7 Eigenvalues and Eigenvectors

7.1 Introduction

The simplest of matrices are the diagonal ones. Thus a linear map will be also easy to handle if its associated matrix is a diagonal matrix. Then again we have seen that the matrix associated depends upon the choice of the bases to some extent. This naturally leads us to the problem of investigating the existence and construction of a suitable basis with respect to which the matrix associated to a given linear transformation is diagonal.

Definition 7.1 *A $n \times n$ matrix A is called* diagonalizable *if there exists an invertible $n \times n$ matrix M such that $M^{-1}AM$ is a diagonal matrix. A linear map $f : V \longrightarrow V$ is called diagonalizable if the matrix associated to f with respect to some basis is diagonal.*

Remark 7.1
(i) Clearly, f is diagonalizable iff the matrix associated to f with respect to some basis (any basis) is diagonalizable.
(ii) Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis. The matrix M_f of a linear transformation f w.r.t. this basis is diagonal iff $f(\mathbf{v}_i) = \lambda_i \mathbf{v}_i,\ 1 \leq i \leq n$ for some scalars λ_i. Naturally a subquestion here is: does there exist such a basis for a given linear transformation?

Definition 7.2 *Given a linear map $f : V \longrightarrow V$ we say $\mathbf{v} \in V$ is an* eigenvector *for f if $\mathbf{v} \neq 0$ and $f(\mathbf{v}) = \lambda \mathbf{v}$ for some $\lambda \in \mathbb{K}$. In that case λ is called as* eigenvalue *of f. For a square matrix A we say λ is an eigenvalue if there exists a non zero column vector $\mathbf{v}$ such that $A\mathbf{v} = \lambda \mathbf{v}$. Of course $\mathbf{v}$ is then called the eigenvector of A corresponding to λ.*

Remark 7.2
(i) It is easy to see that eigenvalues and eigenvectors of a linear transformation are same as those of the associated matrix.
(ii) Even if a linear map is not diagonalizable, the existence of eigenvectors and eigenvalues itself throws some light on the nature of the linear map. Thus the study of eigenvalues becomes extremely important. They arise naturally in the study of differential equations. Here we shall use them to address the problem of diagonalization and then see some geometric applications of diagonalization itself.

7.2 Characteristic Polynomial

Proposition 7.1
(1) Eigenvalues of a square matrix A are solutions of the equation

$$\chi_A(\lambda) = \det(A - \lambda I) = 0.$$

(2)The null space of $A - \lambda I$ is equal to the eigenspace

$$E_A(\lambda) := \{\mathbf{v} \ : \ A\mathbf{v} = \lambda \mathbf{v}\} = \mathcal{N}(A - \lambda I).$$

Proof: (1) If $\mathbf{v}$ is an eigenvector of A then $\mathbf{v} \neq 0$ and $A\mathbf{v} = \lambda \mathbf{v}$ for some scalar λ. Hence $(A - \lambda I)\mathbf{v} = 0$. Thus the nullity of $A - \lambda I$ is positive. Hence rank$(A - \lambda I)$ is less than n. Hence $\det(A - \lambda I) = 0$.
(2) $E_A(\lambda) = \{\mathbf{v} \in V \ : \ A\mathbf{v} = \lambda \mathbf{v}\} = \{\mathbf{v} \in V \ : \ (A - \lambda I)\mathbf{v} = 0\} = \mathcal{N}(A - \lambda I)$. ♠

Definition 7.3 *For any square matrix A, the polynomial $\chi_A(\lambda) = \det(A - \lambda I)$ in λ is called the* **characteristic polynomial** *of A.*

Example 7.1

(1) $A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$. To find the eigenvalues of A, we solve the equation

$$\det(A - \lambda I) = \det \begin{bmatrix} 1-\lambda & 2 \\ 0 & 3-\lambda \end{bmatrix} = (1-\lambda)(3-\lambda) = 0.$$

Hence the eigenvalues of A are 1 and 3. Let us calculate the eigenspaces $E(1)$ and $E(3)$. By definition

$$E(1) = \{\mathbf{v} \mid (A - I)\mathbf{v} = 0\} \text{ and } E(3) = \{\mathbf{v} \mid (A - 3I)\mathbf{v} = 0\}.$$

$A - I = \begin{bmatrix} 0 & 2 \\ 0 & 2 \end{bmatrix}$. *Hence $(x, y)^t \in E(1)$ iff* $\begin{bmatrix} 0 & 2 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2y \\ 2y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. *Hence* $E(1) = L\{(1, 0)\}$.

$$A - 3I = \begin{bmatrix} 1-3 & 2 \\ 0 & 3-3 \end{bmatrix} = \begin{bmatrix} -2 & 2 \\ 0 & 0 \end{bmatrix}. \text{ Suppose } \begin{bmatrix} -2 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Then $\begin{bmatrix} -2x + 2y \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. *This is possible iff $x = y$. Thus $E(3) = L(\{(1, 1)\})$.*

(2) Let $A = \begin{bmatrix} 3 & 0 & 0 \\ -2 & 4 & 2 \\ -2 & 1 & 5 \end{bmatrix}$. *Then* $\det(A - \lambda I) = (3 - \lambda)^2(6 - \lambda)$.

Hence eigenvalues of A are 3 and 6. The eigenvalue $\lambda = 3$ is a double root of the characteristic polynomial of A. We say that $\lambda = 3$ has **algebraic multiplicity** *2. Let us find the eigenspaces $E(3)$ and $E(6)$.*

$\underline{\lambda = 3}$: $A - 3I = \begin{bmatrix} 0 & 0 & 0 \\ -2 & 1 & 2 \\ -2 & 1 & 2 \end{bmatrix}$. *Hence rank* $(A - 3I) = 1$. *Thus nullity* $(A - 3I) = 2$. *By solving the system $(A - 3I)\mathbf{v} = 0$, we find that*

$$\mathcal{N}(A - 3I) = E_A(3) = L(\{(1, 0, 1), (1, 2, 0)\}).$$

The dimension of $E_A(\lambda)$ is called the **geometric multiplicity** *of λ. Hence geometric multiplicity of $\lambda = 3$ is 2.*

$\underline{\lambda = 6}$: $A - 6I = \begin{bmatrix} -3 & 0 & 0 \\ -2 & -2 & 2 \\ -2 & 1 & -1 \end{bmatrix}$. *Hence rank*$(A - 6I) = 2$. *Thus* $\dim E_A(6) = 1$. *(It can be shown that $\{(0, 1, 1)\}$ is a basis of $E_A(6)$.) Thus both the algebraic and geometric multiplicities of the eigenvalue 6 are equal to 1.*

(3) $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. *Then* $\det(A - \lambda I) = (1 - \lambda)^2$. *Thus $\lambda = 1$ has algebraic multiplicity 2.*

$A - I = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. *Hence nullity* $(A - I) = 1$ *and* $E_A(1) = L\{e_1\}$. *In this case the geometric multiplicity is less than the algebraic multiplicity of the eigenvalue* 1.

