Lecture 30

Andrei Antonenko

April 25, 2003

1 Eigenvectors. Eigenvalues

Last lecture we saw, that in order to find vectors, "stretched" by the operator with matrix A, we need to solve the characteristic equation

$$\det(A - \lambda I) = 0, \tag{1}$$

which will give us different λ_i's — coefficients, showing, how the vectors are changed after applying the operator. Now we will give the following definition.

Definition 1.1. *Let V be a vector space, and let $\mathcal{A}$ be a linear operator in vector space V. Then the vector x is called **eigenvector** of the operator $\mathcal{A}$ is there exist a number λ, which is called **eigenvalue** such that*

$$\mathcal{A}(x) = \lambda x.$$

So, our goal is to find eigenvectors, since the following proposition holds:

Proposition 1.2. *Let V be an n-dimensional vector space, and $\mathcal{A}$ be a linear operator. Then if there are n linearly independent eigenvectors, then the matrix of $\mathcal{A}$ is diagonal in the basis, consisting of eigenvectors.*

So far we know how to find λ_i's — eigenvalues of the operator. In order to find eigenvectors, we need to solve the system

$$(A - \lambda_i I)x = 0 \tag{2}$$

for every found eigenvalue λ_i.

We will give an example of computing eigenvalues and eigenvectors.

Example 1.3. *Let $A = \begin{pmatrix} 1 & -3 \\ 1 & 5 \end{pmatrix}$. Let's compute its eigenvalues and eigenvectors.*

$$A - \lambda I = \begin{pmatrix} 1-\lambda & -3 \\ 1 & 5-\lambda \end{pmatrix}$$

$$\begin{aligned}\det(A-\lambda I) &= (1-\lambda)(5-\lambda)+3\\ &= 5-6\lambda+\lambda^2+3\\ &= \lambda^2-6\lambda+8.\end{aligned}$$

The roots of this equation are $\lambda_1 = 2$ *and* $\lambda_2 = 4$. *Now we'll find eigenvectors corresponding to these eigenvalues.*

$\boldsymbol{\lambda = 2}$. *Let's subtract* λ*'s from the diagonal. We'll get the following matrix, and the system:*

$$\begin{pmatrix}-1 & -3\\ 1 & 3\end{pmatrix}, \qquad \begin{cases} -x - 3y = 0\\ x + 3y = 0\end{cases}$$

From this system it follows that $x = -3y$, *so each vector of the form* $(-3c, c)$, *i.e.* $(-3, 1)$ *is an eigenvector corresponding to the eigenvalue* $\lambda = 2$.

$\boldsymbol{\lambda = 4}$. *Let's subtract* λ*'s from the diagonal. We'll get the following matrix, and the system:*

$$\begin{pmatrix}-3 & -3\\ 1 & 1\end{pmatrix}, \qquad \begin{cases} -3x - 3y = 0\\ x + y = 0\end{cases}$$

From this system it follows that $x = -y$, *so each vector of the form* $(-c, c)$, *i.e.* $(-1, 1)$ *is an eigenvector corresponding to the eigenvalue* $\lambda = 4$.

So, in the basis, consisting of the vectors $e'_1 = (-3, 1)$ *and* $e'_2 = (-1, 1)$ *the matrix of the corresponding operator has form*

$$\begin{pmatrix}2 & 0\\ 0 & 4\end{pmatrix}$$

Now we can check our formula $D = C^{-1}AC$, *where* D *is a diagonal form of the matrix, and* C *is a change-of-basis matrix. We have*

$$C = \begin{pmatrix}-3 & -1\\ 1 & 1\end{pmatrix}, \quad C^{-1} = \begin{pmatrix}-1/2 & -1/2\\ 1/2 & 3/2\end{pmatrix},$$

so

$$C^{-1}AC = \begin{pmatrix}-1/2 & -1/2\\ 1/2 & 3/2\end{pmatrix}\begin{pmatrix}1 & -3\\ 1 & 5\end{pmatrix}\begin{pmatrix}-3 & -1\\ 1 & 1\end{pmatrix} = \begin{pmatrix}-1 & -1\\ 2 & 6\end{pmatrix}\begin{pmatrix}-3 & -1\\ 1 & 1\end{pmatrix} = \begin{pmatrix}2 & 0\\ 0 & 4\end{pmatrix}.$$

In the same way we can compute eigenvalues and eigenvectors for larger matrices, but it will require solving equations of degree higher than 2. Since we don't have formulae for such equations, we should guess roots of the characteristic equation.

