Lanczos tridiagonalization, Golub-Kahan bidiagonalization and core problem

Iveta Hnětynková *, **Martin Plešinger** **, and **Zdeněk Strakoš*****

Academy of Sciences of the Czech republic, Institute of Computer Science, Pod Vodárenskou věží 2, 182 07 Prague 8, Czech Republic

1 Introduction

Consider an orthogonally invariant linear approximation problem $Ax \approx b$. In [8] it is proved that the partial upper bidiagonalization of the extended matrix $[b, A]$ determines a *core approximation problem* $A_{11}x_1 \approx b_1$, with all necessary and sufficient information for solving the original problem given by b_1 and A_{11}. It is shown how the core problem can be used in a simple and efficient way for solving different formulations of the original approximation problems. In [3] the core problem formulation is derived from the relationship between the Golub-Kahan bidiagonalization [2] and the Lanczos tridiagonalization [5], and from the known properties of Jacobi matrices. Here we briefly recall the approach from [3], and outline a possible direction for further research.

2 Core problem

Consider estimating x from the real linear approximation problem

$$Ax \approx b\,, \quad A \text{ a nonzero } n \text{ by } m \text{ matrix}, \quad b \text{ a nonzero } n\text{-vector}, \tag{1}$$

where the uninteresting case is excluded by the assumption $A^T b \neq 0$. In the paper [8] it was proposed to transform the original data $[b, A]$ into the form

$$P^T \left[\ b \parallel AQ\ \right] = \left[\begin{array}{c||c|c} b_1 & A_{11} & 0 \\ \hline 0 & 0 & A_{22} \end{array}\right], \quad \text{where} \quad P^{-1} = P^T\,, \quad Q^{-1} = Q^T\,, \tag{2}$$

$b_1 = \beta_1 e_1$ and A_{11} is a lower bidiagonal matrix with *nonzero bidiagonal elements*. The matrix A_{11} is either square, when (1) is compatible, or rectangular, when (1) is incompatible. The original problem is in this way decomposed into the approximation problems $A_{11}x_1 \approx b_1$ and $A_{22}x_2 \approx 0$. It is suggested to solve the first problem, set $x_2 \equiv 0$ and substitute $x \equiv Q\,[x_1^T, 0]^T$ for the solution of (1). The transformation described above has the following remarkable properties, see [3, Theorem 1.1], with the proof given in [8, Theorem 2.2, 3.2, 3.3].

Theorem 2.1 *Let A be a nonzero n by m real matrix and b a nonzero real $n-$vector, $A^T b \neq 0$. Then there exists a decomposition (2), where $b_1 = \beta_1 e_1$ and A_{11} is a lower bidiagonal matrix with nonzero bidiagonal elements. Moreover:*

1: *The matrix A_{11} has full column rank and its singular values are simple.*

2: *The matrix A_{11} has minimal dimensions over all orthogonal transformations giving the block structure (2).*

3: *All components of $b_1 = \beta_1 e_1$ in the left singular vector subspaces of A_{11} are nonzero.*

The proof of Theorem 2.1 in [8] is based on the singular value decomposition of the matrix A. In [3] the relationship between the Golub-Kahan bidiagonalization and the Lanczos tridiagonalization is used for constructing the proof.

3 Lanczos tridiagonalization and core problem properties

Consider the *partial* lower Golub-Kahan bidiagonalization of $[b, A]$ in the following form. Given the initial vectors $v_0 \equiv 0, u_1 \equiv b/\beta_1$, where $\beta_1 \equiv \|b\| \neq 0$ and $\|.\|$ represents the standard Euclidean norm, the algorithm computes for $i = 1, 2, \ldots$

$$\alpha_i v_i = A^T u_i - \beta_i v_{i-1}\,, \quad \|v_i\| = 1\,, \qquad \beta_{i+1} u_{i+1} = A v_i - \alpha_i u_i\,, \quad \|u_{i+1}\| = 1 \tag{3}$$

* Corresponding author: e-mail: hnetynkova@cs.cas.cz, Phone: +420 266 053 787, Fax: +420 286 585 789

** e-mail: mata@cs.cas.cz

*** e-mail: strakos@cs.cas.cz

until $\alpha_i = 0$ or $\beta_{i+1} = 0$, or until $i = \min\{n, m\}$. Denote $U_k \equiv (u_1, \dots, u_k)$, $V_k \equiv (v_1, \dots, v_k)$ the matrices with orthonormal columns, L_k the square lower bidiagonal matrix with the main diagonal $(\alpha_1, \dots, \alpha_k)$ and the subdiagonal $(\beta_2, \dots, \beta_k)$ and $L_{k+} \equiv (L_k^T, \beta_{k+1} e_k)^T$. In the rest of this contribution we consider (1) incompatible; the compatible case is simpler and can be treated analogously, see [3]. Then (3) must stop with some $\alpha_{p+1} = 0$ or $p = m$ giving

$$A^T U_p \;=\; V_p L_p^T\,, \quad A V_p \;=\; U_{p+1}\, L_{p+}\,.$$

Consequently, $U_{p+1}^T [b, AV_p] \;=\; [\beta_1 e_1, L_{p+}] \;\equiv\; [b_1 | A_{11}]$ and $A_{11} x_1 \;\equiv\; L_{p+}\, x_1 \;\approx\; \beta_1 e_1 \;\equiv\; b_1$ is the *incompatible* core problem. The matrices U_{p+1} and V_p represent the first $(p+1)$ and p columns of the matrices P and Q, respectively.

