Homework 2 AM213A

Kevin Silberberg

2015-01-29

1 Question 1

1.1 Part a

Let $A \in \mathbb{C}^{m \times m}$ be invertible and $\lambda \neq 0$ is an eigenvalue of A. In the following we will show that $\frac{1}{\lambda}$ is an eigenvalue of A^{-1} .

Proof.

$$Av = \lambda v \tag{1}$$

$$A^{-1}Av = A^{-1}\lambda v \tag{2}$$

$$Iv = \lambda A^{-1}v \tag{3}$$

$$\frac{1}{\lambda}v = A^{-1}v\tag{4}$$

We are left with the equation (4) which is the definition of the eigenvalue.

1.2 Part b

Let $A, B \in \mathbb{C}^{m \times m}$. In the following we will show that AB and BA have the same eigenvalues.

Proof. Let v be the eigenvectors and λ be the eigenvalues of the matrix AB, such that $ABv = \lambda v$. Let ω be a matrix the same size as v under the transformation Bv, such that $\omega = Bv$.

$$BA\omega = BABv \tag{5}$$

$$= B\lambda v \tag{6}$$

$$= \lambda B v \tag{7}$$

$$=\lambda\omega$$
 (8)

Thus we have that $BA\omega = \lambda \omega$, which is the definition of the eigenvalue.

1.3 Part c

Let us show that $\rho(A) \leq ||A||$, where $\rho(A)$ is the spectral radius of A and the norm is any p-norm.

Recall that the spectral radius of a matrix A is defined as

$$\rho(A) = \sup_{\lambda \in \sigma(A)} (|\lambda|) \tag{9}$$

where $\sigma(A)$ is the spectrum of A or

$$\sigma(A) := \{ \lambda \in \mathbb{C} : A - \lambda I \text{ is not bijective } \}$$
 (10)

Proof. Let λ be an eigenvalue of A, and let $v \neq 0$ be a corresponding eigenvector. Then by the submultiplicativity property of Matrix norms, we have,

$$Av = \lambda v \tag{11}$$

$$\|\lambda v\| = \|Av\| \tag{12}$$

$$|\lambda| \|v\| = \|Av\| \le \|A\| \|v\| \tag{13}$$

(14)

dividing both sides by ||v|| we obtain:

$$|\lambda| \le ||A|| \tag{15}$$

since this holds $\forall \lambda \in \sigma(A)$, we conclude that

$$\rho(A) \le \|A\| \tag{16}$$

1.4 Part d

Let $A \in \mathbb{R}^{m \times m}$. Let us show that A and A^* have the same eigenvalues.

Recall that the eigenvalues of the matrix A can be found by finding the roots of the characteristic polynomial,

$$p_A(\lambda) = \det(A - \lambda I) \tag{17}$$

Let us show that $p_A(\lambda) = p_{A^*}(\lambda)$.

Proof. By the property of the determinant that states that for any given matrix M,

$$\det(M) = \det(M^T) \tag{18}$$

it follows that:

$$\det(A - \lambda I) = \det((A - \lambda I)^T) \tag{19}$$

Since λ is just a scalar and I only has zeros in the off-diagonal entries, we conclude that

$$\det(A - \lambda I) = \det(A^T - \lambda I) \tag{20}$$

$$p_A(\lambda) = p_{A^T}(\lambda) \tag{21}$$

Noting that, since A is real-valued, its adjoint is just its transpose.

Furthermore, for real matrices, if λ is a complex eigenvalue of A, then its complex conjugate $\bar{\lambda}$ is also an eigenvalue. Since A^T is also real, its eigenvalues must satisfy the same property. Therefore, A^T has the same eigenvalues as A.

2 Question 2

Let $A \in \mathbb{C}^{m \times m}$ be Hermitian. Suppose that for non-zero eigenvectors of A, there exists corresponding eigenvalues λ satisfying $Ax = \lambda x$.

2.1 Part a

All eigenvalues of A are real.

