2005

Review of: Newton Methods for Nonlinear Problems: Affine Invariance and Adaptive Algorithms, by P. Deuflhard

Tjalling Ypma
Western Washington University, tjalling.ypma@wwu.edu

written as the phase plane system

$$(4)\qquad \frac{df}{d\xi} = -\frac{g}{a'(f)},$$

$$\frac{dg}{d\xi} = -\frac{(\sigma + b'(f))g}{a'(f)} + c(f).$$

Then the corresponding first-order ODE for integral curves is

$$(5)\qquad \frac{dg}{df} = \sigma + b'(f) - \frac{c(f)a'(f)}{g}.$$

Finally, integration yields the integral equation for $g(f)$ that lies at the heart of this book,

$$(6)\quad g(f) = \sigma f + b(f) - \int_0^f \frac{c(r)a'(r)}{g(r)}\, dr.$$

The authors provide the necessary background for the analysis of integral equations of this form. The benefits of this interesting line of attack are well illustrated through the many examples of (1) that would be problematic studied otherwise. What follows is an extensive presentation of proofs of results on existence, uniqueness, boundedness, solutions with bounded support, and dependence on the wavespeed in terms of properties of $a(u)$, $b(u)$, $c(u)$. The book has a reasonable balance between proofs and more specific applications and examples. It also provides physical and historical contexts for the problems with extensive bibliographic notes. The most global results for (1) are presented in the earlier chapters; later on, things are subdivided into more specific results for the cases of convection-diffusion and reaction-diffusion problems. The book concludes with a chapter detailing special cases of (1) with known closed-form explicit solutions.

I found this book to be a very nice addition to Birkhäuser's Progress in Nonlinear Differential Equations series, and I believe it will be a very useful resource for anyone studying traveling waves in nonlinear parabolic equations. Its focus is limited, but within its scope, its presentation is encyclopedic. It has no results on stability, rates of convergence and long time asymptotic behavior, or problems in multiple dimensions, but the results on traveling waves it does present can be used as the building blocks for further research in all of these areas. Hence I see this book as a valuable reference for most problems relating to PDEs such as (1).

REFERENCES

[1] B. H. Gilding and R. Kersner, *The characterization of reaction-convection-diffusion processes by travelling waves*, J. Differential Equations, 124 (1996), pp. 27–79.

[2] A. A. Samarskii, V. A. Galaktionov, S. P. Kurdyumov, and A. P. Mikhailov, *Blow-up in Quasilinear Parabolic Equations*, de Gruyter Exp. Math. 19, Walter de Gruyter, Berlin, 1995.

[3] A. I. Volpert, V. A. Volpert, and V. A. Volpert, *Traveling Wave Solutions of Parabolic Systems*, Transl. Math. Monogr. 140, AMS, Providence, RI, 1994.

Thomas P. Witelski
Duke University

Newton Methods for Nonlinear Problems: Affine Invariance and Adaptive Algorithms. *By P. Deuflhard.* Springer-Verlag, Berlin, 2004. $99.00. xii+424 pp., hardcover. ISBN 3-540-21099-7.

In the context of solving nonlinear equations, the term "affine invariance" was introduced to describe the fact that when a function $F : R^n \to R^n$ is transformed to $G = AF$, where A is an invertible matrix, then the equation $F(x) = 0$ has the same solutions as $G(x) = 0$, and the Newton iterates $x_{k+1} = x_k - F'(x_k)^{-1}F(x_k)$ remain unchanged when F is replaced by G. The idea was that this property of Newton's method should be reflected in its convergence analysis and practical implementation, not only on aesthetic grounds but also because the resulting algorithms would likely be less sensitive to scaling, conditioning, and other numerical issues.

This simple idea has been expanded by Peter Deuflhard and his coworkers to embrace at least four different forms of invariance, pertaining to three different contexts. The invariance mentioned above is now termed affine covariance, while methods and analyses for the same problem class that are independent of the choice of B under the transformation $x = By$ and $G(y) = F(By)$ are termed affine contravariant. For the class of unconstrained optimization problems $\min f(x)$, setting $x = By$ and $g(y) = f(By)$ results in a function g whose Hessian matrix $g''(y)$ is related to the Hessian matrix $f''(x)$ of f through the relation $g''(y) = B^T f''(x)B$; hence invariance under this transformation is termed affine conjugacy. Finally, for the dynamical system $x' = F(x)$, setting $x = By$ gives $y' = G(y) = B^{-1}F(By)$, for which $G'(y) = B^{-1}F'(x)B$; the corresponding invariance is thus termed affine similarity.

