Linear Algebra with Application

(LAWA 2021)

Lecture 13

Francesco Dolce
francesco.dolce@fjfi.cvut.cz

May 12, 2021

1 The Diagonalization Algorithm

In this section we give an altoritm to diagonalize a square matrix.

Diagonalization Algorithm. *Let $A \in \mathcal{M}_{n,n}(\mathbb{R})$ be a square matrix. To diagonalize A we apply the following steps:*

Step 1. *Find all the eigenvalues of A, which are the roots of the characteristic polynomial $c_A(x)$;*

Step 2. *For each eigenvalue λ compute an eigenvector, by finding the basic solution of the homogenous system $(\lambda I - A)X = O$;*

Step 3. *The matrix A is diagonalizable if and only if there are n basic eigenvectors in total;*

Step 4. *If A is diagonalizable, the $n \times n$-matrix P having these eigenvectors as columns is a diagonalizing matrix for A; that is, P is invertible and $P^{-1}AP$ is diagonal.*

Example 1 Let us apply the previous algorithm to the matrix

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

Step 1. Let us first compute the characteristic polynomial

$$\begin{aligned} c_A(x) &= \det(xI - A) \\ &= \det \begin{pmatrix} x & -1 & -1 \\ -1 & x & -1 \\ -1 & -1 & x \end{pmatrix} \\ &= \det \begin{pmatrix} x-2 & x-2 & x-2 \\ -1 & x & -1 \\ -1 & -1 & x \end{pmatrix} \\ &= \det \begin{pmatrix} x-2 & 0 & 0 \\ -1 & x+1 & 0 \\ -1 & 0 & x+1 \end{pmatrix} \\ &= (x-2)(x+1)^2, \end{aligned}$$

where to compute the determinant we first added the second and the third row to the first row, and then we subtracted the first column from the second and from the third column.

Hence, the equation $c_A(x) = 0$ has two solutions: $\lambda_1 = 2$ and $\lambda_2 = -1$, with the last one having multiplicity two.

Step 2. For $\lambda_1 = 2$, the system

$$(\lambda_1 I - A)X = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} X = O$$

solution

$$X = s \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

where t is an arbitrary number. So the basic solution

$$X_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

is an eigenvector corresponding to $\lambda_1 = 2$.

For $\lambda_2 = -1$, the system

$$(\lambda_2 I - A)X = \begin{pmatrix} -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \end{pmatrix} X = O$$

has general solution

$$X = s\begin{pmatrix}-1\\1\\0\end{pmatrix} + t\begin{pmatrix}-1\\0\\1\end{pmatrix}$$

where s and t are arbitrary numbers. Hence there are two basic solutions

$$X_2 = \begin{pmatrix}-1\\1\\0\end{pmatrix} \quad \text{and} \quad X_3 = \begin{pmatrix}-1\\0\\1\end{pmatrix}$$

corresponding to $\lambda_2 = -1$.

Step 3. Since there are three eigenvectors, X_1, X_2 and X_3, we can deduce that A is diagonalizable.

Step 4. If we take

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

we find that P is invertible and

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

Thus

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

In a general case, an eigenvalue λ of a square matrix A is said to have multiplicity m if it occurs m times as a root of the characteristic polynomial $c_A(x)$. When the homogenous system $(\lambda I - A)X = O$ is solved, any set of basic solutions is called a set of *basic eigenvectors* corresponding to λ. Here the number of basic eigenvectors equals the number of parameters involved in the solution of the system $(\lambda I - A)X = O$.

Theorem 2 *A square matrix A is diagonalizable if and only if it the multiplicity of every eigenvalue λ of A equals the number of basic eigenvectors corresponding to λ (which is the number of parameters in the solution of $(\lambda I - A)X = O$).*

In this case, the basic solutions of the system $(\lambda I - A)X = O$ become columns in the invertible diagonalizing matrix P such that $P^{-1}AP$ is diagonal.

Since for each eigenvalues there is at least a basic eigenvector, we have the following immediate consequence of the previous theorem.

Corollary 3 *If A is a $n \times n$-matrix with n distinct eigenvalues, then A is diagonalizable.*

A good example which illustrate an application of diagonalization is given in the following example.

