Linear Algebra 2

Lecture #19
The Cayley-Hamilton theorem. Diagonalization

Irena Penev

April 13, 2023

- This lecture has four parts:

- This lecture has four parts:
  1. the Cayley-Hamilton theorem;

- This lecture has four parts:
  1. the Cayley-Hamilton theorem;
  2. eigenvectors and eigenvalues of similar matrices;

- This lecture has four parts:
  1. the Cayley-Hamilton theorem;
  2. eigenvectors and eigenvalues of similar matrices;
  3. algebraic and geometric multiplicities revisited (proof of Theorem 2.5 from Lecture Notes 18);

- This lecture has four parts:
  1. the Cayley-Hamilton theorem;
  2. eigenvectors and eigenvalues of similar matrices;
  3. algebraic and geometric multiplicities revisited (proof of Theorem 2.5 from Lecture Notes 18);
  4. diagonal matrices and diagonalization.

- This lecture has four parts:
  1. the Cayley-Hamilton theorem;
  2. eigenvectors and eigenvalues of similar matrices;
  3. algebraic and geometric multiplicities revisited (proof of Theorem 2.5 from Lecture Notes 18);
  4. diagonal matrices and diagonalization.

- In what follows, $\mathbb{F}$ is a fixed field.

1. The Cayley-Hamilton theorem

1. The Cayley-Hamilton theorem

The Cayley-Hamilton theorem

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

1. The Cayley-Hamilton theorem

The Cayley-Hamilton theorem

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

- 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.

The Cayley-Hamilton theorem

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

- For example, for a matrix $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, we have

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

and

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

- Before proving the Cayley-Hamilton theorem, we first need to revisit the adjugate matrix.

- Before proving the Cayley-Hamilton theorem, we first need to revisit the adjugate matrix.
- Recall that for a matrix $A = [\ a_{i,j}\ ]_{n\times n}$ $(n \geq 2)$ in $\mathbb{F}^{n\times n}$ and indices $i, j \in \{1, \dots, n\}$, $A_{i,j}$ is the $(n-1)\times(n-1)$ matrix obtained from A by deleting the i-th row and j-th column, and the cofactor $C_{i,j}$ is defined to be $C_{i,j} := (-1)^{i+j}\det(A_{i,j})$.

- Before proving the Cayley-Hamilton theorem, we first need to revisit the adjugate matrix.
- Recall that for a matrix $A = [\ a_{i,j}\ ]_{n \times n}$ $(n \geq 2)$ in $\mathbb{F}^{n \times n}$ and indices $i, j \in \{1, \dots, n\}$, $A_{i,j}$ is the $(n-1) \times (n-1)$ matrix obtained from A by deleting the i-th row and j-th column, and the cofactor $C_{i,j}$ is defined to be $C_{i,j} := (-1)^{i+j}\det(A_{i,j})$.
- The *cofactor matrix* of A is the matrix $[\ C_{i,j}\ ]_{n \times n}$.

- Before proving the Cayley-Hamilton theorem, we first need to revisit the adjugate matrix.
- Recall that for a matrix $A = [\ a_{i,j}\ ]_{n\times n}$ ($n \geq 2$) in $\mathbb{F}^{n\times n}$ and indices $i, j \in \{1, \dots, n\}$, $A_{i,j}$ is the $(n-1)\times(n-1)$ matrix obtained from A by deleting the i-th row and j-th column, and the cofactor $C_{i,j}$ is defined to be $C_{i,j} := (-1)^{i+j}\det(A_{i,j})$.
- The *cofactor matrix* of A is the matrix $[\ C_{i,j}\ ]_{n\times n}$.
- The *adjugate matrix* (also called the *classical adjoint*) of A, denoted by adj(A), is the transponse of the cofactor matrix of A, i.e.

$$\text{adj}(A) \ := \ [\ C_{i,j}\ ]_{n\times n}{}^T.$$

Theorem 1.1

Let $A \in \mathbb{F}^{n\times n}$ $(n \geq 2)$. Then $\operatorname{adj}(A)\, A = \det(A) I_n = A \operatorname{adj}(A)$. Consequently, if A is invertible, then $A^{-1} = \frac{1}{\det(A)}\operatorname{adj}(A)$.

Proof.

Theorem 1.1

Let $A \in \mathbb{F}^{n\times n}$ ($n \geq 2$). Then $\mathrm{adj}(A)\, A = \det(A) I_n = A\, \mathrm{adj}(A)$. Consequently, if A is invertible, then $A^{-1} = \frac{1}{\det(A)} \mathrm{adj}(A)$.

Proof. Let us first show that the first statement implies the second.

Theorem 1.1

Let $A \in \mathbb{F}^{n\times n}$ ($n \geq 2$). Then $\operatorname{adj}(A)\, A = \det(A) I_n = A \operatorname{adj}(A)$. Consequently, if A is invertible, then $A^{-1} = \frac{1}{\det(A)}\operatorname{adj}(A)$.

Proof. Let us first show that the first statement implies the second. Indeed, if A is invertible, then $\det(A) \neq 0$,

Theorem 1.1

Let $A \in \mathbb{F}^{n \times n}$ ($n \geq 2$). Then $\text{adj}(A)\, A = \det(A) I_n = A\, \text{adj}(A)$. Consequently, if A is invertible, then $A^{-1} = \frac{1}{\det(A)} \text{adj}(A)$.

Proof. Let us first show that the first statement implies the second. Indeed, if A is invertible, then $\det(A) \neq 0$, and so if the first statement holds, then we get that

$$\left(\frac{1}{\det(A)}\text{adj}(A)\right) A \;=\; I_n \;=\; A \left(\frac{1}{\det(A)}\text{adj}(A)\right),$$

and consequently, $A^{-1} = \frac{1}{\det(A)} \text{adj}(A)$.

Theorem 1.1

Let $A \in \mathbb{F}^{n \times n}$ $(n \geq 2)$. Then $\text{adj}(A)\, A = \det(A) I_n = A\, \text{adj}(A)$. Consequently, if A is invertible, then $A^{-1} = \frac{1}{\det(A)} \text{adj}(A)$.

Proof (conitnued). It remains to prove the first statement, i.e. that $\text{adj}(A)\, A = \det(A) I_n = A\, \text{adj}(A)$.

Theorem 1.1

Let $A \in \mathbb{F}^{n\times n}$ $(n \geq 2)$. Then $\text{adj}(A)\, A = \det(A) I_n = A\, \text{adj}(A)$. Consequently, if A is invertible, then $A^{-1} = \frac{1}{\det(A)}\text{adj}(A)$.

Proof (conitnued). It remains to prove the first statement, i.e. that $\text{adj}(A)\, A = \det(A) I_n = A\, \text{adj}(A)$. We will do this by proving that the matrices $\text{adj}(A)\, A$, $\det(A) I_n$, and $A\, \text{adj}(A)$ have the same corresponding entries.

Theorem 1.1

Let $A \in \mathbb{F}^{n\times n}$ $(n \geq 2)$. Then $\operatorname{adj}(A)\, A = \det(A) I_n = A \operatorname{adj}(A)$. Consequently, if A is invertible, then $A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A)$.

Proof (conitnued). It remains to prove the first statement, i.e. that $\operatorname{adj}(A)\, A = \det(A) I_n = A \operatorname{adj}(A)$. We will do this by proving that the matrices $\operatorname{adj}(A)\, A$, $\det(A) I_n$, and $A \operatorname{adj}(A)$ have the same corresponding entries.

Fix indices $i, j \in \{1, \dots, n\}$.

Theorem 1.1

Let $A \in \mathbb{F}^{n\times n}$ $(n \geq 2)$. Then $\mathrm{adj}(A)\, A = \det(A) I_n = A\, \mathrm{adj}(A)$. Consequently, if A is invertible, then $A^{-1} = \frac{1}{\det(A)}\mathrm{adj}(A)$.

Proof (conitnued). It remains to prove the first statement, i.e. that $\mathrm{adj}(A)\, A = \det(A) I_n = A\, \mathrm{adj}(A)$. We will do this by proving that the matrices $\mathrm{adj}(A)\, A$, $\det(A) I_n$, and $A\, \mathrm{adj}(A)$ have the same corresponding entries.

Fix indices $i, j \in \{1, \dots, n\}$. The i, j-th entry of the matrix $\det(A) I_n$ is $\det(A)$ if $i = j$, and is zero if $i \neq j$.

Theorem 1.1

Let $A \in \mathbb{F}^{n\times n}$ ($n \geq 2$). Then $\operatorname{adj}(A)\, A = \det(A) I_n = A \operatorname{adj}(A)$. Consequently, if A is invertible, then $A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A)$.

Proof (conitnued). It remains to prove the first statement, i.e. that $\operatorname{adj}(A)\, A = \det(A) I_n = A \operatorname{adj}(A)$. We will do this by proving that the matrices $\operatorname{adj}(A)\, A$, $\det(A) I_n$, and $A \operatorname{adj}(A)$ have the same corresponding entries.

Fix indices $i, j \in \{1, \ldots, n\}$. The i,j-th entry of the matrix $\det(A) I_n$ is $\det(A)$ if $i = j$, and is zero if $i \neq j$. We must show this holds for the i,j-th entry of the matrices $\operatorname{adj}(A)\, A$ and $A \operatorname{adj}(A)$ as well.

Theorem 1.1

Let $A \in \mathbb{F}^{n\times n}$ $(n \geq 2)$. Then $\operatorname{adj}(A)\, A = \det(A) I_n = A \operatorname{adj}(A)$. Consequently, if A is invertible, then $A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A)$.

Proof (conitnued). Reminder: WTS the i, j-th entry of $\operatorname{adj}(A)\, A$ and $A \operatorname{adj}(A)$ is $\det(A)$ if $i = j$, and is zero if $i \neq j$.

Theorem 1.1

Let $A \in \mathbb{F}^{n\times n}$ $(n \geq 2)$. Then $\text{adj}(A)\, A = \det(A) I_n = A\, \text{adj}(A)$. Consequently, if A is invertible, then $A^{-1} = \frac{1}{\det(A)}\text{adj}(A)$.

Proof (conitnued). Reminder: WTS the i, j-th entry of $\text{adj}(A)\, A$ and $A\, \text{adj}(A)$ is $\det(A)$ if $i = j$, and is zero if $i \neq j$.

We first consider the matrix $\text{adj}(A)\, A$.

Theorem 1.1

Let $A \in \mathbb{F}^{n\times n}$ $(n \geq 2)$. Then $\text{adj}(A)\, A = \det(A) I_n = A\, \text{adj}(A)$. Consequently, if A is invertible, then $A^{-1} = \frac{1}{\det(A)}\text{adj}(A)$.

Proof (conitnued). Reminder: WTS the i, j-th entry of $\text{adj}(A)\, A$ and $A\, \text{adj}(A)$ is $\det(A)$ if $i = j$, and is zero if $i \neq j$.

We first consider the matrix $\text{adj}(A)\, A$. The i-th row of $\text{adj}(A)$ is $\begin{bmatrix} C_{1,i} & \dots & C_{n,i} \end{bmatrix}$, and the j-th column of A is $\begin{bmatrix} a_{1,j} & \dots & a_{n,j} \end{bmatrix}^T$.

Theorem 1.1

Let $A \in \mathbb{F}^{n\times n}$ ($n \geq 2$). Then $\text{adj}(A)\ A = \det(A)I_n = A\ \text{adj}(A)$. Consequently, if A is invertible, then $A^{-1} = \frac{1}{\det(A)}\text{adj}(A)$.

Proof (conitnued). Reminder: WTS the i, j-th entry of $\text{adj}(A)\ A$ and $A\ \text{adj}(A)$ is $\det(A)$ if $i = j$, and is zero if $i \neq j$.

We first consider the matrix $\text{adj}(A)\ A$. The i-th row of $\text{adj}(A)$ is $\begin{bmatrix} C_{1,i} & \dots & C_{n,i} \end{bmatrix}$, and the j-th column of A is $\begin{bmatrix} a_{1,j} & \dots & a_{n,j} \end{bmatrix}^T$. So, the i, j-th entry of $\text{adj}(A)\ A$ is $\sum_{k=1}^{n} a_{k,j} C_{k,i}$.