Remark 7.3
(i) Observe that $\chi_A(\lambda) = \chi_{M^{-1}AM}(\lambda)$. *Thus the characteristic polynomial is an invariant of similarity. Thus the characteristic polynomial of any linear map* $f : V \longrightarrow V$ *is also defined (where* V *is finite dimensional) by choosing some basis for* V, *and then taking the characteristic polynomial of the associated matrix* $\mathcal{M}(f)$ *of* f. *This definition does not depend upon the choice of the basis.*
(ii) If we expand $\det(A - \lambda I)$ *we see that there is a term*

$$(a_{11} - \lambda)(a_{22} - \lambda)\cdots(a_{nn} - \lambda).$$

This is the only term which contributes to λ^n *and* λ^{n-1}. *It follows that the degree of the characteristic polynomial is exactly equal to* n, *the size of the matrix; moreover, the coefficient of the top degree term is equal to* $(-1)^n$. *Thus in general, it has* n *complex roots, some of which may be repeated, some of them real, and so on. All these patterns are going to influence the geometry of the linear map.*
(iii) If A *is a real matrix then of course* $\chi_A(\lambda)$ *is a real polynomial. That however, does not allow us to conclude that it has real roots. So while discussing eigenvalues we should consider even a real matrix as a complex matrix and keep in mind the associated linear map* $\mathbb{C}^n \longrightarrow \mathbb{C}^n$. *The problem of existence of real eigenvalues and real eigenvectors will be discussed soon.*
(iv) Next, the above observation also shows that the coefficient of λ^{n-1} *is equal to*

$$(-1)^{n-1}(a_{11} + \cdots + a_{nn}) = (-1)^{n-1} tr\, A.$$

Lemma 7.1 Suppose A is a real matrix with a real eigenvalue λ. Then there exists a real column vector $\mathbf{v} \neq 0$ such that $A\mathbf{v} = \lambda\mathbf{v}$.

Proof: Start with $A\mathbf{w} = \lambda\mathbf{w}$ where $\mathbf{w}$ is a non zero column vector with complex entries. Write $\mathbf{w} = \mathbf{v} + \imath\mathbf{v}'$ where both $\mathbf{v}, \mathbf{v}'$ are real vectors. We then have

$$A\mathbf{v} + \imath A\mathbf{v}' = \lambda(\mathbf{v} + \imath\mathbf{v}')$$

Compare the real and imaginary parts. Since $\mathbf{w} \neq 0$, at least one of the two $\mathbf{v}, \mathbf{v}'$ must be a non zero vector and we are done. ♠

Proposition 7.2 *Let* A *be an* $n \times n$ *matrix with eigenvalues* $\lambda_1, \lambda_2, \ldots, \lambda_n$. *Then*
(i) $tr\,(A) = \lambda_1 + \lambda_2 + \ldots + \lambda_n$.
(ii) $\det A = \lambda_1\lambda_2\ldots\lambda_n$.

Proof: The characteristic polynomial of A is

$$\det(A - \lambda I) = \det \begin{bmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{bmatrix}$$

$$(-1)^n\lambda^n + (-1)^{n-1}\lambda^{n-1}(a_{11} + \ldots + a_{nn}) + \ldots \tag{48}$$

Put $\lambda = 0$ to get $\det A =$ constant term of $\det(A - \lambda I)$.
Since $\lambda_1, \lambda_2, \ldots, \lambda_n$ are roots of $\det(A - \lambda I) = 0$ we have

$$\det(A - \lambda I) \;=\; (-1)^n(\lambda - \lambda_1)(\lambda - \lambda_2)\ldots(\lambda - \lambda_n). \tag{49}$$

$$\tag{50}$$

$$(-1)^n[\lambda^n - (\lambda_1 + \lambda_2 + \ldots + \lambda_n)\lambda^{n-1} + \ldots + (-1)^n\lambda_1\lambda_2\ldots\lambda_n]. \tag{51}$$

Comparing (49) and 51), we get, the constant term of $\det(A - \lambda I)$ is equal to $\lambda_1\lambda_2\ldots\lambda_n = \det A$ and $tr(A) = a_{11} + a_{22} + \ldots + a_{nn} = \lambda_1 + \lambda_2 + \ldots + \lambda_n$. ♠

Proposition 7.3 *Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ be eigenvectors of a matrix A associated to distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$. Then $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are linearly independent.*

Proof: Apply induction on k. It is clear for $k = 1$. Suppose $k \geq 2$ and $c_1\mathbf{v}_1 + \ldots + c_k\mathbf{v}_k = 0$ for some scalars $c_1, c_2, \ldots, c_k$. Hence $c_1A\mathbf{v}_1 + c_2A\mathbf{v}_2 + \ldots + c_kA\mathbf{v}_k = 0$
Hence

$$c_1\lambda_1\mathbf{v}_1 + c_2\lambda_2\mathbf{v}_2 + \ldots + c_k\lambda_k\mathbf{v}_k = 0$$

Hence

$$\begin{aligned}&\lambda_1(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_k\mathbf{v}_k) - (\lambda_1c_1\mathbf{v}_1 + \lambda_2c_2\mathbf{v}_2 + \ldots + \lambda_kc_k\mathbf{v}_k)\\ &= (\lambda_1 - \lambda_2)c_2\mathbf{v}_2 + (\lambda_1 - \lambda_3)c_3\mathbf{v}_3 + \ldots + (\lambda_1 - \lambda_k)c_k\mathbf{v}_k = 0\end{aligned}$$

By induction, $\mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_k$ are linearly independent. Hence $(\lambda_1 - \lambda_j)c_j = 0$ for $j = 2, 3, \ldots, k$. Since $\lambda_1 \neq \lambda_j$ for $j = 2, 3, \ldots, k$, $c_j = 0$ for $j = 2, 3, \ldots, k$. Hence c_1 is also zero. Thus $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are linearly independent. ♠

Proposition 7.4 *Suppose A is an $n \times n$ matrix. Let A have n distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Let C be the matrix whose column vectors are respectively $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ where $\mathbf{v}_i$ is an eigenvector for λ_i for $i = 1, 2, \ldots, n$. Then*

$$C^{-1}AC = D(\lambda_1, \ldots, \lambda_n) = D$$

the diagonal matrix.

