Scientific Computing HW 4

Ryan Chen

September 24, 2024

1. We follow the notation of the lecture notes. Assume A is $n \times n$ with distinct eigenvalues hence diagonalizable, with eigendecomposition
$$A = R\Lambda R^{-1}$$
Then the columns r_j of R are right eigenvectors, while the rows l_k of $L := R^{-1}$ are left eigenvectors. Moreover, $l_k r_j = 0$ for $k \neq j$. Write $\dot{r}_j$ in the basis of r_j's,
$$\dot{r}_j = \sum_{l=1}^{n} m_{jl} r_l$$
From the lecture notes,
$$\dot{A} r_j + A\dot{r}_j = \dot{\lambda}_j r_j + \lambda_r \dot{r}_j$$
Left–multiply both sides by l_k with $k \neq j$.
$$l_k \dot{A} r_j + \lambda_k m_{jk} = \lambda_j m_{jk} \implies m_{jk} = \frac{l_k \dot{A} r_j}{\lambda_j - \lambda_k}$$
We can assume $m_{jj} = 0$. Then
$$\begin{aligned}
\Delta r_j &= \dot{r}_j \Delta t + O(\|\Delta A\|^2) && \text{from lecture notes} \\
&= \sum_{k \neq j} m_{jk} r_k \Delta t + O(\|\Delta A\|^2) \\
&= \sum_{k \neq j} \frac{l_k \dot{A} r_j}{\lambda_j - \lambda_k} r_k \Delta t + O(\|\Delta A\|^2) \\
&= \sum_{k \neq j} \frac{l_k \Delta A r_j}{\lambda_j - \lambda_k} r_k + O(\|\Delta A\|^2)
\end{aligned}$$
Then
$$\begin{aligned}
\|\Delta r_j\| &\leq \sum_{k \neq j} \frac{|l_k \Delta A r_j|}{|\lambda_j - \lambda_k|} \|r_k\| + O(\|\Delta A\|^2) \\
&\leq \|\Delta A\| \|r_j\| \sum_{k \neq j} \frac{\|l_k\| \|r_k\|}{|\lambda_j - \lambda_k|} + O(\|\Delta A\|^2) \\
&\leq \|\Delta A\| \kappa(R) \|r_j\| \sum_{k \neq j} \frac{1}{|\lambda_j - \lambda_k|} + O(\|\Delta A\|^2) && \|l_k\| \|r_k\| \leq \kappa(R)
\end{aligned}$$

Thus

$$\kappa(r_j; A) = \lim_{\epsilon \to 0} \max_{\|\Delta A\| = \epsilon} \frac{\|\Delta r_j\| \|A\|}{\epsilon \|r_j\|} \leq \|A\| \kappa(R) \sum_{k \neq j} \frac{1}{|\lambda_j - \lambda_k|}$$

We see that $\kappa(r_j; A) \gg 1$ if $\kappa(R) \gg 1$ or $\lambda_k \approx \lambda_j$ for some $k \neq j$.

2. Before answering each part, we cite the following

$$\nabla\big(x^T Ax\big) = (A + A^T)x, \quad \nabla\big(x^T x\big) = 2x$$

We compute

$$\nabla Q(x) = \frac{(x^T x)(A + A^T)x - (x^T Ax)2x}{(x^T x)^2}$$

Thus

$$\nabla Q(x) = 0 \iff (x^T x)(A + A^T)x = 2(x^T Ax)x \iff (A + A^T)x = \frac{2x^T Ax}{x^T x}x$$

(a) Under the condition A is symmetric, the above computation gives

$$\nabla Q(x) = 0 \iff Ax = \frac{x^T Ax}{x^T x}x \iff \left(\frac{x^T Ax}{x^T x}, x\right) \text{ is an eigenpair of } A$$

(b) If A is asymmetric, the above calculation says $\nabla Q(x) = 0$ iff $\left(\frac{2x^T Ax}{x^T x}, x\right)$ is an eigenpair of $A + A^T$.

3. (a) Lemma: If A is nonsingular and (λ, v) is an eigenpair of A then (λ^{-1}, v) is an eigenpair of A^{-1}.
Proof of lemma: From A being nonsingular, $\lambda \neq 0$. Then

$$Av = \lambda v \implies v = \lambda A^{-1} v \implies A^{-1} v = \lambda^{-1} v \quad \square$$

Since μ is not an eigenvalue of A, we know $A - \mu I$ is nonsingular. Then

$$(A - \mu I)v = Av - \mu I v = \lambda v - \mu v = (\lambda - \mu)v$$

hence $((\lambda - \mu), v)$ is an eigenpair of $A - \mu I$. By the lemma, $((\lambda - \mu)^{-1}, v)$ is an eigenpair of $(A - \mu I)^{-1}$.

