UNIT 14 SOME LINEAR ALGEBRA

Structure **Page No**

14.1 INTRODUCTION

In this unit, we discuss some concepts from linear algebra which will be useful in the study of multivariate statistical analysis. We look, in some detail, at idempotent matrices and quadratic forms. In the context of this course we start with the study of real symmetric matrices and the associated quadratic forms in Sec. 14.2. We define a classification for the quadratic forms and develop a method for determining the class to which a given quadratic form belongs.

In Sec.14.3, we study positive definite and semi definite matrices. In this Section we also obtain some characterizations of positive definite and semi definite matrices, and study some of their useful properties. Here we give a method of computing a square root of matrices; this plays an important role in transforming correlated random variables to uncorrelated random variables.

Idempotent matrices and Cochran's theorem play a key role in the distribution of quadratic forms in independent standard normal variables, particularly, in connection with the distribution of quadratic forms to become independent chi-squares. In Sections 14.4 and 14.5, respectively we study the properties of idempotent matrices, and prove the algebraic version of Cochran's theorem.

Singular value decompositions plays a very important role in developing the theory and studying the properties of canonical correlations between two random vectors. In Section 14.6, we study the singular value decomposition.

In this Unit, we shall be using the following notations. Matrices are denoted by capital letters like A, B, C. Vectors are denoted by boldface lower case letters like $\mathbf{x}, \mathbf{y}, \mathbf{z}$. Scalars are denoted by lower case letters like a, b, α. The transpose, rank and trace of a matrix A are denoted by A' or A^t, rank A and $\mathrm{tr}(A)$, respectively. R^n denotes the n-dimensional Euclidean space.

Objectives

After studying this unit, you should be able to

- determine the definiteness of a given quadratic form;
- apply the spectral decomposition in the study of principal components;
- compute a triangular square root of a positive definite matrix;
- apply the properties of positive and semi definite matrices to certain problems;

- apply Cochran's theorem to the distribution of quadratic forms in normal variables;
- apply the singular value decomposition in the development of canonical correlations.

Let us start our discussion with real symmetric matrices.

14.2 REAL SYMMETRIC MATRICES

Real symmetric matrices play a very important role in the study of multivariate statistical analysis. For example, the variance-covariance matrices are real symmetric matrices. They also play a crucial role in the distribution of quadratic forms in correlated normal random variables. We shall denote $(i, j)^{th}$ element of a matrix A by a_{ij}. Then we write $A = (a_{ij})$.

Definition 1: A square matrix $A = (a_{ij})$ of order n is called a **real symmetric matrix** if (i) all the elements of A are real and (ii) $a_{ij} = a_{ji}$ for $i, j = 1, \ldots, n$.

Definition 2: A quadratic form in n variables $x_1, x_2, \ldots, x_n$ is a homogeneous polynomial of degree 2 in these variable.

You also know that there is a unique real symmetric matrix A associated with a given real quadratic form $\mathbf{Q(x)}$, in the sense that $\mathbf{Q(x) = x'Ax}$. This matrix A is called the **matrix of the quadratic form $\mathbf{Q(x)}$**. (For reference, MTE-02, Sec. 14.3)

Example 1: Examine the following matrices for symmetric property

(i) $\begin{bmatrix} 2 & i \\ i & 3 \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix}$ (iii) $\begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}$ (iv) $\begin{bmatrix} -1 & 4 \\ 4 & 5 \end{bmatrix}$

Solution:

(i) is not a real symmetric matrix because all of its elements are not real.

(ii) is not a real symmetric matrix because it is not a square matrix.

(iii) is not a real symmetric matrix because $a_{12} = 3$ and $a_{21} = 2$ therefore $a_{12} \neq a_{21}$.

(iv) is a real symmetric matrix because (a) it is a square matrix (of order 2×2), (b) all of its elements are real and (c) $a_{12} = 4 = a_{21}$.

***

Example 2: Find the matrix of the quadratic form $\mathbf{Q(x)} = 2x_1x_2 + 5x_1x_3 + 3x_2^2 - x_3^2$.

Solution: Since there are three variables x_1, x_2 and x_3, $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$. Let A be the symmetric matrix such that $Q(x) = \mathbf{x'Ax}$. Then A is of order 3×3. Further $a_{21} = a_{12} = 1/2$ (coefficient of $x_1 x_2$) $= 1$. In general, whenever $i \neq j$, $a_{ij} = a_{ji} = 1/2$ (coefficient of $x_i x_j$). Also a_{ii} = the coefficient of x_i^2, $i = 1, 2, 3$. Thus

$$A = \begin{bmatrix} 0 & 1 & 2.5 \\ 1 & 3 & 0 \\ 2.5 & 0 & -1 \end{bmatrix}$$

***

Let us now try an exercise.

E1) Find the matrices of the following quadratic forms:

(i) $x_1^2 - x_2^2$

(ii) $2x_1^2 + 3x_1x_2 + 5x_2^2$

(iii) $3x_1x_2 + 5x_2x_3 - 4x_1x_3$

(iv) $x_1^2 + x_2^2 + x_2x_4$ (in four variables x_1, x_2, x_3 and x_4)

Depending upon the range, every non null quadratic form $Q(\mathbf{x})$ in n variables can be classified into one of the following mutually exclusive and collectively exhaustive classes:
(It is also said to be identification of the definiteness of the quadratic form.)

(a) positive definite (pd) if $Q(\mathbf{x}) > 0$ for all $\mathbf{x} \in \mathbf{R}^n, \mathbf{x} \neq \mathbf{0}$,

(b) positive semidefinite (psd) if
(i) $Q(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \mathbf{R}^n$, and
(ii) $Q(\mathbf{x}) = 0$ for some $\mathbf{x} \neq \mathbf{0}$

(c) negative definite (nd) if $Q(\mathbf{x}) < 0$ for all $\mathbf{x} \neq \mathbf{0}$ (i.e., if $-Q(x)$ is positive definite),

(d) negative semidefinite (nsd) if (i) $Q(\mathbf{x}) \leq 0$ for all $\mathbf{x}$ and (ii) $Q(x) = 0$ for some $\mathbf{x} \neq \mathbf{0}$. (i.e., if $-Q(x)$ is positive semidefinite),

(e) indefinite, if it does not belong to any one of the above classes (a) – (d) (i.e., there exists $\mathbf{x}$ and $\mathbf{y}$ in $\mathbf{R}^n$ such that $Q(\mathbf{x}) > 0$ and $Q(\mathbf{y}) < 0$)).

The quadratic form $Q(\mathbf{x}) \equiv 0$ can be classified into any one of the classes (b) and (d).

Example 3: Classify each of the following quadratic forms using the above classification. Also write down the matrices of the respective quadratic forms.

(i) $x_1^2 - x_2^2$ (ii) $x_1^2 + x_2^2$

(iii) $x_1^2 + x_2^2 + 2x_4^2$ (in four variables x_1, x_2, x_3 and x_4)

(iv) $-x_1^2 - x_2^2$

(v) $-x_1^2 - x_2^2 - 2x_4^2$ (again in four variables)

Solution:

(i) $x_1^2 - x_2^2 = 1$ if $x_1 = 1$ and $x_2 = 0$. Again $x_1^2 - x_2^2 = -1$ if $x_1 = 1$ and $x_2 = 1$.

Thus it is indefinite. The matrix of the quadratic form is $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$.

(ii) $x_1^2 + x_2^2 > 0$ whenever at least one of x_1 and x_2 is not zero. Hence this quadratic form is positive definite. The matrix of the quadratic form is $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, the identity matrix.

(iii) $Q(\mathbf{x}) = x_1^2 + x_2^2 + 2x_4^2 \geq 0$ for all values of x_1, x_2, x_3 and x_4. However, for $x_3 = 1$ and $x_1 = x_2 = x_4 = 0$, the value of $x_1^2 + x_2^2 + 2x_4^2 = 0$. Thus, there is a vector $\mathbf{x} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \neq \mathbf{0}$ such that $Q(\mathbf{x}) = 0$. Hence this quadratic form is

positive semi-definite. The matrix of the quadratic form is $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

We leave it to you to show that the quadratic forms in (iv) and (v) are negative definite and negative semi-definite, respectively. (You can use the quadratic forms in (ii) and (iii) to arrive at this conclusion, and for writing down the matrices of the quadratic forms in (iv) and (v).)

***

In Example 3, we considered quadratic forms whose matrices are diagonal matrices. Here it is easy to identify the definiteness of the quadratic form. In fact, if

$Q(\mathbf{x}) = \sum_{i=1}^{n} \lambda_i x_i^2$ is a quadratic form in n variables $x_1, \ldots, x_n$, then $Q(\mathbf{x})$ is p.d., p.s.d., n.d., n.s.d. or indefinite according as $\lambda_i > 0$ for all i; $\lambda_i \geq 0$ for all i and $\lambda_j = 0$ for some j; $\lambda_i < 0$ for all i; $\lambda_i \leq 0$ for all i and $\lambda_j = 0$ for some j; or if $\lambda_i < 0$ for some i and $\lambda_j > 0$ for some j, respectively.

Now, what if we have a quadratic form $Q(\mathbf{x}) = 2x_1^2 - 3x_1x_2 + x_2^2$ or $Q(\mathbf{x}) = 2x_1^2 + x_2^2 + x_3^2 - 3x_1x_2 - 2x_1x_3 + 4x_2x_3$? (Notice that the matrices of these quadratic forms are $\begin{bmatrix} 2 & -1.5 \\ -1.5 & 1 \end{bmatrix}$ and $\begin{bmatrix} 2 & -1.5 & -1 \\ -1.5 & 1 & 2 \\ -1 & 2 & 1 \end{bmatrix}$, respectively, and are not diagonal matrices.)

In general, consider a quadratic form $Q(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x}$, where A is not a diagonal matrix. How do we determine the definiteness of the quadratic form in such a case? The following results will be useful towards that end.

