Linear Algebra 2

Lecture #21

Eigenvalues and eigenvectors

Irena Penev

April 17, 2024

- This lecture has four parts:

- This lecture has four parts:
  1. Eigenvalues and eigenvectors of linear functions and square matrices

- This lecture has four parts:
  1. Eigenvalues and eigenvectors of linear functions and square matrices
  2. The characteristic polynomial and spectrum

- This lecture has four parts:
  1. Eigenvalues and eigenvectors of linear functions and square matrices
  2. The characteristic polynomial and spectrum
  3. Eigenvalues and invertibility (plus the Invertible Matrix Theorem, version 4)

- This lecture has four parts:
  1. Eigenvalues and eigenvectors of linear functions and square matrices
  2. The characteristic polynomial and spectrum
  3. Eigenvalues and invertibility (plus the Invertible Matrix Theorem, version 4)
  4. The relationship between algebraic and geometric multiplicities of eigenvalues

1. Eigenvalues and eigenvectors of linear functions and square matrices

1. Eigenvalues and eigenvectors of linear functions and square matrices

Definition

Suppose that V is a vector spaces over a field $\mathbb{F}$, and that $f : V \to V$ is a linear function. An *eigenvector* of f is a vector $\mathbf{v} \in V \setminus \{\mathbf{0}\}$ for which there exists a scalar $\lambda \in \mathbb{F}$, called the *eigenvalue* of f associated with the eigenvector $\mathbf{v}$, s.t.

$$f(\mathbf{v}) = \lambda \mathbf{v}.$$

Under these circumstances, we also say that $\mathbf{v}$ is an eigenvector of f associated with the eigenvalue λ.

1. Eigenvalues and eigenvectors of linear functions and square matrices

Definition

Suppose that V is a vector spaces over a field $\mathbb{F}$, and that $f : V \to V$ is a linear function. An *eigenvector* of f is a vector $\mathbf{v} \in V \setminus \{\mathbf{0}\}$ for which there exists a scalar $\lambda \in \mathbb{F}$, called the *eigenvalue* of f associated with the eigenvector $\mathbf{v}$, s.t.

$$f(\mathbf{v}) = \lambda \mathbf{v}.$$

Under these circumstances, we also say that $\mathbf{v}$ is an eigenvector of f associated with the eigenvalue λ.

- So, the eigenvectors of f are those **non-zero** vectors in V that simply get scaled by f, and the eigenvalues are the scalars that the eigenvectors get scaled by.

1. Eigenvalues and eigenvectors of linear functions and square matrices

Definition

Suppose that V is a vector spaces over a field $\mathbb{F}$, and that $f : V \to V$ is a linear function. An *eigenvector* of f is a vector $\mathbf{v} \in V \setminus \{\mathbf{0}\}$ for which there exists a scalar $\lambda \in \mathbb{F}$, called the *eigenvalue* of f associated with the eigenvector $\mathbf{v}$, s.t.

$$f(\mathbf{v}) = \lambda \mathbf{v}.$$

Under these circumstances, we also say that $\mathbf{v}$ is an eigenvector of f associated with the eigenvalue λ.

- So, the eigenvectors of f are those **non-zero** vectors in V that simply get scaled by f, and the eigenvalues are the scalars that the eigenvectors get scaled by.
- By definition, an eigenvector cannot be $\mathbf{0}$, but an eigenvalue may possibly be 0.

Definition

Suppose that V is a vector spaces over a field $\mathbb{F}$, and that $f : V \to V$ is a linear function. An *eigenvector* of f is a vector $\mathbf{v} \in V \setminus \{\mathbf{0}\}$ for which there exists a scalar $\lambda \in \mathbb{F}$, called the *eigenvalue* of f associated with the eigenvector $\mathbf{v}$, s.t.

$$f(\mathbf{v}) = \lambda \mathbf{v}.$$

Under these circumstances, we also say that $\mathbf{v}$ is an eigenvector of f associated with the eigenvalue λ.

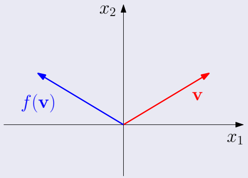
- **Remark:** Note that eigenvectors and eigenvalues are only defined for those linear functions whose domain is the same as the codomain.

Example 8.1.1

Consider the linear function $f : \mathbb{R}^2 \to \mathbb{R}^2$ given by

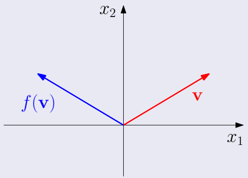
$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_1 \\ x_2 \end{bmatrix}$$

for all $x_1, x_2 \in \mathbb{R}$. So, f is the reflection about the x_2-axis (see the picture below), and its standard matrix is $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$.

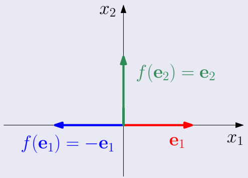


As usual, $\mathbf{e}_1$ and $\mathbf{e}_2$ are the standard basis vectors of $\mathbb{R}^2$. Then (next slide)

Example 8.1.1



- $\mathbf{e}_1$ is an eigenvector of f associated with the eigenvalue $\lambda_1 := -1$, since $f(\mathbf{e}_1) = -\mathbf{e}_1 = \lambda_1 \mathbf{e}_1$;
- $\mathbf{e}_2$ is an eigenvector of f associated with the eigenvalue $\lambda_2 := 1$, since $f(\mathbf{e}_2) = \mathbf{e}_2 = \lambda_2 \mathbf{e}_2$.

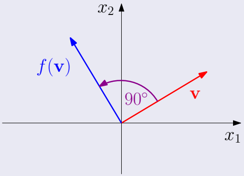


Example 8.1.2

Consider the linear function $f : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix}$$

for all $x_1, x_2 \in \mathbb{R}$. So, f is the counterclockwise rotation by 90° about the origin (see the picture below), and its standard matrix is $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. This function has no eigenvectors (and consequently, it has no eigenvalues), since it does not simply scale any non-zero vector in $\mathbb{R}^2$.



Example 8.1.3

Consider the linear function $f : \mathbb{C}^2 \to \mathbb{C}^2$ given by

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix}$$

for all $x_1, x_2 \in \mathbb{C}$. (This is the same formula as the one from Example 8.1.2, except that we are now working over $\mathbb{C}$, rather than over $\mathbb{R}$.) Then

Example 8.1.3

Consider the linear function $f : \mathbb{C}^2 \to \mathbb{C}^2$ given by

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix}$$

for all $x_1, x_2 \in \mathbb{C}$. (This is the same formula as the one from Example 8.1.2, except that we are now working over $\mathbb{C}$, rather than over $\mathbb{R}$.) Then

- $\mathbf{v}_1 := \begin{bmatrix} i \\ 1 \end{bmatrix}$ is an eigenvalue of f associated with the eigenvalue $\lambda_1 := i$, since $f(\mathbf{v}_1) = \begin{bmatrix} -1 \\ i \end{bmatrix} = i\begin{bmatrix} i \\ 1 \end{bmatrix} = \lambda_1 \mathbf{v}_1$;

Example 8.1.3

Consider the linear function $f : \mathbb{C}^2 \to \mathbb{C}^2$ given by

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix}$$

for all $x_1, x_2 \in \mathbb{C}$. (This is the same formula as the one from Example 8.1.2, except that we are now working over $\mathbb{C}$, rather than over $\mathbb{R}$.) Then

- $\mathbf{v}_1 := \begin{bmatrix} i \\ 1 \end{bmatrix}$ is an eigenvalue of f associated with the eigenvalue $\lambda_1 := i$, since $f(\mathbf{v}_1) = \begin{bmatrix} -1 \\ i \end{bmatrix} = i\begin{bmatrix} i \\ 1 \end{bmatrix} = \lambda_1\mathbf{v}_1$;
- $\mathbf{v}_2 := \begin{bmatrix} -i \\ 1 \end{bmatrix}$ is an eigenvector of f associated with the eigenvalue $\lambda_2 := -i$, since $f(\mathbf{v}_2) = \begin{bmatrix} -1 \\ -i \end{bmatrix} = (-i)\begin{bmatrix} -i \\ 1 \end{bmatrix} = \lambda_2\mathbf{v}_2$.

- Example 8.1.2: $f : \mathbb{R}^2 \to \mathbb{R}^2$, given by

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix} \quad \forall x_1, x_2 \in \mathbb{R}$$

(counterclockwise rotation by 90° about the origin);

- Example 8.1.3: $f : \mathbb{C}^2 \to \mathbb{C}^2$,

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix} \quad \forall x_1, x_2 \in \mathbb{C}.$$

- Example 8.1.2: $f : \mathbb{R}^2 \to \mathbb{R}^2$, given by

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix} \qquad \forall x_1, x_2 \in \mathbb{R}$$

  (counterclockwise rotation by 90° about the origin);
- Example 8.1.3: $f : \mathbb{C}^2 \to \mathbb{C}^2$,

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix} \qquad \forall x_1, x_2 \in \mathbb{C}.$$

- **Remark:** It may be somewhat surprising that the linear function f from Example 8.1.2 has no eigenvectors and no eigenvalues, whereas the one from Example 8.1.3 has them.

- Example 8.1.2: $f : \mathbb{R}^2 \to \mathbb{R}^2$, given by

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix} \qquad \forall x_1, x_2 \in \mathbb{R}$$

(counterclockwise rotation by 90° about the origin);
- Example 8.1.3: $f : \mathbb{C}^2 \to \mathbb{C}^2$,

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix} \qquad \forall x_1, x_2 \in \mathbb{C}.$$

- **Remark:** It may be somewhat surprising that the linear function f from Example 8.1.2 has no eigenvectors and no eigenvalues, whereas the one from Example 8.1.3 has them.
- As we shall see once we learn how to actually compute eigenvalues and eigenvectors (this will involve finding roots of polynomials), the essential difference is that $\mathbb{C}$ is an algebraically closed field, whereas $\mathbb{R}$ is not.

- 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 $\mathbb{F}$.

- 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 $\mathbb{F}$.

- If $\mathbb{F}$ is an algebraically closed field, and $p(x)$ is non-constant polynomial with coefficients in $\mathbb{F}$, then $p(x)$ can be factored into linear terms.

- 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 $\mathbb{F}$.

- If $\mathbb{F}$ is an algebraically closed field, and $p(x)$ is non-constant polynomial with coefficients in $\mathbb{F}$, then $p(x)$ can be factored into linear terms.
- $\mathbb{C}$ is algebraically closed.
- $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{Z}_p$ (where p is a prime number) are **not** algebraically closed.

- For a linear function $f : V \to V$, where V is a vector space over a field $\mathbb{F}$, and for a scalar $\lambda \in \mathbb{F}$, we define

$$E_\lambda(f) \quad := \quad \{\mathbf{v} \in V \mid f(\mathbf{v}) = \lambda\mathbf{v}\}.$$

- For a linear function $f : V \to V$, where V is a vector space over a field $\mathbb{F}$, and for a scalar $\lambda \in \mathbb{F}$, we define

$$E_\lambda(f) \quad := \quad \{\mathbf{v} \in V \mid f(\mathbf{v}) = \lambda \mathbf{v}\}.$$

- Note that $\mathbf{0} \in E_\lambda(f)$, since $f(\mathbf{0}) \stackrel{(*)}{=} \mathbf{0} = \lambda \mathbf{0}$, where (*) follows from Proposition 6.1.4 (since f is linear).

- For a linear function $f : V \to V$, where V is a vector space over a field $\mathbb{F}$, and for a scalar $\lambda \in \mathbb{F}$, we define

$$E_\lambda(f) \quad := \quad \{\mathbf{v} \in V \mid f(\mathbf{v}) = \lambda \mathbf{v}\}.$$

- Note that $\mathbf{0} \in E_\lambda(f)$, since $f(\mathbf{0}) \overset{(*)}{=} \mathbf{0} = \lambda \mathbf{0}$, where (*) follows from Proposition 6.1.4 (since f is linear).
- The set $E_\lambda(f)$ can be defined for any scalar λ, but it is only interesting in the case when λ is an eigenvalue of V, in which case $E_\lambda(f)$ is called the *eigenspace* of f associated with the eigenvalue λ.

