FSF3580, HW1

Ivar Bengtsson, Tianfang Zhang

November 9, 2020

Problem 2

a)

For this problem, we expect the convergence to be $\mathcal{O}(|\lambda_2/\lambda_1|^2)$, where the square is due to A being symmetric. This can be seen using the fact that the vector v in Algorithm 1 in the notes (our eigenvector approximation) converges with rate $|\lambda_2/\lambda_1|$ in conjunction with Theorem 1.1.2, which states that the corresponding Rayleigh quotient is quadratic in accuracy when A is symmetric. Figure 1 shows the convergence of the power method in a semilog plot, in comparison to the expected convergence rate.

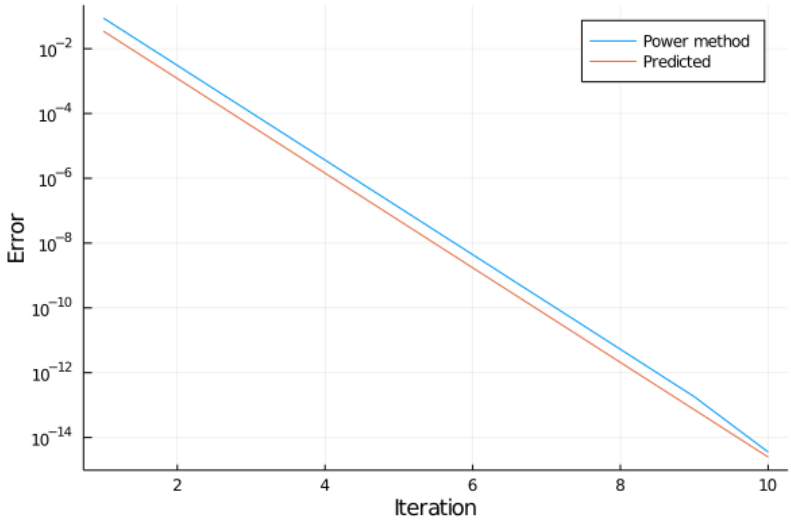

Figure 1: Plot of the error from the true eigenvalue λ_1 using the power method (blue) compared to the expected convergence rate $|\lambda_2/\lambda_1|^2$.

b)

The Rayleigh quotient method is a combination of the inverse iteration method and the power method. The inverse iteration method has convergence $\mathcal{O}(|\lambda_J - \mu|^k/|\lambda_K - \mu|^k)$,

where λ_J and λ_K are the eigenvalues closest and second closest, respectively, to μ. Since the Rayleigh quotient method can be seen as the power method but with μ updated at each iteration, the convergence will be superlinear. This behaviour is confirmed in Figure 2, where we plot the error of the Rayleigh quotient method over the iterations.

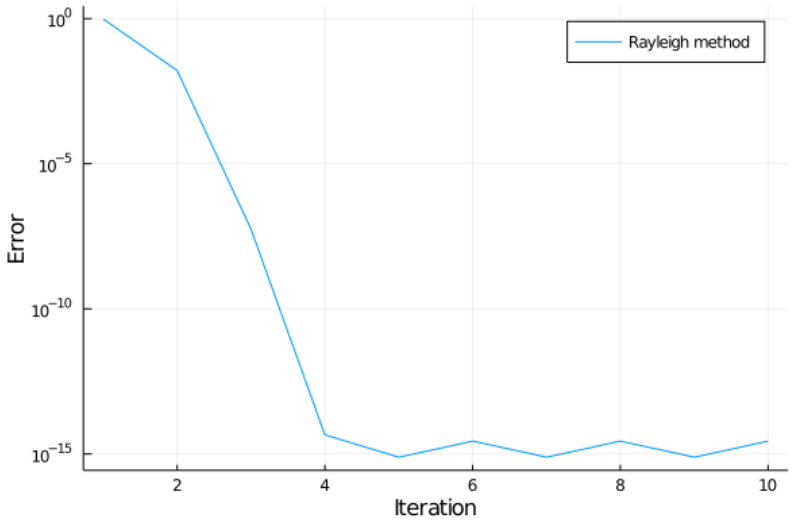

Figure 2: Plot of the error from the true eigenvalue λ_1 using the Rayleigh quotient method.

c)

The change makes A asymmetric. According to the reasoning in a), we would not get a quadratic but a linear accuracy of the Rayleigh quotient, which would make the convergence slower. This is shown in Figure 3, where we note that 5 iterations are required to approximately reach machine precision, compared to 4 iterations when A was symmetric.

