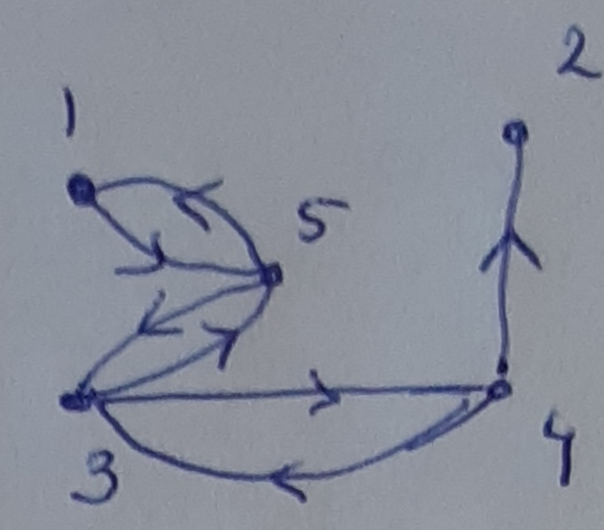
Exc 1 $A = \begin{bmatrix} -2 & 0 & 0 & 0 & 1 \\ 0 & -2 & 0 & 0 & 0 \\ 0 & 0 & -4 & 2 & 2 \\ 0 & 1 & 1 & -1 & 0 \\ 1 & 0 & 3 & 0 & -4 \end{bmatrix}$



a. the graph of A

P_1, P_3, P_4, P_5 are strongly connected

it is not possible to go from P_2 to the other vertices

$\Rightarrow$ A reducible

b. before applying GGT's reorder the vertices and rewrite A

use : $\{P_1, P_3, P_4, P_5, P_2\}$

$$\Rightarrow \tilde{A} = \left[\begin{array}{cccc|c} -2 & 0 & 0 & 1 & 0 \\ 0 & -4 & 2 & 2 & 0 \\ 0 & 1 & -1 & 0 & 1 \\ 1 & 3 & 0 & -4 & 0 \\ \hline 0 & 0 & 0 & 0 & -2 \end{array}\right]$$

$\tilde{A}$ is of form $\begin{bmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{bmatrix}$

$\Rightarrow \lambda = -2$ is eigenvalue of A $\quad \sigma(A) = \sigma(B_{11}) \cup \{-2\}$

- for the remainder look at B_{11}

B_{11} is not symmetric $\Rightarrow$ apply Gerschgorin using both rows and columns

$$B_{11} = \begin{pmatrix} -2 & 0 & 0 & 1 \\ 0 & -4 & 2 & 2 \\ 0 & 1 & -1 & 0 \\ 1 & 3 & 0 & -4 \end{pmatrix} \Rightarrow$$

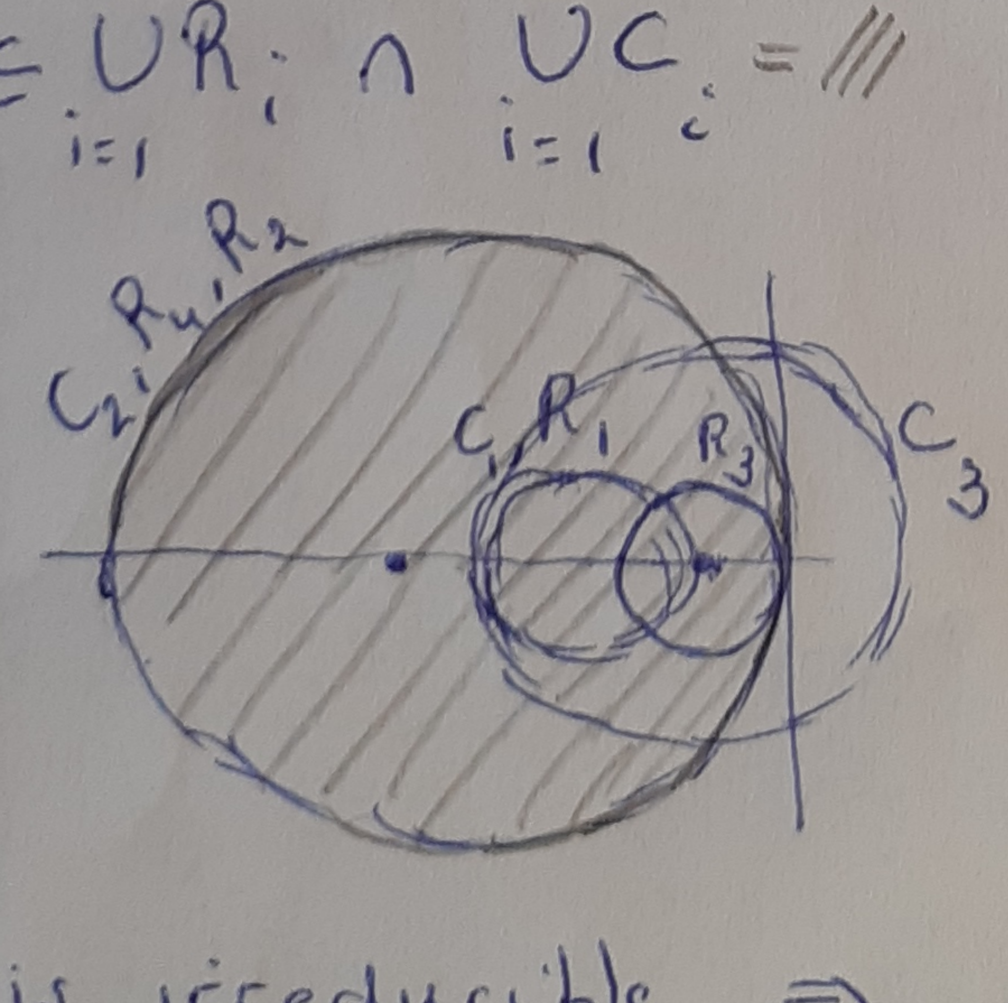
$R_1 = \{ |z+2| \leq 1 \}$ $\quad C_1 = \{ |z+2| \leq 1 \}$

