Symmetric Schemes for Computing the Minimum Eigenvalue of a Symmetric Toeplitz Matrix

Heinrich Voss

Technical University Hamburg–Harburg, Section of Mathematics, D–21071 Hamburg, Federal Republic of Germany, e-mail: voss@tu-harburg.de

Dedicated to Ludwig Elsner on the occasion of his 60th birthday

Abstract

In [8] and [9] W. Mackens and the present author presented two generalizations of a method of Cybenko and Van Loan [4] for computing the smallest eigenvalue of a symmetric, positive definite Toeplitz matrix. Taking advantage of the symmetry or skew symmetry of the corresponding eigenvector both methods are improved considerably.

Keywords. Toeplitz matrix, eigenvalue problem, projection method, symmetry

1 Introduction

Several approaches have been reported in the literature for computing the smallest eigenvalue of a real symmetric, positive definite Toeplitz matrix (RSPDT). This problem is of considerable interest in signal processing. Given the covariance sequence of the observed data, Pisarenko [10] suggested a method which determines the sinusoidal frequencies from the eigenvector of the covariance matrix associated with its minimum eigenvalue.

Cybenko and Van Loan [4] presented an algorithm which is a combination of bisection and Newton's method for the secular equation. Replacing Newton's method by a root finding method based on rational Hermitian interpolation of the secular equation Mackens and the present author in [8] improved this approach substantially. In [9] it was shown that the algorithm from [8] is equivalent to a projection method where in every step the eigenvalue problem is projected to a two dimensional space. This interpretation suggested a further enhancement of Cybenko and Van Loan's method.

If $T_n \in \mathbb{R}^{(n,n)}$ is a RSPDT matrix and E_n denotes the (n,n) flipmatrix with ones in its secondary diagonal and zeros elsewhere, then $E_n^2 = I$ and $T_n = E_n T_n E_n$. Hence $T_n x = \lambda x$ if and only if

$$T_n(E_n x) = E_n T_n E_n^2 x = \lambda E_n x,$$

and x is an eigenvector of T_n if and only if $E_n x$ is. If λ is a simple eigenvalue of T_n then from $\|x\|_2 = \|E_n x\|_2$ we obtain $x = E_n x$ or $x = -E_n x$. We say that an eigenvector x is symmetric and the corresponding eigenvalue λ is even if $x = E_n x$, and x is called skew-symmetric and λ is odd if $x = -E_n x$.

One disadvantage of the approximation schemes in [8] and [9] is that they do not reflect the symmetry properties of the eigenvector corresponding to the minimum eigenvalue. In this paper we present variants which take advantage of the symmetry of the eigenvector and which essentially are of equal cost as the methods considered in [8] and [9].

The symmetry class of the principal eigenvector is known in advance only for a small class of Toeplitz matrices. The following result was given by Trench [11]:

Theorem 1:
Let

$$T_n = (t_{|i-j|})_{i,j=1,\dots,n},\ t_j := \frac{1}{\pi}\int_0^\pi F(\theta)\cos(j\theta)\,d\theta, \quad j = 0,1,2,\dots,n-1,$$

where $F : (0,\pi) \to \mathbb{R}$ *is nonincreasing and* $F(0+) =: M > m := F(\pi-)$. *Then for every* n *the matrix* T_n *has* n *distinct eigenvalues in* (m, M)*, its even and odd spectra are interlaced, and its largest eigenvalue is even.*

If T_n satisfies the conditions of Theorem 1 then for even n the principal eigenvector is odd and vice versa. For general Toeplitz matrices T_n the symmetry class is detected by the algorithm at negligible cost.

The paper is organized as follows. In Section 2 we briefly sketch the algorithms from [8] and [9]. Sections 3 and 4 describe their generalizations if the symmetry class of the principal eigenvector is taken into account. Finally, some concluding remarks are made in Section 5.

2 Nonsymmetric methods

In this section we briefly review the approach to the computation of the smallest eigenvalue of a RSPDT matrix which was presented in [8] and [9].

Let

$$T_n = (t_{|i-j|})_{i,j=1,\dots,n} \in \mathbb{R}^{(n,n)}$$

be a RSPDT matrix. We denote by $T_j \in \mathbb{R}^{(j,j)}$ its j-th principal submatrix, and we assume that its diagonal is normalized by $t_0 = 1$. If $\lambda_1^{(j)} \le \lambda_2^{(j)} \le \dots \le \lambda_j^{(j)}$ are the eigenvalues of T_j then the interlacing property $\lambda_{j-1}^{(k)} \le \lambda_{j-1}^{(k-1)} \le \lambda_j^{(k)}$, $2 \le j \le k \le n$, holds.

