Math 221: LINEAR ALGEBRA

Chapter 8. Orthogonality

§8-2. Orthogonal Diagonalization

Le Chen[1]

Emory University, 2021 Spring

(last updated on 01/25/2021)

[1] Slides are adapted from those by Karen Seyffarth from University of Calgary.

Orthogonal Matrices

Orthogonal Diagonalization and Symmetric Matrices

Quadratic Forms

Orthogonal Matrices

Orthogonal Diagonalization and Symmetric Matrices

Quadratic Forms

Orthogonal Matrices

Orthogonal Matrices

Definition

An $n \times n$ matrix A is a orthogonal if its inverse is equal to its transpose, i.e., $A^{-1} = A^T$.

Orthogonal Matrices

Definition

An $n \times n$ matrix A is a orthogonal if its inverse is equal to its transpose, i.e., $A^{-1} = A^T$.

Example

$$\frac{\sqrt{2}}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \quad \text{and} \quad \frac{1}{7}\begin{bmatrix} 2 & 6 & -3 \\ 3 & 2 & 6 \\ -6 & 3 & 2 \end{bmatrix}$$

are orthogonal matrices (verify).

Theorem

The following are equivalent for an $n \times n$ matrix A.

1. A is orthogonal.
2. The rows of A are orthonormal.
3. The columns of A are orthonormal.

Theorem

The following are equivalent for an $n \times n$ matrix A.

1. A is orthogonal.
2. The rows of A are orthonormal.
3. The columns of A are orthonormal.

Proof.

"(1) $\iff$ (3)": Write $A = [\vec{a}_1, \cdots \vec{a}_n]$.

$$A \text{ is orthogonal} \iff A^T A = I_n \iff \begin{pmatrix} \vec{a}_1^T \\ \vdots \\ \vec{a}_n \end{pmatrix} [\vec{a}_1, \cdots \vec{a}_n] = I_n$$

$$\iff \begin{bmatrix} \vec{a}_1 \cdot \vec{a}_1 & \vec{a}_1 \cdot \vec{a}_2 & \cdots & \vec{a}_1 \cdot \vec{a}_n \\ \vec{a}_2 \cdot \vec{a}_1 & \vec{a}_2 \cdot \vec{a}_2 & \cdots & \vec{a}_2 \cdot \vec{a}_n \\ \vdots & \vdots & \ddots & \vdots \\ \vec{a}_n \cdot \vec{a}_1 & \vec{a}_n \cdot \vec{a}_2 & \cdots & \vec{a}_n \cdot \vec{a}_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

"(1) $\iff$ (2)": Similarly (Try it yourself). ∎

Example

$$A = \begin{bmatrix} 2 & 1 & -2 \\ -2 & 1 & 2 \\ 1 & 0 & 8 \end{bmatrix}$$

has orthogonal columns, but its rows are not orthogonal (verify).

Example

$$A = \begin{bmatrix} 2 & 1 & -2 \\ -2 & 1 & 2 \\ 1 & 0 & 8 \end{bmatrix}$$

has orthogonal columns, but its rows are not orthogonal (verify).

Normalizing the columns of A gives us the matrix

$$A' = \begin{bmatrix} 2/3 & 1/\sqrt{2} & -1/3\sqrt{2} \\ -2/3 & 1/\sqrt{2} & 1/3\sqrt{2} \\ 1/3 & 0 & 4/3\sqrt{2} \end{bmatrix},$$

which has orthonormal columns. Therefore, A' is an orthogonal matrix.

Example

$$A = \begin{bmatrix} 2 & 1 & -2 \\ -2 & 1 & 2 \\ 1 & 0 & 8 \end{bmatrix}$$

has orthogonal columns, but its rows are not orthogonal (verify).

Normalizing the columns of A gives us the matrix

$$A' = \begin{bmatrix} 2/3 & 1/\sqrt{2} & -1/3\sqrt{2} \\ -2/3 & 1/\sqrt{2} & 1/3\sqrt{2} \\ 1/3 & 0 & 4/3\sqrt{2} \end{bmatrix},$$

which has orthonormal columns. Therefore, A' is an orthogonal matrix.

If an $n \times n$ matrix has orthogonal rows (columns), then normalizing the rows (columns) results in an orthogonal matrix.

Example (Orthogonal Matrices: Products and Inverses)

Suppose A and B are orthogonal matrices.

1. Since
$$(AB)(B^TA^T) = A(BB^T)A^T = AA^T = I.$$
and AB is square, $B^TA^T = (AB)^T$ is the inverse of AB, so AB is invertible, and $(AB)^{-1} = (AB)^T$. Therefore, AB is orthogonal.

Example (Orthogonal Matrices: Products and Inverses)

Suppose A and B are orthogonal matrices.

