Linear Algebra 2

Lecture #23

The Jordan normal form. Symmetric matrices and orthogonal diagonalization

Irena Penev

May 6, 2024

- This lecture has two parts:

- This lecture has two parts:
    1. The Jordan normal form

- This lecture has two parts:
    1. The Jordan normal form
    2. Symmetric matrices and orthogonal diagonalization

1. The Jordan normal form

1. The Jordan normal form
- This topic is by far the most technical part of this course.

1. The Jordan normal form

- This topic is by far the most technical part of this course.
- For this reason, we skip the proofs, and we merely state the main theorems, with some examples.

1. The Jordan normal form

- This topic is by far the most technical part of this course.
- For this reason, we skip the proofs, and we merely state the main theorems, with some examples.
    - We will cover subsections 8.6.1 and 8.6.2 from the Lecture Notes.

1. The Jordan normal form

- This topic is by far the most technical part of this course.
- For this reason, we skip the proofs, and we merely state the main theorems, with some examples.
  - We will cover subsections 8.6.1 and 8.6.2 from the Lecture Notes.
  - An outline of the omitted proofs is given in subsection 8.6.3 (please read this).

1. The Jordan normal form

- This topic is by far the most technical part of this course.
- For this reason, we skip the proofs, and we merely state the main theorems, with some examples.
  - We will cover subsections 8.6.1 and 8.6.2 from the Lecture Notes.
  - An outline of the omitted proofs is given in subsection 8.6.3 (please read this).
  - For the intrepid: the full proof of the main results is given in subsections 8.6.4-8.6.6 of the Lecture Notes (long, technical, and strictly optional!).

1. The Jordan normal form

- This topic is by far the most technical part of this course.
- For this reason, we skip the proofs, and we merely state the main theorems, with some examples.
    - We will cover subsections 8.6.1 and 8.6.2 from the Lecture Notes.
    - An outline of the omitted proofs is given in subsection 8.6.3 (please read this).
    - For the intrepid: the full proof of the main results is given in subsections 8.6.4-8.6.6 of the Lecture Notes (long, technical, and strictly optional!).
    - Subsection 8.6.7 of the Lecture Notes contains some more advanced computation, and it is also left as optional reading.

- As we shall see, our main results involving the Jordan normal form work only for algebraically closed fields.

- As we shall see, our main results involving the Jordan normal form work only for algebraically closed fields.
- Reminder:

Definition

An *algebraically closed field* is a field $\mathbb{F}$ that has the property that every non-constant polynomial with coefficients in $\mathbb{F}$ has a root in $\mathbb{F}$.

- As we shall see, our main results involving the Jordan normal form work only for algebraically closed fields.
- Reminder:

Definition

An *algebraically closed field* is a field $\mathbb{F}$ that has the property that every non-constant polynomial with coefficients in $\mathbb{F}$ has a root in $\mathbb{F}$.

- Any polynomial with coefficients in an algebraically closed field can be factored into linear terms.

- As we shall see, our main results involving the Jordan normal form work only for algebraically closed fields.
- Reminder:

Definition

An *algebraically closed field* is a field $\mathbb{F}$ that has the property that every non-constant polynomial with coefficients in $\mathbb{F}$ has a root in $\mathbb{F}$.

- Any polynomial with coefficients in an algebraically closed field can be factored into linear terms.
- The only algebraically closed field that we have seen in this course is the field $\mathbb{C}$ of complex numbers, although other algebraically closed fields exist.
  - Fields $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{Z}_p$ (where p is a prime number) are not algebraically closed.

Definition

Suppose that $\mathbb{F}$ is a field and $A \in \mathbb{F}^{n_1 \times n_1}$ and $B \in \mathbb{F}^{n_2 \times n_2}$ are **square** matrices. Then the *direct sum* of A and B is the $(n_1 + n_2) \times (n_1 + n_2)$ matrix

$$A \oplus B \quad := \quad \left[\begin{array}{c|c} A & O_{n_1 \times n_2} \\ \hline O_{n_2 \times n_1} & B \end{array}\right].$$

More generally, for square matrices $A_1 \in \mathbb{F}^{n_1 \times n_1}, A_2 \in \mathbb{F}^{n_2 \times n_2}, \ldots, A_k \in \mathbb{F}^{n_k \times n_k}$, we define the *direct sum* of $A_1, A_2 \ldots, A_k$ to be the $(n_1 + n_2 + \cdots + n_k) \times (n_1 + n_2 + \cdots + n_k)$ matrix

$$A_1 \oplus A_2 \oplus \cdots \oplus A_k \quad := \quad \left[\begin{array}{c|c|c|c} A_1 & O_{n_1 \times n_2} & \cdots & O_{n_1 \times n_k} \\ \hline O_{n_2 \times n_1} & A_2 & \cdots & O_{n_2 \times n_k} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline O_{n_k \times n_1} & O_{n_k \times n_2} & \cdots & A_k \end{array}\right].$$

- For example:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \oplus \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \oplus \begin{bmatrix} 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 0 & 0 & 0 & 0 \\ 3 & 4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 3 & 0 \\ 0 & 0 & 4 & 5 & 6 & 0 \\ 0 & 0 & 7 & 8 & 9 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Definition

For a field $\mathbb{F}$, a scalar $\lambda_0 \in \mathbb{F}$, and a positive integer t, the *Jordan block* $J_t(\lambda_0)$ is defined to be following $t \times t$ matrix (with entries understood to be in $\mathbb{F}$):

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

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

Definition

A *Jordan matrix* (also called a matrix in *Jordan normal form*) is any matrix that is a direct sum of one or more Jordan blocks.

- Thus, a Jordan matrix is a matrix of the form

$$J_{t_1}(\lambda_1) \oplus J_{t_2}(\lambda_2) \oplus \cdots \oplus J_{t_\ell}(\lambda_\ell) \quad = \quad \left[\begin{array}{c|c|c|c} J_{t_1}(\lambda_1) & O & \ldots & O \\ \hline O & J_{t_2}(\lambda_2) & \ldots & O \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline O & O & \ldots & J_{t_\ell}(\lambda_\ell) \end{array}\right],$$

where $\lambda_1, \ldots, \lambda_\ell$ are scalars in $\mathbb{F}$, $t_1, \ldots, t_\ell$ 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 $J_3(5)$, $J_2(2)$, $J_1(2)$, and $J_3(5)$:

$$J_3(5) \oplus J_2(2) \oplus J_1(2) \oplus J_3(5) \quad = \quad \begin{bmatrix} 5 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 5 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 5 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5 \end{bmatrix}.$$

- **Remark:** Every diagonal matrix is a Jordan matrix.

- **Remark:** Every diagonal matrix is a Jordan matrix.
  - Moreover, note that a Jordan matrix is diagonal iff all its Jordan blocks are of size 1×1.

- **Remark:** Every diagonal matrix is a Jordan matrix.
  - Moreover, note that a Jordan matrix is diagonal iff all its Jordan blocks are of size 1×1.
  - On the other hand, if some Jordan block of a Jordan matrix J is of larger size (i.e. is of size $t \times t$ for some $t \geq 2$), then J will have at least one 1 on the diagonal right above the main diagonal.

- **Remark:** Every diagonal matrix is a Jordan matrix.
  - Moreover, note that a Jordan matrix is diagonal iff all its Jordan blocks are of size 1×1.
  - On the other hand, if some Jordan block of a Jordan matrix J is of larger size (i.e. is of size $t \times t$ for some $t \geq 2$), then J will have at least one 1 on the diagonal right above the main diagonal.

- **Remark:** Not all matrices that have an arbitrary main diagonal, all 0's and 1's on the diagonal right above the main one, and 0's everywhere else, are Jordan matrices.

- **Remark:** Every diagonal matrix is a Jordan matrix.
  - Moreover, note that a Jordan matrix is diagonal iff all its Jordan blocks are of size 1×1.
  - On the other hand, if some Jordan block of a Jordan matrix J is of larger size (i.e. is of size $t \times t$ for some $t \geq 2$), then J will have at least one 1 on the diagonal right above the main diagonal.

- **Remark:** Not all matrices that have an arbitrary main diagonal, all 0's and 1's on the diagonal right above the main one, and 0's everywhere else, are Jordan matrices.
  - For example, the matrix

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

    is **not** a Jordan matrix (because it is not a direct sum of Jordan blocks).

Theorem 8.6.1

Let $\mathbb{F}$ be a field, and let $J_1, J_2 \in \mathbb{F}^{n \times n}$ be Jordan matrices. Then J_1 and J_2 are similar iff they have exactly the same Jordan blocks (counting repetitions, but not counting the order in which the blocks appear in the two matrices).

Theorem 8.6.1

Let $\mathbb{F}$ be a field, and let $J_1, J_2 \in \mathbb{F}^{n \times n}$ be Jordan matrices. Then J_1 and J_2 are similar iff they have exactly the same Jordan blocks (counting repetitions, but not counting the order in which the blocks appear in the two matrices).

- The full proof of Theorem 8.6.1 is given in the Lecture Notes (optional reading).
  - The "$\Longrightarrow$" requires relatively serious work.
  - The "$\Longleftarrow$" part ("if two Jordan matrices have the same Jordan blocks, then they are similar") is easier, and here is the idea.

- By Theorem 4.5.16, similar matrices represent the same linear function, only with respect to (possibly) different bases.

Theorem 4.5.16 (abridged)

Let $\mathbb{F}$ be a field, let $B, C \in \mathbb{F}^{n \times n}$ be matrices, and let V be an n-dimensional vector space over the field $\mathbb{F}$. Then the following are equivalent:

(a) B and C are similar;

(d) there exist bases $\mathcal{B}$ and $\mathcal{C}$ of V and a linear function $f : V \to V$ s.t. $B = {}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}}$ and $C = {}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{C}}$.

- By Theorem 4.5.16, similar matrices represent the same linear function, only with respect to (possibly) different bases.

Theorem 4.5.16 (abridged)

Let $\mathbb{F}$ be a field, let $B, C \in \mathbb{F}^{n \times n}$ be matrices, and let V be an n-dimensional vector space over the field $\mathbb{F}$. Then the following are equivalent:

(a) B and C are similar;

(d) there exist bases $\mathcal{B}$ and $\mathcal{C}$ of V and a linear function $f : V \to V$ s.t. $B = {}_{\mathcal{B}}\left[\ f\ \right]_{\mathcal{B}}$ and $C = {}_{\mathcal{C}}\left[\ f\ \right]_{\mathcal{C}}$.

- A change in the order of Jordan blocks corresponds to a change in the order of basis vectors.

- By Theorem 4.5.16, similar matrices represent the same linear function, only with respect to (possibly) different bases.

Theorem 4.5.16 (abridged)

Let $\mathbb{F}$ be a field, let $B, C \in \mathbb{F}^{n \times n}$ be matrices, and let V be an n-dimensional vector space over the field $\mathbb{F}$. Then the following are equivalent:

(a) B and C are similar;

(d) there exist bases $\mathcal{B}$ and $\mathcal{C}$ of V and a linear function $f : V \to V$ s.t. $B = {}_{\mathcal{B}}\Big[\ f\ \Big]_{\mathcal{B}}$ and $C = {}_{\mathcal{C}}\Big[\ f\ \Big]_{\mathcal{C}}$.

- A change in the order of Jordan blocks corresponds to a change in the order of basis vectors.
- Let's take a look at an example.

- Suppose that V is a finite-dimensional vector space over a field $\mathbb{F}$, that $f : V \to V$ is a linear function, and that $\mathcal{B} = \{\mathbf{a}_1, \dots, \mathbf{a}_{t_1}, \mathbf{b}_1, \dots, \mathbf{b}_{t_2}, \mathbf{c}_1, \dots, \mathbf{c}_{t_3}, \mathbf{d}_1, \dots, \mathbf{d}_{t_4}\}$ (with $t_1, t_2, t_3, t_4 \geq 1$) is a basis of V s.t.

$$
{}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}} = \left[\begin{array}{c:c:c:c} J_{t_1}(\lambda_1) & O & O & O \\ \hdashline O & J_{t_2}(\lambda_2) & O & O \\ \hdashline O & O & J_{t_3}(\lambda_3) & O \\ \hdashline O & O & O & J_{t_4}(\lambda_4) \end{array}\right].
$$

- Suppose that V is a finite-dimensional vector space over a field $\mathbb{F}$, that $f : V \to V$ is a linear function, and that $\mathcal{B} = \{\mathbf{a}_1, \dots, \mathbf{a}_{t_1}, \mathbf{b}_1, \dots, \mathbf{b}_{t_2}, \mathbf{c}_1, \dots, \mathbf{c}_{t_3}, \mathbf{d}_1, \dots, \mathbf{d}_{t_4}\}$ (with $t_1, t_2, t_3, t_4 \geq 1$) is a basis of V s.t.

$$
{}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}} \;=\; \left[\begin{array}{c:c:c:c} J_{t_1}(\lambda_1) & O & O & O \\ \hdashline O & J_{t_2}(\lambda_2) & O & O \\ \hdashline O & O & J_{t_3}(\lambda_3) & O \\ \hdashline O & O & O & J_{t_4}(\lambda_4) \end{array}\right].
$$

- Then for the basis $\mathcal{C} = \{\mathbf{b}_1, \dots, \mathbf{b}_{t_2}, \mathbf{d}_1, \dots, \mathbf{d}_{t_4}, \mathbf{a}_1, \dots, \mathbf{a}_{t_1}, \mathbf{c}_1, \dots, \mathbf{c}_{t_3}\}$ of V, we have the following:

$$
{}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{C}} \;=\; \left[\begin{array}{c:c:c:c} J_{t_2}(\lambda_2) & O & O & O \\ \hdashline O & J_{t_4}(\lambda_4) & O & O \\ \hdashline O & O & J_{t_1}(\lambda_1) & O \\ \hdashline O & O & O & J_{t_3}(\lambda_3) \end{array}\right].
$$

- Suppose that V is a finite-dimensional vector space over a field $\mathbb{F}$, that $f : V \to V$ is a linear function, and that $\mathcal{B} = \{\mathbf{a}_1, \dots, \mathbf{a}_{t_1}, \mathbf{b}_1, \dots, \mathbf{b}_{t_2}, \mathbf{c}_1, \dots, \mathbf{c}_{t_3}, \mathbf{d}_1, \dots, \mathbf{d}_{t_4}\}$ (with $t_1, t_2, t_3, t_4 \geq 1$) is a basis of V s.t.

$$
{}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}} = \left[\begin{array}{c:c:c:c} J_{t_1}(\lambda_1) & O & O & O \\ \hdashline O & J_{t_2}(\lambda_2) & O & O \\ \hdashline O & O & J_{t_3}(\lambda_3) & O \\ \hdashline O & O & O & J_{t_4}(\lambda_4) \end{array}\right].
$$

- Then for the basis $\mathcal{C} = \{\mathbf{b}_1, \dots, \mathbf{b}_{t_2}, \mathbf{d}_1, \dots, \mathbf{d}_{t_4}, \mathbf{a}_1, \dots, \mathbf{a}_{t_1}, \mathbf{c}_1, \dots, \mathbf{c}_{t_3}\}$ of V, we have the following:

$$
{}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{C}} = \left[\begin{array}{c:c:c:c} J_{t_2}(\lambda_2) & O & O & O \\ \hdashline O & J_{t_4}(\lambda_4) & O & O \\ \hdashline O & O & J_{t_1}(\lambda_1) & O \\ \hdashline O & O & O & J_{t_3}(\lambda_3) \end{array}\right].
$$

- By Theorem 4.5.16, matrices ${}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}}$ and ${}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{C}}$ are similar, and so the two Jordan matrices above are similar.

