Lanczos tridiagonalization and core problems

Iveta Hnětynková, Zdeněk Strakoš

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

Received 17 April 2006; accepted 11 May 2006
Available online 14 July 2006
Submitted by R.A. Brualdi

Abstract

The Lanczos tridiagonalization orthogonally transforms a real symmetric matrix \( A \) to symmetric tridiagonal form. The Golub–Kahan bidiagonalization orthogonally reduces a nonsymmetric rectangular matrix to upper or lower bidiagonal form. Both algorithms are very closely related.

The paper [C.C. Paige, Z. Strakoš, Core problems in linear algebraic systems, SIAM J. Matrix Anal. Appl. 27 (2006) 861–875] presents a new formulation of orthogonally invariant linear approximation problems \( Ax \approx b \). 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 further shown how the core problem can be used in a simple and efficient way for solving different formulations of the original approximation problem. Our contribution relates the core problem formulation to the Lanczos tridiagonalization and derives its characteristics from the relationship between the Golub–Kahan bidiagonalization, the Lanczos tridiagonalization and the well-known properties of Jacobi matrices.

© 2006 Elsevier Inc. All rights reserved.

AMS classification: 15A06; 15A18; 15A23; 65F10; 65F20; 65F25

Keywords: Linear approximation problem; Orthogonal transformation; Core problem; Golub–Kahan bidiagonalization; Lanczos tridiagonalization; Jacobi matrix
1. Introduction

Let $A$ be a nonzero $n \times m$ real matrix, and $b$ be a nonzero real $n$-vector. Consider estimating $x$ from the linear approximation problem

$$Ax \approx b,$$

(1.1)

where the uninteresting case is for clarity of exposition excluded by the natural assumption $b \not\perp \mathbb{R}(A)$, that is $A^Tb \neq 0$. In a sequence of papers [1–3] it was proposed to orthogonally transform the original data $b, A$ into the form

$$P^T [b \parallel AQ] = \begin{bmatrix} b_1 & 0 & A_{11} & 0 \\ 0 & 0 & 0 & A_{22} \end{bmatrix},$$

(1.2)

where $P^{-1} = P^T$, $Q^{-1} = Q^T$, $b_1 = \beta_1 e_1$, and $A_{11}$ is a lower bidiagonal matrix with nonzero bidiagonal elements. The matrix $A_{11}$ is either square, when (1.1) is compatible, or rectangular, when (1.1) is incompatible. The matrix $A_{22}$ (which need not be bidiagonalized) and the corresponding block row and/or column in (1.2) can be nonexistent. The transformed data $b_1$ and $A_{11}$ can be computed either directly, using Householder orthogonal transformations (see for example [4, Section 5.4.3, p. 251]), or iteratively, using the Golub–Kahan bidiagonalization [5], see also [6]. The bidiagonalization is stopped at the first zero element, giving the block structure in (1.2). The original problem is in this way decomposed into the approximation problem

$$A_{11}x_1 \approx b_1,$$

(1.3)

which contains the necessary and sufficient information for solving the problem (1.1), and the remaining part $A_{22}x_2 \approx 0$. The problem (1.3) is therefore called a core problem within (1.1). In [3], it was proposed to find $x_1$ from (1.3), set $x_2 \equiv 0$, and substitute for the solution of (1.1)

$$x \equiv Q \begin{bmatrix} x_1 \\ 0 \end{bmatrix}.$$

(1.4)

The (partial) upper bidiagonalization of $[b, A]$ described above represents a fundamental decomposition of data in the linear approximation problem (1.1), with the following remarkable characteristics.

