

Research Article

A Dual Deflation Method for the Computation of Eigenmodes of Positive Definite Matrices

Pravin Singh^{1*}, Shivani Singh², Virath Singh¹

¹Department of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, Durban 4001, South Africa

²Department of Decision Science, University of South Africa, Pretoria 0003, South Africa

E-mail: singhprook@gmail.com

Received: 15 September 2023; **Revised:** 6 December 2023; **Accepted:** 19 December 2023

Abstract: In this paper we advocate a new method to compute the eigenmodes of real symmetric positive definite matrices. Our method involves the zeros of a quartic polynomial, shifted inverse iteration and deflation of two eigenmodes simultaneously.

Keywords: eigenvalues, symmetric positive definite, deflation

MSC: 01A01, 22B22, 31K13

1. Introduction

The importance of eigenvalues and eigenvectors of a matrix cannot be overemphasized. They inherently feature in almost all spheres of the sciences. Recently the eigenvalues of a matrix was used to analyse peak viral load of the SARS-Cov2 virus [1]. The PageRank $PR(\mathbf{A})$ is calculated using the power method to determine the principal eigenvector of the Google matrix [2]. The spread $sp(\mathbf{A})$ is indispensable in graph theory applications [3]. Eigenvalues are especially useful in data science [4]. While the algebra of the characteristic polynomial is interesting, it is by no means an easy task to determine and find its zeros, especially for moderate to large size matrices. For real symmetric matrices the spectrum $\sigma(\mathbf{A})$ is real and so are the eigenvectors. It is thus easier to determine crude bounds for $\sigma(\mathbf{A})$ using the Gerschgorin circle theorem and the Ovals of Cassini [5]. In most cases superior bounds are achieved by using the trace $tr(\mathbf{A})$ ([6–8]). Some applications require just the minimal and maximal eigenvalues and these may be achieved by the well known power iteration [9]. In fact several variants of the latter method exist. For positive definite symmetric matrices $\sigma(\mathbf{A}) > 0$ and the conditioning of the associated linear system is given by the ratio $\frac{\lambda_1}{\lambda_n}$, where λ_1 is the largest eigenvalue and λ_n the smallest one. The symmetry in nature, after mathematical modelling is preserved in some cases, and presents itself as a symmetric linear operator with real eigenvalues. While there exists a plethora of techniques to determine $\sigma(\mathbf{A})$ for special cases, these are most often based on iteration. Together with the application of Rayleigh's theorem [10], this provides a fairly accurate means of determining $\sigma(\mathbf{A})$. In some cases all eigenmodes are desired and deflation [11], being stable, is the preferred technique. It is well known that the eigenbasis represents a canonical axes system, in which the linear operator of a normal linear system is diagonalizable. Here we present a simple algorithm, to approximate $\sigma(\mathbf{A})$, for real positive definite symmetric matrices, with distinct eigenvalues.

2. Theory

Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be a positive definite symmetric matrix, with spectrum $\sigma(\mathbf{A}) = \{\lambda_i\}_{i=1}^n$. The associated normalized eigenvectors are denoted by $\{\mathbf{u}_i\}_{i=1}^n$, $\|\mathbf{u}_i\|_2 = 1$. We shall denote the i th eigenmode by $(\lambda_i, \mathbf{u}_i)$. It is well known that \mathbf{A} is unitarily diagonalizable [10]. Thus it follows from the spectral theorem that

$$\mathbf{I} = \sum_{i=1}^n \mathbf{G}_i \tag{1}$$

$$\mathbf{A} = \sum_{i=1}^n \lambda_i \mathbf{G}_i.$$

Here \mathbf{G}_i is the orthogonal projector onto the nullspace $N(\mathbf{A} - \lambda_i \mathbf{I})$ along the range $R(\mathbf{A} - \lambda_i \mathbf{I})$, $\mathbf{G}_i = \mathbf{u}_i \mathbf{u}_i^T$ and satisfies $\mathbf{G}_i \mathbf{G}_j = \delta_{ij} \mathbf{I}$, where δ_{ij} denote the well known Kronecker delta symbol. Suppose that we have computed the eigenmodes $\{(\lambda_k, \mathbf{u}_k)\}_{k=1}^{2i}$, $i \in \{1, 2, \dots, \lfloor \frac{n}{2} \rfloor - 1\}$, then define the deflated matrix \mathbf{B}_{2i} by

$$\mathbf{B}_{2i} = \mathbf{A} - \sum_{k=1}^{2i} \lambda_k \mathbf{G}_k. \tag{2}$$

