TECHNISCHE UNIVERSITÄT

TUHH

HAMBURG - HARBURG

A Projection Method for Computing the Minimum Eigenvalue of a Symmetric Positive Definite Toeplitz Matrix

Wolfgang Mackens and Heinrich Voss

Report 1 March 1996

Technische Universität Hamburg – Harburg
Arbeitsbereich Mathematik, Kasernenstrasse 12, D – 21073 Hamburg

A Projection Method for Computing the Minimum Eigenvalue of a Symmetric Positive Definite Toeplitz Matrix

Wolfgang Mackens and Heinrich Voss

Technische Universität Hamburg–Harburg, Arbeitsbereich Mathematik, Kasernenstraße 12, D—21073 Hamburg
e-mail: mackens@tu-harburg.de, voss@tu-harburg.de

Abstract

A projection method for computing the minimal eigenvalue of a symmetric and positive definite Toeplitz matrix is presented. It generalizes and accelerates the algorithm considered in [12]. Global and cubic convergence is proved. Randomly generated test problems up to dimension 1024 demonstrate the methods good global behaviour.

Keywords. Toeplitz matrix, eigenvalue problem, projection method

1 Introduction

In this paper we present a projection method for computing the smallest eigenvalue λ_1 of a symmetric and positive definite Toeplitz matrix $\boldsymbol{T}$. This problem is of considerable interest in signal processing. Given the covariance sequence of the observed data, Pisarenko [15] suggested a method which determines the sinusoidal frequencies from the eigenvector of the covariance matrix associated with the minimum eigenvalue of $\boldsymbol{T}$.

Several approaches have been reported in the literature for computing the minimal eigenvalue of a real symmetric Toeplitz matrix. Cybenko and Van Loan [5] introduced an algorithm which is a combination of a bisection method and Newton's method for the secular equation, and which was generalized to the computation of the complete spectrum by Trench [16] and by Noor and Morgera [13]. Hu and Kung [9] considered a safeguarded inverse iteration with shifts and Huckle [10], [11] studied the spectral transformation Lanczos method. In a recent paper [12], [17] the authors presented a generalization of Cybenko and Van Loan's approach where the Newton method is replaced by a root finding method based on rational Hermitian interpolation of the secular equation. For randomly generated test matrices of dimension up to 1024 this method reduced the cost of Cybenko and Van Loan's algorithm by approximately 65%.

In this paper we show that the method considered in [12] is equivalent to a projection method where in every step the eigenvalue problem is projected to a two dimensional space spanned by $(\boldsymbol{T}-\mu_j\boldsymbol{I})^{-1}\boldsymbol{e}^1$, $j=1,2$, where the parameters μ_j are determined in the course of the algorithm. This result suggests generalizations of the method where the problem is projected to subspaces of the same type of increasing dimension. We consider two variants of equal cost per step that are shown to be at least cubically convergent. The increase of the dimension of the projected problem by one requires the same cost as one step of the algorithm in [12]: The solution of one Yule-Walker system and the evaluation of two inner products.

Our paper is organized as follows. In section 2 we briefly sketch the method introduced in [12]. Section 3 interprets the rational Hermitian interpolation as a projection method to a particular two dimensional space and generalizes the method to higher dimensional spaces. In Section 4 we prove that the originating method is at least cubically convergent, Section 5 contains a MATLAB program and in Section 6 we discuss its numerical behaviour. The paper closes with concluding remarks concerning the use of super fast Toeplitz solvers.

2 A method based on rational Hermitian interpolation

In this section we briefly review an approach to the computation of the minimum eigenvalue of a real symmetric, positive definite Toeplitz matrix which was presented in [12] and which is a generalization of a method of Cybenko and Van Loan [5].

Let $\boldsymbol{T} \in \mathbb{R}^{(n,n)}$ be a symmetric positive definite Toeplitz matrix. We assume that its diagonal is normalized, and we consider the following partition:

$$\boldsymbol{T} = \begin{pmatrix} 1 & \boldsymbol{t}^T \\ \boldsymbol{t} & \boldsymbol{G} \end{pmatrix}.$$

It is well known that the eigenvalues of $\boldsymbol{T}$ and of $\boldsymbol{G}$ are real and positive and that they satisfy the interlacing property $\lambda_1 \le \omega_1 \le \lambda_2 \le \ldots \le \omega_{n-1} \le \lambda_n$ where λ_j and ω_j is the jth smallest eigenvalue of $\boldsymbol{T}$ and $\boldsymbol{G}$, respectively.

We assume that $\lambda_1 < \omega_1$. Then λ_1 is the smallest root of the secular equation

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

It is easily seen that f is strictly monotonely increasing and strictly convex in the interval $(0, \omega_1)$, and therefore for every initial value $\mu_0 \in (\lambda_1, \omega_1)$ Newton's method converges monotonely decreasing and quadratically to λ_1. Cybenko and Van Loan [5] combined Newton's method with a bisection method (to be sketched below) to design a method for the computation of the minimum eigenvalue of $\boldsymbol{T}$.

