

Assignment 5

Lokesh Mohanty (SR no: 21014)

November 2022

Problem 1

November 28, 2022

```
[1]: import numpy as np
      from matplotlib import pyplot as plt
```

```
[2]: # Markov Transition Matrix
A = np.array([[0, 0, 1, 0, 0, 0, 0, 0, 0, 0 ],
              [0.5, 0, 0, 0, 0, 0, 0, 0, 0, 0.5 ],
              [0, 0.5, 0, 0, 0, 0, 0.5, 0, 0, 0 ],
              [0, 0, 0, 0, 0, 0, 0, 0, 0, 0 ],
              [0, 0, 0, 1, 0, 0, 0, 0, 0, 0 ],
              [0, 0, 0, 0, 0.5, 0, 0, 0.5, 0, 0 ],
              [0, 0, 0, 0, 0, 0.5, 0, 0, 0, 0 ],
              [0, 0.5, 0, 0, 0.5, 0, 0, 0, 0, 1 ],
              [0.5, 0, 0, 0, 0, 0, 0.5, 0.5, 0, 0 ],
              [0, 0, 0, 0, 0, 0.5, 0, 0, 0.5, 0 ]])

A
```

```
[2]: array([[0. , 0. , 1. , 0. , 0. , 0. , 0. , 0. , 0. , 0. ],
            [0.5, 0. , 0. , 0. , 0. , 0. , 0. , 0. , 0.5, 0. ],
            [0. , 0.5, 0. , 0. , 0. , 0. , 0.5, 0. , 0. , 0. ],
            [0. , 0. , 0. , 0. , 0. , 0. , 0. , 0. , 0. , 0. ],
            [0. , 0. , 0. , 1. , 0. , 0. , 0. , 0. , 0. , 0. ],
            [0. , 0. , 0. , 0. , 0.5, 0. , 0. , 0.5, 0. , 0. ],
            [0. , 0. , 0. , 0. , 0. , 0.5, 0. , 0. , 0. , 0. ],
            [0. , 0.5, 0. , 0. , 0.5, 0. , 0. , 0. , 0. , 1. ],
            [0.5, 0. , 0. , 0. , 0. , 0. , 0.5, 0.5, 0. , 0. ],
            [0. , 0. , 0. , 0. , 0. , 0.5, 0. , 0. , 0.5, 0. ]])
```

```
[3]: v = np.random.randn(10)
      v /= np.linalg.norm(v)

      V = np.copy(v)
      RQ = []

      for k in range(300):
          w = A.dot(v)
          v = w / np.linalg.norm(w)
          RQ.append(np.dot(v,A.dot(v)))
          V = np.vstack((V,v))
```

```

print("Rayleigh quotient : ", RQ[-1])
print("Iterations taken : ", k)
print("Norm of A*v - RQ[-1]*v", np.linalg.norm(A.dot(v) - RQ[-1]*v))

```

```

Rayleigh quotient :  1.0000000000000002
Iterations taken :  299
Norm of A*v - RQ[-1]*v 1.8824747269678055e-16

