Lecture 10: Jacobi method, Bisection method, Divide-and-conquer method

School of Mathematical Sciences, Xiamen University

1. Jacobi method

- The method is based on the fact that a 2×2 real symmetric matrix $\mathbf{A}$ can be diagonalized in the form

$$\mathbf{J}_2 = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}, \quad \mathbf{J}_2^\top \begin{bmatrix} a & d \\ d & b \end{bmatrix} \mathbf{J}_2 = \begin{bmatrix} \times & 0 \\ 0 & \times \end{bmatrix},$$

  where θ satisfies $(b-a)\sin(2\theta) = 2d\cos(2\theta)$.

- Define the $m \times m$ Jacobi rotation matrix $\mathbf{J}(i,j;\theta)$, $i<j$,

$$\mathbf{J} = \mathbf{J}(i,j;\theta) := \mathbf{I} + \begin{bmatrix} \mathbf{e}_i & \mathbf{e}_j \end{bmatrix} \begin{bmatrix} \cos\theta - 1 & \sin\theta \\ -\sin\theta & \cos\theta - 1 \end{bmatrix} \begin{bmatrix} \mathbf{e}_i^\top \\ \mathbf{e}_j^\top \end{bmatrix}$$

$$= \begin{array}{c} \begin{bmatrix} \mathbf{I} & & & & \\ & \cos\theta & & \sin\theta & \\ & & \mathbf{I} & & \\ & -\sin\theta & & \cos\theta & \\ & & & & \mathbf{I} \end{bmatrix} \end{array} \begin{array}{l} \\ (\text{row } i) \\ \\ (\text{row } j) \\ \\ \end{array}$$

  The (i,j) and (j,i) entries of $\mathbf{B} := \mathbf{J}^\top \mathbf{A} \mathbf{J}$ are zeros via appropriate θ.

Remark 1

The Jacobi rotation matrix $\mathbf{J}$ *is orthogonal.*

Remark 2

We have $\|\mathbf{B}\|_{\mathrm{F}} = \|\mathbf{A}\|_{\mathrm{F}}$ ($\because$ $\|\cdot\|_{\mathrm{F}}$ *is invariant under orthogonal* $\mathbf{J}$).

Theorem 3

Suppose that $a_{ij} = a_{ji} \neq 0$ *and* $i < j$. *Let* $\mathbf{J}(i, j; \theta)$ *be the Jacobi rotation matrix such that the* (i, j) *and* (j, i) *entries of* $\mathbf{B} = \mathbf{J}^\top \mathbf{A} \mathbf{J}$ *are zeros. Then, for* $k \neq i, j$, $b_{kk} = a_{kk}$ *and*

$$b_{ii}^2 + b_{jj}^2 = a_{ii}^2 + a_{jj}^2 + a_{ij}^2 + a_{ji}^2.$$

Remark 4 (Jacobi mehtod: $\mathbf{A}^{(k)} = \mathbf{J}(i_k, j_k; \theta_k)^\top \mathbf{A}^{(k-1)} \mathbf{J}(i_k, j_k; \theta_k)$)

At each step a symmetric pair of zeros is introduced into the matrix (note that previous zeros maybe destroyed). The usual effect is that the sum of the squares of magnitude of off-diagonal entries shrink steadily.

2. Bisection method

- Consider an unreduced (all of its $(i+1,i)$ and $(i,i+1)$ entries are nonzero) tridiagonal real symmetric matrix $\mathbf{A}$:

$$\mathbf{A} = \begin{bmatrix} a_1 & b_1 & & & \\ b_1 & a_2 & b_2 & & \\ & b_2 & a_3 & \ddots & \\ & & \ddots & \ddots & b_{m-1} \\ & & & b_{m-1} & a_m \end{bmatrix}, \quad b_j \neq 0.$$

- Let $\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(m)}$ denote its leading square principal submatrices of dimension $1, \cdots, m$.

Proposition 5

For $k = 1, 2, \ldots, m$, the eigenvalues of $\mathbf{A}^{(k)}$ are distinct:

$$\lambda_1^{(k)} > \lambda_2^{(k)} > \cdots > \lambda_k^{(k)}.$$

Proposition 6

The eigenvalues of $\mathbf{A}^{(1)}, \cdots, \mathbf{A}^{(m)}$ *strictly interlace, i.e.,*

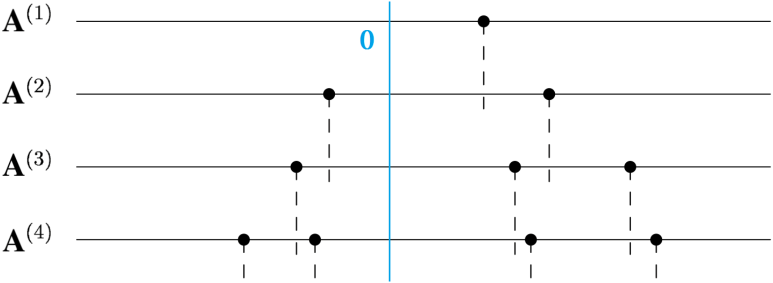
$$\lambda_j^{(k+1)} > \lambda_j^{(k)} > \lambda_{j+1}^{(k+1)},$$

for $k = 1, 2, \ldots, m-1$ *and* $j = 1, 2, \ldots, k$.

