

Michael Patriksson

**NONLINEAR
PROGRAMMING
AND VARIATIONAL
INEQUALITY
PROBLEMS**

A Unified Approach

Kluwer Academic Publishers

Nonlinear Programming and Variational Inequality Problems

A Unified Approach

by

Michael Patriksson

*Chalmers University of Technology,
Gothenburg, Sweden*



KLUWER ACADEMIC PUBLISHERS
DORDRECHT / BOSTON / LONDON

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN 0-7923-5455-9

Published by Kluwer Academic Publishers,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

Sold and distributed in North, Central and South America
by Kluwer Academic Publishers,
101 Philip Drive, Norwell, MA 02061, U.S.A.

In all other countries, sold and distributed
by Kluwer Academic Publishers,
P.O. Box 322, 3300 AH Dordrecht, The Netherlands.

Printed on acid-free paper

All Rights Reserved

©1999 Kluwer Academic Publishers

No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without written permission from the copyright owner

Printed in the Netherlands.

Preface

Since I started working in the area of nonlinear programming and, later on, variational inequality problems, I have frequently been surprised to find that many algorithms, however scattered in numerous journals, monographs and books, and described rather differently, are closely related to each other. This book is meant to help the reader understand and relate algorithms to each other in some intuitive fashion, and represents, in this respect, a consolidation of the field.

The framework of algorithms presented in this book is called *Cost Approximation*. (The preface of the Ph.D. thesis [Pat93d] explains the background to the work that lead to the thesis, and ultimately to this book.) It describes, for a given formulation of a variational inequality or nonlinear programming problem, an algorithm by means of approximating mappings and problems, a principle for the update of the iteration points, and a merit function which guides and monitors the convergence of the algorithm.

One purpose of this book is to offer this framework as an intuitively appealing tool for describing an algorithm. One of the advantages of the framework, or any reasonable framework for that matter, is that two algorithms may be easily related and compared through its use. This framework is particular in that it covers a vast number of methods, while still being fairly detailed; the level of abstraction is in fact the same as that of the original problem statement.

Another purpose of the book is to provide a convergence analysis of the algorithms in the framework. The analysis is performed under different interesting combinations of choices of implementation and under different combinations of assumptions on the problem being solved and the algorithm devised for it. The analysis compares favourably with previous attempts to describe algorithms for nonlinear programs and variational inequality problems in a common framework, and establishes the convergence both of new versions of existing algorithms and of methods previously unpublished. A fairly detailed, and to a large degree non-technical, summary of the contents of the book can be found in Section 1.3.

This book can be used in postgraduate courses in nonlinear optimization. If the focus is on algorithm theory, then the prerequisites to (or the first parts of) such a course should cover the fundamental theory of convex analysis (recommended: Rockafellar [Roc70a] or Hiriart-Urruty and Lemaréchal [HiL93a])

and nonlinear optimization (recommended: Bazaraa et al. [BSS93] or Bertsekas [Ber95]). In this case, a course focusing on Chapters 1–4, 7, and the first two sections of Chapter 9 covers some of the fundamentals of nonlinear optimization and variational inequality problems, with emphasis on the theoretical properties of them in association with the construction of algorithms.

A course oriented more towards the numerical aspects of large-scale nonlinear optimization can be based on this book, then requiring a background in numerical analysis and computing (recommended: Bertsekas and Tsitsiklis [BeT89]). In this case, a course would concentrate mostly on Chapters 5, 6, 8, and 9, which include convergence analyses and adaptations of algorithms to problems whose forms typically are found in large-scale settings.

With over 800 references, the book also serves as a reference source for algorithms for the solution of nonlinear optimization and variational inequality problems.

The idea to write this book formed during and after my stay 1994–1995 as a postdoctoral fellow at the University of Washington in Seattle with Prof. Terry Rockafellar. His initial input is greatly appreciated. The bulk of the book was written in 1997, at Linköping University and at Chalmers University of Technology in Göteborg. The planning of the structure of the book benefited from discussions with Prof. Torbjörn Larsson, Prof. Sakis Migdalas, and especially with Dr. Laura Wynter, to whom my deepest thanks are due.

