Linear Algebra 2: Lecture 20

Irena Penev

Summer 2023

1 The Jordan normal form

Given a field $\mathbb{F}$, a scalar $\lambda_0 \in \mathbb{F}$, and a positive integer k, the *Jordan block* $J_k(\lambda_0)$ is defined to be the matrix

$$J_k(\lambda_0) = \begin{bmatrix} \lambda_0 & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_0 & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda_0 \end{bmatrix}_{k \times k}.$$

Thus, $J_k(\lambda_0)$ is a matrix in $\mathbb{F}^{k \times k}$, it has all λ_0's on the main diagonal, all 1's on the diagonal right above the main diagonal, and 0's everywhere else. For example:

- $J_1(\lambda_0) = \begin{bmatrix} \lambda_0 \end{bmatrix}$;
- $J_2(\lambda_0) = \begin{bmatrix} \lambda_0 & 1 \\ 0 & \lambda_0 \end{bmatrix}$;
- $J_3(\lambda_0) = \begin{bmatrix} \lambda_0 & 1 & 0 \\ 0 & \lambda_0 & 1 \\ 0 & 0 & \lambda_0 \end{bmatrix}$;
- $J_4(\lambda_0) = \begin{bmatrix} \lambda_0 & 1 & 0 & 0 \\ 0 & \lambda_0 & 1 & 0 \\ 0 & 0 & \lambda_0 & 1 \\ 0 & 0 & 0 & \lambda_0 \end{bmatrix}$;
- $J_5(\lambda_0) = \begin{bmatrix} \lambda_0 & 1 & 0 & 0 & 0 \\ 0 & \lambda_0 & 1 & 0 & 0 \\ 0 & 0 & \lambda_0 & 1 & 0 \\ 0 & 0 & 0 & \lambda_0 & 1 \\ 0 & 0 & 0 & 0 & \lambda_0 \end{bmatrix}$.

A *Jordan matrix* (also called a matrix in *Jordan normal form*) is a square matrix that has Jordan blocks along the main diagonal, and has 0's everywhere else. Thus, a Jordan matrix is a matrix of the form

$$\begin{bmatrix} J_{k_1}(\lambda_1) & O & \dots & O \\ O & J_{k_2}(\lambda_2) & \dots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \dots & J_{k_\ell}(\lambda_\ell) \end{bmatrix},$$

where $\lambda_1, \dots, \lambda_\ell$ are scalars in $\mathbb{F}$, $k_1, \dots, k_\ell$ are positive integers, and the O's are zero matrices of appropriate sizes. For instance, the following is a Jordan matrix with four Jordan blocks, namely $J_3(3)$, $J_2(2)$, $J_1(2)$, and $J_3(3)$:

$$\begin{bmatrix} 3 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \end{bmatrix}$$

Remark: Note that any two Jordan matrices that have exactly the same Jordan blocks (counting repetitions) are similar. This is because (by Theorem 4.3 of Lecture Notes 10) similar matrices represent the same linear transformation, only with respect to (possibly) different bases. A change in the order of Jordan blocks corresponds to a change in the order of basis elements. For instance, suppose that V is a finite-dimensional vector space over a field $\mathbb{F}$, that $f : V \to V$ is a linear transformation, and that $\mathcal{B} = \{\mathbf{a}_1, \dots, \mathbf{a}_{k_1}, \mathbf{b}_1, \dots, \mathbf{b}_{k_2}, \mathbf{c}_1, \dots, \mathbf{c}_{k_3}, \mathbf{d}_1, \dots, \mathbf{d}_{k_4}\}$ (with $k_1, k_2, k_3, k_4 \geq 1$) is a basis of V such that

$$_{\mathcal{B}}[f]_{\mathcal{B}} = \begin{bmatrix} J_{k_1}(\lambda_1) & O & O & O \\ O & J_{k_2}(\lambda_2) & O & O \\ O & O & J_{k_3}(\lambda_3) & O \\ O & O & O & J_{k_4}(\lambda_4) \end{bmatrix}.$$

Then for the basis $\mathcal{C} = \{\mathbf{b}_1, \dots, \mathbf{b}_{k_2}, \mathbf{d}_1, \dots, \mathbf{d}_{k_4}, \mathbf{a}_1, \dots, \mathbf{a}_{k_1}, \mathbf{c}_1, \dots, \mathbf{c}_{k_3}\}$, we have

$$_{\mathcal{C}}[f]_{\mathcal{C}} = \begin{bmatrix} J_{k_2}(\lambda_2) & O & O & O \\ O & J_{k_4}(\lambda_4) & O & O \\ O & O & J_{k_1}(\lambda_1) & O \\ O & O & O & J_{k_3}(\lambda_3) \end{bmatrix}.$$

Clearly, the two Jordan matrices above are similar, since $_{\mathcal{B}}[f]_{\mathcal{B}}$ and $_{\mathcal{C}}[f]_{\mathcal{C}}$ are similar.