Example 4 Let us compute A^{100} for

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

As we have already seen in the example in the previous lecture, the matrix A has eigenvalues

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

with corresponding eigenvectors

$$X_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad X_2 = \begin{pmatrix} -2 \\ 1 \\ 3 \end{pmatrix} \quad \text{and} \quad X_3 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

A diagonalizing matrix for A is thus given by the invertible matrix

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

having inverse

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

We thus have

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

Thus, using Theorem 11 in Lecture 11, we have

$$\begin{aligned}
A^{100} &= P \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{pmatrix}^{100} P^{-1} \\
&= P \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3^{100} \end{pmatrix} P^{-1} \\
&= \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & -1 \\ 0 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3^{100} \end{pmatrix} \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{4} & \frac{1}{4} \\ 0 & -\frac{3}{4} & \frac{1}{4} \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1+3^{101}}{4} & \frac{1-3^{100}}{4} \\ 0 & \frac{3-3^{101}}{4} & \frac{3+3^{100}}{4} \end{pmatrix}.
\end{aligned}$$

Example 5 Let us consider the matrix

$$A = \begin{pmatrix} 3 & -4 & 2 \\ 1 & -2 & 2 \\ 1 & -5 & 5 \end{pmatrix}.$$

We can compute A^{20} (to do so we need first to diagonalize the matrix). (`Exercise`)

2 Similar matrices

Let us consider two square matrices A and B of the same size. We say that A and B are *similar* if

$$B = P^{-1}AP$$

for some invertible matrix P. When this is the case, we write $A \sim B$.

Using this terminology, we can say that a square matrix A is diagonalizable if and only if it is similar to a diagonal matrix.

Here are some simply properties of similarity.

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

1. *$A \sim A$.*
2. *If $A \sim B$ then $B \sim A$.*
3. *If $A \sim B$ and $B \sim C$, then $A \sim C$.*

Proof.

- The first point is clear since $A = I^{-1}AI$, and I is invertible.
- If $A \sim B$ then there exists an invertible matrix P such that $B = P^{-1}AP$. Thus $A = PBP^{-1}$, with P^{-1} invertible. That is, $B \sim A$.
- Let P and Q be two invertible matrices such that $B = P^{-1}AP$ and $C = Q^{-1}BQ$. Thus

  $$C = Q^{-1}\left(P^{-1}AP\right)Q = \left(Q^{-1}P^{-1}\right)A\,(PQ) = (PQ)^{-1}A\,(PQ),$$

  with PQ invertible. Hence $A \sim C$.

∎

The properties in the previous proposition are often expressed by saying that the similarity relation $\sim$ is an *equivalence relation* on the set of $n \times n$-matrices.

Proposition 7 *Let A, B be two square matrices such that $A \sim B$. Then*

1. *A is invertible if and only if B is invertible, and in this case $A^{-1} \sim B^{-1}$.*
2. *$A^T \sim B^T$.*

3. $A^k \sim B^k$ *for all* $k \geq 0$.

Example 8 Let A, B be two square matrices such that $A \sim B$. If A is diagonalizable, then B is also diagonalizable. (`Exercise`)

Following the previous example, it is possible to prove that if A is diagonalizable, then so are also the matrices A^T, A^{-1} (if it exists) and A^k for all $k \geq 0$.

The following theorem easily follows from the Product Theorem and the Diagonalization Algorithm.

Theorem 9 *Let* A, B *be two similar matrices. Then*

1. $det(A) = det(B)$.
2. $c_A(x) = c_B(x)$.
3. A *and* B *have the same eigenvalues.*

3 Cayley-Hamilton Theorem

Example 10 Let

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

Then $c_A(A) = 0$.

Indeed, we saw in Example 1 that the characteristic polynomial is

$$c_A(x) = (x-2)(x+1)^2$$

When we *evaluate* this polynomial at A, we obtain $c_A(A) = (A - 2I)(A + I)^2$.

Let us prove that this evaluation equals wero. Recall from Example 1 that $A = PDP^{-1}$ with

$$P = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \quad \text{and} \quad D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

So

$$\begin{aligned}
c_A(A) &= (A-2I)\,(A+I)^2 \\
&= \left(PDP^{-1}-2I\right)\left(PDP^{-1}+I\right)^2 \\
&= \left(PDP^{-1}-2PIP^{-1}\right)\left(PDP^{-1}+PIP^{-1}\right)^2 \\
&= \left(P(D-2I)P^{-1}\right)\left(P(D+I)P^{-1}\right)^2 \\
&= P(D-2I)P^{-1}P(D+I)P^{-1}P(D+I)P^{-1} \\
&= P(D-2I)(D+I)(D+I)P^{-1} \\
&= P\begin{pmatrix}0 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & -1\end{pmatrix}\begin{pmatrix}3 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}^2 P^{-1} \\
&= P\begin{pmatrix}0 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & -1\end{pmatrix}\begin{pmatrix}9 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix} P^{-1} \\
&= P\begin{pmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix} P^{-1} \\
&= O.
\end{aligned}$$

We can generalize the previous example in the following importan theorem

Theorem 11 (Cayley-Hamilton Theorem) *Let A be a square matrix. Thus $c_A(A) = O$.*