**Theorem 1.1.** Let $A$ be a nonzero $n \times m$ real matrix and $b$ a nonzero real $n$-vector, $A^Tb \neq 0$. Then there exists a decomposition

$$P^T [b \parallel AQ] = \begin{bmatrix} b_1 & 0 & A_{11} & 0 \\ 0 & 0 & 0 & A_{22} \end{bmatrix},$$

where $P^{-1} = P^T$, $Q^{-1} = Q^T$, $b_1 = \beta_1 e_1$ and $A_{11}$ is a lower bidiagonal matrix with nonzero bidiagonal elements. Moreover:

**B1.** The matrix $A_{11}$ has full column rank and its singular values are simple. Consequently, any zero singular values or repeats that $A$ has must appear in $A_{22}$.

**B2.** The matrix $A_{11}$ has minimal dimensions, and $A_{22}$ has maximal dimensions, over all orthogonal transformations giving the block structure above, without any additional assumptions on the structure of $A_{11}$ and $b_1$.

**B3.** All components of $b_1 = \beta_1 e_1$ in the left singular vector subspaces of $A_{11}$, i.e., the first elements of all left singular vectors of $A_{11}$, are nonzero.
The proofs of B1–B3 are given in [3, Theorems 2.2, 3.2, 3.3]; they are based on the singular value decomposition of $A$ and on the properties of the upper bidiagonal form $[b_1, A_{11}]$ with positive bidiagonal elements.

As mentioned above, when $A$ is large and sparse, the Golub–Kahan bidiagonalization is suggested in [3] as the algorithm for computing the core problem data $b_1$ and $A_{11}$. At any iteration step, the computed left principal part of $A_{11}$ represents an approximation to the core problem matrix. Practical applications may require stopping the computation before the full decomposition (1.2) is reached. Therefore it is important to study iterative approximations to the core problem decomposition. The Golub–Kahan bidiagonalization is closely related to the Lanczos tridiagonalization [7], which has been thoroughly investigated as a tool for computation of a few dominant eigenvalues. We believe that the knowledge about the partial Lanczos tridiagonalization may prove useful in the future investigation of the partial core problem decomposition. Therefore, as a first step, we summarize in this paper the relationship of the core problem decomposition with the Lanczos tridiagonalization. We prove B1–B3 from the connection between the Golub–Kahan bidiagonalization and the Lanczos tridiagonalization and from well-known properties of the related Jacobi matrices.

We assume, for simplicity of notation, that $A$ and $b$ are real. The extension to complex data is straightforward.

2. Golub and Kahan bidiagonalization and Lanczos tridiagonalization

Consider the partial lower Golub–Kahan bidiagonalization of the $n \times m$ real matrix $A$ 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 $\|\cdot\|$ represents the standard Euclidean norm, the algorithm computes for $i = 1, 2, \ldots$

\[
\begin{align*}
\alpha_i v_i &= A^T u_i - \beta_i v_{i-1}, \quad \|v_i\| = 1, \quad \text{(2.1)} \\
\beta_{i+1} u_{i+1} &= A v_i - \alpha_i u_i, \quad \|u_{i+1}\| = 1 \quad \text{(2.2)}
\end{align*}
\]

until $\alpha_i = 0$ or $\beta_{i+1} = 0$, or until $i = \min\{n, m\}$.

We present, for completeness, the basic properties of the Golub–Kahan bidiagonalization as given in [6]. Consider $\alpha_i \beta_i \neq 0$ for $1 \leq i \leq k + 1$ and denote $U_k \equiv (u_1, \ldots, u_k)$, $V_k \equiv (v_1, \ldots, v_k)$,

\[
L_k \equiv \begin{pmatrix}
\alpha_1 \\
\alpha_2 \\
\vdots \\
\alpha_k \\
\beta_2 \\
\vdots \\
\beta_k
\end{pmatrix}, \quad L_{k+} \equiv \begin{pmatrix}
L_k \\
\beta_{k+1} e_k^T
\end{pmatrix}.
\]