- For a linear function $f : V \to V$, where V is a vector space over a field $\mathbb{F}$, and for a scalar $\lambda \in \mathbb{F}$, we define

$$E_\lambda(f) \quad := \quad \{\mathbf{v} \in V \mid f(\mathbf{v}) = \lambda\mathbf{v}\}.$$

- Note that $\mathbf{0} \in E_\lambda(f)$, since $f(\mathbf{0}) \overset{(*)}{=} \mathbf{0} = \lambda\mathbf{0}$, where (*) follows from Proposition 6.1.4 (since f is linear).
- The set $E_\lambda(f)$ can be defined for any scalar λ, but it is only interesting in the case when λ is an eigenvalue of V, in which case $E_\lambda(f)$ is called the *eigenspace* of f associated with the eigenvalue λ.
- Note that, for an eigenvalue λ of f, the elements of the eigenspace $E_\lambda(f)$ are precisely the zero vector and the eigenvectors of f associated with λ.

- For a linear function $f : V \to V$, where V is a vector space over a field $\mathbb{F}$, and for a scalar $\lambda \in \mathbb{F}$, we define

$$E_\lambda(f) \quad := \quad \{\mathbf{v} \in V \mid f(\mathbf{v}) = \lambda\mathbf{v}\}.$$

- Note that $\mathbf{0} \in E_\lambda(f)$, since $f(\mathbf{0}) \stackrel{(*)}{=} \mathbf{0} = \lambda\mathbf{0}$, where (*) follows from Proposition 6.1.4 (since f is linear).
- The set $E_\lambda(f)$ can be defined for any scalar λ, but it is only interesting in the case when λ is an eigenvalue of V, in which case $E_\lambda(f)$ is called the *eigenspace* of f associated with the eigenvalue λ.
- Note that, for an eigenvalue λ of f, the elements of the eigenspace $E_\lambda(f)$ are precisely the zero vector and the eigenvectors of f associated with λ.
  - By definition, $\mathbf{0}$ cannot be an eigenvector.

- For a linear function $f : V \to V$, where V is a vector space over a field $\mathbb{F}$, and for a scalar $\lambda \in \mathbb{F}$, we define

$$E_\lambda(f) \;\; := \;\; \{\mathbf{v} \in V \mid f(\mathbf{v}) = \lambda \mathbf{v}\}.$$

- Note that $\mathbf{0} \in E_\lambda(f)$, since $f(\mathbf{0}) \overset{(*)}{=} \mathbf{0} = \lambda \mathbf{0}$, where (*) follows from Proposition 6.1.4 (since f is linear).
- The set $E_\lambda(f)$ can be defined for any scalar λ, but it is only interesting in the case when λ is an eigenvalue of V, in which case $E_\lambda(f)$ is called the *eigenspace* of f associated with the eigenvalue λ.
- Note that, for an eigenvalue λ of f, the elements of the eigenspace $E_\lambda(f)$ are precisely the zero vector and the eigenvectors of f associated with λ.
  - By definition, $\mathbf{0}$ cannot be an eigenvector.
- On the other hand, if λ is not an eigenvalue of f, then we simply have that $E_\lambda(f) = \{\mathbf{0}\}$, and we do not refer to $E_\lambda(f)$ as an eigenspace.

Proposition 8.1.4

Let V be a vector space over a field $\mathbb{F}$, and let $f : V \to V$ be a linear function. Then both the following hold:

(a) for all scalars $\lambda \in \mathbb{F}$, $E_\lambda(f)$ is a subspace of V, and this subspace is non-trivial (i.e. contains at least one non-zero vector) iff λ is an eigenvalue of f;

(b) for all distinct scalars $\lambda_1, \lambda_2 \in \mathbb{F}$, we have that $E_{\lambda_1}(f) \cap E_{\lambda_2}(f) = \{\mathbf{0}\}$.

Proof (outline). (a) For $\lambda \in \mathbb{F}$:

Proposition 8.1.4

Let V be a vector space over a field $\mathbb{F}$, and let $f : V \to V$ be a linear function. Then both the following hold:

(a) for all scalars $\lambda \in \mathbb{F}$, $E_\lambda(f)$ is a subspace of V, and this subspace is non-trivial (i.e. contains at least one non-zero vector) iff λ is an eigenvalue of f;

(b) for all distinct scalars $\lambda_1, \lambda_2 \in \mathbb{F}$, we have that $E_{\lambda_1}(f) \cap E_{\lambda_2}(f) = \{\mathbf{0}\}$.

Proof (outline). (a) For $\lambda \in \mathbb{F}$:

- we check that $E_\lambda(f)$ contains $\mathbf{0}$ and is closed under vector addition and scalar multiplication, and we deduce (by Theorem 3.1.7) that $E_\lambda(f)$ is a subspace of V;

Proposition 8.1.4

Let V be a vector space over a field $\mathbb{F}$, and let $f : V \to V$ be a linear function. Then both the following hold:

(a) for all scalars $\lambda \in \mathbb{F}$, $E_\lambda(f)$ is a subspace of V, and this subspace is non-trivial (i.e. contains at least one non-zero vector) iff λ is an eigenvalue of f;

(b) for all distinct scalars $\lambda_1, \lambda_2 \in \mathbb{F}$, we have that $E_{\lambda_1}(f) \cap E_{\lambda_2}(f) = \{\mathbf{0}\}$.

Proof (outline). (a) For $\lambda \in \mathbb{F}$:

- we check that $E_\lambda(f)$ contains $\mathbf{0}$ and is closed under vector addition and scalar multiplication, and we deduce (by Theorem 3.1.7) that $E_\lambda(f)$ is a subspace of V;
- any non-zero vector in $E_\lambda(f)$ is an eigenvector of f associated with λ, and so $E_\lambda(f)$ is non-trivial iff λ is an eigenvalue of f.

Proposition 8.1.4

Let V be a vector space over a field $\mathbb{F}$, and let $f : V \to V$ be a linear function. Then both the following hold:

(a) for all scalars $\lambda \in \mathbb{F}$, $E_\lambda(f)$ is a subspace of V, and this subspace is non-trivial (i.e. contains at least one non-zero vector) iff λ is an eigenvalue of f;

(b) for all distinct scalars $\lambda_1, \lambda_2 \in \mathbb{F}$, we have that $E_{\lambda_1}(f) \cap E_{\lambda_2}(f) = \{\mathbf{0}\}$.

Proof (outline, continued). (b) Fix distinct $\lambda_1, \lambda_2 \in \mathbb{F}$.

Proposition 8.1.4

Let V be a vector space over a field $\mathbb{F}$, and let $f : V \to V$ be a linear function. Then both the following hold:

a. for all scalars $\lambda \in \mathbb{F}$, $E_\lambda(f)$ is a subspace of V, and this subspace is non-trivial (i.e. contains at least one non-zero vector) iff λ is an eigenvalue of f;
b. for all distinct scalars $\lambda_1, \lambda_2 \in \mathbb{F}$, we have that $E_{\lambda_1}(f) \cap E_{\lambda_2}(f) = \{\mathbf{0}\}$.

Proof (outline, continued). (b) Fix distinct $\lambda_1, \lambda_2 \in \mathbb{F}$. Obviously, $\mathbf{0} \in E_{\lambda_1}(f) \cap E_{\lambda_2}(f)$.

Proposition 8.1.4

Let V be a vector space over a field $\mathbb{F}$, and let $f : V \to V$ be a linear function. Then both the following hold:

a. for all scalars $\lambda \in \mathbb{F}$, $E_\lambda(f)$ is a subspace of V, and this subspace is non-trivial (i.e. contains at least one non-zero vector) iff λ is an eigenvalue of f;
b. for all distinct scalars $\lambda_1, \lambda_2 \in \mathbb{F}$, we have that $E_{\lambda_1}(f) \cap E_{\lambda_2}(f) = \{\mathbf{0}\}$.

Proof (outline, continued). (b) Fix distinct $\lambda_1, \lambda_2 \in \mathbb{F}$. Obviously, $\mathbf{0} \in E_{\lambda_1}(f) \cap E_{\lambda_2}(f)$. On the other hand, for $\mathbf{v} \in E_{\lambda_1}(f) \cap E_{\lambda_2}(f)$:

$$\begin{aligned} f(\mathbf{v}) &= \lambda_1\mathbf{v} \text{ (because } \mathbf{v} \in E_{\lambda_1}(f)) \text{ and} \\ f(\mathbf{v}) &= \lambda_2\mathbf{v} \text{ (because } \mathbf{v} \in E_{\lambda_2}(f)) \end{aligned}$$

$$\implies \quad \lambda_1\mathbf{v} = \lambda_2\mathbf{v} \implies \underbrace{(\lambda_1 - \lambda_2)}_{\neq 0}\mathbf{v} = \mathbf{0} \implies \mathbf{v} = \mathbf{0}.$$

So, $E_{\lambda_1}(f) \cap E_{\lambda_2}(f) = \{\mathbf{0}\}$. $\square$

Proposition 8.1.4

Let V be a vector space over a field $\mathbb{F}$, and let $f : V \to V$ be a linear function. Then both the following hold:

(a) for all scalars $\lambda \in \mathbb{F}$, $E_\lambda(f)$ is a subspace of V, and this subspace is non-trivial (i.e. contains at least one non-zero vector) iff λ is an eigenvalue of f;

(b) for all distinct scalars $\lambda_1, \lambda_2 \in \mathbb{F}$, we have that $E_{\lambda_1}(f) \cap E_{\lambda_2}(f) = \{\mathbf{0}\}$.

- **Terminology:** Suppose that V is a vector space over a field $\mathbb{F}$, and that λ is an eigenvalue of a linear function $f : V \to V$.
  - The *geometric multiplicity* of the eigenvalue λ is defined to be $\dim\big(E_\lambda(f)\big)$.
  - So, the geometric multiplicity of an eigenvalue is the dimension of the associated eigenspace.

Definition

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$ be a **square** matrix. An *eigenvector* of A is a vector $\mathbf{v} \in \mathbb{F}^n \setminus \{\mathbf{0}\}$ for which there exists a scalar $\lambda \in \mathbb{F}$, called the *eigenvalue* of A associated with the eigenvector $\mathbf{v}$, s.t.

$$A\mathbf{v} = \lambda\mathbf{v}.$$

Under these circumstances, we also say that $\mathbf{v}$ is an eigenvector of A associated with the eigenvalue λ.

- Eigenvectors are, by definition, non-zero, whereas eigenvalues may possibly be zero.

- For a square matrix $A \in \mathbb{F}^{n \times n}$ (where $\mathbb{F}$ is some field), and for a scalar $\lambda \in \mathbb{F}$, we define

$$E_\lambda(A) \quad := \quad \{\mathbf{v} \in \mathbb{F}^n \mid A\mathbf{v} = \lambda \mathbf{v}\}.$$

If λ is an eigenvalue of A, then $E_\lambda(A)$ is called the *eigenspace* of A associated with the eigenvalue λ.

- For a square matrix $A \in \mathbb{F}^{n \times n}$ (where $\mathbb{F}$ is some field), and for a scalar $\lambda \in \mathbb{F}$, we define

$$E_\lambda(A) \quad := \quad \{\mathbf{v} \in \mathbb{F}^n \mid A\mathbf{v} = \lambda\mathbf{v}\}.$$

  If λ is an eigenvalue of A, then $E_\lambda(A)$ is called the *eigenspace* of A associated with the eigenvalue λ.
  - Note that, for an eigenvalue λ of A, the elements of the eigenspace $E_\lambda(A)$ are precisely the zero vector and the eigenvectors of A associated with λ.

- For a square matrix $A \in \mathbb{F}^{n \times n}$ (where $\mathbb{F}$ is some field), and for a scalar $\lambda \in \mathbb{F}$, we define

$$E_\lambda(A) \quad := \quad \{\mathbf{v} \in \mathbb{F}^n \mid A\mathbf{v} = \lambda\mathbf{v}\}.$$

  If λ is an eigenvalue of A, then $E_\lambda(A)$ is called the *eigenspace* of A associated with the eigenvalue λ.
  - Note that, for an eigenvalue λ of A, the elements of the eigenspace $E_\lambda(A)$ are precisely the zero vector and the eigenvectors of A associated with λ.
  - On the other hand, if λ is not an eigenvalue of A, then we simply have that $E_\lambda(A) = \{\mathbf{0}\}$, and we do not refer to $E_\lambda(A)$ as an eigenspace.