```

```

[4]: residual_norm = np.linalg.norm(A.dot(V.T) - RQ[-1]*V.T, axis=0)

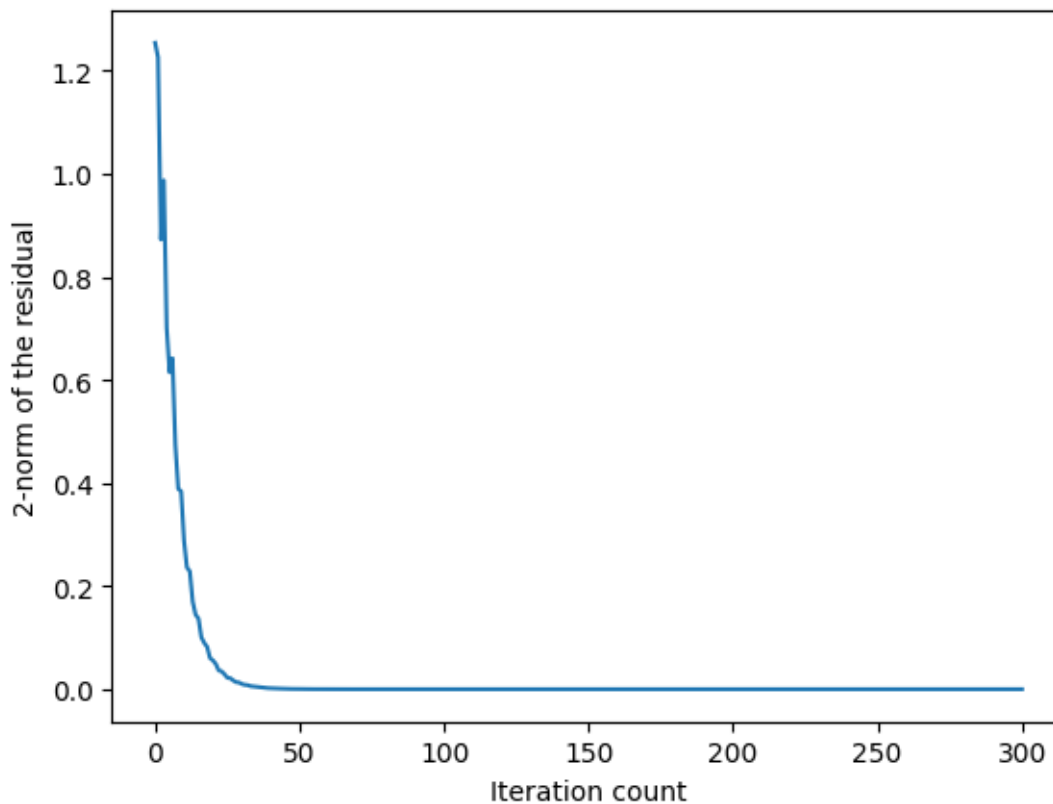
plt.xlabel("Iteration count")
plt.ylabel("2-norm of the residual")
plt.plot(residual_norm)

```

```

[4]: [<matplotlib.lines.Line2D at 0x7f173d0b29b0>]

```



```

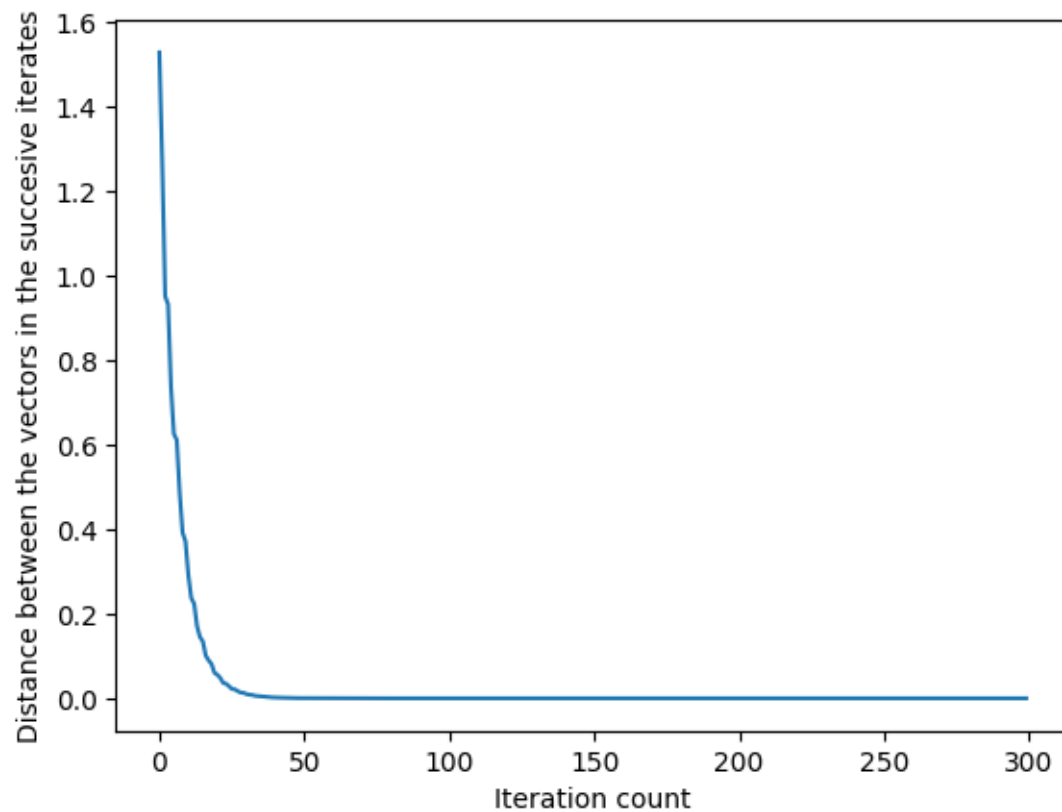
[5]: diff_norm = np.linalg.norm(V[:-1,:] - V[1:,:], axis=1)

plt.xlabel("Iteration count")
plt.ylabel("Distance between the vectors in the successive iterates")