Göteborg (Gothenburg), April, 1998

Michael Patriksson

Notation

\mathbb{R}^n	The n -dimensional space
$\mathbb{R}_+, \mathbb{R}_-$	Non-negative and non-positive orthants of \mathbb{R}^n
\mathbb{R}_{++}	The set $\{x \in \mathbb{R} \mid x > 0\}$
2^X	Set of all subsets of X
$ \mathcal{C} $	Cardinality of a finite set \mathcal{C}
$X \times Y, \prod_{i \in \mathcal{C}} X_i$	Cartesian product of X and Y and of $X_i, i \in \mathcal{C}$
$[a, b]$	Closed interval $\{x \in \mathbb{R} \mid a \leq x \leq b\}$
$\lceil x \rceil$	Upper integer part of x
$\text{int } X, \text{rint } X$	Interior and relative interior of X
$\text{cl } X$	Closure of X
$\text{aff } X, \text{lin } X$	Affine hull and lineality of X
X°	Polar set of X
$\text{conv } X, \text{cone } X$	The convex hull and cone of X
P_X, P_X^Q	Projection onto X , Euclidean and w.r.t. a matrix norm
$[s]_+$	$\max\{s, 0\}$
δ_X	Indicator function of X
N_X, T_X	Normal and tangent cones of X
$\mathcal{F}, \mathcal{F}^*$	A face of a convex set, and the optimal face
$\mathcal{E}_X(d)$	The face of X exposed by the vector d
T_φ	A CA subproblem objective function
∇f	Gradient of f
$\nabla_i f$	Gradient of f with respect to x_i
$\nabla_y \varphi(y, x)$	Gradient of φ with respect to y
$\nabla^X f$	Projected gradient
$\nabla^2 f$	Hessian of f
∇F	Jacobian of F
∂u	Subdifferential of u
$\xi_u(x) \in \partial u(x)$	A subgradient of u at x
$u'(x; d)$	Directional derivative of u at x in the direction of d
u°	Conjugate function of u
Π^{-1}	Inverse mapping of Π

x^t, X^t	Iterates of vectors and sets
x^∞, X^∞	Limit vector and set
y^t, d^t	Subproblem solution and search direction
ℓ_t	Step length
$\{x^t\}, \{x^t\}_{t \in \mathcal{T}}$	Sequence of iterates and a subsequence
f^*	Optimal value of f
$\text{epi } u$	Epigraph of u
$\text{dom } u$	Effective domain of u
$\text{range } \Pi$	Range of Π
$\text{lev}_\alpha f$	Lower level set of f
$f \in C^p$	f is p times continuously differentiable
$f \in C^p$ on X	f is p times continuously differentiable on an open neighbourhood of X
$f \in SC^1$	f is semismooth
u.s.c., l.s.c.	Upper and lower semicontinuity
x_{i-}, x_{i+}	The subvectors $(x_1, \dots, x_{i-1})^T$ and $(x_{i+1}, \dots, x_n)^T$ of x
$x_{\neq i}$	The subvector $(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)^T$ of x
e_i	Unit vector
0^n	The n -dimensional zero vector
I^n	The $n \times n$ -dimensional identity matrix
$\text{symm } Q, \text{diag } Q$	Symmetric part and diagonal part of the matrix Q
$\ \cdot\ $	Euclidean vector norm and induced matrix norm
$\ \cdot\ _Q$	Matrix norm defined by Q
$\ \cdot\ $	Operator norm induced by the Euclidean vector norm
M_f	Lipschitz continuity constant
m_f	Strong convexity (or monotonicity) constant

