Eigenvalues and Eigenvectors

November 3, 2016

Definition (Eigenvalues and Eigenvectors)

The (complex) number λ *is called an* **eigenvalue** *of the* $n \times n$ *matrix* $\mathbf{A}$ *provided there exists a nonzero (complex) vector* $\mathbf{v}$ *such that*

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

in which case the vector $\mathbf{v}$ *is called an* **eigenvector** *of* $\mathbf{A}$.

Example

Consider

$$\mathbf{A} = \begin{pmatrix} 5 & -6 \\ 2 & -2 \end{pmatrix}.$$

We have

$$\mathbf{A}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 & -6 \\ 2 & -2 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix} = 2\begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Thus $\lambda = 2$ is an eigenvalue of $\mathbf{A}$ and $(2,1)^T$ is an eigenvector of $\mathbf{A}$ associated with the eigenvalue $\lambda = 2$. We also have

$$\mathbf{A}\begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 5 & -6 \\ 2 & -2 \end{pmatrix}\begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

Thus $\lambda = 1$ is an eigenvalue of $\mathbf{A}$ and $(3,2)^T$ is an eigenvector of $\mathbf{A}$ associated with the eigenvalue $\lambda = 1$.

Remarks

1. An eigenvalue may be zero but an eigenvector is by definition a nonzero vector.
2. If $\mathbf{v}_1$ and $\mathbf{v}_2$ are eigenvectors of $\mathbf{A}$ associated with eigenvalue λ, then for any scalars c_1 and c_2, $c_1\mathbf{v}_1 + c_2\mathbf{v}_2$ is also an eigenvector of $\mathbf{A}$ associated with eigenvalue λ if it is non-zero.
3. If λ is an eigenvalue of an $n \times n$ matrix $\mathbf{A}$, then the set of all eigenvectors associated with eigenvalue λ together with the zero vector $\mathbf{0}$ form a vector subspace of $\mathbb{R}^n$. It is called the **eigenspace** of $\mathbf{A}$ associated with eigenvalue λ.

Definition (Characteristic equation)

Let $\mathbf{A}$ *be an* $n \times n$ *matrix. The polynomial equation*

$$\det(\mathbf{A} - x\mathbf{I}) = 0$$

of degree n *is called the* **characteristic equation** *of* $\mathbf{A}$*.*

Theorem

Let $\mathbf{A}$ *be an* $n \times n$ *matrix. The following statements for scalar* λ *are equivalent.*

1. λ *is an eigenvalue of* $\mathbf{A}$.
2. *The equation* $(\mathbf{A} - \lambda \mathbf{I})\mathbf{v} = \mathbf{0}$ *has nontrivial solution for* $\mathbf{v}$.
3. $\mathrm{Null}(\mathbf{A} - \lambda \mathbf{I}) \neq \{\mathbf{0}\}$.
4. *The matrix* $\mathbf{A} - \lambda \mathbf{I}$ *is singular.*
5. λ *is a root of the characteristic equation* $\det(\mathbf{A} - x\mathbf{I}) = 0$

Example

Find the eigenvalues and associated eigenvectors of the matrix

$$\mathbf{A} = \begin{pmatrix} 3 & 2 \\ 3 & -2 \end{pmatrix}.$$

Solution: Solving the characteristic equation, we have

$$\begin{aligned} \det(\mathbf{A} - \lambda \mathbf{I}) &= 0 \\ \begin{vmatrix} 3-\lambda & 2 \\ 3 & -2-\lambda \end{vmatrix} &= 0 \\ \lambda^2 - \lambda - 12 &= 0 \\ \lambda &= 4, -3 \end{aligned}$$

When $\lambda = 4$,

$$\begin{aligned} (\mathbf{A} - 4\mathbf{I})\mathbf{v} &= \mathbf{0} \\ \begin{pmatrix} -1 & 2 \\ 3 & -6 \end{pmatrix} \mathbf{v} &= \mathbf{0} \end{aligned}$$

Thus $\mathbf{v} = (2, 1)^T$ is an eigenvector associated with $\lambda = 4$.
When $\lambda = -3$,

$$\begin{aligned} (\mathbf{A} + 3\mathbf{I})\mathbf{v} &= \mathbf{0} \\ \begin{pmatrix} 6 & 2 \\ 3 & 1 \end{pmatrix} \mathbf{v} &= \mathbf{0} \end{aligned}$$

Thus $\mathbf{v} = (1, -3)^T$ is an eigenvector associated with $\lambda = -3$. $\square$

