SF3580
HW 2

Anna Broms & Fredrik Fryklund

2018/11/29

1 Task 2

We observe faster convergence and faster predicted convergence for larger α, see (1) and (2). This is expected, as most egeinvalues are clustered around the origin for $\alpha = 1$. Setting $\alpha = 5$ translates the centre of the cluster away from zero, although some eigenvalues are still close to the origin. Note that we may still have fast convergence, which is the case for $\alpha = 5$, but not for $\alpha = 1$. For the latter 100 iterations is required, which is the maximum rank for a 100×100 matrix. As expected the estimated rates of convergence are not of any use and diverges. For $\alpha = 5$ we have convergence and the predicted rates of convergence are indeed bounds, but not very useful due to the underestimation. Furthermore we see that the predicted rate of convergence for one disc is better than the corresponding for two discs. This is not the case for $\alpha = 10$, 100.

For all α there is an isolated eigenvalue. All predicted rates of convergence are based on either one circle containing all the eigenvalues or two circles, where one is a point consisiting of the outlying eigenvalue. The circles are found in Figure 3–6.

(c)

bs: Samples 10000, evals per sample 1. mean time in table. bs: Samples 3673, evals per sample 1. mean time in table.

2 Task 5

Given a real symmetric matrix A with eigenvalues 10, 10.5 and 100 eigenvalues in the interval $[2, 3]$, we prove a bound for then number of steps needed for CG to reduce the error measured in $\|Ax_n - b\|_{A^{-1}} = \|x_n - x_*\|$ by a factor 10^7. We assume exact arithmetic and no breakdown.

Proof:
Let $\kappa = \lambda_{\max}/\lambda_{\min}$. By Theorem 38.5. in T.B. we have

$$\frac{\|e_n\|}{\|e_0\|} = 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^n. \tag{1}$$

We seek n such that the error has been reduced by a factor 10^7, i.e. n such that (1) is about $1e-7$. For the matrix A the largest lower bound for $\lambda_{\min}$ is 2. The largest eigenvalue $\lambda_{\max} = 10.5$. We now solve the following problem for n,

Table 1: $\alpha = 1$

GMRES								
	$m = 100$		$m = 200$		$m = 500$		$m = 1000$	
	resnorm	time	resnorm	time	resnorm	time	resnorm	time
n = 5	2.583	162.117 μs	3.848	304.663 μs	6.378	1.511 ms	9.0192	3.915 ms
n = 10	2.548	363.392 μs	3.826	591.597 μs	6.326	2.231 ms	9.007	8.300 ms
n = 20	2.428	1.055 ms	3.744	1.439 ms	6.307	5.081 ms	8.967	16.586 ms
n = 50	1.874	5.246 ms	3.415	8.117 ms	6.194	17.627 ms	8.796	48.643 ms
n = 100	$2.212e-12$	28.858 ms	2.794	32.087 ms	5.718	59.167 ms	8.555	133.731 ms
Backslash								
	$m = 100$		$m = 200$		$m = 500$		$m = 1000$	
	resnorm	time	resnorm	time	resnorm	time	resnorm	time
	$1.056e-12$	1.350 ms	$4.632e-14$	6.207 ms	$1.868e-13$	45.407 ms	$7.074e-13$	170.036 ms

Table 2: $\alpha = 100$

GMRES								
	$m = 100$		$m = 200$		$m = 500$		$m = 1000$	
	resnorm	time	resnorm	time	resnorm	time	resnorm	time
n = 5	$6.545e-7$	159.159 μs	$5.636e-6$	291.925 μs	$9.712e-5$	962.453 μs	$6.893e-4$	3.790 ms
n = 10	$1.773e-14$	362.841 μs	$7.944e-13$	564.961 μs	$1.667e-10$	1.905 ms	$7.337e-9$	7.722 ms
n = 20	$2.262e-15$	1.014 ms	$4.983e-15$	1.712 ms	$9.328e-15$	6.269 ms	$2.400e-14$	16.591 ms
n = 50	$2.269e-15$	5.625 ms	$5.044e-15$	8.152 ms	$9.291e-15$	18.087 ms	$2.402e-14$	58.155 ms
n = 100	$2.224e-15$	30.860 ms	$5.050e-15$	32.769 ms	$1.030e-14$	58.455 ms	$1.030e-14$	61.989 ms
Backslash								
	$m = 100$		$m = 200$		$m = 500$		$m = 1000$	
	resnorm	time	resnorm	time	resnorm	time	resnorm	time
	$1.662e-15$	1.313 ms	$3.893e-15$	5.170 ms	$7.977e-15$	35.644 ms	$1.565e-14$	159.975 ms

