UNIT 5 POSITIVE DEFINITENESS

Structure **Page No.**

5.1 INTRODUCTION

In Unit 4, you studied about normal operators, those operators on inner product spaces which commute with their adjoints. To help study such operators, we look at all their eigenvalues, and re-present the operators as a sum of projections on their eigenspaces. This is the spectral decomposition theorem, as you will see in Sec. 5.2. This, in fact, gives us a way of diagonalising these operators.

The next section, Sec. 5.3, is about those self-adjoint operators which have all their eigenvalues positive. Such operators are called positive definite operators, and have applications in diverse areas, some of which you will see in the course MMT-008, 'Probability and Statistics'.

Objectives

After reading this unit, you should be able to

- write the spectral resolution for normal operators and matrices;
- determine when a given operator/matrix is positive definite or positive semi-definite;
- use some basic properties of positive definite (semi-definite) operators/matrices.

5.2 SPECTRAL DECOMPOSITION

In this section we shall prove, and apply, the spectral decomposition theorem for normal operators. Recall that 'spectral' refers to the spectrum of a linear operator/matrix, which is its set of eigenvalues.

We first prove a few other results pertaining to normal operators. To start with, we shall use the Primary Decomposition Theorem (Theorem 3, Unit 2) to prove the following result.

Theorem 1: Let V be a finite-dimensional inner product space and let T be a triangulisable linear operator on V. Then. T is normal if and only if V has an orthonormal basis consisting of eigenvectors of T.

Proof: First let us assume that T is normal. Let the minimal polynomial of T be $(t-\lambda_1)^{m_1}\ldots(t-\lambda_k)^{m_k}$, where $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of T. By the primary decomposition theorem,

$V = \text{Ker}(T-\lambda_1)^{m_1} \oplus \ldots \oplus \text{Ker}(T-\lambda_k)^{m_k}$.

Since T is normal, so is $(T-\lambda_i)$ for each $i = 1, \ldots k$. Therefore, by Theorem 5, Unit 4, $\text{Ker}(T-\lambda_i)^{m_i} = \text{Ker}(T-\lambda_i)$. Therefore,

$V = \text{Ker}(T-\lambda_1)\oplus\ldots\oplus\text{Ker}(T-\lambda_k)$.

Let $\mathcal{B}_i$ be an orthonormal basis of $\text{Ker}(T-\lambda_i)$. Then, it follows from Theorem 5, Unit 4, that $\mathcal{B}=\mathcal{B}_1\cup\ldots\cup\mathcal{B}_k$ is an orthonormal basis of V consisting of eigenvectors of T.

Conversely, let $\mathcal{B}=\{x_1, \ldots, x_n\}$ be an orthonormal basis of V consisting of eigenvectors of T. Then, for each i, $Tx_i=\mu_i x_i$. (Note that $\mu_1, \ldots, \mu_n$ need not be distinct.) Now, take $v\in V$. Then $v=\sum_{i=1}^{n}\alpha_i x_i$ for $\alpha_1, \ldots, \alpha_n \in \mathbf{C}$.

So $(T^*T)v = T^*\left(\sum_{i=1}^{n}\alpha_i Tx_i\right) = T^*\left(\sum_{i=1}^{n}\alpha_i \mu_i x_i\right) = \sum_{i=1}^{n}\alpha_i\mu_i\bar{\mu}_i x_i = (TT^*)v$.

Hence $TT^* = T^*T$, that is, T is normal. ■

What this theorem tells us is that "if T is a normal operator on V, then V has an orthonormal basis with respect to which the matrix of T is diagonal".
Let us, now, use this result for proving the main theorem of this section.

Theorem 2 (Spectral Decomposition Theorem): Let T be a normal operator on a finite-dimensional inner product space V over **C**, with eigenvalues $\lambda_1, \ldots, \lambda_k$ and corresponding eigenspaces $W_1, \ldots, W_k$. Then $T=\lambda_1P_1+\ldots+\lambda_kP_k$, where P_i is the orthogonal projection of T on $W_i \forall i = 1, \ldots, k$.

Proof: Since T is normal, it is diagonalisable. Therefore, the minimal polynomial of T is $m(t)=(t-\lambda_1)\ldots(t-\lambda_k)$.

