THE MATRIX INVERSE EIGENVALUE PROBLEM FOR PERIODIC JACOBI MATRICES

by

D. L. Boley and G. H. Golub

STAN-CS-78-684
DECEMBER 1978

COMPUTER SCIENCE DEPARTMENT
School of Humanities and Sciences
STANFORD UNIVERSITY

SU326 P30-63

THE MATRIX INVERSE EIGEN-VALUE PROBLEM

FOR PERIODIC JACOBI MATRICES

D. L. Boley and G. H. Golub

Computer Science Department
Stanford University
Stanford, CA 94305
(415) 497-3124.

This work was supported in part by Department of Energy contract EY-76-S-03-0326 PA #30 and by National Science Foundation grant MCS75-13497-A01.

Invited paper presented at The Fourth Conference on Basic Problems of Numerical Analysis, (LIBLICE IV), Pilsen, Czech., September 1978.

Abstract

A stable numerical algorithm is presented for generating a periodic Jacobi matrix from two sets of eigenvalues and the product of the off-diagonal elements of the matrix. The algorithm requires a simple generalization of the Lanczos algorithm. It is shown that the matrix is not unique, but the algorithm will generate all possible solutions.

THE MATRIX INVERSE EIGENVALUE PROBLEM FOR PERIODIC JACOBI MATRICES

D.L. Boley and G.H. Golub
Computer Science Department, Stanford University,
Stanford, CA 94305

We are interested in solving the inverse eigenvalue problem for periodic matrices of the form:

$$J = \begin{bmatrix} \alpha_1 & \beta_1 & & & & \beta_n \\ \beta_1 & \alpha_2 & \beta_2 & & 0 & \\ & \beta_2 & \cdot & \cdot & & \\ & & \cdot & \cdot & & \\ & 0 & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & \beta_{n-1} \\ \beta_n & & & & \beta_{n-1} & \alpha_n \end{bmatrix}.$$

Such problems arise in inverse scattering theory problems (cf. [3]). The problem and the algorithm given here are closely related to that discussed in our previous paper [2] and that of Van Moerbeke [3]. For this problem, we consider the matrices

$$\text{(1a)} \qquad J^+ = \begin{bmatrix} \alpha_1 & \beta_1 & & \beta_n \\ \beta_1 & \alpha_2 & \cdot & 0 \\ & \cdot & \cdot & \cdot \\ 0 & & \cdot & \beta_{n-1} \\ \beta_n & & \beta_{n-1} & \alpha_n \end{bmatrix} \equiv \begin{bmatrix} \alpha_1 & (\tilde{b}^+)^T \\ \tilde{b}^+ & K \end{bmatrix}$$

and

(1b)
$$J^- = \begin{bmatrix} \alpha_1 & \beta_1 & & & -\beta_n \\ \beta_1 & \alpha_2 & \cdot & & \\ & \beta_2 & \cdot & & \\ & & & \cdot & \beta_{n-1} \\ -\beta_n & & & \beta_{n-1} & \alpha_n \end{bmatrix} = \begin{bmatrix} \alpha_1 & (\tilde{b}^-)^T \\ \tilde{b}^- & K \end{bmatrix}$$

where $\tilde{b}^+$ and b^- are $(n-1)$-vectors; K is an $(n-1)\times(n-1)$ matrix, and α_1 is a scalar. We assume three sets of eigenvalues are given: we denote the eigenvalues of J^+ by $\lambda_1^+ < \ldots < \lambda_n^+$, those of J^- by $\lambda_1^- < \lambda_2^- < \ldots < \lambda_n^-$, and those of K by $\mu_1 < \cdots < \mu_{n-1}$. We will use the notation $\Lambda^+ = \mathrm{diag}\,(\lambda_1^+, \lambda_2^+, \ldots, \lambda_n^+)$; $\Lambda^- = \mathrm{diag}\,(\lambda_1^-, \ldots, \lambda_n^-)$ and $M = \mathrm{diag}(\mu_1, \mu_2, \ldots, \mu_{n-1})$. We will also assume that the $\{\mu_i\}_{i=1}^{n-1}$ <u>strictly</u> interlace both sets of eigenvalues $\{\lambda_i^+\}_{i=1}^{n}$ and $\{\lambda_i^-\}_{i=1}^{n-1}$, <u>viz</u>

$$\begin{aligned} \lambda_i^+ < \mu_i < \lambda_{i+1}^+ \\ \lambda_i^- < \mu_i < \lambda_{i+1}^- \end{aligned} \qquad i = 1,2,..,n-1 .$$