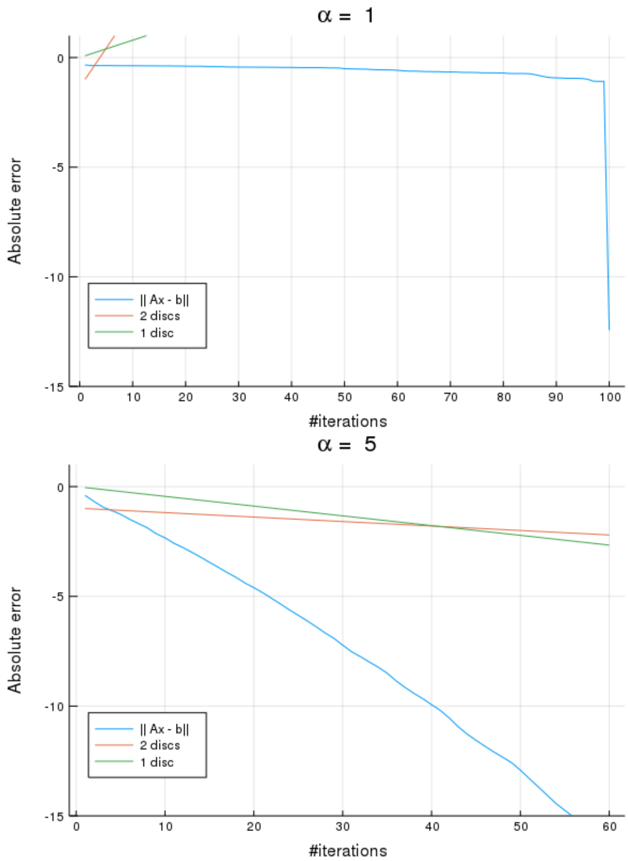

Figure 1: The convergence of the residual for GMRES, plotted against the number of iterations. Here for $\alpha = 1$ and $\alpha = 5$

$$2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^n \leq 10^{-7}$$

$$\begin{aligned}
&\Leftrightarrow n \leq \frac{\log\left(0.5\,10^{-7}\right)}{\log\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)} \\
&\Leftrightarrow n \geq 17.967661670561174 \\
&\Leftrightarrow n \geq 18.
\end{aligned}$$

We confirm this numerically, $n = 17$ gives

$$2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^{18} \approx 2.473e-7$$

and $n = 18$ gives

$$2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^{18} \approx 9.702e-8.$$

