Linear Algebra

Santanu Dey

February 18, 2011

Spectral Theorem

Similarity does not necessarily preserve the distance.

Spectral Theorem

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.

Spectral Theorem

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.

Spectral Theorem

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 unitary matrices.

Spectral Theorem

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 unitary matrices. We shall then see some applications of this to geometry.

Triangularization

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

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

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

Triangularization

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

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

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

Remark

Obviously all diagonalizable matrices are triangularizable.

Triangularization

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

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

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

Remark

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

Triangularization

Proposition

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

Triangularization

Proposition

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

Proof: Let A be an $n \times n$ matrix with complex entries.

Triangularization

Proposition

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

Proof: Let A be an $n \times n$ matrix with complex entries. We have to find a unitary matrix C such that C^*AC is upper triangular.

Triangularization

Proposition

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

Proof: Let A be an $n \times n$ 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.

Triangularization

Proposition

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

Proof: Let A be an $n \times n$ 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.

Triangularization

Proposition

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

Proof: Let A be an $n \times n$ 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 $n - 1$.

Triangularization

Proposition

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

Proof: Let A be an $n \times n$ 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 $\mathbf{v} \in \mathbb{C}^n \setminus \{0\}$ be such that $A\mathbf{v}_1 = \mu\mathbf{v}_1$.

Triangularization

Proposition

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

Proof: Let A be an $n \times n$ 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 $\mathbf{v} \in \mathbb{C}^n \setminus \{0\}$ 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?)

Triangularization

We can choose $\mathbf{v}_1$ to be of norm 1.

Triangularization

We can choose $\mathbf{v}_1$ to be of norm 1. We can then complete it to an orthonormal basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$.

Triangularization

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.)

Triangularization

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 $\mathbf{v}_i,\ 1 \leq i \leq n$.

Triangularization

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 $\mathbf{v}_i,\ 1 \leq i \leq n$. Then as seen earlier, C_1 is a unitary matrix.

Triangularization

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 $\mathbf{v}_i$, $1 \leq i \leq n$. Then as seen earlier, C_1 is a unitary matrix. Put $A_1 = C_1^{-1} A C_1$. Then
$A_1 \mathbf{e}_1 = C_1^{-1} A C_1 \mathbf{e}_1 = C_1^{-1} A \mathbf{v}_1 = \mu C_1^{-1}(\mathbf{v}_1) = \mu \mathbf{e}_1$.

Triangularization

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 $\mathbf{v}_i$, $1 \leq i \leq n$. 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.

Triangularization

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 $\mathbf{v}_i$, $1 \leq i \leq n$. 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

$$A_1 = \left[\begin{array}{ll} \mu & \star \\ 0_{n-1} & B \end{array}\right]$$

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

Triangularization

By induction there exists a $(n-1) \times (n-1)$ unitary matrix M such that $M^{-1}BM$ is an upper triangular matrix.

Triangularization

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

Triangularization

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 $C = C_1 M_1$ is also unitary.

Triangularization

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

Then M_1 is unitary and hence $C = C_1 M_1$ 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

$$\begin{bmatrix} 1 & 0^t_{n-1} \\ 0_{n-1} & M^{-1} \end{bmatrix} \begin{bmatrix} \mu & * \\ 0_{n-1} & B \end{bmatrix} \begin{bmatrix} 1 & 0^t_{n-1} \\ 0_{n-1} & M \end{bmatrix} = \begin{bmatrix} \mu & * \\ 0_{n-1} & M^{-1}BM \end{bmatrix}$$

and hence is upper triangular. ♠

Triangularization

Remark

Assume now that A is a real matrix with all its eigenvalues real.

Triangularization

Remark

Assume now that A is a real matrix with all its eigenvalues real. Then from lemma 13, it follows that we can choose the eigenvector $\mathbf{v}_1$ to be a real vector and then complete this into a orthonormal basis for $\mathbb{R}^n$.

Triangularization

Remark

Assume now that A is a real matrix with all its eigenvalues real. Then from lemma 13, it follows that we can choose the eigenvector $\mathbf{v}_1$ to be a real vector and then complete this into a orthonormal 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.

Triangularization

Remark

Assume now that A is a real matrix with all its eigenvalues real. Then from lemma 13, it follows that we can choose the eigenvector $\mathbf{v}_1$ to be a real vector and then complete this into a orthonormal 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

For a real square matrix A with all its eigenvalues real, there exists an orthogonal matrix C such that $C^t AC$ is upper triangular.

