Final Report

DOE Grant DE-FG02-05ER25672

Schwarz Preconditioners for Krylov Methods: Theory and Practice

PI: Daniel B. Szyld
Temple University

Executive Summary

Several numerical methods were produced and analyzed. The main thrust of the work relates to inexact Krylov subspace methods for the solution of linear systems of equations arising from the discretization of partial differential equations. These are iterative methods, i.e., where an approximation is obtained and at each step. Usually, a matrix-vector product is needed at each iteration. In the inexact methods, this product (or the application of a preconditioner) can be done inexactly. Schwarz methods, based on domain decompositions, are excellent preconditioners for thise systems. We contributed towards their understanding from an algebraic point of view, developed new ones, and studied their performance in the inexact setting. We also worked on combinatorial problems to help define the algebraic partition of the domains, with the needed overlap, as well as PDE-constraint optimization using the above-mentioned inexact Krylov subspace methods.

Work supported by this grant, and earlier work related to the DOE mission

During the six-and-a-half years of support under this grant, we have been very productive and have written numerous papers which directly or indirectly contribute to the mission of DOE. We obtain good results in all areas of the proposed work (inexact Krylov subspace methods, Schwarz preconditioners, combinatorial graph partitions with overlap, applications in control).

Both before and during this period of DOE funding, we have worked in the area of Inexact and Flexible state-of-the-art iterative methods for the solution of linear systems of equations [50], [41], [42], [45], [43], [44], [39]; see also [46]. These methods and their analyses are fundamental for high-performance computing in modern architectures, including those with multicore nodes. Succinctly, these methods consider inexact and variable matrix-vector products or inexact preconditioners which reduce computational costs considerably without any significant penalty in convergence rates. In fact, the theory we have developed will be ripe for further application in upcoming machines, expected to reach the exascale regime [36]. We believe that the only methods that will be viable in this setting are those using inexact matrix-vector products and preconditioning.

More recently, we concluded a long-term project of using inexact methods for control problems constrained by parabolic PDEs [9], [8].

Among the DOE researchers who are currently applying our theory on inexact Krylov subspace methods is Eric Cyr at Sandia National Laboratory, who works on multi-physics preconditioners.

Continuing on the theme of inexact methods, we mention our work (jointly with our postdoc funded by this grant, Fei Xue) [51] on such methods for linear generalized eigenvalue problems (of the form $Av - \lambda Bv = 0$). It is worth repeating that these are the only methods which will survive in the next generation of computer architectures. In [51], we demonstrate how to choose the tolerances for the inexact solution of the the inner linear systems, e.g., in inexact Rayleigh Quotient Iteration, in order to maintain fast convergence toward the desired (interior) eigenpair. We were able to show for this method, as well as for single-vector Jacobi-Davidson that, if the inner tolerances are appropriately chosen, the same local convergence rate can be attained for the inexact method as for the exact counterpart. We should mention that Barry Lee of the Pacific Northwest Laboratory, has expressed interest in implementing for an High Performance Computing environment a related work of Xue (and Elman) on inexact implicitly restarted Arnoldi for the same linear eigenvalue problems [55]. Similarly, we show in [53] how inexactness influences the convergence of Newton-type methods for general *nonlinear* eigenvalue problems, i.e., of fhe form $T(\lambda)v = 0$. More recently, we further analyzed several properties of invariant pairs of nonlinear algebraic eigenvalue problems [52].

Another area of recent work is more combinatorial in nature. It deals with graph partitioning, and growing these partitions to obtain overlapping subgraphs, i.e., subgraphs whose union is the whole graph, but where the subgraphs are not disjoint; they share some nodes, called the "overlap." Our motivation for this work comes from the need for permutations and partitions (with overlap) of sparse matrices, to build preconditioners. In [19] we present a linear-time algorithm (sublinear in the parallel case) to obtain an effective (overlapping) decomposition; so that the preconditioners thus produced work well (and better than many alternatives). The inspiration for this work came from Schwarz preconditioners for discretized PDE problems, where the overlap between the subdomains plays a major role in the convergence rates, which are mesh independent [49], [54]. We point out that our proposal for growing a partition is more effective than the current default in the software PETSc [2] from Argonne; and Barry Smith has already mentioned that this default is scheduled to be revised to incorporate our algorithm.

