Linear Algebra 2

Lecture #22

The Cayley-Hamilton theorem. Diagonalization

Irena Penev

April 24, 2024

- This lecture has three parts:

- This lecture has three parts:
  1. The Cayley-Hamilton theorem

- This lecture has three parts:
    1. The Cayley-Hamilton theorem
    2. Eigenvectors and linear independence. Eigenbases

- This lecture has three parts:
  1. The Cayley-Hamilton theorem
  2. Eigenvectors and linear independence. Eigenbases
  3. Diagonalization

1. The Cayley-Hamilton theorem

1. The Cayley-Hamilton theorem

The Cayley-Hamilton theorem

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n \times n}$ be a square matrix, and let $p_A(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$ be the characteristic polynomial of A. Then

$$A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I_n \quad = \quad O_{n \times n}.$$

- The Cayley-Hamilton theorem essentially states that every square matrix is a root of its own characteristic polynomial.
  - Here, we need to treat the free coefficient of the characteristic polynomial as that coefficient times the identity matrix of the appropriate size.

- For example, for the matrix

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

with entries understood to be in $\mathbb{R}$ or $\mathbb{C}$, we have that

$$p_A(\lambda) = \det(\lambda I_2 - A) = \begin{vmatrix} \lambda - 1 & -2 \\ -3 & \lambda - 4 \end{vmatrix} = \lambda^2 - 5\lambda - 2,$$

and we see that

$$\begin{aligned} A^2 - 5A - 2I_2 &= \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}^2 - 5\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} - 2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 7 & 10 \\ 15 & 22 \end{bmatrix} - \begin{bmatrix} 5 & 10 \\ 15 & 20 \end{bmatrix} - \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

The Cayley-Hamilton theorem

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n \times n}$ be a square matrix, and let $p_A(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$ be the characteristic polynomial of A. Then

$$A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I_n \quad = \quad O_{n \times n}.$$

- The proof of the Cayley-Hamilton theorem relies on the adjugate matrix and on the theorem below.

Theorem 7.8.2

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$ $(n \geq 2)$. Then

$$\text{adj}(A)\, A \quad = \quad A\, \text{adj}(A) \quad = \quad \det(A) I_n.$$

Consequently, if A is invertible, then $A^{-1} = \frac{1}{\det(A)}\text{adj}(A)$.

The Cayley-Hamilton theorem

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n \times n}$ be a square matrix, and let $p_A(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$ be the characteristic polynomial of A. Then

$$A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I_n \quad = \quad O_{n \times n}.$$

Proof.

The Cayley-Hamilton theorem

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n\times n}$ be a square matrix, and let $p_A(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$ be the characteristic polynomial of A. Then

$$A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I_n \quad = \quad O_{n\times n}.$$

Proof. If $n = 1$, then the result is immediate.

- Indeed, suppose that $n = 1$, and consider any matrix $A = \left[\ a_{1,1}\ \right]$ in $\mathbb{F}^{1\times 1}$.
- Then $p_A(\lambda) = \det(\lambda I_1 - A) = \det(\left[\ \lambda - a_{1,1}\ \right]) = \lambda - a_{1,1}$, and we see that $A - a_{1,1}I_1 = O_{1\times 1}$.

Proof (continued). From now on, we assume that $n \geq 2$.

Proof (continued). From now on, we assume that $n \geq 2$. By Theorem 7.8.2 applied to the matrix $\lambda I_n - A$ (where λ is a variable), we get that

$$(\lambda I_n - A)\,\mathrm{adj}(\lambda I_n - A) \;\; = \;\; \det(\lambda I_n - A) I_n.$$

Proof (continued). From now on, we assume that $n \geq 2$. By Theorem 7.8.2 applied to the matrix $\lambda I_n - A$ (where λ is a variable), we get that

$$(\lambda I_n - A)\,\text{adj}(\lambda I_n - A) \quad = \quad \det(\lambda I_n - A) I_n.$$

Now, note that each cofactor of the matrix $\lambda I_n - A$ is a polynomial (in variable λ) of degree at most λ^{n-1}.

Proof (continued). From now on, we assume that $n \geq 2$. By Theorem 7.8.2 applied to the matrix $\lambda I_n - A$ (where λ is a variable), we get that

$$(\lambda I_n - A)\,\mathrm{adj}(\lambda I_n - A) \;\; = \;\; \det(\lambda I_n - A) I_n.$$

Now, note that each cofactor of the matrix $\lambda I_n - A$ is a polynomial (in variable λ) of degree at most λ^{n-1}. Since the entries of $\mathrm{adj}(\lambda I_n - A)$ are precisely the cofactors of $\lambda I_n - A$, it follows that each entry of $\mathrm{adj}(\lambda I_n - A)$ is a polynomial (in the variable λ) of degree at most $n - 1$.

Proof (continued). From now on, we assume that $n \geq 2$. By Theorem 7.8.2 applied to the matrix $\lambda I_n - A$ (where λ is a variable), we get that

$$(\lambda I_n - A)\,\text{adj}(\lambda I_n - A) \;\; = \;\; \det(\lambda I_n - A) I_n.$$

Now, note that each cofactor of the matrix $\lambda I_n - A$ is a polynomial (in variable λ) of degree at most λ^{n-1}. Since the entries of $\text{adj}(\lambda I_n - A)$ are precisely the cofactors of $\lambda I_n - A$, it follows that each entry of $\text{adj}(\lambda I_n - A)$ is a polynomial (in the variable λ) of degree at most $n-1$. So, the matrix $\text{adj}(\lambda I_n - A)$ can be expressed in the form

$$\text{adj}(\lambda I_n - A) \;\; = \;\; \lambda^{n-1} B_{n-1} + \lambda^{n-2} B_{n-2} + \cdots + \lambda B_1 + B_0,$$

for some matrices $B_0, B_1, \dots, B_{n-1} \in \mathbb{F}^{n \times n}$.

Proof (continued). From now on, we assume that $n \geq 2$. By Theorem 7.8.2 applied to the matrix $\lambda I_n - A$ (where λ is a variable), we get that

$$(\lambda I_n - A)\,\text{adj}(\lambda I_n - A) \;\;=\;\; \det(\lambda I_n - A) I_n.$$

Now, note that each cofactor of the matrix $\lambda I_n - A$ is a polynomial (in variable λ) of degree at most λ^{n-1}. Since the entries of $\text{adj}(\lambda I_n - A)$ are precisely the cofactors of $\lambda I_n - A$, it follows that each entry of $\text{adj}(\lambda I_n - A)$ is a polynomial (in the variable λ) of degree at most $n-1$. So, the matrix $\text{adj}(\lambda I_n - A)$ can be expressed in the form

$$\text{adj}(\lambda I_n - A) \;\;=\;\; \lambda^{n-1} B_{n-1} + \lambda^{n-2} B_{n-2} + \dots + \lambda B_1 + B_0,$$

for some matrices $B_0, B_1, \dots, B_{n-1} \in \mathbb{F}^{n \times n}$. Consequently,

$$\underbrace{(\lambda I_n - A)(\underbrace{\lambda^{n-1} B_{n-1} + \lambda^{n-2} B_{n-2} + \dots + \lambda B_1 + B_0}_{=\text{adj}(\lambda I_n - A)})}_{:=\text{LHS}} \;\;=\;\; \underbrace{\det(\lambda I_n - A) I_n}_{:=\text{RHS}}.$$

Proof (continued). Reminder: $n \geq 2$,

$$\underbrace{(\lambda I_n - A)(\underbrace{\lambda^{n-1}B_{n-1} + \lambda^{n-2}B_{n-2} + \cdots + \lambda B_1 + B_0}_{=\mathrm{adj}(\lambda I_n - A)})}_{:=\mathrm{LHS}} = \underbrace{\det(\lambda I_n - A)I_n}_{:=\mathrm{RHS}}.$$

Proof (continued). Reminder: $n \geq 2$,

$$\underbrace{(\lambda I_n - A)(\underbrace{\lambda^{n-1}B_{n-1} + \lambda^{n-2}B_{n-2} + \cdots + \lambda B_1 + B_0}_{=\mathrm{adj}(\lambda I_n - A)})}_{:=\mathrm{LHS}} = \underbrace{\det(\lambda I_n - A)I_n}_{:=\mathrm{RHS}}.$$

For the left-hand-side, we have

$$\begin{aligned} \mathrm{LHS} &= (\lambda I_n - A)(\lambda^{n-1}B_{n-1} + \cdots + \lambda B_1 + B_0) \\ &= \lambda^n B_{n-1} + \lambda^{n-1}(B_{n-2} - AB_{n-1}) + \lambda^{n-2}(B_{n-3} - AB_{n-2}) + \\ &\quad + \cdots + \lambda(B_0 - AB_1) - AB_0. \end{aligned}$$

Proof (continued). Reminder: $n \geq 2$,

$$\underbrace{(\lambda I_n - A)(\underbrace{\lambda^{n-1}B_{n-1} + \lambda^{n-2}B_{n-2} + \cdots + \lambda B_1 + B_0}_{=\text{adj}(\lambda I_n - A)})}_{:=\text{LHS}} = \underbrace{\det(\lambda I_n - A)I_n}_{:=\text{RHS}}.$$

For the left-hand-side, we have

$$\begin{aligned} \text{LHS} &= (\lambda I_n - A)(\lambda^{n-1}B_{n-1} + \cdots + \lambda B_1 + B_0) \\ &= \lambda^n B_{n-1} + \lambda^{n-1}(B_{n-2} - AB_{n-1}) + \lambda^{n-2}(B_{n-3} - AB_{n-2}) + \\ &\quad + \cdots + \lambda(B_0 - AB_1) - AB_0. \end{aligned}$$

For the right-hand-side, we have

$$\begin{aligned} \text{RHS} &= \det(\lambda I_n - A)I_n = p_A(\lambda)I_n \\ &= (\lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-1} + \cdots + a_1\lambda + a_0)I_n \\ &= \lambda^n I_n + \lambda^{n-1}a_{n-1}I_n + \lambda^{n-2}a_{n-2}I_n + \cdots + \lambda a_1 I_n + a_0 I_n. \end{aligned}$$

Proof (continued). Reminder: $n \geq 2$,

$$\underbrace{(\lambda I_n - A)(\underbrace{\lambda^{n-1}B_{n-1} + \lambda^{n-2}B_{n-2} + \cdots + \lambda B_1 + B_0}_{=\text{adj}(\lambda I_n - A)})}_{:=\text{LHS}} = \underbrace{\det(\lambda I_n - A)I_n}_{:=\text{RHS}}.$$

For the left-hand-side, we have

$$\begin{aligned}\text{LHS} &= (\lambda I_n - A)(\lambda^{n-1}B_{n-1} + \cdots + \lambda B_1 + B_0) \\ &= \lambda^n B_{n-1} + \lambda^{n-1}(B_{n-2} - AB_{n-1}) + \lambda^{n-2}(B_{n-3} - AB_{n-2}) + \\ &\quad + \cdots + \lambda(B_0 - AB_1) - AB_0.\end{aligned}$$

For the right-hand-side, we have

$$\begin{aligned}\text{RHS} &= \det(\lambda I_n - A)I_n = p_A(\lambda)I_n \\ &= (\lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-1} + \cdots + a_1\lambda + a_0)I_n \\ &= \lambda^n I_n + \lambda^{n-1}a_{n-1}I_n + \lambda^{n-2}a_{n-2}I_n + \cdots + \lambda a_1 I_n + a_0 I_n.\end{aligned}$$

The corresponding coefficients in front of λ^i (for $i \in \{0, 1, \ldots, n\}$) must be equal on the left-hand-side (LHS) and the right-hand-side (RHS).

Proof (continued). Reminder: $n \geq 2$,

$$\underbrace{(\lambda I_n - A)(\underbrace{\lambda^{n-1}B_{n-1} + \lambda^{n-2}B_{n-2} + \cdots + \lambda B_1 + B_0}_{=\mathrm{adj}(\lambda I_n - A)})}_{:=\mathrm{LHS}} = \underbrace{\det(\lambda I_n - A) I_n}_{:=\mathrm{RHS}}.$$

For the left-hand-side, we have

$$\begin{aligned}\mathrm{LHS} &= (\lambda I_n - A)(\lambda^{n-1}B_{n-1} + \cdots + \lambda B_1 + B_0) \\ &= \lambda^n B_{n-1} + \lambda^{n-1}(B_{n-2} - AB_{n-1}) + \lambda^{n-2}(B_{n-3} - AB_{n-2}) + \\ &\quad + \cdots + \lambda(B_0 - AB_1) - AB_0.\end{aligned}$$

For the right-hand-side, we have

$$\begin{aligned}\mathrm{RHS} &= \det(\lambda I_n - A) I_n = p_A(\lambda) I_n \\ &= (\lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-1} + \cdots + a_1\lambda + a_0) I_n \\ &= \lambda^n I_n + \lambda^{n-1} a_{n-1} I_n + \lambda^{n-2} a_{n-2} I_n + \cdots + \lambda a_1 I_n + a_0 I_n.\end{aligned}$$

The corresponding coefficients in front of λ^i (for $i \in \{0, 1, \ldots, n\}$) must be equal on the left-hand-side (LHS) and the right-hand-side (RHS). This yields the following $n+1$ equations (next slide).

Proof (continued).

$$\begin{aligned}
B_{n-1} &= I_n \\
B_{n-2} - AB_{n-1} &= a_{n-1} I_n \\
B_{n-3} - AB_{n-2} &= a_{n-2} I_n \\
&\vdots \\
B_0 - AB_1 &= a_1 I_n \\
-AB_0 &= a_0 I_n
\end{aligned}$$

Proof (continued).

$$\begin{array}{rcl} B_{n-1} & = & I_n \\ B_{n-2} - AB_{n-1} & = & a_{n-1} I_n \\ B_{n-3} - AB_{n-2} & = & a_{n-2} I_n \\ & \vdots & \\ B_0 - AB_1 & = & a_1 I_n \\ -AB_0 & = & a_0 I_n \end{array}$$

We now multiply the first (top) equation by A^n on the left, the second equation by A^{n-1} on the left, the third equation by A^{n-2} on the left, and so on.

Proof (continued).

$$\begin{aligned} B_{n-1} &= I_n \\ B_{n-2} - AB_{n-1} &= a_{n-1} I_n \\ B_{n-3} - AB_{n-2} &= a_{n-2} I_n \\ &\vdots \\ B_0 - AB_1 &= a_1 I_n \\ -AB_0 &= a_0 I_n \end{aligned}$$

We now multiply the first (top) equation by A^n on the left, the second equation by A^{n-1} on the left, the third equation by A^{n-2} on the left, and so on. This yields the following.

$$\begin{aligned} A^n B_{n-1} &= A^n \\ A^{n-1} B_{n-2} - A^n B_{n-1} &= a_{n-1} A^{n-1} \\ A^{n-2} B_{n-3} - A^{n-1} B_{n-2} &= a_{n-2} A^{n-2} \\ &\vdots \\ AB_0 - A^2 B_1 &= a_1 A \\ -AB_0 &= a_0 I_n \end{aligned}$$

Proof (continued). Reminder:

$$
\begin{aligned}
A^n B_{n-1} &= A^n \\
A^{n-1} B_{n-2} - A^n B_{n-1} &= a_{n-1} A^{n-1} \\
A^{n-2} B_{n-3} - A^{n-1} B_{n-2} &= a_{n-2} A^{n-2} \\
&\vdots \\
AB_0 - A^2 B_1 &= a_1 A \\
-AB_0 &= a_0 I_n
\end{aligned}
$$

Proof (continued). Reminder:

$$
\begin{array}{rcl}
A^n B_{n-1} & = & A^n \\
A^{n-1} B_{n-2} - A^n B_{n-1} & = & a_{n-1} A^{n-1} \\
A^{n-2} B_{n-3} - A^{n-1} B_{n-2} & = & a_{n-2} A^{n-2} \\
& \vdots & \\
AB_0 - A^2 B_1 & = & a_1 A \\
-AB_0 & = & a_0 I_n
\end{array}
$$

We now add up the equations that we obtained.

Proof (continued). Reminder:

$$
\begin{array}{rcl}
A^n B_{n-1} & = & A^n \\
A^{n-1} B_{n-2} - A^n B_{n-1} & = & a_{n-1} A^{n-1} \\
A^{n-2} B_{n-3} - A^{n-1} B_{n-2} & = & a_{n-2} A^{n-2} \\
 & \vdots & \\
A B_0 - A^2 B_1 & = & a_1 A \\
-A B_0 & = & a_0 I_n
\end{array}
$$

We now add up the equations that we obtained.

On the left-hand-side, the sum is obviously $O_{n \times n}$.

Proof (continued). Reminder:

$$\begin{array}{rcl} A^n B_{n-1} & = & A^n \\ A^{n-1} B_{n-2} - A^n B_{n-1} & = & a_{n-1} A^{n-1} \\ A^{n-2} B_{n-3} - A^{n-1} B_{n-2} & = & a_{n-2} A^{n-2} \\ & \vdots & \\ AB_0 - A^2 B_1 & = & a_1 A \\ -AB_0 & = & a_0 I_n \end{array}$$

We now add up the equations that we obtained.

On the left-hand-side, the sum is obviously $O_{n \times n}$.

