APPLIED OPTIMIZATION

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

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

$\Re^n$	The n-dimensional space
$\Re^n_+$, $\Re^n_-$	Non-negative and non-positive orthants of $\Re^n$
$\Re_{++}$	The set $\{ x \in \Re \mid x > 0 \}$
2^X	Set of all subsets of X
$\lvert\mathcal{C}\rvert$	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 $\{ x \in \Re \mid a \leq x \leq b \}$
$\lceil x \rceil$	Upper integer part of x
$\mathtt{int}\, X$, $\mathtt{rint}\, X$	Interior and relative interior of X
$\mathtt{cl}\, X$	Closure of X
$\mathtt{aff}\, X$, $\mathtt{lin}\, X$	Affine hull and lineality of X
X°	Polar set of X
$\mathtt{conv}\, X$, $\mathtt{cone}\, X$	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
$\mathtt{epi}\, u$	Epigraph of u
$\mathtt{dom}\, u$	Effective domain of u
$\mathtt{range}\, \Pi$	Range of Π
$\mathtt{lev}_\alpha f$	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 $(x_1, \ldots, x_{i-1})^{\mathrm{T}}$ and $(x_{i+1}, \ldots, x_n)^{\mathrm{T}}$ of x
$x_{\neq i}$	The subvector $(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)^{\mathrm{T}}$ of x
e_i	Unit vector
0^n	The n-dimensional zero vector
I^n	The $n \times n$-dimensional identity matrix
$\mathtt{symm}\, Q$, $\mathtt{diag}\, Q$	Symmetric part and diagonal part of the matrix Q
$\\|\cdot\\|$	Euclidean vector norm and induced matrix norm
$\\|\cdot\\|_Q$	Matrix norm defined by Q
$\vert\vert\vert\cdot\vert\vert\vert$	Operator norm induced by the Euclidean vector norm
M_f	Lipschitz continuity constant
m_f	Strong convexity (or monotonicity) constant

Contents

Chapter 1

Introduction

1.1 The variational inequality problem

Let $X \subseteq \Re^n$ be a nonempty, closed and convex set, $u : \Re^n \mapsto \Re \cup \{+\infty\}$ a lower semicontinuous (l.s.c.), proper[1] and convex function, and $F : \mathrm{dom}\, u \cap X \mapsto \Re^n$ a vector-valued and continuous mapping on $\mathrm{dom}\, u \cap X$.[2] The problem under study is defined by three operators: the *normal cone* operator for X,

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

the *subdifferential* operator for u,

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

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

[GVIP(F, u, X)]

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

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,

[1]The function u is *proper* if $u(x) < +\infty$ for at least one x and $u(x) > -\infty$ for every x.
[2]The effective domain $\mathrm{dom}\, u$ of u is the subset of $\Re^n$ 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 GVIP(F, u, X), which we will denote by 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 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 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). $\texttt{int}\,(\texttt{dom}\,u) \cap X \neq \emptyset$. □

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]),

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

The assumption is introduced to ensure that $\partial[u + \delta_X](x) = \partial u(x) + \partial\delta_X(x)$, $x \in \texttt{dom}\,u \cap X$ (e.g., [Roc81, Thm. 5C]), and may be replaced by, for example, the symmetric condition that $\texttt{rint}\ (\texttt{dom}\ u) \cap \texttt{rint}\ X \neq \emptyset$, where $\texttt{rint}$ 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 $\texttt{dom}\,u = \Re^n$. □

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

$$F(x^*)^{\mathrm{T}}(x - x^*) + u(x) - u(x^*) \geq 0, \qquad \forall x \in X. \tag{1.4}$$

PROOF. Consider the convex problem

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

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

$$\partial h(x^*) + N_X(x^*) \ni 0^n \tag{1.6}$$

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

$$\partial h(x) = F(x^*) + \partial u(x), \qquad x \in X; \tag{1.7}$$

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

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

$$F(x^*)^{\mathrm{T}}(x - x^*) + u'(x^*; x - x^*) \geq 0, \qquad \forall x \in X. \tag{1.8}$$