Let $\frac{1}{m(t)}=\frac{a_1}{(t-\lambda_1)}+\cdots+\frac{a_k}{(t-\lambda_k)}$, as a sum of partial fractions, where $a_1, \ldots, a_k \in \mathbf{C}$.

Then $1 = a_1m_1(t)+\cdots+a_km_k(t)$, ...(1)

where $m_i(t)=(t-\lambda_1)\ldots(t-\lambda_{i-1})\,(t-\lambda_{i+1})\ldots(t-\lambda_k),\ i=1, \ldots, k$.

Now, let $P_i = a_im_i(T)$ for $i=1, \ldots, k$. Then each P_i is a linear operator on V. Since each P_i is a polynomial in T, P_i and P_j commute. Now, from (1) it follows that $I=P_1+\cdots+P_k$.

Also, for $i\neq j$, $P_iP_j=a_im_i(T)a_jm_j(T)=a_ia_jm_i(T)m_j(T)=\mathbf{0}$, since $m(t)$ divides $m_i(t)m_j(t)$. It follows that $P_i=P_i(P_1+\cdots+P_k)=P_i^2$, that is, P_i **is an orthogonal projection of** V **onto** $\mathbf{Ker(T-\lambda_iI)}$.

(Remember that $V=\text{Ker}(T-\lambda_1I)\oplus\ldots\oplus\text{Ker}(T-\lambda_kI)$.)

$TP_i=(T-\lambda_iI)P_i+\lambda_iP_i=\lambda_iP_i$, since $(T-\lambda_iI)P_i=a_im(T)=\mathbf{0}$. Therefore,

$T=TI=T(P_1+\cdots+P_k)=\lambda_1P_1+\cdots+\lambda_kP_k$.

Hence we have shown that if T is unitarily diagonalisable, then $T=\lambda_1P_1+\cdots+\lambda_kP_k$, where $P_1, \ldots, P_k$ are linear operators on V such that $P_1+\cdots+P_k=I$, $P_i^2=P_i$ and $P_iP_j=0$. ■

This theorem is the basis of spectral theory, initiated by the great German mathematician David Hilbert. It is really a generalisation of finding the principal axes of an ellipsoid.

Let us also look at the matrix version of this theorem.

Theorem 3 (Spectral decomposition): Let $A \in \mathbf{M}_n(\mathbf{C})$ be normal, with eigenvalues $\lambda_1, \ldots, \lambda_n$ and an orthonormal basis of eigenvectors $\{v_1, \ldots, v_n\}$. Then

$$A = \sum_{i=1}^{n} \lambda_i v_i v_i^*.$$

Proof: Let us write E_{ij} for an $n \times n$ matrix whose (i, j)-th entry is 1 and all other entries are 0. You should verify that $E_{ij} = e_i e_j^t$.

Let A be an $n \times n$ normal matrix and U be a unitary matrix such that $U^* A U = \text{diag}(\lambda_1, \ldots, \lambda_n)$. Let the jth column of U be u_j. Then

$$\begin{aligned}
A &= U\text{diag}(\lambda_1, \ldots, \lambda_n)U^* \\
&= U(\lambda_1 E_{11} + \cdots + \lambda_n E_{nn})U^* \\
&= \lambda_1 U E_{11} U^* + \cdots + \lambda_n U E_{nn} U^* \\
&= \lambda_1 U e_1 e_1^t U^* + \cdots + \lambda_n U e_n e_n^t U^* \\
&= \lambda_1 (Ue_1)(Ue_1)^* + \cdots + \lambda_n (Ue_n)(Ue_n)^* \\
&= \lambda_1 P_1 + \cdots + \lambda_n P_n, \text{ where } P_i = (Ue_i)(Ue_i)^* \qquad \ldots(2)
\end{aligned}$$

The matrices $P_1, \ldots, P_n$ are the projection matrices, and satisfy the following properties:

i) For each i, $P_i^2 = P_i$, and for $i \neq j$, $P_i P_j = \mathbf{0}$.

$$\left(\text{because } P_i P_j = (Ue_i)(Ue_i)^*(Ue_j)(Ue_j)^* = UE_{ii}E_{jj}U^* = \begin{cases} \mathbf{0} & \text{if } i \neq j \\ P_i & \text{if } i = j \end{cases}.\right)$$

ii) $P_1 + \cdots + P_n = I_n$.

$$\begin{aligned}
\Big[P_1 + \cdots + P_n &= (Ue_1)(Ue_1^*) + \cdots + (Ue_n)(Ue_n)^* \\
&= U(e_1 e_1^*)U^* + \cdots + U(e_n e_n^*)U^* \\
&= UE_{11}U^* + \cdots + UE_{nn}U^* \\
&= UI_nU^* = UU^* = I_n.\Big]
\end{aligned}$$

Thus, (2) is the spectral decomposition of A. ■

Let us see how a spectral decomposition is actually obtained.

