1. QR factorization gives a sequence of matrices $\{A^{(0)}, A^{(1)}, A^{(2)}, \ldots\}$, where

$$A^{(0)} = \begin{pmatrix} 1 & 2 & 9 \\ 0 & 2 & 1 \\ 1 & 2 & -3 \end{pmatrix}$$

Find the QR factorization of $A^{(0)}$ by Gram-Schmidt process. Also compute $A^{(1)}$. Please show all your steps.

Solution:

$$\mathbf{q}_1 = \frac{\mathbf{a}_1}{||\mathbf{a}_1||} = \begin{pmatrix} \sqrt{2}/2 \\ 0 \\ \sqrt{2}/2 \end{pmatrix}$$

$$\begin{aligned}
\tilde{\mathbf{q}}_2 &= \mathbf{a}_2 - (\mathbf{q}_1^T \mathbf{a}_2)\mathbf{q}_1 \\
&= \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} - 2\sqrt{2} \begin{pmatrix} \sqrt{2}/2 \\ 0 \\ \sqrt{2}/2 \end{pmatrix} \\
&= \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} \\
\mathbf{q}_2 &= \frac{\tilde{\mathbf{q}}_2}{||\tilde{\mathbf{q}}_2||} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \\
\tilde{\mathbf{q}}_3 &= \mathbf{a}_3 - (\mathbf{q}_1^T \mathbf{a}_3)\mathbf{q}_1 - (\mathbf{q}_2^T \mathbf{a}_3)\mathbf{q}_2 \\
&= \begin{pmatrix} 9 \\ 1 \\ -3 \end{pmatrix} - 3\sqrt{2} \begin{pmatrix} \sqrt{2}/2 \\ 0 \\ \sqrt{2}/2 \end{pmatrix} - 1 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} 6 \\ 0 \\ -6 \end{pmatrix} \\
\mathbf{q}_3 &= \frac{\tilde{\mathbf{q}}_3}{||\tilde{\mathbf{q}}_3||} = \begin{pmatrix} \sqrt{2}/2 \\ 0 \\ -\sqrt{2}/2 \end{pmatrix}
\end{aligned}$$

So,

$$Q = \begin{pmatrix} \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \\ 0 & 1 & -1 \\ \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2} \end{pmatrix} \text{ and } R = \begin{pmatrix} \sqrt{2} & 2\sqrt{2} & 3\sqrt{2} \\ 0 & 2 & 1 \\ 0 & 0 & 6\sqrt{2} \end{pmatrix}$$

Therefore,

$$A^{(1)} = RQ = \begin{pmatrix} 4 & 2\sqrt{2} & -2 \\ \frac{\sqrt{2}}{2} & 2 & -\frac{\sqrt{2}}{2} \\ 6 & 0 & -6 \end{pmatrix} \tag{1}$$

2. Let A be a non-singular $n \times n$ real matrix. We apply the QR method on A to obtain a sequence of matrices $\{A^{(j)}\}_{j=0}^{\infty}$, which satisfies:

$$\begin{aligned} A^{(0)} &= A; \\ A^{(j+1)} &= R^{(j)}Q^{(j)} \text{ for } j = 0, 1, 2, ..., \end{aligned}$$

where $A^{(j)} = Q^{(j)}R^{(j)}$ is the QR factorization of $A^{(j)}$. Let k be an integer greater than 2020. Given that the QR factorizations of A^{k-1} and $A^{(k-1)}$ are given by

$$A^{k-1} = Q_1R_1 \text{ and } A^{(k-1)} = Q_2R_2.$$

In this question, all QR factorization is obtained in such a way that the diagonal entries of the upper triangular matrix are positive.

(a) Express A^2 in terms of $Q^{(0)}$, $Q^{(1)}$, $R(0)$, and $R^{(1)}$, and show that A^k can be expressed in terms of $Q^{(0)}, \dots, Q^{(k-1)}$ and $R^{(0)}, \dots, R^{(k-1)}$. Please explain your answer in detail.

(b) Express A in terms of Q_1, Q_2, R_1 and R_2 only. Please explain your answer in detail.