Normal Matrices

Definition

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

Normal Matrices

Definition

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

Remark

(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.

Normal Matrices

Definition

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

Remark

(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 in order that a matrix A is congruent to a diagonal matrix, it is necessary that A is normal.

Normal Matrices

Definition

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

Remark

(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 in order that a matrix A is congruent to a diagonal matrix, it is necessary that A is normal. Amazingly, this condition turns out to be sufficient as well. That is one simple reason to define this concept.

Normal Matrices

Definition

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

Remark

(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 in order that a matrix A is congruent to a diagonal matrix, it is necessary that A is normal. Amazingly, this condition turns out to be sufficient as well. That is one simple reason to define this concept.
(iii) Observe that product of two normal matrices may not be normal.

Normal Matrices

Definition

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

Remark

(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 in order that a matrix A is congruent to a diagonal matrix, it is necessary that A is normal. Amazingly, this condition turns out to be sufficient as well. That is one simple reason to define this concept.
(iii) Observe that product of two normal matrices may not be normal. For example take $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$; $B = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$.

Normal Matrices

Definition

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

Remark

(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 in order that a matrix A is congruent to a diagonal matrix, it is necessary that A is normal. Amazingly, this condition turns out to be sufficient as well. That is one simple 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.

Normal Matrices

Definition

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

Remark

(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 in order that a matrix A is congruent to a diagonal matrix, it is necessary that A is normal. Amazingly, this condition turns out to be sufficient as well. That is one simple 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.

Normal Matrices

Definition

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

Remark

(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 in order that a matrix A is congruent to a diagonal matrix, it is necessary that A is normal. Amazingly, this condition turns out to be sufficient as well. That is one simple 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 $A = \begin{bmatrix} \imath & 0 \\ 0 & -\imath \end{bmatrix}$.

Normal Matrices

Lemma

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

Normal Matrices

Lemma

For a normal matrix A we have, $\|A\mathbf{x}\|^2 = \|A^*\mathbf{x}\|^2$ for all $\mathbf{x} \in \mathbb{C}^n$; in particular, for each i, the norm of i^{th}-row of A is equal to the norm of i^{th}-column.

Normal Matrices

Lemma

For a normal matrix A we have, $\|A\mathbf{x}\|^2 = \|A^*\mathbf{x}\|^2$ for all $\mathbf{x} \in \mathbb{C}^n$; in particular, for each i, the norm of i^{th}-row of A is equal to the norm of i^{th}-column.

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

Normal Matrices

Lemma

For a normal matrix A we have, $\|A\mathbf{x}\|^2 = \|A^*\mathbf{x}\|^2$ for all $\mathbf{x} \in \mathbb{C}^n$; in particular, for each i, the norm of i^{th}-row of A is equal to the norm of i^{th}-column.

Proof: $\|A\mathbf{x}\|^2 = (A\mathbf{x})^*(A\mathbf{x}) = \mathbf{x}^*A^*A\mathbf{x} = \mathbf{x}^*AA^*\mathbf{x} = (A^*\mathbf{x})^*(A^*\mathbf{x}) = \|A^*\mathbf{x}\|^2$.
Taking $\mathbf{x} = \mathbf{e}_i$ we get that norms of i^{th}-column of A is equal to the norm of i^{th}-column of A^*= norm of the i^{th}-row of $\bar{A}$=norm of the i^{th}-row of A. ♠

Normal Matrices

Lemma

For a normal matrix A we have, $\|A\mathbf{x}\|^2 = \|A^*\mathbf{x}\|^2$ for all $\mathbf{x} \in \mathbb{C}^n$; in particular, for each i, the norm of i^{th}-row of A is equal to the norm of i^{th}-column.

Proof: $\|A\mathbf{x}\|^2 = (A\mathbf{x})^*(A\mathbf{x}) = \mathbf{x}^*A^*A\mathbf{x} = \mathbf{x}^*AA^*\mathbf{x} = (A^*\mathbf{x})^*(A^*\mathbf{x}) = \|A^*\mathbf{x}\|^2$.
Taking $\mathbf{x} = \mathbf{e}_i$ we get that norms of i^{th}-column of A is equal to the norm of i^{th}-column of A^*= norm of the i^{th}-row of $\bar{A}$=norm of the i^{th}-row of A. ♠

Lemma

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 $\overline{\mu}$.

