SF3580

Anna Broms & Fredrik Fryklund

2018/11/11

1 Task 2

As can be seen from Figure 1 the rate of convergence is linear for the power method, as expected. For the Rayleigh quotient iteration the rate of convergence is cubic when A is symmetric and quadratic when A is nonsymmetric. This is also expected. The rate of convergence p for the respective settings is approximated to $p \approx 3.03$ and $p \approx 2.2$, where

$$p = \frac{\log |\lambda^{(k+1)} - \lambda|}{|\lambda^{(k)} - \lambda|} \tag{1}$$

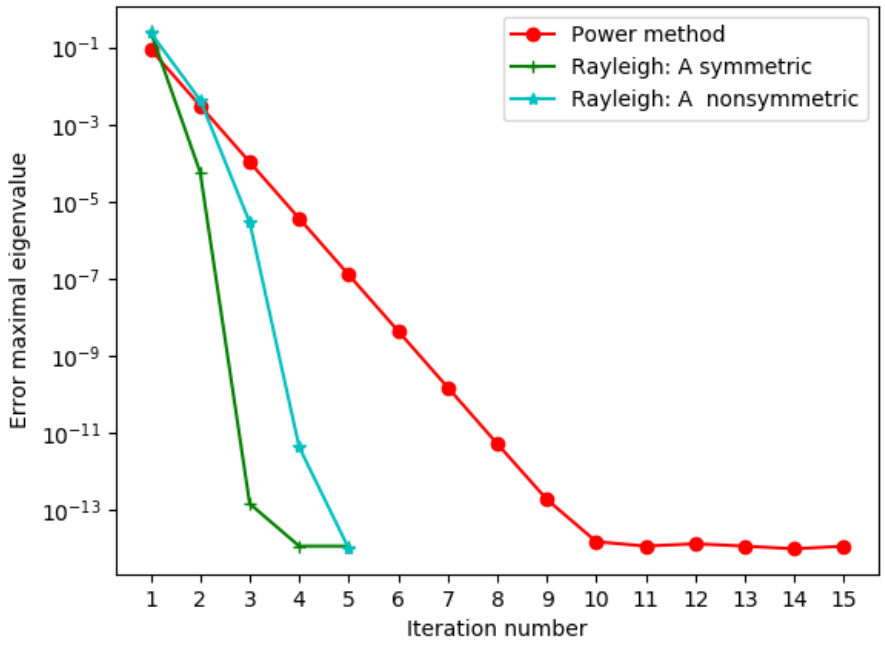

Figure 1: Comparison of eigenvalues after m iterations.

The Rayleigh quotient only uses the symmetric part of A in

$$r(\mathbf{x}) = \mathbf{x}^H A \mathbf{x}$$

assuming $\mathbf{x}$ is normalized. For $a_{13} = 4$ the matrix A is no longer symmetric, i.e. $A \neq A^H$, but any

square matrix can be decomposed into a symmetric part A_s and a nonsymmetric part A_{ns} by

$$A = \underbrace{\frac{1}{2}\left(A + A^H\right)}_{=A_s} + \underbrace{\frac{1}{2}\left(A - A^H\right)}_{=A_{ns}}. \tag{2}$$

Thus

$$r(\mathbf{x}) = \mathbf{x}^H A_s \mathbf{x} + \mathbf{x}^H A_{ns} \mathbf{x} = \mathbf{x}^H A_s \mathbf{x}$$

since

$$\mathbf{x}^H A_{ns} \mathbf{x} = \mathbf{x}^H A \mathbf{x} - \mathbf{x}^H A^H \mathbf{x} = 0.$$

For a nonsymmetric matrix all available information is not used.

2 Task 3

The performance of different versions of Gram-Schmidt orthogonalisations is investigated when combined with the Arnoldi method. We consider the matrix A constructed by

```
Random.seed!(0); A=matrixdepot("wathen",nn,nn)
```
(3)

where we choose $nn = 500$ and use m number of iterations in the Arnoldi method. Results for CPU time and orthogonality of the basis in Q is given in Table 1. It can be concluded that double Gram-Schmidt

Table 1: Comparison for different types of Gram-Schmidt (GS) orthogonalisation in the Arnoldi method: SGS (single GS), MDS (modified GS), DGS (double GS), TGS (triple GS), where *time* is the measured CPU time and *orth* is the orthogonality of the basis in terms of $\|Q_m^T Q_m - I\|$.