(c) Starting from $\mathbf{x}_0$, we apply the Power's method on A as follows:

$$\mathbf{x}_{j+1} = \frac{A\mathbf{x}_j}{||A\mathbf{x}_j||_\infty} \text{ for } j = 0, 1, 2, ...$$

Write $\mathbf{x}_k$ in terms of $\mathbf{x}_0$, Q_1, Q_2, R_1 and R_2 only (without A and k). Please explain your answer in detail.

Solution:

(a)

$$\begin{aligned} A^2 =& Q^{(0)}R^{(0)}Q^{(0)}R^{(0)} \\ =& Q^{(0)}A^{(1)}R^{(0)} \\ =& Q^{(0)}Q^{(1)}R^{(1)}R^{(0)} \end{aligned}$$

By induction, suppose $A^k = Q^{(0)} \dots Q^{(k-1)}R^{(k-1)} \dots R^{(0)}$, then

$$\begin{aligned} A^{k+1} = A^kA =& Q^{(0)} \dots Q^{(k-1)}R^{(k-1)} \dots R^{(0)}(Q^{(0)}R^{(0)}) \\ =& Q^{(0)} \dots Q^{(k-1)}R^{(k-1)} \dots (R^{(0)}Q^{(0)})R^{(0)} \\ =& Q^{(0)} \dots Q^{(k-1)}R^{(k-1)} \dots R^{(1)}Q^{(1)}R^{(1)}R^{(0)} \\ =& Q^{(0)} \dots Q^{(k-1)}R^{(k-1)} \dots Q^{(2)}R^{(2)}R^{(1)}R^{(0)} \\ =& \dots \\ =& Q^{(0)} \dots Q^{(k-1)}Q^{(k)}R^{(k)}R^{(k-1)} \dots R^{(0)} \end{aligned}$$

This prove that $A^k = Q^{(0)} \dots Q^{(k-1)}R^{(k-1)} \dots R^{(0)}$.

(b) Notice that $A^{k-1} = Q_1 R_1$, so

$$A^{k-1} = Q^{(0)} Q^{(1)} \cdots Q^{(k-2)} R^{(k-2)} R^{(k-3)} \cdots R^{(0)} = Q_1 R_1$$

By the uniqueness of the QR factorisation while restricting the diagonal entrie of the upper triangular matrix as positive, we get that

$$Q^{(0)} Q^{(1)} \cdots Q^{(k-2)} = Q_1, R^{(k-2)} R^{(k-3)} \cdots R^{(0)} = R_1$$

In addition,

$$A^k = Q^{(0)} Q^{(1)} \cdots Q^{(k-2)} Q_2 R_2 R^{(k-2)} R^{(k-3)} \cdots R^{(0)} = Q_1 Q_2 R_2 R_1$$

since $Q_2 = Q^{(k-1)}$ and $R_2 = R^{(k-1)}$.
So we get $A = R_1^{-1} Q_2 R_2 R_1$.

(c) It is easy to verify that

$$x_k = \frac{A^k x_0}{\|A^k x_0\|_\infty}$$

So $x_k = \frac{Q_1 Q_2 R_2 R_1 x_0}{\|Q_1 Q_2 R_2 R_1 x_0\|_\infty}$.

3. Let $A \in M_{n\times n}(\mathbb{C})$ be a $n \times n$ complex-valued matrix. Suppose the characteristic polynomial of A is given by: $f_A(t) = (-1)^n (t-\lambda_1)(t-\lambda_2)...(t-\lambda_n)$, where $\lambda_1, ..., \lambda_n$ are eigenvalues of A. Assuming that

$$|\lambda_1| = |\lambda_2| = ... = |\lambda_k| > |\lambda_{k+1}| \geq ... \geq |\lambda_n|,$$

where $k < n$. Suppose $A = QJQ^{-1}$, where J is the Jordan canonical form of A and Q is an invertible matrix. Assuming that the diagonal entries of J are arranged in descending order in terms of their magnitudes. Denote the j-th column of Q by $\mathbf{q}_j$, where $\mathbf{q}_1$, $\mathbf{q}_2$, ..., $\mathbf{q}_k$ are eigenvectors of A associated to λ_1, λ_2,..., λ_k respectively.

