Linear Generalized Nash Equilibrium Problems

Zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaften

Dr. rer. pol.

bei der Fakultät für Wirtschaftswissenschaften
des Karlsruher Instituts für Technologie (KIT)

genehmigte

DISSERTATION

von

M. Sc. Nathan Georg Sudermann-Merx

Tag der mündlichen Prüfung:	13. April 2016
Referent:	Prof. Dr. Oliver Stein
Korreferent:	Prof. Dr. Ralf Werner

List of Figures

List of Tables

Abbreviations

Game theory

NEP	(Classical) Nash equilibrium problem
AGNEP	Affine generalized Nash equilibrium problem
GNEP	Generalized Nash equilibrium problem
LGNEP	Linear generalized Nash equilibrium problem
ETP	Extended transportation problem

Regularity conditions

CCC	Collective cone condition
PCC	Player cone condition
LICQ	Linear independence constraint qualification
CLICQ	Collective linear independence constraint qualification
PLICQ	Player linear independence constraint qualification
CSMFC	Collective strict Mangasarian Fromovitz condition
PSMFC	Player strict Mangasarian Fromovitz condition

Algorithms

PRA	Potential reduction algorithm
PRALP	A combination of PRA and a local LP-Newton method
PSM	Projected subgradient method
RGS	Robust gradient sampling

Further abbreviations

KKT	Karush-Kuhn-Tucker
LP	Linear program
s.t.	Subject to

Notation

Mathematical Notation

$\mathbb{N}$	Natural numbers $1, 2, 3 \ldots$
$\mathbb{R}^n$	n-dimensional Euclidean space
$\mathbb{R}^n_+$	Nonnegative orthant of $\mathbb{R}^n$
$\mathbb{R}^n_{++}$	Positive orthant of $\mathbb{R}^n$
$\mathbb{R}^{M\times N}_R$	Set of $(M \times N)$-matrices of rank R
$\langle x, y\rangle$	Canonical inner product $\langle x, y\rangle = x^T y$ in $\mathbb{R}^n$
$x \circ y$	Hadamard product of the vectors x and y
$\lvert A\rvert$	Cardinality of the set A
A^c	Complement of the set A
$\operatorname{bd} A$	Topological boundary of the set A
$\operatorname{int} A$	Topological interior of the set A
$\operatorname{conv}(A)$	Convex hull of the set A
$\operatorname{vert} A$	Set of vertices of the polyhedral set A
Δ_N	$(N-1)$-dimensional standard simplex
$A \setminus B$	Set difference of A and B
$[a, b]$	Line segment between the vectors a and b
$B(x, \delta)$	Closed ball with center x and radius δ
$\mathcal{L}(x, W)$	Outer linearization cone at W in x
$P_W(x)$	Projection of x onto the set W with respect to the Euclidean norm
$\operatorname{dist}(x, A)$	Distance of the point x to the set A
$\operatorname{rank} B$	Rank of the matrix B
$\operatorname{diag}(w)$	Diagonal matrix with the vector w on its diagonal
$\operatorname{dom} f$	Domain of the function $f : \mathbb{R}^n \to \mathbb{R}$
∇f	Gradient of the function $f : \mathbb{R}^n \to \mathbb{R}$
JF	Jacobian of the mapping $F : \mathbb{R}^n \to \mathbb{R}^p$
$J_x F(x, \lambda)$	Jacobian of the mapping $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$ with respect to x
$f'(x, d)$	One-sided directional derivative of f at x along d
$\partial f(x)$	Clarke subdifferential of f at x
$F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$	Set-valued mapping F

General game theory

N	Number of players
n_ν	Dimension of player ν's strategy space
m_ν	Number of player ν's constraints
n	Dimension of the common unfolded strategy space $n = n_1 + \ldots + n_N$
m	Total number of constraints $m = m_1 + \ldots + m_N$
x^ν	Decision variable of player ν
$x^{-\nu}$	Decisions of all players except player ν
$Q_\nu(x^{-\nu})$	Optimization problem of player ν for a given $x^{-\nu}$
$S_\nu(x^{-\nu})$	Set of optimal points of $Q_\nu(x^{-\nu})$
$KKT_\nu(x^{-\nu})$	Set of KKT multipliers of $Q_\nu(x^{-\nu})$
$D_\nu(x^{-\nu})$	Dual optimization problem of player ν for a given $x^{-\nu}$
Z_ν	Feasible set of player ν's dual problem
O_ν	Set of vertices of Z_ν
$O_\nu(x^{-\nu})$	Index set of active selection functions of player ν at $x^{-\nu}$
$\theta_\nu(\cdot, x^{-\nu})$	Objective function of player ν for fixed $x^{-\nu}$
$X_\nu(x^{-\nu})$	Strategy set of player ν for a given $x^{-\nu}$
$\operatorname{dom} X_\nu$	Domain of the set-valued mapping X_ν
c^ν	Cost vector of player ν
A^ν, $B^{-\nu}$	Matrices defining the constraints of player ν in Part I
$A^{\nu\mu}$	Matrices defining the constraints of player ν in Parts II and III
b^ν	Right hand side of player ν's constraints
φ_ν	Optimal value function of player ν
$\widehat{\varphi}_\nu$	Real-valued global extension of φ_ν
ℓ_{λ^ν}	Selection function of the gap function V with index λ^ν
$I_0^\nu(x)$	Set of active indices of player ν in x
$I_+^\nu(\lambda^\nu)$	Index set of positive multipliers of player ν at λ^ν
W	Common unfolded strategy set of all players
V	Gap function
M	Domain of the gap function V
$\widehat{V}$	Real-valued global extension of V
Ψ	Nikaido-Isoda function
$H(x, \lambda, w)$	A function that is used to reformulate an LGNEP as system of constrained equations
ψ	Potential function in PRA
$P(x)$	Penalty function of the set W
S	Set of all (generalized) Nash equilibria

Extended transportation problem

R	Number of manufacturers
T	Number of consumers
S_r	Supply of manufacturer r
D_t	Demand of consumer t
c_{rt}^{ν}	Cost for the transportation of one unit from manufacturer r to consumer t by forwarder ν
C^{ν}	Maximal transportation costs of forwarder ν
Y	Set of efficiently computable Nash equilibria

Basic economic market model

K	Number of price categories
C^{ν}	Contingent of player ν
p_k^{ν}	Price of player ν in price category k
D_k	Total demand in price category k

Contents

Chapter 1

Introduction

1.1 Existing Literature

In 1951, John Nash investigated the problem of finding equilibria in situations where several competing players try to optimize their objective functions over strategy sets that are independent of the decisions of the remaining players (cf. [53]). We will refer to this situation as the (classical) *Nash equilibrium problem* (NEP). Only a few years later, Kenneth Arrow and Gérard Debreu extended the setting towards a model with coupled strategy sets, that is, strategy sets that may depend on the decisions of the remaining players (cf. [1, 15]). Coupled strategy sets arise in a very natural way if, for instance, players share at least one constraint which could be a common budget or commonly used infrastructure. A Nash equilibrium problem with coupled constraints is called a *generalized Nash equilibrium problem* (GNEP). Despite the early introduction of GNEPs it took over 40 years until GNEPs attracted attention in the operations research community. During this time, that is, until the mid nineties mainly existence results for (generalized) Nash equilibria where available and the numerical computation of equilibria was less developed (cf. [27]). However, the field of operations research had a deep impact on game theory and provided powerful numerical methods for the computation of Nash equilibria. Excellent overviews of theoretical and numerical results as well as numerous applications of GNEPs are given in [27] and [32].

Typically, GNEPs are studied under some general convexity assumptions. In the present work, we assume linearity of the cost functions and the constraints of all players. This special GNEP, the so-called *linear generalized Nash equilibrium problem* (LGNEP), enables a treatment that heavily exploits the linear structure and is therefore not available in the convex nonlin-

ear setting. Moreover, some techniques from nonlinear GNEPs require the nonsingularity of some second order derivatives and are therefore not applicable for *linear* generalized Nash equilibrium problems. LGNEPs arise very naturally in many situations as we shall see in this work. To the best of our knowledge, the articles [26] and [67] of the author contain the first systematic treatments of LGNEPs. Therefore, this work fills a remaining gap in literature and provides a systematic approach of tackling linear Nash equilibrium problems. The closest setting treated in literature considers so-called affine generalized Nash equilibrium problems (AGNEPs) where GNEPs with linear constraints and quadratic cost functions are considered which are convex in the player variables. This setting was introduced explicitly in [62] and also investigated in [21, 23].

This thesis is based on the articles [26] and [67] of the author and refers also to [66] where the author does not consider LGNEPs but smoothness properties of optimal value functions that are closely related to GNEPs.

1.2 Setting and Blanket Assumptions

In a generalized Nash equilibrium problem (GNEP), player ν, $\nu \in \{1, \ldots, N\}$, $N \in \mathbb{N}$, controls a decision vector $x^\nu \in \mathbb{R}^{n_\nu}$, $n_\nu \in \mathbb{N}$, and wishes to solve his optimization problem

$$Q_\nu(x^{-\nu}): \qquad \min_{x^\nu \in \mathbb{R}^{n_\nu}} \theta_\nu(x^\nu, x^{-\nu}) \quad \text{s.t.} \quad x^\nu \in X_\nu(x^{-\nu})$$

which depends on the decisions of the remaining players

$$x^{-\nu} := \begin{pmatrix} x^1 \\ \vdots \\ x^{\nu-1} \\ x^{\nu+1} \\ \vdots \\ x^N \end{pmatrix} \in \mathbb{R}^{n-n_\nu},$$

with $n := n_1 + \ldots + n_N$ and is therefore a parametric optimization problem (see [4, 6, 64] for an introduction to parametric optimization). For fixed $x^{-\nu} \in \mathbb{R}^{n-n_\nu}$, we call $\theta_\nu(\cdot, x^{-\nu}) : \mathbb{R}^{n_\nu} \to \mathbb{R}$ the *cost function* of player ν and denote his *strategy set* by $X_\nu(x^{-\nu}) \subseteq \mathbb{R}^{n_\nu}$. Note that both, the cost function and the strategy set of player ν depend on the decisions of the remaining players $x^{-\nu}$ whereas in a classical Nash equilibrium problem (NEP), we have $X_\nu(x^{-\nu}) \equiv X_\nu$ for some set $X_\nu \in \mathbb{R}^{n_\nu}$. A *generalized Nash equilibrium*

problem (GNEP) consists of finding some

$$\bar{x} \;:=\; (\bar{x}^\nu, \bar{x}^{-\nu}) \;:=\; \begin{pmatrix} \bar{x}^1 \\ \vdots \\ \bar{x}^N \end{pmatrix} \in \mathbb{R}^n,$$

such that $\bar{x}^\nu$ is a global minimal point of $Q_\nu(\bar{x}^{-\nu})$ for each $\nu \in \{1, \ldots, N\}$. Such a point is called a generalized Nash equilibrium and, obviously, in a generalized Nash equilibrium no player has an incentive to deviate unilaterally from his chosen strategy $\bar{x}^\nu$. Since Nash equilibria will play a crucial role in this work we restate this definition in a more explicit way.

Definition 1.2.1 *A vector $\bar{x} \in \mathbb{R}^n$ with $\bar{x}^\nu \in X_\nu(\bar{x}^{-\nu})$ and*

$$\theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}) \;\leq\; \theta_\nu(x^\nu, \bar{x}^{-\nu})$$

for all $x^\nu \in X_\nu(\bar{x}^{-\nu})$ and each $\nu \in \{1, \ldots, N\}$ is called a (generalized) Nash equilibrium.

In the following, we assume linearity of all functions in the whole vector x, that is, for each $\nu \in \{1, \ldots, N\}$ there exist an integer $m_\nu \in \mathbb{N}$, matrices $A^\nu \in \mathbb{R}^{m_\nu \times n_\nu}$, $B^\nu \in \mathbb{R}^{m_\nu \times (n-n_\nu)}$ and vectors $c^\nu \in \mathbb{R}^{n_\nu}$, $b^\nu \in \mathbb{R}^{m_\nu}$, such that player ν's minimization problem can be expressed as

$$Q_\nu(x^{-\nu}): \qquad \min_{x^\nu \in \mathbb{R}^{n_\nu}} \langle c^\nu, x^\nu \rangle \quad \text{s.t.} \quad A^\nu x^\nu + B^\nu x^{-\nu} \leq b^\nu,$$

where we denoted the canonical inner product by $\langle \cdot, \cdot \rangle$, that is, we have $\langle x, y \rangle := x^T y$. Hence, throughout this paper we shall have

$$\theta_\nu(x^\nu, x^{-\nu}) \;=\; \langle c^\nu, x^\nu \rangle$$

and

$$X_\nu(x^{-\nu}) \;=\; \{x^\nu \in \mathbb{R}^{n_\nu} :\; A^\nu x^\nu \leq b^\nu - B^\nu x^{-\nu}\}$$

for all $\nu \in \{1, \ldots, N\}$.

Remark 1.2.2 *The appearance of equality constraints in player ν's optimization problem does not lead to any qualitative changes in our results. This is obvious since it is possible to reformulate linear equations as linear inequalities without destroying the Abadie constraint qualification. However, in Part III of this work we shall introduce a new problem format in order to examine equality constraints in a more elegant way. This is not necessary in the prior parts of this work and therefore omitted.*

Remark 1.2.3 *All constraints of player ν where the corresponding rows in B^ν vanish are so-called* private constraints*, that is, constraints that do not depend on the decisions of the remaining players. In particular, for $B^\nu = 0$ for all $\nu \in \{1, \ldots, N\}$ we arrive at a NEP.*

Remark 1.2.4 *The representation of player ν's constraints with two matrices A^ν and B^ν is very handy but in some situations it will be advantageous to reformulate the constraints of player ν in order to obtain a more specific representation. Therefore, if necessary, we shall introduce suitable matrices $A^{\nu\mu}$, $\nu, \mu \in \{1, \ldots, N\}$, and rewrite the constraints of player ν as*

$$A^{\nu\nu}x^\nu + \sum_{\mu \neq \nu} A^{\nu\mu}x^\mu \leq b^\nu$$

for all $\nu \in \{1, \ldots, N\}$.

Remark 1.2.5 *Note that it would add no generality to augment player ν's cost function θ_ν by a term $\langle e^\nu, x^{-\nu}\rangle$ or even by a nonlinear function $f(x^{-\nu})$ because this would only affect its optimal value but not its optimal point and therefore not the position of the Nash equilibrium.*

Throughout this work we assume the existence of at least one Nash equilibrium. In our main application, the extended transportation problem (ETP) (cf. Part III), this assumption is always fulfilled.

Assumption 1.2.6 *There exists at least one (generalized) Nash equilibrium.*

Usually, sufficient conditions for the existence of Nash equilibria are derived by some boundedness and continuity conditions (cf. [1, 27]). Furthermore, Proposition 1.5.1 characterizes the existence of Nash equilibria with means of an reformulation as an optimization problem.

For each $\nu \in \{1, \ldots, N\}$ we denote by

$$\operatorname{dom} X_\nu \;:=\; \{x^{-\nu} \in \mathbb{R}^{n-n_\nu} :\; X_\nu(x^{-\nu}) \neq \emptyset\}$$

the (effective) domain of the set-valued mapping $X_\nu : \mathbb{R}^{n-n_\nu} \rightrightarrows \mathbb{R}^\nu$ which assigns to each vector $x^{-\nu} \in \mathbb{R}^{n-n_\nu}$ a set $X_\nu(x^{-\nu}) \subseteq \mathbb{R}^{n_\nu}$ (see [2, 60] for an introduction to set-valued analysis). Throughout this article we will use the subsequent assumption, which follows, for example, if the strategy sets $X_\nu(x^{-\nu})$ are bounded for each $x^{-\nu} \in \operatorname{dom} X_\nu$.

Assumption 1.2.7 *For each $\nu \in \{1, \ldots, N\}$ and all $x^{-\nu} \in \operatorname{dom} X_\nu$ the problem $Q_\nu(x^{-\nu})$ is solvable, that is, $Q_\nu(x^{-\nu})$ possesses at least one optimal point.*

1.3 Examples

Linear generalized Nash equilibrium problems arise quite naturally in situations where several competing agents face linear optimization problems while sharing at least one constraint. In Section 1.3.1 and 1.3.2 we consider some examples to illustrate this thought. Furthermore, it is possible to reformulate some classes of nonlinear generalized Nash equilibrium problems as LGNEPs as we shall see in Section 1.3.3.

1.3.1 Basic Economic Market Model

In this section, we introduce a simple economic model which we call *basic economic market model*. The basic economic market model is an LGNEP where some constraints are shared by all players.

Assume that we have N companies offering the same product on a common market. Company $\nu \in \{1, \dots, N\}$ possesses a contingent $C^\nu \geq 0$ of this product and sells it in K different price categories $p_1^\nu, \dots, p_K^\nu$, where the prices are given, that is, company ν acts as price taker. For each price category $k \in \{1, \dots, K\}$ there is a given total demand $D_k \geq 0$ of this good. The optimization problem of company ν is to determine the quantity x_k^ν to sell in price category p_k^ν in order to maximize its profit. Hence, company ν faces its optimization problem

$$\max_{x^\nu \in \mathbb{R}^K} \langle p^\nu, x^\nu \rangle$$

subject to its nonnegativity constraint

$$x^\nu \geq 0$$

and the constraint

$$\sum_{k=1}^{K} x_k^\nu \leq C^\nu$$

which guarantees that the total offering of company ν respects its capacity. These are private constraints for company ν. Further, there is a coupling third constraint

$$\sum_{\nu=1}^{N} x_k^\nu \leq D_k, \ k \in \{1, \dots, K\},$$

assuring that within each price category the accumulated offering of all companies does not exceed its total demand. Obviously, the search of Nash equilibria in this setting yields an LGNEP.

Example 1.3.1 *As an example for the considered market setting one might think of several travel agencies offering seats for the same flight. Usually, these seats are offered in different price categories which depend on the remaining time until departure. These prices may differ within the travel agencies, that is, $p_k^\nu \neq p_k^\mu$ is possible for $\nu \neq \mu$. Typically, the demand D_k is a random variable which, for instance, can be replaced by its expected value or other estimators that are based on historical data.*

Obviously, we arrive at GNEPs if players share at least one constraint but it is also possible that player ν possesses a coupled restriction that is not shared with the remaining players. This is illustrated in the following example where we consider a modification of the economic market model. Assume that a new company $N+1$ is entering the market. In order to find its market position company $N+1$ tries to offer at least the same amount in the cheapest price category as the average of all other companies. Therefore, company $N+1$ has the profit maximization problem

$$\begin{aligned} \max_{x^{N+1}\in\mathbb{R}^K} \langle p^{N+1}, x^{N+1}\rangle \quad \text{s.t.} \quad & x^{N+1} \geq 0, \\ & \sum_{k=1}^{K} x_k^{N+1} \leq C^{N+1}, \\ & \sum_{\nu=1}^{N+1} x_k^\nu \leq D_k,\ k \in \{1,\dots,K\}, \\ & x_1^{N+1} \geq \frac{1}{N}\sum_{\nu=1}^{N} x_1^\nu. \end{aligned}$$

The established N companies have the same optimization problems as in the original model.

Both market models, the basic economic market model with N companies and its modified version with $N+1$ companies will be further investigated in Part II where we shall also use them as test examples for our numerical tests.

1.3.2 Extended Transportation Problem

In the classical transportation problem we have one forwarder who transports a given good from manufacturers to consumers while minimizing his transportation costs. We extend the transportation problem towards a more realistic scenario and introduce several forwarders as depicted in Figure 1.1.

To improve the readability, in the subsequent work we will refer to the transportation problem with several forwarders just as *extended transportation problem* (ETP)

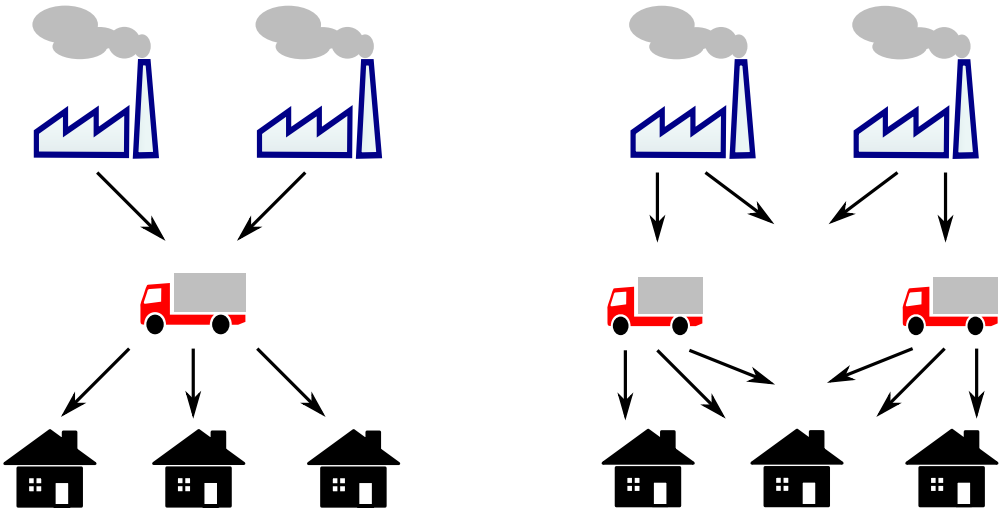

Figure 1.1: The classical transportation problem with one forwarder on the left hand side and the extended transportation problem with several forwarders on the right hand side

More formally, consider N competing forwarding agencies which want to transport one good from R manufacturers to T consumers. Manufacturer $r \in \{1, \dots, R\}$ has a production capacity of $S_r \geq 0$ and consumer $t \in \{1, \dots, T\}$ needs at least $D_t \geq 0$ units of this good with $\sum_{r=1}^{R} S_r = \sum_{t=1}^{T} D_t$. The transportation cost per unit from manufacturer $r \in \{1, \dots, R\}$ to consumer $t \in \{1, \dots, T\}$ by forwarder $\nu \in \{1, \dots, N\}$ is denoted by c_{rt}^{ν}. Further, x_{rt}^{ν} is defined as the number of transported units from manufacturer r to consumer t by forwarder ν.

Each forwarder wants to minimize his transportation costs given the decisions of the remaining forwarders, that is, forwarder ν faces the optimization problem

$$\min_{x^{\nu} \in \mathbb{R}^{R \times T}} \sum_{r=1}^{R} \sum_{t=1}^{T} c_{rt}^{\nu} x_{rt}^{\nu}$$

subject to his constraints concerning the supply

$$\sum_{\ell=1}^{N} \sum_{t=1}^{T} x_{rt}^{\ell} \;=\; S_r, \quad r \in \{1, \dots, R\},$$

as well as his demand constraints

$$\sum_{\ell=1}^{N}\sum_{r=1}^{R} x_{rt}^{\ell} \;=\; D_t, \quad t \in \{1,\ldots,T\},$$

and the nonnegativity condition

$$x_{rt}^{\nu} \;\geq\; 0, \quad r \in \{1,\ldots,R\},\; t \in \{1,\ldots,T\}.$$

The search for equilibria in extended transportation problems yields an equality constrained LGNEP. Therefore, we shall adapt our notation in Part III where we examine ETPs extensively and consider both, theoretical properties of ETPs as well as the numerical computation of their equilibria.

1.3.3 Epigraphical Reformulation of Min-Max Games

Suppose there are N players facing nonsmooth convex piecewise linear optimization problems with coupled linear constraints, that is, player ν's optimization problem is given by

$$\min_{x^\nu \in \mathbb{R}^{n_\nu}} \max_{t=1,\ldots,T} \langle a_t, x^\nu \rangle \quad \text{s.t.} \quad A^\nu x^\nu + B^\nu x^{-\nu} \;\leq\; b^\nu$$

for some $T \in \mathbb{N}$. This can be reformulated using the so-called *epigraphical reformulation* (cf. [65]) and we obtain the equivalent linear optimization problem

$$\begin{aligned} \min_{(x^\nu,\alpha)\in\mathbb{R}^{n_\nu}\times\mathbb{R}} \alpha \qquad \text{s.t.} \qquad & \langle a_t, x^\nu \rangle \leq \alpha, \quad t \in \{1,\ldots,T\}, \\ & A^\nu x^\nu + B^\nu x^{-\nu} \leq b^\nu \end{aligned}$$

for each player $\nu \in \{1,\ldots,N\}$, such that we arrive at an LGNEP.

Example 1.3.2 *Let us modify the ETP from Section 1.3.2. Suppose the transportation from manufacturer r to consumer t may fail due to some logistic problems. The costs of such an error are assumed to be very high. In order to limit the associated costs forwarder ν wants to set the maximal delivery size as low as possible. Additionally, the accumulated costs of player ν must not exceed a given amount C^ν.*

Hence, player ν faces the transportation problem

$$\min_{x^\nu \in \mathbb{R}^{R\times T}} \max_{(r,t)\in\{1,\ldots,R\}\times\{1,\ldots,T\}} x_{rt}^{\nu}$$

subject to the constraints

$$\begin{aligned}
\sum_{r=1}^{R}\sum_{t=1}^{T} c_{rt}^{\nu} x_{rt}^{\nu} &\leq C^{\nu} \\
\sum_{\ell=1}^{N}\sum_{t=1}^{T} x_{rt}^{\ell} &= S_r, \quad r \in \{1,\dots,R\}, \\
\sum_{\ell=1}^{N}\sum_{r=1}^{R} x_{rt}^{\ell} &= D_t, \quad t \in \{1,\dots,T\}, \\
x_{rt}^{\nu} &\geq 0, \quad r \in \{1,\dots,R\},\ t \in \{1,\dots,T\}.
\end{aligned}$$

As described above, the resulting system of piecewise linear optimization problems can be reformulated as an LGNEP.

Remark 1.3.3 *The epigraphical reformulation is a useful modeling tool, in particular, to exploit min-max structures in optimization problems.*

1.4 Structure of the Set of Nash Equilibria

Typically, in LGNEPs where players share at least one constraint there is no *unique* Nash equilibrium. In contrast, there may exist many Nash equilibria and under mild conditions the set of Nash equilibria together with their corresponding dual variables form a Lipschitz manifold whose dimension coincides with the number of shared constraints (cf. [17]). Furthermore, the set of Nash equilibria is formed by the union of facets of a polyhedron since Nash equilibria are exactly the minimal points of a concave function where the minimization is taken over a polyhedral set as we shall see in Section 1.5.1. The following examples illustrate that, even in the linear case, in general the set of all Nash equilibria does not possess any specific topological properties like connectedness of convexity. The set of all Nash equilibria is denoted by S.

Example 1.4.1 *Consider the following LGNEP with two players, that is, $N = 2$. Player one controls the variable x_1 and player two x_2, respectively. The optimization problems of the players are given by*

$$\begin{aligned}
Q_1(x_2): \qquad \min_{x_1 \in \mathbb{R}} -x_1 \quad s.t. \qquad x_1 + x_2 &\leq 1, \\
-x_1 &\leq 0, \\
-x_2 &\leq 0
\end{aligned}$$

for player one and

$$Q_2(x_1): \quad \min_{x_2 \in \mathbb{R}} -x_2 \quad s.t. \quad \begin{aligned} x_1 + x_2 &\leq 1, \\ -x_1 &\leq 0, \\ -x_2 &\leq 0. \end{aligned}$$

for the second player. Direct inspections show that the set of Nash equilibria is the line segment

$$S = \left[\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right]$$

as depicted in Figure 1.2.

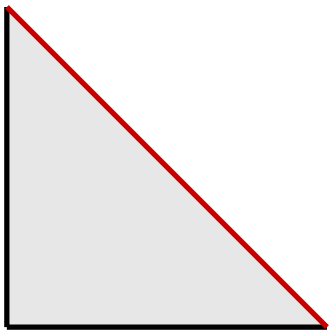

Figure 1.2: An illustration of the common strategy set where the Nash equilibria form the convex red colored set

In Example 1.4.1, the set of all Nash equilibria is a line segment and therefore a convex set. In general, the set of Nash equilibria does not have to be convex as we see in Example 1.4.2.

Example 1.4.2 *Consider the following LGNEP where the optimization problems are given by*

$$Q_1(x_2): \quad \min_{x_1 \in \mathbb{R}} -x_1 \quad s.t. \quad \begin{aligned} \tfrac{1}{2}x_1 + x_2 &\leq 1, \\ x_1 + \tfrac{1}{2}x_2 &\leq 1, \\ -x_1 &\leq 0, \\ -x_2 &\leq 0 \end{aligned}$$

and

$$Q_2(x_1): \quad \min_{x_2 \in \mathbb{R}} -x_2 \quad s.t. \quad \begin{aligned} \tfrac{1}{2}x_1 + x_2 &\leq 1, \\ x_1 + \tfrac{1}{2}x_2 &\leq 1, \\ -x_1 &\leq 0, \\ -x_2 &\leq 0. \end{aligned}$$

Straightforward calculations show that the set of Nash equilibria is given by

$$S = \left[\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} \frac{2}{3} \\ \frac{2}{3} \end{pmatrix}\right] \cup \left[\begin{pmatrix} \frac{2}{3} \\ \frac{2}{3} \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right]$$

which is a connected but nonconvex set as depicted in Figure 1.3.

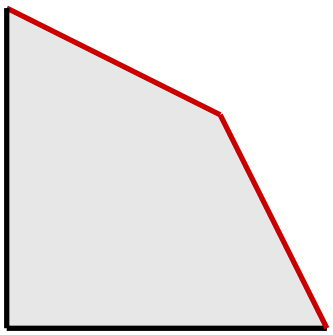

Figure 1.3: An illustration of the common strategy set where the Nash equilibria form the red colored set which is connected but nonconvex

Furthermore, the set of all Nash equilibria of a given LGNEP may even fail to be connected as illustrated in Example 1.4.3.

Example 1.4.3 *Consider an LGNEP which is defined by*

$$Q_1(x_2): \quad \min_{x_1 \in \mathbb{R}} -x_1 \quad s.t. \quad \begin{aligned} x_1 - x_2 &\leq 0, \\ x_1 + x_2 &\leq 1, \\ -2x_1 + x_2 &\leq 0 \end{aligned}$$

and

$$Q_2(x_1): \quad \min_{x_2 \in \mathbb{R}} -x_2 \quad s.t. \quad \begin{aligned} x_1 - x_2 &\leq 0, \\ x_1 + x_2 &\leq 1, \\ -2x_1 + x_2 &\leq 0. \end{aligned}$$

The set of Nash equilibria is formed by

$$S = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} \cup \left[\begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}, \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \end{pmatrix} \right]$$

which is a set that is not connected as illustrated in Figure 1.4.

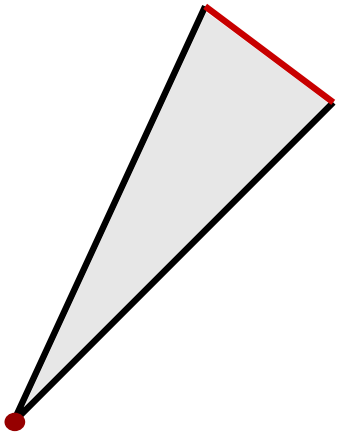

Figure 1.4: An illustration of the common strategy set where the Nash equilibria form the red colored set which is not connected

1.5 Computation of Nash Equilibria

In this section, we discuss some reformulations of GNEPs that yield numerical methods for the computation of generalized Nash equilibria. In Section 1.5.1, we shall see that GNEPs may be reformulated as nonsmooth optimization problems. In Section 1.5.2, we consider a reformulation of the LGNEP as a constrained equation that is based on the concatenated KKT systems of all players. In Section 1.5.3, we briefly examine a reformulation *via* so-called quasi-variational inequalities and, finally, in Section 1.5.4 we shall have a look on further numerical methods for the computation of Nash equilibria in LGNEPs.