m	SGS time	SGS orth	MGS time	MGS orth	DGS time	DGS orth	TGS time	TGS orth
5	188 ms	$4.82 \cdot 10^{-13}$	241 ms	$3.28 \cdot 10^{-15}$	246 ms	$3.81 \cdot 10^{-15}$	298 ms	$3.84 \cdot 10^{-15}$
10	496 ms	$2.15 \cdot 10^{-12}$	705 ms	$1.32 \cdot 10^{-14}$	665 ms	$4.19 \cdot 10^{-15}$	820 ms	$4.23 \cdot 10^{-15}$
20	1.10 s	$1.04 \cdot 10^{-11}$	2.03 s	$6.41 \cdot 10^{-14}$	1.682 s	$5.66 \cdot 10^{-15}$	2.16 s	$5.72 \cdot 10^{-15}$
50	4.05 s	$1.06 \cdot 10^{-10}$	9.95 s	$6.04 \cdot 10^{-13}$	6.69 s	$6.18 \cdot 10^{-15}$	9.49 s	$6.38 \cdot 10^{-15}$
100	12.0 s	$4.77 \cdot 10^{-10}$	37.8 s	$3.29 \cdot 10^{-12}$	21.7 s	$7.82 \cdot 10^{-15}$	31.2 s	$7.87 \cdot 10^{-15}$

performs the best in terms of orthogonalization error, while single Gram-Schmidt is the fastest among the algorithms. Triple Gram-Schmidt performs almost exactly as well as double Gram-Schmidt in terms of error, which is not surprising. It was stated during the lecture that the best result possible to achieve using multiple Gram-Schmidt is indeed obtained for double Gram-Schmidt. Note however that the CPU time required for triple Gram-Schmidt of course is considerably larger than for double Gram-Schmidt. The orthogonalization errors for the modified Gram-Schmidt is smaller than for single Gram-Schmidt but the algorithm is in this example seen to be worse than double Gram-Schmidt both in terms of CPU time and orthogonalization error.

3 Task 4

We investigate a primitive version of the Arnoldi method. Let K_m be a matrix representing the Krylov subspace:

$$K_m = [b, Ab/\|Ab\|, \ldots, A^{m-1}b/\|A^{m-1}b\|] \in \mathbb{R}^{n \times m}. \tag{4}$$

3.1 (a)

The equation

$$\mu K_m^T K_m w = K_m^T A K_m w \tag{5}$$

is stemming from the Galerking method applied to the bilinear form associated with the eigenvalue problem $a(u,v) = u^T Av - \mu u^T v$, *with* $f(v) = 0$. *Prov that* (5) *is identical to the approximation generated by Arnoldi's method for eigenvalue problems.* The Arnoldi method computes an orthogonal basis of K_m such that after m iterations

$$AQ_m = Q_{m+1}H_{\underline{m}}, \tag{6}$$

where Q_m is an orthogonal matrix of size m and $H_{\underline{m}}$ is the corresponding Hessenberg matrix. The eigenvalues of A can then be approximated by the eigenvalues of

$$Q_m^T AQ_m. \tag{7}$$

We use the QR-factorisation of the matrix K_m, such that $K_m = Q_m R$. We first show that $R^T R = K_m^T K_m$. Using orthogonality, we have

$$I = Q_m^T Q_m. \tag{8}$$

Multiplying both sides of (8) with R^T and R respectively yields

$$\begin{aligned} R^T = R^T Q^T Q_m = K_m^T Q_m \Leftrightarrow \\ R^T R = K_m^T Q_m R = K_m^T K_m. \end{aligned} \tag{9}$$

Now, considering (5), we have

$$\begin{aligned} \mu K_m^T K_m w = K_m^T A K_m w \Leftrightarrow \\ \mu R^T R w = K_m^T A K_m w. \end{aligned} \tag{10}$$

Replacing K_m and K_m^T with the QR-factorisation of K_m yields

$$\mu R^T R w = R^T Q_m^T A Q_m R w. \tag{11}$$

Using that R is non-singular, we obtain

$$\mu w = Q_m^T A Q_m w. \tag{12}$$

Thus, the approximation computed from (5) is identical to the eigenvalue approximation obtained from (7), which is what we wanted to show.

3.2 (b)

We compare the eigenvalues computed using the Arnoldi method to eigenvalues computed using (5) after m iterations. Double Gram-Schmidt is used for orthogonalization and the matrix from task 3 is used with $nn = 12$ along with a random starting vector b. The result is visualised in Figure 2.

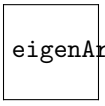

Figure 2: Comparison of eigenvalues after m iterations.

3.3 (c)

In exact arithmetic we expect the results from the two approaches to agree. However, forming the Krylov matrices K_m for larger m gives close to singular matrices. As a result of bad conditioning, the Arnoldi approach is to prefer for computing the eigenvalues of the matrix.

