Linear Algebra with Application

(LAWA 2021)

Lecture 12

Francesco Dolce
francesco.dolce@fjfi.cvut.cz

May 5, 2021

1 Diagonalization (continued)

In the previous lecture we defined diagonalizable and diagonalizing matrices. Here we are going to explain how to diagonalize a matrix. Let us start with an example.

Example 1 Let us diagonalize the matrix

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & -1 \\ 0 & -3 & 0 \end{pmatrix}.$$

By definition, we need to find an invertible matrix P such that

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} = \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3),$$

for certain numbers $\lambda_1, \lambda_2, \lambda_3$. Let us set

$$P = \begin{pmatrix} X_1 & X_2 & X_3 \end{pmatrix}$$

where

$$X_1 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad X_2 = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \quad \text{and} \quad X_3 = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}.$$

This is equivalent to find X_1, X_2 and X_3 such that

$$\begin{aligned}
\begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & -1 \\ 0 & -3 & 0 \end{pmatrix} \begin{pmatrix} X_1 & X_2 & X_3 \end{pmatrix} &= \begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 x_1 & \lambda_2 y_1 & \lambda_3 z_1 \\ \lambda_1 x_2 & \lambda_2 y_2 & \lambda_3 z_2 \\ \lambda_1 x_3 & \lambda_2 y_3 & \lambda_3 x_3 \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 X_1 & \lambda_2 X_2 & \lambda_3 X_3 \end{pmatrix}.
\end{aligned}$$

Comparing columns, it shows that $P^{-1}AP = \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3)$ if and only if $AX_i = \lambda_i X_i$ for $i = 1, 2, 3$. Moreover, if we want that $P = \begin{pmatrix} X_1 & X_2 & X_3 \end{pmatrix}$ is invertible, we need to make sure that $X_i \neq O$.

In the following, we begin to find λ and $X^* \neq O$ such that $AX^* = \lambda X^*$. This is equivalent to asking that the homogenous linear system

$$(A - \lambda I_3)\, X = O \tag{1}$$

has a nontrivial solution $X^* \neq O$. Using the Gaussian algorithm we reduce the matrix $(A - \lambda I_3)$ into a (reduced) row-echelon form B which is equivalent to left multiplication by a certain invertible matrix, say U, that is we have

$$U\,(A - \lambda I_3) = B.$$

By the Product Theorem we have

$$\det(U) \det(A - \lambda I_3) = \det(B).$$

Since $\det(U) \neq 0$ (the matrix is invertible), we have

$$\operatorname{rank}(A - \lambda I_3) = \operatorname{rank}(B) < n \quad \Leftrightarrow \quad \det(B) = 0 \quad \Leftrightarrow \quad \det(A - \lambda I_3) = 0.$$

Then we compute the determinant of $A - \lambda I_3$

$$\begin{aligned}
\det(A - \lambda I_3) &= \det \begin{pmatrix} 1-\lambda & 1 & 1 \\ 0 & 2-\lambda & -1 \\ 0 & -3 & -\lambda \end{pmatrix} \\
&= (1-\lambda) \det \begin{pmatrix} 2-\lambda & -1 \\ -3 & -\lambda \end{pmatrix} \\
&= (1-\lambda)\,(-\lambda(2-\lambda) - 3) \\
&= (1-\lambda)(\lambda - 3)(\lambda + 1).
\end{aligned}$$

For the equation $\det(A - \lambda I_3)$, we obtain three solutions which are

$$\lambda_1 = 1, \quad \lambda_2 = -1 \quad \text{and} \quad \lambda_3 = 3.$$

Then, we substitute each λ_i into the Equation (1) to fin a basic solution for each equation. For example, we solve

$$(A - \lambda_1 I_3) X_1 = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & -3 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = O,$$

which is equivalent to solve

$$\begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

We thus get the general solution

$$X_1^* = s \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}^T$$

where s is an arbitrary number. We can use, for instance, the basic solution

$$X_1 = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}^T$$

which is not a trivial solution as our solution corresponding to $\lambda_1 = 1$. Similarly, we can get

$$X_2 = \begin{pmatrix} -2 & 1 & 3 \end{pmatrix}^T \quad \text{and} \quad X_3 = \begin{pmatrix} 0 & -1 & 1 \end{pmatrix}^T$$

corresponding respectively to to $\lambda_2 = -1$ and $\lambda_3 = 3$. Note that here X_1, X_2, X_3 can be arbitrary nonzero solutions corresponding to $\lambda_1, \lambda_2, \lambda_3$.

Thus we can solve the equation

$$P^{-1}AP = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3) \quad \Leftrightarrow \quad AP = P\mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)$$

by obtaining

$$P = \begin{pmatrix} X_1 & X_2 & X_3 \end{pmatrix} = \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & -1 \\ 0 & 3 & 1 \end{pmatrix},$$

and

$$\mathrm{diag}(\lambda_1, \lambda_2, \lambda_3) = \mathrm{diag}(1, -1, 3).$$

Using the Matrix Inverse Algorithm we can find

$$\left(\begin{array}{ccc|ccc} 1 & -2 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ 0 & 3 & 1 & 0 & 0 & 1 \end{array}\right) \xrightarrow[\substack{R_1 \to R_1 + 2R_2 \\ R_3 \to R_3 - 3R_2}]{iii)} \left(\begin{array}{ccc|ccc} 1 & 0 & -2 & 1 & 2 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 4 & 0 & -3 & 1 \end{array}\right)$$

$$\xrightarrow[R_3 \to \frac{1}{4} R_3]{ii)} \left(\begin{array}{ccc|ccc} 1 & 0 & -2 & 1 & 2 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & -\frac{3}{4} & \frac{1}{4} \end{array}\right)$$

$$\xrightarrow[\substack{R_1 \to R_1 + 2R_3 \\ R_2 \to R_2 + R_3}]{iii)} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & 0 & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 1 & 0 & -\frac{3}{4} & \frac{1}{4} \end{array}\right).$$