Let $\mathbf{x}_0$ be the initial vector defined as $\mathbf{x}_0 = a_1\mathbf{q}_1 + a_2\mathbf{q}_2 + ... + a_n\mathbf{q}_n$, where $a_j \in \mathbb{C}$ for $1 \leq j \leq n$ and $a_i \neq 0$ for $i = 1, 2, ..., k$. Consider the iterative scheme:

$$\mathbf{x}_{j+1} = \frac{A\mathbf{x}_j}{||A\mathbf{x}_j||_\infty} \text{ for } j = 0, 1, 2, ...$$

(a) Suppose $\lambda_1 = \lambda_2 = ... = \lambda_k \in \mathbb{R}$. will $||A\mathbf{x}_j||_\infty$ always converge as $j \to \infty$. If yes, what will it converge to? If not, please give a counter-example and explain your answer with details. Please show the full details of your proof.

(b) In general, if $|\lambda_1| = |\lambda_2| = ... = |\lambda_k|$, will $||A\mathbf{x}_j||_\infty$ always converge $j \to \infty$? If yes, what will it converge to? If not, please give a counter-example and explain your answer with details. Please show the full details of your proof.

solution:
It's easy to find for all $m \in \mathbb{N}^+$

$$\mathbf{x}_m = \frac{A\mathbf{x}_{m-1}}{\|A\mathbf{x}_{m-1}\|_\infty} = \frac{A^2\mathbf{x}_{m-2}}{\|A\mathbf{x}_{m-1}\|_\infty \|A\mathbf{x}_{m-2}\|_\infty} = \cdots = \frac{A^m\mathbf{x}_0}{\prod_{i=0}^{m-1} \|A\mathbf{x}_i\|_\infty}.$$

On the other side, we have $\|\mathbf{x}_m\|_\infty = 1$, so $\prod_{i=0}^{m-1} \|A\mathbf{x}_i\|_\infty = \|A^m\mathbf{x}_0\|_\infty$ and then

$$\mathbf{x}_m = \frac{A^m\mathbf{x}_0}{\|A^m\mathbf{x}_0\|_\infty}.$$

From the definition of $\mathbf{x}_0$,

$$A^m\mathbf{x}_0 = \sum_{i=1}^{n} a_i\lambda_i^m\mathbf{q}_i.$$

(a) Yes. Given $\lambda_1 = \lambda_2 = \cdots = \lambda_k \in \mathbb{R}$ and $|\lambda_1| > |\lambda_{k+1}| \geq |\lambda_{k+2}| \geq \cdots \geq |\lambda_n| > 0$, we can split $A^m\mathbf{x}_0$ into 2 parts,

$$A^m\mathbf{x}_0 = \lambda_1^m \sum_{i=1}^{k} a_i\mathbf{q}_i + \sum_{i=k+1}^{n} a_i\lambda_i^m\mathbf{q}_i = \lambda_1^m\mathbf{y} + \mathbf{z}_m.$$

When m is big enough, it's clear that $|\lambda_1|^m\|\mathbf{y}\|_\infty > \|\mathbf{z}_m\|_\infty$, $\lim_{m\to\infty} \frac{\|\mathbf{z}_m\|_\infty}{|\lambda_1|^m\|\mathbf{y}\|_\infty} = 0$ and

$$|\lambda_1|^m\|\mathbf{y}\|_\infty - \|\mathbf{z}_m\|_\infty \leq \|A^m\mathbf{x}_0\|_\infty \leq |\lambda_1|^m\|\mathbf{y}\|_\infty + \|\mathbf{z}_m\|_\infty.$$

Therefore, for such big m, we have

$$\frac{|\lambda_1|^{m+1}\|\mathbf{y}\|_\infty - \|\mathbf{z}_{m+1}\|_\infty}{|\lambda_1|^m\|\mathbf{y}\|_\infty + \|\mathbf{z}_m\|_\infty} \leq \|A\mathbf{x}_m\|_\infty = \frac{\|A^{m+1}\mathbf{x}_0\|_\infty}{\|A^m\mathbf{x}_0\|_\infty} \leq \frac{|\lambda_1|^{m+1}\|\mathbf{y}\|_\infty + \|\mathbf{z}_{m+1}\|_\infty}{|\lambda_1|^m\|\mathbf{y}\|_\infty - \|\mathbf{z}_m\|_\infty}.$$