Theorem 1 Let \mathbf{B}_{2i} be as defined in (2). Let $\mathbf{x}_0 \in \mathbb{R}^n$ be chosen randomly with $\|\mathbf{x}_0\|_2 = 1$. Choose $p \in \mathbb{N}$ and evaluate

$$\mathbf{x}_j = \mathbf{B}_{2i} \mathbf{x}_{j-1}, \quad j = 1, 2, \dots, p+3. \tag{3}$$

Define $\alpha_p = \|\mathbf{x}_p\|_2^2$, then good approximations to λ_{2i+1} and λ_{2i+2} are the positive zeros of the quartic polynomial

$$\begin{aligned} P(\lambda) &= (\alpha_{p+1}^2 - \alpha_p \alpha_{p+2}) \lambda^4 + (\alpha_p \alpha_{p+3} - \alpha_{p+1} \alpha_{p+2}) \lambda^2 \\ &+ (\alpha_{p+2}^2 - \alpha_{p+1} \alpha_{p+3}). \end{aligned} \tag{4}$$

Proof. Note that from (1) and (2) we have that

$$\begin{aligned} \mathbf{B}_{2i} &= \sum_{k=2i+1}^n \lambda_k \mathbf{G}_k \\ \mathbf{B}_{2i}^p &= \sum_{k=2i+1}^n \lambda_k^p \mathbf{G}_k \\ &= \lambda_{2i+1}^p \mathbf{G}_{2i+1} + \lambda_{2i+2}^p \mathbf{G}_{2i+2} + \boldsymbol{\varepsilon}_p, \end{aligned} \tag{5}$$

where $\boldsymbol{\varepsilon}_p = \sum_{k=2i+3}^n \lambda_k^p \mathbf{G}_k$. From (3) it follows that $\mathbf{x}_p = \mathbf{B}_{2i}^p \mathbf{x}_0$, so that from (5) (ignoring $\boldsymbol{\varepsilon}_p$) we have

$$\begin{aligned}\alpha_p &= \lambda_{2i+1}^{2p} \|\mathbf{G}_{2i+1} \mathbf{x}_0\|_2^2 + \lambda_{2i+2}^{2p} \|\mathbf{G}_{2i+2} \mathbf{x}_0\|_2^2 \\ &= c_{2i+1} \lambda_{2i+1}^{2p} + c_{2i+2} \lambda_{2i+2}^{2p},\end{aligned}\tag{6}$$

where $c_{2i+1} = \|\mathbf{G}_{2i+1} \mathbf{x}_0\|_2^2$ and $c_{2i+2} = \|\mathbf{G}_{2i+2} \mathbf{x}_0\|_2^2$. Thus

$$\alpha_{p+1} = c_{2i+1} \lambda_{2i+1}^{2p+2} + c_{2i+2} \lambda_{2i+2}^{2p+2}\tag{7}$$

$$\alpha_{p+2} = c_{2i+1} \lambda_{2i+1}^{2p+4} + c_{2i+2} \lambda_{2i+2}^{2p+4}\tag{8}$$

$$\alpha_{p+3} = c_{2i+1} \lambda_{2i+1}^{2p+6} + c_{2i+2} \lambda_{2i+2}^{2p+6},\tag{9}$$

Multiply (6) by λ_{2i+1}^2 and subtract from (7). Follow a similar procedure for remaining equations (7)–(9) to obtain

$$\alpha_{p+1} - \lambda_{2i+2}^2 \alpha_p = c_{2i+1} \lambda_{2i+1}^{2p} (\lambda_{2i+1}^2 - \lambda_{2i+2}^2)\tag{10}$$

$$\alpha_{p+2} - \lambda_{2i+2}^2 \alpha_{p+1} = c_{2i+1} \lambda_{2i+1}^{2p+2} (\lambda_{2i+1}^2 - \lambda_{2i+2}^2)\tag{11}$$

$$\alpha_{p+3} - \lambda_{2i+2}^2 \alpha_{p+2} = c_{2i+1} \lambda_{2i+1}^{2p+4} (\lambda_{2i+1}^2 - \lambda_{2i+2}^2)\tag{12}$$

Taking ratios in (10)–(12) results in

$$\begin{aligned}\frac{\alpha_{p+2} - \lambda_{2i+2}^2 \alpha_{p+1}}{\alpha_{p+1} - \lambda_{2i+2}^2 \alpha_p} &= \lambda_{2i+1}^2 \\ &= \frac{\alpha_{p+3} - \lambda_{2i+2}^2 \alpha_{p+2}}{\alpha_{p+2} - \lambda_{2i+2}^2 \alpha_{p+1}}\end{aligned}\tag{13}$$