Theorem 1.1. *Assume that $\mathbb{F}$ is an* ***algebraically closed field****, and let $A \in \mathbb{F}^{n\times n}$ be a matrix. Then there exist scalars $\lambda_1, \ldots, \lambda_\ell \in \mathbb{F}$ and positive integers $k_1, \ldots, k_\ell$ such that $k_1 + \cdots + k_\ell = n$ and such that A is similar to the matrix*

$$\begin{bmatrix} J_{k_1}(\lambda_1) & O & \ldots & O \\ O & J_{k_2}(\lambda_2) & \ldots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \ldots & J_{k_\ell}(\lambda_\ell) \end{bmatrix},$$

called the Jordan normal form *of A. The scalars $\lambda_1, \ldots, \lambda_\ell$ and integers $k_1, \ldots, k_\ell$ are unique up to a reordering of the λ_i's and the corresponding k_i's.*

Proof. Omitted. □

Remarks:

1. Theorem 1.1 only holds for **algebraically closed fields**. The only algebraically closed field that we have seen is $\mathbb{C}$, but others exist.

2. Two matrices are similar if and only if they have the same Jordan normal form, up to a reordering of the Jordan blocks. Indeed, as we saw above, any two Jordan matrices with the same Jordan blocks (counting repetitions) are similar. The other direction follows from the uniqueness part of Theorem 1.1.

3. Since every Jordan matrix is upper triangular, its eigenvalues, together with their algebraic multiplicities, can easily be read off the Jordan matrix itself: the eigenvalues are precisely the entries along the main diagonal of the Jordan matrix, and the algebraic multiplicity of each eigenvalue is the number of times that it appears on the main diagonal. For instance, the eigenvalues of the Jordan matrix

$$\begin{bmatrix} 3 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \end{bmatrix}$$

are 3 (with algebraic multiplicity 6) and 2 (with algebraic multiplicity 3).

4. Perhaps more interestingly, the geometric multiplicity of each eigenvalue of a Jordan matrix J can also easily be read off the Jordan matrix J: the geometric multiplicity of each eigenvalue λ is precisely the number of Jordan blocks of the form $J_k(\lambda)$ that appear along the main diagonal of J.[1] For instance, for the Jordan matrix above, the geometric multiplicity of the eigenvalue 3 is 2, and the geometric multiplicity of the eigenvalue 2 is also 2. Recall that similar matrices have the same eigenvalues, with the same corresponding algebraic (respectively, geometric) multiplicities. So, if we know the Jordan normal form of a matrix A, then we can easily read off the eigenvalues of A, together with their algebraic and geometric multiplicities.

5. For the case when $\mathbb{F}$ is an algebraically closed field, Theorem 1.1 yields another proof of the fact that the geometric multiplicity of an eigenvalue λ of a matrix $A \in \mathbb{F}^{n\times n}$ is no greater than the algebraic multiplicity of λ: for the Jordan normal form J of A, the number of Jordan blocks of the form $J_k(\lambda)$ that appear along the main diagonal of J is no greater than the total number of times that λ appears on the main diagonal of the Jordan matrix J. If $\mathbb{F}$ is not algebraically closed, this can be handled by first extending $\mathbb{F}$ to an algebraically closed field (for example, extending $\mathbb{R}$ to $\mathbb{C}$), and then considering the Jordan normal form of A, now considered as a matrix whose entries come from this algebraically closed field. Any field $\mathbb{F}$ can be extended in this way, but the details are beyond the scope of this course.

Example 1.2. *Let $A_1, A_2, A_3 \in \mathbb{C}^{7\times 7}$ be matrices whose Jordan normal forms are, respectively, the matrices J_1, J_2, J_3 below.*

- $J_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$;

- $J_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$;

[1] Check this!

- $J_3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.

Determine which (if any) of A_1, A_2, A_3 are similar. Then, for each $i \in \{1, 2, 3\}$, find all the eigenvalues of A_i, along with their algebraic and geometric multiplicities.

Solutution. We use colors to indicate the Jordan blocks of the three Jordan matrices.

- $J_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$;

- $J_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$;

- $J_3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.

We now list the Jordan blocks of J_1, J_2, J_3 (including any repetitions:

- the Jordan blocks of J_1 are $J_1(0), J_3(1), J_1(1), J_2(0)$;
- the Jordan blocks of J_2 are $J_1(1), J_2(0), J_1(0), J_3(1)$;
- the Jordan blocks of J_3 are $J_1(0), J_2(1), J_2(1), J_2(0)$.