1.5.1 Gap Function

For $x \in \mathbb{R}^n$ and $\nu \in \{1, \ldots, N\}$ we define player ν's *optimal value function*

$$\varphi_\nu(x^{-\nu}) := \begin{cases} \min_{x^\nu \in X_\nu(x^{-\nu})} \langle c^\nu, x^\nu \rangle, & \text{if } x^{-\nu} \in \operatorname{dom} X_\nu, \\ +\infty, & \text{else,} \end{cases}$$

where the attainment of the minimum in the case $x^{-\nu} \in \operatorname{dom} X_\nu$ is ensured by Assumption 1.2.7, and the function

$$V(x) \ := \ \sum_{\nu=1}^{N} \langle c^\nu, x^\nu \rangle - \varphi_\nu(x^{-\nu}),$$

which, again under Assumption 1.2.7, is real-valued if and only if x is taken from the set

$$M \ := \ \bigcap_{\nu=1,\ldots,N} (\mathbb{R}^{n_\nu} \times \operatorname{dom} X_\nu),$$

whereas we have $V(x) = -\infty$ for $x \notin M$. Particularly, in general, the domain of V does not cover the whole space, which causes numerical and theoretical difficulties (see, e.g. [35] and [66], resp.). Rewriting

$$V(x) \ = \ \max_{y \in X_1(x^{-1}) \times \ldots \times X_N(x^{-N})} \sum_{\nu=1}^{N} \langle c^\nu, x^\nu \rangle - \langle c^\nu, y^\nu \rangle$$

shows that V is a merit function based on the Nikaido-Isoda function

$$\Psi(x, y) \ := \ \sum_{\nu=1}^{N} \langle c^\nu, x^\nu \rangle - \langle c^\nu, y^\nu \rangle$$

of LGNEP ([54]). Borrowing the terminology from the related case of (quasi) variational inequalities ([3, 34]) in the following we shall refer to V as a *gap function* of LGNEP. According to [25], the function V is nonnegative on the *unfolded common strategy set*

$$W := \{x \in \mathbb{R}^n : \ x^\nu \in X_\nu(x^{-\nu}), \ \nu = 1, \ldots, N\} \ \subseteq \ M.$$

These properties yield the following result, which also holds under considerably weaker assumptions (cf., e.g., [39]).

Proposition 1.5.1 *The generalized Nash equilibria of LGNEP are the global minimal points of the (possibly non-smooth) optimization problem*

$$\min V(x) \quad s.t. \quad x \in W$$

with optimal value zero.

As mentioned above the gap function V is an extended-valued, that is, not necessarily real-valued, and possibly nonsmooth function. We approach both difficulties in Part I of this work. Furthermore, in Section 7 we shall discuss some nonsmooth numerical methods that are able to compute Nash equilibria by solving the nonsmooth optimization problem of minimizing V over W. We also apply these nonsmooth optimization algorithms in Part III in order to compute Nash equilibria for the ETP.

1.5.2 Concatenated KKT Systems

As mentioned in Remark 1.2.4, in this chapter we change our notation slightly in order to obtain a more specific representation of player ν's constraints, that is, we introduce suitable matrices $A^{\nu\mu} \in \mathbb{R}^{m_\nu \times n_\mu}$, such that player ν's optimization problem is given by

$$Q_\nu(x^{-\nu}): \qquad \min_{x^\nu \in \mathbb{R}^{n_\nu}} \langle c^\nu, x^\nu \rangle \quad \text{s.t.} \quad A^{\nu\nu}x^\nu + \sum_{\mu \neq \nu} A^{\nu\mu}x^\mu \leq b^\nu$$

for each $\nu \in \{1, \ldots, N\}$ and fixed $x^{-\nu} \in \mathbb{R}^{n-n_\nu}$.

One way to compute a generalized Nash equilibrium is to solve the KKT systems for all players simultaneously. In particular, in the linear case this yields that $\bar{x}$ is a generalized Nash equilibrium, if and only if there exist $\lambda^\nu, w^\nu \in \mathbb{R}^{m_\nu}$ satisfying the KKT conditions

$$\begin{aligned} c^\nu + (A^{\nu\nu})^T \lambda^\nu &= 0, \\ (A^{\nu 1}\, A^{\nu 2}\, \ldots\, A^{\nu N})\, \bar{x} - b^\nu + w^\nu &= 0, \\ w^\nu \geq 0, \quad \lambda^\nu \geq 0, \quad (w^\nu)^T \lambda^\nu &= 0 \end{aligned}$$

for all $\nu \in \{1, \ldots, N\}$. Let us define $m := m_1 + \ldots, m_N$ and vectors

$$c := \begin{pmatrix} c^1 \\ \vdots \\ c^N \end{pmatrix} \in \mathbb{R}^n, \quad b := \begin{pmatrix} b^1 \\ \vdots \\ b^N \end{pmatrix} \in \mathbb{R}^m, \quad \lambda := \begin{pmatrix} \lambda^1 \\ \vdots \\ \lambda^N \end{pmatrix} \in \mathbb{R}^m$$

and

$$w := \begin{pmatrix} w^1 \\ \vdots \\ w^N \end{pmatrix} \in \mathbb{R}^m$$

as well as matrices

$$E := \begin{pmatrix} (A^{11})^T & & 0 \\ & \ddots & \\ 0 & & (A^{NN})^T \end{pmatrix}$$

and

$$A := \begin{pmatrix} A^{11} & \cdots & A^{1N} \\ \vdots & & \vdots \\ A^{N1} & \cdots & A^{NN} \end{pmatrix}.$$

Then the optimality conditions can be written in the form

$$\begin{aligned} c + E\lambda &= 0, && (1.1)\\ A\bar{x} - b + w &= 0, && (1.2)\\ \lambda \geq 0, \quad w \geq 0, \quad w^T\lambda &= 0. && (1.3) \end{aligned}$$

Using the Hadamard product

$$\lambda \circ w := \begin{pmatrix} \lambda_1 w_1 \\ \vdots \\ \lambda_m w_m \end{pmatrix}$$

we define the function $H : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^{n+2m}$ by

$$H(x, \lambda, w) := \begin{pmatrix} c + E\lambda \\ A\bar{x} - b + w \\ \lambda \circ w \end{pmatrix}$$

an arrive at the following result (cf. [23]) where the nonnegative orthant is denoted by $\mathbb{R}^m_+$.

Proposition 1.5.2 *A vector $\bar{x}$ is a generalized Nash equilibrium if and only if there are some $\bar{\lambda}, \bar{w} \in \mathbb{R}^m$, such that $(\bar{x}, \bar{\lambda}, \bar{w})$ is a solution of the constrained equation*

$$H(x, \lambda, w) = 0, \qquad (x, \lambda, w) \in \mathbb{R}^n \times \mathbb{R}^m_+ \times \mathbb{R}^m_+. \tag{1.4}$$

In Chapter 6, we shall discuss some algorithms that compute generalized Nash equilibria by solving the constrained equation (1.4).

1.5.3 Quasi-Variational Inequalities

Let $\bar{x}$ be a generalized Nash equilibrium of an LGNEP. Then finding a Nash equilibrium amounts to computing a vector $\bar{x}$ with

$$\bar{x}^\nu \in X_\nu(\bar{x}^{-\nu})$$

and

$$\langle c^\nu, \bar{x}^\nu \rangle \leq \langle c^\nu, x^\nu \rangle$$

for each $\nu \in \{1, \ldots, N\}$ and all $x^\nu \in X_\nu(\bar{x}^{-\nu})$. It was first noticed in [5] in the context of nonlinear GNEPs that this is equivalent to finding a vector $\bar{x}$ with

$$\langle F(\bar{x}), y - \bar{x} \rangle \geq 0 \tag{1.5}$$

for all $y \in S(\bar{x})$ where we defined the constant function

$$F(x) \ := \ \begin{pmatrix} c^1 \\ \vdots \\ c^N \end{pmatrix} \in \mathbb{R}^n$$

and

$$S(x) \ := \ X_1(x^{-1}) \times \cdots \times X_N(x^{-N}).$$

The inequality in (1.5) is a so-called *quasi-variational inequality* (cf., e.g., [36]) and therefore one might think of computing Nash equilibria by exploiting the theory and numerical results of quasi-variational inequality. Unfortunately, both, theory and numerical results on quasi-variational inequalities are far less developed than for variational inequalities where the feasible set S does not depend on $\bar{x}$ (cf. [27]). Therefore, in this work we do not try to compute Nash equilibria by exploiting their reformulation as quasi-variational inequalities.

1.5.4 Further Numerical Approaches

Besides the reformulation presented in Section 1.5.2, the search for solutions of the concatenated KKT systems of all players can be formulated as a so-called *linear complementarity problem* (cf. [13]). A common method to solve linear complementarity problems is Lemke's method which is a finite method and an extension of the famous Simplex method. In [62], the authors compute equilibria for AGNEPs by Lemke's method and reporte numerical problems which where caused by degeneracies that are intrinsic in GNEPs with shared constraints (cf. [18]). We do not consider this approach in this work for the solution of LGNEPs but its application would be an interesting subject to future research.

Besides these general methods, in concrete applications it is sometimes possible to compute Nash equilibria in a more simple way by exploiting the structure of the problem. This works pretty well for the extended transportation problem as we will see in Part III.

1.6 Normalized Nash Equilibria

Let us consider GNEPs where all coupled constraints are shared by all players. Without loss of generality we may assume

$$m_1 = \ldots = m_N, \ b^1 = \ldots = b^N \quad \text{and} \quad A^{1\nu} = \ldots = A^{N\nu}$$

for each $\nu \in \{1, \ldots, N\}$ for these special GNEPs. This does not reduce any generality since all private constraints may be posed to the remaining players without restricting their strategy spaces. For these kind of GNEPs, a special type of equilibrium, the so-called *normalized* (or *variational*) Nash equilibrium, was introduced in the seminal work [61].

Definition 1.6.1 *Let $(\bar{x}, \bar{\lambda}, \bar{\omega})$ be a solution of (1.1), (1.2) and (1.3). Then $\bar{x}$ is a normalized Nash equilibrium if and only if there exist positive weights $r_1, \ldots, r_N > 0$, such that there is a common KKT multiplier $\hat{\lambda} \in \mathbb{R}^{m_1}$ satisfying $\hat{\lambda} = r_\nu \bar{\lambda}^\nu$ for all $\nu \in \{1, \ldots, N\}$.*

As we shall see in Theorem 1.6.2, normalized Nash equilibria are exactly the optimal points of a linear optimization problem whose feasible set is the unfolded common strategy set

$$W = \{x \in \mathbb{R}^n : \ \hat{A}x \leq \hat{b}\}$$

with $\hat{A} := (A^{11}A^{22} \ldots A^{NN}) \in \mathbb{R}^{m_1 \times n}$ and $\hat{b} := b^1$. Note that in the representation of W we omitted the constraints that occur multiply. In order to construct the associated objective function let $r_1, \ldots, r_N > 0$ and let $c^\nu \in \mathbb{R}^{n_\nu}$ be the cost vector of player ν. Then we define the vector

$$c_{r_1,\ldots,r_N} := \begin{pmatrix} r_1 \cdot c^1 \\ \vdots \\ r_N \cdot c^N \end{pmatrix} \in \mathbb{R}^n$$

and arrive at the following result which states that normalized Nash equilibria are equivalent to solutions of a single linear program. Despite its straightforward proof this is a powerful result since it provides a set of efficiently computable Nash equilibria for a certain class of LGNEPs.

Theorem 1.6.2 *Let $r_1, \ldots, r_N > 0$. Then $\bar{x}$ solves the linear optimization problem*

$$\min_x \langle c_{r_1,\ldots,r_N}, x \rangle \quad s.t. \quad \hat{A}x \leq \hat{b} \tag{1.6}$$

if and only if $\bar{x}$ is a normalized Nash equilibrium with weights $r_1, \ldots, r_N$.

Proof. Inspecting the optimality conditions (1.1), (1.2) and (1.3) we see that $\bar{x}$ is a normalized Nash equilibrium if and only if there are $\hat{\lambda}, \hat{w} \in \mathbb{R}^{m_1}$, such that

$$\begin{aligned} c_{r_1,\ldots,r_N} + \hat{A}^T\hat{\lambda} &= 0, \\ \hat{A}\bar{x} - \hat{b} + \hat{w} &= 0, \\ \hat{\lambda} \geq 0, \quad \hat{w} \geq 0, \quad \hat{w}^T\hat{\lambda} &= 0. \end{aligned}$$

These are exactly the optimality conditions of the single linear program (1.6) and therefore we have shown the desired result. •

Remark 1.6.3 *According to Theorem 1.6.2, the existence of a normalized Nash equilibrium is equivalent to the solvability of the corresponding linear program. Further, note that for the considered class of LGNEPs, which are* non-cooperative *games, normalized Nash equilibria can be interpreted as equilibria in* cooperative *games where the function* $\langle c_{r_1,\ldots,r_N}, \cdot \rangle$ *is the common objective function which is minimized by all players in a cooperative way.*

As we shall see in the following example there are Nash equilibria which are not normalized equilibria, such that, in general, one may not expect to be able to compute *all* Nash equilibria of an arbitrary LGNEP by solving linear programs. This is an example that accompanies the famous nonlinear example from Harker (cf. [36]) and states that also for LGNEPs there are Nash equilibria that are not normalized equilibria.

Example 1.6.4 *In Example 1.4.3, we construct the optimization problem* (1.6) *and obtain*

$$P: \quad \min_{x \in \mathbb{R}^2} -r_1 x_1 - r_2 x_2 \quad s.t. \quad \begin{aligned} x_1 - x_2 &\leq 0, \\ x_1 + x_2 &\leq 1, \\ -2x_1 + x_2 &\leq 0 \end{aligned}$$

with $r_1, r_2 > 0$. *Then it is straightforward to see that each equilibrium in the line segment* $\left[\begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}, \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \end{pmatrix} \right]$ *is a normalized Nash equilibrium since it is also an optimal point of* P *provided a suitable choice of* r_1 *and* r_2. *However, there is no choice of positive parameters* r_1 *and* r_2, *such that* $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ *is an optimal point of* P. *Thus,* $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ *is not a normalized Nash equilibrium of the LGNEP.*

1.7 Thesis Overview

In Part I of this work, we study the gap function which arises as the objective function from a reformulation of the LGNEP as a constrained nonsmooth optimization problem (cf. Section 1.5.1). The gap function is a nonsmooth and extended real-valued function. Both of these properties cause theoretical and numerical problems and, therefore, we are going to study its nondifferentiability structure in detail and introduce a real-valued global extension of

this function. Interestingly, there exist very deep connections between nonsmoothness of the gap function and the regularity of its feasible set in a sense that will be explained later on. Besides these theoretical investigations, in Part I we derive a formula for the subdifferential of the gap function which will be exploited numerically by designing a projected subgradient method in Part II.

Besides the introduction of the aforementioned projected subgradient method, in Part II we investigate an interior point method for the computation of Nash equilibria. The known convergence results from literature for the interior point method require the nonsingularity of some second order derivatives and are therefore not applicable in the linear case. Thus, we present new convergence conditions for this method that are also valid in the linear case. All introduced algorithms are extensively tested on instances of the basic market models.

Finally, in Part III, we introduce a special LGNEP, the extended transportation problem (ETP). The ETP is the first treatment of a very natural extension of the celebrated transportation problem by non-cooperative game theory. We apply the projected subgradient method in order to compute Nash equilibria in the ETP. Furthermore, we show that the computation of a large set of Nash equilibria in the ETP can by done very efficiently. This enables the computation of many Nash equilibria for the ETP and, therefore, the question arises which equilibrium to select in practical application. That is a problem which is also known as the *Equilibrium selection problem* and will be addressed before we end this work with some closing remarks.

Part I

Theory

Chapter 2

Gap Function

2.1 Basic Definitions, Notation and Overview

This part is based on the article [66] of the author. Let us briefly recall all definitions from Section 1.5.1 that we need in order to introduce the gap function V.

We have N competing players and player ν tries to optimize his parametric optimization problem

$$Q_\nu(x^{-\nu}): \qquad \min_{x^\nu \in \mathbb{R}^{n_\nu}} \langle c^\nu, x^\nu \rangle \quad \text{s.t.} \quad x^\nu \in X_\nu(x^{-\nu}).$$

His strategy set depends on the decision of the remaining players $x^{-\nu}$ and is described by

$$X_\nu(x^{-\nu}) \ := \ \{x^\nu \in \mathbb{R}^{n_\nu} : \ A^\nu x^\nu + B^\nu x^{-\nu} \leq b^\nu\}$$

whereas, as mentioned in Remark 1.2.2, in Part III we shall also consider equality constraints explicitly. However, in this part the appearance of equality constraints does not imply any qualitative changes in our results and is therefore omitted.

Note that $X_\nu : \mathbb{R}^{n-n_\nu} \rightrightarrows \mathbb{R}^{n_\nu}$ is a set-valued mapping whose domain $\operatorname{dom} X_\nu$ is defined as

$$\operatorname{dom} X_\nu \ := \ \{x^{-\nu} \in \mathbb{R}^{n-n_\nu} : \ X_\nu(x^{-\nu}) \neq \emptyset\}.$$

We define player ν's *optimal value function*

$$\varphi_\nu(x^{-\nu}) \ := \ \begin{cases} \min_{x^\nu \in X_\nu(x^{-\nu})} \langle c^\nu, x^\nu \rangle, & \text{if } x^{-\nu} \in \operatorname{dom} X_\nu, \\ +\infty, & \text{else,} \end{cases}$$

where the attainment of the minimum in the case $x^{-\nu} \in \operatorname{dom} X_\nu$ is ensured by Assumption 1.2.7. This allows us to define the *gap function* at a given $x = (x^\nu, x^{-\nu}) \in \mathbb{R}^n$ by

$$V(x) \;:=\; \sum_{\nu=1}^{N} \langle c^\nu, x^\nu \rangle - \varphi_\nu(x^{-\nu}),$$

whose domain

$$\operatorname{dom} V \;:=\; \{x \in \mathbb{R}^n :\; V(x) \in \mathbb{R}\}$$

is given by

$$\operatorname{dom} V \;=\; M \;:=\; \bigcap_{\nu=1,\ldots,N} (\mathbb{R}^{n_\nu} \times \operatorname{dom} X_\nu).$$

According to [25], the function V is nonnegative on the unfolded common strategy set

$$\begin{aligned} W &= \{x \in \mathbb{R}^n :\; x^\nu \in X_\nu(x^{-\nu}),\; \nu = 1, \ldots, N\} \\ &= \{x \in \mathbb{R}^n :\; A^\nu x^\nu + B^\nu x^{-\nu} \leq b^\nu,\; \nu = 1, \ldots, N\} \\ &\subseteq M. \end{aligned}$$

Note that, in general, the inclusion $W \subseteq M$ is strict, that is, we may have $\operatorname{dom} V \subseteq W$ and $\operatorname{dom} V \neq W$ as illustrated in Figure 2.1. Further, we denote the set of all (generalized) Nash equilibria of a given GNEP by S.

According to Proposition 1.5.1, generalized Nash equilibria of LGNEP are the global minimal points of the (possibly non-smooth) optimization problem

$$\min V(x) \quad \text{s.t.} \quad x \in W$$

with optimal value zero.

The gap function V gives rise to two main difficulties. First, in general, the domain of V does not cover the whole space which causes numerical and theoretical difficulties (see, e.g. [35] and [66], resp.). Thus, in Section 2.2, we will overcome this difficulty by introducing a global real-valued extension of V. This global extension is only computable if the vertices of some polyhedral sets are known or can be computed efficiently. However, as we shall see in Part II this holds true for some LGNEPs and enables a very stable and fast numerical treatment for these LGNEPs.

Second, V may be a nondifferentiable function. Therefore, in Chapter 3, we study its nonsmoothness in detail and introduce a new regularity condition that even characterizes the nonsmoothness of V under mild conditions. Chapter 3 also provides a basis for an algorithmic treatments of LGNEPs with nonsmooth optimization techniques as we will see in Part II.

2.2 Global Extension of the Gap Function

We will use the following example throughout this article to illustrate our results. Note that the example is stable in the sense that small perturbations in the defining data do not imply qualitative changes in our results.

Example 2.2.1 *Let us consider an LGNEP with two players, each having a one-dimensional strategy set, that is, $n_1 = n_2 = 1$ and therefore $N = n = 2$. To simplify the notation, we denote the decision variable of player one by x_1 and of the second player by x_2, respectively. The players share three constraints given by*

$$\begin{pmatrix} -\frac{1}{2} \\ 1 \\ -1 \end{pmatrix} x_1 + \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} x_2 \ \leq \ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Furthermore, the objective functions are $\theta_1(x_1) = -x_1$ and $\theta_2(x_2) = x_2$. Then we obtain the optimal value functions

$$\varphi_1(x_2) \ = \ \begin{cases} -x_2 - 1 & , \ x_2 \in \operatorname{dom} X_1 = [-1, 3], \\ \infty & , \ \textit{else}, \end{cases}$$

and

$$\varphi_2(x_1) \ = \ \begin{cases} |x_1| - 1 & , \ x_1 \in \operatorname{dom} X_2 = [-\frac{4}{3}, 4], \\ \infty & , \ \textit{else}, \end{cases}$$

so that the gap function is given by

$$V(x) \ = \ \begin{cases} 2 - x_1 - |x_1| + 2x_2 & , \ x \in M = [-\frac{4}{3}, 4] \times [-1, 3], \\ -\infty & , \ \textit{else}. \end{cases}$$

The generalized Nash equilibria form the line segment

$$S \ = \ \left[\begin{pmatrix} 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 4 \\ 3 \end{pmatrix} \right]$$

and the sets W and M as well as the graph of V on W are illustrated in Figures 2.1 and 2.2, respectively.

As Example 2.2.1 illustrates, the minimization of V over W may be a non-smooth optimization problem, so that below we shall study smoothness properties of V. As V is only real-valued on the set M, first we will construct a global extension $\widehat{V}$ of V which is real-valued on $\mathbb{R}^n$ and then examine the smoothness properties of $\widehat{V}$ on $\mathbb{R}^n$. The explicit form of one such global extension is straightforward in Example 2.2.1, but in general its investigation

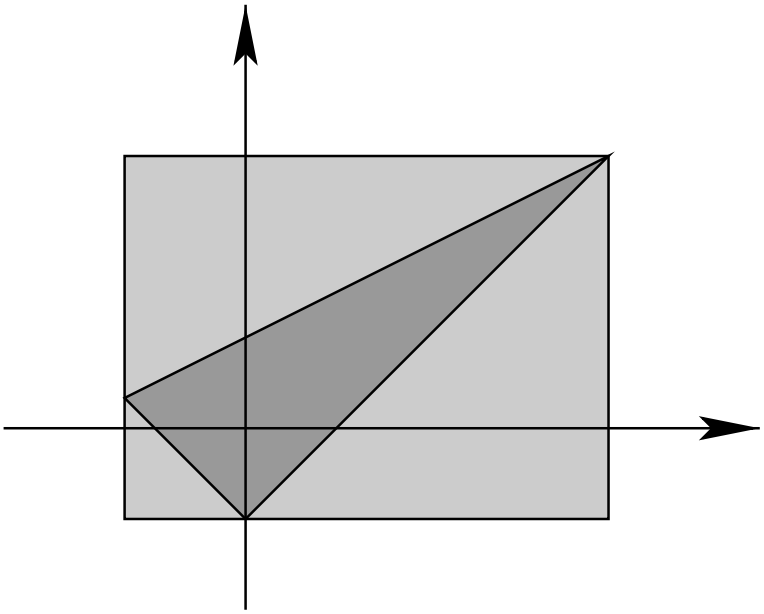

Figure 2.1: Illustration of the sets W and M in Example 2.2.1

is based on a dual representation of V. Therefore, we will now consider the dual problem associated with $Q(x^{-\nu})$.

For $\nu \in \{1, \ldots, N\}$ and $x^{-\nu} \in \operatorname{dom} X_\nu$, the dual problem of $Q_\nu(x^{-\nu})$ is

$$D_\nu(x^{-\nu}): \qquad \max_{\lambda^\nu \in \mathbb{R}^{m_\nu}} \langle \lambda^\nu, B^\nu x^{-\nu} - b^\nu \rangle \quad \text{s.t.} \quad \lambda^\nu \in Z_\nu$$

with

$$Z_\nu \; := \; \{\lambda^\nu \in \mathbb{R}^{m_\nu} : \; c^\nu + (A^\nu)^T \lambda^\nu = 0, \; \lambda^\nu \geq 0\}.$$

Remark 2.2.2 *Note that Z_ν does not depend on $x^{-\nu}$. This is an interesting property since it implies that each LGNEP is equivalent to a 'dual game' where the feasible sets of all players do not depend on the decisions of the remaining players. However, this dual game is* not *a (standard) Nash equilibrium problem since, now, player ν's dual objective function*

$$\langle \cdot, B^\nu x^{-\nu} - b^\nu \rangle$$

depends on the vector $x^{-\nu}$ which is not *a decision vector in this dual game but a fixed parameter. Therefore, the dual game of an LGNEP still depends on the 'primal parameter' x.*

For all $\nu \in \{1, \ldots, N\}$ and $x^{-\nu} \in \operatorname{dom} X_\nu$ strong duality yields

$$\varphi_\nu(x^{-\nu}) \; = \; \max_{\lambda^\nu \in Z_\nu} \langle \lambda^\nu, B^\nu x^{-\nu} - b^\nu \rangle.$$

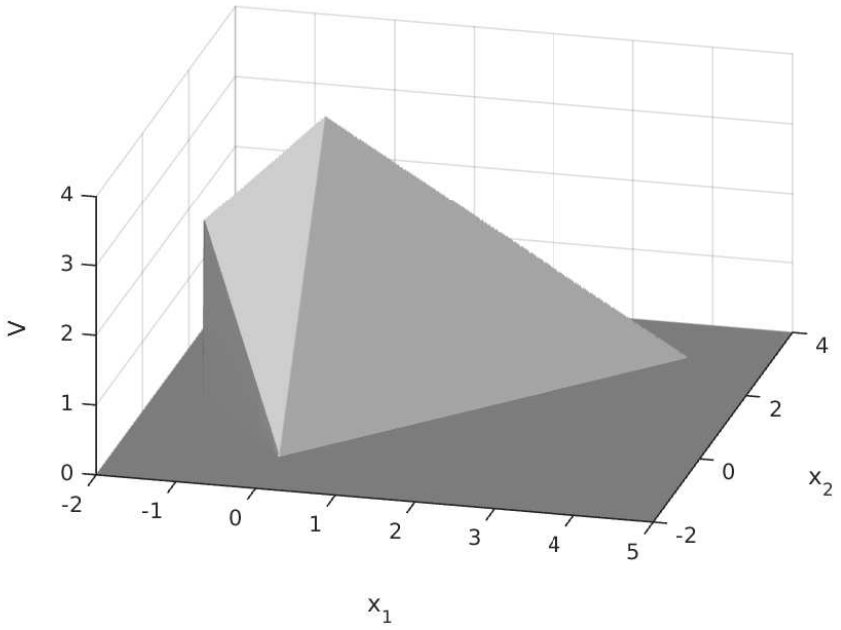

Figure 2.2: The graph of the gap function V on W in Example 2.2.1

Hence, if for each $\lambda^\nu \in Z_\nu$ and any $x \in \mathbb{R}^n$ we define the affine function

$$\ell_{\lambda^\nu}(x) \;:=\; \langle c^\nu, x^\nu\rangle + \langle \lambda^\nu, b^\nu - B^\nu x^{-\nu}\rangle,$$

then for $x \in M$ we obtain

$$\begin{aligned} V(x) &= \sum_{\nu=1}^{N} \langle c^\nu, x^\nu\rangle - \varphi_\nu(x^{-\nu}) \\ &= \sum_{\nu=1}^{N} \langle c^\nu, x^\nu\rangle - \max_{\lambda^\nu \in Z_\nu} \langle \lambda^\nu, B^\nu x^{-\nu} - b^\nu\rangle \\ &= \sum_{\nu=1}^{N} \min_{\lambda^\nu \in Z_\nu} \ell_{\lambda^\nu}(x). \end{aligned}$$

We remark that the latter representation of V, in an extended-valued sense, even holds on all of $\mathbb{R}^n$. In fact, for $x \notin M$ duality implies that $\ell_{\lambda^\nu}(x)$ is not bounded below (in λ^ν) on Z_ν for at least one $\nu \in \{1, \ldots, N\}$, so that the right-hand side formally attains the correct value $-\infty$ at x.

By $O_\nu := \operatorname{vert} Z_\nu$ let us denote the finite vertex set of the polyhedron Z_ν. The set O_ν is nonempty since, due to the nonnegativity condition on λ^ν, the

polyhedron Z_ν cannot contain a line and is, thus, pointed. Consequently we may define the function

$$\widehat{\varphi}_\nu(x^{-\nu}) \;:=\; \max_{\lambda^\nu \in O_\nu} \langle \lambda^\nu, B^\nu x^{-\nu} - b^\nu \rangle$$

for all $x^{-\nu} \in \mathbb{R}^{n-n_\nu}$. Note that $\widehat{\varphi}_\nu$ is a real-valued *global extension* of φ_ν, that is, we have $\widehat{\varphi}_\nu(x^{-\nu}) \in \mathbb{R}$ for all $x^{-\nu} \in \mathbb{R}^{n-n_\nu}$ and furthermore

$$\widehat{\varphi}_\nu(x^{-\nu}) \;=\; \varphi_\nu(x^{-\nu})$$

for all $x^{-\nu} \in \operatorname{dom} X_\nu$ by the vertex theorem of linear programming. This allows us to define a global extension $\widehat{V}$ of V by

$$\begin{aligned}
\widehat{V}(x) \;&:=\; \sum_{\nu=1}^{N} \langle c^\nu, x^\nu \rangle - \widehat{\varphi}_\nu(x^{-\nu}) \\
&=\; \sum_{\nu=1}^{N} \langle c^\nu, x^\nu \rangle - \max_{\lambda^\nu \in O_\nu} \langle \lambda^\nu, B^\nu x^{-\nu} - b^\nu \rangle \\
&=\; \sum_{\nu=1}^{N} \min_{\lambda^\nu \in O_\nu} \ell_{\lambda^\nu}(x)
\end{aligned}$$

for all $x \in \mathbb{R}^n$. Clearly, the functions V and $\widehat{V}$ coincide on M but, in contrast to V, the function $\widehat{V}$ is real-valued on all of $\mathbb{R}^n$. This will be crucial for our subsequent analysis.

Example 2.2.3 *In Example 2.2.1, we have*

$$Z_1 \;=\; \{\lambda^1 \in \mathbb{R}^3 :\; -1 - \frac{\lambda^1_1}{2} + \lambda^1_2 - \lambda^1_3 = 0,\; \lambda^1 \geq 0\}$$

with the singleton vertex set

$$O_1 \;=\; \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$$

for player one, as well as

$$Z_2 \;=\; \{\lambda^2 \in \mathbb{R}^3 :\; 1 + \lambda^2_1 - \lambda^2_2 - \lambda^2_3 = 0,\; \lambda^2 \geq 0\}$$

with the vertex set

$$O_2 \;=\; \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

for player two. Then we obtain the global extensions of the optimal value functions

$$\begin{aligned}\widehat{\varphi}_1(x_2) &= \max_{\lambda^1 \in O_1} \left\langle \lambda^1, \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} x_2 - \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\rangle \\ &= -x_2 - 1\end{aligned}$$

for the first player, and

$$\begin{aligned}\widehat{\varphi}_2(x_1) &= \max_{\lambda^2 \in O_2} \left\langle \lambda^2, \begin{pmatrix} -\frac{1}{2} \\ 1 \\ -1 \end{pmatrix} x_1 - \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\rangle \\ &= \max(x_1 - 1, -x_1 - 1) \\ &= |x_1| - 1\end{aligned}$$

for the second player. Finally, as expected, the global extension of the gap function is given by

$$\widehat{V}(x) \;=\; 2 - x_1 - |x_1| + 2x_2$$

for all $x \in \mathbb{R}^2$. •

The global extension $\widehat{V}$ of V obviously is a piecewise linear concave function, which is formed by means of finitely many affine functions ℓ_{λ^ν}, $\lambda^\nu \in O_\nu$, $\nu \in \{1, \ldots, N\}$. The latter are called the *selection functions* of $\widehat{V}$. For $x \in \mathbb{R}^n$ we call

$$O_\nu(x^{-\nu}) \;:=\; \{\bar{\lambda}^\nu \in O_\nu :\; \ell_{\bar{\lambda}^\nu}(x) = \min_{\lambda^\nu \in O_\nu} \ell_{\lambda^\nu}(x)\}$$

the *index set of active selection functions* of player ν at $x^{-\nu}$. Note that for $\nu \in \{1, \ldots, N\}$ the set $O_\nu(x^{-\nu})$ is nonempty, and that it does not depend on x^ν since the condition $\ell_{\bar{\lambda}^\nu}(x) = \min_{\lambda^\nu \in O_\nu} \ell_{\lambda^\nu}(x)$ is equivalent to

$$\langle \bar{\lambda}^\nu, B^\nu x^{-\nu} - b^\nu \rangle \;=\; \max_{\lambda^\nu \in O_\nu} \langle \lambda^\nu, B^\nu x^{-\nu} - b^\nu \rangle. \tag{2.1}$$

Due to (2.1), we may also write

$$O_\nu(x^{-\nu}) \;=\; \{\lambda^\nu \in O_\nu :\; \langle \lambda^\nu, B^\nu x^{-\nu} - b^\nu \rangle = \widehat{\varphi}_\nu(x^{-\nu})\} \tag{2.2}$$

and have

$$\widehat{V}(x) \;=\; \sum_{\nu=1}^{N} \min_{\lambda^\nu \in O_\nu(x^{-\nu})} \ell_{\lambda^\nu}(x). \tag{2.3}$$

For a given point $\bar{x} \in \mathbb{R}^n$ the sets $O_\nu(\bar{x}^{-\nu})$, $\nu \in \{1, \ldots, N\}$, already determine the *local behavior* of $\widehat{V}$ around $\bar{x}$:

Proposition 2.2.4 *For each $\bar{x} \in \mathbb{R}^n$ there exists a neighborhood U of $\bar{x}$ with*

$$\widehat{V}(x) \;=\; \sum_{\nu=1}^{N} \min_{\lambda^\nu \in O_\nu(\bar{x}^{-\nu})} \ell_{\lambda^\nu}(x)$$

for all $x \in U$.

