Parallel Algorithms of Newton-Krylov-Schwarz Type

David E. Keyes[1]

Parallel implicit solution methods are increasingly important in aerodynamics and other fields leading to large nonlinear systems with sparse Jacobians. Several trends contribute to their importance. Multidisciplinary analysis and optimization require rapidly achievable low residual solutions, since individual component codes are often iterated and their results differenced for sensitivities. Problems possessing multiple space or time scales motivate implicit algorithms, and arise frequently in locally adaptive contexts and in dynamical contexts such as aero-elasticity. Meanwhile, the demand for resolution and prompt turnaround forces consideration of parallelism, and, for cost effectiveness, the high-latency, low-bandwidth parallelism available from workstation clusters. An ICASE program in Newton-Krylov-Schwarz (NKS) solvers responds to this need, in collaborations with academia, national laboratories (NASA and DOE), and industry (Boeing and UTRC).

An NKS method combines a Newton-Krylov (NK) method, such as nonlinear GMRES, with a Krylov-Schwarz (KS) method, such as additive Schwarz. The linkage is the Krylov method, whose most important characteristic, from a computational point of view, is that information about the underlying Jacobian needs to be accessed only in the form of matrix-vector products in a relatively small number of directions. NK methods are suited for problems in which it is unreasonable to compute or store an accurate full-basis representation of the Jacobian. However, if the Jacobian is ill-conditioned, the Krylov method will require an unacceptably large number of iterations. The system can be transformed through the action of a preconditioner whose inverse action approximates that of the Jacobian, but at smaller cost. It is usually in the choice of preconditioning that the battle for low computational cost and scalable parallelism is won or lost. In KS methods, the preconditioning is introduced on a subdomain-by-subdomain basis, providing well load-balanced data locality for parallel implementations over a wide granularity range. This paradigm is credited to Schwarz, who in 1869 used it to prove existence of solutions for elliptic problems on domains not lending themselves to separable coordinate systems. A two-grid-level form of additive Schwarz (see ref. 2 for a reference chain) provides a mesh-independent and granularity-independent convergence rate in elliptically dominated problems, including nonsymmetric and indefinite problems.

An encouraging application of NK technology has arisen in collaborations with W. K. Anderson of the Computational Aerodynamics Branch at NASA Langley. Anderson's state-of-the-art unstructured-grid Euler and Navier-Stokes codes employ a solver common in form to that of many implicit codes. Their objective is to solve the steady-state conservation equations $f(u) = 0$ in the pseudo-transient form $\frac{\partial u}{\partial t} + f(u) = 0$, where the time derivative is approximated by backwards differencing, with a time step that ideally approaches infinity, leaving the steady form of the equation. A left-hand side matrix (not a true Jacobian) is created, in whose construction computational shortcuts are employed, and which may be stabilized by a degree of first-order upwinding that would not be acceptable in the discretization of the residual, itself. We denote this generic distinction in the update equation (2) by subscripting the residual "high" and the left-hand side matrix "low":

$$J_{low}\delta u = -f_{high} \qquad (2)$$

Often, J_{low} is based on a low-accuracy residual:

$$J_{low} = \frac{D}{\delta t} + \frac{\partial f_{low}}{\partial u},$$

where D is a scaling matrix. Inconsistency between the left- and right-hand sides prevents the use of large time steps, δt. Using the built-in capability to solve systems with J_{low} as the preconditioning, we replace (2) with

$$(J_{low})^{-1} J_{high}\delta u = -(J_{low})^{-1} f_{high}, \qquad (3)$$

in which the action of J_{high} on a vector is obtained through directional differencing, namely

$$J_{high}(u^l)v \approx \frac{1}{h}\left[f_{high}(u^l + hv) - f_{high}(u^l)\right], \quad (4)$$

[1]David Keyes received his Ph.D. in Applied Mathematics from Harvard in 1984. He is currently a Senior Research Associate at ICASE, an Associate Professor of Computer Science at Old Dominion University, and the Director of the new Virginia/ICASE/LaRC Program in High Performance Computing and Communication.

where h is a small parameter. Since the operators on both sides of (3) are consistently based on high-order discretizations; time steps can be built up to large values, recovering Newton's method in the limit.