J_1 and J_2 have the same Jordan blocks (counting repetitions), and so A_1 and A_2 are similar. On the other hand, the Jordan blocks of the matrix J_3 are different from those of J_1 and J_2, and so A_3 is not similar to A_1 and A_2. Further, we see from the matrices J_1, J_2, J_3, that A_1, A_2, A_3 all have exactly two eigenvalues: the eigenvalue 0 with algebraic multiplicity 3 and geometric multiplicity 2, and the eigenvalue 1 with algebraic multiplicity 4 and geometric multiplicity 2. □

Remark: By Proposition 2.1 of Lecture Notes 19, similar matrices have the same eigenvalues, with the same corresponding algebraic multiplicities, and the same corresponding geometric multiplicities. However, the solution of Example 1.2 shows that the converse does not hold in general: square matrices that have the same eigenvalues, with the same corresponding algebraic and geometric multiplicities, need not be similar.

Computing the Jordan normal form of an arbitrary square matrix (with entries in some algebraically closed field) is quite complicated in general. However, in some special cases, this can be done relatively easily, as the following example shows.

Example 1.3. *Consider the matrix*

$$A = \begin{bmatrix} 5 & -2 & 2 & -2 & 0 \\ 0 & 6 & -1 & 3 & 2 \\ 2 & 2 & 7 & -2 & -2 \\ 2 & 3 & 1 & 2 & -4 \\ -2 & -2 & -2 & 6 & 11 \end{bmatrix}$$

in $\mathbb{C}^{5\times 5}$*. Compute the Jordan normal form of* A*.*

Solution. First, we compute the characteristic polynomial of A: $p_A(\lambda) = \det(\lambda I_5 - A) = (\lambda - 7)^3(\lambda - 5)^2$. So, A has two eigenvalues: $\lambda_1 = 7$ (with algebraic multiplicity 3) and $\lambda_2 = 5$ (with algebraic multiplicity 2).

We now compute the Jordan blocks associated with each eigenvalue.

Since the eigenvalue $\lambda_1 = 7$ has algebraic multiplicity 3, there are three possibilities: one $J_3(7)$ block; one $J_2(7)$ block and one $J_1(7)$ block; or three $J_1(7)$ blocks. But by row reducing the matrix $7I_5 - A$, we see that $\lambda_1 = 7$ is an eigenvalue of geometric multiplicity 2. Consequently, the number of Jordan blocks associated with the eigenvalue $\lambda_1 = 7$ is two. So, we get blocks $J_2(7)$ and $J_1(7)$.

For the eigenvalue $\lambda_2 = 5$, there are two possibilities: one $J_2(5)$ block, or two $J_1(5)$ blocks. But by row reducing the matrix $5I_5 - A$, we see that $\lambda_1 = 5$ is an eigenvalue of geometric multiplicity 2. Consequently, the number of Jordan blocks associated with the eigenvalue $\lambda_2 = 5$ is two. So, we get two $J_1(5)$ blocks.

Thus, the Jordan normal form of A has four Jordan blocks, namely, $J_2(7), J_1(7), J_1(2), J_1(2)$. The actual Jordan normal form of A can be obtained by arranging these blocks along the main diagonal in any order. For instance, either of the following matrices is the Jordan normal form of A:

- $\begin{bmatrix} 7 & 1 & 0 & 0 & 0 \\ 0 & 7 & 0 & 0 & 0 \\ 0 & 0 & 7 & 0 & 0 \\ 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix}$;
- $\begin{bmatrix} 5 & 0 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 7 & 0 & 0 \\ 0 & 0 & 0 & 7 & 1 \\ 0 & 0 & 0 & 0 & 7 \end{bmatrix}$.

□

1.1 A (very) brief outline of the proof of Theorem 1.1

For convenience, Theorem 1.1 is restated below.

Theorem 1.1. *Assume that $\mathbb{F}$ is an* ***algebraically closed field****, and let $A \in \mathbb{F}^{n\times n}$ be a matrix. Then there exist scalars $\lambda_1, \dots, \lambda_\ell \in \mathbb{F}$ and positive integers $k_1, \dots, k_\ell$ such that $k_1 + \cdots + k_\ell = n$ and such that A is similar to the matrix*

$$\begin{bmatrix} J_{k_1}(\lambda_1) & O & \dots & O \\ O & J_{k_2}(\lambda_2) & \dots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \dots & J_{k_\ell}(\lambda_\ell) \end{bmatrix},$$

called the Jordan normal form *of A. The scalars $\lambda_1, \dots, \lambda_\ell$ and integers $k_1, \dots, k_\ell$ are unique up to a reordering of the λ_i's and the corresponding k_i's.*

Let us very briefly discuss the idea of the proof of Theorem 1.1. The proof proceeds in two stages. In the first stage, it is shown that A is similar to some matrix

$$\begin{bmatrix} A_1 & O & \dots & O \\ O & A_2 & \dots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \dots & A_k \end{bmatrix},$$

where each A_i is a square matrix that has exactly one eigenvalue, and no two of $A_1, \dots, A_k$ share an eigenvalue. This reduces the problem to showing that each square matrix with exactly one eigenvalue is similar to some Jordan matrix.