Proof. Let $\bar{x} \in \mathbb{R}^n$. Then for all $\nu \in \{1, \ldots, N\}$ we have $O_\nu(\bar{x}^{-\nu}) \subseteq Z_\nu$ and therefore

$$\begin{aligned} \widehat{V}(x) &= \sum_{\nu=1}^{N} \min_{\lambda^\nu \in Z_\nu} \ell_{\lambda^\nu}(x) \\ &\leq \sum_{\nu=1}^{N} \min_{\lambda^\nu \in O_\nu(\bar{x}^{-\nu})} \ell_{\lambda^\nu}(x) \end{aligned}$$

for all $x \in \mathbb{R}^n$.

In view of (2.3), for the reverse inequality we show for each $\nu \in \{1, \ldots, N\}$ the existence of a neighborhood U_ν of $\bar{x}$ with $O_\nu(x^{-\nu}) \subseteq O_\nu(\bar{x}^{-\nu})$ for all $x \in U_\nu$. In the case $O_\nu(\bar{x}^{-\nu}) = O_\nu$ this is trivially satisfied. Otherwise, choose any $\bar{\lambda}^\nu \in O_\nu \setminus O_\nu(\bar{x}^{-\nu})$. Then, in view of (2.1), we have

$$\langle \bar{\lambda}^\nu, B^\nu \bar{x}^{-\nu} - b^\nu \rangle \;<\; \max_{\lambda^\nu \in O_\nu(\bar{x}^{-\nu})} \langle \lambda^\nu, B^\nu \bar{x}^{-\nu} - b^\nu \rangle.$$

Continuity and the finiteness of the set $O_\nu(\bar{x}^{-\nu})$ ensure the existence of a neighborhood $U_{\bar{\lambda}^\nu}$ of $\bar{x}$ with

$$\begin{aligned} \langle \bar{\lambda}^\nu, B^\nu x^{-\nu} - b^\nu \rangle &< \max_{\lambda^\nu \in O_\nu(\bar{x}^{-\nu})} \langle \lambda^\nu, B^\nu x^{-\nu} - b^\nu \rangle \\ &\leq \max_{\lambda^\nu \in O_\nu} \langle \lambda^\nu, B^\nu x^{-\nu} - b^\nu \rangle \end{aligned}$$

for all $x \in U_{\bar{\lambda}^\nu}$, which by (2.2) means $\bar{\lambda}^\nu \in O_\nu \setminus O_\nu(x^{-\nu})$. We thus have $O_\nu(x^{-\nu}) \subseteq O_\nu(\bar{x}^{-\nu})$ for all x from the set $U_\nu := \bigcap_{\bar{\lambda}^\nu \in O_\nu \setminus O_\nu(\bar{x}^{-\nu})} U_{\bar{\lambda}^\nu}$, and $U := \bigcap_{\nu=1,\ldots,N} U_\nu$ is the asserted neighborhood of $\bar{x}$. •

Player ν's index set of active selection functions $O_\nu(x^{-\nu})$ at $x^{-\nu}$ is of course intimately related to dual information. In fact, for $\nu \in \{1, \ldots, N\}$ and $x^{-\nu} \in \operatorname{dom} X_\nu$ let

$$S_\nu(x^{-\nu}) \;:=\; \{x^\nu \in X_\nu(x^{-\nu}) :\; \langle c^\nu, x^\nu \rangle = \varphi_\nu(x^{-\nu})\}$$

denote the (nonempty) *set of optimal points* of $Q_\nu(x^{-\nu})$ and

$$KKT_\nu(x^{-\nu}) \;:=\; \{\lambda^\nu \in Z_\nu :\; \langle \lambda^\nu, A^\nu y^\nu + B^\nu x^{-\nu} - b^\nu \rangle = 0\}$$

its set of *Karush-Kuhn-Tucker (KKT) multipliers* for any $y^\nu \in S_\nu(x^{-\nu})$. Note that the latter set may be rewritten as

$$KKT_\nu(x^{-\nu}) \;=\; \{\lambda^\nu \in Z_\nu :\; \langle \lambda^\nu, B^\nu x^{-\nu} - b^\nu \rangle = \varphi_\nu(x^{-\nu})\},$$

so that it does not depend on the actual choice of $y^\nu \in S_\nu(x^{-\nu})$.

Proposition 2.2.5 *The set $O_\nu(x^{-\nu})$ is the set of vertices of $KKT_\nu(x^{-\nu})$ for all $\nu \in \{1, \ldots, N\}$ and $x^{-\nu} \in \operatorname{dom} X_\nu$.*

Proof. The dually optimal set $KKT_\nu(x^{-\nu})$ is a face of Z_ν. Thus, the vertex set of $KKT_\nu(x^{-\nu})$ coincides with the set

$$\begin{aligned} KKT_\nu(x^{-\nu}) \cap \operatorname{vert}(Z_\nu) &= \{\lambda^\nu \in O_\nu :\; \langle \lambda^\nu, B^\nu x^{-\nu} - b^\nu \rangle = \varphi_\nu(x^{-\nu})\} \\ &= O_\nu(x^{-\nu}), \end{aligned}$$

where the last equality holds due to (2.2) and the vertex theorem of linear programming. •

For any $\nu \in \{1, \ldots, N\}$ and $x^{-\nu} \in \operatorname{dom} X_\nu$, Proposition 2.2.5 states that $O_\nu(x^{-\nu})$ coincides with the vertex set of $KKT_\nu(x^{-\nu})$ and, thus, establishes a link between the 'primal' index set of active selection functions $O_\nu(x^{-\nu})$ and the 'dual' set of vertex KKT multipliers $\operatorname{vert} KKT_\nu(x^{-\nu})$. More importantly, it shows that the set-valued mapping $O_\nu : \mathbb{R}^{n-n_\nu} \rightrightarrows \mathbb{R}^{m_\nu}$ is a global extension of the set-valued mapping $\operatorname{vert} KKT_\nu : \operatorname{dom} X_\nu \rightrightarrows \mathbb{R}^{n_\nu}$.

According to Propositions 2.2.4 and 2.2.5, the local behavior of $\widehat{V}$ on M and outside of M is governed by vertex KKT multipliers and their global extensions by active index sets of selection functions, respectively. We will heavily exploit this connection in Section 3.1.

2.3 Directional Derivatives

Let $\bar{x} \in \mathbb{R}^n$ and $d \in \mathbb{R}^n$. The (one-sided) directional derivative of $\widehat{V}$ in $\bar{x}$ along d is defined by

$$\widehat{V}'(\bar{x}, d) \;:=\; \lim_{t \searrow 0} \frac{\widehat{V}(\bar{x} + td) - \widehat{V}(\bar{x})}{t}.$$

The following result immediately follows from the additivity of the directional derivative and the formula for directional derivatives of max-functions from, e.g., [14].

Proposition 2.3.1 *Let $\bar{x} \in \mathbb{R}^n$ and $d \in \mathbb{R}^n$. Then $\widehat{V}'(\bar{x}, d)$ exists and we have*

$$\begin{aligned} \widehat{V}'(\bar{x}, d) &= \sum_{\nu=1}^{N} \min_{\lambda^\nu \in O_\nu(\bar{x}^{-\nu})} \langle \nabla \ell_{\lambda^\nu}(\bar{x}), d \rangle \\ &= \sum_{\nu=1}^{N} \langle c^\nu, d^\nu \rangle - \max_{\lambda^\nu \in O_\nu(\bar{x}^{-\nu})} \langle (B^\nu)^T \lambda^\nu, d^{-\nu} \rangle. \end{aligned}$$

Clearly, if $\widehat{V}$ is locally linear in $\bar{x} \in \mathbb{R}^n$, then $\widehat{V}'(\bar{x}, d)$ is a linear function in d.

Let us briefly compare the assertion of Proposition 2.3.1 with a directional differentiability result that we obtained under considerably weaker assumptions in [66, Prop. 3.10, Prop. 3.19]. For its formulation, consider the *set of active indices* of player ν in $\bar{x}$,

$$I_0^\nu(\bar{x}) := \left\{ i \in \{1, \ldots, m_\nu\} : \; A_i^\nu \bar{x}^\nu + B_i^\nu \bar{x}^{-\nu} = b_i^\nu \right\},$$

where A_i^ν is the i-th row of A^ν and B_i^ν the i-th row of B^ν, respectively. Then we may define the *outer linearization cone*

$$\mathcal{L}(\bar{x}, W) := \{ d \in \mathbb{R}^n : \; (A_i^\nu, B_i^\nu) d \leq 0, \; i \in I_0^\nu(\bar{x}), \; \nu \in \{1, \ldots, N\} \}$$

to W in $\bar{x} \in \operatorname{bd} W$, where $\operatorname{bd} W$ denotes the boundary of W. For $\bar{x} \in \operatorname{int} W$ we put $\mathcal{L}(\bar{x}, W) := \mathbb{R}^n$.

Proposition 2.3.2 ([66]) *Let $\bar{x} \in W$ and $d \in \mathcal{L}(\bar{x}, W)$. Then we have*

$$V'(\bar{x}, d) = \sum_{\nu=1}^{N} \langle c^\nu, d^\nu \rangle - \max_{\lambda^\nu \in KKT_\nu(\bar{x}^{-\nu})} \langle (B^\nu)^T \lambda^\nu, d^{-\nu} \rangle.$$

First note that the weaker assumptions from [66] do not imply the existence of a global extension $\widehat{V}$ of V, which explains the restricted choices of $\bar{x}$ and d in the assumption of Proposition 2.3.2. More importantly, in Proposition 2.3.2 the maximizations in the representation of $V'(\bar{x}, d)$ are taken over the whole sets $KKT_\nu(\bar{x}^{-\nu})$, $\nu \in \{1, \ldots, N\}$, while in Proposition 2.3.1 they are taken over the finite sets $O_\nu(\bar{x}^{-\nu})$ which, in view of $W \subseteq M$ and Proposition 2.2.5, coincide with the vertex sets of $KKT_\nu(\bar{x}^{-\nu})$, $\nu \in \{1, \ldots, N\}$. As the sets $KKT_\nu(\bar{x}^{-\nu})$, $\nu \in \{1, \ldots, N\}$, may be unbounded, the correspondence between the assertions of the two propositions in terms of the vertex theorem is not immediate.

2.4 Subdifferential

Despite the fact that $\widehat{V}$ may be a nondifferentiable function it is still possible to compute some generalized gradients of $\widehat{V}$ at each $x \in \mathbb{R}^n$. This enables the application of nonsmooth subgradient methods as we shall see in Part II.

Let us recall the definition of a generalized gradient in the sense of Clarke (cf. [11]) to which we will just refer as *Clarke subgradient* throughout this work. Further, we denote the convex hull of a set A by $\operatorname{conv}(A)$.

Definition 2.4.1 *Let f be a Lipschitz continuous function and denote the set of its nondifferentiability points by ND. Then we define the* Clarke subdifferential *of f at x by*

$$\partial f(x) \ := \ \operatorname{conv}\left(\lim_{x_i \to x,\ x_i \notin ND} \nabla f(x_i) \right)$$

and each element $s \in \partial f(x)$ is called a Clarke subgradient *of f at x.*

Remark 2.4.2 *According to Rademacher's theorem, the set of nondifferentiability points of Lipschitz continuous functions has Lebesgue measure zero. This ensures the possibility to approach each point x with sequences on smooth parts of a Lipschitz continuous function and, thus, the Clarke subdifferential is well-defined for Lipschitz continuous functions.*

Since $\widehat{V}$ is a piecewise linear function, it is also Lipschitz continuous and therefore its Clarke subdifferential is well-defined and can be computed according to Proposition 2.4.3 where we used the Minkowski sum

$$A + B \ := \ \{a + b : \ a \in A,\ b \in B\}$$

in order to simplify the notation.

Proposition 2.4.3 *The Clarke subdifferential of $\widehat{V}$ at $x \in \mathbb{R}^n$ is given by*

$$\partial \widehat{V}(x) \ = \ \sum_{\nu=1}^{N} \left\{ \begin{pmatrix} c^\nu \\ -b^{-\nu} \end{pmatrix} : \ b^{-\nu} \in \operatorname{conv}\left(\left\{ (B^\nu)^T \lambda^\nu : \ \lambda^\nu \in O_\nu(x^{-\nu}) \right\} \right) \right\}.$$

Proof. Recall that we have

$$\widehat{\varphi}_\nu(x^{-\nu}) \ = \ \max_{\lambda^\nu \in O_\nu} \left\langle \lambda^\nu,\ B^\nu x^{-\nu} - b^\nu \right\rangle.$$

The Clarke subdifferential is homogeneous with respect to also negative scalars. Furthermore, convex functions are subdifferentiable regular (cf. [11]) which implies the validity of a sum rule. Altogether, we have

$$\begin{aligned}
\partial\widehat{V}(x) &= \partial\left(\sum_{\nu=1}^{N}\left(\langle c^\nu, x^\nu\rangle - \widehat{\varphi}_\nu(x^{-\nu})\right)\right) \\
&= \partial\left(-\sum_{\nu=1}^{N}\left(\widehat{\varphi}_\nu(x^{-\nu}) - \langle c^\nu, x^\nu\rangle\right)\right) \\
&= -\partial\left(\sum_{\nu=1}^{N}\widehat{\varphi}_\nu(x^{-\nu}) - \langle c^\nu, x^\nu\rangle\right) \\
&= -\left(\sum_{\nu=1}^{N}\partial\widehat{\varphi}_\nu(x^{-\nu}) - \nabla_x\langle c^\nu, x^\nu\rangle\right) \\
&= \sum_{\nu=1}^{N}\left\{\begin{pmatrix} c^\nu \\ -b^{-\nu}\end{pmatrix} :\ b^{-\nu} \in \partial_{x^{-\nu}}\widehat{\varphi}_\nu(x^{-\nu})\right\}
\end{aligned}$$

and the assertion follows from a standard formula for the subdifferential of convex functions which have a representation as pointwise maximum of finitely many smooth functions (cf. [7]). •

In order to clarify the notation, let us assume for the remainder of this section that player ν's strategy set is given by

$$X_\nu(x^{-\nu}) = \{x^\nu \in \mathbb{R}^{n_\nu} |\ A^{\nu\nu}x^\nu + \sum_{\mu\neq\nu} A^{\nu\mu}x^\mu \leq b^\nu\}$$

with suitable matrices $A^{\nu\mu}$ for all $\nu, \mu \in \{1, \ldots, N\}$. Then we restate Proposition 2.4.3 using this representation of player ν's constraints and obtain a new representation of the Clarke subdifferential which is a little bit lengthy but at the same time more precise. Proposition 2.4.4 is of great importance for this work since it provides exact formulas for the subdifferential which we shall exploit heavily by implementing the corresponding nonsmooth numerical methods.

Proposition 2.4.4 *The Clarke subdifferential of $\widehat{V}$ at $x \in \mathbb{R}^n$ is given by*

$$\begin{aligned}\partial\widehat{V}(x) \;=\; & \left\{ \begin{pmatrix} c^1 \\ -(A^{1,2})^T\lambda^1 \\ \vdots \\ -(A^{1,N})^T\lambda^1 \end{pmatrix} : \; \lambda^1 \in \operatorname{conv}\left(O_1(x^{-1})\right) \right\} \\ & + \left\{ \begin{pmatrix} -(A^{2,1})^T\lambda^2 \\ c^2 \\ -(A^{2,3})^T\lambda^2 \\ \vdots \\ -(A^{2,N})^T\lambda^2 \end{pmatrix} : \; \lambda^2 \in \operatorname{conv}\left(O_2(x^{-2})\right) \right\} \\ & + \ldots + \left\{ \begin{pmatrix} -(A^{N,1})^T\lambda^N \\ \vdots \\ -(A^{N,N-1})^T\lambda^N \\ c^N \end{pmatrix} : \; \lambda^N \in \operatorname{conv}\left(O_N(x^{-N})\right) \right\}.\end{aligned}$$

Now, suppose there is an iterative subgradient based method and we aim to minimize $\widehat{V}$. One evaluation of $\widehat{V}$ at a given iterate x requires the computation of N optimal points $\bar{\lambda}^1, \ldots, \bar{\lambda}^N$. For $x \in M$ this is equivalent to solving the linear optimization problems $D_\nu(x^{-\nu})$ for all $\nu \in \{1, \ldots, N\}$. For $x \notin M$ we solve the *globally extended dual optimization problem*

$$\widehat{D_\nu(x^{-\nu})}: \qquad \max_{\lambda^\nu \in \mathbb{R}^{m_\nu}} \langle \lambda^\nu, B^\nu x^{-\nu} - b^\nu \rangle \quad \text{s.t.} \quad \lambda^\nu \in O_\nu.$$

Despite the fact that the computation of the *whole* subdifferential $\partial\widehat{V}$ at a given iterate x may be a difficult task, using these optimal points $\bar{\lambda}^1, \ldots, \bar{\lambda}^N$, the computation of *one* subgradient s^k of $\widehat{V}$ at x^k can be done easily as stated in the following result which follows immediately from Proposition 2.4.4.

Corollary 2.4.5 *Let $x^k \in \mathbb{R}^n$ and $\bar{\lambda}^\nu$ be an optimal point of player ν's (globally extended) dual problem $D_\nu((x^k)^{-\nu})$ for all $\nu \in \{1, \ldots, N\}$. Then we have*

$$\begin{aligned} s^k \;=\; & \begin{pmatrix} c^1 \\ -(A^{1,2})^T\bar{\lambda}^1 \\ -(A^{1,3})^T\bar{\lambda}^1 \\ \vdots \\ -(A^{1,N})^T\bar{\lambda}^1 \end{pmatrix} + \begin{pmatrix} -(A^{2,1})^T\bar{\lambda}^2 \\ c^2 \\ -(A^{2,3})^T\bar{\lambda}^2 \\ \vdots \\ -(A^{2,N})^T\bar{\lambda}^2 \end{pmatrix} + \ldots + \begin{pmatrix} -(A^{N,1})^T\bar{\lambda}^N \\ -(A^{N,2})^T\bar{\lambda}^N \\ \vdots \\ -(A^{N,N-1})^T\bar{\lambda}^N \\ c^N \end{pmatrix} \\ = \; & \begin{pmatrix} c^1 \\ \vdots \\ c^N \end{pmatrix} + \begin{pmatrix} (A^{1,1})^T\bar{\lambda}^1 \\ \vdots \\ (A^{N,N})^T\bar{\lambda}^N \end{pmatrix} - \sum_{\nu=1}^{N} \begin{pmatrix} -(A^{\nu,1})^T\bar{\lambda}^\nu \\ -(A^{\nu,2})^T\bar{\lambda}^\nu \\ \vdots \\ -(A^{\nu,N})^T\bar{\lambda}^\nu \end{pmatrix} \in \partial\widehat{V}(x^k),\end{aligned}$$

that is s^k is a subgradient of $\widehat{V}$ at x^k.

Remark 2.4.6 *Corollary 2.4.5 will play a crucial rule in Chapter 7 where we shall design a nonsmooth subgradient method in order to compute generalized Nash equilibria.*

Corollary 2.4.7 *Let $x \in \mathbb{R}^n$ and $\bar{\lambda}^\nu$ be an optimal point of player ν's (globally extended) dual problem $D_\nu(x^{-\nu})$ for each $\nu \in \{1, \ldots, N\}$. Then we have*

$$\nabla V(x) \;=\; \begin{pmatrix} c^1 \\ \vdots \\ c^N \end{pmatrix} + \begin{pmatrix} (A^{1,1})^T \bar{\lambda}^1 \\ \vdots \\ (A^{N,N})^T \bar{\lambda}^N \end{pmatrix} - \sum_{\nu=1}^{N} \begin{pmatrix} -(A^{\nu,1})^T \bar{\lambda}^\nu \\ -(A^{\nu,2})^T \bar{\lambda}^\nu \\ \vdots \\ -(A^{\nu,N})^T \bar{\lambda}^\nu \end{pmatrix}.$$

for all $x \in \mathbb{R}^n$ such that $\widehat{V}$ is smooth in x.

Example 2.4.8 *In the basic economic market model (cf. Section 1.3.1) at a given $x \in \mathbb{R}^n$ we have*

$$\begin{aligned} V(x) &= \sum_{\nu=1}^{N} \langle -p^\nu, x^\nu \rangle + \min_{\lambda^\nu \in Z_\nu} \left\langle \lambda^\nu, \; \begin{pmatrix} 0_K \\ C^\nu \\ D \end{pmatrix} - \sum_{\mu \neq \nu} \begin{pmatrix} 0_{K+1} \\ x^\mu \end{pmatrix} \right\rangle \\ &= \sum_{\nu=1}^{N} \min_{\lambda^\nu \in Z_\nu} \left\langle \lambda^\nu, \; \begin{pmatrix} 0_K \\ C^\nu \\ D - \sum_{\mu \neq \nu} x^\mu \end{pmatrix} \right\rangle - \langle p^\nu, x^\nu \rangle \\ &= \sum_{\nu=1}^{N} \min_{\lambda^\nu \in Z_\nu} \left\langle \begin{pmatrix} \lambda^\nu_{K+1} \\ \vdots \\ \lambda^\nu_{2K+1} \end{pmatrix}, \; \begin{pmatrix} C^\nu \\ D - \sum_{\mu \neq \nu} x^\mu \end{pmatrix} \right\rangle - \langle p^\nu, x^\nu \rangle \end{aligned}$$

and the set

$$\begin{aligned} Z_\nu &= \left\{ \lambda^\nu \in \mathbb{R}^{2K+1} \;\middle|\; \begin{pmatrix} -I_K \\ e \\ I_K \end{pmatrix}^T \lambda^\nu = p^\nu, \; \lambda^\nu \geq 0 \right\} \\ &= \left\{ \lambda^\nu \in \mathbb{R}^{2K+1} \;\middle|\; \begin{pmatrix} -1 & & 0 & 1 & 1 & & 0 \\ & \ddots & & \vdots & & \ddots & \\ 0 & & -1 & 1 & 0 & & 1 \end{pmatrix} \lambda^\nu = p^\nu, \; \lambda^\nu \geq 0 \right\}. \end{aligned}$$

Therefore, after having computed the optimal points $\bar{\lambda}^{\nu}_{K+2}, \ldots, \bar{\lambda}^{\nu}_{2K+1}$ *of player* ν*'s dual problem* $D_{\nu}(x^{-\nu})$ *for each* $\nu \in \{1, \ldots, N\}$*, we have*

$$s := \begin{pmatrix} -p^1 - \sum_{\nu \neq 1} \begin{pmatrix} \bar{\lambda}^{\nu}_{K+2} \\ \vdots \\ \bar{\lambda}^{\nu}_{2K+1} \end{pmatrix} \\ \vdots \\ -p^N - \sum_{\nu \neq N} \begin{pmatrix} \bar{\lambda}^{\nu}_{K+2} \\ \vdots \\ \bar{\lambda}^{\nu}_{2K+1} \end{pmatrix} \end{pmatrix} \in \partial \widehat{V}(x).$$

Chapter 3

Smoothness and Regularity Conditions

3.1 Definition and Motivation

Recall that the gap function V is not real-valued outside of the set M. For a smoothness analysis of V on $W \subseteq M$, this may cause technical issues at boundary points of W in cases where these also are boundary points of M (cf. Fig. 2.1). Fortunately, we may analyse its global extension $\widehat{V}$ instead. Since $\widehat{V}$ is piecewise linear on $\mathbb{R}^n$, in the following the notion of 'local linearity' is chosen to describe its smoothness properties at a given reference point.

Definition 3.1.1 *We call the extended gap function* $\widehat{V}$ locally linear *in a point* $\bar{x} \in \mathbb{R}^n$ *if and only if there exist an affine linear function* $A : \mathbb{R}^n \to \mathbb{R}$ *and a neighbourhood* U *of* $\bar{x}$ *with* $\widehat{V}(x) = A(x)$ *for all* $x \in U$.

We will use the terms 'smooth' and 'locally linear' synonymously in the following. Analogously, we shall denote $\widehat{V}$ as 'nonsmooth' in a reference point, if it is not locally linear there.

Using Example 2.2.1, we will motivate why smoothness of the gap function is related with regularity conditions of its feasible set which, at first glance, might be a surprising connection. The observations here are rather informal and aim at presenting the main ideas of the following sections. A more stringent and general formulation of these results can be found in Theorem 3.4.9.

Direct inspections show that the nondifferentiability points of V form the set

$$ND := \left[\begin{pmatrix} 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right]$$

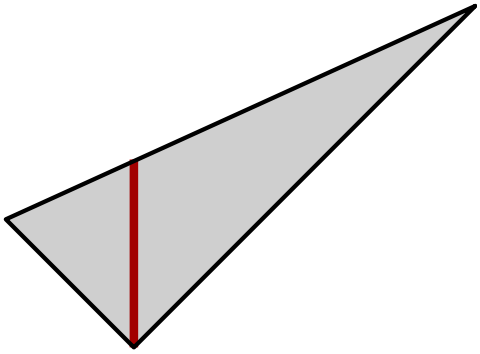

Figure 3.1: The set W and the red colored set of nondifferentiability points of V

as depicted in Figure 3.1. Note that the kinks of V lie on a ray that emerges from a vertex of W while the remaining vertices of W do not 'generate' kinks of V. However, as we shall see in Theorem 3.4.9, the vertex in $(0,-1)^T$ possesses two outstanding properties that are both necessary for 'generating' kinks of V. First, $(0,-1)^T$ is the unfolded optimal point of player two in $X_2(0)$. Second, there is a paraxial ray that emerges from $(0,-1)^T$ and points into the set W. In each of the both remaining vertices of W one of these properties is violated.

The vertex $(0,-1)^T$ of the set W in turn is closely related to regularity conditions in the feasible set of players two: In $(0,-1)^T$ we have *two* active constraints which is caused by the fact that $(0,-1)^T$ is a vertex of a polyhedron in the *two*-dimensional space $\mathbb{R}^2$. This implies that the linear independence constraint qualification is violated in the *one*-dimensional strategy set

$$X_1(0) \;=\; [-1,1]$$

of player one.

In this chapter we shall see that kinks of the gap function V may only occur if some regularity condition is violated in the strategy sets of at least one player, that is, the violation of a regularity condition is *necessary* for the occurrence of nondifferentiability points of V. In Section 3.2, we shall also introduce a new regularity condition, the cone condition, that even *characterizes* the occurrence of nonsmoothness of the gap function V.

3.2 The Cone Condition

From Proposition 2.2.4 we see that locally around any point $\bar{x} \in \mathbb{R}^n$ the function $\widehat{V}$ is constituted as the sum of pointwise minima of affine functions,

indexed by the set of active selection functions $O_\nu(\bar{x}^{-\nu})$. If $O_\nu(\bar{x}^{-\nu})$ is a singleton for all $\nu \in \{1, \ldots, N\}$ then, obviously, $\widehat{V}$ is smooth at $\bar{x}$.

Definition 3.2.1 *For $\nu \in \{1, \ldots, N\}$ and $x^{-\nu} \in \mathbb{R}^{n-n_\nu}$ we say that the* player cone condition (PCC) *is valid in $x^{-\nu}$, iff $O_\nu(x^{-\nu})$ contains at most one element. We say that the* collective cone condition (CCC) *holds in $x \in \mathbb{R}^n$, if PCC holds in $x^{-\nu}$ for all $\nu \in \{1, \ldots, N\}$.*

We emphasize that, as $O_\nu(x^{-\nu})$ is not void, CCC at $x \in \mathbb{R}^n$ actually implies that $O_\nu(x^{-\nu})$ is a singleton for all $\nu \in \{1, \ldots, N\}$.

Remark 3.2.2 *Since, in view of Proposition 2.2.5, for all $\nu \in \{1, \ldots, N\}$ and $x^{-\nu} \in \operatorname{dom} X_\nu$ the set $O_\nu(x^{-\nu})$ coincides with the vertex set of the polyhedron $KKT_\nu(x^{-\nu})$, the player cone condition states for these $x^{-\nu}$ that $KKT_\nu(x^{-\nu})$ possesses at most one vertex or, equivalently, that it is a (translated) convex cone. This explains the terminology in Definition 3.2.1.*

As discussed above, the following result immediately follows from Proposition 2.2.4.

Proposition 3.2.3 *If CCC holds in $\bar{x} \in \mathbb{R}^n$, then $\widehat{V}$ is smooth in $\bar{x}$.*

Next example illustrates the interplay between CCC and smoothness of $\widehat{V}$.

Example 3.2.4 *In Example 2.2.1, by direct inspection it is immediate, that the non-differentiability points of $\widehat{V}$ form the set*

$$ND := \{x \in \mathbb{R}^2 : \ x_1 = 0\}.$$

Since O_1 and, thus, the sets $O_1(x_2)$ are singletons for all $x_2 \in \mathbb{R}$, PCC holds for player one in each $x_2 \in \mathbb{R}$. Furthermore, we have $|O_2(x_1)| > 1$ if and only if $x_1 = 0$, so that PCC holds for player two in $x_1 \in \mathbb{R}$ if and only if $x_1 \neq 0$. Consequently, CCC in $x \in \mathbb{R}^2$ is violated exactly on the set ND and, in this example, CCC even characterizes *the smoothness of $\widehat{V}$.* •

The latter example indicates that the collective cone condition might be a suitable tool to examine the nonsmoothness structure of $\widehat{V}$. Actually, as we will see below, the collective cone condition characterizes the smoothness of $\widehat{V}$ under mild assumptions. Before we show this, we need some auxiliary concepts and results.

For $\nu \in \{1, \ldots, N\}$ and $\lambda^\nu \in O_\nu$ we define the *index set of positive multipliers* of player ν at λ^ν by

$$I_+(\lambda^\nu) := \{i \in \{1, \ldots, m_\nu\} : \ \lambda^\nu_i > 0\}.$$

In the next lemma we obtain an upper bound to the number of non-vanishing multipliers.