So, the right-hand-side must also sum up to $O_{n \times n}$, i.e.

$$A^n + a_{n-1} A^{n-1} + a_{n-2} A^{n-2} + \cdots + a_1 A + a_0 I_n = O_{n \times n}.$$

But this is precisely what we needed to show. □

The Cayley-Hamilton theorem

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n \times n}$ be a square matrix, and let $p_A(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$ be the characteristic polynomial of A. Then

$$A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I_n \quad = \quad O_{n \times n}.$$

The Cayley-Hamilton theorem

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n\times n}$ be a square matrix, and let $p_A(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$ be the characteristic polynomial of A. Then

$$A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I_n \quad = \quad O_{n\times n}.$$

Corollary 8.3.1

Let $\mathbb{F}$ be a field. For all matrices $A \in \mathbb{F}^{n\times n}$:

(a) $A^n \in \text{Span}(I_n, A, A^2, \dots, A^{n-1})$, i.e. A^n is a linear combination of $I_n, A, A^2, \dots, A^{n-1}$;

(b) if A is invertible, then $A^{-1} \in \text{Span}(I_n, A, A^2, \dots, A^{n-1})$, i.e. A^{-1} is a linear combination of $I_n, A, A^2, \dots, A^{n-1}$.

Proof.

The Cayley-Hamilton theorem

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n \times n}$ be a square matrix, and let $p_A(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$ be the characteristic polynomial of A. Then

$$A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I_n \quad = \quad O_{n \times n}.$$

Corollary 8.3.1

Let $\mathbb{F}$ be a field. For all matrices $A \in \mathbb{F}^{n \times n}$:

- (a) $A^n \in \text{Span}(I_n, A, A^2, \dots, A^{n-1})$, i.e. A^n is a linear combination of $I_n, A, A^2, \dots, A^{n-1}$;
- (b) if A is invertible, then $A^{-1} \in \text{Span}(I_n, A, A^2, \dots, A^{n-1})$, i.e. A^{-1} is a linear combination of $I_n, A, A^2, \dots, A^{n-1}$.

Proof. Fix a matrix $A \in \mathbb{F}^{n \times n}$, and consider its characteristic polynomial $p_A(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-2} + \cdots + a_1\lambda + a_0$.

Corollary 8.3.1

(a) $A^n \in \text{Span}(I_n, A, A^2, \ldots, A^{n-1})$, i.e. A^n is a linear combination of $I_n, A, A^2, \ldots, A^{n-1}$;

- Reminder:
  $p_A(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-2} + \cdots + a_1\lambda + a_0.$

Proof of (a).

Corollary 8.3.1

(a) $A^n \in \text{Span}(I_n, A, A^2, \ldots, A^{n-1})$, i.e. A^n is a linear combination of $I_n, A, A^2, \ldots, A^{n-1}$;

- Reminder:
  $p_A(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-2} + \cdots + a_1\lambda + a_0$.

Proof of (a). By the Cayley-Hamilton theorem, we have that

$$A^n + a_{n-1}A^{n-1} + \cdots + a_a A^2 + a_1 A + a_0 I_n \quad = \quad O_{n\times n}.$$

Corollary 8.3.1

(a) $A^n \in \text{Span}(I_n, A, A^2, \ldots, A^{n-1})$, i.e. A^n is a linear combination of $I_n, A, A^2, \ldots, A^{n-1}$;

- Reminder:
  $p_A(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-2} + \cdots + a_1\lambda + a_0$.

Proof of (a). By the Cayley-Hamilton theorem, we have that

$$A^n + a_{n-1}A^{n-1} + \cdots + a_a A^2 + a_1 A + a_0 I_n \;=\; O_{n\times n}.$$

Consequently,

$$A^n \;=\; -a_0 I_n - a_1 A - a_2 A^2 - \cdots - a_{n-1}A^{n-1}.$$

Corollary 8.3.1

(a) $A^n \in \text{Span}(I_n, A, A^2, \ldots, A^{n-1})$, i.e. A^n is a linear combination of $I_n, A, A^2, \ldots, A^{n-1}$;

- Reminder:
  $p_A(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-2} + \cdots + a_1\lambda + a_0$.

Proof of (a). By the Cayley-Hamilton theorem, we have that

$$A^n + a_{n-1}A^{n-1} + \cdots + a_a A^2 + a_1 A + a_0 I_n \quad = \quad O_{n\times n}.$$

Consequently,

$$A^n \quad = \quad -a_0 I_n - a_1 A - a_2 A^2 - \cdots - a_{n-1}A^{n-1}.$$

Thus, A^n is a linear combination of the matrices $I_n, A, A^2, \ldots, A^{n-1}$.

Corollary 8.3.1

(b) if A is invertible, then $A^{-1} \in \text{Span}(I_n, A, A^2, \ldots, A^{n-1})$, i.e. A^{-1} is a linear combination of $I_n, A, A^2, \ldots, A^{n-1}$.

- Reminder:
  $p_A(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-2} + \cdots + a_1\lambda + a_0.$

Proof of (b). Assume that A is invertible.

Corollary 8.3.1

(b) if A is invertible, then $A^{-1} \in \text{Span}(I_n, A, A^2, \dots, A^{n-1})$, i.e. A^{-1} is a linear combination of $I_n, A, A^2, \dots, A^{n-1}$.

- Reminder:
  $p_A(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-2} + \cdots + a_1\lambda + a_0$.

Proof of (b). Assume that A is invertible. Proposition 8.2.11 then guarantees that 0 is not an eigenvalue of A.

Corollary 8.3.1

(b) if A is invertible, then $A^{-1} \in \text{Span}(I_n, A, A^2, \dots, A^{n-1})$, i.e. A^{-1} is a linear combination of $I_n, A, A^2, \dots, A^{n-1}$.

- Reminder:
  $p_A(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-2} + \cdots + a_1\lambda + a_0$.

Proof of (b). Assume that A is invertible. Proposition 8.2.11 then guarantees that 0 is not an eigenvalue of A. Since the eigenvalues of A are precisely the roots of the characteristic polynomial of A, we have that $p_A(0) \neq 0$; since $p_A(0) = a_0$, it follows that $a_0 \neq 0$.

Corollary 8.3.1

(b) if A is invertible, then $A^{-1} \in \text{Span}(I_n, A, A^2, \dots, A^{n-1})$, i.e. A^{-1} is a linear combination of $I_n, A, A^2, \dots, A^{n-1}$.

- Reminder:
  $p_A(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-2} + \cdots + a_1\lambda + a_0$.

Proof of (b). Assume that A is invertible. Proposition 8.2.11 then guarantees that 0 is not an eigenvalue of A. Since the eigenvalues of A are precisely the roots of the characteristic polynomial of A, we have that $p_A(0) \neq 0$; since $p_A(0) = a_0$, it follows that $a_0 \neq 0$.

Now, by the Cayley-Hamilton theorem, we have that

$$A^n + a_{n-1}A^{n-1} + \cdots + a_2A^2 + a_1A + a_0I_n \;=\; O_{n\times n}.$$

Corollary 8.3.1

(b) if A is invertible, then $A^{-1} \in \mathrm{Span}(I_n, A, A^2, \ldots, A^{n-1})$, i.e. A^{-1} is a linear combination of $I_n, A, A^2, \ldots, A^{n-1}$.

Proof of (b) (continued). Reminder: $a_0 \neq 0$,

$$A^n + a_{n-1}A^{n-1} + \cdots + a_2A^2 + a_1A + a_0I_n \quad = \quad O_{n\times n}.$$

Corollary 8.3.1

(b) if A is invertible, then $A^{-1} \in \text{Span}(I_n, A, A^2, \dots, A^{n-1})$, i.e. A^{-1} is a linear combination of $I_n, A, A^2, \dots, A^{n-1}$.

Proof of (b) (continued). Reminder: $a_0 \neq 0$,

$$A^n + a_{n-1}A^{n-1} + \cdots + a_2A^2 + a_1A + a_0I_n \quad = \quad O_{n\times n}.$$

We multiply both sides of the equation by A^{-1} on the right, and we obtain

$$A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_2A + a_1I_n + a_0A^{-1} \quad = \quad O_{n\times n},$$

Corollary 8.3.1

(b) if A is invertible, then $A^{-1} \in \text{Span}(I_n, A, A^2, \dots, A^{n-1})$, i.e. A^{-1} is a linear combination of $I_n, A, A^2, \dots, A^{n-1}$.

Proof of (b) (continued). Reminder: $a_0 \neq 0$,

$$A^n + a_{n-1}A^{n-1} + \cdots + a_2A^2 + a_1A + a_0I_n = O_{n\times n}.$$

We multiply both sides of the equation by A^{-1} on the right, and we obtain

$$A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_2A + a_1I_n + a_0A^{-1} = O_{n\times n},$$

and consequently,

$$a_0A^{-1} = -a_1I_n - a_2A - \cdots - a_{n-1}A^{n-2} - A^{n-1}.$$

Corollary 8.3.1

(b) if A is invertible, then $A^{-1} \in \text{Span}(I_n, A, A^2, \dots, A^{n-1})$, i.e. A^{-1} is a linear combination of $I_n, A, A^2, \dots, A^{n-1}$.

Proof of (b) (continued). Reminder: $a_0 \neq 0$,

$$A^n + a_{n-1}A^{n-1} + \cdots + a_2A^2 + a_1A + a_0I_n \quad = \quad O_{n\times n}.$$

We multiply both sides of the equation by A^{-1} on the right, and we obtain

$$A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_2A + a_1I_n + a_0A^{-1} \quad = \quad O_{n\times n},$$

and consequently,

$$a_0A^{-1} \quad = \quad -a_1I_n - a_2A - \cdots - a_{n-1}A^{n-2} - A^{n-1}.$$

Since $a_0 \neq 0$, this implies that

$$A^{-1} \quad = \quad -\tfrac{a_1}{a_0}I_n - \tfrac{a_2}{a_0}A - \cdots - \tfrac{a_{n-1}}{a_0}A^{n-2} - \tfrac{1}{a_0}A^{n-1}.$$

Corollary 8.3.1

(b) if A is invertible, then $A^{-1} \in \text{Span}(I_n, A, A^2, \dots, A^{n-1})$, i.e. A^{-1} is a linear combination of $I_n, A, A^2, \dots, A^{n-1}$.

Proof of (b) (continued). Reminder: $a_0 \neq 0$,

$$A^n + a_{n-1}A^{n-1} + \cdots + a_2A^2 + a_1A + a_0I_n = O_{n\times n}.$$

We multiply both sides of the equation by A^{-1} on the right, and we obtain

$$A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_2A + a_1I_n + a_0A^{-1} = O_{n\times n},$$

and consequently,

$$a_0A^{-1} = -a_1I_n - a_2A - \cdots - a_{n-1}A^{n-2} - A^{n-1}.$$

Since $a_0 \neq 0$, this implies that

$$A^{-1} = -\frac{a_1}{a_0}I_n - \frac{a_2}{a_0}A - \cdots - \frac{a_{n-1}}{a_0}A^{n-2} - \frac{1}{a_0}A^{n-1}.$$

So, A^{-1} is a linear combination of $I_n, A, A^2, \dots, A^{n-1}$. $\square$

The Cayley-Hamilton theorem

Let $\mathbb{F}$ be a field, let $A \in \mathbb{F}^{n\times n}$ be a square matrix, and let $p_A(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$ be the characteristic polynomial of A. Then

$$A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I_n \quad = \quad O_{n\times n}.$$

Corollary 8.3.1

Let $\mathbb{F}$ be a field. For all matrices $A \in \mathbb{F}^{n\times n}$:

- (a) $A^n \in \text{Span}(I_n, A, A^2, \dots, A^{n-1})$, i.e. A^n is a linear combination of $I_n, A, A^2, \dots, A^{n-1}$;
- (b) if A is invertible, then $A^{-1} \in \text{Span}(I_n, A, A^2, \dots, A^{n-1})$, i.e. A^{-1} is a linear combination of $I_n, A, A^2, \dots, A^{n-1}$.

2. Eigenvectors and linear independence. Eigenbases

2. Eigenvectors and linear independence. Eigenbases

Definition

For a finite-dimensional vector space V over a field $\mathbb{F}$ and a linear function $f : V \to V$, an *eigenbasis* of V associated with f is a basis $\mathcal{B}$ of V s.t. all vectors in $\mathcal{B}$ are eigenvectors of f.

Definition

For an field $\mathbb{F}$ and a matrix $A \in \mathbb{F}^{n \times n}$, an *eigenbasis* of $\mathbb{F}^n$ associated with A is a basis $\mathcal{B}$ of $\mathbb{F}^n$ s.t. all vectors in $\mathcal{B}$ are eigenvectors of A.

2. Eigenvectors and linear independence. Eigenbases

Definition

For a finite-dimensional vector space V over a field $\mathbb{F}$ and a linear function $f : V \to V$, an *eigenbasis* of V associated with f is a basis $\mathcal{B}$ of V s.t. all vectors in $\mathcal{B}$ are eigenvectors of f.

Definition

For an field $\mathbb{F}$ and a matrix $A \in \mathbb{F}^{n\times n}$, an *eigenbasis* of $\mathbb{F}^n$ associated with A is a basis $\mathcal{B}$ of $\mathbb{F}^n$ s.t. all vectors in $\mathcal{B}$ are eigenvectors of A.

- Eigenbases do not always exist, and one of our goals in this section is to determine when they do and do not exist.

2. Eigenvectors and linear independence. Eigenbases

Definition

For a finite-dimensional vector space V over a field $\mathbb{F}$ and a linear function $f : V \to V$, an *eigenbasis* of V associated with f is a basis $\mathcal{B}$ of V s.t. all vectors in $\mathcal{B}$ are eigenvectors of f.

Definition

For an field $\mathbb{F}$ and a matrix $A \in \mathbb{F}^{n \times n}$, an *eigenbasis* of $\mathbb{F}^n$ associated with A is a basis $\mathcal{B}$ of $\mathbb{F}^n$ s.t. all vectors in $\mathcal{B}$ are eigenvectors of A.

- Eigenbases do not always exist, and one of our goals in this section is to determine when they do and do not exist.
- As we shall see (later!), eigenbases play a crucial role in matrix "diagonalization."

Proposition 8.4.1

Let V be a vector space over a field $\mathbb{F}$, let $f : V \to V$ be a linear function, let $\lambda_1, \ldots, \lambda_k \in \mathbb{F}$ be pairwise distinct eigenvalues of f, associated with eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$, respectively. Then $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is a linearly independent set.

Proof.

Proposition 8.4.1

Let V be a vector space over a field $\mathbb{F}$, let $f : V \to V$ be a linear function, let $\lambda_1, \ldots, \lambda_k \in \mathbb{F}$ be pairwise distinct eigenvalues of f, associated with eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$, respectively. Then $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is a linearly independent set.

Proof. We will prove inductively that for all $i \in \{0, \ldots, k\}$, the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_i\}$ is linearly independent.

Proposition 8.4.1

Let V be a vector space over a field $\mathbb{F}$, let $f : V \to V$ be a linear function, let $\lambda_1, \dots, \lambda_k \in \mathbb{F}$ be pairwise distinct eigenvalues of f, associated with eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_k$, respectively. Then $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is a linearly independent set.

Proof. We will prove inductively that for all $i \in \{0, \dots, k\}$, the set $\{\mathbf{v}_1, \dots, \mathbf{v}_i\}$ is linearly independent.

For $i = 0$, we have that $\{\mathbf{v}_1, \dots, \mathbf{v}_i\} = \emptyset$, which is obviously a linearly independent set.

Proposition 8.4.1

Let V be a vector space over a field $\mathbb{F}$, let $f : V \to V$ be a linear function, let $\lambda_1, \dots, \lambda_k \in \mathbb{F}$ be pairwise distinct eigenvalues of f, associated with eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_k$, respectively. Then $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is a linearly independent set.

Proof. We will prove inductively that for all $i \in \{0, \dots, k\}$, the set $\{\mathbf{v}_1, \dots, \mathbf{v}_i\}$ is linearly independent.

For $i = 0$, we have that $\{\mathbf{v}_1, \dots, \mathbf{v}_i\} = \emptyset$, which is obviously a linearly independent set.

Now, fix an index $i \in \{0, \dots, k-1\}$, and assume inductively that the set $\{\mathbf{v}_1, \dots, \mathbf{v}_i\}$ is linearly independent. We must show that $\{\mathbf{v}_1, \dots, \mathbf{v}_i, \mathbf{v}_{i+1}\}$ is linearly independent.

Proposition 8.4.1

Let V be a vector space over a field $\mathbb{F}$, let $f : V \to V$ be a linear function, let $\lambda_1, \dots, \lambda_k \in \mathbb{F}$ be pairwise distinct eigenvalues of f, associated with eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_k$, respectively. Then $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is a linearly independent set.

Proof. We will prove inductively that for all $i \in \{0, \dots, k\}$, the set $\{\mathbf{v}_1, \dots, \mathbf{v}_i\}$ is linearly independent.

For $i = 0$, we have that $\{\mathbf{v}_1, \dots, \mathbf{v}_i\} = \emptyset$, which is obviously a linearly independent set.

Now, fix an index $i \in \{0, \dots, k-1\}$, and assume inductively that the set $\{\mathbf{v}_1, \dots, \mathbf{v}_i\}$ is linearly independent. We must show that $\{\mathbf{v}_1, \dots, \mathbf{v}_i, \mathbf{v}_{i+1}\}$ is linearly independent. Fix scalars $\alpha_1, \dots, \alpha_i, \alpha_{i+1} \in \mathbb{F}$ s.t.

$$\alpha_1 \mathbf{v}_1 + \dots + \alpha_i \mathbf{v}_i + \alpha_{i+1} \mathbf{v}_{i+1} = \mathbf{0}.$$

WTS $\alpha_1 = \dots = \alpha_i = \alpha_{i+1} = 0$.