1. Since
$$(AB)(B^TA^T) = A(BB^T)A^T = AA^T = I.$$
and AB is square, $B^TA^T = (AB)^T$ is the inverse of AB, so AB is invertible, and $(AB)^{-1} = (AB)^T$. Therefore, AB is orthogonal.

2. $A^{-1} = A^T$ is also orthogonal, since
$$(A^{-1})^{-1} = A = (A^T)^T = (A^{-1})^T.$$

Example (Orthogonal Matrices: Products and Inverses)

Suppose A and B are orthogonal matrices.

1. Since
$$(AB)(B^TA^T) = A(BB^T)A^T = AA^T = I.$$
and AB is square, $B^TA^T = (AB)^T$ is the inverse of AB, so AB is invertible, and $(AB)^{-1} = (AB)^T$. Therefore, AB is orthogonal.
2. $A^{-1} = A^T$ is also orthogonal, since
$$(A^{-1})^{-1} = A = (A^T)^T = (A^{-1})^T.$$

Remark (Summary)

If A and B are orthogonal matrices, then AB is orthogonal and A^{-1} is orthogonal.

Orthogonal Diagonalization and Symmetric Matrices

Orthogonal Diagonalization and Symmetric Matrices

Definition

An $n \times n$ matrix A is orthogonally diagonalizable if there exists an orthogonal matrix, P, so that $P^{-1}AP = P^{T}AP$ is diagonal.

Orthogonal Diagonalization and Symmetric Matrices

Definition

An $n \times n$ matrix A is orthogonally diagonalizable if there exists an orthogonal matrix, P, so that $P^{-1}AP = P^{T}AP$ is diagonal.

Theorem (Principal Axis Theorem)

Let A be an $n \times n$ matrix. The following conditions are equivalent.

1. A has an orthonormal set of n eigenvectors.
2. A is orthogonally diagonalizable.
3. A is symmetric.

Proof. ((1) $\Rightarrow$ (2))

Suppose $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n\}$ is an orthonormal set of n eigenvectors of A. Then $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n\}$ is a basis of $\mathbb{R}^n$, and hence $P = \left[\begin{array}{cccc} \vec{x}_1 & \vec{x}_2 & \cdots & \vec{x}_n \end{array} \right]$ is an orthogonal matrix such that $P^{-1}AP = P^TAP$ is a diagonal matrix. Therefore A is orthogonally diagonalizable.

Proof. ((2) $\Rightarrow$ (1))

Suppose that A is orthogonally diagonalizable. Then there exists an orthogonal matrix P such that P^TAP is a diagonal matrix. If P has columns $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n$, then $B = \{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n\}$ is a set of n orthonormal vectors in $\mathbb{R}^n$. Since B is orthogonal, B is independent; furthermore, since $|B| = n = \dim(\mathbb{R}^n)$, B spans $\mathbb{R}^n$ and is therefore a basis of $\mathbb{R}^n$.
Let $P^TAP = \text{diag}(\ell_1, \ell_2, \ldots, \ell_n) = D$. Then $AP = PD$, so

$$A\begin{bmatrix} \vec{x_1} & \vec{x_2} & \cdots & \vec{x_n} \end{bmatrix} = \begin{bmatrix} \vec{x_1} & \vec{x_2} & \cdots & \vec{x_n} \end{bmatrix}\begin{bmatrix} \ell_1 & 0 & \cdots & 0 \\ 0 & \ell_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \ell_n \end{bmatrix}$$

$$\begin{bmatrix} A\vec{x_1} & A\vec{x_2} & \cdots & A\vec{x_n} \end{bmatrix} = \begin{bmatrix} \ell_1\vec{x_1} & \ell_2\vec{x_2} & \cdots & \ell_n\vec{x_n} \end{bmatrix}$$

Thus $A\vec{x}_i = \ell_i\vec{x}_i$ for each i, $1 \leq i \leq n$, implying that B consists of eigenvectors of A. Therefore, A has an orthonormal set of n eigenvectors.

Proof. $((2) \Rightarrow (3))$

Suppose A is orthogonally diagonalizable, that D is a diagonal matrix, and that P is an orthogonal matrix so that $P^{-1}AP = D$.

