Mathematics for Computer Science
Linear Algebra (Part 2)
Eigendecomposition

Karl Southern

Durham University

Febuary 3rd, 2025

Thanks to Andrei Krokhin and William Moses for use of some slides.

Outline

Recap of Last Two Weeks

- For an $n \times n$ matrix, a non-zero vector $\mathbf{x} \in \mathbb{R}^n$ is called an eigenvector of A if $A\mathbf{x} = \lambda\mathbf{x}$.
- In this case, λ is called an eigenvalue of A, and $\mathbf{x}$ is an eigenvector corresponding to λ.

Recap of Last Two Weeks

- For an $n \times n$ matrix, a non-zero vector $\mathbf{x} \in \mathbb{R}^n$ is called an eigenvector of A if $A\mathbf{x} = \lambda\mathbf{x}$.
- In this case, λ is called an eigenvalue of A, and $\mathbf{x}$ is an eigenvector corresponding to λ.
- The polynomial $det(\lambda I - A)$ is called the characteristic polynomial of A and the equation $det(\lambda I - A) = 0$ the characteristic equation of A.
- The eigenvalues of A are the solutions of $det(\lambda I - A) = 0$.

Recap of Last Two Weeks

- For an $n \times n$ matrix, a non-zero vector $\mathbf{x} \in \mathbb{R}^n$ is called an eigenvector of A if $A\mathbf{x} = \lambda\mathbf{x}$.
- In this case, λ is called an eigenvalue of A, and $\mathbf{x}$ is an eigenvector corresponding to λ.
- The polynomial $det(\lambda I - A)$ is called the characteristic polynomial of A and the equation $det(\lambda I - A) = 0$ the characteristic equation of A.
- The eigenvalues of A are the solutions of $det(\lambda I - A) = 0$.
- For an eigenvalue λ_0 of A, the null space of matrix $\lambda_0 I - A$ is the eigenspace of A corresponding to λ_0. The non-zero vectors in this subspace are the eigenvectors of A corresponding to λ_0.

Recap of Last Two Weeks

- For an $n \times n$ matrix, a non-zero vector $\mathbf{x} \in \mathbb{R}^n$ is called an eigenvector of A if $A\mathbf{x} = \lambda\mathbf{x}$.
- In this case, λ is called an eigenvalue of A, and $\mathbf{x}$ is an eigenvector corresponding to λ.
- The polynomial $det(\lambda I - A)$ is called the characteristic polynomial of A and the equation $det(\lambda I - A) = 0$ the characteristic equation of A.
- The eigenvalues of A are the solutions of $det(\lambda I - A) = 0$.
- For an eigenvalue λ_0 of A, the null space of matrix $\lambda_0 I - A$ is the eigenspace of A corresponding to λ_0. The non-zero vectors in this subspace are the eigenvectors of A corresponding to λ_0.
- Eigenvalues and eigenvectors play a major role in PCA (principal component analysis).

Plan for Today

- Similarity of matrices
- Diagonalisation of a matrix and how to find it
- Eigendecomposition of a matrix

Outline

What Is It & Why Does It Matter?

What Is It?

1. Decompose matrix A into PDP^{-1} form.
2. P - eigenvectors, D - diagonal matrix eigenvalues, e.g., $D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$

What Is It & Why Does It Matter?

What Is It?

1. Decompose matrix A into PDP^{-1} form.
2. P - eigenvectors, D - diagonal matrix eigenvalues, e.g., $D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$

Why Care?

1. Faster Self-Matrix Multiplication
   A^k

What Is It & Why Does It Matter?

What Is It?

1. Decompose matrix A into PDP^{-1} form.
2. P - eigenvectors, D - diagonal matrix eigenvalues, e.g., $D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$

Why Care?

1. Faster Self-Matrix Multiplication
   $A^k = PDP^{-1}PDP^{-1}PDP^{-1}PDP^{-1}...$ k times

What Is It & Why Does It Matter?

What Is It?

1. Decompose matrix A into PDP^{-1} form.
2. P - eigenvectors, D - diagonal matrix eigenvalues, e.g., $D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$

Why Care?

1. Faster Self-Matrix Multiplication
   $A^k = PDP^{-1}PDP^{-1}PDP^{-1}PDP^{-1}...$ k times
   $A^k = PDDDD...$ k times P^{-1}

What Is It & Why Does It Matter?

What Is It?

1. Decompose matrix A into PDP^{-1} form.
2. P - eigenvectors, D - diagonal matrix eigenvalues, e.g., $D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$

Why Care?

1. Faster Self-Matrix Multiplication
   $A^k = PDP^{-1}PDP^{-1}PDP^{-1}PDP^{-1}\ldots$ k times
   $A^k = PDDDD\ldots$ k times P^{-1}

   $A^k = PD^kP^{-1}$

What Is It & Why Does It Matter?