Contents

Preface	xi
Notation	xiii
1 Introduction	1
1.1 The variational inequality problem	1
1.1.1 Instances of the problem	2
1.1.2 Example applications	8
1.2 The cost approximation algorithm	13
1.2.1 The subproblem phase	13
1.2.2 The updating phase	20
1.2.3 Discussion	22
1.3 Scope and preview	26
1.3.1 Preview	26
1.3.2 Scope	36
2 Technical preliminaries	39
2.1 Solutions to the variational inequality problem	39
2.2 Solutions to the CA subproblem	40
2.3 A posteriori error bounds and lower bounds	42
2.4 Descent properties	44
2.5 Step length rules	49
3 Instances of the cost approximation algorithm	57
3.1 Classic algorithms	57
3.1.1 Linearization methods	57
3.1.2 Interior point algorithms	59
3.1.3 Jacobi and Gauss–Seidel methods	60
3.2 Regularization and proximal point methods	61
3.2.1 Regularization methods	61
3.2.2 Splitting methods	71
3.3 Decomposition–coordination methods	77
3.3.1 A primal algorithm	78
3.3.2 A primal–dual algorithm	79

3.3.3	An augmented Lagrangean method	80
3.4	Decomposition of optimization problems	81
3.5	Relationships among algorithm frameworks	83
3.6	CA algorithms involving u	87
3.6.1	Subgradient algorithms	87
3.6.2	Perturbed CA algorithms	88
3.7	Continuous CA algorithms	90
3.8	A final remark	92
4	Merit functions for variational inequality problems	95
4.1	Introduction	95
4.2	A class of merit functions for variational inequalities	96
4.3	Properties of the merit function ψ	99
4.4	Instances of the merit function ψ	104
4.4.1	The primal and dual gap functions	104
4.4.2	Some differentiable merit functions	106
4.4.3	Unconstrained and complementarity formulations	107
4.4.4	Merit functions for variational inequality problems with multi-valued operators	110
4.5	Stationarity conditions and descent properties	112
4.5.1	Cost approximating mappings independent of x	112
4.5.2	Cost approximating mappings parameterized by x	115
4.5.3	Combined cost approximating mappings	117
4.5.4	Descent from truncated CA subproblems	120
4.6	Primal-dual variational inequalities	122
4.6.1	Introduction	122
4.6.2	A primal merit function	126
4.6.3	Combined cost approximating mappings	130
5	Convergence of the CA algorithm for nonlinear programs	135
5.1	Convergence under an exact line search	135
5.2	Convergence under the Armijo and modified Armijo rule	138
5.2.1	Application to differentiable optimization	138
5.2.2	Application to non-differentiable optimization	140
5.3	A truncation scheme for the CA subproblem	141
5.3.1	The truncation scheme	141
5.3.2	Convergence under an exact line search	143
5.3.3	Convergence under the Armijo rule	145
5.4	Convergence under various step length rules	146
5.4.1	Exact subproblem solutions	146
5.4.2	Inexact subproblem solutions	151
5.4.3	A perturbed CA algorithm	151
5.4.4	Application to convex problems with explicit constraints	152
5.5	A non-monotone CA algorithm	154
5.5.1	Introduction	154
5.5.2	The non-monotone CA algorithm	155

5.5.3	Convergence of the non-monotone CA algorithm	157
5.6	Linear convergence results	161
5.6.1	Application to differentiable optimization	161
5.6.2	Application to non-differentiable optimization	165
6	Convergence of the CA algorithm for variational inequality problems	169
6.1	Iteration-independent cost approximation	169
6.1.1	Convergence under an exact line search	169
6.1.2	Convergence under different step length rules	170
6.1.3	Linear convergence	172
6.2	Iteration-dependent cost approximation	173
6.2.1	Primal application	173
6.2.2	Primal-dual application	176
6.3	Non-descent CA methods	182
6.3.1	Iteration-independent cost approximation	183
6.3.2	Iteration-dependent cost approximation	185
6.4	Averaging schemes and ergodic sequences	186
6.4.1	Non-expansiveness of the CA subproblem mapping	187
6.4.2	Iterative and ergodic schemes	188
7	Finite identification of active constraints and of solutions	191
7.1	Finite identification of active constraints	191
7.1.1	Facial geometry	192
7.1.2	The projected gradient and stationarity conditions	193
7.1.3	Non-degeneracy	194
7.1.4	Identification results	196
7.2	Finite identification of solutions	203
7.2.1	Sharp solutions	203
7.2.2	Finite termination of the CA algorithm	204
8	Parallel and sequential decomposition CA algorithms	211
8.1	Introduction	211
8.1.1	The problem under study	212
8.1.2	Adapting CA to the Cartesian product structure	214
8.1.3	Scope and preview	215
8.2	Sequential decomposition CA algorithms	217
8.3	Synchronized parallel CA algorithms	218
8.3.1	Synchronized parallel computations	219
8.3.2	The synchronized parallel algorithm	220
8.4	Partially asynchronous parallel decomposition CA algorithms	223
8.4.1	Asynchronous parallel computations	223
8.4.2	The partially asynchronous parallel algorithm	225
8.5	Convergence of the sequential decomposition algorithm	226
8.5.1	On the separability assumption of u	226
8.5.2	Convergence under an exact line search	228