Proof: It is enough to prove $AC = CD$. For $i = 1, 2, \ldots, n$: let $C^i(= \mathbf{v}_i)$ denote the i^{th} column of C etc.. Then

$$(AC)^i = AC^i = A\mathbf{v}_i = \lambda_i\mathbf{v}_i.$$

Similarly,

$$(CD)^i = CD^i = \lambda_i\mathbf{v}_i.$$

Hence $AC = CD$ as required.] ♠

7.3 Relation Between Algebraic and Geometric Multiplicities

Recall that

Definition 7.4 *The* algebraic multiplicity $a_A(\mu)$ *of an eigenvalue* μ *of a matrix* A *is defined to be the multiplicity* k *of the root* μ *of the polynomial* $\chi_A(\lambda)$. *This means that* $(\lambda-\mu)^k$ *divides* $\chi_A(\lambda)$ *whereas* $(\lambda-\mu)^{k+1}$ *does not.*

Definition 7.5 *The* geometric multiplicity *of an eigenvalue* μ *of* A *is defined to be the dimension of the eigenspace* $E_A(\lambda)$*;*

$$g_A(\lambda) := dim\, E_A(\lambda).$$

Proposition 7.5 *Both algebraic multiplicity and the geometric multiplicities are invariant of similarity.*

Proof: We have already seen that for any invertible matrix C, $\chi_A(\lambda) = \chi_{C^{-1}AC}(\lambda)$. Thus the invariance of algebraic multiplicity is clear. On the other hand check that $E_{C^{-1}AC}(\lambda) = C(E_A(\lambda))$. Therefore, $dim\,(E_{C^{-1}AC}(\lambda)) = dim\, C(E_A\lambda)) = dim\,(E_A(\lambda))$, the last equality being the consequence of invertibility of C.

♠

We have observed in a few examples that the geometric multiplicity of an eigenvalue is at most its algebraic multiplicity. This is true in general.

Proposition 7.6 *Let* A *be an* $n\times n$ *matrix. Then the geometric multiplicity of an eigenvalue* μ *of* A *is less than or equal to the algebraic multiplicity of* μ.

Proof: Put $a_A(\mu) = k$. Then $(\lambda-\mu)^k$ divides $\det(A-\lambda I)$ but $(\lambda-\mu)^{k+1}$ does not. Let $g_A(\mu) = g$, be the geometric multiplicity of μ. Then $E_A(\mu)$ has a basis consisting of g eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_g$. We can extend this basis of $E_A(\mu)$ to a basis of $\mathbb{C}^n$, say $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_g, \ldots, \mathbf{v}_n\}$. Let B be the matrix such that $B^j = \mathbf{v}_j$. Then B is an invertible matrix and

$$B^{-1}AB = \left[\begin{array}{c|c} \mu I_g & X \\ \hline 0 & Y \end{array}\right]$$

where X is a $g\times(n-g)$ matrix and Y is an $(n-g)\times(n-g)$ matrix. Therefore,

$$\begin{aligned} \det(A-\lambda I) = \det[B^{-1}(A-\lambda I)B] &= \det(B^{-1}AB-\lambda I) \\ &= (\det(\mu-\lambda)I_g)(\det(C-\lambda I_{n-g}) \\ &= (\mu-\lambda)^g \det(Y-\lambda I_{n-g}). \end{aligned}$$

Thus $g \le k$. ♠

Remark 7.4 *We will now be able to say something about the diagonalizability of a given matrix* A. *Assuming that there exists* B *such that* $B^{-1}AB = D(\lambda_1, \ldots, \lambda_n)$, *as seen in the previous proposition, it follows that* $AB = BD \ldots$ *etc..* $AB^i = \lambda B^i$ *where* B^i *denotes the* i^{th} *column vector of* B. *Thus we need not hunt for* B *anywhere but look for eigenvectors of* A. *Of course* B^i *are linearly independent, since* B *is invertible. Now the problem turns to* *the question whether we have n linearly independent eigenvectors of A so that they can be chosen for the columns of B. The previous proposition took care of one such case, viz., when the eigenvalues are distinct. In general, this condition is not forced on us. Observe that the geometric multiplicity and algebraic multiplicity of an eigenvalue co-incide for a diagonal matrix. Since these concepts are similarity invariants, it is necessary that the same is true for any matrix which is diagonalizable. This turns out to be sufficient also. The following theorem gives the correct condition for diagonalization.*

Theorem 7.1 *A $n \times n$ matrix A is diagonalizable if and only if for each eigenvalue μ of A we have the algebraic and geometric multiplicities are equal: $a_A(\mu) = g_A(\mu)$.*

Proof: We have already seen the necessity of the condition. To prove the converse, suppose that the two multiplicities coincide for each eigenvalue. Suppose that $\lambda_1, \lambda_2, \ldots, \lambda_k$ are all the eigenvalues of A with algebraic multiplicities $n_1, n_2, \ldots, n_k$. Let

$$\begin{aligned} B_1 = \{\mathbf{v}_{11}, \mathbf{v}_{12}, \ldots, \mathbf{v}_{1n_1}\} &= \text{a basis of } E(\lambda_1), \\ B_2 = \{\mathbf{v}_{21}, \mathbf{v}_{22}, \ldots, \mathbf{v}_{2n_2}\} &= \text{a basis of } E(\lambda_2), \\ &\vdots \\ B_k = \{\mathbf{v}_{k1}, \mathbf{v}_{k2}, \ldots, \mathbf{v}_{kn_k}\} &= \text{a basis of } E(\lambda_k). \end{aligned}$$

Use induction on k to show that $B = B_1 \cup B_2 \cup \ldots \cup B_k$ is a linearly independent set. (The proof is exactly similar to the proof of proposition (7.3). Denote the matrix with columns as elements of the basis B also by B itself. Then, check that $B^{-1}AB$ is a diagonal matrix. Hence A is diagonalizable. ♠

7.4 Eigenvalues of Special Matrices

In this section we discuss eigenvalues of special matrices. We will work in the n-dimensional complex vector space $\mathbb{C}^n$. If $\mathbf{u} = (u_1, u_2, \ldots, u_n)^t$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)^t \in \mathbb{C}^n$, we have defined their inner product in $\mathbb{C}^n$ by

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^* \mathbf{v} = \overline{u}_1 v_1 + \overline{u_2} v_2 + \cdots + \overline{u_n} v_n.$$

The length of $\mathbf{u}$ is given by $\|\mathbf{u}\| = \sqrt{|u_1|^2 + \cdots + |u_n|^2}$.

Definition 7.6 *Let A be a square matrix with complex entries. A is called*
(i) **Hermitian** *if $A = A^*$;*
(ii) **Skew Hermitian** *if $A = -A^*$.*

Lemma 7.2 *A is Hermitian iff for all column vectors $\mathbf{v}, \mathbf{w}$ we have*

$$(A\mathbf{v})^* \mathbf{w} = \mathbf{v}^* A \mathbf{w}; \quad i.e., \ (\langle A\mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{v}, A\mathbf{w} \rangle) \tag{52}$$

Proof: If A is Hermitian then $(A\mathbf{v})^* \mathbf{w} = \mathbf{v}^* A^* \mathbf{w} = \mathbf{v}^* A \mathbf{w}$. To see the converse, take $\mathbf{v}, \mathbf{w}$ to be standard basic column vectors. ♠

Remark 7.5
(i) If A is real then $A = A^*$ means $A = A^t$. Hence real symmetric matrices are Hermitian. Likewise a real skew Hermitian matrix is skew symmetric.
(ii) A is Hermitian iff $\imath A$ is skew Hermitian.

