Nonlinear Programming and Variational Inequality Problems—A Unified Approach

Michael Patriksson

Department of Mathematics
Chalmers University of Technology
S-412 96 Gothenburg
Sweden

Revised August 18, 1998

Contents

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common framework, and establishes the convergence both of new versions of existing algorithms and of methods previously unpublished.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
$\\|\\|\cdot\\|\\|$	Operator norm induced by the Euclidean vector norm
M_f	Lipschitz continuity constant
m_f	Strong convexity (or monotonicity) constant

Index