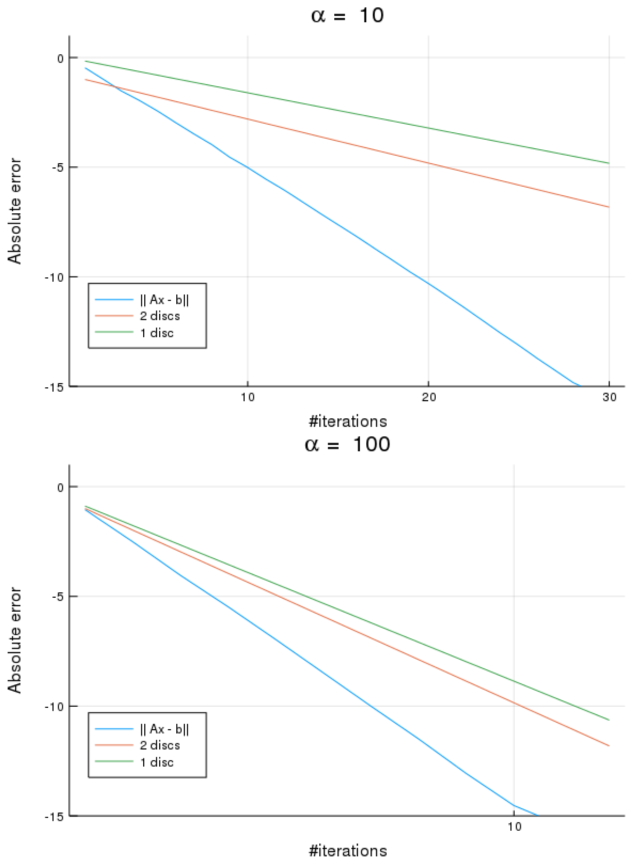

Figure 2: The convergence of the residual for GMRES, plotted against the number of iterations. Here for $\alpha = 10$ and $\alpha = 100$

3 Task 7

(a)

Given that $A = V^{-1}\Lambda V$ we want to show $A^k = A = V^{-1}\Lambda^k V$, which is done by induction. The initial stage is

$$A^2 = V^{-1}\Lambda V V^{-1}\Lambda V = V^{-1}\Lambda^2 V. \tag{2}$$

Assume $A^k = V^{-1}\Lambda^k V$ for some nonzero k, then

$$A^{k+1} = (V^{-1}\Lambda V)^k (V^{-1}\Lambda V) = V^{-1}\Lambda^k V\, V^{-1}\Lambda V = V^{-1}\Lambda^{k+1} V. \tag{3}$$

Thus $A^k = V^{-1}\Lambda^k V$. A simple consequence is that for $p \in P_n^0$ one has

$$p(A) = \sum_{k=1}^{n} a_k A^k = \sum_{k=1}^{n} a_k V^{-1}\Lambda^{k+1} V = V^{-1}\left(\sum_{k=1}^{n} a_k \Lambda^{k+1}\right) V = V^{-1} p\left(\Lambda^{k+1}\right) V \tag{4}$$

with $a_0 = 1$ for $p \in P_n^0$ and knowing that $A^0 = I$. We have

$$\min_{p \in P_n^0} \|p(A)\| \leq \|V\| \|V^{-1}\| \min_{p \in P_n^0} \|p(\Lambda)\| \tag{5}$$

as a consequence of norms being submultiplicative.

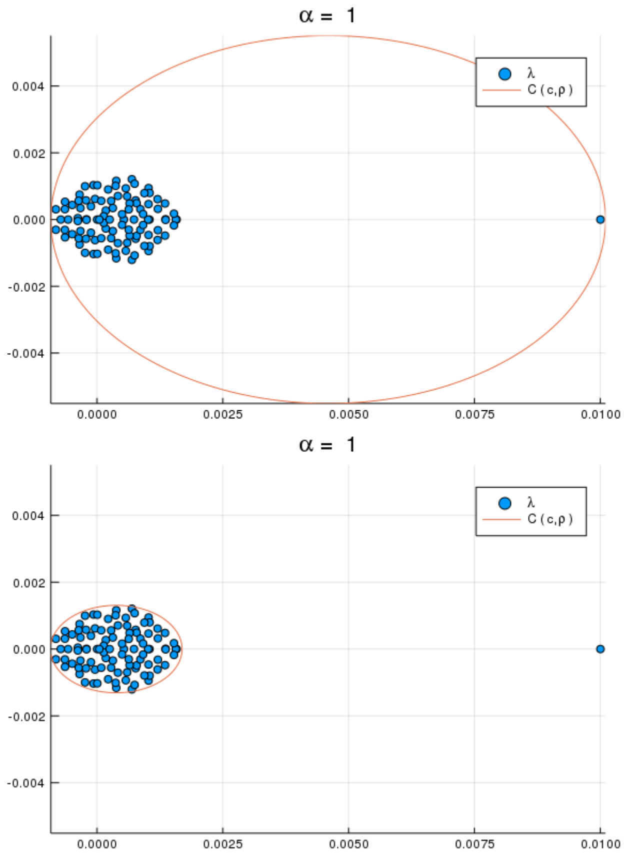

Figure 3: For $\alpha = 1$: The two configurations of discs containing all eigenvalues.