Proof. Let λ be an eigenvalue of A with a corresponding eigenvector $x \neq 0$, satisfying $Ax = \lambda x$. Taking the conjugate transpose of both sides,

$$(\bar{Ax})^T = (\bar{\lambda x})^T. \tag{22}$$

Using the transpose property $(AB)^T = B^TA^T$ for matrices $A, B \in \mathbb{C}^{m \times m}$ and noting that the transpose of a scalar is itself, we obtain

$$\bar{x}^T \bar{A}^T = \bar{\lambda} \bar{x}^T. \tag{23}$$

Since A is Hermitian, $\bar{A}^T = A$, so

$$\bar{x}^T A = \bar{\lambda} \bar{x}^T. \tag{24}$$

Multiplying both sides by x on the right,

$$\bar{\lambda}\bar{x}^T x = \bar{x}^T A x \tag{25}$$

$$= \bar{x}^T \lambda x \tag{26}$$

$$= \lambda \bar{x}^T x. \tag{27}$$

Subtracting $\lambda \bar{x}^T x$ on both sides and factoring out $\bar{x}^T x$ we have,

$$(\bar{\lambda} - \lambda)\bar{x}^T x = 0. \tag{28}$$

Since $\bar{x}^T x$ is positive, it follows that $\bar{\lambda} - \lambda = 0$, implying $\lambda = \bar{\lambda}$. Thus, all eigenvalues of A are real.

2.2 Part b

Let $x, y \neq 0$ be eigenvectors of the Hermitian matrix A corresponding to distinct eigenvalues $\lambda, \mu \in \mathbb{R}$, respectively. Then, the eigenvectors x and y are orthogonal.

Proof. Since A is Hermitian, it satisfies the inner product property:

$$\langle Ax, y \rangle = \langle x, A^*y \rangle,$$
 (29)

where A^* denotes the conjugate transpose of A.

In order to prove (29) we will substitute the definition of the inner product:

$$\langle Ax, y \rangle = \langle x, A^*y \rangle, \tag{30}$$

$$(Ax)^T y = x^T A^* y, (31)$$

$$x^T A y = x^T A^* y. (32)$$

Since A is Hermitian, we have $A^* = A$, which confirms the equality. Continuing from (29), we conclude:

$$\langle Ax, y \rangle = \langle x, Ay \rangle. \tag{33}$$

Using the eigenvector relations $Ax = \lambda x$ and $Ay = \mu y$, it follows that:

$$\langle \lambda x, y \rangle = \langle x, \mu y \rangle, \tag{34}$$

$$\lambda \langle x, y \rangle = \mu \langle x, y \rangle. \tag{35}$$

Rearranging, we obtain:

$$(\lambda - \mu)\langle x, y \rangle = 0. \tag{36}$$

Since $\lambda \neq \mu$, it must be that $\langle x, y \rangle = 0$, proving that x and y are orthogonal.

3 Question 3

Let $Q \in \mathbb{C}^{m \times m}$ be unitary.

3.1 Part a

If λ is an eigenvalue and x its corresponding eigenvector of the matrix Q, then λ satisfies $|\lambda| = 1$.

Proof. Let $Q \in \mathbb{C}^{m \times m}$ be unitary. Recall that $QQ^* = Q^*Q = I$, where I is the identity matrix. The 2-norm of a matrix, also known as the spectral norm, is defined as $\|A\|_2 = \sqrt{\max \sigma(A^*A)}$, which corresponds to the maximum singular value of the matrix A.

It is immediately apparent that the 2-norm of the matrix Q is equal to the eigenvalues of the identity matrix.

$$\|Q\|_2 = \sqrt{\max \sigma(Q^*Q)} \tag{37}$$

$$= \sqrt{\max \sigma(I)} \tag{38}$$

$$=1\tag{39}$$

Thus, the following holds:

$$||Qx||_2 = ||x||_2 \tag{40}$$

Let $\lambda \neq 0$ be an eigenvalue and x be an eigenvector of the unitary matrix Q such that $Qx = \lambda x$. Taking the 2-norm of both sides, we obtain:

$$||Qx||_2 = ||\lambda x||_2 \tag{41}$$

$$||x||_2 = ||\lambda x||_2 \tag{42}$$

$$||x||_2 = |\lambda| ||x||_2 \tag{43}$$

$$|\lambda| = 1 \tag{44}$$

 $||Q||_F = \sqrt{m}.$

3.2 Part b

Proof. Recall that the Frobenius norm is defined as $||A||_F = \sqrt{\operatorname{trace}(A^*A)}$. Also, recall that the trace of a matrix A is the sum of its eigenvalues.

It follows that:

$$\|Q\|_F = \sqrt{\operatorname{trace}(Q^*Q)} \tag{45}$$

$$= \sqrt{\operatorname{trace}(I_m)} \tag{46}$$

$$=\sqrt{m}\tag{47}$$

where I_m is the identity matrix of size $m \times m$ (the same size as Q).

4 Question 4

Let $A \in \mathbb{C}^{m \times m}$ be skew-Hermitian, i.e., $A^* = -A$.

