Nonlinear problems in analysis of Krylov subspace methods

Zdeněk Strakoš[†]

[†] *Institute of Computer Science, Pod Vod. věží 2, 182 07 Praha 8, Czech republic,* `strakos@cs.cas.cz`

Keywords : Krylov subspace methods, nonlinear phenomena, minimization, Gauss-Christoffel quadrature, problem of moments, convergence, spectral decomposition, numerical stability and rounding error analysis

Consider a system of linear algebraic equations $Ax = b$ where A is an n by n real matrix and b a real vector of length n. Unlike in the linear iterative methods based on the idea of splitting of A, the Krylov subspace methods, which are used in computational kernels of various optimization techniques, look for some optimal approximate solution x^n in the subspaces $\mathcal{K}_n(A, b) = \text{span}\{b, Ab, \dots, A^{n-1}b\}, n = 1, 2, \dots$ (here we assume, with no loss of generality, $x^0 = 0$). As a consequence, though the problem $Ax = b$ is linear, Krylov subspace methods *are not*. Their convergence behaviour cannot be viewed as an (unimportant) initial transient stage followed by the subsequent convergence stage. Apart from very simple, and from the point of view of Krylov subspace methods uninteresting cases, it cannot be meaningfully characterized by an asymptotic rate of convergence. In Krylov subspace methods such as the conjugate gradient method (CG) or the generalized minimal residual method (GMRES), the optimality at each step over Krylov subspaces of increasing dimensionality makes any linearized description inadequate.

CG applied to $Ax = b$ with a symmetric positive definite A can be viewed as a method for numerical minimization the quadratic functional $1/2(Ax, x) - (b, x)$. In order to reveal its nonlinear character, we consider CG a matrix formulation of the Gauss-Christoffel quadrature, and show that it essentially solves the classical Stieltjes moment problem. Moreover, though the CG behaviour is fully determined by the spectral decomposition of the problem, the relationship between convergence and spectral information is nothing but simple. We will explain several phenomena where an intuitive commonly used argumentation can lead to wrong conclusions, which can be found in the literature. We also show that rounding error analysis of CG brings fundamental understanding of seemingly unrelated problems in convergence analysis and in theory of the Gauss-Christoffel quadrature [5, 3, 4].

In remaining time we demonstrate that in the unsymmetric case the spectral information is not generally sufficient for description of behaviour of Krylov subspace methods. In particular, given an arbitrary prescribed convergence history of GMRES and an arbitrary prescribed spectrum of the system matrix, there is always a system $Ax = b$ such that GMRES follows the prescribed convergence while A has the prescribed spectrum [2, 1].

References

[1] Arioli, M. and Pták, V. and Strakoš, Z., "Krylov sequences of maximal length and convergence of GMRES", *BIT*, **38** , pp. 636–643 (1998).

[2] Greenbaum, A. and Pták, V. and Strakoš, Z., "Any nonincreasing convergence curve is possible for GMRES", SIAM J. Matrix Anal. Appl., **17**, pp. 465–469 (1996).

[3] Meurant, G. and Strakoš, Z., "The Lanczos and conjugate gradient algorithms in finite precision arithmetic", Acta Numerica, **15**, pp. 471–542 (2006).

[4] O'Leary, D. P. and Strakoš, Z. and and Tichý, P., "On sensitivity of Gauss-Christoffel quadrature", Numerische Mathematik, (published electronically on April 14, 2007).

[5] Strakoš, Z. and Tichý, P., "On error estimation in the conjugate gradient method and why it works in finite precision computations", Electronic Trans. Numer. Anal., **13**, pp. 56–80 (2002).