- We have two main theorems concerning the Jordan normal form.
  - The first (Theorem 8.6.2) concerns matrices.
  - The second (Theorem 8.6.4) concerns linear functions.

- We have two main theorems concerning the Jordan normal form.
  - The first (Theorem 8.6.2) concerns matrices.
  - The second (Theorem 8.6.4) concerns linear functions.
- The two theorems are equivalent in the sense that either one readily implies the other.

- We have two main theorems concerning the Jordan normal form.
  - The first (Theorem 8.6.2) concerns matrices.
  - The second (Theorem 8.6.4) concerns linear functions.
- The two theorems are equivalent in the sense that either one readily implies the other.

Theorem 8.6.2

Assume that $\mathbb{F}$ is an **algebraically closed field**, and let $A \in \mathbb{F}^{n \times n}$ be a square matrix. Then A is similar to a matrix J in Jordan normal form. Moreover, this matrix J is unique up to a reordering of the Jordan blocks.

- **Terminology/Remark:** Suppose that $A \in \mathbb{F}^{n \times n}$ is a matrix, where $\mathbb{F}$ is some algebraically closed field.

- **Terminology/Remark:** Suppose that $A \in \mathbb{F}^{n \times n}$ is a matrix, where $\mathbb{F}$ is some algebraically closed field.
  - Then any Jordan matrix that is similar to A is called a *Jordan normal form* of A.

- **Terminology/Remark:** Suppose that $A \in \mathbb{F}^{n\times n}$ is a matrix, where $\mathbb{F}$ is some algebraically closed field.
    - Then any Jordan matrix that is similar to A is called a *Jordan normal form* of A.
    - As we have seen, reordering the Jordan blocks of a Jordan matrix produces a Jordan matrix that is similar to the original one.

- **Terminology/Remark:** Suppose that $A \in \mathbb{F}^{n \times n}$ is a matrix, where $\mathbb{F}$ is some algebraically closed field.
  - Then any Jordan matrix that is similar to A is called a *Jordan normal form* of A.
  - As we have seen, reordering the Jordan blocks of a Jordan matrix produces a Jordan matrix that is similar to the original one.
  - So, if J is a Jordan normal form of A, then any Jordan matrix obtained from J by merely rearranging the order in which the Jordan blocks appear along the main diagonal is also a Jordan normal form of A.

- **Terminology/Remark:** Suppose that $A \in \mathbb{F}^{n\times n}$ is a matrix, where $\mathbb{F}$ is some algebraically closed field.
  - Then any Jordan matrix that is similar to A is called a *Jordan normal form* of A.
  - As we have seen, reordering the Jordan blocks of a Jordan matrix produces a Jordan matrix that is similar to the original one.
  - So, if J is a Jordan normal form of A, then any Jordan matrix obtained from J by merely rearranging the order in which the Jordan blocks appear along the main diagonal is also a Jordan normal form of A.
  - However, by the uniqueness part of Theorem 8.6.2, this exhausts the possibilities for different Jordan normal forms of A: any two Jordan normal forms of A have exactly the same Jordan blocks (with repetitions taken into account).

Corollary 8.6.3

Let $\mathbb{F}$ be an **algebraically closed field**, and let $A, B \in \mathbb{F}^{n\times n}$. Then A and B are similar iff they have the same Jordan normal form. More precisely, the following are equivalent:

(a) A and B are similar;

(b) there exists a Jordan matrix $J \in \mathbb{F}^{n\times n}$ s.t. both A and B are similar to J;

(c) there exist Jordan matrices $J_A, J_B \in \mathbb{F}^{n\times n}$ s.t. A is similar to J_A, B is similar to J_B, and the Jordan matrices J_A and J_B can be obtained from each other by possibly rearranging the order of the Jordan blocks.

- Proof: This follows more or less immediately from Theorems 8.6.1 and 8.6.2 (details: Lecture Notes).

Corollary 8.6.3

Let $\mathbb{F}$ be an **algebraically closed field**, and let $A, B \in \mathbb{F}^{n \times n}$. Then A and B are similar iff they have the same Jordan normal form. More precisely, the following are equivalent:

(a) A and B are similar;

(b) there exists a Jordan matrix $J \in \mathbb{F}^{n \times n}$ s.t. both A and B are similar to J;

(c) there exist Jordan matrices $J_A, J_B \in \mathbb{F}^{n \times n}$ s.t. A is similar to J_A, B is similar to J_B, and the Jordan matrices J_A and J_B can be obtained from each other by possibly rearranging the order of the Jordan blocks.

- As we know, the field $\mathbb{C}$ is algebraically closed, and so Corollary 8.6.3 applies to matrices in $\mathbb{C}^{n \times n}$.

Corollary 8.6.3

Let $\mathbb{F}$ be an **algebraically closed field**, and let $A, B \in \mathbb{F}^{n \times n}$. Then A and B are similar iff they have the same Jordan normal form. More precisely, the following are equivalent:

(a) A and B are similar;

(b) there exists a Jordan matrix $J \in \mathbb{F}^{n \times n}$ s.t. both A and B are similar to J;

(c) there exist Jordan matrices $J_A, J_B \in \mathbb{F}^{n \times n}$ s.t. A is similar to J_A, B is similar to J_B, and the Jordan matrices J_A and J_B can be obtained from each other by possibly rearranging the order of the Jordan blocks.

- As we know, the field $\mathbb{C}$ is algebraically closed, and so Corollary 8.6.3 applies to matrices in $\mathbb{C}^{n \times n}$.
- On the other hand, $\mathbb{R}$ is **not** algebraically closed, and so we cannot apply Corollary 8.6.4 to matrices in $\mathbb{R}^{n \times n}$, or at least not directly.
  - However, there is a way around this (later!).

Theorem 8.6.4

Let V be a non-trivial, finite-dimensional vector space over an **algebraically closed field** $\mathbb{F}$, and let $f : V \to V$ be a linear function. Then there exists a basis $\mathcal{B}$ s.t. the matrix ${}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}}$ is in Jordan normal form. Moreover, this matrix is unique in the following sense: if $\mathcal{B}_1$ and $\mathcal{B}_2$ are bases of V s.t. both ${}_{\mathcal{B}_1}\big[\ f\ \big]_{\mathcal{B}_1}$ and ${}_{\mathcal{B}_2}\big[\ f\ \big]_{\mathcal{B}_2}$ are in Jordan normal form, then these two matrices are the same up to a reordering of the Jordan blocks.

- **Remarks:**
  1. Theorems 8.6.2 and 8.6.4 only hold for **algebraically closed fields**. The only algebraically closed field that we have seen is $\mathbb{C}$, but others do exist.
  2. Theorem 4.5.16 essentially states that two $n \times n$ matrices are similar iff they represent the same linear function from an n-dimensional vector space to itself, only possibly with respect to different bases.
     - It is then easy to show that Theorems 8.6.2 and Theorems 8.6.4 are equivalent in the sense that either one of them (combined with Theorem 4.5.16) readily implies the other. The details are left as an exercise.

- **Remarks:**
  3. As we saw in the previous lecture, not all square matrices are diagonalizable, i.e. there are square matrices that are not similar to any diagonal matrix.
     - However, as long as we are working over an **algebraically closed field**, Theorem 8.6.2 guarantees that any square matrix is similar to a matrix that is "almost diagonal," namely to its Jordan normal form.
     - However, in the special case when a square matrix A is diagonalizable, the Jordan normal form of A is any diagonal matrix D that is similar to A.

- **Remarks:**
  4. Since every Jordan matrix is upper triangular, its eigenvalues, together with their algebraic multiplicities, can easily be read off from 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

$$J_3(5) \oplus J_2(2) \oplus J_1(2) \oplus J_3(5) = \begin{bmatrix} 5 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 5 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 5 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5 \end{bmatrix}$$

  are 5 (with algebraic multiplicity 6) and 2 (with algebraic multiplicity 3).

- **Remarks:**
  5. Perhaps more interestingly, the geometric multiplicity of each eigenvalue of a Jordan matrix J can also be read off quite easily: the geometric multiplicity of each eigenvalue λ is precisely the number of Jordan blocks of the form $J_t(\lambda)$ that appear along the main diagonal of J. For instance, for the Jordan matrix

$$J_3(5) \oplus J_2(2) \oplus J_1(2) \oplus J_3(5) = \begin{bmatrix} 5 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 5 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 5 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5 \end{bmatrix}$$

  the geometric multiplicity of the eigenvalue 5 is 2, and the geometric multiplicity of the eigenvalue 2 is also 2.

- **Remarks:**
  6. By Theorem 8.2.9, similar matrices have the same eigenvalues, with the same corresponding algebraic multiplicities, and the same corresponding geometric multiplicities.

- **Remarks:**
  6. By Theorem 8.2.9, similar matrices have the same eigenvalues, with the same corresponding algebraic multiplicities, and the same corresponding 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.

- **Remarks:**

  6. By Theorem 8.2.9, similar matrices have the same eigenvalues, with the same corresponding algebraic multiplicities, and the same corresponding 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.
     - However, two square matrices of the same size, and with exactly the same eigenvalues, with the same corresponding algebraic and geometric multiplicities, need not be similar.

- **Remarks:**
  6. By Theorem 8.2.9, similar matrices have the same eigenvalues, with the same corresponding algebraic multiplicities, and the same corresponding 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.
     - However, two square matrices of the same size, and with exactly the same eigenvalues, with the same corresponding algebraic and geometric multiplicities, need not be similar.
     - Indeed, it is easy to construct two Jordan matrices that have different Jordan blocks, but have the same eigenvalues with the same corresponding algebraic and geometric multiplicities. By Theorem 6.8.1, such matrices are **not** similar.

- **Remarks:**
  6. By Theorem 8.2.9, similar matrices have the same eigenvalues, with the same corresponding algebraic multiplicities, and the same corresponding 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.
     - However, two square matrices of the same size, and with exactly the same eigenvalues, with the same corresponding algebraic and geometric multiplicities, need not be similar.
     - Indeed, it is easy to construct two Jordan matrices that have different Jordan blocks, but have the same eigenvalues with the same corresponding algebraic and geometric multiplicities. By Theorem 6.8.1, such matrices are **not** similar.
     - For a concrete example, consider the Jordan matrices $J_2(\lambda) \oplus J_2(\lambda)$ and $J_3(\lambda) \oplus J_1(\lambda)$, where λ is an arbitrary scalar from the field in question; these two matrices have only one eigenvalue, namely λ, with algebraic multiplicity 4 and geometric multiplicity 2, but they have different Jordan blocks and are therefore not similar.

Example 8.6.5

Let $A_1, A_2, A_3 \in \mathbb{C}^{7\times 7}$ be matrices whose Jordan normal forms are J_1, J_2, J_3, respectively, as follows:

- $J_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$;
- $J_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$;
- $J_3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.

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

Solution. We first identify the Jordan blocks of the the three Jordan matrices. In each matrix, we use colors to indicate the Jordan blocks.

Solution. We first identify the Jordan blocks of the the three Jordan matrices. In each matrix, we use colors to indicate the Jordan blocks.

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

Solution. We first identify the Jordan blocks of the the three Jordan matrices. In each matrix, we use colors to indicate the Jordan blocks.

- $J_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} = J_1(0) \oplus J_3(1) \oplus J_1(1) \oplus J_2(0);$

- $J_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} = J_1(1) \oplus J_2(0) \oplus J_1(0) \oplus J_3(1);$

- $J_3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} = J_1(0) \oplus J_2(1) \oplus J_2(1) \oplus J_2(0).$

We see that 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.

Solution (continued). Reminder:

- $J_1 = J_1(0) \oplus J_3(1) \oplus J_1(1) \oplus J_2(0)$;
- $J_2 = J_1(1) \oplus J_2(0) \oplus J_1(0) \oplus J_3(1)$;
- $J_3 = J_1(0) \oplus J_2(1) \oplus J_2(1) \oplus J_2(0)$.

Solution (continued). Reminder:

- $J_1 = J_1(0) \oplus J_3(1) \oplus J_1(1) \oplus J_2(0)$;
- $J_2 = J_1(1) \oplus J_2(0) \oplus J_1(0) \oplus J_3(1)$;
- $J_3 = J_1(0) \oplus J_2(1) \oplus J_2(1) \oplus J_2(0)$.

For each $i \in \{1, 2, 3\}$, we see that the characteristic polynomial of A_i is

$$p_{A_i}(\lambda) \overset{(*)}{=} p_{J_i}(\lambda) \overset{(**)}{=} \lambda^3(\lambda - 1)^4,$$

where (*) follows from the fact that A_i and J_i are similar (we are using Proposition 8.2.9), and (**) from the fact that the Jordan matrix J_i is upper triangular (we are using Proposition 8.2.7).

Solution (continued). Reminder:

- $J_1 = J_1(0) \oplus J_3(1) \oplus J_1(1) \oplus J_2(0)$;
- $J_2 = J_1(1) \oplus J_2(0) \oplus J_1(0) \oplus J_3(1)$;
- $J_3 = J_1(0) \oplus J_2(1) \oplus J_2(1) \oplus J_2(0)$.

For each $i \in \{1, 2, 3\}$, we see that the characteristic polynomial of A_i is

$$p_{A_i}(\lambda) \stackrel{(*)}{=} p_{J_i}(\lambda) \stackrel{(**)}{=} \lambda^3(\lambda - 1)^4,$$

where (*) follows from the fact that A_i and J_i are similar (we are using Proposition 8.2.9), and (**) from the fact that the Jordan matrix J_i is upper triangular (we are using Proposition 8.2.7).

Finally, we see from the matrices J_1, J_2, J_3, that A_1, A_2, A_3 all have spectrum $\{0, 0, 0, 1, 1, 1, 1\}$, and that they all have exactly two eigenvalues:

- the eigenvalue 0 with algebraic multiplicity 3 and geometric multiplicity 2;
- the eigenvalue 1 with algebraic multiplicity 4 and geometric multiplicity 2. □

Theorem 8.6.6

Let $\mathbb{F}$ be an **algebraically closed field**, let $A \in \mathbb{F}^{n\times n}$, and let

$$\{\underbrace{\lambda_1,\dots,\lambda_1}_{m_1},\dots,\underbrace{\lambda_k,\dots,\lambda_k}_{m_k}\}$$

be the spectrum of A, where $\lambda_1,\dots,\lambda_k$ are pairwise distinct eigenvalues of A and $m_1,\dots,m_k$ are positive integers.[a] Then A is similar to a matrix $J \in \mathbb{F}^{n\times n}$ in Jordan normal form that has the following properties:

1. each Jordan block of the Jordan matrix J is of the form $J_t(\lambda_i)$ for some $i \in \{1,\dots,k\}$ and $t \in \{1,\dots,m_i\}$;
2. for each $i \in \{1,\dots,k\}$ and each positive integer r, the Jordan matrix J has exactly $\text{rank}((A-\lambda_i I_n)^{r-1}) - \text{rank}((A-\lambda_i I_n)^r)$ many Jordan blocks $J_t(\lambda_i)$ satisfying $t \geq r$.