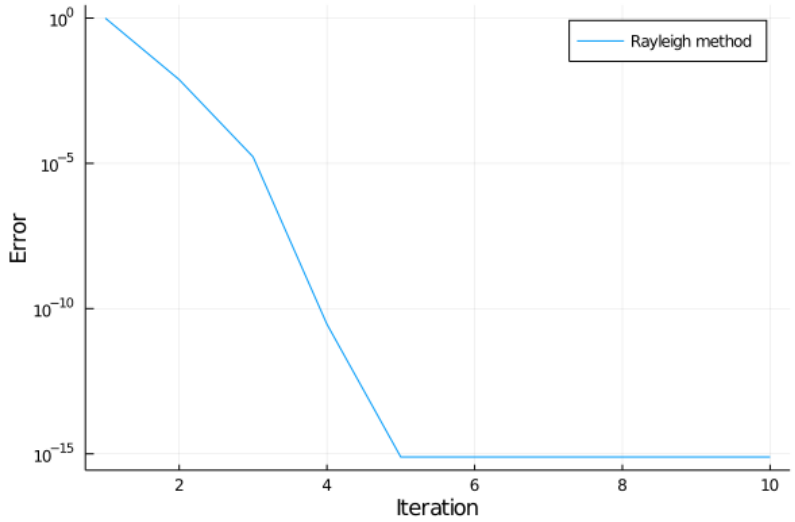

Figure 3: Plot of the error from the true eigenvalue λ_1 using the Rayleigh quotient method, but with A asymmetric.

Problem 4

a)

We know that the approximation stemming from Galerkin method satisfies

$$\mu K_m^T K_m w = K_m^T A K_m w. \tag{2}$$

Since it can be assumed that $\mathcal{K}_m(A, b)$ has dimension m, the columns of K_m are linearly independent and there exists a non-singular R such that $K_m = V_m R$, as suggested in the hint.

Then clearly $K_m^T K_m = (V_m R)^T (V_m R) = R^T V_m^T V_m R = R^T R$ and (2) becomes

$$\mu R^T R w = R^T V_m^T A V_m R w.$$

If we multiply from the left by R^{-T} let $v = Rw$, we see that

$$\mu v = V_m^T A V_m v$$

which is the eigenvalue problem that arises from the Arnoldi methods, showing the equivalence of the primitive approach.

b)

We apply the Arnoldi method and the method of solving (2) directly to the matrix A as generated by

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

with `nn = 500`. The eigenvalue approximations were computed using `eigen` from the package `LinearAlgebra`. The Julia code can be found in `ex4.jl`.

The approximated eigenvalues for a large range of values for m are displayed in figure 4. Figure 5 displays a subset of the data from figure 4 for $m \leq 20$.

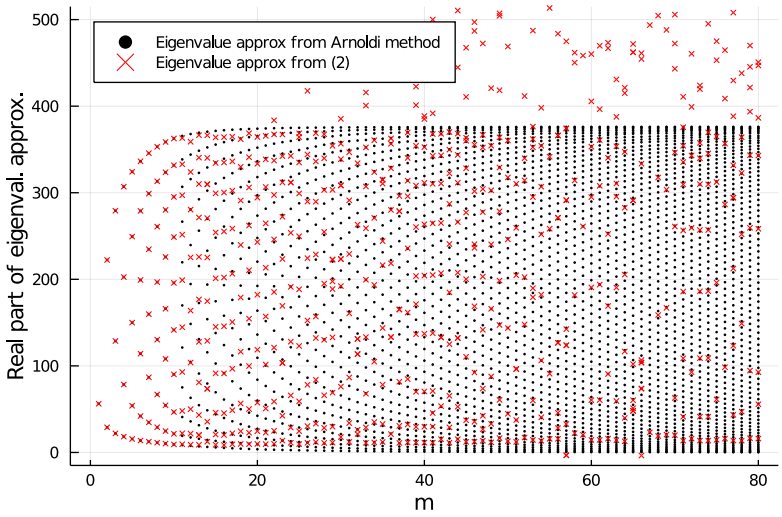

Figure 4: Eigenvalue approximations from the respective methods for m up to $m = 80$.