Proposition 7.7 *Let A be an $n \times n$ Hermitian matrix. Then :*

1. *For any $\mathbf{u} \in \mathbb{C}^n, \mathbf{u}^* A \mathbf{u}$ is a real number.*
2. *All eigenvalues of A are real.*
3. *Eigenvectors of a Hermitian matrix corresponding to distinct eigenvalues are mutually orthogonal.*

Proof: (1) Since $\mathbf{u}^* A \mathbf{u}$ is a complex number, to prove it is real, we prove that $(\mathbf{u}^* A \mathbf{u})^* = \mathbf{u}^* A \mathbf{u}$. But $(\mathbf{u}^* A \mathbf{u})^* = \mathbf{u}^* A^* (\mathbf{u}^*)^* = \mathbf{u}^* A \mathbf{u}$. Hence $\mathbf{u}^* A \mathbf{u}$ is real for all $\mathbf{u} \in \mathbb{C}^n$.

(2) Suppose λ is an eigenvalue of A and $\mathbf{u}$ is an eigenvector for λ. Then

$$\mathbf{u}^* A \mathbf{u} = \mathbf{u}^* (\lambda \mathbf{u}) = \lambda (\mathbf{u}^* \mathbf{u}) = \lambda \|\mathbf{u}\|^2.$$

Since $\mathbf{u}^* A \mathbf{u}$ is real and $\|\mathbf{u}\|$ is a nonzero real number, it follows that λ is real.

(3) Let λ and μ be two distinct eigenvalues of A and $\mathbf{u}$ and $\mathbf{v}$ be corresponding eigenvectors. Then $A\mathbf{u} = \lambda \mathbf{u}$ and $A\mathbf{v} = \mu \mathbf{v}$. Hence

$$\lambda \mathbf{u}^* \mathbf{v} = (\lambda \mathbf{u})^* \mathbf{v} = (A\mathbf{u})^* \mathbf{v} = \mathbf{u}^* (A\mathbf{v}) = \mathbf{u}^* \mu \mathbf{v} = \mu (\mathbf{u}^* \mathbf{v}).$$

Hence $(\lambda - \mu) \mathbf{u}^* \mathbf{v} = 0$. Since $\lambda \neq \mu$, $\quad \mathbf{u}^* \mathbf{v} = 0$. ♠

Corollary 7.1 *Let A be an $n \times n$ skew Hermitian matrix. Then :*

1. *For any $\mathbf{u} \in \mathbb{C}^n, \mathbf{u}^* A \mathbf{u}$ is either zero or a purely imaginary number.*
2. *Each eigenvalue of A is either zero or a purely imaginary number.*
3. *Eigenvectors of A corresponding to distinct eigenvalues are mutually orthogonal.*

Proof: All this follow straight way from the corresponding statement about Hermitian matrix, once we note that A is skew Hermitian implies $\imath A$ is Hermitian and the fact that a complex number c is real iff $\imath c$ is either zero or purely imaginary.

Definition 7.7 *Let A be a square matrix over $\mathbb{C}$. A is called*
(i) **unitary** *if $A^*A = I$;*
(ii) **orthogonal** *if A is real and unitary.*

Thus a real matrix A is orthogonal iff $A^T = A^{-1}$. Also observe that A is unitary iff A^T is unitary iff $\overline{A}$ is unitary.

Example 7.2 *The matrices*

$$U = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \quad \textit{and} \quad V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}$$

are orthogonal and unitary respectively.

Proposition 7.8 *Let A be a square matrix. Then the following conditions are equivalent.*
(i) U is unitary.
(ii) The rows of U form an orthonormal set of vectors.
(iii) The columns of U form an orthonormal set of vectors.
(iv) U preserves the inner product, i.e., for all vectors $\mathbf{x}, \mathbf{y} \in \mathbb{C}^{\mathbf{n}}$, we have $\langle U\mathbf{x}, U\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$.

Proof: Write the matrix U column-wise :

$$U = [\mathbf{u}_1\ \mathbf{u}_2\ \ldots\ \mathbf{u}_n] \quad \text{so that} \quad U^* = \begin{bmatrix} \mathbf{u}_1^* \\ \mathbf{u}_2^* \\ \vdots \\ \mathbf{u}_n^* \end{bmatrix}.$$

Hence

$$\begin{aligned} U^*U &= \begin{bmatrix} \mathbf{u}_1^* \\ \mathbf{u}_2^* \\ \vdots \\ \mathbf{u}_n^* \end{bmatrix} [\mathbf{u}_1\ \mathbf{u}_2\ \ldots\ \mathbf{u}_n] \\ &= \begin{bmatrix} \mathbf{u}_1^*\mathbf{u}_1 & \mathbf{u}_1^*\mathbf{u}_2 & \cdots & \mathbf{u}_1^*\mathbf{u}_n \\ \mathbf{u}_2^*\mathbf{u}_1 & \mathbf{u}_2^*\mathbf{u}_2 & \cdots & \mathbf{u}_2^*\mathbf{u}_n \\ \vdots & & \cdots & \\ \mathbf{u}_n^*\mathbf{u}_1 & \mathbf{u}_n^*\mathbf{u}_2 & \cdots & \mathbf{u}_n^*\mathbf{u}_n \end{bmatrix}. \end{aligned}$$

Thus $U^*U = I$ iff $\mathbf{u}_i^*\mathbf{u}_j = 0$ for $i \neq j$ and $\mathbf{u}_i^*\mathbf{u}_i = 1$ for $i = 1, 2, \ldots, n$ iff the column vectors of U form an orthonormal set. This proves $(i) \iff (ii)$. Since $U^*U = I$ implies $UU^* = I$, the proof of $(i) \iff (iii)$ follows.
To prove $(i) \iff (iv)$ let U be unitary. Then $U^*U = Id$ and hence $\langle U\mathbf{x}, U\mathbf{y}\rangle = \langle \mathbf{x}, U^*U\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$. Conversely, iff U preserves inner product take $\mathbf{x} = \mathbf{e}_i$ and $\mathbf{y} = \mathbf{e}_j$ to get

$$\mathbf{e}_i^*(U^*U)\mathbf{e}_j = \mathbf{e}_i^*\mathbf{e}_j = \delta_{ij}$$

where δ_{ij} are Kronecker symbols ($\delta_{ij} = 1$ if $i = j$; $= 0$ otherwise.) This means the $(i,j)^{th}$ entry of U^*U is δ_{ij}. Hence $U^*U = I_n$. ♠

Remark 7.6 *Observe that the above theorem is valid for an orthogonal matrix also by merely applying it for a real matrix.*

Corollary 7.2 *Let U be a unitary matrix. Then :*
(1) For all $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$, $\langle U\mathbf{x}, U\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$. Hence $\|U\mathbf{x}\| = \|\mathbf{x}\|$.
(2) If λ is an eigenvalue of U then $|\lambda| = 1$.
(3) Eigenvectors corresponding to different eigenvalues are orthogonal.