Thus P is invertible and

$$P^{-1} = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{4} & \frac{1}{4} \\ 0 & -\frac{3}{4} & \frac{1}{4} \end{pmatrix}.$$

In conclusion, we have

$$P^{-1}AP = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{4} & \frac{1}{4} \\ 0 & -\frac{3}{4} & \frac{1}{4} \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & -1 \\ 0 & -3 & 0 \end{pmatrix} \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & -1 \\ 0 & 3 & 1 \end{pmatrix} = \operatorname{diag}(1, -1, 3).$$

We can generalize the previous example to an $n \times n$-matrix. Finding P such that $P^{-1}AP$ is a diagonal matrix is equivalent to find n column vectors $X_1, X_2, \ldots, X_n$ and n numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that

$$AX_i = \lambda_i X_i \quad \text{for each } \ i = 1, 2, \ldots n.$$

Moreover, if $P = \begin{pmatrix} X_1 & X_2 & \cdots & X_n \end{pmatrix}$ is invertible, A is diagonalizable.

2 Eigenvalues and Eigenvectors

Let $A \in \mathcal{M}_{n,n}(\mathbb{R})$. A number λ is called an *eigenvalue* of A if

$$AX = \lambda X$$

for some column $X \neq O$. Such a nonzero column X is called an *eigenvector* of A corresponding to the eigenvalue λ.

Note that the condition $AX = \lambda X$ is automatically satisfied if $X = O$, so the requirement that $X \neq O$ is critical.

The *characteristic polynomial* $c_A(x)$ is defined by

$$c_A(x) = \det(xI - A).$$

A number λ is called a *root* of the characteristic polynomial $c_A(x)$ if $c_A(\lambda) = 0$.

Note that $c_A(\lambda) = 0$ if and only if $-c_A(\lambda) = 0$. For this reason, in the following we will work indifferently with both equations $\det(xI - A) = 0$ and $\det(A - \lambda I) = 0$.

Example 2 Let us consider the matrix

$$A = \begin{pmatrix} 5 & -2 \\ 4 & -1 \end{pmatrix}.$$

Its characteristic polynomial is

$$\begin{aligned} c_A(x) &= \det(xI - A) \\ &= \det\left(x\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 5 & -2 \\ 4 & -1 \end{pmatrix}\right) \\ &= \det\left(\begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix} - \begin{pmatrix} 5 & -2 \\ 4 & -1 \end{pmatrix}\right) \\ &= \det\begin{pmatrix} x-5 & 2 \\ -4 & x+1 \end{pmatrix} \\ &= (x-5)(x+1) - 2(-4) \\ &= x^2 - 4x + 3 \\ &= (x-1)(x-3). \end{aligned}$$

The two roots of the characteristic polynomial are thus $\lambda_1 = 1$ and $\lambda_2 = 3$.

Theorem 3 *Let A be a $n \times n$-matrix.*

1. *The eigenvalues of A are the roots of the characteristic polynomial $c_A(x)$ of A.*

2. *The eigenvectors X corresponding to the eigenvalues λ are the nonzero solutions to the homogenous system of linear equations $(\lambda I - A)X = O$.*

Note that there are many eigenvectors of a square matrix A associated with a given eigenvalue λ. In fact every nonzero solution X of $(\lambda I - A)X = O$ is an eigenvector. Of course the eigenvalue λ is chosen so that there must be nonzero solutions.

The eigenvalues of a real matrix need not to be real numbers.

Example 4 Let us find the eigenvalues of the matrix

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

The characteristic polynomial of the matrix is $\det(xI - A) = x^2 + 1$. So by Theorem 3, the eigenvalues of A are the nonreal complex roots $\lambda_1 = i$ and $\lambda_2 = -i$.

A $n \times n$-matrix has n (possibly complex) eigenvalues, but they may not be distinct.

Example 5 Let us find the eigenvalues of the matrix

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

Its characteristic polynomial is $c_A(x) = (x-1)^2$. So there is only one eigenvalue of A, namely $\lambda_1 = 1$. However, λ_1 is a double root of $c_A(x)$ and we say that $\lambda_1 = 1$ has *multiplicity* 2.

The following result illustrate the previous example

Theorem 6 *Let $A \in \mathcal{M}_{n,n}(\mathbb{R})$.*

1. *A is diagonalizable if and only if it has eigenvectors $X_1, X_2, \ldots, X_n$ such that the matrix*
$$P = \begin{pmatrix} X_1 & X_2 & \cdots & X_n \end{pmatrix}$$
is invertible.

2. *When this is the case, we have*
$$P^{-1}AP = diag(\lambda_1, \lambda_2, \ldots, \lambda_n)$$
where, for each i, λ_i is the eigenvalue of A corresponding to X_i.

Example 7 Let us show that the matrix
$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$
is not diagonalizable.

We know from Example 5 that A has only one eigenvalue $\lambda_1 = 1$, which is of multiplicity 2. But the system of linear equations $(\lambda_1 I - A)X = O$ has general solution
$$X = s\begin{pmatrix} 1 \\ 0 \end{pmatrix},$$
so there is only one basic solution:
$$X_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$
Hence we can only choose
$$P = \begin{pmatrix} s & t \\ 0 & 0 \end{pmatrix}$$
which is never invertible no matter the choice of s and t.