Theorem 1: Consider a quadratic form $Q(\mathbf{x}) = \mathbf{x}'A\mathbf{x}$ where A is symmetric. Make a nonsingular linear transformation of the variables: $\mathbf{y} = \mathbf{T}\mathbf{x}$ (where T is nonsingular). Call the transformed quadratic form as $\boldsymbol{\psi}(\mathbf{y})(= \mathbf{y}'T^{-1'}AT^{-1}\mathbf{y})$. Then the ranges of $Q(\mathbf{x})$ and $\psi(\mathbf{y})$ are the same.

Proof: Let α belongs to the range of $Q(\mathbf{x})$. So, there is a vector $\mathbf{x}_0$ such the $\alpha = Q(\mathbf{x}_0) = \mathbf{x}_0'A\mathbf{x}_0$. Write $y_0 = T\mathbf{x}_0$. Now $\alpha = \mathbf{x}_0'A\mathbf{x}_0 = \mathbf{x}_0'T'T^{-1'}AT^{-1}T\mathbf{x}_0 = \mathbf{y}_0'T^{-1'}AT^{-1}\mathbf{y}_0 = \psi(\mathbf{y}_0)$. Hence α belongs to the range of $\psi(\mathbf{y})$. Thus, the range of $Q(\mathbf{x})$ is a subset of the range of $\psi(\mathbf{y})$. Since T is nonsingular, by reversing the arguments, we can show that the range of $\psi(\mathbf{y})$ is a subset of the range of $Q(\mathbf{x})$. The proof is complete.

What we are saying through Theorem 1 is that the range of a quadratic form is invariant under nonsingular linear transformations. Thus, the definiteness of a quadratic form is invariant under nonsingular linear transformations. (Making a nonsingular linear transformation can also be interpreted as changing the basis.)

Recall that a real square matrix S is called an **orthogonal matrix** if $S' = S^{-1}$. If S and T are orthogonal matrices of the same order, then so is ST. To see this we note that $T'S'ST = T'IT = I$. Similarly, $STT'S' = I$. Hence $(ST)' = T'S'$ is the inverse of ST. Also, you should verify that $\begin{bmatrix} I & 0 \\ 0 & T \end{bmatrix}$ is an orthogonal matrix if T is an orthogonal matrix.

So, now we want to determine the definiteness of a quadratic form $Q(\mathbf{x})$, the matrix of which is not necessarily diagonal. We shall now show that we can make an orthogonal transformation of the variables (i.e., we can make a transformation $\mathbf{y} = P\mathbf{x}$ where P is an orthogonal matrix) such that under this transformation, the quadratic form is transformed into a quadratic form $\sum \lambda_i y_i^2$. Since we know how to determine the definiteness of $\sum \lambda_i y_i^2$, and since the definiteness of $\sum \lambda_i y_i^2$ is the same as that of $Q(\mathbf{x})$, we would then have the definiteness of $Q(\mathbf{x})$.

If A is a real matrix, then it is not necessary that its eigenvalues are real. For example if $A = \begin{bmatrix} 0 & 1 \\ -1 & \end{bmatrix}$, then the eigenvalues are i and $-i$. However, if A is real, and symmetric then all its eigenvalues are real as shown below.

Theorem 2: Let A be a real symmetric matrix. All the eigenvalues of A are real and all the eigenvectors of A can be chosen to be real.

Proof: Let $\lambda + i\mu$ be an eigenvalue of A and let the corresponding eigen vector be $\mathbf{x} + i\mathbf{y}$, where λ, μ are real numbers and $\mathbf{x}, \mathbf{y}$ are real vectors. Clearly at least one of $\mathbf{x}$ and $\mathbf{y}$ is non-zero as $\mathbf{x} + i\mathbf{y}$, being an eigen vector is nonnull. Now,

$$A(\mathbf{x} + i\mathbf{y}) = (\lambda + i\mu)(\mathbf{x} + i\mathbf{y})$$

Equating the real and the imaginary parts on both sides of the above equality, we get,

$$A\mathbf{x} = \lambda\mathbf{x} - \mu\mathbf{y} \tag{1}$$

$$A\mathbf{y} = \lambda\mathbf{y} + \mu\mathbf{x} \tag{2}$$

Premultiplying (1) by $\mathbf{y}'$ and (2) by $\mathbf{x}'$, we get

$$\mathbf{y}'A\mathbf{x} = \lambda\mathbf{y}'\mathbf{x} - \mu\mathbf{y}'\mathbf{y} \tag{3}$$

$$\mathbf{x}'A\mathbf{y} = \lambda\mathbf{x}'\mathbf{y} + \mu\mathbf{x}'\mathbf{x} \tag{4}$$

Since A is symmetric and $\mathbf{y}'A\mathbf{x}$ is a scalar we have $\mathbf{y}'A\mathbf{x} = (\mathbf{y}'A\mathbf{x})' = \mathbf{x}'A'\mathbf{y} = \mathbf{x}'A\mathbf{y}$. Similarly, $\mathbf{y}'\mathbf{x} = \mathbf{x}'\mathbf{y}$. Now subtracting Eqn.(3) from Eqn.(4) we get $\mu(\mathbf{x}'\mathbf{x} + \mathbf{y}\mathbf{y}) = 0$. Since at least one of x and y is non-null, $\mathbf{x}'\mathbf{x} + \mathbf{y}'\mathbf{y} \neq 0$. So, $\mu = 0$.

Hence all the eigenvalues of A are real. Further $A(\mathbf{x} + i\mathbf{y}) = \lambda(\mathbf{x} + i\mathbf{y})$, yields $A\mathbf{x} = \lambda\mathbf{x}$ and $A\mathbf{y} = \lambda\mathbf{y}$. Since at least one of $\mathbf{x}$ and $\mathbf{y}$ is non-zero and $\mathbf{x}$, $\mathbf{y}$ are real, we can choose a non-zero vector among $\mathbf{x}$, $\mathbf{y}$ as an eigenvector of A corresponding to λ. This complete the proof of the theorem.

Now we are ready to prove the result we had mentioned earlier in the following theorem.

Theorem 3: Let A be a real symmetric matrix of order n. Then there exists a real orthogonal matrix P of order n such that $A = P\Lambda P'$, where Λ is a real diagonal matrix.

Proof: We shall prove the theorem by induction on n. Let A be a 1×1 real symmetric matrix, i.e. $A = (\alpha)$, where α is a real number. Clearly $(1)'(\alpha)(1) = (1.\alpha.1) = (\alpha)$. Also (1) is an orthogonal matrix of order 1×1 since $(1)'.(1) = (1.1) = (1)$. So the theorem is true for $n = 1$. Let the theorem be true $n = m$ (a positive integer > 1). Let A be a matrix of order $(m+1) \times (m+1)$. Let x_1 be a normalized eigenvector of A corresponding to eigenvalue λ_1. Then $A\mathbf{x}_1 = \lambda_1 \mathbf{x}_1$. Now $\mathbf{x}_1$ can be extended to an orthonormal basis $\mathbf{x}_1, \ldots, \mathbf{x}_{m+1}$ of $\mathbf{R}^{m+1}$. Write $R = (\mathbf{x}_1 : \ldots : x_{m+1})$. Clearly R is an orthogonal matrix. Now

$$AR = \mathbf{A}(\mathbf{x}_1 : \ldots : \mathbf{x}_{m+1}) = (\mathbf{x}_1 : \ldots : \mathbf{x}_{m+1}) \begin{pmatrix} \lambda_1 & \mathbf{b}'_{12} \\ \mathbf{0} & B_{22} \end{pmatrix} = R \begin{pmatrix} \lambda_1 & \mathbf{b}'_{12} \\ \mathbf{0} & B_{22} \end{pmatrix}$$, where $\mathbf{0}$, $\mathbf{b}'_{12}$ and B_{22} are of order $m \times 1$, $1 \times m$ and $m \times m$, respectively. [This is because, $A\mathbf{x}_2, \ldots, A\mathbf{x}_{m+1}$ are vectors in R^{m+1} and $\mathbf{x}_1, \ldots, \mathbf{x}_{m+1}$ form a basis of R^{m+1}, so $A\mathbf{x}_i$ is a linear combination of $\mathbf{x}_1, \ldots, \mathbf{x}_{m+1}$.]

Therefore $R'AR = \begin{pmatrix} \lambda_1 & \mathbf{b}'_{12} \\ \mathbf{0} & B_{22} \end{pmatrix}$. Since $R'AR$ is real and symmetric it follows that $\mathbf{b}_{12} = \mathbf{0}$ and B_{22} is an $m \times m$ real symmetric matrix. Thus, $R'AR = \begin{pmatrix} \lambda_1 & \mathbf{0} \\ \mathbf{0} & B_{22} \end{pmatrix}$. By induction hypothesis there exists an orthogonal matrix S_1 of order $m \times m$ that $B_{22} = S_1 \Gamma S_1'$ where Γ is a real diagonal matrix. Writing $S = \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & S_1 \end{pmatrix}$, we notice that S is an orthogonal matrix and

$$R'AR = S \begin{pmatrix} \lambda_1 & \mathbf{0} \\ \mathbf{0} & \Gamma \end{pmatrix} S^t \text{ or } A = RS \begin{pmatrix} \lambda_1 & \mathbf{0} \\ \mathbf{0} & \Gamma \end{pmatrix} S'R'$$

Write $P = RS$. Since R and S are orthogonal matrices so is P as noticed earlier. Writing $D = \begin{pmatrix} \lambda & \mathbf{0} \\ \mathbf{0} & \Gamma \end{pmatrix}$, we observe that D is a real diagonal matrix.

Thus, the theorem is true for $n = m + 1$.

Hence the theorem follows by induction on n.

The beauty of Theorem 3 lies in its interpretation. Let A be a real symmetric matrix and let $A = P\Lambda P'$ where P is orthogonal and Λ is a real diagonal matrix. We then have

$$AP = P\Lambda \qquad \text{or} \qquad A(\mathbf{p}_1 : \ldots : \mathbf{p}_n) = (\mathbf{p}_1 : \ldots : \mathbf{p}_n) \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} \qquad \text{or}$$

$AP_i = \lambda_i \mathbf{p}_i, i = 1, \ldots, n$.