Eliminating the variables $x_2, \dots, x_n$ from the system

$$\begin{pmatrix} 1-\lambda & , & t^T \\ t & , & T_{n-1} - \lambda I \end{pmatrix} x = 0$$

that characterizes the eigenvalue of T_n one obtains

$$(1 - \lambda - t^T (T_{n-1} - \lambda I)^{-1} t) x_1 = 0.$$

We assume that $\lambda_1^{(n)} < \lambda_1^{(n-1)}$. Then $x_1 \neq 0$, and $\lambda_1^{(n)}$ is the smallest root of the secular equation

$$f(\lambda) := -1 + \lambda + t^T (T_{n-1} - \lambda I)^{-1} t = 0. \tag{1}$$

f is strictly monotonely increasing and strictly convex in the interval $(0, \lambda_1^{(n-1)})$. Therefore for every initial guess $\mu_0 \in (\lambda_1^{(n)}, \lambda_1^{(n-1)})$ Newton's method converges monotonely decreasing and quadratically to $\lambda_1^{(n)}$. Since

$$f'(\lambda) = 1 + \|(T_{n-1} - \lambda I)^{-1} t\|_2^2$$

a Newton step can be performed in the following way:

$$\text{Solve } (T_{n-1} - \mu_k I) y = -t \text{ for } y\text{, and set } \mu_{k+1} = \mu_k - \frac{-1 + \mu_k - y^T t}{1 + \|y\|_2^2}$$

where the Yule – Walker system $(T_{n-1} - \mu I) y = -t$ can be solved by Durbin's algorithm (cf. [6], p. 195) requiring $2n^2$ flops.

An initial guess μ_0 for Newton's method can be obtained by a bisection process. If μ is not in the spectrum of any of the submatrices $T_j - \mu I$ then Durbin's algorithm for $(T - \mu I)/(1 - \mu)$ determines a lower triangular matrix L such that

$$\frac{1}{1-\mu} L (T - \mu I) L^T = \operatorname{diag}\{1, \delta_1, \dots, \delta_{n-1}\}.$$

Hence, from Sylvester's law of inertia we obtain that

(i) $\mu < \lambda_1^{(n)}$ if $\delta_j > 0$ for $j = 1, \dots, n-1$,

(ii) $\mu \in (\lambda_1^{(n)}, \lambda_1^{(n-1)})$ if $\delta_j > 0$ for $j = 1, \dots, n-2$ and $\delta_{n-1} < 0$

(iii) $\mu > \lambda_1^{(n-1)}$ if $\delta_j < 0$ for some $j \in \{1, \dots, n-2\}$.

Cybenko and Van Loan combined a bisection method with Newton's method for computing the minimum eigenvalue of T_n.

Since the smallest root $\lambda_1^{(n)}$ and the smallest pole $\lambda_1^{(n-1)}$ of the rational function f can be very close to each other usually a large number of bisection steps is needed to get a suitable initial approximation of Newton's method. Moreover, the global convergence behaviour of Newton's method can be quite unsatisfactory. In [8] the approach of Cybenko and Van Loan was improved substantially using a root finding method which is based on a rational model

$$g(\lambda;\mu) := f(0) + f'(0)\lambda + \lambda^2 \frac{b}{c-\lambda},$$

where μ is the current approximation of $\lambda_1^{(n)}$, and b and c are determined such that

$$g(\mu;\mu) = f(\mu),\ \ g'(\mu;\mu) = f'(\mu).$$

It is shown that for $\mu_k \in (\lambda_1^{(n)}, \lambda_1^{(n-1)})$ the function $g(\cdot;\mu_k)$ has exactly one zero $\mu_{k+1} \in (0,\mu_k)$ and that

$$\lambda_1^{(n)} < \mu_{k+1} < \mu_k - f(\mu_k)/f'(\mu_k).$$

Hence, the sequence $\{\mu_k\}$ converges monotonely decreasing to $\lambda_1^{(n)}$, the convergence is quadratic and faster than the convergence of Newton's method. The essential cost of one step are the same as for one Newton step.

In [9] it was shown that the smallest root of $g(\cdot;\mu)$ is the smallest eigenvalue of the projected eigenvalue problem

$$Q^T T_n Q\xi = \lambda Q^T Q \xi \tag{2}$$

where

$$Q = (q(0), q(\mu)) \in \mathbb{R}^{(n,2)}$$

and $q(\nu) := (T_n - \nu I)^{-1} e^1$, $e^1 = (1,0,\ldots,0)^T$. This interpretation suggests generalizations of the method where the problem is projected to subspaces

$$\mathrm{span}\{q(\mu_1),\ldots,q(\mu_k)\}$$

of the same type of increasing order k where the parameters μ_j are constructed in the course of the algorithm. The resulting method was shown to be at least cubically convergent.

The representation in (2) clearly demonstrates a weakness of the approaches in [8] and [9]: Although the eigenvector corresponding to $\lambda_1^{(n)}$ is known to be symmetric or skew-symmetric the trial vectors in the projection method have neither of these properties.