Example

Find the eigenvalues and associated eigenvectors of the matrix

$$\mathbf{A} = \begin{pmatrix} 0 & 8 \\ -2 & 0 \end{pmatrix}.$$

Solution: Solving the characteristic equation, we have

$$\begin{aligned} \begin{vmatrix} -\lambda & 8 \\ -2 & -\lambda \end{vmatrix} &= 0 \\ \lambda^2 + 16 &= 0 \\ \lambda &= \pm 4i \end{aligned}$$

When $\lambda = 4i$,

$$\begin{aligned}(\mathbf{A} - 4i\mathbf{I})\mathbf{v} &= \mathbf{0} \\ \begin{pmatrix} -4i & 8 \\ -2 & -4i \end{pmatrix}\mathbf{v} &= \mathbf{0}\end{aligned}$$

Thus $\mathbf{v} = (2, -i)^T$ is an eigenvector associated with $\lambda = 4i$.
When $\lambda = -4i$,

$$\begin{aligned}(\mathbf{A} + 4i\mathbf{I})\mathbf{v} &= \mathbf{0} \\ \begin{pmatrix} 4i & 8 \\ -2 & 4i \end{pmatrix}\mathbf{v} &= \mathbf{0}\end{aligned}$$

Thus $\mathbf{v} = (2, i)^T$ is an eigenvector associated with $\lambda = -4i$. $\square$

Remark

For any square matrix **A** with real entries, the characteristic polynomial of **A** has real coefficients. Thus if $\lambda = \rho + \mu i$, where $\rho, \mu \in \mathbb{R}$, is a complex eigenvalue of **A**, then $\bar{\lambda} = \rho - \mu i$ is also an eigenvalue of **A**. Furthermore, if $\mathbf{v} = \mathbf{a} + \mathbf{b}i$ is an eigenvector associated with complex eigenvalue λ, then $\bar{\mathbf{v}} = \mathbf{a} - \mathbf{b}i$ is an eigenvector associated with eigenvalue $\bar{\lambda}$.

Example

Find the eigenvalues and associated eigenvectors of the matrix $\mathbf{A} = \begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix}$.

Solution: Solving the characteristic equation, we have

$$\begin{aligned} \begin{vmatrix} 2-\lambda & 3 \\ 0 & 2-\lambda \end{vmatrix} &= 0 \\ (\lambda-2)^2 &= 0 \\ \lambda &= 2, 2 \end{aligned}$$

When $\lambda = 2$,

$$\begin{pmatrix} 0 & 3 \\ 0 & 0 \end{pmatrix} \mathbf{v} = \mathbf{0}$$

Thus $\mathbf{v} = (1, 0)^T$ is an eigenvector associated with $\lambda = 2$. In this example, we call only find one linearly independent eigenvector. □

Example

Find the eigenvalues and associated eigenvectors of the matrix

$$\mathbf{A} = \begin{pmatrix} 2 & -3 & 1 \\ 1 & -2 & 1 \\ 1 & -3 & 2 \end{pmatrix}.$$

Solution: Solving the characteristic equation, we have

$$\begin{aligned} \begin{vmatrix} 2-\lambda & -3 & 1 \\ 1 & -2-\lambda & 1 \\ 1 & -3 & 2-\lambda \end{vmatrix} &= 0 \\ \lambda(\lambda-1)^2 &= 0 \\ \lambda &= 1, 1, 0 \end{aligned}$$

For $\lambda_1 = \lambda_2 = 1$,

$$\begin{pmatrix} 1 & -3 & 1 \\ 1 & -3 & 1 \\ 1 & -3 & 1 \end{pmatrix} \mathbf{v} = \mathbf{0}$$

Thus $\mathbf{v}_1 = (3, 1, 0)^T$ and $\mathbf{v}_2 = (-1, 0, 1)$ are two linearly independent eigenvectors associated with $\lambda = 1$. For $\lambda_3 = 0$,

$$\begin{pmatrix} 2 & -3 & 1 \\ 1 & -2 & 1 \\ 1 & -3 & 2 \end{pmatrix} \mathbf{v} = \mathbf{0}$$

Thus $\mathbf{v}_3 = (1, 1, 1)^T$ is an eigenvector associated with $\lambda = 0$. $\square$

Example

Find the eigenvalues and associated eigenvectors of the matrix

$$\mathbf{A} = \begin{pmatrix} -1 & 1 & 0 \\ -4 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix}.$$