Proposition 8.1.5

Let $\mathbb{F}$ be a field, let $f : \mathbb{F}^n \to \mathbb{F}^n$ be a linear function, and let A be the standard matrix of f. Then f and A have exactly the same eigenvalues and the associated eigenectors. Moreover, for all eigenvalues λ of f and A, we have that $E_\lambda(f) = E_\lambda(A)$.

Proof. This follows immediately from the appropriate definitions. □

- Reminder:

Proposition 8.1.4

Let V be a vector space over a field $\mathbb{F}$, and let $f : V \to V$ be a linear function. Then both the following hold:

(a) for all scalars $\lambda \in \mathbb{F}$, $E_\lambda(f)$ is a subspace of V, and this subspace is non-trivial (i.e. contains at least one non-zero vector) iff λ is an eigenvalue of f;

(b) for all distinct scalars $\lambda_1, \lambda_2 \in \mathbb{F}$, we have that $E_{\lambda_1}(f) \cap E_{\lambda_2}(f) = \{\mathbf{0}\}$.

- Reminder:

Proposition 8.1.4

Let V be a vector space over a field $\mathbb{F}$, and let $f : V \to V$ be a linear function. Then both the following hold:

(a) for all scalars $\lambda \in \mathbb{F}$, $E_\lambda(f)$ is a subspace of V, and this subspace is non-trivial (i.e. contains at least one non-zero vector) iff λ is an eigenvalue of f;

(b) for all distinct scalars $\lambda_1, \lambda_2 \in \mathbb{F}$, we have that $E_{\lambda_1}(f) \cap E_{\lambda_2}(f) = \{\mathbf{0}\}$.

- For square matrices, we have the following analog of Proposition 8.1.4.

Proposition 8.1.6

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$ be a square matrix. Then all the following hold:

(a) for all scalars $\lambda \in \mathbb{F}$, $E_\lambda(A)$ is a subspace of $\mathbb{F}^n$, and this subspace is non-trivial (i.e. contains at least one non-zero vector) iff λ is an eigenvalue of A;

(b) for all distinct scalars $\lambda_1, \lambda_2 \in \mathbb{F}$, we have that $E_{\lambda_1}(A) \cap E_{\lambda_2}(A) = \{\mathbf{0}\}$.

Proof.

Proposition 8.1.6

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$ be a square matrix. Then all the following hold:

- (a) for all scalars $\lambda \in \mathbb{F}$, $E_\lambda(A)$ is a subspace of $\mathbb{F}^n$, and this subspace is non-trivial (i.e. contains at least one non-zero vector) iff λ is an eigenvalue of A;
- (b) for all distinct scalars $\lambda_1, \lambda_2 \in \mathbb{F}$, we have that $E_{\lambda_1}(A) \cap E_{\lambda_2}(A) = \{\mathbf{0}\}$.

Proof. Consider the function $f_A : \mathbb{F}^n \to \mathbb{F}^n$, given by $f_A(\mathbf{v}) = A\mathbf{v}$ for all vectors $\mathbf{v} \in \mathbb{F}^n$. Then f_A is linear (by Proposition 1.10.4), and moreover, A is the standard matrix of f_A.

So, by Proposition 8.1.5, we have that for all $\lambda \in \mathbb{F}$, $E_\lambda(A) = E_\lambda(f_A)$.

The result now follows immediately from Proposition 8.1.4. $\square$

Proposition 8.1.6

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$ be a square matrix. Then all the following hold:

(a) for all scalars $\lambda \in \mathbb{F}$, $E_\lambda(A)$ is a subspace of $\mathbb{F}^n$, and this subspace is non-trivial (i.e. contains at least one non-zero vector) iff λ is an eigenvalue of A;

(b) for all distinct scalars $\lambda_1, \lambda_2 \in \mathbb{F}$, we have that $E_{\lambda_1}(A) \cap E_{\lambda_2}(A) = \{\mathbf{0}\}$.

- **Terminology:** Suppose that $\mathbb{F}$ is a field, and that λ is an eigenvalue of a square matrix $A \in \mathbb{F}^{n \times n}$.
  - The *geometric multiplicity* of the eigenvalue λ is defined to be $\dim\big(E_\lambda(A)\big)$.
  - So, the geometric multiplicity of an eigenvalue is the dimension of the associated eigenspace.

Proposition 8.1.7

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ be a basis of V, and let $f : V \to V$ be a linear function. Then for all $\lambda \in \mathbb{F}$, we have that

$$E_\lambda\Big({}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}} \Big) \ = \ \Big\{ \big[\ \mathbf{v}\ \big]_{\mathcal{B}} \mid \mathbf{v} \in E_\lambda(f) \Big\}.$$

Consequently, the linear function f and the matrix ${}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}}$ have exactly the same eigenvalues, with exactly the same corresponding geometric multiplicities.

- Proof: Lecture Notes.
- Proposition 8.1.7 states that $E_\lambda\Big({}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}} \Big)$ is the image of $E_\lambda(f)$ under the coordinate transformation $\big[\ \cdot\ \big]_{\mathcal{B}}$.

- In view of Propositions 8.1.5 ("linear functions and their standard matrices have the same eigenvalues, eigenvectors, and eigenspaces") and 8.1.7 (previous slide), we see that the study of eigenvalues and eigenvectors of linear functions from a non-trivial, finite-dimensional vector space to itself is essentially equivalent to the study of eigenvalues and eigenvectors of square matrices.
  - The computational tools that we develop for finding eigenvectors and eigenvalues will primarily be for square matrices.
  - On the other hand, some of the theoretical results that we prove will be for linear functions instead, and we will obtain corresponding results for matrices as more or less immediate corollaries.

2. The characteristic polynomial and spectrum

Definition

Given a field $\mathbb{F}$ and a matrix $A \in \mathbb{F}^{n\times n}$, the *characteristic polynomial* of A is defined to be

$$p_A(\lambda) \quad := \quad \det(\lambda I_n - A).$$

The *characteristic equation* of A is the equation

$$\det(\lambda I_n - A) \quad = \quad 0.$$

So, the roots of the characteristic polynomial of A are precisely the solutions of the characteristic equation of A.

Example 8.2.1

Compute the characteristic polynomial of the following matrix in $\mathbb{C}^{3\times 3}$:

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

Solution.

Example 8.2.1

Compute the characteristic polynomial of the following matrix in $\mathbb{C}^{3\times 3}$:

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

Solution. The characteristic polynomial of A is:

$$\begin{aligned} p_A(\lambda) = \det(\lambda I_3 - A) &= \begin{vmatrix} \lambda - 1 & 2 & -3 \\ 1 & \lambda & -2 \\ -2 & 1 & \lambda + 3 \end{vmatrix} \\ &= \lambda^3 + 2\lambda^2 - 9\lambda - 3. \end{aligned}$$

□

- **Remark:** For a field $\mathbb{F}$ and a matrix $A \in \mathbb{F}^{n \times n}$, the characteristic polynomial $p_A(\lambda) = \det(\lambda I_n - A)$ is a polynomial of degree n, with leading coefficient 1, i.e. the coefficient in front of λ^n in $p_A(\lambda)$ is 1.

- **Remark:** For a field $\mathbb{F}$ and a matrix $A \in \mathbb{F}^{n \times n}$, the characteristic polynomial $p_A(\lambda) = \det(\lambda I_n - A)$ is a polynomial of degree n, with leading coefficient 1, i.e. the coefficient in front of λ^n in $p_A(\lambda)$ is 1.
    - In some texts, the characteristic polynomial is defined to be $\det(A - \lambda I_n)$.

- **Remark:** For a field $\mathbb{F}$ and a matrix $A \in \mathbb{F}^{n \times n}$, the characteristic polynomial $p_A(\lambda) = \det(\lambda I_n - A)$ is a polynomial of degree n, with leading coefficient 1, i.e. the coefficient in front of λ^n in $p_A(\lambda)$ is 1.
  - In some texts, the characteristic polynomial is defined to be $\det(A - \lambda I_n)$.
  - By Proposition 7.2.3, we have that $\det(A - \lambda I_n) = (-1)^n \det(\lambda I_n - A)$, and so the polynomials $\det(\lambda I_n - A)$ and $\det(A - \lambda I_n)$ have exactly the same roots, with the same corresponding multiplicities, which is what we will actually care about when it comes to the characteristic polynomial.

- **Remark:** For a field $\mathbb{F}$ and a matrix $A \in \mathbb{F}^{n \times n}$, the characteristic polynomial $p_A(\lambda) = \det(\lambda I_n - A)$ is a polynomial of degree n, with leading coefficient 1, i.e. the coefficient in front of λ^n in $p_A(\lambda)$ is 1.
  - In some texts, the characteristic polynomial is defined to be $\det(A - \lambda I_n)$.
  - By Proposition 7.2.3, we have that $\det(A - \lambda I_n) = (-1)^n \det(\lambda I_n - A)$, and so the polynomials $\det(\lambda I_n - A)$ and $\det(A - \lambda I_n)$ have exactly the same roots, with the same corresponding multiplicities, which is what we will actually care about when it comes to the characteristic polynomial.
  - The main advantage of using $\det(\lambda I_n - A)$ rather than $\det(A - \lambda I_n)$ is that the former polynomial has leading coefficient 1, whereas the latter has leading coefficient $(-1)^n$, which is -1 if n is odd.

Theorem 8.2.2

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n\times n}$, and let $\lambda_0 \in \mathbb{F}$. Then

$$E_{\lambda_0}(A) \;=\; \mathrm{Nul}(\lambda_0 I_n - A) \;=\; \mathrm{Nul}(A - \lambda_0 I_n).$$

Moreover, the following are equivalent:

1. λ_0 is an eigenvalue of A;
2. λ_0 is a root of the characteristic polynomial of A, i.e. $p_A(\lambda_0) = 0$;
3. λ_0 is a solution of the characteristic equation of A, i.e. $\det(\lambda_0 I_n - A) = 0$.

Proof.

Theorem 8.2.2

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n\times n}$, and let $\lambda_0 \in \mathbb{F}$. Then

$$E_{\lambda_0}(A) \;=\; \text{Nul}(\lambda_0 I_n - A) \;=\; \text{Nul}(A - \lambda_0 I_n).$$

Moreover, the following are equivalent:

1. λ_0 is an eigenvalue of A;
2. λ_0 is a root of the characteristic polynomial of A, i.e. $p_A(\lambda_0) = 0$;
3. λ_0 is a solution of the characteristic equation of A, i.e. $\det(\lambda_0 I_n - A) = 0$.

Proof. Obviously, for all $\mathbf{v} \in \mathbb{F}^n$, we have that $(\lambda_0 I_n - A)\mathbf{v} = \mathbf{0}$ iff $(A - \lambda_0 I_n)\mathbf{v} = \mathbf{0}$. So, $\text{Nul}(\lambda_0 I_n - A) = \text{Nul}(A - \lambda_0 I_n)$.

Theorem 8.2.2

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n\times n}$, and let $\lambda_0 \in \mathbb{F}$. Then

$$E_{\lambda_0}(A) \;=\; \text{Nul}(\lambda_0 I_n - A) \;=\; \text{Nul}(A - \lambda_0 I_n).$$

Moreover, the following are equivalent:

1. λ_0 is an eigenvalue of A;
2. λ_0 is a root of the characteristic polynomial of A, i.e. $p_A(\lambda_0) = 0$;
3. λ_0 is a solution of the characteristic equation of A, i.e. $\det(\lambda_0 I_n - A) = 0$.