Proposition 8.4.1

Let V be a vector space over a field $\mathbb{F}$, let $f : V \to V$ be a linear function, let $\lambda_1, \dots, \lambda_k \in \mathbb{F}$ be pairwise distinct eigenvalues of f, associated with eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_k$, respectively. Then $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is a linearly independent set.

Proof (continued). Reminder: $\{\mathbf{v}_1, \dots, \mathbf{v}_i\}$ is linearly independent; $\alpha_1\mathbf{v}_1 + \dots + \alpha_i\mathbf{v}_i + \alpha_{i+1}\mathbf{v}_{i+1} = \mathbf{0}$; WTS $\alpha_1 = \dots = \alpha_i = \alpha_{i+1} = 0$.

Proposition 8.4.1

Let V be a vector space over a field $\mathbb{F}$, let $f : V \to V$ be a linear function, let $\lambda_1, \dots, \lambda_k \in \mathbb{F}$ be pairwise distinct eigenvalues of f, associated with eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_k$, respectively. Then $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is a linearly independent set.

Proof (continued). Reminder: $\{\mathbf{v}_1, \dots, \mathbf{v}_i\}$ is linearly independent; $\alpha_1\mathbf{v}_1 + \dots + \alpha_i\mathbf{v}_i + \alpha_{i+1}\mathbf{v}_{i+1} = \mathbf{0}$; WTS $\alpha_1 = \dots = \alpha_i = \alpha_{i+1} = 0$.

If we multiply both sides of the equation above by λ_{i+1}, we obtain

1. $\lambda_{i+1}\alpha_1\mathbf{v}_1 + \dots + \lambda_{i+1}\alpha_i\mathbf{v}_i + \lambda_{i+1}\alpha_{i+1}\mathbf{v}_{i+1} = \mathbf{0}$.

Proposition 8.4.1

Let V be a vector space over a field $\mathbb{F}$, let $f : V \to V$ be a linear function, let $\lambda_1, \dots, \lambda_k \in \mathbb{F}$ be pairwise distinct eigenvalues of f, associated with eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_k$, respectively. Then $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is a linearly independent set.

Proof (continued). Reminder: $\{\mathbf{v}_1, \dots, \mathbf{v}_i\}$ is linearly independent; $\alpha_1 \mathbf{v}_1 + \dots + \alpha_i \mathbf{v}_i + \alpha_{i+1} \mathbf{v}_{i+1} = \mathbf{0}$; WTS $\alpha_1 = \dots = \alpha_i = \alpha_{i+1} = 0$.

If we multiply both sides of the equation above by λ_{i+1}, we obtain

1. $\lambda_{i+1}\alpha_1 \mathbf{v}_1 + \dots + \lambda_{i+1}\alpha_i \mathbf{v}_i + \lambda_{i+1}\alpha_{i+1} \mathbf{v}_{i+1} = \mathbf{0}$.

If, on the other hand, we apply the function f to both sides and also use the fact that $f(\mathbf{0}) = \mathbf{0}$, then we obtain

2. $f\big(\alpha_1 \mathbf{v}_1 + \dots + \alpha_i \mathbf{v}_i + \alpha_{i+1} \mathbf{v}_{i+1}\big) = \mathbf{0}$.

Proposition 8.4.1

Let V be a vector space over a field $\mathbb{F}$, let $f : V \to V$ be a linear function, let $\lambda_1, \dots, \lambda_k \in \mathbb{F}$ be pairwise distinct eigenvalues of f, associated with eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_k$, respectively. Then $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is a linearly independent set.

Proof (continued). We now compute:

$$\begin{aligned} \mathbf{0} &\stackrel{(2)}{=} f(\alpha_1\mathbf{v}_1 + \dots + \alpha_i\mathbf{v}_i + \alpha_{i+1}\mathbf{v}_{i+1}) \\ &\stackrel{(*)}{=} \alpha_1 f(\mathbf{v}_1) + \dots + \alpha_i f(\mathbf{v}_i) + \alpha_{i+1} f(\mathbf{v}_{i+1}) \\ &\stackrel{(**)}{=} \alpha_1\lambda_1\mathbf{v}_1 + \dots + \alpha_i\lambda_i\mathbf{v}_i + \alpha_{i+1}\lambda_{i+1}\mathbf{v}_{i+1}, \end{aligned}$$

where (*) follows from the linearity of f, and (**) follows from the fact that $\mathbf{v}_1, \dots, \mathbf{v}_i, \mathbf{v}_{i+1}$ are eigenvectors of f associated with eigenvalues $\lambda_1, \dots, \lambda_i, \lambda_{i+1}$, respectively.

Proposition 8.4.1

Let V be a vector space over a field $\mathbb{F}$, let $f : V \to V$ be a linear function, let $\lambda_1, \dots, \lambda_k \in \mathbb{F}$ be pairwise distinct eigenvalues of f, associated with eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_k$, respectively. Then $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is a linearly independent set.

Proof (continued). Reminder:

1. $\lambda_{i+1}\alpha_1\mathbf{v}_1 + \dots + \lambda_{i+1}\alpha_i\mathbf{v}_i + \lambda_{i+1}\alpha_{i+1}\mathbf{v}_{i+1} = \mathbf{0}$;
2. $f\big(\alpha_1\mathbf{v}_1 + \dots + \alpha_i\mathbf{v}_i + \alpha_{i+1}\mathbf{v}_{i+1}\big) = \mathbf{0}$;
3. $\alpha_1\lambda_1\mathbf{v}_1 + \dots + \alpha_i\lambda_i\mathbf{v}_i + \alpha_{i+1}\lambda_{i+1}\mathbf{v}_{i+1} = \mathbf{0}$.

Proposition 8.4.1

Let V be a vector space over a field $\mathbb{F}$, let $f : V \to V$ be a linear function, let $\lambda_1, \dots, \lambda_k \in \mathbb{F}$ be pairwise distinct eigenvalues of f, associated with eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_k$, respectively. Then $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is a linearly independent set.

Proof (continued). Reminder:

1. $\lambda_{i+1}\alpha_1\mathbf{v}_1 + \dots + \lambda_{i+1}\alpha_i\mathbf{v}_i + \lambda_{i+1}\alpha_{i+1}\mathbf{v}_{i+1} = \mathbf{0}$;
2. $f\big(\alpha_1\mathbf{v}_1 + \dots + \alpha_i\mathbf{v}_i + \alpha_{i+1}\mathbf{v}_{i+1}\big) = \mathbf{0}$;
3. $\alpha_1\lambda_1\mathbf{v}_1 + \dots + \alpha_i\lambda_i\mathbf{v}_i + \alpha_{i+1}\lambda_{i+1}\mathbf{v}_{i+1} = \mathbf{0}$.

Combining (1) and (3), we obtain:

$$\begin{aligned} & \alpha_1\lambda_1\mathbf{v}_1 + \dots + \alpha_i\lambda_i\mathbf{v}_i + \alpha_{i+1}\lambda_{i+1}\mathbf{v}_{i+1} \\ = \; & \lambda_{i+1}\alpha_1\mathbf{v}_1 + \dots + \lambda_{i+1}\alpha_i\mathbf{v}_i + \lambda_{i+1}\alpha_{i+1}\mathbf{v}_{i+1}. \end{aligned}$$

Proposition 8.4.1

Let V be a vector space over a field $\mathbb{F}$, let $f : V \to V$ be a linear function, let $\lambda_1, \dots, \lambda_k \in \mathbb{F}$ be pairwise distinct eigenvalues of f, associated with eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_k$, respectively. Then $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is a linearly independent set.

Proof (continued). Reminder:

1. $\lambda_{i+1}\alpha_1\mathbf{v}_1 + \dots + \lambda_{i+1}\alpha_i\mathbf{v}_i + \lambda_{i+1}\alpha_{i+1}\mathbf{v}_{i+1} = \mathbf{0}$;
2. $f(\alpha_1\mathbf{v}_1 + \dots + \alpha_i\mathbf{v}_i + \alpha_{i+1}\mathbf{v}_{i+1}) = \mathbf{0}$;
3. $\alpha_1\lambda_1\mathbf{v}_1 + \dots + \alpha_i\lambda_i\mathbf{v}_i + \alpha_{i+1}\lambda_{i+1}\mathbf{v}_{i+1} = \mathbf{0}$.

Combining (1) and (3), we obtain:

$$\begin{aligned} & \alpha_1\lambda_1\mathbf{v}_1 + \dots + \alpha_i\lambda_i\mathbf{v}_i + \alpha_{i+1}\lambda_{i+1}\mathbf{v}_{i+1} \\ = \quad & \lambda_{i+1}\alpha_1\mathbf{v}_1 + \dots + \lambda_{i+1}\alpha_i\mathbf{v}_i + \lambda_{i+1}\alpha_{i+1}\mathbf{v}_{i+1}. \end{aligned}$$

By subtracting one side from the other and factoring, we get

$$\alpha_1(\lambda_1 - \lambda_{i+1})\mathbf{v}_1 + \dots + \alpha_i(\lambda_i - \lambda_{i+1})\mathbf{v}_i \quad = \quad \mathbf{0}.$$

Proposition 8.4.1

Let V be a vector space over a field $\mathbb{F}$, let $f : V \to V$ be a linear function, let $\lambda_1, \dots, \lambda_k \in \mathbb{F}$ be pairwise distinct eigenvalues of f, associated with eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_k$, respectively. Then $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is a linearly independent set.

Proof (continued). Reminder:

$$\alpha_1(\lambda_1 - \lambda_{i+1})\mathbf{v}_1 + \dots + \alpha_i(\lambda_i - \lambda_{i+1})\mathbf{v}_i \quad = \quad \mathbf{0}$$

Proposition 8.4.1

Let V be a vector space over a field $\mathbb{F}$, let $f : V \to V$ be a linear function, let $\lambda_1, \dots, \lambda_k \in \mathbb{F}$ be pairwise distinct eigenvalues of f, associated with eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_k$, respectively. Then $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is a linearly independent set.

Proof (continued). Reminder:

$$\alpha_1(\lambda_1 - \lambda_{i+1})\mathbf{v}_1 + \cdots + \alpha_i(\lambda_i - \lambda_{i+1})\mathbf{v}_i \quad = \quad \mathbf{0}$$

By the ind. hyp., $\mathbf{v}_1, \dots, \mathbf{v}_i$ are linearly independent,

Proposition 8.4.1

Let V be a vector space over a field $\mathbb{F}$, let $f : V \to V$ be a linear function, let $\lambda_1, \dots, \lambda_k \in \mathbb{F}$ be pairwise distinct eigenvalues of f, associated with eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_k$, respectively. Then $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is a linearly independent set.

Proof (continued). Reminder:

$$\alpha_1(\lambda_1 - \lambda_{i+1})\mathbf{v}_1 + \dots + \alpha_i(\lambda_i - \lambda_{i+1})\mathbf{v}_i \quad = \quad \mathbf{0}$$

By the ind. hyp., $\mathbf{v}_1, \dots, \mathbf{v}_i$ are linearly independent, and it follows that $\alpha_1(\lambda_1 - \lambda_{i+1}) = \dots = \alpha_i(\lambda_i - \lambda_{i+1}) = 0$.

Proposition 8.4.1

Let V be a vector space over a field $\mathbb{F}$, let $f : V \to V$ be a linear function, let $\lambda_1, \dots, \lambda_k \in \mathbb{F}$ be pairwise distinct eigenvalues of f, associated with eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_k$, respectively. Then $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is a linearly independent set.

Proof (continued). Reminder:

$$\alpha_1(\lambda_1 - \lambda_{i+1})\mathbf{v}_1 + \dots + \alpha_i(\lambda_i - \lambda_{i+1})\mathbf{v}_i = \mathbf{0}$$

By the ind. hyp., $\mathbf{v}_1, \dots, \mathbf{v}_i$ are linearly independent, and it follows that $\alpha_1(\lambda_1 - \lambda_{i+1}) = \dots = \alpha_i(\lambda_i - \lambda_{i+1}) = 0$. Since $\lambda_1 - \lambda_{i+1}, \dots, \lambda_i - \lambda_{i+1}$ are all non-zero (because $\lambda_1, \dots, \lambda_i, \lambda_{i+1}$ are pairwise distinct), we deduce that $\alpha_1 = \dots = \alpha_i = 0$.

Proposition 8.4.1

Let V be a vector space over a field $\mathbb{F}$, let $f : V \to V$ be a linear function, let $\lambda_1, \dots, \lambda_k \in \mathbb{F}$ be pairwise distinct eigenvalues of f, associated with eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_k$, respectively. Then $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is a linearly independent set.

Proof (continued). Reminder:

$$\alpha_1(\lambda_1 - \lambda_{i+1})\mathbf{v}_1 + \dots + \alpha_i(\lambda_i - \lambda_{i+1})\mathbf{v}_i \quad = \quad \mathbf{0}$$

By the ind. hyp., $\mathbf{v}_1, \dots, \mathbf{v}_i$ are linearly independent, and it follows that $\alpha_1(\lambda_1 - \lambda_{i+1}) = \dots = \alpha_i(\lambda_i - \lambda_{i+1}) = 0$. Since $\lambda_1 - \lambda_{i+1}, \dots, \lambda_i - \lambda_{i+1}$ are all non-zero (because $\lambda_1, \dots, \lambda_i, \lambda_{i+1}$ are pairwise distinct), we deduce that $\alpha_1 = \dots = \alpha_i = 0$. Plugging this into our equation $\alpha_1\mathbf{v}_1 + \dots + \alpha_i\mathbf{v}_i + \alpha_{i+1}\mathbf{v}_{i+1} = \mathbf{0}$, we get

$$\alpha_{i+1}\mathbf{v}_{i+1} \quad = \quad \mathbf{0}.$$

Proposition 8.4.1

Let V be a vector space over a field $\mathbb{F}$, let $f : V \to V$ be a linear function, let $\lambda_1, \dots, \lambda_k \in \mathbb{F}$ be pairwise distinct eigenvalues of f, associated with eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_k$, respectively. Then $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is a linearly independent set.

Proof (continued). Reminder:

$$\alpha_1(\lambda_1 - \lambda_{i+1})\mathbf{v}_1 + \dots + \alpha_i(\lambda_i - \lambda_{i+1})\mathbf{v}_i \quad = \quad \mathbf{0}$$

By the ind. hyp., $\mathbf{v}_1, \dots, \mathbf{v}_i$ are linearly independent, and it follows that $\alpha_1(\lambda_1 - \lambda_{i+1}) = \dots = \alpha_i(\lambda_i - \lambda_{i+1}) = 0$. Since $\lambda_1 - \lambda_{i+1}, \dots, \lambda_i - \lambda_{i+1}$ are all non-zero (because $\lambda_1, \dots, \lambda_i, \lambda_{i+1}$ are pairwise distinct), we deduce that $\alpha_1 = \dots = \alpha_i = 0$. Plugging this into our equation $\alpha_1\mathbf{v}_1 + \dots + \alpha_i\mathbf{v}_i + \alpha_{i+1}\mathbf{v}_{i+1} = \mathbf{0}$, we get

$$\alpha_{i+1}\mathbf{v}_{i+1} \quad = \quad \mathbf{0}.$$

But $\mathbf{v}_{i+1}$ is an eigenvector of f, and so by definition, $\mathbf{v}_{i+1} \neq \mathbf{0}$.

Proposition 8.4.1

Let V be a vector space over a field $\mathbb{F}$, let $f : V \to V$ be a linear function, let $\lambda_1, \dots, \lambda_k \in \mathbb{F}$ be pairwise distinct eigenvalues of f, associated with eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_k$, respectively. Then $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is a linearly independent set.

Proof (continued). Reminder:

$$\alpha_1(\lambda_1 - \lambda_{i+1})\mathbf{v}_1 + \dots + \alpha_i(\lambda_i - \lambda_{i+1})\mathbf{v}_i \quad = \quad \mathbf{0}$$

By the ind. hyp., $\mathbf{v}_1, \dots, \mathbf{v}_i$ are linearly independent, and it follows that $\alpha_1(\lambda_1 - \lambda_{i+1}) = \dots = \alpha_i(\lambda_i - \lambda_{i+1}) = 0$. Since $\lambda_1 - \lambda_{i+1}, \dots, \lambda_i - \lambda_{i+1}$ are all non-zero (because $\lambda_1, \dots, \lambda_i, \lambda_{i+1}$ are pairwise distinct), we deduce that $\alpha_1 = \dots = \alpha_i = 0$. Plugging this into our equation $\alpha_1\mathbf{v}_1 + \dots + \alpha_i\mathbf{v}_i + \alpha_{i+1}\mathbf{v}_{i+1} = \mathbf{0}$, we get

$$\alpha_{i+1}\mathbf{v}_{i+1} \quad = \quad \mathbf{0}.$$

But $\mathbf{v}_{i+1}$ is an eigenvector of f, and so by definition, $\mathbf{v}_{i+1} \neq \mathbf{0}$. So, $\alpha_{i+1} = 0$. Thus, $\alpha_1 = \dots = \alpha_i = \alpha_{i+1} = 0$.

Proposition 8.4.1

Let V be a vector space over a field $\mathbb{F}$, let $f : V \to V$ be a linear function, let $\lambda_1, \dots, \lambda_k \in \mathbb{F}$ be pairwise distinct eigenvalues of f, associated with eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_k$, respectively. Then $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is a linearly independent set.