(b)

First we show by induction that

$$\begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix}^k = \begin{pmatrix} \lambda_1^k & k\lambda_1^{k-1} \\ 0 & \lambda_1^k \end{pmatrix}. \tag{6}$$

The initial stage is for $k = 2$:

$$\begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix}^2 = \begin{pmatrix} \lambda_1^2 & 2\lambda_1 \\ 0 & \lambda_1^2 \end{pmatrix}. \tag{7}$$

Assume (6) holds for k, then

$$\begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix}^{k+1} = \begin{pmatrix} \lambda_1^k & k\lambda_1^{k-1} \\ 0 & \lambda_1^k \end{pmatrix} \begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix} = \begin{pmatrix} \lambda_1^{k+1} & (k+1)\lambda_1^k \\ 0 & \lambda_1^{k+1} \end{pmatrix}, \tag{8}$$

i.e. the proposition (6) holds for all nonzero k.

Introduce the monomial $p_k(z) = z^k$, then

$$p_k\left(\begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix}\right) = \begin{pmatrix} \lambda_1^k & k\lambda_1^{k-1} \\ 0 & \lambda_1^k \end{pmatrix} = \begin{pmatrix} p_k(\lambda_1) & p_k'(\lambda_1) \\ 0 & p_k(\lambda_1) \end{pmatrix} \tag{9}$$

which holds for all nonzero k from the induction proof above. We now have

$$p\left(\begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix}\right) = \sum_{k=1}^{n} a_k p_k \left(\begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix}\right) = \sum_{k=1}^{n} \begin{pmatrix} a_k p_k(\lambda_1) & a_k p_k'(\lambda_1) \\ 0 & a_k p_k(\lambda_1) \end{pmatrix} = \begin{pmatrix} p(\lambda_1) & p'(\lambda_1) \\ 0 & p(\lambda_1) \end{pmatrix}. \tag{10}$$

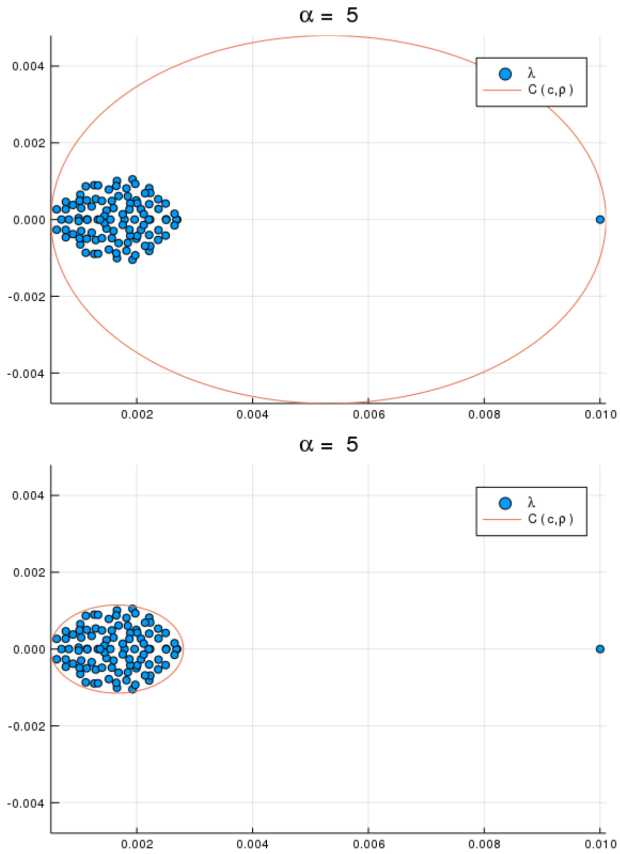

Figure 4: For $\alpha = 5$: The two configurations of discs containing all eigenvalues.

