \nabla_\varepsilon T(0) \end{bmatrix}, \tag{2.49}$$

where the Jacobian

$$J_{\tilde{\mathbf{h}},\mu} = \begin{bmatrix} \nabla_{\tilde{h}} \tilde{c}(h^*,0) & -\tilde{\Lambda}^T \\ \tilde{\Lambda} & 0 \end{bmatrix}, \tag{2.50}$$

is well defined and

$$\nabla_{\tilde{\mathbf{h}}} \tilde{c}(h^*,0) = \tilde{\Delta}^T \nabla_{\mathbf{f}} t(f^*,0) \tilde{\Delta}. \tag{2.51}$$

Applying the general implicit function theorem to Eq. (2.49), the sensitivity formula of column reduction method can be derived:

$$\begin{bmatrix} \nabla_{\varepsilon}\tilde{h} \\ \nabla_{\varepsilon}\mu \end{bmatrix} = \begin{bmatrix} \nabla_{\tilde{h}}\tilde{c}(h^*,0) & -\tilde{\Lambda}^T \\ \tilde{\Lambda} & 0 \end{bmatrix}^{-1} \begin{bmatrix} -\nabla_{\varepsilon}\tilde{c}(h^*,0) \\ \nabla_{\varepsilon}T(0) \end{bmatrix}. \tag{2.52}$$

Let

$$\begin{bmatrix} \nabla_{\tilde{h}}\tilde{c}(h^*,0) & -\tilde{\Lambda}^T \\ \tilde{\Lambda} & 0 \end{bmatrix}^{-1} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}. \tag{2.53}$$

Then the derivatives of the ELI working path-flows with respect to ε at ε=0 are:

$$\nabla_{\varepsilon}\tilde{h} = -B_{11}\nabla_{\varepsilon}\tilde{c}(h^*,0) + B_{12}\nabla_{\varepsilon}T(0). \tag{2.54}$$

In view of $\nabla_{\varepsilon} f = -\tilde{\Delta}B_{11}\tilde{\Delta}^T\nabla_{\varepsilon}t(f^*,0) + \tilde{\Delta}B_{12}\nabla_{\varepsilon}T(0)$ and $\nabla_{\varepsilon}\tilde{h}^c = 0$ as well as $\nabla_{\varepsilon}\tilde{c}(h^*,0) = \tilde{\Delta}^T\nabla_{\varepsilon}t(f^*,0)$, the derivative of arc-flows with respect to ε at ε=0 are eventually obtained as:

$$\nabla_{\varepsilon} f = -\tilde{\Delta}B_{11}\tilde{\Delta}^T\nabla_{\varepsilon}t(f^*,0) + \tilde{\Delta}B_{12}\nabla_{\varepsilon}T(0). \tag{2.55}$$

We also have the derivatives of the equilibrium OD travel cost μ with respect to ε at ε=0 given as:

$$\nabla_{\varepsilon}\mu = -B_{21}\tilde{\Delta}\nabla_{\varepsilon}t(f^*,0) + B_{22}\nabla_{\varepsilon}T(0). \tag{2.56}$$

2.4 Directional Derivative-Based Sensitivity Analysis

The other researches focused on calculating the directional derivatives of the equilibrium arc-flow with respect to perturbation parameters. The papers by Qiu and Magnanti [53], Yen [70], Outrata [49], Patriksson and Rockafellar [52], Patriksson [51], Josefsson and Patriksson [37], and Lu [42] proposed the theoretical development of sensitivity analysis of traffic equilibria and provided sufficient conditions for the existence of directional derivatives.

Qiu and Magnanti [53] proposed an approach for conducting sensitivity analysis of variational inequalities defined on polyhedral sets. In order to overcome the non-uniqueness issue of the path-flow pattern, the conditions imposed in their method do not imply the local uniqueness of the perturbed solution. They generalized the usual definition of differentiability to a point-to-set mapping. Hence, their method can be used to derive sensitivity properties for a number of equilibrium problems including traffic equilibrium problems.

Let the cost function $t(\cdot, \cdot)$ and the set $\Omega^{\perp}$ satisfy assumptions of continuity, convergence, differentiability and positive definite; the derivative in the direction ε_0 solves the following variational inequality:

$VI^{\perp}$: find $f \in \Omega^{\perp}$ satisfying

$$\left[\nabla_f t\left(f^*, \varepsilon^*\right) f + \nabla_\varepsilon t\left(f^*, \varepsilon^*\right)\varepsilon_0\right]^T (f' - f) \geq 0 \quad \text{for any } f' \in \Omega^{\perp}, \tag{2.57}$$

where

$$\Omega^{\perp} = \left\{ f \,\middle|\, f = \Delta h,\ \Lambda h = 0,\ h_p \text{ } UIS \text{ for } p \in J_1, h_p \geq 0 \text{ for } p \in J_2, h_p = 0 \text{ for } p \in J_3, \right\}, \tag{2.58}$$

and

$$J_1 = \left\{ p \,\middle|\, h_p^* \geq 0,\ \sum_{a \in A} \Delta_{ap} c_a\left(f^*, \varepsilon^*\right) = \mu_w^*,\ p \in P_w,\ w \in W \right\}, \tag{2.59}$$

$$J_2 = \left\{ p \,\middle|\, h_p^* = 0,\ \sum_{a \in A} \Delta_{ap} c_a\left(f^*, \varepsilon^*\right) = \mu_w^*,\ p \in P_w,\ w \in W \right\}, \tag{2.60}$$

$$J_3 = \left\{ p \,\middle|\, h_p^* = 0,\ \sum_{a \in A} \Delta_{ap} c_a\left(f^*, \varepsilon^*\right) > \mu_w^*,\ p \in P_w,\ w \in W \right\}. \tag{2.61}$$

Yen [70] considered a network with locally Lipschitz, locally strongly monotone arc cost functions and studied the sensitivity of solutions to a parametric variational inequality with a parametric polyhedral constraint which depends on a pair of perturbation parameters. It is shown that the equilibrium pattern is locally unique and is a locally Lipschitz function with respect to the perturbation of arc cost functions and travel demand.

Outrata [49] considered the traffic equilibrium problem described by the generalized equation and assumed arc cost functions are strongly monotone on the feasible set and with a strictly copositive partial Jacobian matrix. Then the solution mapping of the equilibrium traffic pattern is directional differentiable with respect to the perturbation parameter.

Patriksson and Rockafellar [52] considered variational inequality problems over polyhedral sets and assumed the arc cost functions to be monotone and the negative of the travel demand functions to be strictly monotone. If the Jacobian matrices of the arc cost functions and the negative of the travel demand functions are positive definite, then the traffic equilibria is single-valued, Lipschitz continuous, positively homogeneous and piecewise linear with respect to the perturbations of the arc cost function and the travel demand function. The sensitivity analysis of deterministic and elastic demand can be obtained by solving linearized traffic equilibrium problems.

Patriksson [51] gave a complete study of the uniqueness, continuity and directional differentiability for the sensitivity of the deterministic and the elastic demand traffic equilibrium models. The variational inequality form of the sensitivity problem coincides with the first-order optimality conditions for a similar traffic equilibrium problem. Hence, the directional derivatives can be calculated by a traffic equilibrium solver.

Patriksson and Rockafellar [52] studied the local uniqueness, Lipschitz continuity, and semidifferentiability of the elastic-demand traffic user equilibria under perturbations of the arc cost function and the travel demand function. By assuming the negative of the travel demand function to be strictly monotone in the OD cost, they formulated the problem as a VI in the (x, d) space. Other assumptions in that paper included the requirement that the arc cost function and the negative of the inverse demand function be locally strongly monotone on the affine hull of the critical cone. Recently, Josefsson and Patriksson [37] specialized the analysis to the case of separable arc cost and demand functions, with the latter also being invertible.

Josefsson and Patriksson [37] based on the results of Patriksson [51] and studied the sensitivity analysis of the traffic equilibrium model with separable arc cost and demand functions. They formulated the sensitivity analysis problem into a linear complementary problem equivalent to a convex quadratic program which can be calculated by a state-of-the-art traffic equilibrium solver efficiently. Their method is introduced in the following.

For a variational inequality problem:

$$f\left(\varepsilon, x^*\right)^T \left(x - x^*\right) \geq 0, x \in X,$$

It can formulated as a more natural form as follows:

$$-f\left(\varepsilon, x^*\right) \in N_X\left(x^*\right)$$

where N_X denotes the normal cone to X at x:

$$N_X(x) = \begin{cases} \left\{v \in R^n \middle| v^T (y - x) \geq 0, y \in X\right\} & \text{if} \in X \\ \phi & \text{otherwise} \end{cases}.$$

Let S denote the solution mapping from $\varepsilon \in R^k$ to $x \in R^n$ that

$$S(\varepsilon) = \left\{x^* \in X \middle| -f\left(\varepsilon, x^*\right) \in N_X\left(x^*\right)\right\}, \quad \varepsilon \in R^k.$$

The directional derivative of the solution set $S(\varepsilon)$ is denoted the as $DS(\varepsilon^*|x^*)(\varepsilon')$ in the following:

$$DS\left(\varepsilon^* \mid x^*\right)(\varepsilon') = \left\{x' \in K \mid r(\varepsilon', x')^T (x - x') \geq 0, x \in K\right\}$$

where

$$K = T_x\left(x^*\right) \cap f\left(\varepsilon^*, x^*\right)^{\perp},$$

$$r(\varepsilon', x') = \nabla_\varepsilon f\left(\varepsilon^*, x^*\right)\varepsilon' + \nabla_x f\left(\varepsilon^*, x^*\right)x'.$$

T_X denotes the tangent cone to X and $X = \left\{x \in R^n \mid Ax \geq b; Bx = d\right\}$.

When the formulation is applied to the sensitivity analysis of traffic equilibria, let

$$x = \begin{pmatrix} h \\ f \\ d \end{pmatrix}, \ f(\varepsilon, x) = \begin{pmatrix} 0^\rho \\ t(\varepsilon, f) \\ -\xi(\varepsilon, d) \end{pmatrix}.$$

Accordingly,

$$K = \left\{(h', f', d') \in R^\rho \times R^\alpha \times R^\omega \mid \Gamma^T h' = d'; \Lambda^T h' = f'; h' \in H'\right\},$$

where

$$H' = \left\{ h' \in R^{\rho} \middle| \begin{array}{c} h'_p \text{ is free if } h^*_p > 0 \\ h'_p \le 0 \text{ if } h^*_p = 0 \text{ and } c_p\left(\varepsilon^*, h^*\right) = \mu^*_w \\ h'_p = 0 \text{ if } h^*_p = 0 \text{ and } c_p\left(\varepsilon^*, h^*\right) > \mu^*_w \\ p \in P_w, w \in W \end{array} \right\},$$

and

$$r(\varepsilon', x') = \begin{pmatrix} 0^{\rho} \\ \nabla_{\varepsilon} t\left(\varepsilon^*, v^*\right)\varepsilon' + \nabla_f t\left(\varepsilon^*, f^*\right) f' \\ -\left[\nabla_{\varepsilon} \xi\left(\varepsilon^*, d^*\right)\varepsilon' + \nabla_d \xi\left(\varepsilon^*, d^*\right) d'\right] \end{pmatrix}.$$

Based on the VI form of the equilibrium network flow problem, the directional derivatives of equilibrium network flows can be calculated by solving a quadratic optimization problem [51] as

$$\underset{f'}{\text{Minimize}} \quad \phi'(f') = \left[\nabla_{\varepsilon} t\left(f^*, \varepsilon^*\right)\varepsilon'\right]^T f' + \frac{1}{2}\sum_{a \in A} \frac{\partial t_a\left(f^*_a, \varepsilon^*\right)}{\partial f_a}\left(f'_a\right)^2 \tag{2.62}$$

$$\text{subject to} \quad \Lambda h' = d', \; \Delta h' = f' \tag{2.63}$$

$$h' \in H', \tag{2.64}$$

where

$$H' = \left\{ h' \in R^{\rho} \middle| \begin{array}{l} h'_p \text{ is free if } h^*_p > 0 \\ h'_p \ge 0 \text{ if } h^*_p = 0 \text{ and } c_p\left(h^*, \varepsilon^*\right) = \mu^*_w \\ h'_p = 0 \text{ if } h^*_p = 0 \text{ and } c_p\left(h^*, \varepsilon^*\right) > \mu^*_w \\ p \in P_w, \; w \in W \end{array} \right\}. \tag{2.65}$$

In this model, directional derivatives of link flow, path-flow, and travel demand are denoted by *f*', *h*', and *d*', respectively. The directional derivative can be interpreted as the direction and the rate of change of the equilibrium solution when perturbation parameter is perturbed along the direction ε'. The set *H*' is the set of directional derivatives of path-flow which keep the feasibility and optimality of the original problem in the first order approximation [37, 52].

Lu [42] applied some recently developed sensitivity analysis techniques for generalized equations to analyze the behavior of the equilibrium arc-flow of such a problem when both the arc cost function and the travel demand vary. The semiderivatives can be calculated by solving a linear traffic user equilibrium problem and the derivatives by matrix multiplication together with the solution of a linear equation the dimension of which is at most the number of arcs.

2.5 Discussion

The gradient-based method applied the classical implicit function theorem to the sensitivity analysis of equilibrium network flow problems. Performing the sensitivity analysis either by the row reduction method or the column reduction method, the arc-flow derivatives are identical. In the row reduction method, the arc-flow derivatives are involved with both independent arcs and dependent ones. However, in the original column reduction method, the arc-flow derivatives are only involved with independent paths which did not represent the entirely user behavior when perturbing the system. Compared with the gradient-based method, the directional derivative-based method is developed by the theory of VI and GE. Moreover, directional derivative-based method can deal with more general problem that the equilibrium solution is non-differentiable with respect to perturbation parameters.

CHAPTER 3

The Issues of Nondegeneracy of Sensitivity Analysis

This chapter has been written to clarify the regularity conditions governing application of the Tobin-Friesz method (TFM) for user equilibrium sensitivity analysis presented in Tobin and Friesz [61] and rederived in Cho et al. [16]. We have found that certain statements and numerical examples found in Patriksson [51], Josefsson and Patriksson [52] and Marcotte and Patriksson [48], if taken out of context, leave the impression that the TFM for user equilibrium sensitivity analysis is somehow "wrong" when, in fact, it works quite well provided its application is limited to those problems fulfilling the regularity conditions reviewed in this chapter. The following contents are taken from our paper [17].

3.1 The principal issues

Although the TFM for sensitivity analyses, which was re-derived using alternative arguments by Cho et al. (2000) [16], has been widely used, it has been criticized by Patriksson (2004) [51], Josefsson and Patriksson (2007) [52], Marcotte and Patriksson (2007) [48] and Yang and Bell (2007) [67] on the basis of the following:

Issue 1

To apply the TFM, one must begin with an unperturbed solution that is a nondegenerate extreme point. This regularity condition represents a critical assumption. Example problems have been published that show the TFM may fail when the nondegeneracy assumption is relaxed. However, it is possible to modify the degenerate solutions employed in example 7.3.2 of Josefsson and Patriksson [52], as shown in Section 3.2.2, to create valid initial solutions that are nondegenerate. However, our remarks should not be misconstrued as a claim that such modifications will always be possible, for they will not. However, in the event it is possible to construct a nondegenerate extreme point solution from a degenerate solution, the TFM works.