What Is It?

1. Decompose matrix A into PDP^{-1} form.
2. P - eigenvectors, D - diagonal matrix eigenvalues, e.g., $D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$

Why Care?

1. Faster Self-Matrix Multiplication
   $A^k = PDP^{-1}PDP^{-1}PDP^{-1}PDP^{-1}\ldots$ k times
   $A^k = PDDDD\ldots$ k times P^{-1}
   $A^k = PD^kP^{-1}$ (e.g., $D^k = \begin{pmatrix} \lambda_1^k & 0 & 0 \\ 0 & \lambda_2^k & 0 \\ 0 & 0 & \lambda_3^k \end{pmatrix}$)

What Is It & Why Does It Matter?

What Is It?

1. Decompose matrix A into PDP^{-1} form.
2. P - eigenvectors, D - diagonal matrix eigenvalues, e.g., $D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$

Why Care?

1. Faster Self-Matrix Multiplication
   $A^k = PDP^{-1}PDP^{-1}PDP^{-1}PDP^{-1}\ldots$ k times
   $A^k = PDDDD\ldots$ k times P^{-1}
   $A^k = PD^kP^{-1}$ (e.g., $D^k = \begin{pmatrix} \lambda_1^k & 0 & 0 \\ 0 & \lambda_2^k & 0 \\ 0 & 0 & \lambda_3^k \end{pmatrix}$)
2. Useful in PCA (covariance matrix)

Plan to Learn About It

1. Learn useful concept 1: Similarity of matrices
2. Learn useful concept 2: Diagonalisation of a matrix (what & how)
3. Finally get to eigendecomposition (what is it & conditions for it to exist)

Outline

Similarity of Matrices

Definition

Square matrices A and B are called similar if $A = P^{-1}BP$ for some invertible P.

Note that if $A = P^{-1}BP$ then $B = Q^{-1}AQ$ where $Q = P^{-1}$.

Similarity of Matrices

Definition

Square matrices A and B are called similar if $A = P^{-1}BP$ for some invertible P.

Note that if $A = P^{-1}BP$ then $B = Q^{-1}AQ$ where $Q = P^{-1}$.

Similar matrices have many features in common, including determinant, trace, rank, nullity, characteristic polynomial, eigenvalues, dimensions of corresponding eigenspaces, etc.

Similarity of Matrices

Definition

Square matrices A and B are called similar if $A = P^{-1}BP$ for some invertible P.

Note that if $A = P^{-1}BP$ then $B = Q^{-1}AQ$ where $Q = P^{-1}$.

Similar matrices have many features in common, including determinant, trace, rank, nullity, characteristic polynomial, eigenvalues, dimensions of corresponding eigenspaces, etc.

Lemma

If A and B are similar then $det(A) = det(B)$.

Note: Whilst we have proven that A and B being similar implies that they have the same determinant, the converse is not true.

Properties of Similar Matrices

If A and B are similar then:

1. A is invertible iff B is invertible.
2. A and B have the same rank.
3. A and B have the same characteristic polynomial.
4. A and B have the same eigenvalues.

Recap: Representing a Matrix in Terms of a Basis

1. Suppose we have a vector $\mathbf{u} = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$ in a coordinate system with the standard basis.

Recap: Representing a Matrix in Terms of a Basis

1. Suppose we have a vector $\mathbf{u} = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$ in a coordinate system with the standard basis.
2. We want to represent it in terms of the basis $S = \{\mathbf{v}_1, \mathbf{v}_2\}$, where $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\mathbf{v}_2 = \begin{pmatrix} 0 \\ 4 \end{pmatrix}$.

Recap: Representing a Matrix in Terms of a Basis

1. Suppose we have a vector $\mathbf{u} = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$ in a coordinate system with the standard basis.
2. We want to represent it in terms of the basis $S = \{\mathbf{v}_1, \mathbf{v}_2\}$, where $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\mathbf{v}_2 = \begin{pmatrix} 0 \\ 4 \end{pmatrix}$.
3. In other words, what are a, b such that $\mathbf{u} = a\mathbf{v}_1 + b\mathbf{v}_2$.

Recap: Representing a Matrix in Terms of a Basis

1. Suppose we have a vector $\mathbf{u} = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$ in a coordinate system with the standard basis.
2. We want to represent it in terms of the basis $S = \{\mathbf{v}_1, \mathbf{v}_2\}$, where $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\mathbf{v}_2 = \begin{pmatrix} 0 \\ 4 \end{pmatrix}$.
3. In other words, what are a, b such that $\mathbf{u} = a\mathbf{v}_1 + b\mathbf{v}_2$.
4. Solving, $a = 5$, $b = -2$.