Moreover, A is similar to any Jordan matrix in $\mathbb{F}^{n\times n}$ that satisfies conditions (i) and (ii) above.

[a] Since $\mathbb{F}$ is algebraically closed, we know that $m_1 + \dots + m_k = n$.

- Theorem 8.6.6 does indeed allow us to compute the Jordan normal form of a square matrix A with entries in an algebraically closed field $\mathbb{F}$, as long as we are able to factor its characteristic polynomial into linear terms.
  - Any non-constant polynomial with coefficients in an algebraically closed field $\mathbb{F}$ can be factored into linear terms (with coefficients in $\mathbb{F}$).
  - However, this is merely an existence result: actually computing the linear factors may be extremely difficult or even impossible.
  - If we get stuck factoring the characteristic polynomial into linear terms, then Theorem 8.6.6 is of no use to us (computationally speaking).

- Theorem 8.6.6 does indeed allow us to compute the Jordan normal form of a square matrix A with entries in an algebraically closed field $\mathbb{F}$, as long as we are able to factor its characteristic polynomial into linear terms.
  - Any non-constant polynomial with coefficients in an algebraically closed field $\mathbb{F}$ can be factored into linear terms (with coefficients in $\mathbb{F}$).
  - However, this is merely an existence result: actually computing the linear factors may be extremely difficult or even impossible.
  - If we get stuck factoring the characteristic polynomial into linear terms, then Theorem 8.6.6 is of no use to us (computationally speaking).
- Indeed, condition (i) of Theorem 8.6.6 tells us what sorts of Jordan blocks the Jordan normal form of A may possibly have.

- Theorem 8.6.6 does indeed allow us to compute the Jordan normal form of a square matrix A with entries in an algebraically closed field $\mathbb{F}$, as long as we are able to factor its characteristic polynomial into linear terms.
  - Any non-constant polynomial with coefficients in an algebraically closed field $\mathbb{F}$ can be factored into linear terms (with coefficients in $\mathbb{F}$).
  - However, this is merely an existence result: actually computing the linear factors may be extremely difficult or even impossible.
  - If we get stuck factoring the characteristic polynomial into linear terms, then Theorem 8.6.6 is of no use to us (computationally speaking).
- Indeed, condition (i) of Theorem 8.6.6 tells us what sorts of Jordan blocks the Jordan normal form of A may possibly have.
- Condition (ii) gives us an easy way to compute the number of Jordan blocks of each type.

- Indeed, using the set-up and notation from Theorem 8.6.6, we consider an eigenvalue λ_i of A, and we fix a positive integer r. Then the number of Jordan blocks $J_r(\lambda_i)$ in the Jordan normal form of A is exactly

$$\underbrace{\left(\operatorname{rank}\left((A-\lambda_i I_n)^{r-1}\right)-\operatorname{rank}\left((A-\lambda_i I_n)^{r}\right)\right)}_{\substack{= \quad \text{number of Jordan blocks} \\ J_t(\lambda_i) \text{ satisfying } t \geq r}}-\underbrace{\left(\operatorname{rank}\left((A-\lambda_i I_n)^{r}\right)-\operatorname{rank}\left((A-\lambda_i I_n)^{r+1}\right)\right)}_{\substack{= \quad \text{number of Jordan blocks} \\ J_t(\lambda_i) \text{ satisfying } t \geq r+1}}.$$

- Indeed, using the set-up and notation from Theorem 8.6.6, we consider an eigenvalue λ_i of A, and we fix a positive integer r. Then the number of Jordan blocks $J_r(\lambda_i)$ in the Jordan normal form of A is exactly

$$\underbrace{\Big(\text{rank}\Big((A-\lambda_i I_n)^{r-1}\Big) - \text{rank}\Big((A-\lambda_i I_n)^{r}\Big)\Big)}_{\substack{= \quad \text{number of Jordan blocks} \\ J_t(\lambda_i) \text{ satisfying } t \geq r}} - \underbrace{\Big(\text{rank}\Big((A-\lambda_i I_n)^{r}\Big) - \text{rank}\Big((A-\lambda_i I_n)^{r+1}\Big)\Big)}_{\substack{= \quad \text{number of Jordan blocks} \\ J_t(\lambda_i) \text{ satisfying } t \geq r+1}}.$$

- So, we can compute both the possible types of Jordan blocks that the Jordan normal form of A may have, and the exact number of blocks of each possible type.

- Indeed, using the set-up and notation from Theorem 8.6.6, we consider an eigenvalue λ_i of A, and we fix a positive integer r. Then the number of Jordan blocks $J_r(\lambda_i)$ in the Jordan normal form of A is exactly

$$\underbrace{\left(\operatorname{rank}\Big((A-\lambda_i I_n)^{r-1}\Big) - \operatorname{rank}\Big((A-\lambda_i I_n)^{r}\Big)\right)}_{\substack{= \quad \text{number of Jordan blocks} \\ J_t(\lambda_i) \text{ satisfying } t \geq r}} - \underbrace{\left(\operatorname{rank}\Big((A-\lambda_i I_n)^{r}\Big) - \operatorname{rank}\Big((A-\lambda_i I_n)^{r+1}\Big)\right)}_{\substack{= \quad \text{number of Jordan blocks} \\ J_t(\lambda_i) \text{ satisfying } t \geq r+1}}.$$

- So, we can compute both the possible types of Jordan blocks that the Jordan normal form of A may have, and the exact number of blocks of each possible type.
- We in fact get an exact formula

$$\operatorname{rank}\big((A-\lambda_i I_n)^{r-1}\big) + \operatorname{rank}\big((A-\lambda_i I_n)^{r+1} - 2\operatorname{rank}\big((A-\lambda_i I_n)^{r}\big)$$

for the number of Jordan blocks $J_r(\lambda_i)$ in the Jordan normal form of A.
  - However, it is arguably easier to memorize the formula for the number of Jordan blocks of the form $J_t(\lambda_i)$ satisfying $t \geq r$.

Example 8.6.8

Consider the following matrix in $\mathbb{C}^{10\times 10}$:

$$A \quad := \quad \begin{bmatrix} 3 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 \\ -3 & 1 & 5 & 2 & -2 & -4 & -7 & 4 & -1 & 3 \\ 0 & 1 & 3 & 0 & 0 & 0 & 0 & -1 & 0 & 1 \\ -2 & -1 & 3 & 4 & -1 & -2 & -3 & 2 & -1 & 2 \\ -1 & 0 & 2 & 1 & 2 & -2 & -1 & 1 & 0 & 1 \\ -1 & 0 & 1 & 0 & 0 & 2 & -1 & 0 & 0 & 1 \\ 1 & 1 & -2 & -1 & 1 & 2 & 7 & -2 & 1 & -2 \\ -1 & 0 & 1 & 0 & 0 & -1 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 3 & -1 \\ 1 & 1 & -2 & -1 & 1 & 2 & 5 & -2 & 1 & 0 \end{bmatrix}.$$

Using Theorem 8.6.6, compute the Jordan normal form of A.

Solution.

Example 8.6.8

Consider the following matrix in $\mathbb{C}^{10\times 10}$:

$$A \quad := \quad \begin{bmatrix} 3 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 \\ -3 & 1 & 5 & 2 & -2 & -4 & -7 & 4 & -1 & 3 \\ 0 & 1 & 3 & 0 & 0 & 0 & 0 & -1 & 0 & 1 \\ -2 & -1 & 3 & 4 & -1 & -2 & -3 & 2 & -1 & 2 \\ -1 & 0 & 2 & 1 & 2 & -2 & -1 & 1 & 0 & 1 \\ -1 & 0 & 1 & 0 & 0 & 2 & -1 & 0 & 0 & 1 \\ 1 & 1 & -2 & -1 & 1 & 2 & 7 & -2 & 1 & -2 \\ -1 & 0 & 1 & 0 & 0 & -1 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 3 & -1 \\ 1 & 1 & -2 & -1 & 1 & 2 & 5 & -2 & 1 & 0 \end{bmatrix}.$$

Using Theorem 8.6.6, compute the Jordan normal form of A.

Solution. First of all, we compute the characteristic polynomial of A, and we factor it into linear terms:

$$p_A(\lambda) \quad = \quad \det(\lambda I_{10} - A) \quad = \quad (\lambda - 3)^8(\lambda - 2)^2.$$

So, the eigenvalues of A are $\lambda_1 = 3$ (with alg. mult. 8) and $\lambda_2 = 2$ (with alg. mult. 2).

Example 8.6.8

Consider the following matrix in $\mathbb{C}^{10\times 10}$:

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

Using Theorem 8.6.6, compute the Jordan normal form of A.

Solution. First of all, we compute the characteristic polynomial of A, and we factor it into linear terms:

$$p_A(\lambda) = \det(\lambda I_{10} - A) = (\lambda-3)^8(\lambda-2)^2.$$

So, the eigenvalues of A are $\lambda_1 = 3$ (with alg. mult. 8) and $\lambda_2 = 2$ (with alg. mult. 2). So, all of our Jordan blocks will be of the form $J_t(3)$ and $J_t(2)$ for various positive integers t.

Example 8.6.8

Consider the following matrix in $\mathbb{C}^{10\times 10}$:

$$A \quad := \quad \begin{bmatrix} 3 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 \\ -3 & 1 & 5 & 2 & -2 & -4 & -7 & 4 & -1 & 3 \\ 0 & 1 & 3 & 0 & 0 & 0 & 0 & -1 & 0 & 1 \\ -2 & -1 & 3 & 4 & -1 & -2 & -3 & 2 & -1 & 2 \\ -1 & 0 & 2 & 1 & 2 & -2 & -1 & 1 & 0 & 1 \\ -1 & 0 & 1 & 0 & 0 & 2 & -1 & 0 & 0 & 1 \\ 1 & 1 & -2 & -1 & 1 & 2 & 7 & -2 & 1 & -2 \\ -1 & 0 & 1 & 0 & 0 & -1 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 3 & -1 \\ 1 & 1 & -2 & -1 & 1 & 2 & 5 & -2 & 1 & 0 \end{bmatrix}.$$

Using Theorem 8.6.6, compute the Jordan normal form of A.

Solution. First of all, we compute the characteristic polynomial of A, and we factor it into linear terms:

$$p_A(\lambda) \quad = \quad \det(\lambda I_{10} - A) \quad = \quad (\lambda - 3)^8 (\lambda - 2)^2.$$

So, the eigenvalues of A are $\lambda_1 = 3$ (with alg. mult. 8) and $\lambda_2 = 2$ (with alg. mult. 2). So, all of our Jordan blocks will be of the form $J_t(3)$ and $J_t(2)$ for various positive integers t. We now deal with the two eigenvalues separately, as follows.

Solution (continued). We first deal with the eigenvalue $\lambda_1 = 3$.

Solution (continued). We first deal with the eigenvalue $\lambda_1 = 3$. We compute the matrices $(A - \lambda_1 I_{10})^r$ for $r = 0, 1, 2, 3, \ldots$ along with their ranks.

Solution (continued). We first deal with the eigenvalue $\lambda_1 = 3$. We compute the matrices $(A - \lambda_1 I_{10})^r$ for $r = 0, 1, 2, 3, \ldots$ along with their ranks. We keep computing until we get the same rank twice in a row.

Solution (continued). We first deal with the eigenvalue $\lambda_1 = 3$. We compute the matrices $(A - \lambda_1 I_{10})^r$ for $r = 0, 1, 2, 3, \dots$ along with their ranks. We keep computing until we get the same rank twice in a row. We obtain:

- $\text{rank}\Big((A - \lambda_1 I_{10})^0\Big) = 10$;
- $\text{rank}\Big((A - \lambda_1 I_{10})^1\Big) = 7$;
- $\text{rank}\Big((A - \lambda_1 I_{10})^2\Big) = 4$;
- $\text{rank}\Big((A - \lambda_1 I_{10})^3\Big) = 2$;
- $\text{rank}\Big((A - \lambda_1 I_{10})^4\Big) = 2$.

We have now obtained the same rank twice in a row, and so we can stop.

Solution (continued). We first deal with the eigenvalue $\lambda_1 = 3$. We compute the matrices $(A - \lambda_1 I_{10})^r$ for $r = 0, 1, 2, 3, \dots$ along with their ranks. We keep computing until we get the same rank twice in a row. We obtain:

- $\text{rank}\Big((A - \lambda_1 I_{10})^0\Big) = 10$;
- $\text{rank}\Big((A - \lambda_1 I_{10})^1\Big) = 7$;
- $\text{rank}\Big((A - \lambda_1 I_{10})^2\Big) = 4$;
- $\text{rank}\Big((A - \lambda_1 I_{10})^3\Big) = 2$;
- $\text{rank}\Big((A - \lambda_1 I_{10})^4\Big) = 2$.

We have now obtained the same rank twice in a row, and so we can stop. We compute (next slide):

Solution (continued).

- $\text{rank}\Big((A - \lambda_1 I_{10})^0\Big) - \text{rank}\Big((A - \lambda_1 I_{10})^1\Big) = 3;$
- $\text{rank}\Big((A - \lambda_1 I_{10})^1\Big) - \text{rank}\Big((A - \lambda_1 I_{10})^2\Big) = 3;$
- $\text{rank}\Big((A - \lambda_1 I_{10})^2\Big) - \text{rank}\Big((A - \lambda_1 I_{10})^3\Big) = 2;$
- $\text{rank}\Big((A - \lambda_1 I_{10})^3\Big) - \text{rank}\Big((A - \lambda_1 I_{10})^4\Big) = 0.$

Solution (continued).

- $\mathrm{rank}\Big((A-\lambda_1 I_{10})^0\Big) - \mathrm{rank}\Big((A-\lambda_1 I_{10})^1\Big) = 3$;
- $\mathrm{rank}\Big((A-\lambda_1 I_{10})^1\Big) - \mathrm{rank}\Big((A-\lambda_1 I_{10})^2\Big) = 3$;
- $\mathrm{rank}\Big((A-\lambda_1 I_{10})^2\Big) - \mathrm{rank}\Big((A-\lambda_1 I_{10})^3\Big) = 2$;
- $\mathrm{rank}\Big((A-\lambda_1 I_{10})^3\Big) - \mathrm{rank}\Big((A-\lambda_1 I_{10})^4\Big) = 0$.

By Theorem 8.6.6, the Jordan normal form of A will contain:

- three Jordan blocks $J_t(\lambda_1) = J_t(3)$ with $t \geq 1$;
- three Jordan blocks $J_t(\lambda_1) = J_t(3)$ with $t \geq 2$;
- two Jordan blocks $J_t(\lambda_1) = J_t(3)$ with $t \geq 3$;
- zero Jordan blocks $J_t(\lambda_1) = J_t(3)$ with $t \geq 4$.

Solution (continued). Reminder:

- three Jordan blocks $J_t(\lambda_1) = J_t(3)$ with $t \geq 1$;
- three Jordan blocks $J_t(\lambda_1) = J_t(3)$ with $t \geq 2$;
- two Jordan blocks $J_t(\lambda_1) = J_t(3)$ with $t \geq 3$;
- zero Jordan blocks $J_t(\lambda_1) = J_t(3)$ with $t \geq 4$.