4 Task 6

4.1 (a, b)

We identify three eigenvalues λ_1, λ_2 and λ_3 as outliers in $\mathbb{C}$. A crude estimate for the rate of convergence for eigenvalue λ_i is given by

$$\varepsilon_i^{(m)} \leq \frac{\rho^{m-1}}{|\lambda_i - c|^{m-1}}.$$

Thus the circles with centre c_i and radius ρ_i give the convergence estimate. They are plotted in **??** together with the eigenvalues, where the outliers are marked. The corresponding values are

$$\begin{aligned}
&\lambda_1 \approx -47.016 + i0.166, && c_1 = -5 + i0.5, \rho_1 = 14, && \varepsilon_1^m \leq (0.33)^{m-1} \\
&\lambda_2 = 1.314 + i12.664, && c_2 = -23 - i6, \rho_2 = 25, && \varepsilon_2^m \leq (0.82)^{m-1} \\
&\lambda_3 = 0.986 - i11.898, && c_3 = -23 + i9, \rho_3 = 26, && \varepsilon_2^m \leq (0.82)^{m-1}
\end{aligned}$$

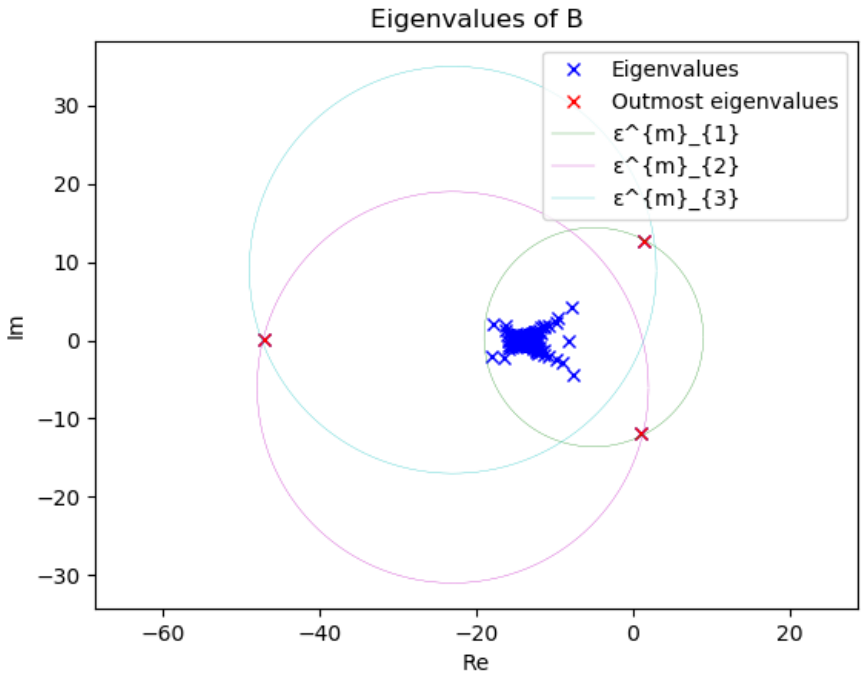

Figure 3: The eigenvalues of B, with the outliers marked and the corresponding circles.

4.2 (c)

We expect the fastest convergence for λ_1, motivated by $\varepsilon_1^m \leq (0.33)^{m-1}$. The eigenvalue error for λ_1 for Arnoldi method, with double GS for $m = 2, 4, 8, 10, 20, 30, 40$ is plotted in 4. About $m = 14$ gives an eigenvalue error less than 10^{-10}.

In 5 we have plotted the evolution of the Ritz values as m increases. As predicted are λ_1, λ_2 and λ_3 the first to converge.

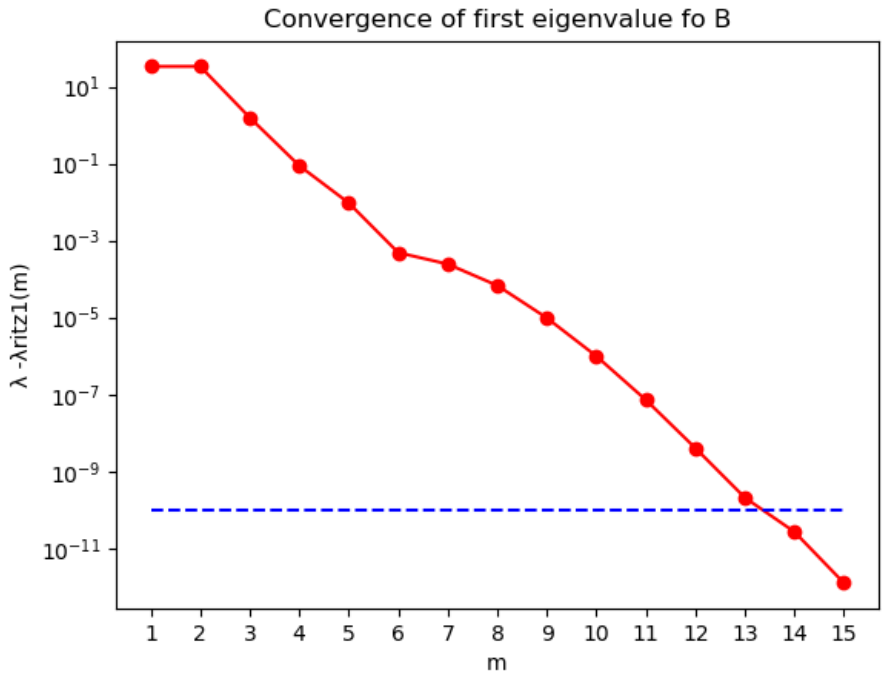

Figure 4: Comparison of eigenvalues after m iterations.