So, suppose that A has exactly one eigenvalue, say λ_0. We may in fact assume that $\lambda_0 = 0$, for otherwise, we consider $A' := A - \lambda_0 I_n$ instead of A. (If we can show that A' is similar to a Jordan matrix J, then A is similar to the Jordan matrix $J + \lambda_0 I_n$.) Now that we have reduced the problem to case when 0 is the only eigenvalue of A, we proceed as follows. It is not hard to show that there exists some positive integer p such that

$$\text{Nul}(A) \subsetneqq \text{Nul}(A^2) \subsetneqq \dots \subsetneqq \text{Nul}(A^p) = \text{Nul}(A^{p+1}) = \dots,$$

and moreover, it can be shown (with a bit more effort) that $\text{Nul}(A^p) = \mathbb{F}^n$. Now, let $f_A : \mathbb{F}^n \to \mathbb{F}^n$ be given by $\mathbf{x} \mapsto A\mathbf{x}$, so that A is the standard matrix of f_A. The goal is now to construct a basis $\mathcal{B}$ such that the matrix ${}_{\mathcal{B}}[f_A]_{\mathcal{B}}$ is in Jordan normal form. First, we consider any basis of $\text{Nul}(A^{p-1})$, and then we extend it to a basis of $\text{Nul}(A^p) = \mathbb{F}^n$ using some vectors $\mathbf{u}_1, \dots, \mathbf{u}_r \in \text{Nul}(A^p) \setminus \text{Nul}(A^{p-1})$. The first vectors of our basis $\mathcal{B}$ will be the vectors $A^{p-1}\mathbf{u}_1, A^{p-2}\mathbf{u}_1, \dots, A\mathbf{u}_1, \mathbf{u}_1, \dots, A^{p-1}\mathbf{u}_k, A^{p-2}\mathbf{u}_k, \dots, A\mathbf{u}_k, \mathbf{u}_k$. If this list has n vectors, then we have our basis $\mathcal{B}$. Otherwise, we will keep adding chains of the form $A^{q-1}\mathbf{u}, A^{q-2}\mathbf{u}, \dots, A\mathbf{u}, \mathbf{u}$. Each such chain $A^{q-1}\mathbf{u}, A^{q-2}\mathbf{u}, \dots, A\mathbf{u}, \mathbf{u}$ will produce the Jordan block $J_q(0)$ in the matrix ${}_{\mathcal{B}}[f_A]_{\mathcal{B}}$.

2 Symmetric matrices and orthogonal diagonalization

Recall that the *standard scalar product* on $\mathbb{C}^{n\times n}$, denoted by $\cdot$, is defined as follows: for all $\mathbf{x} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T$ and $\mathbf{y} = \begin{bmatrix} y_1 & \dots & y_n \end{bmatrix}^T$ in $\mathbb{C}^n$, we set

$$\mathbf{x} \cdot \mathbf{y} = \sum_{k=1}^{n} x_k \overline{y_k}.$$

We shall denote by $||\cdot||$ the norm on $\mathbb{C}^{n\times n}$ induced by the standard scalar product on $\mathbb{C}^{n\times n}$.

For any field $\mathbb{F}$, the matrix $A \in \mathbb{F}^{n\times n}$ is *symmetric* if $A^T = A$. If $\mathbb{F} = \mathbb{C}$, then it turns out that symmetric matrices are less interesting than the so-called "Hermitian matrices." For a matrix $A = [a_{i,j}]_{n\times m}$ in $\mathbb{C}^{n\times m}$, we set $\overline{A} = [\overline{a_{i,j}}]_{n\times m}$, i.e. for the i,j-th entry of $\overline{A}$ is the $\overline{a_{i,j}}$ (the complex conjugate of $a_{i,j}$). The *Hermitian transpose* of a matrix $A \in \mathbb{C}^{n\times m}$ is the matrix $A^* = (\overline{A})^T$. For example, if

$$A = \begin{bmatrix} -1+i & 3 & 2i \\ 1+2i & 4-2i & 3 \end{bmatrix},$$

then

$$\overline{A} = \begin{bmatrix} -1-i & 3 & -2i \\ 1-2i & 4+2i & 3 \end{bmatrix} \quad \text{and} \quad A^* = \begin{bmatrix} -1-i & 1-2i \\ 3 & 4+2i \\ -2i & 3 \end{bmatrix}.$$

A square matrix $A \in \mathbb{C}^{n \times n}$ is *Hermitian* if $A^* = A$. For example, the matrix

$$\begin{bmatrix} -1 & 1+i & 2-i \\ 1-i & 2 & -3+i \\ 2+i & -3-i & 0 \end{bmatrix}$$

is Hermitian. Note that all entries on the main diagonal of a Hermitian matrix are real. Note also that if all entries of a matrix in $\mathbb{C}^{n \times n}$ happen to be real, then that matrix is Hermitian if and only if it is symmetric.