Solution (continued). Reminder:

- three Jordan blocks $J_t(\lambda_1) = J_t(3)$ with $t \geq 1$;
- three Jordan blocks $J_t(\lambda_1) = J_t(3)$ with $t \geq 2$;
- two Jordan blocks $J_t(\lambda_1) = J_t(3)$ with $t \geq 3$;
- zero Jordan blocks $J_t(\lambda_1) = J_t(3)$ with $t \geq 4$.

Keeping in mind that for any positive integer r, the number of Jordan blocks $J_r(\lambda_1) = J_r(3)$ in the Jordan normal form of A is equal to

$$\left(\begin{array}{l} \text{number of Jordan blocks} \\ J_t(\lambda_1) \text{ satisfying } t \geq r \end{array} \right) - \left(\begin{array}{l} \text{number of Jordan blocks} \\ J_t(\lambda_1) \text{ satisfying } t \geq r+1 \end{array} \right),$$

we conclude that the Jordan normal form of A will contain exactly two Jordan blocks $J_3(\lambda_1) = J_3(3)$, and exactly one Jordan block $J_2(\lambda_1) = J_2(3)$. The Jordan normal form of A contains no other Jordan blocks of the form $J_t(\lambda_1) = J_t(3)$.

Solution (continued). Reminder: The Jordan normal form of A will contain exactly two Jordan blocks $J_3(\lambda_1) = J_3(3)$, and exactly one Jordan block $J_2(\lambda_1) = J_2(3)$. The Jordan normal form of A contains no other Jordan blocks of the form $J_t(\lambda_1) = J_t(3)$.

Solution (continued). Reminder: The Jordan normal form of A will contain exactly two Jordan blocks $J_3(\lambda_1) = J_3(3)$, and exactly one Jordan block $J_2(\lambda_1) = J_2(3)$. The Jordan normal form of A contains no other Jordan blocks of the form $J_t(\lambda_1) = J_t(3)$.

A similar computation shows that A contains exactly one Jordan block $J_2(\lambda_2) = J_2(2)$, and it contains no other Jordan blocks of the form $J_t(\lambda_2) = J_t(2)$. (Details: Lecture Notes.)

Solution (continued). Putting everything together, we get that the Jordan normal form of A is the following:

$$\begin{aligned} J &:= J_3(\lambda_1) \oplus J_3(\lambda_1) \oplus J_2(\lambda_1) \oplus J_2(\lambda_2) \\ &= J_3(3) \oplus J_3(3) \oplus J_2(3) \oplus J_2(2) \\ &= \begin{bmatrix} 3 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}. \end{aligned}$$

Solution (continued). Putting everything together, we get that the Jordan normal form of A is the following:

$$
\begin{aligned}
J &:= J_3(\lambda_1) \oplus J_3(\lambda_1) \oplus J_2(\lambda_1) \oplus J_2(\lambda_2) \\
&= J_3(3) \oplus J_3(3) \oplus J_2(3) \oplus J_2(2) \\
&= \begin{bmatrix}
3 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 3 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 3 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 3 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 3 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2
\end{bmatrix}.
\end{aligned}
$$

We remark that we could have written our Jordan blocks in a different order, but in any case, the Jordan blocks would have to be the same as above (counting repetitions). For instance, $J_2(3) \oplus J_3(3) \oplus J_2(2) \oplus J_3(3)$ is also a Jordan normal form of A.

Solution (continued). Reminder:

$$\begin{aligned} J &:= J_3(\lambda_1) \oplus J_3(\lambda_1) \oplus J_2(\lambda_1) \oplus J_2(\lambda_2) \\ &= J_3(3) \oplus J_3(3) \oplus J_2(3) \oplus J_2(2) \\ &= \begin{bmatrix} 3 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}. \end{aligned}$$

Remark: It is acceptable to leave $J_3(3) \oplus J_3(3) \oplus J_2(3) \oplus J_2(2)$ (color coded or not) as a final answer, without exhibiting the actual 10×10 matrix with its 100 entries. It is **not** acceptable to leave $J_3(\lambda_1) \oplus J_3(\lambda_1) \oplus J_2(\lambda_1) \oplus J_2(\lambda_2)$ as a final answer. $\square$

- There is another worked out example in the Lecture Notes (see Example 8.6.9).

- There is another worked out example in the Lecture Notes (see Example 8.6.9).
- **Remark:** Suppose we are given a matrix $A \in \mathbb{F}^{n \times n}$, where $\mathbb{F}$ is an algebraically closed field.

- There is another worked out example in the Lecture Notes (see Example 8.6.9).
- **Remark:** Suppose we are given a matrix $A \in \mathbb{F}^{n \times n}$, where $\mathbb{F}$ is an algebraically closed field.
  - We just saw how we can compute the Jordan normal form of A, that is, how we can find a Jordan matrix $J \in \mathbb{F}^{n \times n}$ that is similar to A.

- There is another worked out example in the Lecture Notes (see Example 8.6.9).
- **Remark:** Suppose we are given a matrix $A \in \mathbb{F}^{n \times n}$, where $\mathbb{F}$ is an algebraically closed field.
  - We just saw how we can compute the Jordan normal form of A, that is, how we can find a Jordan matrix $J \in \mathbb{F}^{n \times n}$ that is similar to A.
  - Could we also compute an invertible matrix $P \in \mathbb{F}^{n \times n}$ for which $J = P^{-1}AP$?

- There is another worked out example in the Lecture Notes (see Example 8.6.9).
- **Remark:** Suppose we are given a matrix $A \in \mathbb{F}^{n \times n}$, where $\mathbb{F}$ is an algebraically closed field.
  - We just saw how we can compute the Jordan normal form of A, that is, how we can find a Jordan matrix $J \in \mathbb{F}^{n \times n}$ that is similar to A.
  - Could we also compute an invertible matrix $P \in \mathbb{F}^{n \times n}$ for which $J = P^{-1}AP$?
  - This is indeed possible, but it is significantly more complicated than just computing a suitable Jordan matrix J.

- There is another worked out example in the Lecture Notes (see Example 8.6.9).
- **Remark:** Suppose we are given a matrix $A \in \mathbb{F}^{n \times n}$, where $\mathbb{F}$ is an algebraically closed field.
  - We just saw how we can compute the Jordan normal form of A, that is, how we can find a Jordan matrix $J \in \mathbb{F}^{n \times n}$ that is similar to A.
  - Could we also compute an invertible matrix $P \in \mathbb{F}^{n \times n}$ for which $J = P^{-1}AP$?
  - This is indeed possible, but it is significantly more complicated than just computing a suitable Jordan matrix J.
  - Unfortunately, any example that illustrates the procedure in full generality (more or less) requires a great deal of long and laborious computation.

- There is another worked out example in the Lecture Notes (see Example 8.6.9).
- **Remark:** Suppose we are given a matrix $A \in \mathbb{F}^{n \times n}$, where $\mathbb{F}$ is an algebraically closed field.
  - We just saw how we can compute the Jordan normal form of A, that is, how we can find a Jordan matrix $J \in \mathbb{F}^{n \times n}$ that is similar to A.
  - Could we also compute an invertible matrix $P \in \mathbb{F}^{n \times n}$ for which $J = P^{-1}AP$?
  - This is indeed possible, but it is significantly more complicated than just computing a suitable Jordan matrix J.
  - Unfortunately, any example that illustrates the procedure in full generality (more or less) requires a great deal of long and laborious computation.
  - For the sufficiently brave, a recipe and a couple of examples are given in subsection 8.6.7 of the Lecture Notes.

- Reminder:

Corollary 8.6.3

Let $\mathbb{F}$ be an **algebraically closed field**, and let $A, B \in \mathbb{F}^{n \times n}$. Then A and B are similar iff they have the same Jordan normal form. More precisely, the following are equivalent:

(a) A and B are similar;

(b) there exists a Jordan matrix $J \in \mathbb{F}^{n \times n}$ s.t. both A and B are similar to J;

(c) there exist Jordan matrices $J_A, J_B \in \mathbb{F}^{n \times n}$ s.t. A is similar to J_A, B is similar to J_B, and the Jordan matrices J_A and J_B can be obtained from each other by possibly rearranging the order of the Jordan blocks.

- Reminder:

Corollary 8.6.3

Let $\mathbb{F}$ be an **algebraically closed field**, and let $A, B \in \mathbb{F}^{n \times n}$. Then A and B are similar iff they have the same Jordan normal form. More precisely, the following are equivalent:

(a) A and B are similar;

(b) there exists a Jordan matrix $J \in \mathbb{F}^{n \times n}$ s.t. both A and B are similar to J;

(c) there exist Jordan matrices $J_A, J_B \in \mathbb{F}^{n \times n}$ s.t. A is similar to J_A, B is similar to J_B, and the Jordan matrices J_A and J_B can be obtained from each other by possibly rearranging the order of the Jordan blocks.

- What if $\mathbb{F}$ is not algebraically closed?

Definition

A field $\mathbb{F}_1$ is a *subfield* of a field $\mathbb{F}_2$ if the following three conditions are satisfied:

- $\mathbb{F}_1 \subseteq \mathbb{F}_2$;
- for all $a, b \in \mathbb{F}_1$, the sum $a + b$ is the same in $\mathbb{F}_1$ and in $\mathbb{F}_2$;
- for all $a, b \in \mathbb{F}_1$, the product ab is the same in $\mathbb{F}_1$ and in $\mathbb{F}_2$.

Definition

A field $\mathbb{F}_1$ is a *subfield* of a field $\mathbb{F}_2$ if the following three conditions are satisfied:

- $\mathbb{F}_1 \subseteq \mathbb{F}_2$;
- for all $a, b \in \mathbb{F}_1$, the sum $a + b$ is the same in $\mathbb{F}_1$ and in $\mathbb{F}_2$;
- for all $a, b \in \mathbb{F}_1$, the product ab is the same in $\mathbb{F}_1$ and in $\mathbb{F}_2$.

- For example, $\mathbb{Q}$ is a subfield of both $\mathbb{R}$ and $\mathbb{C}$, and $\mathbb{R}$ is a subfield of $\mathbb{C}$.

Definition

A field $\mathbb{F}_1$ is a *subfield* of a field $\mathbb{F}_2$ if the following three conditions are satisfied:

- $\mathbb{F}_1 \subseteq \mathbb{F}_2$;
- for all $a, b \in \mathbb{F}_1$, the sum $a + b$ is the same in $\mathbb{F}_1$ and in $\mathbb{F}_2$;
- for all $a, b \in \mathbb{F}_1$, the product ab is the same in $\mathbb{F}_1$ and in $\mathbb{F}_2$.

- For example, $\mathbb{Q}$ is a subfield of both $\mathbb{R}$ and $\mathbb{C}$, and $\mathbb{R}$ is a subfield of $\mathbb{C}$.
- On the other hand, for distinct prime numbers p and q, $\mathbb{Z}_p$ is **not** a subfield of $\mathbb{Z}_q$ (even if $p < q$).

Definition

A field $\mathbb{F}_1$ is a *subfield* of a field $\mathbb{F}_2$ if the following three conditions are satisfied:

- $\mathbb{F}_1 \subseteq \mathbb{F}_2$;
- for all $a, b \in \mathbb{F}_1$, the sum $a + b$ is the same in $\mathbb{F}_1$ and in $\mathbb{F}_2$;
- for all $a, b \in \mathbb{F}_1$, the product ab is the same in $\mathbb{F}_1$ and in $\mathbb{F}_2$.

- For example, $\mathbb{Q}$ is a subfield of both $\mathbb{R}$ and $\mathbb{C}$, and $\mathbb{R}$ is a subfield of $\mathbb{C}$.
- On the other hand, for distinct prime numbers p and q, $\mathbb{Z}_p$ is **not** a subfield of $\mathbb{Z}_q$ (even if $p < q$).
- Moreover, for a prime number p, $\mathbb{Z}_p$ is **not** a subfield of any one of $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$.

Definition

A field $\mathbb{F}_1$ is a *subfield* of a field $\mathbb{F}_2$ if the following three conditions are satisfied:

- $\mathbb{F}_1 \subseteq \mathbb{F}_2$;
- for all $a, b \in \mathbb{F}_1$, the sum $a + b$ is the same in $\mathbb{F}_1$ and in $\mathbb{F}_2$;
- for all $a, b \in \mathbb{F}_1$, the product ab is the same in $\mathbb{F}_1$ and in $\mathbb{F}_2$.

- For example, $\mathbb{Q}$ is a subfield of both $\mathbb{R}$ and $\mathbb{C}$, and $\mathbb{R}$ is a subfield of $\mathbb{C}$.
- On the other hand, for distinct prime numbers p and q, $\mathbb{Z}_p$ is **not** a subfield of $\mathbb{Z}_q$ (even if $p < q$).
- Moreover, for a prime number p, $\mathbb{Z}_p$ is **not** a subfield of any one of $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$.
- It can be shown that any field is a subfield of some algebraically closed field, but the proof of this fact is beyond the scope of this course.
  - However, let us point out that the field $\mathbb{R}$ is a subfield of the algebraically closed field $\mathbb{C}$. ($\mathbb{Q}$ is also a subfield of $\mathbb{C}$.)

Definition

For a field $\mathbb{F}$, we say that $n \times n$ matrices A and B with entries in $\mathbb{F}$ are *similar over* $\mathbb{F}$ if there exists an invertible matrix $P \in \mathbb{F}^{n \times n}$ s.t. $B = P^{-1}AP$.

- This is simply our usual definition of matrix similarity in $\mathbb{F}^{n \times n}$.

Definition

For a field $\mathbb{F}$, we say that $n \times n$ matrices A and B with entries in $\mathbb{F}$ are *similar over* $\mathbb{F}$ if there exists an invertible matrix $P \in \mathbb{F}^{n \times n}$ s.t. $B = P^{-1}AP$.

- This is simply our usual definition of matrix similarity in $\mathbb{F}^{n \times n}$.
- However, if $\mathbb{F}$ is a subfield of some larger field $\widetilde{\mathbb{F}}$, then it makes sense to speak of A and B being (or not being) similar over $\mathbb{F}$, or of them being (or not being) similar over $\widetilde{\mathbb{F}}$.

Definition

For a field $\mathbb{F}$, we say that $n \times n$ matrices A and B with entries in $\mathbb{F}$ are *similar over* $\mathbb{F}$ if there exists an invertible matrix $P \in \mathbb{F}^{n \times n}$ s.t. $B = P^{-1}AP$.

- This is simply our usual definition of matrix similarity in $\mathbb{F}^{n \times n}$.
- However, if $\mathbb{F}$ is a subfield of some larger field $\widetilde{\mathbb{F}}$, then it makes sense to speak of A and B being (or not being) similar over $\mathbb{F}$, or of them being (or not being) similar over $\widetilde{\mathbb{F}}$.
- In fact, it can be shown that the two notions are equivalent.

Definition

For a field $\mathbb{F}$, we say that $n \times n$ matrices A and B with entries in $\mathbb{F}$ are *similar over* $\mathbb{F}$ if there exists an invertible matrix $P \in \mathbb{F}^{n\times n}$ s.t. $B = P^{-1}AP$.

