MATH 323

Linear Algebra

# Lecture 19:
# Eigenvalues and eigenvectors (continued). Diagonalization.

# Eigenvalues and eigenvectors of a matrix

*Definition.* Let $A$ be an $n\times n$ matrix. A number $\lambda \in \mathbb{R}$ is called an **eigenvalue** of the matrix $A$ if $\boxed{A\mathbf{v} = \lambda\mathbf{v}}$ for a nonzero column vector $\mathbf{v} \in \mathbb{R}^n$. The vector $\mathbf{v}$ is called an **eigenvector** of $A$ belonging to (or associated with) the eigenvalue $\lambda$.

If $\lambda$ is an eigenvalue of $A$ then the nullspace $N(A-\lambda I)$, which is nontrivial, is called the **eigenspace** of $A$ corresponding to $\lambda$. The eigenspace consists of all eigenvectors belonging to the eigenvalue $\lambda$ plus the zero vector.

# Characteristic equation

*Definition.* Given a square matrix $A$, the equation $\det(A - \lambda I) = 0$ is called the **characteristic equation** of $A$.

Eigenvalues $\lambda$ of $A$ are roots of the characteristic equation.

If $A$ is an $n \times n$ matrix then $p(\lambda) = \det(A - \lambda I)$ is a polynomial of degree $n$. It is called the **characteristic polynomial** of $A$.

**Theorem** Any $n \times n$ matrix has at most $n$ eigenvalues.

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

Characteristic equation:

$$\begin{vmatrix} 1-\lambda & 1 & -1 \\ 1 & 1-\lambda & 1 \\ 0 & 0 & 2-\lambda \end{vmatrix} = 0.$$

Expand the determinant by the 3rd row:

$$(2-\lambda) \begin{vmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{vmatrix} = 0.$$

$$\big((1-\lambda)^2 - 1\big)(2-\lambda) = 0 \iff -\lambda(2-\lambda)^2 = 0$$

$$\implies \quad \lambda_1 = 0, \;\; \lambda_2 = 2.$$

$$A\mathbf{x} = \mathbf{0} \iff \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Convert the matrix to reduced row echelon form:

$$\begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & -1 \\ 0 & 0 & 2 \\ 0 & 0 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

$$A\mathbf{x} = \mathbf{0} \iff \begin{cases} x + y = 0, \\ z = 0. \end{cases}$$

The general solution is $(-t, t, 0) = t(-1, 1, 0)$, $t \in \mathbb{R}$. Thus $\mathbf{v}_1 = (-1, 1, 0)$ is an eigenvector associated with the eigenvalue 0. The corresponding eigenspace is the line spanned by $\mathbf{v}_1$.

$$(A-2I)\mathbf{x}=\mathbf{0} \iff \begin{pmatrix} -1 & 1 & -1 \\ 1 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\iff \begin{pmatrix} 1 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \iff x-y+z=0.$$

The general solution is $x=t-s,\ \ y=t,\ \ z=s,$ where $t,s\in\mathbb{R}$. Equivalently,

$$\mathbf{x}=(t-s,t,s)=t(1,1,0)+s(-1,0,1).$$

Thus $\mathbf{v}_2=(1,1,0)$ and $\mathbf{v}_3=(-1,0,1)$ are eigenvectors associated with the eigenvalue 2. The corresponding eigenspace is the plane spanned by $\mathbf{v}_2$ and $\mathbf{v}_3$.

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

- The matrix $A$ has two eigenvalues: 0 and 2.
- The eigenvalue 0 is *simple:* the corresponding eigenspace is a line.
- The eigenvalue 2 is of *multiplicity* 2: the corresponding eigenspace is a plane.
- Eigenvectors $\mathbf{v}_1 = (-1, 1, 0)$, $\mathbf{v}_2 = (1, 1, 0)$, and $\mathbf{v}_3 = (-1, 0, 1)$ of the matrix $A$ form a basis for $\mathbb{R}^3$.
- Geometrically, the map $\mathbf{x} \mapsto A\mathbf{x}$ is the projection on the plane $\mathrm{Span}(\mathbf{v}_2, \mathbf{v}_3)$ along the lines parallel to $\mathbf{v}_1$ with the subsequent scaling by a factor of 2.