Proposition 2.1. *For all* $\mathbf{x} \in \mathbb{C}^n$, *we have that* $\mathbf{x}^*\mathbf{x} = ||\mathbf{x}||^2$.

Proof. Fix a vector $\mathbf{x} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T$ in $\mathbb{C}^n$. Then

$$\begin{aligned} \mathbf{x}^*\mathbf{x} &= \begin{bmatrix} \overline{x_1} & \dots & \overline{x_n} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \\ &= \sum_{k=1}^{n} \overline{x_k} x_k \\ &= \sum_{k=1}^{n} x_k \overline{x_k} \\ &= \mathbf{x} \cdot \mathbf{x} \\ &= ||\mathbf{x}||^2. \end{aligned}$$

□

Proposition 2.2. *For all matrices* $A, B \in \mathbb{C}^{n \times m}$ *and scalars* $\alpha \in \mathbb{C}$, *the following hold:*

(a) $(A^*)^* = A$;

(b) $(\alpha A)^* = \overline{\alpha} A^*$;

(c) $(A + B)^* = A^* + B^*$;

(d) $(AB)^* = B^* A^*$.

Proof. Exercise. □

Theorem 2.3. *All eigenvalues of a Hermetrian matrix are real.*

Remark: Recall that, since the field $\mathbb{C}$ is algebraically closed, every matrix in $\mathbb{C}^{n \times n}$ has n complex eigenvalues (with algebraic multiplicities taken into account). So, Theorem 2.3 states that if A is a Hermitian matrix in $\mathbb{C}^{n \times n}$, then all n eigenvalues of A (with algebraic multiplicities taken into account) are real.

Proof. Let $A \in \mathbb{C}^{n\times n}$ be Hermitian matrix, let λ be any eigenvalue of A, and let $\mathbf{x}$ be an associated eigenvector of A. After possibly rescaling (i.e. replacing $\mathbf{x}$ by $\frac{\mathbf{x}}{||\mathbf{x}||}$), we may assume that $\mathbf{x}$ is a unit vector, i.e. that it satisfies $||\mathbf{x}|| = 1$. Then $A\mathbf{x} = \lambda\mathbf{x}$, and we have that

$$\begin{aligned}
\mathbf{x}^*A\mathbf{x} &= \mathbf{x}^*(\lambda\mathbf{x}) && \text{because } A\mathbf{x} = \lambda\mathbf{x} \\
&= \lambda(\mathbf{x}^*\mathbf{x}) \\
&= \lambda||\mathbf{x}||^2 && \text{by Proposition 2.1} \\
&= \lambda && \text{because } ||\mathbf{x}|| = 1.
\end{aligned}$$

But we now have the following:

$$\begin{aligned}
\lambda &= \mathbf{x}^*A\mathbf{x} \\
&= \mathbf{x}^*A^*\mathbf{x} && \text{because } A \text{ is Hermitian} \\
&= \mathbf{x}^*A^*(\mathbf{x}^*)^* && \text{by Proposition 2.2(a)} \\
&= (\mathbf{x}^*A\mathbf{x})^* && \text{by Proposition 2.2(d)} \\
&= \lambda^* && \text{where we consider } \lambda \text{ as a } 1\times 1 \text{ complex matrix} \\
&= \overline{\lambda} && \text{where we consider } \lambda \text{ as a complex number.}
\end{aligned}$$

We have now shown that $\lambda = \overline{\lambda}$, and it follows that λ is a real number. □

Corollary 2.4. *Every symmetric matrix in* $\mathbb{R}^{n\times n}$ *has* n *real eigenvalues (with algebraic multiplicities taken into account). In other words, for every symmetric matrix* $A \in \mathbb{R}^{n\times n}$*, the sum of algebraic multiplicities of its distinct (real) eigenvalues is* n*.*

Proof. Consider any symmetric matrix $A \in \mathbb{R}^{n\times n}$. If we consider A as a matrix in $\mathbb{C}^{n\times n}$, then A is in fact Hermitian,[2] and so by Theorem 2.3, all n eigenvalues (with algebraic multiplicities taken into account) of A are real. □

Recall that a matrix $Q \in \mathbb{R}^{n\times n}$ is *orthogonal* if $Q^TQ = I_n$. By Theorem 2.1 of Lecture Notes 14, the following are equivalent for any matrix Q in $\mathbb{R}^{n\times n}$:

[2]Indeed, $A^* = (\overline{A})^T \overset{(*)}{=} A^T \overset{(**)}{=} A$, where (*) follows from the fact that all entries of A are real (and so $\overline{A} = A$), and (**) follows from the fact that that A is symmetric.

- Q is orthogonal;
- Q is invertible and satisfies $Q^{-1} = Q^T$;
- the columns of Q form an an orthonormal basis of $\mathbb{R}^n$.

