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 these 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 the 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