Proof. ((2) $\Rightarrow$ (3))

Suppose A is orthogonally diagonalizable, that D is a diagonal matrix, and that P is an orthogonal matrix so that $P^{-1}AP = D$. Then $P^{-1}AP = P^{T}AP$, so

$$A = PDP^{T}.$$

Proof. $((2) \Rightarrow (3))$

Suppose A is orthogonally diagonalizable, that D is a diagonal matrix, and that P is an orthogonal matrix so that $P^{-1}AP = D$. Then $P^{-1}AP = P^TAP$, so

$$A = PDP^T.$$

Taking transposes of both sides of the equation:

$$\begin{aligned} A^T = (PDP^T)^T &= (P^T)^T D^T P^T \\ &= PD^TP^T \quad \text{(since } (P^T)^T = P\text{)} \\ &= PDP^T \quad \text{(since } D^T = D\text{)} \\ &= A. \end{aligned}$$

Proof. ((2) $\Rightarrow$ (3))

Suppose A is orthogonally diagonalizable, that D is a diagonal matrix, and that P is an orthogonal matrix so that $P^{-1}AP = D$. Then $P^{-1}AP = P^{T}AP$, so

$$A = PDP^{T}.$$

Taking transposes of both sides of the equation:

$$\begin{aligned} A^{T} = (PDP^{T})^{T} &= (P^{T})^{T}D^{T}P^{T} \\ &= PD^{T}P^{T} \quad \text{(since } (P^{T})^{T} = P\text{)} \\ &= PDP^{T} \quad \text{(since } D^{T} = D\text{)} \\ &= A. \end{aligned}$$

Since $A^{T} = A$, A is symmetric.

Proof. ((3) $\Rightarrow$ (2))

If A is $n \times n$ symmetric matrix, we will prove by induction on n that A is orthogonal diagonalizable. If $n = 1$, A is already diagonalizable. If $n \geq 2$, assume that (3)$\Rightarrow$(2) for all $(n-1) \times (n-1)$ symmetric matrix.

First we know that all eigenvalues are real (because A is symmetric). Let λ_1 be one real eigenvalue and $\vec{x}_1$ be the normalized eigenvector. We can extend $\{\vec{x}_1\}$ to an orthonormal basis of $\mathbb{R}^n$, say $\{\vec{x}_1, \cdots, \vec{x}_n\}$ by adding vectors. Let $P_1 = [\vec{x}_1, \cdots, \vec{x}_n]$. So P is orthogonal.
Now we can apply the technical lemma proved in Section 5.5 to see that

$$P_1^T A P_1 = \begin{bmatrix} \lambda_1 & B \\ \vec{0} & A_1 \end{bmatrix}.$$

Since LHS is symmetric, so does the RHS. This implies that $B = O$ and A_1 is symmetric.

Proof. ((3) ⇒ (2) – continued)

By induction assumption, A_1 is orthogonal diagonalizable, i.e., for some orthogonal matrix Q and diagonal matrix D, $A_1 = QDQ^T$. Hence,

$$P_1^T A P_1 = \begin{bmatrix} \lambda_1 & \vec{0}^T \\ \vec{0} & QDQ^T \end{bmatrix} = \begin{bmatrix} 1 & \vec{0}^T \\ \vec{0} & Q \end{bmatrix} \begin{bmatrix} \lambda_1 & \vec{0}^T \\ \vec{0} & D \end{bmatrix} \begin{bmatrix} 1 & \vec{0}^T \\ \vec{0} & Q^T \end{bmatrix}$$

which is nothing but

$$\begin{aligned} A &= P_1 \begin{bmatrix} 1 & \vec{0}^T \\ \vec{0} & Q \end{bmatrix} \begin{bmatrix} \lambda_1 & \vec{0}^T \\ \vec{0} & D \end{bmatrix} \begin{bmatrix} 1 & \vec{0}^T \\ \vec{0} & Q^T \end{bmatrix} P_1^T \\ &= \left(P_1 \begin{bmatrix} 1 & \vec{0}^T \\ \vec{0} & Q \end{bmatrix} \right) \begin{bmatrix} \lambda_1 & \vec{0}^T \\ \vec{0} & D \end{bmatrix} \left(P_1 \begin{bmatrix} 1 & \vec{0}^T \\ \vec{0} & Q \end{bmatrix} \right)^T . \end{aligned}$$

Finally, it is ready to verify that the matrix

$$P_1 \begin{bmatrix} 1 & \vec{0}^T \\ \vec{0} & Q \end{bmatrix}$$

is a diagonal matrix. This complete the proof of the theorem. ■

Definition

Let A be an $n \times n$ matrix. A set of n orthonormal eigenvectors of A is called a set of principal axes of A.

Problem

Orthogonally diagonalize the matrix

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

Problem

Orthogonally diagonalize the matrix

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

Solution

- $c_A(x) = (x+3)(x-3)^2$, so A has eigenvalues $\lambda_1 = 3$ of multiplicity two, and $\lambda_2 = -3$.

Problem

Orthogonally diagonalize the matrix

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

Solution

- $c_A(x) = (x+3)(x-3)^2$, so A has eigenvalues $\lambda_1 = 3$ of multiplicity two, and $\lambda_2 = -3$.
- $\{\vec{x}_1, \vec{x}_2\}$ is a basis of $E_3(A)$, where $\vec{x}_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ and $\vec{x}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$.