Recap: Representing a Matrix in Terms of a Basis

1. Suppose we have a vector $\mathbf{u} = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$ in a coordinate system with the standard basis.
2. We want to represent it in terms of the basis $S = \{\mathbf{v}_1, \mathbf{v}_2\}$, where $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\mathbf{v}_2 = \begin{pmatrix} 0 \\ 4 \end{pmatrix}$.
3. In other words, what are a, b such that $\mathbf{u} = a\mathbf{v}_1 + b\mathbf{v}_2$.
4. Solving, $a = 5,\ b = -2$.
5. Here, a, b are the coordinates of $\mathbf{u}$ in basis S and the vector $\begin{pmatrix} 5 \\ -2 \end{pmatrix}$ is the coordinate vector of vector $\mathbf{u}$ in basis S.

Similarity and Linear Maps

Theorem

Matrices A and B are similar iff they represent the same linear operator $f : \mathbb{R}^n \to \mathbb{R}^n$, possibly in different bases. *(Proof omitted.)*

Similarity and Linear Maps

Theorem

Matrices A and B are similar iff they represent the same linear operator $f : \mathbb{R}^n \to \mathbb{R}^n$, possibly in different bases. *(Proof omitted.)*

- Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear operator and $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ a basis in $\mathbb{R}^n$.
- Let $[T]_S$ be the $n \times n$ matrix $[T]_S = [(T(\mathbf{v}_1))_S|(T(\mathbf{v}_2))_S| \ldots |(T(\mathbf{v}_n))_S]$ whose columns are the coordinate vectors of vectors $T(\mathbf{v}_i)$ in basis S.

Similarity and Linear Maps

Theorem

Matrices A and B are similar iff they represent the same linear operator $f : \mathbb{R}^n \to \mathbb{R}^n$, possibly in different bases. *(Proof omitted.)*

- Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear operator and $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ a basis in $\mathbb{R}^n$.
- Let $[T]_S$ be the $n \times n$ matrix $[T]_S = [(T(\mathbf{v}_1))_S | (T(\mathbf{v}_2))_S | \ldots | (T(\mathbf{v}_n))_S]$ whose columns are the coordinate vectors of vectors $T(\mathbf{v}_i)$ in basis S.
  This matrix is called the matrix of T in S.

Similarity and Linear Maps

Theorem

Matrices A and B are similar iff they represent the same linear operator $f : \mathbb{R}^n \to \mathbb{R}^n$, possibly in different bases. *(Proof omitted.)*

- Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear operator and $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ a basis in $\mathbb{R}^n$.
- Let $[T]_S$ be the $n \times n$ matrix $[T]_S = [(T(\mathbf{v}_1))_S|(T(\mathbf{v}_2))_S| \ldots |(T(\mathbf{v}_n))_S]$ whose columns are the coordinate vectors of vectors $T(\mathbf{v}_i)$ in basis S. This matrix is called the matrix of T in S.
- We also say that the matrix $[T]_S$ represents T in basis S.

Similarity and Linear Maps

Theorem

Matrices A and B are similar iff they represent the same linear operator $f : \mathbb{R}^n \to \mathbb{R}^n$, possibly in different bases. *(Proof omitted.)*

- Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear operator and $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ a basis in $\mathbb{R}^n$.
- Let $[T]_S$ be the $n \times n$ matrix $[T]_S = [(T(\mathbf{v}_1))_S | (T(\mathbf{v}_2))_S | \ldots | (T(\mathbf{v}_n))_S]$ whose columns are the coordinate vectors of vectors $T(\mathbf{v}_i)$ in basis S. This matrix is called the matrix of T in S.
- We also say that the matrix $[T]_S$ represents T in basis S.
- If S is the standard basis of $\mathbb{R}^n$ then $[T]_S$ is the standard matrix of T.

Similarity and Linear Maps

Theorem

Matrices A and B are similar iff they represent the same linear operator $f : \mathbb{R}^n \to \mathbb{R}^n$, possibly in different bases. *(Proof omitted.)*

- Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear operator and $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ a basis in $\mathbb{R}^n$.
- Let $[T]_S$ be the $n \times n$ matrix $[T]_S = [(T(\mathbf{v}_1))_S | (T(\mathbf{v}_2))_S | \ldots | (T(\mathbf{v}_n))_S]$ whose columns are the coordinate vectors of vectors $T(\mathbf{v}_i)$ in basis S.
  This matrix is called the matrix of T in S.
- We also say that the matrix $[T]_S$ represents T in basis S.
- If S is the standard basis of $\mathbb{R}^n$ then $[T]_S$ is the standard matrix of T.

Theorem: Matrices A and B are similar iff there is a linear operator $T : \mathbb{R}^n \to \mathbb{R}^n$ and two bases S, S' of $\mathbb{R}^n$ such that $A = [T]_S$ and $B = [T]_{S'}$.

Outline

Diagonalisable Matrices

Definition

A matrix A is called diagonalisable if it is similar to a diagonal matrix –
in other words, if there exists an invertible matrix P such that $P^{-1}AP$ is diagonal.
Then P is said to diagonalise A.