Lemma 3.2.5 *Let $\nu \in \{1,\dots,N\}$ and $\bar\lambda^\nu \in O_\nu$. Then we have $|I_+(\bar\lambda^\nu)| \le n_\nu$.*

Proof. Let $\bar\lambda$ be a vertex of $Z_\nu = \{\lambda^\nu \in \mathbb{R}^{m_\nu} :\ c^\nu + (A^\nu)^T\lambda^\nu = 0,\ \lambda^\nu \ge 0\}$. Then, by definition, the rank of the gradients that belong to active constraints equals m_ν. More formally, the rank of the $m_\nu \times (n_\nu + |I_+(\bar\lambda^\nu)^c|)$-matrix $\big(A^\nu \big| e_{I_+(\bar\lambda^\nu)^c}\big)$ is m_ν, where $e_{I_+(\bar\lambda^\nu)^c}$ is a matrix whose columns are the m_ν-dimensional unit vectors e_i, $i \in I_+(\bar\lambda^\nu)^c$. Since the rank of a matrix cannot exceed its number of columns, we have

$$m_\nu \le n_\nu + |I_+(\bar\lambda^\nu)^c| = n_\nu + m_\nu - |I_+(\bar\lambda^\nu)|,$$

which proves the assertion. •

As we will see in Chapter 4.1, the following assumption holds generically, that is, it holds on an open and dense subset of the defining data.

Assumption 3.2.6 *For any $\nu \in \{1,\dots,N\}$ and $J \subseteq \{1,\dots,m_\nu\}$ with $|J| \le n_\nu$ the rows (A^ν_j, B^ν_j), $j \in J$, are linearly independent.*

Note that Assumption 3.2.6 is *unrelated* to LICQ in the unfolded common strategy space W or to player LICQ (cf. Sec. 3.4 below).

Proposition 3.2.7 *Let Assumption 3.2.6 be valid, and let $\widehat V$ be smooth in $\bar x \in \mathbb{R}^n$. Then CCC holds at $\bar x$.*

Proof. Let $\widehat V$ be smooth in $\bar x \in \mathbb{R}^n$. Then due to Proposition 2.3.1 its directional derivative

$$\widehat V'(\bar x, d) \ = \ \sum_{\nu=1}^N \langle c^\nu, d^\nu\rangle - \max_{\lambda^\nu \in O_\nu(\bar x^{-\nu})} \langle (B^\nu)^T\lambda^\nu, d^{-\nu}\rangle$$

is a linear function in d.

As $\widehat V'(\bar x, d)$ is the sum of functions which are concave in d, easy calculations show that each summand $\max_{\lambda^\nu \in O_\nu(\bar x^{-\nu})}\langle (B^\nu)^T\lambda^\nu, d^{-\nu}\rangle$ must be linear in $d^{-\nu}$, that is, there exist vectors $w^\nu \in \mathbb{R}^{n-n_\nu}$ with

$$\max_{\lambda^\nu \in O_\nu(\bar x^{-\nu})} \langle (B^\nu)^T\lambda^\nu, d^{-\nu}\rangle \ = \ \langle w^\nu, d^{-\nu}\rangle, \quad \nu = 1,\dots,N.$$

Now we choose $\nu \in \{1,\dots,N\}$ and $\lambda^\nu \in O_\nu(\bar x^{-\nu})$ arbitrarily, and obtain

$$\langle (B^\nu)^T\lambda^\nu - w^\nu, d^{-\nu}\rangle \ \le \ 0$$

for all $d^{-\nu} \in \mathbb{R}^{n-n_\nu}$, which implies $(B^\nu)^T\lambda^\nu = w^\nu$. Moreover, due to $O_\nu(x^{-\nu}) \subseteq Z_\nu$ we have $(A^\nu)^T\lambda^\nu = -c^\nu$, so that we arrive at

$$(A^\nu, B^\nu)^T\lambda^\nu = \begin{pmatrix} -c^\nu \\ w^\nu \end{pmatrix}. \tag{3.1}$$

Using the submatrix

$$A^\nu_{I_+(\lambda^\nu)} := \begin{pmatrix} \vdots \\ A^\nu_i, \quad i \in I_+(\lambda^\nu) \\ \vdots \end{pmatrix}$$

of A^ν which contains the rows of A^ν corresponding to the positive multipliers at λ^ν, as well as the analogously reduced submatrix $B^\nu_{I_+(\lambda^\nu)}$, the system (3.1) reduces to

$$(A^\nu_{I_+(\lambda^\nu)}, B^\nu_{I_+(\lambda^\nu)})^T\lambda^\nu_{I_+(\lambda^\nu)} = \begin{pmatrix} -c^\nu \\ w^\nu \end{pmatrix}.$$

Due to Lemma 3.2.5 and Assumption 3.2.6, the rows (A^ν_i, B^ν_i), $i \in I_+(\lambda^\nu)$, are linearly independent, and therefore λ^ν is uniquely determined. This implies that $O_\nu(\bar{x}^{-\nu})$ is a singleton, and therefore PCC is valid at $\bar{x}^{-\nu}$ for player ν. As $\nu \in \{1, \ldots, N\}$ was arbitrarily chosen, we have shown that CCC holds at $\bar{x}$. •

Our subsequent main result is a direct consequence of Propositions 3.2.3 and 3.2.7. Since the *non*smoothness points of $\widehat{V}$ are of special interest, we formulate the result as a characterization of their location.

Theorem 3.2.8 *Let Assumption 3.2.6 be valid. Then $\widehat{V}$ is nonsmooth at $\bar{x} \in \mathbb{R}^n$ if and only if CCC is violated at $\bar{x}$.*

Assumption 3.2.6 is nearly trivial, if each player's strategy space is one-dimensional. This yields the next corollary.

Corollary 3.2.9 *Let $n_\nu = 1$ and $(A^\nu_j, B^\nu_j) \neq 0$ for all $j \in \{1, \ldots, m_\nu\}$ and $\nu \in \{1, \ldots, N\}$. Then $\widehat{V}$ is nonsmooth at $\bar{x} \in \mathbb{R}^n$ if and only if CCC is violated at $\bar{x}$.*

Remark 3.2.10 *The condition $(A^\nu_j, B^\nu_j) \neq 0$ from Corollary 3.2.9 is not restrictive, because a constraint with $(A^\nu_j, B^\nu_j) = 0$ either is redundant and can be removed or yields an infeasible LGNEP, where the latter case would contradict Assumption 1.2.7.*

Note that Corollary 3.2.9 explains, in particular, the observation from Example 3.2.4.

3.3 Strict Mangasarian Fromovitz Condition

While the collective cone condition captures nonsmoothness very sharply, it is rather hard to verify, so that it might sometimes be more convenient to work with a different regularity condition instead.

In view of Proposition 2.2.5, for all $\nu \in \{1, \ldots, N\}$ and $x^{-\nu} \in \operatorname{dom} X_\nu$ the set $O_\nu(x^{-\nu})$ coincides with the vertex set of $KKT_\nu(x^{-\nu})$. Hence, every condition that implies unique KKT multipliers at $x \in M$ will also imply smoothness of the extended gap function $\widehat{V}$ there. According to [48] it is possible even to characterize unique KKT multipliers by the strict Mangasarian Fromovitz condition, *if KKT multipliers exist at all*. While we will not use this set of conditions explicitly, it gives rise to the following notion.

Definition 3.3.1 *For $\nu \in \{1, \ldots, N\}$ and $x^{-\nu} \in \operatorname{dom} X_\nu$ we say that the* player strict Mangasarian Fromovitz condition (PSMFC) *is valid in $x^{-\nu}$, iff $KKT_\nu(x^{-\nu})$ contains at most one element. We say that the* collective strict Mangasarian Fromovitz condition (CSMFC) *holds in $x \in M$, if PSMFC holds in $x^{-\nu}$ for all $\nu \in \{1, \ldots, N\}$.*

Remark 3.3.2 *Although Definition 3.3.1 does not explicitly involve the notion of an optimal point $y^\nu \in S_\nu(x^{-\nu})$ to which the set $KKT_\nu(x^{-\nu})$ corresponds, it is well-defined, since Assumption 1.2.7 guarantees $S_\nu(x^{-\nu}) \neq \emptyset$ for any $x^{-\nu} \in \operatorname{dom} X_\nu$, and the set of KKT multipliers does not depend on the actual choice $y^\nu \in S_\nu(x^{-\nu})$.*

Remark 3.3.3 *We refrain from calling PSMFC a* constraint qualification*, since, first, it is not imposed* only *on the constraints and, second, it does not guarantee the existence of KKT multipliers at a local optimal point. The weakest constraint qualification that implies existence* and *uniqueness of KKT multipliers for all objective functions is the linear independence constraint qualification (cf. [71]).*

In our linear setting the Abadie constraint qualification and Assumption 1.2.7 ensure $KKT_\nu(x^{-\nu}) \neq \emptyset$ for all $x^{-\nu} \in \operatorname{dom} X_\nu$ and $\nu \in \{1, \ldots, N\}$. Hence, CSMFC is valid at $x \in M$ if and only if $KKT_\nu(x^{-\nu})$ is a singleton for all $\nu \in \{1, \ldots, N\}$.

The regularity condition CSMFC is sufficient for smoothness of the global extension of the gap function $\widehat{V}$, as we will see in the following proposition. We state the result for $\widehat{V}$ instead of V, because at boundary points of the domain M smoothness of V is not defined.

Proposition 3.3.4 *If CSMFC holds at $\bar{x} \in M$, then $\widehat{V}$ is smooth in $\bar{x}$.*

Proof. The CSMFC at $\bar{x} \in M$ implies unique KKT multipliers, that is, the set $KKT_\nu(\bar{x}^{-\nu})$ is a singleton for each $\nu \in \{1, \ldots, N\}$. This implies CCC at $\bar{x}$, so that the assertion follows from Proposition 3.2.3. •

According to Proposition 3.3.4, CSMFC is sufficient for smoothness of $\widehat{V}$ at $\bar{x} \in M$, but the following example shows that CSMFC is not *necessary*.

Example 3.3.5 *In Example 2.2.1, easy calculations show that PSMFC is violated for player one if and only if* $x_2 = -1$ *or* $x_2 = 3$. *The associated KKT multipliers are given by*

$$KKT_1(-1) \;=\; \left\{ \begin{pmatrix} 0 \\ 1+t \\ t \end{pmatrix},\; t \geq 0 \right\}$$

and

$$KKT_1(3) \;=\; \left\{ \begin{pmatrix} t \\ 1+\frac{1}{2}t \\ 0 \end{pmatrix},\; t \geq 0 \right\}.$$

Analogously, we obtain that PSMFC is violated for player two if and only if $x_1 = -\frac{4}{3}$, $x_1 = 0$ *or* $x_1 = 4$. *The corresponding KKT multipliers are*

$$KKT_2\left(-\frac{4}{3}\right) \;=\; \left\{ \begin{pmatrix} t \\ 0 \\ 1+t \end{pmatrix},\; t \geq 0 \right\}$$

as well as

$$KKT_2(0) \;=\; \left\{ \begin{pmatrix} 0 \\ 1-t \\ t \end{pmatrix},\; t \in [0,1] \right\} \;=\; \left[\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},\; \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right]$$

and

$$KKT_2(4) \;=\; \left\{ \begin{pmatrix} t \\ 1+t \\ 0 \end{pmatrix},\; t \geq 0 \right\}.$$

Altogether, CSMFC is violated exactly in the boundary points of M *and on the line segment* $\left[\begin{pmatrix} 0 \\ -1 \end{pmatrix},\; \begin{pmatrix} 0 \\ 3 \end{pmatrix} \right]$.

The nondifferentiability points of $\widehat{V}$ *in the set* M *form the line segment*

$$ND \;:=\; \left[\begin{pmatrix} 0 \\ -1 \end{pmatrix},\; \begin{pmatrix} 0 \\ 3 \end{pmatrix} \right],$$

that is, they are 'created' by the violation of PSMFC of player two in $x_1 = 0$, *whereas the other points where PSMFC is violated do* not *affect the smoothness of* $\widehat{V}$. *This effect will be further pursued in Theorem 3.3.10 below.*

Note that the phenomena in this example are stable under small perturbations of the defining data, so that not even under generic conditions we may expect to characterize smoothness of $\widehat{V}$ *via* CSMFC . However, if $\bar{x}$ is chosen from the topological interior of W, then, under the generic Assumption 3.3.6, CSMFC is not only sufficient but also necessary for smoothness of V in $\bar{x}$, as we shall see in the following result.

Assumption 3.3.6 *The rows (A_i^ν, B_i^ν), $i \in I_0^\nu(x)$, are linearly independent for any $\nu \in \{1, \dots, N\}$ and $x \in \mathbb{R}^n$ with $A^\nu x^\nu + B^\nu x^{-\nu} \leq b^\nu$.*

Proposition 3.3.7 *Let Assumption 3.3.6 be valid, and let V be smooth in $\bar{x} \in \operatorname{int} W$. Then CSMFC holds at $\bar{x}$.*

Proof. Let V be smooth in $\bar{x} \in \operatorname{int} W$. Then Proposition 2.3.2 implies that its directional derivative

$$V'(\bar{x}, d) \;=\; \sum_{\nu=1}^{N} \langle c^\nu, d^\nu \rangle - \max_{\lambda^\nu \in KKT_\nu(\bar{x}^{-\nu})} \langle (B^\nu)^T \lambda^\nu, d^{-\nu} \rangle$$

is a linear function in d. Following the lines of the proof of Proposition 3.2.7, we obtain that for each $\nu \in \{1, \dots, N\}$ there exists a vector $w^\nu \in \mathbb{R}^{n-n_\nu}$ with

$$(A^\nu, B^\nu)^T \lambda^\nu \;=\; \begin{pmatrix} -c^\nu \\ w^\nu \end{pmatrix} \tag{3.2}$$

for all $\lambda^\nu \in KKT_\nu(\bar{x}^{-\nu})$. Let $y^\nu \in X_\nu(\bar{x}^{-\nu})$. Due to the complementarity condition we have $\lambda_i^\nu = 0$ for all $i \notin I_0^\nu(y^\nu, \bar{x}^{-\nu})$, and therefore system (3.2) reduces to

$$(A^\nu_{I_0^\nu(y^\nu, \bar{x}^{-\nu})}, B^\nu_{I_0^\nu(y^\nu, \bar{x}^{-\nu})})^T \lambda^\nu_{I_0^\nu(y^\nu, \bar{x}^{-\nu})} \;=\; \begin{pmatrix} -c^\nu \\ w^\nu \end{pmatrix}.$$

Finally, due to Assumption 3.3.6, the latter system of equations determines the multipliers λ^ν uniquely, which implies the validity of CSMFC at $\bar{x}$. •

In analogy to Theorem 3.2.8 we may, thus, characterize the nonsmoothness of V at interior points of W by CSMFC. Recall that, in contrast, Theorem 3.2.8 characterizes the nonsmoothness of $\widehat{V}$ at arbitrary points by CCC.

Theorem 3.3.8 *Let Assumption 3.3.6 be valid. The mapping V is nonsmooth at $\bar{x} \in \operatorname{int} W$ if and only if CSMFC is violated at $\bar{x}$.*

It is possible to extend the result from Proposition 3.3.7 to certain boundary points of W as we will see in the next result which, as discussed above, we state for $\widehat{V}$ instead of V, because smoothness of V is not defined at boundary points of M.

Proposition 3.3.9 *Let Assumption 3.3.6 be valid and let $\widehat{V}$ be smooth in $\bar{x} \in W$. Then for all $\nu \in \{1, \ldots, N\}$ such that PSMFC is violated at $\bar{x}^{-\nu}$ we have $(x^\nu, \bar{x}^{-\nu}) \notin \operatorname{int} W$ for all $x^\nu \in X_\nu(\bar{x}^{-\nu})$.*

Proof. We prove the assertion by contraposition. Let $\bar{x} \in W$ and $\nu \in \{1, \ldots, N\}$ such that PSMFC is violated at $\bar{x}^{-\nu}$ and $x^\nu \in X_\nu(\bar{x}^{-\nu})$ with $(x^\nu, \bar{x}^{-\nu}) \in \operatorname{int} W$. Furthermore, for $\lambda \in (0, 1]$ we define

$$x(\lambda) \;:=\; (1-\lambda)\bar{x} + \lambda(x^\nu, \bar{x}^{-\nu}) \;=\; ((1-\lambda)\bar{x}^\nu + \lambda x^\nu, \bar{x}^{-\nu}).$$

Then we have $x(\lambda) \in \operatorname{int} W$ for all $\lambda \in (0, 1]$ (cf. [59, Th. 6.1]), and PSMFC is violated at $x(\lambda)^{-\nu} = \bar{x}^{-\nu}$. According to Proposition 3.3.7, the function $\widehat{V}$ is not smooth in $x(\lambda)$ for all $\lambda \in (0, 1]$. By a standard argument from calculus, this also holds for $\lambda = 0$, that is, $\widehat{V}$ is not smooth in $\bar{x}$. •

The next result follows from Propositions 3.3.4 and 3.3.9.

Theorem 3.3.10 *Let Assumption 3.3.6 be valid and let $\bar{x} \in W$. Furthermore, if PSMFC is violated for some player $\nu \in \{1, \ldots, N\}$ at $\bar{x}^{-\nu}$, let there exist some $x^\nu \in X_\nu(\bar{x}^{-\nu})$ with $(x^\nu, \bar{x}^{-\nu}) \in \operatorname{int} W$. Then $\widehat{V}$ is nonsmooth at $\bar{x}$ if and only if CSMFC is violated at $\bar{x}$.*

The latter theorem has an interesting interpretation for one-dimensional strategy sets, that indicates why not all violations of CSMFC in Example 3.3.5 enforce nonsmoothness of $\widehat{V}$ in $\bar{x}$: Paraxial rays that emerge from optimal points in kinks of the boundary of the set W cause kinks of the function V, if these rays point into the *interior* of W.

3.4 Linear Independence Constraint Qualification

The following constraint qualification is the strongest common regularity condition, but has the advantage that its verification is an easy task. Again, we distinguish between a player regularity condition that may hold in his strategy set and a collective regularity condition that acts on the unfolded common strategy set.

Definition 3.4.1 *For $\nu \in \{1, \ldots, N\}$ and $x^{-\nu} \in \operatorname{dom} X_\nu$ we say that the* player linear independence constraint qualification (PLICQ) *holds in $x^{-\nu}$, iff for some $y^\nu \in S(x^{-\nu})$ the vectors A_i^ν, $i \in I_0^\nu(y^\nu, x^{-\nu})$, are linearly independent. We say that the* collective linear independence constraint qualification (CLICQ) *holds in $x \in M$, if PLICQ holds in $x^{-\nu}$ for all $\nu \in \{1, \ldots, N\}$.*

It is a well-known fact that PLICQ enforces PSMFC, so that the next result immediately follows from Proposition 3.3.4.

Proposition 3.4.2 *If CLICQ holds at $\bar{x} \in M$, then $\widehat{V}$ is smooth in $\bar{x}$.*

While CLICQ implies CSMFC, both conditions even coincide under a mild assumption, as we shall see in Proposition 3.4.6.

Assumption 3.4.3 *For any $\nu \in \{1, \dots, N\}$ and $x^{-\nu} \in \operatorname{dom} X_\nu$, all $y^\nu \in S_\nu(x^{-\nu})$ and all $J \subseteq I_0^\nu(y^\nu, x^{-\nu})$ with $|J| \leq n_\nu$ the rows A_j^ν, $j \in J$, are linearly independent.*

Remark 3.4.4 *Assumption 3.4.3 is unrelated to Assumption 3.2.6 and Assumption 3.3.6.*

Remark 3.4.5 *For $\nu \in \{1, \dots, N\}$ let PLICQ be violated at $x^{-\nu} \in \operatorname{dom} X_\nu$. Then Assumption 3.4.3 implies $|I_0^\nu(y^\nu, x^{-\nu})| > n_\nu$ for all $y^\nu \in S_\nu(x^{-\nu})$. This is one main reason why the violation of PLICQ is sufficient for the existence of kinks that are related with the existence of 'too many' active constraints.*

Proposition 3.4.6 *Let Assumption 3.4.3 be valid, let $\nu \in \{1, \dots, N\}$ and let $\bar{x}^{-\nu} \in \operatorname{dom} X_\nu$. Then PSMFC at $\bar{x}^{-\nu}$ implies PLICQ at $\bar{x}^{-\nu}$.*

Proof. We prove the assertion by contraposition. Let $\nu \in \{1, \dots, N\}$ and $\bar{x}^{-\nu} \in \operatorname{dom} X_\nu$, such that PLICQ is violated at $\bar{x}^{-\nu}$.

On the one hand, the strong theorem of complementarity (cf. [12, Th. A.7]) yields the existence of an optimal point $\bar{y}^\nu \in S_\nu(\bar{x}^{-\nu})$ with positive KKT multipliers $\lambda_i > 0$, $i \in I_0(\bar{y}^\nu, \bar{x}^{-\nu})$. Therefore, due to Remark 3.4.5, we have at least $n_\nu + 1$ positive scalars $\lambda_i > 0$ with

$$-(c^\nu)^T = \sum_{i \in I_0(\bar{y}^\nu, \bar{x}^{-\nu})} \lambda_i A_i^\nu. \tag{3.3}$$

On the other hand, Carathéodory's theorem states, that the conic representation from (3.3) is also available with *at most* n_ν positive multipliers. Therefore, there are two different sets of KKT multipliers, and PSMFC is violated at $\bar{x}^{-\nu}$. •

We summarize our observations in the following result.

Theorem 3.4.7 *Let Assumption 3.4.3 be valid and let $\bar{x} \in M$. Then CLICQ is valid at $\bar{x}$ if and only if CSMFC is valid at $\bar{x}$.*

Theorem 3.4.7 allows us to restate the Theorems 3.3.8 and 3.3.10 in terms of CLICQ which is advantageous, because CLICQ is easier to verify than CSMFC or even CCC.

Theorem 3.4.8 *Let Assumptions 3.3.6 and 3.4.3 be valid. Then V is nonsmooth at $\bar{x} \in \operatorname{int} W$ if and only if CLICQ is violated at $\bar{x}$.*

Theorem 3.4.9 *Let Assumptions 3.3.6 and 3.4.3 be valid and let $\bar{x} \in W$. Furthermore, if PLICQ is violated for some player $\nu \in \{1, \ldots, N\}$ at $\bar{x}^{-\nu}$, let there exist some $x^\nu \in X_\nu(\bar{x}^{-\nu})$ with $(x^\nu, \bar{x}^{-\nu}) \in \operatorname{int} W$. Then $\widehat{V}$ is nonsmooth at $\bar{x}$ if and only if CLICQ is violated at $\bar{x}$.*

Note that Theorem 3.4.9 provides an explanation for the connection between some vertices of W and kinks in the graph of V that we have noticed in Section 3.1 and recall in the following figure. In Figure 3.2 all assumptions

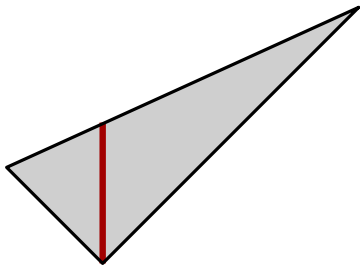

Figure 3.2: The red colored set of nondifferentiability points of V

of Theorem 3.4.9 are valid, such that V is nonsmooth in x if and only if the collective linear independence constraint qualification is violated in x which clearly exhibits the connection between a regularity condition on the set W and nonsmoothness of the function V.

Chapter 4

Further results

4.1 Genericity

The notion of genericity is a powerful concept to distinguish mild from strong assumptions. In this section we will show that our Assumptions 3.2.6, 3.3.6 and 3.4.3 are mild in the sense that they hold generically.

We identify an instance of an LGNEP with the data tuples $(c^\nu, A^\nu, B^\nu, b^\nu)$ in $\mathbb{R}^{n_\nu + m_\nu \cdot n + m_\nu}$, $\nu \in \{1, \dots, N\}$, and say that an assumption $\mathcal{A}$ holds generically, if the set of all tuples $(c^\nu, A^\nu, B^\nu, b^\nu)$, $\nu \in \{1, \dots, N\}$, such that $\mathcal{A}$ is valid constitutes a set that is open and dense in the space $\mathbb{R}^{n+(n+1)m}$, where we put $m := \sum_{\nu=1}^{N} m_\nu$. The openness yields that a generic property $\mathcal{A}$ is stable under small perturbations of the defining data $(c^\nu, A^\nu, B^\nu, b^\nu)$, $\nu \in \{1, \dots, N\}$.

Remark 4.1.1 *Note that sufficient conditions for smoothness like CCC, CSMFC and CLICQ, of course, do* not *hold generically everywhere. This corresponds to the fact that there are nondifferentiability points of the gap function which do* not *vanish under small perturbations of the data.*

In order to show the genericity of an assumption $\mathcal{A}$, we will prove that the set of tuples with the respective undesired properties lie in the union of finitely many smooth manifolds with positive codimensions. Before we start with the proofs we need some definitions and results from [43, 63], which will be our main tools in the subsequent genericity proofs.

For $M, N, R \in \mathbb{N}$ with $R \leq \min(M, N)$ we denote the set of (M, N)-matrices of rank R by

$$\mathbb{R}_R^{M \times N} := \{A \in \mathbb{R}^{M \times N} : \operatorname{rank}(A) = R\}$$

and for $L \subseteq \{1,\dots,M\}$ and $\max(R+|L|-M,0) \le S \le \min(R,|L|)$, we define

$$R^{M\times N}_{R,L,S} \;:=\; \{A \in \mathbb{R}^{M\times N}_R :\; A_{(L)} \in \mathbb{R}^{(M-|L|)\times N}_{R-S}\},$$

where the matrix $A_{(L)}$ results from A by deletion of the rows with indices in L. The above restrictions on S follow from the relations $0 \le R-S \le M-|L|$ and $R-|L| \le R-S \le R$.

Proposition 4.1.2 *a) The set $\mathbb{R}^{M\times N}_R$ is a smooth manifold of codimension $(M-R)\cdot(N-R)$ in $\mathbb{R}^{M\times N}$.*

b) The set $\mathbb{R}^{M\times N}_{R,L,S}$ is a smooth manifold of codimension $(M-R)\cdot(N-R)+S\cdot(M-R+S-|L|)$ in $\mathbb{R}^{M\times N}$.

Proof. The proof of part a) can by found in [43]. For part b) see [63]. •

Proposition 4.1.3 *The Assumptions 3.2.6, 3.3.6 and 3.4.3 hold generically.*

Proof. First, we show the genericity of Assumption 3.2.6. As the splitting of the involved matrices in A- and B-parts is irrelevant for this proof, we show that the set of $(m,n)-$matrices A such that Assumption 3.2.6 is violated, lies in the finite union of smooth manifolds with positive codimensions in $\mathbb{R}^{m\times n}$.

In fact, let $A \in \mathbb{R}^{m\times n}$ be a matrix such that Assumption 3.2.6 is violated. Then, for some $\nu \in \{1,\dots,N\}$ and a submatrix $A^\nu \in \mathbb{R}^{m_\nu\times n}$ of A there exists a set $J \subseteq \{1,\dots,m_\nu\}$ with $|J| \le n_\nu$ such that the rows of A^ν_J are linearly dependent. Let us define $r_\nu := \operatorname{rank}(A^\nu)$. In the case $r_\nu < \min(m_\nu,n)$, according to Proposition 4.1.2a), the matrix A^ν lies in a smooth manifold of codimension $(m_\nu-r_\nu)\cdot(n-r_\nu) > 0$.

On the other hand, let $r_\nu = \min(m_\nu,n)$. Due to the trivial bounds $|J| \le m_\nu$ and $n_\nu \le n$ we obtain $|J| \le \min(m_\nu,n)$ and, thus,

$$\operatorname{rank} A^\nu_J \;<\; |J| \;\le\; \min(m_\nu,n) \;=\; r_\nu.$$

We define $s_\nu := r_\nu - \operatorname{rank} A^\nu_J > 0$ and notice that, according to Proposition 4.1.2b), the matrix A^ν lies in a smooth manifold of codimension

$$\begin{aligned}
\underbrace{(m_\nu-r_\nu)}_{=0}\cdot(n-r_\nu)+s_\nu\cdot(m_\nu-r_\nu+s_\nu-|J^c|) &= s_\nu\cdot(s_\nu-r_\nu+|J|)\\
&= s_\nu\cdot\underbrace{(|J|-\operatorname{rank} A_J)}_{>0}\\
&> 0.
\end{aligned}$$

Since the possible choices of ν, r_ν and s_ν in both cases only yield finitely many manifolds, the matrices that do not fulfill Assumption 3.2.6 lie in the

finite union of smooth manifolds with positive codimensions, and therefore the desired property holds generically.

Assumption 3.3.6 just states that LICQ holds everywhere in the set $\{x \in \mathbb{R}^n : A^\nu x^\nu + B^\nu x^{-\nu} \leq b^\nu\}$ for any $\nu \in \{1, \ldots, N\}$. It is well-known that this property holds generically (cf. [58]).

To show the genericity of Assumption 3.4.3, as in the genericity proof for Assumption 3.2.6, we show that the set of data where it is violated lies in the finite union of smooth manifolds with positive codimensions. In fact, if Assumption 3.4.3 is violated, there exist some $\nu \in \{1, \ldots, N\}$, $x^{-\nu} \in \operatorname{dom} X_\nu$, $y^\nu \in S_\nu(x^{-\nu})$ as well as $J \subseteq I_0^\nu(y^\nu, x^{-\nu})$ with $|J| \leq n_\nu$ such that $\operatorname{rank} A_J^\nu < |J|$ for a submatrix $A^\nu \in \mathbb{R}^{m_\nu \times n_\nu}$ of the data. After dropping the dependence of this condition on $(y^\nu, x^{-\nu})$ and replacing the set $I_0^\nu(y^\nu, x^{-\nu})$ by the larger set $\{1, \ldots, m_\nu\}$, along the lines of the genericity proof of Assumption 3.2.6, one can easily show that also Assumption 3.4.3 holds generically. •

4.2 Error bounds

One way to compute Nash equilibria is to minimize the gap function V over the common strategy set W as stated in Proposition 1.5.1. Since the equilibria are exactly the vectors in W with optimal value zero, a reasonable numerical stopping criterion would be $x \in W$ and $V(x) < \varepsilon$ with a small positive number ε. By doing so we implicitly assume that the distance of an iterate x to the set of all Nash equilibria S is small for small values of ε. This is an issue that is examined within the theory of error bounds (see [57] for an article that may serve as starting point to this field of research).

Let $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ be an arbitrary norm on $\mathbb{R}^n$ and S the set of all Nash equilibria in a given LGNEP. Since V is a nonconvex function we can only expect to obtain a *local* error bound, that is, the existence of a Nash equilibrium $\bar{x}$ and positive scalars $\ell(\bar{x}) > 0$ and $\delta > 0$ with

$$\operatorname{dist}(x, S) \cdot \ell(\bar{x}) \;\leq\; V(x)$$

for all $x \in B(\bar{x}, \delta) \cap W$ where we denote the closed unit ball with radius δ and center $\bar{x}$ with respect to $\|\cdot\|$ by $B(\bar{x}, \delta)$ and define

$$\operatorname{dist}(\bar{x}, S) \;:=\; \inf_{x \in S} \|\bar{x} - x\|.$$

However, error bounds for GNEPs are very rare and we are only aware of [22, 42] where the authors derive error bounds for nonlinear GNEPs in order to apply an interior point method for the computation of Nash equilibria. Therefore, Theorem 4.2.2 seems to be the first result on error bounds related

to the gap function of LGNEPs. However, Theorem 4.2.2 considers only the case of *isolated* Nash equilibria.

Definition 4.2.1 *A Nash equilibrium $\bar{x}$ is called* isolated *if there exists an neighborhood U of $\bar{x}$ with $V(x) > 0$ for all $x \in (U \cap W) \setminus \{\bar{x}\}$.*

Theorem 4.2.2 *Let $\bar{x} \in W$ be an isolated Nash equilibrium. Then there exist scalars $\ell(\bar{x}) > 0$ and $\delta > 0$ with*

$$\ell(\bar{x}) \cdot \operatorname{dist}(x, S) \;\leq\; V(x)$$

for all $x \in B(\bar{x}, \delta) \cap W$.