Normal Matrices

Lemma

For a normal matrix A we have, $\|A\mathbf{x}\|^2 = \|A^*\mathbf{x}\|^2$ for all $\mathbf{x} \in \mathbb{C}^n$; in particular, for each i, the norm of i^{th}-row of A is equal to the norm of i^{th}-column.

Proof: $\|A\mathbf{x}\|^2 = (A\mathbf{x})^*(A\mathbf{x}) = \mathbf{x}^*A^*A\mathbf{x} = \mathbf{x}^*AA^*\mathbf{x} = (A^*\mathbf{x})^*(A^*\mathbf{x}) = \|A^*\mathbf{x}\|^2$. Taking $\mathbf{x} = \mathbf{e}_i$ we get that norms of i^{th}-column of A is equal to the norm of i^{th}-column of A^*= norm of the i^{th}-row of $\bar{A}$=norm of the i^{th}-row of A. ♠

Lemma

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 $(A^* - \overline{\mu} I)(\mathbf{v}) = 0$.
♠

Normal Matrices

Proposition

An upper triangular normal matrix is diagonal.

Normal Matrices

Proposition

An upper triangular normal matrix is diagonal.

Proof: Let A be an upper triangular normal matrix.

Normal Matrices

Proposition

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 $\|A\mathbf{e}_1\|^2 = |a_{11}|^2$.

Normal Matrices

Proposition

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 $\|A^*\mathbf{e}_1\|^2 = |a_{11}|^2 + |a_{12}|^2 + \cdots + |a_{1n}|^2$.

Normal Matrices

Proposition

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 $\|A^*\mathbf{e}_1\|^2 = |a_{11}|^2 + |a_{12}|^2 + \cdots + |a_{1n}|^2$. Hence $a_{12} = a_{13} = \cdots = a_{1n} = 0$.

Normal Matrices

Proposition

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 $\|A^*\mathbf{e}_1\|^2 = |a_{11}|^2 + |a_{12}|^2 + \cdots + |a_{1n}|^2$. Hence $a_{12} = a_{13} = \cdots = a_{1n} = 0$. Inductively, suppose we have shown $a_{ij} = 0$ for $j > i$ for all $1 \leq i \leq k-1$.

Normal Matrices

Proposition

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 $\|A^*\mathbf{e}_1\|^2 = |a_{11}|^2 + |a_{12}|^2 + \cdots + |a_{1n}|^2$. Hence $a_{12} = a_{13} = \cdots = a_{1n} = 0$. Inductively, suppose we have shown $a_{ij} = 0$ for $j > i$ for all $1 \leq i \leq k - 1$. Then it follows that $A\mathbf{e}_k = a_{kk}\mathbf{e}_k$. Exactly as in the first case, this implies that $\|A^*\mathbf{e}_k\|^2 = |a_{k,k}|^2 + |a_{k,k+1}|^2 + \cdots + |a_{k,n}|^2 = |a_{k,k}|^2$. Hence $a_{k,k+1} = \cdots = a_{k,n} = 0$. ♠

Theorem

Spectral Theorem *Given any normal matrix $A \in M(n, \mathbb{C})$, there exists a unitary matrix C such that C^*AC is a diagonal matrix.*

Theorem

Spectral Theorem *Given any normal matrix $A \in M(n, \mathbb{C})$, there exists a unitary matrix C such that C^*AC is a diagonal matrix.*

Corollary

(a) Every Hermitian matrix A is congruent to a diagonal matrix.

Theorem

Spectral Theorem *Given any normal matrix $A \in M(n, \mathbb{C})$, there exists a unitary matrix C such that C^*AC is a diagonal matrix.*

Corollary

(a) Every Hermitian matrix A is congruent to a diagonal matrix.
(b) A real symmetric matrix is real-congruent to a diagonal matrix.

Proof: (a) We simply observe that a Hermitian matrix is normal and apply the above theorem.

Theorem

Spectral Theorem *Given any normal matrix $A \in M(n, \mathbb{C})$, there exists a unitary matrix C such that C^*AC is a diagonal matrix.*

Corollary

(a) Every Hermitian matrix A is congruent to a diagonal matrix.
(b) A real symmetric matrix is real-congruent to a diagonal matrix.

Proof: (a) We simply observe that a Hermitian matrix is normal and apply the above theorem.
(b) We first recall that for a real symmetric matrix, all eigenvalues are real.