```

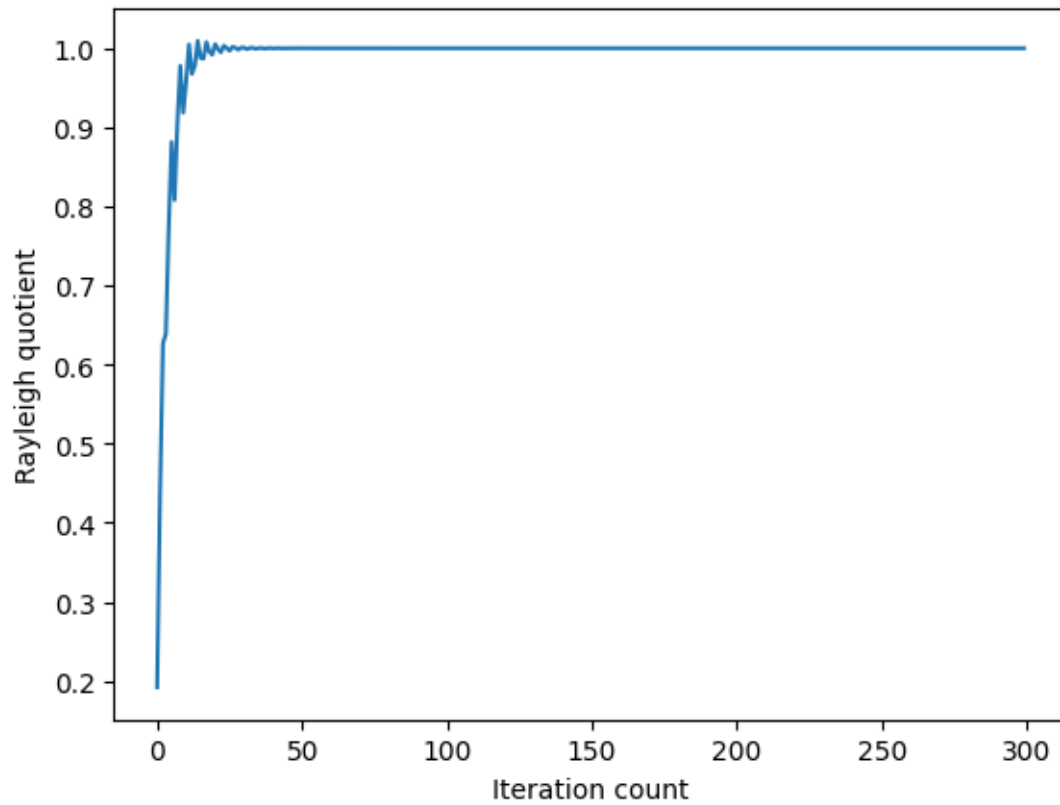
```
plt.plot(diff_norm)
```

[5]: [<matplotlib.lines.Line2D at 0x7f173cfcc6d0>]



```
[6]: plt.xlabel("Iteration count")  
plt.ylabel("Rayleigh quotient")  
plt.plot(RQ)
```

[6]: [<matplotlib.lines.Line2D at 0x7f173ce68040>]



```
[7]: page_rank = np.absolute(v)
page_rank /= np.sum(page_rank)

print("Highest page rank: ", np.argmax(np.absolute(v)))
print("Lowest page rank: ", np.argmax(-np.absolute(v)))
```

```
Highest page rank: 7
Lowest page rank: 3
```

Problem 2

(a) False

Pure QR:

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

Simultaneous Iteration:

$$\mathbf{A}^{(1)} = \mathbf{R}^{(1)}\mathbf{Q}^{(1)} = (\mathbf{Q}^{(1)})^T \mathbf{A} \mathbf{Q}^{(0)} \mathbf{Q}^{(1)}$$

Hence Pure QR algorithm is equivalent to Simultaneous Iteration only when $\mathbf{Q}^{(0)} = \mathbf{I}$.