Proof: (1) We have, $\|U\mathbf{x}\|^2 = \langle U\mathbf{x}, U\mathbf{x}\rangle = \langle \mathbf{x}, \mathbf{x}\rangle = \|\mathbf{x}\|^2$.
(2) If λ is an eigenvalue of U with eigenvector $\mathbf{x}$ then $U\mathbf{x} = \lambda\mathbf{x}$. Hence $\|\mathbf{x}\| = \|U\mathbf{x}\| = |\lambda|\,\|\mathbf{x}\|$. Hence $|\lambda| = 1$.
(3) Let $U\mathbf{x} = \lambda\mathbf{x}$ and $U\mathbf{y} = \mu\mathbf{y}$ where $\mathbf{x}, \mathbf{y}$ are eigenvectors with distinct eigenvalues λ and μ respectively. Then

$$\langle \mathbf{x}, \mathbf{y}\rangle = \langle U\mathbf{x}, U\mathbf{y}\rangle = \langle \lambda\mathbf{x}, \mu\mathbf{y}\rangle = \overline{\lambda}\mu\langle \mathbf{x}, \mathbf{y}\rangle.$$

Hence $\overline{\lambda}\mu = 1$ or $\langle \mathbf{x}, \mathbf{y}\rangle = 0$. Since $\overline{\lambda}\lambda = 1$, we cannot have $\overline{\lambda}\mu = 1$. Hence $\langle \mathbf{x}, \mathbf{y}\rangle = 0$, i.e., $\mathbf{x}$ and $\mathbf{y}$ are orthogonal. ♠

Example 7.3 $U = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ *is an orthogonal matrix. The characteristic polynomial of U is :*

$$D(\lambda) = \det(U - \lambda I) = \det, \begin{bmatrix} \cos\theta - \lambda & -\sin\theta \\ \sin\theta & \cos\theta - \lambda \end{bmatrix} = \lambda^2 - 2\lambda\cos\theta + 1.$$

Roots of $D(\lambda) = 0$ are :

$$\lambda = \frac{2\cos\theta \pm \sqrt{4cos^2\theta - 4}}{2} = \cos\theta \pm \imath\sin\theta = e^{\pm\imath\theta}.$$

Hence $|\lambda| = 1$. Check that eigenvectors are :
for $\underline{\lambda = e^{\imath\theta}}$: $x = \begin{bmatrix} 1 \\ -\imath \end{bmatrix}$ and for $\underline{\lambda = e^{-\imath\theta}}$: $y = \begin{bmatrix} 1 \\ \imath \end{bmatrix}$.
Thus $\mathbf{x}^*\mathbf{y} = [1\ \imath]\begin{bmatrix} 1 \\ \imath \end{bmatrix} = 1 + \imath^2 = 0$. Hence $\mathbf{x} \perp \mathbf{y}$. Normalize the eigenvectors $\mathbf{x}$ and $\mathbf{y}$. Therefore if we take,

$$C = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -\imath & \imath \end{bmatrix}$$

then $C^{-1}UC = D(e^{\imath\theta}, e^{-\imath\theta})$.

7.5 Spectral Theorem and Applications

Similarity does not necessarily preserve the distance. In terms of matrices, this may be noticed in the fact that an arbitrary conjugate $C^{-1}AC$ of a Hermitian matrix may not be Hermitian. Thus the diagonalization problem for special matrices such as Hermitian matrices needs a special treatment viz., we need to restrict C to those matrices which preserve the inner product. In this section we shall first establish an important result that says that a Hermitian matrix can be diagonalized using a unitary transformation. We shall then see some applications of this to geometry.

Closely related to the problem of diagonalization is the problem of triangularization. We shall use this concept as a stepping stone toward the solution of diagonalization problem.

Definition 7.8 *Two $n \times n$ matrices A and B are said to be* congruent *if there exists a unitary matrix C such that $C^*AC = B$.*

Definition 7.9 *We say A is* triangularizable *if there exists an invertible matrix C such that $C^{-1}AC$ is upper triangular.*

Remark 7.7 *Obviously all diagonalizable matrices are triangularizable. The following result says that triangularizability causes least problem:*

Proposition 7.9 *Over the complex numbers every square matrix is congruent to an upper triangular matrix.*

Proof: Let A be a $n \times n$ matrix with complex entries. We have to find a unitary matrix C such that C^*AC is upper triangular. We shall prove this by induction. For $n = 1$ there is nothing to prove. Assume the result for $n - 1$. Let μ be an eigenvalue of A and let $\mathbf{v} \in \mathbb{C}^n$ be such that $A\mathbf{v}_1 = \mu\mathbf{v}_1$. (Here we need to work with complex numbers; real numbers won't do. why?) We can choose $\mathbf{v}_1$ to be of norm 1. We can then complete it to an orthonormal basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. (Here we use Gram-Schmidt.) We then take C_1 to be the matrix whose i^{th} column is $\mathbf{v}_i$. Then as seen earlier, C_1 is a unitary matrix. Put $A_1 = C_1^{-1}AC_1$. Then $A_1\mathbf{e}_1 = C_1^{-1}AC_1\mathbf{e}_1 = C_1^{-1}A\mathbf{v}_1 = \mu C_1^{-1}(\mathbf{v}_1) = \mu\mathbf{e}_1$. This shows that the first column of A_1 has all entries zero except the first one which is equal to μ. Let B be the matrix obtained from A_1 by cutting down the first row and the first column, so that A_1 is a block matrix of the form

$$A_1 = \begin{bmatrix} \mu & \ldots \\ 0_{n-1} & B \end{bmatrix}$$

where 0_{n-1} is the column of zeros of size $n - 1$.

By induction there exists a $(n-1) \times (n-1)$ unitary matrix M such that $M^{-1}BM$ is an upper triangular matrix. Put $M_1 = \left[\begin{array}{c|c} 1 & 0_{n-1}^t \\ \hline 0_{n-1} & M \end{array}\right]$.