Since $\mathbf{p}_1$ is a vector in an orthonormal basis, $\mathbf{p}_1$ is (a non-null vector) of unit norm. Hence λ_1 is an eigenvalue of A and $\mathbf{p}_1$ is an eigenvector of A corresponding to the eigenvalue λ_1. Thus, the diagonal elements of Λ are the eigenvalues of A and the columns of P are the orthonormal eigenvectors of A. Further

$$A = P\Lambda P' = (\mathbf{p}_1 : \ldots : \mathbf{p}_n)\begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}\begin{pmatrix} \mathbf{p}_1' \\ \vdots \\ \mathbf{p}_n' \end{pmatrix} = \lambda_1 \mathbf{p}_1 \mathbf{p}_1' + \cdots + \lambda_n \mathbf{p}_n \mathbf{p}_n'$$

Write $E_i = \mathbf{p}_i \mathbf{p}_i'$, $i = 1, \ldots, n$

Then $E_i E_j = \begin{cases} E_i \text{ if } i = j \\ 0 \text{ otherwise} \end{cases}$

and rank E_i = rank $\mathbf{p}_i \mathbf{p}_i'$ = rank $p_i = 1$.

Thus, we are able to write $A = \sum_{i=1}^{n} \lambda_i E_i$ where $E_1, \ldots, E_n$ are symmetric idempotent matrices of rank 1 such that $E_i E_j = 0$ whenever $i \neq j$. The set $\{\lambda_1, \ldots, \lambda_n\}$ is called the spectrum of A. Since the decomposition mentioned above involves the spectrum and the eigenvectors it is called a spectral decomposition of A.

Example 4: Obtain a spectral decomposition of the matrix $A = \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix}$.

Solution: The characteristic equation of the matrix is $\begin{vmatrix} 4-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = 0$

or $(4-\lambda)(2-\lambda) - 1 = 0$

or $\lambda^2 - 6\lambda + 7 = 0$

The roots are $\dfrac{6 \pm \sqrt{36-28}}{2} = 3 \pm \sqrt{2}$

So the eigenvalues are $3+\sqrt{2}$ and $3-\sqrt{2}$.

Let $\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ be the eigenvector of the given matrix corresponding to the eigenvalue $3+\sqrt{2}$.

Then $[A-(3+\sqrt{2})I]\mathbf{u} = 0$ or $\begin{bmatrix} 1-\sqrt{2} & 1 \\ 1 & -1-\sqrt{2} \end{bmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \mathbf{0}$

Notice that the second column of $\left(A-(3+\sqrt{2})I\right)$ is $\left(-1-\sqrt{2}\right)$ times the first column.

So, $u_1 = 1+\sqrt{2}$ and $u_2 = 1$ satisfy the equation $\left(A-(3+\sqrt{2})I\right)\mathbf{u} = 0$.

To normalize $\mathbf{u}$, we divide it by its norm namely $\sqrt{\left(1+\sqrt{2}\right)^2 + 1^2} = \sqrt{4+2\sqrt{2}}$. Thus, the normalized eigenvector corresponding to the eigenvalue $3+\sqrt{2}$ is

$\dfrac{1}{\sqrt{4+2\sqrt{2}}}\begin{pmatrix} 1+\sqrt{2} \\ 1 \end{pmatrix}$. It can be shown similarly that the normalized eigenvector

corresponding to $3-\sqrt{2}$ is $\dfrac{1}{\sqrt{4+2\sqrt{2}}}\begin{pmatrix} 1 \\ -\left(1+\sqrt{2}\right) \end{pmatrix}$. Hence $A = P\Lambda P'$

where $P = \dfrac{1}{\sqrt{4+2\sqrt{2}}}\begin{pmatrix} 1+\sqrt{2} & 1 \\ 1 & -\left(1+\sqrt{2}\right) \end{pmatrix}$ and $\Lambda = \begin{pmatrix} 3+\sqrt{2} & 0 \\ 0 & 3-\sqrt{2} \end{pmatrix}$, is the required spectral decomposition.

***

Using Theorem 3, we can determine the definiteness of a quadratic form. Consider the quadratic form $Q(\mathbf{x}) = \mathbf{x}'A\mathbf{x}$. Let $A = P\Lambda P'$ be a spectral decomposition of A. Then $Q(\mathbf{x}) = \mathbf{x}'P\Lambda P'\mathbf{x} = \mathbf{y}'\Lambda\mathbf{y}$ where $\mathbf{y} = P'\mathbf{x}$. Since P is nonsingular (in fact, orthogonal) the definiteness of $Q(\mathbf{x})$ is the same as the definiteness of $\mathbf{y}'\Lambda\mathbf{y}$. The definiteness of $\mathbf{y}'\Lambda\mathbf{y}$ is determined by the diagonal elements $\lambda_1,\ldots,\lambda_n$ in, Λ the eigen values of A.

Thus

$$\mathbf{x}'A\mathbf{x} \text{ is } \begin{cases} \text{positive definite if } \lambda_i > 0 \text{ for all i} \\ \text{positive semidefinite if } \lambda_i \geq 0\ \forall \text{i and } \lambda_j = 0 \text{ for some j} \\ \text{negative definite if } \lambda_i < 0 \text{ for all i} \\ \text{negative semidefinite if } \lambda_i \leq 0 \text{ for all i and } \lambda_j = 0 \text{ for some j} \\ \text{indefinite if } \lambda_i > 0 \text{ for some i and } \lambda_j = 0 \text{ for some j} \end{cases}$$

Because of the one-one correspondence between real symmetric matrices and the quadratic forms, we call a real symmetric matrix A as positive definite, positive semidefinite, negative definite, negative semidefinite or indefinite accordingly as the corresponding quadratic form $\mathbf{x}'A\mathbf{x}$ is positive definite, positive semidefinite, negative definite, negative semidefinite or indefinite, respectively.

Now let us illustrate the following example as an application of this.

Example 5: Determine the definiteness of the quadratic forms (i) $2x_1^2 - x_1x_2 + x_2^2$ and (ii) $x_1^2 + x_2^2 + x_3^2 - 3x_1x_2 - 3x_1x_3 - 3x_2x_3$.

Solution: (i) The matrix of the quadratic form is $A = \begin{pmatrix} 2 & -0.5 \\ -0.5 & 1 \end{pmatrix}$.

We see that its characteristic equation $|A - \lambda I| = 0$ is

$(2-\lambda)(1-\lambda) - 0.25 = 0$ or $\lambda^2 - 3\lambda + 1.75 = 0$

Hence the eigenvalues which are the roots of the above equation are $\frac{3 \pm \sqrt{9-7}}{2}$ or

$\frac{3+\sqrt{2}}{2}$ and $\frac{3-\sqrt{2}}{2}$, which are positive.

Hence the quadratic form is positive definite.

(ii) The matrix of the quadratic form is $A = \begin{bmatrix} 1 & -1.5 & -1.5 \\ -1.5 & 1 & -1.5 \\ -1.5 & -1.5 & 1 \end{bmatrix}$.

It is easy to notice that the sum of each row in A is -2.

Hence $A\begin{pmatrix}1\\1\\1\end{pmatrix} = -2\begin{pmatrix}1\\1\\1\end{pmatrix}$. Thus, -2 is an eigenvalue of A. Further, the sum of the eigenvalues which is the same as the trace of A is 3. Hence there must be at least one positive eigenvalue of A. So the quadratic form is indefinite.

***

Now, try the following exercises.

E2) Let A be a real symmetric matrix, a diagonal element of which is negative. Show that A cannot be positive definite or positive semidefinite.

E3) Determine the definiteness of the following quadratic forms:

(i) $x_1^2 - 5x_1x_2 + 7x_2^2$, (ii) $x_1^2 - x_2^2 + x_3^2 - x_1x_2 + 10x_1x_3 - 2x_2x_3$,

(iii) $2x_1^2 + 3x_2^2 + 4x_3^2 + 6x_1x_2$

E4) Let $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. Obtain the spectral decomposition of A. Hence write down A^{100}.

In next section, we shall discuss more about positive definite and non-negative definite matrices.

14.3 POSITIVE DEFINITE AND SEMIDEFINITE MATRICES

In the previous section, we noted that the definiteness of a quadratic form is also attributed to the matrix of the quadratic form. Thus, if $\mathbf{x'Ax}$ is positive definite where A is a real symmetric matrix, then we call A as a positive definite (pd) matrix. A real symmetrix matrix is called a Semidefinite matrix if it is either pd or psd i.e., if $\mathbf{x'Ax \geq 0}$ for all $\mathbf{x}$. Positive definite (pd) and nonnegative definite (nnd) matrices play a very important role in the multivariate analysis. *Unless stated otherwise, when we say a matrix is pd, psd, nnd, nd, nsd we mean that the matrix is real and symmetric. We may not state this fact explicitly each time.*

In this section, we study several properties of positive definite and nonnegative definite matrices. We shall also give an easy way to construct positive definite matrices and orthogonal matrices of order $n \times n$. Let us start with the following very useful theorem.

Theorem 4: a) A matrix A is positive definite if and only if $A = BB'$ for some nonsingular matrix B.

b) A matrix A is nonnegative definite if and only if $A = CC'$ for some matrix C.

Proof: a) If part. Let $A = BB'$ for some nonsingular matrix B. Let $\mathbf{x}$ be a nonnull vector. Then $\mathbf{x'Ax = x'BB'x = y'y} = y_1^2 + \cdots + y_n^2 \geq 0$ where $\mathbf{y = B'x}$. Since B is nonsingular, so is B'.

Let if possible $\mathbf{y = B'x = 0}$. Then $\mathbf{x} = (B')^{-1} y = 0$.

Since $\mathbf{x \neq 0}$, there is a contradiction. So $\mathbf{y \neq 0}$.