(b) False

$$\begin{aligned}\mathbf{A}^{(k)} &= (\mathbf{Q}^{(k)})^T \mathbf{A}^{(k-1)} \mathbf{Q}^{(k)} \\ &= (\tilde{\mathbf{Q}}_k)^T \mathbf{A} \tilde{\mathbf{Q}}_k \\ \tilde{\mathbf{Q}}_k &= \mathbf{Q}^{(1)} \mathbf{Q}^{(2)} \dots \mathbf{Q}^{(k)}\end{aligned}$$

Hence as $k \rightarrow \infty$, $\tilde{\mathbf{Q}}_k$ converges to the eigenvector matrix but not the sequence of matrices $\mathbf{Q}^{(k)}$

(c) False. As at the kth iteration, $\mathbf{A}^{(k)} = \mathbf{Q}^{(k+1)}\mathbf{R}^{(k+1)}$

(d) True. As at each iteration,

- Pure QR:

QR factorization: $\mathbf{A}^{(k-1)} = \mathbf{Q}^{(k)}\mathbf{R}^{(k)}$

Matrix multiplication: $\mathbf{A}^{(k)} = \mathbf{R}^{(k)}\mathbf{Q}^{(k)}$

- Simultaneous iteration:

Matrix multiplication: $\mathbf{Z} = \mathbf{A} \mathbf{Q}^{(k-1)}$

QR factorization: $\mathbf{Q}^{(k)}\mathbf{R}^{(k)} = \mathbf{Z}$

Since both the algorithms require FLOPS of the same order i.e., $O(m^3)$, they have the same computational complexity.

Problem 3

(a) False. As we know that the roots of a polynomial equation of degree greater than or equal to 5 cannot be computed in finite number of steps and since a polynomial equation can always be assumed as the characteristic polynomial of a matrix (whose roots are the eigenvalues of the matrix), we can say that computing eigenvalues and eigenvectors of a matrix is not possible in a finite number of steps for $m \geq 5$.

(b) False. Without reducing to tridiagonal form (phase I) first, each iteration of phase II will cost $O(m^3)$ flops with the total cost being $O(m^4)$. In case the symmetric matrices, there might be improvement in the number flops but it still is not of the order $O(m^3)$.

(c) False. If $\mathbf{v}^{(0)}$ is equal to any eigenvector other than the one corresponding to the largest eigenvalue, then all the sequence of vectors will be the same eigenvector and hence will never converge to the largest eigenvalue.

(d) True. Let λ, \mathbf{v} be a eigenvalue-vector pair of \mathbf{A} , then

$$\implies \mathbf{A}\mathbf{v} = \lambda\mathbf{v}$$

$$\implies \mathbf{F}\mathbf{A}\mathbf{v} = \lambda\mathbf{F}\mathbf{v}$$

$$\implies (\mathbf{F}\mathbf{A} - \lambda\mathbf{F})\mathbf{v} = \mathbf{0}$$

$$\text{Let } \mathbf{u} = \mathbf{F}\mathbf{v} \implies \mathbf{v} = \mathbf{F}^{-1}\mathbf{u} \text{ and we know that } \mathbf{F}\mathbf{F}^{-1} = \mathbf{I}$$

$$\implies (\mathbf{F}\mathbf{A} - \lambda\mathbf{F})\mathbf{F}^{-1}\mathbf{u} = \mathbf{0}$$

$$\implies (\mathbf{F}\mathbf{A}\mathbf{F}^{-1} - \lambda\mathbf{F}\mathbf{F}^{-1})\mathbf{u} = \mathbf{0}$$

$$\implies (\mathbf{F}\mathbf{A}\mathbf{F}^{-1} - \lambda\mathbf{I})\mathbf{u} = \mathbf{0}$$

$$\implies \mathbf{F}\mathbf{A}\mathbf{F}^{-1}\mathbf{u} = \lambda\mathbf{u}$$

Therefore $\mathbf{F}\mathbf{A}\mathbf{F}^{-1}$ and \mathbf{A} have same eigenvalues.