4.3 (d)

4.4 (e)

From Figure 7 we read that about 12 iterations are required for an eigenvalue error about 10^{-10}, i.e. Arnoldi's method with shift-invert is faster than standard Arnoldi's method. This increase in convergence should not be exaggerated, as theory says that they should be about the same. The idea of shift and invert is to transform the spectrum such that eigenvalues close to σ becomes an outlier. But with $\sigma = -11 + i2$ we read from 6 that the outliers are the same as for B, i.e. the transformed spectrum has not changed considerably.

5 Task 7

It is evident from Figure 8–11 that convergence of the approximation of eigenvalues can't be expected. We see that round-off errors heavily affects the convergence, however this is expected for explicit restart.

The idea of restart is to be able to use a smaller m by restarting the Arnoldi method with a new starting vector based on the Ritz vectors from the previous run. In the explicit case the the vector is a linear combination of the Ritz vectors corresponding to the k largest Ritz values in modulus.

In the implicit case the Arnoldi method is not entirely restarted. Instead it only keeps the first p to continue, i.e.

$$AQ_m = Q_{m+1}\underline{H}_m \rightarrow AQ_p = Q_{p+1}\underline{H}_p$$

where Q_p and Q_{p+1} are transformations of Q_m and Q_{m+1}, respectively. From 12–15 we see that the approximation of the eigenvalues indeed converges after 8 and 10 iterations. For the latter we used