Solution: Solving the characteristic equation, we have

$$\begin{aligned} \begin{vmatrix} -1-\lambda & 1 & 0 \\ -4 & 3-\lambda & 0 \\ 1 & 0 & 2-\lambda \end{vmatrix} &= 0 \\ (\lambda-2)(\lambda-1)^2 &= 0 \\ \lambda &= 2, 1, 1 \end{aligned}$$

For $\lambda_1 = 2$,

$$\begin{pmatrix} -3 & 1 & 0 \\ -4 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \mathbf{v} = \mathbf{0}$$

Thus $\mathbf{v} = (0, 0, 1)^T$ is an eigenvector associated with $\lambda = 2$.
For $\lambda_2 = \lambda_3 = 1$,

$$\begin{pmatrix} -2 & 1 & 0 \\ -4 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} \mathbf{v} = \mathbf{0}$$

Thus $(-1, -2, 1)^T$ is an eigenvectors associated with $\lambda = 1$. Note that here $\lambda = 1$ is a double root but we can only find one linearly independent eigenvector associated with $\lambda = 1$. $\square$

Definition (Similar matrices)

Two $n \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$ are said to be **similar** *if there exists an invertible matrix $\mathbf{P}$ such that*

$$\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}.$$

Proposition

Similarity of square matrices is an equivalence relation, that is,

1. *For any square matrix $\mathbf{A}$, we have $\mathbf{A}$ is similar to $\mathbf{A}$;*
2. *If $\mathbf{A}$ is similar to $\mathbf{B}$, then $\mathbf{B}$ is similar to $\mathbf{A}$;*
3. *If $\mathbf{A}$ is similar to $\mathbf{B}$ and $\mathbf{B}$ is similar to $\mathbf{C}$, then $\mathbf{A}$ is similar to $\mathbf{C}$.*

Proof.

1. Since $\mathbf{I}$ is a non-singular matrix and $\mathbf{A} = \mathbf{I}^{-1}\mathbf{A}\mathbf{I}$, we have $\mathbf{A}$ is similar to $\mathbf{A}$.
2. If $\mathbf{A}$ is similar to $\mathbf{B}$, then there exists non-singular matrix $\mathbf{P}$ such that $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. Now $\mathbf{P}^{-1}$ is a non-singular matrix and $(\mathbf{P}^{-1})^{-1} = \mathbf{P}$. There exists non-singular matrix $\mathbf{P}^{-1}$ such that $(\mathbf{P}^{-1})^{-1}\mathbf{B}\mathbf{P}^{-1} = \mathbf{P}\mathbf{B}\mathbf{P}^{-1} = \mathbf{A}$. Therefore $\mathbf{B}$ is similar to $\mathbf{A}$.
3. If $\mathbf{A}$ is similar to $\mathbf{B}$ and $\mathbf{B}$ is similar to $\mathbf{C}$, then there exists non-singular matrices $\mathbf{P}$ and $\mathbf{Q}$ such that $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ and $\mathbf{C} = \mathbf{Q}^{-1}\mathbf{B}\mathbf{Q}$. Now $\mathbf{PQ}$ is a non-singular matrix and $(\mathbf{PQ})^{-1} = \mathbf{Q}^{-1}\mathbf{P}^{-1}$. There exists non-singular matrix $\mathbf{PQ}$ such that $(\mathbf{PQ})^{-1}\mathbf{A}(\mathbf{PQ}) = \mathbf{Q}^{-1}(\mathbf{P}^{-1}\mathbf{A}\mathbf{P})\mathbf{Q} = \mathbf{Q}^{-1}\mathbf{B}\mathbf{Q} = \mathbf{C}$. Therefore $\mathbf{A}$ is similar to $\mathbf{C}$.

