# Extremal Properties of Balanced Tri-Diagonal Matrices

By Peter A. Businger

**Abstract.** If $A$ is a square matrix with distinct eigenvalues and $D$ a nonsingular matrix, then the angles between row- and column-eigenvectors of $D^{-1}AD$ differ from the corresponding quantities of $A$. Perturbation analysis of the eigenvalue problem motivates the minimization of functions of these angles over the set of diagonal similarity transforms; two such functions which are of particular interest are the spectral and the Euclidean condition numbers of the eigenvector matrix $X$ of $D^{-1}AD$. It is shown that for a tri-diagonal real matrix $A$ both these condition numbers are minimized when $D$ is chosen such that the magnitudes of corresponding sub- and super-diagonal elements are equal. ∎

If a tri-diagonal matrix $A$ is such that corresponding sub- and super-diagonal elements have equal magnitude then $A$ is said to be *balanced* or *equilibrated*. Wilkinson [5, p. 424] uses norms of balanced tri-diagonal matrices for error analysis of the eigenvalue problem. He observes that, given a tri-diagonal matrix $A = [a_{ij}]$ all of whose sub- and super-diagonal elements are nonzero, a diagonal matrix $D = \text{diag}(d_1, d_2, \cdots, d_n)$ can be found such that $D^{-1}AD$ is balanced. In fact, such a $D$ is defined by

$$d_{i+1}/d_i = (|a_{i+1,i}|/|a_{i,i+1}|)^{1/2}, \qquad i = 1, 2, \cdots, n-1 .$$

If some sub- or super-diagonal element of $A$ is zero then finding its eigenvalues can be reduced to finding the eigenvalues of submatrices, each of which can be balanced separately.

It is an immediate consequence of Osborne's Lemma 2 [3] that a balanced tri-diagonal matrix $A$ has the extremal property

$$\|A\|_E = \inf_D \|D^{-1}AD\|_E ,$$

where $\|\cdot\|_E$ denotes the *Euclidean matrix norm* (*Schur norm, Frobenius norm*). Our Theorem 1 states the analogous result for the spectral norm; Theorems 2 and 3 show that the eigenvalue problem of a balanced tri-diagonal matrix is optimally conditioned in the sense that no matrix of the form $D^{-1}AD$ has smaller angles between corresponding row- and column-eigenvectors.

We use $\|\cdot\|$ to denote the *Euclidean vector norm,* $\|\cdot\|_2$ for the subordinate matrix bound (the *spectral matrix norm*), $k_2(\cdot)$ for the *spectral condition number* of a nonsingular matrix, and $k_E(\cdot)$ for the *Euclidean condition number* (defined by $k_E(X) = \|X\|_E \, \|X^{-1}\|_E$). Absolute value signs applied to vectors are understood componentwise. $D$, $D_1$, and $D_2$ denote diagonal matrices with positive diagonal elements.

Theorem 1. *If $A$ is a balanced tri-diagonal real matrix then*

$$\|A\|_2 = \inf_D \|D^{-1}AD\|_2 .$$

Received May 9, 1968.

*Proof.* There exists a real diagonal matrix $E$ with $|E| = I$ such that $B = EA$ is symmetric. Since, for all $D$,

$$\|D^{-1}AD\|_2 = \|ED^{-1}AD\|_2 = \|D^{-1}EAD\|_2 = \|D^{-1}BD\|_2,$$

the conclusion follows from the observation that for a symmetric matrix $B$, $\|B\|_2 \leqq \|D^{-1}BD\|_2$ for all $D$.

The following theorem deals with the secants $\|y^H\|\ \|x\|/|y^Hx|$ of the angles between corresponding row- and column-eigenvectors of a matrix.

THEOREM 2. *If $A$ is a balanced tri-diagonal real matrix with distinct eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$, corresponding column-eigenvectors $x_1, x_2, \cdots, x_n$, and corresponding row-eigenvectors $y_1^H, y_2^H, \cdots, y_n^H$ then*

$$\frac{\|y_i^H\|\ \|x_i\|}{|y_i^Hx_i|} = \inf_D \frac{\|y_i^HD\|\ \|D^{-1}x_i\|}{|y_i^Hx_i|}$$

*for $i = 1, 2, \cdots, n$.*

*Proof.* Basing his argument on a theorem due to Stoer and Witzgall [4], Bauer [1] showed that for any vector pair $y^H$ and $x$,

$$\inf_D \frac{\|y^HD\|\ \|D^{-1}x\|}{|y^Hx|} = \frac{|y^H|\ |x|}{|y^Hx|}.$$

Since $A = EA^TE$ for some real diagonal matrix $E$ with $|E| = I$, $y_i = c_iEx_i$ for some scalars $c_i$. Hence

$$\frac{\|y_i^H\|\ \|x_i\|}{|y_i^Hx_i|} = \frac{|y_i^H|\ |x_i|}{|y_i^Hx_i|}$$

for $i = 1, 2, \cdots, n$, which completes the proof.