We mention that this overlap is needed in the Krylov-type algorithms being proposed for minimal communication between cores, which will be important in future high performance architectures; see, e.g., [3], [26], as well as for certain network problems [25].

Schwarz preconditioners are based on domain decomposition methods, and are widely used at many DOE labs. Scientists from five DOE labs presented work at the Twentieth International Conference on Domain Decomposition Methods in San Diego in 2011, and the meeting was in fact co-sponsored by

Sandia and Livermore. We have made several contributions to the algebraic understanding of these preconditioners, and in some instances proved convergence results, or proposed new alternatives, which could not have been done without this algebraic approach [21], [22], [4], [23], [20], [34], [5], [32], [31], [6], [39], [35], [40], [1], [27], [29], [28], [30], [12], [24]. As one example, let us mention that the default Schwarz preconditioner in PETSc (from Argonne) is Restricted Additive Schwarz [7], and its convergence properties are understood thanks to our work [23].

Other recent contributions in the area of modern iterative methods for the solution of linear systems of equations include works on several types of Krylov subspace methods [44], [45], [43], [46], [11], [47], [10], [48], [38], and we know that these have influenced research and production codes at several DOE labs.

More recently, during summer 2011, our student Kirk Soodhalter (who graduated in Spring 2012) was an intern at Sandia, working with Michael Parks on extending recycling strategies for Krylov subspace methods to a block setting, and using this for the solution of linear systems arising from the use of Newton iterations for fluid Density Function Theory problems. This work is ongoing and will include an implementation of block GCRO in Trilinos [37].

Let me mention here that in terms of training future scientists for the needs of DOE, in addition to Kirk Soodhalter, we have two students working with us, and both of them are spending this summer (2013) as interns in DOE labs. Stephen Shank, who is working on methods for matrix equations, is at Sandia working with Ray Tuminaro; and Scott Ladenheim, who is working on multipreconditioners and on saddle-point problems, is at Livermore working with Panayot Vassilevsky.

During this grant, we also produced important results in applied linear algebra. We produced a whole new theory for matrices which are not necessarily nonnegative or M-matrices, but which possess a "Perron-Frobenius property." Namely, the spectral radius is an eigenvalue, it is dominant, and the corresponding eigenvector can be taken to be nonnegative (or positive) [14], [16], [15], [17], [13], [18]. These properties have been used for decades to show important results in numerical and applied linear algebra. For example, they were used to show convergence of certain iterative methods for the solution of discretized elliptic PDEs, or for the existence of the incomplete LU factorization [33]. The new theory now can be used to show the same results in more general settings, for example, with different boundary conditions (which make the underlying matrices lose the M-matrix property).

More recently, we used the Perron-Frobenius property on some matrices originating in a tensor representation of a diffusion process in the context of semi-supervised machine learning [56]. This was used to show convergence of the method, and numerical experiments demonstrated its superiority over the state-of-the-art algorithms.

References

[1] Josep Arnal, Violeta Migallón, José Penadés, and Daniel B. Szyld. Newton additive and multiplicative Schwarz iterative methods. *IMA Journal of Numerical Analysis*, 28:143–161, 2008.

[2] Satish Balay, Kris Buschelman, Victor Eijkhout, William D. Gropp, Dinesh Kaushik, Matthew G. Knepley, Lois Curfman McInnes, Barry F. Smith, and Hong Zhang. PETSc users manual. Technical Report ANL-95/11 - Revision 2.1.5, Argonne National Laboratory, 2004. Available at `http://www.mcs.anl.gov/petsc`.

[3] Grey Ballard, James Demmel, Olga Holtz, and Oded Schwartz. Minimizing communication in numerical linear algebra. *SIAM Journal on Matrix Analysis and Applications*, 32:866–901, 2011.

[4] Michele Benzi, Andreas Frommer, Reinhard Nabben, and Daniel B. Szyld. Algebraic theory of multiplicative Schwarz methods. *Numerische Mathematik*, 89:605–639, 2001.

