Global convergence of iterative solution methods for nonlinear problems

Nonlinear problems $\mathcal{F}(\mathbf{u}) = 0$, where $\mathbf{u} \in \mathbb{R}^n$ arise in many contexts such as in nonlinear (i.e. non-quadratic) optimization and in numerical discretization of nonlinear partial differential equations (PDEs). We are here especially concerned with the latter.

Depending on the differentiability properties of the nonlinear mapping $\mathcal{F}$, there exist several approaches, some with much limited applicability, i.e., having too demanding convergence criteria and others with more general applicability, where convergence can be shown for quite general functions $\mathcal{F}$. To the former belongs the Picard iteration method and the nonlinear Gauss-Seidel method. Such methods require that the nonlinearity of the function is sufficiently weak.

When $\mathcal{F}$ is differentiable, one can apply the Newton method, which can permit strong nonlinearities. Without damping the steplength, this method may also have restricted convergence in the sense that the initial approximation must already be sufficiently close to the nonlinear equation. On the other hand, with proper control of the step lengths, the resulting method, called *damped* Newton iteration method, can converge essentially globally, i.e., for any choice of the initial approximation. Furthermore, it is not required that the Newton corrections are computed exactly at each step, which implies that an *inner* iteration method can be used to solve the arising linear systems. The accuracy with which this takes place can be controlled using some proper stopping criteria, and is normally weak at the initial steps and increasing when one approaches the exact solution to the nonlinear equation. Even though sometimes the convergence in the initial phase can be slow, each step is hence cheaper. One can also couple this process with adaptive mesh refinements.

Using two or more meshes, there exists also alternative, more general approaches to prove convergence. In the two-level Newton method one solves a coarse mesh nonlinear problem (on a mesh Ω_H) exactly, and then interpolates the coarse solution to a fine mesh $\Omega)h$, where the solution is approximated using a $\underline{\text{single}}$ Newton step with the interpolated solution as initial approximation. Under certain conditions, it can be shown that the resulting fine approximation has the same order of accuracy as the exact fine mesh solution. Furthermore, due to the quadratic approximation of a Newton step, there holds a relation $H = O(h^\nu)$, where typically $\nu \geq 2$, sometimes even $\nu = 4$. This implies that the coarse mesh can be chosen very coarse as related to the fine mesh. The rationale behind this method is that the nonlinear iterations on the coarse mesh are computationally much cheaper than they would be on the fine mesh and often the cost is negligible with the cost of the single linear solver on the fine mesh.

However, should need arise, the method can be extended to a multilevel method, where each each arising coarse mesh problem is itself solved as a two-grid problem. Much effort has in fact been devoted to the so-called *nonlinear multigrid methods*. Here, there exist basically two approaches. In the Newton multigrid method, the on each mesh arising linear systems are solved, for instance using a linear multigrid method. It turns out that direct solvers are often very competitive with iterative solvers even on quite fine meshes. This holds in particular for ill-conditioned problems.

In the so-called *full approximation scheme* (FAS) one uses multigrid approaches instead of the Newton method, i.e., fine mesh smoothing followed by a coarse mesh correction for the nonlinear equation on each mesh. It has been shown that for mildly nonlinear problems, nonlinear

MG can converge nearly as fast as linear MG, i.e., applied for a corresponding linear problem.

Another approach used to solve nonlinear problems is the nonlinear conjugate gradient method.

References

[1] Ortega JM, Rheinboldt WC. *Iterative solution of nonlinear equations in several variables*, Academic Press, 1972.

[2] Iserles A. *A first course in the Numerical Analysis of Differential Equations*, Cambridge texts in Applied Mathematics, 1997.

[3] Glowinski R. *Numerical methods for nonlinear variational problems*, New York, Berlin Springer Vlg, 1984.

[4] Axelsson O. On mesh independence and Newton-type methods, *Applied Mathematics*, 1993; **38**:249-265.

[5] Axelsson O, Chronopoulos AT. On nonlinear generalized conjugate gradient methods, *Numerische Mathematik*, 1994; **69**:1-15.

[6] Axelsson O, Layton W. A two-level method for the discretization of nonlinear boundary value problems. *SIAM Journal on Numerical Analysis* 1996; **33**:2359-2374.

[7] Axelsson O, Kaporin I. On the solution of nonlinear equations for nondifferentiable mappings. Hackbusch, Wolfgang and Wittum G. (eds), Fast solvers for flow problems, NNFM 49, Vieweg Verlag, Braunschweig/Wiesbaden, 1995. (Proceedings of the tenth GAMM-Seminar Kiel, Germany, January 14–16, 1994.)

[8] Axelsson O, Neytcheva M. *Supercomputers and Numerical Linear Algebra*. Lecture Notes, 1996.

[9] Baur W, Strassen V. The complexity of Partial Derivatives, *Theoretical Computer Science*, 1983; **22**:317–330.

[10] M. Berz, C. Bischof and G. Corliss and A. Griewank (eds), *Computational Differentiation: Techniques, Application and Tools*, SIAM, Philadelphia, Penn., 1996.

[11] Brown PN, Saad Y. hybrid Krylov methods for nonlinear systems of equations, *SIAM Journal on Scientific and statistical computing*, 1990; **11**:450-481.