Problem 4

(a) Let λ, \mathbf{u} be an eigenvalue-vector pair of \mathbf{A} , then

$$\begin{aligned}\mathbf{A}\mathbf{u} &= \lambda\mathbf{u} \\ \implies \mathbf{A}^T\mathbf{A}\mathbf{u} &= \lambda\mathbf{A}^T\mathbf{u} \\ &= \lambda\mathbf{A}\mathbf{u} \\ &= \lambda^2\mathbf{u}\end{aligned}$$

Since eigenvalues of $\mathbf{A}^T\mathbf{A}$ are square of singular values(σ) of \mathbf{A} ,

$$\begin{aligned}\implies \lambda^2 &= \sigma^2 \\ \implies |\lambda| &= |\sigma| \\ \implies |\lambda| &= \sigma\end{aligned}$$

(b) We know that Rayleigh coefficient of any vector lies between the minimum and maximum eigenvalue and since the largest eigenvalue of a matrix is its second norm,

$$\begin{aligned}\left| \frac{\mathbf{x}^T\mathbf{A}\mathbf{x}}{\mathbf{x}^T\mathbf{x}} \right| &\leq \|\mathbf{A}\| \\ \text{Since } \mathbf{x} \text{ is a unit vector,} \\ \implies |\mathbf{x}^T\mathbf{A}\mathbf{x}| &\leq \|\mathbf{A}\|\end{aligned}$$

(c)

$$\begin{aligned}
& (\mathbf{A} + \delta\mathbf{A})(\mathbf{u} + \delta\mathbf{u}) = (\lambda + \delta\lambda)(\mathbf{u} + \delta\mathbf{u}) \\
\implies & \mathbf{A}\delta\mathbf{u} + \delta\mathbf{A}(\mathbf{u} + \delta\mathbf{u}) = \delta\lambda(\mathbf{u} + \delta\mathbf{u}) + \lambda\delta\mathbf{u} \\
\implies & \mathbf{u}^T\mathbf{A}\delta\mathbf{u} + \mathbf{u}^T\delta\mathbf{A}(\mathbf{u} + \delta\mathbf{u}) = \mathbf{u}^T\delta\lambda(\mathbf{u} + \delta\mathbf{u}) + \lambda\mathbf{u}^T\delta\mathbf{u} \\
\implies & \mathbf{u}^T\delta\mathbf{A}(\mathbf{u} + \delta\mathbf{u}) = \mathbf{u}^T\delta\lambda(\mathbf{u} + \delta\mathbf{u}) \\
\implies & \delta\lambda = \frac{\mathbf{u}^T\delta\mathbf{A}(\mathbf{u} + \delta\mathbf{u})}{\mathbf{u}^T(\mathbf{u} + \delta\mathbf{u})} \\
& \text{we know that } \mathbf{u}^T(\mathbf{u} + \delta\mathbf{u}) \approx \|\mathbf{u}\|\|\mathbf{u} + \delta\mathbf{u}\| \\
& \text{and } \mathbf{u}^T\delta\mathbf{A}(\mathbf{u} + \delta\mathbf{u}) \leq \|\mathbf{u}\|\|\delta\mathbf{A}(\mathbf{u} + \delta\mathbf{u})\| \\
\implies & |\delta\lambda| \leq \frac{\|\mathbf{u}\|\|\delta\mathbf{A}(\mathbf{u} + \delta\mathbf{u})\|}{\|\mathbf{u}\|\|\mathbf{u} + \delta\mathbf{u}\|} \\
& \leq \frac{\|\delta\mathbf{A}(\mathbf{u} + \delta\mathbf{u})\|}{\|(\mathbf{u} + \delta\mathbf{u})\|} \\
& \leq \frac{\|\delta\mathbf{A}\|\|(\mathbf{u} + \delta\mathbf{u})\|}{\|(\mathbf{u} + \delta\mathbf{u})\|} \\
& \leq \|\delta\mathbf{A}\|
\end{aligned}$$