Proof (continued). Reminder:

$$\alpha_1(\lambda_1 - \lambda_{i+1})\mathbf{v}_1 + \dots + \alpha_i(\lambda_i - \lambda_{i+1})\mathbf{v}_i \quad = \quad \mathbf{0}$$

By the ind. hyp., $\mathbf{v}_1, \dots, \mathbf{v}_i$ are linearly independent, and it follows that $\alpha_1(\lambda_1 - \lambda_{i+1}) = \dots = \alpha_i(\lambda_i - \lambda_{i+1}) = 0$. Since $\lambda_1 - \lambda_{i+1}, \dots, \lambda_i - \lambda_{i+1}$ are all non-zero (because $\lambda_1, \dots, \lambda_i, \lambda_{i+1}$ are pairwise distinct), we deduce that $\alpha_1 = \dots = \alpha_i = 0$. Plugging this into our equation $\alpha_1\mathbf{v}_1 + \dots + \alpha_i\mathbf{v}_i + \alpha_{i+1}\mathbf{v}_{i+1} = \mathbf{0}$, we get

$$\alpha_{i+1}\mathbf{v}_{i+1} \quad = \quad \mathbf{0}.$$

But $\mathbf{v}_{i+1}$ is an eigenvector of f, and so by definition, $\mathbf{v}_{i+1} \neq \mathbf{0}$. So, $\alpha_{i+1} = 0$. Thus, $\alpha_1 = \dots = \alpha_i = \alpha_{i+1} = 0$. So, $\{\mathbf{v}_1, \dots, \mathbf{v}_i, \mathbf{v}_{i+1}\}$ is linearly independent. This completes the induction. $\square$

Proposition 8.4.1

Let V be a vector space over a field $\mathbb{F}$, let $f : V \to V$ be a linear function, let $\lambda_1, \dots, \lambda_k \in \mathbb{F}$ be pairwise distinct eigenvalues of f, associated with eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_k$, respectively. Then $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is a linearly independent set.

Proposition 8.4.1

Let V be a vector space over a field $\mathbb{F}$, let $f : V \to V$ be a linear function, let $\lambda_1, \dots, \lambda_k \in \mathbb{F}$ be pairwise distinct eigenvalues of f, associated with eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_k$, respectively. Then $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is a linearly independent set.

Proposition 8.4.2

Let V be a vector space over a field $\mathbb{F}$, let $f : V \to V$ be a linear function, and let $\lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{F}$ be pairwise distinct eigenvalues of f. For each $i \in \{1, \dots, k\}$, let $\mathbf{v}_{i,1}, \dots, \mathbf{v}_{i,t_i}$ be linearly independent eigenvectors of f associated with the eigenvalue λ_i. Then the eigenvectors

$$\mathbf{v}_{1,1}, \dots, \mathbf{v}_{1,t_1}, \mathbf{v}_{2,1}, \dots, \mathbf{v}_{2,t_2}, \dots, \mathbf{v}_{k,1}, \dots, \mathbf{v}_{k,t_k}$$

are linearly independent.

Proof.

Proof. Fix $\alpha_{1,1}, \ldots, \alpha_{1,t_1}, \alpha_{2,1}, \ldots, \alpha_{2,t_2}, \ldots, \alpha_{k,1}, \ldots, \alpha_{k,t_k} \in \mathbb{F}$ s.t.

$$\sum_{i=1}^{k} \left(\alpha_{i,1} \mathbf{v}_{i,1} + \cdots + \alpha_{i,t_i} \mathbf{v}_{i,t_i} \right) = \mathbf{0}.$$

Proof. Fix $\alpha_{1,1}, \ldots, \alpha_{1,t_1}, \alpha_{2,1}, \ldots, \alpha_{2,t_2}, \ldots, \alpha_{k,1}, \ldots, \alpha_{k,t_k} \in \mathbb{F}$ s.t.

$$\sum_{i=1}^{k} \left(\alpha_{i,1}\mathbf{v}_{i,1} + \cdots + \alpha_{i,t_i}\mathbf{v}_{i,t_i}\right) = \mathbf{0}.$$

$\forall i \in \{1, \ldots, k\}$: set $\mathbf{v}_i := \alpha_{i,1}\mathbf{v}_{i,1} + \cdots + \alpha_{i,t_i}\mathbf{v}_{i,t_i}$, that is

- $\mathbf{v}_1 := \alpha_{1,1}\mathbf{v}_{1,1} + \cdots + \alpha_{1,t_1}\mathbf{v}_{1,t_1}$;
- $\mathbf{v}_2 := \alpha_{2,1}\mathbf{v}_{2,1} + \cdots + \alpha_{2,t_2}\mathbf{v}_{2,t_2}$;

  $\vdots$
- $\mathbf{v}_k := \alpha_{k,1}\mathbf{v}_{k,1} + \cdots + \alpha_{k,t_k}\mathbf{v}_{k,t_k}$.

Proof. Fix $\alpha_{1,1}, \dots, \alpha_{1,t_1}, \alpha_{2,1}, \dots, \alpha_{2,t_2}, \dots, \alpha_{k,1}, \dots, \alpha_{k,t_k} \in \mathbb{F}$ s.t.

$$\sum_{i=1}^{k} \left(\alpha_{i,1}\mathbf{v}_{i,1} + \dots + \alpha_{i,t_i}\mathbf{v}_{i,t_i} \right) = \mathbf{0}.$$

$\forall i \in \{1, \dots, k\}$: set $\mathbf{v}_i := \alpha_{i,1}\mathbf{v}_{i,1} + \dots + \alpha_{i,t_i}\mathbf{v}_{i,t_i}$, that is

- $\mathbf{v}_1 := \alpha_{1,1}\mathbf{v}_{1,1} + \dots + \alpha_{1,t_1}\mathbf{v}_{1,t_1}$;
- $\mathbf{v}_2 := \alpha_{2,1}\mathbf{v}_{2,1} + \dots + \alpha_{2,t_2}\mathbf{v}_{2,t_2}$;

  $\vdots$

- $\mathbf{v}_k := \alpha_{k,1}\mathbf{v}_{k,1} + \dots + \alpha_{k,t_k}\mathbf{v}_{k,t_k}$.

So, $\mathbf{v}_1 + \mathbf{v}_2 + \dots + \mathbf{v}_k = \mathbf{0}$. Now, note that for each $i \in \{1, \dots, k\}$, the vector $\mathbf{v}_i$ is a linear combination of vectors in $E_{\lambda_i}(f)$;

Proof. Fix $\alpha_{1,1},\dots,\alpha_{1,t_1},\alpha_{2,1},\dots,\alpha_{2,t_2},\dots,\alpha_{k,1},\dots,\alpha_{k,t_k}\in\mathbb{F}$ s.t.

$$\sum_{i=1}^{k}\Big(\alpha_{i,1}\mathbf{v}_{i,1}+\cdots+\alpha_{i,t_i}\mathbf{v}_{i,t_i}\Big) \quad = \quad \mathbf{0}.$$

$\forall i\in\{1,\dots,k\}$: set $\mathbf{v}_i := \alpha_{i,1}\mathbf{v}_{i,1}+\cdots+\alpha_{i,t_i}\mathbf{v}_{i,t_i}$, that is

- $\mathbf{v}_1 := \alpha_{1,1}\mathbf{v}_{1,1}+\cdots+\alpha_{1,t_1}\mathbf{v}_{1,t_1}$;
- $\mathbf{v}_2 := \alpha_{2,1}\mathbf{v}_{2,1}+\cdots+\alpha_{2,t_2}\mathbf{v}_{2,t_2}$;

  $\vdots$
- $\mathbf{v}_k := \alpha_{k,1}\mathbf{v}_{k,1}+\cdots+\alpha_{k,t_k}\mathbf{v}_{k,t_k}$.

So, $\mathbf{v}_1+\mathbf{v}_2+\cdots+\mathbf{v}_k=\mathbf{0}$. Now, note that for each $i\in\{1,\dots,k\}$, the vector $\mathbf{v}_i$ is a linear combination of vectors in $E_{\lambda_i}(f)$; since $E_{\lambda_i}(f)$ is a subsapce of V, it follows that $\mathbf{v}_i\in E_{\lambda_i}(f)$.

Proof. Fix $\alpha_{1,1}, \dots, \alpha_{1,t_1}, \alpha_{2,1}, \dots, \alpha_{2,t_2}, \dots, \alpha_{k,1}, \dots, \alpha_{k,t_k} \in \mathbb{F}$ s.t.

$$\sum_{i=1}^{k} \left(\alpha_{i,1}\mathbf{v}_{i,1} + \dots + \alpha_{i,t_i}\mathbf{v}_{i,t_i} \right) = \mathbf{0}.$$

$\forall i \in \{1, \dots, k\}$: set $\mathbf{v}_i := \alpha_{i,1}\mathbf{v}_{i,1} + \dots + \alpha_{i,t_i}\mathbf{v}_{i,t_i}$, that is

- $\mathbf{v}_1 := \alpha_{1,1}\mathbf{v}_{1,1} + \dots + \alpha_{1,t_1}\mathbf{v}_{1,t_1}$;
- $\mathbf{v}_2 := \alpha_{2,1}\mathbf{v}_{2,1} + \dots + \alpha_{2,t_2}\mathbf{v}_{2,t_2}$;

  $\vdots$
- $\mathbf{v}_k := \alpha_{k,1}\mathbf{v}_{k,1} + \dots + \alpha_{k,t_k}\mathbf{v}_{k,t_k}$.

So, $\mathbf{v}_1 + \mathbf{v}_2 + \dots + \mathbf{v}_k = \mathbf{0}$. Now, note that for each $i \in \{1, \dots, k\}$, the vector $\mathbf{v}_i$ is a linear combination of vectors in $E_{\lambda_i}(f)$; since $E_{\lambda_i}(f)$ is a subsapce of V, it follows that $\mathbf{v}_i \in E_{\lambda_i}(f)$. Consequently, $\forall i \in \{1, \dots, k\}$: $\mathbf{v}_i$ is either $\mathbf{0}$ or an eigenvector of f associated with the eigenvalue λ_i. WTS $\mathbf{v}_1 = \mathbf{v}_2 = \dots = \mathbf{v}_k = \mathbf{0}$.

Proof. Fix $\alpha_{1,1}, \dots, \alpha_{1,t_1}, \alpha_{2,1}, \dots, \alpha_{2,t_2}, \dots, \alpha_{k,1}, \dots, \alpha_{k,t_k} \in \mathbb{F}$ s.t.

$$\sum_{i=1}^{k} \left(\alpha_{i,1}\mathbf{v}_{i,1} + \dots + \alpha_{i,t_i}\mathbf{v}_{i,t_i} \right) \quad = \quad \mathbf{0}.$$

$\forall i \in \{1, \dots, k\}$: set $\mathbf{v}_i := \alpha_{i,1}\mathbf{v}_{i,1} + \dots + \alpha_{i,t_i}\mathbf{v}_{i,t_i}$, that is

- $\mathbf{v}_1 := \alpha_{1,1}\mathbf{v}_{1,1} + \dots + \alpha_{1,t_1}\mathbf{v}_{1,t_1}$;
- $\mathbf{v}_2 := \alpha_{2,1}\mathbf{v}_{2,1} + \dots + \alpha_{2,t_2}\mathbf{v}_{2,t_2}$;

  $\vdots$

- $\mathbf{v}_k := \alpha_{k,1}\mathbf{v}_{k,1} + \dots + \alpha_{k,t_k}\mathbf{v}_{k,t_k}$.

So, $\mathbf{v}_1 + \mathbf{v}_2 + \dots + \mathbf{v}_k = \mathbf{0}$. Now, note that for each $i \in \{1, \dots, k\}$, the vector $\mathbf{v}_i$ is a linear combination of vectors in $E_{\lambda_i}(f)$; since $E_{\lambda_i}(f)$ is a subsapce of V, it follows that $\mathbf{v}_i \in E_{\lambda_i}(f)$. Consequently, $\forall i \in \{1, \dots, k\}$: $\mathbf{v}_i$ is either $\mathbf{0}$ or an eigenvector of f associated with the eigenvalue λ_i. WTS $\mathbf{v}_1 = \mathbf{v}_2 = \dots = \mathbf{v}_k = \mathbf{0}$. Suppose otherwise.

Proof. Fix $\alpha_{1,1}, \dots, \alpha_{1,t_1}, \alpha_{2,1}, \dots, \alpha_{2,t_2}, \dots, \alpha_{k,1}, \dots, \alpha_{k,t_k} \in \mathbb{F}$ s.t.

$$\sum_{i=1}^{k} \Big(\alpha_{i,1}\mathbf{v}_{i,1} + \dots + \alpha_{i,t_i}\mathbf{v}_{i,t_i} \Big) \quad = \quad \mathbf{0}.$$

$\forall i \in \{1, \dots, k\}$: set $\mathbf{v}_i := \alpha_{i,1}\mathbf{v}_{i,1} + \dots + \alpha_{i,t_i}\mathbf{v}_{i,t_i}$, that is

- $\mathbf{v}_1 := \alpha_{1,1}\mathbf{v}_{1,1} + \dots + \alpha_{1,t_1}\mathbf{v}_{1,t_1}$;
- $\mathbf{v}_2 := \alpha_{2,1}\mathbf{v}_{2,1} + \dots + \alpha_{2,t_2}\mathbf{v}_{2,t_2}$;

  $\vdots$
- $\mathbf{v}_k := \alpha_{k,1}\mathbf{v}_{k,1} + \dots + \alpha_{k,t_k}\mathbf{v}_{k,t_k}$.

So, $\mathbf{v}_1 + \mathbf{v}_2 + \dots + \mathbf{v}_k = \mathbf{0}$. Now, note that for each $i \in \{1, \dots, k\}$, the vector $\mathbf{v}_i$ is a linear combination of vectors in $E_{\lambda_i}(f)$; since $E_{\lambda_i}(f)$ is a subsapce of V, it follows that $\mathbf{v}_i \in E_{\lambda_i}(f)$. Consequently, $\forall i \in \{1, \dots, k\}$: $\mathbf{v}_i$ is either $\mathbf{0}$ or an eigenvector of f associated with the eigenvalue λ_i. WTS $\mathbf{v}_1 = \mathbf{v}_2 = \dots = \mathbf{v}_k = \mathbf{0}$. Suppose otherwise. By symmetry, WMA $\exists \ell \in \{1, \dots, k\}$ s.t. $\mathbf{v}_1, \dots, \mathbf{v}_\ell$ are all non-zero (and are consequently eigenvectors of f associated with $\lambda_1, \dots, \lambda_\ell$), while $\mathbf{v}_{\ell+1}, \dots, \mathbf{v}_k$ are all zero.

Proof (continued). So,

$$\mathbf{v}_1 + \cdots + \mathbf{v}_\ell = \mathbf{0},$$

and it follows that $\{\mathbf{v}_1, \ldots, \mathbf{v}_\ell\}$ is a linearly dependent set. But this contradicts Proposition 8.4.1.

Proof (continued). So,

$$\mathbf{v}_1 + \cdots + \mathbf{v}_\ell \quad = \quad \mathbf{0},$$

and it follows that $\{\mathbf{v}_1, \ldots, \mathbf{v}_\ell\}$ is a linearly dependent set. But this contradicts Proposition 8.4.1.

We have now shown that $\mathbf{v}_1 = \cdots = \mathbf{v}_k = \mathbf{0}$.

Proof (continued). So,

$$\mathbf{v}_1 + \cdots + \mathbf{v}_\ell \quad = \quad \mathbf{0},$$

and it follows that $\{\mathbf{v}_1, \ldots, \mathbf{v}_\ell\}$ is a linearly dependent set. But this contradicts Proposition 8.4.1.

We have now shown that $\mathbf{v}_1 = \cdots = \mathbf{v}_k = \mathbf{0}$. So, for all indices $i \in \{1, \ldots, k\}$, we have that

$$\alpha_{i,1}\mathbf{v}_{i,1} + \cdots + \alpha_{i,t_i}\mathbf{v}_{i,t_i} \quad = \quad \mathbf{0};$$

Proof (continued). So,

$$\mathbf{v}_1 + \cdots + \mathbf{v}_\ell \quad = \quad \mathbf{0},$$

and it follows that $\{\mathbf{v}_1, \dots, \mathbf{v}_\ell\}$ is a linearly dependent set. But this contradicts Proposition 8.4.1.

We have now shown that $\mathbf{v}_1 = \cdots = \mathbf{v}_k = \mathbf{0}$. So, for all indices $i \in \{1, \dots, k\}$, we have that

$$\alpha_{i,1}\mathbf{v}_{i,1} + \cdots + \alpha_{i,t_i}\mathbf{v}_{i,t_i} \quad = \quad \mathbf{0};$$

since vectors $\mathbf{v}_{i,1}, \dots, \mathbf{v}_{i,t_i}$ are linearly independent, it follows that $\alpha_{i,1} = \cdots = \alpha_{i,t_i} = 0$.

Proof (continued). So,

$$\mathbf{v}_1 + \cdots + \mathbf{v}_\ell \;=\; \mathbf{0},$$

and it follows that $\{\mathbf{v}_1, \dots, \mathbf{v}_\ell\}$ is a linearly dependent set. But this contradicts Proposition 8.4.1.