Proof (continued). Further, we compute (next slide):

$$\begin{aligned} E_{\lambda_0}(A) &= \{\mathbf{v} \in \mathbb{F}^n \mid A\mathbf{v} = \lambda_0\mathbf{v}\} \\ &= \{\mathbf{v} \in \mathbb{F}^n \mid A\mathbf{v} = \lambda_0 I_n\mathbf{v}\} \\ &= \{\mathbf{v} \in \mathbb{F}^n \mid (\lambda_0 I_n - A)\mathbf{v} = \mathbf{0}\} \\ &= \text{Nul}(\lambda_0 I_n - A). \end{aligned}$$

Theorem 8.2.2

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n \times n}$, and let $\lambda_0 \in \mathbb{F}$. Then

$$E_{\lambda_0}(A) \;=\; \text{Nul}(\lambda_0 I_n - A) \;=\; \text{Nul}(A - \lambda_0 I_n).$$

Moreover, the following are equivalent:

1. λ_0 is an eigenvalue of A;
2. λ_0 is a root of the characteristic polynomial of A, i.e. $p_A(\lambda_0) = 0$;
3. λ_0 is a solution of the characteristic equation of A, i.e. $\det(\lambda_0 I_n - A) = 0$.

Proof (continued). We have now shown that

$$E_{\lambda_0}(A) \;=\; \text{Nul}(\lambda_0 I_n - A) \;=\; \text{Nul}(A - \lambda_0 I_n).$$

Theorem 8.2.2

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n \times n}$, and let $\lambda_0 \in \mathbb{F}$. Then

$$E_{\lambda_0}(A) \quad = \quad \text{Nul}(\lambda_0 I_n - A) \quad = \quad \text{Nul}(A - \lambda_0 I_n).$$

Moreover, the following are equivalent:

1. λ_0 is an eigenvalue of A;
2. λ_0 is a root of the characteristic polynomial of A, i.e. $p_A(\lambda_0) = 0$;
3. λ_0 is a solution of the characteristic equation of A, i.e. $\det(\lambda_0 I_n - A) = 0$.

Proof (continued). We have now shown that

$$E_{\lambda_0}(A) \quad = \quad \text{Nul}(\lambda_0 I_n - A) \quad = \quad \text{Nul}(A - \lambda_0 I_n).$$

It remains to show that (1), (2), and (3) are equivalent. The fact that (2) and (3) are equivalent follows immediately from the appropriate definitions. It remains to prove that (1) and (3) are equivalent.

Theorem 8.2.2

Moreover, the following are equivalent:

1. λ_0 is an eigenvalue of A;
3. λ_0 is a solution of the characteristic equation of A, i.e. $\det(\lambda_0 I_n - A) = 0$.

Proof (continued). Reminder: $E_{\lambda_0}(A) = \mathrm{Nul}(\lambda_0 I_n - A)$.

$$\underbrace{\lambda_0 \text{ is an eigenvalue of } A}_{(1)} \overset{\text{Prop. 8.1.6}}{\iff} E_{\lambda_0}(A) \neq \{\mathbf{0}\}$$

$$\iff \underbrace{\mathrm{Nul}(\lambda_0 I_n - A)}_{=E_{\lambda_0}(A)} \neq \{\mathbf{0}\}$$

$$\overset{\text{IMT}}{\iff} \text{the matrix } \lambda_0 I_n - A \text{ is not invertible}$$

$$\overset{\text{IMT}}{\iff} \underbrace{\det(\lambda_0 I_n - A) = 0}_{(3)}$$

□

Theorem 8.2.2

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n \times n}$, and let $\lambda_0 \in \mathbb{F}$. Then

$$E_{\lambda_0}(A) \;=\; \mathrm{Nul}(\lambda_0 I_n - A) \;=\; \mathrm{Nul}(A - \lambda_0 I_n).$$

Moreover, the following are equivalent:

1. λ_0 is an eigenvalue of A;
2. λ_0 is a root of the characteristic polynomial of A, i.e. $p_A(\lambda_0) = 0$;
3. λ_0 is a solution of the characteristic equation of A, i.e. $\det(\lambda_0 I_n - A) = 0$.

- By Theorem 8.2.2, the eigenvalues of a square matrix are precisely the roots of its characteristic polynomial.

- By Theorem 8.2.2, the eigenvalues of a square matrix are precisely the roots of its characteristic polynomial.
- For a field $\mathbb{F}$, a matrix $A \in \mathbb{F}^{n \times n}$, and an eigenvalue λ_0 of A, the *algebraic multiplicity* of the eigenvalue λ_0 is its multiplicity as a root of the characteristic polynomial of A, or in other words, it is the largest integer k such that $(\lambda - \lambda_0)^k \mid p_A(\lambda)$, i.e. such that $(\lambda - \lambda_0)^k$ divides the polynomial $p_A(\lambda)$.

- By Theorem 8.2.2, the eigenvalues of a square matrix are precisely the roots of its characteristic polynomial.
- For a field $\mathbb{F}$, a matrix $A \in \mathbb{F}^{n \times n}$, and an eigenvalue λ_0 of A, the *algebraic multiplicity* of the eigenvalue λ_0 is its multiplicity as a root of the characteristic polynomial of A, or in other words, it is the largest integer k such that $(\lambda - \lambda_0)^k \mid p_A(\lambda)$, i.e. such that $(\lambda - \lambda_0)^k$ divides the polynomial $p_A(\lambda)$.
- Since $\deg(p_A(\lambda)) = n$, the sum of algebraic multiplicities of the eigenvalues of the matrix $A \in \mathbb{F}^{n \times n}$ is at most n; if the field $\mathbb{F}$ is algebraically closed, then the sum of algebraic multiplicities of the eigenvalues of A is exactly n.

- By Theorem 8.2.2, the eigenvalues of a square matrix are precisely the roots of its characteristic polynomial.
- For a field $\mathbb{F}$, a matrix $A \in \mathbb{F}^{n \times n}$, and an eigenvalue λ_0 of A, the *algebraic multiplicity* of the eigenvalue λ_0 is its multiplicity as a root of the characteristic polynomial of A, or in other words, it is the largest integer k such that $(\lambda - \lambda_0)^k \mid p_A(\lambda)$, i.e. such that $(\lambda - \lambda_0)^k$ divides the polynomial $p_A(\lambda)$.
- Since $\deg(p_A(\lambda)) = n$, the sum of algebraic multiplicities of the eigenvalues of the matrix $A \in \mathbb{F}^{n \times n}$ is at most n; if the field $\mathbb{F}$ is algebraically closed, then the sum of algebraic multiplicities of the eigenvalues of A is exactly n.
  - Indeed, if $\mathbb{F}$ is algebraically closed, then the characteristic polynomial $p_A(\lambda)$ can be written as a product of linear factors, and there are n of those factors.

- By Theorem 8.2.2, the eigenvalues of a square matrix are precisely the roots of its characteristic polynomial.
- For a field $\mathbb{F}$, a matrix $A \in \mathbb{F}^{n \times n}$, and an eigenvalue λ_0 of A, the *algebraic multiplicity* of the eigenvalue λ_0 is its multiplicity as a root of the characteristic polynomial of A, or in other words, it is the largest integer k such that $(\lambda - \lambda_0)^k \mid p_A(\lambda)$, i.e. such that $(\lambda - \lambda_0)^k$ divides the polynomial $p_A(\lambda)$.
- Since $\deg(p_A(\lambda)) = n$, the sum of algebraic multiplicities of the eigenvalues of the matrix $A \in \mathbb{F}^{n \times n}$ is at most n; if the field $\mathbb{F}$ is algebraically closed, then the sum of algebraic multiplicities of the eigenvalues of A is exactly n.
    - Indeed, if $\mathbb{F}$ is algebraically closed, then the characteristic polynomial $p_A(\lambda)$ can be written as a product of linear factors, and there are n of those factors.
    - If $\mathbb{F}$ is not algebraically closed, we might or might not be able to factor $p_A(\lambda)$ in this way, which is why the sum of algebraic multiplicities of the eigenvalues of A is at most n (possibly strictly smaller than n).

Theorem 8.2.3

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$. Then the geometric multiplicity of any eigenvalue of A is no greater than the algebraic multiplicity of that eigenvalue.

- Proof: Later!
- Schematically, Theorem 8.2.3 states that for an eigenvalue λ of A:

$$\text{geometric multiplicity of } \lambda \quad \leq \quad \text{algebraic multiplicity of } \lambda.$$

Theorem 8.2.3

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$. Then the geometric multiplicity of any eigenvalue of A is no greater than the algebraic multiplicity of that eigenvalue.

- Proof: Later!
- Schematically, Theorem 8.2.3 states that for an eigenvalue λ of A:

$$\text{geometric multiplicity of } \lambda \quad \leq \quad \text{algebraic multiplicity of } \lambda.$$

- For now, we have only stated Theorem 8.2.3. We will not use this theorem before proving it.

- The *spectrum* of a square matrix $A \in \mathbb{F}^{n \times n}$ is the multiset of all eigenvalues of A, with algebraic multiplicities taken into account.
    - This means that the number of times that an eigenvalue appears in the spectrum is equal to the algebraic multiplicity of that eigenvalue. The order in which we list the eigenvalues in the spectrum does not matter, but repetitions do matter.

- The *spectrum* of a square matrix $A \in \mathbb{F}^{n\times n}$ is the multiset of all eigenvalues of A, with algebraic multiplicities taken into account.
    - This means that the number of times that an eigenvalue appears in the spectrum is equal to the algebraic multiplicity of that eigenvalue. The order in which we list the eigenvalues in the spectrum does not matter, but repetitions do matter.
- For example, if a matrix $A \in \mathbb{C}^{5\times 5}$ has eigenvalues 1 (with algeraic multiplicity 1), $1+i$ (with algebraic multiplicity 2), and $1-i$ (with algebraic multiplicity 2), then the spectrum of A is $\{1, 1+i, 1+i, 1-i, 1-i\}$.

- The *spectrum* of a square matrix $A \in \mathbb{F}^{n \times n}$ is the multiset of all eigenvalues of A, with algebraic multiplicities taken into account.
    - This means that the number of times that an eigenvalue appears in the spectrum is equal to the algebraic multiplicity of that eigenvalue. The order in which we list the eigenvalues in the spectrum does not matter, but repetitions do matter.
- For example, if a matrix $A \in \mathbb{C}^{5 \times 5}$ has eigenvalues 1 (with algeraic multiplicity 1), $1 + i$ (with algebraic multiplicity 2), and $1 - i$ (with algebraic multiplicity 2), then the spectrum of A is $\{1, 1+i, 1+i, 1-i, 1-i\}$.
- In general, the spectrum of a matrix $A \in \mathbb{F}^{n \times n}$ (where $\mathbb{F}$ is a field) has at most n elements; if the field $\mathbb{F}$ is algebraically closed, then the spectrum of A has exactly n elements.

Example 8.2.4

Consider the following matrix in $\mathbb{C}^{3\times 3}$:

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

(a) Compute the characteristic polynomial $p_A(\lambda)$ of the matrix A.

(b) Compute all the eigenvalues of A and their algebraic multiplicities, and compute the spectrum of A.

(c) For each eigenvalue λ of A, compute a basis of the eigenspace $E_\lambda(A)$ and specify the geometric multiplicity of the eigenvalue λ.

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

Solution. (a) The characteristic polynomial of A is:

$$\begin{aligned} p_A(\lambda) &= \det(\lambda I_3 - A) \\ &= \begin{vmatrix} \lambda - 4 & 0 & 2 \\ -2 & \lambda - 5 & -4 \\ 0 & 0 & \lambda - 5 \end{vmatrix} \\ &= (\lambda - 4)(\lambda - 5)^2 \qquad \text{via Laplace expansion along 2nd column} \\ &= \lambda^3 - 14\lambda^2 + 65\lambda - 100. \end{aligned}$$

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

Solution. (a) The characteristic polynomial of A is:

$$\begin{aligned} p_A(\lambda) &= \det(\lambda I_3 - A) \\ &= \begin{vmatrix} \lambda - 4 & 0 & 2 \\ -2 & \lambda - 5 & -4 \\ 0 & 0 & \lambda - 5 \end{vmatrix} \\ &= (\lambda - 4)(\lambda - 5)^2 \qquad \text{via Laplace expansion along 2nd column} \\ &= \lambda^3 - 14\lambda^2 + 65\lambda - 100. \end{aligned}$$

- **Remark:** We did not really need to expand in the last line.
  - We only really care about the roots of the characteristic polynomial, and it is more convenient to have a form that is already factored.
  - So, $p_A(\lambda) = (\lambda - 4)(\lambda - 5)^2$ is a "better" answer than $p_A(\lambda) = \lambda^3 - 14\lambda^2 + 65\lambda - 100$, although they are both correct.