[5] Rafael Bru, Francisco Pedroche, and Daniel B. Szyld. Overlapping additive and multiplicative Schwarz iterations for H-matrices. *Linear Algebra and its Applications*, 393:91–105, 2004.

[6] Rafael Bru, Francisco Pedroche, and Daniel B. Szyld. Additive Schwarz iterations for Markov chains. *SIAM Journal on Matrix Analysis and Applications*, 27:445–458, 2005.

[7] Xiao-Chuan Cai and Marcus Sarkis. A restricted additive Schwarz preconditioner for general sparse linear systems. *SIAM Journal on Scientific Computing*, 21:792–797, 1999.

[8] Xiuhong Du, Eldad Haber, Maria Karampataki, and Daniel B. Szyld. Varying iteration accuracy using inexact conjugate gradients in control problems governed by PDE's. In *Proceedings of the 2nd Annual International Conference on Computational Mathematics, Computational Geometry and Statistics (CMCGS 2013).*, pages 29–38, Singapore, 2013. Global and Technology Forum.

[9] Xiuhong Du, Marcus Sarkis, Christian E. Schaerer, and Daniel B. Szyld. Inexact and truncated parareal-in-time krylov subspace methods for parabolic optimal control problems. *Electronic Transactions on Numerical Analysis*, 40:36–57, 2013.

[10] Xiuhong Du and Daniel B. Szyld. Inexact GMRES for singular linear system. *BIT Numerical Mathematics*, 48:511–531, 2008.

[11] Xiuhong Du and Daniel B. Szyld. A note on the mesh independence of convergence bounds for additive Schwarz preconditioned GMRES. *Numerical Linear Algebra with Applications*, 15:547–557, 2008.

[12] Olivier Dubois, Martin J. Gander, Sébastien Loisel, Amik St-Cyr, and Daniel B. Szyld. The optimized Schwarz method with a coarse grid correction. *SIAM Journal on Scientific Computing*, 34:A421–A458, 2012.

[13] Abed Elhashash, Uriel G. Rothblum, and Daniel B. Szyld. Paths of matrices with the strong Perron-Frobenius property converging to a given matrix with the Perron-Frobenius property. *Electronic Journal on Linear Algebra*, 19:90–97, 2009.

[14] Abed Elhashash and Daniel B. Szyld. Perron-Frobenius properties of general matrices. Technical Report 07-01-10, Department of Mathematics, Temple University, January 2007. Revised November 2007.

[15] Abed Elhashash and Daniel B. Szyld. Generalizations of M-matrices which may not have a nonnegative inverse. *Linear Algebra and its Applications*, 429:2435–2450, 2008.

[16] Abed Elhashash and Daniel B. Szyld. On general matrices having the Perron-Frobenius property. *Electronic Journal on Linear Algebra*, 17:389–413, 2008.

[17] Abed Elhashash and Daniel B. Szyld. Two characterizations of matrices with the Perron-Frobenius property. *Numerical Linear Algebra with Applications*, 16:863–869, 2009.

[18] Abed Elhashash and Daniel B. Szyld. Matrix functions preserving sets of generalized nonnegative matrices. *Electronic Journal on Linear Algebra*, 20:673–690, 2010.

[19] David Fritzsche, Andreas Frommer, Stephen D. Shank, and Daniel B. Szyld. Overlapping blocks by growing a partition with applications to preconditioning. *SIAM Journal on Scientific Computing*, 35:A453–A473, 2013.

[20] Andreas Frommer, Reinhard Nabben, and Daniel B. Szyld. An algebraic convergence theory for restricted additive and multiplicative Schwarz methods. In N. Debit, M. Garbey, R. Hoppe, D. Keyes, Y. Kuznetzov, and J. Périaux, editors, *Domain Decomposition Methods in Science and Engineering, Thirteenth International Conference on Domain Decomposition Methods, Lyon, France*, pages 371–377, Barcelona, 2002. CIMNE, UPC.

[21] Andreas Frommer, Hartmut Schwandt, and Daniel B. Szyld. Asynchronous weighted additive Schwarz methods. *Electronic Transactions on Numerical Analysis*, 5:48–61, 1997.