(d)

$$\begin{aligned}
\kappa &= \max_{\delta\mathbf{A}} \frac{\frac{|\delta\lambda|}{|\lambda|}}{\frac{\|\delta\mathbf{A}\|}{\|\mathbf{A}\|}} = \max_{\delta\mathbf{A}} \frac{|\delta\lambda|}{|\lambda|} \frac{\|\mathbf{A}\|}{\|\delta\mathbf{A}\|} \\
&= \max_{\delta\mathbf{A}} \frac{|\delta\lambda|}{\|\delta\mathbf{A}\|} \frac{\|\mathbf{A}\|}{|\lambda|} \\
&= \frac{\|\mathbf{A}\|}{|\lambda|} \geq 1 \\
\implies & \kappa \geq 1
\end{aligned}$$

Problem 5

(a) Let λ and \mathbf{u} be any eigenvalue-eigenvector pair of the matrix $(\alpha\mathbf{I} - \mathbf{M})$ where $\alpha \in \mathbb{R} \setminus \{0\}$, then

$$\begin{aligned} (\alpha\mathbf{I} - \mathbf{M})\mathbf{u} &= \lambda\mathbf{u} \\ \implies \mathbf{M}\mathbf{u} &= (\alpha - \lambda)\mathbf{u} \\ \implies \mathbf{M}^2\mathbf{u} &= (\alpha - \lambda)\mathbf{M}\mathbf{u} \\ \text{Since } \mathbf{M}^2 &= 0 \implies \mathbf{0} = (\alpha - \lambda)^2\mathbf{M}\mathbf{u} \end{aligned}$$

Since $\alpha \neq 0$ and $(\alpha - \lambda)^2 = 0$ for all eigen values, $\lambda = \alpha$. Which implies that the algebraic multiplicity of α is n .

\therefore True

(b) Let λ and \mathbf{u} be any eigenvalue-eigenvector pair of the matrix $(\alpha\mathbf{I} + \mathbf{M})$ where $\alpha \in \mathbb{R} \setminus \{0\}$, then

$$\begin{aligned} (\mathbf{M} + \alpha\mathbf{I})\mathbf{u} &= \lambda\mathbf{u} \\ \implies \mathbf{M}\mathbf{u} &= (\lambda - \alpha)\mathbf{u} \\ \implies \mathbf{M}^2\mathbf{u} &= (\lambda - \alpha)\mathbf{M}\mathbf{u} \\ \text{Since } \mathbf{M}^2 &= 0 \implies \mathbf{0} = (\lambda - \alpha)^2\mathbf{M}\mathbf{u} \end{aligned}$$

Since $\alpha \neq 0$ and $(\lambda - \alpha)^2 = 0$ for all eigen values, $\lambda = \alpha$. Which implies that the algebraic multiplicity of α is n . Hence there is only one unique eigenvalue.

\therefore False (since $k = 1$)

(c) Let $\lambda_1, \lambda_2, \lambda_3$ and $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ be three eigenvalue-eigenvector pairs of the matrix \mathbf{A} such that $\lambda_1 = \lambda_2 = \lambda_3$ and $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ are distinct.

$$\begin{aligned} \implies (\mathbf{A} - \lambda\mathbf{I})\mathbf{u}_1 &= 0 \\ (\mathbf{A} - \lambda\mathbf{I})\mathbf{u}_2 &= 0 \\ (\mathbf{A} - \lambda\mathbf{I})\mathbf{u}_3 &= 0 \end{aligned}$$

This implies that there are 3 vectors in the null space $\mathbf{A} - \lambda\mathbf{I}$. Hence the dimension of its null space is 3 and the dimension of its column space is $m - 3$.

\therefore False

(d) Since \mathbf{A} is diagonalizable, it can be written as $\mathbf{A} = \mathbf{M}\mathbf{D}\mathbf{M}^{-1}$ where \mathbf{D} is a diagonal matrix such that $\mathbf{D}, \mathbf{M} \in \mathbb{R}^{m \times m}$. Let λ, \mathbf{u} be any eigenvalue-vector pair of \mathbf{A} , then

$$\begin{aligned}\mathbf{A}\mathbf{u} &= \lambda\mathbf{u} \\ \implies \mathbf{M}\mathbf{D}\mathbf{M}^{-1}\mathbf{u} &= \lambda\mathbf{u} \\ \implies \mathbf{D}(\mathbf{M}^{-1}\mathbf{u}) &= \lambda(\mathbf{M}^{-1}\mathbf{u})\end{aligned}$$

Since \mathbf{M} is full rank, $\mathbf{M}^{-1}\mathbf{u}$ is unique for every \mathbf{u} . Hence the matrix \mathbf{D} and \mathbf{A} have same eigenvalues and rank and since \mathbf{D} is diagonal, its eigenvalues are its diagonal elements. And since the rank of a diagonal matrix is the number of non-zero diagonal elements, rank of \mathbf{A} is equal to the number of non-zero eigenvalues of \mathbf{A} .