Quadratic forms and their diagonalization

Definition

Let $A = (a_{ij})$ be an $n \times n$ real matrix.

Quadratic forms and their diagonalization

Definition

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

$$Q_A(\mathbf{x}) := Q_A((x_1, x_2, \ldots, x_n)^t) := \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$$

is called the **quadratic form** associated with A.

Quadratic forms and their diagonalization

Definition

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

$$Q_A(\mathbf{x}) := Q_A((x_1, x_2, \ldots, x_n)^t) := \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$$

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

Quadratic forms

Proposition

$$Q(\mathbf{x}) = [x_1, x_2, \ldots, x_n]A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \mathbf{x}^t A\mathbf{x} \text{ where } \mathbf{x} = (x_1, x_2, \ldots, x_n)^t.$$

Quadratic forms

Proposition

$$Q(\mathbf{x}) = [x_1, x_2, \ldots, x_n] A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \mathbf{x}^t A \mathbf{x} \text{ where } \mathbf{x} = (x_1, x_2, \ldots, x_n)^t.$$

Proof:

$$[x_1, x_2, \ldots, x_n] A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = [x_1, x_2, \ldots, x_n] \begin{bmatrix} \displaystyle\sum_{j=1}^{n} a_{1j} x_j \\ \vdots \\ \displaystyle\sum_{j=1}^{n} a_{nj} x_j \end{bmatrix}$$

Quadratic forms

Proposition

$$Q(\mathbf{x}) = [x_1, x_2, \ldots, x_n]A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \mathbf{x}^t A\mathbf{x} \text{ where } \mathbf{x} = (x_1, x_2, \ldots, x_n)^t.$$

Proof:

$$[x_1, x_2, \ldots, x_n]A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = [x_1, x_2, \ldots, x_n] \begin{bmatrix} \sum_{j=1}^{n} a_{1j}x_j \\ \vdots \\ \sum_{j=1}^{n} a_{nj}x_j \end{bmatrix}$$

$$= \sum_{j=1}^{n} a_{1j}x_jx_1 + \cdots + \sum_{j=1}^{n} a_{nj}x_jx_n$$

Quadratic forms

Proposition

$$Q(\mathbf{x}) = [x_1, x_2, \ldots, x_n] A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \mathbf{x}^t A \mathbf{x} \text{ where } \mathbf{x} = (x_1, x_2, \ldots, x_n)^t.$$

Proof:

$$[x_1, x_2, \ldots, x_n] A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = [x_1, x_2, \ldots, x_n] \begin{bmatrix} \sum_{j=1}^{n} a_{1j} x_j \\ \vdots \\ \sum_{j=1}^{n} a_{nj} x_j \end{bmatrix}$$

$$= \sum_{j=1}^{n} a_{1j} x_j x_1 + \cdots + \sum_{j=1}^{n} a_{nj} x_j x_n = \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij} x_i x_j = Q(x).$$

♠

Quadratic forms

Example

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

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

Quadratic forms

Example

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

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

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

Quadratic forms

Example

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

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

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

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

Quadratic forms

Example

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

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

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

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

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

Proposition

For any $n \times n$ matrix A and the column vector $\mathbf{x} = (x_1, x_2, \ldots x_n)^t$

$$\mathbf{x}^t A\mathbf{x} = \mathbf{x}^t B\mathbf{x} \quad \textit{where} \quad B = \frac{1}{2}(A + A^t).$$

Proposition

For any $n \times n$ *matrix* A *and the column vector* $\mathbf{x} = (x_1, x_2, \ldots x_n)^t$

$$\mathbf{x}^t A\mathbf{x} = \mathbf{x}^t B\mathbf{x} \quad \textit{where} \quad B = \frac{1}{2}(A + A^t).$$

Hence every quadratic form is associated with a symmetric matrix.

Proposition

For any $n \times n$ matrix A and the column vector $\mathbf{x} = (x_1, x_2, \ldots x_n)^t$

$$\mathbf{x}^t A\mathbf{x} = \mathbf{x}^t B\mathbf{x} \quad \textit{where} \quad B = \frac{1}{2}(A + A^t).$$

Hence every quadratic form is associated with a symmetric matrix.

Proof: $\mathbf{x}^t A\mathbf{x}$ is a 1×1 matrix. Hence $\mathbf{x}^t A\mathbf{x} = (\mathbf{x}^t A\mathbf{x})^t = \mathbf{x}^t A^t\mathbf{x}$.