Hence $\mathbf{x'Ax} = y'y > 0$. The choice of $\mathbf{x}$ being arbitrary, it follows, that A is positive definite. Only if part. Let A be positive definite. Then all its eigenvalues are strictly positive. Let $A = P\Lambda P'$ be a spectral decomposition of A. Let $\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n}$ be positive square roots of $\lambda_1, \ldots, \lambda_n$, respectively. Write

$$\Lambda^{\frac{1}{2}} = \begin{pmatrix} \sqrt{\lambda_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sqrt{\lambda_n} \end{pmatrix}$$

Then $B = P\Lambda^{1/2}P'$ is symmetric and $BB' = P\Lambda^{1/2}P'P\Lambda^{1/2}P' = P\Lambda^{1/2}\Lambda^{1/2}P' = P\Lambda P' = A$.

Further since $\lambda_1, \ldots, \lambda_n$ are strictly positive, so are $\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n}$. Now $|B| = |P\Lambda^{1/2}P'| = |P||P'||\Lambda^{1/2}| = |PP'||\Lambda^{1/2}| = |I||\Lambda^{1/2}| = \sqrt{\lambda_1 \ldots \lambda_n} > 0$. Hence B is nonsingular.

[Notice that $B = P\Lambda^{1/2}P'$ is a spectral decomposition of B. Now since diagonal elements of $\Lambda^{1/2}$ are strictly positive it follows that B is pd. In fact, we proved that more stronger result that if A is pd then $A = BB'$ for some symmetric pd matrix B.]

(b) It suffices to prove the statement for positive semidefinite matrices as an nnd matrix is either pd or psd. (For the pd matrices we already proved the statement in (a).)

If part. Let $A = CC'$. Then $\mathbf{x}'A\mathbf{x} = \mathbf{x}'CC'\mathbf{x} = \mathbf{u}'\mathbf{u} \geq 0$ where $\mathbf{u} = C'\mathbf{x}$. Hence A is nnd.

Only if part. Notice that since A is psd all its eigenvalues are nonnegative. Let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r > \lambda_{r+1} = \ldots = \lambda_n = 0$ be the eigenvalues of A. We can write a spectral decomposition of A as

$$A = P\begin{pmatrix} \Lambda_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}P'$$

where P is an orthogonal matrix and Λ_1 is a diagonal pd matrix of order $r \times r$. Write

$$\Lambda_1^{1/2} = \begin{pmatrix} \sqrt{\lambda_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sqrt{\lambda_r} \end{pmatrix}.$$

Then

$\mathbf{C} = P\begin{pmatrix} \Lambda_1^{1/2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}P'$ is symmetric and $CC^t = A$. This complete the proof.

A matrix B such that $A = BB'$ is called a square root of A. Given A, B is not unique since $A = BPP'B'$ where P is any orthogonal matrix. In Theorem 4 we gave a method of computing a square root if we know the spectral decomposition of A. However, obtaining spectral decomposition is not easy in general. We shall now discuss a method of obtaining a square root of a positive definite matrix.

Let us start with an example.

Example 6: Obtain a square root of the positive definite matrix $A = \begin{pmatrix} 4 & 1 & 2 \\ 1 & 3 & 1 \\ 2 & 1 & 5 \end{pmatrix}$

Solution: We shall obtain a lower triangular matrix B such that $A = BB'$. Write

$B = \begin{pmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & 0 \\ b_{31} & b_{32} & b_{33} \end{pmatrix}$. We shall solve for b_{ij} $j = i, \ldots, 3$, $i = 1, 2, 3$ such that $A = BB'$.

Write

$$\begin{pmatrix} 4 & 1 & 2 \\ 1 & 3 & 1 \\ 2 & 1 & 5 \end{pmatrix} = \begin{pmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & 0 \\ b_{31} & b_{32} & b_{33} \end{pmatrix}\begin{pmatrix} b_{11} & b_{21} & b_{31} \\ 0 & b_{22} & b_{32} \\ 0 & 0 & b_{33} \end{pmatrix}$$

Equating the elements on both sides, we get

$a_{11} = 4 = b_{11}^2$ or $b_{11} = 2$ (You can choose either $+2$ or -2 but choose and fix one of them.)

$a_{12} = 1 = b_{11}\, b_{21}$ so $b_{21} = \dfrac{1}{b_{11}} = 1/2$

$a_{13} = 2 = b_{11}\ b_{31}$ so $b_{31} = 1$

$a_{22} = 3 = b_{21}^2 + b_{22}^2$ or $b_{22}^2 = 3 - \frac{1}{4} = \frac{11}{4}$. So $b_{22} = \sqrt{\frac{11}{4}}$.

$a_{23} = 1 = b_{21}\ b_{31} + b_{22}\ b_{32}$ or $b_{22}\ b_{32} = 1 - 1/2.1 = 1/2$ or $b_{32} = \frac{1}{2}\sqrt{\frac{4}{11}} = \frac{1}{\sqrt{11}}$.

$a_{33} = 5 = b_{31}^2 + b_{32}^2 + b_{33}^2$ or $b_{33}^2 = 5 - 1^2 - \left(\frac{1}{\sqrt{11}}\right)^2 = 5 - 1 - \frac{1}{11} = \frac{43}{11}$ or $b_{33} = \sqrt{\frac{43}{11}}$.

Thus, $B = \begin{pmatrix} 2 & 0 & 0 \\ \frac{1}{2} & \sqrt{\frac{11}{4}} & 0 \\ 1 & \frac{1}{\sqrt{11}} & \sqrt{\frac{43}{11}} \end{pmatrix}$ is square root of A.

***

For the given matrix A in Example 6, we could obtain a lower triangular matrix B such that $A = BB'$. Can we always do this? Let us examine how we went about in solving for the elements of B. First we solved for the first column of B, then for the second column and so on. Also observe that each time we just had to solve one equation in one unknown to obtain the elements of B. Could there have been some hitch? What if the computed value for b_{22}^2 or later for b_{33}^2 turns out to be negative? If it happens to be so, we would not have been able to solve for B. It can be shown that if A is positive definite then the above situation never arises.

Let us now try an exercise.

E 5) Compute a lower triangular square root in each of the following cases.

(i) $\begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix}$ (ii) $\begin{pmatrix} 9 & 3 & 3 \\ 3 & 5 & 1 \\ 3 & 1 & 6 \end{pmatrix}$

Let us illustrate few more examples to understand the concept of definiteness.

Example 7: Let $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{12}' & A_{22} \end{pmatrix}$ be an $n \times n$ positive definite matrix where A_{11} and A_{22} are square matrices of order $r \times r$ and $(n-r) \times (n-r)$ respectively for some $r (1 \le r \le n-1)$. Show that A_{11} is positive definite.

Solution: Let $\mathbf{x}$ be a $r \times 1$ nonnull vector.

Then $\mathbf{x}' A_{11} \mathbf{x} = (\mathbf{x}' : \mathbf{0})_{1\times n} \begin{pmatrix} A_{11} & A_{12} \\ A_{12}' & A_{22} \end{pmatrix}_{n\times n} \begin{pmatrix} \mathbf{x} \\ \mathbf{0} \end{pmatrix}_{n\times 1} = (\mathbf{x}' : \mathbf{0}) A \begin{pmatrix} x \\ 0 \end{pmatrix} > \mathbf{0}$, since $\begin{pmatrix} \mathbf{x} \\ \mathbf{0} \end{pmatrix}$ is a nonnull vector. Hence A_{11} is pd.

***

Example 8: Let A be a positive definite matrix. Then show that $|A| > \mathbf{0}$.

Solution: Since A is positive definite, by theorem 4(a), $A = BB'$ for some nonsingular matrix B. So

$$|A| = |BB'| = |B| . |B'| = |B|^2 > 0$$

Let $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix}$ be a pd matrix.

Write $A_i = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1i} \\ a_{12} & a_{22} & \cdots & a_{2i} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1i} & a_{2i} & \cdots & a_{ii} \end{pmatrix}$, $i = 1, \ldots, n$.

The matrices $A_i, i = 1, \ldots, n$ are called leading principal submatrices of A. Combining Examples 6 and 7 we have $|A_i| > 0$ for $i = 1, \ldots, n$ if A is p.d. Is the converse true? Next theorem has the answer of this which is stated below without proof.

***

Theorem 5: Let A be a real symmetric matrix of order $n \times n$. Let $A_i, i = 1, \ldots, n$ be as defined above. Then A is positive definite if and only if $|A_i| > 0$ for $A_i, i = 1, \ldots, n$.

Example 9: Let A be a symmetric positive definite matrix. Show that RAR' is pd where R is any nonsingular matrix.

Solution: By Theorem 4(a) $A = BB'$, where B is nonsingular. Then $RAR' = RBB'R' = CC'$ where $C = RB$. Further, C is nonsingular since R and B are nonsingular. Hence RAR' is pd.

***

Example 10: A symmetric matrix A is positive definite if RAR' is pd for some nonsingular matrix R.

Solution: Let $\mathbf{x} \neq \mathbf{0}$. Consider $\mathbf{x}'A\mathbf{x} = \mathbf{x}'R^{-1}RAR'R^{-1'}\mathbf{x} = \mathbf{y}'RAR'\mathbf{y}$ where $\mathbf{y} = R^{-1'}\mathbf{x} \neq \mathbf{0}$ since $\mathbf{x} \neq \mathbf{0}$. Hence $\mathbf{x}'A\mathbf{x} = \mathbf{y}'RAR'\mathbf{y} > 0$ since $\mathbf{y} \neq \mathbf{0}$ and RAR' is pd. Hence A is pd.

***

Now try the following exercises.

E6) Let C and D be symmetric matrices of order $r \times r$ and $(n-r) \times (n-r)$ respectively. Show that $\begin{pmatrix} C & \mathbf{0} \\ \mathbf{0} & D \end{pmatrix}$ is pd if and only if C and D are pd.

E7) Let $A = \begin{pmatrix} A_{11} & A_{12} \\ A'_{12} & A_{22} \end{pmatrix}$ where A_{11} and A_{22} are square. Show that if A is pd, then A_{22} is pd.