Proof: See Golub and van Loan's book: Theorem 8.4.1, Page 468, Matrix computations, 4th edition.

- The interlacing property makes it possible to count the exact number of negative eigenvalues of a real symmetric tridiagonal matrix. For example,

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & & \\ 1 & 0 & 1 & \\ & 1 & 2 & 1 \\ & & 1 & -1 \end{bmatrix}, \qquad \begin{aligned} \det(\mathbf{A}^{(1)}) &= 1, \\ \det(\mathbf{A}^{(2)}) &= -1, \\ \det(\mathbf{A}^{(3)}) &= -3, \\ \det(\mathbf{A}^{(4)}) &= 4. \end{aligned}$$



Remark 7

In general, for any unreduced tridiagonal real symmetric $\mathbf{A}$, *the number of negative eigenvalues is equal to the number of sign changes in the sequence*

$$1,\ \det(\mathbf{A}^{(1)}),\ \det(\mathbf{A}^{(2)}),\ \cdots,\ \det(\mathbf{A}^{(m)}),$$

which is known as a Sturm sequence. Here, we define a "sign change" to mean a transition from $+$ *or* 0 *to* $-$ *or from* $-$ *or* 0 *to* $+$ *but not from* $+$ *or* $-$ *to* 0.

Remark 8

By shifting $\mathbf{A}$ *by a multiple of the identity, we can determine the number of eigenvalues in any interval* $[a, b)$*: it is the number of eigenvalues in* $(-\infty, b)$ *minus the number in* $(-\infty, a)$*; i.e., we only need consider two matrices* $\mathbf{A} - b\mathbf{I}$ *and* $\mathbf{A} - a\mathbf{I}$*.*

Remark 9

The determinants of the matrices $\{\mathbf{A}^{(k)}\}$ *are related by a three-term recurrence relation:*

$$\det(\mathbf{A}^{(k)}) = a_k \det(\mathbf{A}^{(k-1)}) - b_{k-1}^2 \det(\mathbf{A}^{(k-2)}).$$

Introducing the shift by $z\mathbf{I}$ *and writing* $p^{(k)}(z) = \det(\mathbf{A}^{(k)} - z\mathbf{I})$*, we get*

$$p^{(k)}(z) = (a_k - z)p^{(k-1)}(z) - b_{k-1}^2 p^{(k-2)}(z),$$

where $p^{(-1)}(z) = 0,\ p^{(0)}(z) = 1.$

3. Secular equation

Proposition 10

Let $\mathbf{D} \in \mathbb{R}^{m\times m}$ be a diagonal matrix with distinct diagonal entries $\{d_j\}$ and $\mathbf{w} \in \mathbb{R}^m$ be a vector with $w_j \neq 0$ for all j. Assume $\beta \in \mathbb{R}$ and $\beta \neq 0$. The eigenvalues of $\mathbf{D} + \beta\mathbf{w}\mathbf{w}^\top$ are the roots of the rational function

$$f(\lambda) = 1 + \beta \sum_{j=1}^{m} \frac{w_j^2}{d_j - \lambda}.$$

Proof.

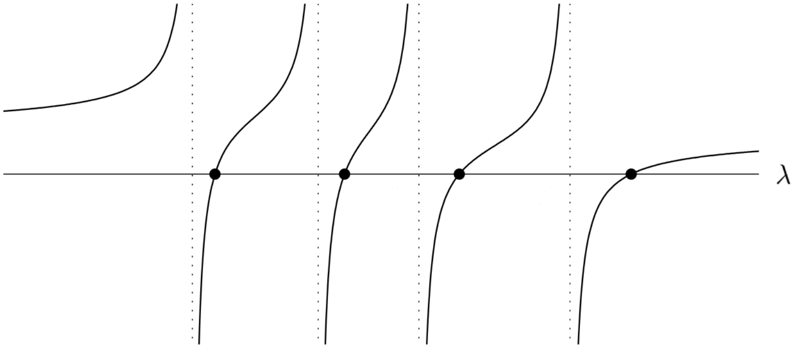
Suppose $\mathbf{q}$ is an eigenvector of $\mathbf{D} + \beta\mathbf{w}\mathbf{w}^\top$. The statement follows from $\mathbf{w}^\top\mathbf{q} \neq 0$, $\lambda \neq d_j$ (why?) and $\mathbf{w}^\top\mathbf{q}(1 + \beta\mathbf{w}^\top(\mathbf{D} - \lambda\mathbf{I})^{-1}\mathbf{w}) = 0$. $\square$

Remark 11

The equation $f(\lambda) = 0$ is known as the secular equation.

Exercise: Assume $\mathbf{D}$, β and $\mathbf{w}$ are those in Proposition 10. If λ is an eigenvalue of $\mathbf{D} + \beta\mathbf{w}\mathbf{w}^\top$, then $(\mathbf{D} - \lambda\mathbf{I})^{-1}\mathbf{w}$ is a corresponding eigenvector.