Proposition

For any $n \times n$ matrix A and the column vector $\mathbf{x} = (x_1, x_2, \ldots x_n)^t$

$$\mathbf{x}^t A\mathbf{x} = \mathbf{x}^t B\mathbf{x} \quad \textit{where} \quad B = \frac{1}{2}(A + A^t).$$

Hence every quadratic form is associated with a symmetric matrix.

Proof: $\mathbf{x}^t A\mathbf{x}$ is a 1×1 matrix. Hence $\mathbf{x}^t A\mathbf{x} = (\mathbf{x}^t A\mathbf{x})^t = \mathbf{x}^t A^t\mathbf{x}$. Hence

$$\mathbf{x}^t A\mathbf{x} = \frac{1}{2}\mathbf{x}^t A\mathbf{x} + \frac{1}{2}\mathbf{x}^t A^t\mathbf{x} = \mathbf{x}^t \frac{1}{2}(A + A^t)\mathbf{x} = \mathbf{x}^t B\mathbf{x}.$$

$\square$

Quadratic forms and their diagonalization

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

Quadratic forms and their diagonalization

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

Theorem

Let $\mathbf{x}^t A \mathbf{x}$ *be a quadratic form associated with a real symmetric matrix* A*.*

Quadratic forms and their diagonalization

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

Theorem

Let $\mathbf{x}^t A\mathbf{x}$ be a quadratic form associated with a real symmetric matrix A. Let U be an orthogonal matrix such that $U^t AU = \text{ diag } (\lambda_1, \lambda_2, \ldots, \lambda_n)$.

Quadratic forms and their diagonalization

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

Theorem

Let $\mathbf{x}^t A\mathbf{x}$ be a quadratic form associated with a real symmetric matrix A. Let U be an orthogonal matrix such that $U^t AU =$ diag $(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Then

$$\mathbf{x}^t A\mathbf{x} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2,$$

where $\mathbf{y} = (y_1, \ldots, y_n)$ are define by

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = U \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = U\mathbf{y}.$$

Quadratic forms and their diagonalization

Proof: Since $\mathbf{x} = U\mathbf{y}$,

$$\mathbf{x}^t A\mathbf{x} = (U\mathbf{y})^t A(U\mathbf{y}) = \mathbf{y}^t(U^t AU)\mathbf{y}.$$

Quadratic forms and their diagonalization

Proof: Since $\mathbf{x} = U\mathbf{y}$,

$$\mathbf{x}^t A\mathbf{x} = (U\mathbf{y})^t A(U\mathbf{y}) = \mathbf{y}^t(U^t AU)\mathbf{y}.$$

Since $U^t AU = \text{ diag } (\lambda_1, \lambda_2, \ldots, \lambda_n)$, we get

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

♠

Quadratic forms: An example

Example

Let us determine the orthogonal matrix U which reduces the quadratic form $Q(\mathbf{x}) = 2x_1^2 + 4xy + 5x_2^2$ to a diagonal form.

Quadratic forms: An example

Example

Let us determine the orthogonal matrix U which reduces the quadratic form $Q(\mathbf{x}) = 2x_1^2 + 4xy + 5x_2^2$ to a diagonal form. We write

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

The symmetric matrix A can be diagonalized.

Quadratic forms: An example

Example

Let us determine the orthogonal matrix U which reduces the quadratic form $Q(\mathbf{x}) = 2x_1^2 + 4xy + 5x_2^2$ to a diagonal form. We write

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

The symmetric matrix A can be diagonalized. The eigenvalues of A are $\lambda_1 = 1$ and $\lambda_2 = 6$.

Quadratic forms: An example

Example

Let us determine the orthogonal matrix U which reduces the quadratic form $Q(\mathbf{x}) = 2x_1^2 + 4xy + 5x_2^2$ to a diagonal form. We write

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

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

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

Quadratic forms: An example

Example

Let us determine the orthogonal matrix U which reduces the quadratic form $Q(\mathbf{x}) = 2x_1^2 + 4xy + 5x_2^2$ to a diagonal form. We write

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

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

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

Hence $U = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}$.

Quadratic forms: An example

Example

Let us determine the orthogonal matrix U which reduces the quadratic form $Q(\mathbf{x}) = 2x_1^2 + 4xy + 5x_2^2$ to a diagonal form. We write

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

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

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

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:

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

Quadratic forms: An example

Example

Let us determine the orthogonal matrix U which reduces the quadratic form $Q(\mathbf{x}) = 2x_1^2 + 4xy + 5x_2^2$ to a diagonal form. We write

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

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

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

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 $U^t AU =$ diag $(1, 6)$.