Since

$$f'(\lambda) = 1 + \|(\boldsymbol{G} - \lambda \boldsymbol{I})^{-1}\boldsymbol{t}\|_2^2 \qquad (2)$$

a Newton step can be performed in the following way:

$$\text{Solve } (\boldsymbol{G} - \mu_k \boldsymbol{I})\boldsymbol{w} = -\boldsymbol{t} \quad \text{for} \quad \boldsymbol{w},$$
$$\text{and set } \mu_{k+1} := \mu_k - \frac{-1 + \mu_k - \boldsymbol{w}^T\boldsymbol{t}}{1 + \|\boldsymbol{w}\|_2^2}$$

where the Yule-Walker system

$$(\boldsymbol{G} - \mu \boldsymbol{I})\boldsymbol{w} = -\boldsymbol{t} \qquad (3)$$

can be solved by Durbin's algorithm (cf. [7], p. 184 ff) requiring $2n^2$ flops.

The global convergence behaviour of Newton's method usually is not satisfactory since the smallest root λ_1 and the smallest pole ω_1 of the rational function f can be very close to each other. In this situation the initial steps of Newton's method are extremely slow.

The convergence can be improved considerably if an iteration method is based on a better model of the rational function f than its tangent in Newton's method. In terms of condensation methods (cf. [8]) the secular equation f can be interpreted as the exact condensation of the eigenvalue problem $\boldsymbol{Tx} = \lambda \boldsymbol{x}$ where $x_2, \dots, x_n$ are chosen to be slaves and x_1 is the only master. Using spectral informations of the slave problem $(\boldsymbol{G} - \mu \boldsymbol{I})\boldsymbol{v} = \boldsymbol{0}$ the function f can be written as (cf. [8])

$$f(\lambda) = f(0) + f'(0)\lambda + \lambda^2 \sum_{j=1}^{n-1} \frac{\alpha_j^2}{\omega_j - \lambda}$$

where α_j, $j = 1, \dots, n-1$, are real numbers depending on the eigenvectors of $\boldsymbol{G}$.

If we are given an approximation $\mu \in (0, \omega_1)$ we therefore approximate f by a rational function

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

where b and c are determined by the Hermitian interpolation conditions

$$g(\mu; \mu) = f(\mu) \quad \text{and} \quad g'(\mu; \mu) = f'(\mu)$$

and base a method on this approximation. Theorem 1 from [12] contains the basic properties of $g(\cdot; \mu)$.

Theorem 1:
Let $\mu \in (0, \omega_1)$ and let

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

where b and c are determined such that the interpolation conditions $g(\mu; \mu) = f(\mu)$ and $g'(\mu; \mu) = f'(\mu)$ are satisfied.

Then it holds that

(i) $\quad b > 0$ *and* $c > \mu$,

(ii) $\quad g(\lambda_1;\mu) < 0$ *for* $\mu \neq \lambda_1$.

From Theorem 1 we deduce the following improvement of the method of Cybenko and Van Loan:
Let $\mu_n \in (\lambda_1, \omega_1)$ be a given approximation to λ_1, then the function $g(\cdot;\mu_n)$ is strictly convex in the interval $(0,\mu_n)$. Since

$$g(\lambda_1;\mu_n) < 0 = f(\lambda_1) < f(\mu_n) = g(\mu_n;\mu_n),$$

$g(\mu;\mu_n)$ has exactly one zero $\mu_{n+1} \in (\lambda_1, \mu_n)$.

From the convexity of $g(\cdot;\mu_n)$ we obtain

$$g(\mu;\mu_n) > g(\mu_n;\mu_n) + g'(\mu_n;\mu_n)(\mu - \mu_n) = f(\mu_n) + f'(\mu_n)(\mu - \mu_n)$$

for every $\mu \in (\lambda_1, \mu_n)$, and thus μ_{n+1} always is a better approximation to λ_1 than the Newton iterate with initial guess μ_n. Hence, for $\mu_0 \in (\lambda_1, \omega_1)$ the method which defines μ_{n+1} as the unique root of the rational Hermitian interpolation $g(\cdot;\mu_n)$ in $(0,\mu_n)$ converges monotonely decreasing to λ_1, and it is guaranteed to be faster than Newton's method.

Notice that the cost of Newton's method and of the method defined above are nearly identical. One has to solve one Yule-Walker system ($2n^2$ flops) and to evaluate two inner products to obtain $f(\mu_n)$ and $f'(\mu_n)$. The determination of b and c and the solution of a quadratic equation to obtain μ_{n+1} require only $O(1)$ flops and can be neglected.

An initial value $\mu_0 \in (\lambda_1, \omega_1)$ can be obtained by the bisection process that was introduced by Cybenko and Van Loan. If μ is not in the spectrum of any of the principal submatrices of $\boldsymbol{T} - \mu\boldsymbol{I}$ then Durbin's algorithm applied to $(\boldsymbol{T} - \mu\boldsymbol{I})/(1-\mu)$ determines a lower triangular matrix

$$\boldsymbol{L} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ \ell_{21} & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ \ell_{n1} & \ell_{n2} & \dots & 1 \end{pmatrix}$$

such that

$$\frac{1}{1-\mu}\boldsymbol{L}(\boldsymbol{T} - \mu\boldsymbol{I})\boldsymbol{L}^T = \boldsymbol{D} := \mathrm{diag}\{1, E_1, \dots, E_{n-1}\}. \qquad (4)$$

If $\tilde{\boldsymbol{L}}$ is obtained from $\boldsymbol{L}$ by dropping the last row and last column then obviously

$$\frac{1}{1-\mu}\tilde{\boldsymbol{L}}(\boldsymbol{G} - \mu\boldsymbol{I})\tilde{\boldsymbol{L}}^T = \tilde{\boldsymbol{D}} := \mathrm{diag}\{1, E_1, \dots, E_{n-2}\}$$

Hence, from Sylvester's law of inertia one gets

(i) $\mu < \lambda_1$, if $E_j > 0$ for $j = 1, \dots, n-1$,

(ii) $\mu \in [\lambda_1, \omega_1)$, if $E_j > 0$ for $j = 1, \dots, n-2$ and $E_{n-1} \le 0$,

(iii) and $\mu > \omega_1$, if $E_j < 0$ for some $j \in \{1, \dots, n-2\}$.