$\square$

Proposition

1. *The only matrix similar to the zero matrix* $\mathbf{0}$ *is the zero matrix.*
2. *The only matrix similar to the identity matrix* $\mathbf{I}$ *is the identity matrix.*
3. *If* $\mathbf{A}$ *is similar to* $\mathbf{B}$*, then* $a\mathbf{A}$ *is similar to* $a\mathbf{B}$ *for any real number* a*.*
4. *If* $\mathbf{A}$ *is similar to* $\mathbf{B}$*, then* $\mathbf{A}^k$ *is similar to* $\mathbf{B}^k$ *for any non-negative integer* k*.*
5. *If* $\mathbf{A}$ *and* $\mathbf{B}$ *are similar non-singular matrices, then* $\mathbf{A}^{-1}$ *is similar to* $\mathbf{B}^{-1}$*.*
6. *If* $\mathbf{A}$ *is similar to* $\mathbf{B}$*, then* $\mathbf{A}^T$ *is similar to* $\mathbf{B}^T$*.*
7. *If* $\mathbf{A}$ *is similar to* $\mathbf{B}$*, then* $\det(\mathbf{A}) = \det(\mathbf{B})$*.*
8. *If* $\mathbf{A}$ *is similar to* $\mathbf{B}$*, then* $\mathrm{tr}(\mathbf{A}) = \mathrm{tr}(\mathbf{B})$ *where* $\mathrm{tr}(\mathbf{A}) = a_{11} + a_{22} + \cdots + a_{nn}$ *is the* **trace***, i.e., the sum of the entries in the diagonal, of* $\mathbf{A}$*.*
9. *If* $\mathbf{A}$ *is similar to* $\mathbf{B}$*, then* $\mathbf{A}$ *and* $\mathbf{B}$ *have the same characteristic equation.*

Proof.

(1) Suppose $\mathbf{A}$ is similar to $\mathbf{0}$, then there exists non-singular matrix $\mathbf{P}$ such that $\mathbf{0} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. Hence $\mathbf{A} = \mathbf{P}\mathbf{0}\mathbf{P}^{-1} = \mathbf{0}$.

(4) If $\mathbf{A}$ is similar to $\mathbf{B}$, then there exists non-singular matrix $\mathbf{P}$ such that $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. We have

$$\begin{aligned}
\mathbf{B}^k &= \overbrace{(\mathbf{P}^{-1}\mathbf{A}\mathbf{P})(\mathbf{P}^{-1}\mathbf{A}\mathbf{P})\cdots(\mathbf{P}^{-1}\mathbf{A}\mathbf{P})}^{k \text{ copies}} \\
&= \mathbf{P}^{-1}\mathbf{A}(\mathbf{P}\mathbf{P}^{-1})\mathbf{A}\mathbf{P}\cdots\mathbf{P}^{-1}\mathbf{A}(\mathbf{P}\mathbf{P}^{-1})\mathbf{A}\mathbf{P} \\
&= \mathbf{P}^{-1}\mathbf{A}\mathbf{I}\mathbf{A}\mathbf{I}\cdots\mathbf{I}\mathbf{A}\mathbf{I}\mathbf{A}\mathbf{P} \\
&= \mathbf{P}^{-1}\mathbf{A}^k\mathbf{P}
\end{aligned}$$

Therefore $\mathbf{A}^k$ is similar to $\mathbf{B}^k$.

$\square$

Proof.

(7) If $\mathbf{A}$ is similar to $\mathbf{B}$, then there exists non-singular matrix $\mathbf{P}$ such that $\mathbf{B} = \mathbf{P}^{-1}\mathbf{AP}$. Thus $\det(\mathbf{B}) = \det(\mathbf{P}^{-1}\mathbf{AP}) = \det(\mathbf{P}^{-1})\det(\mathbf{A})\det(\mathbf{P}) = \det(\mathbf{A})$ since $\det(\mathbf{P}^{-1}) = \det(\mathbf{P})^{-1}$.

(8) If $\mathbf{A}$ is similar to $\mathbf{B}$, then there exists non-singular matrix $\mathbf{P}$ such that $\mathbf{B} = \mathbf{P}^{-1}\mathbf{AP}$. Thus $\mathrm{tr}(\mathbf{B}) = \mathrm{tr}((\mathbf{P}^{-1}\mathbf{A})\mathbf{P}) = \mathrm{tr}(\mathbf{P}(\mathbf{P}^{-1}\mathbf{A})) = \mathrm{tr}(\mathbf{A})$. (Note: It is well-known that $\mathrm{tr}(\mathbf{PQ}) = \mathrm{tr}(\mathbf{QP})$ for any square matrices $\mathbf{P}$ and $\mathbf{Q}$. But in general, it is not always true that $\mathrm{tr}(\mathbf{PQR}) = \mathrm{tr}(\mathbf{QPR})$.)

$\square$

Definition

An $n \times n$ matrix $\mathbf{A}$ is said to be **diagonalizable** *if there exists a nonsingular (may be complex) matrix $\mathbf{P}$ such that*

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$$

is a diagonal matrix. We say that $\mathbf{P}$ diagonalizes $\mathbf{A}$. In other words, $\mathbf{A}$ is diagonalizable if it is similar to a diagonal matrix.

Theorem

Let $\mathbf{A}$ be an $n \times n$ matrix. Then $\mathbf{A}$ is diagonalizable if and only if $\mathbf{A}$ has n linearly independent eigenvectors.