In what follows, we will repeatedly use the fact that the three statements above are equivalent, without mentioning Theorem 2.1 of Lecture Notes 14 explicitly.

Let us say that a matrix $A \in \mathbb{R}^{n\times n}$ is *orthogonally diagonalizable* if there exists a diagonal matrix D and an orthogonal matrix Q, both in $\mathbb{R}^{n\times n}$, such that $D = Q^T AQ$ (equivalently: $A = QDQ^T$).

Theorem 2.5. *A matrix in* $\mathbb{R}^{n\times n}$ *is symmetric if and only if it is orthogonally diagonalizable.*

Proof. Let us first show that orthogonally diagonalizable matrices are symmeteric. Fix any orthogonally diagonalizable matrix $A \in \mathbb{R}^{n\times n}$. Let D be a diagonal and Q an orthogonal matrix, both in $\mathbb{R}^{n\times n}$, such that $D = Q^T AQ$. Then $A = QDQ^T$, and we see that

$$A^T \quad = \quad (QDQ^T)^T \quad = \quad (Q^T)^T D^T Q^T \quad \overset{(*)}{=} \quad QDQ^T \quad = \quad A,$$

where in (*), we used the fact that $D^T = D$, since D is diagonal. Thus, A is symmetric.

It remains to prove the reverse implication: symmetric matrices in $\mathbb{R}^{n\times n}$ are orthogonally diagonalizable. We proceed by induction on n.

For $n = 1$, the result is immediate: indeed, if $A \in \mathbb{R}^{1\times 1}$, then A is diagonal, and we can take $D := A$ and $Q := I_1$.

Now, fix a positive integer n, and assume inductively that the claim holds for symmetric matrices in $\mathbb{R}^{n\times n}$. Fix any symmetric matrix $A \in \mathbb{R}^{(n+1)\times(n+1)}$. By Corollary 2.4, A has $n + 1$ real eigenvalues (with algebraic multiplicities taken into account). Let $\lambda_0 \in \mathbb{R}$ be an eigenvalue of A, and let $\mathbf{x}_0 \in \mathbb{R}^n$ be an associated eigenvector of A. After possibly rescaling (i.e. replacing $\mathbf{x}_0$ by $\frac{\mathbf{x}_0}{||\mathbf{x}_0||}$), we may assume that $||\mathbf{x}_0|| = 1$. Now, using Corollary 2.5(d) of Lecture Notes 12, we let $\{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n\}$ be an orthonormal basis of $\mathbb{R}^{n+1}$.[3] Set $S := \begin{bmatrix} \mathbf{x}_0 & \mathbf{x}_1 & \dots & \mathbf{x}_n \end{bmatrix}$; then S is an orthogonal matrix.

Since $\mathbf{x}_0$ is an eigenvector of A associated with the eigenvalue λ_0, we know $A\mathbf{x}_0 = \lambda_0\mathbf{x}_0$, and consequently, $(\lambda_0 I_{n+1} - A)\mathbf{x}_0 = \mathbf{0}$. Since $\mathbf{x}_0$ is the first column of S, it follows that the first column of $(\lambda_0 I_{n+1} - A)S$ is $\mathbf{0}$, and consequently, the first column of $S^T(\lambda_0 I_{n+1} - A)S$ is also $\mathbf{0}$. Moreover,

[3] Indeed, $\{\mathbf{x}_0\}$ is an orthonormal basis of the subspace $U := \text{Span}(\mathbf{x}_0)$ of $\mathbb{R}^{n+1}$, and so by Corollary 2.5(d) of Lecture Notes 12, $\{\mathbf{x}_0\}$ can be extended to an orthonormal basis of $\mathbb{R}^{n+1}$.

$S^T(\lambda_0 I_{n+1} - A)S$ is a symmetric matrix,[4] and consequently, there exists a symmetric matrix $A_0 \in \mathbb{R}^{n\times n}$ such that

$$S^T(\lambda_0 I_{n+1} - A)S = \left[\begin{array}{c|c} 0 & \mathbf{0}^T \\ \hline \mathbf{0} & A_0 \end{array}\right].$$

By the induction hypothesis, there exists a diagonal matrix D_0 and an orthogonal matrix Q_0, both in $\mathbb{R}^{n\times n}$, such that $D_0 = Q_0^T A_0 Q_0$, and consequently, $A_0 = Q_0 D_0 Q_0^T$. Now, set

- $\widetilde{D} := \left[\begin{array}{c|c} 0 & \mathbf{0}^T \\ \hline \mathbf{0} & D_0 \end{array}\right]_{(n+1)\times(n+1)}$;
- $R := \left[\begin{array}{c|c} 1 & \mathbf{0}^T \\ \hline \mathbf{0} & Q_0 \end{array}\right]_{(n+1)\times(n+1)}$;
- $D := \lambda_0 I_{n+1} - \widetilde{D}$;
- $Q := SR$.