c)

From a), we know that the approximation of eigenvalues from (2) is identical to the approximation from the Arnoldi method. Therefore, we expect the two methods to produce the same results in exact arithmetic. As seen in 5 this seems largely true for low values of m. However, as m increases, rounding errors start to affect the result and the performance of solving (2) directly quickly deteriorates. Because of this, and because the approximation converges with m, the Arnoldi method is better.

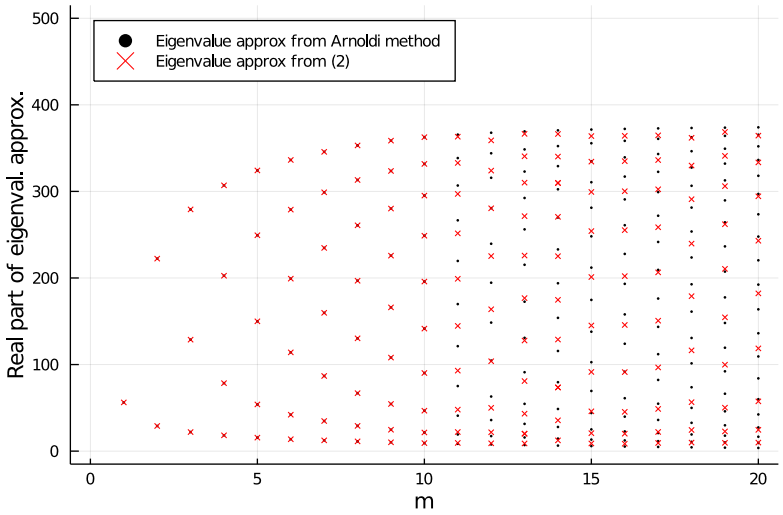

Figure 5: Eigenvalue approximations from the respective methods for m up to $m = 20$.

Problem 6

a)

The convergence theory for the Arnoldi method states that the method will converge to outer eigenvalues faster. Figure 6 shows the eigenvalues of B and highlights the outer eigenvalues which will have the fastest convergence.

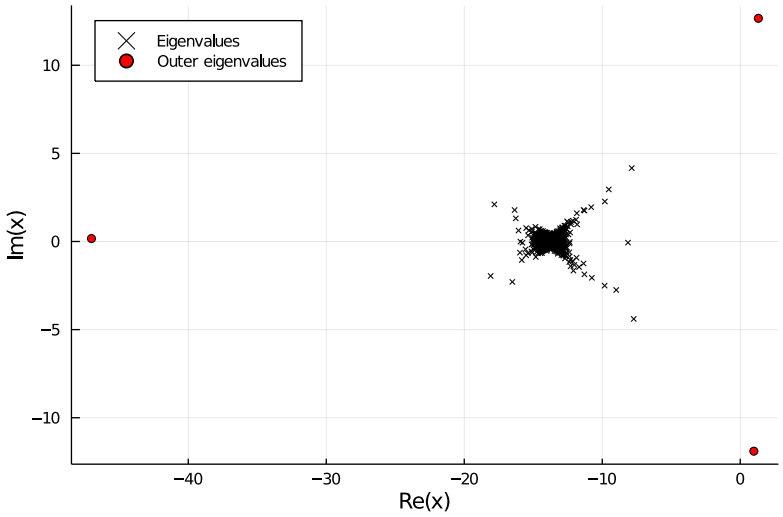

Figure 6: The eigenvalues of B.

b)