Diagonalisable Matrices

Definition

A matrix A is called diagonalisable if it is similar to a diagonal matrix –
in other words, if there exists an invertible matrix P such that $P^{-1}AP$ is diagonal.
Then P is said to diagonalise A.

Note that A is diagonalisable if it decomposes as $A = PDP^{-1}$ where P is invertible and $D = diag(\lambda_1, \dots, \lambda_n)$ is diagonal.

$$D = diag(\lambda_1, \dots, \lambda_n) \text{ is shorthand for } D = \begin{pmatrix} \lambda_1 & 0 & . & . & . & 0 & 0 \\ 0 & \lambda_2 & . & . & . & 0 & 0 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & 0 & . & . & . & \lambda_{n-1} & 0 \\ 0 & 0 & . & . & . & 0 & \lambda_n \end{pmatrix}$$

A Characterisation of Diagonalisable Matrices

Theorem (Thm 14.1)

An $n \times n$ matrix is diagonalisable iff it has n linearly independent eigenvectors.

A Characterisation of Diagonalisable Matrices

Theorem (Thm 14.1)

An $n \times n$ matrix is diagonalisable iff it has n linearly independent eigenvectors.

Proof.

($\Rightarrow$). Assume that there is an invertible matrix P and a diagonal matrix $D = diag(\lambda_1, \dots \lambda_n)$ such that $D = P^{-1}AP$, or $AP = PD$. Denote the column vectors of P by $\mathbf{p}_1, \dots, \mathbf{p}_n$, so $P = [\mathbf{p}_1 | \dots | \mathbf{p}_n]$. Then

$$AP = A[\mathbf{p}_1 | \dots | \mathbf{p}_n] = [A\mathbf{p}_1 | \dots | A\mathbf{p}_n].$$

On the other hand,

$$PD = [\lambda_1 \mathbf{p}_1 | \dots | \lambda_n \mathbf{p}_n].$$

Since $AP = PD$, we conclude that $A\mathbf{p}_i = \lambda_i \mathbf{p}_i$ for all $1 \leq i \leq n$. Since P is invertible, its rank is n and so the vectors $\mathbf{p}_1, \dots, \mathbf{p}_n$ are linearly independent. Then none of $\mathbf{p}_1, \dots, \mathbf{p}_n$ is $\mathbf{0}$, so each of them is an eigenvector. □

Proof Cont'd

Theorem (Thm 14.1)

An $n \times n$ matrix is diagonalisable iff it has n linearly independent eigenvectors.

Proof.

($\Leftarrow$). Assume that A has n linearly independent eigenvectors $\mathbf{p}_1, \ldots, \mathbf{p}_n$ and let $\lambda_1, \ldots, \lambda_n$ be the corresponding eigenvalues (not necessarily distinct). Define

$$P = [\mathbf{p}_1 | \ldots | \mathbf{p}_n] \quad \text{and} \quad D = diag(\lambda_1, \ldots, \lambda_n).$$

Then

$$AP \;=\; A[\mathbf{p}_1 | \ldots | \mathbf{p}_n] \;=\; [A\mathbf{p}_1 | \ldots | A\mathbf{p}_n] \;=\; [\lambda_1 \mathbf{p}_1 | \ldots | \lambda_n \mathbf{p}_n] \;=\; PD.$$

The columns of P are linearly independent, so its rank is n and it is invertible.
Finally $AP = PD$ is equivalent to $D = P^{-1}AP$. $\square$

An Algorithm for Diagonalisation

Given a matrix A, this algorithm diagonalises it (or reports that this is impossible):

An Algorithm for Diagonalisation

Given a matrix A, this algorithm diagonalises it (or reports that this is impossible):

1. Find the eigenvalues of A (e.g. by solving its characterstic equation).
2. Find a basis in each eigenspace of A and merge these bases into one set S.

An Algorithm for Diagonalisation

Given a matrix A, this algorithm diagonalises it (or reports that this is impossible):

1. Find the eigenvalues of A (e.g. by solving its characterstic equation).
2. Find a basis in each eigenspace of A and merge these bases into one set S.
3. If S has fewer than n vectors then A is not diagonalisable.
4. Else, form the matrix $P = [\mathbf{p}_1|\ldots|\mathbf{p}_n]$ where $S = \{\mathbf{p}_1, \ldots, \mathbf{p}_n\}$.

An Algorithm for Diagonalisation

Given a matrix A, this algorithm diagonalises it (or reports that this is impossible):

1. Find the eigenvalues of A (e.g. by solving its characterstic equation).
2. Find a basis in each eigenspace of A and merge these bases into one set S.
3. If S has fewer than n vectors then A is not diagonalisable.
4. Else, form the matrix $P = [\mathbf{p}_1 | \ldots | \mathbf{p}_n]$ where $S = \{\mathbf{p}_1, \ldots, \mathbf{p}_n\}$.