For the left one,

$$\lim_{m\to\infty} \frac{|\lambda_1|^{m+1}\|\mathbf{y}\|_\infty - \|\mathbf{z}_{m+1}\|_\infty}{|\lambda_1|^m\|\mathbf{y}\|_\infty + \|\mathbf{z}_m\|_\infty} = |\lambda_1| \cdot \frac{1 - |\lambda_1| \cdot \lim_{m\to\infty} \frac{\|\mathbf{z}_{m+1}\|_\infty}{|\lambda_1|^{m+1}\|\mathbf{y}\|_\infty}}{1 + \lim_{m\to\infty} \frac{\|\mathbf{z}_m\|_\infty}{|\lambda_1|^m\|\mathbf{y}\|_\infty}} = |\lambda_1|.$$

Similarly, $\lim_{m\to\infty} \frac{|\lambda_1|^{m+1}\|\mathbf{y}\|_\infty + \|\mathbf{z}_{m+1}\|_\infty}{|\lambda_1|^m\|\mathbf{y}\|_\infty - \|\mathbf{z}_m\|_\infty} = |\lambda_1|$, which means $\lim_{m\to\infty} \|A\mathbf{x}_m\|_\infty = |\lambda_1|$.

(b) No. Suppose

$$J = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, Q = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}, A = QJQ^{-1} = \begin{pmatrix} \frac{3}{2} & \frac{1}{2} & -\frac{1}{2} \\ 2 & 0 & -2 \\ \frac{3}{2} & -\frac{3}{2} & -\frac{1}{2} \end{pmatrix},$$

Then for A, we have $\mathbf{q}_1 = (1, 1, 0)^T$, $\mathbf{q}_2 = (0, 1, 1)^T$, $\mathbf{q}_3 = (1, 0, 1)^T$, $\lambda_1 = 2$, $\lambda_2 = -2$ and $\lambda_3 = 1$.

Let $\mathbf{x}_0 = \mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3$,

$$A^m \mathbf{x}_0 = (2^m + 1, 2^m + (-2)^m, (-2)^m + 1)^T.$$

When m is odd, $\|A^m \mathbf{x}_0\|_\infty = 2^m + 1$ and when m is even, $\|A^m \mathbf{x}_0\|_\infty = 2^{m+1}$, hence

$$\|A\mathbf{x}_m\|_\infty = \frac{\|A^{m+1}\mathbf{x}_0\|_\infty}{\|A^m \mathbf{x}_0\|_\infty} = \begin{cases} 2 - \frac{2}{2^m+1}, m \text{ is odd} \\ \frac{1}{2} + \frac{1}{2^{m+1}}, m \text{ is even} \end{cases}$$

which means $\|A\mathbf{x}_m\|_\infty$ diverges.

4. Let A be an $n \times n$ complex matrix, whose eigenvalues satisfy:

$$|\lambda_1| > |\lambda_2| > |\lambda_3| > \cdots > |\lambda_n| > 0$$

Also, we define the following:

$$\begin{aligned} \cos \angle(x, y) &= \frac{|\langle x, y \rangle|}{\|x\|\|y\|}; \\ \sin \angle(x, y) &= \sqrt{1 - \cos^2 \angle(x, y)}; \\ \tan \angle(x, y) &= \frac{\sin \angle(x, y)}{\cos \angle(x, y)}. \end{aligned}$$

If $\cos \angle(x, y) = 0$, then let $\tan \angle(x, y) = \infty$. Here, $\langle x, y \rangle = \sum_{i=1}^n x_i \bar{y}_i$, where x_i and y_i are the i-th entries of $x \in \mathbb{C}^n$ and $y \in \mathbb{C}^n$ respectively.