Theorem 1.1

Let $A \in \mathbb{F}^{n\times n}$ ($n \geq 2$). Then $\text{adj}(A)\, A = \det(A) I_n = A\, \text{adj}(A)$. Consequently, if A is invertible, then $A^{-1} = \frac{1}{\det(A)}\text{adj}(A)$.

Proof (conitnued). Reminder: WTS the i, j-th entry of $\text{adj}(A)\, A$ and $A\, \text{adj}(A)$ is $\det(A)$ if $i = j$, and is zero if $i \neq j$.

We first consider the matrix $\text{adj}(A)\, A$. The i-th row of $\text{adj}(A)$ is $\begin{bmatrix} C_{1,i} & \dots & C_{n,i} \end{bmatrix}$, and the j-th column of A is $\begin{bmatrix} a_{1,j} & \dots & a_{n,j} \end{bmatrix}^T$. So, the i, j-th entry of $\text{adj}(A)\, A$ is $\sum_{k=1}^{n} a_{k,j} C_{k,i}$.

Now, let B_1 be the matrix obtained by replacing the i-th column of A by the j-th column of A.

Theorem 1.1

Let $A \in \mathbb{F}^{n\times n}$ $(n \geq 2)$. Then $\text{adj}(A)\, A = \det(A) I_n = A\, \text{adj}(A)$. Consequently, if A is invertible, then $A^{-1} = \frac{1}{\det(A)}\text{adj}(A)$.

Proof (conitnued). Reminder: WTS the i,j-th entry of $\text{adj}(A)\, A$ and $A\, \text{adj}(A)$ is $\det(A)$ if $i = j$, and is zero if $i \neq j$.

We first consider the matrix $\text{adj}(A)\, A$. The i-th row of $\text{adj}(A)$ is $\begin{bmatrix} C_{1,i} & \dots & C_{n,i} \end{bmatrix}$, and the j-th column of A is $\begin{bmatrix} a_{1,j} & \dots & a_{n,j} \end{bmatrix}^T$. So, the i,j-th entry of $\text{adj}(A)\, A$ is $\sum_{k=1}^{n} a_{k,j} C_{k,i}$.

Now, let B_1 be the matrix obtained by replacing the i-th column of A by the j-th column of A. Then $\det(B_1) = \sum_{k=1}^{n} a_{k,j} C_{k,i}$ (via Laplace expansion along the i-th column of B_1).

Theorem 1.1

Let $A \in \mathbb{F}^{n\times n}$ $(n \geq 2)$. Then $\text{adj}(A)\ A = \det(A) I_n = A\ \text{adj}(A)$. Consequently, if A is invertible, then $A^{-1} = \frac{1}{\det(A)}\text{adj}(A)$.

Proof (conitnued). Reminder: WTS the i,j-th entry of $\text{adj}(A)\ A$ and $A\ \text{adj}(A)$ is $\det(A)$ if $i = j$, and is zero if $i \neq j$.

We first consider the matrix $\text{adj}(A)\ A$. The i-th row of $\text{adj}(A)$ is $\begin{bmatrix} C_{1,i} & \dots & C_{n,i} \end{bmatrix}$, and the j-th column of A is $\begin{bmatrix} a_{1,j} & \dots & a_{n,j} \end{bmatrix}^T$. So, the i,j-th entry of $\text{adj}(A)\ A$ is $\sum_{k=1}^{n} a_{k,j} C_{k,i}$.

Now, let B_1 be the matrix obtained by replacing the i-th column of A by the j-th column of A. Then $\det(B_1) = \sum_{k=1}^{n} a_{k,j} C_{k,i}$ (via Laplace expansion along the i-th column of B_1). But if $i = j$, then $\det(B_1) = \det(A)$ (because $B_1 = A$), and if $i \neq j$, then $\det(B_1) = 0$ (because B_1 has two identical columns, namely, the i-th and j-th column).

Theorem 1.1

Let $A \in \mathbb{F}^{n\times n}$ $(n \geq 2)$. Then $\text{adj}(A)\,A = \det(A)I_n = A\,\text{adj}(A)$. Consequently, if A is invertible, then $A^{-1} = \frac{1}{\det(A)}\text{adj}(A)$.

Proof (conitnued). Reminder: WTS the i, j-th entry of $\text{adj}(A)\,A$ and $A\,\text{adj}(A)$ is $\det(A)$ if $i = j$, and is zero if $i \neq j$.

We now consider the matrix $A\,\text{adj}(A)$.

Theorem 1.1

Let $A \in \mathbb{F}^{n\times n}$ ($n \geq 2$). Then $\operatorname{adj}(A)\, A = \det(A) I_n = A \operatorname{adj}(A)$. Consequently, if A is invertible, then $A^{-1} = \frac{1}{\det(A)}\operatorname{adj}(A)$.

Proof (conitnued). Reminder: WTS the i,j-th entry of $\operatorname{adj}(A)\, A$ and $A \operatorname{adj}(A)$ is $\det(A)$ if $i = j$, and is zero if $i \neq j$.

We now consider the matrix $A \operatorname{adj}(A)$. The i-th row of the matrix A is $\begin{bmatrix} a_{i,1} & \dots & a_{i,n} \end{bmatrix}$ and the j-th column of $\operatorname{adj}(A)$ is $\begin{bmatrix} C_{j,1} & \dots & C_{j,n} \end{bmatrix}^T$.

Theorem 1.1

Let $A \in \mathbb{F}^{n\times n}$ $(n \geq 2)$. Then $\text{adj}(A)\, A = \det(A) I_n = A\, \text{adj}(A)$. Consequently, if A is invertible, then $A^{-1} = \frac{1}{\det(A)}\text{adj}(A)$.

Proof (conitnued). Reminder: WTS the i, j-th entry of $\text{adj}(A)\, A$ and $A\, \text{adj}(A)$ is $\det(A)$ if $i = j$, and is zero if $i \neq j$.

We now consider the matrix $A\, \text{adj}(A)$. The i-th row of the matrix A is $\begin{bmatrix} a_{i,1} & \dots & a_{i,n} \end{bmatrix}$ and the j-th column of $\text{adj}(A)$ is $\begin{bmatrix} C_{j,1} & \dots & C_{j,n} \end{bmatrix}^T$. So, the i, j-th entry of $A\, \text{adj}(A)$ is $\sum_{k=1}^{n} a_{i,k} C_{j,k}$.

Theorem 1.1

Let $A \in \mathbb{F}^{n\times n}$ $(n \geq 2)$. Then $\text{adj}(A)\,A = \det(A)I_n = A\,\text{adj}(A)$.
Consequently, if A is invertible, then $A^{-1} = \frac{1}{\det(A)}\text{adj}(A)$.

Proof (conitnued). Reminder: WTS the i, j-th entry of $\text{adj}(A)\,A$ and $A\,\text{adj}(A)$ is $\det(A)$ if $i = j$, and is zero if $i \neq j$.

We now consider the matrix $A\,\text{adj}(A)$. The i-th row of the matrix A is $\begin{bmatrix} a_{i,1} & \dots & a_{i,n} \end{bmatrix}$ and the j-th column of $\text{adj}(A)$ is $\begin{bmatrix} C_{j,1} & \dots & C_{j,n} \end{bmatrix}^T$. So, the i, j-th entry of $A\,\text{adj}(A)$ is $\sum_{k=1}^{n} a_{i,k}C_{j,k}$.

Now, let B_2 be the matrix obtained by replacing the j-th row of A by the i-th row of A.

Theorem 1.1

Let $A \in \mathbb{F}^{n\times n}$ $(n \geq 2)$. Then $\operatorname{adj}(A)\, A = \det(A) I_n = A \operatorname{adj}(A)$. Consequently, if A is invertible, then $A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A)$.

Proof (conitnued). Reminder: WTS the i,j-th entry of $\operatorname{adj}(A)\, A$ and $A \operatorname{adj}(A)$ is $\det(A)$ if $i = j$, and is zero if $i \neq j$.

We now consider the matrix $A \operatorname{adj}(A)$. The i-th row of the matrix A is $\begin{bmatrix} a_{i,1} & \dots & a_{i,n} \end{bmatrix}$ and the j-th column of $\operatorname{adj}(A)$ is $\begin{bmatrix} C_{j,1} & \dots & C_{j,n} \end{bmatrix}^T$. So, the i,j-th entry of $A \operatorname{adj}(A)$ is $\sum_{k=1}^{n} a_{i,k} C_{j,k}$.

Now, let B_2 be the matrix obtained by replacing the j-th row of A by the i-th row of A. Then $\det(B_2) = \sum_{k=1}^{n} a_{i,k} C_{j,k}$ (via Laplace expansion along the j-th row of B_2).

Theorem 1.1

Let $A \in \mathbb{F}^{n\times n}$ $(n \geq 2)$. Then $\text{adj}(A)\, A = \det(A) I_n = A\, \text{adj}(A)$. Consequently, if A is invertible, then $A^{-1} = \frac{1}{\det(A)}\text{adj}(A)$.

Proof (conitnued). Reminder: WTS the i, j-th entry of $\text{adj}(A)\, A$ and $A\, \text{adj}(A)$ is $\det(A)$ if $i = j$, and is zero if $i \neq j$.

We now consider the matrix $A\, \text{adj}(A)$. The i-th row of the matrix A is $\begin{bmatrix} a_{i,1} & \dots & a_{i,n} \end{bmatrix}$ and the j-th column of $\text{adj}(A)$ is $\begin{bmatrix} C_{j,1} & \dots & C_{j,n} \end{bmatrix}^T$. So, the i, j-th entry of $A\, \text{adj}(A)$ is $\sum_{k=1}^{n} a_{i,k} C_{j,k}$.

Now, let B_2 be the matrix obtained by replacing the j-th row of A by the i-th row of A. Then $\det(B_2) = \sum_{k=1}^{n} a_{i,k} C_{j,k}$ (via Laplace expansion along the j-th row of B_2). But if $i = j$, then $\det(B_2) = \det(A)$ (because $B_2 = A$), and if $i \neq j$, then $\det(B_2) = 0$ (because B_2 has two identical rows, namely, the i-th and j-th row). Q.E.D.

The Cayley-Hamilton theorem

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

Proof.

The Cayley-Hamilton theorem

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

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

The Cayley-Hamilton theorem

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

Proof. If $n = 1$, then the result is immediate. Indeed, suppose that $n = 1$, and consider any matrix $A = [a_{1,1}]$ in $\mathbb{F}^{1 \times 1}$.

The Cayley-Hamilton theorem

Let $A \in \mathbb{F}^{n \times n}$ be a matrix, and let $p_A(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_1\lambda + a_0$ be its characteristic polynomial. Then $A^n + a_{n-1}A^{n-1} + \dots + a_1 A + a_0 I_n = O_{n \times n}$.

Proof. If $n = 1$, then the result is immediate. Indeed, suppose that $n = 1$, and consider any matrix $A = [a_{1,1}]$ in $\mathbb{F}^{1 \times 1}$. Then $p_A(\lambda) = \det(\lambda I_1 - A) = \det(\left[\begin{array}{c} \lambda - a_{1,1} \end{array}\right]) = \lambda - a_{1,1}$, and we see that $A - a_{1,1} I_1 = O_{1 \times 1}$.

The Cayley-Hamilton theorem

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

Proof. If $n = 1$, then the result is immediate. Indeed, suppose that $n = 1$, and consider any matrix $A = [a_{1,1}]$ in $\mathbb{F}^{1 \times 1}$. Then $p_A(\lambda) = \det(\lambda I_1 - A) = \det(\begin{bmatrix} \lambda - a_{1,1} \end{bmatrix}) = \lambda - a_{1,1}$, and we see that $A - a_{1,1} I_1 = O_{1 \times 1}$.

The Cayley-Hamilton theorem

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

Proof. If $n = 1$, then the result is immediate. Indeed, suppose that $n = 1$, and consider any matrix $A = [a_{1,1}]$ 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 $A - a_{1,1}I_1 = O_{1 \times 1}$.