[12] Chandra R, Eisenstat SC, Schultz MH. The modified conjugate residual method for partial differential equations, in Advances in Computer Methods for Partial Differential Equations II, ed. R. Vichnevetsky, IMACS, 1977:13–19.

[13] Chen W. A Newton method without evaluation of nonlinear function values, manuscript, 1995.

[14] Chronopoulos AT. Nonlinear CG-like iterative methods, *Journal of Computational and Applied Mathematics*, 1992; **40**:73-89.

[15] Daniel JW. The conjugate gradient method for linear and nonlinear operator equations, *SIAM journal on Numerical Analysis*, 1967; 4:10-26.

[16] Dembo RS, Eisenstat SC, Steihaug T. inexact Newton methods, *SIAM journal on Numerical Analysis*, 1982; 19:400-408.

[17] Dembo RS, Steihaug T. Truncated-Newton algorithms for large-scale unconstrained optimization, *Mathematical Programming*, 1983; **26**:190-212.

[18] Dennis JE, More JJ. Quasi-Newton methods, motivation and theory, *SIAM Review*, 1977; **19**:46-89

[19] Faragó I, Karátson J. *Numerical solution of nonlinear elliptic problems via preconditioning operators. Theory and applications*, Huntington, NY: Nova Science Publishers, 2002.

[20] Fischer H. Automatic differentiation of characterizing sequences, *JCAM*, 28 (1989), 181–185.

[21] Fischer H. Automatic differentiation. Parallel computations of function, gradient and Hessian matrix, *Parallel Computing*, 1990; 13:101–110.

[22] Fischer H. The key idea and an illustrative example. Fischer, Herbert (ed.) et al., Applied mathematics and parallel computing. Festschrift for Klaus Ritter. Heidelberg: Physica. 121-139 (1996).

[23] Fokkema DR, Sleijpen LG, van der Vorst HA. Accelerated inexact Newton schemes for large systems of nonlinear equations, *SIAM Journal on Scientific Computing*, 1998; **19**:657-674.

[24] Gilbert JC, Nocedal J. Global Convergence Properties of Conjugate Gradient Methods for Optimization. *SIAM J. Opt*, 1992; 2:21-42.

[25] Glowinski R, Keller HB, Reinhart L. Continuation conjugate gradient methods for the least squares solution of nonlinear boundary value problems, *SIAM Journal on Scientific and Statistical Computing*, 1985; **4**:793-832.

[26] Griewank A. On automatic differentiation. In Iri M, Tanabe K. (eds) *Mathematical Programming*, kluwer Academic Publishers, Dordrecht, 1989.

[27] Griewank A. Direct calculation Newton steps without accumulating Jacobians, in Coleman TF, Yuying Li (eds), *Large scale numerical optimization*, SIAM 1990, 115-137.

[28] Hackbusch W, Reusken A. On global multigrid convergence for nonlinear problems. *Robust multi-grid methods*, Proc. 4th GAMM-Semin., Kiel/FRG 1988, *Notes on Numerical Fluid Mechanics* 1989; 23:105-113.

[29] Van Emden Henson. Multigrid methods for nonlinear problems: an overview. Lawrence Livermore National Laboratory, Technical Report UCRL-JC-150259.

[30] Iri M, Tsuchiya T. Hoshi M. Automatic Computation of Partial Derivatives and Rounding Error Estimates with Applications to Large-Scale Systems of Nonlinear Equations, *J. Comp. Appl. Math.*, 1988; 24:365–392.

[31] Kaporin I. Conjugate gradient-like nonlinear least squares solvers based on curvilinear searches over Krylov subspaces. Moscow State university, Moscow, 1993. Manuscript.

[32] Kaporin I, Axelsson O. On a class of nonlinear equation solvers based on the residual norm reduction over a sequence of affine subspaces. *SIAM J. Sci. Comput.* 16, No.1, 228-249 (1995).

[33] Linnainmaa S. Taylor expansion of the accumulated rounding error, *BIT*, 16 (1976), 146–160.

[34] Loyce M. Adams, Nazareth JL (editors). *Linear and Nonlinear Conjugate-Gradient Related Methods*, SIAM, August 1996, ISBN 0898713765.

[35] Nash SG. A survey of truncated-Newton methods, *Journal of computational and Applied Mathematics*, 2000; **124**:45-59.

[36] Powell MJD. Nonconvex Minimization Calculations and the Conjugate Gradient Method, *Lecture Notes in Mathematics*, 1984; 1066:122-141.

[37] Shanno DF. Conjugate Gradient Methods with Inexact Searches, *Math. Oper. Res.*, 1978; 3:244-256.

[38] Steihaug T. Damped inexact quasi-Newton methods, Technical report MASC TR 81-3, Department of Mathematical Sciences, Rice University, Houston, May 1984.

[39] Tijsken E, Roose D, Ramon H, De Baerdemaeker J. Automatic differentiation for solving nonlinear partial differential equations: an efficient operator overloading approach, *Numerical Algorithms*, 2002; **30**:259-301.

[40] Wang Deren, Bai Zhongzhi. Parallel nonlinear multisplitting relaxation methods, *Applied Mathematics*, 1995; 10B:251-266.

[41] Zinčenko. Some approximate methods of solving equations with nondifferentiable operators, *Doklady Academii Nauk Ukraini, RSR*, 1963, 156-161, 852-856.