We can show that there is a close relationship between the characteristic polynomials of J^+ and J^-, and thus it is sufficient to have given the single scalar quantity

$$\beta^* = \beta_1 \beta_2 \cdots \beta_n$$

in place of the n eigenvalues $\lambda_1^-, \cdots, \lambda_n^-$. Using equation (5.2) in [1], we can write

$$\det\,(J^+ - \lambda I) = p(\lambda) - \beta_n^2\, r(\lambda) + 2(-1)^{n-1} \beta^*$$

and

$$\det (J^- - \lambda I) = p(\lambda) - \beta_n^2 r(\lambda) - 2(-1)^{n-1}\beta^*$$

where $p(\lambda)$ is the characteristic polynomial of the matrix obtained from J^+ or J^- by setting $\beta_n = 0$, and $r(\lambda)$ is the characteristic polynomial of the submatrix consisting of rows and columns $2,3,\ldots,n-1$ of the matrix J^+ or J^-. By subtracting these two expressions we obtain

$$(2) \qquad \det (J^- - \lambda I) = \det (J^+ - \lambda I) + 4(-1)^n \beta^* .$$

We will see the values λ_i^- appear only as the product $(\lambda_1^- - \lambda)\ldots(\lambda_n^- - \lambda)$, so that if we are given the values $\lambda_1^+,\ldots,\lambda_n^+$ as well as β^*, then we do not need explicitly the values $\lambda_1^-,\ldots,\lambda_n^-$.

Note that the roles of λ_i^+ and λ_i^- are essentially interchangeable at several stages, and we have made the choice to use λ_i^+ as much as possible.

We let $Q = [\underset{\sim}{q}_1,\ldots,\underset{\sim}{q}_n]$ be the matrix of eigenvectors of J^+ so that

$$\Lambda^+ = Q^T J^+ Q .$$

It is useful to write Q in terms of rows as well:

$$(3a) \qquad Q \equiv \begin{bmatrix} \underset{\sim}{r}_1^T \\ \cdot \\ \underset{\sim}{r}_n^T \end{bmatrix}$$

We let P be the matrix of eigenvectors of K , and write P only in terms of its rows:

(3b) $$P = \begin{bmatrix} \underset{\sim}{p}_1^T \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \underset{\sim}{p}_{n-1}^T \end{bmatrix}$$

In section 1 of [2] it is shown how we may compute the first row $\underset{\sim}{r}_1^T$ of Q from $\{\lambda_i^+\}_{i=1}^n$ and $\{\mu_i\}_{i=1}^{n-1}$. The final result obtained there is

(4) $$q_{1j}^2 = \frac{\prod_{k=1}^{n-1} (\mu_k - \lambda_j^+)}{\prod_{\substack{k=1 \\ k\neq j}}^{n} (\lambda_k^+ - \lambda_j^+)} \qquad (j = 1,\ldots,n) .$$

We can pick the signs of the q_{1j} arbitrarily, since changing one sign is equivalent to just multiplying the corresponding eigenvector by -1 .

Consider the matrix

$$L^+ \equiv \begin{bmatrix} \alpha_1 & (\hat{\underset{\sim}{b}}^+)^T \\ \hat{b}^+ & M \end{bmatrix} = \begin{bmatrix} 1 & \underset{\sim}{0} \\ 0 & P^T \end{bmatrix} J^+ \begin{bmatrix} 1 & 0^T \\ 0 & P \end{bmatrix}$$

where we partition J^+ as in (1a), $\hat{\underset{\sim}{b}}^+ = P^T b^+$, and M is the matrix of eigenvalues of K . We define L- similarly': Since the eigenvalues of

L^+ are $\{\lambda_1^+,\ldots,\lambda_n^+\}$, we have that

$$\det(L^+ - \lambda I) = \prod_{k=1}^{n} (\lambda_k^+ - \lambda) .$$

If we evaluate the determinant and equate the two sides of the equation, we obtain the result:

$$(5a) \qquad (\hat{b}_k^+)^2 = - \frac{\prod_{j=1}^{n} (\lambda_j^+ - \mu_k)}{\prod_{\substack{j=1 \\ j\neq k}}^{n=1} (\mu_j - \mu_k)} \qquad (k=1,\ldots,n-1).$$

Similarly, we obtain

$$(5b) \qquad (\hat{b}_k^-)^2 = - \frac{\prod_{j=1}^{n} (\lambda_j^- - \mu_k)}{\prod_{\substack{j=1 \\ j\neq k}}^{n-1} (\mu_j - \mu_k)} \qquad (k=1,\ldots,n-1),$$

but using equation (2) we can rewrite (5b) as:

$$(5c) \qquad (\hat{b}_k^-)^2 = - \frac{\prod_{j=1}^{n} (\lambda_j^+ - \mu_k) + 4(-1)^n \beta^*}{\prod_{\substack{j=1 \\ j\neq k}}^{n-1} (\mu_j - \mu_k)}$$