The error in the eigenvector x_i associated with eigenvalue λ_i can be characterized by $||(I - QQ^*)x_i||$, which in turn adheres to the bound given by

$$||(I - QQ^*)x_i|| \leq \xi_i \varepsilon_i^{(m)}$$

where ξ_i and $\varepsilon_i^{(m)}$ are scalars where $\varepsilon_i^{(m)}$ relates to the positioning of the eigenvalue λ_i in relation to the other eigenvalues. In particular, if we construct a disc with center c_i and radius ρ_i that contains all eigenvalues except λ_i, then

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

Applying the theory to the outer eigenvalues in a), denoted here as λ_1, λ_2 and λ_3, allows for the estimation of convergence factors. The discs constructed are displayed in figure 7.

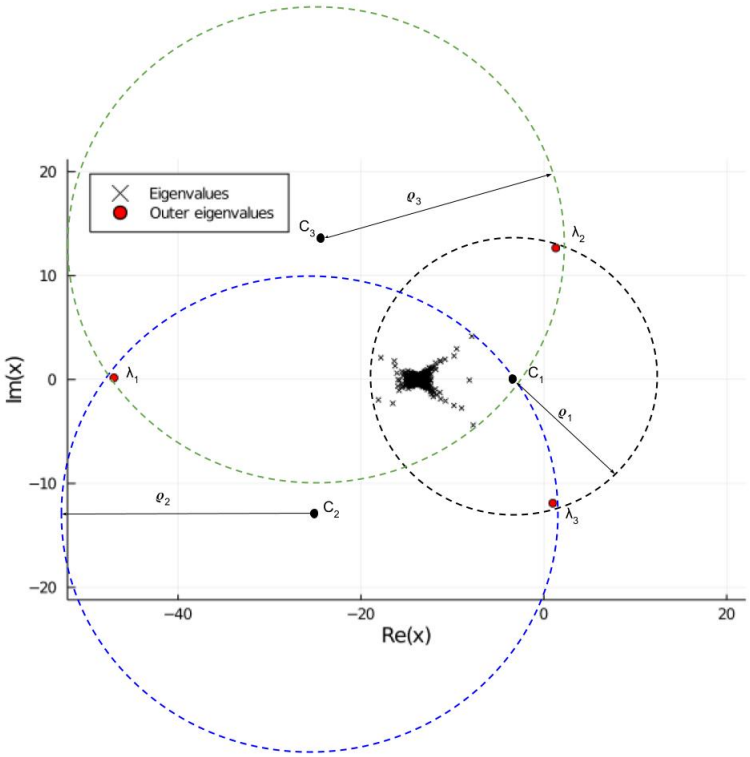

Figure 7: Illustration of discs to estimate convergence bounds for the Arnoldi method.

Clearly, the fastest convergence can be expected for λ_1, as all other eigenvalues may be enclosed in a relatively small disc. Between λ_2 and λ_3 it is tougher to determine which will converge faster.

We estimate the respective convergence factors to be:

$$\alpha_1 = \frac{\rho_1}{|\lambda_1 - c_1|} \approx \frac{15}{42} \approx 0.36,$$
$$\alpha_2 = \frac{\rho_2}{|\lambda_2 - c_2|} \approx \frac{23}{36} \approx 0.64,$$
$$\alpha_3 = \frac{\rho_3}{|\lambda_3 - c_3|} \approx \frac{23}{35} \approx 0.66.$$

c)

Applying the Arnoldi method to B with $m = 2, 4, 8, 10, 20, 30, 40$ yields the results displayed in figures 8, 9 and 10. The red circles indicate the locations of the ritz values that we expect to converge the fastest to the corresponding eigenvalues of B.