Moreover, the rederived method by Cho et al. [16] can sometimes be applied to cases where an unperturbed solution is a nondegenerate extreme point as shown in Section 3.2.6 for examples 7.3.2 and 7.3.3 of Josefsson and Patriksson [52].

Issue 2

From Tobin and Friesz [61], it is clear that existence and invertibility of the Jacobian matrix of the path cost vector, a submatrix of the entire Jacobian matrix formed from the Kuhn-Tucker conditions (including the complementary slackness conditions), are crucial to the validity of the TFM. Differentiability of the relevant functions assures existence of that Jacobian, and the stipulation of differentiability as a regularity condition is not at issue. However, Bell and Iida [7] correctly observed that the Jacobian of the path cost vector is not invertible when the number of paths is larger than the number of arcs, as example 6 of Yang and Bell [67] illustrates. In Section 3.3, we show that sometimes it is possible for the whole Jacobian matrix to be invertible even though the aforementioned submatrix (the Jacobian matrix of the path cost vector) is not invertible. Moreover, in assessing the TFM, one should not forget that the alternative derivation of identical sensitivity analysis formulae in Cho et al. [16] was performed to overcome the potential noninvertibility of the Jacobian matrix of the path cost vector, provided the appropriate derivatives needed to express the Jacobian may be calculated.

Issue 3

Violation of traditional strict complementarity may occur. Because one of the regularity conditions stipulated by Tobin and Friesz [61] is strict complementarity, it is reasonable to expect that the violation of some form of strict complementarity would generally prevent sensitivity analysis based on the TFM. However, in this chapter, we show by numerical example that sensitivity analysis may sometimes be performed using the TFM even if strict complementarity is violated, provided there is differentiability at the unperturbed user equilibrium solution. Example 5 in Patriksson [51] is an instance wherein both strict complementarity and differentiability fail to hold. In fact, sensitivity analysis of such a non-differentiable problem cannot be conducted using the TFM. However, when we consider example 7.3.1 of Josefsson and Patriksson [52], which does not fulfill strict complementarity but meets the differentiability standard, we find the TFM may be used successfully, as we illustrate in Section 3.2.2.

In the following sections, we discuss some numerical examples appearing in the literature on user equilibrium sensitivity analysis about the issues mentioned above.

3.2 Counterexamples in Patriksson (2004), Josefsson and Patriksson (2007) and Marcotte and Patriksson (2007)

In Patriksson [51], there is an illustrative example explaining that the TFM can provide an inaccurate result. Also, Josefsson and Patriksson [52] proposed three counterexamples which depicted some pitfalls of the TFM and which were re-used in Marcotte and Patriksson [48]. However, in the aforementioned examples, the authors either ignored the requirements of the TFM or applied the TFM to non-differentiable examples which violated the method's basic assumptions. As such the examples are not bonafide counter examples. In the discussion below, we will carry out TFM calculations for some of the aforementioned examples while enforcing the regularity conditions of TFM.

3.2.1 Example 5 in Patriksson (2004)

In Patriksson [51], a 5 node, 7 arc network with 2 origin-destination (OD) pairs and 6 paths is depicted and repeated here as Fig. 1. There are two fixed travel demands for OD pairs (1, 4), and (3, 5). There are three paths corresponding to each OD pair: $p_1 = \{1, 3\}$, $p_2 = \{1, 7, 4\}$ and

$p_3 = \{2, 4\}$ for OD pair (1,4) and $p_4 = \{5, 2\}$, $p_5 = \{5, 1, 7\}$ and $p_6 = \{6, 7\}$ for OD pair (3, 5). The arc cost functions are

$t_1(f_1) = 10 f_1$

$t_2(f_2) = 0.5 f_2$

$t_3(f_3) = 3 + 10 f_3$

$t_4(f_4) = 1 + 10 f_4$

$t_5(f_5) = f_5$

$t_6(f_6) = 2 + f_6$

$t_7(f_7) = f_7$

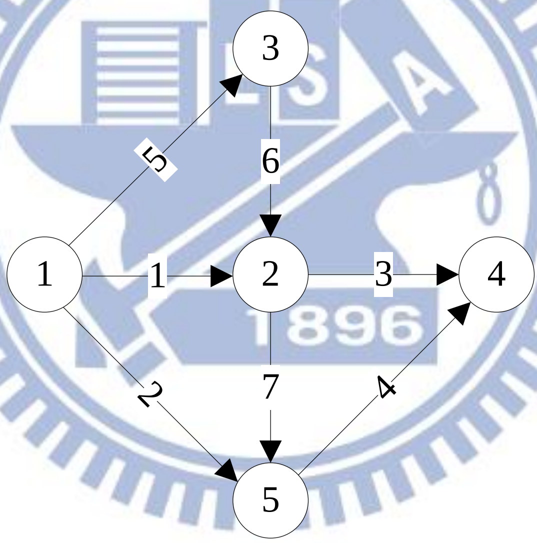

Fig. 1: The network of Patriksson's example 5

The travel demands are subject to perturbations expressed as the following vector: $\varepsilon = [\varepsilon_{14} \quad \varepsilon_{35}]^T$. When $\varepsilon = 0$, the equilibrium arc flow solution is $f^*(0) = [0,2,0,1,1,0,0]^T$. However, as mentioned in Patriksson [51], the solution does not obey strict complementarity. Also, the equilibrium solution is not differentiable. Therefore, this example problem is not one for which the TFM is applicable. That is, the TFM was never intended to apply to such a problem, and the example is not a counterexample.

3.2.2 Example 7.3.1 in Josefsson and Patriksson (2007)

This example, proposed by Josefsson and Patriksson [52], considers the 4 node, 5 arc network with 2 OD pairs and 4 paths depicted in Fig. 2. There are fixed demands of 2 and 1 units of flow for OD pairs (1, 2) and (4, 2), respectively. There are four paths corresponding to the two OD pairs: $p_1 = \{1\}$, $p_2 = \{2, 3\}$, $p_3 = \{4\}$ and $p_4 = \{3, 5\}$. The arc cost functions are

$$t_1(f_1, \varepsilon) = 2f_1 + \varepsilon$$

$$t_2(f_2) = f_2$$

$$t_3(f_3) = 1$$

$$t_4(f_4) = f_4 + 2$$

$$t_5(f_5) = f_5$$

When $\varepsilon = 0$, the equilibrium arc flow solution is $f^*(0) = [1,1,2,0,1]^T$. (Note that the arc flow solution, namely $f^*(0) = [1,1,1,1,1]^T$, given in Josefsson and Patriksson [52] is incorrect.) Since the solution violates strict complementary, the TFM should was never intended for application to this problem. However, since strict complementarity is not a necessary condition for differentiability, it is possible that the problem is differentiable in the neighborhood of $\varepsilon = 0$ and that the TFM may be applied, as is next discussed.

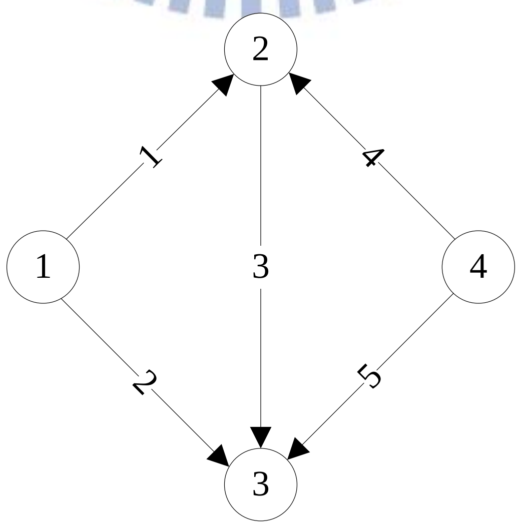

Fig. 2: The network of Josefsson and Patriksson's example 7.3.1

We consider perturbations of the travel demands; that is, $\varepsilon = [\varepsilon_{12} \quad \varepsilon_{42}]^T$. The unperturbed solution is $f^*(0) = [1,1,2,0,1]^T$, and the restricted arc-path and OD-path incidence matrices are

$$\Delta = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \text{ and } \Lambda = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

In the restricted arc-path and OD-path incidence matrices, arc 4 is eliminated in the restricted network. The rows of the restricted arc-path incidence matrix correspond to arcs 1, 2, 3 and 5. The columns of arc-path incidence matrix correspond to paths 1, 2 and 4, respectively. The corresponding path flow solution is $h^*(0) = [1,1,1]^T$ >0. The rank of $[\Delta^T, \Lambda^T]$ is equal to the number of paths with positive flow, which means that $h^*(0)$ is a nondegenerate extreme point. It follows that

$$\begin{bmatrix} \nabla_h c(h^*,0) & -\Lambda^T \\ \Lambda & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} -\nabla_\varepsilon c(h^*,0) \\ \nabla_\varepsilon T(0) \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Therefore, the sensitivity of equilibrium arc flows can be obtained by Eq. (2.25) ~ (2.27) as

$$\nabla_\varepsilon f^* = \begin{bmatrix} -1/3 \\ 1/3 \\ 1/3 \\ 0 \end{bmatrix}$$

which is identical to the solution reported in Josefsson and Patriksson [52].

3.2.3 Example 7.3.2 in Josefsson and Patriksson (2007)

In another example, Josefsson and Patriksson [52] considered a 3 node, 4 arc network with the single OD pair (1, 3) and 4 paths, as depicted in Fig. 3. There is a fixed demand of 2 units of flow for OD pair (1, 3). Also there are four paths: $p_1 = \{1, 3\}$, $p_2 = \{1, 4\}$, $p_3 = \{2, 3\}$ and $p_4 = \{2, 4\}$. Furthermore, the arc cost functions are

$$t_1(f_1, \varepsilon) = f_1 + \varepsilon$$

$$t_2(f_2) = f_2$$

$$t_3(f_3) = f_3$$

$$t_4(f_4) = f_4$$

where ε is a scalar perturbation parameter of the cost function of arc 1. When $\varepsilon = 0$, the equilibrium arc flow solution is $f^*(0) = [1,1,1,1]^T$. Thus, the restricted arc-path and OD-path incidence matrices are

$$\Delta = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \text{ and } \Lambda = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}.$$

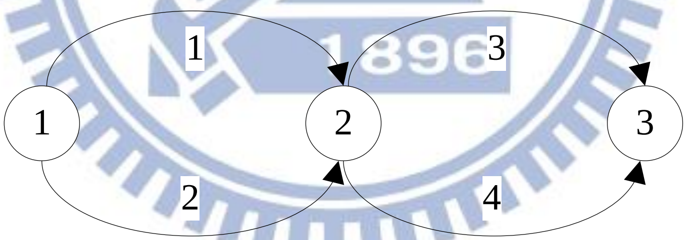

Fig. 3: The network of Josefsson and Patriksson's example 7.3.2

In this example, the rank of $[\Delta^T, \Lambda^T]$ is 3. As the analysis in Josefsson and Patriksson [52] establishes, the possible number of paths having non-zero flow is either 2 or 4. It is impossible to find a nondegenerate path flow solution with only 3 non-zero flows. In other words, it is impossible to satisfy the requirements of the TFM in this example. Therefore, the TFM should not be applied in this example to calculate the sensitivity information even though the gradient exists, and the example does not constitute a counterexample.

3.2.4 Example 7.3.3 in Josefsson and Patriksson (2007)

In this example, Josefsson and Patriksson [52] considered the 3 node, 3 arc network with 3 OD pairs and 4 paths depicted in Fig. 4. There are three fixed demands of 1 unit of flow for each of the OD pairs (1, 2), (1, 3) and (3, 2). There are four paths corresponding to the three OD pairs denoted by $p_1 = \{1\}$, $p_2 = \{2, 3\}$, $p_3 = \{2\}$ and $p_4 = \{3\}$. The arc cost functions are given by

$$t_1(f_1, \varepsilon) = 2f_1 + \varepsilon$$

$$t_2(f_2) = f_2$$

$$t_3(f_3) = f_3$$

where ε is again a scalar perturbation parameter. When $\varepsilon = 0$, the equilibrium arc flow solution is $f^*(0) = [1,1,1]^T$. Thus, the restricted arc-path and OD-path incidence matrices are

$$\Delta = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \text{ and } \Lambda = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

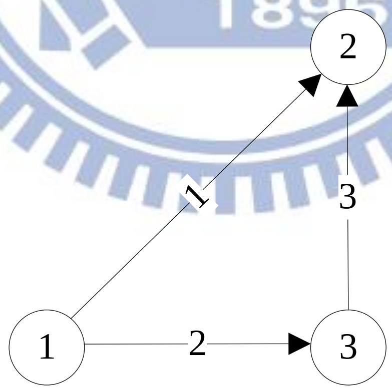

Fig. 4: The network of Josefsson and Patriksson's example 7.3.3

In this example, the rank of $[\Delta^T, \Lambda^T]$ is 4. Path 3 is the only path connecting OD pair (1, 3); so the flow on path 3 is exactly equal to the associated demand. Similarly, the flow on path 4 is exactly equal to the demand between OD pair (3, 2). Therefore, the flow on path 1 must be 1 and the flow on path 2 is zero. The flows just described form the unique equilibrium path flow solution, and the rank of $[\Delta^T, \Lambda^T]$ cannot possibly equal the number of paths with positive flow. Once again, this example does not satisfy the nondegenerate extreme point condition for the TFM. In addition, this example is non-differentiable, which fact also violates the TFM regularity conditions. Thus, the TFM should never be considered for application to this example, and the example, again, does not constitute a counterexample.

3.2.5 Modifying Example Problems to satisfy the TFM Conditions

Example 7.3.2 in Josefsson and Patriksson [52] violates the regularity assumptions on which the TFM is predicated due to its symmetric arc cost functions and network topology. If we make the cost functions asymmetric, as we shall illustrate, the possible number of paths having non-zero flow is no longer restricted to either 2 or 4. As depicted in Fig. 3. the network contains three nodes, four arcs, one OD pair and four paths. This example is identical to example 7.3.2 in Josefsson and Patriksson [52] except for the cost function of arc 2, which is now $t_2(f_2) = 1 + f_2$.

When $\varepsilon = 0$, the equilibrium arc flow solution is $f^*(0) = [3/2, 1/2, 1, 1]^T$. Thus, the restricted arc-path and OD-path incidence matrices are

$$\Delta = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \text{ and } \Lambda = [1 \quad 1 \quad 1 \quad 1].$$

According to the equilibrium arc flow solution, a path flow solution $h^*(0) = [1, 1/2, 0, 1/2]^T$ can be obtained. In this case, the rank of $[\Delta^T, \Lambda^T]$ is equal to the number of paths with positive flow, which implies that h^* is a nondegenerate extreme point. Due to the flow on path 3 being zero, we may eliminate path 3 to generatethe modified $h^{0*}(0) = [1, 1/2, 1/2]^T$ which contains only those path variables having positive path flows. Conformally defined with respect to h^{0*}, the modified arc-path and OD-path incidence matrices, denoted by Δ^0 and Λ^0 respectively, are

$$\Delta^0 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \text{ and } \Lambda^0 = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}.$$

It follows that

$$\begin{bmatrix} \nabla_h c^0(h^*,0) & -\Lambda^{0T} \\ \Lambda^0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 & -1 \\ 1 & 2 & 1 & -1 \\ 0 & 1 & 2 & -1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} -\nabla_\varepsilon c^0(h^*,0) \\ \nabla_\varepsilon T(0) \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ 0 \\ 0 \end{bmatrix}$$

Therefore, we have the following derivative of arc flows with respect to the perturbation (of the cost function for arc 1):

$$\nabla_\varepsilon f^* = \begin{bmatrix} -1/2 \\ 1/2 \\ 0 \\ 0 \end{bmatrix}$$

Clearly, all requirements of the TFM are satisfied and the derivative of differentiable solutions with respect to perturbation parameters can be calculated by the TFM. This contrasts to the example 7.3.3 in Josefsson and Patriksson [52] and the example 5 in Patriksson [51], for which non-differentiability is encountered in violation of the assumptions intrinsic to the TFM.