Note that the signs of $\hat{b}_k^+$ and $\hat{b}_k^-$ are unspecified. Any combination of signs will give a valid matrix, and using a different set of signs may or may not give a different matrix. We see that $\underset{\sim}{b}^+ = [\beta_1, 0, \ldots, 0, +\beta_n]^T$ and

$\underset{\sim}{b}^- = [\beta_1,0,\ldots,0,-\beta_n]^T$ from (1), so that

$$\underset{\sim}{b}^+ - b^- = 2\,\beta_n\,\underset{\sim}{e}_{n-1}$$

where $\underset{\sim}{e}_{n-1} = [0,\ldots,0,1]^T$. Note that all the vectors in the above equation are (n-1)- vectors. But, using the definition of L, we may write

$$(6) \qquad \hat{\underset{\sim}{b}}^+ - \hat{\underset{\sim}{b}}^- = P^T(\underset{\sim}{b}^+ - \underset{\sim}{b}^-) = 2\,\beta_n\,P^T\underset{\sim}{e}_{n-1} = 2\,\beta_n\,\underset{\sim}{p}_{n-1}\ .$$

Since $\underset{\sim}{p}_n$ is of norm unity- we need not know β_n explicitly; we just have to scale the vector $\hat{b}^+ - \hat{\underset{\sim}{b}}^-$ to be of norm unity.

It is useful to partition the eigenvector $\underset{\sim}{q}_k$ of J^+ as follows

$$\underset{\sim}{q}_k = \begin{bmatrix} x_k \\ \underset{\sim}{y}_k \end{bmatrix}\ .$$

where $x_k = q_{1k}$ is a scalar, and $\underset{\sim}{y}_k$ is an (n-1)- vector. Since $J^+\underset{\sim}{q}_k = \lambda_k^+\underset{\sim}{q}_k$, we have

$$\underset{\sim}{b}^+ x_k + K\underset{\sim}{y}_k = \lambda_k^+\underset{\sim}{y}_k$$

or

$$(K - \lambda_k^+ I)\underset{\sim}{y}_k = \underset{\sim}{b}^+ x_k\ .$$

Using the decomposition $K = PMP^T$, we obtain

$$y_k = -P(M - \lambda_k^+ I)^{-1} P^T b + x_k$$

$$= -P(M - \lambda_k^+ I)^{-1} \hat{b}^+ x_k$$

Thus, the last row of Q is

$$q_{n,k} \equiv y_{n-1,k} = -q_{1k} \sum_{j=1}^{n-1} \frac{p_{n-1,j}\, \hat{b}_j^+}{(\mu_j - \lambda_k^+)} \tag{7}$$

where all the quantities on the right hand side are either given or are computed using (4), (5a) and (6).

Now we are in a position to generate the matrix using a modified Lanczos process. We need two initial vectors:

$$z_1 = r_1 \qquad \text{(computed using (4))}$$

and

$$z_n = r_n \qquad \text{(computed using (7))},$$

using the notation in (3a). We have for these two vectors

$$\|z_1\|_2 = \|z_n\|_2 = 1 \quad \text{and} \quad z_1^T z_n = 0 \; .$$

Note that we have obtained two initial vectors for the Lanczos process in a manner very similar to that used in our previous paper for the five diagonal case [2]. The Lanczos process itself is a little different but is based on the same ideas.

To derive the Algorithm we use the relationships

$$Z^T A Z = J^+ \quad \text{and} \quad A Z = Z J^+$$

where $Z = [z_1, z_2, \ldots, z_n]$ is an orthogonal matrix. We then arrive at the following procedure, using the identification

$$A = \Lambda^{+} = \text{diag}\,(\lambda_1^{+}, \lambda_2^{+}, \ldots, \lambda_n^{+}) .$$

(8) Modified Lanczos Procedure.

1. Set $\beta_0 = \beta_n = z_n^T A z_1$

$$\alpha_n = z_n^T A z_n$$

and

set $z_0 = z_n$.

2. For $k = 1,2,\ldots,n-1$

Set $\alpha_k = z_k^T A z_k$

$$v_k = Az_k - \alpha_k z_k - \beta_{k-1} z_{k-1}$$

$$\beta_k = \|v_k\|_2$$

If $k \neq n-1$ then set $z_{k+1} = (\beta_k)^{-1} v_k$.