(b) We assume A is symmetric with distinct eigenvalues, so we can pick orthonormal eigenvectors v_i.
First write

$$\kappa((A - \mu I)^{-1}, v) = \left\| (A - \mu I)^{-1} \right\| \frac{\|v\|}{\|(A - \mu I)^{-1} v\|}$$

Now we calculate each RHS factor. By construction, $\|v\| = 1$.
Since A has eigenpairs (λ_i, v_i), by part (a), $(A - \mu I)^{-1}$ has eigenpairs $((\lambda_i - \mu)^{-1}, v_i)$. From $\mu \approx \lambda_1$ and A having distinct eigenvalues, the eigenvalue of $(A - \mu I)^{-1}$ with largest absolute value is $(\lambda_1 - \mu)^{-1}$, hence $\left\| (A - \mu I)^{-1} \right\| = |\lambda_1 - \mu|^{-1}$.
Since $(A - \mu I)^{-1}$ has eigenpairs $((\lambda_i - \mu)^{-1}, v_i)$,

$$(A - \mu I)^{-1} v = \left[1 - \sum_{i=2}^{n} \delta_i^2 \right]^{1/2} (\lambda_1 - \mu)^{-1} v_1 + \sum_{i=2}^{n} \delta_i (\lambda_i - \mu)^{-1} v_i$$