Proof. Let $\mathbf{P}$ be an $n \times n$ matrix and write

$$\mathbf{P} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix}.$$

where $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are the column vectors of $\mathbf{P}$. First observe that $\mathbf{P}$ is nonsingular if and only if $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are linearly independent.

Furthermore

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \text{ is a diagonal matrix.}$$

$$\Leftrightarrow \quad \mathbf{A}\mathbf{P} = \mathbf{P}\mathbf{D} \text{ for some diagonal matrix } \mathbf{D} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

$$\Leftrightarrow \quad \mathbf{A}\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix} = \begin{bmatrix} \lambda_1\mathbf{v}_1 & \lambda_2\mathbf{v}_2 & \cdots & \lambda_n\mathbf{v}_n \end{bmatrix}.$$

$$\Leftrightarrow \quad \begin{bmatrix} \mathbf{A}\mathbf{v}_1 & \mathbf{A}\mathbf{v}_2 & \cdots & \mathbf{A}\mathbf{v}_n \end{bmatrix} = \begin{bmatrix} \lambda_1\mathbf{v}_1 & \lambda_2\mathbf{v}_2 & \cdots & \lambda_n\mathbf{v}_n \end{bmatrix}.$$

$$\Leftrightarrow \quad \mathbf{v}_k \text{ is an eigenvector of } \mathbf{A} \text{ associated with eigenvalue } \lambda_k \text{ for } k = 1, 2, \cdots, n.$$

Therefore $\mathbf{P}$ diagonalizes $\mathbf{A}$ if and only if $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are linearly independent eigenvectors of $\mathbf{A}$. □

Example

Diagonalize the matrix

$$\mathbf{A} = \begin{pmatrix} 3 & 2 \\ 3 & -2 \end{pmatrix}.$$

Solution: We have seen in Example 8 that **A** has eigenvalues $\lambda_1 = 4$ and $\lambda_2 = -3$ associated with linearly independent eigenvectors $\mathbf{v}_1 = (2,1)^T$ and $\mathbf{v}_2 = (1,-3)^T$ respectively. Thus the matrix

$$\mathbf{P} = \begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix}$$

diagonalizes **A** and

$$\begin{aligned} \mathbf{P}^{-1}\mathbf{A}\mathbf{P} &= \begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix}^{-1} \begin{pmatrix} 3 & 2 \\ 3 & -2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix} \\ &= \begin{pmatrix} 4 & 0 \\ 0 & -3 \end{pmatrix}. \end{aligned}$$

Example

Diagonalize the matrix

$$\mathbf{A} = \begin{pmatrix} 0 & 8 \\ -2 & 0 \end{pmatrix}.$$

Solution: We have seen in Example 10 that $\mathbf{A}$ has eigenvalues $\lambda_1 = 4i$ and $\lambda_2 = -4i$ associated with linearly independent eigenvectors $\mathbf{v}_1 = (2, -i)^T$ and $\mathbf{v}_2 = (2, i)^T$ respectively. Thus the matrix

$$\mathbf{P} = \begin{pmatrix} 2 & 2 \\ -i & i \end{pmatrix}$$

diagonalizes $\mathbf{A}$ and

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix} 4i & 0 \\ 0 & -4i \end{pmatrix}.$$

Example

We have seen in Example 14 that

$$\mathbf{A} = \begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix}$$

has only one linearly independent eigenvector. Therefore it is not diagonalizable.

Example

Diagonalize the matrix $\mathbf{A} = \begin{pmatrix} 2 & -3 & 1 \\ 1 & -2 & 1 \\ 1 & -3 & 2 \end{pmatrix}$.

Solution: We have seen in Example 10 that $\mathbf{A}$ has eigenvalues $\lambda_1 = \lambda_2 = 1$ and $\lambda_3 = 0$. For $\lambda_1 = \lambda_2 = 1$, there are two linearly independent eigenvectors $\mathbf{v}_1 = (3, 1, 0)^T$ and $\mathbf{v}_2 = (-1, 0, 1)^T$. For $\lambda_3 = 0$, there associated one linearly independent eigenvector $\mathbf{v}_3 = (1, 1, 1)^T$. The three vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly independent eigenvectors. Thus the matrix

$$\mathbf{P} = \begin{pmatrix} 3 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

diagonalizes $\mathbf{A}$ and

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Example

Show that the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

is not diagonalizable.