E8) Let A be as in E7. Show that if A is nnd, then A_{11} and A_{22} are nnd.

Now let us take up some more theorems.

Theorem 6: Let $A = \begin{pmatrix} A_{11} & A_{12} \\ A^t_{12} & A_{22} \end{pmatrix}$ be a partition of A, where A_{11} and A_{22} are square. A is pd, if and only if A_{11} and $A_{11} - A'_{12} A_{11}^{-1} A_{12}$ are pd.

Proof: For the 'if' part A_{11} is pd. For the 'only if' part A is pd and hence A_{11} is pd by Theorem 5. Thus for both parts A_{11}^{-1} exists. It is easy to see that

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{12}^t & A_{22} \end{pmatrix} = \begin{pmatrix} I & 0 \\ A_{12}'A_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} A_{11} & \mathbf{0} \\ \mathbf{0} & A_{22} - A_{12}'A_{11}^{-1}A_{12} \end{pmatrix} \begin{pmatrix} I & A_{11}^{-1}A_{12} \\ \mathbf{0} & I \end{pmatrix}$$

Hence $A = R \begin{pmatrix} A_{11} & \mathbf{0} \\ \mathbf{0} & A_{22} - A_{12}'A_{11}^{-1}A_{12} \end{pmatrix} R'$

where $R = \begin{pmatrix} I & 0 \\ A_{12}'A_{11}^{-1} & I \end{pmatrix}$. Notice that $|R| = |I|.|I| = 1$

Hence R is nonsingular. By Examples 8 and 9, and E6 it follows that A is pd if and only if A_{11} and $A_{22} - A_{12}' A_{11}^{-1} A_{12}$ are pd. This completes the proof.

As we stated in the beginning that we shall give an easy way to construct pd matrices and orthogonal matrices. We shall do so now.

Theorem 7: Let A be a symmetric matrix of order $n \times n$ with positive diagonal elements and let $a_{ii} > \sum_{j \neq i} |a_{ij}|$ $i = 1, \ldots, n$.

Then A is positive definite.

The proof is beyond the scope of this notes.
Using the above theorem, it is easy to see that the matrix A in Example 5 and those in E5 are pd.

Theorem 8: Let $\mathbf{u}$ be a vector with unit norm. Then $I - 2\mathbf{uu}'$ is a symmetric orthogonal matrix.

Proof: $(I - 2\mathbf{uu}')' = I' - 2(\mathbf{u}')'\mathbf{u}' = I - 2\mathbf{uu}'$
Hence $I - 2\mathbf{uu}'$ is symmetric.
Further, $(I - 2\mathbf{uu}')\ (I - 2\mathbf{uu}') = I - 2\mathbf{uu}' - 2\mathbf{uu}' + 4\mathbf{uu}'\mathbf{uu}' = I - 4\mathbf{uu}' + 4\mathbf{uu}' = I$ since $\mathbf{u}'\mathbf{u} = 1$. So $I - 2\mathbf{uu}'$ is orthogonal.

Example 11: Let A be a positive semidefinite matrix of order $n \times n$ and let $a_{ii} = 0$. Then show that $a_{ij} = 0$ for $j = 1, \ldots, n$.

Solution: Consider $\mathbf{x} = \alpha\, \mathbf{e}_i + \mathbf{e}_j$ where $\mathbf{e}_i$ and $\mathbf{e}_j$ are i^{th} and j^{th} columns of the identity matrix of order $n \times n$ and α is a real number.

Now $\quad \mathbf{x}'A\mathbf{x} = \alpha^2 \mathbf{e}_i' A \mathbf{e}_i + \mathbf{e}_j' A \mathbf{e}_j + 2\alpha \mathbf{e}_i' A \mathbf{e}_j$

$$= \alpha^2 \mathbf{a}_{ii} + \mathbf{a}_{jj} + 2\alpha \mathbf{a}_{ij} = \mathbf{a}_{jj} + 2\alpha \mathbf{a}_{ij}$$

Let if possible $a_{ij} \neq 0$. Choose $\alpha = \dfrac{-(a_{jj} + 1)}{2a_{ij}}$. Then $\mathbf{x}'A\mathbf{x} = a_{jj} - (a_{jj} + 1) = -1 < 0$ which is a contradiction since A is psd. Hence $a_{ij} = 0$. Choice of j being arbitrary, the result follows.

Now let us try the following exercises.

E9) Let A be a nnd matrix. Show that $\mathbf{x}'A\mathbf{x} = 0$ if and only if $A\mathbf{x} = \mathbf{0}$.

E10) Show that $(1 - \rho)I + \rho 1 1'$ is pd if and only if $-\dfrac{1}{n-1} < \rho < 1$ where n is the order of the matrix and $1' = (1, \ldots, 1)$.

E11) Construct a 3×3 symmetric nondiagonal positive definite matrix A such that $a_{11} = 2, a_{22} = 5, a_{33} = 4$.

E12) Let A and B be nnd matrices of the same order. Show that (i) $A + B$ is nnd; (ii) the column space of A is a subspace of the column space of $A + B$.

So far we have discussed about various types of real symmetric matrices and other types of real symmetric matrices. Now let us discuss about idempotent matrices.

14.4 IDEMPOTENT MATRICES

A square matrix A is said to be idempotent if $A^2 = A$. Can you quickly come up with some examples of idempotent matrices? Yes, you are right! O and I are idempotent matrices. In fact, the only nonsingular idempotent matrix is I. Why? This is so because $A^2 = A$ and A is nonsingular implies $A = I$ (premultiply both sides of $A^2 = A$ by A^{-1}.) Similarly, the only rank 0 square matrix, namely O is idempontent. What about idempotent matrices of order $n \times n$ of rank r $(1 \le r \le n-1)$? $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ is an example of an idempotent matrix of rank r.

Further, if A is an idempotent matrix and P is a nonsingular matrix of the same order, then PAP^{-1}. $PAP^{-1} = PA^2P^{-1} = PAP^{-1}$.

Thus, PAP^{-1} is an idempotent matrix.

Hence $P\begin{pmatrix} I_r & 0 \\ \mathbf{0} & \mathbf{0} \end{pmatrix} P^{-1}$ is an idempotent matrix of rank r for every nonsingular matrix P. We shall now show that every idempotent matrix of rank r is of the form $P\begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} P^{-1}$ for some nonsingular matrix P.

Theorem 9: Let A be an $n \times n$ matrix of rank r $(1 \le r \le n-1)$. Then A is idempotent if and only if $A = P\begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} P^{-1}$ for some nonsingular matrix P.

Proof: 'If' part has already been proved above. For the 'only if' part, let A be an : $n \times n$ idempotent matrix of rank r $(1 \le r \le n-1)$. Let $A = (\mathbf{a}_1 : \mathbf{a}_2 : \ldots : \mathbf{a}_n)$. We have

$$(\mathbf{a}_1 : \ldots : \mathbf{a}_n) = A = A^2 = A(\mathbf{a}_1 : \ldots : \mathbf{a}_n)$$

Hence $A\mathbf{a}_i = \mathbf{a}_i, i = 1, \ldots, n$, Since rank $A = r$, there exist r linearly independent columns $\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_r}$ of A. Thus

$$A\,\mathbf{a}_{i_j} = \mathbf{a}_{i_j}, \; j = 1, \ldots, r. \qquad (1)$$

Again, $A(I - A) = \mathbf{0}$. Hence the set of columns of $(I - A)$ is a subspace of the null space $N(A)$ of A. We know that dimension of $N(A) = n - \text{rank } A = n - r$. So, rank of $I - A$ is at most $n - r$. On the other hand, since $I = A + (I - A)$, $n =$ rank $I \le$ rank $A +$ rank $(I - A)$.

Hence rank $(I - A) \ge n -$ rank A. Thus, rank $(I - A) = n -$ rank A. This, coupled with the fact the column space of $(I - A) \subset N(A)$, shows that the column space of

$(I-A)$ is the same as the null space of A. Let $\mathbf{e}_{l_1}-\mathbf{a}_{l_1},\ldots,\mathbf{e}_{l_{n-r}}-\mathbf{a}_{l_{n-r}}$ be linearly independent columns of $I-A$. Then

$$A\left(\mathbf{e}_{l_k}-\mathbf{a}_{l_k}\right)=0,\ k=1,\ldots,n-r \qquad (2)$$

Consider $P=\left(\mathbf{a}_{i_1}:\ldots:\mathbf{a}_{i_r}:\mathbf{e}_{l_1}-\mathbf{a}_{l_1}:\ldots:\mathbf{e}_{l_{n-r}}-\mathbf{a}_{l_{n-r}}\right)$. Clearly, P is an $n\times n$ matrix. Let $P\mathbf{x}=\mathbf{0}$. Then, $x_1\mathbf{a}_{i_1}+x_2\mathbf{a}_{i_2}+\ldots+x_r\mathbf{a}_{i_r}+x_{r+1}\left(\mathbf{e}_{l_1}-\mathbf{a}_{l_1}\right)+\ldots+x_n\left(\mathbf{e}_{l_{n-r}}-\mathbf{a}_{l_{n-r}}\right)=\mathbf{0}$.

Now,

$$\begin{aligned}\mathbf{0}=AP\mathbf{x}&=x_1A\mathbf{a}_{i_1}+x_2A\mathbf{a}_{i_2}+\ldots+x_rA\mathbf{a}_{i_r}+x_{r+1}A\left(\mathbf{e}_{l_1}-\mathbf{a}_{l_1}\right)+\ldots+x_nA\left(\mathbf{e}_{l_{n-r}}-\mathbf{a}_{l_{n-r}}\right)\\&=x_1\mathbf{a}_{i_1}+\ldots+x_r\mathbf{a}_{i_r}\quad\text{in view of Eqns. (1) and (2).}\end{aligned}$$

This implies $x_1=x_2=x_r=0$ since $a_{i_1},\ldots,a_{i_r}$ are linearly independent. This in turn implies $x_{r+1}=\ldots,x_n=0$ since $\mathbf{e}_{l_i}-\mathbf{a}_{l_i},\ i=1,\ldots,n-r$ are linearly independently.