Since D_0 is diagonal, so are $\widetilde{D}$ and D. Since Q_0 is orthogonal, so is R.[5] Further, since R and S are orthogonal, Proposition 2.2 of Lecture Notes 14 guarantees that $Q := SR$ is also orthogonal. Our goal is to show that $D = Q^T A Q$.

First of all, we have that

$$\begin{aligned} S^T(\lambda_0 I_{n+1} - A)S &= \left[\begin{array}{c|c} 0 & \mathbf{0}^T \\ \hline \mathbf{0} & A_0 \end{array}\right] \\ &= \left[\begin{array}{c|c} 0 & \mathbf{0}^T \\ \hline \mathbf{0} & Q_0 D_0 Q_0^T \end{array}\right] \\ &= \left[\begin{array}{c|c} 1 & \mathbf{0}^T \\ \hline \mathbf{0} & Q_0 \end{array}\right] \left[\begin{array}{c|c} 0 & \mathbf{0}^T \\ \hline \mathbf{0} & D_0 \end{array}\right] \left[\begin{array}{c|c} 1 & \mathbf{0}^T \\ \hline \mathbf{0} & Q_0 \end{array}\right]^T \\ &= R\widetilde{D}R^T, \end{aligned}$$

---

[4]Indeed, $\left(S^T(\lambda_0 I_{n+1} - A)S\right)^T = S^T(\lambda_0 I_{n+1} - A)^T(S^T)^T = S^T(\lambda_0 I_{n+1}^T - A^T)S \overset{(*)}{=} S^T(\lambda_0 I_{n+1} - A)S$, where in (*), we used the fact that $A^T = A$ (since A is symmetric).

[5]Indeed,

$$\begin{aligned} R^T R &= \left[\begin{array}{c|c} 1 & \mathbf{0}^T \\ \hline \mathbf{0} & Q_0 \end{array}\right]^T \left[\begin{array}{c|c} 1 & \mathbf{0}^T \\ \hline \mathbf{0} & Q_0 \end{array}\right] \\ &= \left[\begin{array}{c|c} 1 & \mathbf{0}^T \\ \hline \mathbf{0} & Q_0^T \end{array}\right] \left[\begin{array}{c|c} 1 & \mathbf{0}^T \\ \hline \mathbf{0} & Q_0 \end{array}\right] \\ &= \left[\begin{array}{c|c} 1 & \mathbf{0}^T \\ \hline \mathbf{0} & Q^T Q_0 \end{array}\right] \\ &\overset{(*)}{=} \left[\begin{array}{c|c} 1 & \mathbf{0}^T \\ \hline \mathbf{0} & I_n \end{array}\right] \\ &= I_{n+1}, \end{aligned}$$

where (*) follows from the fact that Q is orthogonal.

and consequently,

$$\begin{aligned}
\widetilde{D} &= R^T S^T (\lambda_0 I_{n+1} - A) SR \\
&= (SR)^T (\lambda_0 I_{n+1} - A) SR \\
&= Q^T (\lambda_0 I_{n+1} - A) Q \\
&= Q^T (\lambda_0 I_{n+1}) Q - Q^T A Q \\
&= \lambda_0 Q^T Q - Q^T A Q \\
&= \lambda_0 I_{n+1} - Q^T A Q.
\end{aligned}$$

It follows that

$$Q^T A Q \;=\; \lambda_0 I_{n+1} - \widetilde{D} \;=\; D,$$

and we are done. □

The proof of Theorem 2.5 gives us a recipe of sorts for orthogonally diagonalizing a symmetric matrix in $\mathbb{R}^{n\times n}$, but this recipe is rather impractical! There is, however, an easier way. First, we need a corollary of Theorem 2.5 (see Corollary 2.6 below). For $U, W \subseteq \mathbb{R}^n$, we write $U \perp W$ if $\mathbf{u} \perp \mathbf{w}$ for all $\mathbf{u} \in U$ and $\mathbf{w} \in W$.

Corollary 2.6. *Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix, and suppose that λ_1 and λ_2 are distinct eigenvalues of A. Then $E_{\lambda_1} \perp E_{\lambda_2}$.*

Proof (outline). Using Theorem 2.5, we fix a diagonal matrix D and an orthogonal matrix Q, both in $\mathbb{R}^{n\times n}$, such that $D = Q^T AQ$. Then the main diagonal of D is formed by the eigenvalues of A. Now, suppose the eigenvalue λ_1 appears (precisely) in entries $i_1, \dots, i_{k_1}$ of the main diagonal of D; then columns number $i_1, \dots, i_{k_1}$ of Q form a basis $\mathcal{B}_1$ of E_{λ_1}.[6] Similarly, suppose that the eigenvalue λ_2 appears (precisely) in entries $j_1, \dots, j_{k_2}$ of the main diagonal of D; then columns number $j_1, \dots, j_{k_2}$ of Q form a basis $\mathcal{B}_2$ of E_{λ_2}. But since Q is orthogonal, we know that its columns are orthonormal. So, vectors in $\mathcal{B}_1$ are orthogonal to vectors in $\mathcal{B}_2$, and it readily follows that $E_{\lambda_1} \perp E_{\lambda_2}$. □