The most effective solver to date for systems based on J_{low} is multigrid, typically V-cycle MG on a family of non-nested coarser grids. However, the construction of such a family is extremely time-consuming in two dimensions and much more of an art than a science in three dimensional problems with complex, multiply connected geometry, such as a multiple-element airfoil. NKS methods may permit solution of the fine-grid discretization alone in comparable time, as illustrated for a two-dimensional Navier-Stokes problem on an unstructured grid [ref. 1] in Fig. 1. The baseline method is a single-grid point relaxation process. The second method is MG with four non-nested coarse grids in a FAS V-cycle. The third matrix-free NK method uses ILU as a preconditioner and restarted GMRES as the accelerator.

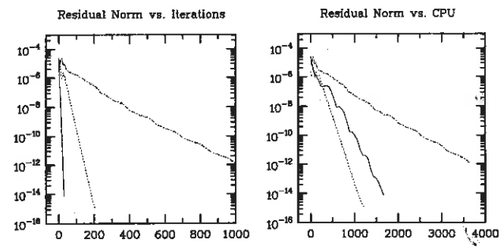

Figure 1: Norm of steady-state residual vs. iterations (left) and vs. Cray CPU time (right) for a single grid Gauss-Seidel scheme (solid), a 4-V-cycle multigrid scheme (dashed), and an NK scheme (dotted).

The left plot of residual norm versus "outer iterations," shows that MG and NK can achieve close to machine-zero residual reductions in vastly fewer iterations than can single-grid relaxation. Since each NK iteration involves a set of Krylov subiterations, it trails multigrid by a factor of about 1.5 in execution time, as shown in the right plot of residual norm versus sequential time. A small sequential time disadvantage is tolerable since the NK method does all of its computation without user generation of coarse unstructured grids and lends itself to easily parallelizable preconditioners. This work has been extended to three-dimensional problems in conjunction with E. J. Nielsen and R. W. Walters of Virginia Tech.

Related projects include a two-level Schwarz preconditioned version of Boeing's TRANAIR [ref. 3], a low Mach number combustion simulation based on NKS (with D. A. Knoll and P. R. McHugh of INEL), an unstructured Euler code parallelized on the Paragon (with Venkatakrishnan of ICASE), an Ethernet Sparcstation cluster implementation of a KS solver for a structured-grid Euler Jacobians (with M. D. Tidriri of ICASE, W. D. Gropp of Argonne, and J. S. Mounts (deceased) of United Technologies), and an NKS solver for a model full potential equation parallelized on the SP2 (with X.-C. Cai of UC-Boulder and Gropp).

A variety of CFD applications are (or have inner) nonlinear elliptically-dominated problems amenable to solution by NKS algorithms, which are characterized by relatively low storage requirements (for an implicit method) and locally concentrated data dependencies. The main disadvantage of NKS algorithms is the large number of parameters that require tuning. Each component (Newton, Krylov, and Schwarz) has its own set of parameters. Parametric tuning is important to performance, but robust choices are not difficult. Nonsymmetry, nonlinearity, and multicomponent structure hinder theoretical development, but experimental development continues in a variety of settings.

References

D.E. Keyes, "Aerodynamic Applications of Newton-Krylov-Schwarz Solvers, Proceedings of the 14th International Conference on Numerical Methods in Fluid Dynamics (R. Narasimha, et al., eds.)," Springer, New York, 1995 (to appear).

D.E. Keyes and J. Xhu, eds., " Domain Decomposition Methods in Scientific and Engineering Computing," Amer. Math. Soc., Providence, RI, 1995.

R.G. Melvin, D.P. Young, D.E. Keyes, C.C. Ashcraft, M.B. Bieterman, C.L. Hilmes, W.P. Huffman and F.T. Johnson, "A Two-Level Iterative Method Applied to Aerodynamic Sensitivity Calculations," BCSTECH-94-047, Boeing Computer Services, Seattle, Dec. 1994.