$R_2 = \{ |z+4| \leq 4 \}$ $\quad C_2 = \{ |z+4| \leq 4 \}$

$R_3 = \{ |z+1| \leq 1 \}$ $\quad C_3 = \{ |z+1| \leq 2 \}$

$R_4 = \{ |z+4| \leq 4 \}$ $\quad C_4 = \{ |z+4| \leq 3 \}$

$$\sigma(B_{11}) \subseteq \bigcup_{i=1}^{4} R_i \cap \bigcup_{i=1}^{4} C_i = ///$$



$\Rightarrow \sigma(B_{11}) \subseteq \{ z \in \mathbb{C} \mid |z+4| \leq 4 \}$

B_{11} is irreducible $\Rightarrow$ an eigenvalue cannot lie on the boundary of S_R unless it belongs to the boundary of every circle R_i

for instance: since $R_1 \subset R_4 \Rightarrow \lambda \in B_{11}$ are not on the boundary of $\{ z \in \mathbb{C} \mid |z+4| \leq 4 \}$

$\Rightarrow \lambda = -2$ eigenvalue of A, the other 4 eigenvalues are in $\{ z \in \mathbb{C} \mid |z+4| < 4 \}$ $\Rightarrow$ $Re(\lambda) < 0$ for $\lambda \in A$

Exc 2 A real symmetric

x, $\|x\|_2 = 1$

θ Ritz pair obtained from the Lanczos method

Show $\min_{\lambda \in \sigma(A)} |\lambda - \theta| \leq \|Ax - \theta x\|_2$

A real, symmetric $\Rightarrow \exists Q$ orthogonal s.t $Q^T A Q = D$
$D_{ii} = \lambda_i \quad \lambda_i \in \sigma(A)$

hence $\|Ax - \theta x\|_2 = \|Q^T(Ax - \theta x)\|_2$

↑ length preservation after multiplication with Q^T because Q orthogonal

$= \|Q^T(AQy - \theta Qy)\|_2 = \|Dy - \theta y\|_2$

↑ set $x = Qy$
i.e $\exists y$ s.t $x \equiv Qy$

$= \sqrt{\sum_{i=1}^{n} |\lambda_i - \theta|^2 |y_i|^2}$

$\geq \sqrt{\sum_{i=1}^{n} (\min_{\lambda \in \sigma(A)} |\lambda - \theta|)^2 |y_i|^2}$

$= \min_{\lambda \in \sigma(A)} |\lambda - \theta| \sqrt{\sum_{i=1}^{n} |y_i|^2} = \min_{\lambda \in \sigma(A)} |\lambda - \theta| \|y\|_2$

$= \min_{\lambda \in \sigma(A)} |\lambda - \theta| \|x\|_2 = \min_{\lambda \in \sigma(A)} |\lambda - \theta|$