3.2.6 Applying Cho-Smith-Friesz Method on Counterexamples

We illustrate in this section that the Cho-Smith-Friesz method(CSFM) extends the circumstances under which the TFM is applicable.

3.2.6.1 Example 7.3.2 in Josefsson and Patriksson (2007)

This example has been stated previously. Recall that, when $\varepsilon = 0$, the equilibrium arc flow solution is $f^*(0) = [1,1,1,1]^T$ and the restricted arc-path and OD-path incidence matrices are

$$\Delta = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \text{ and } \Lambda = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}.$$

In this example, the rank of $[\Delta^T, \Lambda^T]$ is 3. By inspection, Λ has one independent row. Also the rows of Δ that correspond to arcs 1 and 3 are linearly independent, so that we may partition Δ according to

$$\Delta = [\Delta_1 \quad \Delta_2]^T. \tag{3.1}$$

where

$$\Delta_1 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix} \text{ and } \Delta_2 = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}.$$

The matrices referred to in Eq. (2.52) are the followings:

$$\nabla_f t(f(0),0) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$M = \begin{bmatrix} -1 & -1 & 0 & 0 \\ 0 & 0 & -1 & -1 \end{bmatrix}$$

$$\nabla_\varepsilon t(f(0),0) = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Delta_2 N_2 \nabla_\varepsilon T(0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Therefore, Eq. (2.52) reduces to the following:

$$\nabla_\varepsilon f^* = \begin{bmatrix} -1/2 \\ 1/2 \\ 0 \\ 0 \end{bmatrix}$$

which is identical to the solution in Josefsson and Patriksson [52]. Thereby, we see that the CSFM may sometimes be able to deal with cases wherein h^* is not a nondegenerate path solution.

3.2.6.2 Example 7.3.3 in Josefsson and Patriksson (2007)

When $\varepsilon = 0$, the equilibrium arc flow solution is $f^*(0) = [1,1,1]^T$. Thus, the restricted arc-path and OD-path incidence matrices are

$$\Delta = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \text{ and } \Lambda = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

In this example, the rank of $[\Delta^T, \Lambda^T]$ is 3. By inspection Λ has three independent rows. Moreover, the rows of Δ are linearly independent. Thus

$$\Delta = \Delta_2$$

based on the notation introduced in Eq. (3.1). Therefore, the matrices of Eq. (2.52) may be expressed as

$$\nabla_f t(f(0),0) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$M = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

$$\nabla_\varepsilon t(f(0),0) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\Delta_2 N_2 \nabla_\varepsilon T(0) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Furthermore, Eq. (2.52) yields

$$\nabla_\varepsilon f^* = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

which is the directional derivative when $\varepsilon = 1$ employed in the Josefsson-Patriksson solution. However, the directional derivative when $\varepsilon = -1$ cannot be obtained by the CSFM because it corresponds to a non-differentiable circumstance, and thereby violates the regularity conditions needed for application of the CSFM method.

3.3 Counterexamples in Yang and Bell (2007)

Subsequent to Bell and Iida [7], non-invertibility of the Jacobian of the path cost occurring in Eq. (2.52) of the TFM was also observed by Yang and Bell [67]. The Jacobian of the path cost is in general not invertible so that the sensitivity formulae of the TFM would seem to fail in such a circumstance. However, Yang and Bell [67] presented an example for which the Jacobian of the path cost is not invertible yet the requisite information for sensitivity analysis exists. As we have already indicated, this is not a surprise, since the invertibility of the Jacobian of path cost is a regularity condition for the TFM that one may relax when it is realized that the sensitivity analysis formulae may be derived via the CSFM. More generally, the TFM sensitivity formulae remain applicable when the Jacobian of the path cost is not invertible. This understanding, as we now reiterate, is established by the analysis of Cho et al. [16] and the summary thereof presented in Section 3.3, because Cho et al. [16] derive those formulae without reference to the network's topology or the presumption of invertibility of the Jacobian of path cost. Our remarks immediately above are not criticisms of the manuscripts cited but rather are meant to establish connections among the various papers on the subject of equilibrium sensitivity analysis that have appeared over a considerable period of time in various journals and books.

3.3.1 Example 6 in Yang and Bell (2007)

In this example, Yang and Bell [67] considered a 3 node, 2 arc network with 3 OD pairs and 3 paths which is depicted in Fig. 5. There are two fixed demands of 5 units of flow for each of the OD pairs (1, 2) and (2, 3). The demand for OD pair (1, 3) is perturbed and becomes $5+\varepsilon_{13}$. The three paths are $p_1=\{1\}$, $p_2=\{2\}$ and $p_3=\{1,2\}$. The arc cost functions are given by

$$t_1(f_1,\varepsilon_1)=f_1+1+\varepsilon_1$$

$$t_2(f_2)=f_2+1$$

It is helpful to define

$$\varepsilon = \begin{bmatrix} \varepsilon_{13} \\ \varepsilon_{1} \end{bmatrix}.$$

When $\varepsilon = 0$, the equilibrium arc flow solution is $f^{*}(0) = [10,10]^{T}$. Thus, the restricted arc-path and OD-path incidence matrices are

$$\Delta = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \text{ and } \Lambda = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

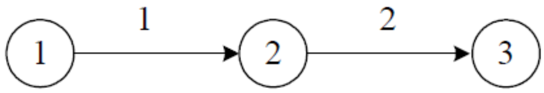

Fig. 5: The network of Yang and Bell's Example 6

According to the equilibrium arc flow solution, a path flow solution $h^{*}(0) = [5,5,5]^{T}$ is obtained. The rank of $[\Delta^{T}, \Lambda^{T}]$ is equal to the number of paths with positive flow, which implies that h^{*} are all positive. By inspection, we have

$$\begin{bmatrix} \nabla_{h} c^{0}\left(h^{*},0\right) & -\Lambda^{0T} \\ \Lambda^{0} & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & -1 & 0 & 0 \\ 0 & 1 & 1 & 0 & -1 & 0 \\ 1 & 1 & 2 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}. \tag{3.2}$$

Yang and Bell [67] note that the inverse of the Jacobian of the path cost matrix,

$$\nabla_{h} c^{0}\left(h^{*},0\right) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}^{-1} \tag{3.3}$$

does not exist, and it would seem that the TFM is not applicable. However, the inverse of the entire matrix presented on the right-hand side of Eq. (3.2) is easily shown to be the following:

$$\begin{bmatrix} \nabla_h c^0\left(h^*,0\right) & -\Lambda^{0T} \\ \Lambda^0 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 1 & 0 & 1 \\ 0 & -1 & 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 1 & 1 & 2 \end{bmatrix}. \tag{3.4}$$

Also we note that

$$\begin{bmatrix} -\nabla_\varepsilon c^0\left(h^*,0\right) \\ \nabla_\varepsilon T(0) \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 0 \\ -1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}$$

According to Eqs. (2.25)-(2.27), the derivative of arc flows with respect to ε is

$$\nabla_\varepsilon f^* = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \tag{3.5}$$

which is identical to the solution in Yang and Bell [67]. The derivative of equilibrium costs with respect to ε is

$$\nabla_\varepsilon \mu^* = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 2 \end{bmatrix} \tag{3.6}$$

which is also identical with the solution in Yang and Bell [67]. Thus, the TFM may remain applicable even though the Jacobian of the path cost matrix is not invertible. More precisely, the non-invertibility problem encountered in expression Eq. (3.3) need not be inherited by inverse matrix in Eq. (2.25).

3.4 Discussion

This section has reviewed several prior articles in which some defects of the Tobin-Friesz method (TFM) are indicated and some examples are shown to explain that the TFM may fail or lead to an incorrect solution. However, we found this is not because the TFM is incorrect but because some of the so-called counter examples were purposely fabricated to violate the assumptions intrinsic to the TFM and the related Cho-Smith-Friesz method (CSFM). That is, the counter examples reported in the literature are not true counter examples for the TFM because they violate the regularity conditions on which the method is based.

Nonetheless, we readily acknowledge that the TFM must be applied with care. The reward for exercising such care is the rather simple formulae that characterize the TFM. In this paper, we have seen that, sometimes when the regularity conditions of the TFM are violated, the equilibrium problem of interest may be modified to allow the method's application; this is especially so when cost symmetries are the complicating aspect of a given problem. We also observed that sometimes the CSFM may be employed in lieu of the TFM to deal with noninvertible Jacobians.

CHAPTER 4

Extension of Gradient-Based Sensitivity Analysis Method

The sensitivity analysis method is the most popular algorithm in solving bilevel programming problem in the field of transportation science. In the sensitivity analysis method, the reaction function of the lower level problem is usually approximated by the first-order sensitivity information of equilibrium network flows. Due to the non-convexity and non-differentiability of the bilevel problem, the high-order sensitivity analysis method is expected to solve the problem more efficiently. Therefore, the second-order sensitivity formula of equilibrium network flows is one of the objectives in this research.

Based on the first-order gradient-based sensitivity formula proposed by Cho et al. [16], the second-order sensitivity formula is regarded as taking the derivative of the first-order formula with respect to perturbation parameters. Since the first-order gradient-based sensitivity formula is usually a matrix, we have to introduce the theory of matrix differential calculus to derive the second-order sensitivity formula.

In this section, we firstly review the main result of the row reduction gradient-based sensitivity method which is the first-order gradient-based sensitivity formula. Subsequently, some definitions and theorems of matrix differential calculus are introduced which applied to derive the second-order gradient-based sensitivity formula.

4.1 Existence of High-Order Sensitivity of Gradient-Based Method

In a general mathematical programming, the existence of high-order sensitivity of a local solution has been proposed by Fiacco [29]. Consider the problem $P(\varepsilon)$ of determining a local solution $x(\varepsilon)$ of

$$\underset{x}{\text{Minimize}} \quad g(x,\varepsilon) \tag{4.1}$$

$$\text{subject to} \quad p_i(x,\varepsilon) \geq 0 \quad (i = 1, \ldots, m), \tag{4.2}$$

$$q_j(x,\varepsilon) \quad (j=1, ..., n), \tag{4.3}$$

where $x \in R^n$ and ε is a parameter vector in R^{ε}.

In order to show the existence of high-order sensitivity, we consider the following conditions:

Condition 4.1

The functions defining $P(\varepsilon)$ are $(p+1)$th order continuously differentiable in x and if their gradients with respect to x and the constraints are pth order continuously differentiable in ε in a neighborhood of $(x^*, 0)$, with $p \geq 1$.

Condition 4.2

The second-order sufficient for a local minimum of $P(0)$ hold at x^*, with associated Lagrange multipliers u^* and v^*.

Condition 4.3

The gradients $\nabla p_i(x^*,0)$ (for i such that $p_i(x^*,0)=0$) and $\nabla q_j(x^*,0)$ (all j) are linearly independent.

Condition 4.4

$u_i^* > 0$ when $p_i(x^*,0)=0$ (i = 1, …, m) (i.e., strict complementary slackness holds).

Based on condition 4.1~4.4, we can show the existence of high-order sensitivity proposed by Fiacco [28] as follows:

Theorem 4.1 (Existence of high-order sensitivity, [29])

If condition 4.1~4.4 hold, then $y(\varepsilon) \equiv [x(\varepsilon), u(\varepsilon), v(\varepsilon)]^T \in C^p$ in a neighborhood of $\varepsilon = 0$. If the problem functions are analytic in (x, ε) in a neighborhood of $(x^*, 0)$, then $y(\varepsilon)$ is analytic in a neighborhood of $\varepsilon = 0$.

Hence, if condition 4.1~4.4 hold with p=2, then the second-order sensitivity of the solution of $P(\varepsilon)$ exists. Moreover, if condition 4.1~4.4 hold with p=1, then x^* is a locally unique solution to variational inequality problem which can derive the gradient-based sensitivity analysis method [61]. In order to derive the second-order sensitivity formula, we start with the row reduction gradient-based method [16] and assume $c(\cdot,\varepsilon)$ is 3rd order continuously differentiable in arc-flow f and $c(\cdot,\varepsilon)$ and $T(\varepsilon)$ are twice continuously differentiable in ε.

4.2 Preliminary Definitions and Theorems for Second-Order Sensitivity

In this section, the Kronecker product and the *vec* operator are introduced [44]. The Kronecker product maps two matrices $A=(a_{ij})$ and $B=(b_{st})$ into a matrix $C=(a_{ij}b_{st})$. The *vec* operator transforms a matrix into a vector by stacking its columns one underneath the other. According to the definitions, the chain rule and the second-order Taylor expansion with respect to a matrix are derived which can be used to provide the second-order sensitivity analysis formula of equilibrium network flows.

Definition 4.1 (Kronecker product)

Let U be an $m\times n$ matrix and V be a $p\times q$ matrix, then the Kronecker product of U and V, denoted by $U \otimes V$, is an $mp\times nq$ matrix defined by

$$U \otimes V = \begin{bmatrix} u_{11}V & \cdots & u_{1n}V \\ \vdots & \ddots & \vdots \\ u_{m1}V & \cdots & u_{mn}V \end{bmatrix}. \tag{4.4}$$

Definition 4.2 (*vec* operator)

Let U be an $m \times n$ matrix and U_j is the j-th column of U, then *vec* U is the $mn \times 1$ vector

$$vec\,U = \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_n \end{bmatrix}. \tag{4.5}$$

Based on the definition of *vec* operator, the derivative of matrix functions of matrices can be defined as follows.

Definition 4.3

Let U be an $m \times n$ real matrix function of a $p \times q$ matrix of real variables ε. The derivative of U with respect to s is the $mn \times pq$ matrix

$$\nabla_{\varepsilon} U = \frac{\partial\, vec\, U}{\partial (vec\ \varepsilon)^T}. \tag{4.6}$$

In order to derive the second-order sensitivity formula, we have to define the chain rule for matrix functions first. Hence, Theorem 4.2 is introduced to provide the formula of chain rule for matrix functions.

Theorem 4.2 (Chain rule for matrix functions, [43])

Let D be a subset of $R^{m \times n}$, and assume that $U : D \to R^{p \times q}$ is differentiable at an interior point y of D. Let P be a subset of $R^{p \times q}$ such that $U(x) \in P$ for all $x \in D$, and assume that $V : P \to R^{r \times s}$ is differentiable at an interior point $z = U(y)$ of P. Then the composite function $F : D \to R^{r \times s}$ defined by $F(x) = V(U(x))$ is differentiable at y, and

$$\nabla_y F = (\nabla_z V)(\nabla_y U). \tag{4.7}$$

Theorem 4.3 (Derivative of simple product of matrices, [43])

Let $U : D \to R^{m \times r}$ and $V : D \to R^{r \times n}$ be two matrix functions defined and differentiable on an open set D in $R^{p\times q}$. Then the simple product UV is differentiable on D and the Jacobian matrix is the $mn\times pq$ matrix

$$\nabla_{\varepsilon}(UV)=\frac{\partial\, vec\, UV}{\partial(vec\ \varepsilon)^{T}}=\left(V^{T}\otimes I_{m}\right)\nabla_{\varepsilon}U+\left(I_{n}\otimes U\right)\nabla_{\varepsilon}V, \tag{4.8}$$

where I_m and I_n are the identity matrices of size m and n, respectively.