Problem

Orthogonally diagonalize the matrix

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

Solution

- $c_A(x) = (x+3)(x-3)^2$, so A has eigenvalues $\lambda_1 = 3$ of multiplicity two, and $\lambda_2 = -3$.
- $\{\vec{x}_1, \vec{x}_2\}$ is a basis of $E_3(A)$, where $\vec{x}_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ and $\vec{x}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$.
- $\{\vec{x}_3\}$ is a basis of $E_{-3}(A)$, where $\vec{x}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

Problem

Orthogonally diagonalize the matrix

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

Solution

- $c_A(x) = (x+3)(x-3)^2$, so A has eigenvalues $\lambda_1 = 3$ of multiplicity two, and $\lambda_2 = -3$.
- $\{\vec{x}_1, \vec{x}_2\}$ is a basis of $E_3(A)$, where $\vec{x}_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ and $\vec{x}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$.
- $\{\vec{x}_3\}$ is a basis of $E_{-3}(A)$, where $\vec{x}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.
- $\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$ a linearly independent set of eigenvectors of A, and a basis of $\mathbb{R}^3$.

Solution (continued)

- Orthogonalize $\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$ using the Gram-Schmidt orthogonalization algorithm.

Solution (continued)

- Orthogonalize $\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$ using the Gram-Schmidt orthogonalization algorithm.
- Let $\vec{f_1} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \vec{f_2} = \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}$ and $\vec{f_3} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Then $\{\vec{f_1}, \vec{f_2}, \vec{f_3}\}$ is an orthogonal basis of $\mathbb{R}^3$ consisting of eigenvectors of A.

Solution (continued)

- Orthogonalize $\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$ using the Gram-Schmidt orthogonalization algorithm.
- Let $\vec{f_1} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \vec{f_2} = \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}$ and $\vec{f_3} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Then $\{\vec{f_1}, \vec{f_2}, \vec{f_3}\}$ is an orthogonal basis of $\mathbb{R}^3$ consisting of eigenvectors of A.
- Since $||\vec{f_1}|| = \sqrt{2}$, $||\vec{f_2}|| = \sqrt{6}$, and $||\vec{f_3}|| = \sqrt{3}$,

$$P = \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \end{bmatrix}$$

is an orthogonal diagonalizing matrix of A,

Solution (continued)

- Orthogonalize $\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$ using the Gram-Schmidt orthogonalization algorithm.
- Let $\vec{f_1} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$, $\vec{f_2} = \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}$ and $\vec{f_3} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Then $\{\vec{f_1}, \vec{f_2}, \vec{f_3}\}$ is an orthogonal basis of $\mathbb{R}^3$ consisting of eigenvectors of A.
- Since $||\vec{f_1}|| = \sqrt{2}$, $||\vec{f_2}|| = \sqrt{6}$, and $||\vec{f_3}|| = \sqrt{3}$,

$$P = \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \end{bmatrix}$$

is an orthogonal diagonalizing matrix of A, and

$$P^T AP = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -3 \end{bmatrix}.$$

■

Theorem

If A is a symmetric matrix, then the eigenvectors of A corresponding to distinct eigenvalues are orthogonal.

Theorem

If A is a symmetric matrix, then the eigenvectors of A corresponding to distinct eigenvalues are orthogonal.

Proof.

Suppose λ and μ are eigenvalues of A, $\lambda \neq \mu$, and let $\vec{x}$ and $\vec{y}$, respectively, be corresponding eigenvectors, i.e., $A\vec{x} = \lambda\vec{x}$ and $A\vec{y} = \mu\vec{y}$. Consider $(\lambda - \mu)\vec{x} \cdot \vec{y}$.

Theorem

If A is a symmetric matrix, then the eigenvectors of A corresponding to distinct eigenvalues are orthogonal.

Proof.

Suppose λ and μ are eigenvalues of A, $\lambda \neq \mu$, and let $\vec{x}$ and $\vec{y}$, respectively, be corresponding eigenvectors, i.e., $A\vec{x} = \lambda\vec{x}$ and $A\vec{y} = \mu\vec{y}$. Consider $(\lambda - \mu)\vec{x} \cdot \vec{y}$.

$$(\lambda - \mu)\vec{x} \cdot \vec{y} = \lambda(\vec{x} \cdot \vec{y}) - \mu(\vec{x} \cdot \vec{y})$$

Theorem

If A is a symmetric matrix, then the eigenvectors of A corresponding to distinct eigenvalues are orthogonal.

Proof.

