MATH 423

Linear Algebra II

Lecture 22: Diagonalization (continued). Matrix polynomials.

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.

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

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 $\mathbb{F}^n$ formed by eigenvectors of A.

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

Diagonalization of a matrix

The **diagonalization** of an $n \times n$ matrix A consists of finding a diagonal matrix B and an invertible matrix U such that $A = UBU^{-1}$. Suppose we have such a representation. Let $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ be consecutive columns of the matrix U. These are linearly independent vectors (since U is invertible), hence they form a basis β for $\mathbb{F}^n$. Then U is the transition matrix from β to the standard basis α. Consider a linear operator L_A on $\mathbb{F}^n$ given by $L_A(\mathbf{x}) = A\mathbf{x}$. We have $[L_A]_\alpha = A$. Therefore

$$[L_A]_\beta = [\mathrm{id}]^\beta_\alpha [L_A]^\alpha_\alpha [\mathrm{id}]^\alpha_\beta = U^{-1}AU = U^{-1}(UBU^{-1})U = B.$$

Thus the matrix of L_A relative to the basis β is diagonal, which implies that β consists of eigenvectors of L_A (i.e., of A).

Conversely, suppose there exists a basis $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ for $\mathbb{F}^n$ formed by eigenvectors of the matrix A: $A\mathbf{v}_i = \lambda_i \mathbf{v}_i$, $1 \le i \le n$. Then $A = UBU^{-1}$, where $U = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)$ and $B = \mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$.

Problem. Diagonalize the matrix $A = \begin{pmatrix} 4 & 3 \\ 0 & 1 \end{pmatrix}$.

We need to find a diagonal matrix B and an invertible matrix U such that $A = UBU^{-1}$.
Suppose that $\mathbf{v}_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$ is a basis for $\mathbb{R}^2$ formed by eigenvectors of A, i.e., $A\mathbf{v}_i = \lambda_i \mathbf{v}_i$ for some $\lambda_i \in \mathbb{R}$. Then we can take

$$B = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \qquad U = \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix}.$$

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

Problem. Diagonalize the matrix $A = \begin{pmatrix} 4 & 3 \\ 0 & 1 \end{pmatrix}$.

Characteristic equation of A: $\begin{vmatrix} 4-\lambda & 3 \\ 0 & 1-\lambda \end{vmatrix} = 0.$

$(4-\lambda)(1-\lambda) = 0 \implies \lambda_1 = 4,\ \lambda_2 = 1.$

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:

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

Problem. Let $A = \begin{pmatrix} 4 & 3 \\ 0 & 1 \end{pmatrix}$. Find A^5.

We know that $A = UBU^{-1}$, where

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

Then $A^5 = UBU^{-1}UBU^{-1}UBU^{-1}UBU^{-1}UBU^{-1}$

$$= UB^5U^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1024 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1024 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1024 & 1023 \\ 0 & 1 \end{pmatrix}.$$

Problem. Let $A = \begin{pmatrix} 4 & 3 \\ 0 & 1 \end{pmatrix}$. Find A^k ($k \geq 1$).

We know that $A = UBU^{-1}$, where

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

Then

$$A^k = UB^kU^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 4^k & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 4^k & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 4^k & 4^k - 1 \\ 0 & 1 \end{pmatrix}.$$

Matrix polynomials

Definition. Given an $n \times n$ matrix A, we let

$$A^2 = AA, \quad A^3 = AAA, \ \ldots, \quad A^k = \underbrace{AA\ldots A}_{k \text{ times}}, \ \ldots$$

Also, let $A^1 = A$ and $A^0 = I_n$.

Associativity of matrix multiplication implies that all powers A^k are well defined and $A^j A^k = A^{j+k}$ for all $j, k \geq 0$. In particular, all powers of A commute.

Definition. For any polynomial

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

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

Example. $p(x) = x^2 - 3x + 1, \;\; C = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$

$$p(C) = C^2 - 3C + I = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}^2 - 3\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix} - \begin{pmatrix} 6 & 3 \\ 3 & 3 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Thus $C^2 - 3C + I = O$.

Remark. $p(x)$ is the characteristic polynomial of the matrix C.

Properties of matrix polynomials

Suppose A is a square matrix, $p(x), p_1(x), p_2(x)$ are polynomials, and r is a scalar. Then

$$
\begin{aligned}
p(x) = p_1(x)+p_2(x) \quad &\Longrightarrow \quad p(A) = p_1(A) + p_2(A) \\
p(x) = rp_1(x) \quad &\Longrightarrow \quad p(A) = rp_1(A) \\
p(x) = p_1(x)p_2(x) \quad &\Longrightarrow \quad p(A) = p_1(A)p_2(A) \\
p(x) = p_1(p_2(x)) \quad &\Longrightarrow \quad p(A) = p_1(p_2(A))
\end{aligned}
$$

In particular, matrix polynomials $p_1(A)$ and $p_2(A)$ always commute.

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

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

Examples.

- $(A - I)(A + I) = A^2 - I$
- $(A + I)^2 = A^2 + 2A + I$
- $(A - I)^2 = A^2 - 2A + I$
- $(A + I)^3 = A^3 + 3A^2 + 3A + I$
- $(A - I)^3 = A^3 - 3A^2 + 3A - I$
- $(A - I)(A^2 + A + I) = A^3 - I$
- $(A + I)(A^2 - A + I) = A^3 + I$

Remark. On the other hand, the matrix equality $(A - B)(A + B) = A^2 - B^2$ holds only if $AB = BA$.

Problem. Let $A = \begin{pmatrix} 4 & 3 \\ 0 & 1 \end{pmatrix}$. Find $I + 2A - A^3$.

We have $A = UBU^{-1}$, where $B = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}$, $U = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$.

Then $A^2 = UBU^{-1}UBU^{-1} = UB^2U^{-1}$,
$A^3 = A^2A = UB^2U^{-1}UBU^{-1} = UB^3U^{-1}$.
Further, $I + 2A - A^3 = UIU^{-1} + 2UBU^{-1} - UB^3U^{-1}$
$= U(I + 2B - B^3)U^{-1}$. That is, $p(A) = Up(B)U^{-1}$, where
$p(x) = 1 + 2x - x^3$. Thus

$$p(A) = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} p(4) & 0 \\ 0 & p(1) \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -55 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -55 & -57 \\ 0 & 2 \end{pmatrix}.$$

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