MATH 423

Linear Algebra II

# Lecture 23: Diagonalization (continued). The Cayley-Hamilton theorem.

# Matrix polynomials

*Definition.* For any $n \times n$ matrix $A$ and any polynomial

$$p(x) = c_0 x^m + c_1 x^{m-1} + \cdots + c_{m-1} x + c_m,$$

let $p(A) = c_0 A^m + c_1 A^{m-1} + \cdots + c_{m-1} A + c_m I_n$.

**Theorem 1** If $A = \mathrm{diag}(s_1, s_2, \ldots, s_n)$ then

$$p(A) = \mathrm{diag}\big(p(s_1), p(s_2), \ldots, p(s_n)\big).$$

**Theorem 2** If $A = UBU^{-1}$, then $p(A) = Up(B)U^{-1}$ for any polynomial $p(x)$.

*Problem.* Let $A = \begin{pmatrix} 4 & 3 \\ 0 & 1 \end{pmatrix}$. Find a matrix $C$ such that $C^2 = A$.

We know from the previous lecture that $A = UBU^{-1}$, where

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

Suppose that $D^2 = B$ for some matrix $D$. Let $C = UDU^{-1}$.
Then $C^2 = UDU^{-1}UDU^{-1} = UD^2U^{-1} = UBU^{-1} = A$.

We can take $D = \begin{pmatrix} \sqrt{4} & 0 \\ 0 & \sqrt{1} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$.

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

**Proposition** An eigenvector of a matrix $A$ is also an eigenvector of any matrix polynomial $p(A)$. The associated eigenvalue for $p(A)$ is $p(\lambda)$, where $\lambda$ is the eigenvalue for $A$.

*Sketch of the proof:* Suppose that $A\mathbf{v} = \lambda\mathbf{v}$, where $\mathbf{v} \neq \mathbf{0}$. Then $A^k\mathbf{v} = \lambda^k\mathbf{v}$ for $k = 0, 1, 2, \dots$
$\implies p(A)\mathbf{v} = p(\lambda)\mathbf{v}$ for any polynomial $p(x)$.

**Cayley-Hamilton Theorem** Consider the characteristic polynomial $p(\lambda) = \det(A - \lambda I)$. Then $p(A) = O$.

*Remark.* Notice that $p(A) \neq \det(A - AI)$ **!!!**

# Characterizations of a direct sum

Suppose $V_1, V_2, \dots, V_k$ are nontrivial subspaces of a vector space $V$ and let $W = V_1 + V_2 + \cdots + V_k$.

**Theorem** The following conditions are equivalent:

**(i)** the subspaces $V_1, V_2, \dots, V_k$ form a direct sum: $W = V_1 \oplus V_2 \oplus \cdots \oplus V_k$;

**(ii)** if $\mathbf{v}_i$ is any nonzero vector from $V_i$ for $1 \le i \le k$, then $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ are linearly independent vectors;

**(iii)** if $S_i$ is any basis for $V_i$, $1 \le i \le k$, then these bases are disjoint and the union $S_1 \cup S_2 \cup \cdots \cup S_k$ is a linearly independent set;

**(iv)** if $S_i$ is any basis for $V_i$, $1 \le i \le k$, then these bases are disjoint and the union $S_1 \cup S_2 \cup \cdots \cup S_k$ is a basis for $W$.

In the case $\dim W < \infty$, there is one more equivalent condition: **(v)** $\dim W = \sum_{i=1}^{k} \dim V_i$.

# How to find a basis of eigenvectors

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

**Corollary 1** The sum of the eigenspaces $\mathcal{E}_{\lambda_1}, \mathcal{E}_{\lambda_2}, \dots, \mathcal{E}_{\lambda_k}$ of the operator $L$ is direct.

**Corollary 2** Let $\lambda_1, \lambda_2, \dots, \lambda_k$ be all eigenvalues of a linear operator $L : V \to V$. For any $1 \le i \le k$, let $S_i$ be a basis for the eigenspace $\mathcal{E}_{\lambda_i}$. 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, $L$ is diagonalizable if and only if $S$ is a basis for $V$.

**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 $\mathbb{F}^n$ consisting of eigenvectors of $A$; **(ii)** all eigenspaces of $A$ are one-dimensional.

*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 eigenspace $\mathcal{E}_0$ is one-dimensional; it has a basis $S_1 = \{\mathbf{v}_1\}$, where $\mathbf{v}_1 = (-1, 1, 0)$.
- The eigenspace $\mathcal{E}_2$ 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

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

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

Eigenvalues of $A$ are roots of its characteristic polynomial

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

Let us expand the determinant by the 1st row.

Expand the determinant by the 1st row:

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

Expand both $3\times3$ determinants by the 1st row:

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

$$= \lambda^2(\lambda^2+1) + (\lambda^2+1) = (\lambda^2+1)^2.$$

Since there are no real eigenvalues, $A$ is not diagonalizable in $\mathbb{R}^4$. How about $\mathbb{C}^4$?

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

The eigenvalues are $i$ and $-i$.

One can show that both eigenspaces of $A$ are one-dimensional. The eigenspace for $i$ is spanned by $(0, 0, i, 1)$ and the eigenspace for $-i$ spanned by $(0, 0, -i, 1)$. It follows that the matrix $A$ is not diagonalizable in $\mathbb{C}^4$.

There is also an indirect way to show that $A$ is not diagonalizable in $\mathbb{C}^4$. Assume the contrary. Then $A = UXU^{-1}$, where $U$ is an invertible matrix with complex entries and

$$X = \begin{pmatrix} i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & 0 & 0 & -i \end{pmatrix}$$

(note that $X$ should have the same characteristic polynomial as $A$). This would imply that $A^2 = UX^2U^{-1}$. But $X^2 = -I$ so that $A^2 = U(-I)U^{-1} = -I$.

One can easily check that, in fact, $A^2 \neq -I$.