An upper bound of λ_1 to start the bisection process can be obtained in the following way. Let $\boldsymbol{w} := -\boldsymbol{G}^{-1}\boldsymbol{t}$ be the solution of the Yule-Walker system. Then

$$\boldsymbol{q} := \frac{1}{1 + \boldsymbol{t}^T\boldsymbol{w}} \begin{pmatrix} 1 \\ \boldsymbol{w} \end{pmatrix} = \boldsymbol{T}^{-1}\boldsymbol{e}^1$$

is the first iterate of the inverse iteration with shift parameter 0 starting with the unit vector $\boldsymbol{e}^1$ which can be expected to be not too bad an approximation of the eigenvector corresponding to the smallest eigenvalue λ_1. The Rayleigh quotient

$$R(\boldsymbol{q}) := \frac{\boldsymbol{q}^T\boldsymbol{T}\boldsymbol{q}}{\boldsymbol{q}^T\boldsymbol{q}} = \frac{1 + \boldsymbol{t}^T\boldsymbol{w}}{1 + \|\boldsymbol{w}\|_2^2} \tag{5}$$

is an upper bound of λ_1 which should be not too bad either.

3 Rational Hermitian interpolation and projection

The root finding method based on rational Hermitian interpolation of the last section can be interpreted as a projection method. To see this we first prove the following

Lemma 2:
Let $\boldsymbol{e}^1$ be the unit vector containing a 1 in its first component, and for λ not in the spectrum of $\boldsymbol{T}$ and of $\boldsymbol{G}$ let

$$\boldsymbol{q}(\lambda) := -f(\lambda)(\boldsymbol{T} - \lambda\boldsymbol{I})^{-1}\boldsymbol{e}^1.$$

Then

$$\boldsymbol{q}(\lambda) = \begin{pmatrix} 1 \\ \boldsymbol{w}(\lambda) \end{pmatrix}, \quad \text{where } \boldsymbol{w}(\lambda) := -(\boldsymbol{G} - \lambda\boldsymbol{I})^{-1}\boldsymbol{t}, \tag{6}$$

and it holds that

$$\boldsymbol{q}(\lambda)^T\boldsymbol{T}\boldsymbol{q}(\mu) = \begin{cases} -f(\lambda) + \lambda f'(\lambda) & \text{for} \quad \lambda = \mu \\ \dfrac{\mu f(\lambda) - \lambda f(\mu)}{\lambda - \mu} & \text{for} \quad \lambda \ne \mu \end{cases}$$

and

$$\boldsymbol{q}(\lambda)^T\boldsymbol{q}(\mu) = \begin{cases} f'(\lambda) & \text{for} \quad \lambda = \mu \\ \dfrac{f(\lambda) - f(\mu)}{\lambda - \mu} & \text{for} \quad \lambda \ne \mu. \end{cases}$$

Proof: Equation (6) follows immediately from

$$\begin{pmatrix} 1-\lambda & \boldsymbol{t}^T \\ \boldsymbol{t} & \boldsymbol{G}-\lambda\boldsymbol{I} \end{pmatrix} \begin{pmatrix} 1 \\ \boldsymbol{w}(\lambda) \end{pmatrix} = \begin{pmatrix} 1-\lambda+\boldsymbol{t}^T\boldsymbol{w}(\lambda) \\ \boldsymbol{t}+(\boldsymbol{G}-\lambda\boldsymbol{I})\boldsymbol{w}(\lambda) \end{pmatrix} = -f(\lambda)\begin{pmatrix} 1 \\ \boldsymbol{0} \end{pmatrix}.$$

If λ and μ are not in the spectrum of $\boldsymbol{T}$ then

$$\begin{aligned} \boldsymbol{q}(\lambda)^T\boldsymbol{T}\boldsymbol{q}(\mu) &= -f(\mu)\boldsymbol{q}(\lambda)^T(\boldsymbol{T}-\mu\boldsymbol{I}+\mu\boldsymbol{I})(\boldsymbol{T}-\mu\boldsymbol{I})^{-1}\boldsymbol{e}^1 \\ &= -f(\mu)\boldsymbol{q}(\lambda)^T\boldsymbol{e}^1-\mu f(\mu)\boldsymbol{q}(\lambda)^T(\boldsymbol{T}-\mu\boldsymbol{I})^{-1}\boldsymbol{e}^1 \\ &= -f(\mu)+\mu\boldsymbol{q}(\lambda)^T\boldsymbol{q}(\mu). \end{aligned}$$

From the symmetry of $\boldsymbol{T}$ we obtain

$$\boldsymbol{q}(\lambda)^T\boldsymbol{T}\boldsymbol{q}(\mu) = -f(\lambda)+\lambda\boldsymbol{q}(\lambda)^T\boldsymbol{q}(\mu). \tag{7}$$