So, assume that $n \geq 2$. By Theorem 1.1 applied to the matrix $\lambda I_n - A$ (where λ is a variable), we get that

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

The Cayley-Hamilton theorem

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

Proof. If $n = 1$, then the result is immediate. Indeed, suppose that $n = 1$, and consider any matrix $A = [a_{1,1}]$ in $\mathbb{F}^{1\times 1}$. Then $p_A(\lambda) = \det(\lambda I_1 - A) = \det(\left[\begin{array}{c} \lambda - a_{1,1} \end{array}\right]) = \lambda - a_{1,1}$, and we see that $A - a_{1,1}I_1 = O_{1\times 1}$.

So, assume that $n \geq 2$. By Theorem 1.1 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}, and consequently, each entry of $\text{adj}(\lambda I_n - A)$ is a polynomial (in the variable λ) of degree at most $n-1$.

The Cayley-Hamilton theorem

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

Proof (continued). So, the matrix $\mathrm{adj}(\lambda I_n - A)$ can be expressed in the form

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

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

The Cayley-Hamilton theorem

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

Proof (continued). So, the matrix $\mathrm{adj}(\lambda I_n - A)$ can be expressed in the form

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

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

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

The Cayley-Hamilton theorem

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

Proof (continued). For the left-hand-side, we have

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

The Cayley-Hamilton theorem

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

Proof (continued). 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

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

The Cayley-Hamilton theorem

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

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

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

The Cayley-Hamilton theorem

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

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

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

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. This yields the following (next slide).

The Cayley-Hamilton theorem

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

Proof (continued).

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

The Cayley-Hamilton theorem

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

Proof (continued).

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

We now add up the equations that we obtained. On the left-hand-side, the sum is obviously $O_{n\times n}$.

The Cayley-Hamilton theorem

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

Proof (continued).

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

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.

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

Q.E.D.

The Cayley-Hamilton theorem

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

The Cayley-Hamilton theorem

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

Corollary 1.2

For all matrices $A \in \mathbb{F}^{n\times n}$, both the following hold:

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

Corollary 1.2

For all matrices $A \in \mathbb{F}^{n\times n}$, both the following hold:

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

(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 the matrices $I_n, A, A^2, \ldots, A^{n-1}$.

Proof.

Corollary 1.2

For all matrices $A \in \mathbb{F}^{n \times n}$, both the following hold:

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

(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 the matrices $I_n, A, A^2, \dots, A^{n-1}$.

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

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

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

Consequently,

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

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

Corollary 1.2

For all matrices $A \in \mathbb{F}^{n \times n}$, both the following hold:

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

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

Corollary 1.2

For all matrices $A \in \mathbb{F}^{n \times n}$, both the following hold:

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

(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 the matrices $I_n, A, A^2, \dots, A^{n-1}$.

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

(b) Since A is invertible, we know that 0 is not an eigenvalue of A (by Problem 3 of HW#5). Since the eigenvalues of A are precisely the roots of the characteristic polynomial of A, $a_0 = p_A(0) \neq 0$.

Corollary 1.2

For all matrices $A \in \mathbb{F}^{n\times n}$, both the following hold:

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

(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 the matrices $I_n, A, A^2, \dots, A^{n-1}$.

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

(b) Since A is invertible, we know that 0 is not an eigenvalue of A (by Problem 3 of HW#5). Since the eigenvalues of A are precisely the roots of the characteristic polynomial of A, $a_0 = p_A(0) \neq 0$.

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

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

We multiply both sides of the equation by A^{-1} on the right, and we obtain $A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_2A + a_1I_n + a_0A^{-1} = O_{n\times n}$.

Corollary 1.2

For all matrices $A \in \mathbb{F}^{n \times n}$, both the following hold:

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

(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 the matrices $I_n, A, A^2, \dots, A^{n-1}$.

Proof (continued). Reminder:
$A^{n-1} + a_{n-1}A^{n-2} + \dots + a_2 A + a_1 I_n + a_0 A^{-1} = O_{n \times n}$.

Corollary 1.2

For all matrices $A \in \mathbb{F}^{n \times n}$, both the following hold:

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

Proof (continued). Reminder:
$A^{n-1} + a_{n-1}A^{n-2} + \dots + a_2A + a_1I_n + a_0A^{-1} = O_{n \times n}$.

Consequently,

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

Corollary 1.2

For all matrices $A \in \mathbb{F}^{n\times n}$, both the following hold:

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

(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 the matrices $I_n, A, A^2, \dots, A^{n-1}$.

Proof (continued). Reminder:
$A^{n-1} + a_{n-1}A^{n-2} + \dots + a_2 A + a_1 I_n + a_0 A^{-1} = O_{n\times n}$.

Consequently,

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

Since $a_0 \neq 0$, this implies that

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

Corollary 1.2

For all matrices $A \in \mathbb{F}^{n \times n}$, both the following hold:

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

(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 the matrices $I_n, A, A^2, \dots, A^{n-1}$.

Proof (continued). Reminder:
$A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_2 A + a_1 I_n + a_0 A^{-1} = O_{n \times n}$.

Consequently,

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

Since $a_0 \neq 0$, this implies that

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

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

2. Eigenvalues and eigenvectors of similar matrices

2. Eigenvalues and eigenvectors of similar matrices

- Matrices $A, B \in \mathbb{F}^{n \times n}$ are *similar* if there exists an invertible matrix $P \in \mathbb{F}^{n \times n}$ such that $B = P^{-1}AP$.

2. Eigenvalues and eigenvectors of similar matrices

- Matrices $A, B \in \mathbb{F}^{n \times n}$ are *similar* if there exists an invertible matrix $P \in \mathbb{F}^{n \times n}$ such that $B = P^{-1}AP$.
- Obviously, if $B = P^{-1}AP$, then $A = PBP^{-1} = (P^{-1})^{-1}AP^{-1}$, and so matrix similarity is a symmetric relation on $\mathbb{F}^{n \times n}$.

2. Eigenvalues and eigenvectors of similar matrices

- Matrices $A, B \in \mathbb{F}^{n\times n}$ are *similar* if there exists an invertible matrix $P \in \mathbb{F}^{n\times n}$ such that $B = P^{-1}AP$.
- Obviously, if $B = P^{-1}AP$, then $A = PBP^{-1} = (P^{-1})^{-1}AP^{-1}$, and so matrix similarity is a symmetric relation on $\mathbb{F}^{n\times n}$.
- It is also easy to check that it is reflexive and transitive, and it follows that matrix similarity is an equivalence relation on $\mathbb{F}^{n\times n}$.

2. Eigenvalues and eigenvectors of similar matrices

- Matrices $A, B \in \mathbb{F}^{n\times n}$ are *similar* if there exists an invertible matrix $P \in \mathbb{F}^{n\times n}$ such that $B = P^{-1}AP$.
- Obviously, if $B = P^{-1}AP$, then $A = PBP^{-1} = (P^{-1})^{-1}AP^{-1}$, and so matrix similarity is a symmetric relation on $\mathbb{F}^{n\times n}$.
- It is also easy to check that it is reflexive and transitive, and it follows that matrix similarity is an equivalence relation on $\mathbb{F}^{n\times n}$.
- By Theorem 4.3 of Lecture Notes 10, two matrices in $\mathbb{F}^{n\times n}$ are similar if and only if they represent the same linear transformation from an n-dimensional vector space V over $\mathbb{F}$ to itself, but possibly with respect to different bases of V.

Proposition 2.1

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

Proof.

Proposition 2.1

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

Proof. Since A and B are similar, there exists an invertible matrix $P \in \mathbb{F}^{n\times n}$ such that $B = P^{-1}AP$. To see that A and B have the same characteristic polynomial, we compute:

$$\begin{aligned} p_B(\lambda) &= \det(\lambda I_n - B) \\ &= \det(\lambda I_n - P^{-1}AP) \\ &= \det(P^{-1}(\lambda I_n - A)P) \\ &= \underbrace{\det(P^{-1})}_{=\frac{1}{\det(P)}} \underbrace{\det(\lambda I_n - A)}_{=p_A(\lambda)} \det(P) \quad = \quad p_A(\lambda). \end{aligned}$$

Proposition 2.1

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

Proof (continued). So, A and B have the same characteristic polynomial, and consequently, the same eigenvalues with the same corresponding algebraic multiplicities.

Proposition 2.1

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

Proof (continued). So, A and B have the same characteristic polynomial, and consequently, the same eigenvalues with the same corresponding algebraic multiplicities.

Now, fix an eigenvalue λ of A and B, and let m_A and m_B be the geometric multiplicities of λ as an eigenvalue of A and B, respectively.

Proposition 2.1

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

Proof (continued). So, A and B have the same characteristic polynomial, and consequently, the same eigenvalues with the same corresponding algebraic multiplicities.

Now, fix an eigenvalue λ of A and B, and let m_A and m_B be the geometric multiplicities of λ as an eigenvalue of A and B, respectively. WTS $m_A = m_B$. Since matrix similarity is symmetric, it suffices to show that $m_A \leq m_B$.

Proposition 2.1

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

Proof (continued). So, A and B have the same characteristic polynomial, and consequently, the same eigenvalues with the same corresponding algebraic multiplicities.

Now, fix an eigenvalue λ of A and B, and let m_A and m_B be the geometric multiplicities of λ as an eigenvalue of A and B, respectively. WTS $m_A = m_B$. Since matrix similarity is symmetric, it suffices to show that $m_A \leq m_B$. Let $\{\mathbf{v}_1, \dots, \mathbf{v}_{m_A}\}$ be a basis for the eigenspace $E_\lambda(A)$.

Proposition 2.1

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

Proof (continued). So, A and B have the same characteristic polynomial, and consequently, the same eigenvalues with the same corresponding algebraic multiplicities.

Now, fix an eigenvalue λ of A and B, and let m_A and m_B be the geometric multiplicities of λ as an eigenvalue of A and B, respectively. WTS $m_A = m_B$. Since matrix similarity is symmetric, it suffices to show that $m_A \leq m_B$. Let $\{\mathbf{v}_1, \dots, \mathbf{v}_{m_A}\}$ be a basis for the eigenspace $E_\lambda(A)$. Then for all $i \in \{1, \dots, m_A\}$, $P^{-1}\mathbf{v}_i$ is an eigenvector of B associated with λ, since (next slide)

Proposition 2.1

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

Proof (continued).

$$\begin{aligned}
B(P^{-1}\mathbf{v}_i) &= (\underbrace{P^{-1}AP}_{=B})(P^{-1}\mathbf{v}_i) \\
&= P^{-1}A(\underbrace{PP^{-1}}_{=I_n})\mathbf{v}_i \\
&= P^{-1}(A\mathbf{v}_i) \\
&= P^{-1}(\lambda\mathbf{v}_i) && \text{because } \mathbf{v}_i \in E_\lambda(A) \\
&= \lambda(P^{-1}\mathbf{v}_i).
\end{aligned}$$

Proposition 2.1

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

Proof (continued). Reminder: $P^{-1}\mathbf{v}_1, \dots, P^{-1}\mathbf{v}_{m_A}$ are eigenvectors of B associated with the eigenvalue λ_i.

Proposition 2.1

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

Proof (continued). Reminder: $P^{-1}\mathbf{v}_1, \ldots, P^{-1}\mathbf{v}_{m_A}$ are eigenvectors of B associated with the eigenvalue λ_i.

WTS $\{P^{-1}\mathbf{v}_1, \ldots, P^{-1}\mathbf{v}_{m_A}\}$ is linearly independent.

Proposition 2.1

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

Proof (continued). Reminder: $P^{-1}\mathbf{v}_1, \ldots, P^{-1}\mathbf{v}_{m_A}$ are eigenvectors of B associated with the eigenvalue λ_i.