Suppose λ and μ are eigenvalues of A, $\lambda \neq \mu$, and let $\vec{x}$ and $\vec{y}$, respectively, be corresponding eigenvectors, i.e., $A\vec{x} = \lambda\vec{x}$ and $A\vec{y} = \mu\vec{y}$. Consider $(\lambda - \mu)\vec{x} \cdot \vec{y}$.

$$\begin{aligned}(\lambda - \mu)\vec{x} \cdot \vec{y} &= \lambda(\vec{x} \cdot \vec{y}) - \mu(\vec{x} \cdot \vec{y}) \\ &= (\lambda\vec{x}) \cdot \vec{y} - \vec{x} \cdot (\mu\vec{y})\end{aligned}$$

Theorem

If A is a symmetric matrix, then the eigenvectors of A corresponding to distinct eigenvalues are orthogonal.

Proof.

Suppose λ and μ are eigenvalues of A, $\lambda \neq \mu$, and let $\vec{x}$ and $\vec{y}$, respectively, be corresponding eigenvectors, i.e., $A\vec{x} = \lambda\vec{x}$ and $A\vec{y} = \mu\vec{y}$. Consider $(\lambda - \mu)\vec{x} \cdot \vec{y}$.

$$
\begin{aligned}
(\lambda - \mu)\vec{x} \cdot \vec{y} &= \lambda(\vec{x} \cdot \vec{y}) - \mu(\vec{x} \cdot \vec{y}) \\
&= (\lambda\vec{x}) \cdot \vec{y} - \vec{x} \cdot (\mu\vec{y}) \\
&= (A\vec{x}) \cdot \vec{y} - \vec{x} \cdot (A\vec{y})
\end{aligned}
$$

Theorem

If A is a symmetric matrix, then the eigenvectors of A corresponding to distinct eigenvalues are orthogonal.

Proof.

Suppose λ and μ are eigenvalues of A, $\lambda \neq \mu$, and let $\vec{x}$ and $\vec{y}$, respectively, be corresponding eigenvectors, i.e., $A\vec{x} = \lambda\vec{x}$ and $A\vec{y} = \mu\vec{y}$. Consider $(\lambda - \mu)\vec{x} \cdot \vec{y}$.

$$\begin{aligned}
(\lambda - \mu)\vec{x} \cdot \vec{y} &= \lambda(\vec{x} \cdot \vec{y}) - \mu(\vec{x} \cdot \vec{y}) \\
&= (\lambda\vec{x}) \cdot \vec{y} - \vec{x} \cdot (\mu\vec{y}) \\
&= (A\vec{x}) \cdot \vec{y} - \vec{x} \cdot (A\vec{y}) \\
&= (A\vec{x})^T\vec{y} - \vec{x}^T(A\vec{y})
\end{aligned}$$

Theorem

If A is a symmetric matrix, then the eigenvectors of A corresponding to distinct eigenvalues are orthogonal.

Proof.

Suppose λ and μ are eigenvalues of A, $\lambda \neq \mu$, and let $\vec{x}$ and $\vec{y}$, respectively, be corresponding eigenvectors, i.e., $A\vec{x} = \lambda\vec{x}$ and $A\vec{y} = \mu\vec{y}$. Consider $(\lambda - \mu)\vec{x} \cdot \vec{y}$.

$$\begin{aligned}
(\lambda - \mu)\vec{x} \cdot \vec{y} &= \lambda(\vec{x} \cdot \vec{y}) - \mu(\vec{x} \cdot \vec{y}) \\
&= (\lambda\vec{x}) \cdot \vec{y} - \vec{x} \cdot (\mu\vec{y}) \\
&= (A\vec{x}) \cdot \vec{y} - \vec{x} \cdot (A\vec{y}) \\
&= (A\vec{x})^T\vec{y} - \vec{x}^T(A\vec{y}) \\
&= \vec{x}^T A^T \vec{y} - \vec{x}^T A\vec{y}
\end{aligned}$$

Theorem

If A is a symmetric matrix, then the eigenvectors of A corresponding to distinct eigenvalues are orthogonal.

Proof.

Suppose λ and μ are eigenvalues of A, $\lambda \neq \mu$, and let $\vec{x}$ and $\vec{y}$, respectively, be corresponding eigenvectors, i.e., $A\vec{x} = \lambda\vec{x}$ and $A\vec{y} = \mu\vec{y}$. Consider $(\lambda - \mu)\vec{x} \cdot \vec{y}$.

$$\begin{aligned}
(\lambda - \mu)\vec{x} \cdot \vec{y} &= \lambda(\vec{x} \cdot \vec{y}) - \mu(\vec{x} \cdot \vec{y}) \\
&= (\lambda\vec{x}) \cdot \vec{y} - \vec{x} \cdot (\mu\vec{y}) \\
&= (A\vec{x}) \cdot \vec{y} - \vec{x} \cdot (A\vec{y}) \\
&= (A\vec{x})^T\vec{y} - \vec{x}^T(A\vec{y}) \\
&= \vec{x}^T A^T \vec{y} - \vec{x}^T A\vec{y} \\
&= \vec{x}^T A\vec{y} - \vec{x}^T A\vec{y} \qquad \text{since A is symmetric}
\end{aligned}$$

Theorem

If A is a symmetric matrix, then the eigenvectors of A corresponding to distinct eigenvalues are orthogonal.