Let us now explain how a symmetric matrix $A \in \mathbb{R}^{n\times n}$ can be orthogonally diagonalized. First, we compute the characteristic polynomial of A, factor it, and find all the eigenvalues of A. Then, for each eigenvalue λ of A, we compute a basis $\mathcal{B}_\lambda$ for its eigenspace. Then, we apply the Gram-Schmidt orthogonalization process to each $\mathcal{B}_\lambda$ in order to obtain an orthonormal basis $\mathcal{C}_\lambda$ of E_λ. We now form the diagonal matrix D by listing all the eigenvalues of

[6] This essentially follows from the proof of Theorem 4.2 of Lecture Notes 19.

A on the main diagonal D (respecting the algebraic/geometric multiplicities), and we form Q by placing the vectors from the corresponding orthonormal bases in the corresponding columns. Since all the $\mathcal{C}_\lambda$'s are orthonormal sets, and since they are moreover orthogonal to each other (by Corollary 2.6), we see that the columns of Q form an orthonormal set; since Q is an $n \times n$ matrix, its columns in fact form an orthonormal basis of $\mathbb{R}^n$, and it follows that Q is orthogonal.

Example 2.7. *Orthogonally diagonalize the following symmetric matrix in $\mathbb{R}^{3\times 3}$:*

$$A \quad = \quad \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}.$$

Proof. First, we compute the characteristic polynomial of A:

$$\begin{aligned} p_A(\lambda) &= \det(\lambda I_3 - A) \\ &= \begin{vmatrix} \lambda - 3 & 2 & -4 \\ 2 & \lambda - 6 & -2 \\ -4 & -2 & \lambda - 3 \end{vmatrix} \\ &= \lambda^3 - 12\lambda^2 + 21\lambda + 98 \\ &= (\lambda + 2)(\lambda - 7)^2. \end{aligned}$$

Thus, A has two eigenvectors: $\lambda_1 = -2$ (of algebraic multiplicity 1) and $\lambda_2 = 7$ (of algebraic multiplicity 2). We now compute a basis $\mathcal{B}_1 = \{\begin{bmatrix} -2 & -1 & 2 \end{bmatrix}^T\}$ of E_{λ_1} and a basis $\mathcal{B}_2 = \{\begin{bmatrix} -1 & 2 & 0 \end{bmatrix}^T, \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}^T\}$ of E_{λ_2}. Next, we apply the Gram-Schmidt orthogonalization process to $\mathcal{B}_1$ and $\mathcal{B}_2$. This yields an orthonormal basis $\mathcal{C}_1 = \{\begin{bmatrix} -\frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix}^T\}$ of E_{λ_1}, and an orthonormal basis $\mathcal{C}_2 = \{\begin{bmatrix} -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} & 0 \end{bmatrix}^T, \begin{bmatrix} \frac{4}{3\sqrt{5}} & \frac{2}{3\sqrt{5}} & \frac{5}{3\sqrt{5}} \end{bmatrix}^T\}$ of E_{λ_2}. We now set

$$D \quad := \quad \begin{bmatrix} -2 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 7 \end{bmatrix} \qquad \text{and} \qquad Q \quad := \quad \begin{bmatrix} -2/3 & -1/\sqrt{5} & 4/(3\sqrt{5}) \\ -1/3 & 2/\sqrt{5} & 2/(3\sqrt{5}) \\ 2/3 & 0 & 5/(3\sqrt{5}) \end{bmatrix}.$$

Now D is diagonal, Q is orthogonal, and $D = Q^T A Q$. □

We now summarize the results of this section in what is sometimes called the "spectral theorem for symmetric matrices."

The spectral theorem for symmetric matrices. *Every symmetric matrix $A \in \mathbb{R}^{n\times n}$ satisfies all the following:*

(a) A is orthogonally diagonalizable;

(b) the eigenspaces of A are pairwise orthogonal;

(c) A has n pairwise orthogonal eigenvectors.

Proof. (a) follows from Theorem 2.5, and (b) follows from Corollary 2.6. It remains to prove (c). By (a), there exists a diagonal matrix D and an orthogonal matrix Q, both in $\mathbb{R}^{n \times n}$, such that $D = Q^T AQ$. The columns of Q are all eigenvectors of A (this follows from the proof of Theorem 4.2 of Lecture Notes 19), and since Q is orthogonal, its columns form an orthonormal basis of $\mathbb{R}^n$. Thus, (c) holds. □