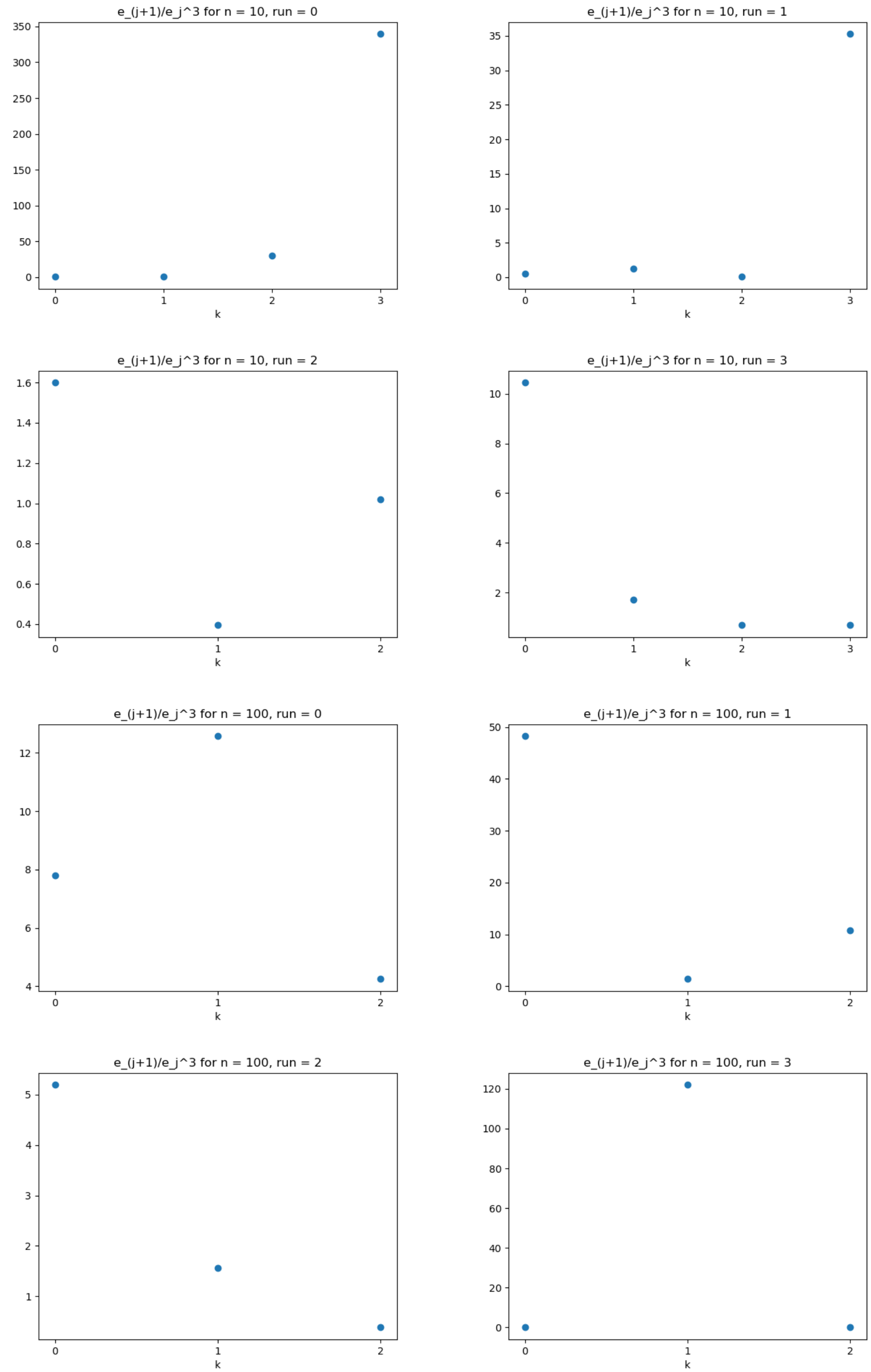
Thus

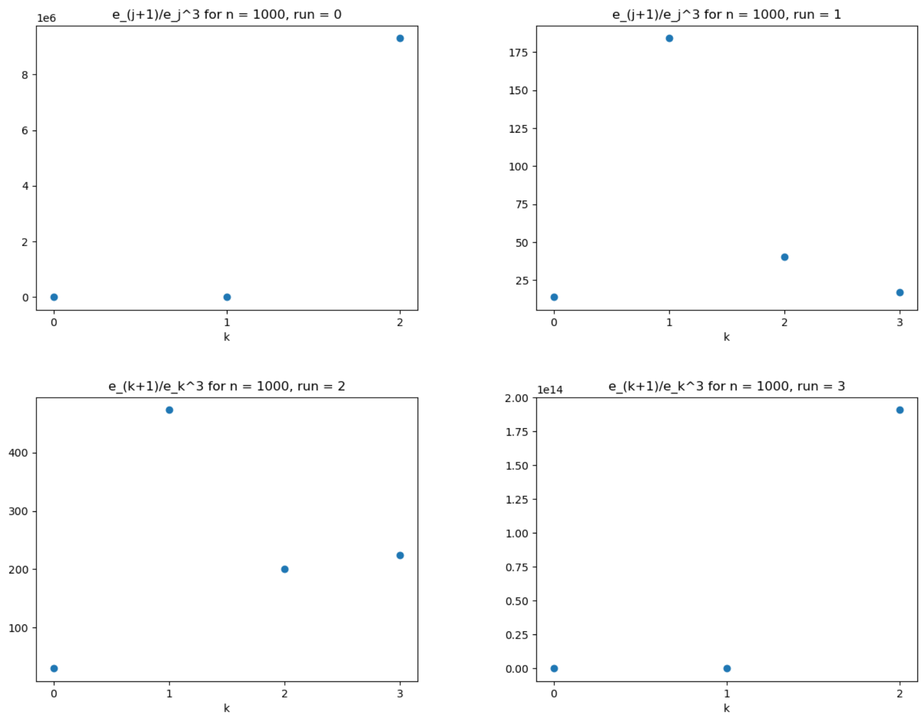
$$\begin{aligned}
\left\| (A - \mu I)^{-1} v \right\| &= \left\{ \left[1 - \sum_{i=2}^{n} \delta_i^2 \right] (\lambda_1 - \mu)^{-2} + \sum_{i=2}^{n} \delta_i^2 (\lambda_i - \mu)^{-2} \right\}^{1/2} && v_i\text{'s are orthonormal} \\
&= \left\{ (\lambda_1 - \mu)^{-2} - (\lambda_1 - \mu)^{-2} \sum_{i=2}^{n} \delta_i^2 + \sum_{i=2}^{n} \delta_i^2 (\lambda_i - \mu)^{-2} \right\}^{1/2} \\
&= \left\{ (\lambda_1 - \mu)^{-2} - \sum_{i=2}^{n} \delta_i^2 \left[(\lambda_1 - \mu)^{-2} - (\lambda_i - \mu)^{-2} \right] \right\}^{1/2}
\end{aligned}$$

Putting together the calculations,

$$\begin{aligned}
\kappa((A - \mu I)^{-1}, v) &= \left((\lambda_1 - \mu)^2 \right)^{-1/2} \left\{ (\lambda_1 - \mu)^{-2} - \sum_{i=2}^{n} \delta_i^2 \left[(\lambda_1 - \mu)^{-2} - (\lambda_i - \mu)^{-2} \right] \right\}^{1/2} \\
&= \left\{ 1 - \sum_{i=2}^{n} \delta_i^2 \left[1 - \left(\frac{\lambda_1 - \mu}{\lambda_i - \mu} \right)^2 \right] \right\}^{-1/2} \\
&\approx \left\{ 1 - \sum_{i=2}^{n} \delta_i^2 \right\}^{-1/2} && \mu \approx \lambda_1
\end{aligned}$$

(c) Code: `https://github.com/RokettoJanpu/scientific-computing-1-redux/blob/main/hw4.ipynb`


e_(j+1)/e_j^3 for n = 10, run = 0
350
300
250
200
150
100
50
0
0
1
2
3
k
e_(j+1)/e_j^3 for n = 10, run = 1
35
30
25
20
15
10
5
0
0
1
2
3
k
e_(j+1)/e_j^3 for n = 10, run = 2
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0
1
2
k
e_(j+1)/e_j^3 for n = 10, run = 3
10
8
6
4
2
0
1
2
3
k
e_(j+1)/e_j^3 for n = 100, run = 0
12
10
8
6
4
0
1
2
k
e_(j+1)/e_j^3 for n = 100, run = 1
50
40
30
20
10
0
0
1
2
k
e_(j+1)/e_j^3 for n = 100, run = 2
5
4
3
2
1
0
1
2
k
e_(j+1)/e_j^3 for n = 100, run = 3
120
100
80
60
40
20
0
0
1
2
k



Indeed, we do not have enough iterates to observe an eventual "leveling out" of the sequence $\frac{e_{j+1}}{e_j^3}$ as promised by cubic convergence.