WTS $\{P^{-1}\mathbf{v}_1, \ldots, P^{-1}\mathbf{v}_{m_A}\}$ is linearly independent. Fix scalars $\alpha_1, \ldots, \alpha_{m_A} \in \mathbb{F}$ such that $\alpha_1 P^{-1}\mathbf{v}_1 + \cdots + \alpha_{m_A} P^{-1}\mathbf{v}_{m_A} = \mathbf{0}$.

Proposition 2.1

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

Proof (continued). Reminder: $P^{-1}\mathbf{v}_1, \ldots, P^{-1}\mathbf{v}_{m_A}$ are eigenvectors of B associated with the eigenvalue λ_i.

WTS $\{P^{-1}\mathbf{v}_1, \ldots, P^{-1}\mathbf{v}_{m_A}\}$ is linearly independent. Fix scalars $\alpha_1, \ldots, \alpha_{m_A} \in \mathbb{F}$ such that $\alpha_1 P^{-1}\mathbf{v}_1 + \cdots + \alpha_{m_A} P^{-1}\mathbf{v}_{m_A} = \mathbf{0}$. By multiplying both sides on the left by P, we obtain
$\alpha_1 \mathbf{v}_1 + \cdots + \alpha_{m_A}\mathbf{v}_{m_A} = \mathbf{0}$.

Proposition 2.1

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

Proof (continued). Reminder: $P^{-1}\mathbf{v}_1, \dots, P^{-1}\mathbf{v}_{m_A}$ are eigenvectors of B associated with the eigenvalue λ_i.

WTS $\{P^{-1}\mathbf{v}_1, \dots, P^{-1}\mathbf{v}_{m_A}\}$ is linearly independent. Fix scalars $\alpha_1, \dots, \alpha_{m_A} \in \mathbb{F}$ such that $\alpha_1 P^{-1}\mathbf{v}_1 + \dots + \alpha_{m_A} P^{-1}\mathbf{v}_{m_A} = \mathbf{0}$. By multiplying both sides on the left by P, we obtain $\alpha_1 \mathbf{v}_1 + \dots + \alpha_{m_A}\mathbf{v}_{m_A} = \mathbf{0}$. Since $\{\mathbf{v}_1, \dots, \mathbf{v}_{m_A}\}$ is a basis of $E_\lambda(A)$, and is therefore linearly independent, we see that $\alpha_1 = \dots = \alpha_{m_A} = 0$.

Proposition 2.1

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

Proof (continued). Reminder: $P^{-1}\mathbf{v}_1, \dots, P^{-1}\mathbf{v}_{m_A}$ are eigenvectors of B associated with the eigenvalue λ_i.

WTS $\{P^{-1}\mathbf{v}_1, \dots, P^{-1}\mathbf{v}_{m_A}\}$ is linearly independent. Fix scalars $\alpha_1, \dots, \alpha_{m_A} \in \mathbb{F}$ such that $\alpha_1 P^{-1}\mathbf{v}_1 + \dots + \alpha_{m_A} P^{-1}\mathbf{v}_{m_A} = \mathbf{0}$. By multiplying both sides on the left by P, we obtain $\alpha_1\mathbf{v}_1 + \dots + \alpha_{m_A}\mathbf{v}_{m_A} = \mathbf{0}$. Since $\{\mathbf{v}_1, \dots, \mathbf{v}_{m_A}\}$ is a basis of $E_\lambda(A)$, and is therefore linearly independent, we see that $\alpha_1 = \dots = \alpha_{m_A} = 0$. So, $\{P^{-1}\mathbf{v}_1, \dots, P^{-1}\mathbf{v}_{m_A}\}$ is linearly independent.

Proposition 2.1

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

Proof (continued). Reminder: $P^{-1}\mathbf{v}_1, \dots, P^{-1}\mathbf{v}_{m_A}$ are eigenvectors of B associated with the eigenvalue λ_i.

WTS $\{P^{-1}\mathbf{v}_1, \dots, P^{-1}\mathbf{v}_{m_A}\}$ is linearly independent. Fix scalars $\alpha_1, \dots, \alpha_{m_A} \in \mathbb{F}$ such that $\alpha_1 P^{-1}\mathbf{v}_1 + \dots + \alpha_{m_A} P^{-1}\mathbf{v}_{m_A} = \mathbf{0}$. By multiplying both sides on the left by P, we obtain $\alpha_1\mathbf{v}_1 + \dots + \alpha_{m_A}\mathbf{v}_{m_A} = \mathbf{0}$. Since $\{\mathbf{v}_1, \dots, \mathbf{v}_{m_A}\}$ is a basis of $E_\lambda(A)$, and is therefore linearly independent, we see that $\alpha_1 = \dots = \alpha_{m_A} = 0$. So, $\{P^{-1}\mathbf{v}_1, \dots, P^{-1}\mathbf{v}_{m_A}\}$ is linearly independent. Since $\{P^{-1}\mathbf{v}_1, \dots, P^{-1}\mathbf{v}_{m_A}\} \subseteq E_\lambda(B)$, it follows that $m_A \leq \dim(E_\lambda(B)) = m_B$. Q.E.D.

Proposition 2.2

Let $B, C \in \mathbb{F}^{n \times n}$ be similar matrices, with $C = P^{-1}BP$ for some invertible matrix $P \in \mathbb{F}^{n \times n}$. Then for all non-negative integers m, we have that $C^m = P^{-1}B^mP$ and $B^m = PC^mP^{-1}$.

Proof.

Proposition 2.2

Let $B, C \in \mathbb{F}^{n \times n}$ be similar matrices, with $C = P^{-1}BP$ for some invertible matrix $P \in \mathbb{F}^{n \times n}$. Then for all non-negative integers m, we have that $C^m = P^{-1}B^mP$ and $B^m = PC^mP^{-1}$.

Proof. Obviously, each of the two equalities implies the other, and so it is enough to prove one of them.

Proposition 2.2

Let $B, C \in \mathbb{F}^{n\times n}$ be similar matrices, with $C = P^{-1}BP$ for some invertible matrix $P \in \mathbb{F}^{n\times n}$. Then for all non-negative integers m, we have that $C^m = P^{-1}B^mP$ and $B^m = PC^mP^{-1}$.

Proof. Obviously, each of the two equalities implies the other, and so it is enough to prove one of them.

Let us prove the first one.

Proposition 2.2

Let $B, C \in \mathbb{F}^{n \times n}$ be similar matrices, with $C = P^{-1}BP$ for some invertible matrix $P \in \mathbb{F}^{n \times n}$. Then for all non-negative integers m, we have that $C^m = P^{-1}B^mP$ and $B^m = PC^mP^{-1}$.

Proof. Obviously, each of the two equalities implies the other, and so it is enough to prove one of them.

Let us prove the first one. We proceed by induction on m.

Proposition 2.2

Let $B, C \in \mathbb{F}^{n \times n}$ be similar matrices, with $C = P^{-1}BP$ for some invertible matrix $P \in \mathbb{F}^{n \times n}$. Then for all non-negative integers m, we have that $C^m = P^{-1}B^mP$ and $B^m = PC^mP^{-1}$.

Proof. Obviously, each of the two equalities implies the other, and so it is enough to prove one of them.

Let us prove the first one. We proceed by induction on m. First, we have that $C^0 = I_n$ and that $P^{-1}B^0P = P^{-1}I_nP = I_n$, and so $C^0 = P^{-1}B^0P$.

Proposition 2.2

Let $B, C \in \mathbb{F}^{n\times n}$ be similar matrices, with $C = P^{-1}BP$ for some invertible matrix $P \in \mathbb{F}^{n\times n}$. Then for all non-negative integers m, we have that $C^m = P^{-1}B^mP$ and $B^m = PC^mP^{-1}$.

Proof. Obviously, each of the two equalities implies the other, and so it is enough to prove one of them.

Let us prove the first one. We proceed by induction on m. First, we have that $C^0 = I_n$ and that $P^{-1}B^0P = P^{-1}I_nP = I_n$, and so $C^0 = P^{-1}B^0P$. Thus, the claim holds for $m = 0$.

Proposition 2.2

Let $B, C \in \mathbb{F}^{n \times n}$ be similar matrices, with $C = P^{-1}BP$ for some invertible matrix $P \in \mathbb{F}^{n \times n}$. Then for all non-negative integers m, we have that $C^m = P^{-1}B^mP$ and $B^m = PC^mP^{-1}$.

Proof (continued). Now, fix a non-negative integer m, and assume inductively that $C^m = P^{-1}B^mP$.

Proposition 2.2

Let $B, C \in \mathbb{F}^{n \times n}$ be similar matrices, with $C = P^{-1}BP$ for some invertible matrix $P \in \mathbb{F}^{n \times n}$. Then for all non-negative integers m, we have that $C^m = P^{-1}B^mP$ and $B^m = PC^mP^{-1}$.

Proof (continued). Now, fix a non-negative integer m, and assume inductively that $C^m = P^{-1}B^mP$. Then

$$\begin{aligned} C^{m+1} &= C^mC \\ &= (\underbrace{P^{-1}B^mP}_{=C^m})(\underbrace{P^{-1}BP}_{=C}) \qquad \text{by the induction hypothesis} \\ &= P^{-1}B^m(\underbrace{PP^{-1}}_{=I_n})BP \\ &= P^{-1}B^{m+1}P. \end{aligned}$$

This completes the induction. Q.E.D.

3. Algebraic and geometric multiplicities revisited

Theorem 2.5 of Lecture Notes 18

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

Proof.

3. Algebraic and geometric multiplicities revisited

Theorem 2.5 of Lecture Notes 18

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

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

3. Algebraic and geometric multiplicities revisited

Theorem 2.5 of Lecture Notes 18

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

Proof. Suppose that λ_0 is an eigenvalue of A of geometric multiplicity k. WTS the eigenvalue λ_0 has algebraic multiplicity at least k.

3. Algebraic and geometric multiplicities revisited

Theorem 2.5 of Lecture Notes 18

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

Proof. Suppose that λ_0 is an eigenvalue of A of geometric multiplicity k. WTS the eigenvalue λ_0 has algebraic multiplicity at least k. In view of Proposition 2.1, it is enough to exhibit a matrix $B \in \mathbb{F}^{n\times n}$ similar to A, and such that λ_0 is an eigenvalue of B of algebraic multiplicity at least k.

3. Algebraic and geometric multiplicities revisited

Theorem 2.5 of Lecture Notes 18

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

Proof. Suppose that λ_0 is an eigenvalue of A of geometric multiplicity k. WTS the eigenvalue λ_0 has algebraic multiplicity at least k. In view of Proposition 2.1, it is enough to exhibit a matrix $B \in \mathbb{F}^{n \times n}$ similar to A, and such that λ_0 is an eigenvalue of B of algebraic multiplicity at least k.

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

3. Algebraic and geometric multiplicities revisited

Theorem 2.5 of Lecture Notes 18

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

Proof. Suppose that λ_0 is an eigenvalue of A of geometric multiplicity k. WTS the eigenvalue λ_0 has algebraic multiplicity at least k. In view of Proposition 2.1, it is enough to exhibit a matrix $B \in \mathbb{F}^{n \times n}$ similar to A, and such that λ_0 is an eigenvalue of B of algebraic multiplicity at least k.

Since the geometric multiplicity of the eigenvalue λ_0 of A is k, we see that the eigenspace E_{λ_0} has a k-element basis, say $\{\mathbf{p}_1, \dots, \mathbf{p}_k\}$. We now extend $\{\mathbf{p}_1, \dots, \mathbf{p}_k\}$ to a basis $\mathcal{P} = \{\mathbf{p}_1, \dots, \mathbf{p}_k, \mathbf{p}_{k+1}, \dots, \mathbf{p}_n\}$ of $\mathbb{F}^n$.

Theorem 2.5 of Lecture Notes 18

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

Proof (continued). Reminder: $\{\mathbf{p}_1, \dots, \mathbf{p}_k\}$ is a basis of E_{λ_0}; $\mathcal{P} = \{\mathbf{p}_1, \dots, \mathbf{p}_k, \mathbf{p}_{k+1}, \dots, \mathbf{p}_n\}$ of $\mathbb{F}^n$.