Example 8.2.4

Consider the following matrix in $\mathbb{C}^{3\times 3}$:

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

(a) Compute the characteristic polynomial $p_A(\lambda)$ of the matrix A.

(b) Compute all the eigenvalues of A and their algebraic multiplicities, and compute the spectrum of A.

(c) For each eigenvalue λ of A, compute a basis of the eigenspace $E_\lambda(A)$ and specify the geometric multiplicity of the eigenvalue λ.

Solution (continued). Reminder: (a) $p_A(\lambda) = (\lambda - 4)(\lambda - 5)^2$.

Example 8.2.4

Consider the following matrix in $\mathbb{C}^{3\times 3}$:

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

(a) Compute the characteristic polynomial $p_A(\lambda)$ of the matrix A.

(b) Compute all the eigenvalues of A and their algebraic multiplicities, and compute the spectrum of A.

(c) For each eigenvalue λ of A, compute a basis of the eigenspace $E_\lambda(A)$ and specify the geometric multiplicity of the eigenvalue λ.

Solution (continued). Reminder: (a) $p_A(\lambda) = (\lambda - 4)(\lambda - 5)^2$.

(b) From part (a), we see that A has two eigenvalues, namely, the eigenvalue $\lambda_1 = 4$ (with algebraic multiplicity 1), and the eigenvalue $\lambda_2 = 5$ (with algebraic multiplicity 2). So, the spectrum of A is $\{4, 5, 5\}$.

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

Solution (continued). Reminder: the eigenvalues of A are $\lambda_1 = 4$ and $\lambda_2 = 5$.

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

Solution (continued). Reminder: the eigenvalues of A are $\lambda_1 = 4$ and $\lambda_2 = 5$.

(c) For each $i \in \{1, 2\}$, we have that

$$E_{\lambda_i}(A) \;\; = \;\; \mathrm{Nul}(\lambda_i I_3 - A),$$

which is precisely the set of all solutions of the characteristic equation

$$(\lambda_i I_3 - A)\mathbf{x} \;\; = \;\; \mathbf{0}.$$

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

Solution (continued). Reminder: the eigenvalues of A are $\lambda_1 = 4$ and $\lambda_2 = 5$.

(c) For each $i \in \{1, 2\}$, we have that

$$E_{\lambda_i}(A) \quad = \quad \text{Nul}(\lambda_i I_3 - A),$$

which is precisely the set of all solutions of the characteristic equation

$$(\lambda_i I_3 - A)\mathbf{x} \quad = \quad \mathbf{0}.$$

Let us now compute a basis of each of the two eigenspaces.

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

Solution (continued). (c) For $\lambda_1 = 4$, we have that

$$\lambda_1 I_3 - A = \begin{bmatrix} \lambda_1 - 4 & 0 & 2 \\ -2 & \lambda_1 - 5 & -4 \\ 0 & 0 & \lambda_1 - 5 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 2 \\ -2 & -1 & -4 \\ 0 & 0 & -1 \end{bmatrix},$$

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

Solution (continued). (c) For $\lambda_1 = 4$, we have that

$$\lambda_1 I_3 - A = \begin{bmatrix} \lambda_1 - 4 & 0 & 2 \\ -2 & \lambda_1 - 5 & -4 \\ 0 & 0 & \lambda_1 - 5 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 2 \\ -2 & -1 & -4 \\ 0 & 0 & -1 \end{bmatrix},$$

and that

$$\text{RREF}(\lambda_1 I_3 - A) = \begin{bmatrix} 1 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

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

Solution (continued). (c) For $\lambda_1 = 4$, we have that

$$\lambda_1 I_3 - A = \begin{bmatrix} \lambda_1 - 4 & 0 & 2 \\ -2 & \lambda_1 - 5 & -4 \\ 0 & 0 & \lambda_1 - 5 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 2 \\ -2 & -1 & -4 \\ 0 & 0 & -1 \end{bmatrix},$$

and that

$$\text{RREF}(\lambda_1 I_3 - A) = \begin{bmatrix} 1 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Consequently, the general solution of the equation $(\lambda_1 I_3 - A)\mathbf{x} = \mathbf{0}$ is

$$\mathbf{x} = \begin{bmatrix} -t/2 \\ t \\ 0 \end{bmatrix} = t \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix} = \tfrac{t}{2} \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}, \qquad \text{with } t \in \mathbb{C}.$$

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

Solution (continued). (c) For $\lambda_1 = 4$, we have that

$$\lambda_1 I_3 - A = \begin{bmatrix} \lambda_1 - 4 & 0 & 2 \\ -2 & \lambda_1 - 5 & -4 \\ 0 & 0 & \lambda_1 - 5 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 2 \\ -2 & -1 & -4 \\ 0 & 0 & -1 \end{bmatrix},$$

and that

$$\mathrm{RREF}(\lambda_1 I_3 - A) = \begin{bmatrix} 1 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Consequently, the general solution of the equation $(\lambda_1 I_3 - A)\mathbf{x} = \mathbf{0}$ is

$$\mathbf{x} = \begin{bmatrix} -t/2 \\ t \\ 0 \end{bmatrix} = t \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix} = \tfrac{t}{2} \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}, \qquad \text{with } t \in \mathbb{C}.$$

So, $\left\{ \begin{bmatrix} -1 & 2 & 0 \end{bmatrix}^T \right\}$ is a basis of the eigespace $E_{\lambda_1}(A) = \mathrm{Nul}(A - \lambda_1 I_n)$,

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

Solution (continued). (c) For $\lambda_1 = 4$, we have that

$$\lambda_1 I_3 - A = \begin{bmatrix} \lambda_1 - 4 & 0 & 2 \\ -2 & \lambda_1 - 5 & -4 \\ 0 & 0 & \lambda_1 - 5 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 2 \\ -2 & -1 & -4 \\ 0 & 0 & -1 \end{bmatrix},$$

and that

$$\text{RREF}(\lambda_1 I_3 - A) = \begin{bmatrix} 1 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Consequently, the general solution of the equation $(\lambda_1 I_3 - A)\mathbf{x} = \mathbf{0}$ is

$$\mathbf{x} = \begin{bmatrix} -t/2 \\ t \\ 0 \end{bmatrix} = t \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix} = \tfrac{t}{2} \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}, \qquad \text{with } t \in \mathbb{C}.$$

So, $\left\{ \begin{bmatrix} -1 & 2 & 0 \end{bmatrix}^T \right\}$ is a basis of the eigespace $E_{\lambda_1}(A) = \text{Nul}\big(A - \lambda_1 I_n\big)$, and we see that the eigenvalue $\lambda_1 = 4$ has geometric multiplicity 1.

Example 8.2.4

Consider the following matrix in $\mathbb{C}^{3\times 3}$:

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

(a) Compute the characteristic polynomial $p_A(\lambda)$ of the matrix A.

(b) Compute all the eigenvalues of A and their algebraic multiplicities, and compute the spectrum of A.

(c) For each eigenvalue λ of A, compute a basis of the eigenspace $E_\lambda(A)$ and specify the geometric multiplicity of the eigenvalue λ.

Solution (continued). (c) Similarly, for $\lambda_2 = 5$, we get that $\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \right\}$ is a basis of the eigenspace $E_{\lambda_2}(A) = \text{Nul}(A - \lambda_2 I_n)$, and we see that the eigenvalue $\lambda_2 = 5$ has geometric multiplicity 2 (details: Lecture Notes). $\square$

- Reminder:

Proposition 7.3.1

Let $\mathbb{F}$ be a field, and let $A = \left[\ a_{i,j}\ \right]_{n\times n}$ be a triangular matrix in $\mathbb{F}^{n\times n}$. Then

$$\det(A) \;=\; \prod_{i=1}^{n} a_{i,i} \;=\; a_{1,1}a_{2,2}\dots a_{n,n},$$

that is, $\det(A)$ is equal to the product of entries on the main diagonal of A.

$$\begin{bmatrix} * & * & * & \dots & * & * \\ 0 & * & * & \dots & * & * \\ 0 & 0 & * & \dots & * & * \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & * & * \\ 0 & 0 & 0 & \dots & 0 & * \end{bmatrix}$$

upper triangular matrix

$$\begin{bmatrix} * & 0 & 0 & \dots & 0 & 0 \\ * & * & 0 & \dots & 0 & 0 \\ * & * & * & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ * & * & * & \dots & * & 0 \\ * & * & * & \dots & * & * \end{bmatrix}$$

lower triangular matrix

Proposition 8.2.7

Let $\mathbb{F}$ be a field, and let $A = \left[\; a_{i,j} \; \right]_{n \times n}$ be a triangular matrix in $\mathbb{F}^{n \times n}$. Then the characteristic polynomial of A is

$$p_A(\lambda) \;=\; \prod_{i=1}^{n} (\lambda - a_{i,i}) \;=\; (\lambda - a_{1,1})(\lambda - a_{2,2}) \dots (\lambda - a_{n,n}),$$

the eigenvalues of A are precisely the entries of A on its main diagonal, and moreover, the algebraic multiplicity of each eigenvalue is precisely the number of times that it appears on the main diagonal of A.[a] Consequently, the spectrum of A is $\{a_{1,1}, a_{2,2}, \dots, a_{n,n}\}$, i.e. the multiset formed precisely by the main diagonal entries of A, with each number appearing in the spectrum of A the same number of times as on the main diagonal of A.

[a] However, the geometric multiplicity may possibly be smaller.

- For example, for the matrix

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

in $\mathbb{C}^{5\times 5}$, we have the following:
  - the characteristic polynomial of A is:

$$\begin{aligned} p_A(\lambda) &= (\lambda - 1)(\lambda - 2)(\lambda - 1)(\lambda - 3)(\lambda - 3) \\ &= (\lambda - 1)^2(\lambda - 2)(\lambda - 3)^2; \end{aligned}$$

  - the spectrum of A is $\{1, 1, 2, 3, 3\}$.

Definition

Let $\mathbb{F}$ be a field. Given matrices $A, B \in \mathbb{F}^{n \times n}$, we say that A is *similar* to B if there exists an invertible matrix $P \in \mathbb{F}^{n \times n}$ s.t. $B = P^{-1}AP$.

Definition

Let $\mathbb{F}$ be a field. Given matrices $A, B \in \mathbb{F}^{n \times n}$, we say that A is *similar* to B if there exists an invertible matrix $P \in \mathbb{F}^{n \times n}$ s.t. $B = P^{-1}AP$.

Theorem 4.5.16

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;

(b) for all bases $\mathcal{B}$ of V and linear functions $f : V \to V$ s.t. $B = {}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$, there exists a basis $\mathcal{C}$ of V s.t. $C = {}_{\mathcal{C}}[\ f\ ]_{\mathcal{C}}$;

(c) for all bases $\mathcal{C}$ of V and linear functions $f : V \to V$ s.t. $C = {}_{\mathcal{C}}[\ f\ ]_{\mathcal{C}}$, there exists a basis $\mathcal{B}$ of V s.t. $B = {}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$;

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

Theorem 8.2.9

Let $\mathbb{F}$ be a field, and let $A, B \in \mathbb{F}^{n \times n}$ be similar matrices. Then A and B have the same **characteristic polynomial**, as well as the same **eigenvalues**, with the same corresponding **algebraic multiplicities**, and the same corresponding **geometric multiplicities**. Moreover, A and B have the same **spectrum**.

- **Warning:** Similar matrices A and B need not have the same eigenspaces, that is, for an eigenvalue λ of A and B:

$$E_\lambda(A) \not\asymp E_\lambda(B)$$

Theorem 8.2.9

Let $\mathbb{F}$ be a field, and let $A, B \in \mathbb{F}^{n \times n}$ be similar matrices. Then A and B have the same **characteristic polynomial**, as well as the same **eigenvalues**, with the same corresponding **algebraic multiplicities**, and the same corresponding **geometric multiplicities**. Moreover, A and B have the same **spectrum**.

Proof.

Theorem 8.2.9

Let $\mathbb{F}$ be a field, and let $A, B \in \mathbb{F}^{n \times n}$ be similar matrices. Then A and B have the same **characteristic polynomial**, as well as the same **eigenvalues**, with the same corresponding **algebraic multiplicities**, and the same corresponding **geometric multiplicities**. Moreover, A and B have the same **spectrum**.