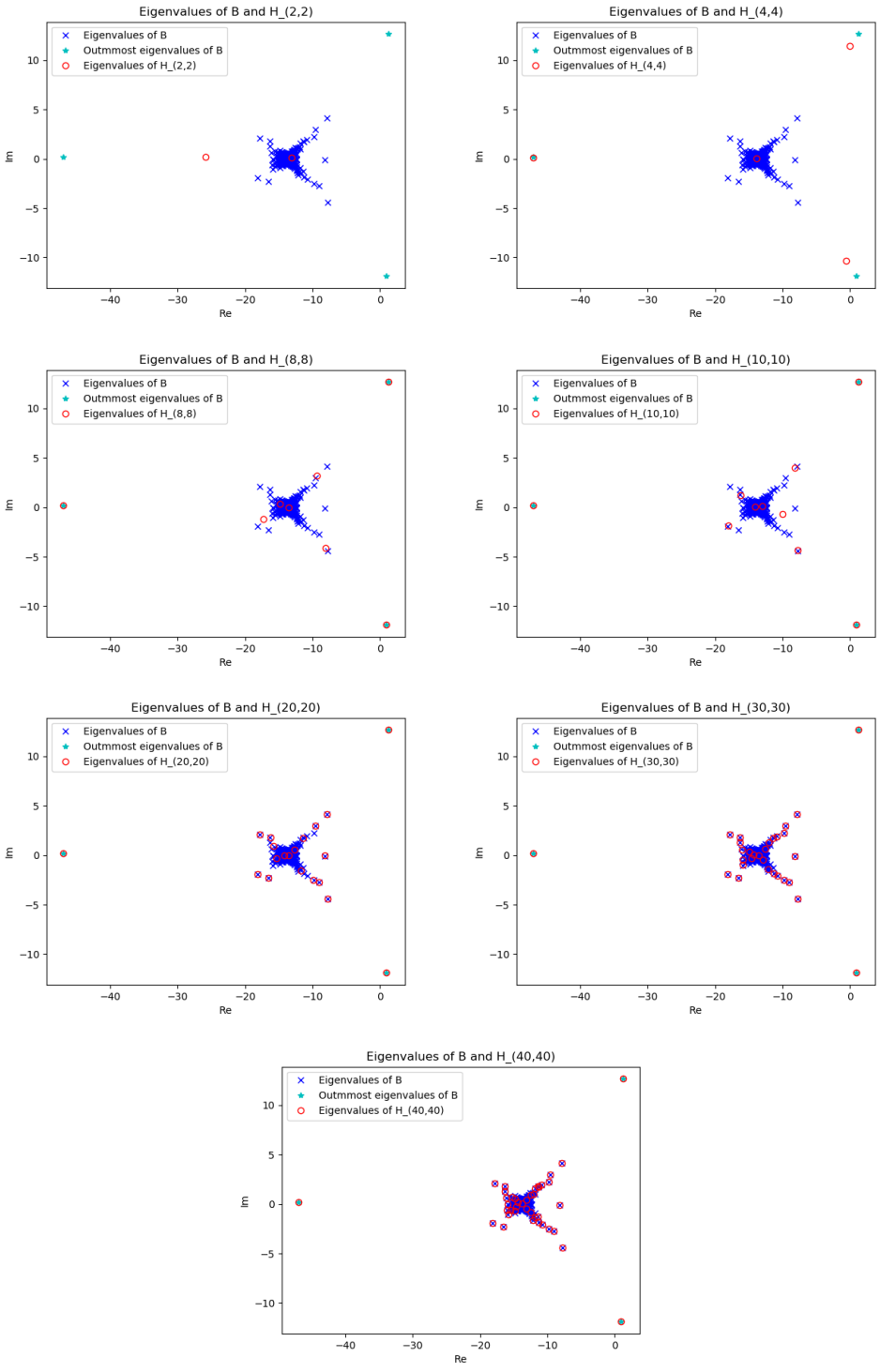

Figure 5: The eigenvalues of B, with the outliers marked, and the Ritz values for the Arnoldi method, with double GS for $m = 2, 4, 8, 10, 20, 30, 40$.

$m = 16$, instead of 20, as the method breaks down for 20. For $p = 18$ one of the approximate eigenvalues started to deviate at the 18:th iteration, which could be due to immediate break down.

We see that implicit restart indeed circumvents the instabilities that polute the explicit restart method, and should thus be used instead.

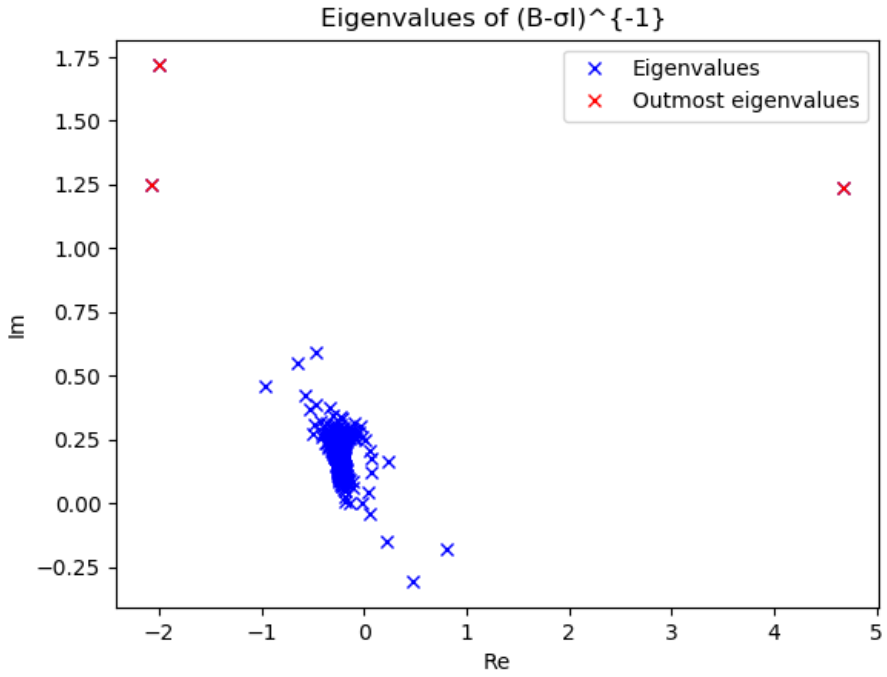

Figure 6: The eigenvalues of $(B - \sigma I)^{-1}$, with the outliers marked.