Example 1: Write the spectral decomposition of the self-adjoint operator on $\mathbf{C}^3$ whose matrix representation with respect to the standard basis is given by

$$A = \begin{bmatrix} 4 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 4 \end{bmatrix}.$$

Solution: The eigenvalues of A are 3, 4, 6. An orthonormal basis of eigenvectors is

$$\left\{\frac{1}{\sqrt{3}}\begin{bmatrix}1\\1\\1\end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\\-1\end{bmatrix}, \frac{1}{\sqrt{6}}\begin{bmatrix}1\\-2\\1\end{bmatrix}\right\}.$$

$$\text{So } A = \frac{3}{3}\begin{bmatrix}1\\1\\1\end{bmatrix}\begin{bmatrix}1 & 1 & 1\end{bmatrix} + \frac{4}{2}\begin{bmatrix}1\\0\\-1\end{bmatrix}\begin{bmatrix}1 & 0 & -1\end{bmatrix} + \frac{6}{6}\begin{bmatrix}1\\-2\\1\end{bmatrix}\begin{bmatrix}1 & -2 & 1\end{bmatrix}$$

So, the spectral decomposition is

$$A = 3P_1 + 4P_2 + 6P_3, \text{ where } P_1 = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \ P_2 = \frac{1}{2}\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix},$$

$$P_3 = \frac{1}{6}\begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix}$$

You should verify that $P_1 + P_2 + P_3 = I_3$.

Example 2: Write the spectral decomposition of the operator $T : \mathbf{R}^3 \to \mathbf{R}^3 : T(x, y, z) = (6x - 2y - z, -2x + 6y - z, -x - y + 5z)$.

Solution: The matrix of T is $\begin{bmatrix} 6 & -2 & -1 \\ -2 & 6 & -1 \\ -1 & -1 & 5 \end{bmatrix}$. The eigenvalues of T are 3, 6, 8.

So, the minimum polynomial of T is $m(t) = (t-3)(t-6)(t-8)$.

Now, $\dfrac{1}{m(t)} = \dfrac{1}{15(t-3)} - \dfrac{1}{6(t-6)} + \dfrac{1}{10(t-8)}$.

So $1 = \dfrac{1}{15}(t-6)(t-8) - \dfrac{1}{6}(t-3)(t-8) + \dfrac{1}{10}(t-3)(t-6)$.

Take $P_1 = \dfrac{1}{15}(A - 6I)(A - 8I) = \dfrac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$,

$P_2 = -\dfrac{1}{6}(A - 3I)(A - 8I) = \dfrac{1}{6}\begin{bmatrix} 1 & 1 & -2 \\ 1 & 1 & -2 \\ -2 & -2 & 4 \end{bmatrix}$,

$P_3 = \dfrac{1}{10}(A - 3I)(A - 6I) = \dfrac{1}{2}\begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

Then $T = 3P_1 + 6P_2 + 8P_3$ is the spectral decomposition.

Example 3: Write the spectral decomposition for $A = \begin{bmatrix} 2 & 0 & -1 \\ 0 & 3 & 0 \\ -1 & 0 & 2 \end{bmatrix}$.

Solution: Let $U = \begin{bmatrix} 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 1 & 0 & 0 \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$. Then $U^*AU = \text{diag}(3, 1, 3)$.

Now $u_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $u_2 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}$ and $u_3 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix}$.

Thus, $P_1 = u_1u_1^* = E_{22}$, $P_2 = u_2u_2^* = \dfrac{1}{2}\begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}$ and $P_3 = u_3u_3^* = \dfrac{1}{2}\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix}$.

Hence, the spectral decomposition of $A = 3P_1 + P_2 + 3P_3$.