Proof. Let $\bar{x} \in W$ be an arbitrary Nash equilibrium. Then there exists an $\delta > 0$, such that V can be expressed as pointwise minimum of finitely many active selection functions, which implies that for all $x \in B(\bar{x}, \delta) \cap W$ we have

$$\begin{aligned} V(x) &= \underbrace{V(\bar{x})}_{=0} + V'(\bar{x}, x - \bar{x}) \\ &= V'(\bar{x}, x - \bar{x}) \\ &= \|x - \bar{x}\| \cdot V'(\bar{x}, \frac{x - \bar{x}}{\|x - \bar{x}\|}). \end{aligned}$$

This means we arrive at

$$\begin{aligned} V(x) &= V'\left(\bar{x}, \frac{x - \bar{x}}{\|x - \bar{x}\|}\right) \cdot \underbrace{\|x - \bar{x}\|}_{\geq \operatorname{dist}(x,S)} \\ &\geq V'\left(\bar{x}, \frac{x - \bar{x}}{\|x - \bar{x}\|}\right) \cdot \operatorname{dist}(x, S) \\ &\geq \min_{d \in \mathcal{L}(\bar{x}, W) \cap B_=(0,1)} V'(\bar{x}, d) \cdot \operatorname{dist}(x, S) \end{aligned}$$

where $\mathcal{L}(\bar{x}, W)$ is the outer linearization cone in $\bar{x}$ at W and $B_=(0, 1)$ is the unit sphere.

Let us define

$$\ell(\bar{x}) \;:=\; \min_{d \in \mathcal{L}(\bar{x}, W) \cap B_=(0,1)} V'(\bar{x}, d).$$

Since $\bar{x}$ is a Nash equilibrium and therefore a global minimal point of V we arrive at $\ell(\bar{x}) \geq 0$. Assume that we have $\ell(\bar{x}) = 0$. This implies the existence of a direction $\bar{d} \in \mathcal{L}(\bar{x}, W) \cap B_=(0, 1)$ and a scalar $\bar{t} > 0$ with

$$V\left(\bar{x} + t\bar{d}\right) \;=\; 0$$

for all $t \in [0, \bar{t}]$ which is a contradiction becase $\bar{x}$ is an *isolated* Nash equilibrium. Therefore, we have $\ell(\bar{x}) > 0$ which shows the assertion. •

However, this is only a first result concerning the theory of error bounds applied to the gap function of LGNEPs and especially the case of nonisolated Nash equilibria deserves special attention which we leave for future research.

Part II

Algorithms

Chapter 5

Related Literature and Overview

This part ist based on the article [26] of this author.

In the literature on nonlinear generalized Nash equilibrium problems, the following reformulation techniques have been used to tackle GNEPs by existing algorithms:

- The solution of the concatenated Karush-Kuhn-Tucker (KKT) conditions of all players which was used to discuss local convergence properties for suitable Newton methods in [30, 56], and in [22, 23] to develop a robust potential reduction algorithm for GNEPs.
- The use of the Nikaido-Isoda function in order to get a constrained or unconstrained, typically nonsmooth, optimization reformulation of the GNEP (cf. [24, 25, 39]). Note that the Nikaido-Isoda function has also been used to design relaxation methods like, e.g., in [40, 46, 68].
- A transformation of GNEPs into easier Nash equilibrium problems with the help of penalty terms as, e.g., in [28, 31, 33].
- A quasi-variational inequality reformulation for GNEPs (see, e.g., [47]).

In this part, we compare methods which are based on the first three approaches described above. We do not use the quasi-variational inequality reformulation, since in our opinion there is a lack of efficient numerical methods for this problem class and the recent promising algorithm from [29] is similar to the potential reduction algorithm we will consider.

To be more specific, in Chapter 6, we examine the interior point method PRA that was originally designed for the computation of Nash equilibria in nonlinear GNEPs. Since these convergence results are not available in the

linear setting, in Theorem 6.2.1 we present some new convergence conditions for PRA that are tailored to the linear case.

In Chapter 7, we examine some subgradient methods that solve LGNEPs by minimizing the nonsmooth function V over the unfolded common strategy set W. Unfortunately, we were not able to prove convergence for the resulting methods but at least it is possible to show that V possesses no local minimal points in the basic economic market which indicates the promising performance of the subgradient methods.

In Chapter 8, we report numerical results that we obtain from applying PRA and the subgradient methods to some test examples. Due to the absence of test examples of LGNEPs in the existing literature we test the algorithms on different instances of the basic economic market model. Even for test instances with more than ten thousand variables the numerical results are very convincing.

In order to obtain a more detailed representation of the LGNEP we introduce matrices $A^{\nu\mu} \in \mathbb{R}^{m_\nu \times n_\mu}$ and vectors $c^\nu \in \mathbb{R}^{n_\nu}$, $b^\nu \in \mathbb{R}^{m_\nu}$, such that player ν's optimization problem is given by

$$Q_\nu(x^{-\nu}): \qquad \min_{x^\nu \in \mathbb{R}^{n_\nu}} \langle c^\nu, x^\nu \rangle \quad \text{s.t.} \quad A^{\nu\nu}x^\nu + \sum_{\mu \neq \nu} A^{\nu\mu}x^\mu \leq b^\nu,$$

for all $\nu \in \{1, \ldots, N\}$ where the decisions of the remaining players $x^\mu \in \mathbb{R}^{m_\mu}$, $\mu \neq \nu$, are given. Thus, player ν's strategy set is denoted by

$$X_\nu(x^{-\nu}) := \left\{ x^\nu \in \mathbb{R}^{n_\nu} : \; A^{\nu\nu}x^\nu \leq \; b^\nu - \sum_{\mu \neq \nu} A^{\nu\mu}x^\mu \right\}.$$

for all $\nu \in \{1, \ldots, N\}$.

Chapter 6

Interior Point Method

In this section, we will discuss two algorithms that are able to compute generalized Nash equilibria by solving the concatenated KKT conditions as described in Section 1.5.2.

6.1 Concatenated KKT Systems

As mentioned in Section 1.5.2, $\bar{x}$ is a generalized Nash equilibrium, if and only if there exist $\lambda^\nu, w^\nu \in \mathbb{R}^{m_\nu}$ satisfying the KKT conditions

$$\begin{aligned} c^\nu + (A^{\nu\nu})^T \lambda^\nu &= 0, \\ (A^{\nu 1}\, A^{\nu 2}\, \dots\, A^{\nu N})\, \bar{x} - b^\nu + w^\nu &= 0, \\ w^\nu \geq 0, \quad \lambda^\nu \geq 0, \quad (w^\nu)^T \lambda^\nu &= 0 \end{aligned}$$

for all $\nu \in \{1, \dots, N\}$.

Remark 6.1.1 *In LGNEPs, the Abadie constraint qualification is always valid in the strategy sets of all players which implies that solutions of the concatenated KKT conditions are exactly the Nash equilibria* without *posing any further assumptions. This is an important difference to nonlinear GNEPs where we have to assume further player constraint qualifications, as, e.g., player LICQ, in order to obtain a one to one correspondence between solutions of the KKT systems and Nash equilibria. However, these player constraint qualifications are* not *mild assumptions since they may be violated in a stable way which yields theoretical and numerical difficulties as mentioned in Part I of this work. The stable violation of constraint qualification is caused by the fact that each player faces a* parametric *optimization problem (cf. [43]).*

The above system is equivalent to the constrained system

$$H(x, \lambda, w) = 0, \qquad (x, \lambda, w) \in \mathbb{R}^n \times \mathbb{R}^m_+ \times \mathbb{R}^m_+. \tag{6.1}$$

with

$$H(x, \lambda, w) \ := \ \begin{pmatrix} c + E\lambda \\ A\bar{x} - b + w \\ \lambda \circ w \end{pmatrix}$$

as we have seen in Proposition 1.5.2 for suitable defined matrices E and A and vectors c, b, w and λ. In the following sections we discuss two algorithms that compute Nash equilibria in LGNEPs by solving the constrained equation (6.1).

6.2 Potential Reduction Algorithm

The following algorithm was introduced in [26] for the solution of LGNEPs. In Section 8, we shall compare this algorithm numerically with the subgradient method from Section 7.5.

For a numerical solution of LGNEPs we consider the potential reduction algorithm (PRA) from [23] which was designed for nonlinear GNEPs. This algorithm exploits the constrained equation reformulation given in Proposition 1.5.2. Setting

$$Z_I := \{(x, \lambda, w) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m : \ Ax - b + w > 0, \ \lambda > 0, \ w > 0\}$$

and $z := (x, \lambda, w)$, we define the potential function

$$\psi(z) := \zeta \log(\|H(z)\|^2) - \sum_{i=1}^{m} (\log(A_i x - b_i + w_i) + \log(\lambda_i w_i))$$

for all $z \in \mathbb{R}^n \times \mathbb{R}^m_{++} \times \mathbb{R}^m_{++}$ and some $\zeta > m$, where $\|\cdot\|$ denotes the Euclidean norm and $\mathbb{R}^m_{++}$ the positive orthant. Below, we restate a version of the PRA from [23], where the linear equation system is solved exactly in each iteration.

We denote the n-dimensional vector of zeros by 0_n, the $2m$-dimensional vector of ones by 1_{2m} and the Jacobian of H by JH.

Algorithm 1: Potential Reduction Algorithm (PRA)

(S.0): Choose $z^0 \in \mathbb{R}^n \times \mathbb{R}^m_{++} \times \mathbb{R}^m_{++}$, β, $\gamma \in (0,1)$, $\varepsilon \geq 0$ and set $k := 0$, $a^T := (0_n^T, 1_{2m}^T)$.
(S.1): If $\|H(z^k)\| \leq \varepsilon$: STOP.
(S.2): Choose $\lambda_k \in [0,1)$ and compute $d^k \in \mathbb{R}^{n+2m}$, such that

$$JH(z^k)\, d^k = -H(z^k) + \lambda_k \frac{a^T H(z^k)}{\|a\|^2} a.$$

(S.3): Compute a stepsize $t_k := \max\{\beta^\ell : \ \ell = 0, 1, 2, \ldots\}$ such that

$$\begin{array}{rcll} z^k + t_k d^k & \in & Z_I & \text{and} \\ \psi(z^k + t_k d^k) & \leq & \psi(z^k) + \gamma t_k \nabla\psi(z^k)^T d^k. & \end{array}$$

(S.4): Set $z^{k+1} := z^k + t_k d^k$, $k := k+1$ and go to (S.1).

The application of PRA requires the Jacobian matrix $JH(z)$ to be nonsingular on Z_I which is also the only assumption in the convergence theorem (every accumulation point is a solution) of PRA in [23]. But all the conditions for nonsingularity given in [23] require the nonsingularity of a matrix $J_x F(x,\lambda)$ which consists of second order derivatives of the cost and constraint functions. By the linearity of these functions in our setting, all the nonsingularity results of [23] are therefore not applicable to LGNEPs. Thus, we will develop a nonsingularity condition tailored to LGNEPs. Note that the definitions of the matrices E and A in Section 6.1 correspond to the matrices $E(x)$ and $J_x g(x)$ in [23] which, in our linear context, are independent of x. Furthermore, for positive vectors $\lambda, w \in \mathbb{R}^m_{++}$ we define the diagonal matrices

$$\Lambda := \operatorname{diag}(\lambda) \quad \text{and} \quad W := \operatorname{diag}(w)$$

and have to find conditions that guarantee the nonsingularity of

$$JH(x,\lambda,w) = JH(\lambda,w) = \begin{pmatrix} 0 & E & 0 \\ A & 0 & I_m \\ 0 & W & \Lambda \end{pmatrix}$$

for all $\lambda, w \in \mathbb{R}^m_{++}$. Note that this matrix is independent of the x-part of the considered point. Clearly, a necessary nonsingularity condition is that E has full row rank and A has full column rank, that is n. Therefore we assume for the remaining part $m \geq n$ which, for instance, follows from a boundedness assumption for the LGNEP. To prove a sufficient condition we introduce for

an index set $I \subseteq \{1, \ldots, m\}$ the notation $E_{\bullet I}$ and $A_{I\bullet}$ for the matrices E and A, where all columns and rows, respectively, with indices not in I are dropped. Further, for $I, J \subseteq \{1, \ldots, m\}$ and a matrix $M \in \mathbb{R}^{m\times m}$ we denote by M_{JI} the matrix where all rows with indices not in J and all columns with indices not in I are dropped.

Theorem 6.2.1 *Assume that* $\det(E_{\bullet I}A_{I\bullet}) \geq 0$ *for all* $I \subseteq \{1, \ldots, m\}$ *with* $|I| = n$, *and that the determinant is positive for one of these index sets. Then* $JH(\lambda, w)$ *is nonsingular for all* $\lambda, w \in \mathbb{R}^m_{++}$.

Proof. The matrix

$$JH(\lambda, w) = \begin{pmatrix} 0 & E & 0 \\ A & 0 & I_m \\ 0 & W & \Lambda \end{pmatrix}$$

is nonsingular if and only if the matrix

$$\begin{pmatrix} E & 0 \\ W & -\Lambda A \end{pmatrix}$$

is nonsingular. Since $w \in \mathbb{R}^m_{++}$ the matrix W is a positive definite diagonal matrix, and hence nonsingular. Therefore the matrix $JH(\lambda, w)$ is nonsingular if and only if

$$EW^{-1}\Lambda A$$

is nonsingular. Let $I(n)$ denote all subsets of $\{1, \ldots, m\}$ with exactly n elements. Then applying the Binet-Cauchy formula for the determinant of this matrix twice yields

$$\begin{aligned} \det(EW^{-1}\Lambda A) &= \sum_{I\subseteq I(n)} \det((EW^{-1}\Lambda)_{\bullet I}) \det(A_{I\bullet}) \\ &= \sum_{I\subseteq I(n)} \sum_{J\subseteq I(n)} \det(E_{\bullet J}) \det((W^{-1}\Lambda)_{JI}) \det(A_{I\bullet}). \end{aligned}$$

Since $W^{-1}\Lambda$ is a diagonal matrix we have $\det((W^{-1}\Lambda)_{JI}) = 0$ for $J \neq I$. Hence we have

$$\det(EW^{-1}\Lambda A) = \sum_{I\subseteq I(n)} \det((W^{-1}\Lambda)_{II}) \det(E_{\bullet I}A_{I\bullet}).$$

Since $\lambda, w \in \mathbb{R}^m_{++}$, we have $\det((W^{-1}\Lambda)_{II}) > 0$ for all $I \in I(n)$. Hence our assumption yields $\det(EW^{-1}\Lambda A) > 0$ and hence the nonsingularity of $JH(\lambda, w)$ for all $\lambda, w \in \mathbb{R}^m_{++}$. •

Obviously, we can also assume that $\det(E_{\bullet I}A_{I\bullet}) \leq 0$ for all $I \subseteq \{1, \ldots, m\}$ with $|I| = n$, and that the determinant is negative for one of these index sets.

In the case of LGNEPs with shared constraints we obtain a nice corollary.

Corollary 6.2.2 *Assume that the matrix* $\hat{A} = (A^{11}A^{22}\dots A^{NN}) \in \mathbb{R}^{m_1\times n}$ *has full column rank for an LGNEP with shared constraints. Then* $JH(\lambda, w)$ *is nonsingular for all* $\lambda, w \in \mathbb{R}^m_{++}$.

Proof. In view of Theorem 6.2.1 it suffices to prove that $\det(E_{\bullet I}A_{I\bullet}) \geq 0$ for all $I \subseteq \{1,\dots,m\}$ with $|I| = n$, and that the determinant is positive for one of these index sets.

Let an arbitrary $I \subseteq \{1,\dots,m\}$ with $|I| = n$ be given. Then we define the splitting $I = I_1 \cup \dots \cup I_N$ where I_ν contains all those indices corresponding to chosen constraints for player ν. Exploiting the structure of A and E we obtain

$$\begin{aligned} E_{\bullet I}A_{I\bullet} &= \begin{pmatrix} ((A^{11})^T)_{\bullet I_1} & & \\ & \ddots & \\ & & ((A^{NN})^T)_{\bullet I_N} \end{pmatrix} \begin{pmatrix} (A^{11})_{I_1\bullet} & \dots & (A^{NN})_{I_N\bullet} \\ \vdots & & \vdots \\ (A^{11})_{I_1\bullet} & \dots & (A^{NN})_{I_N\bullet} \end{pmatrix} \\ &= \begin{pmatrix} ((A^{11})^T)_{\bullet I_1}(A^{11})_{I_1\bullet} & \dots & ((A^{11})^T)_{\bullet I_1}(A_{NN})_{I_N\bullet} \\ \vdots & & \vdots \\ ((A^{NN})^T)_{\bullet I_N}(A^{11})_{I_1\bullet} & \dots & ((A^{NN})^T)_{\bullet I_N}(A^{NN})_{I_N\bullet} \end{pmatrix} \\ &= \begin{pmatrix} ((A^{11})^T)_{\bullet I_1} \\ \vdots \\ ((A^{NN})^T)_{\bullet I_N} \end{pmatrix} \big((A^{11})_{I_1\bullet} \dots (A^{NN})_{I_N\bullet}\big). \end{aligned}$$

This shows that for all $I \subseteq \{1,\dots,m\}$ with $|I| = n$, the matrix $E_{\bullet I}A_{I\bullet}$ is positive semidefinite, and hence a P_0-matrix, implying $\det(E_{\bullet I}A_{I\bullet}) \geq 0$. Further the full column rank assumption guarantees that we can find at least one index set $\tilde{I}$ with $|\tilde{I}| = n$ such that the matrix

$$\big((A^{11})_{\tilde{I}_1\bullet} \dots (A^{NN})_{\tilde{I}_N\bullet}\big)$$

is nonsingular. Thus $E_{\bullet\tilde{I}}A_{\tilde{I}\bullet}$ is positive definite, and hence a P-matrix, which implies $\det(E_{\bullet\tilde{I}}A_{\tilde{I}\bullet}) > 0$. Then Theorem 6.2.1 shows the assertion. •

Let us consider the nonsingularity condition in the context of Example 6.2.3.

Example 6.2.3 *Consider the 2-player game whose cost functions possess a maximum structure*

$$\begin{aligned} &\min_{x_1} \quad \max\{x_1 - 2x_2, -x_1 - x_2\} \quad s.t.\ x_1 + x_2 \leq 1,\ x_1 \geq 0 \\ &\min_{x_2} \quad \max\{x_2 - x_1, x_1 - x_2\} \quad s.t.\ x_1 + x_2 \leq 1,\ x_2 \geq 0. \end{aligned}$$

Here we have nondifferentiable but convex cost functions and we are not aware of efficient algorithms that solve GNEPs of this particular class. However, we can reformulate the problem exploiting the maximum structure in the cost function. Applying the epigraphical reformulation (cf. Sec. 1.3.3) we obtain the equivalent LGNEP

$$\begin{aligned} &\min_{\alpha_1,x_1} \alpha_1 \quad s.t. \quad x_1 + x_2 \leq 1,\, x_1 \geq 0,\, x_1 - 2x_2 \leq \alpha_1,\, -x_1 - x_2 \leq \alpha_1, \\ &\min_{\alpha_2,x_2} \alpha_2 \quad s.t. \quad x_1 + x_2 \leq 1,\, x_2 \geq 0,\, x_2 - x_1 \leq \alpha_2,\, x_1 - x_2 \leq \alpha_2. \end{aligned}$$

Example 6.2.4 *In order to apply the PRA for the reformulated LGNEP from Example 6.2.3, one has to guarantee nonsingularity of the matrix*

$$JH(\lambda, w) = \begin{pmatrix} 0 & E & 0 \\ A & 0 & I_8 \\ 0 & W & \Lambda \end{pmatrix},$$

with

$$E := \begin{pmatrix} 0 & 0 & -1 & -1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 & 1 & -1 & 1 & -1 \end{pmatrix}$$

and

$$A = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & -1 & 0 & 0 \\ -1 & 1 & 0 & -2 \\ -1 & -1 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & -1 & -1 & 1 \\ 0 & 1 & -1 & -1 \end{pmatrix}.$$

For these matrices it is (with some computational effort) possible to check that the conditions of Theorem 6.2.1 are satisfied and hence $JH(\lambda, w)$ is nonsingular for all $\lambda, w \in \mathbb{R}^m_{++}$. The potential reduction algorithm requires 15 iterations to compute the solution

$$(\alpha_1, x_1, \alpha_2, x_2) = (-3.342, 2.513, 1.420, 3.237) \cdot 10^{-9}.$$

Let us further mention that our theoretical conditions are not always satisfied. For a slight change in the cost function of player 1 to

$$\max\{x_1 - 2x_2, -x_1 + x_2\},$$

we can still get an LGNEP with non-shared constraints as in Example 6.2.3 and we have the same E as before but a slightly different matrix

$$\tilde{A} = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & -1 & 0 & 0 \\ -1 & 1 & 0 & -2 \\ -1 & -1 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & -1 & -1 & 1 \\ 0 & 1 & -1 & -1 \end{pmatrix},$$

which differs from A only in the last entry in the fourth row. Now we have for the index set $I = \{3, 4, 7, 8\}$

$$\begin{aligned} \det(E_{\bullet I}\tilde{A}_{I\bullet}) &= \det\left(\begin{pmatrix} -1 & -1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix}\begin{pmatrix} -1 & 1 & 0 & -2 \\ -1 & -1 & 0 & 1 \\ 0 & -1 & -1 & 1 \\ 0 & 1 & -1 & -1 \end{pmatrix}\right) \\ &= -8 < 0, \end{aligned}$$

and the condition of Theorem 6.2.1 is violated. Moreover, it is possible to find $\lambda, w \in \mathbb{R}^m_{++}$ such that $JH(\lambda, w)$ is indeed singular. However, the potential reduction algorithm is still able to solve the problem numerically in 14 iterations.

6.3 A Hybrid Algorithm

Based on a combination of the robust potential reduction algorithm (PRA) and a local LP-Newton-method, a hybrid algorithm for GNEPs was developed in [22] which was called PRALP and is presented in Algorithm 2. In an initial phase this hybrid algorithm is equal to the potential reduction algorithm and hence solves a linear equation system in each iteration. When the current iterate seems to be close to a solution the algorithm switches to the LP-Newton part and solves a linear program to obtain the next iterate. If the convergence is fast enough, the algorithm continues solving linear programs until convergence, otherwise it returns to the last iterate of the potential reduction part and performs a further step there. The algorithm inherits the convergence properties of the potential reduction algorithm and is therefore applicable to LGNEPs under the conditions developed in the previous section. Furthermore, under an additional assumption on the multipliers,

Algorithm 2: Hybrid Algorithm for GNEPs (PRALP)

(S.0): Choose $z^0 := \hat{z} \in Z_I$, $\beta, \eta, \theta \in (0,1)$, $\zeta > m$, $0 < \tau_{\min} \leq \tau_{\max}$,$\tau_0 \in [\tau_{\min}, \tau_{\max}]$, $\varepsilon \geq 0$,
and set $k := 0$, $a^T := (0_n^T, 1_{2m}^T)$.

(S.1): If $\|H(z^k)\| \leq \varepsilon$ then STOP.
If $\|H(z^k)\| \leq \tau_k$ set $\lambda_k := 0$ and go to (S.4), else go to (S.2).

(S.2): Choose $\lambda_k \in [0,1)$ and compute $d^k \in \mathbb{R}^{n+2m}$ such that

$$JH(z^k)\, d^k = -H(z^k) + \lambda_k \frac{a^T H(z^k)}{\|a\|^2}\, a.$$

(S.3): Compute a stepsize $t_k := \max\left\{\beta^i \mid i = 0, 1, 2, \dots\right\}$ such that

$$z^k + t_k d^k \in Z_I \qquad \text{and}$$
$$\psi(z^k + t_k d^k) \leq \psi(z^k) + \eta t_k \nabla\psi(z^k)^T d^k.$$

Set $z^{k+1} := z^k + t_k d^k$, $\hat{z} := z^{k+1}$, $\tau_{k+1} := \tau_k$, $k := k+1$, go to (S.1).

(S.4): Compute a solution $(\tilde{z}^{k+1}, \gamma^{k+1})$ of the linear program

$$\begin{aligned} \min_{z,\gamma} \gamma \qquad \text{s.t.} \qquad & z \in \Omega, \\ & \|H(z^k) + JH(z^k)(z - z^k)\|_\infty \leq \gamma \|H(z^k)\|_\infty^2, \\ & \|z - z^k\|_\infty \leq \gamma \|H(z^k)\|_\infty. \end{aligned}$$

If $\|H(\tilde{z}^{k+1})\| \leq \theta \|H(z^k)\|$ then
set $z^{k+1} := \tilde{z}^{k+1}$, $\tau_{k+1} := \tau_k$, $k := k+1$, go to (S.1),
else set $z^{k+1} := \hat{z}$, $k := k+1$, choose $\tau_{k+1} \in [\tau_{\min}, \tau_{\max}]$, go to (S.2).

which is weaker than strict complementarity, PRALP is locally quadratic convergent. Hence, the additional effort in solving a linear program instead of linear equation systems pays off if one requires to compute solutions with high precision.

Finally, we would like to mention that there is also a nonsmooth variant of the hybrid algorithm presented in [20] which, however, has similar convergence conditions and numerical performance as PRALP.

Chapter 7

Subgradient Method

7.1 Nikaido-Isoda Based Approaches

A different approach to solve GNEPs is an optimization reformulation that was introduced in [39] for the computation of generalized Nash equilibria and uses the so-called Nikaido-Isoda function from [54] which, for LGNEPs, is defined by

$$\Psi(x,y) \;:=\; \sum_{\nu=1}^{N} \langle c^\nu, x^\nu\rangle - \langle c^\nu, y^\nu\rangle.$$

Additionally, we define player ν's optimal value function by

$$\varphi_\nu(x^{-\nu}) \;:=\; \begin{cases} \min\limits_{x^\nu \in X_\nu(x^{-\nu})} \langle c^\nu, x^\nu\rangle, & \text{if } X_\nu(x^{-\nu}) \neq \emptyset, \\ \infty, & \text{else,} \end{cases}$$

where Assumption 1.2.7 ensures the attainment of the minimum in case of nonempty strategy sets. For LGNEPs strong duality yields

$$\varphi_\nu(x^{-\nu}) \;=\; \max_{\lambda^\nu \in Z_\nu} \left\langle \lambda^\nu,\; \sum_{\mu\neq\nu} A^{\nu\mu}x^\mu - b^\nu \right\rangle$$

with

$$Z_\nu \;:=\; \left\{\lambda^\nu \in \mathbb{R}^{m_\nu} :\; c^\nu + (A^{\nu\nu})^T\lambda^\nu = 0,\; \lambda^\nu \geq 0\right\}.$$

Note that in contrast to the primal feasible set $X_\nu(x^{-\nu})$, the feasible set Z_ν of player ν's dual problem does not depend on $x^{-\nu}$. According to [39], generalized Nash equilibria are exactly the roots of the gap function

$$\begin{aligned} V(x) \;&:=\; \max_{y\in X_1(x^{-1})\times\ldots\times X_N(x^{-N})} \Psi(x,y) \\ &=\; \sum_{\nu=1}^{N} \langle c^\nu, x^\nu\rangle - \max_{\lambda^\nu\in Z_\nu} \left\langle \lambda^\nu,\; \sum_{\mu\neq\nu} A^{\nu\mu}x^\mu - b^\nu \right\rangle \end{aligned}$$

on the set

$$W \;:=\; \{x \in \mathbb{R}^n :\; Ax \leq b\}.$$

The gap function V is an extended-valued piecewise linear concave function and $V(x)$ is real-valued if and only if $X_\nu(x^{-\nu}) \neq \emptyset$ for all $\nu \in \{1, \ldots, N\}$. As mentioned in Proposition 1.5.1, a vector $\bar{x}$ is a generalized Nash equilibrium if and only if $\bar{x}$ is a global minimal point of the *constrained* nonsmooth optimization problem

$$\min V(x) \quad \text{s.t.} \quad x \in W$$

with optimal value $V(\bar{x}) = 0$.

In order to get an *unconstrained* optimization reformulation we have to deal with the fact that V may be not real-valued all over $\mathbb{R}^n$. To overcome this difficulty we construct a real-valued global extension of the gap function V which can be done in at least two different ways: In [24, 25], a projection based real-valued extension of V is proposed. However, this approach uses a regularized version of the Nikaido-Isoda function that we try to avoid since this would destroy our linear structure. Therefore, we apply a global real-valued extension of V that was suggested in Section 2.2 and is applicable if the set of vertices of the polyhedron Z_ν is known or easily computable. This holds true for some LGNEPs as we shall see in the Lemmata 7.4.1 and 7.4.2. Therefore, let

$$O_\nu := \operatorname{vert}(Z_\nu)$$

be the finite set of all vertices of Z_ν. Then, by the vertex theorem of linear programming, the function

$$\widehat{V}(x) \;=\; \sum_{\nu=1}^{N} \langle c^\nu, x^\nu \rangle - \max_{\lambda^\nu \in O_\nu} \left\langle \lambda^\nu,\; \sum_{\mu \neq \nu} A^{\nu\mu} x^\mu - b^\nu \right\rangle$$

is a real-valued global extension of V on $\mathbb{R}^n$, that is, $\widehat{V}$ is real-valued all over $\mathbb{R}^n$ and $\widehat{V}(x) = V(x)$ for all $x \in W$.

Denoting the i-th row of $A^{\nu\mu}$ by $A_i^{\nu\mu}$ we define the ℓ_1-penalty term

$$P(x) := \sum_{\nu=1}^{N} \sum_{i=1}^{m_\nu} \max \left\{ 0,\; A_i^{\nu\nu} x^\nu + \sum_{\mu \neq \nu} A_i^{\nu\mu} x^\mu - b_i^\nu \right\}$$

which is a penalty function for W and arrive at the following result that was indicated in [67].

Proposition 7.1.1 *There exists some $\rho > 0$, such that $\bar{x}$ is a generalized Nash equilibrium if and only if $\bar{x}$ is a global minimal point of the unconstrained nonsmooth optimization problem*

$$\min \widehat{V}_\rho(x) \;:=\; \widehat{V}(x) + \rho P(x)$$

with optimal value $\widehat{V}_\rho(\bar{x}) = 0$.

Proof. The function $\widehat{V}$ is piecewise affine-linear and by standard arguments one could find a finite penalty parameter for each of its linear pieces. Since the polyhedral sets O_ν have only a finite number of vertices, $\widehat{V}$ has only a finite number of linear pieces and setting ρ as the maximum of the associated penalty parameters, we have that $\widehat{V}_\rho(x)$ is an exact penalty function. This implies $\widehat{V}_\rho(x) > 0$ for all $x \in \mathbb{R}^n \setminus W$ and together with the fact that we have $\widehat{V}_\rho(x) = V(x)$ on W, the assertion follows from Proposition 1.5.1. •

The algorithms in Section 7.5 and 7.6 compute generalized Nash equilibria by solving the constrained and unconstrained nonsmooth optimization problems from Proposition 1.5.1 and 7.1.1.

7.2 Structural Properties of Subgradient Methods

The optimization problem

$$\min V(x) \quad \text{s.t.} \quad x \in W$$

is a linearly constrained nonsmooth piecewise-linear concave optimization problem which, in principle, can be tackled by each algorithm from nonsmooth, nonconvex optimization. However, due to the fact that we cannot exploit second order information in the linear case, we restrict ourselves to subgradient methods which only use first order information.

One major drawback of general nonsmooth methods is that an ordinary 'pointwise' subdifferential does not yield a numerical stopping criterion. One possibility to overcome this difficulty is the use of an epsilon subdifferential that contains information of subderivatives from a neighborhood of the point of interest as it is done, e.g., in bundle methods or the robust gradient sampling method (RGS) from [8]. Usually, the approximated epsilon subdifferential is used to compute a search direction or to check a stopping criterion at the cost of an quadratic optimization problem at each iteration (cf. [8, 49]). In particular for high dimensional problem data this computation becomes numerically expensive.

However, in contrast to general nonsmooth optimization problems, we can exploit that the optimal function value is known to be zero, which provides the handy stopping criterion $V(x^k) < \varepsilon$, and we do not have to construct a numerically expensive epsilon subdifferential.

Since our objective function V is piecewise-linear, the set of nondifferentiability points has Lebesgue measure zero, that is, V is differentiable almost everywhere. Therefore, a subgradient based method that works with numerical precision is likely to compute gradients instead of subgradients at each iteration. Hence, together with our stopping criterion $V(x^k) < \varepsilon$, in this case we would basically apply a smooth method to a nonsmooth optimization problem. The idea of applying smooth methods to nonsmooth nonconvex problems is known to work very well but it seems to be a very challenging task to obtain convergence results (cf., e.g., [50, 51]).