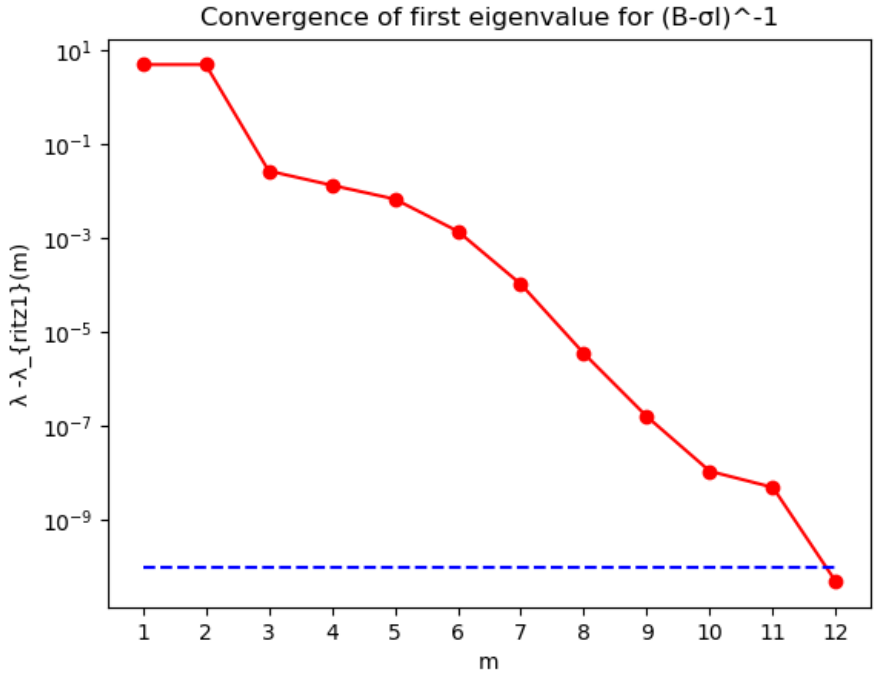

Figure 7: Comparison of eigenvalues after m iterations with shift and invert.

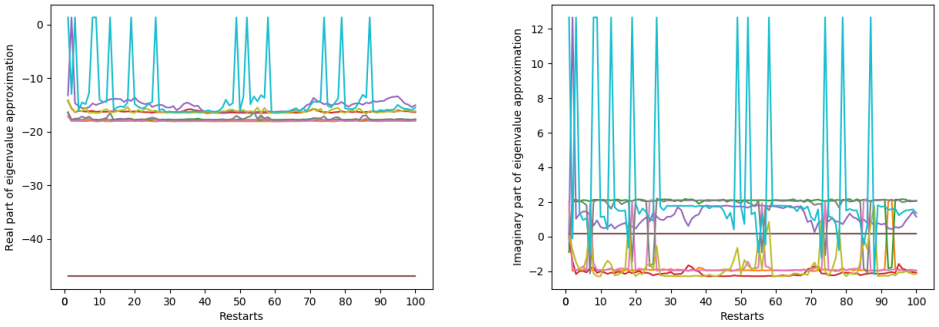

Figure 8: The evolution of eigenvalue approximation for λ_i, $i = 1, \ldots, m$ with explicit restart for $k = 5$ and $m = 10$.

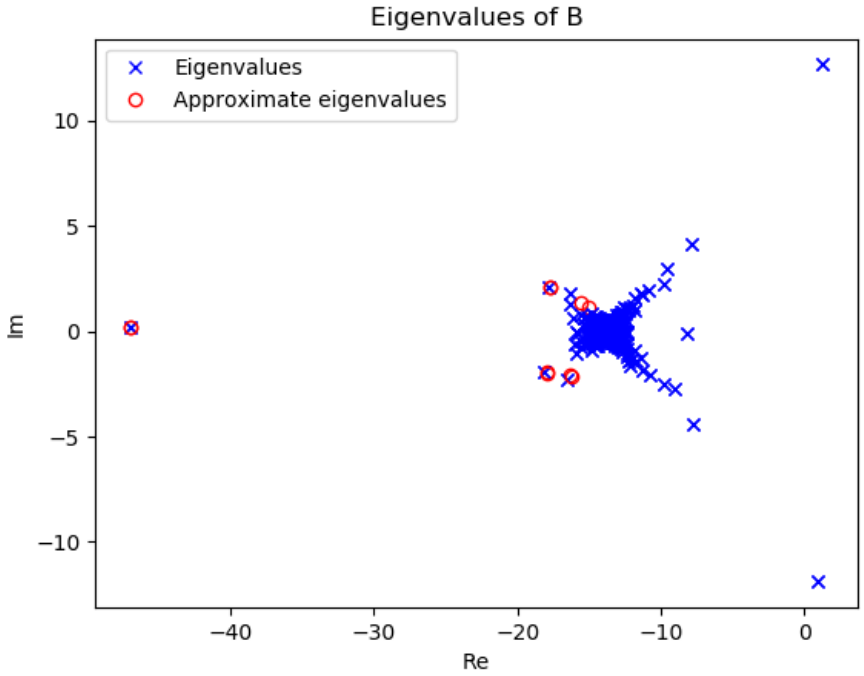

Figure 9: Explicit restart: Plot of eigenvalue approximation for λ_i, $i = 1, \ldots, m$ for $k = 5$ and $m = 10$.