Now try some exercises.

E1) Write the spectral decomposition of the linear operator on C^3 whose matrix decomposition with respect to the standard basis is given by $A = \begin{bmatrix} 5 & 1 & 1 \\ 1 & 1 & 3 \\ 1 & 3 & 1 \end{bmatrix}$.

E2) Construct a spectral decomposition of $\begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$.

E3) Write the spectral decomposition for $\begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}$.

We now look at those linear operators (and matrices) that play a role in functional analysis which is somewhat analogous to the role played by positive real numbers in classical analysis.

5.3 POSITIVE DEFINITE OPERATORS

Recall that a linear operator T on an inner product space is called **self-adjoint** or **hermitian** if $T^* = T$. Now, if T is hermitian, and λ is an eigenvalue of T with $Tx = \lambda x,\ x \neq \dot{0}$, then

$\lambda < x, x > = < Tx, x > = < x, T^* x > = < x, Tx > = \bar{\lambda} < x, x >$.

$(\lambda - \bar{\lambda}) < x, x > = 0 \Rightarrow \lambda = \bar{\lambda} \Rightarrow \lambda$ is real.

So, a hermitian linear operator has real eigenvalues only.

We now look at those self-adjoint operators which have **positive eigenvalues or non-negative eigenvalues**. You would be using such operators while doing multivariate analysis and in many mathematical models. Let us define them formally.

Definition: A self-adjoint operator T on a vector space V is called

i) **positive definite** if $< Tx, x > > 0$ for all $x \neq 0$.

ii) **positive semi-definite** if $< Tx, x > \geq 0\ \forall x \in V$.

iii) **negative definite** if $< Tx, x > < 0 \forall x \neq 0$.

iv) **indefinite** if T is neither positive nor negative semi-definite.

Can you think of some obvious examples? The identity operator is positive definite, and the zero operator is positive semi-definite. –I is negative definite. You can think of other examples, and you will see more as you go through this section.

As always, we have analogous definitions for positive (or negative) definite and semi-definite matrices. Can you guess what they are? Here's an exercise about this.

E4) Define positive definite and positive semi-definite matrices. Also give one example of a positive semi-definite matrix, which is not positive definite, with justification.

Positive definite matrices have considerable theoretical and computational applications. They are used, for instance, in optimisation algorithms and linear regression models.

Notation: We denote ' T is a positive definite (semi-definite) operator' by $\mathbf{T > 0\ (T \geq 0)}$. Similar notation is used for positive definite/semi-definite matrices.

Let us now consider some properties of positive definite and positive semi-definite operators and matrices.

Theorem 4: Let T be a self-adjoint operator on a finite-dimensional inner product space V. The following statements are equivalent.

i) T is positive definite.

ii) All eigenvalues of T are positive.

iii) There is a unique positive definite operator S such that $T = S^2$. (This operator S is called **the square root of T**.)

iv) There is an invertible linear operator R such that $T = R^*R$.

Proof: (i)⇒(ii): If λ is an eigenvalue of T, then you already know that λ is a real number. Now, if x is an eigenvector corresponding to λ, then $< Tx, x > > 0 \Rightarrow \lambda < x, x > > 0 \Rightarrow \lambda > 0$.

(ii)⇒(i): Since all the eigenvalues of T are real and T is self-adjoint, by Theorem 2, Unit 4, it follows that T is diagonalisable. So, V has an orthonormal basis consisting of eigenvectors of T. Let $\mathbb{B} = \{x_1, \ldots, x_n\}$ be such a basis. Then $Tx_i = \mu_i x_i$, where $\mu_i > 0$. (Note that the μ_is need not be distinct.) For $x \neq 0$, write $x = \sum_{i=1}^{n} \alpha_i x_i$, α_i being scalars and not all $\alpha_i = 0$.

Then, $< Tx, x > = \left\langle \sum_{i=1}^{n} \alpha_i Tx_i, \sum_{i=1}^{n} \alpha_i x_i \right\rangle = \sum_{j=1}^{n}\sum_{i=1}^{n} \mu_i \alpha_i \bar{\alpha}_j < x_i, x_j > = \sum_{i=1}^{n} \mu_i \,|\, \alpha_i \,|^2$, since $\{x_i\}_i$ is an orthonormal set.