Theorem 4.4 ([43])

Let f: $D\to R^m$ be a function defined on a set D in R^n. Let r be an interior point of D, and let $B(\varepsilon_0; r)$ be an n-ball lying in D. Let ε be a point in R^n with $\|\varepsilon\| < r$, so that $\varepsilon_0+\varepsilon\in B(\varepsilon_0; r)$. If f is twice differentiable at ε_0, then the second-order Taylor expansion of function f at $\varepsilon_0+\varepsilon$ is

$$f(\varepsilon_0+\varepsilon)=f(\varepsilon_0)+df(\varepsilon_0;\varepsilon)+\frac{1}{2}d^2f(\varepsilon_0;\varepsilon), \tag{4.9}$$

where $df(\varepsilon_0;\ \varepsilon)$ and $d^2f(\varepsilon_0;\ \varepsilon)$ are the first differential and the second differential of f at ε_0, respectively, and

$$df(\varepsilon_0;\varepsilon)=\nabla_{\varepsilon}f(\varepsilon_0)\cdot(\varepsilon-\varepsilon_0), \tag{4.10}$$

$$d^2f(\varepsilon_0;\varepsilon)=\left((\varepsilon-\varepsilon_0)^T\otimes I_m\right)\cdot\nabla_{\varepsilon}^2f(\varepsilon_0)\cdot(\varepsilon-\varepsilon_0). \tag{4.11}$$

4.3 Second-Order Sensitivity Formula for Network Equilibrium Flows

To derive the second-order sensitivity formula for equilibrium network flows, it is intuitively to take derivative of Eq. (2.44) with respect to s. For convenience, let

$$\begin{bmatrix}\nabla_f t(f(\varepsilon),\varepsilon) & -M^T\\ M & 0\end{bmatrix}^{-1}=U(f(\varepsilon),\varepsilon), \tag{4.12}$$

$$\begin{bmatrix}-\nabla_{\varepsilon}t(f(\varepsilon),\varepsilon)\\ \Delta_2N_2\nabla_{\varepsilon}T(\varepsilon)\end{bmatrix}=V(f(\varepsilon),\varepsilon), \tag{4.13}$$

where U is an $(\alpha+\alpha_2)\times(\alpha+\alpha_2)$ matrix, and V is an $(\alpha+\alpha_2)\times\rho$ matrix, respectively.

Lemma 4.1

The second-order sensitivity for equilibrium network flows is

$$\begin{bmatrix}\nabla_{\varepsilon}^2 f\\ \nabla_{\varepsilon}^2\mu\end{bmatrix}=\left(V^T\otimes I_{(\alpha+\alpha_2)}\right)\nabla_{\varepsilon}U+\left(I_k\otimes U\right)\nabla_{\varepsilon}V, \tag{4.14}$$

where

$$U\left(f(\varepsilon),\varepsilon\right)=\begin{bmatrix}\nabla_f t\left(f(\varepsilon),\varepsilon\right) & -M^T \\ M & 0\end{bmatrix}^{-1}, \tag{4.15}$$

$$V\left(f(\varepsilon),\varepsilon\right)=\begin{bmatrix}-\nabla_\varepsilon t\left(f(\varepsilon),\varepsilon\right) \\ \Delta_2 N_2 \nabla_\varepsilon T(\varepsilon)\end{bmatrix}. \tag{4.16}$$

Proof

Since the first-order sensitivity is the product of Eq. (4.15) and (4.16), the second-order sensitivity can be obtained by taking derivative of the product with respect to *s* directly. According to Theorem 4.3, the formula of the second-order sensitivity is expressed as Eq. (4.14) and the proof is complete.

4.4 Application to Solve Stackelberg Games with Sensitivity Analysis Method

In this section, we briefly introduced the formulation of a Stackelberg game or a leader-follower game between road users and the administration sector [31]. With the sensitivity information, two algorithms are proposed to solve a Stackelberg game [15].

4.4.1 Formulation of Stackelberg Games in the Field of Transportation

In the field of transportation, Fisk [31] pointed out a Stackelberg game can be represented as a bilevel problem where the upper level problem aims to find the optimal signal setting or capacity enhancement of arcs which maximizes system performance, and the lower level problem aims to solve the user equilibrium flows, respectively. The solution algorithm for calculating optimal strategy in general road networks should take the anticipating the reactions of road users into account. However, the iterative optimization assignment procedure which solves strategy and equilibrium flows iteratively cannot be expected to converge to the true solution and might lead to a decline in network performance. By contrast, the sensitivity analysis-based algorithm evaluates the influence factors as the derivatives of the reaction functions with respect to the upper-level decision variables. The derivative information is obtained by implementing sensitivity analysis for a given solution of the user equilibrium problem. With this information, the linear approximation of the reaction function can be obtained and applied to sensitivity analysis-based algorithm.

Due to the nonlinearity of the perturbed solutions in equilibrium constraints, the nonlinear approximation of the reaction function is expected to solve a Stackelberg game more efficiently. Therefore, this research tries to establish the theory of high-order sensitivity analysis of network equilibrium flows which can be applied to solve a Stackelberg game with a nonlinear approximation of the reaction function.

Consider a signal optimization problem where the aim of the regulating agency is to minimize a network performance function P1 such as total travel time or gas consumption, with fixed OD travel demand, where travelers selecting routes on the network in an optimal user fashion. Notably, D denotes the set of feasible signal control variables. For any given $\varepsilon \in D$ a user optimal arc flow solution $f(\varepsilon) \in \Omega$ exists and the problem of the regulator is to solve

$$\begin{aligned} P1: \quad & \min_{\varepsilon \in D} P\big(f(\varepsilon), \varepsilon\big) \\ & s.t. \quad \text{U.E.} \end{aligned} \tag{4.17}$$

In the general problem, the signal variables that can be set by the controlling agent include green and cycle times, and offsets. By specifying the cost functions t_a for each network arc a in terms of these variables, and assuming that the behavioral hypothesis for route choice follows the first principle of Wardrop [62], problem P1 can be presented as

$$P_2: \quad \min_{\varepsilon \in D} \sum_{a \in A} t_a\big(f(\varepsilon), \varepsilon\big) f_a(\varepsilon) \tag{4.18}$$

$$s.t. \quad t\big(f(\varepsilon), \varepsilon\big) \cdot (u - f) \geq 0, \ \forall u \in \Omega. \tag{4.19}$$

If $t(\varepsilon, f)$ is strictly monotone, then for each $\varepsilon \geq 0$, Eq. (4.19) has a unique solution, and function $f(\varepsilon)$ is (continuously) differentiable at every point $\varepsilon \geq 0$. Thus, P2 can be rewritten as P3.

$$
\begin{aligned}
P_3: \quad & \min_{\varepsilon \in D} Z(s) = \sum_a t_a\left(f(\varepsilon), \varepsilon\right) f_a(\varepsilon) \\
& s.t. \quad \varepsilon \geq 0 .
\end{aligned} \tag{4.20}
$$

Also, given $R(f(\varepsilon), \varepsilon) = \min t(f(\varepsilon), \varepsilon)\cdot(u\text{-}f)$, then P2 is equivalent to P4.

$$
\begin{aligned}
P_4: \quad & \min_{\varepsilon \in D} Z(s) = \sum_a t_a\left(f(\varepsilon), \varepsilon\right) f_a(\varepsilon) \\
& s.t. \quad R\left(f(\varepsilon), \varepsilon\right) = 0 .
\end{aligned} \tag{4.21}
$$

4.4.2 A Sensitivity Analysis-Based Linear Approximation Heuristic Algorithm

The Iterative Optimization Assignment (IOA) method is proceeded as follows. First, fix ε and solve Eq. (4.19) for f, then fix f and solve Eq. (4.18) for s, continuing this process until $\varepsilon^{k+1} - \varepsilon^{k} \rightarrow 0$ or $f^{k+1} - f^{k} \rightarrow 0$. The final solution (f^N, ε^N) is termed the Nash solution. Notably, that the solution obtained using the IOA algorithm is not necessarily an optimal solution of the equilibrium network control problem [31]. The sensitivity analysis of equilibrium network flows was used to solve the equilibrium network signal design problem.

The challenge in solving problem P2 is that, since the lower level of the problem cannot be represented in closed form, it is impossible to obtain an explicit reaction function that can be plugged into the upper level. In the sensitivity analysis-based linear approximation heuristic algorithm, the sensitivity information is used to create a linear approximation of the reaction function and is then inserted into the upper level problem, iterating until the solutions converge (abbreviated as LAA). This heuristic algorithm was firstly proposed by Cho and Lo to solve the continuous equilibrium network design problem [15].

The heuristic is detailed as follows:

(A1)

Step 0: Determine a fixed small value $\delta > 0$ and an initial value ε^0. Set $k = 0$.

Step 1: Solve Eq. (4.19) given ε^k and yielding f^k.

Step 2: Calculate the sensitivity information $\nabla_\varepsilon f$ by Eq. (2.44).

Step 3: Using $\nabla_\varepsilon f$, Taylor expansion and Theorem 4.4, form the linear approximation f^{k+1},

$f^{k+1} = f^k + \nabla_\varepsilon f \cdot (\varepsilon^{k+1} - \varepsilon^k)$. Since f^k, ε^k and $\nabla_\varepsilon f$ are known, f^{k+1} can be replaced by a function of ε^{k+1}. Thus, $f^{k+1} = A + B\varepsilon^{k+1}$.

Step 4: Reformulate Eq. (4.20) as

$$\min_{\varepsilon \in D} \sum_a t_a\left(A + B\varepsilon^{k+1}, s\right) \cdot \left(A + B\varepsilon^{k+1}\right)$$
$$s.t. \quad \varepsilon \geq 0 .$$

Step 5: Solve the problem in step 4 using any software package which can solve the optimal solution for ε^{k+1}. If $|\varepsilon^{k+1} - \varepsilon^k| \leq \delta$, then stop, otherwise set $k = k + 1$ and go to step 1.

4.4.3 A Sensitivity Analysis-Based Nonlinear Approximation Heuristic Algorithm

In the sensitivity analysis-based linear approximation algorithm, the reaction function of the lower level is based on approximation by a linear function. In this section, the reaction function of the lower level problem is based on approximation by a nonlinear function, and is plugged into the upper level problem and iterated until the solutions converge (abbreviated as NLAA).

(A2)

Step 0: Determine a fixed small value $\delta > 0$ and an initial value ε^0. Set $k = 0$.

Step 1: Solve Eq. (4.19) given ε^k and yielding f^k.

Step 2: Calculate the sensitivity information $\nabla_\varepsilon f$ and $\nabla^2_\varepsilon f$ by Eq. (2.44) and (4.14).

Step 3: Using $\nabla_\varepsilon f$, $\nabla^2_\varepsilon f$ Taylor expansion and Theorem 3, form the nonlinear approximation f^{k+1},

$$f^{k+1} = f^k + \nabla_\varepsilon f \cdot \left(\varepsilon^{k+1} - \varepsilon^k\right) + \left(\left(\varepsilon^{k+1} - \varepsilon^k\right)^T \otimes I_m\right) \cdot \nabla^2_\varepsilon f \cdot \left(\varepsilon^{k+1} - \varepsilon^k\right).$$

Since f^k, ε^k, $\nabla_\varepsilon f$ and $\nabla^2_\varepsilon f$ are known, f^{k+1} can be replaced by a function of ε^{k+1}. Thus, $f^{k+1} = A + B\varepsilon^{k+1} + C(\varepsilon^{k+1})^2$.

Step 4: Reformulate Eq. (4.20) as

$$\min_{\varepsilon \in D} \sum_a t_a \left(A + B\varepsilon^{k+1} + C\left(\varepsilon^{k+1}\right)^2, \varepsilon\right) \cdot \left(A + B\varepsilon^{k+1} + C\left(\varepsilon^{k+1}\right)^2\right)$$
$$s.t. \quad \varepsilon \geq 0 .$$

Step 5: Solve the problem in step 4 using any software package which can solve the optimal solution for ε^{k+1}. If $|\varepsilon^{k+1} - \varepsilon^k| \leq \delta$, then stop, otherwise set $k = k + 1$ and go to step 1.

In addition to describe the algorithm in more detail, we will provide a proof that if this algorithm converges; it converges to an optimal solution of problem P2.

Lemma 4.2

If algorithm A2 converges, it converges to a critical point of P2.

Proof

If the sequence s^k converges to s^*, $s^k \to s^*$, then we know that:

(1) If we set $s^0 = s^*$, $f^0 = f^*$ then $s^1 = s^*$, $f^1 = f^*$; and

(2) Let

$$\hat{Z}(s) = \sum_{a \in A} t_a \left(A + Bs + Cs^2, s\right) f_a(s)$$

Then, by the Karush-Kuhn-Tucker necessary conditions for optimality of vectors $s^* \geq 0$, we know the following must be true:

(i)

$$\frac{\partial \hat{Z}(s^*)}{\partial s^i} = 0 \quad \text{if } s^{*i} > 0, i = 1, ..., n,$$

(ii)

$$\frac{\partial \hat{Z}(s^*)}{\partial s^i} \geq 0 \quad \text{if } s^{*i} = 0, i = 1, ..., n.$$

So, taking the derivative with respect to s, we get

$$\frac{\partial}{\partial s^i} \left[\sum_a t_a \left(A + Bs + Cs^2, s\right)\left(A + Bs + Cs^2\right) \right]\Bigg|_{s=s^*}$$

$$= \sum_a t_a\left(A + Bs + Cs^2, s\right)\cdot\left(B + 2Cs\right)\Big|_{s=s*}$$

$$+ \sum_a \left[\frac{\partial t_a\left(A + Bs + Cs^2, s\right)}{\partial\left(A + Bs + Cs^2\right)}\cdot\left(B + 2Cs\right) + \frac{\partial t_a\left(A + Bs + Cs^2, s\right)}{\partial s}\right]\cdot\left(A + Bs + Cs^2\right)\Big|_{s=s*}.$$

Further, we know that

$$B = D_s f\big|_{s=s*},$$

$$C = \frac{1}{2}\cdot H_s f\big|_{s=s*},$$

and

$$f(s) = A + Bs + Cs^2.$$

So, substituting Eq. (42) we know

$$\frac{\partial}{\partial s^i}\left[\sum_a t_a\left(A + Bs + Cs^2, s\right)\left(A + Bs + Cs^2\right)\right]\Bigg|_{s=s*}$$

$$= \sum_a t_a\left(f(s), s\right)\cdot\left(D_s f + H_s f\cdot s\right)\Big|_{s=s*}$$

$$+ \sum_a \left[\frac{\partial t_a\left(f(s), s\right)}{\partial f}\cdot\left(D_s f + H_s f\cdot s\right) + \frac{\partial t_a\left(f(s), s\right)}{\partial s}\right]\cdot f_a\Big|_{s=s*,\ f=f*}$$

$$= \sum_a t_a\left(f(s), s\right)\cdot D_s f(s)\Big|_{s=s*} + \sum_a \left[\frac{\partial t_a\left(f(s), s\right)}{\partial f}\cdot D_s f(s) + \frac{\partial t_a\left(f(s), s\right)}{\partial s}\right]\cdot f_a\Big|_{s=s*,\ f=f*}$$

$$= \frac{\partial}{\partial s}\left[\sum_a t_a\left(f(s), s\right)\cdot f_a(s)\right]\Bigg|_{s=s*,\ f=f*}$$