Conic Sections

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

$$ax^2 + bxy + cy^2 + dx + ey + f = 0. \tag{10}$$

Conic Sections

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

$$ax^2 + bxy + cy^2 + dx + ey + f = 0. \tag{10}$$

It can be proved that this equation represents one of the following:
(i) the empty set

Conic Sections

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

$$ax^2 + bxy + cy^2 + dx + ey + f = 0. \tag{10}$$

It can be proved that this equation represents one of the following:
(i) the empty set
(ii) a single point

Conic Sections

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

$$ax^2 + bxy + cy^2 + dx + ey + f = 0. \tag{10}$$

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

Conic Sections

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

$$ax^2 + bxy + cy^2 + dx + ey + f = 0. \tag{10}$$

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

Conic Sections

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

$$ax^2 + bxy + cy^2 + dx + ey + f = 0. \tag{10}$$

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 or

Conic Sections

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

$$ax^2 + bxy + cy^2 + dx + ey + f = 0. \tag{10}$$

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 or
(vi) a parabola.
The second degree part of (10)

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

is a quadratic form.

Conic Sections

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

$$ax^2 + bxy + cy^2 + dx + ey + f = 0. \tag{10}$$

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 or
(vi) a parabola.
The second degree part of (10)

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

is a quadratic form. This determines the type of the conic.

Conic Sections

We can write the equation (10) into matrix form:

$$[x, y] \left[\begin{array}{cc} a & b/2 \\ b/2 & c \end{array} \right] \left[\begin{array}{c} x \\ y \end{array} \right] + [d, e] \left[\begin{array}{c} x \\ y \end{array} \right] + f = 0 \qquad (11)$$

Conic Sections

We can write the equation (10) into matrix form:

$$[x, y] \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + [d, e] \begin{bmatrix} x \\ y \end{bmatrix} + f = 0 \tag{11}$$

Write $A = \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix}$.

Conic Sections

We can write the equation (10) into matrix form:

$$[x, y] \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + [d, e] \begin{bmatrix} x \\ y \end{bmatrix} + f = 0 \tag{11}$$

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.

Conic Sections

We can write the equation (10) into matrix form:

$$[x, y] \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + [d, e] \begin{bmatrix} x \\ y \end{bmatrix} + f = 0 \tag{11}$$

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

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

to diagonalize the quadratic form $Q(x, y)$ to the diagonal form $\lambda_1 u^2 + \lambda_2 v_2^2$.

Conic Sections

We can write the equation (10) into matrix form:

$$[x, y] \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + [d, e] \begin{bmatrix} x \\ y \end{bmatrix} + f \quad = \quad 0 \tag{11}$$

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

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

to diagonalize the quadratic form $Q(x, y)$ to the diagonal form $\lambda_1 u^2 + \lambda_2 v_2^2$. 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

$$\begin{aligned} 0 &= [u, v] \text{ diag } (\lambda_1, \lambda_2)[u, v]^T + (BU)[u, v]^T + f \\ &= \lambda_1 u^2 + \lambda_2 v^2 + [d, e][\mathbf{v}_1, \mathbf{v}_2][u, v]^T + f. \end{aligned} \tag{12}$$

Conic Sections

If the conic determined by (12) is not degenerate i.e., not an empty set, a point, nor line(s) then the signs of λ_1 and λ_2 determine whether it is a parabola, an hyperbola or an ellipse.

Conic Sections

If the conic determined by (12) is not degenerate i.e., not an empty set, a point, nor line(s) then the signs of λ_1 and λ_2 determine whether it is a parabola, an hyperbola or an ellipse. The equation (10) will represent

Conic Sections

If the conic determined by (12) is not degenerate i.e., not an empty set, a point, nor line(s) then the signs of λ_1 and λ_2 determine whether it is a parabola, an hyperbola or an ellipse. The equation (10) will represent

(1) ellipse if $\lambda_1\lambda_2 > 0$

Conic Sections

If the conic determined by (12) is not degenerate i.e., not an empty set, a point, nor line(s) then the signs of λ_1 and λ_2 determine whether it is a parabola, an hyperbola or an ellipse. The equation (10) will represent

(1) ellipse if $\lambda_1\lambda_2 > 0$

(2) hyperbola if $\lambda_1\lambda_2 < 0$

Conic Sections

If the conic determined by (12) is not degenerate i.e., not an empty set, a point, nor line(s) then the signs of λ_1 and λ_2 determine whether it is a parabola, an hyperbola or an ellipse. The equation (10) will represent

(1) ellipse if $\lambda_1 \lambda_2 > 0$

(2) hyperbola if $\lambda_1 \lambda_2 < 0$

(3) parabola if $\lambda_1 \lambda_2 = 0$.