Hence $< Tx, x > > 0 \,\forall\, x \neq 0$.

(ii)⇒(iii): If $T = \lambda_1 P_1 + \cdots + \lambda_k P_k$ is the spectral resolution of T, then $\lambda_i > 0$ and $S = \sqrt{\lambda_1} P_1 + \cdots + \sqrt{\lambda_k} P_k$ is a self-adjoint operator with all its eigenvalues positive. So, by the equivalence of (i) and (ii), it follows that S is a positive definite operator. Also, $S^2 = T$.

The uniqueness of S requires some more background. So, we will not be proving it here.

(iii)⇒(iv): Take $R = S$. Then R is invertible and $T = R^*R$.

(iv)⇒(i): Let $T = R^*R$. Then T is self-adjoint. Also, for any $x \neq 0$, $< Tx, x > = < R^*Rx, x > = < Rx, Rx > = \| Rx \|^2 > 0$.

Hence T is positive definite. ■

A similar characterisation holds for positive semi-definite operators, which we state below. The proof is left as an exercise.

Theorem 5: Let T be a self-adjoint operator on a finite-dimensional inner product space V. The following statements are equivalent.

i) T is positive semi-definite.

ii) All eigenvalues of T are non-negative.

iii) There is a unique positive semi-definite operator S such that $T = S^2$. (This operator S is called the square root of T.)

iv) There is a linear operator R such that $T = R^*R$. ■

As before, you can reword the properties in Theorems 4 and 5 for positive definite and positive semi-definite matrices. The results are given below.

Theorem 6: Let $A \in \mathbf{M}_n(\mathbf{C})$ be a hermitian matrix. Then the following are equivalent:

i) A is positive definite (*positive semi-definite, respectively*).

ii) All the eigenvalues of A are strictly positive (*non-negative, respectively*).

iii) $A = S^2$ for a unique positive definite (*positive semi-definite, respectively*) matrix S.

iv) $\exists B \in \mathbf{M}_n(\mathbf{C})$ s.t. $A = B^*B$, where B is non-singular (*B need not be non-singular, respectively*).

The proof of Theorem 6 is along the same lines as the proof for Theorem 4. ■

Let us do a related example now.

Example 4: Find the square root of $A = \begin{bmatrix} 4 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 4 \end{bmatrix}$, given in Example 1, Unit 4.

Solution: The spectral resolution of A is $A = 3P_1 + 4P_2 + 6P_3$, where

$$P_1 = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, P_2 = \frac{1}{2}\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}, \text{ and } P_3 = \frac{1}{6}\begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix}.$$

Therefore, the square root of A is $\sqrt{A} = \sqrt{3}P_1 + 2P_2 + \sqrt{6}P_3$

$$= \frac{1}{\sqrt{6}}\begin{bmatrix} \sqrt{2}+\sqrt{6}+1 & \sqrt{2}-2 & \sqrt{2}-\sqrt{6}+1 \\ \sqrt{2}-2 & \sqrt{2}+4 & \sqrt{2}-2 \\ \sqrt{2}-\sqrt{6}+1 & \sqrt{2}-2 & \sqrt{2}+\sqrt{6}+1 \end{bmatrix}.$$

Try some exercises now.

E5) Prove Theorems 5 and 6.

E6) Find the square root of $\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$.

E7) Show that every positive definite operator/matrix is invertible.

E8) Is the inverse of a positive definite operator/matrix also positive definite? Give reasons for your answer.

Let us now consider another property that characterises positive definite matrices. This uses the concept of minors of a matrix. You would recall that if A is an $n \times n$ matrix, a **principal minor** of A of size k is the determinant of the $k \times k$ sub-matrix of A which is obtained from A by taking any k rows and **the same** k columns.

For example, if $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{bmatrix}$, a principal minor of size 2 is $\det\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} = -1$,

which is obtained by taking the first two rows and columns. Another principal minor of size 2 of A is $\det\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} = 8$, obtained by taking the 2^{nd} and 4^{th} rows as well as the 2^{nd} and 4^{th} columns of A.

If we take the minor formed by the first k rows and k columns of A, we get its **leading principal minor** of size k. For instance, $\det[1]$, $\det\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$ and

$\det\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 1 \end{bmatrix}$ are the leading principal minors of A of sizes 1, 2 and 3, respectively.