Equation (13) gives the quartic (4) in λ_{2i+2} . From symmetry it is not difficult to see that λ_{2i+1} satisfies the same quartic. Let $z = \lambda^2$ in (4) and write the quartic as

$$az^2 + bz^2 + c,$$

where

$$a = (\alpha_{p+1}^2 - \alpha_p \alpha_{p+2})\tag{14}$$

$$\begin{aligned}
b &= (\alpha_p \alpha_{p+3} - \alpha_{p+1} \alpha_{p+2}) \\
c &= (\alpha_{p+2}^2 - \alpha_{p+1} \alpha_{p+3}).
\end{aligned}
\tag{15}$$

Then from (6)–(9) it can easily be shown that

$$\begin{aligned}
a &= -c_{2i+1} c_{2i+2} \lambda_{2i+1}^{2p} \lambda_{2i+2}^{2p} (\lambda_{2i+1}^2 - \lambda_{2i+2}^2)^2 \\
b &= c_{2i+1} c_{2i+2} \lambda_{2i+1}^{2p} \lambda_{2i+2}^{2p} (\lambda_{2i+1}^2 - \lambda_{2i+2}^2)^2 (\lambda_{2i+1}^2 + \lambda_{2i+2}^2) \\
c &= -c_{2i+1} c_{2i+2} \lambda_{2i+1}^{2p+2} \lambda_{2i+2}^{2p+2} (\lambda_{2i+1}^2 - \lambda_{2i+2}^2)^2
\end{aligned}$$

Thus it follows that the discriminant is given by

$$c_{2i+1}^2 c_{2i+2}^2 \lambda_{2i+1}^{4p} \lambda_{2i+2}^{4p} (\lambda_{2i+1}^2 - \lambda_{2i+2}^2)^6$$

and is positive. Since $a < 0$, the larger root is given by

$$\lambda_{2i+1} = \sqrt{\frac{-b - \sqrt{b^2 - 4ac}}{2a}} \tag{16}$$

and the smaller root is given by

$$\lambda_{2i+2} = \sqrt{\frac{-b + \sqrt{b^2 - 4ac}}{2a}} \tag{17}$$

We now write α_p from (6) as

$$\begin{aligned}
\alpha_p &= c_{2i+1} \lambda_{2i+1}^{2p} + \lambda_{2i+2}^{2p} \left(c_{2i+2} + c_{2i+3} \left(\frac{\lambda_{2i+3}}{\lambda_{2i+2}} \right)^{2p} \right) \\
&= c_{2i+1} \lambda_{2i+1}^{2p} + c_{2i+2} \lambda_{2i+2}^{2p} + \mathcal{O}(\varepsilon^{2p}),
\end{aligned}$$

where $\varepsilon = \frac{\lambda_{2i+3}}{\lambda_{2i+2}} < 1$. Replacing α_p by $\alpha_p + \mathcal{O}(\varepsilon^{2p})$ in (14) we obtain

$$\begin{aligned}
 a' &= [\alpha_{p+1} + \mathcal{O}(\varepsilon^{2p+2})]^2 - [\alpha_p + \mathcal{O}(\varepsilon^{2p})] [\alpha_{p+2} + \mathcal{O}(\varepsilon^{2p+4})] \\
 &= a + \mathcal{O}(\varepsilon^{2p}).
 \end{aligned}$$

Similarly

$$b' = b + \mathcal{O}(\varepsilon^{2p})$$

$$c' = c + \mathcal{O}(\varepsilon^{2p+2}).$$

Thus from (4) we see that a more accurate quartic polynomial is given by

$$\begin{aligned}
 \tilde{P}(\lambda) &= a'\lambda^4 + b'\lambda^2 + c' \\
 &= P(\lambda) + \mathcal{O}(\varepsilon^{2p}) [\lambda^4 + \lambda^2 + \mathcal{O}(\varepsilon^2)]
 \end{aligned} \tag{18}$$

From (18) it is evident that such a small perturbation will have minimal effect on the zeros of $P(\lambda)$. □

We shall denote by $\sigma_{app}(\mathbf{A})$ the approximate eigenvalues obtained from (16) and (17). We now present a simple algorithm to implement our method. We shall denote by $\sigma_{alg}(\mathbf{A})$ the eigenvalues determined using Algorithm 1.