8.5.3	Convergence of a truncated algorithm	229
8.5.4	Essentially cyclic decomposition CA algorithms	231
8.5.5	Linear convergence	236
8.6	Convergence of the synchronized parallel algorithms	237
8.6.1	Convergence of a truncated algorithm	238
8.6.2	Convergence under different step length rules	238
8.7	Convergence of the partially asynchronous algorithm	239
8.7.1	Convergence results	239
8.7.2	Qualitative analysis	243
8.8	Variational inequality problems over Cartesian product sets . . .	244
9	A column generation/simplicial decomposition algorithm	253
9.1	Column generation approaches	253
9.1.1	Background	253
9.1.2	Inner representation	254
9.1.3	Simplicial decomposition	256
9.2	The column generation CA algorithm	262
9.2.1	Updating the inner approximation	262
9.2.2	The column generation CA algorithm	263
9.2.3	Instances	264
9.3	Convergence for nonlinear programs	267
9.3.1	Set augmentation	267
9.3.2	Convergence under exact solutions of RMP	270
9.3.3	Convergence of a truncated algorithm	272
9.3.4	An algorithm with general column dropping rules	273
9.4	Convergence for variational inequality problems	273
9.4.1	Convergence under an exact solution of RMP	273
9.4.2	An algorithm with general column dropping rules	274
9.4.3	A primal-dual application	275
	A Definitions	277
	References	283
	Index	325

Chapter 1

Introduction

1.1 The variational inequality problem

Let $X \subseteq \mathbb{R}^n$ be a nonempty, closed and convex set, $u : \mathbb{R}^n \mapsto \mathbb{R} \cup \{+\infty\}$ a lower semicontinuous (l.s.c.), proper¹ and convex function, and $F : \text{dom } u \cap X \mapsto \mathbb{R}^n$ a vector-valued and continuous mapping on $\text{dom } u \cap X$.² The problem under study is defined by three operators: the *normal cone* operator for X ,

$$N_X(x) := \begin{cases} \{z \in \mathbb{R}^n \mid z^T(y - x) \leq 0, \quad \forall y \in X\}, & x \in X, \\ \emptyset, & x \notin X; \end{cases} \quad (1.1)$$

the *subdifferential* operator for u ,

$$\partial u(x) := \{ \xi_u \in \mathbb{R}^n \mid u(y) \geq u(x) + \xi_u^T(y - x), \quad \forall y \in \mathbb{R}^n \};$$

and the mapping F . Consider the problem of finding a vector $x^* \in \mathbb{R}^n$ such that

$$[\text{GVIP}(F, u, X)]$$

$$F(x^*) + \partial u(x^*) + N_X(x^*) \ni 0^n. \quad (1.2)$$

This problem is known as a *generalized variational inequality* ([FaP82]), as a *nonlinear variational inequality* ([Noo75, Noo82a, Noo82b, Noo91b]), and also as a *generalized equation* ([Rob79, Rob82, Rob83]). This problem, and its various special cases, has a large variety of applications in the mathematical and engineering sciences, for example in partial differential equations ([HaS66, DuL72, EkT76, CGL80, KiS80, GLT81, BaC84, Cra84, Rod87, KiO88]), equilibrium problems in games, economics and transportation analysis ([Kar69a,

¹The function u is *proper* if $u(x) < +\infty$ for at least one x and $u(x) > -\infty$ for every x .

²The effective domain $\text{dom } u$ of u is the subset of \mathbb{R}^n for which $u(x) < +\infty$.