Thus, $P\mathbf{x}=\mathbf{0}$ implies $\mathbf{x}=\mathbf{0}$ or the columns of P are linearly independent. So rank $P=n$ or in other words, P is nonsingular.

Further,
$$\begin{aligned}AP&=A\left(\mathbf{a}_{i_1}:\ldots:\mathbf{a}_{i_r}:\mathbf{e}_{l_1}-\mathbf{a}_{l_1}:\ldots:\mathbf{e}_{l_{n-r}}-\mathbf{a}_{l_{n-r}}\right)\\&=(a_{i_1}:\ldots:a_{i_r}:0:\ldots:0)\\&=\left(\mathbf{a}_{i_1}:\ldots:\mathbf{a}_{i_r}:\mathbf{e}_{l_1}-\mathbf{a}_{l_1}:\ldots:\mathbf{e}_{l_{n-r}}-\mathbf{a}_{l_{n-r}}\right)\begin{pmatrix}I_r&0\\0&0\end{pmatrix}\\&=P\begin{pmatrix}I_r&0\\0&0\end{pmatrix}.\end{aligned}$$

Thus, we have $A=P\begin{pmatrix}I_r&0\\0&0\end{pmatrix}P^{-1}$. This completes the proof.

Let A be an idempotent matrix of order $n\times n$ with rank r. From Theorem 9, the following statements are clear.

(a) A is similar to a diagonal matrix.

(b) A has at most two distinct eigenvalues 1 and 0. Eigenvalue 1 is with algebraic multiplicity r and 0 with algebraic multiplicity $n-r$.

Finding the rank of a matrix in general is not very easy. However, it is quite easy for idempotent matrices. We start with a definition.

Definition: The trace of a square matrix A of order $n\times n$ is defined as the sum of its diagonal elements and is denoted by $\mathrm{tr}(A)$. Thus, $\mathrm{tr}(A)=\sum_{i=1}^{n}a_{ii}$

Example 12: Let A and B be are square matrices of the same order. Show that (i) $\mathrm{tr}(cA)=c.\mathrm{tr}(A)$ when c is a real number; (ii) $\mathrm{tr}(A+B)=\mathrm{tr}(A)+\mathrm{tr}(B)$.

Solution: (i) $\mathrm{tr}(CA)=\sum_{i=1}^{n}caii=c\sum_{i=1}^{n}aii=c\,\mathrm{tr}(A)$

(ii) $\mathrm{tr}(A+B)=\sum_{i=1}^{n}(aii+bii)=\sum_{i=1}^{n}aii+\sum_{i=1}^{n}bii=\mathrm{tr}(A)+\mathrm{tr}(B)$

***

Example 13: Let A and B be matrices of order $m\times n$, and $n\times m$ respectively. Show that $tr(AB) = tr(BA)$.

Solution: The $(i, i)^{th}$ element of AB is given by $\sum_{j=1}^{n} a_{ij}b_{ji}$. Hence

$tr(AB) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}b_{ji} = \sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}b_{ji} = \sum_{j=1}^{n}\sum_{i=1}^{m} b_{ji}a_{ij} = tr(BA)$, since $\sum_{i=1}^{m} b_{ji}a_{ij}$ is the $(j, j)^{th}$ element of AB.

***

We are now ready to prove the following theorems.

Theorem 10: Let A be an idempotent matrix of any order $n\times n$. Then rank $A = tr(A)$.

Proof: If rank A is 0 or n, then always $A = tr(A)$. Let rank $A = r$ when $1 \le r \le n-1$. Then by Theorem 9, there exists a nonsingular matrix P such that $A = P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}P^{-1}$.

Now $tr(A) = tr\left[P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}P^{-1}\right] = tr\left[\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}P^{-1}P\right] = tr\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} = r =$ rank A.

We state below another result on idempotent matrices.

Theorem 11: A square matrix A of order $n\times n$ is idempotent if and only if rank $(I-A) = n -$ rank A.

Theorem 12: Let A be a real symmetric and idempotent matrix of rank r. Then there exists an orthogonal matrix S such that $A = S\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}S'$. Hence A is nonnegative definite.

Proof: Left as an exercise.

Let us now have an example.

Example 14: Let A and B be idempotent matrices of the same order. Then show that $A+B$ is idempotent if and only if $AB = BA = \mathbf{0}$.

Solution: 'If' part is trivial. For the 'only if' part, let A, B and $A+B$ be idempotent. Then $A+B = (A+B)(A+B) = A^2 + B^2 + AB + BA = A + B + AB + BA$. So, $AB + BA = \mathbf{0}$. Pre-multiplying by A, we get $AB + ABA = \mathbf{0}$. Now post multiplying the previous equality by A we get $ABA + ABA = \mathbf{0}$ or $ABA = \mathbf{0}$. Hence $AB = \mathbf{0}$ and as a consequence $BA = \mathbf{0}$.

***

Try the following Exercises.

E13) Let A be a 2×2 idempotent matrix. Can a_{11} be equal to 2?

E14) Let A and B be idempotent matrices. Then show that $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ is also idempotent.

E15) Show that if A and B are idempotent and the column space of A is contained in the column space of B, the $BA = A$.

In this section, we shall discuss Cochran's theorem.

14.5 COCHRAN'S THEOREM

Cochran's theorem concerns with the probability distributions of quadratic forms in independent standard normal variables. This is a very important theorem which allows us to decompose sum of squares into several quadratic forms and identify their distributions and establish their independence.

Let $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ be a vector of n independent standard normal variables. Let

$A_1, A_2, \ldots, A_k$ be real symmetric (nonrandom) matrices such that $\mathbf{x'x} = \mathbf{x}' A_1\mathbf{x} + \mathbf{x}'A_2\mathbf{x} + \cdots + \mathbf{x}' A_k\mathbf{x}$. We know that $\mathbf{x'x}$ is distributed as chi-square with n degrees of freedom. Cochran's theorem asserts that $x'A_i\, x, i = 1, \ldots, k$ are distributed as independent chi-squares if and only if $\sum_{i=1}^{n} \text{rank } A_i = n$. In this section, we prove an algebraic version of this result. In the next unit, we shall prove the statistical version.

Theorem 13: Let $A_1, A_2, \ldots, A_k$ be real symmetric matrices such that $A_1 + A_2 + \cdots + A_k = I$. The following are equivalent:

(a) A_i is idempotent, $i = 1, \ldots, k$

(b) $\sum_{i=1}^{k} \text{rank } A_i = n$

(c) $A_i\, A_j = \mathbf{0}$ whenever $i \neq j$

Proof: $(a) \Rightarrow (b): n = \text{rank } I = \text{tr}(I) = \text{tr}(A_1 + \cdots + A_k) = \sum_{i=1}^{k} \text{tr}(A_i) = \sum_{i=1}^{k} \text{rank } A_i$

(since $A_1, \ldots, A_k$ are idempotent rank $A_i = \text{tr}(A_i)$ by Theorem 10.)

$(b) \Rightarrow (c)$: Let rank $A_i = \mathbf{r}_i$. Then by (b), $\sum_{i=1}^{k} \mathbf{r}_i = n$. Since A_i is a real symmetric matrix there exists a matrix P_i of order $n \times r_i$ such that $A_i = P_i\Delta_i P_i'$, $P_i' P_i = I_{r_i}$, and Δ_i is a real nonsingular diagonal matrix, $i = 1, \ldots, k$. (Take the help of Theorem 3)

So, $I = A_1 + \cdots + A_k = \sum_{i=1}^{k} P_i\Delta_i P_i' = P\Delta P'$

where $P = (P_1 : P_2 : \ldots : P_k)$ and $\Delta = \begin{pmatrix} \Delta_1 & 0 & 0 & 0 \\ 0 & \Delta_2 & 0 & 0 \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \ldots & \Delta_k \end{pmatrix}$ Notice that the number of

columns in P is $\sum_{1}^{k} r_i$ which equals n by hypothesis. Hence P is a square matrix of order $n \times n$. So, $n \geq \text{rank } P \geq \text{rank } (P\Delta P') = \text{rank } I = n$. Hence P is a nonsingular matrix. Similarly, Δ is also nonsingular. Since $P\Delta P' = I$ and $P\Delta$ is a square matrix,

$P' = (P\Delta)^{-1}$. In other words, $P'P\Delta = I$ or $P'P = \Delta^{-1}$ which is a diagonal matrix. So $P_i'P_j = \mathbf{0}$ whenever $i \neq j$. Hence $A_i\ A_j = P_i\Delta_i P_i' P_j \Delta_j P_j' = \mathbf{0}$ whenever $i \neq j$.

$(c) \Rightarrow (a)$: For each i, $A_i = A_i\ I = A_i\,(A_1 + \cdots + A_k) = A_i^2$ since $A_i\ A_j = \mathbf{0}$ whenever $i \neq j$. Thus, A_i is idempotent for each i. Theorem is thus proved.

We now prove an algebraic version of another useful result in connection with the distribution of quadratic forms in normal variables.

Theorem 14: Let A and B be symmetric idempotent matrices and let $B - A$ be non negative definite. Then $B - A$ is also a symmetric idempotent matrix.

Proof: Since A is symmetric idempotent, it is nnd. Since $B - A$ is nnd, the column space of A is contained in the column space of B. So $BA = A$. Then $(B - A)A = 0 = A(B - A)$. Since $B(I - B) = (I - B)B = \mathbf{0}$, $A(I - B) = (I - B)A = \mathbf{0}$ and $(I - B)(B - A) = \mathbf{0} = (B - A)(I - B)$. Now $A + (B - A) + (I - B) = I$. By $(c) \Rightarrow (a)$ of Theorem 13 it follows that $B - A$ is idempotent.

After Cochran's theorem, we shall discuss the singular value decomposition in this section.