- This is simply our usual definition of matrix similarity in $\mathbb{F}^{n\times n}$.
- However, if $\mathbb{F}$ is a subfield of some larger field $\widetilde{\mathbb{F}}$, then it makes sense to speak of A and B being (or not being) similar over $\mathbb{F}$, or of them being (or not being) similar over $\widetilde{\mathbb{F}}$.
- In fact, it can be shown that the two notions are equivalent.
- More precisely, it can be shown that if $\mathbb{F}$ is a subfield of $\widetilde{\mathbb{F}}$, then $n \times n$ matrices A and B, with entries in $\mathbb{F}$, are similar over $\mathbb{F}$ iff they are similar over $\widetilde{\mathbb{F}}$, that is, the following are equivalent:
    - there exists an invertible matrix $P \in \mathbb{F}^{n\times n}$ s.t. $B = P^{-1}AP$;
    - there exists an invertible matrix $P \in \widetilde{\mathbb{F}}^{n\times n}$ s.t. $B = P^{-1}AP$.

Definition

For a field $\mathbb{F}$, we say that $n \times n$ matrices A and B with entries in $\mathbb{F}$ are *similar over* $\mathbb{F}$ if there exists an invertible matrix $P \in \mathbb{F}^{n \times n}$ s.t. $B = P^{-1}AP$.

- This is simply our usual definition of matrix similarity in $\mathbb{F}^{n \times n}$.
- However, if $\mathbb{F}$ is a subfield of some larger field $\widetilde{\mathbb{F}}$, then it makes sense to speak of A and B being (or not being) similar over $\mathbb{F}$, or of them being (or not being) similar over $\widetilde{\mathbb{F}}$.
- In fact, it can be shown that the two notions are equivalent.
- More precisely, it can be shown that if $\mathbb{F}$ is a subfield of $\widetilde{\mathbb{F}}$, then $n \times n$ matrices A and B, with entries in $\mathbb{F}$, are similar over $\mathbb{F}$ iff they are similar over $\widetilde{\mathbb{F}}$, that is, the following are equivalent:
  - there exists an invertible matrix $P \in \mathbb{F}^{n \times n}$ s.t. $B = P^{-1}AP$;
  - there exists an invertible matrix $P \in \widetilde{\mathbb{F}}^{n \times n}$ s.t. $B = P^{-1}AP$.
- We will not prove this in full generality, since it would involve theory that is beyond the scope of this course.

- For the special case of $\mathbb{R}$ and $\mathbb{C}$, we give a proof (nice and not very hard) in the Lecture Notes.

Theorem 8.6.7

Two $n \times n$ matrices with real entries are similar over $\mathbb{R}$ iff they are similar over $\mathbb{C}$.

- For the special case of $\mathbb{R}$ and $\mathbb{C}$, we give a proof (nice and not very hard) in the Lecture Notes.

Theorem 8.6.7

Two $n \times n$ matrices with real entries are similar over $\mathbb{R}$ iff they are similar over $\mathbb{C}$.

- What does this have to do with the Jordan normal form?

- Suppose that we need to check if two $n \times n$ matrices, call them A and B, with entries in some field $\mathbb{F}$, are similar (over $\mathbb{F}$).

- Suppose that we need to check if two $n \times n$ matrices, call them A and B, with entries in some field $\mathbb{F}$, are similar (over $\mathbb{F}$).
- We first extend $\mathbb{F}$ to an algebraically closed field $\widetilde{\mathbb{F}}$.
  - For example, we extend $\mathbb{R}$ to $\mathbb{C}$.

- Suppose that we need to check if two $n \times n$ matrices, call them A and B, with entries in some field $\mathbb{F}$, are similar (over $\mathbb{F}$).
- We first extend $\mathbb{F}$ to an algebraically closed field $\widetilde{\mathbb{F}}$.
    - For example, we extend $\mathbb{R}$ to $\mathbb{C}$.
- Then the following are equivalent:
    - A and B are similar over $\mathbb{F}$;
    - A and B are similar over $\widetilde{\mathbb{F}}$;
    - A and B have the same Jordan normal form in $\widetilde{\mathbb{F}}^{n \times n}$ (up to a reordering of the Jordan blocks).

  (The equivalence of the second and third item above follows from Corollary 8.6.3.)

- Suppose that we need to check if two $n \times n$ matrices, call them A and B, with entries in some field $\mathbb{F}$, are similar (over $\mathbb{F}$).
- We first extend $\mathbb{F}$ to an algebraically closed field $\widetilde{\mathbb{F}}$.
  - For example, we extend $\mathbb{R}$ to $\mathbb{C}$.
- Then the following are equivalent:
  - A and B are similar over $\mathbb{F}$;
  - A and B are similar over $\widetilde{\mathbb{F}}$;
  - A and B have the same Jordan normal form in $\widetilde{\mathbb{F}}^{n \times n}$ (up to a reordering of the Jordan blocks).

  (The equivalence of the second and third item above follows from Corollary 8.6.3.)
- So, if we can compute the Jordan normal forms of A and B in $\widetilde{\mathbb{F}}^{n \times n}$, then we can immediately determine if A and B are similar over $\mathbb{F}$.

- Suppose that we need to check if two $n \times n$ matrices, call them A and B, with entries in some field $\mathbb{F}$, are similar (over $\mathbb{F}$).
- We first extend $\mathbb{F}$ to an algebraically closed field $\widetilde{\mathbb{F}}$.
  - For example, we extend $\mathbb{R}$ to $\mathbb{C}$.
- Then the following are equivalent:
  - A and B are similar over $\mathbb{F}$;
  - A and B are similar over $\widetilde{\mathbb{F}}$;
  - A and B have the same Jordan normal form in $\widetilde{\mathbb{F}}^{n \times n}$ (up to a reordering of the Jordan blocks).

  (The equivalence of the second and third item above follows from Corollary 8.6.3.)
- So, if we can compute the Jordan normal forms of A and B in $\widetilde{\mathbb{F}}^{n \times n}$, then we can immediately determine if A and B are similar over $\mathbb{F}$.
- Of course, actually computing the Jordan normal forms of A and B (in $\widetilde{\mathbb{F}}^{n \times n}$) may be very difficult or even impossible, essentially because we might not succeed in factoring the characteristic polynomials $p_A(\lambda)$ and $p_B(\lambda)$.

2. Symmetric matrices and orthogonal diagonalization

2. Symmetric matrices and orthogonal diagonalization

- Reminder: The complex conjugate of a complex number z is denoted by $\overline{z}$.

2. Symmetric matrices and orthogonal diagonalization

- Reminder: The complex conjugate of a complex number z is denoted by $\overline{z}$.
- Reminder: The *standard scalar product* in $\mathbb{R}^n$ or $\mathbb{C}^n$, denoted by $\cdot$, is defined as follows:

2. Symmetric matrices and orthogonal diagonalization

- Reminder: The complex conjugate of a complex number z is denoted by $\overline{z}$.
- Reminder: The *standard scalar product* in $\mathbb{R}^n$ or $\mathbb{C}^n$, 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 $\mathbb{R}^n$:

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

2. Symmetric matrices and orthogonal diagonalization

- Reminder: The complex conjugate of a complex number z is denoted by $\overline{z}$.
- Reminder: The *standard scalar product* in $\mathbb{R}^n$ or $\mathbb{C}^n$, 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 $\mathbb{R}^n$:
$$\mathbf{x} \cdot \mathbf{y} = \sum_{k=1}^{n} x_k y_k;$$
  - 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 $\mathbb{C}^n$:
$$\mathbf{x} \cdot \mathbf{y} = \sum_{k=1}^{n} x_k \overline{y_k}.$$

2. Symmetric matrices and orthogonal diagonalization

- Reminder: The complex conjugate of a complex number z is denoted by $\overline{z}$.
- Reminder: The *standard scalar product* in $\mathbb{R}^n$ or $\mathbb{C}^n$, 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 $\mathbb{R}^n$:
$$\mathbf{x} \cdot \mathbf{y} = \sum_{k=1}^{n} x_k y_k;$$
  - 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 $\mathbb{C}^n$:
$$\mathbf{x} \cdot \mathbf{y} = \sum_{k=1}^{n} x_k \overline{y_k}.$$
- In what follows, we shall denote by $|| \cdot ||$ the norm induced by the standard scalar product $\cdot$ in $\mathbb{R}^n$ or $\mathbb{C}^n$ (as appropriate).

2. Symmetric matrices and orthogonal diagonalization

- Reminder: The complex conjugate of a complex number z is denoted by $\overline{z}$.
- Reminder: The *standard scalar product* in $\mathbb{R}^n$ or $\mathbb{C}^n$, 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 $\mathbb{R}^n$:
  $$\mathbf{x} \cdot \mathbf{y} = \sum_{k=1}^{n} x_k y_k;$$
  - 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 $\mathbb{C}^n$:
  $$\mathbf{x} \cdot \mathbf{y} = \sum_{k=1}^{n} x_k \overline{y_k}.$$
- In what follows, we shall denote by $|| \cdot ||$ the norm induced by the standard scalar product $\cdot$ in $\mathbb{R}^n$ or $\mathbb{C}^n$ (as appropriate).
- In particular, orthogonality and orthonormality will always be assumed to be with respect to the standard scalar product and the induced norm.

- For any field $\mathbb{F}$, a matrix $A \in \mathbb{F}^{n \times n}$ is *symmetric* if $A^T = A$.

- For any field $\mathbb{F}$, a 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 any field $\mathbb{F}$, a 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 $A = \left[\ a_{i,j}\ \right]_{n \times m}$ in $\mathbb{C}^{n \times m}$, we set $\overline{A} = \left[\ \overline{a_{i,j}}\ \right]_{n \times m}$, i.e. $\overline{A}$ is an $n \times m$ matrix s.t. for all indices $i \in \{1, \dots, n\}$ and $j \in \{1, \dots, m\}$, the i, j-th entry of $\overline{A}$ is $\overline{a_{i,j}}$ (the complex conjugate of $a_{i,j}$).

- For any field $\mathbb{F}$, a 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 $A = \left[\begin{array}{c} a_{i,j} \end{array}\right]_{n \times m}$ in $\mathbb{C}^{n \times m}$, we set $\overline{A} = \left[\begin{array}{c} \overline{a_{i,j}} \end{array}\right]_{n \times m}$, i.e. $\overline{A}$ is an $n \times m$ matrix s.t. for all indices $i \in \{1, \dots, n\}$ and $j \in \{1, \dots, m\}$, the i, j-th entry of $\overline{A}$ is $\overline{a_{i,j}}$ (the complex conjugate of $a_{i,j}$).
- The *Hermitian transpose* of a A is the matrix $A^* = (\overline{A})^T$.

- For any field $\mathbb{F}$, a 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 $A = \left[\ a_{i,j}\ \right]_{n \times m}$ in $\mathbb{C}^{n \times m}$, we set $\overline{A} = \left[\ \overline{a_{i,j}}\ \right]_{n \times m}$, i.e. $\overline{A}$ is an $n \times m$ matrix s.t. for all indices $i \in \{1, \dots, n\}$ and $j \in \{1, \dots, m\}$, the i,j-th entry of $\overline{A}$ is $\overline{a_{i,j}}$ (the complex conjugate of $a_{i,j}$).
- The *Hermitian transpose* of a A is the matrix $A^* = (\overline{A})^T$.
- For example, for

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

we have the following:

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

- For any field $\mathbb{F}$, a 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 $A = \begin{bmatrix} a_{i,j} \end{bmatrix}_{n \times m}$ in $\mathbb{C}^{n \times m}$, we set $\overline{A} = \begin{bmatrix} \overline{a_{i,j}} \end{bmatrix}_{n \times m}$, i.e. $\overline{A}$ is an $n \times m$ matrix s.t. for all indices $i \in \{1, \dots, n\}$ and $j \in \{1, \dots, m\}$, the i,j-th entry of $\overline{A}$ is $\overline{a_{i,j}}$ (the complex conjugate of $a_{i,j}$).
- The *Hermitian transpose* of a A is the matrix $A^* = (\overline{A})^T$.
- For example, for

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

 we have the following:

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

- A square matrix $A \in \mathbb{C}^{n \times n}$ is *Hermitian* if $A^* = A$.

- A square matrix $A \in \mathbb{C}^{n \times n}$ is *Hermitian* if $A^* = A$.

- 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.

- 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.

- 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 iff it is symmetric.

Proposition 8.7.1

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

Proof.

Proposition 8.7.1

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

Proof. For any vector $\mathbf{x} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T$ in $\mathbb{C}^n$, we have that

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

which is what we needed. $\square$

Proposition 8.7.1

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

Proof. For any vector $\mathbf{x} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T$ in $\mathbb{C}^n$, we have that

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

which is what we needed. $\square$

Proposition 8.7.2

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

a. $(A^*)^* = A$;
b. $(A + B)^* = A^* + B^*$;
c. $(\alpha A)^* = \overline{\alpha} A^*$;
d. $(AB)^* = B^* A^*$.

- Proof: exercise.
- Note that this is very similar to the properties of the ordinary transpose.

Theorem 8.7.3

All eigenvalues of a Hermetrian matrix are real.

- **Remark:** The field $\mathbb{C}$ is algebraically closed, and consequently, every matrix in $\mathbb{C}^{n \times n}$ has n complex eigenvalues (with algebraic multiplicities taken into account). So, Theorem 8.7.3 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.

Theorem 8.7.3

All eigenvalues of a Hermetrian matrix are real.

Proof.

Theorem 8.7.3

All eigenvalues of a Hermetrian matrix are real.

Proof. Let $A \in \mathbb{C}^{n \times n}$ be a Hermitian matrix, let λ be any eigenvalue of A, and let $\mathbf{x}$ be an associated eigenvector of A.

Theorem 8.7.3

All eigenvalues of a Hermetrian matrix are real.

Proof. Let $A \in \mathbb{C}^{n\times n}$ be a Hermitian matrix, let λ be any eigenvalue of A, and let $\mathbf{x}$ be an associated eigenvector of A. After possibly normalizing the eigenvector $\mathbf{x}$ (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 $||\mathbf{x}|| = 1$.

Theorem 8.7.3

All eigenvalues of a Hermetrian matrix are real.

Proof. Let $A \in \mathbb{C}^{n \times n}$ be a Hermitian matrix, let λ be any eigenvalue of A, and let $\mathbf{x}$ be an associated eigenvector of A. After possibly normalizing the eigenvector $\mathbf{x}$ (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 compute:

$$\begin{array}{rcll} \mathbf{x}^* A\mathbf{x} & = & \mathbf{x}^*(\lambda\mathbf{x}) & \text{because } A\mathbf{x} = \lambda\mathbf{x} \\ & = & \lambda(\mathbf{x}^*\mathbf{x}) & \\ & = & \lambda||\mathbf{x}||^2 & \text{by Proposition 8.7.1} \\ & = & \lambda & \text{because } ||\mathbf{x}|| = 1. \end{array}$$

Theorem 8.7.3

All eigenvalues of a Hermetrian matrix are real.

Proof (continued). Reminder: $\mathbf{x}^* A\mathbf{x} = \lambda$.