We now use the principal minors of positive definite matrices to give useful characterisations, **without proof**.

Theorem 7: Let A be a hermitian $n \times n$ matrix. Then

i) A is positive definite if and only if all its principal minors are positive.

ii) A is positive semi-definite if and only if all its principal minors are non-negative.

iii) A is positive definite if and only if all its leading principal minors are positive. ■

Using this result we can see that the matrix $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{bmatrix}$ is not positive definite,

because its leading principal minor of size 2, $\det\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} = -1$.

Note: It is **not true** that A is positive semi-definite if all its leading principal minors are non-negative. For instance, take $A = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}$. Then both its leading principal minors are 0, but the matrix is not positive semi-definite.

Let us now consider some useful properties of positive definite matrices.

This was proved by Andre-Louis Cholesky (1875-1918), a Ukrainian-French mathematician.

Theorem 8 (Cholesky decomposition): A is positive definite iff $A = BB^*$ for some non-singular lower triangular matrix B.

Proof: Firstly, if $A = BB^*$, by Theorem 6 (iv), A is positive definite. Conversely, if A is positive definite, then in the process of diagonalising A, we get $A = LDL^*$, where L is a lower triangular matrix and D is diagonal. Take $B = L\sqrt{D}$. Then B is non-singular and lower triangular, and $A = BB^*$. ■

The Cholesky decomposition is very useful for solving systems of linear equations, and is twice as efficient as the LU decomposition you have studied in your undergraduate Numerical Analysis course. It is commonly used in the Monte Carlo method for simulating systems with multiple correlated variables.

This is named after the French mathematician Jacques Hadamard (1865-1963).

Theorem 9 (Hadamard's Inequality): i) For a positive definite $n \times n$ matrix $A = [a_{ij}]$, $\det A \le a_{11} a_{22} \ldots a_{nn}$.

ii) For any $n \times n$ matrix $B = [b_{ij}]$, $|\det B|^2 \le \prod_{i=1}^{n} \left(\sum_{j=1}^{n} |b_{ij}|^2 \right)$.

Proof: i) We prove it by induction on n. If $n = 1$, then $\det A = a_{11}$, and the statement is true.

Assume that it is true for $n = m$. Now take A to be an $(m+1) \times (m+1)$ matrix.

Write $A = \begin{bmatrix} A_1 & x \\ x^t & a_{m+1,m+1} \end{bmatrix}$, where $x = \left[a_{m+1,1} \ldots a_{m+1,m} \right]^t$.

Note that A_1 is also positive definite, and so $\det A_1 \le a_{11} \ldots a_{mm}$.

Define $\alpha = \begin{bmatrix} I_m & 0 \\ -x^t A_1^{-1} & 1 \end{bmatrix}$. Then $\det \alpha = 1$.

Also $\alpha A = \begin{bmatrix} A_1 & x \\ 0^t & a_{m+1,m+1} - x^t A_1^{-1} x \end{bmatrix}$.

So $\det A = \det \alpha A = \det A_1 (a_{m+1,m+1} - x^t A_1^{-1} x)$

$\le a_{m+1,m+1} \det A_1$, since A_1^{-1} is also positive definte.

$= a_{11} a_{22} \ldots a_{m+1,m+1}$.

Hence the statement (i) is true for all $n \in \mathbf{N}$.

ii) If $\det B = 0$, the statement is trivially true.

If $\det B \ne 0$, $A = B^* B$ is positive definite. Now, applying (i) to this A, we get the result. ■

Remark: From your earlier studies, you know that $|\det B|$ represents the volume of an n-dimensional solid whose rows are given by the row vectors that form the rows of B. So, Hadamard's inequality tells us that this volume is less than or equal to the volume of the n-dimensional rectangular solid whose sides have the same lengths.

Some exercises now.

E9) For which values of $x \in \mathbf{C}$ will $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & x \\ 1 & \bar{x} & 1 \end{bmatrix}$ be positive semi-definite?

E10) Check if $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & -1 \end{bmatrix}$ is positive semi-definite. Is this matrix positive definite? Why?

E11) Show that $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ is positive semi-definite. Also find its square root.