2 Properties of eigenvectors of an operator

Before start studying properties of eigenvectors and eigenvalues we will recall some definitions.

Let V be a vector space, and let $\mathcal{A}$ be a linear operator.

Definition 2.1. *The vector x is called an* ***eigenvector*** *of $\mathcal{A}$ if there exists number λ such that*

$$\mathcal{A}(x) = \lambda x.$$

Such number λ is called an ***eigenvalue****.*

To determine eigenvalues we used characteristic polynomial.

Definition 2.2. *Let A be a square $n \times n$-matrix. The* ***characteristic polynomial*** *of A is*

$$p_A(\lambda) = (-1)^n \det(A - \lambda I) = \det(\lambda I - A).$$

We will prove in the next theorem 2.4, that it is uniquely defined by an operator, i.e. if we take two different matrices of operator, the characteristic polynomial will be the same for both of them.

Remark 2.3. *$(-1)^n$ before determinant is needed to have positive sign before λ^n in the polynomial. But sometimes we will omit $(-1)^n$ before determinant. We need only roots of this polynomial, so change of the sign doesn't affect them.*

If we have an operator, we may wish to define a characteristic polynomial of it as a characteristic polynomial of its matrix in some basis. The problem is that we don't know which basis should we choose. The following theorem shows that the choose of basis doesn't matter.

Theorem 2.4. *If A and B are 2 matrices of a linear operator, i.e. there exists an invertible matrix C such that*

$$B = C^{-1}AC,$$

then characteristic polynomials of A and B are the same.

Proof.

$$\begin{aligned}
p_B(\lambda) &= \det(B - \lambda I) \\
&= \det(C^{-1}AC - \lambda I) \\
&= \det(C^{-1}AC - C^{-1}\lambda IC) \\
&= \det(C^{-1}(A - \lambda I)C) \\
&= \det((A - \lambda I)CC^{-1}) \\
&= \det(A - \lambda I) \\
&= p_A(\lambda).
\end{aligned}$$

□

This theorem allows us to define a characteristic polynomial of the operator without choosing a particular basis.

Now our goal is to understand whether the operator is diagonalizable or not. Of course we can compute its eigenvalues. If there are n different eigenvalues, then the following theorem will show us that in this case there will be n linearly independent eigenvectors, and the basis with respect to which the operator is diagonal is just a basis, which consists of the eigenvectors.

Theorem 2.5. *Eigenvectors corresponding to different eigenvalues are linearly independent.*

Proof. The proof goes by induction. Let $\lambda_1, \lambda_2, \dots, \lambda_k$ be eigenvalues and corresponding eigenvectors are linearly independent, i.e. if $e_1, e_2, \dots, e_n$ are eigenvectors such that $\mathcal{A}(e_i) = \lambda_i e_i$ for all $i = 1, \dots, k$, and

$$d_1 e_1 + d_2 e_2 + \cdots + d_k e_k = \mathbf{0}$$

then $d_i = 0$ for all i's.

Let we add another eigenvalue λ_{k+1} and corresponding eigenvector e_{k+1}, such that $\mathcal{A}(e_{k+1}) = \lambda_{k+1} e_{k+1}$. We'll prove that vectors $e_1, e_2, \dots, e_k, e_{k+1}$ are still linearly independent. Let's consider a linear combination of them which is equal to $\mathbf{0}$:

$$c_1 e_1 + c_2 e_2 + \cdots + c_k e_k + c_{k+1} e_{k+1} = \mathbf{0}. \tag{3}$$

Now we can apply a linear operator to both sides of this equality:

$$\mathcal{A}(c_1 e_1) + \mathcal{A}(c_2 e_2) + \cdots + \mathcal{A}(c_k e_k) + \mathcal{A}(c_{k+1} e_{k+1}) = \mathbf{0}.$$

