RESIDUAL SMOOTHING FOR KRYLOV SPACE SOLVERS: DOES IT HELP AT ALL?

Martin H. Gutknecht
Seminar for Applied Mathematics, ETH Zurich
ETH-Zentrum HG, CH-8092 Zürich, Switzerland
mhg@sam.math.ethz.ch, http://www.sam.math.ethz.ch/~mhg

Abstract

Some Lanczos-type solvers for non-Hermitian linear systems, such as BiCG and (Bi)CGS, are prone to an erratic convergence behavior: error norms and residual norms may increase and then again decrease by several orders of magnitude within just a few iteration steps.

By combining these methods with smoothing processes that replace the iterates by smoothed ones, one can make the residual and error norm plots look much smoother. Even monotone convergence can be enforced. But does this really mean that the solution is found sooner? Or that the error can be made smaller? Does the answer depend on how the smoothing process is implemented?

RESIDUAL SMOOTHING FOR KRYLOV SPACE SOLVERS: DOES IT HELP AT ALL?

Martin H. Gutknecht
Seminar for Applied Mathematics, ETH Zurich
ETH-Zentrum HG, CH-8092 Zürich, Switzerland
`mhg@sam.math.ethz.ch`, `http://www.sam.math.ethz.ch/~mhg`

Abstract

The erratic convergence behavior of basic Lanczos-type solvers, such as the BiCG and BiCGS methods, often gives rise to criticism. Since plots of the residual norm are the most often used measure when algorithms are compared, a method sells well if its residual norm converges quickly and smoothly. Thus, it is not surprising that there is an interest in smoothing processes that modify the BiCG or BiCGS iterates so that the residual norm plot becomes smoother and, hopefully, faster decreasing. One might also hope that the ultimate accuracy that can be attained with a method is improved by smoothing, since in the smoothed sequence far outlying iterates and residuals are avoided.

However, these hopes are questionable, and the whole attitude of promoting smoothing is a bit dubious. In fact, what really counts in practice is that a method finds the solution (up to a certain error) as quickly as possible, and since the error cannot be checked, the residual has to be monitored instead. The smoothness of the convergence does not matter from that point of view.

It can be seen that smoothing processes can speed up the convergence slightly, but only in the case when the primary process does not converge well. On the other hand, they hide some useful information: a peak in the residual norm plot indicates a temporary stagnation of convergence, while in the smoothed residual plot we cannot distinguish temporary from permanent stagnation. Moreover, there is normally virtually no improvement in the ultimate accuracy that can be attained.

Most of our considerations concern the smoothing process of Schönauer [6] and Weiss [7, 8], and its enhanced version due to Zhou and Walker [9]. This process in known to transform in theory Conjugate Gradient (CG) iterates into Conjugate Residual (CR) iterates, and iterates of the Full Orthogonalization Method (FOM or Arnoldi) into Generalized Minimum Residual (GMRes) iterates. The relevant theoretical relationship between the convergence behavior of CG and CR, as well as FOM and GMRes is by now well understood through the work of Brown, Cullum, Freund, Greenbaum, Gutknecht, Weiss, and others; see, in particular, Cullum and Greenbaum [1], and Gutknecht [2]. This relationship approximately carries over to BiCG and QMR, and makes us understand why peaks in the residual norm plot of BiCG are matched by plateaux in the one of QMR.

The new results in this talk on the limitations of smoothing processes (which also apply to MinRes and, in a certain sense, to QMR) and on the influence of roundoff are joint work with Miroslav Rozložník [5, 3], see also [4].

References

[1] J. Cullum and A. Greenbaum, Relations between Galerkin and norm-minimizing iterative methods for solving linear systems, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 223–247. Received Apr. 2, 1993.

[2] M. H. Gutknecht, Lanczos-type solvers for nonsymmetric linear systems of equations, Acta Numerica, 6 (1997), pp. 271–397.

[3] M. H. Gutknecht and M. Rozložník, By how much can residual minimization accelerate the convergence of orthogonal residual methods?, Numerical Algorithms. Submitted.

[4] ——, A framework for generalized conjugate gradient methods — with special emphasis on contributions by Rüdiger Weiss, Appl. Numer. Math. To appear.

[5] ——, Residual smoothing techniques: do they improve the limiting accuracy of iterative solvers?, BIT, (2001), pp. ...–...

[6] W. Schönauer, *Scientific Computing on Vector Computers*, Elsevier, Amsterdam, 1987.

[7] R. Weiss, *Convergence behavior of generalized conjugate gradient methods*, PhD thesis, University of Karlsruhe, 1990.

[8] ——, Properties of generalized conjugate gradient methods, J. Numer. Linear Algebra Appl., 1 (1994), pp. 45–63. Received June 1991.

[9] L. Zhou and H. F. Walker, Residual smoothing techniques for iterative methods, SIAM J. Sci. Comput., 15 (1994), pp. 297–312.