↑ $\|y\|_2 = \|Q^T x\|_2 = \|x\|_2$ ↑ $\|x\|_2 = 1$

Exc3

a) • $\lambda^{(i)}$ converges to 1

• convergence rate, two possible approaches

approach 1: rate = $\dfrac{\text{exact error iterate}^{(i+1)}}{\text{exact error iterate}^{(i)}}$: $\dfrac{\lambda^{(4)}-1}{\lambda^{(3)}-1} = \dfrac{0{,}00696}{0{,}03569} \approx 0{,}195$

$\dfrac{\lambda^{(5)}-1}{\lambda^{(4)}-1} = \dfrac{0.00140}{0{,}00696} \approx 0{,}201$ $\quad \dfrac{\lambda^{(6)}-1}{\lambda^{(5)}-1} = 0{,}2$ $\quad \dfrac{\lambda^{(7)}-1}{\lambda^{(6)}-1} \approx 0{,}214$

$\left(\dfrac{\lambda_8^{(8)}-1}{\lambda^{(9)}-1} = 0{,}17 \right)$ $\Rightarrow$ conv. rate $\approx 0{,}2$

note: (1) looking at only one value of $^{(i)}$ is not sufficient, to make sure the convergence rate is stable over some iterates, you need to compute some successive rates

(2) do no take i too large, conv. rate will be inaccurate since results are not printed in sufficient digits

approach 2: rate $\approx \dfrac{\lambda^{(i+1)} - \lambda^{(i)}}{\lambda^{(i)} - \lambda^{(i-1)}}$ $\quad \dfrac{\lambda^{(5)}-\lambda^{(4)}}{\lambda^{(4)}-\lambda^{(3)}} \approx 0{,}19$ $\quad \dfrac{\lambda^{(6)}-\lambda^{(5)}}{\lambda^{(5)}-\lambda^{(4)}} \approx 0{,}20$

$\dfrac{\lambda^{(7)}-\lambda^{(6)}}{\lambda^{(6)}-\lambda^{(5)}} \approx 0{,}196$ $\quad \dfrac{\lambda^{(8)}-\lambda^{(7)}}{\lambda^{(7)}-\lambda^{(6)}} \approx 0{,}23$

$\Rightarrow$ conv. rate $\approx 0{,}2$

b) • power method applied to $(A-7I)^{-1}$ converges to $\lambda = 1$ of $(A-7I)^{-1}$

relation : $\lambda = \dfrac{1}{\mu - 7}$ with $\mu \in \sigma(A)$, $\lambda \in \sigma((A-7I)^{-1})$

$\Rightarrow \mu = \dfrac{1}{\lambda} + 7 = \dfrac{1}{1} + 7 = 8$

method converges to eigenvalue 8 of A (i.e. closest eigenvalue of A to 7)

• theoretical convergence rate = $\dfrac{|\lambda_2|}{|\lambda_1|}$ $\left[\text{th. conv. rate, if } \lambda_1, \lambda_2 \text{ real} = \text{sign}\left(\dfrac{\lambda_2}{\lambda_1}\right) \dfrac{|\lambda_2|}{|\lambda_1|} \right]$

with λ_1 largest eigenvalue of $(A-7I)^{-1}$ and λ_2 one but largest eigenvalue

$\lambda_1 = 1 \Rightarrow \dfrac{|\lambda_2|}{1} \approx 0.2 \Rightarrow |\lambda_2| \approx 0.2 \Rightarrow \lambda_2 \approx 0.2$

conv rate $\approx 0{,}2$

• since (see a.) conv. rate > 0 in each step, λ_1 and λ_2 same sign $\Rightarrow \lambda_2 > 0$

• λ_2 is real, if not $\lambda_2 = \bar{\lambda}_3$, i.e $|\lambda_2| = |\lambda_3|$ and conv. rate would not be constant over (i)

$\Rightarrow \mu = \dfrac{1}{\lambda_2} + 7 = \dfrac{1}{0.2} + 7 = 12$ second eigenvalue of A relevant for convergence (i.e second closest eigenvalue of A to 7)

or $\dfrac{1}{|\mu_2 - 7|} \Big/ \dfrac{1}{|\mu_1 - 7|} = \dfrac{|\mu_1 - 7|}{|\mu_2 - 7|} \approx 0.2$ $\quad \mu_1 = 8 \Rightarrow \mu_2 = \ldots$