## Eigenvalues and eigenvectors of an operator

*Definition.* Let $V$ be a vector space and $L : V \to V$ be a linear operator. A number $\lambda$ is called an **eigenvalue** of the operator $L$ if $\boxed{L(\mathbf{v}) = \lambda\mathbf{v}}$ for a nonzero vector $\mathbf{v} \in V$. The vector $\mathbf{v}$ is called an **eigenvector** of $L$ associated with the eigenvalue $\lambda$. (If $V$ is a functional vector space then eigenvectors are usually called **eigenfunctions**.)

If $V = \mathbb{R}^n$ then the linear operator $L$ is given by $L(\mathbf{x}) = A\mathbf{x}$, where $A$ is an $n \times n$ matrix (and $\mathbf{x}$ is regarded a column vector). In this case, eigenvalues and eigenvectors of the operator $L$ are precisely eigenvalues and eigenvectors of the matrix $A$.

# Eigenspaces

Let $L : V \to V$ be a linear operator.

For any $\lambda \in \mathbb{R}$, let $V_\lambda$ denotes the set of all solutions of the equation $L(\mathbf{x}) = \lambda\mathbf{x}$.

Then $V_\lambda$ is a *subspace* of $V$ since $V_\lambda$ is the *kernel* of a linear operator given by $\mathbf{x} \mapsto L(\mathbf{x}) - \lambda\mathbf{x}$.

$V_\lambda$ minus the zero vector is the set of all eigenvectors of $L$ associated with the eigenvalue $\lambda$. In particular, $\lambda \in \mathbb{R}$ is an eigenvalue of $L$ if and only if $V_\lambda \neq \{\mathbf{0}\}$.

If $V_\lambda \neq \{\mathbf{0}\}$ then it is called the **eigenspace** of $L$ corresponding to the eigenvalue $\lambda$.

*Example.* $V = C^\infty(\mathbb{R})$, $D : V \to V$, $Df = f'$.

A function $f \in C^\infty(\mathbb{R})$ is an eigenfunction of the operator $D$ belonging to an eigenvalue $\lambda$ if $f'(x) = \lambda f(x)$ for all $x \in \mathbb{R}$.

It follows that $f(x) = ce^{\lambda x}$, where $c$ is a nonzero constant.

Thus each $\lambda \in \mathbb{R}$ is an eigenvalue of $D$.
The corresponding eigenspace is spanned by $e^{\lambda x}$.

*Example.* $V = C^\infty(\mathbb{R})$, $L : V \to V$, $Lf = f''$.

$Lf = \lambda f \iff f''(x) - \lambda f(x) = 0$ for all $x \in \mathbb{R}$.

It follows that each $\lambda \in \mathbb{R}$ is an eigenvalue of $L$ and the corresponding eigenspace $V_\lambda$ is two-dimensional. Note that $L = D^2$, hence $Df = \mu f \implies Lf = \mu^2 f$.

If $\lambda > 0$ then $V_\lambda = \mathrm{Span}(e^{\mu x}, e^{-\mu x})$, where $\mu = \sqrt{\lambda}$.

If $\lambda < 0$ then $V_\lambda = \mathrm{Span}\big(\sin(\mu x), \cos(\mu x)\big)$, where $\mu = \sqrt{-\lambda}$.

If $\lambda = 0$ then $V_\lambda = \mathrm{Span}(1, x)$.

Suppose $L : V \to V$ is a linear operator on a **finite-dimensional** vector space $V$.

Let $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ be a basis for $V$ and $g : V \to \mathbb{R}^n$ be the corresponding coordinate mapping. Let $A$ be the matrix of $L$ with respect to this basis. Then

$$L(\mathbf{v}) = \lambda \mathbf{v} \iff A\, g(\mathbf{v}) = \lambda\, g(\mathbf{v}).$$

Hence the eigenvalues of $L$ coincide with those of the matrix $A$. Moreover, the associated eigenvectors of $A$ are coordinates of the eigenvectors of $L$.