(a) Suppose $A = QDQ^*$ where Q is unitary (i.e. $Q^*Q = I$, where Q^* is the conjugate transpose of Q) and D is diagonal, prove that

$$\cos \angle(Q^* x, Q^* y) = \cos \angle(x, y).$$

(b) Consider the power method in the form

$$x^{(n)} = Ax^{(n-1)}$$

Assume that $\cos \angle(x^{(0)}, e_1) \neq 0$, where e_1 is the eigenvector associated to the dominant eigenvalue λ_1. Prove that

$$\tan \angle(x^{(m+1)}, e_1) \leq \frac{|\lambda_2|}{|\lambda_1|} \tan \angle(x^{(m)}, e_1).$$

(c) For some $\mu \in \mathbb{R}$, let $A - \mu I$ be invertible. Assume

$$|\lambda_1 - \mu| < |\lambda_2 - \mu| \leq \cdots \leq |\lambda_n - \mu|.$$

Under the same notations and assumptions in (b), consider the shifted inverse power iteration

$$x^{(m)} = (A - \mu I)^{-1} x^{(m-1)}.$$

Using part (b) or otherwise, prove that

$$\tan \angle(x^{(m+1)}, e_1) \le \frac{|\lambda_1 - \mu|}{|\lambda_2 - \mu|} \tan \angle(x^{(m)}, e_1).$$

(d) Suppose A is a real matrix, and for $x \in \mathbb{R}^n \setminus \{0\}$, define the Rayleigh quotient $R(x, A) = \frac{x^* A x}{x^* x}$. Let r be a nonzero eigenvector of A for the eigenvalue λ, show that

$$|R(x, A) - \lambda| \le \rho(A - \lambda I) \sin \angle(x, r) \le \rho(A - \lambda I) \tan \angle(x, r)$$

(Hint: show that $\sin \angle(x, y) = \min \left\{ \frac{\|x - \alpha y\|}{\|x\|} : \alpha \in \mathbb{R} \right\}$ and note that $(A - \lambda I) r = 0$)

solution:

(a)

$$\cos \angle(Q^* x, Q^* y) = \frac{|\langle Q^* x, Q^* y \rangle|}{\|Q^* x\| \|Q^* y\|} = \frac{|\langle Q Q^* x, y \rangle}{\|x\| \|y\|} = \frac{|\langle x, y \rangle|}{\|x\| \|y\|} = \cos \angle(x, y)$$

(b) Denote $\hat{x}^{(m)} = Q^* x^{(m)}$, and δ_1 be the first column of I_n

$$\cos^2 \angle(x^{(m)}, e_1) = \cos^2 \angle(\hat{x}^{(m)}, \delta_1) = \frac{|\langle \hat{x}^{(m)}, \delta_1 \rangle|^2}{\|\hat{x}^{(m)}\|^2 \|\delta_1\|^2} = \frac{|\hat{x_1}^{(m)}|^2}{\|\hat{x}^{(m)}\|^2}$$

$$\sin^2 \angle(x^{(m)}, e_1) = 1 - \cos^2 \angle(x^{(m)}, e_1) = \frac{\sum_{j=2}^n |\hat{x_j}^{(m)}|^2}{\|\hat{x}^{(m)}\|^2}$$

$$\tan^2 \angle(x^{(m)}, e_1) = \frac{\sin^2 \angle(x^{(m)}, e_1)}{\cos^2 \angle(x^{(m)}, e_1)} = \frac{\sum_{j=2}^n |\hat{x_j}^{(m)}|^2}{|\hat{x_1}^{(m)}|^2}$$

Hence,

$$\tan^2 \angle(x^{(m+1)}, e_1) = \frac{\sum_{j=2}^n |\hat{x_j}^{(m+1)}|^2}{|\hat{x_1}^{(m+1)}|^2} = \frac{\sum_{j=2}^n |\lambda_j \hat{x_j}^{(m)}|^2}{|\lambda_1 \hat{x_1}^{(m)}|^2}$$
$$\le \left(\frac{|\lambda_2|}{|\lambda_1|} \right)^2 \frac{\sum_{j=2}^n |\hat{x_j}^{(m)}|^2}{|\hat{x_1}^{(m)}|^2}$$