4.1 Part a

The eigenvalues of A are purely imaginary.

Proof. Let λ be an eigenvalue and v be an eigenvector of the skew-Hermitian matrix A, such that $Av = \lambda v$.

$$Av = \lambda v \tag{48}$$

$$v^* A v = v^* \lambda v \tag{49}$$

$$= v^* A v \tag{50}$$

$$= (A^*v)^* v \tag{51}$$

$$= (-Av)^* v \tag{52}$$

$$= (-\lambda v)^* v \tag{53}$$

$$=v^*\left(-\bar{\lambda}\right)v\tag{54}$$

$$-\bar{\lambda}v^*v = \lambda v^*v \tag{55}$$

Consider the complex number $\lambda = \alpha + \beta i$ and note that $-\bar{\lambda} = -\alpha + \beta i$.

For the equality $-\bar{\lambda} = \lambda$ to hold,

$$-\alpha + \beta i = \alpha + \beta i \tag{56}$$

The real part of the complex number λ must be zero. Therefore, the eigenvalues of A are purely imaginary.

4.2 Part b

The matrix I - A is nonsingular.

Proof. Let λ be an eigenvalue and v be an eigenvector of the matrix A, which is skew-Hermitian, such that $Av = \lambda v$.

Consider the matrix (I - A) acting on the eigenvector v:

$$(I - A)v = Iv - Av (57)$$

$$= v - \lambda v \tag{58}$$

$$= (1 - \lambda) v. \tag{59}$$

Hence, the eigenvalues of the matrix (I - A) are $(1 - \lambda)$.

Recall the definition of a nonsingular matrix: for any square matrix B, if $det(B) \neq 0$, then B is said to be nonsingular. Additionally, recall that the determinant of any square matrix B is the product of its eigenvalues.

It follows that

$$\det(I - A) \neq 0,\tag{60}$$

because in Part a, we have already shown that the matrix A has eigenvalues with zero real part. Thus, the product of all eigenvalues of the matrix (I - A), which we have shown to be $(1 - \lambda) \forall \lambda \in \sigma(A)$, is guaranteed to be nonzero.

5 Question 5

We say that $A, B \in \mathbb{C}^{m \times m}$ are unitarily equivalent if $A = QBQ^*$ for some unitary matrix $Q \in \mathbb{C}^{m \times m}$.

5.1 Part a

If A and B are unitarily equivalent, then they have the same singular values.

Proof. Let $A, B \in \mathbb{C}^{m \times m}$ be unitarily equivalent.

We begin by proving that if A and B are unitarily equivalent, then they have the same eigenvalues. We will then use this result to show that they have the same singular values.

$$A = QBQ^* (61)$$

$$A - \lambda I = QBQ^* - \lambda I \tag{62}$$

$$= Q\left(B - \lambda I\right)Q^* \tag{63}$$

$$\det(A - \lambda I) = \det(Q(B - \lambda I)Q^*)$$
(64)

$$= \det(Q) \cdot \det(B - \lambda I) \cdot \det(Q^*)$$
(65)

Recall from Question 3, Part (a), that we proved if Q is unitary, then $|\lambda|=1$ for all $\lambda \in \sigma(Q)$. Thus, the determinant of Q is the product of these eigenvalues, which equals 1.

Therefore, we obtain the equality

$$\det(A - \lambda I) = \det(B - \lambda I),\tag{66}$$

implying that the eigenvalues of two unitarily equivalent matrices are the same.

Now consider the matrix A^*A :

$$A^*A = (QBQ^*)^* QBQ^* \tag{67}$$

$$= QB^*Q^*QBQ^* \tag{68}$$

$$= QB^*BQ^*. (69)$$

Therefore, the matrix A^*A is unitarily similar to the matrix B^*B . Since the eigenvalues of the matrix A^*A are the singular values of A, and the eigenvalues of B^*B are the singular values of B, it follows that the singular values of A and B are equal.

5.2 Part b

If A and B have the same singular values, then they are **not** necessarily unitarily equivalent.

Proof. Let

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \tag{70}$$

We will show that the eigenvalues of the self-adjoint matrices of A and B are the same, which implies that the singular values are also the same.

$$A^*A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad B^*B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \tag{71}$$

Both self-adoint matrices have eigenvalues $\lambda_{1,2} = \{0,1\}$. Hence the singular values of the matrix A and B are the same.