Kar69b, Smi79, Daf80, BKS80, AhH82, BaC84, Flo86, Mat87, Zha88, Daf90, ZhD91, Pat94b]), and nonlinear programming ([Roc69a, Sta69, Kar69b, MaS72, Aus76, Roc80, BeT89, HaP90b]).

The set of solutions to $\text{GVIP}(F, u, X)$, which we will denote by $\text{SOL}(F, u, X)$, is nonempty under conditions that are stated in Section 2.1.

We next give a flavour of the large variety of problems that can be modelled as special cases of $\text{GVIP}(F, u, X)$, and introduce the names of the most important ones that we will study in detail.

1.1.1 Instances of the problem

Under the following assumption, the problem $\text{GVIP}(F, u, X)$ can be equivalently stated in terms of the function u rather than its subdifferential mapping.

ASSUMPTION 1.1 (A regularity assumption). $\text{rint}(\text{dom } u) \cap X \neq \emptyset$. □

REMARK 1.2 (Observations). Recall that the normal cone mapping N_X associated with the convex set X is the subdifferential mapping of the *indicator function* δ_X for X ([Roc70a, p. 215]),

$$\delta_X(x) := \begin{cases} 0, & x \in X, \\ +\infty, & x \notin X. \end{cases} \quad (1.3)$$

The assumption is introduced to ensure that $\partial[u + \delta_X](x) = \partial u(x) + \partial \delta_X(x)$, $x \in \text{dom } u \cap X$ (e.g., [Roc81, Thm. 5C]), and may be replaced by, for example, the symmetric condition that $\text{rint}(\text{dom } u) \cap \text{rint } X \neq \emptyset$, where rint denotes relative interior ([Roc70a, Thm. 23.8]); it can be further weakened whenever u is a polyhedral function or X is polyhedral.

Note, finally, that the assumption is fulfilled whenever $\text{dom } u = \mathbb{R}^n$. □

PROPOSITION 1.3 [Pat97] (Equivalent variational inequality formulation). *Under Assumption 1.1, the problem $\text{GVIP}(F, u, X)$ is equivalent to the problem of finding an $x^* \in X$ such that*

$$F(x^*)^T(x - x^*) + u(x) - u(x^*) \geq 0, \quad \forall x \in X. \quad (1.4)$$

PROOF. Consider the convex problem

$$\underset{x \in X}{\text{minimize}} \ h(x) := F(x^*)^T x + u(x), \quad (1.5)$$

where $x^* \in X$. It is clear that (1.4) is equivalent to x^* being a globally optimal solution to this problem. By virtue of Assumption 1.1, we may characterize x^* by the inclusion

$$\partial h(x^*) + N_X(x^*) \ni 0^n \quad (1.6)$$

([Roc70a, Thm. 27.4]). Further, Assumption 1.1 implies that

$$\partial h(x) = F(x^*) + \partial u(x), \quad x \in X; \quad (1.7)$$

combining (1.6) and (1.7) yields the desired result. □

COROLLARY 1.4 (Equivalent variational inequality formulation). *Under Assumption 1.1, the problem $\text{GVIP}(F, u, X)$ is equivalent to the problem of finding an $x^* \in X$ such that*

$$F(x^*)^T(x - x^*) + u'(x^*; x - x^*) \geq 0, \quad \forall x \in X. \quad (1.8)$$

PROOF. The result follows from utilizing that $h'(x^*; x - x^*) \geq 0$ for all $x \in X$ constitutes the necessary and sufficient optimality conditions of x^* in (1.5). \square

EXAMPLE 1.5 (System of variational inequalities over a Cartesian product set). Let the feasible set of $\text{GVIP}(F, u, X)$ be described by a *Cartesian product* of feasible sets,

$$X = \prod_{i \in \mathcal{C}} X_i, \quad X_i \subseteq \mathbb{R}^{n_i}, \quad \sum_{i \in \mathcal{C}} n_i = n, \quad (1.9)$$

for some finite index set \mathcal{C} , where each set X_i is nonempty, closed and convex. Furthermore, the function u is assumed to be separable with respect to the partition of \mathbb{R}^n defined by (1.9), that is, u is of the form

$$u(x) := \sum_{i \in \mathcal{C}} u_i(x_i),$$

where $u_i : \mathbb{R}^{n_i} \mapsto \mathbb{R} \cup \{+\infty\}$ is a lower semicontinuous, proper and convex function for each $i \in \mathcal{C}$.