Then M_1 is unitary and hence $C = C_1M_1$ is also unitary. Clearly $C^{-1}AC = M_1^{-1}C_1^{-1}AC_1M_1 = M_1^{-1}A_1M_1$ which is of the form

$$\begin{bmatrix} 1 & 0_{n-1}^t \\ 0_{n-1} & M^{-1} \end{bmatrix} \begin{bmatrix} \mu & \ldots \\ 0_{n-1} & B \end{bmatrix} \begin{bmatrix} 1 & 0_{n-1}^t \\ 0_{n-1} & M \end{bmatrix}$$

and hence is upper triangular. ♠

Remark 7.8 *Assume now that A is a real matrix with all its eigenvalues real. Then from lemma 7.1, it follows that we can choose the eigenvector $\mathbf{v}_1$ to be a real vector and then complete this into a basis for $\mathbb{R}^n$. Thus the matrix C_1 corresponding to this basis will have real entries. By induction M will have real entries and hence the product $C = MC_1$ will also have real entries. Thus we have proved:*

Proposition 7.10 *For a real square matrix A with all its eigenvalues real, there exists an orthogonal matrix C such that C^tAC is upper triangular.*

Definition 7.10 *A square matrix A is called* normal *if $A^*A = AA^*$.*

Remark 7.9
(i) Normality is congruence invariant. This means that if C is unitary and A is normal then $C^{-1}AC$ is also normal. This is easy to verify.
(ii) Any diagonal matrix is normal. Therefore it follows that normality is necessary for diagonalization. Amazingly, it turns out to be sufficient. That is the reason to define this concept.
(iii) Observe that product of two normal matrices may not be normal. For example take $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$; $B = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$.
(iv) Certainly Hermitian matrices are normal. Of course there are normal matrices which are not Hermitian. For example take $A = \begin{bmatrix} \imath & 0 \\ 0 & -\imath \end{bmatrix}$.

Lemma 7.3 *For a normal matrix A we have $\|A\mathbf{x}\|^2 = \|A^*\mathbf{x}\|^2$ for all $\mathbf{x} \in \mathbb{C}^n$.*

Proof: $\langle A\mathbf{x}, A\mathbf{x}\rangle = (A\mathbf{x})^*(A\mathbf{x}) = \mathbf{x}^*A^*A\mathbf{x} = \mathbf{x}^*AA^*\mathbf{x} = (A^*\mathbf{x})^*(A^*\mathbf{x}) = \langle A^*\mathbf{x}, A^*\mathbf{x}\rangle = \|A^*\mathbf{x}\|^2$. ♠

Lemma 7.4 *If A is normal, then $\mathbf{v}$ is an eigenvector of A with eigenvalue μ iff $\mathbf{v}$ is an eigenvector of A^* with eigenvalue $\overline{\mu}$.*

Proof: Observe that if A is normal then $A - \mu I$ is also normal. Now $(A - \mu I)(\mathbf{v}) = 0$ iff $\|(A - \mu I)(\mathbf{v})\| = 0$ iff $\|(A - \mu I)^*\mathbf{v}\| = 0$ iff $(A^* - \overline{\mu} I)(\mathbf{v}) = 0$. ♠

Proposition 7.11 *An upper triangular normal matrix is diagonal.*

Proof: Let A be an upper triangular normal matrix. Inductively we shall show that $a_{ij} = 0$ for $j > i$. We have $A\mathbf{e}_1 = a_{11}\mathbf{e}_1$. Hence $\|A\mathbf{e}_1\|^2 = |a_{11}|^2$. On the other hand this is equal to $\|A^*\mathbf{e}_1\| = |a_{11}|^2 + |a_{12}|^2 + \cdots + |a_{1n}|^2$. Hence $a_{12} = a_{13} = \cdots = a_{1n} = 0$. Inductively suppose we have shown $a_{ij} = 0$ for $j > i$ for all $1 \le i \le k-1$. Then it follows that $A\mathbf{e}_k = a_{kk}\mathbf{e}_k$. Exactly as in the first case, this implies that $\|A^*\mathbf{e}_k\| = |a_{k,k}|^2 + |a_{k,k+1}|^2 + \cdots + |a_{k,n}|^2 = |a_{k,k}|^2$. Hence $a_{k,k+1} = \cdots = a_{k,n} = 0$. ♠

Why all this fuss? Well for one thing, we now have a big theorem. All we have to do is to combine propositions 7.9 and 7.11.

Theorem 7.2 Spectral Theorem *Given any normal matrix A, there exists a unitary matrix C such that C^*AC is a diagonal matrix.*

Corollary 7.3 *Every Hermitian matrix A is congruent to a diagonal matrix. A real symmetric matrix is real-congruent to a diagonal matrix.*

Proof: For the first statement we simply observe that a Hermitian matrix is normal and apply the above theorem. For the second statement, we first recall that for a real symmetric matrix, all eigenvalues are real. Hence the proposition 7.10 is applicable and so we can choose C to be an orthogonal matrix. Along with proposition 7.11, this gives the result. ♠

Quadratic forms and their diagonalization

Definition 7.11 *Let $A = (a_{ij})$ be an $n \times n$ real matrix . The function $Q : \mathbb{R}^n \to \mathbb{R}$ defined by :*

$$Q(x) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j, \quad X = (x_1, x_2, \ldots, x_n)^t \in \mathbb{R}^n$$

is called the **quadratic form** *associated with A. If $A =$ diag $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ then $Q(X) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2$ is called a* **diagonal form**.

Proposition 7.12 $Q(X) = [x_1, x_2, \ldots, x_n]A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = X^tAX$ *where* $X = (x_1, x_2, \ldots, x_n)^t$.

Proof:

$$\begin{aligned}
[x_1, x_2, \ldots, x_n]A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} &= [x_1, x_2, \ldots, x_n] \begin{bmatrix} \sum_{j=1}^{n} a_{1j}x_j \\ \vdots \\ \sum_{j=1}^{n} a_{nj}x_j \end{bmatrix} \\
&= \sum_{j=1}^{n} a_{1j}x_jx_1 + \cdots + \sum_{i=1}^{n} a_{nj}x_jx_n \\
&= \sum_{j=1}^{n}\sum_{i=1}^{n} a_{ij}x_ix_j \\
&= Q(x).
\end{aligned}$$

♠

Example 7.4

(1) $A = \begin{bmatrix} 1 & 1 \\ 3 & 5 \end{bmatrix}, \quad X = \begin{bmatrix} x \\ y \end{bmatrix}$. *Then*

$$X^tAX = [x, y] \begin{bmatrix} 1 & 1 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = [x, y] \begin{bmatrix} x + y \\ 3x + 5y \end{bmatrix} = x_1^2 + 4xy + 5y^2.$$

(2) Let $B = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}, \quad X = \begin{bmatrix} x \\ y \end{bmatrix}$. *Then*

$$X^tBX = [x, y] \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = [x, y] \begin{bmatrix} x + 2y \\ 2x + 5y \end{bmatrix} = x_1^2 + 4xy + 5x_2^2.$$

Notice that A and B give rise to same $Q(x)$ and $B = \frac{1}{2}(A + A^t)$ is a symmetric matrix.

Proposition 7.13 *For any $n \times n$ matrix A and the column vector $X = (x_1, x_2, \ldots, x_n)^t$,*

$$X^tAX = X^tBX \quad \text{where } B = \frac{1}{2}(A + A^t).$$

Hence every quadratic form is associated with a symmetric matrix.