Theorem 2.5 of Lecture Notes 18

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

Proof (continued). Reminder: $\{\mathbf{p}_1, \dots, \mathbf{p}_k\}$ is a basis of E_{λ_0}; $\mathcal{P} = \{\mathbf{p}_1, \dots, \mathbf{p}_k, \mathbf{p}_{k+1}, \dots, \mathbf{p}_n\}$ of $\mathbb{F}^n$.

Let $f_A : \mathbb{F}^n \to \mathbb{F}^n$ be the linear transformation whose standard matrix is A (i.e. $f_A(\mathbf{v}) = A\mathbf{v}$ for all $\mathbf{v} \in \mathbb{F}^n$), so that $A = {}_{\mathcal{E}_n}[f_A]_{\mathcal{E}_n}$, where $\mathcal{E}_n = \{\mathbf{e}_1^n, \dots, \mathbf{e}_n^n\}$ is the standard basis of $\mathbb{F}^n$. Consider the matrix $B := {}_{\mathcal{P}}[f_A]_{\mathcal{P}}$. Then A and B are similar, since (next slide):

Theorem 2.5 of Lecture Notes 18

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

Proof (continued). Reminder: $\{\mathbf{p}_1, \dots, \mathbf{p}_k\}$ is a basis of E_{λ_0}; $\mathcal{P} = \{\mathbf{p}_1, \dots, \mathbf{p}_k, \mathbf{p}_{k+1}, \dots, \mathbf{p}_n\}$ of $\mathbb{F}^n$.

$$\begin{aligned} A &= {}_{\mathcal{E}_n}[f_A]_{\mathcal{E}_n} \\ &= {}_{\mathcal{E}_n}[Id_{\mathbb{F}^n} \circ f_A \circ Id_{\mathbb{F}^n}]_{\mathcal{E}_n} \\ &= {}_{\mathcal{E}_n}[Id_{\mathbb{F}^n}]_{\mathcal{P}} \; {}_{\mathcal{P}}[f_A]_{\mathcal{P}} \; {}_{\mathcal{P}}[Id_{\mathbb{F}^n}]_{\mathcal{E}_n} \\ &= ({}_{\mathcal{P}}[Id_{\mathbb{F}^n}]_{\mathcal{E}_n})^{-1} \; \underbrace{{}_{\mathcal{P}}[f_A]_{\mathcal{P}}}_{=B} \; {}_{\mathcal{P}}[Id_{\mathbb{F}^n}]_{\mathcal{E}_n}, \end{aligned}$$

Theorem 2.5 of Lecture Notes 18

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

Proof (continued). Reminder: $\{\mathbf{p}_1, \dots, \mathbf{p}_k\}$ is a basis of E_{λ_0}; $\mathcal{P} = \{\mathbf{p}_1, \dots, \mathbf{p}_k, \mathbf{p}_{k+1}, \dots, \mathbf{p}_n\}$ of $\mathbb{F}^n$.

$$\begin{aligned}
A &= {}_{\mathcal{E}_n}[f_A]_{\mathcal{E}_n} \\
&= {}_{\mathcal{E}_n}[Id_{\mathbb{F}^n} \circ f_A \circ Id_{\mathbb{F}^n}]_{\mathcal{E}_n} \\
&= {}_{\mathcal{E}_n}[Id_{\mathbb{F}^n}]_{\mathcal{P}} \ {}_{\mathcal{P}}[f_A]_{\mathcal{P}} \ {}_{\mathcal{P}}[Id_{\mathbb{F}^n}]_{\mathcal{E}_n} \\
&= ({}_{\mathcal{P}}[Id_{\mathbb{F}^n}]_{\mathcal{E}_n})^{-1} \ \underbrace{{}_{\mathcal{P}}[f_A]_{\mathcal{P}}}_{=B} \ {}_{\mathcal{P}}[Id_{\mathbb{F}^n}]_{\mathcal{E}_n},
\end{aligned}$$

(Alternatively, the similarity of A and B follows from Theorem 4.3 of Lecture Notes 10.)

Proof (continued). It now remains to show that λ_0 is an eigenvalue of B of algebraic multiplicity at least k.

Proof (continued). It now remains to show that λ_0 is an eigenvalue of B of algebraic multiplicity at least k. For this, we observe that

$$\begin{aligned}
B &= {}_{\mathcal{P}}[f_A]_{\mathcal{P}} \\
&= \begin{bmatrix} [f_A(\mathbf{p}_1)]_{\mathcal{P}} & \dots & [f_A(\mathbf{p}_k)]_{\mathcal{P}} & [f_A(\mathbf{p}_{k+1})]_{\mathcal{P}} & \dots & [f_A(\mathbf{p}_n)]_{\mathcal{P}} \end{bmatrix} \\
&= \begin{bmatrix} [A\mathbf{p}_1]_{\mathcal{P}} & \dots & [A\mathbf{p}_k]_{\mathcal{P}} & [f_A(\mathbf{p}_{k+1})]_{\mathcal{P}} & \dots & [f_A(\mathbf{p}_n)]_{\mathcal{P}} \end{bmatrix} \\
&= \begin{bmatrix} [\lambda_0\mathbf{p}_1]_{\mathcal{P}} & \dots & [\lambda_0\mathbf{p}_k]_{\mathcal{P}} & [f_A(\mathbf{p}_{k+1})]_{\mathcal{P}} & \dots & [f_A(\mathbf{p}_n)]_{\mathcal{P}} \end{bmatrix} \\
&= \begin{bmatrix} \lambda_0\mathbf{e}_1^n & \dots & \lambda_0\mathbf{e}_k^n & [f_A(\mathbf{p}_{k+1})]_{\mathcal{P}} & \dots & [f_A(\mathbf{p}_n)]_{\mathcal{P}} \end{bmatrix} \\
&= \left[\begin{array}{c|ccc} \lambda_0 I_k & & & \\ \hline O_{(n-k)\times k} & [f_A(\mathbf{p}_{k+1})]_{\mathcal{P}} & \dots & [f_A(\mathbf{p}_n)]_{\mathcal{P}} \end{array} \right].
\end{aligned}$$

Proof (continued). We now have that

$$p_B(\lambda) \;=\; \det(\lambda I_n - B) \;=\; \left| \begin{array}{c|c} (\lambda - \lambda_0) I_k & \\ \hline O_{(n-k)\times k} & C \end{array} \right|,$$

where

$$C \;=\; \left[\begin{array}{ccc} \lambda \mathbf{e}_{k+1} - [f_A(\mathbf{p}_{k+1})]_{\mathcal{P}} & \ldots & \lambda \mathbf{e}_n - [f_A(\mathbf{p}_{k+1})]_{\mathcal{P}} \end{array} \right]_{(n-k)\times n}.$$

Proof (continued). We now have that

$$p_B(\lambda) = \det(\lambda I_n - B) = \left| \begin{array}{c|c} (\lambda - \lambda_0) I_k & C \\ \hline O_{(n-k)\times k} & \end{array} \right|,$$

where

$$C = \begin{bmatrix} \lambda \mathbf{e}_{k+1} - [f_A(\mathbf{p}_{k+1})]_{\mathcal{P}} & \ldots & \lambda \mathbf{e}_n - [f_A(\mathbf{p}_{k+1})]_{\mathcal{P}} \end{bmatrix}_{(n-k)\times n}.$$

Thus, $p_B(\lambda)$ has the form

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

where the red matrix in the upper-left corner is of size $k \times k$.

Proof (continued). Reminder:

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

where the red matrix in the upper-left corner is of size $k \times k$.

Proof (continued). Reminder:

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

where the red matrix in the upper-left corner is of size $k \times k$.

By iteratively performing Laplace expansion along the first column, we see that $p_B(\lambda)$ has a factor $(\lambda - \lambda_0)^k$, and consequently, λ_0 is an eigenvalue of B of algebraic multiplicity at least k. Q.E.D.

Theorem 2.5 of Lecture Notes 18

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

4. Diagonal matrices and diagonalization

4. Diagonal matrices and diagonalization

- A matrix $A = [\ a_{i,j}\ ]_{n\times n}$ in $\mathbb{F}^{n\times n}$ is *diagonal* if $a_{i,j} = 0$ for all distinct $i, j \in \{1, \ldots, n\}$.

4. Diagonal matrices and diagonalization

- A matrix $A = [\ a_{i,j}\ ]_{n \times n}$ in $\mathbb{F}^{n \times n}$ is *diagonal* if $a_{i,j} = 0$ for all distinct $i, j \in \{1, \ldots, n\}$.
    - In other words, a square matrix is diagonal if all its entries off the main diagonal are zero (and the main diagonal is arbitrary).

4. Diagonal matrices and diagonalization

- A matrix $A = [\ a_{i,j}\ ]_{n\times n}$ in $\mathbb{F}^{n\times n}$ is *diagonal* if $a_{i,j} = 0$ for all distinct $i, j \in \{1, \dots, n\}$.
    - In other words, a square matrix is diagonal if all its entries off the main diagonal are zero (and the main diagonal is arbitrary).
- Given scalars $\lambda_1, \dots, \lambda_n \in \mathbb{F}$, we will denote by $D(\lambda_1, \dots, \lambda_n)$ the diagonal matrix in which $\lambda_1, \dots, \lambda_n$ appear on the main diagonal (in that order), i.e.

$$
\begin{aligned}
D(\lambda_1, \lambda_2, \dots, \lambda_n) \quad &:= \quad \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} \\
&= \quad \begin{bmatrix} \lambda_1\mathbf{e}_1 & \lambda_2\mathbf{e}_2 & \dots & \lambda_n\mathbf{e}_n \end{bmatrix}.
\end{aligned}
$$

Porposition 4.1

Let $\lambda_1, \dots, \lambda_n \in \mathbb{F}^n$, and set $D := D(\lambda_1, \dots, \lambda_n)$. Then both the following hold:

(a) For all matrices $P = \left[\begin{array}{ccc} \mathbf{p}_1 & \dots \mathbf{p}_n \end{array} \right]$ in $\mathbb{F}^{n \times n}$, we have that $PD = \left[\begin{array}{ccc} \lambda_1 \mathbf{p}_1 & \dots \lambda_n \mathbf{p}_n \end{array} \right]$;

(b) For all non-negative integers m, we have that $D^m = D(\lambda_1^m, \dots, \lambda_n^m)$.

Proof.

Porposition 4.1

Let $\lambda_1, \dots, \lambda_n \in \mathbb{F}^n$, and set $D := D(\lambda_1, \dots, \lambda_n)$. Then both the following hold:

- (a) For all matrices $P = \begin{bmatrix} \mathbf{p}_1 & \dots \mathbf{p}_n \end{bmatrix}$ in $\mathbb{F}^{n \times n}$, we have that $PD = \begin{bmatrix} \lambda_1 \mathbf{p}_1 & \dots \lambda_n \mathbf{p}_n \end{bmatrix}$;
- (b) For all non-negative integers m, we have that $D^m = D(\lambda_1^m, \dots, \lambda_n^m)$.

Proof. (a) Fix a matrix $P = \begin{bmatrix} \mathbf{p}_1 & \dots & \mathbf{p}_n \end{bmatrix}$ in $\mathbb{F}^{n \times n}$. Then

$$\begin{aligned} PD &= P \begin{bmatrix} \lambda_1 \mathbf{e}_1 & \dots & \lambda_n \mathbf{e}_n \end{bmatrix} \\ &= \begin{bmatrix} P(\lambda_1 \mathbf{e}_1) & \dots & P(\lambda_n \mathbf{e}_n) \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 (P\mathbf{e}_1) & \dots & \lambda_n (P\mathbf{e}_n) \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 \mathbf{p}_1 & \dots & \lambda_n \mathbf{p}_n \end{bmatrix}. \end{aligned}$$

Porposition 4.1

Let $\lambda_1, \dots, \lambda_n \in \mathbb{F}^n$, and set $D := D(\lambda_1, \dots, \lambda_n)$. Then both the following hold:

(a) For all matrices $P = \left[\begin{array}{ccc} \mathbf{p}_1 & \dots & \mathbf{p}_n \end{array} \right]$ in $\mathbb{F}^{n \times n}$, we have that $PD = \left[\begin{array}{ccc} \lambda_1 \mathbf{p}_1 & \dots & \lambda_n \mathbf{p}_n \end{array} \right]$;

(b) For all non-negative integers m, we have that $D^m = D(\lambda_1^m, \dots, \lambda_n^m)$.