3 Exploiting symmetry in rational interpolation

In this section we discuss a variant of the approximation scheme from [8] that exploits the symmetry and skew-symmetry of the corresponding eigenvector, respectively.

To take into account the symmetry properties of the eigenvector we eliminate the variables $x_2, \ldots, x_{n-1}$ from the system

$$\begin{pmatrix} 1-\lambda & , & \tilde{t}^T & , & t_{n-1} \\ \tilde{t} & , & T_{n-2}-\lambda I & , & E_{n-2}\tilde{t} \\ t_{n-1} & , & \tilde{t}^T E_{n-2} & , & 1-\lambda \end{pmatrix} x = 0 \tag{3}$$

where $\tilde{t} = (t_1, \ldots, t_{n-2})^T$.

Then every eigenvalue λ of T_n which is not in the spectrum of T_{n-2} is an eigenvalue of the two dimensional nonlinear eigenvalue problem

$$\begin{pmatrix} 1-\lambda-\tilde{t}^T(T_{n-2}-\lambda I)^{-1}\tilde{t} & , & t_{n-1}-\tilde{t}^T(T_{n-2}-\lambda I)^{-1}E_{n-2}\tilde{t} \\ t_{n-1}-\tilde{t}^T E_{n-2}(T_{n-2}-\lambda I)^{-1}\tilde{t} & , & 1-\lambda-\tilde{t}^T(T_{n-2}-\lambda I)^{-1}\tilde{t} \end{pmatrix} \begin{pmatrix} x_1 \\ x_n \end{pmatrix} = 0. \tag{4}$$

Moreover, if such a λ is an even eigenvalue of T_n, then $(1,1)^T$ is the corresponding eigenvector of problem (4), and if λ is an odd eigenvalue of T_n then $(1,-1)^T$ is the corresponding eigenvector of system (4).

Hence, if the smallest eigenvalue $\lambda_1^{(n)}$ is even, then it is the smallest root of the rational function

$$g_+(\lambda) := -1 - t_{n-1} + \lambda + \tilde{t}^T(T_{n-2}-\lambda I)^{-1}(\tilde{t} + E_{n-2}\tilde{t}), \tag{5}$$

and if $\lambda_1^{(n)}$ is an odd eigenvalue of T_n then it is the smallest root of

$$g_-(\lambda) := -1 + t_{n-1} + \lambda + \tilde{t}^T(T_{n-2}-\lambda I)^{-1}(\tilde{t} - E_{n-2}\tilde{t}). \tag{6}$$

If the symmetry class of the principal eigenvector is known in advance then a straight forward generalization of the scheme in [8] can be based on (5) or (6), respectively. In the general case it is the minimum of the smallest roots of g_+ and g_-, and the symmetry class must be detected by the method itself.

The elimination of $x_2, \ldots, x_{n-1}$ is nothing else but exact condensation of the eigenvalue problem $Tx = \lambda x$ where x_1 and x_n are chosen to be masters and $x_2, \ldots, x_{n-1}$ are the slaves. If $\phi^1, \ldots, \phi^{n-2}$ denotes an orthonormal set of eigenvectors of the slave problem

$$T_{n-2}\phi^j = \lambda_j^{(n-2)}\phi^j, \quad j = 1, \ldots, n-2,$$

then the functions g_+ and g_- can be written as (cf. [7])

$$g_\pm(\lambda) = g_\pm(0) + g'_\pm(0)\lambda + \lambda^2 \sum_{j=1}^{n-2} \frac{\alpha_{\pm,j}^2}{\lambda_j^{(n-2)} - \lambda} \tag{7}$$