7.3 Local and Global Minima

As illustrated in the following Examples 7.3.1 and 7.3.2, the gap function V may have strict local minimal points as well as plateaus. Therefore, in general, we cannot expect a gradient based local method to find a global minimal point of the gap function V or the penalty function $\widehat{V}_\rho$ in order to compute a Nash equilibrium as described in Section 1.5.1. Nonetheless, in contrast to these general LGNEPs there are no local minimal points in our economic market example with shared constraints that are not solutions of the LGNEP as we shall see in Lemma 7.3.3.

Example 7.3.1 *Consider the 2-player game with shared constraints defined by*

$$\begin{aligned} &\min_{x_1} -x_1 \quad s.t. \quad x_1 \in [0,2], x_1 - 2x_2 \leq 1, -2x_1 + x_2 \leq 1, \\ &\min_{x_2} -x_2 \quad s.t. \quad x_2 \in [0,2], x_1 - 2x_2 \leq 1, -2x_1 + x_2 \leq 1. \end{aligned}$$

Here we can find

$$\begin{aligned} \varphi_1(x_2) &= \min_{x_1 \in X_1(x_2)} -x_1 &&= -\min\{2, 1 + 2x_2\}, \\ \varphi_2(x_1) &= \min_{x_2 \in X_2(x_1)} -x_2 &&= -\min\{2, 1 + 2x_1\}, \end{aligned}$$

and therefore we have

$$V(x) = -x_1 - x_2 + \min\{2, 1 + 2x_2\} + \min\{2, 1 + 2x_1\}.$$

In a feasible neighborhood of $(0,0)$ *we have*

$$V(x) = -x_1 - x_2 + 1 + 2x_2 + 1 + 2x_1 = 2 + x_1 + x_2$$

and hence $(0,0)$ *is a local minimum which is not a solution, since* $V(0,0) = 2 > 0$.

Example 7.3.2 *Consider the 2-player game with shared constraints defined by*

$$\begin{aligned}
&\min_{x_1} -x_1 \quad s.t. \quad x_1 + x_2 \leq 1, -x_1 + x_2 \leq 1, x_1 - x_2 \leq 1, -x_1 - x_2 \leq 1,\\
&\min_{x_2} -x_2 \quad s.t. \quad x_1 + x_2 \leq 1, -x_1 + x_2 \leq 1, x_1 - x_2 \leq 1, -x_1 - x_2 \leq 1.
\end{aligned}$$

Here we can find

$$\begin{aligned}
\varphi_1(x_2) &= \min_{x_1 \in X_1(x_2)} -x_1 = -\min\{1 - x_2, 1 + x_2\},\\
\varphi_2(x_1) &= \min_{x_2 \in X_2(x_1)} -x_2 = -\min\{1 - x_1, 1 + x_1\},
\end{aligned}$$

and therefore we have

$$\begin{aligned}
V(x) &= -x_1 - x_2 + \min\{1 - x_2, 1 + x_2\} + \min\{1 - x_1, 1 + x_1\}\\
&= 2 - x_1 - x_2 - |x_1| - |x_2|.
\end{aligned}$$

In this example we have

$$V(x) = 2 \quad \forall x \in \{x \in \mathbb{R}^2 \mid x_1 < 0, x_2 < 0, -x_1 - x_2 \leq 1\},$$

and therefore we have a part of the feasible set where V *is constant.*

Lemma 7.3.3 *Assume the basic market model from Section 1.3.1 with non-negative and ordered price vectors, that is, with* $0 < p_1^\nu \leq p_2^\nu \leq \ldots \leq p_K^\nu$ *for all* $\nu \in \{1, \ldots, N\}$. *Further assume that for each* $i = 1, \ldots, K$, *there is at least one* $\nu \in \{1, \ldots, N\}$ *with* $p_{i-1}^\nu < p_i^\nu$ *(where* $p_0^\nu := 0$*). Then every local minimum of* V *on* W *is a solution of the LGNEP.*

Proof. See [26]. •

The existence of local minimal points in the modified economic market example is more involved and we have not been able to obtain a positive or negative result. However, our subgradient based methods work numerically very well for the basic economic markets and its modified version as we shall see in Section 8.

7.4 The Computational Effort of Subgradient Methods

In order to apply a subgradient method we need to compute one subgradient of V in each iteration. Recall that V may not be defined outside of W, but it is possible to replace V by its global extension $\widehat{V}$ for all $x \notin W$. The subdifferential in the sense of Clarke (cf. [11]) of $\widehat{V}$ can be computed explicitly as already mentioned in [67]. For completeness, we recall the result whose proof follows from calculation rules of Clarke subdifferentials. We denote for given $x^{-\nu}$ by $O_\nu(x^{-\nu})$ the set of optimal vertex points of player ν's dual problem

$$\max \left\langle \lambda^\nu, \sum_{\mu \neq \nu} A^{\nu\mu} x^\mu - b^\nu \right\rangle \quad \text{s.t.} \quad \lambda^\nu \in Z_\nu$$

and use the Minkowski sum $A + B$ for two sets $A, B \subseteq \mathbb{R}^n$.

Recall that according to Proposition 2.4.4 the Clarke subdifferential of $\widehat{V}$ at $x \in \mathbb{R}^n$ is given by

$$\begin{aligned}
\partial \widehat{V}(x) \;=\; & \left\{ \begin{pmatrix} c^1 \\ -(A^{1,2})^T \lambda^1 \\ \vdots \\ -(A^{1,N})^T \lambda^1 \end{pmatrix} : \; \lambda^1 \in \operatorname{conv}\left(O_1(x^{-1})\right) \right\} \\
& + \left\{ \begin{pmatrix} -(A^{2,1})^T \lambda^2 \\ c^2 \\ -(A^{2,3})^T \lambda^2 \\ \vdots \\ -(A^{2,N})^T \lambda^2 \end{pmatrix} : \; \lambda^2 \in \operatorname{conv}\left(O_2(x^{-2})\right) \right\} \\
& + \ldots + \left\{ \begin{pmatrix} -(A^{N,1})^T \lambda^N \\ \vdots \\ -(A^{N,N-1})^T \lambda^N \\ c^N \end{pmatrix} : \; \lambda^N \in \operatorname{conv}\left(O_N(x^{-N})\right) \right\}.
\end{aligned}$$

By definition, each evaluation of $\widehat{V}$ requires the computation of N optimal points $\bar{\lambda}^1, \ldots, \bar{\lambda}^N$. Having these optimal points, the computation of one subgradient of $\widehat{V}$ is an easy task as we have seen in Corollary 2.4.5. Therefore, the main computational effort is the evaluation of $\widehat{V}$. For $x \in W$, the computation of the optimal points $\bar{\lambda}^1, \ldots, \bar{\lambda}^N$ is equivalent to solving N linear programs whereas for $x \notin W$, we need to compute $\widehat{V}(x)$ with help of the vertex set O_ν of the polyhedron Z_ν, which, in general, is a difficult task.

However, if the vertex set O_ν of Z_ν is known, the evaluation of $\widehat{V}$ becomes cheap. For our economic market models from Section 1.3.1 the sets O_ν can be computed using the following Lemmata. Note that for known vertex sets O_ν also in the case $x \in W$, the numerical solution of N linear optimization problems can be replaced by an enumeration of the known vertices of Z_ν which is possible since the number of vertices increases only linearly in K.

Lemma 7.4.1 *The set* $Z := \{\lambda \in \mathbb{R}^{2K+1} : \ A^T\lambda = p,\ \lambda \geq 0\}$ *with*

$$A^T := \begin{pmatrix} 1 & & 0 & 1 & -1 & & 0 \\ & \ddots & & \vdots & & \ddots & \\ 0 & & 1 & 1 & 0 & & -1 \end{pmatrix} \in \mathbb{R}^{K\times(K+1+K)}$$

possesses at most $K+1$ *different vertices which are given by*

$$\lambda^{(i)} \ = \ \begin{pmatrix} \max\{0, p - p_i \cdot e\} \\ p_i \\ \max\{0, p_i \cdot e - p\} \end{pmatrix}$$

for all $i \in \{1, \ldots, K\}$ *and*

$$\lambda^{(K+1)} \ = \ \begin{pmatrix} \max\{0, p\} \\ 0 \\ \max\{0, -p\} \end{pmatrix}$$

where the maximum is taken componentwise.

Proof. First, we rearrange the equations to obtain

$$\begin{pmatrix} \lambda_{K+2} \\ \vdots \\ \lambda_{2K+1} \end{pmatrix} = \begin{pmatrix} 1 & & 0 & 1 \\ & \ddots & & \vdots \\ 0 & & 1 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_{K+1} \end{pmatrix} - \begin{pmatrix} p_1 \\ \vdots \\ p_K \end{pmatrix},$$

and, hence, finding the vertices of Z becomes equivalent to finding the vertices of

$$\begin{aligned} \tilde{Z} \ &= \ \left\{ \tilde{\lambda} \in \mathbb{R}^{K+1} : \ \begin{pmatrix} 1 & & 0 & 1 \\ & \ddots & & \vdots \\ 0 & & 1 & 1 \end{pmatrix} \tilde{\lambda} \geq p,\ \tilde{\lambda} \geq 0 \right\} \\ &= \ \left\{ \tilde{\lambda} \in \mathbb{R}^{K+1} : \ \tilde{\lambda}_{K+1} \geq 0,\ \tilde{\lambda}_i \geq \max\{0, p_i - \tilde{\lambda}_{K+1}\}\, \forall i = 1, \ldots, K \right\}. \end{aligned}$$

Since at any vertex at least $K+1$ linearly independent constraints are active, we consider the following cases:

- $\tilde{\lambda}_{K+1} = 0$: Then we must have $\tilde{\lambda}_i = \max\{0, p_i\}$ for all $i \in \{1, \dots, K\}$ in order to get $K+1$ linearly independent active constraints, which is clear for $p_i \neq 0$, and follows for $p_i = 0$ from the fact that $\tilde{\lambda}_i \geq 0, \tilde{\lambda}_{K+1} \geq 0$ and $\tilde{\lambda}_i + \tilde{\lambda}_{K+1} \geq 0$ are linearly dependent. This results in the vertex

$$\tilde{\lambda}^{(K+1)} = \begin{pmatrix} \max\{0, p\} \\ 0 \end{pmatrix}.$$

- $\tilde{\lambda}_{K+1} \neq 0$ and $\tilde{\lambda}_{K+1} \neq p_i$ for all $i \in \{1, \dots, K\}$: Then we can have at most K active constraints $\lambda_i = \max\{0, p_i - \tilde{\lambda}_{K+1}\}$ and hence we do not have a vertex in this case.

- $\tilde{\lambda}_{K+1} \neq 0$ and $\tilde{\lambda}_{K+1} = p_i$ for some $i \in \{1, \dots, K\}$: Then we have the two independent active constraints $\tilde{\lambda}_i = 0$ and $\tilde{\lambda} + \tilde{\lambda}_{K+1} = p_i$. If for some $j \neq i$ we have $p_j = p_i$, then $\tilde{\lambda}_j = \max\{0, p_j - \tilde{\lambda}_{K+1}\}$ brings two new active constraints, but the number of linearly independent active constraints increases only by one. Further, also if we have $p_j \neq p_i$ we can increase the number of linearly independent active constraints by one if we set $\tilde{\lambda}_j = \max\{0, p_j - \tilde{\lambda}_{K+1}\} = \max\{0, p_j - p_i\}$. Altogether this means that we have only one way to obtain $K+1$ linearly independent constraints, and hence only one vertex

$$\tilde{\lambda}^{(i)} = \begin{pmatrix} \max\{0, p - p_i \cdot e\} \\ p_i \end{pmatrix}.$$

This shows that we get one vertex from the first case and at most K (if all p_i are different) vertices from the third case. Having the vertices from $\tilde{Z}$ we get those of Z using the equations, and hence we have at most $K+1$ vertices

$$\lambda^{(i)} = \begin{pmatrix} \max\{0, p - p_i \cdot e\} \\ p_i \\ \max\{0, p_i \cdot e - p\} \end{pmatrix}$$

for all $i \in \{1, \dots, K\}$ and

$$\lambda^{(K+1)} = \begin{pmatrix} \max\{0, p\} \\ 0 \\ \max\{0, -p\} \end{pmatrix}$$

•

In the economic market with nonshared constraints the vertex computation for the players $1, \dots, N$ does not change but player $N+1$ faces the additional constraint

$$-N \cdot x_1^{N+1} \leq -\sum_{\nu=1}^{N} x_1^{\nu}.$$

Furtunately, the vertex set O_{N+1} of Z_{N+1} is still computable as we see in the following Lemma.

Lemma 7.4.2 *For the matrix*

$$A^T := \begin{pmatrix} 1 & & 0 & 1 & -1 & & 0 & -N \\ & \ddots & & \vdots & & \ddots & & 0 \\ 0 & & 1 & 1 & 0 & & -1 & 0 \end{pmatrix} \in \mathbb{R}^{K\times(K+1+K+1)}$$

the set

$$Z := \{\lambda \in \mathbb{R}^{2K+2} |\ A^T\lambda = p,\ \lambda \geq 0\}$$

possesses at most $K+2$ different vertices which are given by

$$\lambda^{(i)} = \begin{pmatrix} \max\{0, p - p_i \cdot e\} \\ p_i \\ \max\{0, p_i \cdot e - p\} \\ 0 \end{pmatrix}$$

for all $i \in \{1, \ldots, K\}$ as well as

$$\lambda^{(K+1)} = \begin{pmatrix} \max\{0, p\} \\ 0 \\ \max\{0, -p\} \\ 0 \end{pmatrix}$$

and

$$\lambda^{(K+2)} = \begin{pmatrix} 0 \\ \max\{p_2, 0\} \\ \vdots \\ \max\{p_K, 0\} \\ 0 \\ 0 \\ \max\{p_2, 0\} - p_2 \\ \vdots \\ \max\{p_K, 0\} - p_K \\ \frac{-p_1}{N} \end{pmatrix}$$

provided that $p_1 < 0$.

Proof. First we can rearrange the equations to obtain

$$\begin{pmatrix} \lambda_{K+2} \\ \vdots \\ \lambda_{2K+1} \end{pmatrix} = \begin{pmatrix} 1 & & 0 & 1 & -N \\ & \ddots & & \vdots & 0 \\ 0 & & 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_{K+1} \\ \lambda_{2K+2} \end{pmatrix} - \begin{pmatrix} p_1 \\ \vdots \\ p_K \end{pmatrix},$$

and finding the vertices of Z becomes equivalent to finding the vertices of

$$\begin{aligned}
\tilde{Z} &= \left\{ \tilde{\lambda} \in \mathbb{R}^{K+2} \middle| \begin{pmatrix} 1 & & 0 & 1 & -N \\ & \ddots & & \vdots & 0 \\ 0 & & 1 & 1 & 0 \end{pmatrix} \tilde{\lambda} \geq p,\ \tilde{\lambda} \geq 0 \right\} \\
&= \{\tilde{\lambda} \in \mathbb{R}^{K+2} |\ \tilde{\lambda} \geq 0,\ \tilde{\lambda}_1 \geq p_1 - \tilde{\lambda}_{K+1} + N \cdot \lambda_{2K+2}, \\
&\quad \lambda_i \geq p_i - \tilde{\lambda}_{K+1} \forall i = 2, \ldots, K\} \\
&= \big\{\tilde{\lambda} \in \mathbb{R}^{K+2} \big|\ \tilde{\lambda}_{2K+2} \geq 0, \tilde{\lambda}_{K+1} \geq 0, \tilde{\lambda}_1 \geq \max\{0, p_1 - \tilde{\lambda}_{K+1} + N \cdot \lambda_{2K+2}\} \\
&\quad \tilde{\lambda}_i \geq \max\{0, p_i - \tilde{\lambda}_{K+1}\}\, \forall i = 2, \ldots, K \big\}.
\end{aligned}$$

Since at any vertex at least $K + 2$ linearly independent constraints are active, we consider the following cases:

- Let $\tilde{\lambda}_{2K+2} = 0$. This yields at most $K + 1$ different vertices

$$\tilde{\lambda}^{(i)} = \begin{pmatrix} \max\{0, p - p_i \cdot e\} \\ p_i \\ 0 \end{pmatrix}$$

  for all $i \in \{1, \ldots, K\}$ and

$$\tilde{\lambda}^{(K+1)} = \begin{pmatrix} \max\{0, p\} \\ 0 \\ 0 \end{pmatrix}$$

  as seen in Lemma 7.4.1.

- Let $\tilde{\lambda}_{2K+2} \neq 0$ and $\tilde{\lambda}_{K+1} = 0$. Then we have $\tilde{\lambda}_i = \max\{p_i, 0\}$, $i = 2, \ldots, K$, which yields $K - 1$ independent equalities. The only way to obtain two more linearly independent constraints is the case $\lambda_1 = 0$ and $\lambda_{2K+2} = -p_1/N$, and the vertex

$$\tilde{\lambda}^{(K+2)} = \begin{pmatrix} 0 \\ \max\{p_2, 0\} \\ \vdots \\ \max\{p_K, 0\} \\ 0 \\ \frac{-p_1}{N} \end{pmatrix}$$

  is only feasible, if we have $p_1 < 0$.

- Let $\tilde{\lambda}_{2K+2} \neq 0$, $\tilde{\lambda}_{K+1} \neq 0$ and $\tilde{\lambda}_{K+1} \neq p_i$ for all $i = 2, \ldots, K$. Then there is no possibility to obtain more than K active constraints and therefore we do not have a vertex in this case.

- Let $\tilde{\lambda}_{2K+2} \neq 0$, $\tilde{\lambda}_{K+1} \neq 0$ and $\tilde{\lambda}_{K+1} = p_i$ for some $i = 2, \ldots, K$. Then the constraint $\tilde{\lambda}_i = \max\{0, p_i - \tilde{\lambda}_{K+2}\}$ yields two active independent constraints. Following the arguments in the proof of Lemma 7.4.1, we obtain from each constraint

$$\tilde{\lambda}_j = \max\{0, p_j - \tilde{\lambda}_{K+2}\} = \max\{0, p_j - p_i\}$$

  with $j \in \{2, \ldots, K\} \setminus \{i\}$ only one more independent equation. However, also in this case it is only possible to obtain at most $K+1$ vertices, such that there is no vertex in this case.

Finally, we obtain at most $K + 2$ different vertices which are given by

$$\lambda^{(i)} = \begin{pmatrix} \max\{0, p - p_i \cdot e\} \\ p_i \\ \max\{0, p_i \cdot e - p\} \\ 0 \end{pmatrix}$$

for all $i \in \{1, \ldots, K\}$ as well as

$$\lambda^{(K+1)} = \begin{pmatrix} \max\{0, p\} \\ 0 \\ \max\{0, -p\} \\ 0 \end{pmatrix}$$

and

$$\lambda^{(K+2)} = \begin{pmatrix} 0 \\ \max\{p_2, 0\} \\ \vdots \\ \max\{p_K, 0\} \\ 0 \\ 0 \\ \max\{p_2, 0\} - p_2 \\ \vdots \\ \max\{p_K, 0\} - p_K \\ \frac{-p_1}{N} \end{pmatrix}$$

provided that $p_1 < 0$. •

7.5 Projected Subgradient Method

In order to solve the constrained optimization reformulation from Proposition 1.5.1 we suggest the following simple projected subgradient method (PSM), where we denote the projection of a vector x onto the set W with respect to the Euclidean norm by $P_W(x)$.

Algorithm 3: Projected Subgradient Method (PSM)

(S.0): Choose $x^0 \in W$, $\varepsilon \geq 0$, and set $k := 0$.
(S.1): If $V(x^k) \leq \varepsilon$: STOP.
(S.2): Compute a subgradient $s^k \in \partial V(x^k)$.
(S.3): Set $\alpha_k := \frac{1}{\sqrt{k+1}}$.
(S.4): Compute $x^{k+1} = P_W[x^k - \alpha_k s^k]$.
(S.5): Set $k := k + 1$ and go to (S.1).

Note that we choose a nonsummable diminishing a priori step size in (S.3) that does not require a function evaluation of V at the point $x^k - \alpha_k s^k$ in order to avoid the evaluation of V outside of W, where it is possibly undefined. The effort in (S.4) is the computation of the projection which is a quadratic program whenever $x^k - \alpha_k s^k \notin W$. The remaining effort is the evaluation of V in (S.1) since the computation of a subgradient in (S.2) is cheap using the formula from Corollary 2.4.5.

7.6 Penalty Approach

Here we want to exploit the unconstrained optimization reformulation of an LGNEP obtained by a penalty approach as stated in Proposition 7.1.1 and suggested in [67]. In principle, if one can estimate a sufficiently large penalty parameter, this unconstrained optimization problem can be tackled by an arbitrary method from unconstrained nonsmooth, nonconvex optimization. In Section 8, we will use the following algorithms in order to solve the penalty reformulation of LGNEPs.

- The robust gradient sampling (RGS) algorithm from [8] which guarantees convergence to a Clarke starionary point but requires sampling and the solution of a quadratic program at each iteration.
- The following ordinary subgradient method which exploits prior knowledge about the optimal value in both, the step size with the so-called Polyak step size rule, and the stopping criterion as discussed in the beginning of Section 7.

Algorithm 4: Penalty Method

(S.0): Choose $x^0 \in W$, $\varepsilon_1, \varepsilon_2 \geq 0, \rho > 0$, and set $k := 0$.
(S.1): If $\widehat{V}(x^k) \leq \varepsilon_1$ and $P(x) \leq \varepsilon_2$: STOP.
(S.2): Compute a subgradient $s^k \in \partial \widehat{V}_\rho(x^k)$.
(S.3): Set $\alpha_k := \frac{\widehat{V}_\rho(x^k)}{(s^k)^T s^k}$.
(S.4): Set $x^{k+1} := x^k - \alpha_k s^k$ as well as $k := k + 1$ and go to (S.1).

In contrast to the evaluation of V which requires the solution of linear programs that may not have a solution for $x \notin W$, the evaluation of $\widehat{V}$ is possible for all $x \in \mathbb{R}^n$ but requires the computation of the vertex sets $O_\nu = \mathrm{vert}(X_\nu)$. While this is difficult in general, we are able to compute the vertex sets O_ν for the economic market models from Section 1.3.1 by the Lemmata 7.4.1 and 7.4.2.

Using Corollary 2.4.5 we are able to compute an element of $\partial \widehat{V}(x)$ and since $\partial P(x)$ is computable by standard arguments, we can find explicit formulas for a subgradient $s \in \partial \widehat{V}_\rho(x)$ that we have to compute in (S.2). Therefore, if we are able to compute the vertex set O_ν the penalty method is very cheap.

Chapter 8

Numerical results

Due to the lack of an explicit treatment of LGNEPs in the existing literature, we were not able to find a set of existing test examples and, therefore, the numerical tests in this section are based on the economic market models that were introduced in Section 1.3.1.

8.1 Construction of Test Examples

In order to construct numerical examples for the economic problems with N players and K categories as described in Section 1.3.1, we created 25 instances of the market model in different dimensions by doing the following:

- The price vectors p_k^ν are the sum of two parts: a random integral between 1 and 100, which is equal for all players $\nu \in \{1, \ldots, N\}$ and an individual random integral between 1 and 10, which may be different for each player and models the individual profit.
- The capacities $C^\nu = 100 \cdot K$ are assumed to be equal for all players $\nu \in \{1, \ldots, N\}$.
- The demand D_k for the product in the price category k is the sum of $50 \cdot N$ and a random integral between N and $40 \cdot N$, which is different for all categories $k \in \{1, \ldots, K\}$.

We also construct the modified basic economic problems from Section 1.3.1 by the rules above with randomly generated but ordered price vectors.

Although the LGNEPs in our basic economic market model do not have local minimal points (cf. Lemma 7.3.3), LGNEPs resulting from these economic market models have indeed nonsmooth concave gap functions and the

structure of the solution set is not simple, as one can see for low dimensional problems by direct inspection.

8.2 Implementation Details

All our implementations are done in Matlab® and the tests are run on a personal computer with four cores à 3.2 GHz and 12 GB RAM.

As we have seen in Section 1.6, the computation of normalized Nash equilibria of an LGNEP is equivalent to the solution of a single linear program. We used the dual simplex method implemented in the Matlab® function *linprog* to find normalized solutions for different instances of our economic market model. Note that we used the dual simplex method in *linprog* since it outperformed the interior point algorithm in *linprog* in both, running time and accuracy of the results.

In PRA and PRALP we set the starting point $x^0 = 0, \lambda^0 = 10, w^0 = \max\{10, 5 - g(x^0)\}$, and as stopping criterion we use that the merit function, defined in [22], is less or equal $\varepsilon = 10^{-8}$. In PRA we use the parameters $\gamma = 0.01, \beta = 0.5$ and a constant $\lambda_k = 0.1$. The hybrid method PRALP is run with the parameters as in [20], i.e.,

$$\beta = 0.5,\ \eta = 0.01,\ \theta = 0.9,\ \zeta = 2m,\ \tau_{\min} = 10^{-11},\ \tau_{\max} = \tau_0 = 10^{-3},$$

the constant choice $\lambda_k = 0.1$ in (S.2), and the update strategy $\tau_{k+1} = \max\{\tau_{\min}, \theta\|H(z^k)\|\}$. For further implementation details we refer to [20]. The linear programs in PRALP were solved with the dual simplex method from *linprog*.

The projected subgradient method (PSM) is run with $x^0 = 0$ and $\varepsilon = 10^{-6}$ and the quadratic programs are solved by the Matlab® function *quadprog*. In the case of unknown vertices O_ν, the linear optimization problems are solved with the primal simplex method in *linprog*. The robust gradient sampling (RGS) algorithm is run with standard settings and starting point $x^0 = 0$. For the penalty method we choose $x^0 = 0$, $\varepsilon_1 = \varepsilon_2 = 10^{-4}$ and $\rho = 1000$.

8.3 Discussion of Numerical Results

8.3.1 Normalized Solutions

Table 8.1 summarizes the results for the computation of normalized equilibria which is done by solving a linear program with *linprog* as described in Section 1.6. In order to compare it with a different approach, we also test

the potential reduction algorithm (PRA), Algorithm 1, which computes by [19, Lemma 6.7] normalized Nash equilibria with weights $r_1 = \ldots = r_N = 1$, if equal starting vectors for all components of λ and w are chosen.

The first column is the number of players N, the second the number of categories K. Then we have three columns for the potential reduction algorithm (PRA): The first, 'LS', is the number of equation systems solved. The second, 'term', is the final merit function value when the algorithm terminates, where the merit function is defined for the KKT system of the GNEP using the minimum function as complementarity function, see [23] for details. The third column is the running 'time' of the method (which is in all tables given in seconds). Furthermore, the last two columns in Table 8.1 are for the test runs for the direct solution of the linear program with *linprog*. We report the number of iterations 'iter' and the running 'time'.

As Table 8.1 reveals for the computation of normalized Nash equilibria the direct solution of the linear program with the Matlab function *linprog* is very efficient and except for some problems with small dimension much faster than PRA, even though it solves much more linear equation systems, since it requires the solution of two of them at each iteration.

8.3.2 General Solutions

As we have seen in the previous section the computation of normalized solutions by *linprog* is very efficient. However, since these are only defined for GNEPs with shared constraints and even there not all solutions are normalized, we present here numerical results for different methods to compute not necessarily normalized solutions.

First, we use PRA again. Choosing the starting vectors λ and w, such that not all partial vectors of λ and w coincide this algorithm computes not necessarily normalized solutions. However, its performance is very similar to the case with equal starting vectors, as reported in Table 8.1, and therefore we do not report new results here. Second, we test the hybrid method PRALP, Algorithm 2. We report the number of linear equation systems 'LS', the number of linear programs 'LP', the final merit function value 'term', and the running 'time'. As a third algorithm we use the projected subgradient method (PSM), Algorithm 3. Here we report the number of linear programs 'LP', the number of quadratic programs 'QP', the final value of the merit function V, and the running 'time'. Note that the stopping criterion $V(x) \leq 10^{-6}$ yields results that are for most of the problems slightly more precise as those for PRA with a stopping tolerance of 10^{-8}. Compared to PRALP with the same stopping tolerance 10^{-8}, the precision of PSM with 10^{-6} is worse for small and middle problem sizes and better for larger problems. Finally,

we solve the penalty reformulation with two algorithms, the robust gradient sampling (RGS) from [8] and our penalty method Algorithm 4.

As one can see in the Tables 8.2 and 8.3, PRA and PSM are able to solve all instances of the test problems, and we have only one failure for PRALP for the largest problem in each class which was caused by the fact that the resulting linear problem was to large to be solved by *linprog* with the provided memory. Therefore, all algorithms seem to be suitable to solve LGNEPs. Let us mention that slightly negative values of V correspond to iterates that are slightly not feasible but within the feasibility tolerance used by the quadratic program to obtain the projection on W. PRA solves the test examples faster than PRALP, which, however, achieves a higher accuracy than PRA. This phenomenon is due to the solution of linear programs in PRALP which, on one hand, takes more time than the solution of linear equation systems but, on the other hand, improves the precision of the results. Comparing the running time between PRA and PSM, we see that PRA is often faster for the lower and middle dimensional cases whereas PSM is much faster than PRA for high dimensional problems. In our opinion, all three algorithms are recommendable for the solution of LGNEPs.

In Table 8.4, we test the general PSM against 'PSMvert' which is a version of PSM that exploits the knowledge of the vertex sets in order to avoid the numerical solution of linear optimization problems in each iteration of PSM. We see that especially in lower and middle dimensional examples, the enumeration of the known vertices saves a lot of time compared with the solution of linear programs. However, for the largest examples, this advantage becomes negligible since then most of the running time is spent with solving the quadratic problem to compute the projection onto W. We see that both algorithms are able to solve LGNEPs with up to ten thousand variables in a reasonable time.

Furthermore, with known vertex sets we can apply the penalty method Algorithm 4 and RGS to the penalty reformulation of LGNEPs. In Table 8.5, we report the number of iterations 'iter', the final value of $\hat{V}_\rho$, and the running 'time' for both methods. The penalty method performs much better than RGS which can be explained by the fact that it exploits prior knowledge of our optimal value as explained in the beginning of Section 7. However, both algorithms are not comparable to PRA, PRALP or PSM, in both, accuracy and running time. Hence, in conclusion, we do not recommend the penalty reformulation for the solution of LGNEPs.

8.4 Conclusions

In this part we analysed algorithms for the solution of LGNEPs. We provided a new convergence condition for the potential reduction algorithm applied to LGNEPs since the existing ones do not cover the linear case. The algorithm turns out to be robust and for small and middle dimensional problems also fast. Due to the lack of LGNEPs in the literature, our numerical results were based on some economic market models, introduced in Section 1.3.1. These problems were shown to have the favorable property that all local minima are also global in some optimization reformulation of the LGNEP. Therefore, we proposed a projected subgradient method for the solution of the LGNEPs which turned out to be very efficient and fast and, in particular, outperforms all other algorithms for larger dimensional problems.

The use of a penalty method as indicated in [67] which is only possible if one is able to compute the vertices of some polyhedral sets is, in our opinion, not competitive to the other algorithms.

Using linear functions as approximations for nonlinear ones, we think that the numerical solution of LGNEPs can play an important role also for the solution of nonlinear problems and it is a future research topic to find appropriate approximation procedures for GNEPs by LGNEPs.