Proof: X^tAX is a 1×1 matrix. Hence $X^tAX = X^tA^tX = (X^tAX)^t$. Hence

$$X^tAX = \frac{1}{2}X^tAX + \frac{1}{2}X^tA^tX = X^t\frac{1}{2}\left(A + A^t\right)X = X^tBX.$$

♠

We now show how the spectral theorem helps us in converting a quadratic form into a diagonal form.

Theorem 7.3 *Let X^tAX be a quadratic form associated with a real symmetric matrix A. Let U be an orthogonal matrix such that $U^tAU =$ diag $(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Then*

$$X^tAX = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2,$$

where

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = U \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = UY.$$

Proof: Since $X = UY$,

$$X^tAX = (UY)^tA(UY) = Y^t(U^tAU)Y.$$

Since $U^tAU =$ diag $(\lambda_1, \lambda_2, \ldots, \lambda_n)$, we get

$$\begin{aligned} X^tAX &= [y_1, y_2, \ldots, y_n] \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \\ &= \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2. \end{aligned}$$

♠

Example 7.5 *Let us determine the orthogonal matrix U which reduces the quadratic form $Q(X) = 2x_1^2 + 4xy + 5x_2^2$ to a diagonal form. We write*

$$Q(X) = [x, y] \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = X^tAX.$$

The symmetric matrix A can be diagonalized. The eigenvalues of A are $\lambda_1 = 1$ and $\lambda_2 = 6$. An orthonormal set of eigenvectors for λ_1 and λ_2 is

$$\mathbf{v}_1 = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ -1 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Hence $U = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}$. *The change of variables equations are* $\begin{bmatrix} x \\ y \end{bmatrix} = U\begin{bmatrix} u \\ v \end{bmatrix}$. *The diagonal form is:*

$$[u,v]\begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix}[u,v]^T = u^2 + 6v^2.$$

Check that $U^tAU = $ *diag* $(1,6)$.

7.6 Conic Sections and quadric surfaces

A conic section is the locus in the Cartesian plane $\mathbb{R}^2$ of an equation of the form

$$ax^2 + bxy + cy^2 + dx + ey + f = 0. \tag{53}$$

It can be proved that this equation represents one of the following: (i) the empty set (ii) a single point (iii) one or two straight lines (iv) an ellipse (v) an hyperbola and (vi) a parabola. The second degree part of (53)

$$Q(x,y) = ax^2 + bxy + cy^2$$

is a quadratic form. This determines the type of the conic. We can write the equation (53) into matrix form after setting $x = x, y = y$:

$$[x,y]\begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + [d,e]\begin{bmatrix} x \\ y \end{bmatrix} + f \;=\; 0 \tag{54}$$

Write $A = \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix}$. Let $U = [\mathbf{v}_1, \mathbf{v}_2]$ be an orthogonal matrix whose column vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are eigenvectors of A with eigenvalues λ_1 and λ_2. Apply the change of variables

$$X = \begin{bmatrix} x \\ y \end{bmatrix} = U\begin{bmatrix} u \\ v \end{bmatrix}$$

to diagonalize the quadratic form $Q(x,y)$ to the diagonal form $\lambda_1 y_1^2 + \lambda_2 y_2^2$. The orthonormal basis $\{\mathbf{v}_1, \mathbf{v}_2\}$ determines a new set of coordinate axes with respect to which the locus of the equation $[x,y]A[x,y]^T + B[x,y]^T + f = 0$ with $B = [d,e]$ is same as the locus of the equation

$$\begin{aligned} 0 &= [u,v] \text{ diag } (\lambda_1,\lambda_2)[u,v]^T + (BU)[u,v]^T + f \\ &= \lambda_1 u^2 + \lambda_2 v^2 + [d,e][\mathbf{v}_1,\mathbf{v}_2][u,v]^T + f. \end{aligned} \tag{55}$$

If the conic determined by (55) is not degenerate i.e., not an empty set, a point, nor line(s) then signs of λ_1 and λ_2 determine whether it is a parabola, an hyperbola or an ellipse. The equation (53) will represent (1) ellipse if $\lambda_1\lambda_2 > 0$ (2) hyperbola if $\lambda_1\lambda_2 < 0$ (3) parabola if $\lambda_1\lambda_2 = 0$

Example 7.6
(1) $2x^2 + 4xy + 5y^2 + 4x + 13y - 1/4 = 0$.

We have earlier diagonalized the quadratic form $2x^2+4xy+5y^2$. *The associated symmetric matrix, the eigenvectors and eigenvalues are displayed in the equation of diagonalization :*

$$U^tAU = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix}\frac{1}{\sqrt{5}}\begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix}.$$

Set $t = 1/\sqrt{5}$ for convenience. Then the change of coordinates equations are :

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2t & t \\ -t & 2t \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix},$$

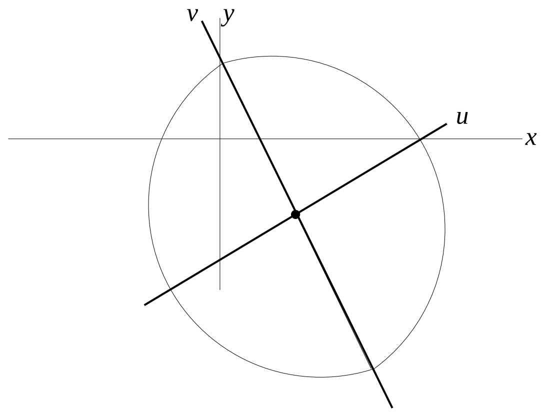
i.e., $x = t(2u + v)$ *and* $y = t(-u + 2v)$. *Substitute these into the original equation to get*

$$u^2 + 6v^2 - \sqrt{5}u + 6\sqrt{5}v - \frac{1}{4} = 0.$$

Complete the square to write this as

$$(u - \frac{1}{2}\sqrt{5})^2 + 6(v + \frac{1}{2}\sqrt{5})^2 = 9.$$

This is an equation of ellipse with center $(\frac{1}{2}\sqrt{5}, -\frac{1}{2}\sqrt{5})$ *in the uv-plane. The u-axis and v-axis are determined by the eigenvectors* $\mathbf{v}_1$ *and* $\mathbf{v}_2$ *as indicated in the following figure :*



(2) $2x^2 - 4xy - y^2 - 4x + 10y - 13 = 0$. *Here, the matrix* $A = \begin{bmatrix} 2 & -2 \\ -2 & -1 \end{bmatrix}$ *gives the quadratic part of the equation. We write the equation in matrix form as*

$$[x, y]\begin{bmatrix} 2 & -2 \\ -2 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + [-4, 10]\begin{bmatrix} x \\ y \end{bmatrix} - 13 = 0.$$