Therefore for $\lambda \neq \mu$ we get

$$\boldsymbol{q}(\lambda)^T\boldsymbol{q}(\mu) = \frac{f(\lambda)-f(\mu)}{\lambda-\mu}$$

and substituting this expression into equation (7) yields

$$\boldsymbol{q}(\lambda)^T\boldsymbol{T}\boldsymbol{q}(\mu) = \frac{\mu f(\lambda)-\lambda f(\mu)}{\lambda-\mu}.$$

Moreover, for $\lambda = \mu$ we get from (6) and (2)

$$\|\boldsymbol{q}(\lambda)\|_2^2 = 1+\|\boldsymbol{w}(\lambda)\|_2^2 = 1+\|(\boldsymbol{G}-\lambda\boldsymbol{I})\boldsymbol{t}\|_2^2 = f'(\lambda)$$

and from equation (7)

$$\boldsymbol{q}(\lambda)^T\boldsymbol{T}\boldsymbol{q}(\lambda) = -f(\lambda)+\lambda\|\boldsymbol{q}(\lambda)\|_2^2 = -f(\lambda)+\lambda f'(\lambda).$$

∎

Theorem 3:
For $\mu \in (0,\omega_1)$, $\mu \neq \lambda_1$, let $g(\cdot;\mu)$ be the rational Hermitian interpolation of f considered in Theorem 1 and denote by $\hat{\mu}$ the unique root of $g(\lambda;\mu)=0$ in (λ_1, c).

Then $\hat{\mu}$ is the smallest eigenvalue of the projected eigenvalue problem

$$\boldsymbol{Q}^T\boldsymbol{T}\boldsymbol{Q}\boldsymbol{\xi} = \lambda\boldsymbol{Q}^T\boldsymbol{Q}\boldsymbol{\xi}, \tag{8}$$

where

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

and $\boldsymbol{q}(\lambda) \in \mathbb{R}^n$ is defined as in Lemma 2.

Proof: From Lemma 2 with $\lambda = 0$ we immediately get

$$\boldsymbol{Q}^T\boldsymbol{T}\boldsymbol{Q} = \begin{pmatrix} -f(0) & -f(0) \\ -f(0) & -f(\mu)+\mu f'(\mu) \end{pmatrix}, \boldsymbol{Q}^T\boldsymbol{Q} = \begin{pmatrix} f'(0) & \frac{1}{\mu}(f(\mu)-f(0)) \\ \frac{1}{\mu}(f(\mu)-f(0)) & f'(\mu) \end{pmatrix},$$

from which we obtain the characteristic polynomial of the projected eigenvalue problem (8) to be

$$\begin{aligned}\chi(\lambda) &= f(0)(f(\mu) - \mu f'(\mu) - f(0)) \\ &+\lambda\Big(f(0)f'(\mu) + f'(0)f(\mu) - \mu f'(0)f'(\mu) - \frac{2}{\mu}f(0)(f(\mu) - f(0))\Big) \\ &+\lambda^2\Big(f'(0)f'(\mu) - \frac{1}{\mu^2}(f(\mu) - f(0))^2\Big).\end{aligned}$$

The interpolation conditions on g yield

$$b = \frac{1}{\mu^2}\frac{(f(\mu) - f(0) - \mu f'(0))^2}{f'(0) + f'(\mu) - \frac{2}{\mu}(f(\mu) - f(0))}, \quad c = \frac{f(0) - f(\mu) + \mu f'(\mu)}{f'(0) + f'(\mu) - \frac{2}{\mu}(f(\mu) - f(0))},$$

and an easy calculation shows that the equation $g(\lambda) = 0$ is equivalent to $\chi(\lambda) = 0$. ∎

Theorem 2 suggests the following generalization of the method introduced in [12]:

(i) Choose parameters $\mu_1, \ldots, \mu_k$ (which are not in the spectrum of $\boldsymbol{T}$ and of $\boldsymbol{G}$) and solve the linear systems

$$(\boldsymbol{T} - \mu_k \boldsymbol{I})\boldsymbol{q}^k = -f(\mu_k)\boldsymbol{e}^1 \quad \text{for } \boldsymbol{q}^k.$$

(ii) Determine the smallest eigenvalue of the projected problem

$$\boldsymbol{Q}_k^T \boldsymbol{T} \boldsymbol{Q}_k \boldsymbol{\xi} = \lambda \boldsymbol{Q}_k^T \boldsymbol{Q}_k \boldsymbol{\xi}$$

where

$$\boldsymbol{Q}_k = (\boldsymbol{q}^1, \ldots, \boldsymbol{q}^k) \in \mathbb{R}^{(n,k)}.$$

From Lemma 2 the entries of the projected matrices $\boldsymbol{A}_k := \boldsymbol{Q}_k^T \boldsymbol{T} \boldsymbol{Q}_k$ and $\boldsymbol{B}_k := \boldsymbol{Q}_k^T \boldsymbol{Q}_k$ are found to be

$$a_{ij} = \begin{cases} -f(\mu_i) + \mu_i f'(\mu_i) & \text{for} \quad i = j \\ \dfrac{\mu_i f(\mu_j) - \mu_j f(\mu_i)}{\mu_j - \mu_i} & \text{for} \quad i \neq j \end{cases} \tag{9}$$

and

$$b_{ij} = \begin{cases} f'(\mu_i) & \text{for} \quad i = j \\ \dfrac{f(\mu_i) - f(\mu_j)}{\mu_i - \mu_j} & \text{for} \quad i \neq j \end{cases} \tag{10}$$

By (1) and (2), increasing the dimension of the projected problem by one essentially requires the same cost as one step of the algorithm in [12]: The solution of one Yule-Walker system (3), and the evaluation of the two scalar products $\|\boldsymbol{w}^k\|_2^2$ and $\boldsymbol{t}^T \boldsymbol{w}^k$.

