Math 304–504

Linear Algebra

Lecture 31:
Bases of eigenvectors.
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 λ.

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

Example. $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} a - \lambda & b \\ c & d - \lambda \end{vmatrix} \\ &= (a - \lambda)(d - \lambda) - bc \\ &= \lambda^2 - (a + d)\lambda + (ad - bc). \end{aligned}$$

Example. $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$

$$\det(A - \lambda I) = \begin{vmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ a_{21} & a_{22} - \lambda & a_{23} \\ a_{31} & a_{32} & a_{33} - \lambda \end{vmatrix}$$

$$= -\lambda^3 + c_1\lambda^2 - c_2\lambda + c_3,$$

where $c_1 = a_{11} + a_{22} + a_{33}$ (the *trace* of A),

$c_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix},$

$c_3 = \det A.$

Eigenvalues and eigenvectors of an operator

Definition. Let V be a vector space and $L : V \to V$ be a linear operator. A number λ 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 λ.

(If V is a functional space then eigenvectors are also 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.
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_λ denotes the set of all solutions of the equation $L(\mathbf{x}) = \lambda\mathbf{x}$.

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

V_λ minus the zero vector is the set of all eigenvectors of L associated with the eigenvalue λ. 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 λ.

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 λ 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_λ is two-dimensional.

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

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

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

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

Proposition 1 Suppose $\mathbf{v}_1$ and $\mathbf{v}_2$ are eigenvectors of L associated with different eigenvalues λ_1 and λ_2. Then $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent.

Proof: Assume that $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly dependent. Then $\mathbf{v}_1 = t\mathbf{v}_2$ for some $t \in \mathbb{R}$. It follows that

$$L(\mathbf{v}_1) = L(t\mathbf{v}_2) = tL(\mathbf{v}_2) = t(\lambda_2\mathbf{v}_2) = \lambda_2(t\mathbf{v}_2) = \lambda_2\mathbf{v}_1.$$

But $L(\mathbf{v}_1) = \lambda_1\mathbf{v}_1 \implies \lambda_1\mathbf{v}_1 = \lambda_2\mathbf{v}_1$
$\implies (\lambda_1 - \lambda_2)\mathbf{v}_1 = \mathbf{0} \implies \mathbf{v}_1 = \mathbf{0}$, a contradiction.

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

Proposition 2 If $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ are eigenvectors of L associated with distinct eigenvalues λ_1, λ_2, and λ_3, then they are linearly independent.

Proof: Suppose that $t_1\mathbf{v}_1 + t_2\mathbf{v}_2 + t_3\mathbf{v}_3 = \mathbf{0}$ for some $t_1, t_2, t_3 \in \mathbb{R}$. Then

$$L(t_1\mathbf{v}_1 + t_2\mathbf{v}_2 + t_3\mathbf{v}_3) = \mathbf{0},$$
$$t_1L(\mathbf{v}_1) + t_2L(\mathbf{v}_2) + t_3L(\mathbf{v}_3) = \mathbf{0},$$
$$t_1\lambda_1\mathbf{v}_1 + t_2\lambda_2\mathbf{v}_2 + t_3\lambda_3\mathbf{v}_3 = \mathbf{0}.$$

It follows that $t_1(\lambda_1 - \lambda_3)\mathbf{v}_1 + t_2(\lambda_2 - \lambda_3)\mathbf{v}_2 = \mathbf{0}$. By the above $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent. Hence $t_1(\lambda_1 - \lambda_3) = t_2(\lambda_2 - \lambda_3) = 0$
$\implies t_1 = t_2 = 0 \implies t_3 = 0.$

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 Let A be an $n \times n$ matrix such that the characteristic equation $\det(A - \lambda I) = 0$ has n distinct real roots. Then $\mathbb{R}^n$ has a basis consisting of eigenvectors of A.

Proof: Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be distinct real roots of the characteristic equation. Any λ_i is an eigenvalue of A, hence there is an associated eigenvector $\mathbf{v}_i$. By the theorem, vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are linearly independent. Therefore they form a basis for $\mathbb{R}^n$.

Theorem 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 $\lambda_1, \ldots, \lambda_k$.

Characteristic polynomial of an operator

Let L be 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. Let A be the matrix of L with respect to this basis.

Definition. The characteristic polynomial 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.

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 obtain

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

Diagonalization

Let L be a linear operator on a finite-dimensional vector space V.

Definition. The operator L is **diagonalizable** if it has a diagonal matrix with respect to some basis for V. Equivalently, if V has a basis consisting of eigenvectors of L.

Theorem Suppose A is a square matrix that admits a basis of eigenvectors. Then A is similar to a diagonal matrix B, i.e., $A = UBU^{-1}$ for an invertible matrix U.

Definition. A square matrix A is **diagonalizable** if it is similar to a diagonal matrix. Otherwise the matrix A is called **defective**.

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

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

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 corresponding to 0 is spanned by $\mathbf{v}_1 = (-1, 1, 0)$.
- The eigenspace corresponding to 2 is spanned by $\mathbf{v}_2 = (1, 1, 0)$ and $\mathbf{v}_3 = (-1, 0, 1)$.
- Eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ form a basis for $\mathbb{R}^3$.

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

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