*Definition.* The characteristic polynomial $p(\lambda) = \det(A - \lambda I)$ of the matrix $A$ is called the **characteristic polynomial** of the operator $L$.

Then eigenvalues of $L$ are roots of its characteristic polynomial.

**Theorem.** The characteristic polynomial of the operator $L$ is well defined. That is, it does not depend on the choice of a basis.

*Proof:* Let $B$ be the matrix of $L$ with respect to a different basis $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. Then $A = UBU^{-1}$, where $U$ is the transition matrix from the basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ to $\mathbf{u}_1, \ldots, \mathbf{u}_n$. We have to show that $\det(A - \lambda I) = \det(B - \lambda I)$ for all $\lambda \in \mathbb{R}$. We obtain

$$\det(A - \lambda I) = \det(UBU^{-1} - \lambda I)$$
$$= \det\big(UBU^{-1} - U(\lambda I)U^{-1}\big) = \det\big(U(B - \lambda I)U^{-1}\big)$$
$$= \det(U)\,\det(B - \lambda I)\,\det(U^{-1}) = \det(B - \lambda I).$$

## Basis of eigenvectors

Let $V$ be a finite-dimensional vector space and $L : V \to V$ be a linear operator. Let $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ be a basis for $V$ and $A$ be the matrix of the operator $L$ with respect to this basis.

**Theorem** The matrix $A$ is diagonal if and only if vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ are eigenvectors of $L$.

If this is the case, then the diagonal entries of the matrix $A$ are the corresponding eigenvalues of $L$.

$$L(\mathbf{v}_i) = \lambda_i \mathbf{v}_i \iff A = \begin{pmatrix} \lambda_1 & & & O \\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_n \end{pmatrix}$$

**Theorem** If $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are eigenvectors of a linear operator $L$ associated with distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$, then $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are linearly independent.

**Corollary 1** If $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct real numbers, then the functions $e^{\lambda_1 x}, e^{\lambda_2 x}, \ldots, e^{\lambda_k x}$ are linearly independent.

*Proof:* Consider a linear operator $D : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ given by $Df = f'$. Then $e^{\lambda_1 x}, \ldots, e^{\lambda_k x}$ are eigenfunctions of $D$ associated with distinct eigenvalues $\lambda_1, \ldots, \lambda_k$. By the theorem, the eigenfunctions are linearly independent.

# How to find a basis of eigenvectors

**Corollary 2** Suppose $\lambda_1, \lambda_2, \ldots, \lambda_k$ are all eigenvalues of a linear operator $L : V \to V$. For any $1 \leq i \leq k$, let $S_i$ be a basis for the eigenspace associated to the eigenvalue $\lambda_i$. Then these bases are disjoint and the union $S = S_1 \cup S_2 \cup \cdots \cup S_k$ is a linearly independent set.

Moreover, if the vector space $V$ admits a basis consisting of eigenvectors of $L$, then $S$ is such a basis.

**Corollary 3** Let $A$ be an $n \times n$ matrix such that the characteristic equation $\det(A - \lambda I) = 0$ has $n$ distinct roots. Then **(i)** there is a basis for $\mathbb{R}^n$ consisting of eigenvectors of $A$; **(ii)** all eigenspaces of $A$ are one-dimensional.

# Diagonalization

**Theorem 1** Let $L$ be a linear operator on a finite-dimensional vector space $V$. Then the following conditions are equivalent:

- the matrix of $L$ with respect to some basis is diagonal;
- there exists a basis for $V$ formed by eigenvectors of $L$.

The operator $L$ is **diagonalizable** if it satisfies these conditions.