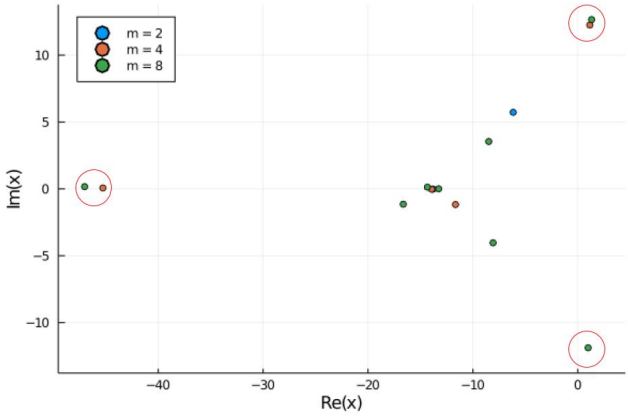

Figure 8: Ritz values from the Arnoldi method for $m = 2, 4, 8$.

From table 1 we see that the eigenvalue reaches 10^{-10} somewhere between $m = 10$ and $m = 20$.

d)

We denote the eigenvalue closest to $-9.8 + 2i$ as $\hat{\lambda} \approx -9.812 + 2.270i$. We will refer to this eigenvalue as the target eigenvalue.

To implement the Arnoldi method with shift and invert, we construct $A = (B - \sigma I)^{-1}$ where σ is the shift. The relationship between a Ritz value of B, μ_i, and the corresponding eigenvalue approximation of B, $\tilde{\lambda}_i$, is then $\tilde{\lambda}_i = \sigma + \frac{1}{\mu_i} \iff \frac{1}{\tilde{\lambda}_i - \sigma} = \mu_i$.

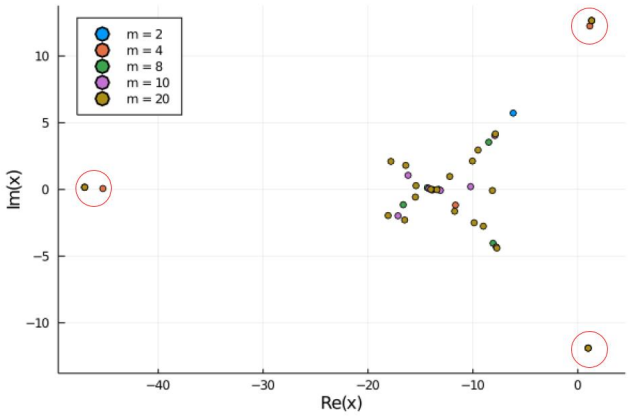

Figure 9: Ritz values from the Arnoldi method for $m = 2, 4, 8, 10, 20$.

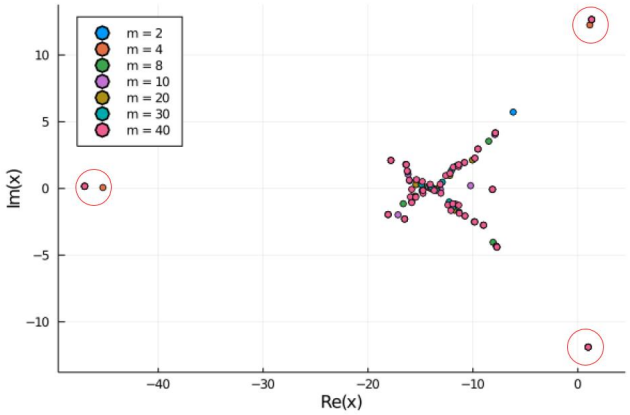

Figure 10: Ritz values from the Arnoldi method for $m = 2, 4, 8, 10, 20, 30, 40$.

m	error λ_1	error λ_2	error λ_3
$m = 2$	33.08	21.1	19.08
$m = 4$	0.2244	24.48	18.54
$m = 8$	0.000185	0.02124	0.001363
$m = 10$	$7.61 \cdot 10^{-7}$	0.001488	0.0001177
$m = 20$	$4.994 \cdot 10^{-14}$	$9.932 \cdot 10^{-11}$	$1.026 \cdot 10^{-12}$
$m = 30$	$2.133 \cdot 10^{-13}$	$2.386 \cdot 10^{-14}$	$1.812 \cdot 10^{-14}$
$m = 40$	$1.208 \cdot 10^{-13}$	$5.107 \cdot 10^{-15}$	$5.515 \cdot 10^{-15}$

Table 1: Differences between Ritz values and corresponding eigenvalues for different values of m.