So, we know if conditions (i) and (ii) are satisfied, then the following should also be satisfied

(iii)

$$\frac{\partial Z(s*)}{\partial s^i} = 0 \quad \text{if } s*^i > 0, i = 1, \ldots, n,$$

(iv)

$$\frac{\partial Z(s*)}{\partial s^i} \geq 0 \quad \text{if } s*^i = 0, i = 1, \ldots, n.$$

CHAPTER 5

Extension of Directional Derivative-Based Sensitivity Analysis Method

5.1 Formulation of Derivative-Based Sensitivity Analysis Method

Consider a parametric variational inequality (VI) with a parametric polyhedral constraint as following [70]

$$f(x_0,\varepsilon_0)^T(x-x_0)\geq 0,\ \ \forall x\in K(\lambda_0), \tag{5.1}$$

$$K(\lambda)=\{x\in R^n | Ax\geq\lambda,\ x\geq 0\}, \tag{5.2}$$

$$\Gamma=\{\lambda\in R^r | K(\lambda)\neq 0\},\quad E=\{\varepsilon\in R^m\}, \tag{5.3}$$

where f: $R^n\times E\rightarrow R^n$ is a given function, x_0 is a solution of this VI problem and $(\varepsilon_0, \lambda_0)\in E\times\Gamma$ are given parameters. Let S be the mapping that assigns each ε the set $S(\varepsilon)$ of solutions to the VI problem:

$$S(\varepsilon)=\{x^*\in C\,|\,0\in f(x^*,\varepsilon)+N_{K(\lambda)}(x^*)\}. \tag{5.4}$$

For a convex set C in R^n, the normal and tangent cones to C at a point $x\in C$ are denoted by $N_C(x)$ and $T_C(x)$, respectively, defined by

$$N_C(x)=\begin{cases}\{z\in R^n\,|\,z^T(y-x)\leq 0,\ \forall y\in C\}, & x\in C\\ 0 & ,\ x\notin C\end{cases} \tag{5.7}$$

and

$$T_C(x)=\mathrm{cl}\{z\in R^n\,|\,\text{for some }\lambda>0,\ x+\lambda z\in C\}. \tag{5.8}$$

With the definition of convex cone, the form of VI Eq. (5.1) is equivalent to

$$0\in f(x_0,\varepsilon_0)+N_{K(\lambda)}(x_0). \tag{5.9}$$

Based on the VI form of the equilibrium network flow problem, the directional derivatives of equilibrium network flows can be calculated by solving a quadratic optimization problem [37] as

$$\underset{f'}{\text{Minimize}} \quad \phi'(f') = \left[\nabla_{\varepsilon} t\left(f^*, \varepsilon^*\right)\varepsilon'\right]^T f' + \frac{1}{2}\sum_{a \in A} \frac{\partial t_a\left(f_a^*, \varepsilon^*\right)}{\partial f_a}\left(f_a'\right)^2$$

$$\text{subject to} \quad \Lambda h' = d', \ \Delta h' = f'$$

$$h' \in H',$$

where

$$H' = \left\{ h' \in R^{|P|} \left| \begin{array}{l} h_p' \text{ is free if } h_p^* > 0 \\ h_p' \geq 0 \text{ if } h_p^* = 0 \text{ and } c_p\left(h^*, \varepsilon^*\right) = \pi_w^* \\ h_p' = 0 \text{ if } h_p^* = 0 \text{ and } c_p\left(h^*, \varepsilon^*\right) > \pi_w^* \\ p \in P_w, \ w \in W \end{array} \right. \right\}.$$

In this model, directional derivatives of link flow, path-flow, and travel demand are denoted by *f*', *h*', and *d*', respectively. The directional derivative can be interpreted as the direction and the rate of change of the equilibrium solution when perturbation parameter is perturbed along the direction ε'. The set *H*' is the set of directional derivatives of path-flow which keep the feasibility and optimality of the original problem in the first order approximation. We only summarize main results here and readers are encouraged to refer the original papers [37, 52] for more details.

In previous works, link cost functions are assumed continuously differentiable. This assumption is relaxed by adopting piecewise linear link cost functions in this study. A quadratic programming model with complementary constraints is proposed to calculate directional derivatives of equilibrium network flows with piecewise linear link cost functions.

5.2 Sensitivity Analysis without Continuously Differentiable Assumption

In the sensitivity analysis without continuously differentiable assumption of cost functions, two basic assumptions needs to be satisfied to guarantee the perturbed solution is Lipschitz continuous so that the directional derivative of perturbed solution exists. Consider two assumptions as follows:

Assumption 5.1

Assume g is locally Lipschitz at (x_0, ε_0), that is

$$\|g(x',\varepsilon')-g(x,\varepsilon)\| \le l(\|x'-x\|+\|\varepsilon'-\varepsilon\|),\ \forall x,x'\in X;\ \varepsilon,\varepsilon'\in E\cap U, \tag{5.10}$$

where X and U are the neighborhoods of x_0 and ε_0 respectively and $l > 0$ is a constant.

Assumption 5.2

Assume $g(\cdot,\varepsilon)$ is locally strongly monotone around x_0 with a common coefficient for all $\varepsilon\in E\cap U$, that is

$$\langle g(x',\varepsilon)-g(x,\varepsilon), x'-x\rangle \ge \alpha\|x'-x\|^2,\ \forall x,x'\in X;\ \varepsilon\in E\cap U, \tag{5.11}$$

where $\alpha>0$ is a constant.

Theorem 5.1 (Theorem 5.1 from [70])

Let $K(\lambda)$ be defined by Eq. (5.2) and x_0 be a solution of Eq. (5.1), where $(\varepsilon_0, \lambda_0)\in E\times\Gamma$ is a given pair of parameters. If Assumption 5.1 and Assumption 5.2 are satisfied, then there exist constants $k_\varepsilon>0$ and $k_\lambda>0$, neighborhoods U of ε_0 and V of λ_0 such that:

(1) For every $(\varepsilon, \lambda)\in(E\cap U)\times(\Gamma\cap V)$ there exists a unique solution of Eq. (5.1) in X, denoted by $x(\varepsilon, \lambda)$;

(2) For all (ε', λ'), $(\varepsilon, \lambda)\in(E\cap U)\times(\Gamma\cap V)$,

$$\|x(\varepsilon',\lambda')-x(\varepsilon,\lambda)\| \le k_\varepsilon\|\varepsilon'-\varepsilon\|+k_\lambda\|\lambda'-\lambda\|.$$

From Theorem 5.1, we can know that when cost function, $t(\cdot,\varepsilon)$, is locally Lipschitz at (x_0, ε_0) and locally strongly monotone, the perturbed solution is Lipschitz continuous. That is, in a traffic network with locally Lipschitz, locally strongly monotone function of costs on arcs, the equilibrium arcs flow is locally unique and is a locally Lipschitz function of the perturbation of costs on arcs and of the vector of demand [70]. Thus, in this case, the directional derivative of the equilibrium network flows exists.

Consider a piecewise linear function are usually used to approximate an arbitrary continuous function. In the domain [$f_{a,1}$, $f_{a,m}$], a series of grid points, $f_{a,1}$, $f_{a,2}$, …, $f_{a,m}$, is introduced which divides the domain into m-1 intervals, [$t_{a,i}$, $t_{a,i+1}$], i=1, 2, …, m-1. In each interval i, a linear function, $t_{a,i}(f_a,\varepsilon)$, is introduced. In general, the formulation of a piecewise linear function $t_a(v_a,\varepsilon)$ can be represented as the following equation.

$$t_a(f_a,\varepsilon)=\begin{cases} t_{a,1}(f_a,\varepsilon), & f_{a,1}\le f_a \le f_{a,2} \\ t_{a,2}(f_a,\varepsilon), & f_{a,2}\le f_a \le f_{a,3} \\ \quad\vdots & \\ t_{a,m-1}(f_a,\varepsilon), & f_{a,m-1}\le f_a \le f_{a,m} \end{cases}, \forall a\in A, \tag{5.12}$$

where $t_a(f_a,\varepsilon)$ is divided into m-1 segments and has m-2 inflection points, $f_{a,2}$, $f_{a,3}$, …, $f_{a,m-1}$. Generally, $f_{a,m}$ can be regarded as the capacity of arc a.

Recall that the sensitivity analysis of traffic equilibria can be formulated as a first-order approximation of original VI:

$$DS(\varepsilon^* \mid x^*)(\varepsilon')=\{x'\in K \mid r(\varepsilon',x')^T(x-x')\ge 0, x\in K\} \tag{5.13}$$

where

$$K=\{(h',f',d')\in R^\rho\times R^\alpha\times R^\omega \mid \Gamma^T h'=d';\Lambda^T h'=f';h'\in H'\}, \tag{5.14}$$

$$H'=\left\{h'\in R^\rho \left| \begin{array}{c} h'_p \text{ is free if } h^*_p>0 \\ h'_p\le 0 \text{ if } h^*_p=0 \text{ and } c_p(\varepsilon^*,h^*)=\mu^*_w \\ h'_p=0 \text{ if } h^*_p=0 \text{ and } c_p(\varepsilon^*,h^*)>\mu^*_w \\ p\in P_w, w\in W \end{array}\right.\right\}, \tag{5.15}$$

and

$$r(\varepsilon',x')=\begin{pmatrix} 0^\rho \\ \nabla_\varepsilon t(\varepsilon^*,v^*)\varepsilon'+\nabla_f t(\varepsilon^*,f^*)f' \\ -[\nabla_\varepsilon \xi(\varepsilon^*,d^*)\varepsilon'+\nabla_d \xi(\varepsilon^*,d^*)d'] \end{pmatrix}. \tag{5.16}$$

In this study, the original link cost function which is continuously differentiable is replaced by the piecewise linear cost function, $t_a(f_a,\varepsilon)$. The piecewise linear cost function is assumed continuous. Moreover, $t_a(f_a,\varepsilon)$ is non-differentiable at each inflection point but is continuously differentiable in elsewhere. In the sensitivity analysis of traffic equilibria with piecewise linear cost function problem, Eq. (5.16) has to be modified as

$$r(\varepsilon', x') = \begin{pmatrix} 0^{\rho} \\ \nabla_{\varepsilon} t(\varepsilon^*, v^*)\varepsilon' + \partial_f t(\varepsilon^*, f^*) f' \\ -\left[\nabla_{\varepsilon}\xi(\varepsilon^*, d^*)\varepsilon' + \nabla_d \xi(\varepsilon^*, d^*) d'\right] \end{pmatrix}, \tag{5.16}$$

where $\partial_f t(\cdot,\cdot)$ is the subgradient of the piecewise linear cost function w.r.t. arc flow.

In order to depict the congestion phenomenon, $t_a(f_a,\varepsilon)$ is assumed to be strictly increasing with respect to arc-flow v_a. Then the directional derivative of the equilibrium network flows exists according to Theorem 5.1, and it can be obtained by solving a convex quadratic programming with complementarity constraint as following.

QPCC model:

Given an ε^*, if f^* is an optimal solution of the perturbed traffic equilibria, then the directional derivative

$$f'(\varepsilon^*;\varepsilon') = \lim_{s\to 0^+} s^{-1}\left[f(\varepsilon^* + s\varepsilon') - f^*\right] \tag{5.17}$$

is the optimal solution of the following convex quadratic programming with complementarity constraint (QPCC):

$$\begin{aligned}
\underset{x'}{\text{Minimize}} \quad \phi'(x') = & \sum_{a\in A}\left[\varepsilon' \frac{\partial t_{a,i}(f_a^*,\varepsilon^*)}{\partial \varepsilon_a} f'_{a,1} + \varepsilon' \frac{\partial t_{a,i+1}(f_a^*,\varepsilon^*)}{\partial \varepsilon_a} f'_{a,2}\right] \\
& + \frac{1}{2}\sum_{a\in A}\left[\frac{\partial t_{a,i}(f_a^*,\varepsilon^*)}{\partial f_a}(f'_{a,1})^2 + \frac{\partial t_{a,i+1}(f_a^*,\varepsilon^*)}{\partial f_a}(f'_{a,2})^2\right]
\end{aligned}$$

$$\begin{aligned}
\text{subject to} \quad & \Lambda h' = 0,\ \ \Delta h' = f' \\
& f' = (-f'_{a,1} + f'_{a,2} : a \in A) \\
& 0 \le f'_{a,1} \perp f'_{a,2} \ge 0,\ \ \forall a \in A \\
& h' \in H',
\end{aligned}$$

CHAPTER 6

Numerical Examples

This chapter demonstrates some numerical examples for the proposed approaches of gradient-based sensitivity analysis method and directional derivative-based sensitivity analysis method, respectively. For gradient-based sensitivity analysis method, we perform the second-order sensitivity analysis method in two numerical examples. The first one is a signal control problem for a tiny network [31]. In this example, the exact solution is known, so that we can justify the accuracy and efficiency of the nonlinear approximation sensitivity-based algorithm (NLAA). The second one is a signal control problem of a simplified real network. We compare the efficiency between linear approximation sensitivity-base method (LAA) and NLAA in this example.

6.1 Numerical Examples of Second-Order Sensitivity Analysis Method

6.1.1 Example 1

The first example is chosen from Fisk [31]. The network topology is shown in Fig. 6. The set of OD pairs is {(1, 2), (3, 4)} and a signal exists at the intersection of arc 1 and 3. The cost functions used are

$$t_1 = \frac{f_1}{\varepsilon_1}, \qquad t_2 = 2 f_2, \qquad t_3 = \frac{2 f_3}{\varepsilon_3}, \tag{5.1}$$

where ε_a denotes the green time on arc a and the cycle time, $\varepsilon_1 + \varepsilon_3$, is equal to 20.