COROLLARY. *$A$ has an eigenvector matrix $X = [x_1, x_2, \cdots, x_n]$ such that $k_E(X) = \inf_{D_1,D_2} k_E(D_1^{-1}XD_2)$.*

*Proof.* By Theorem 2, each term in the sum on the right of the relationship

$$\inf_{D_1,D_2} k_E(D_1^{-1}XD_2) = \inf_{D_1} \sum_{i=1}^{n} \frac{\|y_i^HD_1\|\ \|D_1^{-1}x_i\|}{|y_i^Hx_i|}$$

is minimized when $D_1 = I$. This implies the corollary.

THEOREM 3. *If $A$ is a balanced tri-diagonal real matrix with distinct eigenvalues then $A$ has an eigenvector matrix $X = [x_1, x_2, \cdots, x_n]$ such that $k_2(X) = \inf_{D_1,D_2} k_2(D_1^{-1}XD_2)$.*

*Proof.* Bauer [2] showed that

$$\inf_{D_1,D_2} k_2(D_1^{-1}XD_2) \geqq \rho(E_1X^{-1}E_2X)$$

for all diagonal matrices $E_1$ and $E_2$ for which $|E_1| = |E_2| = I$ ($\rho$ denotes the spectral radius). Hence it suffices for us to obtain equality for some eigenvector matrix $X$ of $A$ and for some such $E_1$ and $E_2$.

Let $Q$ be a unitary matrix such that if $Z = XQ$ then $J = Z^{-1}AZ$ is the direct sum of 1 by 1 and 2 by 2 matrices. (The latter are of the form

$$\begin{bmatrix} \lambda & \mu \\ -\mu & \lambda \end{bmatrix}$$

and correspond to conjugate complex pairs of eigenvalues $\lambda \pm i\mu$.) If the permutation matrix $P$ is chosen such that $\overline{X} = XP$, invariance of $k_2$ implies that for all $D_1$ and $D_2$

$$\begin{aligned} k(D_1^{-1}XD_2) &= k\overline{(D_1^{-1}XD_2)} = k(D_1^{-1}\overline{X}D_2) \\ &= k(D_1^{-1}XPD_2) = k(D_1^{-1}X(PD_2P^T)) \,. \end{aligned}$$

Hence no generality is lost if we assume that those pairs of diagonal elements of $D_2$ are equal which correspond to a complex conjugate pair of eigenvectors. Under this assumption

$$k(D_1^{-1}XD_2) = k(D_1^{-1}XD_2Q) = k(D_1^{-1}XQD_2) \,,$$

which allows us to replace the problem of minimizing $k(D_1^{-1}XD_2)$ by that of finding $\inf_{D_1,D_2} k(D_1^{-1}ZD_2)$. Now $Z^{-1}AZ = J$ implies

$$Z^TA^TZ^{-T} = J^T = E_1JE_1$$

for some real diagonal matrix $E_1$ such that $|E_1| = I$. Hence, if $A^T = E_2AE_2$, it follows that $E_2Z^{-T}E_1 = ZD_2$ for some diagonal matrix $D_2$. Thus there exists a matrix $Z_0$ such that $Z_0^{-1}AZ_0 = J$ as well as $Z_0^{-1} = E_1Z_0^TE_2$. Hence

$$k(Z_0) = \|Z_0\|_2\|E_1Z_0^TE_2\|_2 = \|Z_0\|_2^2 = \rho(Z_0^TZ_0) = \rho(E_1Z_0^{-1}E_2Z_0) \,.$$

The result of Bauer stated at the beginning of this proof now establishes the theorem.

Bell Telephone Laboratories, Incorporated
Murray Hill, New Jersey 07974

1. F. L. Bauer, "Some aspects of scaling invariance," *Colloq. Internat. C.N.R.S.*, No. 165, pp. 37–47.
2. F. L. Bauer, "Optimally scaled matrices," *Numer. Math.*, v. 5, 1963, pp. 73–87. MR **28** #2629.
3. E. E. Osborne, "On pre-conditioning of matrices," *J. Assoc. Comput. Mach.*, v. 7, 1960, pp. 338–345. MR **26** #892.
4. J. Stoer & Ch. Witzgall, "Transformations by diagonal matrices in a normed space," *Numer. Math.*, v. 4, 1962, pp. 158–171. MR **27** #154.
5. J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965. MR **32** #1894.