But now we have the following:

$$\begin{array}{rcll}
\lambda & = & \mathbf{x}^* A\mathbf{x} & \\
 & = & \mathbf{x}^* A^*\mathbf{x} & \text{because } A \text{ is Hermitian} \\
 & = & \mathbf{x}^* A^*(\mathbf{x}^*)^* & \text{by Proposition 8.7.2(a)} \\
 & = & (\mathbf{x}^* A\mathbf{x})^* & \text{by Proposition 8.7.2(d)} \\
 & = & \lambda^* & \text{where we consider } \lambda \text{ as} \\
 & & & \text{a } 1 \times 1 \text{ complex matrix} \\
 & = & \overline{\lambda} & \text{where we consider } \lambda \text{ as} \\
 & & & \text{a complex number.}
\end{array}$$

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

Theorem 8.7.3

All eigenvalues of a Hermetrian matrix are real.

Theorem 8.7.3

All eigenvalues of a Hermetrian matrix are real.

Corollary 8.7.4

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.

Theorem 8.7.3

All eigenvalues of a Hermetrian matrix are real.

Corollary 8.7.4

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 Hermitian, and so Theorem 8.7.3 guarantees that all complex eigenvalues of A are in fact real. Finally, the fact that A has n complex eigenvalues follows from the fact that $\mathbb{C}$ is algebraically closed. $\square$

Definition

A matrix $Q \in \mathbb{R}^{n \times n}$ is *orthogonal* if it satisfies $Q^T Q = I_n$.

Theorem 6.8.1

Let $Q \in \mathbb{R}^{n \times n}$. Then the following are equivalent:

(a) Q is orthogonal (i.e. satisfies $Q^T Q = I_n$);

(b) Q is invertible and satisfies $Q^{-1} = Q^T$;

(c) $QQ^T = I_n$;

(d) Q^T is orthogonal;

(e) Q is invertible and Q^{-1} is orthogonal;

(f) the columns of Q form an orthonormal basis of $\mathbb{R}^n$;

(g) the columns of Q^T form an orthonormal basis of $\mathbb{R}^n$.

Definition

A matrix $Q \in \mathbb{R}^{n \times n}$ is *orthogonal* if it satisfies $Q^T Q = I_n$.

Theorem 6.8.1

Let $Q \in \mathbb{R}^{n \times n}$. Then the following are equivalent:

(a) Q is orthogonal (i.e. satisfies $Q^T Q = I_n$);

(b) Q is invertible and satisfies $Q^{-1} = Q^T$;

(c) $QQ^T = I_n$;

(d) Q^T is orthogonal;

(e) Q is invertible and Q^{-1} is orthogonal;

(f) the columns of Q form an orthonormal basis of $\mathbb{R}^n$;

(g) the columns of Q^T form an orthonormal basis of $\mathbb{R}^n$.

- In what follows, we will repeatedly use the fact that the three red statements above are equivalent, without explicitly mentioning Theorem 6.8.1.

Definition

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 $\mathbb{R}^{n \times n}$, s.t. $D = Q^T A Q$.

Definition

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 $\mathbb{R}^{n\times n}$, s.t. $D = Q^T AQ$.

- Since orthogonal matrices Q are invertible and satisfy $Q^T = Q^{-1}$, we see that orthogonally diagonalizable matrices are, in particular, diagonalizable in the usual sense.

Definition

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 $\mathbb{R}^{n\times n}$, s.t. $D = Q^T AQ$.

- Since orthogonal matrices Q are invertible and satisfy $Q^T = Q^{-1}$, we see that orthogonally diagonalizable matrices are, in particular, diagonalizable in the usual sense.
- Our goal is to prove the following theorem:

Theorem 8.7.6

A matrix in $\mathbb{R}^{n\times n}$ is orthogonally diagonalizable iff it is symmetric.

Definition

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 $\mathbb{R}^{n \times n}$, s.t. $D = Q^T AQ$.

- Since orthogonal matrices Q are invertible and satisfy $Q^T = Q^{-1}$, we see that orthogonally diagonalizable matrices are, in particular, diagonalizable in the usual sense.
- Our goal is to prove the following theorem:

Theorem 8.7.6

A matrix in $\mathbb{R}^{n \times n}$ is orthogonally diagonalizable iff it is symmetric.

- The proof proceeds by induction on n, and in the induction step, it will be convenient to reduce the problem to the case when the matrix has an eigenvalue 0. To this end, we will use the following technical proposition (next slide).

Proposition 8.7.5

Let $A \in \mathbb{R}^{n \times n}$ and $\lambda_0 \in \mathbb{R}$. Then all the following hold:

(a) λ_0 is an eigenvalue of A iff 0 is an eigenvalue of $A - \lambda_0 I_n$, and moreover, $E_{\lambda_0}(A) = E_0(A - \lambda_0 I_n)$;[a]

(b) A is symmetric iff $A - \lambda_0 I_n$ is symmetric;

(c) A is diagonalizable iff $A - \lambda_0$ is digonalizable;

(d) A is orthogonally disagonalizable iff $A - \lambda_0 I_n$ is orthogonally diagonalizable.

---

[a]Here, $E_{\lambda_0}(A) = E_0(A - \lambda_0 I_n)$ holds even if λ_0 is not an eigenvalue of A. In that case, we simply have that $E_{\lambda_0}(A) = E_0(A - \lambda_0 I_n) = \{\mathbf{0}\}$.

Proof.

Proposition 8.7.5

Let $A \in \mathbb{R}^{n\times n}$ and $\lambda_0 \in \mathbb{R}$. Then all the following hold:

(a) λ_0 is an eigenvalue of A iff 0 is an eigenvalue of $A - \lambda_0 I_n$, and moreover, $E_{\lambda_0}(A) = E_0(A - \lambda_0 I_n)$;[a]

(b) A is symmetric iff $A - \lambda_0 I_n$ is symmetric;

(c) A is diagonalizable iff $A - \lambda_0$ is digonalizable;

(d) A is orthogonally disagonalizable iff $A - \lambda_0 I_n$ is orthogonally diagonalizable.

---

[a] Here, $E_{\lambda_0}(A) = E_0(A - \lambda_0 I_n)$ holds even if λ_0 is not an eigenvalue of A. In that case, we simply have that $E_{\lambda_0}(A) = E_0(A - \lambda_0 I_n) = \{\mathbf{0}\}$.

Proof. (a) For all $\mathbf{v} \in \mathbb{R}^n$, we have that $A\mathbf{v} = \lambda_0\mathbf{v}$ iff $(A - \lambda_0 I_n)\mathbf{v} = \mathbf{0} = 0\mathbf{v}$,

Proposition 8.7.5

Let $A \in \mathbb{R}^{n \times n}$ and $\lambda_0 \in \mathbb{R}$. Then all the following hold:

- (a) λ_0 is an eigenvalue of A iff 0 is an eigenvalue of $A - \lambda_0 I_n$, and moreover, $E_{\lambda_0}(A) = E_0(A - \lambda_0 I_n)$;[a]
- (b) A is symmetric iff $A - \lambda_0 I_n$ is symmetric;
- (c) A is diagonalizable iff $A - \lambda_0$ is digonalizable;
- (d) A is orthogonally disagonalizable iff $A - \lambda_0 I_n$ is orthogonally diagonalizable.

---

[a] Here, $E_{\lambda_0}(A) = E_0(A - \lambda_0 I_n)$ holds even if λ_0 is not an eigenvalue of A. In that case, we simply have that $E_{\lambda_0}(A) = E_0(A - \lambda_0 I_n) = \{\mathbf{0}\}$.

Proof. (a) For all $\mathbf{v} \in \mathbb{R}^n$, we have that $A\mathbf{v} = \lambda_0 \mathbf{v}$ iff $(A - \lambda_0 I_n)\mathbf{v} = \mathbf{0} = 0\mathbf{v}$, and so $\mathbf{v} \in E_{\lambda_0}(A)$ iff $\mathbf{v} \in E_0(A - \lambda_0 I_n)$.

Proposition 8.7.5

Let $A \in \mathbb{R}^{n\times n}$ and $\lambda_0 \in \mathbb{R}$. Then all the following hold:

- (a) λ_0 is an eigenvalue of A iff 0 is an eigenvalue of $A - \lambda_0 I_n$, and moreover, $E_{\lambda_0}(A) = E_0(A - \lambda_0 I_n)$;[a]
- (b) A is symmetric iff $A - \lambda_0 I_n$ is symmetric;
- (c) A is diagonalizable iff $A - \lambda_0$ is digonalizable;
- (d) A is orthogonally disagonalizable iff $A - \lambda_0 I_n$ is orthogonally diagonalizable.

[a] Here, $E_{\lambda_0}(A) = E_0(A - \lambda_0 I_n)$ holds even if λ_0 is not an eigenvalue of A. In that case, we simply have that $E_{\lambda_0}(A) = E_0(A - \lambda_0 I_n) = \{\mathbf{0}\}$.

Proof. (a) For all $\mathbf{v} \in \mathbb{R}^n$, we have that $A\mathbf{v} = \lambda_0 \mathbf{v}$ iff $(A - \lambda_0 I_n)\mathbf{v} = \mathbf{0} = 0\mathbf{v}$, and so $\mathbf{v} \in E_{\lambda_0}(A)$ iff $\mathbf{v} \in E_0(A - \lambda_0 I_n)$. Thus, $E_{\lambda_0}(A) = E_0(A - \lambda_0 I_n)$.

Proposition 8.7.5

Let $A \in \mathbb{R}^{n \times n}$ and $\lambda_0 \in \mathbb{R}$. Then all the following hold:

- (a) λ_0 is an eigenvalue of A iff 0 is an eigenvalue of $A - \lambda_0 I_n$, and moreover, $E_{\lambda_0}(A) = E_0(A - \lambda_0 I_n)$;[a]
- (b) A is symmetric iff $A - \lambda_0 I_n$ is symmetric;
- (c) A is diagonalizable iff $A - \lambda_0$ is digonalizable;
- (d) A is orthogonally disagonalizable iff $A - \lambda_0 I_n$ is orthogonally diagonalizable.

[a]Here, $E_{\lambda_0}(A) = E_0(A - \lambda_0 I_n)$ holds even if λ_0 is not an eigenvalue of A. In that case, we simply have that $E_{\lambda_0}(A) = E_0(A - \lambda_0 I_n) = \{\mathbf{0}\}$.

Proof. (a) For all $\mathbf{v} \in \mathbb{R}^n$, we have that $A\mathbf{v} = \lambda_0 \mathbf{v}$ iff $(A - \lambda_0 I_n)\mathbf{v} = \mathbf{0} = 0\mathbf{v}$, and so $\mathbf{v} \in E_{\lambda_0}(A)$ iff $\mathbf{v} \in E_0(A - \lambda_0 I_n)$. Thus, $E_{\lambda_0}(A) = E_0(A - \lambda_0 I_n)$.

In particular, $E_{\lambda_0}(A)$ is non-trivial iff $E_0(A - \lambda_0 I_n)$ is non-trivial,

Proposition 8.7.5

Let $A \in \mathbb{R}^{n \times n}$ and $\lambda_0 \in \mathbb{R}$. Then all the following hold:

- (a) λ_0 is an eigenvalue of A iff 0 is an eigenvalue of $A - \lambda_0 I_n$, and moreover, $E_{\lambda_0}(A) = E_0(A - \lambda_0 I_n)$;[a]
- (b) A is symmetric iff $A - \lambda_0 I_n$ is symmetric;
- (c) A is diagonalizable iff $A - \lambda_0$ is digonalizable;
- (d) A is orthogonally disagonalizable iff $A - \lambda_0 I_n$ is orthogonally diagonalizable.

[a] Here, $E_{\lambda_0}(A) = E_0(A - \lambda_0 I_n)$ holds even if λ_0 is not an eigenvalue of A. In that case, we simply have that $E_{\lambda_0}(A) = E_0(A - \lambda_0 I_n) = \{\mathbf{0}\}$.

Proof. (a) For all $\mathbf{v} \in \mathbb{R}^n$, we have that $A\mathbf{v} = \lambda_0 \mathbf{v}$ iff $(A - \lambda_0 I_n)\mathbf{v} = \mathbf{0} = 0\mathbf{v}$, and so $\mathbf{v} \in E_{\lambda_0}(A)$ iff $\mathbf{v} \in E_0(A - \lambda_0 I_n)$. Thus, $E_{\lambda_0}(A) = E_0(A - \lambda_0 I_n)$.

In particular, $E_{\lambda_0}(A)$ is non-trivial iff $E_0(A - \lambda_0 I_n)$ is non-trivial, and consequently (by definition, or alternatively, by Prop. 8.1.6(a)), λ_0 is an eigenvalue of A iff 0 is an eigenvalue of $A - \lambda_0 I_n$.

Proposition 8.7.5

Let $A \in \mathbb{R}^{n\times n}$ and $\lambda_0 \in \mathbb{R}$. Then all the following hold:

(a) λ_0 is an eigenvalue of A iff 0 is an eigenvalue of $A - \lambda_0 I_n$, and moreover, $E_{\lambda_0}(A) = E_0(A - \lambda_0 I_n)$;[a]

(b) A is symmetric iff $A - \lambda_0 I_n$ is symmetric;

(c) A is diagonalizable iff $A - \lambda_0$ is digonalizable;

(d) A is orthogonally disagonalizable iff $A - \lambda_0 I_n$ is orthogonally diagonalizable.

---

[a]Here, $E_{\lambda_0}(A) = E_0(A - \lambda_0 I_n)$ holds even if λ_0 is not an eigenvalue of A. In that case, we simply have that $E_{\lambda_0}(A) = E_0(A - \lambda_0 I_n) = \{\mathbf{0}\}$.

Proof (continued). (b) First, we note that

$$(A - \lambda_0 I_n)^T \quad = \quad A^T - \lambda_0 I_n^T \quad = \quad A^T - \lambda_0 I_n;$$

Proposition 8.7.5

Let $A \in \mathbb{R}^{n\times n}$ and $\lambda_0 \in \mathbb{R}$. Then all the following hold:

(a) λ_0 is an eigenvalue of A iff 0 is an eigenvalue of $A - \lambda_0 I_n$, and moreover, $E_{\lambda_0}(A) = E_0(A - \lambda_0 I_n)$;[a]

(b) A is symmetric iff $A - \lambda_0 I_n$ is symmetric;

(c) A is diagonalizable iff $A - \lambda_0$ is digonalizable;

(d) A is orthogonally disagonalizable iff $A - \lambda_0 I_n$ is orthogonally diagonalizable.

[a] Here, $E_{\lambda_0}(A) = E_0(A - \lambda_0 I_n)$ holds even if λ_0 is not an eigenvalue of A. In that case, we simply have that $E_{\lambda_0}(A) = E_0(A - \lambda_0 I_n) = \{\mathbf{0}\}$.

Proof (continued). (b) First, we note that

$$(A - \lambda_0 I_n)^T \;=\; A^T - \lambda_0 I_n^T \;=\; A^T - \lambda_0 I_n;$$

consequently, $(A - \lambda_0 I_n)^T = A - \lambda_0 I_n$ iff $A^T = A$, i.e. $A - \lambda_0 I_n$ is symmetric iff A is symmetric.

Proposition 8.7.5

Let $A \in \mathbb{R}^{n \times n}$ and $\lambda_0 \in \mathbb{R}$. Then all the following hold:

(c) A is diagonalizable iff $A - \lambda_0$ is digonalizable;

(d) A is orthogonally disagonalizable iff $A - \lambda_0 I_n$ is orthogonally diagonalizable.