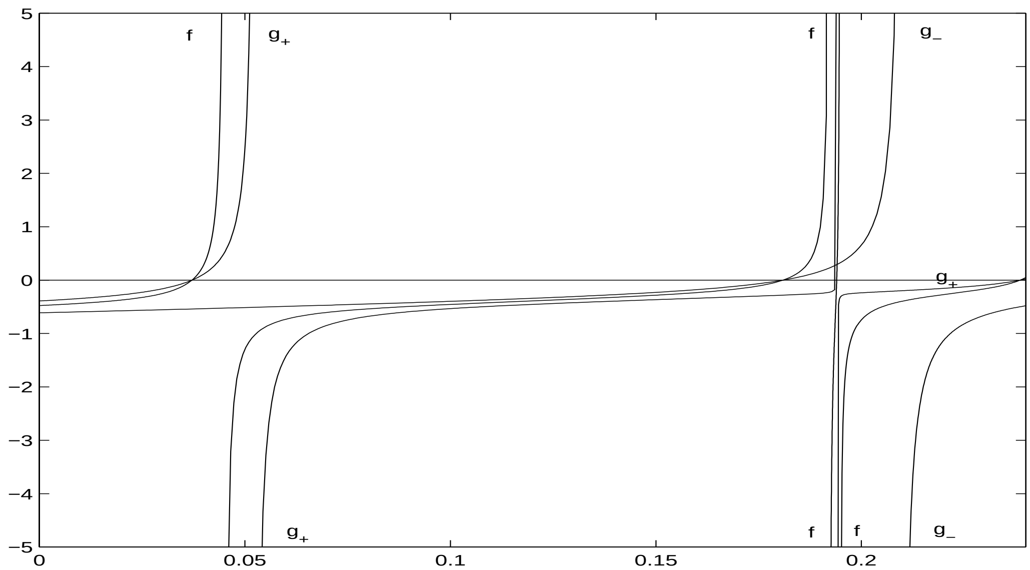

Fig 1: Graphs of f, g_+ and g_-

where

$$g_\pm(0) = -1 \mp t_{n-1} + \tilde{t}^T T_{n-2}^{-1}(\tilde{t} \pm E_{n-2}\tilde{t}),$$

$$g'_\pm(0) = 1 + \tilde{t}^T T_{n-2}^{-2}(\tilde{t} \pm E_{n-2}\tilde{t}) = 1 + 0.5\|T_{n-2}^{-1}(\tilde{t} \pm E_{n-2}\tilde{t})\|_2^2,$$

and

$$\alpha_{\pm,j} = \frac{1}{\lambda_j^{(n-1)}}(\phi^j)^T(\tilde{t} \pm E_{n-2}\tilde{t}).$$

Hence, the zeros of g_+ and g_- are the even and odd eigenvalues of T_n, and the poles of g_+ and g_- are the even and odd eigenvalues of T_{n-2}, respectively. Figure 1 shows the graphs of the functions f, g_+ and g_- for a Toeplitz matrix of dimension 32.

If we are given an approximation μ of $\lambda_1^{(n)}$ then equation (7) suggests the following rational Hermitian approximation of $g_\pm(\lambda)$:

$$h_\pm(\lambda;\mu) = g_\pm(0) + g'_\pm(0)\lambda + \lambda^2 \frac{b_\pm}{c_\pm - \lambda} \tag{8}$$

where the parameters $b_\pm$ and $c_\pm$ are determined from the Hermitian interpolation conditions

$$h_\pm(\mu;\mu) = g_\pm(\mu),\ \ h'_\pm(\mu;\mu) = g'_\pm(\mu). \tag{9}$$

Following the lines of the proof of Theorem 1 in [8] one gets the basic properties of $h_\pm$.

Theorem 2:
Let $\omega_\pm$ be the smallest pole of $g_\pm$, let $\mu \in [0, \omega_\pm)$, and let $h_\pm$ be defined by equations (8) and (9). Then it holds that

(i) $\quad b_\pm > 0 \quad$ *and* $\quad c_\pm > \mu$,
whence $h_\pm$ *is strictly monotonely increasing and strictly convex in* $[0, c_\pm)$.

(ii) $\quad h_\pm(\lambda_1^{(n)}) < 0$.

From Theorem 2 we deduce the following method for computing the smallest eigenvalue of a RSPDT matrix T_n. Set $\alpha = 0$ as a lower bound of the smallest eigenvalue and let the variable τ monitor whether $\lambda_1^{(n)}$ is even or odd or the type of $\lambda_1^{(n)}$ is not yet known:

$$\tau = \begin{cases} -1 & \text{if } \lambda_1^{(n)} \text{is odd} \\ 0 & \text{if the type of } \lambda_1^{(n)} \text{ is unknown} \\ 1 & \text{if } \lambda_1^{(n)} \text{is even} \end{cases} \tag{10}$$

To obtain an upper bound of $\lambda_1^{(n)}$ solve the Yule – Walker system $T_{n-2}z = -\tilde{t}$, and let $z_+ := z + E_{n-2}z$ and $z_- := z - E_{n-2}z$. Then

$$g_\pm(0) = -1 \mp t_{n-1} - \tilde{t}^T z_\pm,\ \ g_\pm'(0) = 1 + 0.5\|z_\pm\|_2^2,$$

and from the monotonicity and convexity of g_+ and g_- in $[0, \lambda_1^{(n)}]$ it follows that

$$\beta := \min\{-g_+(0)/g_+'(0), -g_-(0)/g_-'(0)\}$$

is an upper bound of $\lambda_1^{(n)}$.

Choose $\mu_0 \in (0, \beta]$, set $k := 0$, and do the following steps until convergence of the sequence $\{\mu_k\}$:

1. Solve $(T_{n-2} - \mu_k I)y = -\tilde{t}$ using Durbin's algorithm and determine whether $\mu_k \geq \lambda_1^{(n-2)}$ or not.