Proof.

Suppose λ and μ are eigenvalues of A, $\lambda \neq \mu$, and let $\vec{x}$ and $\vec{y}$, respectively, be corresponding eigenvectors, i.e., $A\vec{x} = \lambda\vec{x}$ and $A\vec{y} = \mu\vec{y}$. Consider $(\lambda - \mu)\vec{x} \cdot \vec{y}$.

$$\begin{aligned}
(\lambda - \mu)\vec{x} \cdot \vec{y} &= \lambda(\vec{x} \cdot \vec{y}) - \mu(\vec{x} \cdot \vec{y}) \\
&= (\lambda\vec{x}) \cdot \vec{y} - \vec{x} \cdot (\mu\vec{y}) \\
&= (A\vec{x}) \cdot \vec{y} - \vec{x} \cdot (A\vec{y}) \\
&= (A\vec{x})^T\vec{y} - \vec{x}^T(A\vec{y}) \\
&= \vec{x}^T A^T \vec{y} - \vec{x}^T A\vec{y} \\
&= \vec{x}^T A\vec{y} - \vec{x}^T A\vec{y} \qquad \text{since A is symmetric} \\
&= 0.
\end{aligned}$$

Theorem

If A is a symmetric matrix, then the eigenvectors of A corresponding to distinct eigenvalues are orthogonal.

Proof.

Suppose λ and μ are eigenvalues of A, $\lambda \neq \mu$, and let $\vec{x}$ and $\vec{y}$, respectively, be corresponding eigenvectors, i.e., $A\vec{x} = \lambda\vec{x}$ and $A\vec{y} = \mu\vec{y}$. Consider $(\lambda - \mu)\vec{x} \cdot \vec{y}$.

$$\begin{aligned}
(\lambda - \mu)\vec{x} \cdot \vec{y} &= \lambda(\vec{x} \cdot \vec{y}) - \mu(\vec{x} \cdot \vec{y}) \\
&= (\lambda\vec{x}) \cdot \vec{y} - \vec{x} \cdot (\mu\vec{y}) \\
&= (A\vec{x}) \cdot \vec{y} - \vec{x} \cdot (A\vec{y}) \\
&= (A\vec{x})^T\vec{y} - \vec{x}^T(A\vec{y}) \\
&= \vec{x}^T A^T \vec{y} - \vec{x}^T A\vec{y} \\
&= \vec{x}^T A\vec{y} - \vec{x}^T A\vec{y} \qquad \text{since A is symmetric} \\
&= 0.
\end{aligned}$$

Since $\lambda \neq \mu$, $\lambda - \mu \neq 0$, and therefore $\vec{x} \cdot \vec{y} = 0$, i.e., $\vec{x}$ and $\vec{y}$ are orthogonal.
■

Remark (Diagonalizing a Symmetric Matrix)

Let A be a symmetric $n \times n$ matrix.

1. Find the characteristic polynomial and distinct eigenvalues of A.
2. For each distinct eigenvalue λ of A, find an orthonormal basis of $E_A(\lambda)$, the eigenspace of A corresponding to λ. This requires using the Gram-Schmidt orthogonalization algorithm when $\dim(E_A(\lambda)) \geq 2$.
3. By the previous theorem, the eigenvectors of distinct eigenvalues produce orthogonal eigenvectors, so the result is an orthonormal basis of $\mathbb{R}^n$.

Problem

Orthogonally diagonalize the matrix

$$A = \begin{bmatrix} 3 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 2 \end{bmatrix}.$$

Problem

Orthogonally diagonalize the matrix

$$A = \begin{bmatrix} 3 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 2 \end{bmatrix}.$$

Solution

1. Since row sum is 5, $\lambda_1 = 5$ is one eigenvalue, corresponding eigenvector should be $(1, 1, 1)^T$. After normalization it should be

$$\vec{v}_1 = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}$$

Solution (continued)

2. Since last two rows are identical, $\det(A) = 0$, so $\lambda_2 = 0$ is another eigenvalue, corresponding eigenvector should be $(0, 1, -1)^T$. After normalization it should be

$$\vec{v}_2 = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

Solution (continued)

3. Since $\text{tr}(A) = 7 = \lambda_1 + \lambda_2 + \lambda_3$, we see that $\lambda_3 = 7 - 5 - 0 = 2$. Its eigenvector should be orthogonal to both $\vec{v}_1$ and $\vec{v}_2$, hence, $\vec{v}_3 = (2, -1, -1)$. After normalization,

$$\begin{pmatrix} \frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$$

Solution (continued)

