EIGENVALUES OF TRIDIAGONAL MATRICES

\[ A = \begin{bmatrix}
\alpha_1 & \beta_1 & 0 & 0 & \cdots & 0 \\
\beta_1 & \alpha_2 & \beta_2 & 0 \\
0 & \beta_2 & \alpha_3 & \beta_3 & \ddots \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
0 & \cdots & \beta_{n-2} & \alpha_{n-1} & \beta_{n-1} \\
\end{bmatrix} \]

To calculate the characteristic polynomial, we define the sequence of polynomials

\[ f_k(\lambda) = \det \begin{bmatrix}
\alpha_1 - \lambda & \beta_1 & 0 & \cdots & 0 \\
\beta_1 & \alpha_2 - \lambda & \beta_2 & & \ddots \\
0 & \cdots & \ddots & \ddots & \ddots \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
0 & \cdots & \alpha_{k-1} - \lambda & \beta_{k-1} \\
\end{bmatrix} \]

for \( k \geq 1 \), with \( f_1(\lambda) = \alpha_1 - \lambda \). Also introduce \( f_0(\lambda) \equiv 1 \). We will assume all \( \beta_i \neq 0 \).
We can obtain a recursive relationship for these polynomials. Expand the determinant by minors, using the last row. With that, there is a need for one further expansion in the last column of one of the reduced determinants. Then

\[ f_k(\lambda) = (\alpha_k - \lambda) f_{k-1}(\lambda) - \beta_{k-1}^2 f_{k-2}(\lambda), \quad k \geq 1 \]  

(1)

The characteristic polynomial for the original matrix \( T \) is \( f_n(\lambda) \), and we want to compute its zeros.

Note that we can use (1) to evaluate \( f_n(\lambda) \). What is the cost?

Assume the quantities \( \{\beta_k^2\} \) have been prepared already. Then given a value of \( \lambda \), \( f_1(\lambda) \) costs 1 operation; and

\[ f_2(\lambda) = (\alpha_k - \lambda) f_1(\lambda) - \beta_1^2 \]

costs 3 operations. All of the remaining polynomials \( f_k(\lambda) \) cost 4 operations, \( k = 3, \ldots, n \). Thus there is a total operations cost of \( 4(n-1) \). This is more efficient than if we were to obtain \( f_n(\lambda) \) explicitly.
We can solve
\[ f_n(\lambda) = 0 \]
by using a rootfinding method. Since it is difficult to obtain the derivative, the secant method is the natural choice for the root finding. Where are the roots located? We can use the Gerschgorin circle theorem to obtain a bounding interval. But there turns out to be a better approach.

The sequence \( \{f_0, f_1, \ldots, f_n\} \) forms a Sturm sequence of polynomials; and such sequences have special properties. Given a point \( b \), calculate
\[
\{f_0(b), f_1(b), \ldots, f_n(b)\}
\]
and observe the signs of these quantities. If some \( f_j(\lambda) = 0 \), then choose the sign of \( f_j(\lambda) \) to be opposite to that of \( f_{j-1}(\lambda) \). It can be shown that
\[ f_j(\lambda) = 0 \quad \Rightarrow \quad f_{j-1}(\lambda) \neq 0 \]
Having obtain a sequence of signs from (2), let \( s(\lambda) \) denote the number of agreements of sign between consecutive members of the sign sequence.
EXAMPLE

From the text on p. 620, let

\[
T = \begin{bmatrix}
2 & 1 & 0 & 0 & 0 & 0 \\
1 & 2 & 1 & 0 & 0 & 0 \\
0 & 1 & 2 & 1 & 0 & 0 \\
0 & 0 & 1 & 2 & 1 & 0 \\
0 & 0 & 0 & 1 & 2 & 1 \\
0 & 0 & 0 & 0 & 1 & 2 \\
\end{bmatrix}
\]

Then

\[f_0(\lambda) \equiv 1, \quad f_1(\lambda) = 2 - \lambda\]

\[f_k(\lambda) = (2 - \lambda) f_{k-1}(\lambda) - f_{k-2}(\lambda), \quad k = 2, \ldots, 6\]

Then for \(\lambda = 1\),

\[\{f_0(1), f_1(1), \ldots, f_6(1)\} = \{1, 1, 0, -1, -1, 0, 1\}\]

The sign sequence is

\[\{+, +, -, -, -, +, +\}\]

Then \(s(1) = 4\).
Theorem: The number of roots greater than $\lambda = a$ is given by $s(a)$. For $a < b$, the number of roots in the interval $a < \lambda \leq b$ is given by $s(a) - s(b)$.

It can also be shown that with our assumption that all $\beta_i \neq 0$ that the matrix $T$ will have $n$ distinct eigenvalues

$$\lambda_1 < \lambda_2 < \cdots < \lambda_n$$

However, these may be located very close to one another. We can use the above theorem to separate the roots of $f_n(\lambda) = 0$ into disjoint subintervals; and then we can use a guaranteed rootfinder such as Brent’s zero program to converge quickly to the root in each such subinterval. This is a practical method to find the roots; although it is most widely used when only a few eigenvalues are desired, say for example, the 5 largest ones.
EXAMPLE (continuation)

Recall the earlier example for

\[
T = \begin{bmatrix}
2 & 1 & 0 & 0 & 0 & 0 \\
1 & 2 & 1 & 0 & 0 & 0 \\
0 & 1 & 2 & 1 & 0 & 0 \\
0 & 0 & 1 & 2 & 1 & 0 \\
0 & 0 & 0 & 1 & 2 & 1 \\
0 & 0 & 0 & 0 & 1 & 2 \\
\end{bmatrix}
\]

\[
\{f_0(1), f_1(1), \ldots, f_6(1)\} = \{1, 1, 0, -1, -1, 0, 1\}
\]

with \(s(1) = 4\). For \(\lambda = 3\),

\[
\{f_0(3), f_1(3), \ldots, f_6(3)\} = \{1, -1, 0, 1, -1, 0, 1\}
\]

and \(s(3) = 2\). Note that neither \(\lambda = 1\) nor \(\lambda = 3\) is a root; and \(s(1) - s(3) = 2\). Therefore there are 2 roots in the interval \(1 < \lambda < 3\).

This example is carried further in the text on pages 622-623.