In the next section the parameters μ_k in the projection method are chosen such that we get safe and fast convergence to the minimum eigenvalue of $\boldsymbol{T}$.

4 A model projection method of global and cubic convergence

We first consider a model algorithm for computing the smallest eigenvalue of the Toeplitz matrix $\boldsymbol{T}$.

Let $\mu_1 := 0$ and $\lambda_\ell = 0$ (λ_ℓ denotes the currently best known lower bound of λ_1). Determine the solution $\boldsymbol{w}$ of the linear system $\boldsymbol{G}\boldsymbol{w} = -\boldsymbol{t}$, compute $f(0) = -1 - \boldsymbol{t}^T\boldsymbol{w}$ and $f'(0) = 1 + \boldsymbol{w}^T\boldsymbol{w}$, and set

$$\boldsymbol{A}_1 := (-f(0)) \in \mathbb{R}^{(1,1)} \quad \text{and} \quad \boldsymbol{B}_1 := (f'(0)) \in \mathbb{R}^{(1,1)}$$

and

$$\lambda_u := \frac{-f(0)}{f'(0)}.$$

Since λ_u is the value of the Rayleigh quotient of $\boldsymbol{T}$ at $\boldsymbol{q}^1 := -f(0)\boldsymbol{T}^{-1}\boldsymbol{e}^1$ it is an upper bound of λ_1.
Choose any $\mu_2 \in (0, \lambda_u]$ and set $k := 2$.

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

(i) Solve

$$(\boldsymbol{G} - \mu_k \boldsymbol{I})\boldsymbol{w} = -\boldsymbol{t}$$

(e.g. by Durbin's algorithm) and determine (e.g. in the course of Durbin's algorithm) which of the intervals $(0, \lambda_1)$, (λ_1, ω_1) and (ω_1, ∞), respectively, contains the parameter μ_k (We do not take into account the very unlikely situation that $\mu_k \in \{\lambda_1, \omega_1\}$; in the first case a lucky break down would have occurred, and the algorithm could be stopped, in the latter case μ_k would be perturbed and the algorithm would be continued).

(ii) If $\mu_k > \omega_1$ then set

$$\lambda_u := \min\{\lambda_u\,,\, \mu_k\} \quad \text{and} \quad \mu_k := 0.5(\lambda_\ell + \lambda_u)$$

else

compute $f(\mu_k) = -1 + \mu_k - \boldsymbol{t}^T\boldsymbol{w}$ and $f'(\mu_k) = 1 + \boldsymbol{w}^T\boldsymbol{w}$, update the matrices $\boldsymbol{A}_k$ and $\boldsymbol{B}_k$ and compute the smallest eigenvalue μ_{k+1} of the projected problem

$$\boldsymbol{A}_k \boldsymbol{\xi} = \mu \boldsymbol{B}_k \boldsymbol{\xi}. \tag{11}$$

If $\mu_k < \lambda_1$ then set $\lambda_\ell := \mu_k$.
$k := k + 1$.

Convergence of the sequence $\{\mu_k\}$ to λ_1 is obtained by comparison with the method of section 2.

Theorem 4: *The model algorithm converges eventually monotonely decreasing to* λ_1*, and the convergence is eventually faster than that of the algorithm based on rational Hermitian interpolation (cf. Section 2), i.e. there exists* $m \in \mathbb{N}$ *such that* $\mu_{k+1} \le \mu_k$ *for every* $k \ge m$ *and if* $\tilde{\mu}_{k+1} \in (0, \mu_k)$ *denotes the unique root of* $g(\mu; \mu_k) = 0$ *where* g *is the rational function considered in Theorem 1 then it holds that*

$$\lambda_1 \le \mu_{k+1} \le \tilde{\mu}_{k+1}.$$

Proof: Obviously, after a finite number of initial steps we obtain $\mu_2 < \omega_1$. Moreover, since the minimal eigenvalue of any projected problem is an upper bound of λ_1 after a finite number of steps we arrive at $\mu_m \in (\lambda_1, \omega_1)$.

For $k \ge m$ let $\mu_k \in (\lambda_1, \omega_1)$. By Theorem 3 $\tilde{\mu}_{k+1}$ is the minimal eigenvalue of the projected eigenvalue problem

$$\tilde{\boldsymbol{Q}}_k^T \boldsymbol{T} \tilde{\boldsymbol{Q}}_k \boldsymbol{\eta} = \mu \tilde{\boldsymbol{Q}}_k^T \tilde{\boldsymbol{Q}}_k \boldsymbol{\eta}$$

where $\tilde{\boldsymbol{Q}}_k = (\boldsymbol{q}(0), \boldsymbol{q}(\mu_k)) \in \mathbb{R}^{(n,2)}$. Since the columns of $\tilde{\boldsymbol{Q}}_k$ are columns of $\boldsymbol{Q}_k$, too, we obtain from Rayleigh's principle

$$\lambda_1 \le \mu_{k+1} \le \tilde{\mu}_{k+1}.$$

∎

From Theorem 4 and [12] we obtain that the model algorithm converges at least quadratically to λ_1. Comparing it to the Rayleigh quotient iteration one even gets at least cubic convergence.