**Theorem 2** Let $A$ be an $n\times n$ matrix. Then the following conditions are equivalent:

- $A$ is the matrix of a diagonalizable operator;
- $A$ is similar to a diagonal matrix, i.e., it is represented as $A = UBU^{-1}$, where the matrix $B$ is diagonal;
- there exists a basis for $\mathbb{R}^n$ formed by eigenvectors of $A$.

The matrix $A$ is **diagonalizable** if it satisfies these conditions.

*Example.* $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$.

- The matrix $A$ has two eigenvalues: 1 and 3.
- The eigenspace of $A$ associated with the eigenvalue 1 is the line spanned by $\mathbf{v}_1 = (-1, 1)$.
- The eigenspace of $A$ associated with the eigenvalue 3 is the line spanned by $\mathbf{v}_2 = (1, 1)$.
- Eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$ form a basis for $\mathbb{R}^2$.

Thus the matrix $A$ is diagonalizable. Namely, $A = UBU^{-1}$, where

$$B = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}, \qquad U = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Notice that $U$ is the transition matrix from the basis $\mathbf{v}_1, \mathbf{v}_2$ to the standard basis.

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

- The matrix $A$ has two eigenvalues: 0 and 2.
- The eigenspace for 0 is one-dimensional; it has a basis $S_1 = \{\mathbf{v}_1\}$, where $\mathbf{v}_1 = (-1, 1, 0)$.
- The eigenspace for 2 is two-dimensional; it has a basis $S_2 = \{\mathbf{v}_2, \mathbf{v}_3\}$, where $\mathbf{v}_2 = (1, 1, 0)$, $\mathbf{v}_3 = (-1, 0, 1)$.
- The union $S_1 \cup S_2 = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a linearly independent set, hence it is a basis for $\mathbb{R}^3$.

Thus the matrix $A$ is diagonalizable. Namely, $A = UBU^{-1}$, where

$$B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \qquad U = \begin{pmatrix} -1 & 1 & -1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

There are **two obstructions** to existence of a basis consisting of eigenvectors. They are illustrated by the following examples.

*Example 1.* $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

$\det(A - \lambda I) = (\lambda - 1)^2$. Hence $\lambda = 1$ is the only eigenvalue. The associated eigenspace is the line $t(1, 0)$.

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

$\det(A - \lambda I) = \lambda^2 + 1$.

$\implies$ no real eigenvalues or eigenvectors

(However there are *complex* eigenvalues/eigenvectors.)

To *diagonalize* an $n \times n$ matrix $A$ is to find a diagonal matrix $B$ and an invertible matrix $U$ such that $A = UBU^{-1}$.

Suppose there exists a basis $\mathbf{v}_1, \dots, \mathbf{v}_n$ for $\mathbb{R}^n$ consisting of eigenvectors of $A$. That is, $A\mathbf{v}_k = \lambda_k \mathbf{v}_k$, where $\lambda_k \in \mathbb{R}$.
Then $A = UBU^{-1}$, where $B = \mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ and $U$ is a transition matrix whose columns are vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$.

*Example.* $A = \begin{pmatrix} 4 & 3 \\ 0 & 1 \end{pmatrix}$. $\quad \det(A - \lambda I) = (4 - \lambda)(1 - \lambda)$.

Eigenvalues: $\lambda_1 = 4$, $\lambda_2 = 1$.
Associated eigenvectors: $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$.

Thus $A = UBU^{-1}$, where

$$B = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}, \qquad U = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

Suppose we have a problem that involves a square matrix $A$ in the context of matrix multiplication.

Also, suppose that the case when $A$ is a diagonal matrix is simple. Then the diagonalization may help in solving this problem (or may not). Namely, it may reduce the case of a diagonalizable matrix to that of a diagonal one.

An example of such a problem is, given a square matrix $A$, to find its power $A^k$:

$$A = \begin{pmatrix} s_1 & & & O \\ & s_2 & & \\ & & \ddots & \\ O & & & s_n \end{pmatrix} \implies A^k = \begin{pmatrix} s_1^k & & & O \\ & s_2^k & & \\ & & \ddots & \\ O & & & s_n^k \end{pmatrix}$$