For each value of m, we let the approximation of $\hat{\lambda}$ be the $\tilde{\lambda}_i$ closest to $\hat{\lambda}$. Table 2 displays the approximation error for $m = 10, 20, 30$.

	Standard AM	$\sigma = -10$	$\sigma = -7 + 2i$	$\sigma = -9.8 + 1.5i$
$m = 10$	2.5	0.0377	0.01044	$1.087 \cdot 10^{-5}$
$m = 20$	0.6	0.001174	$2.571 \cdot 10^{-7}$	$1.006 \cdot 10^{-13}$
$m = 30$	$2 \cdot 10^{-4}$	$6.871 \cdot 10^{-8}$	$9.541 \cdot 10^{-14}$	$5.429 \cdot 10^{-14}$

Table 2: Eigenvalue approximation error for the standard Arnoldi method as well as the shift and invert Arnoldi method for some values of σ and m.

Clearly, the closer the shift σ is to the target eigenvalue, the faster is the convergence. The reason for this is that the smallest $|\hat{\lambda} - \sigma|$ will mean that the corresponding eigenvalue of A will be the largest (and in this case also most extreme) and that the approximation of this eigenvalue will have the fastest convergence using the Arnoldi method.

Problem 8

a)

We claim that (λ, u, v) is a triple of singular value, left-singular vector and right-singular vector (which we call *singular triple*) to A if and only if $(\lambda, (u, v))$ is an eigenpair to C (or, equivalently, $(\lambda, (v, u))$ is an eigenpair to B). Note that if there is a λ and $u \in \mathbb{R}^m$, $v \in \mathbb{R}^n$ satisfying

$$C \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 & A \\ A^{\mathrm{T}} & 0 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \lambda \begin{bmatrix} u \\ v \end{bmatrix},$$

we get from $C^2(u, v) = \lambda^2(u, v)$ that

$$AA^{\mathrm{T}}u = \lambda^2 u, \quad A^{\mathrm{T}}Av = \lambda^2 v,$$

—that is, that (λ, u) and (λ, η) are eigenpairs to AA^{T} and $A^{\mathrm{T}}A$, respectively. Thus, if $(\lambda, (u, v))$ is an eigenpair to C, then (λ, u, v) is a singular triple to A. Conversely, if (λ, u, v) is a singular triple to A, then $(\lambda^2, (u, v))$ is an eigenpair to C^2. Since C is symmetric, it must be diagonalizable, and (μ, ξ) is an eigenpair to C if and only if (μ^2, ξ) is an eigenpair to C^2. Hence, (u, v) is an eigenvector to C with eigenvalue λ.

b)

This is visualized in Figure 11. We note that the picture is very recognizable despite being a rank-1 approximation. In particular, the basis vectors are common for the four frames loaded, which explains the "double vision" effect in the picture.

Figure 11: The rank-1 approximation to the first frame in the India movie.

c)

We implement the Lanczos method to find the Hessenberg matrix H and the orthogonal matrix Q (see 1.3.2 in notes) corresponding to $A^{\mathrm{T}}A$ (which is much smaller than AA^{T} since A is "tall" in our case). Calling `eigs` from Arpack on H, we can obtain the largest eigenvalue λ_1 of H and the corresponding eigenvector ξ_1 and thus $(\lambda_1, Q\xi_1)$ as an eigenpair to $A^{\mathrm{T}}A$. Letting $v_1 = Q\xi_1/\|Q\xi_1\|$, this is the sought right-singular vector, and $\sigma_1 = \sqrt{\lambda_1}$ is the corresponding singular value. We also obtain the left-singular vector u_1 as $u_1 = Av_1/\sigma_1$, since

$$Av_1 = \sum_{i=1}^{\operatorname{rank} A} \sigma_i u_i v_i^{\mathrm{T}} v_1 = \sum_{i=1}^{\operatorname{rank} A} \sigma_i u_i \delta_{i1} = \sigma_1 u_1.$$

Figure 12 shows the rank-1 approximation of the first frame in the Market movie.

Figure 12: The rank-1 approximation to the first frame in the Market movie.