3. Since $\text{tr}(A) = 7 = \lambda_1 + \lambda_2 + \lambda_3$, we see that $\lambda_3 = 7 - 5 - 0 = 2$. Its eigenvector should be orthogonal to both $\vec{v}_1$ and $\vec{v}_2$, hence, $\vec{v}_3 = (2, -1, -1)$. After normalization,

$$\begin{pmatrix} \frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$$

Hence, we have

$$\begin{bmatrix} 3 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{bmatrix} \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$

$\blacksquare$

Orthogonal Matrices

Orthogonal Diagonalization and Symmetric Matrices

Quadratic Forms

Quadratic Forms

Definitions

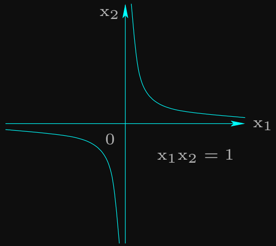
Let q be a real polynomial in variables x_1 and x_2 such that

$$q(x_1, x_2) = ax_1^2 + bx_1x_2 + cx_2^2.$$

Then q is called a quadratic form in variables x_1 and x_2. The term bx_1x_2 is called the cross term. The graph of the equation $q(x_1, x_2) = 1$, is call a conic in variables x_1 and x_2.

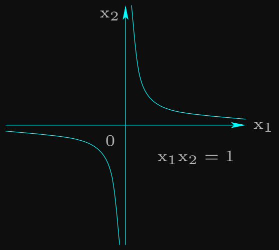
Example

Below is the graph of the equation $x_1x_2 = 1$.



Example

Below is the graph of the equation $x_1x_2 = 1$.



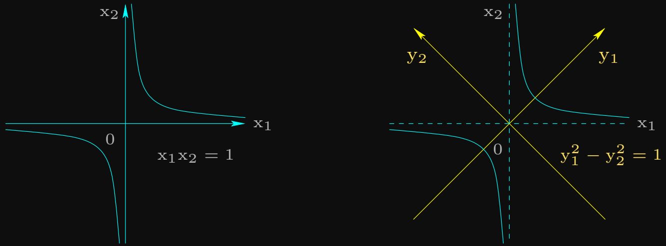
Let y_1 and y_2 be new variables such that

$$x_1 = y_1 + y_2 \quad \text{and} \quad x_2 = y_1 - y_2,$$

i.e., $y_1 = \frac{x_1+x_2}{2}$ and $y_2 = \frac{x_1-x_2}{2}$. Then $x_1x_2 = y_1^2 - y_2^2$, and $y_1^2 - y_2^2$ is a quadratic form with no cross terms, called a diagonal quadratic form;

Example

Below is the graph of the equation $x_1x_2 = 1$.



Let y_1 and y_2 be new variables such that

$$x_1 = y_1 + y_2 \quad \text{and} \quad x_2 = y_1 - y_2,$$

i.e., $y_1 = \frac{x_1+x_2}{2}$ and $y_2 = \frac{x_1-x_2}{2}$. Then $x_1x_2 = y_1^2 - y_2^2$, and $y_1^2 - y_2^2$ is a quadratic form with no cross terms, called a diagonal quadratic form; y_1 and y_2 are called principal axes of the quadratic form x_1x_2.

Principal axes of a quadratic form can be found by using orthogonal diagonalization.

Principal axes of a quadratic form can be found by using orthogonal diagonalization.

Problem

Find principal axes of the quadratic form $q(x_1, x_2) = x_1^2 + 6x_1x_2 + x_2^2$, and transform $q(x_1, x_2)$ into a diagonal quadratic form.

Principal axes of a quadratic form can be found by using orthogonal diagonalization.

Problem

Find principal axes of the quadratic form $q(x_1, x_2) = x_1^2 + 6x_1x_2 + x_2^2$, and transform $q(x_1, x_2)$ into a diagonal quadratic form.

Solution

Express $q(x_1, x_2)$ as a matrix product:

$$q(x_1, x_2) = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & 6 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \quad (1)$$

We want a 2×2 symmetric matrix. Since $6x_1x_2 = 3x_1x_2 + 3x_2x_1$, we can rewrite (1) as

$$q(x_1, x_2) = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \quad (2)$$

Setting $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and $A = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}$, $q(x_1, x_2) = \vec{x}^T A \vec{x}$.

We now orthogonally diagonalize A.

Solution (continued)

$$c_A(z) = \begin{vmatrix} z-1 & -3 \\ -3 & z-1 \end{vmatrix} = (z-4)(z+2),$$

so A has eigenvalues $\lambda_1 = 4$ and $\lambda_2 = -2$.