Exc 4

a. $K^m(A,v) = \text{span}\{v, Av, A^2v, \dots, A^{m-1}v\}$

- orthogonalize using Gram-Schmidt process:

$$A[v_1 \cdots v_m] = [v_1 \cdots v_m \; v_{m+1}] H_{m+1\,m}$$

$$\Rightarrow AV_m = V_{m+1} H_{m+1\,m} \quad \text{with } H \text{ Hessenberg } (*)$$

- from this we see

$$V_m^T A V_m = V_m^T V_{m+1} H_{m+1\,m} = V_m^T [V_m \; v_{m+1}] H_{m+1\,m}$$

$$= [V_m^T V_m \;\; V_m^T v_{m+1}] H_{m+1\,m} = [I \;\; 0] H_{m+1\,m}$$

$$\Rightarrow V_m^T A V_m = H_{m\,m}$$

- suppose we want to find the eigenvector in the sub-space spanned by the columns of $V_m = [v_1 \cdots v_m]$.

i.e $\quad A V_m \hat{x} = \theta V_m \hat{x} \qquad (x = V_m \hat{x})$

- in general this is overdetermined, therefor condiscider: premultiplying by V_m^T gives

$$V_m^T A V_m \hat{x} = \theta V_m^T V_m \hat{x} = \theta \hat{x}$$

$$\Rightarrow H_{m\,m} \hat{x} = \theta \hat{x}$$

- so during the orthogonolization process $(*)$ we can compute the eigenvalues and eigenvectors of $H_{m,m}$ which give approximate eigenvalues and eigenvectors $V_m \hat{x}$ of A

b. $K^m(A,v) = K^m(A - \alpha I, v)$.

{ giving one of the two was sufficient
Here we give 2 important properties of this for its convergence

I. the consequence is that not only the eigenvector corresponding to the biggest eigenvalue of A will converge in the Krylov subspace, but also an eigenvector corresponding to an eigenvalue that can be made the biggest eigenvalue through choosing an appropriate α.

For the second property see last page

b make plausible $K^m(A, v) = K^m(A-\alpha I, v)$

- for $m=1$ $K^m(A, v) = \text{span}\{v\}$
  $K^m(A-\alpha I, v) = \text{span}\{v\}$ $\Rightarrow K^1(A,v) = K^1(A-\alpha I, v)$

- for $m=2$ $K^m(A, v) = \text{span}\{v, Av\}$

$$K^m(A-\alpha I, v) = \text{span}\{v, (A-\alpha I)v\}$$
$$= \text{span}\{v, (A-\alpha I)v + \beta v\} \text{ for arbitrary } \beta$$
$$= \text{span}\{v, (A-\alpha I)v + \alpha v\} \text{ set } \beta = \alpha$$
$$= \text{span}\{v, Av\} = K^m(A, v)$$

this can be repeated for any m

hence, it is plausible that $K^m(A, v) = K^m(A-\alpha I, v)$

- precise proof is not required for this exercise, but could be like:

note: $(*)$ $(A-\alpha I)^m v = \sum_{k=0}^{m} A^k (-\alpha)^{m-k} v = \sum_{k=0}^{m} (-\alpha)^{m-k} A^k v$

using induction: assume $K^{m-1}(A-\alpha I, v) = K^{m-1}(A, v)$

we have: $K^m(A-\alpha I) = \text{span}\{v, (A-\alpha I)v, \ldots, (A-\alpha I)^{m-1}v, (A-\alpha I)^m v\}$

$$= \text{span}\{v, Av, \ldots, A^{m-1}v, (A-\alpha I)^m v\}$$

↑ induction hypothesis

↳ use $(*)$
$$= \text{span}\{v, Av, \ldots, A^{m-1}v, \sum_{k=0}^{m} (-\alpha)^{m-k} A^k v\}$$
$$= \text{span}\{v, Av, \ldots, A^{m-1}v, \sum_{k=0}^{m} (-\alpha)^{m-k} A^k v + \beta_0 v + \beta_1 Av + \ldots + \beta_{m-1} A^{m-1} v\}$$

arbitrary β_i

take $\beta_i = -(-\alpha)^{m-i}$
$$= \text{span}\{v, Av, \ldots, A^{m-1}v, A^m v\}$$
$$= K^m(A-\alpha I)$$

2 Second important property of its convergence (see also previous page)
The eigenvalues of A can be shifted such that the convergence rate (i.e. the ratio of the second largest and the largest eigenvalues of the shifted matrix) can be made optimal. But note that this is only for analysis of convergence. One cannot control to which eigenvalue one converges as with shift and invere in Power method.