(For numerical stability only);

Orthogonalize z_{k+1} with respect to $z_0,\ldots,z_k$ using, say, the modified Gram-Schmidt procedure or Householder transformations.

The matrix J^{+} and J^{-} will be defined by (1) using the $\alpha_1,\alpha_2,\ldots,\alpha_n$, $\beta_1,\beta_2,\ldots,\beta_n$ computed in the Lanczos procedure. There is an ambiguity in the signs of the $\{\beta_i\}_{i=1}^{n}$ in that we can switch the sign of β_i while changing the sign of z_i without affecting the result of the procedure. Because an even number of sign changes in the $\{\beta_i\}_{i=1}^{n}$ in J^{+} will not affect the eigenvalues (see Appendix II), we may assume that $\beta_1,\beta_2,\ldots,\beta_{n-1}$ are all positive and set the sign of β_n to match that of β^{*}, adjusting the vector $\underset{\sim}{z}_n$ accordingly.

To summarize the Algorithm:

We are given two sets of eigenvalues

$$\{\lambda_i^+\}_{i=1}^n \quad , \quad \{\mu_i\}_{i=1}^{n-1}$$

and the scalar β^* .

1. Compute $\underset{\sim}{z}_1 = \underset{\sim}{r}_1 = [q_{11}, q_{12}, \ldots, q_{1n}]^T$ by (4) .

2. Compute $\hat{\underset{\sim}{b}}^+$ and $\hat{b}^-$ by (5) .

3. Compute $\underset{\sim}{p}_{n\ 1} = [p_{n-1,1}, \ldots, p_{n-1,n-1}]^T$ by (6).

4. Compute $\underset{\sim}{z}_n = \underset{\sim}{r}_n = [q_{n1}, \ldots, q_{nn}]^T$ by (7).

5. Apply the modified Lanczos procedure (8) using Λ^+ and initial vectors $\underset{\sim}{z}_1$ and $\underset{\sim}{z}_n$.

Acknowledgement.

The authors are very pleased to thank Dr. Warren Ferguson of the University of Arizona for several interesting and enlightening discussions. The work of G.H. Golub was initiated while visiting Professor H. Keller of the California Institute of Technology.

References.

[1] A. Björck and G.H. Golub: Eigenproblems for matrices associated with periodic boundary conditions. SIAM Review 19 (1977), 5-16.

[2] D. Boley and G.H. Golub: Inverse eigenvalue problems for band matrices. Proceedings of the Dundee Conference on Numerical Analysis, Springer-Verlag (1977).

[3] P. van Moerbeke: The spectrum of Jacobi matrices. Invent. Math. 37 (1976), 45-81.

Appendix I.

It is possible to show that the Lanczos process described by (8), if it reaches completion (with $\beta_k \neq 0$ for all k) will generate orthonormal vectors for any real symmetric A and thus the matrix J will be a periodic matrix. For notational convenience, we write $\underset{\sim}{z}_n = \underset{\sim}{z}_0$. We assume that $\underset{\sim}{z}_0$ and $\underset{\sim}{z}_1$ are given with $\|\underset{\sim}{z}_0\|_2 = \|\underset{\sim}{z}_1\|_2 = 1$ and that

$$\underset{\sim}{z}_0^T \underset{\sim}{z}_1 = 0 .$$

To generate the periodic matrix, we use the. recurrence relationship

$$(9) \qquad \beta_{j-1} \underset{\sim}{z}_j = A\underset{\sim}{z}_{j-1} - \alpha_{j-1} \underset{\sim}{z}_{j-1} - \beta_{j-2} \underset{\sim}{z}_{j-2}$$

for $j = 2,3,\ldots,n-1$.

Let

$$\beta_0 = \beta_n = \underset{\sim}{z}_0^T \underset{\sim}{A} z_1$$

and

$$\alpha_1 = \underset{\sim}{z}_1^T A \underset{\sim}{z}_1 .$$

Note this immediately implies

$$\beta_1 \underset{\sim}{z}_0^T \underset{\sim}{z}_2 = 0 \quad \text{and} \quad \beta_1 \underset{\sim}{z}_1^T \underset{\sim}{z}_2 = 0.$$

The parameter β_1 is computed so that $\|\underset{\sim}{z}_2\|_2 = 1$. Let us assume

$$(10a) \quad \underset{\sim}{z}_i^T \underset{\sim}{z}_j = 0 \qquad \text{for } i \neq j ,\ i,j = 0,1,\ldots,k$$

and

$$(10b) \quad \|\underset{\sim}{z}_i\|_2 = 1 \qquad \text{for } i = 0,1,\ldots,k .$$

We calculate

(11a) $$\alpha_k = \mathbf{z}_k^T A \mathbf{z}_k$$

(11b) $$\mathbf{v}_{k+1} = A \mathbf{z}_k - \alpha_k \mathbf{z}_k - \beta_{k-1} \mathbf{z}_{k-1}$$

(11c) $$\beta_k = \|\mathbf{v}_{k+1}\|_2$$

(11d) $$\mathbf{z}_{k+1} = (\beta_k)^{-1} \mathbf{v}_{k+1} \quad .$$

Note that $\|\mathbf{z}_{k+1}\|_2 = 1$ providing $\|\mathbf{v}_{k+1}\|_2 \neq 0$.