Theorem 5: *The sequence* $\{\mu_k\}$ *constructed by the model algorithm converges at least cubically to* λ_1.

Proof: We first note that for μ, κ, $\mu \ne \kappa$, which are not in the spectrum of $\boldsymbol{T}$

$$\frac{1}{\mu - \kappa}\Big((\boldsymbol{T} - \mu \boldsymbol{I})^{-1} - (\boldsymbol{T} - \kappa \boldsymbol{I})^{-1}\Big) = (\boldsymbol{T} - \mu \boldsymbol{I})^{-1} (\boldsymbol{T} - \kappa \boldsymbol{I})^{-1}. \tag{12}$$

To prove this equation just multiply it by $\boldsymbol{T} - \mu \boldsymbol{I}$ from the left and by $\boldsymbol{T} - \kappa \boldsymbol{I}$ from the right.

Let $V_k := \operatorname{span}\{\boldsymbol{q}^1, \dots, \boldsymbol{q}^k\}$ where $\boldsymbol{q}^j = (\boldsymbol{T} - \mu_j \boldsymbol{I})^{-1} \boldsymbol{e}^1$, and denote the Rayleigh quotient of $\boldsymbol{T}$ at $\boldsymbol{x} \in \mathbb{R}^n \setminus \{\boldsymbol{0}\}$ by $R(\boldsymbol{x})$.

By construction $\mu_k = \min\{R(\boldsymbol{x}) : \boldsymbol{x} \in V_{k-1}\}$. Let $\boldsymbol{x}^k \in V_{k-1}$ such that $R(\boldsymbol{x}^k) = \mu_k$ and denote by

$$\boldsymbol{u}^k := (\boldsymbol{T} - \mu_k \boldsymbol{I})^{-1} \boldsymbol{x}^k$$

the result of 1 step of the Rayleigh quotient iteration with initial guess $\boldsymbol{x}^k$. From the cubic convergence of the Rayleigh quotient iteration (cf. Parlett [14], p. 72 ff) we get the existence of some $C > 0$ such that

$$0 \le R(\boldsymbol{u}^k) - \lambda_1 \le C(\mu_k - \lambda_1)^3. \tag{13}$$

From

$$\boldsymbol{x}^k := \sum_{j=1}^{k-1} \alpha_j (\boldsymbol{T} - \mu_j \boldsymbol{I})^{-1} \boldsymbol{e}^1 \in V_{k-1}$$

and equation (12) we obtain

$$\begin{aligned}
\boldsymbol{u}^k &= (\boldsymbol{T} - \mu_k \boldsymbol{I})^{-1} \boldsymbol{x}^k = \sum_{j=1}^{k-1} \alpha_j (\boldsymbol{T} - \mu_k \boldsymbol{I})^{-1} (\boldsymbol{T} - \mu_j \boldsymbol{I})^{-1} \boldsymbol{e}^1 \\
&= \sum_{j=1}^{k-1} \alpha_j \frac{1}{\mu_k - \mu_j} \Big((\boldsymbol{T} - \mu_k \boldsymbol{I})^{-1} - (\boldsymbol{T} - \mu_j \boldsymbol{I})^{-1} \Big) \boldsymbol{e}^1 \in V_k.
\end{aligned}$$

Hence

$$\mu_{k+1} = \min\{R(\boldsymbol{x}) \,:\, \boldsymbol{x} \in V_k\} \le R(\boldsymbol{u}^k),$$

and inequality (13) yields

$$\mu_{k+1} - \lambda_1 \le R(\boldsymbol{u}^k) - \lambda_1 \le C(\mu_k - \lambda_1)^3,$$

i.e. the sequence $\{\mu_k\}$ converges at least cubically to λ_1 ∎

5 An implementable projection method

In our final algorithm we introduce two modifications which improve the performance of the method.

Especially if the dimension n of the problem is very large the gap between the smallest eigenvalue λ_1 of $\boldsymbol{T}$ and the smallest eigenvalue ω_1 of the submatrix $\boldsymbol{G}$ can be very small. In this situation it may happen that the model algorithm bounces between upper bounds $\mu_k > \omega_1$ which are obtained from projected problems and lower bounds $\mu_{k+1} < \lambda_1$ from bisection steps several times before entering the interval (λ_1, ω_1) and then converging monotonely to λ_1.

To break a tie like this we introduced the following modification: Updating the matrices $\boldsymbol{A}_k$ and $\boldsymbol{B}_k$ we have already evaluated $f(\mu_k)$ and $f'(\mu_k)$. Hence, along with the minimal eigenvalue μ_{k+1} of the projected problem (11) we can obtain the Newton iterate $\hat{\mu}_{k+1}$ of $f(\mu) = 0$ with initial guess μ_k at negligible cost. μ_{k+1} and $\hat{\mu}_{k+1}$ are approximations to λ_1 with errors

$$\hat{\mu}_{k+1} - \lambda_1 = O(|\lambda_1 - \mu_k|^2)\ , \quad \mu_{k+1} - \lambda_1 = O(|\lambda_1 - \mu_k|^3).$$

Hence the relative difference $(\hat{\mu}_{k+1} - \mu_{k+1})/\mu_{k+1}$ estimates the relative error of a Newton step with initial guess μ_k, and therefore is an indicator whether μ_k is close to λ_1 or not. For

$$\frac{\hat{\mu}_{k+1} - \mu_{k+1}}{\mu_{k+1}} < 0.1$$

we continue the projection method with the parameter μ_{k+1} otherwise we choose $\mu_{k+1} := 0.1\mu_k + 0.9\mu_{k+1}$.