Proof. Let us first show that A and B have the same eigenvalues with the same corresponding geometric multiplicities.

Theorem 8.2.9

Let $\mathbb{F}$ be a field, and let $A, B \in \mathbb{F}^{n \times n}$ be similar matrices. Then A and B have the same **characteristic polynomial**, as well as the same **eigenvalues**, with the same corresponding **algebraic multiplicities**, and the same corresponding **geometric multiplicities**. Moreover, A and B have the same **spectrum**.

Proof. Let us first show that A and B have the same eigenvalues with the same corresponding geometric multiplicities.

Since A and B are similar, Theorem 4.5.16 guarantees that there exists a linear function $f : \mathbb{F}^n \to \mathbb{F}^n$ and bases $\mathcal{A}$ and $\mathcal{B}$ of $\mathbb{F}^n$ s.t. $A = {}_{\mathcal{A}}[\ f\ ]_{\mathcal{A}}$ and $B = {}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$.

Theorem 8.2.9

Let $\mathbb{F}$ be a field, and let $A, B \in \mathbb{F}^{n \times n}$ be similar matrices. Then A and B have the same **characteristic polynomial**, as well as the same **eigenvalues**, with the same corresponding **algebraic multiplicities**, and the same corresponding **geometric multiplicities**. Moreover, A and B have the same **spectrum**.

Proof. Let us first show that A and B have the same eigenvalues with the same corresponding geometric multiplicities.

Since A and B are similar, Theorem 4.5.16 guarantees that there exists a linear function $f : \mathbb{F}^n \to \mathbb{F}^n$ and bases $\mathcal{A}$ and $\mathcal{B}$ of $\mathbb{F}^n$ s.t. $A = {}_{\mathcal{A}}\big[\ f\ \big]_{\mathcal{A}}$ and $B = {}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}}$.

But then by Proposition 8.1.7, the linear function f and the matrix $A = {}_{\mathcal{A}}\big[\ f\ \big]_{\mathcal{A}}$ have exactly the same eigenvalues, with exactly the same corresponding geometric multiplicities, and the same holds for f and the matrix $B = {}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}}$.

Theorem 8.2.9

Let $\mathbb{F}$ be a field, and let $A, B \in \mathbb{F}^{n\times n}$ be similar matrices. Then A and B have the same **characteristic polynomial**, as well as the same **eigenvalues**, with the same corresponding **algebraic multiplicities**, and the same corresponding **geometric multiplicities**. Moreover, A and B have the same **spectrum**.

Proof. Let us first show that A and B have the same eigenvalues with the same corresponding geometric multiplicities.

Since A and B are similar, Theorem 4.5.16 guarantees that there exists a linear function $f : \mathbb{F}^n \to \mathbb{F}^n$ and bases $\mathcal{A}$ and $\mathcal{B}$ of $\mathbb{F}^n$ s.t. $A = {}_{\mathcal{A}}[\ f\ ]_{\mathcal{A}}$ and $B = {}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$.
But then by Proposition 8.1.7, the linear function f and the matrix $A = {}_{\mathcal{A}}[\ f\ ]_{\mathcal{A}}$ have exactly the same eigenvalues, with exactly the same corresponding geometric multiplicities, and the same holds for f and the matrix $B = {}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$. So, A and B have exactly the same eigenvalues with exactly the same corresponding geometric multiplicities.

Theorem 8.2.9

Let $\mathbb{F}$ be a field, and let $A, B \in \mathbb{F}^{n \times n}$ be similar matrices. Then A and B have the same **characteristic polynomial**, as well as the same **eigenvalues**, with the same corresponding **algebraic multiplicities**, and the same corresponding **geometric multiplicities**. Moreover, A and B have the same **spectrum**.

Proof (continued). It now remains to show that A and B have the same characteristic polynomial, since this will (by definition) imply that A and B have the same spectrum, and in particular, that the eigenvalues of A and B have the same corresponding algebraic multiplicities.

Theorem 8.2.9

Let $\mathbb{F}$ be a field, and let $A, B \in \mathbb{F}^{n \times n}$ be similar matrices. Then A and B have the same **characteristic polynomial**, as well as the same **eigenvalues**, with the same corresponding **algebraic multiplicities**, and the same corresponding **geometric multiplicities**. Moreover, A and B have the same **spectrum**.

Proof (continued). It now remains to show that A and B have the same characteristic polynomial, since this will (by definition) imply that A and B have the same spectrum, and in particular, that the eigenvalues of A and B have the same corresponding algebraic multiplicities.

Since A and B are similar, we know that there exists an invertible matrix $P \in \mathbb{F}^{n \times n}$ s.t. $B = P^{-1}AP$.

Theorem 8.2.9

Let $\mathbb{F}$ be a field, and let $A, B \in \mathbb{F}^{n \times n}$ be similar matrices. Then A and B have the same **characteristic polynomial**, as well as the same **eigenvalues**, with the same corresponding **algebraic multiplicities**, and the same corresponding **geometric multiplicities**. Moreover, A and B have the same **spectrum**.

Proof (continued). It now remains to show that A and B have the same characteristic polynomial, since this will (by definition) imply that A and B have the same spectrum, and in particular, that the eigenvalues of A and B have the same corresponding algebraic multiplicities.

Since A and B are similar, we know that there exists an invertible matrix $P \in \mathbb{F}^{n \times n}$ s.t. $B = P^{-1}AP$. We now compute (next slide):

Theorem 8.2.9

Let $\mathbb{F}$ be a field, and let $A, B \in \mathbb{F}^{n \times n}$ be similar matrices. Then A and B have the same **characteristic polynomial**, as well as the same **eigenvalues**, with the same corresponding **algebraic multiplicities**, and the same corresponding **geometric multiplicities**. Moreover, A and B have the same **spectrum**.

Proof (continued).

$$\begin{aligned}
p_B(\lambda) &= \det\big(\lambda I_n - B\big) \\
&= \det\big(\lambda I_n - P^{-1}AP\big) \\
&= \det\big(P^{-1}(\lambda I_n - A)P\big) \\
&= \det(P^{-1}) \det\big(\lambda I_n - A\big) \det(P) && \text{by Theorem 7.5.2} \\
&= \tfrac{1}{\det(P)} \det\big(\lambda I_n - A\big) \det(P) && \text{by Corollary 7.5.3} \\
&= \det\big(\lambda I_n - A\big) \\
&= p_A(\lambda).
\end{aligned}$$

□

Theorem 8.2.9

Let $\mathbb{F}$ be a field, and let $A, B \in \mathbb{F}^{n \times n}$ be similar matrices. Then A and B have the same **characteristic polynomial**, as well as the same **eigenvalues**, with the same corresponding **algebraic multiplicities**, and the same corresponding **geometric multiplicities**. Moreover, A and B have the same **spectrum**.

- **Remark:** The converse of Theorem 8.2.9 is false: two matrices in $\mathbb{F}^{n \times n}$ (where $\mathbb{F}$ is a field) that have the same characteristic polynomial, as well as the same eigenvalues, with the same corresponding algebraic multiplicities, and the same corresponding geometric multiplicities, need not be similar.
    - We will see examples of this when we study the "Jordan normal form."

Definition

The *trace* of a square matrix $A = \left[\ a_{i,j}\ \right]_{n \times n}$ with entries in some field $\mathbb{F}$ is defined to be $\operatorname{trace}(A) := \sum_{i=1}^{n} a_{i,i}$, i.e. the trace of A is the sum of entries on the main diagonal of A.

- For example, for the matrix

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

in $\mathbb{C}^{3\times 3}$, we have that $\operatorname{trace}(A) = 1 + 5 + 9 = 15$.

Theorem 8.2.10

Let $\mathbb{F}$ be a field, let $A = \left[\ a_{i,j} \ \right]_{n \times n}$ be a matrix in $\mathbb{F}^{n \times n}$, and assume that $\{\lambda_1, \dots, \lambda_n\}$ is the spectrum of A. Then

(a) $\det(A) = \lambda_1 \dots \lambda_n$;

(b) $\text{trace}(A) = \lambda_1 + \dots + \lambda_n$.

Proof (outline).

Theorem 8.2.10

Let $\mathbb{F}$ be a field, let $A = \left[\ a_{i,j}\ \right]_{n \times n}$ be a matrix in $\mathbb{F}^{n \times n}$, and assume that $\{\lambda_1, \dots, \lambda_n\}$ is the spectrum of A. Then

(a) $\det(A) = \lambda_1 \dots \lambda_n$;

(b) $\text{trace}(A) = \lambda_1 + \dots + \lambda_n$.

Proof (outline). (a) Compute $p_A(0)$ in two different ways.
(b) Compute the coefficient in front of λ^{n-1} in $p_A(\lambda)$ in two different ways. (Details: Lecture Notes.) $\square$

Theorem 8.2.10

Let $\mathbb{F}$ be a field, let $A = \left[\ a_{i,j}\ \right]_{n\times n}$ be a matrix in $\mathbb{F}^{n\times n}$, and assume that $\{\lambda_1, \dots, \lambda_n\}$ is the spectrum of A. Then

(a) $\det(A) = \lambda_1 \dots \lambda_n$;

(b) $\text{trace}(A) = \lambda_1 + \dots + \lambda_n$.

Proof (outline). (a) Compute $p_A(0)$ in two different ways.
(b) Compute the coefficient in front of λ^{n-1} in $p_A(\lambda)$ in two different ways. (Details: Lecture Notes.) $\square$

- **Warning:** Theorem 8.2.10 only applies if the spectrum of the matrix $A \in \mathbb{F}^{n\times n}$ contains n eigenvalues (counting algebraic multiplicities)!
  - This will always be the case if the field $\mathbb{F}$ is **algebraically closed** (for example, if $\mathbb{F} = \mathbb{C}$), but need not be the case otherwise.

3. Eigenvalues and invertibility (plus the Invertible Matrix Theorem, version 4)

3. Eigenvalues and invertibility (plus the Invertible Matrix Theorem, version 4)

Proposition 8.2.11

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$. Then A is invertible iff 0 is **not** an eigenvalue of A.

Proof.

3. Eigenvalues and invertibility (plus the Invertible Matrix Theorem, version 4)

Proposition 8.2.11

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$. Then A is invertible iff 0 is **not** an eigenvalue of A.

Proof. It suffices to show that 0 is an eigenvalue of A iff A is not invertible.

3. Eigenvalues and invertibility (plus the Invertible Matrix Theorem, version 4)

Proposition 8.2.11

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$. Then A is invertible iff 0 is **not** an eigenvalue of A.

Proof. It suffices to show that 0 is an eigenvalue of A iff A is not invertible. We have the following sequence of equivalent statements:

$$\begin{aligned}
0 \text{ is eigenvalue of } A \quad & \overset{\text{Thm. 8.2.2}}{\iff} \quad \det(0I_n - A) = 0 \\
& \iff \quad \det(-A) = 0 \\
& \overset{\text{Prop. 7.2.3}}{\iff} \quad (-1)^n \det(A) = 0 \\
& \iff \quad \det(A) = 0 \\
& \overset{\text{Thm. 7.4.1}}{\iff} \quad A \text{ is not invertible}
\end{aligned}$$

□

Proposition 8.2.11

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$. Then A is invertible iff 0 is **not** an eigenvalue of A.

- We now add the eigenvalue condition from Proposition 8.2.11 to our previous version of the Invertible Matrix Theorem to obtain the fourth and final version of that theorem (next three slides).
    - It uses all 26 letters of the English alphabet!

The Invertible Matrix Theorem (version 4)

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n\times n}$ be a **square** matrix. Further, let $f : \mathbb{F}^n \to \mathbb{F}^n$ be given by $f(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^n$.[a] Then the following are equivalent:

(a) A is invertible (i.e. A has an inverse);

(b) A^T is invertible;

(c) $\text{RREF}(A) = I_n$;

(d) $\text{RREF}(\left[\begin{array}{c|c} A & I_n \end{array}\right]) = \left[\begin{array}{c|c} I_n & B \end{array}\right]$ for some matrix $B \in \mathbb{F}^{n\times n}$;

(e) $\text{rank}(A) = n$;

(f) $\text{rank}(A^T) = n$;

(g) is a product of elementary matrices;

---

[a] Since f is a matrix transformation, Proposition 1.10.4 guarantees that f is linear. Moreover, A is the standard matrix of f.