   $\#$ The set S contains n vectors and is linearly independent (will prove this),

   $\#$ so S is a basis for $\mathbb{R}^n$. Hence, P is invertible.

An Algorithm for Diagonalisation

Given a matrix A, this algorithm diagonalises it (or reports that this is impossible):

1. Find the eigenvalues of A (e.g. by solving its characterstic equation).
2. Find a basis in each eigenspace of A and merge these bases into one set S.
3. If S has fewer than n vectors then A is not diagonalisable.
4. Else, form the matrix $P = [\mathbf{p}_1 | \dots | \mathbf{p}_n]$ where $S = \{\mathbf{p}_1, \dots, \mathbf{p}_n\}$.

   $\#$ The set S contains n vectors and is linearly independent (will prove this),

   $\#$ so S is a basis for $\mathbb{R}^n$. Hence, P is invertible.
5. The matrix $D = P^{-1}AP$ is diagonal, $D = diag(\lambda_1, \dots, \lambda_n)$ where, for each i, λ_i is the eigenvalue corresponding to $\mathbf{p}_i$.

An Algorithm for Diagonalisation

Given a matrix A, this algorithm diagonalises it (or reports that this is impossible):

1. Find the eigenvalues of A (e.g. by solving its characterstic equation).
2. Find a basis in each eigenspace of A and merge these bases into one set S.
3. If S has fewer than n vectors then A is not diagonalisable.
4. Else, form the matrix $P = [\mathbf{p}_1| \dots |\mathbf{p}_n]$ where $S = \{\mathbf{p}_1, \dots, \mathbf{p}_n\}$.

   # The set S contains n vectors and is linearly independent (will prove this),

   # so S is a basis for $\mathbb{R}^n$. Hence, P is invertible.
5. The matrix $D = P^{-1}AP$ is diagonal, $D = diag(\lambda_1, \dots, \lambda_n)$ where, for each i, λ_i is the eigenvalue corresponding to $\mathbf{p}_i$.

Remark:

- If an $n \times n$ matrix has n distinct eigenvalues then it is diagonalisable. Why? We'll see later on.

Example 14.1

For $k \neq 0$, is the following matrix (corresponding to shear) diagonalisable?

$$A = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$$

Example 14.2

Diagonalise the following matrix, if possible,

$$A = \begin{pmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{pmatrix}.$$

Outline

Eigendecomposition

- Let A be a diagonalisable matrix.

Eigendecomposition

- Let A be a diagonalisable matrix.
- We have $AP = PD$,

Eigendecomposition

- Let A be a diagonalisable matrix.
- We have $AP = PD$,

  where P is invertible and its columns $\mathbf{p}_1, \ldots, \mathbf{p}_n$ are eigenvectors of A, and

  $D = diag(\lambda_1, \ldots, \lambda_n)$ and each λ_i is the eigenvalue of A corresponding to $\mathbf{p}_i$.

Eigendecomposition

- Let A be a diagonalisable matrix.
- We have $AP = PD$,

  where P is invertible and its columns $\mathbf{p}_1, \dots, \mathbf{p}_n$ are eigenvectors of A, and

  $D = diag(\lambda_1, \dots, \lambda_n)$ and each λ_i is the eigenvalue of A corresponding to $\mathbf{p}_i$.
- Then we have a decomposition $A = PDP^{-1}$.

  Since the factors in this decomposition are made of eigenvectors and eigenvalues, it is called an eigendecomposition of A.

Eigendecomposition

- Let A be a diagonalisable matrix.
- We have $AP = PD$,

  where P is invertible and its columns $\mathbf{p}_1, \ldots, \mathbf{p}_n$ are eigenvectors of A, and $D = diag(\lambda_1, \ldots, \lambda_n)$ and each λ_i is the eigenvalue of A corresponding to $\mathbf{p}_i$.
- Then we have a decomposition $A = PDP^{-1}$.

  Since the factors in this decomposition are made of eigenvectors and eigenvalues, it is called an eigendecomposition of A.

  (Can use the diagonalisation algorithm to find it.)

Eigendecomposition

- Let A be a diagonalisable matrix.
- We have $AP = PD$,

  where P is invertible and its columns $\mathbf{p}_1, \dots, \mathbf{p}_n$ are eigenvectors of A, and

  $D = diag(\lambda_1, \dots, \lambda_n)$ and each λ_i is the eigenvalue of A corresponding to $\mathbf{p}_i$.
- Then we have a decomposition $A = PDP^{-1}$.

  Since the factors in this decomposition are made of eigenvectors and eigenvalues, it is called an eigendecomposition of A.

  (Can use the diagonalisation algorithm to find it.)
- Eigendecomposition of matrix A (of size $n \times n$) exists iff A can be diagonalised.