In some examples it happened that the projected mass matrix $\boldsymbol{Q}_k^T \boldsymbol{Q}_k$ from (10) was not positive definite (at least numerically). This situation occurred when the current parameter μ_k was already a very accurate approximation to λ_1. It was due to the fact that some parameters μ_j in use were close to each other. Hence the angle between the corresponding columns $\boldsymbol{q}^j$ of $\boldsymbol{Q}_k$ was very small and $\boldsymbol{Q}_k$ was very badly conditioned.

One way out was the direct calculation of the inner products $(\boldsymbol{q}^j)^T \boldsymbol{q}^i$. We preferred to replace the projection to $V_k = \text{span}\{\boldsymbol{q}^1, \ldots, \boldsymbol{q}^k\}$ by a projection to the 2–dimensional space $\tilde{V}_k = \text{span}\{\boldsymbol{q}^\ell, \boldsymbol{q}^k\}$, where $\boldsymbol{q}^\ell = -f(\mu_\ell)(\boldsymbol{T} - \mu_\ell)^{-1}\boldsymbol{e}^1$ and μ_ℓ is the maximal lower bound of λ_1 produced in the algorithm. This modification clearly destroys the cubic convergence of the method. Notice however, that in all examples that we considered after a modified step our accuracy requirement (relative error of μ_k less than 10^{-6}) was satisfied.

In the following we give a MATLAB program for the determination of the smallest eigenvalue of a symmetric and positive definite Toeplitz matrix based on the considerations above.

Therein $[w, where] = durbin(\mu)$ denotes a function which for a given test parameter μ returns the integer variable

$$\text{where} = \begin{cases} 0 & , \quad \text{if } \mu \in (0, \lambda_1) \\ 1 & , \quad \text{if } \mu \in [\lambda_1, \omega_1) \\ 2 & , \quad \text{if } \mu \in (\omega_1, \infty) \end{cases}$$

and for $\mu \in (0, \omega_1)$ additionally the solution $\boldsymbol{w}$ of the Yule-Walker system $(\boldsymbol{G} - \mu \boldsymbol{I})\boldsymbol{w} = \boldsymbol{t}$. Notice that in the case $\mu > \omega_1$ the Durbin algorithm is terminated as soon as a negative diagonal element E_j is detected. Hence, for $\mu \in (0, \omega_1)$ a call of *durbin* needs $2n^2$ flops, for $\mu > \omega_1$ it needs less than $2n^2$ flops.

The procedure $[a, b, min\_ev, boole] = pro\_ev(a, b, mu, f, df, k, nu)$ updates the projected matrices a and b, and it returns the minimal eigenvalue min_ev of the projected problem (11). If the projected mass matrix b is positive definite, then the boolean variable $boole$ is set to 1, otherwise it is set to 0. In the latter case min_ev is the smallest eigenvalue of the two dimensional projected problem corresponding to the modification explained above.

$\kappa = quadroot(\mu_{k+1}, \nu)$ returns the unique root in (μ_ν, μ_{k+1}) of the quadratic polynomial p satisfying the Hermitian interpolation conditions

$$p(\mu_\nu) = f(\mu_\nu) \ , \ p'(\mu_\nu) = f'(\mu_\nu) \ , \ p(\mu_{k+1}) = f(\mu_{k+1}).$$

It was shown in [12] that κ is a lower bound of λ_1 if $\lambda_1 \in (\mu_\nu, \mu_{k+1})$.

```
[w,where]=durbin(0);
mu(1)=0;
f(1)=-1-t'*w;
df(1)=1+w'*w;
a(1,1)=-f(1);
b(1,1)=df(1);
nu=1;                  %mu(nu) is the maximal lower bound of lambda_1
lau=a(1,1)/b(1,1);     %lau holds the minimal upper bound of lambda_1
mu(2)=lau/(4+0.02*n);  %for the choice of mu(2) see [12]
k=1;
h=1;
while h > mu(k+1)*1.e-6
  [y,where]=durbin(mu(k+1));
  if where == 2
    lau=min(lau,mu(k+1));
    mu(k+1)=0.5*(lau+mu(nu));
  else
    k=k+1;
    f(k)=-1+mu(k)-y'*t;
    df(k)=1+y'*y;
    [a,b,min_ev,boole]=pro_ev(a,b,mu,f,df,k,nu);
    lau=min(lau,min_ev);
    mu(k+1)=lau;
    if where == 0
      h=lau-mu(k);
      if boole == 1
        nu=k;
        end
      la_newt=mu(k)-f(k)/df(k);
      if (la_newt-lau)>0.1*lau
        mu(k+1)=0.9*lau+0.1*mu(k);
        end;
    else
      ga=quadroot(mu(k+1),nu)
      h=mu(k+1)-ga;
      end;
  if boole == 0
    mu(k)=mu(k+1);
    k=k-1;
    end;
  end;
end;
```

Notice that we include a vector $\boldsymbol{q}(\mu_k) = -f(\mu_k)(\boldsymbol{T} - \mu_k \boldsymbol{I})^{-1}\boldsymbol{e}^1$ into the basis of the subspace we are projecting on only if $\mu_k \in (0, \omega_1)$. For these parameters the Toeplitz matrix $\boldsymbol{G} - \mu_k \boldsymbol{I}$ is positive definite and by [4] Durbin's algorithm for the solution of $(\boldsymbol{G} - \mu_k \boldsymbol{I})\boldsymbol{w} = -\boldsymbol{t}$ is stable.