This is equivalent to

$$c_1 \mathcal{A}(e_1) + c_2 \mathcal{A}(e_2) + \cdots + c_k \mathcal{A}(e_k) + c_{k+1} \mathcal{A}(e_{k+1}) = \mathbf{0},$$

and since they are eigenvectors, i.e. $\mathcal{A}(e_i) = \lambda_i e_i$, we have

$$c_1 \lambda_1 e_1 + c_2 \lambda_2 e_2 + \cdots + c_k \lambda_k e_k + c_{k+1} \lambda_{k+1} e_{k+1} = \mathbf{0}. \tag{4}$$

Now, let's multiply the equality (3) by λ_{k+1}, and subtract from (4). We'll have:

$$c_1(\lambda_1 - \lambda_{k+1}) e_1 + c_2(\lambda_2 - \lambda_{k+1}) e_2 + \cdots + c_k(\lambda_k - \lambda_{k+1}) e_k = \mathbf{0}.$$

(note, that we don't have term with e_{k+1} anymore!). But $\lambda_{k+1} \neq \lambda_i,\ i = 1, \dots, k$. So, if $c_i \neq 0$, for all i's, we got a nontrivial linear combination of $e_1, e_2, \dots, e_k$ which is equal to zero, and vectors $e_1, e_2, \dots, e_k$ are not linearly independent. But they are linearly independent! Thus, all c_i's are equal to 0, and vectors $e_1, e_2, \dots, e_k, e_{k+1}$ are linearly independent. □

So, now we can specify the main corollary of this theorem.

Corollary 2.6. *Let $\mathcal{A}$ be a linear operator in the space V. If the characteristic polynomial of A has n different roots, then $\mathcal{A}$ is diagonalizable with respect to basis, which consists of eigenvectors.*

Now we will see, that even if there are no n different roots, then there may exist a basis of eigenvectors.

Example 2.7. *Let* $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. *Then* $A - \lambda I = \begin{pmatrix} 1-\lambda & 0 \\ 0 & 1-\lambda \end{pmatrix}$, *so*

$$p_A(\lambda) = (\lambda - 1)^2.$$

Thus, there exist only one eigenvalue $\lambda = 1$. *Subtracting* $\lambda = 1$ *from diagonal elements of* A *we get zero matrix. So, each vector is an eigenvector of* A, *and thus of course there exists a basis, consisting of eigenvectors, i.e* $e_1 = (1, 0)$, *and* $e_2 = (0, 1)$.

Example 2.8. *Let* $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. *Then* $A - \lambda I = \begin{pmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{pmatrix}$, *so*

$$p_A(\lambda) = (\lambda - 1)^2.$$

Thus, there exist only one eigenvalue $\lambda = 1$. *Subtracting* $\lambda = 1$ *from diagonal elements of* A *we get* $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. *The corresponding system is*

$$\begin{cases} 0x_1 + x_2 = 0 \\ 0x_1 + 0x_2 = 0 \end{cases}$$

So, there are no 2 linearly independent eigenvectors, because all eigenvectors have form $(c, 0)$. *So, there is no basis, consisting of eigenvectors, and thus this operator is not diagonalizable.*

3 Formulae for the characteristic polynomials of 2×2- and 3×3-matrices

Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then the characteristic polynomial is equal to:

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

Recall, that by $\operatorname{tr} A$ we denote the sum of diagonal elements of A, trace of the matrix A.

In the same way we can get a formula for the characteristic polynomial of a 3×3-matrix. Here, by A_{11}, A_{22} and A_{33} we denote the cofactors of a_{11}, a_{22}, and a_{33} correspondingly. So, if

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix},$$

then we have

$$A_{11} = \begin{vmatrix} e & f \\ h & i \end{vmatrix}, \; A_{22} = \begin{vmatrix} a & c \\ g & i \end{vmatrix}, \; A_{33} = \begin{vmatrix} a & b \\ d & e \end{vmatrix},$$

and the characteristic polynomial is equal to

$$p_A(\lambda) = \lambda^3 - (\operatorname{tr} A)\lambda^2 + (A_{11} + A_{22} + A_{33})\lambda - \det A.$$