Algorithm 1: Eigenmode Computation

- 1: choose $\mathbf{x}_0 \in \mathbb{R}^n$ randomly, $\|\mathbf{x}_0\|_2 = 1$, and p (we use $p = 4$)
 - 2: for $j = 1$ to $p + 3$ do
 - 3: $\mathbf{x}_j = \mathbf{A}\mathbf{x}_{j-1}$
 - 4: end for
 - 5: for $j = 0$ to 3 do
 - 6: $\alpha_{p+j} = \|\mathbf{x}_{p+j}\|_2^2$
 - 7: end for
 - 8: evaluate a, b and c from (14)-(15)
 - 9: calculate λ_1 and λ_2 from (16) and (17)
 - 10: use two shifted inverse iterations to determine \mathbf{u}_1
 - 11: use two shifted inverse iterations to determine \mathbf{u}_2
 - 12: refine λ_1 using the Rayleigh quotient $\langle \mathbf{A}\mathbf{u}_1, \mathbf{u}_1 \rangle$
 - 13: refine λ_2 using the Rayleigh quotient $\langle \mathbf{A}\mathbf{u}_2, \mathbf{u}_2 \rangle$
 - 14: replace \mathbf{A} by $\mathbf{A} - \lambda_1\mathbf{G}_1 - \lambda_2\mathbf{G}_2$
 - 15: repeat from 2 until all modes are computed
-

3. Examples

Example 1 Consider the real symmetric positive definite matrix [12]

$$\mathbf{A} = \begin{bmatrix} n & n-1 & n-2 & \cdots & 2 & 1 \\ n-1 & n-1 & n-2 & \cdots & 2 & 1 \\ n-2 & n-2 & n-2 & \cdots & 2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 2 & 2 & 2 & 2 & 2 & 2 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

The exact eigenvalues are given by [12]

$$\lambda_k = \frac{1}{2} \left[1 - \frac{\cos(2k-1)}{2n+1} \right]^{-1}, \quad k = 1, 2, \dots, n.$$

The exact eigenvalues $\sigma(\mathbf{A})$, $\sigma_{app}(\mathbf{A})$ as well as $\sigma_{alg}(\mathbf{A})$, for $n = 10$ are summarized in Table 1. The zeros of the quartic give good approximations to the eigenvalues, which are further most accurately refined by the deflation and inverse iteration Algorithm 1. The maximum error is approximately 2.4×10^{-7} , while the minimum error is 9.2×10^{-14} .

Table 1. Eigenvalues: example 1

$\sigma(\mathbf{A})$	$\sigma_{app}(\mathbf{A})$	$\sigma_{alg}(\mathbf{A})$
44.7660686527149565	44.7660276589360180	44.7660686527150489
5.0489173395223066	5.0472178683045188	5.0489173395195417
1.8730230604249107	1.8730142725780632	1.8730230604265663
1.0000000000000000	0.9937216569951369	0.9999998610173921
0.6431041321077906	0.6408851806344559	0.6431038908003185
0.4652330878085649	0.4625709642967941	0.4652331556900618
0.3662088746157992	0.3662088746160075	0.3662088746160205
0.3079785283699041	0.3076864231992443	0.3079785366735136
0.2737867616392449	0.2730371005920817	0.2737867905243412
0.2556795627964359	0.2556795662515789	0.2556795662515796

Example 2 Consider the matrix [12]

$$\mathbf{A} = \begin{bmatrix} 10 & 1 & 2 & 3 & 4 \\ 1 & 9 & -1 & 2 & 3 \\ 2 & -1 & 7 & 3 & -5 \\ 3 & 2 & 3 & 12 & -1 \\ 4 & -3 & -5 & -1 & 15 \end{bmatrix},$$

where n is odd. The distinct eigenvalues are calculated quite accurately using Julia 1.8 and summarized in Table 2. The maximum error is approximately 2.6×10^{-5} , while the minimum error is 2.0×10^{-10} . The last eigenvalue is calculated by few iterations of the power method since n is odd.

Table 2. Eigenvalues: example 2

$\sigma(\mathbf{A})$	$\sigma_{app}(\mathbf{A})$	$\sigma_{alg}(\mathbf{A})$
19.1754202772797377	19.1054892564292267	19.1753940849438997
15.8089207643904945	15.7818151319803164	15.8089207641914840
9.3655549201061383	9.3655548599178182	9.3655552314781527
6.9948378304964773	6.9948270999095241	6.9948381621132745
1.6552662077271696		1.6552606029685881

Example 3 As a final example we consider the matrix \mathbf{A} given by $\mathbf{A} = \text{diag}(0, 1, \dots, n-1) + 0.5\mathbf{e}\mathbf{e}^t$, where $n = 10$ and $\mathbf{e} = [1, 1, \dots, 1]^t$. The exact eigenvalues are evaluated using Julia 1.8 and presented in Table 3. Here we also present two errors. Error1 denotes the error after applying Algorithm 1, while Error 2 denotes the error after applying a refined version of Algorithm 1. This refined version uses inverse iteration until $\|\mathbf{u} - \mathbf{u}'\|_\infty \leq 10^{-6}$, where \mathbf{u} and \mathbf{u}' are two successive eigenvector iterates. The eigenvalues are almost uniformly spread in $(0, 12)$. Once again the efficiency of our algorithm is verified.