14.6 SINGULAR VALUE DECOMPOSITION

In Theorem 3, we showed that if A is a real symmetric matrix, there exists an orthogonal matrix P such that $A = P\Delta P'$. We also showed that the diagonal elements of Δ are the eigenvalues and the column of P are the orthonormal eigenvectors of A. What about a real $m \times n$ matrix A? We know that every $m \times n$ matrix A of rank $r(1 \leq r \leq \min\{m, n\})$ can be written as $A = R\begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}S$ when R and S are nonsingular.

Can we replace the nonsingular matrices by orthogonal matrices if we can relax I_r to a positive definite diagonal matrix? If so, what interpretation can we give to the orthogonal matrices and the diagonal elements of the diagonal matrix? We study these details in this section.

Theorem 15: Let A be a real matrix of order $m \times n$ with rank $r(1 \leq r \leq \min\{m, n\})$. Then there exists orthogonal matrices U and V of orders $m \times m$ and $n \times n$ respectively such that $A = U\begin{pmatrix} \Delta & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}V'$ where Δ is a positive definite diagonal matrix.

Proof: Notice AA' and $A'A$ are nonnegative definite matrices (Why? See Theorem 4). Let $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ be orthogonal eigenvectors of AA' corresponding to the eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r > \lambda_{r+1} = \ldots = \lambda_m = 0$. So $AA'\mathbf{u}_i = \lambda_i \mathbf{u}_i$, $i = 1, \ldots, m$. Write $\mathbf{v}_i = \frac{1}{\sqrt{\lambda_i}} A'u_i$, $i = 1, \ldots, r$, where $\sqrt{\lambda_i}$ is the positive square root of λ_i. Then for $i, j = 1, \ldots, r$

$$\mathbf{v}_i'\mathbf{v}_j = \frac{1}{\sqrt{\lambda_i}\sqrt{\lambda_j}} u_i' AA' u_j = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$$

Thus, $\mathbf{v}_1,\ldots,\mathbf{v}_r$ are orthonormal vectors. Extend $\mathbf{v}_1,\ldots,\mathbf{v}_r$ to an orthonormal basis $\mathbf{v}_1,\ldots,\mathbf{v}_n$ of R^n. Write $U=(\mathbf{u}_1:\ldots:\mathbf{u}_m)$ and $V=(\mathbf{v}_1:\ldots:\mathbf{v}_n)$. Clearly U and V are orthogonal matrices.
Also $AA'\mathbf{u}_i=\mathbf{0}$ for $i=r+1,\ldots,m$. Hence $u_i'AA'u_i=0$ or $A'u_i=0$ for $i=r+1,\ldots,m$.

Since $\mathbf{u}_1\mathbf{u}_1'+\ldots+\mathbf{u}_m\mathbf{u}_m'=I$, we have

$$\begin{aligned} A &= (\mathbf{u}_1\mathbf{u}_1'+\ldots+\mathbf{u}_m\mathbf{u}_m')A \\ &= (\mathbf{u}_1\mathbf{u}_1'+\ldots+\mathbf{u}_r\mathbf{u}_r')A \text{ since } \mathbf{u}_i'A=\mathbf{0} \text{ for } i=r+1,\ldots,m. \\ &= \sum_{i=1}^{r}\sqrt{\lambda_i}\,\mathbf{u}_i\mathbf{v}_i' \end{aligned}$$

Denote $\delta_i=\sqrt{\lambda_i}$, $i=1,\ldots,r$ and $\Delta=\mathrm{diag}(\delta_1,\delta_2,\ldots,\delta_r$

It follows that $A=\sum_{i=1}^{r}\sqrt{\lambda_i}\,\mathbf{u}_i\mathbf{v}_i'=U\begin{pmatrix}\Delta & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}V'$. The proof is complete.

We shall now interpret the columns of U and V and the diagonal elements of Δ in the above form.

Let $A=U\begin{pmatrix}\Delta & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}V'$, where U and V are orthogonal and Δ is a positive definite diagonal matrix.

Then rank of A is the same as rank of Δ which in turn is the number of rows in Δ.
Now

$$AA'=U\begin{pmatrix}\Delta & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\begin{pmatrix}\Delta & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}'V'=U\begin{pmatrix}\Delta^2 & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}V'.$$

Thus, $U\begin{pmatrix}\Delta^2 & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}V'$ is a spectral decomposition of AA'. Hence the diagonal elements of Δ^2 are the nonzero eigenvalues of AA' and the columns of U are the orthonormal eigenvectors of $\mathbf{AA}'$. To be more specific

$$AA'u_i=\begin{cases}\delta_i^2\mathbf{u}_i \text{ for } i=1,\ldots,r\\ \mathbf{0} \text{ for } i=r+1,\ldots,m\end{cases}$$

Again $A'A=V\begin{pmatrix}\Delta^2 & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}U'$, which is a spectral decomposition of AA'. Hence

$$A'Av_i=\begin{cases}\delta_i^2\mathbf{v}_i \text{ for } i=1,\ldots,r\\ \mathbf{0} \text{ for } i=r+1,\ldots,n\end{cases}$$

Thus, the diagonal elements of Δ and the columns of U and V relate to the eigenvalues and eigenvectors of AA' and $A'A$. *The diagonal elements of* Δ *are called the singular values and the columns of* U *and* V *are called the singular vectors of* A. *The decomposition* $A=U\begin{pmatrix}\Delta & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}V'$ *is called the singular value decomposition of* A.

Example 15: Let $A = \frac{1}{3}\begin{pmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ -2 & -2 & 1 \end{pmatrix}\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\frac{1}{2}\begin{pmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ -1 & -1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{pmatrix}$ be a singular value decomposition of A. What are the eigenvalues of AA' and $A'A$? Identify the corresponding eigenvectors. What is the rank of A?

Solution: $A = U\begin{pmatrix} \Delta & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}V'$

where $U = \frac{1}{3}\begin{pmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ -2 & -2 & 1 \end{pmatrix}$ and $V = \frac{1}{2}\begin{pmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ -1 & -1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{pmatrix}$ are orthogonal matrices,

and $\Delta = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$

The eigenvalues of AA' are $\delta_1^2 = 4$, $\delta_2^2 = 1$ and 0. The corresponding eigenvectors are the first, second and third columns respectively of U, namely $\frac{1}{3}\begin{pmatrix} 1 \\ -2 \\ -2 \end{pmatrix}, \frac{1}{3}\begin{pmatrix} -2 \\ 1 \\ -2 \end{pmatrix}$, and $\frac{1}{3}\begin{pmatrix} -2 \\ -2 \\ 1 \end{pmatrix}$ respectively.

***

Let us now sum up whatever we have studied in this unit.

14.7 SUMMARY

In this unit, we have covered the following points:

1. Definition of a real symmetric matrix
2. Classification of quadratic forms
3. Spectral decomposition of a real symmetric matrix
4. A method of determining the definiteness of a quadratic form
5. Properties of positive definite and nonnegative definite matrices
6. A method of finding a triangular square root of a pd matrix
7. Properties of idempotent matrices
8. Cochran's Theorem
9. Singular Value Decomposition.

14.8 SOLUTIONS/ANSWERS

E1. i) Coefficient of x_1^2, x_2^2, and x_1x_2 are respectively $1, -1$ and 0. So the matrix of the quadratic form $x_1^2 - x_2^2$ is $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$.

ii) Coefficients of x_1^2, x_2^2, and x_1x_2 are respectively $2, 5$ and 3. So the matrix of the quadratic form is $\begin{bmatrix} 2 & 1.5 \\ 1.5 & 5 \end{bmatrix}$.

iii) Coefficients of $x_1^2, x_2^2, x_3^2, x_1x_2, x_1x_3$ and x_2x_3 are respectively $0, 0, 0, 3, -4, 5$. So the matrix of the quadratic form $3x_1x_2 + 5x_2x_3 - 4x_1x_3$ is

$$\begin{bmatrix} 0 & 1.5 & -2 \\ 1.5 & 0 & 2.5 \\ -2 & 2.5 & 0 \end{bmatrix}.$$

iv) Coefficients of $x_1^2, x_2^2, x_3^2, x_4^2, x_1x_2, x_1x_3, x_1x_4, x_2x_3, x_2x_4$ and x_3x_4 are respectively $1, 1, 0, 0, 0, 0, 0, 0, 1, 0$. So the matrix of the quadratic form $x_1^2 + x_2^2 + x_2x_4$ is

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0.5 \\ 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

E2. Suppose $a_{ii} < 0$. Let $\mathbf{e}_i$ denotes the i^{th} column of the identity matrix. Then $\mathbf{e}_i' \ \mathbf{A}\mathbf{e}_i = a_{ii} < 0$. Hence A cannot be pd or psd.

E3. i) The matrix of the quadratic form $x_1^2 - 5x_1x_2 + 7x_2^2$ is $\mathbf{A} = \begin{bmatrix} 1 & -2.5 \\ -2.5 & 7 \end{bmatrix}$.

The eigenvalues of A are the roots of the characteristic equation $|A - \lambda I| = 0$ or $(1-\lambda)(7-\lambda) - 6.25 = 0$.

The Characteristic equation can be rewritten as $\lambda^2 - 8\lambda + 0.75 = 0$.

Hence the eigenvalues are $\dfrac{8+\sqrt{64-3}}{2}$ and $\dfrac{8-\sqrt{64-3}}{2}$ which are both positive. Hence the quadratic form $x_1^2 - 5x_1x_2 + 7x_2^2$ is positive definite.

ii) For $x_1 = 1$ and $x_2 = x_3 = 0$, the value of the quadratic form is 1. Again for $x_2 = 1, x_1 = x_3 = 0$, the value of the quadratic form is -1. Hence the quadratic form is indefinite.

iii) The matrix of the quadratic form is $A = \begin{bmatrix} 2 & 3 & 0 \\ 3 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix}$. The characteristic equation of A is $(4-\lambda)((2-\lambda)(3-\lambda) - 9) = 0$. So 4 is a root of the above equation. The remaining two eigenvalues are the roots of the equation $(2-\lambda)(3-\lambda) - 9 = 0$ or $\lambda^2 - 5\lambda - 3 = 0$. So the eigenvalues are $\dfrac{5 \pm \sqrt{25-12}}{2}$ and 4. Thus, all the three roots are positive. Hence the given quadratic form is positive definite.