We have now shown that $\mathbf{v}_1 = \cdots = \mathbf{v}_k = \mathbf{0}$. So, for all indices $i \in \{1, \dots, k\}$, we have that

$$\alpha_{i,1}\mathbf{v}_{i,1} + \cdots + \alpha_{i,t_i}\mathbf{v}_{i,t_i} \;=\; \mathbf{0};$$

since vectors $\mathbf{v}_{i,1}, \dots, \mathbf{v}_{i,t_i}$ are linearly independent, it follows that $\alpha_{i,1} = \cdots = \alpha_{i,t_i} = 0$.

Since this holds for all indices $i \in \{1, \dots, k\}$, we deduce that the eigenvectors

$$\mathbf{v}_{1,1}, \dots, \mathbf{v}_{1,t_1}, \mathbf{v}_{2,1}, \dots, \mathbf{v}_{2,t_2}, \dots, \mathbf{v}_{k,1}, \dots, \mathbf{v}_{k,t_k}$$

are linearly independent, which is what we needed to show. □

Proposition 8.4.1

Let V be a vector space over a field $\mathbb{F}$, let $f : V \to V$ be a linear function, let $\lambda_1, \dots, \lambda_k \in \mathbb{F}$ be pairwise distinct eigenvalues of f, associated with eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_k$, respectively. Then $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is a linearly independent set.

Proposition 8.4.2

Let V be a vector space over a field $\mathbb{F}$, let $f : V \to V$ be a linear function, and let $\lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{F}$ be pairwise distinct eigenvalues of f. For each $i \in \{1, \dots, k\}$, let $\mathbf{v}_{i,1}, \dots, \mathbf{v}_{i,t_i}$ be linearly independent eigenvectors of f associated with the eigenvalue λ_i. Then the eigenvectors

$$\mathbf{v}_{1,1}, \dots, \mathbf{v}_{1,t_1}, \mathbf{v}_{2,1}, \dots, \mathbf{v}_{2,t_2}, \dots, \mathbf{v}_{k,1}, \dots, \mathbf{v}_{k,t_k}$$

are linearly independent.

Theorem 8.4.3

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, and set $n := \dim(V)$. Let $f : V \to V$ be a linear function, let $\lambda_1, \dots, \lambda_k$ be all (distinct) the eigenvalues of f, and let $\mathcal{B}_1, \dots, \mathcal{B}_k$ be bases of the associated eigenspaces $E_{\lambda_1}(f), \dots, E_{\lambda_k}(f)$, respectively. Set $\mathcal{B} := \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$. Then all the following hold:

- (a) $\mathcal{B}$ is a linearly independent set of eigenvectors of f;
- (b) $\dim\big(E_{\lambda_1}(f)\big) + \dots + \dim\big(E_{\lambda_k}(f)\big) \leq n$, i.e. the sum of geometric multiplicities of the eigenvalues of f is at most n;
- (c) V has an eigenbasis associated with f iff the sum of geometric multiplicities of the eigenvalues of f is n, and in this case, $\mathcal{B}$ is such an eigenbasis;
- (d) V has an eigenbasis associated with f iff the sum of algebraic multiplicities of the eigenvalues of f is n, and the geometric multiplicity of each eigenvalue is equal to its algebraic multiplicity; in this case, $\mathcal{B}$ is an eigenbasis of V associated with the linear function f.

Theorem 8.4.3

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, and set $n := \dim(V)$. Let $f : V \to V$ be a linear function, let $\lambda_1, \dots, \lambda_k$ be all (distinct) the eigenvalues of f, and let $\mathcal{B}_1, \dots, \mathcal{B}_k$ be bases of the associated eigenspaces $E_{\lambda_1}(f), \dots, E_{\lambda_k}(f)$, respectively. Set $\mathcal{B} := \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$. Then all the following hold:

- (a) $\mathcal{B}$ is a linearly independent set of eigenvectors of f;
- (b) $\dim\big(E_{\lambda_1}(f)\big) + \dots + \dim\big(E_{\lambda_k}(f)\big) \leq n$, i.e. the sum of geometric multiplicities of the eigenvalues of f is at most n;

Proof. Part (a) follows immediately from Proposition 8.4.2.

Theorem 8.4.3

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, and set $n := \dim(V)$. Let $f : V \to V$ be a linear function, let $\lambda_1, \dots, \lambda_k$ be all (distinct) the eigenvalues of f, and let $\mathcal{B}_1, \dots, \mathcal{B}_k$ be bases of the associated eigenspaces $E_{\lambda_1}(f), \dots, E_{\lambda_k}(f)$, respectively. Set $\mathcal{B} := \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$. Then all the following hold:

- (a) $\mathcal{B}$ is a linearly independent set of eigenvectors of f;
- (b) $\dim\big(E_{\lambda_1}(f)\big) + \dots + \dim\big(E_{\lambda_k}(f)\big) \leq n$, i.e. the sum of geometric multiplicities of the eigenvalues of f is at most n;

Proof. Part (a) follows immediately from Proposition 8.4.2.

Part (b) follows from (a) and from the fact that, by Theorem 3.2.17(a), any linearly independent set of vectors in an n-dimensional vector space contains at most n vectors.

Theorem 8.4.3

(c) V has an eigenbasis associated with f iff the sum of geometric multiplicities of the eigenvalues of f is n, and in this case, $\mathcal{B}$ is such an eigenbasis;

Proof (continued). Let us prove (c).

Theorem 8.4.3

(c) V has an eigenbasis associated with f iff the sum of geometric multiplicities of the eigenvalues of f is n, and in this case, $\mathcal{B}$ is such an eigenbasis;

Proof (continued). Let us prove (c). Suppose first that the sum of geometric multiplicities of the eigenvalues of f is equal to n.

Theorem 8.4.3

(c) V has an eigenbasis associated with f iff the sum of geometric multiplicities of the eigenvalues of f is n, and in this case, $\mathcal{B}$ is such an eigenbasis;

Proof (continued). Let us prove (c). Suppose first that the sum of geometric multiplicities of the eigenvalues of f is equal to n. Then $\mathcal{B}$ is a linearly independent set of size n in the n-dimensional vector space V.

Theorem 8.4.3

(c) V has an eigenbasis associated with f iff the sum of geometric multiplicities of the eigenvalues of f is n, and in this case, $\mathcal{B}$ is such an eigenbasis;

Proof (continued). Let us prove (c). Suppose first that the sum of geometric multiplicities of the eigenvalues of f is equal to n. Then $\mathcal{B}$ is a linearly independent set of size n in the n-dimensional vector space V. So, by Corollary 3.2.20(a), $\mathcal{B}$ is a basis of V.

Theorem 8.4.3

(c) V has an eigenbasis associated with f iff the sum of geometric multiplicities of the eigenvalues of f is n, and in this case, $\mathcal{B}$ is such an eigenbasis;

Proof (continued). Let us prove (c). Suppose first that the sum of geometric multiplicities of the eigenvalues of f is equal to n. Then $\mathcal{B}$ is a linearly independent set of size n in the n-dimensional vector space V. So, by Corollary 3.2.20(a), $\mathcal{B}$ is a basis of V. Since all vectors in $\mathcal{B}$ are eigenvectors of f, it follows that $\mathcal{B}$ is an eigenbasis of V associated with f.

Theorem 8.4.3

- (c) V has an eigenbasis associated with f iff the sum of geometric multiplicities of the eigenvalues of f is n, and in this case, $\mathcal{B}$ is such an eigenbasis;

Proof (continued). Suppose, conversely, that V has an eigenbasis $\mathcal{C}$ associated with f;

Theorem 8.4.3

- (c) V has an eigenbasis associated with f iff the sum of geometric multiplicities of the eigenvalues of f is n, and in this case, $\mathcal{B}$ is such an eigenbasis;

Proof (continued). Suppose, conversely, that V has an eigenbasis $\mathcal{C}$ associated with f; since $\dim(V) = n$, we see that $|\mathcal{C}| = n$.

Theorem 8.4.3

- (c) V has an eigenbasis associated with f iff the sum of geometric multiplicities of the eigenvalues of f is n, and in this case, $\mathcal{B}$ is such an eigenbasis;

Proof (continued). Suppose, conversely, that V has an eigenbasis $\mathcal{C}$ associated with f; since $\dim(V) = n$, we see that $|\mathcal{C}| = n$. Since all vectors in $\mathcal{C}$ are eigenvecors of f, we see that they all belong to $E_{\lambda_1}(f) \cup \cdots \cup E_{\lambda_k}(f)$. But since the basis $\mathcal{C}$ of V is, in particular, linearly independent, we see that it cannot contain more than $\dim\Big(E_{\lambda_i}(f)\Big)$ many vectors from $E_{\lambda_i}(f)$ for any index $i \in \{1, \dots, k\}$.

Theorem 8.4.3

(c) V has an eigenbasis associated with f iff the sum of geometric multiplicities of the eigenvalues of f is n, and in this case, $\mathcal{B}$ is such an eigenbasis;

Proof (continued). Suppose, conversely, that V has an eigenbasis $\mathcal{C}$ associated with f; since $\dim(V) = n$, we see that $|\mathcal{C}| = n$. Since all vectors in $\mathcal{C}$ are eigenvecors of f, we see that they all belong to $E_{\lambda_1}(f) \cup \dots \cup E_{\lambda_k}(f)$. But since the basis $\mathcal{C}$ of V is, in particular, linearly independent, we see that it cannot contain more than $\dim\Big(E_{\lambda_i}(f)\Big)$ many vectors from $E_{\lambda_i}(f)$ for any index $i \in \{1, \dots, k\}$. So, $|\mathcal{C}| \leq \dim(E_{\lambda_1}(f)) + \dots + \dim(E_{\lambda_k}(f))$.

Theorem 8.4.3

(c) V has an eigenbasis associated with f iff the sum of geometric multiplicities of the eigenvalues of f is n, and in this case, $\mathcal{B}$ is such an eigenbasis;

Proof (continued). Suppose, conversely, that V has an eigenbasis $\mathcal{C}$ associated with f; since $\dim(V) = n$, we see that $|\mathcal{C}| = n$. Since all vectors in $\mathcal{C}$ are eigenvecors of f, we see that they all belong to $E_{\lambda_1}(f) \cup \cdots \cup E_{\lambda_k}(f)$. But since the basis $\mathcal{C}$ of V is, in particular, linearly independent, we see that it cannot contain more than $\dim\Big(E_{\lambda_i}(f)\Big)$ many vectors from $E_{\lambda_i}(f)$ for any index $i \in \{1, \dots, k\}$. So, $|\mathcal{C}| \leq \dim(E_{\lambda_1}(f)) + \cdots + \dim(E_{\lambda_k}(f))$. But now we have that

$$n \;=\; |\mathcal{C}| \;\leq\; \dim(E_{\lambda_1}(f)) + \cdots + \dim(E_{\lambda_k}(f)) \;\overset{\text{(b)}}{\leq}\; n,$$

Theorem 8.4.3

(c) V has an eigenbasis associated with f iff the sum of geometric multiplicities of the eigenvalues of f is n, and in this case, $\mathcal{B}$ is such an eigenbasis;

Proof (continued). Suppose, conversely, that V has an eigenbasis $\mathcal{C}$ associated with f; since $\dim(V) = n$, we see that $|\mathcal{C}| = n$. Since all vectors in $\mathcal{C}$ are eigenvecors of f, we see that they all belong to $E_{\lambda_1}(f) \cup \cdots \cup E_{\lambda_k}(f)$. But since the basis $\mathcal{C}$ of V is, in particular, linearly independent, we see that it cannot contain more than $\dim\Big(E_{\lambda_i}(f)\Big)$ many vectors from $E_{\lambda_i}(f)$ for any index $i \in \{1, \dots, k\}$. So, $|\mathcal{C}| \leq \dim(E_{\lambda_1}(f)) + \cdots + \dim(E_{\lambda_k}(f))$. But now we have that

$$n = |\mathcal{C}| \leq \dim(E_{\lambda_1}(f)) + \cdots + \dim(E_{\lambda_k}(f)) \overset{\text{(b)}}{\leq} n,$$

and it follows that $\dim(E_{\lambda_1}(f)) + \cdots + \dim(E_{\lambda_k}(f)) = n$,

Theorem 8.4.3

- (c) V has an eigenbasis associated with f iff the sum of geometric multiplicities of the eigenvalues of f is n, and in this case, $\mathcal{B}$ is such an eigenbasis;

Proof (continued). Suppose, conversely, that V has an eigenbasis $\mathcal{C}$ associated with f; since $\dim(V) = n$, we see that $|\mathcal{C}| = n$. Since all vectors in $\mathcal{C}$ are eigenvecors of f, we see that they all belong to $E_{\lambda_1}(f) \cup \cdots \cup E_{\lambda_k}(f)$. But since the basis $\mathcal{C}$ of V is, in particular, linearly independent, we see that it cannot contain more than $\dim\Big(E_{\lambda_i}(f)\Big)$ many vectors from $E_{\lambda_i}(f)$ for any index $i \in \{1, \dots, k\}$. So, $|\mathcal{C}| \leq \dim(E_{\lambda_1}(f)) + \cdots + \dim(E_{\lambda_k}(f))$. But now we have that

$$n \;=\; |\mathcal{C}| \;\leq\; \dim(E_{\lambda_1}(f)) + \cdots + \dim(E_{\lambda_k}(f)) \;\overset{\text{(b)}}{\leq}\; n,$$

and it follows that $\dim(E_{\lambda_1}(f)) + \cdots + \dim(E_{\lambda_k}(f)) = n$, i.e. the sum of geometric multiplicities of the eigenvalues of f is n. This proves (c).

Theorem 8.4.3

(d) V has an eigenbasis associated with f iff the sum of algebraic multiplicities of the eigenvalues of f is n, and the geometric multiplicity of each eigenvalue is equal to its algebraic multiplicity; in this case, $\mathcal{B}$ is an eigenbasis of V associated with the linear function f.

Proof (continued). It remains to prove (d). If the sum of algebraic multiplicities of the eigenvalues of f is equal to n, and the geometric multiplicity of each eigenvalue is equal to its algebraic multiplicity, then obviously, the sum of geometric multiplicities of f is equal to n,

Theorem 8.4.3

(d) V has an eigenbasis associated with f iff the sum of algebraic multiplicities of the eigenvalues of f is n, and the geometric multiplicity of each eigenvalue is equal to its algebraic multiplicity; in this case, $\mathcal{B}$ is an eigenbasis of V associated with the linear function f.

Proof (continued). It remains to prove (d). If the sum of algebraic multiplicities of the eigenvalues of f is equal to n, and the geometric multiplicity of each eigenvalue is equal to its algebraic multiplicity, then obviously, the sum of geometric multiplicities of f is equal to n, and so by (c), V has an eigenbasis associated with f, and $\mathcal{B}$ is one such eigenbasis.

Theorem 8.4.3

(d) V has an eigenbasis associated with f iff the sum of algebraic multiplicities of the eigenvalues of f is n, and the geometric multiplicity of each eigenvalue is equal to its algebraic multiplicity; in this case, $\mathcal{B}$ is an eigenbasis of V associated with the linear function f.

Proof (continued). It remains to prove (d). If the sum of algebraic multiplicities of the eigenvalues of f is equal to n, and the geometric multiplicity of each eigenvalue is equal to its algebraic multiplicity, then obviously, the sum of geometric multiplicities of f is equal to n, and so by (c), V has an eigenbasis associated with f, and $\mathcal{B}$ is one such eigenbasis.

For the converse, assume that V has an eigenbasis $\mathcal{C}$ associated with f.

Theorem 8.4.3

(d) V has an eigenbasis associated with f iff the sum of algebraic multiplicities of the eigenvalues of f is n, and the geometric multiplicity of each eigenvalue is equal to its algebraic multiplicity; in this case, $\mathcal{B}$ is an eigenbasis of V associated with the linear function f.

Proof (continued). It remains to prove (d). If the sum of algebraic multiplicities of the eigenvalues of f is equal to n, and the geometric multiplicity of each eigenvalue is equal to its algebraic multiplicity, then obviously, the sum of geometric multiplicities of f is equal to n, and so by (c), V has an eigenbasis associated with f, and $\mathcal{B}$ is one such eigenbasis.

For the converse, assume that V has an eigenbasis $\mathcal{C}$ associated with f. Let $\lambda_1, \dots, \lambda_k$ be the eigenvalues of f, with geometric multiplicities $g_1, \dots, g_k$, respectively, and algebraic multiplicities $a_1, \dots, a_k$, respectively.

Theorem 8.4.3

(d) V has an eigenbasis associated with f iff the sum of algebraic multiplicities of the eigenvalues of f is n, and the geometric multiplicity of each eigenvalue is equal to its algebraic multiplicity; in this case, $\mathcal{B}$ is an eigenbasis of V associated with the linear function f.

Proof (continued). By (c), we have that $g_1 + \cdots + g_k = n$.

Theorem 8.4.3

(d) V has an eigenbasis associated with f iff the sum of algebraic multiplicities of the eigenvalues of f is n, and the geometric multiplicity of each eigenvalue is equal to its algebraic multiplicity; in this case, $\mathcal{B}$ is an eigenbasis of V associated with the linear function f.

Proof (continued). By (c), we have that $g_1 + \cdots + g_k = n$. On the other hand, the characteristic polynomial of f is of degree n, we see that the sum of algebraic multiplicitis of f is at most n, i.e. $a_1 + \cdots + a_k \leq n$.