E12) Let A be an $n \times n$ positive definite matrix. Let λ_0 and λ_1 be the minimum and the maximum eigenvalues of A. Show that for any $x \in \mathbf{C}^n$ such that $x^*x = 1, \lambda_0 \leq x^*Ax \leq \lambda_1$.
(**Hint:** Let U be unitary such that U^*AU is a diagonal matrix, diag $(\mu_1, \ldots, \mu_n)$. Write $y = Ux$. Then $y^*y = 1$ and $x^*Ax = \sum_{i=1}^{n} \mu_i |y_i|^2$.)

You will be studying some more about positive definite and semi-definite operators in MMT-008. Let us now end this unit with a look at the points covered in it.

5.4 SUMMARY

In this unit we have discussed the following points.

1. **Spectral Decomposition Theorem**:

i) If T is a normal operator on $\mathbf{C}^n$ with distinct eigenvalues $\lambda_1, \ldots, \lambda_k$ and corresponding eigenspaces $W_1, \ldots, W_k$, then $T = \sum_{i=1}^{k} \lambda_i P_i$, where P_i is the orthogonal projection of T on W_i.

ii) (Matrix analogue) If $A \in \mathbf{M}_n(\mathbf{C})$ is normal with eigenvalues $\lambda_1, \ldots, \lambda_n$ and an orthonormal basis of corresponding eigenvectors $\{v_1, \ldots, v_n\}$, then $A = \sum_{i=1}^{n} \lambda_i v_i v_i^*$.

2. The definition, and some properties, of positive and negative definite/semi-definite linear operators and matrices.

3. T is a positive definite (semi-definite) linear operator iff all its eigenvalues are positive (non-negative).

   Analogous properties are true for matrices.

4. $A \in M_n(C)$ is positive definite (semi-definite) iff all its principal minors are positive (non-negative).

5. $A \in M_n(C)$ is positive definite iff all its **leading** principal minors are positive.

6. **The Cholesky decomposition:** $A \in M_n(C)$ is positive definite iff $A = LL^*$, where L is a non-singular lower triangular matrix.
7. **Hadamard's Inequality**: If $A = \left[a_{ij}\right]$ is positive definite, then $\det A \le a_{11}a_{22} \ldots a_{nn}$.

5.5 SOLUTIONS/ANSWERS

E1) The eigenvalues of A are −2, 3, 6.
The unit eigenvectors corresponding to these eigenvalues are

$$u_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, u_2 = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}, u_3 = \frac{1}{\sqrt{6}}\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \text{ respectively.}$$

Then the spectral decomposition of A is $-2u_1 u_1^* + 3u_2 u_2^* + 6u_3 u_3^*$

$$= -2\left(\frac{1}{2}\right)\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 0 \end{bmatrix} + 3\left(\frac{1}{3}\right)\begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{bmatrix} + 6\left(\frac{1}{6}\right)\begin{bmatrix} 4 & 2 & 2 \\ 2 & 1 & 1 \\ 2 & 1 & 1 \end{bmatrix}.$$

E2) The eigenvalues of the matrix (say A) are 7, 7, −2.
It has the linearly independent (l.i) eigenvectors as

$$v_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, v_2 = \frac{1}{\sqrt{5}}\begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}, v_3 = \frac{1}{\sqrt{6}}\begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix}.$$

Now, note that these are l.i but not orthonormal. Here $v_3 \perp v_1$ and $v_3 \perp v_2$, but v_1 is not orthogonal to v_2. So, we use the Gram-Schmidt process to obtain an orthonormal basis of W_7, corresponding to $\{v_1, v_2\}$.

This is $\{w_1, w_2\}$, where $w_1 = v_1$, and $w_2 = \frac{1}{\sqrt{18}}\begin{bmatrix} -1 \\ 4 \\ 1 \end{bmatrix}$.

Then an orthonormal basis of eigenvectors is $\{w_1, w_2, v_3\}$.
So the spectral decomposition of A is $7w_1 w_1^* + 7w_2 w_2^* - 2v_3 v_3^*$.
You can calculate this and **verify** that the sum is A.