The Invertible Matrix Theorem (version 4, continued)

(h) the homogeneous matrix-vector equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution (i.e. the solution $\mathbf{x} = \mathbf{0}$);

(i) there exists some vector $\mathbf{b} \in \mathbb{F}^n$ s.t. the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ has a unique solution;

(j) for all vectors $\mathbf{b} \in \mathbb{F}^n$, the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ has a unique solution;

(k) for all vectors $\mathbf{b} \in \mathbb{F}^n$, the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ has at most one solution;

(l) for all vectors $\mathbf{b} \in \mathbb{F}^n$, the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is consistent;

(m) f is one-to-one;

(n) f is onto;

(o) f is an isomorphism;

The Invertible Matrix Theorem (version 4, continued)

(p) there exists a matrix $B \in \mathbb{F}^{n \times n}$ s.t. $BA = I_n$ (i.e. A has a left inverse);

(q) there exists a matrix $C \in \mathbb{F}^{n \times n}$ s.t. $AC = I_n$ (i.e. A has a right inverse);

(r) the columns of A are linearly independent;

(s) the columns of A span $\mathbb{F}^n$ (i.e. $\text{Col}(A) = \mathbb{F}^n$);

(t) the columns of A form a basis of $\mathbb{F}^n$;

(u) the rows of A are linearly independent;

(v) the rows of A span $\mathbb{F}^{1 \times n}$ (i.e. $\text{Row}(A) = \mathbb{F}^{1 \times n}$);

(w) the rows of A form a basis of $\mathbb{F}^{1 \times n}$;

(x) $\text{Nul}(A) = \{\mathbf{0}\}$ (i.e. $\dim(\text{Nul}(A)) = 0$);

(y) $\det(A) \neq 0$;

(z) 0 is not an eigenvalue of A.

- Reminder:
  - Suppose that V is a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, and that $f : V \to V$ is a linear function. Then we define the *determinant* of f to be

$$\det(f) \quad := \quad \det\Big({}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}} \Big),$$

where $\mathcal{B}$ is any basis of V.

- Reminder:
  - Suppose that V is a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, and that $f : V \to V$ is a linear function. Then we define the *determinant* of f to be

$$\det(f) \quad := \quad \det\Big({}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}} \Big),$$

    where $\mathcal{B}$ is any basis of V.
  - As we explained in section 7.5, the reason that $\det(f)$ is well defined is because, by Theorem 4.5.16, all matrices of the form ${}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}}$ are similar, and therefore (by Corollary 7.5.4) have the same determinant.

Definition

Let V is a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$. The *characteristic polynomial* of a linear function $f : V \to V$ is defined to be the polynomial

$$p_f(\lambda) \quad := \quad \det\big(\lambda \mathrm{Id}_V - f\big) \quad = \quad \det\Big({}_{\mathcal{B}}\Big[\ \lambda \mathrm{Id}_V - f \ \Big]_{\mathcal{B}} \Big),$$

where $\mathcal{B}$ is **any** basis of V.[a]

---

[a] As usual, Id_V is the identity function on V, i.e. it is the function $\mathrm{Id}_V : V \to V$ given by $\mathrm{Id}_V(\mathbf{v}) = \mathbf{v}$ for all $\mathbf{v} \in V$.

Definition

Let V is a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$. The *characteristic polynomial* of a linear function $f : V \to V$ is defined to be the polynomial

$$p_f(\lambda) \;:=\; \det\big(\lambda \mathrm{Id}_V - f\big) \;=\; \det\Big({}_{\mathcal{B}}\big[\; \lambda \mathrm{Id}_V - f \;\big]_{\mathcal{B}} \Big),$$

where $\mathcal{B}$ is **any** basis of V.[a]

[a] As usual, Id_V is the identity function on V, i.e. it is the function $\mathrm{Id}_V : V \to V$ given by $\mathrm{Id}_V(\mathbf{v}) = \mathbf{v}$ for all $\mathbf{v} \in V$.

- As per our discussion above, the polynomial $p_f(\lambda)$ depends only on f, and not on the particular choice of the basis $\mathcal{B}$.

Definition

Let V is a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$. The *characteristic polynomial* of a linear function $f : V \to V$ is defined to be the polynomial

$$p_f(\lambda) \ := \ \det\big(\lambda \mathrm{Id}_V - f\big) \ = \ \det\Big({}_{\mathcal{B}}\big[\ \lambda \mathrm{Id}_V - f\ \big]_{\mathcal{B}} \Big),$$

where $\mathcal{B}$ is **any** basis of V.[a]

[a] As usual, Id_V is the identity function on V, i.e. it is the function $\mathrm{Id}_V : V \to V$ given by $\mathrm{Id}_V(\mathbf{v}) = \mathbf{v}$ for all $\mathbf{v} \in V$.

- As per our discussion above, the polynomial $p_f(\lambda)$ depends only on f, and not on the particular choice of the basis $\mathcal{B}$.
- The *characteristic equation* of f is the equation

  $$\det\big(\lambda \mathrm{Id}_V - f\big) \ = \ 0.$$

  So, the roots of the characteristic polynomial of f are precisely the solutions of the characteristic equation of f.

Proposition 8.2.12

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, let $\mathcal{B}$ be any basis of V, let $f : V \to V$ be a linear function, and set $B := {}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$. Then $p_f(\lambda) = p_B(\lambda)$.

Proof.

Proposition 8.2.12

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, let $\mathcal{B}$ be any basis of V, let $f : V \to V$ be a linear function, and set $B := {}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$. Then $p_f(\lambda) = p_B(\lambda)$.

Proof. We compute:

$$\begin{aligned}
p_f(\lambda) &= \det\big(\lambda \mathrm{Id}_V - f\big) && \text{by definition} \\
&= \det\Big({}_{\mathcal{B}}[\ \lambda \mathrm{Id}_V - f\ ]_{\mathcal{B}}\Big) && \text{by definition} \\
&= \det\Big(\lambda\, {}_{\mathcal{B}}[\ \mathrm{Id}_V\ ]_{\mathcal{B}} - {}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}\Big) && \text{by Theorem 4.5.3} \\
&= \det\big(\lambda I_n - B\big) \\
&= p_B(\lambda) && \text{by definition.}
\end{aligned}$$

□

- Reminder:

Theorem 8.2.2

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n \times n}$, and let $\lambda_0 \in \mathbb{F}$. Then

$$E_{\lambda_0}(A) \;=\; \mathrm{Nul}(\lambda_0 I_n - A) \;=\; \mathrm{Nul}(A - \lambda_0 I_n).$$

Moreover, the following are equivalent:

1. λ_0 is an eigenvalue of A;
2. λ_0 is a root of the characteristic polynomial of A, i.e. $p_A(\lambda_0) = 0$;
3. λ_0 is a solution of the characteristic equation of A, i.e. $\det(\lambda_0 I_n - A) = 0$.

- Reminder:

Theorem 8.2.2

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n\times n}$, and let $\lambda_0 \in \mathbb{F}$. Then

$$E_{\lambda_0}(A) \ = \ \mathrm{Nul}(\lambda_0 I_n - A) \ = \ \mathrm{Nul}(A - \lambda_0 I_n).$$

Moreover, the following are equivalent:

1. λ_0 is an eigenvalue of A;
2. λ_0 is a root of the characteristic polynomial of A, i.e. $p_A(\lambda_0) = 0$;
3. λ_0 is a solution of the characteristic equation of A, i.e. $\det(\lambda_0 I_n - A) = 0$.

- Analogously, we have the following (next slide):

Theorem 8.2.13

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, let $f : V \to V$ be a linear function, and let $\lambda_0 \in \mathbb{F}$. Then

$$E_{\lambda_0}(f) \;=\; \mathrm{Ker}(\lambda_0 \mathrm{Id}_V - f) \;=\; \mathrm{Ker}(f - \lambda_0 \mathrm{Id}_V).$$

Moreover, the following are equivalent:

1. λ_0 is an eigenvalue of f;
2. λ_0 is a root of the characteristic polynomial of f, i.e. $p_f(\lambda_0) = 0$;
3. λ_0 is a solution of the characteristic equation of f, i.e. $\det(\lambda_0 \mathrm{Id}_V - f) = 0$.

- Proof: Lecture Notes. (Similar to the proof of Theorem 8.2.2.)

Proposition 8.2.14

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, let $f : V \to V$ be a linear function, and let $\mathcal{B}$ be any basis of V. Then f and ${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$ have the same **characteristic polynomial**, and the same **spectrum**. Moreover, f and ${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$ have exactly the same **eigenvalues**, with exactly the same corresponding **geometric multiplicities**, and exactly the same corresponding **algebraic multiplicities**.

Proof.

Proposition 8.2.14

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, let $f : V \to V$ be a linear function, and let $\mathcal{B}$ be any basis of V. Then f and ${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$ have the same **characteristic polynomial**, and the same **spectrum**. Moreover, f and ${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$ have exactly the same **eigenvalues**, with exactly the same corresponding **geometric multiplicities**, and exactly the same corresponding **algebraic multiplicities**.

Proof. The fact that f and ${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$ have the same **eigenvalues**, with the same **geometric multiplicities**, follows immediately from Proposition 8.1.7.

Proposition 8.2.14

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, let $f : V \to V$ be a linear function, and let $\mathcal{B}$ be any basis of V. Then f and ${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$ have the same **characteristic polynomial**, and the same **spectrum**. Moreover, f and ${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$ have exactly the same **eigenvalues**, with exactly the same corresponding **geometric multiplicities**, and exactly the same corresponding **algebraic multiplicities**.

Proof. The fact that f and ${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$ have the same **eigenvalues**, with the same **geometric multiplicities**, follows immediately from Proposition 8.1.7.

The fact that they have the same **characteristic polynomial** (and consequently the same **spectrum**) follows immediately from Proposition 8.2.12.

Proposition 8.2.14

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, let $f : V \to V$ be a linear function, and let $\mathcal{B}$ be any basis of V. Then f and ${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$ have the same **characteristic polynomial**, and the same **spectrum**. Moreover, f and ${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$ have exactly the same **eigenvalues**, with exactly the same corresponding **geometric multiplicities**, and exactly the same corresponding **algebraic multiplicities**.

Proof. The fact that f and ${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$ have the same **eigenvalues**, with the same **geometric multiplicities**, follows immediately from Proposition 8.1.7.

The fact that they have the same **characteristic polynomial** (and consequently the same **spectrum**) follows immediately from Proposition 8.2.12. Since f and ${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$ have the same spectrum, their eigenvalues have the same **algebraic multiplicities**. □

Proposition 8.2.14

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, let $f : V \to V$ be a linear function, and let $\mathcal{B}$ be any basis of V. Then f and ${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$ have the same **characteristic polynomial**, and the same **spectrum**. Moreover, f and ${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$ have exactly the same **eigenvalues**, with exactly the same corresponding **geometric multiplicities**, and exactly the same corresponding **algebraic multiplicities**.

- As a special case for linear functions of the form $f : \mathbb{F}^n \to \mathbb{F}^n$ (where $\mathbb{F}$ is a field) and their standard matrices, we have the following proposition (next slide).

Proposition 8.2.15

Let $\mathbb{F}$ be a field, let $f : \mathbb{F}^n \to \mathbb{F}^n$ be a linear function, and let A be the standard matrix of f. Then f and A have the same characteristic polynomial and the same spectrum. Moreover, for each eigenvalue λ of f and A, all the following hold:

- the algebraic multiplicity of λ as an eigenvalue of f is the same as the algebraic multiplicity of λ as an eigenvalue of A;
- the geometric multiplicity of λ as an eigenvalue of f is the same as the geometric multiplicity of λ as an eigenvalue of A;
- $E_\lambda(f) = E_\lambda(A)$.

Proof.

Proposition 8.2.15

Let $\mathbb{F}$ be a field, let $f : \mathbb{F}^n \to \mathbb{F}^n$ be a linear function, and let A be the standard matrix of f. Then f and A have the same characteristic polynomial and the same spectrum. Moreover, for each eigenvalue λ of f and A, all the following hold:

- the algebraic multiplicity of λ as an eigenvalue of f is the same as the algebraic multiplicity of λ as an eigenvalue of A;
- the geometric multiplicity of λ as an eigenvalue of f is the same as the geometric multiplicity of λ as an eigenvalue of A;
- $E_\lambda(f) = E_\lambda(A)$.