[22] Andreas Frommer and Daniel B. Szyld. Weighted max norms, splittings, and overlapping additive Schwarz iterations. *Numerische Mathematik*, 83:259–278, 1999.

[23] Andreas Frommer and Daniel B. Szyld. An algebraic convergence theory for restricted additive Schwarz methods using weighted max norms. *SIAM Journal on Numerical Analysis*, 39:463–479, 2001.

[24] Martin J. Gander, Sébastien Loisel, and Daniel B. Szyld. An optimal block iterative method and preconditioner for banded matrices with applications to PDEs on irregular domains. *SIAM Journal on Matrix Analysis and Applications*, 33:653–680, 2012.

[25] David F. Gleich. Techniques for local and global centrality estimation. Talk presented at the Workshop 11451 on Data Mining, Networks and Dynamics, Schloss Dagstuhl, Leibniz Center for Computer Science, Wadern, Germany, 6–11 November 2011.

[26] Mark Hoemmen. *Communication-avoiding Krylov subspace methods*. PhD thesis, EECS Department, University of California, Berkeley, 2010. Also available as UC Berkeley Technical Report UCB/EECS-2010-37.

[27] Sébastien Loisel, Reinhard Nabben, and Daniel B. Szyld. On hybrid multigrid-Schwarz algorithms. *Journal on Scientific Computing*, 36:165–175, 2008.

[28] Sébastien Loisel and Daniel B. Szyld. A maximum principle for L^2-trace norms with an application to optimized Schwarz methods. In Michel Bercovier, Martin Gander, Ralf Kornhuber, and Olof B. Widlund, editors, *Domain Decomposition Methods in Science and Engineering XVIII*, volume 70 of *Lecture Notes in Computational Science and Engineering*, pages 163–170, Berlin, Heidelberg, New York, 2009. Springer.

[29] Sébastien Loisel and Daniel B. Szyld. On the convergence of optimized Schwarz methods by way of matrix analysis. In Michel Bercovier, Martin Gander, Ralf Kornhuber, and Olof B. Widlund, editors, *Domain Decomposition Methods in Science and Engineering XVIII*, volume 70 of *Lecture Notes in Computational Science and Engineering*, pages 363–370, Berlin, Heidelberg, New York, 2009. Springer.

[30] Sébastien Loisel and Daniel B. Szyld. On the convergence of algebraic optimizable Schwarz methods with applications to elliptic problems. *Numerische Mathematik*, 114:697–728, 2010.

[31] Ivo Marek and Daniel B. Szyld. Algebraic analysis of Schwarz methods for singular systems. In R. Kornhuber, R. H. W. Hoppe, J. Périaux, O. Pironneau, O. B. Widlund, and J. Xu, editors, *Domain Decomposition Methods in Science and Engineering*, volume 40 of *Lecture notes in Computer Science and Engineering*, pages 647–652. Springer, 2004.

[32] Ivo Marek and Daniel B. Szyld. Algebraic Schwarz methods for the numerical solution of Markov chains. *Linear Algebra and its Applications*, 386:67–81, 2004.

[33] J.A. Meijerink and H. van der Vorst. An iterative solution method for linear systems of which the coefficient matrix is a symmetric M-matrix. *Mathematics of Computation*, 31:148–162, 1977.

[34] Reinhard Nabben and Daniel B. Szyld. Convergence theory of restricted multiplicative Schwarz methods. *SIAM Journal on Numerical Analysis*, 40:2318–2336, 2003.

[35] Reinhard Nabben and Daniel B. Szyld. Schwarz iterations for symmetric positive semidefinite problems. *SIAM Journal on Matrix Analysis and Applications*, 29:98–116, 2006.

[36] U.S. Department of Energy. Scientific grand challenges, crosscutting technologies for computing at the exascale. Report on a DOE Workshop held in Washington, DC, 2–4 February 2010.

[37] Michael L. Parks, Kirk Soodhalter, and Daniel B. Szyld. Block Krylov subspace recycling, 2013. In Preparation.