Then (2.1)–(2.2) can be rewritten in the matrix form

\[
\begin{align*}
A^T U_k &= V_k L_k^T, \\ AV_k &= [U_k, u_{k+}] L_{k+},
\end{align*}
\]

giving

\[
U_k^T A V_k = (A^T U_k)^T V_k = L_k V_k^T V_k = U_k^T [U_k, u_{k+}] L_{k+} = U_k^T U_k L_k + U_k^T u_{k+} \beta_{k+1} e_k^T
\]
and thus
\[ L_k V_k^T V_k = U_k^T U_k L_k + U_k^T u_{k+1} \beta_{k+1} e_k^T. \]  
(2.5)

Similarly, (2.1) gives for \( i = k + 1 \)
\[ A^T [U_k, u_{k+1}] = V_k L_{k+1}^T + v_{k+1} \alpha_{k+1} e_{k+1}^T \]  
(2.6)
and therefore
\[ V_k^T A^T [U_k, u_{k+1}] = V_k^T V_k L_{k+1}^T + V_k^T v_{k+1} \alpha_{k+1} e_{k+1}^T = (AV_k)^T [U_k, u_{k+1}] = L_{k+1}^T [U_k, u_{k+1}] [U_k, u_{k+1}], \]
which yields
\[ L^T_{k+1} [U_k, u_{k+1}] = V_k^T V_k L_{k+1}^T + V_k^T v_{k+1} \alpha_{k+1} e_{k+1}^T. \]  
(2.7)
As a direct consequence one gets the following fundamental property: the vectors \( u_1, u_2, \ldots, u_{k+1} \) respectively \( v_1, v_2, \ldots, v_{k+1} \) are orthonormal. Indeed, the induction assumption that \( U_k^T U_k = I \), \( V_k^T V_k = I \) gives from (2.5)
\[ L_k = L_k + U_k^T u_{k+1} \beta_{k+1} e_k^T, \]
and thus \( U_k^T u_{k+1} = 0 \), because \( \beta_{k+1} \neq 0 \). Similarly, (2.7) and \( \alpha_{k+1} \neq 0 \) yield
\[ L^T_{k+1} = L^T_{k+1} + V_k^T v_{k+1} \alpha_{k+1} e_{k+1}^T, \]
that gives \( V_k^T v_{k+1} = 0 \).

Summarizing, the Golub–Kahan bidiagonalization (2.1)–(2.2) of the \( n \times m \) matrix \( A \) with \( u_1 = b/\|b\| \) results in one of the two following situations, which will be distinguished throughout the paper:

**Case 1.** \( \alpha_i \beta_i \neq 0 \) for \( i = 1, \ldots, p \); \( \beta_{p+1} = 0 \) or \( p = n \). Then (2.3) gives
\[ U_p^T A V_p = L_p, \]  
(2.8)
\[ U_p^T [b, A V_p] = \begin{bmatrix} \beta_1 & \alpha_1 & \beta_2 & \alpha_2 & \ldots & \ldots & \beta_p & \alpha_p \end{bmatrix} \equiv [b_1 | A_{11}] \text{ here} \]  
(2.9)
and \( A_{11} x_1 = L_p x_1 \approx \beta_1 e_1 = b_1 \) is the compatible core problem. The matrices \( U_p, V_p \) represent the first \( p \) columns of the matrices \( P, Q \), respectively, see (1.2).

**Case 2.** \( \alpha_i \beta_i \neq 0 \) for \( i = 1, \ldots, p \), and \( \beta_{p+1} \neq 0 \); \( \alpha_{p+1} = 0 \) or \( p = m \). Then (2.4) gives
\[ [U_p, u_{p+1}]^T A V_p = L_{p+1}, \]  
(2.10)
\[ [U_p, u_{p+1}]^T [b, A V_p] = \begin{bmatrix} \beta_1 & \alpha_1 & \beta_2 & \alpha_2 & \ldots & \ldots & \beta_p & \alpha_p \\ \beta_{p+1} \end{bmatrix} \equiv [b_1 | A_{11}] \text{ here} \]  
(2.11)
and $A_{11}x_1 = L_{p+1}x_1 \approx \beta_1 e_1 \equiv b_1$ 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.