(c)

Let A be a block diagonal matrix, such that

$$A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_m \end{pmatrix} \tag{11}$$

where A_i are Jordan block matrices. Due to the block structure we have

$$p(A) = \begin{pmatrix} p(A_1) & & & \\ & p(A_2) & & \\ & & \ddots & \\ & & & p(A_m) \end{pmatrix}. \tag{12}$$

Each block $p(A_i)$ has a singular value decomposition $p(A_i) = U_i S_i V_H^*$, where U_i and V_i are unitary matrices. S_i is a diagonal matrix with the singular values σ as elements. We can now write $p(A)$ as follows.

$$p(A) = \underbrace{\begin{pmatrix} U_1 & & & \\ & U_2 & & \\ & & \ddots & \\ & & & U_m \end{pmatrix}}_{U:=} \underbrace{\begin{pmatrix} S_1 & & & \\ & S_2 & & \\ & & \ddots & \\ & & & S_m \end{pmatrix}}_{S:=} \underbrace{\begin{pmatrix} V_1^H & & & \\ & V_2^H & & \\ & & \ddots & \\ & & & V_m^H \end{pmatrix}}_{V^H:=} \tag{13}$$

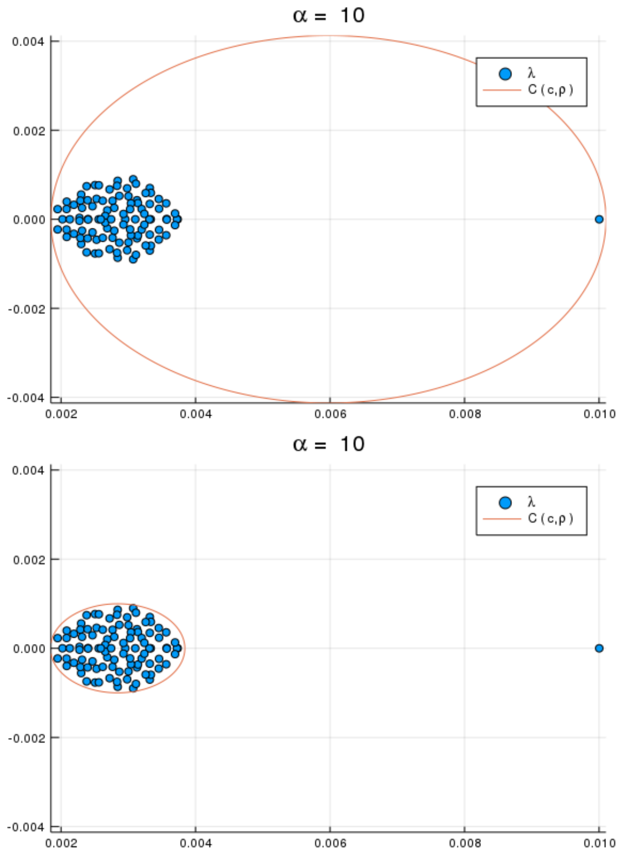

Figure 5: For $\alpha = 10$: The two configurations of discs containing all eigenvalues.

due to the rules of multiplication for block diagonal matrices. The final result follows from the definition of the operator norm $\|\cdot\|_2$:

$$\begin{aligned}
\|p(A)\|_2 &= \sigma_{\max}(p(A)) = \max S = \max_{i=1,\ldots,m} (\max S_i)) \\
&= \max_{i=1,\ldots,m} (\sigma_{\max}(p(A_i))) = \max_{i=1,\ldots,m} (\|p(A_i)\|_2) \\
&= \max \left(\left\| \begin{pmatrix} p(\lambda_1) & p'(\lambda_1) \\ 0 & p(\lambda_1) \end{pmatrix} \right\|_2, |p(\lambda_3)|, \ldots, |p(\lambda_m)| \right)
\end{aligned}$$