The mapping F is in general not separable with respect to the given partition of \mathbb{R}^n ; otherwise, the problem $\text{GVIP}(F, u, X)$ would decompose into a number of independent problems of the form $\text{GVIP}(F_i, u_i, X_i)$. [We can therefore argue that the given problem generalizes $\text{GVIP}(F, u, X)$.]

Several examples from this class of variational inequality problems will be given in this and the next section; Chapter 8 is devoted to algorithms that are designed to utilize such a problem structure, and contains further examples. \square

EXAMPLE 1.6 (Nash equilibrium problem). Let $X := \prod_{i=1}^N X_i$ be the product of individual nonempty, closed and convex strategy sets $X_i \subseteq \mathbb{R}^{n_i}$, $\sum_{i=1}^N n_i = n$. We define a *penalty function* $f_i : X \mapsto \mathbb{R}$ for each player, defined on the joint strategy space, and assumed convex and in C^1 on X_i . Further, we let $x \mapsto u(x) := \sum_{i=1}^N u_i(x_i)$ be a l.s.c., proper and convex separable loss function on X . A Nash equilibrium of the non-cooperative N -person game associated with this data is described by a point $x^* \in X$ which, for each $i \in \{1, \dots, N\}$, satisfies

$$f_i(x_{\neq i}^*, x_i^*) + u_i(x_i^*) = \text{minimum}_{x_i \in X_i} \{f_i(x_{\neq i}^*, x_i) + u_i(x_i)\}, \quad (1.10)$$

that is, the players' strategies are optimal with respect to their individual penalty (disutility) functions, based on the strategies of the other players. The optimality conditions for (1.10) define an instance of $\text{GVIP}(F, u, X)$ of

the form described in Example 1.5, in which $x = (x_1, \dots, x_N)$, $X = \prod_{i=1}^N X_i$, and $F = (\nabla_{x_1} f_1, \dots, \nabla_{x_N} f_N)$.

The theory of non-cooperative N -person games was first studied by Cournot (for $N = 2$) and Nash [Nas50, Nas51]; results on the existence and uniqueness of Nash equilibria are given in [Ros65, HaS66, LiS67, Kar72, Fri77, GaM80, Goo80], and applications and computational approaches are given in [Kar69b, GaM80, Pan85, Coh87, BeT89, HaP90b]. \square

REMARK 1.7 (Non-unique representation of $\text{GVIP}(F, u, X)$). $\text{GVIP}(F, u, X)$ is not stated uniquely in terms of the three-tuple $[F, \partial u, N_X]$. For example, the set X can be represented by adding to u the indicator function δ_X of X , defined in (1.3). This infinite penalty function is l.s.c., proper and convex (e.g., [Phe89, p. 40]), as is $u + \delta_X$ (see [vTi84, Sec. 5.4] and [Roc70a, Thm. 5.2]), $\partial \delta_X \equiv N_X$ holds ([Roc70a, p. 215]), and (as stated in Remark 1.2), $\partial[u + \delta_X](x) = \partial u(x) + N_X(x)$, $x \in \text{dom } u \cap X$, holds under Assumption 1.1. So, any convex constraint can be placed either in the description of X or as an infinite penalty added to the description of u , and under Assumption 1.1, therefore, there is no loss of generality in expressing $\text{GVIP}(F, u, X)$ as the *generalized equation*

$$[\text{GE}(F, u)]$$

$$F(x^*) + \partial u(x^*) \ni 0^n. \quad (1.11)$$

(In other words, letting $u := u + \delta_X$.) This problem is a special case of the problem of finding a zero of the sum of two operators (see [Bré73, LiM79, Tse91a, EcB92], and the references cited therein).

It goes without saying that the problem class $\text{GE}(F, u)$ defines a proper subset of $\text{GVIP}(F, u, X)$ whenever $\text{dom } u = \mathbb{R}^n$ and $X \neq \mathbb{R}^n$; this case is, however, not treated separately.