2. If $\mu_k \geq \lambda_1^{(n-2)}$ then do a bisection step:

$$\beta := \mu_k,\ \ \mu_{k+1} := 0.5(\alpha + \beta)$$

otherwise obtain new bounds of $\lambda_1^{(n)}$ in the following way:

   – if $\tau > -1$ then determine $g_+(\mu_k)$. If $\tau = 1$ and $g_+(\mu_k) < 0$ then $\alpha := \mu_k$ is an improved lower bound

   – if $\tau < 1$ then determine $g_-(\mu_k)$. If $\tau = -1$ and $g_-(\mu_k) < 0$ then $\alpha := \mu_k$ is an improved lower bound

   – If $\tau = 0$ and $g_+(\mu_k) < 0$ and $g_-(\mu_k) < 0$ then $\alpha := \mu_k$ is an improved lower bound of $\lambda_1^{(n)}$

   – if $\tau = 0$ and $g_-(\mu_k) < 0 < g_+(\mu_k)$ then $\lambda_1^{(n)} < \mu_k$ is the smallest root of g_+. Set $\tau := 1$

   – if $\tau = 0$ and $g_+(\mu_k) < 0 < g_-(\mu_k)$ then $\lambda_1^{(n)} < \mu_k$ is the smallest root of g_-. Set $\tau := -1$

- if $\tau > -1$ compute $g'_+(\mu_k)$ and determine the smallest root ρ_+ of $g_+(\cdot;\mu_k)$; else set $\rho_+ = 1$.
- if $\tau < 1$ compute $g'_-(\mu_k)$ and determine the smallest root ρ_- of $g_-(\cdot;\mu_k)$; else set $\rho_- = 1$.
- $\mu_{k+1} := \min\{\rho_+, \rho_-, \beta\}$

(iii) k:=k+1

To check the convergence we use the following lower bound of $\lambda_1^{(n)}$ of [8].

Lemma 3
Let $0 \le \alpha < \lambda_1^{(n)} < \mu < \lambda_1^{(n-1)}$, and let $\lambda_1^{(n)}$ be the smallest positive root of $g_\circ$, $\circ \in \{+,-\}$. Let p be the quadratic polynomial satisfying the interpolation conditions

$$p(\alpha) = g_\circ(\alpha),\ p'(\alpha) = g'_\circ(\alpha),\ p(\mu) = g_\circ(\mu).$$

Then p has a unique root $\kappa \in (\alpha,\mu)$ and $\kappa \le \lambda_1^{(n)}$.

The convergence behaviour is the same as for the nonsymmetric method: $\mu_{k_0} \in (\lambda_1^{(n)}, \lambda_1^{(n-2)})$ for some k_0. For $k \ge k_0$ the sequence $\{\mu_k\}$ converges quadratically and monotonely decreasing to $\lambda_1^{(n)}$, and it converges faster than Newton's method for $g_\circ$, where $\circ \in \{+,-\}$ such that $g_\circ(\lambda_1^{(n)}) = 0$. Notice that $\lambda_1^{(n-1)} \le \lambda_1^{(n-2)}$. Hence, the symmetric method usually will need a smaller number of bisection steps to reach its monotonely decreasing phase than its nonsymmetric counterpart.

To test the improvement upon the nonsymmetric method we considered Toeplitz matrices

$$T = m \sum_{k=1}^{n} \eta_k T_{2\pi\theta_k} \tag{11}$$

where m is chosen such that the diagonal of T is normalized to 1,

$$T_\theta = (t_{ij}) = (\cos(\theta(i-j)))$$

and η_k and θ_k are uniformly distributed random numbers taken from $[0,1]$ (cf. Cybenko, Van Loan [4]).

Table 1 contains the average number of flops and the average number of Durbin steps needed to determine the smallest eigenvalue in 100 test problems with each of the dimensions $n = 32, 64, 128, 256, 512$ and $n = 1024$ for the methods based on rational Hermitian interpolation. The iteration was terminated if Lemma 3 guaranteed the relative error to be less than 10^{-6}.

dimension	non-symmetric method from [8]		symmetric method	
	flops	steps	flops	steps
32	$1.071\ E04$	4.55	$9.087\ E03$ (84.9%)	3.75
64	$4.545\ E04$	5.19	$3.653\ E04$ (80.4%)	4.12
128	$1.695\ E05$	5.01	$1.407\ E05$ (83.0%)	4.14
256	$7.310\ E05$	5.50	$6.046\ E05$ (82.7%)	4.55
512	$3.297\ E06$	6.25	$2.597\ E06$ (78.8%)	4.92
1024	$1.352\ E07$	6,44	$1.065\ E07$ (78.8%)	5.08

Tab. 1. Rational Hermitian interpolation

4 A symmetric projection method

The root finding method of the last section can be interpreted as a projection method where in each step the eigenvalue problem is projected to a 2 dimensional space. Similarly as in [9] this follows easily from