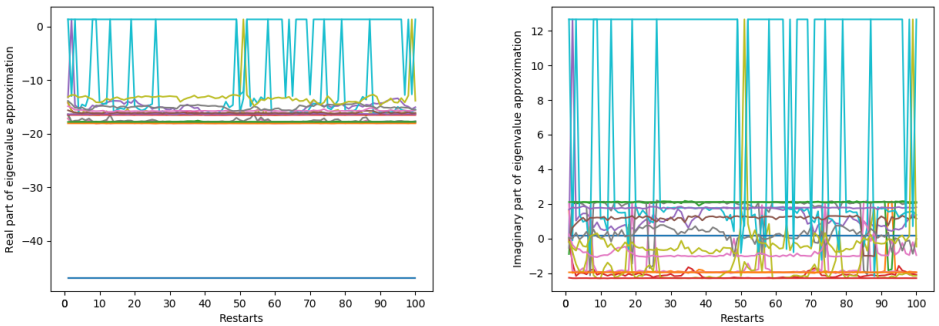

Figure 10: The evolution of eigenvalue approximation for λ_i, $i = 1, \ldots, m$ with explicit restart for $k = 10$ and $m = 20$.

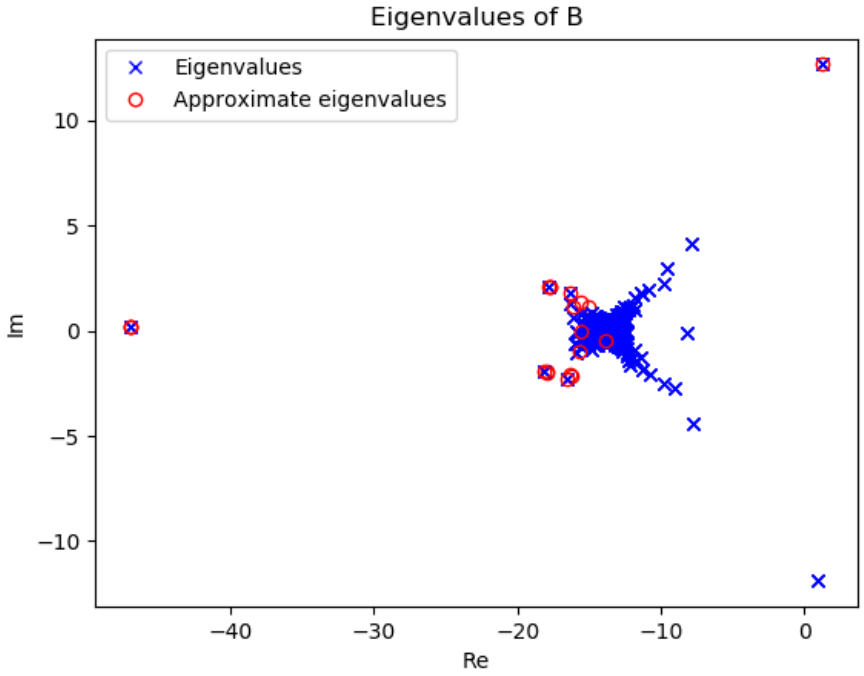

Figure 11: Explicit restart: Plot of eigenvalue approximation for λ_i, $i = 1, \ldots, m$ for $k = 10$ and $m = 20$.

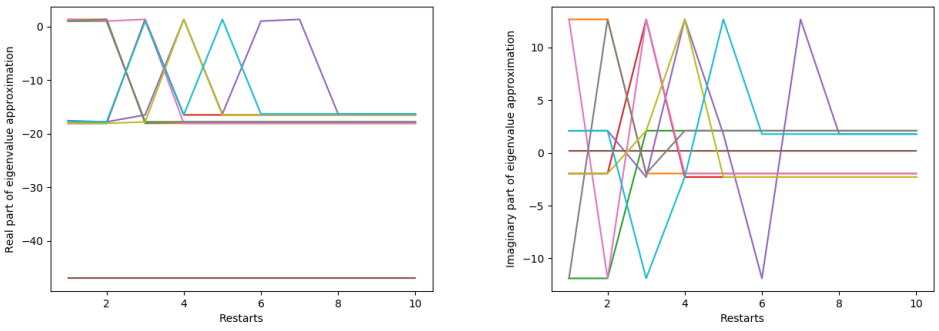

Figure 12: The evolution of eigenvalue approximation for λ_i, $i = 1, \ldots, m$ with implicit restart for $k = 5$, $m = 10$ and $p = 10$.