Proof. Since A is the standard matrix of f, we have that $A = {}_{\mathcal{E}_n}\left[\ f\ \right]_{\mathcal{E}_n}$, where $\mathcal{E}_n$ is the standard basis of $\mathbb{F}^n$.

Proposition 8.2.15

Let $\mathbb{F}$ be a field, let $f : \mathbb{F}^n \to \mathbb{F}^n$ be a linear function, and let A be the standard matrix of f. Then f and A have the same characteristic polynomial and the same spectrum. Moreover, for each eigenvalue λ of f and A, all the following hold:

- the algebraic multiplicity of λ as an eigenvalue of f is the same as the algebraic multiplicity of λ as an eigenvalue of A;
- the geometric multiplicity of λ as an eigenvalue of f is the same as the geometric multiplicity of λ as an eigenvalue of A;
- $E_\lambda(f) = E_\lambda(A)$.

Proof. Since A is the standard matrix of f, we have that $A = {}_{\mathcal{E}_n}[\ f\ ]_{\mathcal{E}_n}$, where $\mathcal{E}_n$ is the standard basis of $\mathbb{F}^n$. The result now follows immediately from Propositions 8.1.5 and 8.2.14. □

4. The relationship between algebraic and geometric multiplicities of eigenvalues

4. The relationship between algebraic and geometric multiplicities of eigenvalues

- Let's now prove Theorem 8.2.3!

Theorem 8.2.3

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$. Then the geometric multiplicity of any eigenvalue of A is no greater than the algebraic multiplicity of that eigenvalue.

- Schematically, Theorem 8.2.3 states that for an eigenvalue λ of A:

$$\text{geometric multiplicity of } \lambda \quad \leq \quad \text{algebraic multiplicity of } \lambda.$$

4. The relationship between algebraic and geometric multiplicities of eigenvalues

- Let's now prove Theorem 8.2.3!

Theorem 8.2.3

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$. Then the geometric multiplicity of any eigenvalue of A is no greater than the algebraic multiplicity of that eigenvalue.

- Schematically, Theorem 8.2.3 states that for an eigenvalue λ of A:

$$\text{geometric multiplicity of } \lambda \quad \leq \quad \text{algebraic multiplicity of } \lambda.$$

- In fact, it will be a bit more convenient to prove this theorem for linear functions first (see Theorem 8.2.17 below), and to then derive Theorem 8.2.3 as in immediate corollary.

Theorem 8.2.17

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, and let $f : V \to V$ be a linear function. Then the geometric multiplicity of any eigenvalue of f is no greater than the algebraic multiplicity of that eigenvalue.

Proof.

Theorem 8.2.17

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, and let $f : V \to V$ be a linear function. Then the geometric multiplicity of any eigenvalue of f is no greater than the algebraic multiplicity of that eigenvalue.

Proof. Suppose that λ_0 is an eigenvalue of f of geometric multiplicity k.

Theorem 8.2.17

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, and let $f : V \to V$ be a linear function. Then the geometric multiplicity of any eigenvalue of f is no greater than the algebraic multiplicity of that eigenvalue.

Proof. Suppose that λ_0 is an eigenvalue of f of geometric multiplicity k. We must show that the eigenvalue λ_0 has algebraic multiplicity at least k, that is, that $(\lambda - \lambda_0)^k \mid p_f(\lambda)$.

Theorem 8.2.17

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, and let $f : V \to V$ be a linear function. Then the geometric multiplicity of any eigenvalue of f is no greater than the algebraic multiplicity of that eigenvalue.

Proof. Suppose that λ_0 is an eigenvalue of f of geometric multiplicity k. We must show that the eigenvalue λ_0 has algebraic multiplicity at least k, that is, that $(\lambda - \lambda_0)^k \mid p_f(\lambda)$.

The goal is to find a basis $\mathcal{B}$ of V for which it can easily be shown that $(\lambda - \lambda_0)^k$ divides the polynomial $p_B(\lambda)$, where $B = {}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$; this is enough because, by Proposition 8.2.12, $p_f(\lambda) = p_B(\lambda)$.

Proof (continued). Reminder: λ_0 is an eigenvalue of f; WTS there exists a basis $\mathcal{B}$ of V s.t. $(\lambda - \lambda_0)^k \mid p_B(\lambda)$, where $B = {}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$.

Proof (continued). Reminder: λ_0 is an eigenvalue of f; WTS there exists a basis $\mathcal{B}$ of V s.t. $(\lambda - \lambda_0)^k \mid p_B(\lambda)$, where $B = {}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$.

Since the geometric multiplicity of the eigenvalue λ_0 of f is k, we see that the eigenspace $E_{\lambda_0}(f)$ has a k-element basis, say $\{\mathbf{b}_1, \dots, \mathbf{b}_k\}$.

Proof (continued). Reminder: λ_0 is an eigenvalue of f; WTS there exists a basis $\mathcal{B}$ of V s.t. $(\lambda - \lambda_0)^k \mid p_B(\lambda)$, where $B = {}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$.

Since the geometric multiplicity of the eigenvalue λ_0 of f is k, we see that the eigenspace $E_{\lambda_0}(f)$ has a k-element basis, say $\{\mathbf{b}_1, \dots, \mathbf{b}_k\}$. In particular, $\{\mathbf{b}_1, \dots, \mathbf{b}_k\}$ is a linearly independent set of vectors in V, and so by Theorem 8.2.19, it can be extended to a basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_k, \mathbf{b}_{k+1}, \dots, \mathbf{b}_n\}$ of V.

Proof (continued). Reminder: λ_0 is an eigenvalue of f; WTS there exists a basis $\mathcal{B}$ of V s.t. $(\lambda - \lambda_0)^k \mid p_B(\lambda)$, where $B = {}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$.

Since the geometric multiplicity of the eigenvalue λ_0 of f is k, we see that the eigenspace $E_{\lambda_0}(f)$ has a k-element basis, say $\{\mathbf{b}_1, \dots, \mathbf{b}_k\}$. In particular, $\{\mathbf{b}_1, \dots, \mathbf{b}_k\}$ is a linearly independent set of vectors in V, and so by Theorem 8.2.19, it can be extended to a basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_k, \mathbf{b}_{k+1}, \dots, \mathbf{b}_n\}$ of V. We now compute:

$$\begin{array}{rcl} B & := & {}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}} \\ & \overset{(*)}{=} & \left[\ [\ f(\mathbf{b}_1)\ ]_{\mathcal{B}} \ \dots \ [\ f(\mathbf{b}_k)\ ]_{\mathcal{B}} \ \ [\ f(\mathbf{b}_{k+1})\ ]_{\mathcal{B}} \ \dots \ [\ f(\mathbf{b}_n)\ ]_{\mathcal{B}}\ \right] \\ & \overset{(**)}{=} & \left[\ [\ \lambda_0\mathbf{b}_1\ ]_{\mathcal{B}} \ \dots \ [\ \lambda_0\mathbf{b}_k\ ]_{\mathcal{B}} \ \ [\ f(\mathbf{b}_{k+1})\ ]_{\mathcal{B}} \ \dots \ [\ f(\mathbf{b}_n)\ ]_{\mathcal{B}}\ \right] \\ & = & \left[\ \lambda_0\mathbf{e}_1^n \ \dots \ \lambda_0\mathbf{e}_k^n \ \ [\ f(\mathbf{b}_{k+1})\ ]_{\mathcal{B}} \ \dots \ [\ f(\mathbf{b}_n)\ ]_{\mathcal{B}}\ \right] \\ & = & \left[\begin{array}{c|c} \lambda_0 I_k & \\ \hline O_{(n-k)\times k} & \end{array} [\ f(\mathbf{b}_{k+1})\ ]_{\mathcal{B}} \ \dots \ [\ f(\mathbf{b}_n)\ ]_{\mathcal{B}} \right], \end{array}$$

where (*) follows from Theorem 4.5.1, and (**) follows from the fact that $\mathbf{b}_1, \dots, \mathbf{b}_k \in E_{\lambda_0}(f)$.

Proof (continued). Reminder: λ_0 is an eigenvalue of f; WTS there exists a basis $\mathcal{B}$ of V s.t. $(\lambda - \lambda_0)^k \mid p_B(\lambda)$, where $B = {}_{\mathcal{B}}[\, f \,]_{\mathcal{B}}$.

We showed:

$$B \;:=\; {}_{\mathcal{B}}[\, f \,]_{\mathcal{B}} \;=\; \left[\begin{array}{c|ccc} \begin{array}{c} \lambda_0 I_k \\ \hline O_{(n-k)\times k} \end{array} & [\, f(\mathbf{b}_{k+1}) \,]_{\mathcal{B}} & \dots & [\, f(\mathbf{b}_n) \,]_{\mathcal{B}} \end{array}\right].$$

Proof (continued). Reminder: λ_0 is an eigenvalue of f; WTS there exists a basis $\mathcal{B}$ of V s.t. $(\lambda - \lambda_0)^k \mid p_B(\lambda)$, where $B = {}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}}$.

We showed:

$$B \ := \ {}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}} \ = \ \left[\begin{array}{c|ccc} \lambda_0 I_k & \big[\ f(\mathbf{b}_{k+1})\ \big]_{\mathcal{B}} & \dots & \big[\ f(\mathbf{b}_n)\ \big]_{\mathcal{B}} \\ \hline O_{(n-k)\times k} & & & \end{array}\right].$$

Thus, $p_B(\lambda)$ is of the form

$$p_B(\lambda) \ = \ \left|\begin{array}{cccc|cccc} \lambda-\lambda_0 & 0 & \dots & 0 & * & * & \dots & * \\ 0 & \lambda-\lambda_0 & \dots & 0 & * & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda-\lambda_0 & * & * & \dots & * \\ \hline 0 & 0 & \dots & 0 & * & * & \dots & * \\ 0 & 0 & \dots & 0 & * & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & * & * & \dots & * \end{array}\right|,$$

where the red submatrix in the upper-left corner (to the left of the vertical dotted line, and above the horizontal dotted line) is of size $k \times k$. By iteratively performing Laplace expansion along the first column, we see that $p_B(\lambda)$ has a factor $(\lambda - \lambda_0)^k$. □

Theorem 8.2.17

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, and let $f : V \to V$ be a linear function. Then the geometric multiplicity of any eigenvalue of f is no greater than the algebraic multiplicity of that eigenvalue.

Theorem 8.2.17

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, and let $f : V \to V$ be a linear function. Then the geometric multiplicity of any eigenvalue of f is no greater than the algebraic multiplicity of that eigenvalue.

Theorem 8.2.3

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$. Then the geometric multiplicity of any eigenvalue of A is no greater than the algebraic multiplicity of that eigenvalue.

Proof.

Theorem 8.2.17

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, and let $f : V \to V$ be a linear function. Then the geometric multiplicity of any eigenvalue of f is no greater than the algebraic multiplicity of that eigenvalue.

Theorem 8.2.3

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$. Then the geometric multiplicity of any eigenvalue of A is no greater than the algebraic multiplicity of that eigenvalue.

Proof. Let $f_A : \mathbb{F}^n \to \mathbb{F}^n$ be given by $f_A(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^n$. Then f_A is linear (by Prop. 1.10.4), and its standard matrix is A.

Theorem 8.2.17

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, and let $f : V \to V$ be a linear function. Then the geometric multiplicity of any eigenvalue of f is no greater than the algebraic multiplicity of that eigenvalue.

Theorem 8.2.3

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$. Then the geometric multiplicity of any eigenvalue of A is no greater than the algebraic multiplicity of that eigenvalue.

Proof. Let $f_A : \mathbb{F}^n \to \mathbb{F}^n$ be given by $f_A(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^n$. Then f_A is linear (by Prop. 1.10.4), and its standard matrix is A.

By Proposition 8.2.15, A and f_A have exactly the same eigenvalues, with the same corresponding geometric multiplicities, and the same corresponding algebraic multiplicities. The result now follows from Theorem 8.2.17 applied to the linear function f_A. $\square$