Theorem 8.4.3

(d) V has an eigenbasis associated with f iff the sum of algebraic multiplicities of the eigenvalues of f is n, and the geometric multiplicity of each eigenvalue is equal to its algebraic multiplicity; in this case, $\mathcal{B}$ is an eigenbasis of V associated with the linear function f.

Proof (continued). By (c), we have that $g_1 + \cdots + g_k = n$. On the other hand, the characteristic polynomial of f is of degree n, we see that the sum of algebraic multiplicitis of f is at most n, i.e. $a_1 + \cdots + a_k \leq n$. But by Theorem 8.2.17, the geometric multiplicity of an eigenvalue of f is no greater than the algebraic multiplicity of that eigenvalue, that is, $g_i \leq a_i$ for all indices $i \in \{1, \dots, n\}$.

Theorem 8.4.3

(d) V has an eigenbasis associated with f iff the sum of algebraic multiplicities of the eigenvalues of f is n, and the geometric multiplicity of each eigenvalue is equal to its algebraic multiplicity; in this case, $\mathcal{B}$ is an eigenbasis of V associated with the linear function f.

Proof (continued). By (c), we have that $g_1 + \cdots + g_k = n$. On the other hand, the characteristic polynomial of f is of degree n, we see that the sum of algebraic multiplicitis of f is at most n, i.e. $a_1 + \cdots + a_k \leq n$. But by Theorem 8.2.17, the geometric multiplicity of an eigenvalue of f is no greater than the algebraic multiplicity of that eigenvalue, that is, $g_i \leq a_i$ for all indices $i \in \{1, \dots, n\}$. We now have that

$$n = g_1 + \cdots + g_k \leq a_1 + \cdots + a_k \leq n,$$

Theorem 8.4.3

(d) V has an eigenbasis associated with f iff the sum of algebraic multiplicities of the eigenvalues of f is n, and the geometric multiplicity of each eigenvalue is equal to its algebraic multiplicity; in this case, $\mathcal{B}$ is an eigenbasis of V associated with the linear function f.

Proof (continued). By (c), we have that $g_1 + \cdots + g_k = n$. On the other hand, the characteristic polynomial of f is of degree n, we see that the sum of algebraic multiplicitis of f is at most n, i.e. $a_1 + \cdots + a_k \leq n$. But by Theorem 8.2.17, the geometric multiplicity of an eigenvalue of f is no greater than the algebraic multiplicity of that eigenvalue, that is, $g_i \leq a_i$ for all indices $i \in \{1, \dots, n\}$. We now have that

$$n \;=\; g_1 + \cdots + g_k \;\leq\; a_1 + \cdots + a_k \;\leq\; n,$$

and we deduce that $a_1 + \cdots + a_k = n$ and that $g_i = a_i$ for all $i \in \{1, \dots, k\}$. This proves (d). $\square$

Theorem 8.4.3

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, and set $n := \dim(V)$. Let $f : V \to V$ be a linear function, let $\lambda_1, \dots, \lambda_k$ be all (distinct) the eigenvalues of f, and let $\mathcal{B}_1, \dots, \mathcal{B}_k$ be bases of the associated eigenspaces $E_{\lambda_1}(f), \dots, E_{\lambda_k}(f)$, respectively. Set $\mathcal{B} := \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$. Then all the following hold:

- (a) $\mathcal{B}$ is a linearly independent set of eigenvectors of f;
- (b) $\dim\big(E_{\lambda_1}(f)\big) + \dots + \dim\big(E_{\lambda_k}(f)\big) \leq n$, i.e. the sum of geometric multiplicities of the eigenvalues of f is at most n;
- (c) V has an eigenbasis associated with f iff the sum of geometric multiplicities of the eigenvalues of f is n, and in this case, $\mathcal{B}$ is such an eigenbasis;
- (d) V has an eigenbasis associated with f iff the sum of algebraic multiplicities of the eigenvalues of f is n, and the geometric multiplicity of each eigenvalue is equal to its algebraic multiplicity; in this case, $\mathcal{B}$ is an eigenbasis of V associated with the linear function f.

Corollary 8.4.4

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, and set $n := \dim(V)$. If a linear function $f : V \to V$ has n distinct eigenvalues, then V has an eigenbasis associated with f.

Proof.

Corollary 8.4.4

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, and set $n := \dim(V)$. If a linear function $f : V \to V$ has n distinct eigenvalues, then V has an eigenbasis associated with f.

Proof. Let $f : V \to V$ be a linear function that has n distinct eigenvalues, say $\lambda_1, \dots, \lambda_n$.

Corollary 8.4.4

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, and set $n := \dim(V)$. If a linear function $f : V \to V$ has n distinct eigenvalues, then V has an eigenbasis associated with f.

Proof. Let $f : V \to V$ be a linear function that has n distinct eigenvalues, say $\lambda_1, \ldots, \lambda_n$.

By the definition of an eigenvalue, we have that $\dim\big(E_{\lambda_i}(f)\big) \geq 1$ for all $i \in \{1, \ldots, n\}$.

Corollary 8.4.4

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, and set $n := \dim(V)$. If a linear function $f : V \to V$ has n distinct eigenvalues, then V has an eigenbasis associated with f.

Proof. Let $f : V \to V$ be a linear function that has n distinct eigenvalues, say $\lambda_1, \ldots, \lambda_n$.

By the definition of an eigenvalue, we have that $\dim\big(E_{\lambda_i}(f)\big) \geq 1$ for all $i \in \{1, \ldots, n\}$. Consequently, $\dim\big(E_{\lambda_1}(f)\big) + \cdots + \dim\big(E_{\lambda_n}(f)\big) \geq n$.

Corollary 8.4.4

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, and set $n := \dim(V)$. If a linear function $f : V \to V$ has n distinct eigenvalues, then V has an eigenbasis associated with f.

Proof. Let $f : V \to V$ be a linear function that has n distinct eigenvalues, say $\lambda_1, \dots, \lambda_n$.

By the definition of an eigenvalue, we have that $\dim\big(E_{\lambda_i}(f)\big) \geq 1$ for all $i \in \{1, \dots, n\}$. Consequently, $\dim\big(E_{\lambda_1}(f)\big) + \cdots + \dim\big(E_{\lambda_n}(f)\big) \geq n$.

On the other hand, Theorem 8.4.3(b) guarantees that $\dim\big(E_{\lambda_1}(f)\big) + \cdots + \dim\big(E_{\lambda_n}(f)\big) \leq n$.

Corollary 8.4.4

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, and set $n := \dim(V)$. If a linear function $f : V \to V$ has n distinct eigenvalues, then V has an eigenbasis associated with f.

Proof. Let $f : V \to V$ be a linear function that has n distinct eigenvalues, say $\lambda_1, \dots, \lambda_n$.

By the definition of an eigenvalue, we have that $\dim(E_{\lambda_i}(f)) \geq 1$ for all $i \in \{1, \dots, n\}$. Consequently,
$\dim(E_{\lambda_1}(f)) + \cdots + \dim(E_{\lambda_n}(f)) \geq n$.

On the other hand, Theorem 8.4.3(b) guarantees that
$\dim(E_{\lambda_1}(f)) + \cdots + \dim(E_{\lambda_n}(f)) \leq n$.

Thus, $\dim(E_{\lambda_1}(f)) + \cdots + \dim(E_{\lambda_n}(f)) = n$,

Corollary 8.4.4

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, and set $n := \dim(V)$. If a linear function $f : V \to V$ has n distinct eigenvalues, then V has an eigenbasis associated with f.

Proof. Let $f : V \to V$ be a linear function that has n distinct eigenvalues, say $\lambda_1, \dots, \lambda_n$.

By the definition of an eigenvalue, we have that $\dim(E_{\lambda_i}(f)) \geq 1$ for all $i \in \{1, \dots, n\}$. Consequently, $\dim(E_{\lambda_1}(f)) + \dots + \dim(E_{\lambda_n}(f)) \geq n$.

On the other hand, Theorem 8.4.3(b) guarantees that $\dim(E_{\lambda_1}(f)) + \dots + \dim(E_{\lambda_n}(f)) \leq n$.

Thus, $\dim(E_{\lambda_1}(f)) + \dots + \dim(E_{\lambda_n}(f)) = n$, and so by Theorem 8.4.3(c), V has an eigenbasis associated with f. □

- We would now like to "translate" Theorem 8.4.3 and Corollary 8.4.4 into the language of matrices.

- We would now like to “translate” Theorem 8.4.3 and Corollary 8.4.4 into the language of matrices.
- Given a field $\mathbb{F}$ and a square matrix $A \in \mathbb{F}^{n \times n}$, we can define $f_A : \mathbb{F}^n \to \mathbb{F}^n$ by setting $f_A(\mathbf{v}) = A\mathbf{v}$ for all $\mathbf{v} \in \mathbb{F}^n$.
    - So, f_A is linear, and its standard matrix is A.

- We would now like to "translate" Theorem 8.4.3 and Corollary 8.4.4 into the language of matrices.
- Given a field $\mathbb{F}$ and a square matrix $A \in \mathbb{F}^{n \times n}$, we can define $f_A : \mathbb{F}^n \to \mathbb{F}^n$ by setting $f_A(\mathbf{v}) = A\mathbf{v}$ for all $\mathbf{v} \in \mathbb{F}^n$.
    - So, f_A is linear, and its standard matrix is A.
- We can apply Theorem 8.4.3 and Corollary 8.4.4 to the linear function f_A, and then get the same result for A "for free."
    - Next two slides!

Theorem 8.4.5

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n\times n}$. Let $\lambda_1, \dots, \lambda_k$ be all the (distinct) eigenvalues of A, and let $\mathcal{B}_1, \dots, \mathcal{B}_k$ be bases of the associated eigenspaces $E_{\lambda_1}(A), \dots, E_{\lambda_k}(A)$, respectively. Set $\mathcal{B} := \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$. Then all the following hold:

a. $\mathcal{B}$ is a linearly independent set of eigenvectors of A;
b. $\dim\big(E_{\lambda_1}(A)\big) + \dots + \dim\big(E_{\lambda_k}(A)\big) \leq n$, i.e. the sum of geometric multiplicities of the eigenvalues of A is at most n;
c. $\mathbb{F}^n$ has an eigenbasis associated with A iff the sum of geometric multiplicities of the eigenvalues of A is n, and in this case, $\mathcal{B}$ is such an eigenbasis;
d. $\mathbb{F}^n$ has an eigenbasis associated with A iff the sum of algebraic multiplicities of the eigenvalues of A is n, and the geometric multiplicity of each eigenvalue is equal to its algebraic multiplicity; in this case, $\mathcal{B}$ is an eigenbasis of $\mathbb{F}^n$ associated with the matrix A.

Corollary 8.4.6

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$. If A has n distinct eigenvalues, then $\mathbb{F}^n$ has an eigenbasis associated with A.

3. Diagonalization

3. Diagonalization

Definition

For a field $\mathbb{F}$, a square matrix $D \in \mathbb{F}^{n\times n}$ is *diagonal* if all its entries off the main diagonal are zero (the entries on the main diagonal may or may not be zero). For scalars $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{F}$, $D(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is the $n \times n$ matrix with $\lambda_1, \lambda_2, \ldots, \lambda_n$ on the main diagonal (appearing in that order) and 0's everywhere else, i.e.

$$\begin{aligned} D(\lambda_1, \lambda_2, \ldots, \lambda_n) &:= \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1\mathbf{e}_1 & \ldots & \lambda_n\mathbf{e}_n \end{bmatrix}, \end{aligned}$$

where as usual, $\mathbf{e}_1, \ldots, \mathbf{e}_n$ are the standard basis vectors of $\mathbb{F}^n$.

3. Diagonalization

Definition

For a field $\mathbb{F}$, a square matrix $D \in \mathbb{F}^{n\times n}$ is *diagonal* if all its entries off the main diagonal are zero (the entries on the main diagonal may or may not be zero). For scalars $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{F}$, $D(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is the $n \times n$ matrix with $\lambda_1, \lambda_2, \ldots, \lambda_n$ on the main diagonal (appearing in that order) and 0's everywhere else, i.e.

$$\begin{aligned} D(\lambda_1, \lambda_2, \ldots, \lambda_n) &:= \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 \mathbf{e}_1 & \ldots & \lambda_n \mathbf{e}_n \end{bmatrix}, \end{aligned}$$

where as usual, $\mathbf{e}_1, \ldots, \mathbf{e}_n$ are the standard basis vectors of $\mathbb{F}^n$.

- Note that diagonal matrices are, in particular, triangular.

3. Diagonalization

Definition

For a field $\mathbb{F}$, a square matrix $D \in \mathbb{F}^{n\times n}$ is *diagonal* if all its entries off the main diagonal are zero (the entries on the main diagonal may or may not be zero). For scalars $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{F}$, $D(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is the $n \times n$ matrix with $\lambda_1, \lambda_2, \ldots, \lambda_n$ on the main diagonal (appearing in that order) and 0's everywhere else, i.e.

$$\begin{aligned} D(\lambda_1, \lambda_2, \ldots, \lambda_n) &:= \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 \mathbf{e}_1 & \ldots & \lambda_n \mathbf{e}_n \end{bmatrix}, \end{aligned}$$

where as usual, $\mathbf{e}_1, \ldots, \mathbf{e}_n$ are the standard basis vectors of $\mathbb{F}^n$.

- Note that diagonal matrices are, in particular, triangular.
- So, Propositions 7.3.1 and 8.2.7 (next slide) apply.

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.

Proposition 7.3.1

Let $\mathbb{F}$ be a field, and let $A = \left[\begin{array}{c} a_{i,j} \end{array} \right]_{n \times n}$ 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.

Proposition 8.2.7

Let $\mathbb{F}$ be a field, and let $A = \left[\begin{array}{c} a_{i,j} \end{array} \right]_{n \times n}$ 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. 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.

- Thus, for scalars $\lambda_1, \dots, \lambda_n \in \mathbb{F}$ (where $\mathbb{F}$ is a field), and for the diagonal matrix $D := D(\lambda_1, \dots, \lambda_n)$, we have the following:
  - $\det(D) = \lambda_1 \dots \lambda_n$;
  - $p_D(\lambda) = (\lambda - \lambda_1) \dots (\lambda - \lambda_n)$.

- Thus, for scalars $\lambda_1, \dots, \lambda_n \in \mathbb{F}$ (where $\mathbb{F}$ is a field), and for the diagonal matrix $D := D(\lambda_1, \dots, \lambda_n)$, we have the following:
  - $\det(D) = \lambda_1 \dots \lambda_n$;
  - $p_D(\lambda) = (\lambda - \lambda_1) \dots (\lambda - \lambda_n)$.

- We now state three simple propositions about diagonal matrices.
  - The proofs are easy and we omit them here.
  - However, the proofs can be found in the Lecture Notes.

Proposition 8.5.1

Let $\mathbb{F}$ be a field, let $\lambda_1, \dots, \lambda_n \in \mathbb{F}$ ($n \geq 1$) be arbitrary scalars, and set $D := D(\lambda_1, \dots, \lambda_n)$. Then both the following hold:

- (a) for all vectors $\mathbf{x} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T$ in $\mathbb{F}^n$, we have that

$$D\mathbf{x} \;=\; \begin{bmatrix} \lambda_1 x_1 \\ \vdots \\ \lambda_n x_n \end{bmatrix};$$

- (b) for all matrices $A = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_n \end{bmatrix}$ in $\mathbb{F}^{m \times n}$, we have that

$$AD \;=\; \begin{bmatrix} \lambda_1 \mathbf{a}_1 & \dots & \lambda_n \mathbf{a}_n \end{bmatrix}.$$

- Proof: Lecture Notes (easy!).

Proposition 8.5.2

Let $\mathbb{F}$ be a field, and let $\lambda_1, \ldots, \lambda_n, \mu_1, \ldots, \mu_n \in \mathbb{F}$ ($n \geq 1$) be arbitrary scalars. Then

$$D(\lambda_1, \ldots, \lambda_n)\, D(\mu_1, \ldots, \mu_n) \quad = \quad D(\lambda_1\mu_1, \ldots, \lambda_n\mu_n).$$

- Proof: Lecture Notes (easy!)

Proposition 8.5.2

Let $\mathbb{F}$ be a field, and let $\lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_n \in \mathbb{F}$ $(n \geq 1)$ be arbitrary scalars. Then

$$D(\lambda_1, \dots, \lambda_n)\, D(\mu_1, \dots, \mu_n) \;=\; D(\lambda_1\mu_1, \dots, \lambda_n\mu_n).$$

- Proof: Lecture Notes (easy!)

Proposition 8.5.3

Let $\mathbb{F}$ be a field, let $\lambda_1, \dots, \lambda_n \in \mathbb{F}$ $(n \geq 1)$, and set $D := D(\lambda_1, \dots, \lambda_n)$. Then both the following hold:

(a) for all non-negative integers m, we have that $D^m = D(\lambda_1^m, \dots, \lambda_n^m)$;

(b) D is invertible iff $\lambda_1, \dots, \lambda_n$ are all non-zero, and in this case, we have that $D^m = D(\lambda_1^m, \dots, \lambda_n^m)$ for all integers m.

- Proof: Lecture Notes (easy!)

Theorem 8.5.4

Let V be a non-trivial, finite-dimensional vector space, let $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be a basis of V, and let $f : V \to V$ be a linear function. Then $\mathcal{B}$ is an eigenbasis of V associated with f iff the matrix ${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$ is diagonal. Moreover, in this case, we have that

$$ {}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}} \ = \ D(\lambda_1, \dots, \lambda_n), $$

where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of f associated with the eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$, respectively.