Additionally, the travel demand T_1 from node 1 to node 2 is 10, and the travel demand T_2 from node 3 to node 4 is 10. Table 2 lists the arc cost functions $t_a(f_a, \varepsilon_a)$ and the system objective function $Z(\varepsilon)$. Moreover, this example has an analytical optimal solution:

$$g_1 = 7.7306, \qquad g_3 = 12.2694, \qquad f_1 = 8.4533, \qquad f_2 = 1.5467.$$

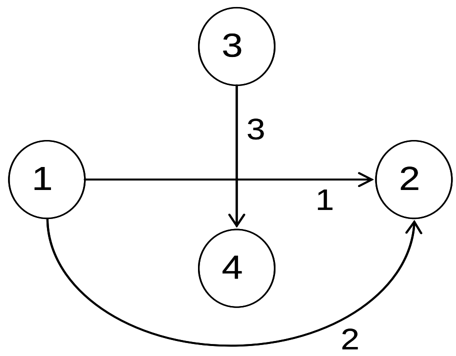

Fig. 6: The network topology of the example 1

Table 2: Arc cost functions and the system objective function in the example 1

$t_a(f_a,\varepsilon_a)=P_a+Q_a(f_a/\varepsilon_a)$		
$Z(\varepsilon)=\sum_a \left(t_a(f_a,\varepsilon_a)\cdot f_a\right)$		
Arc number	P_a	Q_a
1	2	1
2	0	2
3	0	2

In this example, ε_3 can be replaced by 20-ε_1, and ε_1 is the only perturbation parameter (control variable) should be considered. Therefore, $\|\varepsilon\|$ is equal to 1. Together with Eq. (11), (24) and (25), the first-order sensitivity with respect to ε_1 can be rewritten as

$$\begin{bmatrix}\nabla_\varepsilon f\\ \nabla_\varepsilon \mu\end{bmatrix}=U\cdot V$$

$$=\begin{bmatrix}\frac{\varepsilon_1}{1+2\varepsilon_1} & \frac{-\varepsilon_1}{1+2\varepsilon_1} & 0 & \frac{-2\varepsilon_1}{1+2\varepsilon_1} & 0\\ \frac{-\varepsilon_1}{1+2\varepsilon_1} & \frac{\varepsilon_1}{1+2\varepsilon_1} & 0 & \frac{-\varepsilon_1}{1+2\varepsilon_1} & 0\\ 0 & 0 & 0 & 0 & -1\\ \frac{2\varepsilon_1}{1+2\varepsilon_1} & \frac{1}{1+2\varepsilon_1} & 0 & \frac{2}{1+2\varepsilon_1} & 0\\ 0 & 0 & 1 & 0 & \frac{-1}{-20+\varepsilon_1}\end{bmatrix}\begin{bmatrix}\frac{f_1}{\varepsilon_1^2}\\ 0\\ \frac{-2f_3}{(20-\varepsilon_1)^2}\\ 0\\ 0\end{bmatrix}=\begin{bmatrix}\frac{f_1}{\varepsilon_1(1+2\varepsilon_1)}\\ \frac{-f_1}{\varepsilon_1(1+2\varepsilon_1)}\\ 0\\ \frac{2f_1}{\varepsilon_1(1+2\varepsilon_1)}\\ \frac{-2f_3}{(20-\varepsilon_1)^2}\end{bmatrix},$$

$$(5.2)$$

where

$$\nabla_\varepsilon f = \begin{bmatrix} \dfrac{f_1}{\varepsilon_1(1+2\varepsilon_1)} \\ \dfrac{-f_1}{\varepsilon_1(1+2\varepsilon_1)} \\ 0 \end{bmatrix}, \quad \nabla_\varepsilon \mu = \begin{bmatrix} \dfrac{2f_1}{\varepsilon_1(1+2\varepsilon_1)} \\ \dfrac{-2f_3}{(20-\varepsilon_1)^2} \end{bmatrix}. \tag{5.3}$$

In Eq. (5.3), the sensitivity information of arc flow represents the change of arc flow on arc a respectively when the control variable s_1 increases one unit. Since $f_1 \geq 0$, $\varepsilon_1 > 0$, the equilibrium flow on arc 1 will increase when s_1 increases one unit. In the meanwhile, the equilibrium flow on arc 2 will decrease. Because OD pair (3, 4) has only one path (arc 3), ε_1 will not affect the equilibrium flow on arc 3.

From Lemma 4.1, the second-order sensitivity with respect to control variable is

$$\begin{bmatrix} \nabla_\varepsilon^2 f \\ \nabla_\varepsilon^2 \mu \end{bmatrix} = \left(V^T \otimes I_{(\alpha+\alpha_2)}\right)\nabla_\varepsilon U + (I_k \otimes U)\nabla_\varepsilon V\,. \tag{5.4}$$

By Theorem 4.2, $\nabla_\varepsilon U$ can be derived by the chain rule for matrix functions as follows:

$$\begin{aligned} \nabla_\varepsilon U &= \left(\nabla_{f,\mu} U\right)\begin{bmatrix} \nabla_\varepsilon f \\ \nabla_\varepsilon \mu \end{bmatrix} + \left(\nabla_\varepsilon U\right) \\ &= \begin{bmatrix} \dfrac{\partial\, vec\, U}{\partial\, vec\ f} & \dfrac{\partial\, vec\, U}{\partial\, vec\ \mu} \end{bmatrix}\begin{bmatrix} \nabla_\varepsilon f \\ \nabla_\varepsilon \mu \end{bmatrix} + \begin{bmatrix} \dfrac{\partial\, vec\, U}{\partial\, vec\ \varepsilon_1} \end{bmatrix}. \end{aligned} \tag{5.5}$$

In this example, the matrix U is only dependent on s_1. Hence, $\nabla_{f,\mu} U = 0$ and Eq. (5.5) can be rewritten as

$$\nabla_{\varepsilon} U = \left[\frac{\partial\, vec\, U}{\partial\, vec\, \varepsilon_1} \right] = \begin{bmatrix} \frac{\partial U_1}{\partial \varepsilon_1} \\ \frac{\partial U_2}{\partial \varepsilon_1} \\ \frac{\partial U_3}{\partial \varepsilon_1} \\ \frac{\partial U_4}{\partial \varepsilon_1} \\ \frac{\partial U_5}{\partial \varepsilon_1} \end{bmatrix}, \tag{5.6}$$

where U_j represents the j-th column of matrix U, and

$$\frac{\partial U_1}{\partial \varepsilon_1} = \begin{bmatrix} \frac{1}{(1+2\varepsilon_1)^2} \\ \frac{-1}{(1+2\varepsilon_1)^2} \\ 0 \\ \frac{2}{(1+2\varepsilon_1)^2} \\ 0 \end{bmatrix}, \frac{\partial U_2}{\partial \varepsilon_1} = \begin{bmatrix} \frac{-1}{(1+2\varepsilon_1)^2} \\ \frac{1}{(1+2\varepsilon_1)^2} \\ 0 \\ \frac{-2}{(1+2\varepsilon_1)^2} \\ 0 \end{bmatrix}, \frac{\partial U_3}{\partial \varepsilon_1} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \frac{\partial U_4}{\partial \varepsilon_1} = \begin{bmatrix} \frac{-2}{(1+2\varepsilon_1)^2} \\ \frac{2}{(1+2\varepsilon_1)^2} \\ 0 \\ \frac{-4}{(1+2\varepsilon_1)^2} \\ 0 \end{bmatrix}, \frac{\partial U_5}{\partial \varepsilon_1} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \frac{2}{(-20+\varepsilon_1)^2} \end{bmatrix}. \tag{5.7}$$

Similarly, $\nabla_{\varepsilon} V$ can be derived as

$$\nabla_{\varepsilon} V = \left(\nabla_{f,\mu} V\right)\begin{bmatrix}\nabla_{\varepsilon} f \\ \nabla_{\varepsilon}\mu\end{bmatrix} + \left(\nabla_{\varepsilon} V\right)$$

$$= \begin{bmatrix}\dfrac{\partial\, vec\, V}{\partial\, vec\ f} & \dfrac{\partial\, vec\, V}{\partial\, vec\ \mu}\end{bmatrix}\begin{bmatrix}\nabla_{\varepsilon} f \\ \nabla_{\varepsilon}\mu\end{bmatrix} + \begin{bmatrix}\dfrac{\partial\, vec\, V}{\partial\, vec\ \varepsilon_1}\end{bmatrix}$$

$$= \begin{bmatrix}\dfrac{1}{\varepsilon_1^2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \dfrac{-2}{(20-\varepsilon_1)^2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0\end{bmatrix}\begin{bmatrix}\dfrac{f_1}{\varepsilon_1(1+2\varepsilon_1)} \\ \dfrac{-f_1}{\varepsilon_1(1+2\varepsilon_1)} \\ 0 \\ \dfrac{2f_1}{\varepsilon_1(1+2\varepsilon_1)} \\ \dfrac{-2f_3}{(20-\varepsilon_1)^2}\end{bmatrix} + \begin{bmatrix}\dfrac{-2f_1}{\varepsilon_1^3} \\ 0 \\ \dfrac{-4f_3}{(20-\varepsilon_1)^3} \\ 0 \\ 0\end{bmatrix} = \begin{bmatrix}\dfrac{f_1}{\varepsilon_1^3(1+2\varepsilon_1)} - \dfrac{2f_1}{\varepsilon_1^3} \\ 0 \\ \dfrac{-4f_3}{(20-\varepsilon_1)^3} \\ 0 \\ 0\end{bmatrix}. \tag{5.8}$$

According to Definition 4.1, Eq. (5.5), (5.6) and (5.7), Eq. (5.4) can be rewritten as

$$\begin{bmatrix}\nabla_{\varepsilon}^2 f \\ \nabla_{\varepsilon}^2\mu\end{bmatrix} = \left(V^T \otimes I_5\right)\nabla_{\varepsilon} U + \left(I_1 \otimes U\right)\nabla_{\varepsilon} V$$

$$= \begin{bmatrix}\dfrac{f_1}{\varepsilon_1^2}\cdot I_5 \\ 0\cdot I_5 \\ \dfrac{-2f_3}{(20-\varepsilon_1)^2}\cdot I_5 \\ 0\cdot I_5 \\ 0\cdot I_5\end{bmatrix}^T \begin{bmatrix}\dfrac{\partial U_1}{\partial\varepsilon_1} \\ \dfrac{\partial U_2}{\partial\varepsilon_1} \\ \dfrac{\partial U_3}{\partial\varepsilon_1} \\ \dfrac{\partial U_4}{\partial\varepsilon_1} \\ \dfrac{\partial U_5}{\partial\varepsilon_1}\end{bmatrix} + \left(1\cdot U\right)\begin{bmatrix}\dfrac{f_1}{\varepsilon_1^3(1+2\varepsilon_1)} - \dfrac{2f_1}{\varepsilon_1^3} \\ 0 \\ \dfrac{-4f_3}{(20-\varepsilon_1)^3} \\ 0 \\ 0\end{bmatrix} = \begin{bmatrix}\dfrac{-4f_1}{\varepsilon_1(1+2\varepsilon_1)^2} \\ \dfrac{4f_1}{\varepsilon_1(1+2\varepsilon_1)^2} \\ 0 \\ \dfrac{-8f_1}{\varepsilon_1(1+2\varepsilon_1)^2} \\ \dfrac{4f_3}{(-20+\varepsilon_1)^3}\end{bmatrix}. \tag{5.9}$$

Therefore,

$$\nabla_\varepsilon^2 f = \begin{bmatrix} \dfrac{-4f_1}{\varepsilon_1(1+2\varepsilon_1)^2} \\ \dfrac{4f_1}{\varepsilon_1(1+2\varepsilon_1)^2} \\ 0 \end{bmatrix}, \quad \nabla_\varepsilon^2 \mu = \begin{bmatrix} \dfrac{-8f_1}{\varepsilon_1(1+2\varepsilon_1)^2} \\ \dfrac{4f_3}{(-20+\varepsilon_1)^3} \end{bmatrix}. \tag{5.20}$$

Hence, we have the second-order sensitivity analysis information of the equilibrium arc-flows and equilibrium costs, respectively in Eq. (5.20). With the first-order and second-order sensitivity information in Eq. (5.8) and (5.20), respectively, the first differential and the second differential of equilibrium arc flow f_a can be obtained by Eq. (4.10) and (4.11). At iteration k,

$$\begin{aligned} B\varepsilon^{k+1} &= df(\varepsilon_0;\varepsilon) = \nabla_\varepsilon f\Big|_{f=f^k,\varepsilon_1=\varepsilon_1^k} \cdot \left(\varepsilon_1^{k+1}-\varepsilon_1^k\right) \\ &= \begin{bmatrix} \dfrac{f_1^k}{\varepsilon_1^k\left(1+2\varepsilon_1^k\right)} \\ \dfrac{-f_1^k}{\varepsilon_1^k\left(1+2\varepsilon_1^k\right)} \\ 0 \end{bmatrix} \cdot \left(\varepsilon_1^{k+1}-\varepsilon_1^k\right) = \begin{bmatrix} \dfrac{f_1^k\left(\varepsilon_1^{k+1}-\varepsilon_1^k\right)}{\varepsilon_1^k\left(1+2\varepsilon_1^k\right)} \\ \dfrac{-f_1^k\left(\varepsilon_1^{k+1}-\varepsilon_1^k\right)}{\varepsilon_1^k\left(1+2\varepsilon_1^k\right)} \\ 0 \end{bmatrix}, \end{aligned} \tag{5.21}$$

$$\begin{aligned} C\left(\varepsilon^{k+1}\right)^2 &= \frac{1}{2}d^2 f(\varepsilon_0;\varepsilon) = \frac{1}{2}\left(\left(\varepsilon_1^{k+1}-\varepsilon_1^k\right)^T \otimes I_3\right)\cdot \nabla_\varepsilon^2 f\Big|_{f=f^k,\varepsilon_1=\varepsilon_1^k} \cdot \left(\varepsilon_1^{k+1}-\varepsilon_1^k\right) \\ &= \frac{1}{2}\left(\left(\varepsilon_1^{k+1}-\varepsilon_1^k\right)\otimes \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right)\cdot \begin{bmatrix} \dfrac{-4f_1^k}{\varepsilon_1^k\left(1+2\varepsilon_1^k\right)^2} \\ \dfrac{4f_1^k}{\varepsilon_1^k\left(1+2\varepsilon_1^k\right)^2} \\ 0 \end{bmatrix} \cdot \left(\varepsilon_1^{k+1}-\varepsilon_1^k\right) = \begin{bmatrix} \dfrac{-4f_1^k\left(\varepsilon_1^{k+1}-\varepsilon_1^k\right)^2}{2\varepsilon_1^k\left(1+2\varepsilon_1^k\right)^2} \\ \dfrac{4f_1^k\left(\varepsilon_1^{k+1}-\varepsilon_1^k\right)^2}{2\varepsilon_1^k\left(1+2\varepsilon_1^k\right)^2} \\ 0 \end{bmatrix}. \end{aligned}$$

(5.22)

In this example, both LAA and NLAA are implemented in the MATLAB environment. Set $\delta = 0.001$ and the initial $\varepsilon_1 = 10$, Table 3 lists the computational results of LAA and NLAA approaches, and it shows that NLAA is more efficient than LAA.

Table 3: Computational results of LAA and NLAA in the example 1

	LAA			NLAA		
Iteration	f_1	ε_1	Z	f_1	ε_1	Z
1	8.4469	7.6365	47.2379	8.4513	7.7005	47.2358
2	8.4537	7.7367	47.2355	8.4533	7.7305	47.2355
3	8.4532	7.7302	47.2355	8.4533	7.7306	47.2355
4	8.4533	7.7304	47.2355			

6.1.2 Example 2

This example is a simplified real network which represents the afternoon rush hour traffic between the working area Hsinchu Science-based Industrial Park (HSIP) and residential area Jhubei city. The network topology follows Fig. 7. In this period, there is a huge amount of travel demand from HSIP (node 1) to Jhubei city (node 16). There are two parallel paths from HSIP to Jhubei city. One is freeway (arc 2-arc 4-arc 16), and the other is highway with 5 signal-controlled intersections (arc 1-arc 6-arc 8-arc 10-arc 12-arc 14). The objective of this problem is to find the optimal signal settings which minimize the system cost. Table 4 lists the origin-destination demand. The arc cost functions $t_a(f_a, \varepsilon_a)$ and the system objective function $Z(s)$ are listed in Table 5. For the signal-controlled intersections, the arcs entering the same intersection share the same cycle time and the minimum green time for each approach is 10 sec.