Eigendecomposition

- Let A be a diagonalisable matrix.
- We have $AP = PD$,

  where P is invertible and its columns $\mathbf{p}_1, \ldots, \mathbf{p}_n$ are eigenvectors of A, and

  $D = diag(\lambda_1, \ldots, \lambda_n)$ and each λ_i is the eigenvalue of A corresponding to $\mathbf{p}_i$.
- Then we have a decomposition $A = PDP^{-1}$.

  Since the factors in this decomposition are made of eigenvectors and eigenvalues, it is called an eigendecomposition of A.

  (Can use the diagonalisation algorithm to find it.)
- Eigendecomposition of matrix A (of size $n \times n$) exists iff A can be diagonalised.
- Eigendecomposition of matrix A iff A has n linearly independent eigenvectors.

The Return of Algebraic Multiplicity of Eigenvalues

Theorem (Thm 14.2)

If vectors $\mathbf{v}_1 \ldots, \mathbf{v}_k$ *are eigenvectors of a matrix* A *corresponding to (pairwise) distinct eigenvalues* $\lambda_1, \ldots, \lambda_k$ *then the set* $\{\mathbf{v}_1 \ldots, \mathbf{v}_k\}$ *is linearly independent.*

The Return of Algebraic Multiplicity of Eigenvalues

Theorem (Thm 14.2)

If vectors $\mathbf{v}_1 \ldots, \mathbf{v}_k$ *are eigenvectors of a matrix* A *corresponding to (pairwise) distinct eigenvalues* $\lambda_1, \ldots, \lambda_k$ *then the set* $\{\mathbf{v}_1 \ldots, \mathbf{v}_k\}$ *is linearly independent.*

Proof.

Let $r \leq k$ be the largest number such that $\{\mathbf{v}_1 \ldots, \mathbf{v}_r\}$ is linearly independent.
Assume for contradiction that $r < k$, so $\{\mathbf{v}_1 \ldots, \mathbf{v}_r, \mathbf{v}_{r+1}\}$ is linearly dependent:

$$c_1\mathbf{v}_1 + \ldots + c_r\mathbf{v}_r + c_{r+1}\mathbf{v}_{r+1} = \mathbf{0}$$

where not all $c_1, \ldots, c_r, c_{r+1}$ are 0.
Since $\{\mathbf{v}_1 \ldots, \mathbf{v}_r\}$ is linearly independent, we conclude that $c_{r+1} \neq 0$.
Since $\mathbf{v}_{r+1}$ is an eigenvector, we conclude that $c_i \neq 0$ for some $i \leq r$.
Continued on next slide ... □

Proof Continued

Proof.

We assumed that $\{\mathbf{v}_1 \ldots, \mathbf{v}_r\}$ is linearly independent, but $\{\mathbf{v}_1 \ldots, \mathbf{v}_r, \mathbf{v}_{r+1}\}$ is not:

$$c_1\mathbf{v}_1 + \ldots + c_r\mathbf{v}_r + c_{r+1}\mathbf{v}_{r+1} = \mathbf{0} \tag{1}$$

We derived that $c_{r+1} \neq 0$ and $c_i \neq 0$ for some $i \leq r$.
Left multiply both sides of equation (1) by A and use $A\mathbf{v}_i = \lambda_i\mathbf{v}_i$:

$$c_1\lambda_1\mathbf{v}_1 + \ldots + c_r\lambda_r\mathbf{v}_r + c_{r+1}\lambda_{r+1}\mathbf{v}_{r+1} = \mathbf{0}. \tag{2}$$

Now multiply both sides of (1) by λ_{r+1} and subtract that from (2) to obtain

$$c_1(\lambda_1 - \lambda_{r+1})\mathbf{v}_1 + \ldots + c_r(\lambda_r - \lambda_{r+1})\mathbf{v}_r = \mathbf{0}. \tag{3}$$

So $c_i(\lambda_i - \lambda_{r+1}) = 0$, and hence $c_i = 0$, for all $i \leq r$, a contradiction. $\square$

The Return of Algebraic Multiplicity of Eigenvalues

1. From Theorem earlier, we know that $n \times n$ matrix is diagonisable iff it has n linearly independent eigenvectors.

The Return of Algebraic Multiplicity of Eigenvalues

1. From Theorem earlier, we know that $n \times n$ matrix is diagonisable iff it has n linearly independent eigenvectors.
2. A matrix's eigendecomposition exists iff it is diagonisable.

The Return of Algebraic Multiplicity of Eigenvalues

1. From Theorem earlier, we know that $n \times n$ matrix is diagonisable iff it has n linearly independent eigenvectors.
2. A matrix's eigendecomposition exists iff it is diagonisable.
3. **Implication:** if we have a matrix of size $n \times n$ that has n distinct eigenvalues, then we know that it has an eigendecomposition.
4. **Recall:** n distinct eigenvalues means that algebraic multiplicity of all eigenvalues is 1.