(c) Let $\hat{A} = (A - \mu I)^{-1}$ with eigenvalues $\hat{\lambda}_j = (\lambda_j - \mu)^{-1}$. Note

$$Q^* \hat{A} Q = Q^* (A - \mu I)^{-1} Q = (D - \mu I)^{-1} = \operatorname{diag}(\hat{\lambda}_j)$$

By assumption, we have $|\hat{\lambda}_1| > |\hat{\lambda}_2| \geq \cdots \geq |\hat{\lambda}_n|$. Thus we have

$$\tan \angle(x^{(m+1)}, e_1) \leq \frac{|\hat{\lambda}_2|}{|\hat{\lambda}_1|} \tan \angle(x^{(m)}, e_1)$$

This completes the proof.

(d) We let $\alpha_0 := \langle x, y \rangle / \|y\|^2$ and observe that

$$\langle x - \alpha_0 y, y \rangle = \langle x, y \rangle - \frac{\langle x, y \rangle}{\|y\|^2} \langle y, y \rangle = \langle x, y \rangle - \langle x, y \rangle = 0,$$

i.e., $x - \alpha_0 y$ and y are perpendicular vectors. For $\alpha \in \mathbb{R}$ and $\beta := \alpha - \alpha_0$, this implies

$$\begin{aligned}
\|x - \alpha y\|^2 &= \|x - \alpha_0 y - \beta y\|^2 \\
&= \langle (x - \alpha_0 y) - \beta y, (x - \alpha_0 y) - \beta y \rangle \\
&= \|x - \alpha_0 y\|^2 - \langle \beta y, (x - \alpha_0 y) \rangle - \langle (x - \alpha_0 y), \beta y \rangle + |\beta|^2 \|y\|^2 \\
&= \|x - \alpha_0 y\|^2 + |\beta|^2 \|y\|^2,
\end{aligned}$$

i.e., the right-hand side attains its minimum for $\alpha = \alpha_0$. Due to

$$\begin{aligned}
\|x - \alpha_0 y\|^2 &= \|x\|^2 - \bar{\alpha}_0 \langle x, y \rangle - \alpha_0 \langle y, x \rangle + |\alpha_0|^2 \|y\|^2 \\
&= \|x\|^2 - \frac{|\langle x, y \rangle|^2}{\|y\|^2} - \frac{|\langle x, y \rangle|^2}{\|y\|^2} + \frac{|\langle x, y \rangle|^2}{\|y\|^2} \\
&= \|x\|^2 \left(1 - \frac{|\langle x, y \rangle|^2}{\|x\|^2 \|y\|^2} \right) \\
&= \|x\|^2 (1 - \cos^2 \angle(x, y)) = \|x\|^2 \sin^2 \angle(x, y),
\end{aligned}$$

this minimum has to be $\sin \angle(x, y)$.

Since r is an eigenvector, we have $(A - \lambda I) r = 0$, and we can use the Cauchy-Schwarz inequality $|\langle x, y \rangle| \leq \|x\| \, \|y\|$ and the compatibility inequality of the spectral norm $\|Ax\| \leq \rho(A) \, \|x\|$ to find

$$\begin{aligned}
|R(x, A) - \lambda| &= \left| \frac{\langle Ax, x \rangle}{\langle x, x \rangle} - \frac{\langle \lambda x, x \rangle}{\langle x, x \rangle} \right| = \frac{|\langle (A - \lambda I) x, x \rangle|}{\langle x, x \rangle} \\
&= \frac{|\langle (A - \lambda I)(x - \alpha r), x \rangle|}{\langle x, x \rangle} \leq \frac{\|(A - \lambda I)(x - \alpha r)\| \, \|x\|}{\|x\|^2} \\
&\leq \frac{\rho(A - \lambda I) \, \|x - \alpha r\| \, \|x\|}{\|x\|^2} \\
&= \rho(A - \lambda I) \frac{\|x - \alpha r\|}{\|x\|} \quad \text{for all } \alpha \in \mathbb{R}.
\end{aligned}$$

Hence, $|R(x, A) - \lambda| \leq \rho(A - \lambda I) \min_{\alpha} \frac{\|x - \alpha r\|}{\|x\|} = \rho(A - \lambda I) \sin \angle(x, r)$