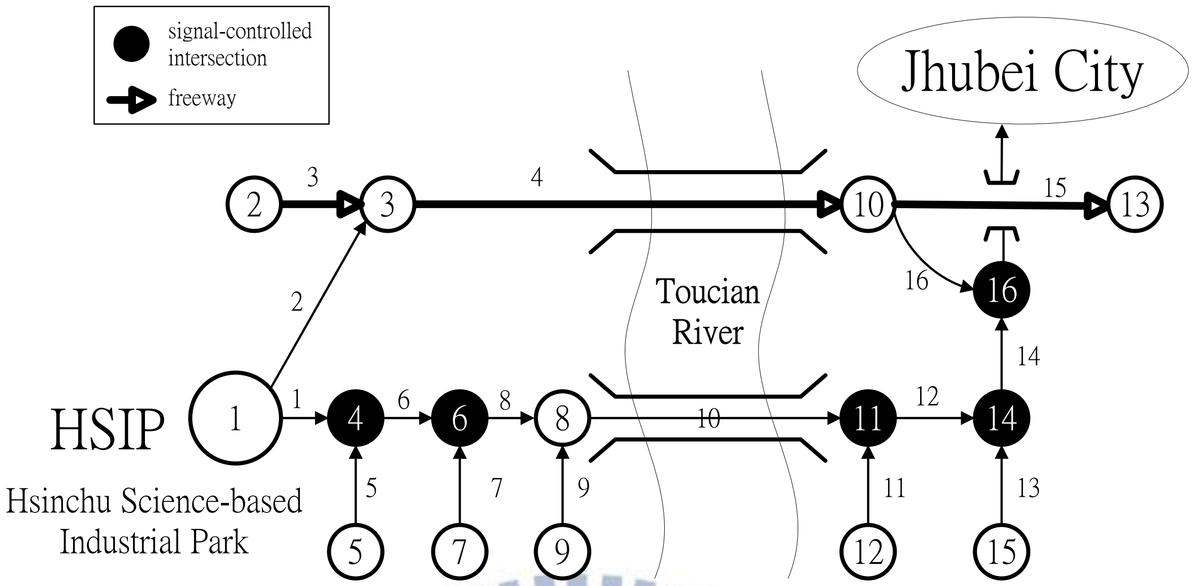

Fig. 7: The network topology of the example 2

Table 4: Origin-destination demand table in the example 2 (unit: veh/hr)

		Destination						
		4	**6**	**8**	**11**	**13**	**14**	**16**
Origin	**1**	50	275	475	400	1250	275	2250
	2	0	0	0	0	2550	0	1400
	5	0	150	250	200	0	150	250
	7	0	0	500	400	0	300	450
	9	0	0	0	325	0	225	350
	12	0	0	0	0	0	125	175
	15	0	0	0	0	0	0	900

Table 5: Arc cost functions and the system objective function in the example 2

Arc number	t_{0a} (min)	α_a	β_a	C_a(veh/min)	Cyc_a (sec)
Signalized arc cost function: $t_a(f_a,\varepsilon_a)=t_{0a}\left(1+\alpha_a\left(\frac{f_a}{C_a(\varepsilon_a/Cyc_a)}\right)^{\beta_a}\right)$					
Non-signalized arc cost function: $t_a(f_a)=t_{0a}\left(1+\alpha_a\left(\frac{f_a}{C_a}\right)^{\beta_a}\right)$					
System objective function: $Z(s)=\sum_a\left(t_a(f_a,\varepsilon_a)\cdot f_a\right)$					
1	1.8545	0.9200	3.5800	56.6667	300
2	0.8667	0.8600	4.3400	40.0000	-
3	1.8000	1.2700	3.9600	115.0000	-
4	2.2364	1.2700	3.9600	115.0000	-
5	0.2945	1.2100	2.3900	28.3333	300
6	0.1964	1.4200	2.3200	85.0000	300
7	0.3818	0.8600	4.3400	85.0000	300
8	1.0154	1.2700	3.9600	68.3333	-
9	1.0000	1.2100	2.3900	20.0000	-
10	1.0154	1.2700	3.9600	68.3333	180
11	0.3273	0.9200	3.5800	56.6667	180
12	0.9818	1.4200	2.3200	85.0000	150
13	0.6545	0.8600	4.3400	113.3333	150
14	1.2000	1.5000	2.4400	113.3333	150
15	3.8727	1.2700	3.9600	115.0000	-
16	0.4909	0.8600	4.3400	40.0000	150

In this example, we set $\delta = 0.1$ and the initial $\varepsilon_a = Cyc_a/2$ for each signalized arc. Table 6 lists the computational results of LAA and NLAA respectively. Two parallel paths from node 1 to node 16 (2-4-16 and 1-6-8-10-12-14) have the same equilibrium travel time 13.4069 min. Compared with LAA, Table 6 shows that NLAA only takes 12% iteration number to attain the same level of precision. Figure 3 shows the convergence curves of LAA and NLAA respectively. The convergence rate of LAA is slower than NLAA due to the zigzag effect. Compared with the example 1, NLAA has more improvement in the speed of convergence than in example 2. It may imply that NLAA is more efficient to deal with more nonlinear problems.

Table 6: Computational results of LAA and NLAA in the example 2

	LAA			NLAA		
Arc number	ε_a (sec)	f_a (veh/min)	t_a (min)	ε_a (sec)	f_a (veh/min)	t_a (min)
1	195.9103	31.4891	2.8118	195.8902	31.4909	2.8123
2	-	51.4275	2.8272	-	51.4258	2.8269
3	-	65.8333	2.0510	-	65.8333	2.0510
4	-	117.2609	5.3043	-	117.2591	5.3041
5	104.0897	16.6667	1.5529	104.1098	16.6667	1.5523
6	181.5056	47.3225	0.4263	181.5054	47.3242	0.4264
7	118.4944	27.5000	0.5199	118.4946	27.5000	0.5199
8	-	67.7391	2.2611	-	67.7409	2.2612
9	-	15.0000	1.6084	-	15.0000	1.6084
10	167.0451	62.3225	2.2192	167.0943	62.3242	2.2179
11	12.9549	5.0000	0.9517	12.9057	5.0000	0.9603
12	123.8799	45.2391	1.4849	123.8775	45.2409	1.4849
13	26.1201	15.0000	0.8256	26.1225	15.0000	0.8255
14	45.4083	42.3225	4.2042	45.4106	42.3242	4.2041
15	-	63.3333	4.3361	-	63.3333	4.3361
16	104.5917	53.9275	5.2760	104.5894	53.9258	5.2759
Iteration number	101			12		
Objective Value (Z)	2188.2886			2188.2235		

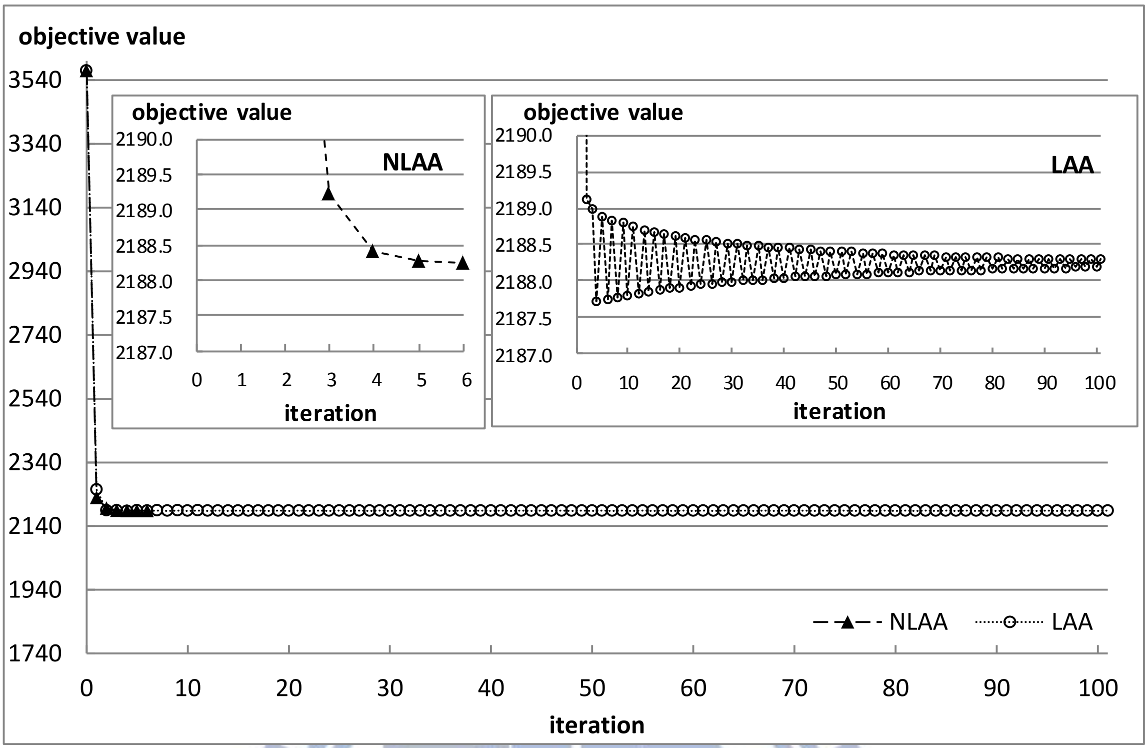

Fig. 8: The convergence curves of LAA and NLAA in the example 2

6.2 Numerical Examples of Sensitivity Analysis with Piecewise Linear Cost Functions

To demonstrate the applicability of the proposed method, we consider a simple network depicted in Fig. 9. The network contains two nodes, two links, on OD pair and two paths. There is a fixed demand of 5 units of flow for OD pair (1, 2). There are two paths corresponding to the OD pair denoted by h_1={1} and h_2={2}. Piecewise linear link cost functions are given by

$$t_1(f_1,\varepsilon_1)=\begin{cases}(1/2)f_1+9/2+\varepsilon_1, & 0\le f_1\le 2\\ f_1+7/2+\varepsilon_1, & 2\le f_1\le 4,\\ 2f_1-1/2+\varepsilon_1, & 4\le f_1\le 6\end{cases}\quad t_2(f_2)=\begin{cases}(1/2)f_2+4, & 0\le f_1\le 3\\ 2f_2-1/2, & 3\le f_1\le 6\end{cases}, \tag{5.23}$$

where the cost function of link 1 has a perturbation parameter ε_1. When ε_1=0, the equilibrium link flow solution is f^*=[2, 3]^T. With the change of ε_1, the equilibrium link flow solution at different ε_1 can be calculated and the trajectory of equilibrium solution with respect to ε_1 is shown in Fig. 10. From Fig. 10, the equilibrium solution is non-differentiable when ε_1=0.

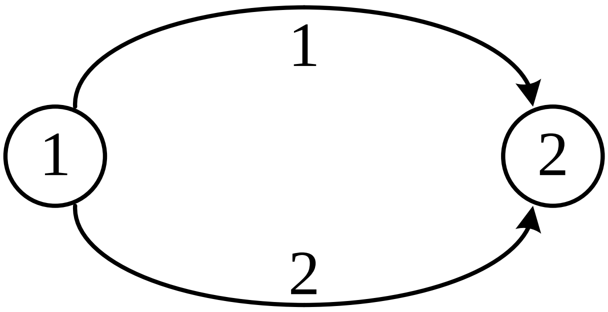

Fig. 9: The network topology of the example 3

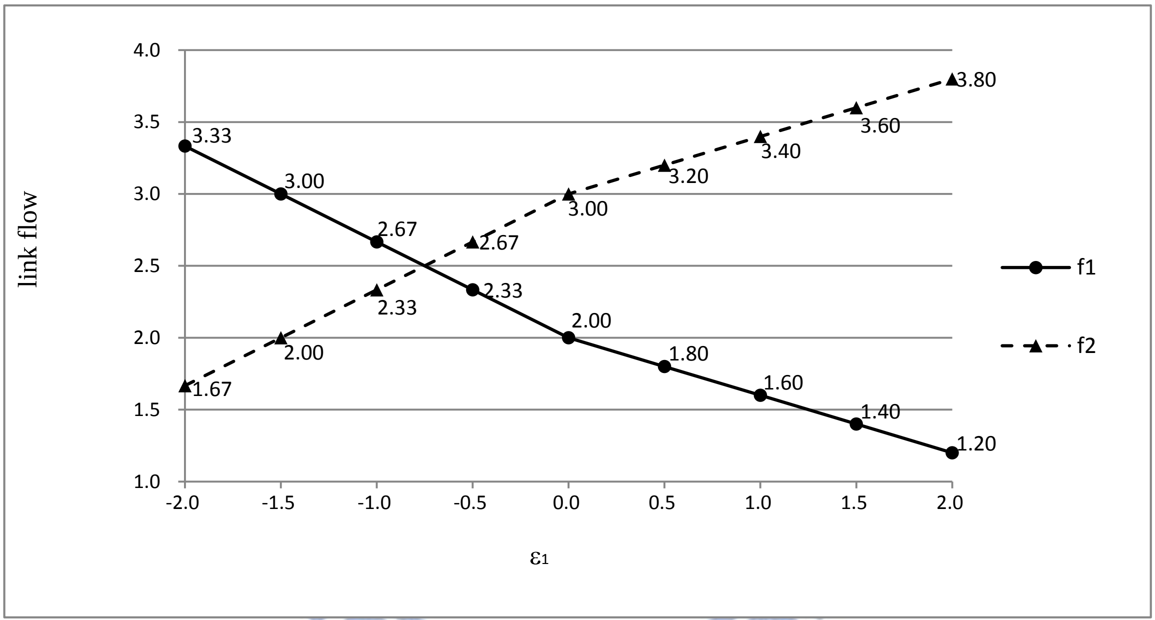

Fig. 10: The trajectory of the equilibrium solution with respect to different ε_1

When ε_1'= 1, the directional derivative of link flow is f'=[-2/5, 2/5]^T. When ε_1'= -1, the directional derivative of link flow is f'=[2/3, -2/3]^T. When ε_1'= 1, the proposed quadratic optimization problem with complementarity constraints is given by

$$\underset{f',h'}{\text{Minimize}} \quad \phi'(f') = 1\cdot f'_{1,1} + 1\cdot f'_{1,2} + \frac{1}{2}\left[\frac{1}{2}\left(f'_{1,1}\right)^2 + 1\cdot\left(f'_{1,2}\right)^2\right] + \frac{1}{2}\left[\frac{1}{2}\left(f'_{2,1}\right)^2 + 2\cdot\left(f'_{2,2}\right)^2\right]$$

$$\text{subject to} \quad \Lambda h' = 0,\ \Delta h' = f'$$

$$f'_1 = f'_{1,1} + f'_{1,2},\ \ f'_2 = f'_{2,1} + f'_{2,2}$$

$$f'_{1,1}\cdot f'_{1,2} = 0,\ \ f'_{2,1}\cdot f'_{2,2} = 0$$

$$f'_{1,1}, f'_{2,1} < 0, \quad f'_{1,2}, f'_{2,2} > 0 \tag{5.24}$$

Solving the optimization problem, the correct directional derivative of link flow, f'=[-2/5, 2/5]^T is obtained. When ε_1'= -1, the correct result, f'=[2/3, -2/3]^T is also obtained.