The Return of Algebraic Multiplicity of Eigenvalues

1. From Theorem earlier, we know that $n \times n$ matrix is diagonisable iff it has n linearly independent eigenvectors.
2. A matrix's eigendecomposition exists iff it is diagonisable.
3. **Implication:** if we have a matrix of size $n \times n$ that has n distinct eigenvalues, then we know that it has an eigendecomposition.
4. **Recall:** n distinct eigenvalues means that algebraic multiplicity of all eigenvalues is 1.
5. **Important Note:** This does not mean that a matrix that doesn't have n distinct eigenvalues doesn't have an eigendecomposition.

Example 14.3

Give the Eigendecomposition of A if one exists.

$$A = \begin{pmatrix} -2 & 0 & 3 \\ -8 & 2 & 8 \\ 0 & 0 & 1 \end{pmatrix}.$$

Outline

Example exam question

(d) Consider the matrix $A = \begin{pmatrix} -3 & 1 & 2 \\ 1 & -3 & 2 \\ 2 & 2 & 0 \end{pmatrix}$.

i. Diagonalise A **[6 Marks]**

ii. A matrix B has an eigendecomposition of $BP = PD$, where , $B = \begin{pmatrix} -2 & 2 & 1 \\ 2 & -2 & 1 \\ 4 & 4 & -2 \end{pmatrix}, P = \begin{pmatrix} 0.5 & -1 & -0.5 \\ 0.5 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -4 \end{pmatrix}$.
Are B and A similar, give a justification. **[2 Marks]**

What We Learnt Today & Next Week's Plan

Today:

- Similarity of matrices
- Diagonalisable matrices ($=$ similar to a diagonal one)
- A characterisation of diagonalisable matrices
- An algorithm for diagonalisation
- Eigendecomposition of matrices

Next Week:

- Inner product spaces.

Example 14.1

For $k \neq 0$, is the following matrix (corresponding to shear) diagonalisable?

$$A = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$$

Example 14.1

For $k \neq 0$, is the following matrix (corresponding to shear) diagonalisable?

$$A = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$$

Solution. First compute the characteristic polynomial $det(\lambda I - A)$:

$$det(\lambda I - A) = \begin{vmatrix} \lambda - 1 & -k \\ 0 & \lambda - 1 \end{vmatrix} = (\lambda - 1)^2.$$

The characteristic equation is $(\lambda - 1)^2 = 0$, so A has only one eigenvalue $\lambda_1 = 1$.

Example 14.1

For $k \neq 0$, is the following matrix (corresponding to shear) diagonalisable?

$$A = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$$

Solution. First compute the characteristic polynomial $det(\lambda I - A)$:

$$det(\lambda I - A) = \begin{vmatrix} \lambda - 1 & -k \\ 0 & \lambda - 1 \end{vmatrix} = (\lambda - 1)^2.$$

The characteristic equation is $(\lambda - 1)^2 = 0$, so A has only one eigenvalue $\lambda_1 = 1$.

The corresponding eigenspace is the nullspace of $\lambda_1 I - A = I - A = \begin{pmatrix} 0 & -k \\ 0 & 0 \end{pmatrix}$.

Does it have two linearly independent vectors? (Is $nullity(I - A) = 2$?)

Example 14.1

For $k \neq 0$, is the following matrix (corresponding to shear) diagonalisable?

$$A = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$$

Solution. First compute the characteristic polynomial $det(\lambda I - A)$:

$$det(\lambda I - A) = \begin{vmatrix} \lambda - 1 & -k \\ 0 & \lambda - 1 \end{vmatrix} = (\lambda - 1)^2.$$

The characteristic equation is $(\lambda - 1)^2 = 0$, so A has only one eigenvalue $\lambda_1 = 1$.

The corresponding eigenspace is the nullspace of $\lambda_1 I - A = I - A = \begin{pmatrix} 0 & -k \\ 0 & 0 \end{pmatrix}$.

Does it have two linearly independent vectors? (Is $nullity(I - A) = 2$?)

Since $rank(I - A) + nullity(I - A) = 2$ (by the Dimension Theorem for matrices) and $rank(I - A) = 1$, we conclude that $nullity(I - A) = 1$, and therefore A is **not** diagonalisable.

Example 14.2

Diagonalise the following matrix, if possible,

$$A = \begin{pmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{pmatrix}.$$

Example 14.2

Diagonalise the following matrix, if possible,

$$A = \begin{pmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{pmatrix}.$$

Solution. We have

$$det(\lambda I - A) = \begin{vmatrix} \lambda & 0 & 2 \\ -1 & \lambda - 2 & -1 \\ -1 & 0 & \lambda - 3 \end{vmatrix} = \lambda^3 - 5\lambda^2 + 8\lambda - 4.$$

The characteristic equation is $\lambda^3 - 5\lambda^2 + 8\lambda - 4 = 0$, can factor it (as in last lecture): $(\lambda - 1)(\lambda - 2)^2 = 0$. The eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = 2$.