Solution: One can show that there is at most one linearly independent eigenvector. Alternatively, one can argue in the following way. The characteristic equation of $\mathbf{A}$ is $(r-1)^2 = 0$. Thus $\lambda = 1$ is the only eigenvalue of $\mathbf{A}$. Hence if $\mathbf{A}$ is diagonalizable by $\mathbf{P}$, then $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{I}$. But then $\mathbf{A} = \mathbf{P}\mathbf{I}\mathbf{P}^{-1} = \mathbf{I}$ which leads to a contradiction. $\square$

Theorem

Suppose that eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ *are associated with the distinct eigenvalues* $\lambda_1, \lambda_2, \cdots, \lambda_k$ *of a matrix* $\mathbf{A}$. *Then* $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ *are linearly independent.*

Proof. We prove the theorem by induction on k. The theorem is obviously true when $k = 1$. Now assume that the theorem is true for any set of $k - 1$ eigenvectors. Suppose

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}.$$

Multiplying $\mathbf{A} - \lambda_k\mathbf{I}$ to the left on both sides, we have

$$\begin{aligned} c_1(\mathbf{A} - \lambda_k\mathbf{I})\mathbf{v}_1 + c_2(\mathbf{A} - \lambda_k\mathbf{I})\mathbf{v}_2 + \cdots + c_{k-1}(\mathbf{A} - \lambda_k\mathbf{I})\mathbf{v}_{k-1} + c_k(\mathbf{A} - \lambda_k\mathbf{I})\mathbf{v}_k &= \mathbf{0} \\ c_1(\lambda_1 - \lambda_k)\mathbf{v}_1 + c_2(\lambda_2 - \lambda_k)\mathbf{v}_2 + \cdots + c_{k-1}(\lambda_{k-1} - \lambda_k)\mathbf{v}_{k-1} &= \mathbf{0} \end{aligned}$$

Note that $(\mathbf{A} - \lambda_k \mathbf{I})\mathbf{v}_k = \mathbf{0}$ since $\mathbf{v}_k$ is an eigenvector associated with λ_k. From the induction hypothesis, $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_{k-1}$ are linearly independent. Thus

$$c_1(\lambda_1 - \lambda_k) = c_2(\lambda_2 - \lambda_k) = \cdots = c_{k-1}(\lambda_{k-1} - \lambda_k) = 0.$$

Since $\lambda_1, \lambda_2, \cdots, \lambda_k$ are distinct, $\lambda_1 - \lambda_k, \lambda_2 - \lambda_k, \cdots, \lambda_{k-1} - \lambda_k$ are all nonzero. Hence

$$c_1 = c_2 = \cdots = c_{k-1} = 0.$$

It follows then that c_k is also equal to zero because $\mathbf{v}_k$ is a nonzero vector. Therefore $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ are linearly independent. $\square$

Theorem

If the $n \times n$ matrix $\mathbf{A}$ has n distinct eigenvalues, then it is diagonalizable.

Theorem

Let $\mathbf{A}$ be an $n \times n$ matrix and $\lambda_1, \lambda_2, \cdots, \lambda_k$ be the roots of the characteristic equation of $\mathbf{A}$ of multiplicity $n_1, n_2, \cdots, n_k$ respectively, i.e., the characteristic equation of $\mathbf{A}$ is

$$(x-\lambda_1)^{n_1}(x-\lambda_2)^{n_2}\cdots(x-\lambda_n)^{n_k}=0.$$

Then $\mathbf{A}$ is diagonalizable if and only if for each $1 \leq i \leq k$, there exists n_i linearly independent eigenvectors associated with eigenvalue λ_i.

Remark

Let $\mathbf{A}$ be a square matrix. A (complex) number λ is an eigenvalue of $\mathbf{A}$ if and only if λ is a root of the characteristic equation of $\mathbf{A}$. The multiplicity of λ being a root of the characteristic equation is called the **algebraic multiplicity** of λ. The dimension of the eigenspace associated to eigenvalue λ, that is, the maximum number of linearly independent eigenvectors associated with eigenvalue λ, is called the **geometric multiplicity** of λ. It can be shown that for each eigenvalue λ, we have

$$1 \leq m_g \leq m_a$$

where m_g and m_a are the geometric and algebraic multiplicity of λ respectively. Therefore we have $\mathbf{A}$ is diagonalizable if and only if the algebraic multiplicity and the geometric multiplicity are equal for all eigenvalues of $\mathbf{A}$.

Theorem

A square matrix is diagonlizable if and only if its minimal polynomial has no repeated factor.