8.5 Tables

		PRA			linprog	
N	K	LS	$term$	$time$	iter	$time$
2	3	30	2.7e-09	0.14	9	0.04
2	10	30	1.8e-09	0.13	26	0.14
2	30	36	1.2e-09	0.20	66	0.14
2	50	42	1.4e-09	0.36	111	0.13
2	100	44	1.1e-09	1.06	248	0.14
5	3	24	1.2e-09	0.10	19	0.13
5	10	38	1.8e-09	0.17	65	0.14
5	30	38	1.8e-09	0.51	183	0.17
5	50	50	2.3e-09	1.73	307	0.18
5	100	49	2.7e-09	7.44	603	0.17
10	3	29	4.8e-09	0.13	39	0.13
10	10	34	7.6e-09	0.31	135	0.14
10	30	40	2.3e-09	2.04	362	0.16
10	50	49	8.2e-09	7.27	589	0.16
10	100	54	8.8e-09	45.53	1167	0.21
30	3	31	1.6e-09	0.28	108	0.14
30	10	38	5.9e-09	2.01	367	0.15
30	30	44	6.9e-09	28.58	996	0.19
30	50	43	1.5e-09	104.09	1739	0.29
30	100	45	6.3e-09	744.48	3776	0.76
50	3	34	5.6e-09	0.55	170	0.14
50	10	40	1.3e-09	6.54	585	0.16
50	30	39	9.1e-09	97.10	1685	0.29
50	50	42	5.7e-09	421.40	2793	0.52
50	100	47	1.1e-09	3369.20	5607	1.58

Table 8.1: Numerical results with PRA and *linprog* for the computation of normalized Nash equilibria

		PRA			PRALP				PSM			
N	K	LS	$term$	$time$	LS	LP	$term$	$time$	LP	QP	$V(x)$	$time$
2	3	30	2.7e-09	0.14	26	2	3.2e-15	0.45	6	2	1.8e-12	0.31
2	10	30	1.8e-09	0.13	26	2	1.2e-14	0.27	6	2	0.0e+00	0.33
2	30	36	1.2e-09	0.20	32	2	8.2e-15	0.39	10	4	3.8e-10	0.58
2	50	42	1.4e-09	0.36	38	2	2.1e-11	0.70	10	4	0.0e+00	0.79
2	100	44	1.1e-09	1.06	40	2	1.8e-10	1.85	8	3	5.8e-11	1.13
5	3	24	1.2e-09	0.10	20	2	3.3e-15	0.33	20	2	3.6e-12	0.54
5	10	38	1.8e-09	0.17	34	2	3.4e-14	0.45	25	4	1.7e-08	0.68
5	30	38	1.8e-09	0.51	34	2	4.9e-14	1.10	25	4	3.4e-09	1.07
5	50	50	2.3e-09	1.73	47	2	6.9e-14	3.23	20	3	4.7e-09	1.30
5	100	49	2.7e-09	7.44	46	2	6.6e-09	11.79	20	3	5.7e-07	2.39
10	3	29	4.8e-09	0.13	25	2	5.1e-14	0.30	40	3	8.5e-08	0.88
10	10	34	7.6e-09	0.31	31	2	5.3e-14	0.65	40	3	7.1e-09	0.92
10	30	40	2.3e-09	2.04	37	2	3.8e-14	3.63	40	3	-8.7e-11	1.46
10	50	49	8.2e-09	7.27	46	2	7.6e-12	11.85	50	4	0.0e+00	2.90
10	100	54	8.8e-09	45.53	53	5	9.2e-09	78.83	50	4	9.3e-10	5.92
30	3	31	1.6e-09	0.28	27	2	1.6e-13	0.59	180	4	1.1e-08	3.18
30	10	38	5.9e-09	2.01	35	2	2.6e-13	3.62	150	4	4.8e-09	3.23
30	30	44	6.9e-09	28.58	44	7	6.9e-09	60.57	150	4	-4.5e-09	8.41
30	50	43	1.5e-09	104.09	42	4	5.3e-10	167.17	150	4	5.1e-07	10.51
30	100	45	6.3e-09	744.48	45	6	3.1e-09	1439.62	180	5	2.8e-08	51.56
50	3	34	5.6e-09	0.55	31	2	4.3e-13	1.54	300	4	1.8e-09	5.45
50	10	40	1.3e-09	6.54	36	2	7.5e-11	10.85	250	4	9.1e-08	5.79
50	30	39	9.1e-09	97.10	39	4	9.1e-09	156.02	250	3	9.1e-09	20.37
50	50	42	5.7e-09	421.40	39	3	5.1e-09	582.19	250	4	9.8e-08	62.17
50	100	47	1.1e-09	3369.20	-				250	4	1.1e-07	53.63

Table 8.2: Numerical results with PRA, PRALP and PSM for the basic economic model

		PRA			PRALP				PSM			
N	K	LS	$term$	$time$	LS	LP	$term$	$time$	LP	QP	$V(x)$	$time$
3	3	25	5.9e-09	0.08	21	2	9.6e-15	0.65	12	2	-1.1e-11	0.90
3	10	41	1.2e-09	0.15	37	2	1.8e-14	0.35	9	2	2.6e-09	0.38
3	30	40	7.0e-09	0.29	37	2	1.4e-14	0.61	12	3	5.0e-07	0.60
3	50	44	6.3e-09	0.61	41	2	1.3e-14	1.15	12	3	3.9e-09	0.91
3	100	50	1.6e-09	2.60	46	2	5.4e-10	4.32	12	3	1.3e-09	1.55
6	3	33	8.0e-09	0.12	30	2	2.7e-14	0.40	24	3	-1.8e-11	0.63
6	10	35	1.6e-09	0.18	31	2	3.2e-14	0.51	24	3	-7.3e-11	0.65
6	30	46	1.1e-09	0.88	42	2	3.0e-14	1.71	24	3	1.3e-07	1.00
6	50	43	6.6e-09	2.19	40	2	3.3e-14	3.97	24	3	6.8e-08	1.45
6	100	53	3.7e-09	12.07	50	2	6.5e-12	19.12	30	4	5.8e-08	3.52
11	3	33	2.3e-09	0.14	29	2	5.2e-14	0.32	44	3	-2.5e-10	0.94
11	10	35	3.2e-09	0.36	32	2	2.9e-14	0.74	44	3	1.3e-10	0.98
11	30	46	3.1e-09	2.76	43	2	5.7e-14	4.94	44	3	1.5e-07	1.62
11	50	47	1.3e-09	8.97	43	2	2.5e-11	14.38	55	4	-3.5e-10	3.10
11	100	52	6.6e-09	55.67	52	5	1.2e-09	95.38	55	4	-1.2e-09	6.49
31	3	39	1.7e-09	0.35	35	2	1.5e-09	0.70	155	3	4.3e-10	2.76
31	10	41	2.2e-09	2.37	38	2	1.7e-13	4.30	124	3	3.4e-10	2.92
31	30	41	5.6e-09	29.61	41	7	5.6e-09	64.11	155	4	9.6e-09	8.78
31	50	50	4.9e-09	139.75	50	6	4.9e-09	232.52	155	4	6.8e-08	10.13
31	100	51	2.7e-09	925.73	51	5	2.7e-09	1635.19	155	4	-1.1e-08	23.02
51	3	31	8.5e-09	0.55	28	2	5.4e-09	1.61	153	2	1.0e-08	2.75
51	10	37	3.8e-09	6.51	34	2	4.0e-09	10.90	204	3	3.8e-07	4.68
51	30	45	2.1e-09	120.94	45	5	2.1e-09	201.53	255	4	5.2e-10	21.96
51	50	44	4.9e-09	473.18	44	5	4.9e-09	816.16	255	4	5.8e-10	67.16
51	100	50	8.5e-09	3878.45	-				357	6	1.7e-07	407.03

Table 8.3: Numerical results with PRA, PRALP and PSM for the modified basic economic market model

		PSMvert				PSM			
N	K	$iter$	QP	$V(x)$	$time$	LP	QP	$V(x)$	$time$
2	3	3	2	1.8e-12	0.12	6	2	1.8e-12	0.31
2	10	3	2	7.3e-12	0.11	6	2	0.0e+00	0.33
2	30	5	4	3.9e-10	0.16	10	4	3.8e-10	0.58
2	50	5	4	5.8e-11	0.17	10	4	0.0e+00	0.79
2	100	4	3	0.0e+00	0.18	8	3	5.8e-11	1.13
5	3	4	2	3.6e-12	0.12	20	2	3.6e-12	0.54
5	10	5	4	1.7e-08	0.16	25	4	1.7e-08	0.68
5	30	5	4	3.4e-09	0.20	25	4	3.4e-09	1.07
5	50	4	3	4.8e-09	0.22	20	3	4.7e-09	1.30
5	100	4	3	5.7e-07	0.31	20	3	5.7e-07	2.39
10	3	4	3	8.5e-08	0.14	40	3	8.5e-08	0.88
10	10	4	3	7.1e-09	0.18	40	3	7.1e-09	0.92
10	30	4	3	-1.2e-10	0.24	40	3	-8.7e-11	1.46
10	50	5	4	5.8e-11	0.41	50	4	0.0e+00	2.90
10	100	5	4	9.3e-10	0.90	50	4	9.3e-10	5.92
30	3	5	3	2.3e-08	0.21	180	4	1.1e-08	3.18
30	10	5	4	2.5e-09	0.65	150	4	4.8e-09	3.23
30	30	5	4	-4.2e-09	4.19	150	4	-4.5e-09	8.41
30	50	5	4	5.1e-07	3.49	150	4	5.1e-07	10.51
30	100	6	5	2.7e-08	32.21	180	5	2.8e-08	51.56
50	3	5	3	-2.6e-09	0.35	300	4	1.8e-09	5.45
50	10	5	4	8.9e-08	1.59	250	4	9.1e-08	5.79
50	30	5	3	9.1e-09	13.85	250	3	9.1e-09	20.37
50	50	5	4	9.7e-08	49.44	250	4	9.8e-08	62.17
50	100	5	4	1.1e-07	26.05	250	4	1.1e-07	53.63
100	3	4	2	2.6e-08	0.81	400	2	2.6e-08	6.46
100	10	4	3	2.2e-08	5.92	400	3	2.2e-08	12.41
100	30	5	4	8.9e-08	94.78	500	4	8.9e-08	110.47
100	50	5	4	7.3e-07	354.68	500	4	7.3e-07	379.44
100	100	6	5	9.3e-07	2848.33	600	5	9.3e-07	2887.73

Table 8.4: Numerical results with (PSMvert) and without (PSM) explicit knowledge of some vertex sets

		RGS			Algorithm 4		
N	K	*iter*	$\widehat{V}_\rho(x)$	*time*	*iter*	$\widehat{V}_\rho(x)$	*time*
2	3	304	4.6e-05	3.17	379	9.6e-05	0.17
2	10	482	5.3e-05	9.86	13115	1.9e-04	5.51
5	3	500	3.8e-05	10.44	943	9.6e-05	0.55
5	10	915	1.2e-04	82.41	6369	1.4e-04	5.31
10	3	554	1.2e-03	39.34	2625	1.1e-04	3.03
10	10	1154	5.1e-04	436.20	1956	1.3e-04	3.27
30	3	1008	2.9e-04	901.80	11170	9.7e-05	54.06
30	10	2252	5.9e-03	11839.31	50487	1.6e-04	334.26

Table 8.5: Numerical results with RGS and the penalty method Algorithm 4 for small instances of the basic economic market model

Part III

The Extended Transportation Problem

Chapter 9

The Extended Transportation Problem

9.1 Motivation

Since its first mathematical formulation in the 18th century (cf. [52]), the transportation problem is one of the most famous problems in operations research. In the classical transportation problem we have one forwarder who transports a given good from manufacturers to producers while minimizing his transportation costs. We extend the transportation problem towards a more realistic scenario and introduce several forwarders as illustrated in Figure 9.1. In the following lines we will refer to the transportation problem with several forwarders as *extended transportation problem* (ETP).

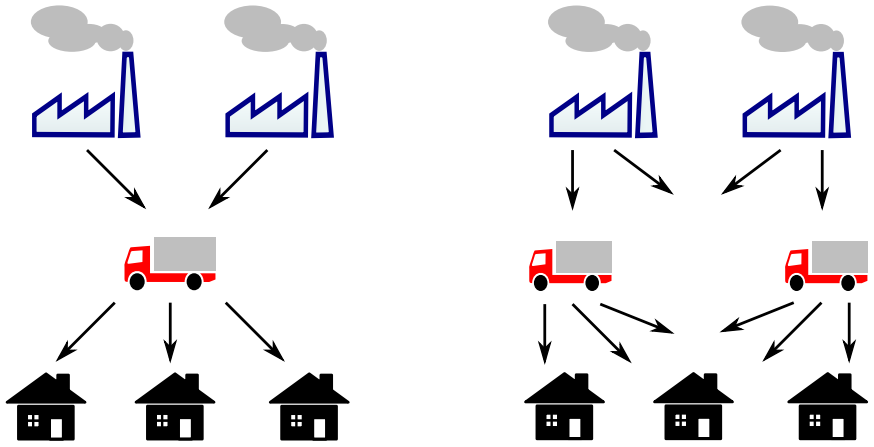

Figure 9.1: The classical transportation problem on the left hand side and the extended transportation problem on the right hand side

In the resulting setting each forwarder wants to minimize his transportation costs while sharing the supply and demand constraints with the remaining forwarders. This is exactly the situation which is addressed in noncooperative game theory. Therefore one may ask about existence and computation of Nash equilibria which are configurations where no forwarder wants to deviate from his strategy given the decision of the remaining forwarders. This question is not only of theoretical interest: Suppose, there is one owner of different factories who negotiates contracts with several competing forwarders. If the contract conditions are set such that the resulting configuration is a Nash equilibrium no forwarder has an incentive to deviate from the contract conditions. This yields a very stable situation and is therefore preferable.

9.2 The Model

Consider N competing forwarding agencies which want to transport one good from R manufacturers to T consumers. Manufacturer r has a production capacity of $S_r \geq 0$, $r \in \{1, \ldots, R\}$, and consumer t needs at least $D_t \geq 0$, $t \in \{1, \ldots, T\}$, units of this good with $\sum_{r=1}^{R} S_r = \sum_{t=1}^{T} D_t$. The unitary transportation cost from manufacturer r to consumer t by forwarder $\nu \in \{1, \ldots, N\}$ is denoted by c_{rt}^{ν} and x_{rt}^{ν} is defined as the number of transported units from manufacturer r to consumer t by forwarder ν. Each forwarder wants to minimize his transportation costs given the decisions of the remaining forwarders, that is forwarder ν faces the optimization problem

$$\min_{x^{\nu} \in \mathbb{R}^{R \times T}} \sum_{r=1}^{R} \sum_{t=1}^{T} c_{rt}^{\nu} x_{rt}^{\nu}$$

subject to his constraints concerning the supply

$$\sum_{\ell=1}^{N} \sum_{t=1}^{T} x_{rt}^{\ell} \;=\; S_r,\; r \in \{1, \ldots, R\},$$

as well as his demand constraints

$$\sum_{\ell=1}^{N} \sum_{r=1}^{R} x_{rt}^{\ell} \;=\; D_t,\; t \in \{1, \ldots, T\}$$

and the nonnegativity condition

$$x_{rt}^{\nu} \;\geq\; 0,\; r \in \{1, \ldots, R\},\; t \in \{1, \ldots, T\}.$$

The search for equilibria in the ETP yields a linear generalized Nash equilibrium problem. We pose the following assumption in order to avoid trivial ETPs.

Assumption 9.2.1 *We have $D_t > 0$ for at least one $t \in \{1, \dots, T\}$.*

The following example will be used throughout this part to illustrate our thoughts.

Example 9.2.2 *Consider the ETP where we have two manufacturers producing one good which is delivered by two forwarders to two consumers, that is, $N = R = T = 2$. The first manufacturer offers one unit of this good and the second one wants to sell two units, thus, we set $S_1 = 1$ and $S_2 = 2$. The demand of the first consumer is given by $D_1 = 2$ and consumer two needs two units, that is, $D_2 = 1$.*

Furthermore, the costs of forwarder $\nu \in \{1, 2\}$ for transporting one good from manufacturer $r \in \{1, 2\}$ to consumer $t \in \{1, 2\}$ are denoted by c^ν_{rt}. The cost matrices $C^\nu = (c^\nu_{rt})$ are given by

$$C^1 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

and

$$C^2 = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

In order to obtain a clear representation of player ν's optimization problem we vectorize it and arrive at

$$Q_\nu(x^{-\nu}): \qquad \min_{x^\nu \in \mathbb{R}^{R \cdot T}} \langle c^\nu, x^\nu \rangle \quad \text{s.t.} \quad Ax^\nu + \sum_{\mu \neq \nu} Ax^\mu = b,\ x^\nu \geq 0$$

given the decisions x^μ, $\mu \neq \nu$, of the remaining players with

$$\begin{aligned} x^\nu &= (x^\nu_{11}, \dots, x^\nu_{1T}, x^\nu_{21}, \dots, x^\nu_{2T}, \dots, x^\nu_{R1}, \dots x^\nu_{RT})^T \in \mathbb{R}^{R \cdot T}, \\ c^\nu &= (c^\nu_{11}, \dots, c^\nu_{1T}, c^\nu_{21}, \dots, c^\nu_{2T}, \dots, c^\nu_{R1}, \dots c^\nu_{RT})^T \in \mathbb{R}^{R \cdot T}, \\ b &= (S_1, \dots, S_R, D_1, \dots, D_T)^T \in \mathbb{R}^{R+T} \end{aligned}$$

and

$$A = \begin{pmatrix} e^T & 0 & \cdots & 0 \\ 0 & e^T & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & e^T \\ I_T & I_T & \cdots & I_T \end{pmatrix} \in \mathbb{R}^{(R+T) \times (R \cdot T)}$$

for all $\nu \in \{1, \dots, N\}$ with $e = (1, \dots, 1)^T \in \mathbb{R}^T$ and the T-dimensional identity matrix I_T.

Remark 9.2.3 *Note that the representation of player ν's constraints in the ETP differs from the model of general inequality constrained LGNEPs presented in the previous parts of this work. Of course, it would be possible to represent the constraints in the ETP in the format*

$$A^{\nu\nu}x^{\nu} + \sum_{\mu \neq \nu} A^{\nu\mu}x^{\mu} \leq b^{\nu}.$$

However, the chosen model enables an explicit and uncluttered treatment of player ν's equality constraints and is therefore preferable.

Example 9.2.4 *In Example 9.2.2 we have*

$$c^1 = \begin{pmatrix} 1 \\ 2 \\ 2 \\ 1 \end{pmatrix}, \; c^2 = \begin{pmatrix} 2 \\ 1 \\ 1 \\ 2 \end{pmatrix}, \; b = \begin{pmatrix} 1 \\ 2 \\ -2 \\ -1 \end{pmatrix}$$

and

$$A \; = \; \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ -1 & 0 & -1 & 0 \\ 0 & -1 & 0 & -1 \end{pmatrix}.$$

However, hereinafter, in order to improve the readability we will prefer to work with the non-vectorized version of Example 9.2.2.

Note that $Q_\nu(x^{-\nu})$ is solvable for all $x^{-\nu} \in \operatorname{dom} X_\nu$ with

$$\operatorname{dom} X_\nu \; = \; \left\{ x^{-\nu} \in \mathbb{R}^{R \cdot T \cdot (N-1)} : \; \exists x^{\nu} \text{ with } Ax^{\nu} + \sum_{\mu \neq \nu} Ax^{\mu} = b, \; x^{\nu} \geq 0 \right\}$$

since the classical transportation problem is solvable. This implies that Assumption 1.2.7 is always valid in the ETP. Furthermore, strong duality yields that also player ν's dual optimization problem $D_\nu(x^{-\nu})$ is solvable which we will need later on in the construction of the gap function V for the ETP.

Also the existence of Nash equilibria in the ETP is always ensured as we shall see in Theorem 10.1.2, such that it is not necessary to refer to Assumption 1.2.6.

9.3 Overview

In Chapter 10, we investigate the numerical computation of Nash equilibria in the ETP. It is possible to compute an $(N-1)$-dimensional set of Nash equilibria very efficiently, such that the question arises which equilibrium one should select in practical applications. This issue, the so-called *Equilibrium selection problem*, is addressed in Chapter 11 where we examine several criteria for the selection of specific Nash equilibria in terms of auxiliary optimization problems and show that for $N = 2$ we even obtain closed form solutions for these optimization problems. However, not each Nash equilibrium of the ETP lies in the efficiently computable $(N-1)$-dimensional set, such that we also apply the projected subgradient method for the computation of arbitrary Nash equilibria in the ETP.

Chapter 10

Computation of Nash Equilibria

10.1 Efficient Computation of Special Nash Equilibria

It is possible to compute an $(N-1)$-dimensional set of Nash equilibria in the ETP very efficiently by solving N linear optimization problems as we shall see in Theorem 10.1.2. Before that we need to introduce some notation.

Let $y^\nu \in \mathbb{R}^{R \cdot T}$ be an optimal point of player ν's optimization problem for $x^{-\nu} = 0$, that is an optimal point of $Q_\nu(0)$. This optimal point exists since the classical transportation problem is solvable. Further, define the vectors

$$\widehat{y}^1 := \begin{pmatrix} y^1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ \widehat{y}^2 := \begin{pmatrix} 0 \\ y^2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ \ldots \ , \ \widehat{y}^N := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ y^N \end{pmatrix} \in \mathbb{R}^{R \cdot T \cdot N}$$

and the set

$$Y \ := \ \operatorname{conv}(\widehat{y}^1, \ldots, \widehat{y}^N).$$

We define the dimension of a convex set by the dimension of its affine hull, that is, the smallest affine subspace that contains Y.

Proposition 10.1.1 *Let Assumption 9.2.1 be valid. Then Y is a $(N-1)$-dimensional set.*

Proof. The affine hull of Y is spanned by the $N-1$ vectors $\widehat{y}^1 - \widehat{y}^2, \ldots, \widehat{y}^1 - \widehat{y}^N$ which are linearly independent due to Assumption 9.2.1. •

In Theorem 10.1.2, we see that each element of Y is a Nash equilibrium and, therefore, we obtain a large set of Nash equilibria which can be computed very efficiently.

Theorem 10.1.2 *Each element of Y is a Nash equilibrium in the ETP.*

Proof. The concatenated KKT systems of all players are given by

$$\begin{aligned} c^\nu + A^T\mu^\nu - \lambda^\nu &= 0 && (10.1)\\ Ax^\nu + \sum_{\mu\neq\nu} Ax^\mu &= b && (10.2)\\ x^\nu &\geq 0 && (10.3)\\ \lambda^\nu &\geq 0 && (10.4)\\ (\lambda^\nu)^T x^\nu &= 0 && (10.5) \end{aligned}$$

for all $\nu \in \{1,\ldots,N\}$. Now take an arbitrary $y \in Y$, that is, there exist nonnegative scalars $\sigma_1,\ldots,\sigma_N \geq 0$ with $\sum_{i=1}^N \sigma_i = 1$ and

$$y \;=\; \sum_{i=1}^N \sigma_i \widehat{y}^i \;=\; \begin{pmatrix} \sigma_1 y^1 \\ \vdots \\ \sigma_N y^N \end{pmatrix}.$$

Note that y is a Nash equilibrium if and only if $\sigma_\nu y^\nu$ solves (10.1)-(10.5) for all $\nu \in \{1,\ldots,N\}$. Since y^ν is an optimal point of $Q_\nu(0)$, there exist vectors $\bar\lambda^\nu \geq 0$ and $\bar\mu^\nu$ with

$$\begin{aligned} c^\nu + A^T\bar\mu^\nu - \bar\lambda^\nu &= 0\\ Ay^\nu &= b\\ y^\nu &\geq 0\\ \bar\lambda^\nu &\geq 0\\ (\bar\lambda^\nu)^T y^\nu &= 0 \end{aligned}$$

for all $\nu \in \{1,\ldots,N\}$. Now, we take the vector $\sigma_\nu y^\nu$ together with the dual variables $\bar\lambda^\nu$ and $\bar\mu^\nu$ of y^ν and plug them into (10.1)-(10.5) for all $\nu \in \{1,\ldots,N\}$. It is easy to see that $(\sigma_\nu y^\nu, \bar\lambda^\nu, \bar\mu^\nu)$ satisfies (10.1), (10.3),(10.4), and (10.5) for all $\nu \in \{1,\ldots,N\}$. Equation (10.2) is also valid for $(\sigma_\nu y^\nu, \bar\lambda^\nu, \bar\mu^\nu)$ since we have

$$A(\sigma_\nu y^\nu) + \sum_{\mu\neq\nu} A(\sigma_\mu y^\mu) \;=\; \sum_{\nu=1}^N \sigma_\nu \underbrace{Ay^\nu}_{=b} \;=\; b\underbrace{\sum_{\nu=1}^N \sigma_\nu}_{=1} \;=\; b.$$

Altogether, we have shown that each element of Y is a Nash equilibrium in the extended transportation problem. •

As we have seen above, we have the $(N-1)$-dimensional set Y which can be computed very easily by solving N classical transportation problems which are linear optimization problems. However, there exist Nash equilibria in the transportation problem that are not elements of Y as we shall see in the following examples.

Example 10.1.3 *In Example 9.2.2, the vector*

$$\bar{x} \;=\; (1,0,0,1,0,0,1,0)^T$$

is a Nash equilibrium since the transportation plan

$$\bar{X}^1 \;:=\; (\bar{x}^1_{rt}) \;=\; \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

is optimal for player one if player two delivers according to

$$\bar{X}^2 \;:=\; (\bar{x}^2_{rt}) \;=\; \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Vice versa, $\bar{X}^2$ is optimal for player two if player one follows the transportation plan $\bar{X}^1$.

Furthermore, the optimization problems $Q_1(0)$ and $Q_2(0)$ have the unique optimal transportation plans

$$y^1 \;=\; \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

and

$$y^2 \;=\; \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}$$

which we did not vectorize in order to improve the readability. Hence, we have

$$Y \;=\; \left\{ \begin{pmatrix} \lambda_1 & 0 \\ \lambda_1 & \lambda_1 \\ 0 & \lambda_2 \\ 2\lambda_2 & 0 \end{pmatrix} \in \mathbb{R}^{4\times 2} :\; \lambda_1+\lambda_2 = 1,\; \lambda_1,\lambda_2 \geq 0 \right\}.$$

We have seen above that the concatenated matrix

$$\bar{x} \;=\; \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 1 & 0 \end{pmatrix}$$

is a Nash equilibrium of this ETP. However, it is straightforward to see that $\bar{x}$ *is not an element of the set* Y*, such that there exist Nash equilibria which are not part of the efficiently computable set* Y*.*

10.2 Projected Subgradient Method

As we have seen in Example 10.1.3, it is not possible to compute each Nash equilibrium of the extended transportation problem by computing the set Y. Therefore, besides the computation of the set Y there is a need to apply some additional methods which are able to compute arbitrary Nash equilibria and not just equilibria in Y.

Remark 10.2.1 *It is, trivially, possible to compute* each *Nash equilibrium of an arbitrary LGNEP by the projected subgradient method. In order to see this recall that all generalized Nash equilibria are global minimal points of the concave function* V *over the polyhedron* W *and, therefore, the Nash equilibria of LGNEPs lie at the boundary of the polyhedron* W*. Hence, initializing PSM with a starting point in the normal cone of the desired Nash equilibrium yields a direct solution in one iteration. However, of course, in practice the desired Nash equilibrium is not known in advance.*

In contrast to the setting in the previous parts of this work, the ETP is an LGNEP with *equality* constraints. Therefore, let us briefly recall the main results and techniques which are necessary to apply the projected subgradient method (PSM) from Chapter 7.5 to the extended transportation problem in the modified format of the ETP.

For given $x^{-\nu}$ player ν's dual problem is given by

$$D_\nu(x^{-\nu}): \qquad \max_{\lambda^\nu \in \mathbb{R}^{R+T}} \langle b - \sum_{\mu \neq \nu} Ax^\mu,\ \lambda^\nu \rangle \quad \text{s.t.} \quad A^T\lambda^\nu \leq c^\nu.$$

We define player ν's optimal value function by

$$\varphi_\nu(x^{-\nu}) \ := \ \max_{\lambda^\nu \in Z_\nu} \langle b - \sum_{\mu \neq \nu} Ax^\mu,\ \lambda^\nu \rangle$$

with

$$Z_\nu \ := \ \{\lambda^\nu \in \mathbb{R}^{R+T} :\ A^T\lambda^\nu \leq c^\nu\}.$$

Furthermore, we define the gap function

$$V(x) \ := \ \sum_{\nu=1}^{N} \langle c^\nu, x^\nu \rangle - \varphi_\nu(x^{-\nu})$$

which does not have to be real-valued outside of W. According to Proposition 1.5.1, the Nash equilibria are exactly the global minimal points of the possibly nonsmooth optimization problem

$$P: \qquad \min V(x) \quad \text{s.t.} \quad x \in W$$

with optimal value zero. Since, in general, V is a nonsmooth function, we use the projected subgradient method which was introduced in Section 7.5 in order to solve P.

Furthermore, let O_ν denote the set of vertices of Z_ν. According to [44, Ex. 6.1.9], the set O_ν is nonempty, that is Z_ν possesses at least one vertex. Then due to the vertex theorem of linear programming we may extend φ_ν to the real-valued function

$$\widehat{\varphi}_\nu(x^{-\nu}) := \max_{\lambda^\nu \in O_\nu} \langle b - \sum_{\mu \neq \nu} Ax^\mu,\ \lambda^\nu \rangle$$

and define

$$O_\nu(x^{-\nu}) := \{\lambda^\nu \in O_\nu :\ \langle b - \sum_{\mu \neq \nu} Ax^\mu,\ \lambda^\nu \rangle = \widehat{\varphi}_\nu(x^{-\nu})\}.$$

If we adapt the result in Proposition 2.4.4 to the ETP we arrive at

$$\begin{aligned} \partial \widehat{V}(x) \ = \ & \left\{ \begin{pmatrix} c^1 \\ A^T\lambda^1 \\ \vdots \\ A^T\lambda^1 \end{pmatrix} :\ \lambda^1 \in \operatorname{conv}\left(O_1(x^{-1})\right) \right\} \\ & + \left\{ \begin{pmatrix} A^T\lambda^2 \\ c^2 \\ A^T\lambda^2 \\ \vdots \\ A^T\lambda^2 \end{pmatrix} :\ \lambda^2 \in \operatorname{conv}\left(O_2(x^{-2})\right) \right\} \\ & + \ldots + \left\{ \begin{pmatrix} A^T\lambda^N \\ \vdots \\ A^T\lambda^N \\ c^N \end{pmatrix} :\ \lambda^N \in \operatorname{conv}\left(O_N(x^{-N})\right) \right\} \end{aligned}$$

and after having evaluated the gap function $\widehat{V}$ at a given point x the vector

$$s := \begin{pmatrix} c^1 \\ A^T\bar{\lambda}^1 \\ \vdots \\ A^T\bar{\lambda}^1 \end{pmatrix} + \begin{pmatrix} A^T\bar{\lambda}^2 \\ c^2 \\ A^T\bar{\lambda}^2 \\ \vdots \\ A^T\bar{\lambda}^2 \end{pmatrix} + \ldots \begin{pmatrix} A^T\bar{\lambda}^{N-1} \\ \vdots \\ A^T\bar{\lambda}^{N-1} \\ c^{N-1} \\ A^T\bar{\lambda}^{N-1} \end{pmatrix} + \begin{pmatrix} A^T\bar{\lambda}^N \\ \vdots \\ A^T\bar{\lambda}^N \\ c^N \end{pmatrix}$$

is a Clarke subgradient of $\widehat{V}$ at x in the ETP. Furthermore, obviously, we have

$$\nabla V(x) = \begin{pmatrix} c^1 \\ \vdots \\ c^N \end{pmatrix} - \begin{pmatrix} A^T\bar{\lambda}^1 \\ \vdots \\ A^T\bar{\lambda}^N \end{pmatrix} + \sum_{\nu=1}^{N} \begin{pmatrix} A^T\bar{\lambda}^\nu \\ \vdots \\ A^T\bar{\lambda}^\nu \end{pmatrix}$$

for all x where V is smooth.

10.3 Numerical Results

We generated random test examples for the ETP for different combinations of N, R and T and applied PSM to these test instances using a Matlab® implementation. We choose the origin as starting point and

$$V(x) < 10^{-6}$$

as stopping criterion and report the needed number of iterations *iter*, the final function value of the gap function *Vterm* as well as the computation time *time*. We see that PSM did not terminate in 6 test examples. This happened since PSM started cycling and we suspect that PSM got stuck in a local minimal point of V since we were able to overcome this difficulty by choosing different starting points. However, note that PSM was able to compute an equilibrium for $N = 50$, $R = 10$ and $T = 50$ in about ten minutes which is a nonsmooth nonconvex optimization problem with 25000 variables.