Proof (continued). (b) readily follows from (a) via an easy induction on m. (Details: Lecture Notes.)

Definition

A matrix $A \in \mathbb{F}^{n \times n}$ is *diagonalizable* if it is similar to some diagonal matrix in $\mathbb{F}^{n \times n}$.

Definition

A matrix $A \in \mathbb{F}^{n\times n}$ is *diagonalizable* if it is similar to some diagonal matrix in $\mathbb{F}^{n\times n}$.

- To *diagonalize* a matrix $A \in \mathbb{F}^{n\times n}$ means to find a diagonal matrix D and an invertible matrix P, both in $\mathbb{F}^{n\times n}$, such that $D = P^{-1}AP$.

Definition

A matrix $A \in \mathbb{F}^{n \times n}$ is *diagonalizable* if it is similar to some diagonal matrix in $\mathbb{F}^{n \times n}$.

- To *diagonalize* a matrix $A \in \mathbb{F}^{n \times n}$ means to find a diagonal matrix D and an invertible matrix P, both in $\mathbb{F}^{n \times n}$, such that $D = P^{-1}AP$.

Theorem 4.2

A matrix $A \in \mathbb{F}^{n \times n}$ is diagonalizable if and only if $\mathbb{F}^n$ has a basis formed by eigenvectors of A.

Definition

A matrix $A \in \mathbb{F}^{n \times n}$ is *diagonalizable* if it is similar to some diagonal matrix in $\mathbb{F}^{n \times n}$.

- To *diagonalize* a matrix $A \in \mathbb{F}^{n \times n}$ means to find a diagonal matrix D and an invertible matrix P, both in $\mathbb{F}^{n \times n}$, such that $D = P^{-1}AP$.

Theorem 4.2

A matrix $A \in \mathbb{F}^{n \times n}$ is diagonalizable if and only if $\mathbb{F}^n$ has a basis formed by eigenvectors of A.

- We will prove Theorem 2.4.

Definition

A matrix $A \in \mathbb{F}^{n \times n}$ is *diagonalizable* if it is similar to some diagonal matrix in $\mathbb{F}^{n \times n}$.

- To *diagonalize* a matrix $A \in \mathbb{F}^{n \times n}$ means to find a diagonal matrix D and an invertible matrix P, both in $\mathbb{F}^{n \times n}$, such that $D = P^{-1}AP$.

Theorem 4.2

A matrix $A \in \mathbb{F}^{n \times n}$ is diagonalizable if and only if $\mathbb{F}^n$ has a basis formed by eigenvectors of A.

- We will prove Theorem 2.4.
- First, however, it may be useful to recall when $\mathbb{F}^n$ has a basis formed by eigenvectors of $A^{n \times n}$.

Definition

A matrix $A \in \mathbb{F}^{n\times n}$ is *diagonalizable* if it is similar to some diagonal matrix in $\mathbb{F}^{n\times n}$.

- To *diagonalize* a matrix $A \in \mathbb{F}^{n\times n}$ means to find a diagonal matrix D and an invertible matrix P, both in $\mathbb{F}^{n\times n}$, such that $D = P^{-1}AP$.

Theorem 4.2

A matrix $A \in \mathbb{F}^{n\times n}$ is diagonalizable if and only if $\mathbb{F}^n$ has a basis formed by eigenvectors of A.

- We will prove Theorem 2.4.
- First, however, it may be useful to recall when $\mathbb{F}^n$ has a basis formed by eigenvectors of $A^{n\times n}$.
- This is precisely described in Corollary 2.13 of Lecture Notes 18.

Corollary 2.13 of Lecture Notes 18

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$. Then the following are equivalent:

(i) $\mathbb{F}^n$ has a basis formed by eigenvectors of A;

(ii) the sum of algebraic multiplicities of all distinct eigenvalues A is equal to n, and the geometric multiplicity of each eigenvalue of A is equal to the algebraic multiplicity of that eigenvalue;

(iii) the sum of geometric multiplicities of all distinct eigenvalues A is equal to n.

Corollary 2.13 of Lecture Notes 18

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$. Then the following are equivalent:

1. (i) $\mathbb{F}^n$ has a basis formed by eigenvectors of A;
2. (ii) the sum of algebraic multiplicities of all distinct eigenvalues A is equal to n, and the geometric multiplicity of each eigenvalue of A is equal to the algebraic multiplicity of that eigenvalue;
3. (iii) the sum of geometric multiplicities of all distinct eigenvalues A is equal to n.

Theorem 4.2

A matrix $A \in \mathbb{F}^{n \times n}$ is diagonalizable if and only if $\mathbb{F}^n$ has a basis formed by eigenvectors of A.

Corollary 2.13 of Lecture Notes 18

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$. Then the following are equivalent:

1. $\mathbb{F}^n$ has a basis formed by eigenvectors of A;
2. the sum of algebraic multiplicities of all distinct eigenvalues A is equal to n, and the geometric multiplicity of each eigenvalue of A is equal to the algebraic multiplicity of that eigenvalue;
3. the sum of geometric multiplicities of all distinct eigenvalues A is equal to n.

Theorem 4.2

A matrix $A \in \mathbb{F}^{n \times n}$ is diagonalizable if and only if $\mathbb{F}^n$ has a basis formed by eigenvectors of A.

- Let's prove Theorem 4.2!

Theorem 4.2

A matrix $A \in \mathbb{F}^{n \times n}$ is diagonalizable if and only if $\mathbb{F}^n$ has a basis formed by eigenvectors of A.

Proof.

Theorem 4.2

A matrix $A \in \mathbb{F}^{n\times n}$ is diagonalizable if and only if $\mathbb{F}^n$ has a basis formed by eigenvectors of A.

Proof. Fix a matrix $A \in \mathbb{F}^{n\times n}$.

Theorem 4.2

A matrix $A \in \mathbb{F}^{n \times n}$ is diagonalizable if and only if $\mathbb{F}^n$ has a basis formed by eigenvectors of A.

Proof. Fix a matrix $A \in \mathbb{F}^{n \times n}$. Suppose first that A is diagonalizable, and fix matrices $P, D \in \mathbb{F}^{n \times n}$ such that D is diagonal, P is invertible, and $D = P^{-1}AP$.

Theorem 4.2

A matrix $A \in \mathbb{F}^{n \times n}$ is diagonalizable if and only if $\mathbb{F}^n$ has a basis formed by eigenvectors of A.

Proof. Fix a matrix $A \in \mathbb{F}^{n \times n}$. Suppose first that A is diagonalizable, and fix matrices $P, D \in \mathbb{F}^{n \times n}$ such that D is diagonal, P is invertible, and $D = P^{-1}AP$. Since P is invertible, its columns form a basis of $\mathbb{F}^{n \times n}$.

Theorem 4.2

A matrix $A \in \mathbb{F}^{n\times n}$ is diagonalizable if and only if $\mathbb{F}^n$ has a basis formed by eigenvectors of A.

Proof. Fix a matrix $A \in \mathbb{F}^{n\times n}$. Suppose first that A is diagonalizable, and fix matrices $P, D \in \mathbb{F}^{n\times n}$ such that D is diagonal, P is invertible, and $D = P^{-1}AP$. Since P is invertible, its columns form a basis of $\mathbb{F}^{n\times n}$. So, it is enough to show that every column of P is an eigenvector of A. Set $P = [\ \mathbf{p}_1 \quad \dots \quad \mathbf{p}_n\ ]$ and $D = D(\lambda_1, \dots, \lambda_n)$.

Theorem 4.2

A matrix $A \in \mathbb{F}^{n\times n}$ is diagonalizable if and only if $\mathbb{F}^n$ has a basis formed by eigenvectors of A.

Proof. Fix a matrix $A \in \mathbb{F}^{n\times n}$. Suppose first that A is diagonalizable, and fix matrices $P, D \in \mathbb{F}^{n\times n}$ such that D is diagonal, P is invertible, and $D = P^{-1}AP$. Since P is invertible, its columns form a basis of $\mathbb{F}^{n\times n}$. So, it is enough to show that every column of P is an eigenvector of A. Set $P = \begin{bmatrix} \mathbf{p}_1 & \dots & \mathbf{p}_n \end{bmatrix}$ and $D = D(\lambda_1, \dots, \lambda_n)$. Since $D = P^{-1}AP$, we have that $PD = AP$. For each $i \in \{1, \dots, n\}$, the i-th column of PD is $\lambda_i \mathbf{p}_i$ (by Proposition 4.1(a)), and the i-th column of AP is $A\mathbf{p}_i$, and consequently, $A\mathbf{p}_i = \lambda_i \mathbf{p}_i$, i.e. λ_i is an eigenvalue of A, and $\mathbf{p}_i$ is an eigenvector of A associated with λ_i.

Theorem 4.2

A matrix $A \in \mathbb{F}^{n\times n}$ is diagonalizable if and only if $\mathbb{F}^n$ has a basis formed by eigenvectors of A.

Proof. Fix a matrix $A \in \mathbb{F}^{n\times n}$. Suppose first that A is diagonalizable, and fix matrices $P, D \in \mathbb{F}^{n\times n}$ such that D is diagonal, P is invertible, and $D = P^{-1}AP$. Since P is invertible, its columns form a basis of $\mathbb{F}^{n\times n}$. So, it is enough to show that every column of P is an eigenvector of A. Set $P = [\ \mathbf{p}_1 \quad \dots \quad \mathbf{p}_n\ ]$ and $D = D(\lambda_1, \dots, \lambda_n)$. Since $D = P^{-1}AP$, we have that $PD = AP$. For each $i \in \{1, \dots, n\}$, the i-th column of PD is $\lambda_i \mathbf{p}_i$ (by Proposition 4.1(a)), and the i-th column of AP is $A\mathbf{p}_i$, and consequently, $A\mathbf{p}_i = \lambda_i \mathbf{p}_i$, i.e. λ_i is an eigenvalue of A, and $\mathbf{p}_i$ is an eigenvector of A associated with λ_i. So, the columns of P are eigenvectors of A.

Theorem 4.2

A matrix $A \in \mathbb{F}^{n \times n}$ is diagonalizable if and only if $\mathbb{F}^n$ has a basis formed by eigenvectors of A.

Proof (continued). Suppose now that $\mathbb{F}^n$ has a basis, say, $\mathcal{P} = \{\mathbf{p}_1, \ldots, \mathbf{p}_n\}$, formed by eigenvectors of A.

Theorem 4.2

A matrix $A \in \mathbb{F}^{n \times n}$ is diagonalizable if and only if $\mathbb{F}^n$ has a basis formed by eigenvectors of A.

Proof (continued). Suppose now that $\mathbb{F}^n$ has a basis, say, $\mathcal{P} = \{\mathbf{p}_1, \ldots, \mathbf{p}_n\}$, formed by eigenvectors of A. WTS A is diagonalizable.

Theorem 4.2

A matrix $A \in \mathbb{F}^{n \times n}$ is diagonalizable if and only if $\mathbb{F}^n$ has a basis formed by eigenvectors of A.

Proof (continued). Suppose now that $\mathbb{F}^n$ has a basis, say, $\mathcal{P} = \{\mathbf{p}_1, \dots, \mathbf{p}_n\}$, formed by eigenvectors of A. WTS A is diagonalizable.

For each $i \in \{1, \dots, n\}$, let λ_i be the eigenvalue of A associated with the eigenvector $\mathbf{p}_i$, so that $A\mathbf{p}_i = \lambda_i \mathbf{p}_i$. Set $D = D(\lambda_1, \dots, \lambda_n)$ and $P = \begin{bmatrix} \mathbf{p}_1 & \dots & \mathbf{p}_n \end{bmatrix}$.