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). □

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

$$X = \prod_{i \in C} X_i, \qquad X_i \subseteq \Re^{n_i}, \qquad \sum_{i \in C} n_i = n, \tag{1.9}$$

for some finite index set 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 $\Re^n$ defined by (1.9), that is, u is of the form

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

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

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

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 $X_i \subseteq \Re^{n_i}$, $\sum_{i=1}^N n_i = n$. We define a *penalty function* $f_i : X \mapsto \Re$ 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 $i \in \{1, \ldots, N\}$, satisfies

$$f_i(x^*_{\neq i}, x^*_i) + u_i(x^*_i) = \operatorname*{minimum}_{x_i \in X_i} \{f_i(x^*_{\neq i}, x_i) + u_i(x_i)\}, \tag{1.10}$$

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

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]. □

REMARK 1.7 (Non-unique representation of GVIP(F, u, X)). 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), $\partial[u + \delta_X](x) = \partial u(x) + N_X(x)$, $x \in \mathrm{dom}\, u \cap X$, 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 GVIP(F, u, X) as the *generalized equation*

[GE(F, u)]

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

(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 GE(F, u) defines a proper subset of GVIP(F, u, X) whenever $\mathrm{dom}\, u = \Re^n$ and $X \neq \Re^n$; 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 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 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 GVIP(F, u, X), the algorithms for solving 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 GVIP(F, u, X) to obtain new and more general results. In particular, we will consider representations based on projecting GVIP(F, u, X) onto different solution spaces and representations involving the introduction of constraint multipliers. □

If $u \equiv 0$ in GVIP(F, u, X) (note that Assumption 1.1 then is satisfied trivially), which is equivalent to assuming that in GE(F, u), $u \equiv \delta_X$ for some nonempty, closed and convex set $X \subseteq \Re^n$, we obtain the *variational inequality problem* of finding $x^* \in \Re^n$ such that

[VIP(F, X)]

$$F(x^*) + N_X(x^*) \ni 0^n, \tag{1.12}$$

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

$$F(x^*)^{\mathrm{T}}(x - x^*) \geq 0, \qquad \forall x \in X. \tag{1.13}$$

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 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

$$[(1/\gamma)x^* + F(x^*) - (1/\gamma)x^*]^{\mathrm{T}}(x - x^*) \geq 0, \qquad \forall x \in X$$
$$\Longleftrightarrow$$
$$[(x^* - \gamma F(x^*)) - x^*]^{\mathrm{T}}(x - x^*) \leq 0, \qquad \forall x \in X;$$

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 $Q \in \Re^{n \times n}$, a similar technique shows that a solution x^* to VIP(F, X) is characterized by the inequality

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

which is equivalent to x^* being the projection of the vector $x^* - \gamma Q^{-1}F(x^*)$ onto X according to the vector norm $\|z\|_Q := \sqrt{z^{\mathrm{T}}Qz}$ 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 VIP(F, X). □

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

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

$$x_{kr} \geq 0, \qquad \forall k \in \mathcal{C}, \tag{1.14b}$$

or, compactly,

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

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

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

Assume further that each route $r \in \mathcal{R}_k$, $k \in \mathcal{C}$, is associated with a route cost (or, travel time) function $F_{kr} : \Re_+^{|\mathcal{R}|} \mapsto \Re_{++}$, 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):

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

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

These conditions for a feasible route flow (the *Wardrop* [War52] *equilibrium conditions*) are equivalent to 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, FlH95]. □

EXAMPLE 1.10 (Saddle point problem). Let $V \subseteq \Re^n$ and $W \subseteq \Re^m$ be closed convex sets, and $\Pi : V \times W \mapsto \Re$ 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])

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

$$\Pi(v^*, w) \leq \Pi(v^*, w^*) \leq \Pi(v, w^*), \qquad \forall (v, w) \in V \times W. \tag{1.16}$$

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

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

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

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

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