Let $t = 1/\sqrt{5}$. *The eigenvalues of* A *are* $\lambda_1 = 3, \lambda_2 = -2$. *An orthonormal set of eigenvectors is* $\mathbf{v}_1 = t(2, -1)^t$ *and* $\mathbf{v}_2 = t(1, 2)^t$. *Put*

$$U = t\begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix} \text{ and } \begin{bmatrix} x \\ y \end{bmatrix} = U\begin{bmatrix} u \\ v \end{bmatrix}.$$

The transformed equation becomes

$$3u^2 - 2v^2 - 4t(2u + v) + 10t(-u + 2v) - 13 = 0$$

or

$$3u^2 - 2v^2 - 18tu + 16tv - 13 = 0.$$

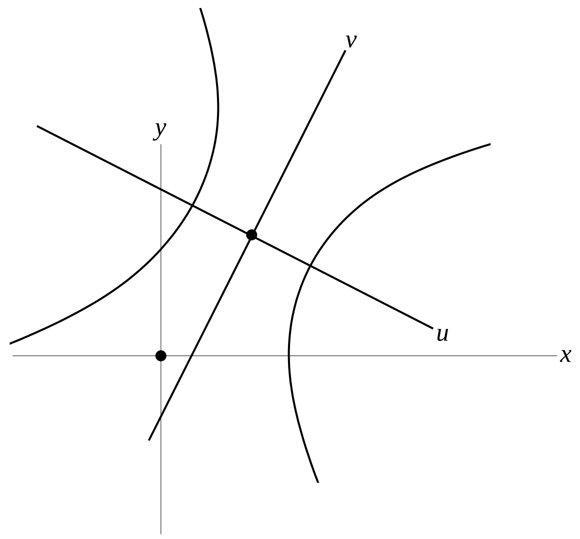
Complete the square in u and v to get

$$3(u - 3t)^2 - 2(v - 4t)^2 = 12$$

or

$$\frac{(u-3t)^2}{4} - \frac{(v-4t)^2}{6} = 1.$$

This represents a hyperbola with center $(3t, 4t)$ *in the uv-plane. The eigenvectors* $\mathbf{v}_1$ *and* $\mathbf{v}_2$ *determine the directions of positive u and v axes.*



(3) $9x^2 + 24xy + 16xy^2 - 20x + 15y = 0$

The symmetric matrix for the quadratic part is $A = \begin{bmatrix} 9 & 12 \\ 12 & 16 \end{bmatrix}$. *The eigenvalues are* $\lambda_1 = 25, \lambda_2 = 0$. *An orthonormal set of eigenvectors is* $\mathbf{v}_1 = a(3,4)^t, \mathbf{v}_2 = a(-4,3)^t$ *where* $a = 1/5$. *An orthogonal diagonalizing matrix is* $U = a\begin{bmatrix} 3 & -4 \\ 4 & 3 \end{bmatrix}$. *The equations of change of coordinates are*

$$\begin{bmatrix} x \\ y \end{bmatrix} = U \begin{bmatrix} u \\ v \end{bmatrix} \quad \text{i.e., } x = a(3u-4v),\ y = a(4u+3v).$$

The equation in uv-plane is $u^2 + v = 0$. *This is an equation of parabola with its vertex at the origin.*

Quadric Surfaces

Let A be a 3×3 real symmetric matrix. The locus of the equation

$$[x, y, z]A\begin{bmatrix} x \\ y \\ z \end{bmatrix} + [b_1, b_2, b_3]\begin{bmatrix} x \\ y \\ z \end{bmatrix} + c = 0 \qquad (56)$$

in three variables is called a **quadric surface**. We can carry out an analysis similar to the one for quadratic forms in two variables to bring (56) into standard form. Degenerate cases may arise. But the primary cases are :

Equation	Surface	signs of eigenvalues of A
$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}$	ellipsoid	all three positive
$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z}{c_2}=0$	elliptic paraboloid	two positive, one negative
$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=0$	elliptic cone	two positive, one negative
$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$	1-sheeted hyperboloid	two positive, one negative
$\frac{x^2}{a^2}-\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$	2-sheeted hyperboloid	one positive, two negative
$\frac{x^2}{a^2}-\frac{y^2}{b^2}-\frac{z}{c}=0$	hyperbolic paraboloid	one positive, one negative, one zero.

Example 7.7

(1) $7x^2+7y^2-2z^2+20yz-20zx-2xy=36.$

The matrix form is:

$$[x,y,z]\begin{bmatrix} 7 & -1 & -10 \\ -1 & 7 & 10 \\ -10 & 10 & -2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}=36.$$

Let A *be the* 3×3 *matrix appearing in this equation. The eigenvalues of* A *are* $\lambda_1=6, \lambda_2=-12$ *and* $\lambda_2=18$*. An orthonormal set of eigenvectors is given by the column vectors of the orthogonal matrix*

$$U=\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \end{bmatrix} \text{ and } U^tAU=\begin{bmatrix} 6 & 0 & 0 \\ 0 & -12 & 0 \\ 0 & 0 & 18 \end{bmatrix}.$$

consider the change of coordinates given by

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}=U\begin{bmatrix} u \\ v \\ w \end{bmatrix}.$$

This change of coordinates transforms the given equation into the form

$$6u^2-12v^2+18w^2=36$$

or

$$\frac{u^2}{6}-\frac{v^2}{3}+\frac{w^2}{2}=1$$

This is a hyperboloid of one sheet.

(2) Consider the quadric

$$x^2+y^2+z^2+4yz-4zx+4xy=27.$$

The symmetric matrix for this is

$$A=\begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & 2 \\ -2 & 2 & 1 \end{bmatrix}.$$

The eigenvalues of A are $3, 3, -3$. Since A is diagonalizable the eigenspace $E(3)$ is 2-dimensional. We find a basis of $E(3)$ first. The eigenvectors for $\lambda = 3$ are solutions of

$$\begin{bmatrix} -2 & 2 & -2 \\ 2 & -2 & 2 \\ -2 & 2 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Hence we obtain the equation $x - y + z = 0$. Hence

$$\begin{aligned} E(3) &= \{(y - z, y, z) \mid y, z \in \mathbb{R}\} \\ &= L(\{u_1 = (0, 1, 1)^t,\ u_2 = (-1, 1, 2)^t\}). \end{aligned}$$

Now we apply Gram-Schmidt process to get an orthonormal basis of $E(3)$:

$$v_1 = (0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})^t \quad \textit{and} \quad v_2 = (-\sqrt{\frac{2}{3}}, -\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}).$$

A unit eigenvector for $\lambda = -3$ is $v_3 = (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$. We know that $\langle v_3, v_1 \rangle = \langle v_3, v_2 \rangle = 0$ since A is a symmetric matrix. The orthogonal matrix for diagonalization is $U = [v_1, v_2, v_3]$ written column wise. The quadric under the change of coordinates

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = U \begin{bmatrix} u \\ v \\ w \end{bmatrix}$$

reduces to $3u^2 + 3v^2 - 3w^2 = 27$. This is a hyperboloid of one sheet.