Theorem

All eigenvalues of a symmetric matrix are real.

Theorem

Any symmetric matrix is diagonalizable (by orthogonal matrix).

Exercise

Diagonalize the following matrices.

(a) $\begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}$ (b) $\begin{pmatrix} 3 & -2 \\ 4 & -1 \end{pmatrix}$ (c) $\begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 6 & 11 & 6 \end{pmatrix}$ (d) $\begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix}$

Exercise

Show that that following matrices are not diagonalizable.

(a) $\begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} -1 & 1 & 0 \\ -4 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix}$ (c) $\begin{pmatrix} -3 & 3 & -2 \\ -7 & 6 & -3 \\ 1 & -1 & 2 \end{pmatrix}$

Suggested answer

1(a) $P = \begin{pmatrix} 3 & -1 \\ 4 & 1 \end{pmatrix}$, $D = \begin{pmatrix} 5 & 0 \\ 0 & -2 \end{pmatrix}$,

(b) $P = \begin{pmatrix} 1+i & 1-i \\ 2 & 2 \end{pmatrix}$, $D = \begin{pmatrix} 1+i & 0 \\ 0 & 1-i \end{pmatrix}$,

(c) $P = \begin{pmatrix} 1 & 1 & 1 \\ -1 & -2 & -3 \\ 1 & 4 & 9 \end{pmatrix}$, $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$,

(d) $P = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}$, $D = \begin{pmatrix} 5 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

Suggested answer

2(a) The characteristic equation of the matrix is $(\lambda - 2)^2 = 0$. There is only one eigenvalue $\lambda = 2$. For eigenvalue $\lambda = 2$, the eigenspace is span$((1, 1)^T)$ which is of dimension 1. Therefore the matrix is not diagonalizable.

2(b) There are two eigenvalues 2 and 1. The algebraic multiplicity of eigenvalue $\lambda = 1$ is 2 but the geometric multiplicity of it is 1. Therefore the matrix is not diagonalizable.

2(c) There are two eigenvalues 2 and 1. The algebraic multiplicity of eigenvalue $\lambda = 2$ is 2 but the geometric multiplicity of it is 1. Therefore the matrix is not diagonalizable.

Power of matrices

Let $\mathbf{A}$ be an $n \times n$ matrix and $\mathbf{P}$ be a matrix diagonalizes $\mathbf{A}$, i.e.,

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$$

is a diagonal matrix. Then

$$\begin{aligned} \mathbf{A}^k &= \left(\mathbf{P}\mathbf{D}\mathbf{P}^{-1}\right)^k \\ &= \mathbf{P}\mathbf{D}^k\mathbf{P}^{-1} \end{aligned}$$

Example

Find $\mathbf{A}^5$ *if* $\mathbf{A} = \begin{pmatrix} 3 & 2 \\ 3 & -2 \end{pmatrix}$.

Solution: From Example 26,

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D} = \begin{pmatrix} 4 & 0 \\ 0 & -3 \end{pmatrix} \text{ where } \mathbf{P} = \begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix}.$$

Thus

$$\begin{aligned} \mathbf{A}^5 &= \mathbf{P}\mathbf{D}^5\mathbf{P}^{-1} \\ &= \begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 0 & -3 \end{pmatrix}^5 \begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} 1024 & 0 \\ 0 & -243 \end{pmatrix} \frac{1}{-7} \begin{pmatrix} -3 & -1 \\ -1 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 843 & 362 \\ 543 & -62 \end{pmatrix} \end{aligned}$$

Example

Find $\mathbf{A}^5$ *if* $\mathbf{A} = \begin{pmatrix} 4 & -2 & 1 \\ 2 & -2 & 3 \\ 2 & -2 & 3 \end{pmatrix}$.

Solution: Diagonalizing **A**, we have

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \text{ where } \mathbf{P} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix}.$$

Thus

$$\begin{aligned} \mathbf{A}^5 &= \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}^5 \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} 243 & 0 & 0 \\ 0 & 32 & 0 \\ 0 & 0 & 32 \end{pmatrix} \begin{pmatrix} 2 & -2 & 1 \\ -2 & 3 & -1 \\ -1 & 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 454 & -422 & 211 \\ 422 & -390 & 211 \\ 422 & -422 & 243 \end{pmatrix} \end{aligned}$$

Example

Consider a metropolitan area with a constant total population of 1 million individuals. This area consists of a city and its suburbs, and we want to analyze the changing urban and suburban populations. Let C_k denote the city population and S_k the suburban population after k years. Suppose that each year 15% of the people in the city move to the suburbs, whereas 10% of the people in the suburbs move to the city. Then it follows that

$$\begin{aligned} C_{k+1} &= 0.85C_k + 0.1S_k \\ S_{k+1} &= 0.15C_k + 0.9S_k \end{aligned}$$

Find the urban and suburban populations after a long time.