N	R	T	$iter$	$Vterm$	$time$
2	2	10	5	1.8e-14	0.23
2	2	30	6	2.1e-10	0.42
2	2	50	5	2.9e-08	0.51
2	5	10	15	2.6e-08	0.73
2	5	30	-	-	-
2	5	50	10	1.0e-12	1.51
2	10	10	-	-	-
2	10	30	-	-	-
2	10	50	-	-	-
10	2	10	6	2.2e-08	1.05
10	2	30	10	1.6e-08	2.43
10	2	50	8	2.4e-09	3.11
10	5	10	8	4.0e-12	1.71
10	5	30	12	1.5e-08	5.48
10	5	50	11	7.0e-09	9.13
10	10	10	10	5.2e-09	3.13
10	10	30	14	8.5e-13	12.24
10	10	50	13	1.4e-09	23.15
50	2	10	15	2.2e-09	17.84
50	2	30	-	-	-
50	2	50	-	-	-
50	5	10	15	1.0e-06	22.37
50	5	30	16	5.3e-07	76.19
50	5	50	18	4.5e-07	160.90
50	10	10	13	2.9e-10	34.73
50	10	30	17	9.3e-10	197.02
50	10	50	16	5.3e-08	636.30

Table 10.1: Numerical results with PSM for the extended transportation problem

Chapter 11

Equilibrium Selection

Recall that in the transportation problem with N forwarders it is possible to compute the $(N-1)$-dimensional set Y of Nash equilibria by solving N linear optimization problems which can be done very efficiently. Thus, the question arises which equilibrium one should select. This problem is known in the economics literature as *Equilibrium Selection* or *Nash Selection Problem* (see [37, 38, 45] for possible entry points in this field of research from an economic perspective). In this chapter, we will concentrate on selection techniques that are based on the idea of minimizing an objective function f over the set Y, that is, we consider the optimization problem

$$P_{sel}: \qquad \min_{y} f(y) \quad \text{s.t.} \quad y \in Y$$

with optimal point $y^\star$. In the following sections, we shall discuss several choices of objective functions f that may be reasonable.

Recall that we have

$$Y \;=\; \left\{ y \in \mathbb{R}^{R\cdot T\cdot N} : \exists \lambda \in \Delta_N \text{ with } y = \sum_{i=1}^{N} \lambda_i \widehat{y}^i \right\}$$

where $\Delta_N \in \mathbb{R}^N$ denotes the $(N-1)$-dimensional standard simplex, that is, we have

$$\Delta_N \;:=\; \{\lambda \in \mathbb{R}^N :\; \lambda \geq 0,\; \sum_{i=1}^{N} \lambda_i = 1\}.$$

Therefore, it is possible to transform P_{sel} into an optimization problem over the standard simplex

$$P_{sel}: \qquad \min_{\lambda} f(\sum_{i=1}^{N} \lambda_i \widehat{y}^i) \quad \text{s.t.} \quad \lambda \in \Delta_N.$$

Note that this is an optimization problem with $N-1$ degrees of freedom since it is possible to eliminate one variable from the representation above.

Example 11.0.1 *We will apply the subsequent results to our running example in order to illustrate our thoughts. In Example 9.2.2 the set Y was given by*

$$Y \;=\; \left\{ y : \exists \lambda \in [0,1] \text{ with } y = \lambda \cdot \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} + (1-\lambda) \cdot \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \\ 2 & 0 \end{pmatrix} \right\}$$

which is a one-dimensional set.

Remark 11.0.2 *Of course it is also possible to consider known equilibria that are not elements of the set Y in the selection problem P_{sel}. Therefore, in the following, we will also evaluate the different objective functions on the equilibrium*

$$\bar{x} \;=\; \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 1 & 0 \end{pmatrix}$$

which is an equilibrium in Example 10.1.3 that is not part of the set Y.

11.1 Minimal Total Costs

In order to select the 'best' Nash equilibrium from the set Y, perhaps the most obvious thought would be to minimize the sum of the cost functions of all forwarders over Y, that is, to set

$$f_1(y) \;:=\; \langle c, y \rangle$$

with $c = (c^1, \ldots, c^N)^T$ and to solve

$$P_1 : \qquad \min_y \langle c, y \rangle \quad \text{s.t.} \quad y \in Y$$

which is a linear optimization problem. There is a very efficient way to solve P_1 without employing a numerical method of linear programming since the set of vertices of Y possesses only N elements which are given by

$$\text{vert}Y \;=\; \left\{ \widehat{y}^1, \ldots, \widehat{y}^N \right\}.$$

Therefore, due to the vertex theorem of linear programming, we can solve P_1 by enumeration of the N vertices of Y which can be done very efficiently even for a large number of forwarders.

Example 11.1.1 *The vertices of Y in Example 9.2.2 are given by*

$$\widehat{y}^1 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \widehat{y}^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \\ 2 & 0 \end{pmatrix}.$$

An evaluation of the sum of both cost functions f at these vertices yields $f_1(\widehat{y}^1) = 4$ and $f_1(\widehat{y}^2) = 3$, such that we would select the Nash equilibrium $\widehat{y}^2$ since it is the equilibrium in Y that minimizes the sum of both cost functions. Further, we have $f_1(\bar{x}) = 3$, that is $\bar{x}$ is also an optimal choice.

The advantage of this approach is that it yields a decision criterion which is easy to interpret and very efficiently computable. One disadvantage in this approach is that we will always arrive at an equilibrium that is a vertex of Y because these equilibria are exactly the points in which all goods are delivered by only one forwarder, such that the remaining forwarders transport nothing. Therefore, in Section 11.2 we shall consider an approach that aims at obtaining a Nash equilibrium which possesses a uniform distribution of all goods within the forwarders.

11.2 Uniform Distribution of Goods

Another approach of choosing a Nash equilibrium from the set Y may be to distribute the good from manufacturers to the consumers in a way that the maximal delivery size is minimal. This has at least two advantages: First, this approach is very risk averse and may be preferable if the costs of transportation errors is very high. Second, it enforces a uniform transportation over all forwarders and possible combinations. This is also the big difference to the approach in Section 11.1 where all units were delivered by only one forwarder. Thus, we set

$$f_2(y) \ := \ \max_{r,t,\nu} y^{\nu}_{r,t}$$

and obtain the optimization problem

$$P_2 : \qquad \min_{y} \max_{r,t,\nu} y^{\nu}_{r,t} \quad \text{s.t.} \quad y \in Y$$

which, using the epigraphical reformulation, can be reformulated, such that P_2 is equivalent to the linear optimization problem

$$\begin{aligned} \min_{(y,\alpha)} \alpha \qquad \text{s.t.} \qquad & y \in Y, \\ & y^{\nu}_{r,t} \leq \alpha, \ r = \{1, \ldots, R\}, \ t = \{1, \ldots, T\}, \ \nu = \{1, \ldots, N\}. \end{aligned}$$

Note that the polyhedron Y is given in the so-called *V-representation* of polyhedral sets (cf. [72]), that is, as convex hull of its vertices. However, standard solvers of linear programming require the so-called *H-representation* of polyhedral feasible sets, that is, a representation as intersection of finitely many half spaces. Fortunately, as mentioned in the beginning of Chapter 11, we can overcome this difficulty by transforming the linear optimization problem into 'λ-variables' and arrive at

$$\begin{aligned} \min_{(\lambda,\alpha)} \alpha \qquad \text{s.t.} \qquad & \lambda \geq 0, \\ & \sum_{i=1}^{N} \lambda_i = 1, \\ & \sum_{i=1}^{N} \lambda_i \widehat{y}^i_{r,t} \leq \alpha, r = \{1,\ldots,R\},\ t = \{1,\ldots,T\}, \end{aligned}$$

which is an optimization problem whose representation is accessible to standard solvers in linear optimization.

Example 11.2.1 *In Example 9.2.2, direct inspections show that finding the equilibrium in Y with the smallest component is equivalent to determining $\lambda \geq 0$, such that*

$$\lambda \cdot 1 \ = \ (1-\lambda) \cdot 2$$

since 1 and 2 are the largest components of $\widehat{y}^1$ and $\widehat{y}^2$, respectively, and the order is not influenced by scaling with λ. Therefore, we arrive at $\lambda = \frac{2}{3}$, that is, we would select the Nash equilibrium

$$y^\star \ = \ \begin{pmatrix} \frac{2}{3} & 0 \\ \frac{2}{3} & \frac{2}{3} \\ 0 & \frac{1}{3} \\ \frac{2}{3} & 0 \end{pmatrix}$$

from the set Y which, indeed, possesses a very uniform distribution of the good over all forwarders and possible combinations. Further, we have

$$f_2(\bar{x}) = 1 \ > \ \frac{2}{3} = f_2(y^\star),$$

that is, in this situation we would not choose the equilibrium $\bar{x}$.

11.3 Minimizing the Sum of Squares

A more technical approach addresses the fact that one may not expect to obtain one unique equilibrium by the decision rules in the Sections 11.1 and

11.2 since they result from solving linear optimization problems that may have nonisolated solutions. A standard approach to tackle this issue is to choose *norm minimal* solutions, that is, solutions with minimal Euclidean norm of. These solutions are unique since they can be obtained as optimal points of an unconstrained strictly convex quadratic optimization problem. Thus, we set

$$f_3(y) \ := \ \|y\|_2^2$$

and obtain the convex quadratic optimization problem

$$P_3 : \qquad \min_y y^T y \quad \text{s.t.} \quad y \in Y.$$

which possesses a unique solution.

Example 11.3.1 *In Example 9.2.2 solving P_3 is equivalent to computing a scalar $\lambda \in [0,1]$, such that*

$$\begin{aligned} \left\| \lambda \cdot \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} + (1-\lambda) \cdot \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \\ 2 & 0 \end{pmatrix} \right\|_F^2 &= \left\| \begin{pmatrix} \lambda & 0 \\ \lambda & \lambda \\ 0 & 1-\lambda \\ 2(1-\lambda) & 0 \end{pmatrix} \right\|_F^2 \\ &= 3\lambda^2 + 5(1-\lambda)^2 \end{aligned}$$

is minimal where we used the Frobenius norm

$$\|A\|_F \ := \ \sum_{i,j} a_{i,j}^2.$$

Thus, we compute $\lambda = \frac{5}{8}$ and choose the Nash equilibrium

$$y^\star \ = \ \begin{pmatrix} \frac{5}{8} & 0 \\ \frac{5}{8} & \frac{5}{8} \\ 0 & \frac{3}{8} \\ \frac{6}{8} & 0 \end{pmatrix}$$

from the set Y with $f_3(y^\star) = \frac{15}{8}$. Further, we have $f_3(\bar{x}) = 3$, that is, the equilibrium $\bar{x}$ is not an optimal choice.

Note that the obtained norm minimal solution is not sparse at all. However, in the following section we will see that it is possible to compute the sparsest equilibria in the set Y very efficiently.

11.4 Sparse Equilibria

One may also think of situations where it is advantageous to select a sparse equilibrium, that is, an equilibrium with as little non-zero components as possible. This could be preferable in situations where each non-zero entry y_{rt}^{ν} is associated with high fix costs that occur for 'activating' the combination x_{rt}^{ν} which for instance might be the costs for constructing new infrastructure or legal fees for designing new contracts.

Selecting the sparsest equilibrium is equivalent to minimizing the so-called *zero norm* of a given vector $y \in \mathbb{R}^n$ which is defined by

$$\|y\|_0 \;:=\; \sum_{i=1}^{n} |y_i|^0 \;=\; |\{1 \le i \le n :\; y_i \neq 0\}|$$

where we defined $0^0 := 0$, that is, we set

$$f_4(y) \;:=\; \|y\|_0$$

and obtain the ℓ_0-minimization problem

$$P_4 : \qquad \min_y \|y\|_0 \quad \text{s.t.} \quad y \in Y$$

In general, the ℓ_0-minimization problem is computationally intractable and is therefore replaced by minimizing the ℓ_1-norm of y (cf., e.g., [10, 16, 9]).

However, in the ETP, the situation differs greatly from the general case. First, the ℓ_1-minimization does not make any sense for the ETP since each equilibrium of Y possesses the same ℓ_1-norm which follows from the following, more general, result.

Proposition 11.4.1 *There exists a scalar $D \in \mathbb{R}$ with*

$$\|x\|_1 \;=\; D$$

for all $x \in W$, that is, each $x \in W$ possesses the same ℓ_1-norm.

Proof. Let $x \in W$. Then we have

$$\begin{aligned}
\|x\|_1 &= \sum_{\nu=1}^{N}\sum_{t=1}^{T}\sum_{r=1}^{R} x_{rt}^{\nu} \\
&= \sum_{t=1}^{T} \underbrace{\sum_{\nu=1}^{N}\sum_{r=1}^{R} x_{rt}^{\nu}}_{=D_t} \\
&= \sum_{t=1}^{T} D_t,
\end{aligned}$$

that is, the assertion is shown with $D := \sum_{t=1}^{T} D_t$. •

Second, for the extended transportation problem the ℓ_0-minimization is tractable since it can be reduced to selecting the most sparsest vertex $\widehat{y}^\nu$, $\nu \in \{1, \ldots, N\}$, of the set Y. This can be seen easily, since the number of zeros of each element in Y cannot be higher than the number of zeros of a vertex $\widehat{y}^\nu$, $\nu \in \{1, \ldots, N\}$ of Y because the creation of a 'new zero' as convex combination of nonnegative elements is not possible.

Example 11.4.2 *In Example 9.2.2, we have $\|y\|_1 = 3 = D_1 + D_2$ for all $y \in Y$. Further, $\widehat{y}^2$ is an optimal point of P_4 since it belongs to the sparsest player solution. The equilibrium $\bar{x}$ is not as sparse at $y^\star = \widehat{y}^2$ and therefore not preferable.*

11.5 Closed Form Solutions for $N = 2$

Two player games form an important field of research in the existing literature on game theory (cf. [55, 70]). Also for ETPs, the case $N = 2$ deserves some special attention since for $N = 2$ the set of efficiently computable Nash equilibria Y is a one-dimensional line segment which implies that the selection problem P_{sel} is a *one*-dimensional optimization problem. In contrast to $\mathbb{R}^n$ for $n > 1$, the one-dimensional strategy space $\mathbb{R}$ is an ordered field which sometimes enables the possibility to find closed form solutions.

Remark 11.5.1 *Note that $N = 2$ is the only restriction in this section. Particularly, each player may possess a large number of decision variables and constraints.*

In Sections 11.1 and 11.4 we selected the desired equilibrium out of a set of N equilibria which just shrinks to a set of two equilibria in the case $N = 2$. However, as described in Sections 11.2 and 11.3, the computation of the equilibrium that possesses a uniform distribution of goods or minimizes the sum of squares, respectively, requires the solution of a linear or a quadratic optimization problem, respectively. In the case $N = 2$ we can avoid the numerical solution of these optimization problems and obtain closed analytical solutions instead as stated in the following propositions.

As mentioned at the beginning of Chapter 11, the problem of selecting a Nash equilibrium y out of the set Y can be transformed into determining a suitable weight $\lambda \in \Delta_N$. In the following, we will denote the desired Nash equilibrium by $y^\star \in \mathbb{R}^{2 \cdot T \cdot R}$ and the corresponding weight by $\lambda^\star \in \mathbb{R}$.

Proposition 11.5.2 *Let $\widehat{y}^1_{max}, \widehat{y}^2_{max} \in \mathbb{R}$ be the largest entries of the vectors $\widehat{y}^1$ and $\widehat{y}^2$, respectively, and define*

$$\lambda^\star \ := \ \frac{\widehat{y}^2_{max}}{\widehat{y}^1_{max} + \widehat{y}^2_{max}}.$$

Then the vector

$$y^\star \ = \ \lambda^\star \widehat{y}^1 + (1 - \lambda^\star)\widehat{y}^2$$

is the Nash equilibrium that solves P_2, that is, the equilibrium which enforces a uniform distribution of all goods.

Proof. It is straightforward to see that in order to solve P_2 we have to choose λ, such that we arrive at

$$\lambda \widehat{y}^1_{max} \ = \ (1 - \lambda)\widehat{y}^2_{max}$$

since the position of the maximal component is not influenced by the multiplication of the corresponding vectors with the scalar $\lambda \geq 0$. Further, Assumption 9.2.1 implies

$$\widehat{y}^1_{max} + \widehat{y}^2_{max} \ \neq \ 0,$$

such that we have

$$\lambda^\star \ := \ \frac{\widehat{y}^2_{max}}{\widehat{y}^1_{max} + \widehat{y}^2_{max}}.$$

•

Proposition 11.5.3 *Let us define*

$$\lambda^\star := \frac{\|\widehat{y}^2\|_F^2}{\|\widehat{y}^1\|_F^2 + \|\widehat{y}^2\|_F^2}.$$

Then the vector

$$y^\star \ = \ \lambda^\star \widehat{y}^1 + (1 - \lambda^\star)\widehat{y}^2$$

is the Nash equilibrium that solves P_3, that is, the equilibrium with minimal Euclidean norm.

Proof. It is easy to see that the vectors $\widehat{y}^1$ and $\widehat{y}^2$ are orthogonal. Therefore, we have

$$f(\lambda) \ := \ \|\lambda \widehat{y}^1 + (1 - \lambda)\widehat{y}^2\|_F^2 \ = \ \lambda^2 \|\widehat{y}^1\|_F^2 + (1 - \lambda)^2 \|\widehat{y}^2\|_F^2$$

and differentiating f with respect to λ yields

$$f'(\lambda) \ = \ 2\lambda \|\widehat{y}^1\|_F^2 - 2(1 - \lambda)\|\widehat{y}^2\|_F^2.$$

Then we have

$$\begin{aligned} f'(\lambda) = 0 \quad &\Longleftrightarrow \quad 2\lambda\|\widehat{y}^1\|_F^2 = 2(1-\lambda)\|\widehat{y}^2\|_F^2 \\ &\Longleftrightarrow \quad \lambda(\|\widehat{y}^1\|_F^2 + \|\widehat{y}^2\|_F^2) = \|\widehat{y}^2\|_F^2 \\ &\Longleftrightarrow \quad \lambda^\star = \frac{\|\widehat{y}^2\|_F^2}{\|\widehat{y}^1\|_F^2 + \|\widehat{y}^2\|_F^2} \end{aligned}$$

where

$$\|\widehat{y}^1\|_F^2 + \|\widehat{y}^2\|_F^2 \ \neq \ 0$$

is ensured by Assumption 9.2.1. Since f is a strongly convex function, we have shown that $\lambda^\star$ is the *unconstrained* global minimal point of f on $\mathbb{R}$. In particular, we also have

$$\lambda^\star \in [0, 1]$$

which implies that $\lambda^\star$ is also an optimal point of the *constrained* optimization problem P_3 which proves the desired assertion. •

Chapter 12

Outlook and Conclusions

This work contains the first systematic treatment of linear generalized Nash equilibrium problems (LGNEPs). We have examined theoretical aspects of LGNEPs as well as numerical algorithms and our main application - the extended transportation problem.

We have seen that there is a deep intrinsic connection between smoothness properties of a gap function that arises from a reformulation of the LGNEP as optimization problem and some regularity conditions of its feasible set. In particular, we have introduced a new regularity condition, the so-called *cone condition*. In contrast to stronger conditions like LICQ oder SMFC, the cone condition does not enforce unique KKT multipliers and seems to be a suitable tool to deal with nonunique KKT multipliers. Multiple KKT multipliers are known to cause severe theoretical and numerical problems (see [41] for a recent discussion). In this context, it would be interesting to study the interplay between the cone condition and these phenomena in a more general setting, which we leave for future research.

We have also considered LGNEPs from a numerical point of view and applied some algorithms in order to compute Nash equilibria for LGNEPs. Particularly, we designed a nonsmooth optimization method for the solution of LGNEPs which showed very promising performance and is based on subdifferentials in the sense of Clarke. It is also possible to compute exact formulas for the Fréchet or the Mordukhovich subdifferential (cf. [69]) and design numerical methods that are based on these subdifferentials. However, this is left to future research. Additionaly, it may also be interesting to apply a finite algorithm like Lemke's method in order to solve the concatenated KKT systems of LGNEPs.

Bibliography

[1] K.J. Arrow, G. Debreu: *Existence of an equilibrium for a competitive economy*. Econometrica 22, 265-290 (1954)

[2] J.-P. Aubin, H. Frankowska: *Set-valued analysis*. Springer Science & Business Media, Berlin (2009)

[3] A. Auslender: *Optimisation: Méthodes Numériques*. Masson, Paris (1976)

[4] G. Bank, J. Guddat, D. Klatte, B. Kummer, D. Tammer: *Non-linear Parametric Optimization.* Birkhäuser Verlag, Basel, (1983)

[5] A. Bensoussan: *Points de Nash dans le cas de fonctionnelles quadratiques et jeux differentiels lineaires a N personnes.* SIAM Journal on Control and Optimization 12, 460499 (1974)

[6] F. Bonnans, A. Shapiro: *Perturbation Analysis of Optimization Problems.* Springer Series in Operations Research. Springer-Verlag, New York (2000)

[7] J.M. Borwein, A.S. Lewis: *Convex Analysis and Nonlinear Optimization - Theory and Examples*. Springer, New York (2006)

[8] J.V. Burke, A.S. Lewis, M.L. Overton: *A robust gradient sampling algorithm for nonsmooth, nonconvex optimization*. SIAM Journal on Optimization 15, 751-779 (2005)

[9] E. Candès, T. Tao: *Decoding by linear programming*. IEEE Transactions on Information Theory 51, 4203-4215 (2005)

[10] E. Candès, J. Romberg, T. Tao: *Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information*. IEEE Transactions on Information Theory 52, 489-509 (2006)

[11] F.H. Clarke: *Optimization and Nonsmooth Analysis*. SIAM, Philadelphia (1990)

[12] W.W. Cooper, L.M. Seiford, K. Tone: *Data Envelopment Analysis.* Springer, New York (2007)

[13] R.W. Cottle, J.-S. Pang, R.E. Stone: *The Linear Complementarity Problem.* SIAM, Philadelphia, 1992

[14] J.M. Danskin: *The Theory of Max-Min and its Applications to Weapons Allocation Problems.* Springer, New York (1967)

[15] G. Debreu: *A social equilibrium existence theorem.* Proceedings of the National Academy of Sciences 38, 886-893 (1952)

[16] D. Donoho: *For most large underdetermined systems of linear equations, the minimal ell-1 norm solution is also the sparsest solution.* Communications on Pure and Applied Mathematics, 59, 797-829 (2006)

[17] D. Dorsch, H.Th. Jongen, V. Shikhman: *On structure and computation of generalized Nash equilibria.* SIAM Journal on Optimization 23, 452-474 (2013)

[18] D. Dorsch, H.Th. Jongen, V. Shikhman: *On intrinsic complexity of Nash equilibrium problems and bilevel optimization.* Journal of Optimization Theory and Applications 159, 606-634 (2013)

[19] A. Dreves: *Globally convergent algorithms for the solution of generalized Nash equilibrium problems.* Dissertation, University of Würzburg (2012), available online at
`http://opus.bibliothek.uni-wuerzburg.de/volltexte/2012/6982`.

[20] A. Dreves: *Improved error bound and a hybrid method for generalized Nash equilibrium problems.* Computational Optimization and Applications (online 2014), DOI 10.1007/s10589-014-9699-z.

[21] A. Dreves: *Finding all solutions of affine generalized Nash equilibrium problems with one-dimensional strategy sets.* Mathematical Methods of Operations Research 80, 139-159 (2014)

[22] A. Dreves, F. Facchinei, A. Fischer, and M. Herrich: *A new error bound result for Generalized Nash equilibrium problems and its algorithmic application.* Computational Optimization and Applications 59, 63-84 (2015)

[23] A. Dreves, F. Facchinei, C. Kanzow, S. Sagratella: *On the solution of the KKT conditions of generalized Nash equilibrium problems.* SIAM Journal on Optimization 21, 1082-1108 (2011)

[24] A. Dreves, C. Kanzow: *Nonsmooth optimization reformulations characterizing all solutions of jointly convex generalized Nash equilibrium problems*. Computational Optimization and Applications 50, 23-48 (2011)

[25] A. Dreves, C. Kanzow, O. Stein: *Nonsmooth optimization reformulations of player convex generalized Nash equilibrium problems*. Journal of Global Optimization 53, 587-614 (2012)

[26] A. Dreves, N. Sudermann-Merx: *Solving linear generalized Nash equilibrium problems numerically*. Preprint, available at Optimization Online (2015)

[27] F. Facchinei, C. Kanzow: *Generalized Nash equilibrium problems*. Annals of Operations Research 175, 177-211 (2010)

[28] F. Facchinei, C. Kanzow: *Penalty methods for the solution of generalized Nash equilibrium problems*. SIAM Journal on Optimization 20, 2228-2253 (2010)

[29] F. Facchinei, C. Kanzow, S. Karl, and S. Sagratella: *The semismooth Newton method for the solution of quasi-variational inequalities*. Computational Optimization and Applications 62, 85-109 (2015)

[30] F. Facchinei, A. Fischer, V. Piccialli: *Generalized Nash equilibrium problems and Newton methods*. Mathematical Programming 117, 163-194 (2009)

[31] F. Facchinei, J.-S. Pang: *Exact penalty functions for generalized Nash problems*. In G. Di Pillo and M. Roma (eds): Large Scale Nonlinear Optimization. Springer, Berlin, 115-126 (2006)

[32] A. Fischer, M. Herrich, K. Schönefeld: *Generalized Nash equilibrium problems - recent advances and challenges*. Pesquisa Operacional 34, 521-558 (2014)

[33] M. Fukushima: *Restricted generalized Nash equilibria and controlled penalty algorithm*. Computational Management Science 8, 201-218 (2011)

[34] F. Giannessi: *Separation of sets and gap functions for quasi-variational inequalities*. In: Giannessi, F., Maugeri, A. (eds.): in Variational Inequality and Network Equilibrium Problems, pp. 101-121. Plenum Press, New York (1995)

[35] N. Harms, C. Kanzow, O. Stein: *On differentiability properties of player convex generalized Nash equilibrium problems*. Optimization 64, 365-388 (2015)

[36] D.T. Harker: *Generalized Nash games and quasi-variational inequalities*. European Journal of Operational Research 54, 81-94 (1991)

[37] J.C. Harsanyi: *A new theory of equilibrium selection for games with complete information*. Games and Economic Behavior 8, 91-122 (1995)

[38] J.C. Harsanyi, R. Selten: *A general theory of equilibrium selection in games*. Games and Economic Behavior 34, 154-158 (1988)

[39] A. v. Heusinger, C. Kanzow: *Optimization reformulations of the generalized Nash equilibrium problem using Nikaido-Isoda-type functions*. Computational Optimization and Applications 43, 353-377 (2009)

[40] A. v. Heusinger, C. Kanzow: *Relaxation methods for generalized Nash equilibrium problems with inexact line search*. Journal of Optimization Theory and Applications 143, 159-183 (2009)

[41] A.F. Izmailov, M.V. Solodov: *Critical Lagrange multipliers: what we currently know about them, how they spoil our lives, and what we can do about it*. TOP 23, 1-26 (2015)

[42] A. F. Izmailov, M. V. Solodov: *On error bounds and Newton-type methods for generalized Nash equilibrium problems*. Computational Optimization and Applications 59, 201-218 (2014)

[43] H.Th. Jongen, P. Jonker, F. Twilt: *Nonlinear Optimization in Finite Dimensions*. Kluwer, Dordrecht (2000)

[44] H.Th. Jongen, K. Meer, E. Triesch: *Optimization Theory*. Kluwer, Dordrecht (2004)

[45] Y. Kim: *Equilibrium selection in n-person coordination games*. Games and Economic Behavior 15, 203-227 (1996)

[46] J.B. Krawczyk, S. Uryasev: *Relaxation algorithms to find Nash equilibria with economic applications*. Environmental Modeling & Assessment 5, 63-73 (2000)

[47] K. Kubota, M. Fukushima: *Gap function approach to the generalized Nash equilibrium problem*. Journal of Optimization Theory and Applications 144, 511-531 (2010)

[48] J. Kyparisis: *On uniqueness of Kuhn-Tucker multipliers in nonlinear programming*. Mathematical Programming 32, 242-246 (1985)

[49] C. Lemaréchal: *An extension of Davidon methods to non differentiable problems*. Mathematical Programming Study 3, 95-109 (1975)

[50] A.S. Lewis, M.L. Overton: *Nonsmooth optimization via quasi-Newton methods*. Mathematical Programming 141, 135-163 (2013)

[51] A.S. Lewis, S. Zhang: *Nonsmoothness and a variable metric method*. Journal of Optimization Theory and Applications 165, 151-171 (2013)

[52] G. Monge: *Mémoire sur la théorie des déblais et des remblais. Histoire de l'Académie Royale des Sciences de Paris, avec les Mémoires de Mathématique et de Physique pour la même année,* 666-704 (1781)

[53] J. Nash: *Non-cooperative games*. Annals of Mathematics 54, 286-295 (1951)

[54] H. Nikaido, K. Isoda: *Note on noncooperative convex games*. Pacific Journal of Mathematics 5, 807-815 (1955)

[55] M.J. Osborne, A. Rubinstein: *A Course in Game Theory*. MIT Press, Cambridge, Massachusetts (1994)

[56] J.-S. Pang: *Computing generalized Nash equilibria*. Technical Report, Department of Mathematical Sciences, The Johns Hopkins University, Baltimore, MD, October 2002.

[57] J.-S. Pang: *Error bounds in mathematical programming*. Mathematical Programming 79, 299-332 (1997)

[58] D. Ralph, O. Stein: *The C-Index: A new stability concept for quadratic programs with complementarity constraints*. Mathematics of Operations Research 36, 503-526 (2011)

[59] R.T. Rockafellar: *Convex Analysis*. Princeton University Press, Princeton (1970)

[60] R.T. Rockafellar: *Variational Analysis*. Springer, Berlin (2009)

[61] J.B. Rosen: *Existence and uniqueness of equilibrium points for concave N-person games*. Econometrica 33, 520-534 (1965)

[62] D.A. Schiro, J.-S. Pang, U.V. Shanbhag: *On the solution of affine generalized Nash equilibrium problems with shared constraints by Lemke's method*. Mathematical Programming 142, 1-46 (2013)

[63] O. Stein: *Bi-Level Strategies in Semi-Infinite Programming*. Kluwer, Boston (2003)

[64] O. Stein: *Parametrische Optimierung*. Lecture Notes, Karlsruhe Institute of Technology (2014)

[65] O. Stein: *Nichtlineare Optimierung*. Lecture Notes, Karlsruhe Institute of Technology (2015)

[66] O. Stein, N. Sudermann-Merx: *On smoothness properties of optimal value functions at the boundary of their domain under complete convexity*. Mathematical Methods of Operations Research 79, 327-352 (2014)

[67] O. Stein, N. Sudermann-Merx: *The cone condition and nonsmoothness in linear generalized Nash games*. Journal of Optimization Theory and Applications (online 2015), DOI 10.1007/s10957-015-0779-8.

[68] S. Uryasev, R.Y. Rubinstein: *On relaxation algorithms in computation of noncooperative equilibria*. IEEE Transactions on Automatic Control 39, 1263-1267 (1994)

[69] N.T. Van Hang, N.D. Yen: *On the problem of minimizing a difference of polyhedral convex functions under linear constraints*. Journal of Optimization Theory and Applications, online first, DOI: 10.1007/s10957-015-0769-x (2015)

[70] V.V. Vazirani: *Algorithmic game theory* (Vol. 1). T. Roughgarden, E. Tardos (Eds.), Cambridge University Press, Cambridge (2007)

[71] G. Wachsmuth: *On LICQ and the uniqueness of Lagrange multipliers*. Operations Research Letters 41, 78-80 (2013)

[72] G.M. Ziegler: *Lectures on Polytopes*. Springer, New York (1995)

Erklärung über verwendete Hilfsmittel

Ich versichere wahrheitsgemäß, die Dissertation bis auf die in der Abhandlung angegebene Hilfe selbständig angefertigt, alle benutzten Hilfsmittel vollständig und genau angegeben und genau kenntlich gemacht zu haben, was aus Arbeiten anderer und aus eigenen Veröffentlichungen unverändert oder mit Abänderungen entnommen wurde.

Karlsruhe, den 26.01.2016