Proof.

Theorem 8.5.4

Let V be a non-trivial, finite-dimensional vector space, let $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be a basis of V, and let $f : V \to V$ be a linear function. Then $\mathcal{B}$ is an eigenbasis of V associated with f iff the matrix ${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$ is diagonal. Moreover, in this case, we have that

$$ {}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}} \;=\; D(\lambda_1, \dots, \lambda_n), $$

where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of f associated with the eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$, respectively.

Proof. Suppose first that $\mathcal{B}$ is an eigenbasis of V associated with f.

Theorem 8.5.4

Let V be a non-trivial, finite-dimensional vector space, let $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be a basis of V, and let $f : V \to V$ be a linear function. Then $\mathcal{B}$ is an eigenbasis of V associated with f iff the matrix ${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$ is diagonal. Moreover, in this case, we have that

$$ {}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}} = D(\lambda_1, \dots, \lambda_n), $$

where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of f associated with the eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$, respectively.

Proof. Suppose first that $\mathcal{B}$ is an eigenbasis of V associated with f. Then, by definition, vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ are eigenvectors of f, and we let $\lambda_1, \dots, \lambda_n$, respectively, be the associated eigenvalues.

Theorem 8.5.4

Let V be a non-trivial, finite-dimensional vector space, let $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be a basis of V, and let $f : V \to V$ be a linear function. Then $\mathcal{B}$ is an eigenbasis of V associated with f iff the matrix ${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$ is diagonal. Moreover, in this case, we have that

$${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}} \;=\; D(\lambda_1, \dots, \lambda_n),$$

where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of f associated with the eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$, respectively.

Proof. Suppose first that $\mathcal{B}$ is an eigenbasis of V associated with f. Then, by definition, vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ are eigenvectors of f, and we let $\lambda_1, \dots, \lambda_n$, respectively, be the associated eigenvalues. Then $f(\mathbf{v}_i) = \lambda_i \mathbf{v}_i$ for all indices $i \in \{1, \dots, n\}$, and we have the following (next slide):

Theorem 8.5.4

Let V be a non-trivial, finite-dimensional vector space, let $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be a basis of V, and let $f : V \to V$ be a linear function. Then $\mathcal{B}$ is an eigenbasis of V associated with f iff the matrix ${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$ is diagonal. Moreover, in this case, we have that

$$ {}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}} \ = \ D(\lambda_1, \dots, \lambda_n), $$

where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of f associated with the eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$, respectively.

Proof (continued).

$$\begin{aligned} {}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}} &= \left[\ [\ f(\mathbf{v}_1)\ ]_{\mathcal{B}} \ \dots \ [\ f(\mathbf{v}_1)\ ]_{\mathcal{B}}\ \right] \qquad \text{by Theorem 4.5.1} \\ &= \left[\ [\ \lambda_1\mathbf{v}_1\ ]_{\mathcal{B}} \ \dots \ [\ \lambda_n\mathbf{v}_n\ ]_{\mathcal{B}}\ \right] \\ &= \left[\ \lambda_1\mathbf{e}_1 \ \dots \ \lambda_n\mathbf{e}_n\ \right] \\ &= D(\lambda_1, \dots, \lambda_n). \end{aligned}$$

Theorem 8.5.4

Let V be a non-trivial, finite-dimensional vector space, let $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be a basis of V, and let $f : V \to V$ be a linear function. Then $\mathcal{B}$ is an eigenbasis of V associated with f iff the matrix ${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$ is diagonal. Moreover, in this case, we have that

$$ {}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}} \quad = \quad D(\lambda_1, \dots, \lambda_n), $$

where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of f associated with the eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$, respectively.

Proof (continued). Conversely, suppose that the matrix ${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$ is diagonal, and let $\lambda_1, \dots, \lambda_n$ be the entries of this matrix on the main diagonal, so that

$$ {}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}} \quad = \quad D(\lambda_1, \dots, \lambda_n) \quad = \quad \begin{bmatrix} \lambda_1\mathbf{e}_1 & \dots & \lambda_n\mathbf{e}_n \end{bmatrix}. $$

Theorem 8.5.4

Let V be a non-trivial, finite-dimensional vector space, let $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be a basis of V, and let $f : V \to V$ be a linear function. Then $\mathcal{B}$ is an eigenbasis of V associated with f iff the matrix ${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$ is diagonal. Moreover, in this case, we have that

$${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}} \quad = \quad D(\lambda_1, \dots, \lambda_n),$$

where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of f associated with the eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$, respectively.

Proof (continued). Conversely, suppose that the matrix ${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$ is diagonal, and let $\lambda_1, \dots, \lambda_n$ be the entries of this matrix on the main diagonal, so that

$${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}} \quad = \quad D(\lambda_1, \dots, \lambda_n) \quad = \quad \left[\begin{array}{ccc} \lambda_1\mathbf{e}_1 & \dots & \lambda_n\mathbf{e}_n \end{array}\right].$$

We will show that the basis vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ are eigenvectors of f with associated eigenvalues $\lambda_1, \dots, \lambda_n$, respectively.

Theorem 8.5.4

Let V be a non-trivial, finite-dimensional vector space, let $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be a basis of V, and let $f : V \to V$ be a linear function. Then $\mathcal{B}$ is an eigenbasis of V associated with f iff the matrix ${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$ is diagonal. Moreover, in this case, we have that

$$ {}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}} = D(\lambda_1, \dots, \lambda_n), $$

where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of f associated with the eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$, respectively.

Proof (continued). Conversely, suppose that the matrix ${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$ is diagonal, and let $\lambda_1, \dots, \lambda_n$ be the entries of this matrix on the main diagonal, so that

$$ {}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}} = D(\lambda_1, \dots, \lambda_n) = \left[\ \lambda_1\mathbf{e}_1 \quad \dots \quad \lambda_n\mathbf{e}_n\ \right]. $$

We will show that the basis vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ are eigenvectors of f with associated eigenvalues $\lambda_1, \dots, \lambda_n$, respectively. Fix any index $i \in \{1, \dots, n\}$; WTS $f(\mathbf{v}_i) = \lambda_i\mathbf{v}_i$.

Theorem 8.5.4

Let V be a non-trivial, finite-dimensional vector space, let $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be a basis of V, and let $f : V \to V$ be a linear function. Then $\mathcal{B}$ is an eigenbasis of V associated with f iff the matrix ${}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}}$ is diagonal. Moreover, in this case, we have that

$$ {}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}} \;=\; D(\lambda_1, \dots, \lambda_n), $$

where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of f associated with the eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$, respectively.

Proof (continued). Conversely, suppose that the matrix ${}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}}$ is diagonal, and let $\lambda_1, \dots, \lambda_n$ be the entries of this matrix on the main diagonal, so that

$$ {}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}} \;=\; D(\lambda_1, \dots, \lambda_n) \;=\; \big[\ \lambda_1\mathbf{e}_1 \quad \dots \quad \lambda_n\mathbf{e}_n\ \big]. $$

We will show that the basis vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ are eigenvectors of f with associated eigenvalues $\lambda_1, \dots, \lambda_n$, respectively. Fix any index $i \in \{1, \dots, n\}$; WTS $f(\mathbf{v}_i) = \lambda_i\mathbf{v}_i$. Since $\mathbf{v}_i$ is the i-th basis vector of $\mathcal{B}$, we have that $\big[\ \mathbf{v}_i\ \big]_{\mathcal{B}} = \mathbf{e}_i$.

Theorem 8.5.4

Let V be a non-trivial, finite-dimensional vector space, let $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be a basis of V, and let $f : V \to V$ be a linear function. Then $\mathcal{B}$ is an eigenbasis of V associated with f iff the matrix ${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$ is diagonal. Moreover, in this case, we have that

$${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}} \quad = \quad D(\lambda_1, \dots, \lambda_n),$$

where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of f associated with the eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$, respectively.

Proof (continued). Conversely, suppose that the matrix ${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}}$ is diagonal, and let $\lambda_1, \dots, \lambda_n$ be the entries of this matrix on the main diagonal, so that

$${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}} \quad = \quad D(\lambda_1, \dots, \lambda_n) \quad = \quad \begin{bmatrix} \lambda_1\mathbf{e}_1 & \dots & \lambda_n\mathbf{e}_n \end{bmatrix}.$$

We will show that the basis vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ are eigenvectors of f with associated eigenvalues $\lambda_1, \dots, \lambda_n$, respectively. Fix any index $i \in \{1, \dots, n\}$; WTS $f(\mathbf{v}_i) = \lambda_i\mathbf{v}_i$. Since $\mathbf{v}_i$ is the i-th basis vector of $\mathcal{B}$, we have that $[\ \mathbf{v}_i\ ]_{\mathcal{B}} = \mathbf{e}_i$. We now compute (next slide):

Theorem 8.5.4

Let V be a non-trivial, finite-dimensional vector space, let $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be a basis of V, and let $f : V \to V$ be a linear function. Then $\mathcal{B}$ is an eigenbasis of V associated with f iff the matrix ${}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}}$ is diagonal. Moreover, in this case, we have that

$$ {}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}} \ = \ D(\lambda_1, \dots, \lambda_n), $$

where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of f associated with the eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$, respectively.

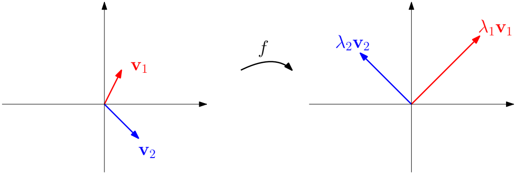
Proof (continued).

$$ \begin{aligned} \big[\ f(\mathbf{v}_i)\ \big]_{\mathcal{B}} \ &= \ {}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}}\big[\ \mathbf{v}_i\ \big]_{\mathcal{B}} \ = \ \big[\ \lambda_1\mathbf{e}_1 \ \ \dots \ \ \lambda_n\mathbf{e}_n\ \big]\,\mathbf{e}_i \\ &\overset{(*)}{=} \ \lambda_i\mathbf{e}_i \ = \ \lambda_i\big[\ \mathbf{v}_i\ \big]_{\mathcal{B}} \ \overset{(**)}{=} \ \big[\ \lambda_i\mathbf{v}_i\ \big]_{\mathcal{B}}, \end{aligned} $$

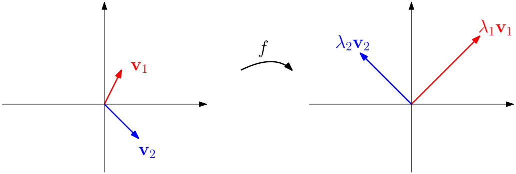
where (*) follows from Proposition 1.4.4, and (**) follows from the linearity of $\big[\ \cdot\ \big]_{\mathcal{B}}$. Since $\big[\ \cdot\ \big]_{\mathcal{B}}$ is an isomorphism (and in particular, one-to-one), it follows that $f(\mathbf{v}_i) = \lambda_i\mathbf{v}_i$, which is what we needed to show. □

- **Remark:** Suppose that V is a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$.

- **Remark:** Suppose that V is a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$.
  - By Theorems 4.3.2 and 8.5.4, linear functions from V to V that have a diagonal matrix are precisely those that can be defined starting from some basis, and then scaling each of the basis elements.

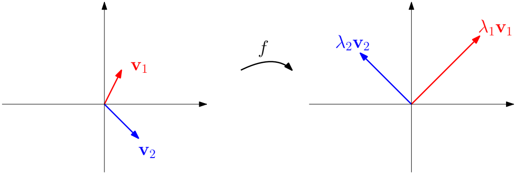


- **Remark:** Suppose that V is a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$.
  - By Theorems 4.3.2 and 8.5.4, linear functions from V to V that have a diagonal matrix are precisely those that can be defined starting from some basis, and then scaling each of the basis elements.



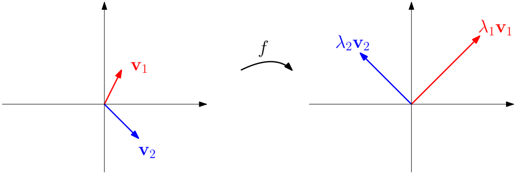
  - Indeed, suppose that $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is some basis of V, and that $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ are some scalars.

- **Remark:** Suppose that V is a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$.
  - By Theorems 4.3.2 and 8.5.4, linear functions from V to V that have a diagonal matrix are precisely those that can be defined starting from some basis, and then scaling each of the basis elements.



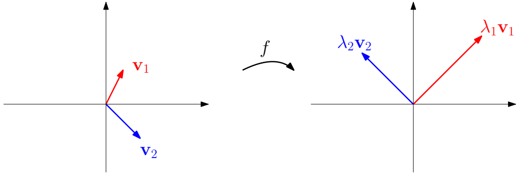
  - Indeed, suppose that $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is some basis of V, and that $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ are some scalars.
  - By Theorem 4.3.2, there exists a unique linear function $f : V \to V$ such that $f(\mathbf{v}_i) = \lambda_i \mathbf{v}_i$.

- **Remark:** Suppose that V is a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$.
    - By Theorems 4.3.2 and 8.5.4, linear functions from V to V that have a diagonal matrix are precisely those that can be defined starting from some basis, and then scaling each of the basis elements.



    - Indeed, suppose that $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is some basis of V, and that $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ are some scalars.
    - By Theorem 4.3.2, there exists a unique linear function $f : V \to V$ such that $f(\mathbf{v}_i) = \lambda_i \mathbf{v}_i$.
    - But then by Theorem 8.5.4, ${}_{\mathcal{B}}\big[\ f\ \big]_{\mathcal{B}} = D(\lambda_1, \ldots, \lambda_n)$.

- **Remark:** Suppose that V is a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$.
  - By Theorems 4.3.2 and 8.5.4, linear functions from V to V that have a diagonal matrix are precisely those that can be defined starting from some basis, and then scaling each of the basis elements.



  - Indeed, suppose that $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is some basis of V, and that $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ are some scalars.
  - By Theorem 4.3.2, there exists a unique linear function $f : V \to V$ such that $f(\mathbf{v}_i) = \lambda_i \mathbf{v}_i$.
  - But then by Theorem 8.5.4, ${}_{\mathcal{B}}[\ f\ ]_{\mathcal{B}} = D(\lambda_1, \ldots, \lambda_n)$.
  - By Theorem 8.5.4, the converse also holds.

Definition

A matrix $A \in \mathbb{F}^{n \times n}$ (where $\mathbb{F}$ is a field) is *diagonalizable* if it is similar to a diagonal matrix. To *diagonalize* a diagonalizable matrix A means to compute a diagonal matrix D and an invertible matrix P such that $D = P^{-1}AP$ (equivalently: $A = PDP^{-1}$).

Theorem 8.5.6

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n\times n}$ be a matrix. Then A is diagonalizable if and only if $\mathbb{F}^n$ has an eigenbasis associated with A. Moreover, if $\mathcal{P} = \{\mathbf{p}_1, \dots, \mathbf{p}_n\}$ is any eigenbasis of $\mathbb{F}^n$ associated with A, and $\lambda_1, \dots, \lambda_n$ are the eigenvalues of A associated with the eigenvectors $\mathbf{p}_1, \dots, \mathbf{p}_n$, respectively, then

$$D = P^{-1}AP \qquad \text{and} \qquad A = PDP^{-1},$$

where $D = D(\lambda_1, \dots, \lambda_n)$ and $P = \begin{bmatrix} \mathbf{p}_1 & \dots & \mathbf{p}_n \end{bmatrix}$.

- Proof: Lecture Notes.

Theorem 8.5.6

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$ be a matrix. Then A is diagonalizable if and only if $\mathbb{F}^n$ has an eigenbasis associated with A. Moreover, if $\mathcal{P} = \{\mathbf{p}_1, \dots, \mathbf{p}_n\}$ is any eigenbasis of $\mathbb{F}^n$ associated with A, and $\lambda_1, \dots, \lambda_n$ are the eigenvalues of A associated with the eigenvectors $\mathbf{p}_1, \dots, \mathbf{p}_n$, respectively, then

$$D = P^{-1}AP \qquad \text{and} \qquad A = PDP^{-1},$$

where $D = D(\lambda_1, \dots, \lambda_n)$ and $P = \begin{bmatrix} \mathbf{p}_1 & \dots & \mathbf{p}_n \end{bmatrix}$.

- Proof: Lecture Notes.
  - Theorem 8.5.6 can be obtained as a corollary of Theorem 8.5.4 (try it!).
  - However, in the Lecture Notes, there is a proof "from scratch" (i.e. one that uses matrices only).

Theorem 8.5.6

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n\times n}$ be a matrix. Then A is diagonalizable if and only if $\mathbb{F}^n$ has an eigenbasis associated with A. Moreover, if $\mathcal{P} = \{\mathbf{p}_1, \dots, \mathbf{p}_n\}$ is any eigenbasis of $\mathbb{F}^n$ associated with A, and $\lambda_1, \dots, \lambda_n$ are the eigenvalues of A associated with the eigenvectors $\mathbf{p}_1, \dots, \mathbf{p}_n$, respectively, then

$$D \quad = \quad P^{-1}AP \qquad \text{and} \qquad A \quad = \quad PDP^{-1},$$

where $D = D(\lambda_1, \dots, \lambda_n)$ and $P = \begin{bmatrix} \mathbf{p}_1 & \dots & \mathbf{p}_n \end{bmatrix}$.

Corollary 8.5.7

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n\times n}$. If A has n distinct eigenvalues, then A is diagonalizable.