Proof (continued). (c) Suppose first that A is diagonalizable.

Proposition 8.7.5

Let $A \in \mathbb{R}^{n \times n}$ and $\lambda_0 \in \mathbb{R}$. Then all the following hold:

(c) A is diagonalizable iff $A - \lambda_0$ is digonalizable;

(d) A is orthogonally disagonalizable iff $A - \lambda_0 I_n$ is orthogonally diagonalizable.

Proof (continued). (c) Suppose first that A is diagonalizable. Then there exist a diagonal matrix D and an invertible matrix P, both in $\mathbb{R}^{n \times n}$, s.t. $D = P^{-1}AP$.

Proposition 8.7.5

Let $A \in \mathbb{R}^{n \times n}$ and $\lambda_0 \in \mathbb{R}$. Then all the following hold:

- (c) A is diagonalizable iff $A - \lambda_0$ is digonalizable;
- (d) A is orthogonally disagonalizable iff $A - \lambda_0 I_n$ is orthogonally diagonalizable.

Proof (continued). (c) Suppose first that A is diagonalizable. Then there exist a diagonal matrix D and an invertible matrix P, both in $\mathbb{R}^{n \times n}$, s.t. $D = P^{-1}AP$. But then

$$\begin{aligned} P^{-1}(A - \lambda_0 I_n)P &= P^{-1}AP - P^{-1}(\lambda_0 I_n)P \\ &= \underbrace{P^{-1}AP}_{=D} - \lambda_0 \underbrace{P^{-1}P}_{=I_n} \\ &= D - \lambda_0 I_n, \end{aligned}$$

and obviously, $D - \lambda_0 I_n$ is diagonal.

Proposition 8.7.5

Let $A \in \mathbb{R}^{n \times n}$ and $\lambda_0 \in \mathbb{R}$. Then all the following hold:

(c) A is diagonalizable iff $A - \lambda_0$ is digonalizable;

(d) A is orthogonally disagonalizable iff $A - \lambda_0 I_n$ is orthogonally diagonalizable.

Proof (continued). (c) Suppose first that A is diagonalizable. Then there exist a diagonal matrix D and an invertible matrix P, both in $\mathbb{R}^{n \times n}$, s.t. $D = P^{-1}AP$. But then

$$\begin{aligned} P^{-1}(A - \lambda_0 I_n)P &= P^{-1}AP - P^{-1}(\lambda_0 I_n)P \\ &= \underbrace{P^{-1}AP}_{=D} - \lambda_0 \underbrace{P^{-1}P}_{=I_n} \\ &= D - \lambda_0 I_n, \end{aligned}$$

and obviously, $D - \lambda_0 I_n$ is diagonal. So, $A - \lambda_0 I_n$ is diagonalizable. The proof of the converse is analogous.

Proposition 8.7.5

Let $A \in \mathbb{R}^{n\times n}$ and $\lambda_0 \in \mathbb{R}$. Then all the following hold:

a. λ_0 is an eigenvalue of A iff 0 is an eigenvalue of $A - \lambda_0 I_n$, and moreover, $E_{\lambda_0}(A) = E_0(A - \lambda_0 I_n)$;[a]
b. A is symmetric iff $A - \lambda_0 I_n$ is symmetric;
c. A is diagonalizable iff $A - \lambda_0$ is digonalizable;
d. A is orthogonally disagonalizable iff $A - \lambda_0 I_n$ is orthogonally diagonalizable.

[a] Here, $E_{\lambda_0}(A) = E_0(A - \lambda_0 I_n)$ holds even if λ_0 is not an eigenvalue of A. In that case, we simply have that $E_{\lambda_0}(A) = E_0(A - \lambda_0 I_n) = \{\mathbf{0}\}$.

Proof (continued). (d) This is completely analogous to the proof of (c), except that instead of P and P^{-1} (where $P \in \mathbb{R}^{n\times n}$ is an invertible matrix), we have Q and Q^T (where $Q \in \mathbb{R}^{n\times n}$ is an orthogonal matrix). $\square$

Theorem 8.7.6

A matrix in $\mathbb{R}^{n\times n}$ is orthogonally diagonalizable iff it is symmetric.

Proof.

Theorem 8.7.6

A matrix in $\mathbb{R}^{n \times n}$ is orthogonally diagonalizable iff it is symmetric.

Proof. Let us first show that orthogonally diagonalizable matrices are symmeteric.

Theorem 8.7.6

A matrix in $\mathbb{R}^{n\times n}$ is orthogonally diagonalizable iff it is symmetric.

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 $\mathbb{R}^{n\times n}$, s.t. $D = Q^T AQ$.

Theorem 8.7.6

A matrix in $\mathbb{R}^{n\times n}$ is orthogonally diagonalizable iff it is symmetric.

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 $\mathbb{R}^{n\times n}$, s.t. $D = Q^T AQ$. Then $A = QDQ^T$, and we see that

$$A^T \;=\; (QDQ^T)^T \;=\; (Q^T)^T D^T Q^T \;\stackrel{(*)}{=}\; QDQ^T \;=\; A,$$

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

Theorem 8.7.6

A matrix in $\mathbb{R}^{n\times n}$ is orthogonally diagonalizable iff it is symmetric.

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

Theorem 8.7.6

A matrix in $\mathbb{R}^{n\times n}$ is orthogonally diagonalizable iff it is symmetric.

Proof (continued). 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 $Q := I_1$ to obtain $D = Q^T AQ$.

Theorem 8.7.6

A matrix in $\mathbb{R}^{n\times n}$ is orthogonally diagonalizable iff it is symmetric.

Proof (continued). 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 $Q := I_1$ to obtain $D = Q^T AQ$.

Now, fix a positive integer n, and assume inductively that every symmetric matrix in $\mathbb{R}^{n\times n}$ is orthogonally diagonalizable.

Theorem 8.7.6

A matrix in $\mathbb{R}^{n\times n}$ is orthogonally diagonalizable iff it is symmetric.

Proof (continued). 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 $Q := I_1$ to obtain $D = Q^T AQ$.

Now, fix a positive integer n, and assume inductively that every symmetric matrix in $\mathbb{R}^{n\times n}$ is orthogonally diagonalizable. Fix any symmetric matrix $A \in \mathbb{R}^{(n+1)\times(n+1)}$; we must show that A is orthogonally diagonalizable.

Theorem 8.7.6

A matrix in $\mathbb{R}^{n\times n}$ is orthogonally diagonalizable iff it is symmetric.

Proof (continued). 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 $Q := I_1$ to obtain $D = Q^T AQ$.

Now, fix a positive integer n, and assume inductively that every symmetric matrix in $\mathbb{R}^{n\times n}$ is orthogonally diagonalizable. Fix any symmetric matrix $A \in \mathbb{R}^{(n+1)\times(n+1)}$; we must show that A is orthogonally diagonalizable. By Corollary 8.7.4, A has $n + 1$ real eigenvalues (with algebraic multiplicities taken into account).

Theorem 8.7.6

A matrix in $\mathbb{R}^{n\times n}$ is orthogonally diagonalizable iff it is symmetric.

Proof (continued). 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 $Q := I_1$ to obtain $D = Q^T AQ$.

Now, fix a positive integer n, and assume inductively that every symmetric matrix in $\mathbb{R}^{n\times n}$ is orthogonally diagonalizable. Fix any symmetric matrix $A \in \mathbb{R}^{(n+1)\times(n+1)}$; we must show that A is orthogonally diagonalizable. By Corollary 8.7.4, A has $n + 1$ real eigenvalues (with algebraic multiplicities taken into account). Let $\lambda_0 \in \mathbb{R}$ be an eigenvalue of A.

Theorem 8.7.6

A matrix in $\mathbb{R}^{n \times n}$ is orthogonally diagonalizable iff it is symmetric.

Proof (continued). 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 $Q := I_1$ to obtain $D = Q^T AQ$.

Now, fix a positive integer n, and assume inductively that every symmetric matrix in $\mathbb{R}^{n \times n}$ is orthogonally diagonalizable. Fix any symmetric matrix $A \in \mathbb{R}^{(n+1) \times (n+1)}$; we must show that A is orthogonally diagonalizable. By Corollary 8.7.4, A has $n + 1$ real eigenvalues (with algebraic multiplicities taken into account). Let $\lambda_0 \in \mathbb{R}$ be an eigenvalue of A. In view of Proposition 8.7.5, we may assume that $\lambda_0 = 0$, for otherwise, we simply consider $A - \lambda_0 I_n$ instead of A.

Theorem 8.7.6

A matrix in $\mathbb{R}^{n \times n}$ is orthogonally diagonalizable iff it is symmetric.

Proof (continued). Let $\mathbf{x}_0 \in \mathbb{R}^n$ be an eigenvector of A associated with the eigenvalue 0, so that $A\mathbf{x}_0 = \mathbf{0}$.

Theorem 8.7.6

A matrix in $\mathbb{R}^{n \times n}$ is orthogonally diagonalizable iff it is symmetric.

Proof (continued). Let $\mathbf{x}_0 \in \mathbb{R}^n$ be an eigenvector of A associated with the eigenvalue 0, so that $A\mathbf{x}_0 = \mathbf{0}$. After possibly normalizing the eigenvector $\mathbf{x}_0$ (i.e. replacing $\mathbf{x}_0$ by $\frac{\mathbf{x}_0}{||\mathbf{x}_0||}$), we may assume that $||\mathbf{x}_0|| = 1$.

Theorem 8.7.6

A matrix in $\mathbb{R}^{n \times n}$ is orthogonally diagonalizable iff it is symmetric.

Proof (continued). Let $\mathbf{x}_0 \in \mathbb{R}^n$ be an eigenvector of A associated with the eigenvalue 0, so that $A\mathbf{x}_0 = \mathbf{0}$. After possibly normalizing the eigenvector $\mathbf{x}_0$ (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 6.3.11(d), we let $\{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n\}$ be an orthonormal basis of $\mathbb{R}^{n+1}$.

- 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 6.3.11(d), $\{\mathbf{x}_0\}$ can be extended to an orthonormal basis of $\mathbb{R}^{n+1}$.

Theorem 8.7.6

A matrix in $\mathbb{R}^{n \times n}$ is orthogonally diagonalizable iff it is symmetric.

Proof (continued). Let $\mathbf{x}_0 \in \mathbb{R}^n$ be an eigenvector of A associated with the eigenvalue 0, so that $A\mathbf{x}_0 = \mathbf{0}$. After possibly normalizing the eigenvector $\mathbf{x}_0$ (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 6.3.11(d), we let $\{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n\}$ be an orthonormal basis of $\mathbb{R}^{n+1}$.

- 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 6.3.11(d), $\{\mathbf{x}_0\}$ can be extended to an orthonormal basis of $\mathbb{R}^{n+1}$.

Set $S := \begin{bmatrix} \mathbf{x}_0 & \mathbf{x}_1 & \dots & \mathbf{x}_n \end{bmatrix}$; then S is an orthogonal matrix.

Theorem 8.7.6

A matrix in $\mathbb{R}^{n\times n}$ is orthogonally diagonalizable iff it is symmetric.

Proof (continued). Let $\mathbf{x}_0 \in \mathbb{R}^n$ be an eigenvector of A associated with the eigenvalue 0, so that $A\mathbf{x}_0 = \mathbf{0}$. After possibly normalizing the eigenvector $\mathbf{x}_0$ (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 6.3.11(d), we let $\{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n\}$ be an orthonormal basis of $\mathbb{R}^{n+1}$.

- 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 6.3.11(d), $\{\mathbf{x}_0\}$ can be extended to an orthonormal basis of $\mathbb{R}^{n+1}$.

Set $S := \begin{bmatrix} \mathbf{x}_0 & \mathbf{x}_1 & \dots & \mathbf{x}_n \end{bmatrix}$; then S is an orthogonal matrix. Now, since A is symmetric, so is S^TAS; indeed,

$$(S^TAS)^T = S^TA^TS \overset{(*)}{=} S^TAS,$$

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

Theorem 8.7.6

A matrix in $\mathbb{R}^{n\times n}$ is orthogonally diagonalizable iff it is symmetric.

Proof (continued). Reminder: $A \in \mathbb{R}^{(n+1)\times(n+1)}$ is symmetric, $A\mathbf{x}_0 = \mathbf{0}$; $||\mathbf{x}_0|| = 1$; $S := \left[\begin{array}{cccc} \mathbf{x}_0 & \mathbf{x}_1 & \dots & \mathbf{x}_n \end{array}\right]$ is orthogonal; S^TAS is symmetric.

Theorem 8.7.6

A matrix in $\mathbb{R}^{n \times n}$ is orthogonally diagonalizable iff it is symmetric.

Proof (continued). Reminder: $A \in \mathbb{R}^{(n+1)\times(n+1)}$ is symmetric, $A\mathbf{x}_0 = \mathbf{0}$; $||\mathbf{x}_0|| = 1$; $S := \begin{bmatrix} \mathbf{x}_0 & \mathbf{x}_1 & \dots & \mathbf{x}_n \end{bmatrix}$ is orthogonal; $S^T AS$ is symmetric.

Moreover, it is easy to see that the first (i.e. leftmost) column of $S^T AS$ is $\mathbf{0}$; indeed:

$$\begin{aligned} S^T AS &= S^T A \begin{bmatrix} \mathbf{x}_0 & \mathbf{x}_1 & \dots & \mathbf{x}_n \end{bmatrix} \\ &= \begin{bmatrix} S^T A\mathbf{x}_0 & S^T A\mathbf{x}_1 & \dots & S^T A\mathbf{x}_n \end{bmatrix} && \text{by the definition of matrix multiplication} \\ &= \begin{bmatrix} S^T \mathbf{0} & S^T A\mathbf{x}_1 & \dots & S^T A\mathbf{x}_n \end{bmatrix} && \text{because } A\mathbf{x}_0 = \mathbf{0} \\ &= \begin{bmatrix} \mathbf{0} & S^T A\mathbf{x}_1 & \dots & S^T A\mathbf{x}_n \end{bmatrix}. \end{aligned}$$

Theorem 8.7.6

A matrix in $\mathbb{R}^{n \times n}$ is orthogonally diagonalizable iff it is symmetric.

Proof (continued). Reminder: $A \in \mathbb{R}^{(n+1)\times(n+1)}$ is symmetric, $A\mathbf{x}_0 = \mathbf{0}$; $||\mathbf{x}_0|| = 1$; $S := \begin{bmatrix} \mathbf{x}_0 & \mathbf{x}_1 & \dots & \mathbf{x}_n \end{bmatrix}$ is orthogonal; $S^T AS$ is symmetric.

Moreover, it is easy to see that the first (i.e. leftmost) column of $S^T AS$ is $\mathbf{0}$; indeed:

$$\begin{aligned}
S^T AS &= S^T A \begin{bmatrix} \mathbf{x}_0 & \mathbf{x}_1 & \dots & \mathbf{x}_n \end{bmatrix} \\
&= \begin{bmatrix} S^T A\mathbf{x}_0 & S^T A\mathbf{x}_1 & \dots & S^T A\mathbf{x}_n \end{bmatrix} && \text{by the definition of matrix multiplication} \\
&= \begin{bmatrix} S^T \mathbf{0} & S^T A\mathbf{x}_1 & \dots & S^T A\mathbf{x}_n \end{bmatrix} && \text{because } A\mathbf{x}_0 = \mathbf{0} \\
&= \begin{bmatrix} \mathbf{0} & S^T A\mathbf{x}_1 & \dots & S^T A\mathbf{x}_n \end{bmatrix}.
\end{aligned}$$

We now know that $S^T AS \in \mathbb{R}^{(n+1)\times(n+1)}$ is a symmetric matrix, and that its leftmost column is $\mathbf{0}$.