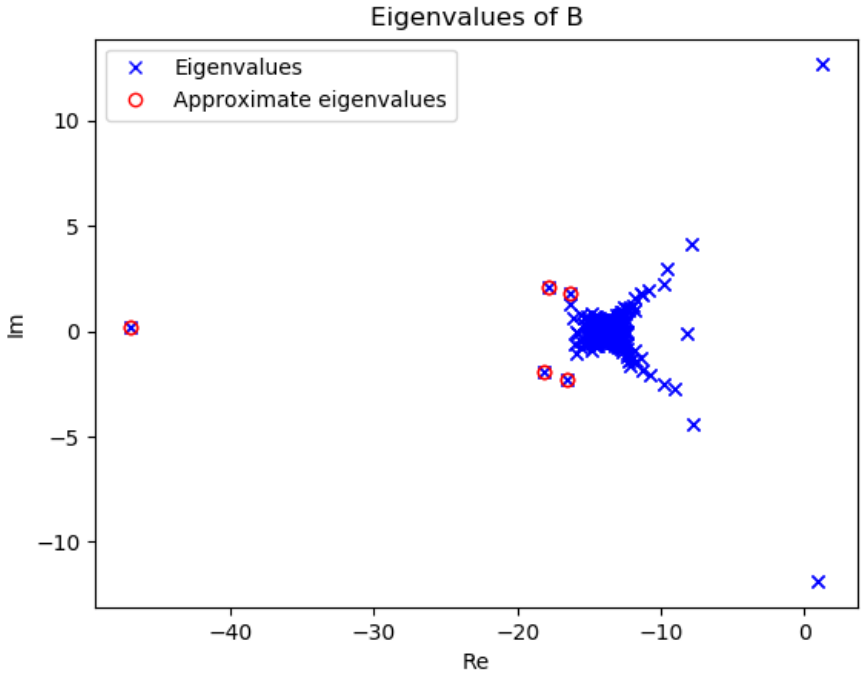

Figure 13: Implicit restart: Plot of eigenvalue approximation for λ_i, $i = 1, \ldots, m$ for $k = 5$, $m = 10$ and $p = 10$.

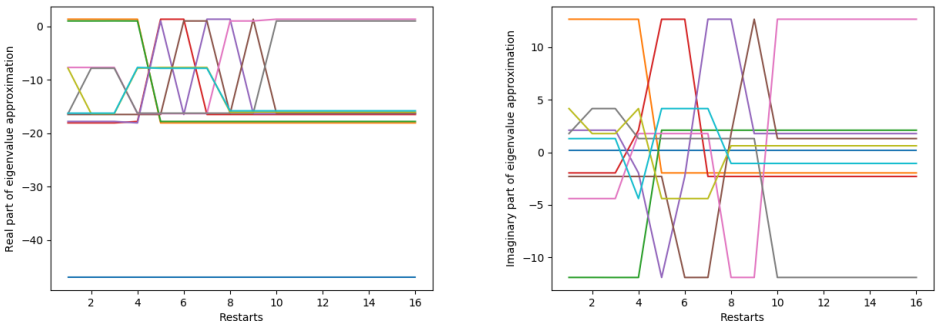

Figure 14: The evolution of eigenvalue approximation for λ_i, $i = 1, \ldots, m$ with implicit restart for $k = 10$, $m = 16$ and $p = 20$.

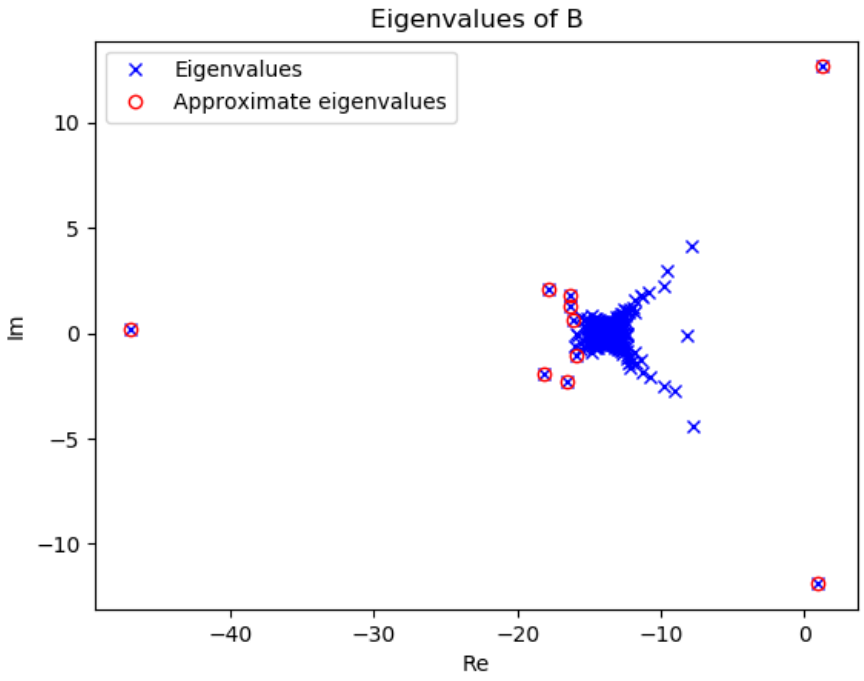

Figure 15: Implicit restart: Plot of eigenvalue approximation for λ_i, $i = 1, \ldots, m$ for $k = 10$, $m = 16$ and $p = 20$.