For clarity of exposition we review the situations when the bidiagonalization is not stopped until the maximum number of steps is reached. If $p = n = m$, then $U_p = U_n = P$, $V_p = V_m = Q$ and

$$P^T \begin{bmatrix} b & \parallel AQ \parallel \end{bmatrix} = \begin{bmatrix} b_1 & \parallel A_{11} \parallel \end{bmatrix}.$$

If $p = n < m$, then $U_p = U_n = P$, and completing $V_p$ by $(m - n)$ additional columns into the orthogonal matrix $Q$ gives

$$P^T \begin{bmatrix} b & \parallel AQ \parallel \end{bmatrix} = \begin{bmatrix} b_1 & \parallel A_{11} \parallel & 0 \end{bmatrix}.$$

If $p = m < n$, then $V_p = V_m = Q$, and completing $U_{p+1}$ by $(n - m - 1)$ additional columns into the orthogonal matrix $P$ gives

$$P^T \begin{bmatrix} b & \parallel AQ \parallel \end{bmatrix} = \begin{bmatrix} b_1 & 0 & 0 \\ 0 & \parallel A_{11} \parallel & 0 \end{bmatrix}.$$

The bidiagonalization algorithm is closely connected with the Lanczos tridiagonalization (see [7]). Let $B$ be a $t \times t$ real symmetric matrix. Given the initial vector $w_1$, $\|w_1\| = 1; w_0 \equiv 0, \delta_1 \equiv 0$, the algorithm computes for $i = 1, 2, \ldots$

$$y_i = Bw_i - \delta_i w_{i-1},$$

$$\gamma_i = (y_i, w_i),$$

$$\delta_{i+1} w_{i+1} = y_i - \gamma_i w_i, \quad \|w_{i+1}\| = 1$$

until $\delta_{i+1} = 0$, or until $i + 1 = t$. Consider $\delta_i \neq 0$ for $1 \leq i \leq k + 1$ and denote $W_k \equiv (w_1, \ldots, w_k)$,

$$T_k \equiv \begin{pmatrix} \gamma_1 & \delta_2 & & \\ \delta_2 & \gamma_2 & & \\ & \ddots & \ddots & \delta_k \\ & & \delta_k & \gamma_k \end{pmatrix}.$$

Then $W_k$ has orthonormal columns and $T_k$ represents the symmetric tridiagonal matrix with positive elements on the subdiagonal (Jacobi matrix). The Lanczos algorithm can be written in the matrix form

$$BW_k = W_k T_k + \delta_{k+1} w_{k+1} e_k^T, \quad W_k^T w_{k+1} = 0. \quad (2.12)$$

Given a real symmetric $B$, (2.12) is fully determined by the starting vector $w_1$. Moreover, by the well-known properties of Jacobi matrices:

**J1.** $T_k$ has distinct eigenvalues (see, e.g., [8, Lemma 7.7.1, p. 134]);

**J2.** If $B$ is real symmetric positive semidefinite and $w_1 \perp \ker (B)$, then all eigenvalues of $T_k$ are positive;

**J3.** The first (as well as the last) components of all eigenvectors of $T_k$ are nonzero (see, e.g., [8, Theorem 7.9.3, p. 140]).

Note that **J2** follows from the fact that the final Jacobi matrix $T_l$, for which $BW_l = W_l T_l$, must be nonsingular (and, using the assumption in **J2**, symmetric positive definite) and from the interlacing property (see [8, Theorem 10.1.1, p. 203]).
The relationship between the Lanczos tridiagonalization and the Golub–Kahan bidiagonalization can be described in several ways, see [9, pp. 662–663], [10, pp. 513–515], [11, p. 611], [5, pp. 212–214] and also [6, pp. 199–200], [12, pp. 44–48], [13, pp. 115–118]. Consider \( \alpha_i \beta_i \neq 0 \) for \( 1 \leq i \leq k + 1 \). The Lanczos tridiagonalization applied to the augmented matrix

\[
B \equiv \begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix}
\]