We also experimented with a modification where all vectors that were produced by Durbin's algorithm were included into the basis. Although for parameters $\mu_k > \omega_1$ Durbin's algorithm is unstable (cf. [4], [3]) we did not observe any stability problems.

6 Numerical Experiments

To test the projection methods we considered Toeplitz matrices

$$\boldsymbol{T} = m \sum_{k=1}^{n} \eta_k \boldsymbol{T}_{2\pi\theta_k} \tag{14}$$

where m is chosen such that $\boldsymbol{T}$ has normalized diagonal,

$$\boldsymbol{T}_\theta = (t_{ij}) = (\cos(\theta(i-j)))$$

and η_k and θ_k are uniformly distributed random numbers taken from $[0, 1]$ (cf. Cybenko, Van Loan [5]).

Table 1 contains the number of flops needed for 100 test problems with each of the dimensions $n = 64$, 128, 256, 512 and 1024 for three methods: The method based on rational Hermitian interpolation, the projection method where the vector $\boldsymbol{q}(\mu_j)$ was included into the basis of V_k only if $\mu_j \in [0, \omega_1)$ (stable projection) and the projection method where every $\boldsymbol{q}(\mu_j)$ was considered (complete projection). Although in the last case stability of Durbin's algorithm is not guaranteed we did not observe unstable behaviour. The iteration was terminated if the relative error was less than 10^{-6}.

dimension	rational approximation	stable projection	complete projection
64	4.41 $E6$	4.09 $E6$ (92.9%)	3.90 $E6$ (88.4%)
128	1.68 $E7$	1.55 $E7$ (92.5%)	1.50 $E7$ (89.3%)
256	7.30 $E7$	6.66 $E7$ (91.3%)	6.23 $E7$ (85.3%)
512	3.22 $E8$	2.77 $E8$ (86.0%)	2.53 $E8$ (78.4%)
1024	1.33 $E9$	1.14 $E9$ (86.1%)	1.05 $E9$ (78.7%)

Tab. 1. number of flops for 100 test examples

7 Concluding remarks

We have presented an algorithm for the computation of the minimum eigenvalue of a symmetric and positive definite Toeplitz matrix which improves the method of

Cybenko and Van Loan considerably. Realistic and rigorous error bounds are obtained at negligible cost. In our numerical tests we used Durbin's algorithm to solve Yule-Walker systems and to determine the diagonal matrix in the decomposition (4). These informations can be gained from superfast Toeplitz solvers (cf. [1], [2], [6]) 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.), Rational Approximation and its Applications in Mathematics and Physics. Lecture Notes in Mathematics 1237, pp. 315 — 330, Berlin 1987

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

[3] J.R. Bunch, Stability of methods for solving Toeplitz systems of equations. SIAM J. Sci. Stat. Comput. 6 : 349 — 364 (1985)

[4] G. Cybenko, The numerical stability of the Levinson-Durbin algorithm for Toeplitz systems of equations. SIAM J. Sci. Stat. Comput. 1 : 303 — 309 (1980)

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

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

[7] G.H. Golub and C.F. Van Loan, Matrix Computations. 2nd edition. The John Hopkins University Press, Baltimore and London, 1989.

[8] 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.

[9] Y.H. Hu and S.-Y. Kung, Toeplitz eigensystem solver. IEEE Trans. Acoustics, Speech, Signal Processing 33 : 1264 — 1271 (1985)

[10] T. Huckle, Computing the minimum eigenvalue of a symmetric positive definite Toeplitz matrix with spectral transformation Lanczos method. In J. Albrecht, L. Collatz, P. Hagedorn, W. Velte (eds.), Numerical Treatment of Eigenvalue Problems, vol. 5, Birkhäuser Verlag, Basel 1991, pp. 109 — 115

[11] T. Huckle, Circulant and skewcirculant matrices for solving Toeplitz matrices. SIAM J. Matr. Anal. Appl. 13 : 767 — 777 (1992)

[12] W. Mackens and H. Voss, The minimum eigenvalue of a symmetric positive definite Toeplitz matrix and rational Hermitian interpolation. To appear in SIAM J. Matr. Anal. Appl.

[13] F. Noor and S.D. Morgera, Recursive and iterative algorithms for computing eigenvalues of Hermitian Toeplitz matrices. IEEE Trans. Signal Processing 41 : 1272 — 1280 (1993)

[14] B.N. Parlett, The Symmetric Eigenvalue Problem. Prentice—Hall, Englewood Cliffs 1980

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

[16] W.F. Trench, Numerical solution of the eigenvalue problem for Hermitian Toeplitz matrices. SIAM J. Matr. Anal. Appl. 10 : 135 — 146 (1989)

[17] H. Voss and W. Mackens, Computing the minimal eigenvalue of a symmetric Toeplitz matrix. ZAMM 76, Proceedings of ICIAM/GAMM95, Issue 2: Applied Analysis, O. Mahrenholtz, R. Mennicken (eds.), pp. 701 — 702 (1996)