\therefore True

Problem 6

(a) Given that, $\mathbf{A}\mathbf{w} = \mathbf{v}$

$$\begin{aligned}
 \mathbf{v} &= (\mathbf{A} + \delta\mathbf{A})(\mathbf{w} + \delta\mathbf{w}) \\
 \implies \mathbf{A}\mathbf{w} &= (\mathbf{A} + \delta\mathbf{A})\mathbf{w} + (\mathbf{A} + \delta\mathbf{A})\delta\mathbf{w} \\
 \implies 0 &= \delta\mathbf{A}\mathbf{w} + (\mathbf{A} + \delta\mathbf{A})\delta\mathbf{w} \\
 \implies \delta\mathbf{w} &= -(\mathbf{A} + \delta\mathbf{A})^{-1}\delta\mathbf{A}\mathbf{w}
 \end{aligned}$$

(b)

$$\begin{aligned}
 \mathbf{v} &= a_1\mathbf{q}_1 + a_2\mathbf{q}_2 + \dots + a_m\mathbf{q}_m \\
 \mathbf{w} &= \mathbf{A}^{-1}\mathbf{v} \\
 &= a_1\mathbf{A}^{-1}\mathbf{q}_1 + a_2\mathbf{A}^{-1}\mathbf{q}_2 + \dots + a_m\mathbf{A}^{-1}\mathbf{q}_m \\
 &= \frac{a_1}{\lambda_1} \left[\mathbf{q}_1 + \frac{a_2\lambda_1}{a_1\lambda_2}\mathbf{q}_2 + \dots + \frac{a_m\lambda_1}{a_1\lambda_m}\mathbf{q}_m \right] \\
 &\approx \frac{a_1}{\lambda_1}\mathbf{q}_1 \\
 \implies \mathbf{w} &\text{ is approximately in the direction of } \mathbf{q}_1 \\
 \implies \frac{\mathbf{w}}{\|\mathbf{w}\|} &\approx \mathbf{q}_1
 \end{aligned}$$

(c) Given that $(\mathbf{A} + \delta\mathbf{A})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}(\delta\mathbf{A})\mathbf{A}^{-1} + O(\epsilon_M^2)$,

$$\begin{aligned}
 \mathbf{v} &= (\mathbf{A} + \delta\mathbf{A})(\mathbf{w} + \delta\mathbf{w}) \\
 \implies (\mathbf{A} + \delta\mathbf{A})^{-1}\mathbf{v} &= \mathbf{w} + \delta\mathbf{w} \\
 \implies (\mathbf{A}^{-1} - \mathbf{A}^{-1}(\delta\mathbf{A})\mathbf{A}^{-1} + O(\epsilon_M^2))\mathbf{A}\mathbf{w} &= \mathbf{w} + \delta\mathbf{w} \\
 \implies \mathbf{w} + \delta\mathbf{w} &\approx (\mathbf{I} - \mathbf{A}^{-1}(\delta\mathbf{A}))\mathbf{w} \\
 \implies \delta\mathbf{w} &\approx -\mathbf{A}^{-1}(\delta\mathbf{A})\mathbf{w}
 \end{aligned}$$

Since λ_1 is very less than the other eigenvalues, $\|\mathbf{A}^{-1}\|$ will be very high and hence \mathbf{A}^{-1} will tend to extend the vector \mathbf{w} in the direction of \mathbf{q}_1 and since \mathbf{w} is already approximately in the direction of \mathbf{q}_1 , $\delta\mathbf{w}$ is also approximately in the direction of \mathbf{q}_1 . Since both \mathbf{w} and $\delta\mathbf{w}$ are in the direction of \mathbf{q}_1 , $\tilde{\mathbf{w}}$ is also in the direction of \mathbf{q}_1 . Hence $\frac{\tilde{\mathbf{w}}}{\|\tilde{\mathbf{w}}\|} \approx \mathbf{q}_1$.