This book focuses on providing convergence analyses of various standard numerical methods for these problem classes, with both the formulation of the methods and the analysis couched in the appropriate invariant format. This analysis is then used to develop invariant computational estimates of quantities which can be used to monitor the convergence behavior of the corresponding algorithm. A typical example is the analysis of Newton's method itself. First an affine covariant form of the classic Newton–Mysovskikh theorem is proved, characterizing the convergence through an invariant majorizing sequence. Then properties of some invariant computable estimates of terms of this majorizing sequence are derived. These quantities are then incorporated as convergence monitors and termination criteria in an affine covariant implementation of Newton's method.

The book is broad in scope, covering a wide range of topics generally associated with the solution of nonlinear equations. This includes the special situations arising in the context of stiff differential equations and boundary value problems. For example, the extensive portion of the book that centers on the use of Newton's method and its variants for solving nonlinear equations includes an analysis of inexact Newton methods and quasi-Newton methods, while damped Newton methods, trust region algorithms, and continuation methods are among the more widely convergent classes of methods analyzed. The analysis extends to the invariant computational control of the iterative linear equation solvers (such as GMRES and PCG) used within the context of the nonlinear equation solver. The word "adaptive" in the title of the book refers to the invariant adjustment of steplengths and grids in the context of solving differential equations. Arguably the scope of the book is too wide, with the sheer volume of material demanding sometimes very terse presentations and cross-referencing in the quest for brevity, at the expense of clarity and readability.

The invariant (re)formulations of various standard numerical methods, and the corresponding (re)statements of the related convergence theorems, are an attractive feature of this book, not least because they provide alternative perspectives on an otherwise classic landscape. These features also sometimes permit derivations of standard results that are somewhat more general and transparent than usual. The analysis of invariant computable convergence quantities is also of considerable value, and many of the results are quite readily implementable. An interesting feature of the analysis is what the author calls "bit counting lemmas," in which properties of these computable quantities are derived on the basis of the relative accuracy of a priori or a posteriori estimates of other underlying theoretical quantities. The term "bit counting" refers to the fact that these analyses cover even the case when there is only one correct bit in the estimate.

This book is certainly not an easy read. Much of it is inevitably devoted to the often-tedious pursuit of algebraic inequalities. While the algebra is sometimes enlivened by some deft trickery, there is cause to be grateful that the painful manipulation has been done by someone else. A considerable amount of detective work, and occasionally frustration, is involved in piecing together some of the arguments, since the notation is not always clearly defined and the many backward references are sometimes obscure. Occasionally awkward phrasing makes some statements hard to interpret, and the book is not error-free.

To a large extent this book is an organized compendium and synthesis of the work of the author and his students and coworkers over the last 30 years. The references go back to the author's doctoral dissertation dating from 1972, and a significant percentage of the 200-odd references relate to publications of the author and co-workers. A number of these publications are technical reports or doctoral theses, which this book thus makes more accessible and places within a broader context. Unfortunately this book will not free one from the need to consult these references: as the complexity of the material increases, so does the sketchiness of the discussion, and a good deal of the last part of the book is only partially intelligible without access to the original publications.

A considerable amount of background knowledge is required to appreciate much of the content of this book. Such a background would certainly have to include at least a passing acquaintance with most of the content of the introductory texts by Kelley [3, 4, 5], including some knowledge of iterative linear equation solvers based on Krylov subspaces. A stronger background would include the classic text of Dennis and Schnabel [1], Greenbaum's introduction to Krylov subspace methods [2], and some elements of the standard reference by Ortega and Rheinboldt [6]. To the extent that this book can be regarded as a text it is restricted to an advanced audience, or at least an audience with an experienced guide. The relatively few exercises at the end of each chapter are generally not for the faint-hearted.

The book is thus primarily a research monograph, but it is also useful as a reference from which one can scavenge good ideas and interesting theoretical approaches, as well as providing readily implementable techniques with firm theoretical foundations to replace the ad hoc devices so often incorporated into numerical algorithms. It is a pity that these virtues are not supported by a clearer exposition.

REFERENCES

[1] J. E. Dennis, Jr., and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, Englewood Cliffs, NJ, 1983; reprinted as Classics Appl. Math. 16, SIAM, Philadelphia, 1996.

[2] A. Greenbaum, *Iterative Methods for Solving Linear Systems*, SIAM, Philadelphia, 1997.

[3] C. T. Kelley, *Iterative Methods for Linear and Nonlinear Equations*, SIAM, Philadelphia, 1995.

[4] C. T. Kelley, *Iterative Methods for Optimization*, SIAM, Philadelphia, 1999.

[5] C. T. Kelley, *Solving Nonlinear Equations with Newton's Method*, SIAM, Philadelphia, 2003.

[6] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970; reprinted as Classics Appl. Math. 30, SIAM, Philadelphia, 2000.

Tjalling Ypma
Western Washington University