Given a real symmetric matrix B and a starting vector w_1, $\|w_1\| = 1$, the Lanczos tridiagonalization algorithm can be written in the matrix form

$$B W_k \;=\; W_k T_k \;+\; \delta_{k+1} w_{k+1} e_k^T\,, \quad W_k^T w_{k+1} = 0\,, \tag{4}$$

where $W_k e_1 = w_1$, W_k has orthonormal columns and T_k is a symmetric tridiagonal matrix with positive subdiagonal elements, a Jacobi matrix. Defining $B \equiv A^T A, w_1 \equiv v_1 = A^T b / \|A^T b\|$, it can be shown (see, e.g., [1]) that the Lanczos tridiagonalization (4) produces the matrices $W_k \equiv V_k$, $T_k \equiv L_{k+}^T L_{k+}$ and $\delta_{k+1} \equiv \alpha_{k+1}\beta_{k+1}$. Thus, the matrix L_{p+} can be linked to the Jacobi matrix

$$T_p \;\equiv\; L_{p+}^T L_{p+} \;=\; \begin{pmatrix} \alpha_1^2 + \beta_2^2 & \alpha_2\beta_2 & & \\ \alpha_2\beta_2 & \alpha_2^2 + \beta_3^2 & \ddots & \\ & \ddots & \ddots & \alpha_p\beta_p \\ & & \alpha_p\beta_p & \alpha_p^2 + \beta_{p+1}^2 \end{pmatrix}; \quad L_{p+} \;=\; \begin{pmatrix} \alpha_1 & & & \\ \beta_2 & \alpha_2 & & \\ & \ddots & \ddots & \\ & & \beta_p & \alpha_p \\ & & & \beta_{p+1} \end{pmatrix}. \tag{5}$$

This fact is used in [3], together with the properties of Jacobi matrices (see, e.g., [9]), for proving Theorem 2.1.

The presented relationship may be found useful in applications of the core problem formulation. In large ill-posed problems the outer Golub-Kahan bidiagonalization can be combined with an inner regularization applied to the problem $L_{k+} y \approx \beta_1 e_1$. Here stopping criteria are typically based on estimation of the L-curve, the discrepancy principle or generalized cross validation (see, e.g., [6], [7]). From the core problem point of wiev one should ask whether and when the matrix L_{k+} for $k < p$ (possibly $k \ll p$) can be considered a *sufficiently good approximation* to the core matrix L_{p+}. When $p \ll m$, one must ask how to *numerically* indicate the separation of the core problem, since in finite precision computation α_{p+1} will hardly be identicaly zero. Similarly, one can ask when the tridiagonal matrix T_k, $k < p$, sufficiently approximates the matrix T_p discussed above. It might be useful to study in this context perturbation theory of Jacobi matrices, in particular the specific perturbations when the off-diagonal element $\delta_{k+1} = \alpha_{k+1}\beta_{k+1}$ is replaced by zero, see [4].

We believe that the presented relationships, together with known results on Jacobi matrices, can be used in further investigation of effective stopping criteria in regularization methods.

Acknowledgements This work was supported by the National Program of Research "Information Society" under project 1ET400300415, and by the Institutional Research Plan AVOZ10300504. The research of I. Hnětynková was partly supported by the Grant GAUK 6/2005/R.

References

[1] A. Björck, A bidiagonalization algorithm for solving large and sparse ill-posed systems of linear equations, BIT **28**, pp. 659–670 (1988).
[2] G. H. Golub, W. Kahan, Calculating the singular values and pseudo-inverse of a matrix, SIAM J. Numer. Anal. Ser. B **2**, pp. 205–224 (1965).
[3] I. Hnětynková, Z. Strakoš, Lanczos tridiagonalization and core problems, to appear in Lin. Alg. Appl.
[4] E.-X. Jiang, Perturbation in eigenvalues of a symmetric tridiagonal matrix, Lin. Alg. Appl. **399**, pp. 91–107 (2005).
[5] C. Lanczos, An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, J. Res. Nat. Bur. Standards **45**, pp. 255–282 (1950).
[6] D. Sima, S. Van Huffel, Using core formulations for ill-posed linear systems, PAMM **5**, pp. 795–796 (2005).
[7] D. Sima, Regularization techniques in model fitting and parameter estimation, PhD thesis, Faculty of Engineering, K.U.Leuven (2006).
[8] C. C. Paige, Z. Strakoš, Core problems in linear algebraic systems, SIAM J. Matrix Anal. Appl. **27**, pp. 861–875 (2006).
[9] B. N. Parlett, The Symmetric Eigenvalue Problem, second ed., (SIAM, Philadelphia, 1998), chap. 7.