Table 3. Eigenvalues and errors: example 3

$\sigma(\mathbf{A})$	Error 1	Error 2
11.0360760794846389	5.3×10^{-15}	1.8×10^{-15}
8.5182673343854827	2.6×10^{-11}	8.9×10^{-15}
7.4337623653080769	6.6×10^{-11}	2.7×10^{-15}
6.3828680351978209	1.4×10^{-4}	1.7×10^{-14}
5.3452852286876587	1.2×10^{-4}	1.1×10^{-14}
4.3142619092190175	3.3×10^{-4}	3.0×10^{-14}
3.2864482991237578	2.4×10^{-4}	4.3×10^{-14}
2.2594654170579243	3.1×10^{-4}	9.0×10^{-16}
1.2305280809347721	4.5×10^{-5}	1.1×10^{-15}
0.1930372506008435	1.3×10^{-5}	3.0×10^{-16}

4. Conclusion

For a rather modest value of $p = 4$ and two shifted inverse iterations, we are able to achieve fairly accurate results for approximating the spectrum $\sigma(\mathbf{A})$. Much better results are attained, by using more shifted inverse iterations, to accurately determine the eigenvectors. Thus the dual deflation process is much more accurate and the precision propagates as we determine the less dominant eigenvalues. Our process is advocated for finding especially the first few eigenmodes as this gives good results for minimum effort. In this regard tolerances for the accuracy may be preset when using the inverse shifted iteration, however as is well known, too many inverse iterations implies additional cost. There is the possibility of simultaneously calculating $K > 2$ eigenmodes simultaneously, however this leads to the solution of a degree $2K$ polynomial. In this case it is not possible to get an analytic solution as in the $K = 2$ case. Thus additional numerical effort is needed to find the zeros. We have therefore restricted ourselves to the quartic polynomial case.

Conflict of interest

The authors declare no competing financial interest.

References

- [1] Frank TD. Eigenvalue analysis of SARS-CoV-2 viral load data: illustration for eight COVID-19 patients. *International Journal of Data Science and Analytics*. 2023; 15(3): 281-290.
- [2] Gu C, Xie F, Zhang K. A two-step matrix splitting iteration for computing the PageRank. *Journal of Computational and Applied Mathematics*. 2015; 278: 19-28.
- [3] Gregory DA, Hershkowitz D. The spread of the spectrum of a graph. *Linear Algebra and Its Applications*. 2001; 332: 23-35.
- [4] Piedmont RL. Eigenvalues. *Encyclopedia of Quality of Life and Well-Being Research*. Dordrecht, The Netherlands: Springer; 2014. p.1847-1848.
- [5] Brauer A. Limits for the characteristic roots of a matrix. *Duke Mathematical Journal*. 1958; 25: 583-590.
- [6] Wolkowicz H, Styan GPH. Bounds for eigenvalues using traces. *Linear Algebra and its Applications*. 1980; 29: 471-506.
- [7] Sharma R, Kumar R, Saini R. Note on Bounds for Eigenvalues using Traces. *Functional Analysis*. 2014; 1: 1-13.
- [8] Singh P, Singh V, Singh S. New bounds for the maximal eigenvalues of positive definite matrices. *International Journal of Applied Mathematics*. 2022; 35: 685-691.
- [9] Faddeev DK, Faddeeva VN. *Computational Methods of Linear Algebra*. San Francisco, CA, USA; London, UK: WH Freeman and Company; 1963.
- [10] Horn R, Johnson CA. *Matrix Analysis*. Cambridge, UK: Cambridge University Press; 2012.
- [11] Singh KV, Ram YM. Numerical deflation of the transcendental eigenvalue problem. *Mechanical Systems and Signal Processing*. 2017; 83: 522-532.
- [12] Gregory RT, Karney DL. *A Collection of Matrices for Testing Computational Algorithms*. 2nd ed. New York, NY, USA: Robert E. Krieger Publishing Company; 1978.