(d)

It is clear that

$$p(z) = (\alpha_n + \beta_n z) \frac{(c-z)^{n-1}}{c^{n-1}} \tag{14}$$

satisfies $p \in P_n$. This immediately gives $\alpha_n = 1$. We now study

$$p'(z) = \frac{c\left(1 - \frac{z}{c}\right)^n (\alpha_n - \alpha_n n + \beta_n(c - nz))}{(c-z)^2} = \frac{c\left(1 - \frac{z}{c}\right)^n (1 - n + \beta_n(c - nz))}{(c-z)^2}. \tag{15}$$

Thus

$$p'(\lambda_1) = 0 \Leftrightarrow \frac{c\left(1 - \frac{\lambda_1}{c}\right)^n (1 - n + \beta_n(c - n\lambda_1))}{(c-\lambda_1)^2} = 0 \Leftrightarrow (1 - n + \beta_n(c - n\lambda_1)) = 0, \tag{16}$$

that is

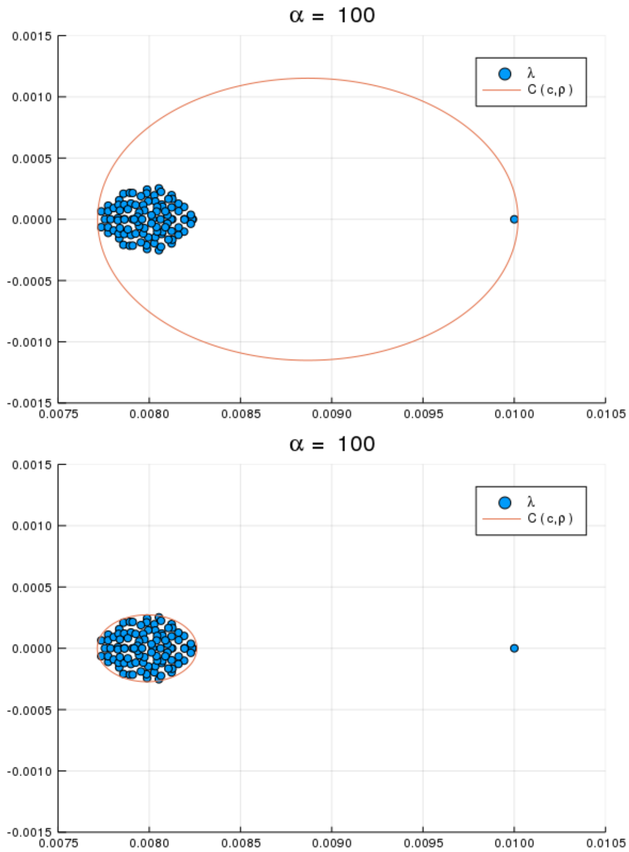

Figure 6: For $\alpha = 100$: The two configurations of discs containing all eigenvalues.

$$\beta_n = \frac{n-1}{c - n\lambda_1}. \tag{17}$$

In turn this assumes that $c \neq n\lambda_1$ for $n > 1$.

(e)

Assuming x_n is the n:th iterate generated by GMRES-iterate, we have by lemma 2.1.3 from that lecture notes that

$$\begin{aligned}
&\|Ax_n - b\|_2 = \min_{x \in \mathcal{K}n(A,b)} \|Ax - b\|_2 = \min_{p \in P_n^0} \|p(A)b\| \leq \|V\|\|V^{-1}\| \min_{p \in P_n^0} \|p(\Lambda)\|\|b\| \\
\Leftrightarrow &\frac{\|Ax_n - b\|_2}{\|b\|} \leq \|V\|\|V^{-1}\| \min_{p \in P_n^0} \|p(\Lambda)\|
\end{aligned}$$

due to the result in 7 (a). Let

$$q(z) = \left(1 + z\frac{n-1}{c - n\lambda_1}\right) \frac{(c-z)^{n-1}}{c^{n-1}}. \tag{18}$$

from the previous task, which by construction is an element of P_n^0. Thus

$$\min_{p \in P_n^0} \|p(\Lambda)\| \leq \|q(\Lambda)\| = \max\left(\left\|\begin{pmatrix} q(\lambda_1) & q'(\lambda_1) \\ 0 & q(\lambda_1) \end{pmatrix}\right\|_2, |q(\lambda_3)|, \ldots, |q(\lambda_m)|\right). \tag{19}$$