Furthermore, the decomposition of F and ∂u is not unique: adding the gradient mapping ∇h of an arbitrary convex function h to ∂u and subtracting it from F leaves $\text{GVIP}(F, u, X)$ unaltered.

Due to the non-uniqueness of the decompositions of N_X and ∂u and of F and ∂u , there is a large freedom-of-choice in representing an instance of $\text{GVIP}(F, u, X)$ in terms of these mappings. This is important because the algorithms that we shall deal with are defined by different approximations of the three-tuple $[F, \partial u, N_X]$. Hence, depending on the representation of $\text{GVIP}(F, u, X)$, the algorithms for solving $\text{GVIP}(F, u, X)$ will vary as well, both in terms of their interpretation and properties as well as in terms of their convergence requirements. We will make use of the possibility to change the representation of $\text{GVIP}(F, u, X)$ to obtain new and more general results. In particular, we will consider representations based on projecting $\text{GVIP}(F, u, X)$ onto different solution spaces and representations involving the introduction of constraint multipliers. \square

If $u \equiv 0$ in $\text{GVIP}(F, u, X)$ (note that Assumption 1.1 then is satisfied trivially), which is equivalent to assuming that in $\text{GE}(F, u)$, $u \equiv \delta_X$ for some

nonempty, closed and convex set $X \subseteq \mathbb{R}^n$, we obtain the *variational inequality problem* of finding $x^* \in \mathbb{R}^n$ such that

$$[\text{VIP}(F, X)]$$

$$F(x^*) + N_X(x^*) \ni 0^n, \quad (1.12)$$

or, in its more familiar form [utilizing (1.1)] of finding $x^* \in X$ such that

$$F(x^*)^T(x - x^*) \geq 0, \quad \forall x \in X. \quad (1.13)$$

$\text{VIP}(F, X)$ is also known as the *stationary point problem* ([Eav78b]), and x^* as a stationary point.

REMARK 1.8 (Projection characterization). We note for future reference a characterization of the solutions x^* to $\text{VIP}(F, X)$ in terms of a fixed-point involving the projection of a vector defined by x^* onto X . Introducing $\gamma > 0$, we may write (1.13) equivalently as the inequality

$$\begin{aligned} [(1/\gamma)x^* + F(x^*) - (1/\gamma)x^*]^T(x - x^*) &\geq 0, \quad \forall x \in X \\ \iff [(x^* - \gamma F(x^*)) - x^*]^T(x - x^*) &\leq 0, \quad \forall x \in X; \end{aligned}$$

this inequality shows that $x^* - \gamma F(x^*) - x^*$ belongs to the normal cone to X at x^* , which is equivalent to the statement that x^* is the Euclidean projection of the vector $x^* - \gamma F(x^*)$ onto X , or, in other words, $x^* = P_X[x^* - \gamma F(x^*)]$. More generally, introducing a symmetric and positive definite matrix $Q \in \mathbb{R}^{n \times n}$, a similar technique shows that a solution x^* to $\text{VIP}(F, X)$ is characterized by the inequality

$$[(x^* - \gamma Q^{-1}F(x^*)) - x^*]^T Q(x - x^*) \leq 0, \quad \forall x \in X,$$

which is equivalent to x^* being the projection of the vector $x^* - \gamma Q^{-1}F(x^*)$ onto X according to the vector norm $\|z\|_Q := \sqrt{z^T Q z}$ defined by Q , that is, $x^* = P_X^Q[x^* - \gamma Q^{-1}F(x^*)]$.