Recall the definition of a Normal Matrix. A square matrix $M \in \mathbb{C}^{m \times m}$ is said to be normal, if it commutes with its conjugate transpose M^* . That is,

$$M \text{ is normal } \iff M^*M = MM^*.$$
 (72)

An equivalent definition is that M is diagonalizable by a unitary matrix.

Let us assume for the sake of contradiction that the matrices A and B are unitarily equivalent. Then there exists a unitary matrix $Q \in \mathbb{C}^{2\times 2}$ such that

$$B = Q^* A Q \tag{73}$$

for unitary equivalence to hold, it is a requirement that if the matrix A is a normal matrix, then the matrix B must also be normal.

$$AA^* = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = A^*A = A \tag{74}$$

Hence, the matrix A is normal.

$$BB^* = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = B^*B \tag{75}$$

We have arrived at a contradiction. The matrix A is normal, but B is not. Therefore even though both matrices have the same singular values, they are not unitarily equivalent.

6 Question 6

Find the relative condition number of the following functions and discuss if there is any concern of being ill-conditioned. If so, discuss when.

6.1 Part a

$$f_1(x_1, x_2) = x_1 + x_2 (76)$$

For the continuously differentiable function f_1 , the relative condition number κ can be calculated by the following formula:

$$\kappa(x_1, x_2) = \frac{\|J\|_{\infty} \|x\|_{\infty}}{|f_1(x_1, x_2)|} \tag{77}$$

Let us find the Jacobian of the matrix, and its max norm (since the Jacobian matrix in this case is a row vector it is simply the sum of the absolute value of all the entries).

$$J = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} 1 & 1 \end{bmatrix} \tag{78}$$

$$||J||_{\infty} = 2 \tag{79}$$

Thus the condition number is

$$\kappa(x_1, x_2) = \frac{2 \max\{|x_1|, |x_2|\}}{|x_1 + x_2|} \tag{80}$$

The function is ill-conditioned when $x_1 \to -x_2$. Interestingly though the function is well-conditioned if you restrict x_1 and x_2 to be strictly positive, or that $x_1, x_2 \in (0, \infty)$. Then as you can see in the following plot, at no point does the κ diverge.

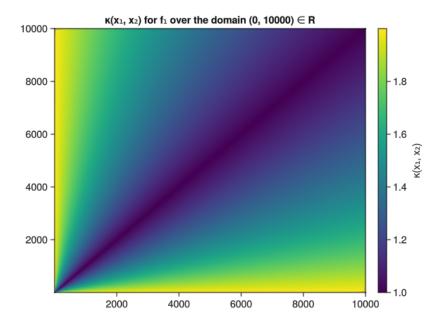


Figure 1: Plot of the condition number $\kappa(x_1,x_2)$ for f_1 over a strictly positive domain $x_1,x_2\in (0,10000).$

6.2Part b

$$f_2(x_1, x_2) = x_1 x_2 \tag{81}$$

The Jocabian matrix of the function f_2 is

$$J = \begin{bmatrix} \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix}$$

$$= \begin{bmatrix} x_2 & x_1 \end{bmatrix}$$
(82)

$$= \begin{bmatrix} x_2 & x_1 \end{bmatrix} \tag{83}$$

The max-norm of the Jacobian is the sum of the absolute values of the entries. Hence,

$$||J||_{\infty} = |x_2| + |x_1| \tag{84}$$

Therefore $\exists x_1,x_2 \in \mathbb{R} : \kappa$ the condition number is

$$\kappa(x_1,x_2) = \frac{(|x_2| + |x_1|) \max{(|x_1|,|x_2|)}}{|x_1x_2|} \tag{85}$$

The equation is well-conditioned as long as both $x_1 \ll 1$ and $x_2 \ll 1$ at the same time. Another way to say this is when $x_1 \cdot x_2 \to 0$ the function f_2 approaches a regiene of being ill-conditioned. This can be clearly see in the following heatmap of the log of the condition number with $x_1, x_2 \in (-1, 1)$.