Theorem 8.7.6

A matrix in $\mathbb{R}^{n \times n}$ is orthogonally diagonalizable iff it is symmetric.

Proof (continued). So, there exists a symmetric matrix $A_0 \in \mathbb{R}^{n \times n}$ s.t.

$$S^T A S = \left[\begin{array}{c:c} 0 & \mathbf{0}^T \\ \hdashline \mathbf{0} & A_0 \end{array}\right].$$

Theorem 8.7.6

A matrix in $\mathbb{R}^{n\times n}$ is orthogonally diagonalizable iff it is symmetric.

Proof (continued). So, there exists a symmetric matrix $A_0 \in \mathbb{R}^{n\times n}$ s.t.

$$S^T A S \quad = \quad \left[\begin{array}{c:c} 0 & \mathbf{0}^T \\ \hdashline \mathbf{0} & A_0 \end{array}\right].$$

By the induction hypothesis, A_0 is orthogonally diagonalizable, i.e. there exist a diagonal matrix D_0 and an orthogonal matrix Q_0, both in $\mathbb{R}^{n\times n}$, s.t. $D_0 = Q_0^T A_0 Q_0$.

Theorem 8.7.6

A matrix in $\mathbb{R}^{n \times n}$ is orthogonally diagonalizable iff it is symmetric.

Proof (continued). So, there exists a symmetric matrix $A_0 \in \mathbb{R}^{n \times n}$ s.t.

$$S^T A S = \left[\begin{array}{c|c} 0 & \mathbf{0}^T \\ \hline \mathbf{0} & A_0 \end{array}\right].$$

By the induction hypothesis, A_0 is orthogonally diagonalizable, i.e. there exist a diagonal matrix D_0 and an orthogonal matrix Q_0, both in $\mathbb{R}^{n \times n}$, s.t. $D_0 = Q_0^T A_0 Q_0$. Now, set

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

Theorem 8.7.6

A matrix in $\mathbb{R}^{n\times n}$ is orthogonally diagonalizable iff it is symmetric.

Proof (continued). So, there exists a symmetric matrix $A_0 \in \mathbb{R}^{n\times n}$ s.t.

$$S^T AS = \left[\begin{array}{c|c} 0 & \mathbf{0}^T \\ \hline \mathbf{0} & A_0 \end{array}\right].$$

By the induction hypothesis, A_0 is orthogonally diagonalizable, i.e. there exist a diagonal matrix D_0 and an orthogonal matrix Q_0, both in $\mathbb{R}^{n\times n}$, s.t. $D_0 = Q_0^T A_0 Q_0$. Now, set

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

Clearly, D is diagonal (because D_0 is diagonal), and R is orthogonal (because Q_0 is orthogonal).

Theorem 8.7.6

A matrix in $\mathbb{R}^{n\times n}$ is orthogonally diagonalizable iff it is symmetric.

Proof (continued). So, there exists a symmetric matrix $A_0 \in \mathbb{R}^{n\times n}$ s.t.

$$S^T A S \;=\; \left[\begin{array}{c|c} 0 & \mathbf{0}^T \\ \hline \mathbf{0} & A_0 \end{array}\right].$$

By the induction hypothesis, A_0 is orthogonally diagonalizable, i.e. there exist a diagonal matrix D_0 and an orthogonal matrix Q_0, both in $\mathbb{R}^{n\times n}$, s.t. $D_0 = Q_0^T A_0 Q_0$. Now, set

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

Clearly, D is diagonal (because D_0 is diagonal), and R is orthogonal (because Q_0 is orthogonal). Since R and S are orthogonal, Proposition 6.8.3 guarantees that $Q := SR$ is also orthogonal.

Theorem 8.7.6

A matrix in $\mathbb{R}^{n\times n}$ is orthogonally diagonalizable iff it is symmetric.

Proof (continued). So, there exists a symmetric matrix $A_0 \in \mathbb{R}^{n\times n}$ s.t.

$$S^TAS = \left[\begin{array}{c|c} 0 & \mathbf{0}^T \\ \hline \mathbf{0} & A_0 \end{array}\right].$$

By the induction hypothesis, A_0 is orthogonally diagonalizable, i.e. there exist a diagonal matrix D_0 and an orthogonal matrix Q_0, both in $\mathbb{R}^{n\times n}$, s.t. $D_0 = Q_0^T A_0 Q_0$. Now, set

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

Clearly, D is diagonal (because D_0 is diagonal), and R is orthogonal (because Q_0 is orthogonal). Since R and S are orthogonal, Proposition 6.8.3 guarantees that $Q := SR$ is also orthogonal. Finally, we compute (next slide):

Theorem 8.7.6

A matrix in $\mathbb{R}^{n\times n}$ is orthogonally diagonalizable iff it is symmetric.

Proof (continued).

$$
\begin{aligned}
Q^T A Q &= (SR)^T A (SR) \\
&= R^T (S^T A S) R \\
&= \left[\begin{array}{c:c} 1 & \mathbf{0}^T \\ \hdashline \mathbf{0} & Q_0^T \end{array}\right] \left[\begin{array}{c:c} 0 & \mathbf{0}^T \\ \hdashline \mathbf{0} & A_0 \end{array}\right] \left[\begin{array}{c:c} 1 & \mathbf{0}^T \\ \hdashline \mathbf{0} & Q_0 \end{array}\right] \\
&= \left[\begin{array}{c:c} 0 & \mathbf{0}^T \\ \hdashline \mathbf{0} & Q_0^T A_0 Q_0 \end{array}\right] \\
&= \left[\begin{array}{c:c} 0 & \mathbf{0}^T \\ \hdashline \mathbf{0} & D_0 \end{array}\right] \\
&= D,
\end{aligned}
$$

and we are done. □

Theorem 8.7.6

A matrix in $\mathbb{R}^{n \times n}$ is orthogonally diagonalizable iff it is symmetric.

Theorem 8.7.6

A matrix in $\mathbb{R}^{n \times n}$ is orthogonally diagonalizable iff it is symmetric.

- By combing Theorem 8.7.6 with what we know about diagonalizability, eigenbases, and orthogonality, we obtain the spectral theorem for symmetric matrices (next slide).
  - The full proof is in the Lecture Notes, but it is essentially just a compilation of the various results we have already seen.

The spectral theorem for symmetric matrices

For every matrix $A \in \mathbb{R}^{n \times n}$, the following are equivalent:

(a) A is symmetric;
(b) A is orthogonally diagonalizable;
(c) $\mathbb{R}^n$ has an orthonormal eigenbasis associated with A;
(d) $\mathbb{R}^n$ has an orthogonal eigenbasis associated with A;
(e) $\mathbb{R}^n$ has an eigenbasis associated with A, and the eigenspaces of A are pairwise orthogonal;
(f) A has n pairwise orthogonal eigenvectors.[a]

[a] This means that some n eigenvectors of A are pairwise orthogonal. It does **not** mean that A has exactly n eigenvectors (which happen to be orthogonal).

- By Theorem 8.7.6, every symmetric matrix in $\mathbb{R}^{n\times n}$ can be orthogonally diagonalized, and in fact, the proof of Theorem 8.7.6 gives us a recipe of sorts for orthogonally diagonalizing such a matrix.

- By Theorem 8.7.6, every symmetric matrix in $\mathbb{R}^{n \times n}$ can be orthogonally diagonalized, and in fact, the proof of Theorem 8.7.6 gives us a recipe of sorts for orthogonally diagonalizing such a matrix.
- However, that recipe is not particularly practical, and we are better off using the spectral theorem instead.

- By Theorem 8.7.6, every symmetric matrix in $\mathbb{R}^{n \times n}$ can be orthogonally diagonalized, and in fact, the proof of Theorem 8.7.6 gives us a recipe of sorts for orthogonally diagonalizing such a matrix.
- However, that recipe is not particularly practical, and we are better off using the spectral theorem instead.
- Suppose we are given a symmetric matrix $A \in \mathbb{R}^{n \times n}$, which we wish to orthogonally diagonalize.

- By Theorem 8.7.6, every symmetric matrix in $\mathbb{R}^{n \times n}$ can be orthogonally diagonalized, and in fact, the proof of Theorem 8.7.6 gives us a recipe of sorts for orthogonally diagonalizing such a matrix.
- However, that recipe is not particularly practical, and we are better off using the spectral theorem instead.
- Suppose we are given a symmetric matrix $A \in \mathbb{R}^{n \times n}$, which we wish to orthogonally diagonalize.
- So, our goal is to construct a diagonal matrix D and an orthogonal matrix Q, both in $\mathbb{R}^{n \times n}$, s.t. $D = Q^T AQ$.

- By Theorem 8.7.6, every symmetric matrix in $\mathbb{R}^{n \times n}$ can be orthogonally diagonalized, and in fact, the proof of Theorem 8.7.6 gives us a recipe of sorts for orthogonally diagonalizing such a matrix.
- However, that recipe is not particularly practical, and we are better off using the spectral theorem instead.
- Suppose we are given a symmetric matrix $A \in \mathbb{R}^{n \times n}$, which we wish to orthogonally diagonalize.
- So, our goal is to construct a diagonal matrix D and an orthogonal matrix Q, both in $\mathbb{R}^{n \times n}$, s.t. $D = Q^T AQ$.
- We proceed as follows (next two slides).

1. First, we compute the characteristic polynomial of A, we factor it, and we find all the (real) eigenvalues of A along with their algebraic multiplicities.

1. First, we compute the characteristic polynomial of A, we factor it, and we find all the (real) eigenvalues of A along with their algebraic multiplicities.
2. Since A is orthogonally diagonalizable (and in particular, diagonalizable), Theorems 8.4.5(d) and 8.5.6 together guarantee that $\mathbb{R}^n$ has an eigenbasis associated with A, and moreover, that the sum of algebraic multiplicities of the eigenvalues of A is n, and that the geometric multiplicity of each eigenvalue is equal to its algebraic multiplicity.

1. First, we compute the characteristic polynomial of A, we factor it, and we find all the (real) eigenvalues of A along with their algebraic multiplicities.
2. Since A is orthogonally diagonalizable (and in particular, diagonalizable), Theorems 8.4.5(d) and 8.5.6 together guarantee that $\mathbb{R}^n$ has an eigenbasis associated with A, and moreover, that the sum of algebraic multiplicities of the eigenvalues of A is n, and that the geometric multiplicity of each eigenvalue is equal to its algebraic multiplicity.
3. Next, for each eigenvalue λ of A, we compute a basis $\mathcal{B}_\lambda$ of the eigenspace $E_\lambda(A)$, and then we apply the Gram-Schmidt orthogonalization process to $\mathcal{B}_\lambda$ in order to obtain an orthonormal basis $\mathcal{C}_\lambda$ of $E_\lambda(A)$.

1. First, we compute the characteristic polynomial of A, we factor it, and we find all the (real) eigenvalues of A along with their algebraic multiplicities.
2. Since A is orthogonally diagonalizable (and in particular, diagonalizable), Theorems 8.4.5(d) and 8.5.6 together guarantee that $\mathbb{R}^n$ has an eigenbasis associated with A, and moreover, that the sum of algebraic multiplicities of the eigenvalues of A is n, and that the geometric multiplicity of each eigenvalue is equal to its algebraic multiplicity.
3. Next, for each eigenvalue λ of A, we compute a basis $\mathcal{B}_\lambda$ of the eigenspace $E_\lambda(A)$, and then we apply the Gram-Schmidt orthogonalization process to $\mathcal{B}_\lambda$ in order to obtain an orthonormal basis $\mathcal{C}_\lambda$ of $E_\lambda(A)$.
4. In view of the spectral theorem, we see that the union $\mathcal{C}$ of the $\mathcal{C}_\lambda$'s is an orthonormal eigenbasis of $\mathbb{R}^n$ associated with A.

5. We now form the diagonal matrix D by placing the eigenvalues of A on the main diagonal of D (while respecting the algebraic/geometric multiplicity of each eigenvalue), and we form Q by arranging the vectors of our orthonormal eigenbasis $\mathcal{C}$ into a matrix (while respecting the order from D). Since the columns of Q form an orthonormal basis of $\mathbb{R}^n$, we see that Q is orthogonal, and so $Q^{-1} = Q^T$. But now Theorem 8.5.6 guarantees that $D = Q^{-1}AQ = Q^TAQ$.

Example 8.7.7

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

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

Solution.

Example 8.7.7

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

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

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

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

Solution.

Solution. Reminder: $p_A(\lambda) = (\lambda + 2)(\lambda - 7)^2$.

Thus, A has two eigenvalues:

- $\lambda_1 = -2$ (with algebraic multiplicity 1),
- and $\lambda_2 = 7$ (with algebraic multiplicity 2).

Solution. Reminder: $p_A(\lambda) = (\lambda + 2)(\lambda - 7)^2$.

Thus, A has two eigenvalues:

- $\lambda_1 = -2$ (with algebraic multiplicity 1),
- and $\lambda_2 = 7$ (with algebraic multiplicity 2).

We now compute a basis $\mathcal{B}_1 = \{\begin{bmatrix} -2 & -1 & 2 \end{bmatrix}^T\}$ of $E_{\lambda_1}(A)$ and a basis $\mathcal{B}_2 = \{\begin{bmatrix} -1 & 2 & 0 \end{bmatrix}^T, \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}^T\}$ of $E_{\lambda_2}(A)$.

Solution. Reminder: $p_A(\lambda) = (\lambda + 2)(\lambda - 7)^2$.

Thus, A has two eigenvalues:

- $\lambda_1 = -2$ (with algebraic multiplicity 1),
- and $\lambda_2 = 7$ (with algebraic multiplicity 2).

We now compute a basis $\mathcal{B}_1 = \{\begin{bmatrix} -2 & -1 & 2 \end{bmatrix}^T\}$ of $E_{\lambda_1}(A)$ and a basis $\mathcal{B}_2 = \{\begin{bmatrix} -1 & 2 & 0 \end{bmatrix}^T, \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}^T\}$ of $E_{\lambda_2}(A)$.

Next, we apply the Gram-Schmidt orthogonalization process to $\mathcal{B}_1$ and $\mathcal{B}_2$.

Solution. Reminder: $p_A(\lambda) = (\lambda + 2)(\lambda - 7)^2$.

Thus, A has two eigenvalues:

- $\lambda_1 = -2$ (with algebraic multiplicity 1),
- and $\lambda_2 = 7$ (with algebraic multiplicity 2).

We now compute a basis $\mathcal{B}_1 = \{\begin{bmatrix} -2 & -1 & 2 \end{bmatrix}^T\}$ of $E_{\lambda_1}(A)$ and a basis $\mathcal{B}_2 = \{\begin{bmatrix} -1 & 2 & 0 \end{bmatrix}^T, \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}^T\}$ of $E_{\lambda_2}(A)$.

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

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

Now D is diagonal, Q is orthogonal, and $D = Q^T AQ$. $\square$