We now show

$$\mathbf{z}_{k+1}^T \mathbf{z}_j = 0 \qquad \text{for } j < k \quad .$$

Since $\mathbf{z}_k^T \mathbf{z}_{k-1} = 0$ and $\|\mathbf{z}_k\|_2 = 1$,

$$\mathbf{z}_{k+1}^T \mathbf{z}_k = 0$$

when α_k is calculated by (11a). Now

(12) $$\beta_k \mathbf{z}_j^T \mathbf{z}_{k+1} = \mathbf{z}_j^T A \mathbf{z}_k - \alpha_k \mathbf{z}_j^T \mathbf{z}_k - \beta_{k-1} \mathbf{z}_j^T \mathbf{z}_{k-1}$$

so that for $j < k-1$

(13) $$\beta_k \mathbf{z}_j^T \mathbf{z}_{k+1} = \mathbf{z}_j^T A \mathbf{z}_k - \beta_{k-1} \mathbf{z}_j^T \mathbf{z}_{k-1}$$

But $\mathbf{z}_j^T A \mathbf{z}_k = \mathbf{z}_k^T A \mathbf{z}_j$ so that (13) becomes, using (9)

$$\beta_k \mathbf{z}_j^T \mathbf{z}_{k+1} = \mathbf{z}_k^T(\beta_j \mathbf{z}_{j+1} + \alpha_j \mathbf{z}_j + \beta_{j-1} \mathbf{z}_{j-1}) - \beta_{k-1} \mathbf{z}_j^T \mathbf{z}_{k-1}$$

$$= \beta_j \mathbf{z}_k^T \mathbf{z}_{j+1} + 0 \cdot \alpha_j + 0 \cdot \beta_{j-1} - \beta_{k-1} \mathbf{z}_j^T \mathbf{z}_{k-1} \ .$$

Therefore

$$\beta_k \underset{\sim}{z}_{k-1}^T \underset{\sim}{z}_{k+1} = \beta_{k-1} \underset{\sim}{z}_k^T \underset{\sim}{z}_k - \beta_{k-1} \underset{\sim}{z}_{k-1}^T \underset{\sim}{z}_{k-1} = 0$$

and for $j \leq k-2$,

$$\beta_k \underset{\sim}{z}_j^T \underset{\sim}{z}_{k+1} = 0 \cdot \beta_j - 0 \cdot \beta_{k-1} = 0 .$$

<u>Appendix II.</u>

We include here a short discussion of the signs of $\beta_i (i=1,2,\ldots,n)$. We wish to find what sign changes can be made that will leave the eigenvalues unaffected. In order to do this, we will try to see what sign changes can be made by similiarity transformations.

Define the matrix

$$V_{ij} = \text{diag}\,\{v_1, v_2, \ldots, v_n\} , \qquad 1 \leq i \leq j \leq n ,$$

where

$$v_1 = v_2 = \ldots = v_i = 1 ,$$

$$v_{i+1} = \ldots = v_j = -1 ,$$

and

$$v_{j+1} = \ldots = v_n = 1 .$$

Then $V_{ij} = V_{ij}^{-1}$ for any i and j .

Consider the matrix

$$\tilde{J} = V_{ij}^{-1} J^+ V_{ij} ,$$

We can see that $\tilde{J}$ will be identical to J^+ except that β_i and β_j are replaced by $-\beta_i$ and $-\beta_j$; however, $\beta^+ = \prod_{i=1}^{n} \beta_i$ is unchanged. Thus, we may toggle the signs on any two β_k without affecting the eigenvalues. It follows then, that making any even number of sign change will leave the eigenvalues unaffected.

If we make an odd number of sign changes then by using the above argument we can see that we get a matrix with the same eigenvalues as J^-. Thus by toggling the signs of the β_i we can get only two sets of eigenvalues, those of J^+ called $\{\lambda_i^+\}_{i=1}^n$ and those of J^- called $\{\lambda_i^-\}_{i=1}^n$. Whichever set we get is determined by how we set the sign of β^*.