E3) From Example 3, Unit 4, the eigenvalues are 5, 2, 2.
The corresponding set of linearly independent eigenvectors are

$$\frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}.$$

These are not orthonormal. To orthonormalise them, we apply the Gram-Schmidt process to the basis of W_2. Then we get the vectors as

$$u_1 = \frac{1}{\sqrt{3}}\begin{bmatrix}1\\1\\1\end{bmatrix},\ u_2 = \frac{1}{\sqrt{2}}\begin{bmatrix}0\\1\\-1\end{bmatrix},\ u_3 = \frac{1}{\sqrt{3}}\begin{bmatrix}\sqrt{2}\\-\frac{1}{\sqrt{2}}\\-\frac{1}{\sqrt{2}}\end{bmatrix}.$$

Then the spectral decomposition for the matrix is $5u_1u_1^* + 2u_2u_2^* + 2u_3u_3^*$.

E4) An $n \times n$ Hermitian matrix A is called

i) positive definite if $x^*Ax > 0\ \forall x \neq 0,\ x \in \mathbf{C}^n$.

ii) positive semi-definite if $x^*Ax \geq 0\ \forall x \in \mathbf{C}^n$.

An example is $A = \begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}$. Here $x^*Ax = x_1^2 \geq 0\ \forall \begin{bmatrix}x_1\\x_2\end{bmatrix} \in \mathbf{C}^2$.

This is not positive definite since $x^*Ax = 0$ for $x = \begin{bmatrix}0\\1\end{bmatrix}$.

E5) The proof of Theorem 5 is on the same lines as that of Theorem 4, replacing '> 0' by '≥ 0'.

Similarly, you can follow the matrix version of the proof of Theorem 4 to get the proof of Theorem 6.

E6) Its eigenvalues are 4, 1, 1.

Accordingly, (as in E3) the spectral resolution is $4u_1u_1^* + u_2u_2^* + u_3u_3^*$,

$$\text{where } u_1 = \frac{1}{\sqrt{3}}\begin{bmatrix}1\\1\\1\end{bmatrix},\ u_2 = \frac{1}{\sqrt{2}}\begin{bmatrix}0\\1\\-1\end{bmatrix},\ u_3 = \frac{1}{\sqrt{3}}\begin{bmatrix}\sqrt{2}\\-\frac{1}{\sqrt{2}}\\-\frac{1}{\sqrt{2}}\end{bmatrix}.$$

$\therefore$ the square root is $S = 2u_1u_1^* + u_2u_2^* + u_3u_3^*$.

Check that S^2 is the given matrix.

E7) Let T be a positive definite operator on V. Since all its eigenvalues are positive, the constant term in its characteristic equation is non-zero. Hence, by the Cayley-Hamilton theorem, T is invertible.

E8) Firstly, since A is Hermitian, so is A^{-1}.

Next, if $A > 0$, then $A^{-1} > 0$

This is because λ is an eigenvalue of A^{-1}

iff λ^{-1} is an eigenvalue of A

iff $\lambda^{-1} > 0$

iff $\lambda > 0$.

E9) $\begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & x \\ 1 & \bar{x} & 1 \end{vmatrix} = x + \bar{x} - x\bar{x} - 1$

So, the matrix is psd iff $x + \bar{x} - x\bar{x} - 1 \geq 0$

i.e., iff $(a-1)^2 + b^2 \leq 0$, where $x = a + ib$

i.e., iff $a = 1, b = 0$.

i.e., iff $x = 1$.

E10) $|A| = -2$. So, by Theorem 7, A is neither pd nor psd.

E11) By E9, it is psd. Its sq. root is $S = \sqrt{3}\, u_1 u_1^*$, where $u_1 = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

You should **verify it** by actually checking that $S^2 = T$.

E12) $A = U^*DU$, where U is unitary, $D = \text{diag}(\mu_1, \mu_2, \ldots, \mu_n)$, where the μ_is are the eigenvalues of A.

So $\lambda_0 \leq \mu_i \leq \lambda_1 \ \forall\, i = 1, \ldots, n$.

Now $x^*Ax = y^*Dy$, where $y = Ux$

$$= \sum_{i=1}^{n} \mu_i |y_i|^2$$

$$\leq \lambda_1 \sum_{i=1}^{n} |y_i|^2$$

$$= \lambda_1 y^* y$$

$$= \lambda_1, \text{since } y^*y = x^* U^*Ux = 1.$$

Similarly, you can show that $\lambda_0 \leq x^*Ax$.