with the starting vector \( w_1 \equiv (u_1, 0)^T \) yields in \( 2k \) steps the orthogonal matrix

\[
W_{2k} = \begin{pmatrix} u_1 & 0 & \cdots & u_k & 0 \\ 0 & v_1 & \cdots & 0 & v_k \end{pmatrix}
\]

and the Jacobi matrix \( T_{2k} \) with the zero main diagonal and the subdiagonals equal to \( (\alpha_1, \beta_2, \ldots, \beta_k, \alpha_k) \). Note the equivalence of (2.12) (using this \( B \) and \( W \)) with (2.1)–(2.2). Furthermore, (2.3) multiplied by \( A \) and combined with (2.4) gives

\[
AA^T U_k = A V_k L_k^T = [U_k, u_{k+1}] L_k + L_k^T = U_k L_k L_k^T + \alpha_k \beta_{k+1} u_{k+1} e_k^T.
\]

(2.13)

where

\[
L_k L_k^T = \begin{pmatrix}
\alpha_1^2 & \alpha_1 \beta_2 \\
\alpha_1 \beta_2 & \alpha_2^2 + \beta_2^2 & \ddots \\
\vdots & \ddots & \ddots & \alpha_{k-1} \beta_k \\
\alpha_{k-1} \beta_k & \alpha_k^2 + \beta_k^2 
\end{pmatrix}
\]

In short, (2.13) represents \( k \) steps of the Lanczos tridiagonalization of the matrix \( AA^T \) with the starting vector \( u_1 = b/\beta_1 = b/\|b\| \). Here we have \( B^{(1)} \equiv AA^T, W^{(1)} \equiv U_k, T^{(1)} \equiv L_k L_k^T \) and \( \delta^{(1)}_{k+1} \equiv \alpha_k \beta_{k+1} \). Similarly, (2.4) together with (2.6) gives

\[
A^T A V_k = A^T [U_k, u_{k+1}] L_{k+} = V_k L_{k+} L_{k+}^T + \alpha_{k+1} \beta_{k+1} v_{k+1} e_k^T,
\]

(2.14)

where

\[
L_{k+} L_{k+}^T = L_k^T L_k + \beta_{k+1}^2 e_k e_k^T = \begin{pmatrix}
\alpha_1^2 + \beta_2^2 & \alpha_2 \beta_2 \\
\alpha_2 \beta_2 & \alpha_2^2 + \beta_3^2 & \ddots \\
\vdots & \ddots & \ddots & \alpha_{k-1} \beta_k \\
\alpha_{k-1} \beta_k & \alpha_k^2 + \beta_{k+1}^2 
\end{pmatrix}
\]

The identity (2.14) represents \( k \) steps of the Lanczos tridiagonalization of the matrix \( A^T A \) with the starting vector \( v_1 = A^T u_1 / \alpha_1 = A^T b / \|A^T b\| \). Here we have \( B^{(2)} \equiv A^T A, W^{(2)} \equiv V_k, T^{(2)} \equiv L_{k+}^T L_{k+} \) and \( \delta^{(2)}_{k+1} \equiv \alpha_k \beta_{k+1} \).

3. The core problem characteristics

In this section, we prove Theorem 1.1 by relating the characteristics \( B_1 \)–\( B_3 \) of the core problem to the well known properties of the Lanczos tridiagonalization and the Jacobi matrices. We distinguish two cases described above.

**Case 1.** \( \alpha_i \beta_i \neq 0 \) for \( i = 1, \ldots, p \); \( \beta_{p+1} = 0 \) or \( p = n \) (i.e. \( n \leq m \)), see (2.8) and (2.9). The square matrix \( A_{11} \equiv L_p \) represents a Cholesky factor of \( T^{(1)}_p \equiv L_p L_p^T \), which we see by (2.13)
results from the Lanczos tridiagonalization of $B^{(1)} = AA^T$ with the starting vector $u_1 = b/\|b\|$, which stops exactly in $p$ steps, i.e.