Conic Sections

If the conic determined by (12) is not degenerate i.e., not an empty set, a point, nor line(s) then the signs of λ_1 and λ_2 determine whether it is a parabola, an hyperbola or an ellipse. The equation (10) will represent

(1) ellipse if $\lambda_1\lambda_2 > 0$

(2) hyperbola if $\lambda_1\lambda_2 < 0$

(3) parabola if $\lambda_1\lambda_2 = 0$.

Conic Sections: Examples

Example

(1) $2x^2 + 4xy + 5y^2 + 4x + 13y - 1/4 = 0$.
We have earlier diagonalized the quadratic form $2x^2 + 4xy + 5y^2$.

Conic Sections: Examples

Example

(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 :*

Conic Sections: Examples

Example

(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 :*

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

Conic Sections: Examples

Example

(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 :*

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

Put $t = 1/\sqrt{5}$ *for convenience.*

Conic Sections: Examples

Example

(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 :*

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

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

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

Conic Sections: Examples

Example

(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 :*

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

Put $t = 1/\sqrt{5}$ *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., $\quad x = t(2u + v)$ *and* $y = t(-u + 2v)$.

Conic Sections: Examples

Example

(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 :*

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

Put $t = 1/\sqrt{5}$ *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*

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

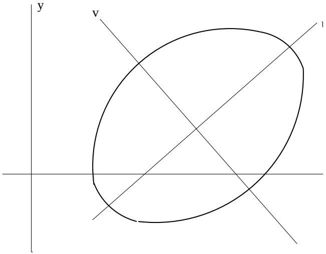
Complete the square to write this as

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

This is an equation of ellipse with center $(\frac{1}{2}\sqrt{5}, -\frac{1}{2}\sqrt{5})$ *in the uv-plane.*

Conic Sections: Examples

The u-axis and v-axis are determined by the eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$ as indicated in the following figure :



Conic Sections: Examples

Example

(2) $2x^2 - 4xy - y^2 - 4x + 10y - 13 = 0.$

Conic Sections: Examples

Example

(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.*

Conic Sections: Examples

Example

(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*

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

Conic Sections: Examples

Example

(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 $t = 1/\sqrt{5}$.

Conic Sections: Examples

Example

(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* $\lambda_1 = 3, \lambda_2 = -2$.

Conic Sections: Examples

Example

(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*

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

Conic Sections: Examples

Example

(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

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

Conic Sections: Examples

which is the same as

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

Conic Sections: Examples

which is the same as

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

Complete the square in u and v to get

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

Conic Sections: Examples

which is the same as

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

Complete the square in u and v to get

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

or

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

Conic Sections: Examples

which is the same as

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

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.

Conic Sections: Examples

which is the same as

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

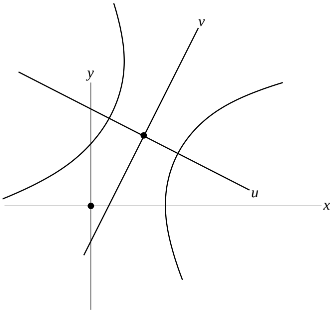
Complete the square in u and v to get

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

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.



Conic Sections: Examples

Example

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

Conic Sections: Examples

Example

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

The symmetric matrix for the quadratic part is $A = \begin{bmatrix} 9 & 12 \\ 12 & 16 \end{bmatrix}$.

Conic Sections: Examples

Example

(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* $\lambda_1 = 25, \lambda_2 = 0.$

Conic Sections: Examples

Example

(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* $a = 1/5$.

Conic Sections: Examples

Example

(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* $U = a \begin{bmatrix} 3 & -4 \\ 4 & 3 \end{bmatrix}$.

Conic Sections: Examples

Example

(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*

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

Conic Sections: Examples

Example

(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*

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

The equation in uv-plane is $u^2 + v = 0$.

Conic Sections: Examples

Example

(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*

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

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