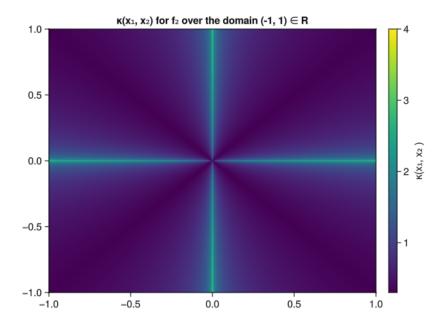


Figure 2: Plot of the condition number $\kappa(x_1,x_2)$ for f_2 over $x_1,x_2\in(-1,1)$. The colorbar is the $\log_{10}(\kappa)$

Part c 6.3

$$f_3(x) = (x-2)^9 (86)$$

For a continuously differentiable function f of a single variable. The condition number κ is

$$\kappa(x) = \left| \frac{xf'(x)}{f(x)} \right| \tag{87}$$

thus for the function f_3 we have that the condition number is:

$$\kappa(x) = \left| \frac{9x(x-2)^8}{(x-2)^9} \right|$$

$$= \left| \frac{9x}{x-2} \right|$$
(88)

$$= \left| \frac{9x}{x-2} \right| \tag{89}$$

the function is ill-conditioned as $x \to 2$ and approaches 0 as $x \to 0$ and is well-conditioned everywhere else.

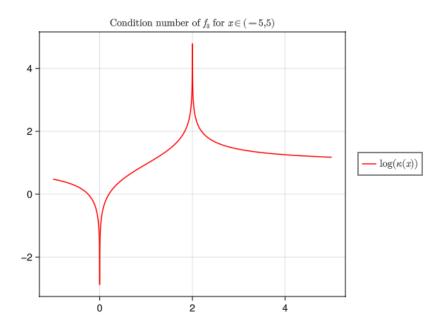


Figure 3: Line plot of the log transform of the condition number $\kappa(x)$ for the function f_3 for $x \in (-1,5)$

Additionally, the condition number approaches the value 10 as $|x| \to \infty$.

7 Question 7

Note that the function $f(x)=(x-2)^9$ in Part c of Question 6 can also be expressed as $g(x)=x^9-18x^8+144x^7-672x^6+2016x^5-4032x^4+5376x^3-4608x^2+2304x-512$. Mathematically, the two functions f and g are identical.

7.1 Part a, b

Plot f(x) by evaluating its values at the discrete points 1.920, 1.921, 1.922, ..., 2.080, which are equally spaced with a distance of 0.001. Then, over-plot g(x) at the same set of discrete points.

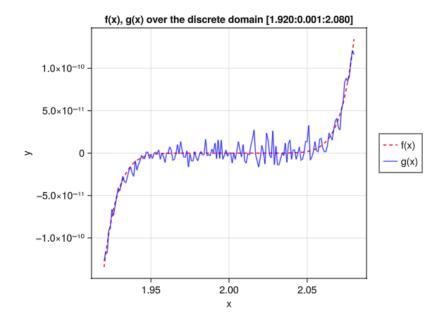


Figure 4: Line plot of the function f(x) in red and g(x) in blue over the discrete domain $\{1.920, 1.921, \dots, 2.080\}$

7.2 Part c

Draw your conclusions from the results of Part c in Question 6 and Parts a and b in this problem.

7.2.1 Solution

Let us analyze the problem by considering the number of operations each function undergoes. For the function $f(x) = (x-2)^9$

$$(\times, \div): 8$$

 $(+, -): 1$

For the function $g(x) = x^9 - 18x^8 + 144x^7 - 672x^6 + 2016x^5 - 4032x^4 + 5376x^3 - 4608x^2 + 2304x - 512$

Here we are assuming that once x^9 is calculated the intermediate values are cached.

$$(\times, \div): 16$$

 $(+,-): 9$

The function f(x) is ill-conditioned in the domain $x \in \{1.920, ..., 2.080\}$ because (as shown in Question 6 Part a) the difference inside of the parenthesis has $x_1 \to 2$ when $x_2 = 2$, more over this new value which is close to zero is being multiplied with itself many times and (as we have shown in Question 6 part b) multiplication is ill-conditioned when $x_1 \cdot x_2 \to 0$.

For the function g(x) we are less concerned with multiplication because x > 1 for every operation which is well-conditioned, but the increased number of subtractions/additions leads to an oscillatory result due to the increased risk of catastrophic cancellation and increased rounding errors.

Interestingly though, if we increase the size of the data type used to plot the two functions from Float64 \rightarrow Float128 they are identical as seen in the figure below.

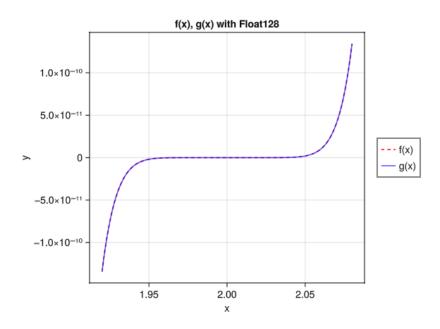


Figure 5: The same plot as Figure 4, but with quadruple precision.