Proof.

Theorem 8.5.6

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$ be a matrix. Then A is diagonalizable if and only if $\mathbb{F}^n$ has an eigenbasis associated with A. Moreover, if $\mathcal{P} = \{\mathbf{p}_1, \dots, \mathbf{p}_n\}$ is any eigenbasis of $\mathbb{F}^n$ associated with A, and $\lambda_1, \dots, \lambda_n$ are the eigenvalues of A associated with the eigenvectors $\mathbf{p}_1, \dots, \mathbf{p}_n$, respectively, then

$$D = P^{-1}AP \qquad \text{and} \qquad A = PDP^{-1},$$

where $D = D(\lambda_1, \dots, \lambda_n)$ and $P = \begin{bmatrix} \mathbf{p}_1 & \dots & \mathbf{p}_n \end{bmatrix}$.

Corollary 8.5.7

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$. If A has n distinct eigenvalues, then A is diagonalizable.

Proof. Assume that A has n distinct eigenvalues. By Corollary 8.4.6, $\mathbb{F}^n$ has an eigenbasis associated with A. So, by Theorem 8.5.6, A is diagonalizable. □

- Theorems 8.4.5 and 8.5.6 together give us a recipe for determining whether a matrix $A \in \mathbb{F}^{n \times n}$ is diagonalizable, and if so, for diagonalizing it (i.e. for finding a diagonal matrix D and an invertible matrix P, both in $\mathbb{F}^{n \times n}$, such that $D = P^{-1}AP$).
- We proceed as follows (next two slides).

1. We compute the characteristic polynomial $p_A(\lambda)$ and its roots. By Theorem 8.2.2, the roots of $p_A(\lambda)$ are the eigenvalues of A, and we can read off the algebraic multiplicities of those eigenvalues from the polynomial $p_A(\lambda)$.
    - Computing the roots of $p_A(\lambda)$ is the computationally tricky part, since there is no formula for computing the roots of a high-degree polynomial. If we cannot figure out how to compute the roots of $p_A(\lambda)$, then we are stuck: the matrix A may or may not be diagonalizable, but computationally, we cannot diagonalize it.

1. We compute the characteristic polynomial $p_A(\lambda)$ and its roots. By Theorem 8.2.2, the roots of $p_A(\lambda)$ are the eigenvalues of A, and we can read off the algebraic multiplicities of those eigenvalues from the polynomial $p_A(\lambda)$.
    - Computing the roots of $p_A(\lambda)$ is the computationally tricky part, since there is no formula for computing the roots of a high-degree polynomial. If we cannot figure out how to compute the roots of $p_A(\lambda)$, then we are stuck: the matrix A may or may not be diagonalizable, but computationally, we cannot diagonalize it.
2. If the sum of algebraic multiplicities of the eigenvalues of A is less than n, then by Theorem 8.4.5, $\mathbb{F}^n$ does not have an eigenbasis associated with A, and so by Theorem 8.5.6, A is not diagonalizable.

1. We compute the characteristic polynomial $p_A(\lambda)$ and its roots. By Theorem 8.2.2, the roots of $p_A(\lambda)$ are the eigenvalues of A, and we can read off the algebraic multiplicities of those eigenvalues from the polynomial $p_A(\lambda)$.
    - Computing the roots of $p_A(\lambda)$ is the computationally tricky part, since there is no formula for computing the roots of a high-degree polynomial. If we cannot figure out how to compute the roots of $p_A(\lambda)$, then we are stuck: the matrix A may or may not be diagonalizable, but computationally, we cannot diagonalize it.
2. If the sum of algebraic multiplicities of the eigenvalues of A is less than n, then by Theorem 8.4.5, $\mathbb{F}^n$ does not have an eigenbasis associated with A, and so by Theorem 8.5.6, A is not diagonalizable.
3. From now on, we assume that the sum of algebraic multiplicities of the eigenvalues of A, call them $\lambda_1, \dots, \lambda_k$, is n. We then compute a basis $\mathcal{B}_i$ for each eigenspace $E_{\lambda_i}(A)$, which allows us to compute the geometric multiplicities of all the eigenvalues of A.

4. If the geometric multiplicity of some eigenvalue of A is smaller than its algebraic multiplicity, then by Theorem 8.4.5, $\mathbb{F}^n$ does not have an eigenbasis associated with A, and so by Theorem 8.5.6, A is not diagonalizable.

4. If the geometric multiplicity of some eigenvalue of A is smaller than its algebraic multiplicity, then by Theorem 8.4.5, $\mathbb{F}^n$ does not have an eigenbasis associated with A, and so by Theorem 8.5.6, A is not diagonalizable.
5. From now on, we assume that the geometric multiplicity of each eigenvalue of A is equal to its algebraic multiplicity. Theorem 8.4.5 then guarantees that $\mathbb{F}^n$ has an eigenbasis associated with A, and moreover, that $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$ is one such eigenbasis.

4. If the geometric multiplicity of some eigenvalue of A is smaller than its algebraic multiplicity, then by Theorem 8.4.5, $\mathbb{F}^n$ does not have an eigenbasis associated with A, and so by Theorem 8.5.6, A is not diagonalizable.
5. From now on, we assume that the geometric multiplicity of each eigenvalue of A is equal to its algebraic multiplicity. Theorem 8.4.5 then guarantees that $\mathbb{F}^n$ has an eigenbasis associated with A, and moreover, that $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$ is one such eigenbasis.
6. By Theorem 8.5.6, A is diagonalizable. We now follow the recipe from Theorem 8.5.6 to actually diagonalize A.

4. If the geometric multiplicity of some eigenvalue of A is smaller than its algebraic multiplicity, then by Theorem 8.4.5, $\mathbb{F}^n$ does not have an eigenbasis associated with A, and so by Theorem 8.5.6, A is not diagonalizable.
5. From now on, we assume that the geometric multiplicity of each eigenvalue of A is equal to its algebraic multiplicity. Theorem 8.4.5 then guarantees that $\mathbb{F}^n$ has an eigenbasis associated with A, and moreover, that $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ is one such eigenbasis.
6. By Theorem 8.5.6, A is diagonalizable. We now follow the recipe from Theorem 8.5.6 to actually diagonalize A.
7. We form the matrix P whose columns are precisely the vectors in the eigenbasis $\mathcal{B}$. We form the diagonal matrix D, where on the main diagonal we place the eigenvalues of A, taking care that, for each $i \in \{1, \dots, n\}$, the i-th entry on the main diagonal of D is the eigenvalue associated with the i-th column of P (which is, by construction, an eigenvector of A). Now $D = P^{-1}AP$.

Example 8.5.8.

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

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

Determine whether A is diagonalizable, and if so, diagonalize it.

Solution.

Example 8.5.8.

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

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

Determine whether A is diagonalizable, and if so, diagonalize it.

Solution. The matrix A is precisely the matrix from Example 8.2.4. In that example, we determined that A has two eigenvalues:

- $\lambda_1 = 4$ (with algebraic multiplicity 1 and geometric multiplicity 1);
- $\lambda_2 = 5$ (with algebraic multiplicity 2 and geometric multiplicity 2).

Example 8.5.8.

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

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

Determine whether A is diagonalizable, and if so, diagonalize it.

Solution. The matrix A is precisely the matrix from Example 8.2.4. In that example, we determined that A has two eigenvalues:

- $\lambda_1 = 4$ (with algebraic multiplicity 1 and geometric multiplicity 1);
- $\lambda_2 = 5$ (with algebraic multiplicity 2 and geometric multiplicity 2).

Since the sum of algebraic multiplicities of the eigenvalues of A is 3, and since the geometric multiplicity of each eigenvalue of A is equal to its algebraic multiplicity, we see that the 3×3 matrix A is indeed diagonalizable.

Solution (continued). Reminder: $\lambda_1 = 4$, $\lambda_2 = 5$.

In Example 8.2.4, we saw that:

- $\left\{ \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} \right\}$ is a basis of the eigespace $E_{\lambda_1}(A)$;
- $\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \right\}$ is a basis of the eigenspace $E_{\lambda_2}(A)$.

Solution (continued). Reminder: $\lambda_1 = 4$, $\lambda_2 = 5$.

In Example 8.2.4, we saw that:

- $\left\{ \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} \right\}$ is a basis of the eigespace $E_{\lambda_1}(A)$;
- $\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \right\}$ is a basis of the eigenspace $E_{\lambda_2}(A)$.

So, we set

$$D \;:=\; \begin{bmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{bmatrix} \quad \text{and} \quad P \;:=\; \begin{bmatrix} -1 & 0 & -2 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and we see that $D = P^{-1}AP$. $\square$

Example 8.5.9

Consider the following matrix in $\mathbb{C}^{5\times 5}$:

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

Determine whether A is diagonalizable, and if so, diagonalize it.

Solution.

Example 8.5.9

Consider the following matrix in $\mathbb{C}^{5\times 5}$:

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

Determine whether A is diagonalizable, and if so, diagonalize it.

Solution. The matrix A is precisely the matrix from Example 8.2.8.

Example 8.5.9

Consider the following matrix in $\mathbb{C}^{5\times 5}$:

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

Determine whether A is diagonalizable, and if so, diagonalize it.

Solution. The matrix A is precisely the matrix from Example 8.2.8. In that example, we determined that A has three eigenvalues:

- $\lambda_1 = 1$ (with alg. mult. 2 and geom. mult. 2);
- $\lambda_2 = 2$ (with alg. mult. 1 and geom. mult. 1);
- $\lambda_3 = 3$ (with alg. mult. 2 and geom. mult. 1).

Example 8.5.9

Consider the following matrix in $\mathbb{C}^{5\times 5}$:

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

Determine whether A is diagonalizable, and if so, diagonalize it.

Solution. The matrix A is precisely the matrix from Example 8.2.8. In that example, we determined that A has three eigenvalues:

- $\lambda_1 = 1$ (with alg. mult. 2 and geom. mult. 2);
- $\lambda_2 = 2$ (with alg. mult. 1 and geom. mult. 1);
- $\lambda_3 = 3$ (with alg. mult. 2 and geom. mult. 1).

Since the geometric multiplicity of the eigenvalue $\lambda_3 = 3$ is strictly smaller than the algebraic multiplicity, we see that A is not diagonalizable. □

- Suppose that we have successfully diagonalized a square matrix $A \in \mathbb{F}^{n \times n}$ (where $\mathbb{F}$ is a field), that is, that we have computed a diagonal matrix D and an invertible matrix P, both in $\mathbb{F}^{n \times n}$, such that $D = P^{-1}AP$.

- Suppose that we have successfully diagonalized a square matrix $A \in \mathbb{F}^{n \times n}$ (where $\mathbb{F}$ is a field), that is, that we have computed a diagonal matrix D and an invertible matrix P, both in $\mathbb{F}^{n \times n}$, such that $D = P^{-1}AP$.
- Then we can easily read off the spectrum and a basis of each eigenspace of A, as Proposition 8.5.12 (next slide) shows.

- Suppose that we have successfully diagonalized a square matrix $A \in \mathbb{F}^{n \times n}$ (where $\mathbb{F}$ is a field), that is, that we have computed a diagonal matrix D and an invertible matrix P, both in $\mathbb{F}^{n \times n}$, such that $D = P^{-1}AP$.
- Then we can easily read off the spectrum and a basis of each eigenspace of A, as Proposition 8.5.12 (next slide) shows.
- This proposition essentially summarizes various facts about diagonalizable matrices that we have proven already, but it is convenient to have them stated in one proposition.

- Suppose that we have successfully diagonalized a square matrix $A \in \mathbb{F}^{n \times n}$ (where $\mathbb{F}$ is a field), that is, that we have computed a diagonal matrix D and an invertible matrix P, both in $\mathbb{F}^{n \times n}$, such that $D = P^{-1}AP$.
- Then we can easily read off the spectrum and a basis of each eigenspace of A, as Proposition 8.5.12 (next slide) shows.
- This proposition essentially summarizes various facts about diagonalizable matrices that we have proven already, but it is convenient to have them stated in one proposition.
  - The proof is in the Lecture Notes. Here, we omit it.

Proposition 8.5.12

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$. Assume that $D = P^{-1}AP$, where $D = D(\lambda_1, \dots, \lambda_n)$ is a diagonal and $P = \begin{bmatrix} \mathbf{p}_1 & \dots & \mathbf{p}_n \end{bmatrix}$ an invertible matrix, both in $\mathbb{F}^{n \times n}$. Then the characteristic polynomial of A is

$$p_A(\lambda) \;=\; \prod_{i=1}^{n} (\lambda - \lambda_i) \;=\; (\lambda - \lambda_1) \dots (\lambda - \lambda_n),$$

and the spectrum of A is $\{\lambda_1, \dots, \lambda_n\}$. Moreover, for each eigenvalue λ_0 of A,[a] the algebraic and geometric multiplicity of λ_0 are both equal to the number of times that λ_0 appears on the main diagonal of D, and moreover, if λ_0 appears precisely in positions $i_1, \dots, i_k$ of the main diagonal of D, then the corresponding columns of P (i.e. vectors $\mathbf{p}_{i_1}, \dots, \mathbf{p}_{i_k}$) form a basis of the eigenspace $E_{\lambda_0}(A)$. Finally, $\{\mathbf{p}_1, \dots, \mathbf{p}_n\}$ is an eigenbasis of $\mathbb{F}^n$ associated with the matrix A.

---

[a] So, $\lambda_0 \in \{\lambda_1, \dots, \lambda_n\}$, since $\{\lambda_1, \dots, \lambda_n\}$ is the spectrum of A.

Example 8.5.13

Consider the following matrices in $\mathbb{C}^{6\times 6}$ (color coded for emphasis):

$$D = \begin{bmatrix} 5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 \end{bmatrix}, \qquad P = \begin{bmatrix} 1 & 3 & 8 & 8 & 3 & 4 \\ 2 & 8 & 0 & 0 & 0 & 2 \\ 5 & 4 & 6 & 4 & 5 & 0 \\ 0 & 5 & 8 & 5 & 4 & 3 \\ 1 & 0 & 8 & 0 & 3 & 0 \\ 0 & 2 & 0 & 3 & 0 & 2 \end{bmatrix}.$$

It can be checked that P is invertible (for example, we can compute that $\det(P) = -1020 \neq 0$, and so by Theorem 7.4.1, P is invertible). We now set $A = PDP^{-1}$, so that $D = P^{-1}AP$. Then by Proposition 8.5.12, all the following hold (next three slides):

$$D = \begin{bmatrix} 5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 \end{bmatrix}, \qquad P = \begin{bmatrix} 1 & 3 & 8 & 8 & 3 & 4 \\ 2 & 8 & 0 & 0 & 0 & 2 \\ 5 & 4 & 6 & 4 & 5 & 0 \\ 0 & 5 & 8 & 5 & 4 & 3 \\ 1 & 0 & 8 & 0 & 3 & 0 \\ 0 & 2 & 0 & 3 & 0 & 2 \end{bmatrix}.$$

Example 8.5.13 (continued)

- the characteristic polynomial of A is

$$p_A(\lambda) = (\lambda - 3)(\lambda - 4)^3(\lambda - 5)^2;$$

- the spectrum of A is $\{5, 4, 5, 3, 4, 4\}$, which we can optionally reorder as $\{3, 4, 4, 4, 5, 5\}$;
- the eigenvalues of A are 3 (with algebraic and geometric multiplicity 1), 4 (with algebraic and geometric multiplicity 3), and 5 (with algebraic and geometric multiplicity 2);

$$D = \begin{bmatrix} 5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 \end{bmatrix}, \qquad P = \begin{bmatrix} 1 & 3 & 8 & 8 & 3 & 4 \\ 2 & 8 & 0 & 0 & 0 & 2 \\ 5 & 4 & 6 & 4 & 5 & 0 \\ 0 & 5 & 8 & 5 & 4 & 3 \\ 1 & 0 & 8 & 0 & 3 & 0 \\ 0 & 2 & 0 & 3 & 0 & 2 \end{bmatrix}.$$

Example 8.5.13 (continued)

- we can read off bases of the eigenspaces $E_3(A)$, $E_4(A)$, and $E_5(A)$, as follows:
  - a basis of $E_3(A)$ is $\left\{ \begin{bmatrix} 8 \\ 0 \\ 4 \\ 5 \\ 0 \\ 3 \end{bmatrix} \right\}$,
  - a basis of $E_4(A)$ is $\left\{ \begin{bmatrix} 3 \\ 8 \\ 4 \\ 5 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 5 \\ 4 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 2 \\ 0 \\ 3 \\ 0 \\ 2 \end{bmatrix} \right\}$,
  - a basis of $E_5(A)$ is $\left\{ \begin{bmatrix} 1 \\ 2 \\ 5 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 8 \\ 0 \\ 6 \\ 8 \\ 8 \\ 0 \end{bmatrix} \right\}$;

$$D = \begin{bmatrix} 5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 \end{bmatrix}, \qquad P = \begin{bmatrix} 1 & 3 & 8 & 8 & 3 & 4 \\ 2 & 8 & 0 & 0 & 0 & 2 \\ 5 & 4 & 6 & 4 & 5 & 0 \\ 0 & 5 & 8 & 5 & 4 & 3 \\ 1 & 0 & 8 & 0 & 3 & 0 \\ 0 & 2 & 0 & 3 & 0 & 2 \end{bmatrix}.$$

Example 8.5.13 (continued)

- the columns of P form an eigenbasis of $\mathbb{C}^n$ associated with the matrix A.