Problem 7

(a) $\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 2 & 3 & 3 \end{bmatrix}$. Characteristic equation: $|\mathbf{A} - \lambda\mathbf{I}| = 0$

$$\implies (1 - \lambda)(2 - \lambda)(3 - \lambda) = 0$$

$$\implies \lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 3$$

$$\boxed{\therefore \text{The eigenvalues are: } 1, 2, 3}$$

(b) Given starting vector $\mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

$$\implies \mathbf{q}_1 = \frac{\mathbf{b}}{\|\mathbf{b}\|} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{aligned} \implies \mathbf{q}_2 &= \frac{\mathbf{v}_1}{h_{21}} = \frac{\mathbf{A}\mathbf{q}_1 - \mathbf{q}_1}{\|\mathbf{A}\mathbf{q}_1 - \mathbf{q}_1\|} \\ &= \frac{1}{\|\mathbf{A}\mathbf{q}_1 - \mathbf{q}_1\|} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 2 & 3 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 1 & 2 \end{bmatrix} \end{aligned}$$

(c) $K_2 = [\mathbf{b}|\mathbf{A}\mathbf{b}] = \mathbf{Q}_2\mathbf{R}_2$, where $\mathbf{Q}_2 = [\mathbf{a}_1|\mathbf{a}_2] = \begin{bmatrix} 1 & 0 \\ 0 & 1/\sqrt{5} \\ 0 & 2/\sqrt{5} \end{bmatrix}$

$$\mathbf{H}(\text{projection matrix}) = \mathbf{Q}_2^T \mathbf{A} \mathbf{Q}_2$$

$$\begin{aligned} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 2 & 3 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1/\sqrt{5} \\ 0 & 2/\sqrt{5} \end{bmatrix} \\ \implies \mathbf{H} &= \begin{bmatrix} 1 & 0 \\ \sqrt{5} & 4 \end{bmatrix} \end{aligned}$$

The eigenvalues of \mathbf{H} are

$$\begin{aligned} (1 - \lambda)(4 - \lambda) + 0 &= 0 \\ \implies \lambda &= 1, 4 \end{aligned}$$

Absolute error between smallest Ritz value and smallest eigenvalue of \mathbf{A} is $|1 - 1| = 0$.
 Absolute error between largest Ritz value and largest eigenvalue of \mathbf{A} is $|4 - 3| = 1$.

(d)

$$\begin{aligned} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 2 & 3 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} &= \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \\ \implies x &= 1 \\ \implies x + 2y = 0 &\implies y = -1/2 \\ \implies 2x + 3y + 3z = 0 &\implies z = -1/6 \\ \implies x^* &= \begin{bmatrix} 1 \\ -1/2 \\ -1/6 \end{bmatrix} \end{aligned}$$

Minimize $\|\mathbf{A}\mathbf{c} - \mathbf{b}\|$ over \mathbf{c} such that $\mathbf{c} \in K_2$

Least square solution($\tilde{\mathbf{x}}$):

$$\begin{aligned} \tilde{\mathbf{x}} &= \mathbf{A}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{b} \\ \mathbf{X} = \mathbf{A}K_2 &= \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 2 & 3 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 2 & 11 \end{bmatrix} \\ \implies \tilde{\mathbf{x}} &= \mathbf{A} \begin{bmatrix} 6 & 26 \\ 26 & 131 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 2 & 3 & 3 \end{bmatrix} \begin{bmatrix} 0.95 \\ -0.18 \end{bmatrix} \\ \therefore \tilde{\mathbf{x}} &= \begin{bmatrix} 0.77 \\ -0.18 \\ -0.37 \end{bmatrix} \end{aligned}$$

2 Norm of error between \mathbf{x}^* and $\tilde{\mathbf{x}}$ is

$$\left\| \begin{bmatrix} 1 \\ -1/2 \\ -1/6 \end{bmatrix} - \begin{bmatrix} 0.77 \\ -0.18 \\ -0.37 \end{bmatrix} \right\| = 0.44$$