Theorem 4:
Let e^1 and e^n be the unit vector containing a 1 in its first and last component, respectively, and for λ not in the spectrum of T_n and T_{n-2} let

$$p_\pm(\lambda) := -g_\pm(\lambda)(T_n - \lambda I)^{-1}(e^1 \pm e^n).$$

Then

$$p_\pm(\lambda) = \begin{pmatrix} 1 \\ z_\pm(\lambda) \\ \pm 1 \end{pmatrix}, \quad \text{where } z_\pm(\lambda) := -(T_{n-2} - \lambda I)^{-1}(\tilde{t} \pm E_{n-2}\tilde{t}), \tag{12}$$

and it holds that

$$p_\pm(\lambda)^T T p_\pm(\mu) = 2 \begin{cases} -g_\pm(\lambda) + \lambda g'_\pm(\lambda) & \text{for} \quad \lambda = \mu, \\ -g_\pm(\lambda) + \lambda \dfrac{g_\pm(\lambda) - g_\pm(\mu)}{\lambda - \mu} & \text{for} \quad \lambda \neq \mu, \end{cases}$$

and

$$p_\pm(\lambda)^T p_\pm(\mu) = 2 \begin{cases} g'_\pm(\lambda) & \text{for} \quad \lambda = \mu, \\ \dfrac{g_\pm(\lambda) - g_\pm(\mu)}{\lambda - \mu} & \text{for} \quad \lambda \neq \mu. \end{cases}$$

Proof: Equation (12) follows immediately from

$$\begin{pmatrix} 1-\lambda & \tilde{t}^T & t_{n-1} \\ \tilde{t} & T_{n-2} - \lambda I & E_{n-2}\tilde{t} \\ t_{n-1} & \tilde{t}^T E_{n-2} & 1-\lambda \end{pmatrix} \begin{pmatrix} 1 \\ z_\pm(\lambda) \\ \pm 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 - \lambda + \tilde{t}^T z_\pm(\lambda) \pm t_{n-1} \\ \tilde{t} + (T_{n-2} - \lambda I) z_\pm(\lambda) \pm E_{n-2}\tilde{t} \\ t_{n-1} + \tilde{t}^T E_{n-2} z_\pm(\lambda) \pm 1 \mp \lambda \end{pmatrix} = -g_\pm(\lambda) \begin{pmatrix} 1 \\ 0 \\ \pm 1 \end{pmatrix}.$$

If μ is not in the spectrum of T_n then

$$\begin{aligned} p_\pm(\lambda)^T T_n p_\pm(\mu) &= -g_\pm(\mu) p_\pm(\lambda)^T (T_n - \mu I + \mu I)(T_n - \mu I)^{-1}(e^1 \pm e^n) \\ &= -g_\pm(\mu) p_\pm(\lambda)^T (e^1 \pm e^n) + \mu p_\pm(\lambda)^T p_\pm(\mu) \\ &= -2g_\pm(\mu) + \mu p_\pm(\lambda)^T p_\pm(\mu), \end{aligned}$$

and for λ not in the spectrum of T_n the symmetry of T_n yields

$$p_\pm(\lambda)^T T_n p_\pm(\mu) = -2g_\pm(\lambda) + \lambda p_\pm(\lambda)^T p_\pm(\mu). \tag{13}$$

Hence for $\lambda \neq \mu$

$$p_\pm(\lambda)^T p_\pm(\mu) = 2\frac{g_\pm(\lambda) - g_\pm(\mu)}{\lambda - \mu}$$

and from eqn. (13) we get

$$p_\pm(\lambda)^T T_n p_\pm(\mu) = -2g_\pm(\lambda) + 2\frac{g_\pm(\lambda) - g_\pm(\mu)}{\lambda - \mu}\lambda.$$

Finally, for $\lambda = \mu$ one obtains from eqns. (12), (5) and (6)

$$\|p_\pm(\lambda)\|_2^2 = 2 + \|z_\pm(\lambda)\|_2^2 = 2g'_\pm(\lambda)$$

and from eqn. (13)

$$p_\pm(\lambda)^T T_n p_\pm(\lambda) = -2g_\pm(\lambda) + 2\lambda g'_\pm(\lambda). \qquad \square$$

Theorem 4 suggests the following type of projection method for computing the smallest eigenvalue of a RSPDT matrix T_n:

(i) Choose parameters $\mu_1, \ldots, \mu_k$ (not in the spectrum of T_n) and solve the linear systems

$$(T_n - \mu_k I) p_\pm(\mu_k) = -g_\pm(\mu_k)(e^1 \pm e^n)$$

(ii) Determine the smallest eigenvalues $\rho_\pm$ of the projected problems

$$(Q_k^\pm)^T T_n Q_k^\pm y = \lambda (Q_k^\pm)^T Q_k^\pm y \tag{14}$$

where

$$Q_k^\pm := (p_\pm(\mu_1), \ldots, p_\pm(\mu_k)) \in \mathbb{R}^{(n,k)}.$$

(iii) $\lambda = \min\{\rho_+, \rho_-\}$

By Theorem 4 the entries of the projected matrices $A_k^\pm := (Q_k^\pm)^T T_n Q_k^\pm$ and $B_k^\pm := (Q_k^\pm)^T Q_k^\pm$ are given by (we divided all entries by 2)

$$a_{ij}^\pm = \begin{cases} -g_\pm(\mu_i) + \mu_i g'_\pm(\mu_i) & \text{if} \quad i = j, \\ -g_\pm(\mu_i) + \dfrac{g_\pm(\mu_i) - g_\pm(\mu_j)}{\mu_i - \mu_j}\mu_i & \text{if} \quad i \neq j, \end{cases} \tag{15}$$

and

$$
b_{ij}^{\pm} = \begin{cases} g'_{\pm}(\mu_i) & \text{if } \ i = j, \\ \dfrac{g_{\pm}(\mu_i) - g_{\pm}(\mu_j)}{\mu_i - \mu_j} & \text{if } \ i \neq j. \end{cases} \tag{16}
$$

In the algorithm to follow we will construct the parameters μ_j in the course of the method. Increasing the dimension of the projected problem by one (adding one parameter) essentially requires the solution of one Yule – Walker system and a small number of level one operations to compute $g_+(\mu_k)$ and $g_-(\mu_k)$. Then the matrices $A_k^{\pm}$ and $B_k^{\pm}$ can be updated easily from the matrices of the previous step.

Symmetric projection method:
Let $\alpha = 0$, $\mu_1 = 0$, and define τ as in eqn. (10). Solve the linear system $T_{n-2}z = -\tilde{t}$, and set $z_{\pm} := z \pm E_{n-2}z$. Compute

$$
g_{\pm}(0) = -1 \mp t_{n-1} - \tilde{t}^T z_{\pm}, \ g'_{\pm}(0) = 1 + 0.5\|z_{\pm}\|_2^2,
$$

set

$$
A_1^{\pm} := (-g_{\pm}(0)) \in \mathbb{R}^{(1,1)} \quad \text{and} \quad B_1^{\pm} := (g'_{\pm}(0)) \in \mathbb{R}^{(1,1)},
$$

and

$$
\beta := \min\{-g_+(0)/g'_+(0), -g_-(0)/g'_-(0)\}.
$$

Choose any $\mu_2 \in (0, \beta]$ and set $k := 2$.

Repeat the following steps until convergence of the sequence $\{\mu_k\}$:

(i) Solve the system

$$
(T_{n-2} - \mu_k I)z = -\tilde{t}
$$

by Durbin's algorithm and determine whether $\mu_k < \lambda_1^{(n-2)}$ or $\mu_k \geq \lambda_1^{(n-2)}$.

(ii) If $\mu_k \geq \lambda_1^{(n-2)}$ then set

$$
\beta := \min\{\beta, \mu_k\} \quad \text{and} \quad \mu_k := 0.5(\alpha + \beta)
$$

else

- if $\tau > -1$ then determine $g_+(\mu_k)$. If $\tau = 1$ and $g_+(\mu_k) < 0$ then $\alpha := \mu_k$ is an improved lower bound
- if $\tau < 1$ then determine $g_-(\mu_k)$. If $\tau = -1$ and $g_-(\mu_k) < 0$ then $\alpha := \mu_k$ is an improved lower bound
- If $\tau = 0$ and $g_+(\mu_k) < 0$ and $g_-(\mu_k) < 0$ then $\alpha := \mu_k$ is an improved lower bound of $\lambda_1^{(n)}$
- if $\tau = 0$ and $g_-(\mu_k) < 0 < g_+(\mu_k)$ then $\lambda_1^{(n)} < \mu_k$ is the smallest root of g_+. Set $\tau := 1$
- if $\tau = 0$ and $g_+(\mu_k) < 0 < g_-(\mu_k)$ then $\lambda_1^{(n)} < \mu_k$ is the smallest root of g_-. Set $\tau := -1$

- if $\tau > -1$ compute $g'_+(\mu_k)$, update the matrices A_k^+ and B_k^+ and determine the smallest eigenvalue ρ_+ of the k dimensional projected problem; else set $\rho_+ = 1$.
- if $\tau < 1$ compute $g'_-(\mu_k)$, update the matrices A_k^- and B_k^- and determine the smallest eigenvalue ρ_- of the k dimensional projected problem; else set $\rho_- = 1$.
- $\mu_{k+1} := \min\{\rho_+, \rho_-, \beta\}$
- test for convergence using Lemma 3
- $k := k + 1$

The convergence properties are obtained in the same way as in [9]: Since for $\mu \in (0, \lambda_1^{(n-2)})$ the smallest positive root of $g_\pm(\cdot;\mu)$ is the smallest eigenvalue of the projected problem (14) (for $k = 2$, $\mu_1 = 0$ and $\mu_2 = \mu$) the symmetric projection method converges eventually monotonely decreasing and faster than the symmetric method from Section 3. Comparing it to the Rayleigh quotient iteration it can even be shown to be cubically convergent (cf. [9], Theorem 5).