Theorem 4.2

A matrix $A \in \mathbb{F}^{n \times n}$ is diagonalizable if and only if $\mathbb{F}^n$ has a basis formed by eigenvectors of A.

Proof (continued). Suppose now that $\mathbb{F}^n$ has a basis, say, $\mathcal{P} = \{\mathbf{p}_1, \dots, \mathbf{p}_n\}$, formed by eigenvectors of A. WTS A is diagonalizable.

For each $i \in \{1, \dots, n\}$, let λ_i be the eigenvalue of A associated with the eigenvector $\mathbf{p}_i$, so that $A\mathbf{p}_i = \lambda_i \mathbf{p}_i$. Set $D = D(\lambda_1, \dots, \lambda_n)$ and $P = \left[\begin{array}{ccc} \mathbf{p}_1 & \dots & \mathbf{p}_n \end{array}\right]$. The matrix D is diagonal (by construction), and the matrix P is invertible because its columns form a basis of $\mathbb{F}^n$. We will show that $D = P^{-1}AP$.

Theorem 4.2

A matrix $A \in \mathbb{F}^{n \times n}$ is diagonalizable if and only if $\mathbb{F}^n$ has a basis formed by eigenvectors of A.

Proof (continued). Suppose now that $\mathbb{F}^n$ has a basis, say, $\mathcal{P} = \{\mathbf{p}_1, \dots, \mathbf{p}_n\}$, formed by eigenvectors of A. WTS A is diagonalizable.

For each $i \in \{1, \dots, n\}$, let λ_i be the eigenvalue of A associated with the eigenvector $\mathbf{p}_i$, so that $A\mathbf{p}_i = \lambda_i \mathbf{p}_i$. Set $D = D(\lambda_1, \dots, \lambda_n)$ and $P = \begin{bmatrix} \mathbf{p}_1 & \dots & \mathbf{p}_n \end{bmatrix}$. The matrix D is diagonal (by construction), and the matrix P is invertible because its columns form a basis of $\mathbb{F}^n$. We will show that $D = P^{-1}AP$.
Note that for each $i \in \{1, \dots, n\}$, the i-th column of AP is $A\mathbf{p}_i$, the i-th column of PD is $\lambda_i \mathbf{p}_i$ (by Proposition 4.1(a)), and we know that $A\mathbf{p}_i = \lambda_i \mathbf{p}_i$.

Theorem 4.2

A matrix $A \in \mathbb{F}^{n \times n}$ is diagonalizable if and only if $\mathbb{F}^n$ has a basis formed by eigenvectors of A.

Proof (continued). Suppose now that $\mathbb{F}^n$ has a basis, say, $\mathcal{P} = \{\mathbf{p}_1, \dots, \mathbf{p}_n\}$, formed by eigenvectors of A. WTS A is diagonalizable.

For each $i \in \{1, \dots, n\}$, let λ_i be the eigenvalue of A associated with the eigenvector $\mathbf{p}_i$, so that $A\mathbf{p}_i = \lambda_i \mathbf{p}_i$. Set $D = D(\lambda_1, \dots, \lambda_n)$ and $P = \begin{bmatrix} \mathbf{p}_1 & \dots & \mathbf{p}_n \end{bmatrix}$. The matrix D is diagonal (by construction), and the matrix P is invertible because its columns form a basis of $\mathbb{F}^n$. We will show that $D = P^{-1}AP$.
Note that for each $i \in \{1, \dots, n\}$, the i-th column of AP is $A\mathbf{p}_i$, the i-th column of PD is $\lambda_i \mathbf{p}_i$ (by Proposition 4.1(a)), and we know that $A\mathbf{p}_i = \lambda_i \mathbf{p}_i$. So, AP and PD have the same corresponding columns, and consequently, $AP = PD$.

Theorem 4.2

A matrix $A \in \mathbb{F}^{n \times n}$ is diagonalizable if and only if $\mathbb{F}^n$ has a basis formed by eigenvectors of A.

Proof (continued). Suppose now that $\mathbb{F}^n$ has a basis, say, $\mathcal{P} = \{\mathbf{p}_1, \dots, \mathbf{p}_n\}$, formed by eigenvectors of A. WTS A is diagonalizable.

For each $i \in \{1, \dots, n\}$, let λ_i be the eigenvalue of A associated with the eigenvector $\mathbf{p}_i$, so that $A\mathbf{p}_i = \lambda_i \mathbf{p}_i$. Set $D = D(\lambda_1, \dots, \lambda_n)$ and $P = \left[\begin{array}{ccc} \mathbf{p}_1 & \dots & \mathbf{p}_n \end{array} \right]$. The matrix D is diagonal (by construction), and the matrix P is invertible because its columns form a basis of $\mathbb{F}^n$. We will show that $D = P^{-1}AP$.
Note that for each $i \in \{1, \dots, n\}$, the i-th column of AP is $A\mathbf{p}_i$, the i-th column of PD is $\lambda_i \mathbf{p}_i$ (by Proposition 4.1(a)), and we know that $A\mathbf{p}_i = \lambda_i \mathbf{p}_i$. So, AP and PD have the same corresponding columns, and consequently, $AP = PD$. We multiply both sides by P^{-1} on the left, and we obtain $P^{-1}AP = D$; since D is diagonal, it follows that A is diagonalizable. Q.E.D.

Theorem 4.2

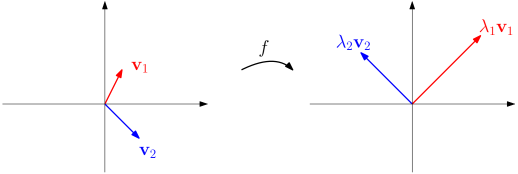
A matrix $A \in \mathbb{F}^{n \times n}$ is diagonalizable if and only if $\mathbb{F}^n$ has a basis formed by eigenvectors of A.

- If $D = P^{-1}AP$ for some diagonal matrix $D = D(\lambda_1, \ldots, \lambda_n)$ and invertible matrix $P = [\ \mathbf{p}_1 \quad \ldots \quad \mathbf{p}_n\ ]$, then A and D are the matrices of the same linear transformation, but with respect to different bases.

Theorem 4.2

A matrix $A \in \mathbb{F}^{n\times n}$ is diagonalizable if and only if $\mathbb{F}^n$ has a basis formed by eigenvectors of A.

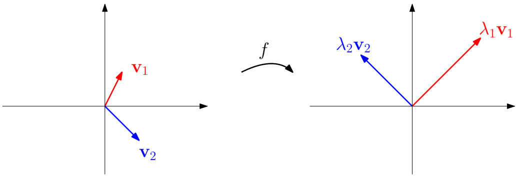
- If $D = P^{-1}AP$ for some diagonal matrix $D = D(\lambda_1, \ldots, \lambda_n)$ and invertible matrix $P = [\ \mathbf{p}_1 \quad \ldots \quad \mathbf{p}_n\ ]$, then A and D are the matrices of the same linear transformation, but with respect to different bases.
- Consider the unique linear transformation $f : \mathbb{F}^n \to \mathbb{F}^n$ such that $f(\mathbf{p}_1) = \lambda_1\mathbf{p}_1, \ldots, f(\mathbf{p}_n) = \lambda_n\mathbf{p}_n$.



Theorem 4.2

A matrix $A \in \mathbb{F}^{n\times n}$ is diagonalizable if and only if $\mathbb{F}^n$ has a basis formed by eigenvectors of A.

- If $D = P^{-1}AP$ for some diagonal matrix $D = D(\lambda_1, \ldots, \lambda_n)$ and invertible matrix $P = [\ \mathbf{p}_1 \quad \ldots \quad \mathbf{p}_n\ ]$, then A and D are the matrices of the same linear transformation, but with respect to different bases.
- Consider the unique linear transformation $f : \mathbb{F}^n \to \mathbb{F}^n$ such that $f(\mathbf{p}_1) = \lambda_1\mathbf{p}_1, \ldots, f(\mathbf{p}_n) = \lambda_n\mathbf{p}_n$.



- Then $A = {}_{\mathcal{E}_n}[f]_{\mathcal{E}_n}$ and $D = {}_{\mathcal{P}}[f]_{\mathcal{P}}$, where $\mathcal{E}_n$ is the standard basis of $\mathbb{F}^n$ and $\mathcal{P} = \{\mathbf{p}_1, \ldots, \mathbf{p}_n\}$.

Theorem 4.2

A matrix $A \in \mathbb{F}^{n \times n}$ is diagonalizable if and only if $\mathbb{F}^n$ has a basis formed by eigenvectors of A.

Theorem 4.2

A matrix $A \in \mathbb{F}^{n \times n}$ is diagonalizable if and only if $\mathbb{F}^n$ has a basis formed by eigenvectors of A.

Corollary 4.3

Let $A \in \mathbb{F}^{n \times n}$ be a matrix that has n distinct eigenvalues. Then A is diagonalizable.

Proof.

Theorem 4.2

A matrix $A \in \mathbb{F}^{n \times n}$ is diagonalizable if and only if $\mathbb{F}^n$ has a basis formed by eigenvectors of A.

Corollary 4.3

Let $A \in \mathbb{F}^{n \times n}$ be a matrix that has n distinct eigenvalues. Then A is diagonalizable.

Proof. Since A has n distinct eigenvalues, and each of those eigenvalues has geometric multiplicity at least 1 (by the definition of an eigenvalue), we see that the sum of geometric multiplicities of the distinct eigenvalues of A is at least n.

Theorem 4.2

A matrix $A \in \mathbb{F}^{n \times n}$ is diagonalizable if and only if $\mathbb{F}^n$ has a basis formed by eigenvectors of A.

Corollary 4.3

Let $A \in \mathbb{F}^{n \times n}$ be a matrix that has n distinct eigenvalues. Then A is diagonalizable.

Proof. Since A has n distinct eigenvalues, and each of those eigenvalues has geometric multiplicity at least 1 (by the definition of an eigenvalue), we see that the sum of geometric multiplicities of the distinct eigenvalues of A is at least n. But then the sum of geometric multiplicities of the distinct eigenvalues of A must be exactly n.

Theorem 4.2

A matrix $A \in \mathbb{F}^{n \times n}$ is diagonalizable if and only if $\mathbb{F}^n$ has a basis formed by eigenvectors of A.

Corollary 4.3

Let $A \in \mathbb{F}^{n \times n}$ be a matrix that has n distinct eigenvalues. Then A is diagonalizable.

Proof. Since A has n distinct eigenvalues, and each of those eigenvalues has geometric multiplicity at least 1 (by the definition of an eigenvalue), we see that the sum of geometric multiplicities of the distinct eigenvalues of A is at least n. But then the sum of geometric multiplicities of the distinct eigenvalues of A must be exactly n. Now Corollary 2.13 of Lecture Notes 18 guarantees that $\mathbb{F}^n$ has a basis formed by eigenvectors of A, and so by Theorem 4.2, A is diagonalizable.

- The theory we have developed so far (including both the statement and the proof of Theorem 4.2) gives us a recipe for determining whether a matrix $A \in \mathbb{F}^{n\times n}$ is diagonalizable, and if so, diagonalizing it.
- We proceed as follows (the steps are in numerals, and justifications of individual steps are given in bullet points underneath the steps in question).

1. We compute the characteristic polynomial $p_A(\lambda)$, factor it, and compute all the eigenvalues of A together with their algebraic multiplicities.
2. If the algebraic multiplicities of the eigenvalues of A add up to less than n, then A is not diagonalizable.
    - **Note:** The sum of algebraic multiplicities of the eigenvalues of A will always be at most n, and if the field $\mathbb{F}$ is algebraically closed (for example, if $\mathbb{F} = \mathbb{C}$), then it will be equal to n. If $\mathbb{F}$ is **not** algebraically closed (for example, if $\mathbb{F}$ is $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{Z}_p$ for some prime number p), then it is possible that the sum of algebraic multiplicities of the eigenvalues of A is less than n. In this case, Corollary 2.13 of Lecture Notes 18 guarantees that $\mathbb{F}$ does not have a basis formed by eigenvectors of A, and consequently (by Theorem 4.2), A is not diagonalizable.