E4. The eigenvalues of $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ are the roots of the characteristic equation $(2-\lambda)^2 - 1 = 0$ or $\lambda^2 - 4\lambda + 3 = 0$ or $(\lambda - 3)(\lambda - 1) = 0$. So the eigenvalues are $\lambda_1 = 3$ and $\lambda_2 = 1$. Let $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ be an eigenvector corresponding to λ_1. Then

$$2x_1 + x_2 = 3x_1$$
$$x_1 + 2x_2 = 3x_2$$

or $-x_1 + x_2 = 0$
$$x_1 - x_2 = 0$$

Thus, $x_1 = x_2$. So the normalized eigenvector corresponding to the eigenvalue 3 is $\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$.

It can similarly be shown that $\frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}$ is the normalized eigenvector corresponding to the eigenvalue 1 orthogonal to the first eigenvector. So the spectral decomposition of A is

$$A = 3.\frac{1}{2}\begin{pmatrix}1\\1\end{pmatrix}(1 \quad 1) + 1.\frac{1}{2}\begin{pmatrix}1\\-1\end{pmatrix}(1 \quad -1)$$

or $\frac{1}{2}\begin{pmatrix}1 & 1\\1 & -1\end{pmatrix}\begin{pmatrix}3 & 0\\0 & 1\end{pmatrix}\left(\frac{1}{2}\begin{pmatrix}1 & 1\\1 & -1\end{pmatrix}\right)$

If $A_{n\times m} = P\Lambda P'$ is a spectral decomposition of A, then
$A^2 = P\Lambda P'P\Lambda P' = P\Lambda^2P'$.

By induction it can be shown that $A^k = P\Lambda^kP'$ for $k = 1, 2, \ldots$

Thus, if $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A, then $\lambda_1^k, \ldots, \lambda_n^k$ are the eigenvalues of A^k. The eigenvectors of A^k can be taken to be the same as the eigenvectors of A.

Hence $A^{100} = \frac{1}{2}\begin{pmatrix}1 & 1\\1 & -1\end{pmatrix}\begin{pmatrix}3^{100} & 0\\0 & 1\end{pmatrix}\left(\frac{1}{2}\begin{pmatrix}1 & 1\\1 & -1\end{pmatrix}\right)$.

E5. i) Let $\begin{pmatrix}4 & 1\\1 & 2\end{pmatrix} = \begin{pmatrix}b_{11} & 0\\b_{21} & b_{22}\end{pmatrix}\begin{pmatrix}b_{11} & b_{21}\\0 & b_{22}\end{pmatrix}$

So $b_{11}^2 = 4$ or $b_{11} = 2$

$b_{21}\, b_{11} = 1$ or $b_{21} = \frac{1}{b_{11}} = \frac{1}{2}$

$b_{21}^2 + b_{22}^2 = 2$ or $b_{22}^2 = 2 - \frac{1}{4} = \frac{7}{4}$

so $b_{22} = \frac{\sqrt{7}}{2}$

Thus, the required lower triangular square root is $\begin{pmatrix}2 & 0\\ \frac{1}{2} & \frac{\sqrt{7}}{2}\end{pmatrix}$.

ii) Let $\begin{pmatrix}9 & 3 & 3\\3 & 5 & 1\\3 & 1 & 6\end{pmatrix} = \begin{pmatrix}b_{11} & 0 & 0\\b_{21} & b_{22} & 0\\b_{31} & b_{32} & b_{33}\end{pmatrix}\begin{pmatrix}b_{11} & b_{21} & b_{31}\\0 & b_{22} & b_{32}\\0 & 0 & b_{33}\end{pmatrix}$

so $b_{11}^2 = 9$ or $b_{11} = 3$

$b_{11}\, b_{21} = 3$ or $b_{21} = 1$

$b_{11}\, b_{31} = 3$ or $b_{31} = 1$

$b_{21}^2 + b_{22}^2 = 5$ or $b_{22}^2 = 5 - 1$ or $b_{22} = 2$

$b_{21}\, b_{31} + b_{22}\, b_{32} = 1$ or $b_{22}\, b_{32} = 1 - 1 = 0$ or $b_{32} = 0$

$b_{31}^2 + b_{32}^2 + b_{33}^2 = 6$ or $b_{33}^2 = 6 - 0 - 1 = 5$ or $b_{33} = \sqrt{5}$.

Thus, the required triangular square root is $\begin{pmatrix} 3 & 0 & 0 \\ 1 & 2 & 0 \\ 1 & 0 & \sqrt{5} \end{pmatrix}$.

E6. 'If' part: let C and D be positive definite. Then

$(\mathbf{x}' : \mathbf{y}')\begin{pmatrix} C & \mathbf{0} \\ \mathbf{0} & D \end{pmatrix}\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \mathbf{x}'C\mathbf{x} + \mathbf{y}'D\mathbf{y} > \mathbf{0}$ whenever at least one of $\mathbf{x}$ and $\mathbf{y}$ is non-null. Hence $\begin{pmatrix} C & \mathbf{0} \\ \mathbf{0} & D \end{pmatrix}$ is pd.

'Only if' part: Let $\begin{pmatrix} C & \mathbf{0} \\ \mathbf{0} & D \end{pmatrix}$ be pd consider $\mathbf{0} < (\mathbf{x}' : \mathbf{0})\begin{pmatrix} C & \mathbf{0} \\ \mathbf{0} & D \end{pmatrix}\begin{pmatrix} \mathbf{x} \\ \mathbf{0} \end{pmatrix} = \mathbf{x}'C\mathbf{x}$ whenever $\mathbf{x} \neq \mathbf{0}$. Hence C is pd.

E7. We know that $\mathbf{0} < (\mathbf{0} : \mathbf{y}')\begin{pmatrix} A_{11} & A_{12} \\ A_{12}' & A_{22} \end{pmatrix}\begin{pmatrix} \mathbf{0} \\ \mathbf{y} \end{pmatrix} = \mathbf{y}'A_{22}\mathbf{y}$ whenever $\mathbf{y} \neq \mathbf{0}$ ($\mathbf{y}$ is chosen so that $\mathbf{y}'A\mathbf{y}$ is conformable). Hence A_{22} is pd.

E8. Use the same procedure as in E7.

E9. Let A be an nnd matrix. Then there exists B such that $A = BB'$. Hence $\mathbf{0} = \mathbf{x}'A\mathbf{x} = \mathbf{x}'BB'\mathbf{x}$ implies that $B'\mathbf{x} = \mathbf{0}$. So $A\mathbf{x} = BB'\mathbf{x} = \mathbf{0}$.

E10. Nonzero eigenvalues of AA' and $A'A$ are the same. Hence nonzero eigenvalues of $\rho 1 1'$ are the same as the eigenvalues of the 1×1 matrix $\rho 1' 1 = n\rho$. The eigenvalues of $\rho 1' 1 = n\rho$. So, the eigenvalues of $\rho 1 1'$ are $n\rho$ and $[0, 0, \ldots, 0\ (n-1)$ times].

Let λ be an eigenvalue of $\rho I I^t$. Let the corresponding eigenvector be $\mathbf{x}$. Then $(1-\rho)1 + \rho 1 1')x = (1-\rho)x + \lambda x = (1 - \rho + \lambda)x$.

Thus, the eigenvalues of $(1-\rho)1 + \rho 1 1'$ are $(1-\rho) + n\rho, (1-\rho), \ldots, (1-\rho)$ ($(n-1)$ times).

Hence $(1-\rho)1 + \rho 1 1'$ is pd if and only if $1 + (n-1)\,\rho > 0$ and $1 - \rho > 0$ or $-\dfrac{1}{n-1} < \rho < 1$.

E11. We use Theorem 7 for this purpose. Thus, $\begin{pmatrix} 2 & 1 & 0.5 \\ 1 & 5 & 3 \\ 0.5 & 3 & 4 \end{pmatrix}$ is a pd matrix with diagonal elements 2, 5 and 4, respectively.

E12. i) $\mathbf{x}'(A+B)\mathbf{x} = \mathbf{x}'A\mathbf{x} + \mathbf{x}'B\mathbf{x} \geq 0$ since $\mathbf{x}'A\mathbf{x}$ and $\mathbf{x}'B\mathbf{x}$ are nonnegative. Hence $A + B$ is nnd.

ii) Write $A = CC'$ and $B = DD'$. Hence $A + B = CC' + DD' = (C : D)\begin{pmatrix} C' \\ D' \end{pmatrix}$.

Hence column space of A = Column space of $C \subseteq$ column space of $(C:D)$ = column space of $A+B$.

E13. Let $\begin{pmatrix} 2 & b \\ c & d \end{pmatrix}$ be an idempotent matrix. Then $\begin{pmatrix} 2 & b \\ c & d \end{pmatrix}\begin{pmatrix} 2 & b \\ c & d \end{pmatrix} = \begin{pmatrix} 2 & b \\ c & d \end{pmatrix}$

$4+bc=2$. b cannot be 0 since $4+bc=2$. Also $bc=-2$

$2b+bd=b$. So $2+d=1$ or $d=-1$

$2c+cd=c$

$bc+d^2=d$

$bc+1=-1$

or $bc=-2$. Choose $b=-2$ and $c=1$

(in fact choose b any nonzero number and $c=\frac{-2}{b}$)

Thus, $\begin{pmatrix} 2 & -2 \\ 1 & -1 \end{pmatrix}$ is idempotent. Hence there is an idempotent matrix with $a_{11}=2$.

E14. $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = \begin{pmatrix} A^2 & 0 \\ 0 & B^2 \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$

So $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ is idempotent if A and B are idempotent.

E15. Since the column space of A is contained in the column space of B, we get $A = BD$ for some D.

Now $BA = B.BD = BD = A$.

—x—