CHAPTER 7

Conclusions

Sensitivity analysis of equilibrium network flows is useful in various fields, such as bilevel network design problems, road pricing and origin-destination matrix estimation problems. By performing the sensitivity analysis of equilibrium network flows, one can predict the direction of the variation in the equilibrium pattern when parameters of cost and demand functions change. With sensitivity information, the linear approximation of the reaction function can be obtained and applied to solve Stackelberg game by a sensitivity analysis-based algorithm.

Due to the nonlinearity and convexity of the problem, the nonlinear approximation of the reaction function is expected to solve the problem more efficiently. Therefore, this research has established the theory of high-order sensitivity analysis of network equilibrium flows based on the gradient based sensitivity method and solved the problem with a nonlinear approximation of the reaction function.

Based on the first-order sensitivity formula and the matrix calculus, this study first presents the general form of the second-order sensitivity formula for equilibrium network flows. With the second-order sensitivity formula, the reaction function can be approximated more accurately by a nonlinear function. From HSIP to Jhubei city, a simplified real network example demonstrates the speed of convergence between LAA and NLAA. The NLAA has significant improvement in solving the equilibrium network signal control problem with complicated arc cost functions; in this example the NLAA only takes 6% iterations to attain the same level of precision.

This study focuses on the NLAA and a simplified delay formula is adopted to reflect the influence of traffic congestion. Practically, a traffic propagation model, such as TRANSYT model, should be included when solving the equilibrium network signal control problem. Since the derivatives of TRANSYT model have been obtained explicitly, it can be extended to second order derivatives and applied to NLAA in the future research.

Compared with LAA, the number of multiplications for matrix multiplication is greatly increasing in NLAA due to the Kronecker product operation. NLAA has polynomial complexity with the network size and the number of perturbation parameters because of the property of the Kronecker product. In this research, the computation time of NLAA is two times larger than LAA. There still has opportunity to improve the computing efficiency through adopting effective Kronecker-product algorithms.

This research has also extended the applicability of directional derivative-based sensitivity analysis method. To generalize the directional derivative-based sensitivity analysis, the continuous differentiability assumption of the arc cost function is relaxed by introducing piecewise linear arc cost functions instead. Based on the original directional derivative-based method, a quadratic optimization problem with complementarity constraints has been proposed to solve the sensitivity information of the network equilibrium problem with piecewise linear arc cost functions.

References

[1] Aashtiani, H.Z. and Magnanti T.L., “Equilibria on a congested transportation network,” SIAM J. Algebraic and Discrete Methods 2, pp. 213-226, 1981.

[2] Abdulaal, M. and L.J. LeBlanc, “Continuous equilibrium network design models,” Transportation Research 13B, pp. 19-32, 1979.

[3] Allsop, R.E., “Some possibilities for using traffic control to influence trip distribution and route choice,” Proceedings of the 6th International Symposium on Transportation and Traffic Theory, pp. 345-373, 1974.

[4] Allsop, R.E., “Evolving applications of mathematical optimization in design and operation of individual signal-controlled road junctions,” Mathematics in Transport Planning and Control (J.D. Griffiths ed.), Clarendon Press, Oxford, U.K., pp. 1–24, 1992.

[5] Baskan, O., “Determining Optimal Link Capacity Expansions in Road Networks Using Cuckoo Search Algorithm with Lévy Flights,” Journal of Applied Mathematics, Vol. 2013, Article ID 718015, 11 pages, 2013.

[6] Beckmann, M., C. McGuire and C. Winsten, Studies in the Economics of Transportation, CT: Yale University Press, 1956.

[7] Bell, M.G.H. and Y. Iida, Transportation network analysis. Wiley & Sons, New Jersey, 1997.

[8] Braess, D., “Uber ein paradoxon aus der Verkehrsplanung,” Mathematical Methods of Operations Research 12, pp. 258-268, 1968.

[9] Ceylan, H. and M.G.H. Bell, “Reserve capacity for a road network under optimized fixed time traffic signal control,” Journal of Intelligent Transportation Systems, Vol. 8, pp. 87-99, 2004.

[10] Chao, G.S. and T.L. Friesz, “Spacial Price Equilibrium Sensitivity Analysis,” Transportation Research 18B, pp.423-440, 1984.

[11] Chiou, S.W., “TRANSYT derivatives for area traffic control optimisation with network equilibrium flows,” Transportation Research Part B, Vol. 37, pp. 263–290, 2003.

[12] Chiou, S.W., “Bilevel programming for the continuous transport network design problem,” Transportation Research Part 39B, pp. 361–383, 2005.

[13] Chiou, S.W., “A generalized iterative scheme for network design problem,” Applied Mathematics and Computation, Vol. 188, pp. 1115–1123, 2007.

[14] Chiou, S.W., “A subgradient optimization model for continuous road network design problem,” Applied Mathematical Modelling, Vol. 33, pp. 1386–1396, 2009.

[15] Cho, H.J. and S.C. Lo, “Solving bilevel network design problem using a linear reaction function without nondegeneracy assumption,” Transportation Research Record, 1667, pp. 96–106, 1999.

[16] Cho, H.J., T.E. Smith and T.L. Friesz, “A reduction method for local sensitivity analyses of network equilibrium arc-flows,” Transportation Research 34B, pp. 31-51, 2000.
[17] Chung, B.D., H.J. Cho, T.L. Friesz, H. Huang and T. Yao, “Sensitivity analysis of user equilibrium flows revisited,” Networks and Spatial Economics, Vol. 12, pp. 183-207, 2014.
[18] Clark, S.D. and D.P. Watling, “Probit-based sensitivity analysis for general traffic networks,” Transportation Research Record, No. 1733, pp.88-95, 2000.
[19] Clark, S.D. and D.P. Watling, “Sensitivity analysis of the probit-based stochastic user equilibrium assignment model,” Transportation Research 36B, pp.617-635, 2002.
[20] Dafermos, S.C., “Traffic equilibrium and variational inequalities,” Transportation Science 14, pp. 42-54, 1980.
[21] Dafermos, S.C., “Sensitivity analysis in variational inequalities,” Mathematics of Operations Research, Vol. 13, pp.421-434, 1988.
[22] Dafermos, S.C. and A. Nagurney, “Sensitivity analysis for the general spatial economic equilibrium problem,” Operations Research 32, pp. 1069-1086, 1984a.
[23] Dafermos, S.C., Nagurney, A., “Sensitivity analysis for the asymmetric network equilibrium problem,” Mathematical Programming 28, pp. 174-184, 1984b.
[24] Dafermos, S.C. and A. Nagurney, “On some traffic equilibrium theory paradoxes,” Transportation Research 18B (2), pp. 101-110, 1984c.
[25] Dantzig, G.B., R.P. Harvey, Z.F. Lansdowne, D.W. Robinson and S.F. Maier, “Formulating and solving the network design problem by decomposition,” Transportation Research, 13B, pp. 5–17, 1979.
[26] De Silva, A.H., “Sensitivity formulas for nonlinear factorable programming and their application to the solution of an implicitly defined optimization model of US crude oil production,” Dissertation, George Washington University, Washington, D.C., 1978.
[27] Fang, S.C., “Fixed point models for equilibrium problems on transportation networks,” 1979, unpublished paper.
[28] Fiacco. A.V., “Sensitivity Analysis for Nonlinear Programming Using Penalty Methods,” Mathematical Programming, Vol. 10, pp. 287-311, 1976.
[29] Fiacco, A.V., Introduction to sensitivity and stability analysis in nonlinear programming, New York: Academic Press, 1983.
[30] Fiacco, A.V. and G.P. McCormick, Nonlinear Programming: Sequential Unconstrained Minimization Techniques, John Wiley & Sons, New York, 1968.
[31] Fisk, C.S., “Game theory and transportation systems modeling,” Transportation Research 18B, pp. 301-313, 1984.
[32] Fisk, C. and D. Boyce, “Alternative variational inequality formulation of the network equilibrium-travel choice problem,” Transportation Science 17, pp. 454-463, 1983.
[33] Friesz, T.L., “Transportation network equilibrium, design and aggregation: key

developments and research opportunities," Transportation Research 19A, pp. 413-427, 1985.
[34] Friesz, T.L. and P.T. Harker, "Properties of the iterative optimization-equilibrium algorithm," Civil Engineering System, Vol. 2, pp. 142-154, 1985.
[35] Friesz, T.L., R.L. Tobin, H.-J. Cho and N.J. Mehta, "Sensitivity analysis based heuristic algorithms for mathematical programs with variational inequality constraints," Mathematical Programming 48, pp. 265–284, 1990.
[36] Gartner, N.H., S.B. Gershwin, J.D.C. Little and P. Ross, "Pilot study of computer-based urban traffic management," Transportation Research Part B, vol. 14, pp. 203-217, 1980.
[37] Josefsson, M. and M. Patriksson, "Sensitivity analysis of separable traffic equilibrium equilibria with application to bilevel optimization in network design," Transportation Research 41B, pp.4-31, 2007.
[38] Kinderlehrer, D. and G. Stampacchia, An Introduction to Variational Inequalities and Applications Academic Press, New York, 1980.
[39] Kolstad, C.D., "A review of the literature on bilevel mathematical programming," Los Alamos National Laboratory Report, LA-10284-MS, Los Alamos, 1985.
[40] Kyparisis, J., "Sensitivity analysis framework for variational inequalities," Mathematical Programming, Vol. 38, pp.203-213, 1987.
[41] LeBlanc, L.J., "An algorithm for the discrete network design problem," Transportation Science 9, pp. 183-199, 1975.
[42] Lu, S., "Sensitivity of Static Traffic User Equilibrium with Perturbations in Arc Cost Function and Travel Demand," Transportation Science 42, pp. 105-123, 2008.
[43] Magnus, J.R. and H. Neudecker, "Matrix differential calculus with applications to simple, Hadamard, and Kronecker products," Journal of Mathematical Psychology 29, pp. 474-492, 1985.
[44] Magnus, J.R. and H. Neudecker, Matrix Differential Calculus with Applications in Statistics and Econometrics, Wiley, Chichester, UK, 1999.
[45] Mao, W., F. Qin, Y. Hu, and Z. Tan, "Joint Implementation of Signal Control and Congestion Pricing in Transportation Network," Journal of Applied Mathematics, Vol. 2013, Article ID 315752, 6 pages, 2013.
[46] Marcott, P., "Network optimization with continuous control parameters," Transportation Science 17, pp. 181-197, 1983.
[47] Marcotte, P. "Network design problem with congestion effects: A case of bilevel programming," Mathematical Programming, 34, pp. 142-146, 1986.
[48] Marcotte, P. and M. Patriksson, Handbooks in operations research and management science: transportation, In: Barnhart C, Laporte G (ed) Traffic equilibrium, Elsevier, Netherlands, pp. 623-714, 2007.
[49] Outrata, J.V., "On a special class of mathematical programs with equilibrium

constraints," Recent Advances in Optimization: Proceedings of the Eighth French-German Conference on Optimization, Vol. 452, Lecture Notes in Economics and Mathematical Systems, Springer-Verlag, Berlin, pp. 246–260, 1997.

[50] Pang, J. S. and C.S. Yu, "Linearized simplical decomposition methods for computing traffic equilibrium networks," Networks 14(3), pp. 427-438, 1984.

[51] Patriksson, M., "Sensitivity analysis of traffic equilibria," Transportation Science, Vol. 38, pp.258-281, 2004.

[52] Patriksson, M. and R.T. Rockafellar, "Sensitivity analysis of aggregated variational inequality problems, with application to traffic equilibria," Transportation Science, Vol. 37, pp.56-68, 2003.

[53] Qui, Y. and T.L. Magnanti, "Sensitivity analysis for variational inequalities defined on polyhedral sets," Mathematics of Operations Research, Vol. 14, pp.410-432, 1989.

[54] Silcock, J.P. and A.P. Sang, "SIGSIGN: A phase-based optimisation program for individual signal-controlled junction," Traffic Engineering and Control, vol. 31, pp. 291–298, 1990.

[55] Smith, M.J., "The Existence, Uniqueness and Stability of Traffic Equilibria," Transportation Research 13B, pp. 295-304, 1979.

[56] Suwansirikul, C., T.L. Friesz, and R.L. Tobin "Equilibrium Decomposed Optimization: A Heuristic for the Continuous Equilibrium Network Design Problem," Transportation Science, Vol. 21, pp. 254-263, 1987.

[57] Tam, M.L., and W.H.K. Lam, "Maximum car ownership under constraints of road capacity and parking space," Transportation Research 34A, pp.145–170, 2000.

[58] Tang, K. and H. Nakamura, "Investigating operational benefits of group-based signal control in Japan through a stochastic approach," International Journal of Intelligent Transportation Systems Research, Vol. 9, pp. 71-81, 2011.

[59] Tobin, R.L., "Sensitivity analysis for variational inequalities," Journal of Optimization Theory and Applications, Vol. 48, pp.191-204, 1986.

[60] Tobin, R.L., "Sensitivity analysis for general spatial price equilibrium," Journal of Regional Science 27, pp. 77-102, 1987.

[61] Tobin, R.L., and T.L. Friesz, "Sensitivity analysis for equilibrium network flow," Transportation Science, Vol. 22, pp.242-250, 1988.

[62] Wardrop, J., "Some theoretical aspects of road traffic research," Proceedings of the Institute of Civil Engineers part II, pp. 325-378, 1952.

[63] Wong, S.C., "Group-based optimization of signal timings using the TRANSYT traffic model," Transportation Research Part 30B, pp. 217–244, 1996.

[64] Wong, S.C. and C. Yang, "An iterative group-based signal optimization scheme for traffic equilibrium networks," Journal of Advanced Transportation, Vol. 33, pp. 201-217, 1999.

[65] Yang, H. and M.G.H. Bell, “Traffic restraint, road pricing and network equilibrium,” Transportation Research 31B, pp. 303-314, 1997.

[66] Yang, H, “Sensitivity analysis for the elastic-demand network equilibrium problem with applications,” Transportation Research 31B, pp. 55-70, 1997.

[67] Yang, H. and M.G.H. Bell, “Sensitivity analysis of network traffic equilibria revisited: the corrected approach,” Proceedings of the 4th IMA International Conference on Mathematics in Transport, pp. 373-396, 2007.

[68] Yang, C. and A. Chen, “Sensitivity analysis of the combined travel demand model with applications,” European Journal of Operational Research, Vol. 198, pp.909-921, 2009.

[69] Yang, H., Q. Meng and M.G.H. Bell, “Simultaneous estimation of the origin destination matrices and travel-cost coefficient for congested networks in a stochastic equilibrium,” Transportation Science, Vol. 35, pp.107-123, 2001.

[70] Yen, N.D., “Lipschitz continuity of solutions of variational inequalities with a parametric polyhedral constraint,” Mathematics of Operations Research 20, pp. 695–708, 1995.

[71] Ying, J.Q. and T. Miyagi, “Sensitivity analysis for stochastic user equilibrium network flows - A dual approach,” Transportation Science, Vol. 35, pp.124-133, 2001.

[72] Ying, J.Q. and H. Yang, “Sensitivity analysis of stochastic user equilibrium flow in a bi-modal network with application to optimal pricing,” Transportation Research 39B, pp.769-795, 2005.