[38] Hassane Sadok and Daniel B. Szyld. A new look at CMRH and its relation to GMRES. *BIT Numerical Mathematics*, 52:485–501, 2012.

[39] Marcus Sarkis and Daniel B. Szyld. A proposal for a dynamically adapted inexact additive Schwarz preconditioner. In Olof Widlund and David Keyes, editors, *Domain Decomposition Methods in Science and Engineering XVI*, volume 55 of *Lecture Notes in Computational Science and Engineering*, pages 333–337, Berlin, Heidelberg, New York, 2006. Springer.

[40] Marcus Sarkis and Daniel B. Szyld. Optimal left and right additive Schwarz preconditioning for minimal residual methods with Euclidean and energy norms. *Computer Methods in Applied Mechanics and Engineering*, 196:1612–1621, 2007.

[41] Valeria Simoncini and Daniel B. Szyld. Flexible inner-outer Krylov subspace methods. *SIAM Journal on Numerical Analysis*, 40:2219–2239, 2003.

[42] Valeria Simoncini and Daniel B. Szyld. Theory of inexact Krylov subspace methods and applications to scientific computing. *SIAM Journal on Scientific Computing*, 25:454–477, 2003.

[43] Valeria Simoncini and Daniel B. Szyld. The effect of non-optimal bases on the convergence of Krylov subspace methods. *Numerische Mathematik*, 100:711–733, 2005.

[44] Valeria Simoncini and Daniel B. Szyld. On the occurrence of superlinear convergence of exact and inexact Krylov subspace methods. *SIAM Review*, 47:247–272, 2005.

[45] Valeria Simoncini and Daniel B. Szyld. Relaxed Krylov subspace approximation. *PAMM*, 5:797–800, 2005.

[46] Valeria Simoncini and Daniel B. Szyld. Recent computational developments in Krylov subspace methods for linear systems. *Numerical Linear Algebra with Applications*, 14:1–59, 2007.

[47] Valeria Simoncini and Daniel B. Szyld. New conditions for non-stagnation of minimal residual methods. *Numerische Mathematik*, 109:477–487, 2008.

[48] Valeria Simoncini and Daniel B. Szyld. Interpreting IDR as a Petrov-Galerkin method. *SIAM Journal on Scientific Computing*, 32:1898–1912, 2010.

[49] Barry F. Smith, Petter E. Bjørstad, and William D. Gropp. *Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations*. Cambridge University Press, Cambridge, New York, Melbourne, 1996.

[50] Daniel B. Szyld and Judith A. Vogel. A flexible quasi-minimal residual method with inexact preconditioning. *SIAM Journal on Scientific Computing*, 23:363–380, 2001.

[51] Daniel B. Szyld and Fei Xue. Efficient preconditioned inner solves for inexact Rayleigh quotient iteration and their connections to the single-vector Jacobi-Davidson method. *SIAM Journal on Matrix Analysis and Applications*, 32:993–1018, 2011.

[52] Daniel B. Szyld and Fei Xue. Several properties of invariant pairs of nonlinear algebraic eigenvalue problems. Technical Report 12-02-09, Department of Mathematics, Temple University, February 2012. Revised February 2013.

[53] Daniel B. Szyld and Fei Xue. Local convergence analysis of several inexact Newton-type algorithms for general nonlinear eigenvalue problems. *Numerische Mathematik*, 123:333–362, 2013.

[54] Andrea Toselli and Olof Widlund. *Domain Decomposition Methods - Algorithms and Theory*, volume 34 of *Series in Computational Mathematics*. Springer, Berlin, Heidelberg, New York, 2005.

[55] Fei Xue and Howard C. Elman. Fast inexact implicitly restarted Arnoldi method for generalized eigenvalue problems with spectral transformation. Technical report, Department of Computer Science, University of Maryland, February 2010. Submitted to *SIAM Journal on Matrix Analysis and Applications*, in revision.

[56] Xingwei Yang, Daniel B. Szyld, and Longin Jan Latecki. Diffusion on a tensor product graph for semi-supervised learning and interactive image segmentation. *Advances in Imaging and Electron Physics*, 169:147–172, 2011.