Example 14.2

Diagonalise the following matrix, if possible,

$$A = \begin{pmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{pmatrix}.$$

Solution. We have

$$det(\lambda I - A) = \begin{vmatrix} \lambda & 0 & 2 \\ -1 & \lambda - 2 & -1 \\ -1 & 0 & \lambda - 3 \end{vmatrix} = \lambda^3 - 5\lambda^2 + 8\lambda - 4.$$

The characteristic equation is $\lambda^3 - 5\lambda^2 + 8\lambda - 4 = 0$, can factor it (as in last lecture): $(\lambda - 1)(\lambda - 2)^2 = 0$. The eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = 2$.

The corresponding eigenspaces are the nullspaces of $I - A$ and $2I - A$, resp.

Example 14.2

Diagonalise the following matrix, if possible,

$$A = \begin{pmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{pmatrix}.$$

Solution. We have

$$det(\lambda I - A) = \begin{vmatrix} \lambda & 0 & 2 \\ -1 & \lambda - 2 & -1 \\ -1 & 0 & \lambda - 3 \end{vmatrix} = \lambda^3 - 5\lambda^2 + 8\lambda - 4.$$

The characteristic equation is $\lambda^3 - 5\lambda^2 + 8\lambda - 4 = 0$, can factor it (as in last lecture): $(\lambda - 1)(\lambda - 2)^2 = 0$. The eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = 2$.

The corresponding eigenspaces are the nullspaces of $I - A$ and $2I - A$, resp.
Using the algorithms for finding a basis in a nullspace, get the following:

$$\lambda_1 = 1 : \mathbf{p}_1 = (-2, 1, 1); \quad \lambda_1 = 2 : \mathbf{p}_2 = (-1, 0, 1) \text{ and } \mathbf{p}_3 = (0, 1, 0).$$

Example 14.2 Cont'd

The eigenvalues of A are $\lambda_1 = 1$ and $\lambda_2 = 2$, and the bases of nullspaces are

$$\lambda_1 = 1 : \mathbf{p}_1 = (-2, 1, 1); \quad \lambda_1 = 2 : \mathbf{p}_2 = (-1, 0, 1) \text{ and } \mathbf{p}_3 = (0, 1, 0).$$

Example 14.2 Cont'd

The eigenvalues of A are $\lambda_1 = 1$ and $\lambda_2 = 2$, and the bases of nullspaces are

$$\lambda_1 = 1 : \mathbf{p}_1 = (-2, 1, 1); \quad \lambda_1 = 2 : \mathbf{p}_2 = (-1, 0, 1) \text{ and } \mathbf{p}_3 = (0, 1, 0).$$

Hence, A is diagonalisable, and one possible matrix that diagonalises it is

$$P = \begin{pmatrix} -2 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

One can check that $P^{-1}AP$ is

$$\begin{pmatrix} -2 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{pmatrix} \begin{pmatrix} -2 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Example 14.2 Cont'd

The eigenvalues of A are $\lambda_1 = 1$ and $\lambda_2 = 2$, and the bases of nullspaces are

$$\lambda_1 = 1 : \mathbf{p}_1 = (-2, 1, 1); \quad \lambda_1 = 2 : \mathbf{p}_2 = (-1, 0, 1) \text{ and } \mathbf{p}_3 = (0, 1, 0).$$

Hence, A is diagonalisable, and one possible matrix that diagonalises it is

$$P = \begin{pmatrix} -2 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

One can check that $P^{-1}AP$ is

$$\begin{pmatrix} -2 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{pmatrix} \begin{pmatrix} -2 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Remark: The order of the $\mathbf{p}_i$'s in P can be changed arbitrarily, this will result in changing the order of the λ_i's in D accordingly.

Example 14.3

Give the Eigendecomposition of A if one exists.

$$A = \begin{pmatrix} -2 & 0 & 3 \\ -8 & 2 & 8 \\ 0 & 0 & 1 \end{pmatrix}.$$

Solution:

1) Find eigenvalues: $\lambda_1 = -2, \lambda_2 = 2, \lambda_3 = 1$
2) Find eigenvectors: $\mathbf{v}_1 = (1, 2, 0), \mathbf{v}_2 = (0, 1, 0), \mathbf{v}_3 = (1, 0, 1)$
3) $D = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ $P = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

Example 14.3 cont

4) Invert P, $P^{-}1 = \begin{pmatrix} 1 & 0 & -1 \\ -2 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$

Giving:

$$\begin{pmatrix} -2 & 0 & 3 \\ -8 & 2 & 8 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 \\ -2 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

The End