Solution (continued)

$$c_A(z) = \begin{vmatrix} z-1 & -3 \\ -3 & z-1 \end{vmatrix} = (z-4)(z+2),$$

so A has eigenvalues $\lambda_1 = 4$ and $\lambda_2 = -2$.

$$\vec{z}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \vec{z}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

are eigenvectors corresponding to $\lambda_1 = 4$ and $\lambda_2 = -2$, respectively. Normalizing these eigenvectors gives us the orthogonal matrix

$$P = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \ \text{ such that } \ P^T A P = \begin{bmatrix} 4 & 0 \\ 0 & -2 \end{bmatrix} = D.$$

Solution (continued)

$$c_A(z) = \begin{vmatrix} z-1 & -3 \\ -3 & z-1 \end{vmatrix} = (z-4)(z+2),$$

so A has eigenvalues $\lambda_1 = 4$ and $\lambda_2 = -2$.

$$\vec{z}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \vec{z}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

are eigenvectors corresponding to $\lambda_1 = 4$ and $\lambda_2 = -2$, respectively. Normalizing these eigenvectors gives us the orthogonal matrix

$$P = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \text{ such that } P^T A P = \begin{bmatrix} 4 & 0 \\ 0 & -2 \end{bmatrix} = D.$$

Thus $A = PDP^T$, and

$$q(x_1, x_2) = \vec{x}^T A \vec{x} = \vec{x}^T (PDP^T)\vec{x} = (\vec{x}^T P) D (P^T \vec{x}) = (P^T \vec{x})^T D (P^T \vec{x}).$$

Solution (continued)

Let

$$\vec{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = P^T\vec{x} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} x_1 + x_2 \\ x_2 - x_1 \end{bmatrix}.$$

Then

$$q(y_1, y_2) = \vec{y}^T D\vec{y} = \begin{bmatrix} y_1 & y_2 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = 4y_1^2 - 2y_2^2.$$

Solution (continued)

Let

$$\vec{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = P^T\vec{x} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} x_1 + x_2 \\ x_2 - x_1 \end{bmatrix}.$$

Then

$$q(y_1, y_2) = \vec{y}^T D\vec{y} = \begin{bmatrix} y_1 & y_2 \end{bmatrix}\begin{bmatrix} 4 & 0 \\ 0 & -2 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = 4y_1^2 - 2y_2^2.$$

Therefore, the principal axes of $q(x_1, x_2) = x_1^2 + 6x_1x_2 + x_2^2$ are

$$y_1 = \frac{1}{\sqrt{2}}(x_1 + x_2)$$

and

$$y_2 = \frac{1}{\sqrt{2}}(x_2 - x_1),$$

yielding the diagonal quadratic form

$$q(y_1, y_2) = 4y_1^2 - 2y_2^2.$$

■

Problem

Find principal axes of the quadratic form

$$q(x_1, x_2) = 7x_1^2 - 4x_1x_2 + 4x_2^2,$$

and transform $q(x_1, x_2)$ into a diagonal quadratic form.

Problem

Find principal axes of the quadratic form

$$q(x_1, x_2) = 7x_1^2 - 4x_1x_2 + 4x_2^2,$$

and transform $q(x_1, x_2)$ into a diagonal quadratic form.

Solution (Final Answer)

$q(x_1, x_2)$ has principal axes

$$\begin{aligned} y_1 &= \frac{1}{\sqrt{5}}(-2x_1 + x_2), \\ y_2 &= \frac{1}{\sqrt{5}}(x_1 + 2x_2). \end{aligned}$$

yielding the diagonal quadratic form

$$q(y_1, y_2) = 8y_1^2 + 3y_2^2.$$

Theorem (Triangulation Theorem – Schur Decomposition)

Let A be an $n \times n$ matrix with n real eigenvalues. Then there exists an orthogonal matrix P such that $P^T AP$ is upper triangular.

Corollary

Let A be an $n \times n$ matrix with real eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, not necessarily distinct. Then $\det(A) = \lambda_1 \lambda_2 \cdots \lambda_n$ and $\operatorname{tr}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$.

Corollary

Let A be an $n \times n$ matrix with real eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, not necessarily distinct. Then $\det(A) = \lambda_1 \lambda_2 \cdots \lambda_n$ and $\operatorname{tr}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$.

Proof.

By the theorem, there exists an orthogonal matrix P such that $P^T A P = U$, where U is an upper triangular matrix. Since P is orthogonal, $P^T = P^{-1}$, so U is similar to A; thus the eigenvalues of U are $\lambda_1, \lambda_2, \ldots, \lambda_n$. Furthermore, since U is (upper) triangular, the entries on the main diagonal of U are its eigenvalues, so $\det(U) = \lambda_1 \lambda_2 \cdots \lambda_n$ and $\operatorname{tr}(U) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$. Since U and A are similar, $\det(A) = \det(U)$ and $\operatorname{tr}(A) = \operatorname{tr}(U)$, and the result follow. ■