$$AA^T U_p = U_p L_p L_p^T.$$  \hspace{1cm} (3.1)

Consider the singular value decomposition $L_p = R \Sigma S^T$, where $\Sigma = \text{diag}(\sigma_1, \ldots, \sigma_p)$, $R^{-1} = R^T$, $S^{-1} = S^T$. Then

$$T_p^{(1)} = L_p L_p^T = R \Sigma^2 R^T$$

is the spectral decomposition of the matrix $T_p^{(1)}$, $\sigma_i^2$ are its eigenvalues and $r_i \equiv Re_i$ its eigenvectors, $i = 1, \ldots, p$. Consequently, from $J1$ the singular values of $L_p$ are distinct. $L_p$ is square with positive elements on its diagonal. Therefore all its singular values must be positive, which proves $B1$. Moreover $B3$ follows from $J3$, since

$$b^T r_i = \beta_1 e_1^T r_i \neq 0 \text{ for } i = 1, \ldots, p.$$ 

The minimality property $B2$ can be proved by contradiction. For some $\tilde{P}, \tilde{Q}$, let $\tilde{P}^{-1} = \tilde{P}^T$, $\tilde{Q}^{-1} = \tilde{Q}^T$,

$$\tilde{P} \begin{bmatrix} b & A \tilde{Q} \end{bmatrix} = \begin{bmatrix} \tilde{b}_1 & \tilde{A}_{11} & 0 \\ 0 & 0 & \tilde{A}_{22} \end{bmatrix},$$

where $\tilde{A}_{11}$ is a $q \times q$ matrix with $q < p$. (The system (1.2) is compatible, see (2.9), and therefore, for example by considering the QR factorization of $\tilde{A}_{11}$, we can with no loss of generality assume that $\tilde{A}_{11}$ is square.) Substituting

$$A = \tilde{P} \begin{bmatrix} \tilde{A}_{11} & 0 \\ 0 & \tilde{A}_{22} \end{bmatrix} \tilde{Q}^T$$

into the Lanczos tridiagonalization (3.1) gives

$$\tilde{P} \begin{bmatrix} \tilde{A}_{11} & 0 \\ 0 & \tilde{A}_{22} \end{bmatrix} \begin{bmatrix} \tilde{A}_{11}^T & 0 \\ 0 & \tilde{A}_{22}^T \end{bmatrix} \tilde{P}^T U_p = U_p T_p^{(1)},$$

i.e.

$$\begin{bmatrix} \tilde{A}_{11} \tilde{A}_{11}^T & 0 \\ 0 & \tilde{A}_{22} \tilde{A}_{22}^T \end{bmatrix} \tilde{P}^T U_p = (\tilde{P}^T U_p) T_p^{(1)} \hspace{1cm} (3.2)$$

with

$$\tilde{P}^T u_1 = \tilde{P}^T b/\|b\| = \begin{bmatrix} \tilde{b}_1/\|b\| \\ 0 \end{bmatrix}.$$ 

Since $\tilde{A}_{11} \tilde{A}_{11}^T$ is the $q \times q$ matrix and $\tilde{b}_1$ is the vector of length $q$, the Lanczos tridiagonalization represented by (3.2) must stop in at most $q$ steps, and $T_p^{(1)}$ must have $\delta_{q+1}^{(1)} = 0$, which contradicts the fact that $T_p^{(1)}$ is a Jacobi matrix.