8 Question 8

Theorem 1 (Hermitian Matrix). If $A \in \mathbb{C}^{m \times m}$ is Hermitian, then A has real eigenvalues λ_i , where i = 1, ..., m, which are not necessarily distinct, and the m corresponding eigenvectors u_i form an orthonormal basis for \mathbb{C}^m .

Definition 1 (Positive Definite Matrix). A matrix A is called positive definite if and only if $\langle Ax, x \rangle > 0$ for all $x \neq 0 \in \mathbb{C}^m$.

Theorem 2 (Principal Axis Theorem). For any Hermitian matrix $A \in \mathbb{C}^{m \times m}$, there exists a unitary matrix U and a diagonal matrix D such that $A = UDU^*$, where D contains the eigenvalues λ_i of A along the diagonal entries.

Suppose A is Hermitian.

A is positive definite if and only if $\lambda_i > 0$ for all $\lambda_i \in \Lambda(A)$, the spectrum of A.

Proof. Assume that A is a positive definite hermitian matrix and has an eigenvalue λ such that $Ax = \lambda x$, where x is the eigenvector associated with the eigenvalue λ .

First, let us show that if A is positive definite, then all eigenvalues of A are strictly positive.

Case 1: $\lambda = 0$. The matrix A is not positive definite because this would mean there exists an eigenvector x such that

$$Ax = 0, (90)$$

$$x^*Ax = 0. (91)$$

By Definition 1, for A to be positive definite, we must have $x^*Ax > 0$ for all $x \neq 0$. This contradiction implies that $\lambda \neq 0$.

Case 2: $\lambda < 0$. Since $x \neq 0$, we have

$$Ax = \lambda x, \tag{92}$$

$$x^*Ax = x^*\lambda x,\tag{93}$$

$$= \lambda x^* x, \tag{94}$$

$$=\lambda \|x\|_2^2. \tag{95}$$

Since the inner product x^*x is a strictly positive real number and $\lambda < 0$, it follows that $x^*Ax < 0$, contradicting Definition 1.

Thus, the only way for a matrix A to be positive definite is if all of its eigenvalues are strictly positive.

Second, let us show that if all eigenvalues of a matrix A are strictly positive, then the matrix is positive definite.

Given that A is Hermitian, by Theorem 1, all the eigenvalues of A are real (proved in question 2), and there exists an orthonormal basis of eigenvectors

$$U:=\{u_i\in\mathbb{C}^m|i=1,2,\cdots,m\}.$$

Since all eigenvalues λ_i are positive, by Theorem 2, the diagonal matrix D has positive entries.

In order to show that A is positive definite, we need to verify that $\langle Ax, x \rangle > 0$ for all non-zero $x \in \mathbb{C}^m$.

Consider any vector $x \neq 0 \in \mathbb{C}^m$. We can write x in terms of the orthonormal eigenbasis U of A as:

$$x = \sum_{i=1}^{m} \alpha_i u_i \tag{96}$$

where $\alpha_i = \langle x, u_i \rangle$ and similarly $\bar{\alpha_i} = \langle u_i, x \rangle.$

Let us now compute the inner product $\langle Ax, x \rangle$

$$\langle Ax, x \rangle = \left\langle A \sum_{i=1}^{m} \alpha_i u_i, \sum_{j=1}^{m} \alpha_j u_j \right\rangle \tag{97}$$

$$= \left\langle \sum_{i=1}^{m} \alpha_i A u_i, \sum_{j=1}^{m} \alpha_i u_j \right\rangle \tag{98}$$

$$= \left\langle \sum_{i=1}^{m} \alpha_i \lambda_i u_i, \sum_{j=1}^{m} \alpha_j u_j \right\rangle \tag{99}$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{m} \alpha_i \lambda_i \bar{\alpha_j} \left\langle u_i, u_j \right\rangle \tag{100}$$

Since we know that the eigenvectors in U form an orthonormal basis and are all orthogonal to each other,

$$\sum_{i=1}^{m} \sum_{j=1}^{m} \langle u_i, u_j \rangle = \delta_{ij} \tag{101}$$

Hence, the only non-zero contributions to the sum are when i=j and we can re-write (101) as:

$$\langle Ax, x \rangle = \sum_{i=1}^{m} \lambda_i \|\alpha_i\|_2^2 > 0 \tag{102}$$

Therefore, assuming that A is Hermitian, A is positive definite if and only if $\lambda_i > 0$ for all $\lambda_i \in \Lambda(A)$, the spectrum of A.