We tested the symmetric projection method using the RSPDT matrices from (11). In the algorithm above we took into account only vectors $p_\pm(\mu_j)$ if $\mu_j < \lambda_1^{(n-2)}$. In this case Durbin's algorithm is known to be stable (cf. [3]). Additionally we considered a projection method (complete projection) where $p_\pm(\mu_j)$ was included into the projection scheme even if a bisection step was performed since $\mu_j > \lambda_1^{(n-2)}$. Although in the latter case Durbin's algorithm is not guaranteed to be stable we did not observe unstable behaviour. We compared the methods to the nonsymmetric counterpart of the method from Section 3 based on rational Hermitian interpolation.

dimension	stable projection		complete projection	
	flops	steps	flops	steps
32	1.124 $E04$ (105.0%)	3.69	1.117 $E04$ (104.4%)	3.60
64	3.776 $E04$ (83.1%)	3.97	3.574 $E04$ (78.6%)	3.72
128	1.399 $E05$ (82.5%)	4.04	1.330 $E05$ (78.5%)	3.81
256	5.863 $E05$ (80.2%)	4.39	5.425 $E05$ (74.2%)	4.03
512	2.410 $E06$ (73.1%)	4.56	2.202 $E06$ (66.8%)	4.15
1024	9.982 $E06$ (73.8%)	4.76	8.879 $E06$ (65.7%)	4.21

Tab. 2. Symmetric projection method

5 Concluding remarks

We have presented symmetric versions of the methods introduced in [8] and [9] for computing the smallest eigenvalue of a real symmetric and positive definite Toeplitz matrix which improve their nonsymmetric counterparts considerably. In our numerical tests we used Durbin's algorithm to solve Yule – Walker systems and to determine the location of parameters in the spectrum of T_{n-2}. This information can be gained from superfast Toeplitz solvers (cf. [1], [2], [5]) as well. Hence the computational complexity can be reduced to $O(n \log^2 n)$ operations.

References

[1] G.S. Ammar and W.B. Gragg, The generalized Schur algorithm for the superfast solution of Toeplitz systems. in J. Gilewicz, M. Pindor, W. Siemaszko, eds., Lecture Notes in Mathematics 1237, Berlin, 1987, pp. 315 — 330,

[2] G.S. Ammar and W.B. Gragg, Numerical experience with a superfast real Toeplitz solver. Lin. Alg. Appl., 121 (1989), pp. 185 — 206

[3] G. Cybenko, The Numerical Stability of the Levinson-Durbin Algorithm for Toeplitz Systems of Equations. SIAM J. Sci. Stat. Comput. 1 (1980), pp. 303 — 309

[4] G. Cybenko and C. Van Loan, Computing the minimum eigenvalue of a symmetric positive definite Toeplitz matrix. SIAM J. Sci. Stat. Comput., 7 (1986), pp. 123 — 131

[5] F. de Hoog, A new algorithm for solving Toeplitz systems of equations. Lin. Alg. Appl., 88/89 (1987), pp. 123 — 138

[6] G.H. Golub and C.F. Van Loan, Matrix Computations. 3rd edition The John Hopkins University Press, Baltimore and London, 1996.

[7] T. Hitziger, W. Mackens, and H. Voss, A condensation-projection method for generalized eigenvalue problems. in H. Power and C.A. Brebbia, eds., High Performance Computing in Engineering 1, Computational Mechanics Publications, Southampton, 1995, pp. 239 — 282.

[8] W. Mackens and H. Voss, The minimum eigenvalue of a symmetric positive definite Toeplitz matrix and rational Hermitian interpolation. SIAM J. Matr. Anal. Appl. 18 (1997), pp. 521 — 534

[9] W. Mackens and H. Voss, A projection method for computing the minimum eigenvalue of a symmetric positive definite Toeplitz matrix. To appear in Lin. Alg. Appl.

[10] V.F. Pisarenko, The retrieval of harmonics from a covariance function. Geophys. J. R. astr. Soc., 33 (1973), pp. 347 — 366

[11] W.F. Trench, Interlacement of the even and odd spectra of real symmetric Toeplitz matrices. Lin. Alg. Appl. 195 (1993), pp. 59 — 68