We note finally that the projection characterizations shown here will reappear in the construction of iterative algorithms for $\text{VIP}(F, X)$. \square

EXAMPLE 1.9 (Traffic equilibrium). Let $\mathcal{G} = (\mathcal{N}, \mathcal{A})$ denote an urban traffic network of nodes (intersections and centroids) and directed links (road sections). A subset \mathcal{C} of $\mathcal{N} \times \mathcal{N}$ defines a set of commodities, associated with pairs k of origins and destinations of trips. It is assumed that the demand for transportation between any pair k of nodes in \mathcal{C} is known; we denote this number by d_k . Letting x_{kr} , $r \in \mathcal{R}_k$, be the flow on route r for commodity (OD pair) k , the set of feasible route flows is described by the constraints

$$\sum_{r \in \mathcal{R}_k} x_{kr} = d_k, \quad \forall k \in \mathcal{C}, \quad (1.14a)$$

$$x_{kr} \geq 0, \quad \forall k \in \mathcal{C}, \quad (1.14b)$$

or, compactly,

$$X := \{x \in \mathbb{R}^{|\mathcal{R}|} \mid \Gamma^T x = d; \quad x \geq 0^{|\mathcal{R}|}\},$$

where $\mathcal{R} := \cup_k \mathcal{R}_k$ and Γ^T is the route-OD pair incidence matrix such that

$$\gamma_{kr} = \begin{cases} 1, & \text{route } r \text{ joins OD pair } k, \\ 0, & \text{otherwise,} \end{cases} \quad r \in \mathcal{R}_k, \quad k \in \mathcal{C}.$$

Assume further that each route $r \in \mathcal{R}_k$, $k \in \mathcal{C}$, is associated with a route cost (or, travel time) function $F_{kr} : \mathbb{R}_+^{|\mathcal{R}|} \mapsto \mathbb{R}_{++}$, which measures the disutility of traversing that route as a function of the volume of traffic on the network. Under the assumption that a traveller chooses the route to her destination which minimizes her cost (travel time) given the current network conditions, a steady-state is characterized by an equilibrium situation in which no traveller can reduce her cost by changing route. Therefore, all routes which are used between any given OD pair have the same, minimal, cost. This equilibrium situation is described as follows (where π_k takes the role of the minimal travel cost in OD pair k):

$$x_{kr}[F_{kr}(x) - \pi_k] = 0, \quad r \in \mathcal{R}_k, \quad k \in \mathcal{C}, \quad (1.15a)$$

$$F_{kr}(x) - \pi_k \geq 0, \quad r \in \mathcal{R}_k, \quad k \in \mathcal{C}. \quad (1.15b)$$

These conditions for a feasible route flow (the *Wardrop* [War52] *equilibrium conditions*) are equivalent to $\text{VIP}(F, X)$, with $F := (F_{kr})_{r \in \mathcal{R}_k, k \in \mathcal{C}}$.³ Note finally that this is an instance of the problem of Example 1.5.

The area of transportation planning was at the forefront in the early developments in algorithms for finite-dimensional variational inequality problems took place in the 1980s; for overviews of the field, see [She85, HaP90b, Nag93, Pat94b, FIH95]. \square

EXAMPLE 1.10 (Saddle point problem). Let $V \subseteq \mathbb{R}^n$ and $W \subseteq \mathbb{R}^m$ be closed convex sets, and $\Pi : V \times W \mapsto \mathbb{R}$ be a continuous function on $V \times W$. The saddle problem associated with Π is to find $(v^*, w^*) \in V \times W$ such that (e.g., [vNe28, Dan67, Roc70a, DeM74a])

$$[\text{SPP}(\Pi, V \times W)]$$

$$\Pi(v^*, w) \leq \Pi(v^*, w^*) \leq \Pi(v, w^*), \quad \forall (v, w) \in V \times W. \quad (1.16)$$

(This problem is closely related to the min-max and max-min problems

$$\underset{v \in V}{\text{minimize maximum}} \Pi(v, w); \quad \underset{w \in W}{\text{maximize minimum}} \Pi(v, w);$$

see, e.g., [Roc70a, Part VII].) Necessary conditions for a saddle point at (v^*, w^*) are that ([Rob76, Rob82])

$$\nabla_v \Pi(v^*, w^*) + N_V(v^*) \ni 0^n; \quad -\nabla_w \Pi(v^*, w^*) + N_W(w^*) \ni 0^m. \quad (1.17)$$

³The Wardrop conditions arise as the primal-dual optimality conditions of the linear program $\text{minimize}_{x \in X} F(x)^T x$ equivalent to $\text{VIP}(F, X)$, see [Pat94b, Thm 3.14.a].