1. We compute the characteristic polynomial $p_A(\lambda)$, factor it, and compute all the eigenvalues of A together with their algebraic multiplicities.
2. If the algebraic multiplicities of the eigenvalues of A add up to less than n, then A is not diagonalizable.
   - **Note:** The sum of algebraic multiplicities of the eigenvalues of A will always be at most n, and if the field $\mathbb{F}$ is algebraically closed (for example, if $\mathbb{F} = \mathbb{C}$), then it will be equal to n. If $\mathbb{F}$ is **not** algebraically closed (for example, if $\mathbb{F}$ is $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{Z}_p$ for some prime number p), then it is possible that the sum of algebraic multiplicities of the eigenvalues of A is less than n. In this case, Corollary 2.13 of Lecture Notes 18 guarantees that $\mathbb{F}$ does not have a basis formed by eigenvectors of A, and consequently (by Theorem 4.2), A is not diagonalizable.
3. From now on, assume that the algebraic multiplicities of the eigenvalues of A add up to n. For each eigenvalue λ of A, find a basis of the eigenspace E_λ and the geometric multiplicity of λ.

4. If the geometric multiplicity of some eigenvalue of A is less than its algebraic multiplicity, then A is not diagonalizable.
    - **Note:** If the geometric multiplicity of some eigenvalue of A is less than its algebraic multiplicity, then Corollary 2.13 of Lecture Notes 18 guarantees that $\mathbb{F}^n$ has no basis formed by eigenvectors of A, and so by Theorem 4.2, A is not diagonalizable.

4. If the geometric multiplicity of some eigenvalue of A is less than its algebraic multiplicity, then A is not diagonalizable.
   - **Note:** If the geometric multiplicity of some eigenvalue of A is less than its algebraic multiplicity, then Corollary 2.13 of Lecture Notes 18 guarantees that $\mathbb{F}^n$ has no basis formed by eigenvectors of A, and so by Theorem 4.2, A is not diagonalizable.
5. From now on, assume that the geometric multiplicity of each eigenvalue of A is equal to its algebraic multiplicity, so that the geometric multiplicities of the eigenvalues of A add up to n. In this case, A is diagonalizable.
   - **Note:** Indeed, if the geometric multiplicities of the eigenvectors of A add up to n, then Corollary 2.13 of Lecture Notes 18 guarantees that $\mathbb{F}^n$ has a basis formed by eigenvectors of A, and so by Theorem 4.2, A is diagonalizable.

6. Let $\lambda_1, \dots, \lambda_\ell$ be the distinct eigenvalues of A, with geometric multiplicities $n_1, \dots, n_\ell$, respectively. (So, $n_1 + \dots + n_\ell = n$.) We now form the diagonal matrix
$D = D(\underbrace{\lambda_1, \dots, \lambda_1}_{n_1}, \dots, \underbrace{\lambda_\ell, \dots, \lambda_\ell}_{n_\ell})$, and we form the matrix P
as follows: the first n_1 columns of P form a basis of the eigenspace E_{λ_1}, the next n_2 columns of P form a basis of the eigenspace E_{λ_2}, and so on, until the last n_ℓ columns of P form a basis of the eigenspace E_{λ_ℓ}. Now $D = PAP^{-1}$

   - **Note:** Theorem 2.12 of Lecture Notes 18 guarantees that the columns of P form a linearly independent set; since P has n columns, it follows that the columns of P form a basis of $\mathbb{F}^n$. Now the correctness of our construction follows from the proof of Theorem 4.2.

Example 4.4

Consider the matrix

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

in $\mathbb{C}^{3\times 3}$. Determine whether A is diagonalizable, and if so, diagonalize it.

Solution.

Example 4.4

Consider the matrix

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

in $\mathbb{C}^{3\times 3}$. Determine whether A is diagonalizable, and if so, diagonalize it.

Solution. The matrix A is precisely the matrix from Example 2.6 of Lecture Notes 18. In that example, we determined that A has two eigenvalues, namely, $\lambda_1 = 4$ (with algebraic multiplicity 1 and geometric multiplicity 1) and $\lambda_2 = 5$ (with algebraic multiplicity 2 and geometric multiplicity 2). Since the sum of geometric multiplicities is 3, we see that the 3×3 matrix A is indeed diagonalizable.

Example 4.4

Consider the matrix

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

in $\mathbb{C}^{3\times 3}$. Determine whether A is diagonalizable, and if so, diagonalize it.

Solution (continued). In Example 2.6 of Lecture Notes 18, we saw that $\{[\ -1 \quad 2 \quad 0\ ]^T\}$ is a basis of the eigespace E_{λ_1}, and that $\{[\ 0 \quad 1 \quad 0\ ]^T, [\ -2 \quad 0 \quad 1\ ]^T\}$ is a basis of the eigenspace E_{λ_2}.

Example 4.4

Consider the matrix

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

in $\mathbb{C}^{3\times 3}$. Determine whether A is diagonalizable, and if so, diagonalize it.

Solution (continued). In Example 2.6 of Lecture Notes 18, we saw that $\{[\begin{matrix} -1 & 2 & 0 \end{matrix}]^T\}$ is a basis of the eigespace E_{λ_1}, and that $\{[\begin{matrix} 0 & 1 & 0 \end{matrix}]^T, [\begin{matrix} -2 & 0 & 1 \end{matrix}]^T\}$ is a basis of the eigenspace E_{λ_2}. 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 $D = P^{-1}AP$.

Example 4.5

Consider the matrix

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

in $\mathbb{C}^{5\times 5}$. Determine whether A is diagonalizable, and if so, diagonalize it.

Solution

Example 4.5

Consider the matrix

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

in $\mathbb{C}^{5\times 5}$. Determine whether A is diagonalizable, and if so, diagonalize it.

Solution The matrix A is precisely the matrix from Example 2.10 of Lecture Notes 18. In that example, we determined that A has three eigenvalues, namely $\lambda_1 = 1$ (with algebraic multiplicity 2 and geometric multiplicity 2), $\lambda_2 = 2$ (with algebraic multiplicity 1 and geometric multiplicity 1), and $\lambda_3 = 3$ (with algebraic multiplicity 2 and geometric multiplicity 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.

- One reason we care about diagonalizability is because diagonalizable matrices are easy to exponentiate.

- Suppose, furthermore, that we have diagonalized A as $D = P^{-1}AP$, where $D = D(\lambda_1, \ldots, \lambda_n)$ is a diagonal matrix and P an invertible matrix in $\mathbb{F}^{n \times n}$.

- One reason we care about diagonalizability is because diagonalizable matrices are easy to exponentiate.
- Indeed, suppose we are given a diagonalizable matrix $A \in \mathbb{F}^{n \times n}$.
- Suppose, furthermore, that we have diagonalized A as $D = P^{-1}AP$, where $D = D(\lambda_1, \ldots, \lambda_n)$ is a diagonal matrix and P an invertible matrix in $\mathbb{F}^{n \times n}$.

- One reason we care about diagonalizability is because diagonalizable matrices are easy to exponentiate.
- Indeed, suppose we are given a diagonalizable matrix $A \in \mathbb{F}^{n \times n}$.
- Suppose, furthermore, that we have diagonalized A as $D = P^{-1}AP$, where $D = D(\lambda_1, \ldots, \lambda_n)$ is a diagonal matrix and P an invertible matrix in $\mathbb{F}^{n \times n}$.
- Then by Proposition 2.2, for all non-negative integers m, we have that $A^m = PD^mP^{-1}$.

- One reason we care about diagonalizability is because diagonalizable matrices are easy to exponentiate.
- Indeed, suppose we are given a diagonalizable matrix $A \in \mathbb{F}^{n\times n}$.
- Suppose, furthermore, that we have diagonalized A as $D = P^{-1}AP$, where $D = D(\lambda_1, \ldots, \lambda_n)$ is a diagonal matrix and P an invertible matrix in $\mathbb{F}^{n\times n}$.
- Then by Proposition 2.2, for all non-negative integers m, we have that $A^m = PD^mP^{-1}$.
- On the other hand, by Proposition 4.1(b), we have that $D^m = D(\lambda_1^m, \ldots, \lambda_n^m)$.

- One reason we care about diagonalizability is because diagonalizable matrices are easy to exponentiate.
- Indeed, suppose we are given a diagonalizable matrix $A \in \mathbb{F}^{n \times n}$.
- Suppose, furthermore, that we have diagonalized A as $D = P^{-1}AP$, where $D = D(\lambda_1, \ldots, \lambda_n)$ is a diagonal matrix and P an invertible matrix in $\mathbb{F}^{n \times n}$.
- Then by Proposition 2.2, for all non-negative integers m, we have that $A^m = PD^mP^{-1}$.
- On the other hand, by Proposition 4.1(b), we have that $D^m = D(\lambda_1^m, \ldots, \lambda_n^m)$.
- So, A^m is easy to compute.

Example 4.6

Consider the matrix

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

in $\mathbb{C}^{3\times 3}$. Find a formula for A^m (where m is an arbitrary non-negative integer).

Solution.

Example 4.6

Consider the matrix

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

in $\mathbb{C}^{3\times 3}$. Find a formula for A^m (where m is an arbitrary non-negative integer).

Solution. This is the matrix from Example 4.4.

Example 4.6

Consider the matrix

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

in $\mathbb{C}^{3\times 3}$. Find a formula for A^m (where m is an arbitrary non-negative integer).

Solution. This is the matrix from Example 4.4. In that example, we computed matrices $D, P \in \mathbb{C}^{3\times 3}$ such that D is diagonal, P is invertible, and $D = P^{-1}AP$. The matrices in question were

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

Example 4.6

Consider the matrix

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

in $\mathbb{C}^{3\times 3}$. Find a formula for A^m (where m is an arbitrary non-negative integer).

Solution (continued). We then compute

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

Solution (continued). Then for all non-negative integers m, we have the following:

$$\begin{aligned} A^m &\overset{(*)}{=} PD^mP^{-1} \\ &\overset{(**)}{=} \underbrace{\begin{bmatrix} -1 & 0 & -2 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{=P} \underbrace{\begin{bmatrix} 4^m & 0 & 0 \\ 0 & 5^m & 0 \\ 0 & 0 & 5^m \end{bmatrix}}_{=D^m} \underbrace{\begin{bmatrix} -1 & 0 & -2 \\ 2 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix}}_{=P^{-1}} \\ &\overset{(***)}{=} \begin{bmatrix} 4^m & 0 & 2\cdot 4^m - 2\cdot 5^m \\ -2\cdot 4^m + 2\cdot 5^m & 5^m & -4^{m+1} + 4\cdot 5^m \\ 0 & 0 & 5^m \end{bmatrix} \\ &= \begin{bmatrix} 4^m & 0 & 2(4^m - 5^m) \\ 2(5^m - 4^m) & 5^m & 4(5^m - 4^m) \\ 0 & 0 & 5^m \end{bmatrix}, \end{aligned}$$

where (*) follows from Proposition 2.2, (**) follows from Proposition 4.1, and (***) follows by simple matrix multiplication.

Example 4.6

Consider the matrix

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

in $\mathbb{C}^{3\times 3}$. Find a formula for A^m (where m is an arbitrary non-negative integer).

Solution (continued). Reminder:

$$A^m = \begin{bmatrix} 4^m & 0 & 2(4^m - 5^m) \\ 2(5^m - 4^m) & 5^m & 4(5^m - 4^m) \\ 0 & 0 & 5^m \end{bmatrix}$$

for all non-negative integers m.

Example 4.6

Consider the matrix

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

in $\mathbb{C}^{3\times 3}$. Find a formula for A^m (where m is an arbitrary non-negative integer).

Solution (continued). Reminder:

$$A^m = \begin{bmatrix} 4^m & 0 & 2(4^m - 5^m) \\ 2(5^m - 4^m) & 5^m & 4(5^m - 4^m) \\ 0 & 0 & 5^m \end{bmatrix}$$

for all non-negative integers m.

Optionally, we can verify by induction on m that our formula is correct (details: Lecture Notes).