- Plot of the function $f(\lambda)$ for a problem of dimension 4. The poles of $f(\lambda)$ are the eigenvalues $\{d_j\}$ of $\mathbf{D}$, and the roots of $f(\lambda)$ (solid dots) are the eigenvalues of $\mathbf{D} + \beta\mathbf{w}\mathbf{w}^\top$. These roots can be determined rapidly.



Proposition 12

Let $\mathbf{D} \in \mathbb{R}^{m\times m}$ *be a diagonal matrix and* $\mathbf{w} \in \mathbb{R}^m$ *be a nonzero vector. Assume* $\beta \in \mathbb{R}$ *and* $\beta \neq 0$. *Then there exist a permutation matrix* $\mathbf{P}$ *and an orthogonal matrix* $\mathbf{V}$ *such that*

$$\mathbf{P}^\top \mathbf{V}^\top (\mathbf{D} + \beta \mathbf{w}\mathbf{w}^\top)\mathbf{V}\mathbf{P} = \begin{bmatrix} \mathbf{D}_1 + \beta \mathbf{w}_1 \mathbf{w}_1^\top & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_2 \end{bmatrix},$$

where $\mathbf{D}_1 \in \mathbb{R}^{r\times r}$ *is a diagonal matrix with distinct diagonal entries,* $\mathbf{D}_2 \in \mathbb{R}^{(m-r)\times(m-r)}$ *is a diagonal matrix, and* $\mathbf{w}_1 \in \mathbb{R}^r$ *is a vector with nonzero entries. More precisely,*

$$\mathbf{V}^\top \mathbf{D}\mathbf{V} = \mathbf{D}, \quad \mathbf{P}^\top \mathbf{D}\mathbf{P} = \begin{bmatrix} \mathbf{D}_1 & \\ & \mathbf{D}_2 \end{bmatrix}, \quad \mathbf{P}^\top \mathbf{V}^\top \mathbf{w} = \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{0} \end{bmatrix}.$$

Exercise: Prove Proposition 12.

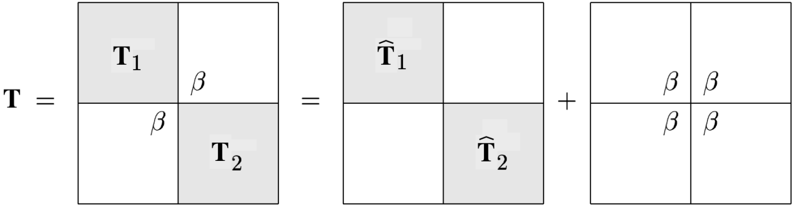
4. Divide-and-conquer

Remark 13

A symmetric tridiagonal matrix can be written as the sum of a 2×2 block diagonal matrix with tridiagonal blocks and a rank-one correction.

- Let $\mathbf{T} \in \mathbb{R}^{m \times m}$ be symmetric, tridiagonal, and unreduced. For any n in the range $1 \leq n < m$, we can write

$$\mathbf{T} = \begin{bmatrix} \widehat{\mathbf{T}}_1 & \\ & \widehat{\mathbf{T}}_2 \end{bmatrix} + \beta \begin{bmatrix} \mathbf{e}_n \\ \mathbf{e}_1 \end{bmatrix} \begin{bmatrix} \mathbf{e}_n^\top & \mathbf{e}_1^\top \end{bmatrix}.$$



- Suppose that the eigen decompositions $\widehat{\mathbf{T}}_1 = \mathbf{Q}_1\mathbf{D}_1\mathbf{Q}_1^\top$ and $\widehat{\mathbf{T}}_2 = \mathbf{Q}_2\mathbf{D}_2\mathbf{Q}_2^\top$ have been computed ($\mathbf{D}_1$ and $\mathbf{D}_2$ are diagonal, and $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are orthogonal). Then we have

$$\mathbf{T} = \begin{bmatrix} \mathbf{Q}_1 & \\ & \mathbf{Q}_2 \end{bmatrix} \left(\begin{bmatrix} \mathbf{D}_1 & \\ & \mathbf{D}_2 \end{bmatrix} + \beta \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} \begin{bmatrix} \mathbf{u}^\top & \mathbf{v}^\top \end{bmatrix} \right) \begin{bmatrix} \mathbf{Q}_1^\top & \\ & \mathbf{Q}_2^\top \end{bmatrix},$$

where $\mathbf{u} := \mathbf{Q}_1^\top \mathbf{e}_n$ and $\mathbf{v} := \mathbf{Q}_2^\top \mathbf{e}_1$. The problem is reduced to find the eigenvalues of a diagonal matrix plus a rank-one correction.

Remark 14

Suppose that the eigenvalues of $\widehat{\mathbf{T}}_1$ and $\widehat{\mathbf{T}}_2$ are known. A nonlinear but rapid calculation can be used to get from the eigenvalues of $\widehat{\mathbf{T}}_1$ and $\widehat{\mathbf{T}}_2$ to those of $\mathbf{T}$ itself by the secular equation. The divide-and-conquer algorithm is based on recursive use of this idea.