Recall that $q'(\lambda_1) = 0$ and that the matrix 2-norm of diagonal matrix is the largest element in modulus. The expression above can be simplified as

$$\min_{p\in P_n^0} \|p(\Lambda)\| \le \max\left(|q(\lambda_1)|, |q(\lambda_3)|, \dots, |q(\lambda_m)|\right) = \max_{\lambda_i}\left(1+\lambda_i\frac{n-1}{c-n\lambda_1}\right)\frac{(c-\lambda_i)^{n-1}}{c^{n-1}}. \tag{20}$$

It is given that all eigenvalues are contained in the disc centered at c with radius ρ. By taking the modulus the inequality (20) and assumin $\lambda_1 \neq 0$ we get

$$\min_{p\in P_n^0} \|p(\Lambda)\| \le \max_{\lambda_i}\left|1+\lambda_i\frac{n-1}{c-n\lambda_1}\right|\frac{\rho^{n-1}}{|c^{n-1}|} \le \max_{\lambda_i}\frac{\overbrace{|c-\lambda_i|}^{\le\rho}+n\overbrace{|\lambda_i-\lambda_1|}^{\le 2\rho}}{|c-n\lambda_1|}\frac{\rho^{n-1}}{|c^{n-1}|} \le \gamma_n\frac{\rho^n}{|c^n|}$$

with

$$\gamma_n = \frac{\frac{1}{n}+2}{\left||\frac{1}{n}|-|\frac{\lambda_1}{c}|\right|}. \tag{21}$$

We already claimed that $c \neq n\lambda_1$, thus the denominator is nonzero for all n. In the limit we have

$$\lim_{n\to\infty}\gamma_n = 2\frac{|c|}{|\lambda_1|} \tag{22}$$

which is bounded. Combining all the results above gives

$$\frac{\|Ax_n-b\|_2}{\|b\|} \le \|V\|\|V^{-1}\|\gamma_n\frac{\rho^n}{|c^n|}.$$

If $\lambda_1 = 0$ then $\beta_n = (n-1)/c$ and the corresponding bound for (20) is

$$\min_{p\in P_n^0} \|p(\Lambda)\| \le \max_{\lambda_i}\left|1+\lambda_i\frac{n-1}{c}\right|\frac{\rho^{n-1}}{|c^{n-1}|} \le \max_{\lambda_i}\left(\overbrace{|c-\lambda_i|}^{\le\rho}+n|\lambda_i|\right)\frac{\rho^{n-1}}{|c^n|} \le \gamma_n\frac{\rho^n}{|c^n|}.$$

However, now

$$\gamma_n = \max_{\lambda_i}\left(1+n\frac{|\lambda_i|}{\rho}\right), \tag{23}$$

which is not a bounded sequence.

(f)

For nonzero λ_1 we have convergence, but the speed is influenced by γ_n. Roughly, the further the centre c is from λ_1 the better. For many iterations we approximately get

$$\gamma_n\frac{\rho^n}{|c^n|} \approx \frac{2\rho}{|\lambda_1|}\frac{\rho^{n-1}}{|c^{n-1}|}. \tag{24}$$

Thus the rate of convergence is the same, but the factor $\frac{2\rho}{|\lambda_1|}$ may be large. So if the double eigenvalues lie close to zero and the other eigenvalues lies far away from the origin then the factor will be large.

For $\lambda_1 = 0$ the sequence γ_n is not bounded. Note that this does not mean that GMRES will diverge, only that the estimate gives no information.

(e)

We discussed with Aku Kammonen and Parikshit Upadhyaya.