**Case 2.** $\alpha_i \beta_i \neq 0$ for $i = 1, \ldots, p$, and $\beta_{p+1} \neq 0$; $\alpha_{p+1} = 0$ or $p = m$ (i.e. $n \geq m$), see (2.10) and (2.11). The rectangular matrix $A_{11} \equiv L_{p+}$ can be linked to the matrix $T_p^{(2)} \equiv L_p^T L_{p+}$, which we see by (2.14) results from the Lanczos tridiagonalization of $B^{(2)} \equiv A^T A$ with the starting vector $v_1 = A^T b/\|A^T b\|$. It stops exactly in $p$ steps, i.e.

$$A^T A V_p = V_p L_p^T L_{p+}.$$  \hspace{1cm} (3.3)
Consider the singular value decomposition $L_{p+} = R \Sigma S^T$, where $R$ is now a rectangular matrix with orthonormal columns, $S^{-1} = S^T$. Then

$$T_p^{(2)} = L_p^T L_{p+} = S \Sigma^2 S^T$$

is the spectral decomposition of the matrix $T_p^{(2)}$, $\sigma_i^2$ are its eigenvalues and $s_i \equiv S e_i$ its eigenvectors, $i = 1, \ldots, p$. Similarly to the previous case, from J1 it follows that the singular values of $L_{p+}$ are distinct. Since by construction $v_1$ does not have any nonzero component in the nullspace of $A^T A$, J2 yields that the singular values of $L_{p+}$ are positive, which proves B1. Moreover, $e_1^T s_i \neq 0$ by J3, $i = 1, \ldots, p$. Considering $L_{p+} S = R \Sigma$ and the fact that $L_{p+}$ is lower bidiagonal with nonzero bidiagonal elements, $e_1^T r_i \neq 0, i = 1, \ldots, p$. Consequently $b_1^T r_i = \beta_1 e_1^T r_i \neq 0, i = 1, \ldots, p$, which proves B3.

The minimality property B2 can be proved by contradiction, similarly to Case 1. For some $\hat{P}$, $\hat{Q}$, $\hat{P}^{-1} = \hat{P}^T$, $\hat{Q}^{-1} = \hat{Q}^T$, let

$$\hat{P}^T \begin{bmatrix} b & \| A \hat{Q} \| \end{bmatrix} = \begin{bmatrix} \hat{b}_1 & \hat{A}_{11} & 0 \\ 0 & 0 & \hat{A}_{22} \end{bmatrix},$$

where $\hat{A}_{11}$ is a $(q + 1) \times q$ matrix with $q < p$. (The system (1.2) is incompatible and therefore we can with no loss of generality assume that $\hat{A}_{11}$ is rectangular of the given dimensions.) Substituting

$$A = \hat{P} \begin{bmatrix} \hat{A}_{11} & 0 \\ 0 & \hat{A}_{22} \end{bmatrix} \hat{Q}^T$$

into the Lanczos tridiagonalization (3.3) gives

$$\begin{bmatrix} \hat{A}_{11}^T \hat{A}_{11} & 0 \\ 0 & \hat{A}_{22}^T \hat{A}_{22} \end{bmatrix} (\hat{Q}^T V_p) = (\hat{Q}^T V_p) T_p^{(2)}$$

(3.4)

with

$$\hat{Q}^T v_1 = \hat{Q}^T A^T b / \| A^T b \| = \begin{bmatrix} \hat{A}_{11}^T & 0 \\ 0 & \hat{A}_{22}^T \end{bmatrix} \hat{P}^T b / \| A^T b \| = \begin{bmatrix} \hat{A}_{11}^T \hat{b}_1 / \| A^T b \| \\ 0 \end{bmatrix},$$

which leads to a contradiction exactly in the same way as in Case 1.

4. Concluding remarks

Core problems within orthogonally invariant linear approximation problems can be computed via the Golub–Kahan bidiagonalization, see [1–3]. We have shown how this fact can be used, together with the well-known relationship with the Lanczos tridiagonalization and the properties of Jacobi matrices, for an alternative derivation of the core problem characteristics given in [3, Theorems 2.2, 3.2, 3.3]. In our paper the Golub–Kahan bidiagonalization and the Lanczos tridiagonalization are used as mathematical tools for constructing proofs. The relationships presented here may be found useful in applications of the core problem formulation.

Acknowledgments

We thank Martin Plešinger and Reinhard Nabben for valuable remarks and discussion, and to Chris Paige for careful reading the text and for several suggestions which significantly improved this manuscript.
References