Solution: Let $\mathbf{x}_k = (C_k, S_k)^T$ be the population vector after k years. Then

$$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} = \mathbf{A}^2\mathbf{x}_{k-2} = \cdots = \mathbf{A}^k\mathbf{x}_0,$$

where

$$\mathbf{A} = \begin{pmatrix} 0.85 & 0.1 \\ 0.15 & 0.9 \end{pmatrix}.$$

Solving the characteristic equation, we have

$$\begin{aligned} \begin{vmatrix} 0.85-\lambda & 0.1 \\ 0.15 & 0.9-\lambda \end{vmatrix} &= 0 \\ \lambda^2 - 1.75\lambda + 0.75 &= 0 \\ \lambda &= 1, 0.75 \end{aligned}$$

Hence the eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = 0.75$. By solving $\mathbf{A} - \lambda\mathbf{I} = \mathbf{0}$, the associated eigenvectors are $\mathbf{v}_1 = (2,3)^T$ and $\mathbf{v}_2 = (-1,1)^T$ respectively.

Thus

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D} = \begin{pmatrix} 1 & 0 \\ 0 & 0.75 \end{pmatrix},$$

where

$$\mathbf{P} = \begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix}.$$

When k is very large

$$\begin{aligned}
\mathbf{A}^k &= \mathbf{P}\mathbf{D}^k\mathbf{P}^{-1} \\
&= \begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0.75 \end{pmatrix}^k \begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix}^{-1} \\
&= \begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0.75^k \end{pmatrix} \frac{1}{5} \begin{pmatrix} 1 & 1 \\ -3 & 2 \end{pmatrix} \\
&\simeq \frac{1}{5} \begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -3 & 2 \end{pmatrix} \\
&= \frac{1}{5} \begin{pmatrix} 2 & 3 \\ 2 & 3 \end{pmatrix}
\end{aligned}$$

Therefore

$$\begin{aligned}
\mathbf{x}_k &= \mathbf{A}^k \mathbf{x}_0 \\
&\simeq \begin{pmatrix} 0.4 & 0.6 \\ 0.4 & 0.6 \end{pmatrix} \begin{pmatrix} C_0 \\ S_0 \end{pmatrix} \\
&= (C_0 + S_0) \begin{pmatrix} 0.4 \\ 0.6 \end{pmatrix} \\
&= \begin{pmatrix} 0.4 \\ 0.6 \end{pmatrix}
\end{aligned}$$

That mean whatever the initial distribution of population is, the long-term distribution consists of 40% in the city and 60% in the suburbs. $\square$

An $n \times n$ matrix is called a **stochastic matrix** if it has nonnegative entries and the sum of the elements in each column is one. A Markov process is a stochastic process having the property that given the present state, future states are independent of the past states. A Markov process can be described by a **Morkov chain** which consists of a sequence of vectors $\mathbf{x}_k$, $k = 0, 1, 2, \cdots$, satisfying

$$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} = \mathbf{A}\mathbf{x}_0.$$

The vector $\mathbf{x}_k$ is called the **state vector** and $\mathbf{A}$ is called the **transition matrix**.

PageRank

The PageRank algorithm was developed by Larry Page and Sergey Brin and is used to rank the web sites in the Google search engine. It is a probability distribution used to represent the likelihood that a person randomly clicking links will arrive at any particular page. The linkage of the web sites in the web can be represented by a linkage matrix $\mathbf{A}$. The probability distribution of a person to arrive the web sites are given by $\mathbf{A}^k\mathbf{x}_0$ for a sufficiently large k and is independent of the initial distribution $\mathbf{x}_0$.

Example

Consider a small web consisting of three pages P, Q and R, where page P links to the pages Q and R, page Q links to page R and page R links to page P and Q. Assume that a person has a probability of 0.5 to stay on each page and the probability of going to other pages are evenly distributed to the pages which are linked to it. Find the page rank of the three web pages.

$$\begin{tikzcd}
R \arrow[d] \arrow[dr] & \\
P \arrow[u] \arrow[r] & Q \arrow[ul]
\end{tikzcd}$$

Solution: Let p_k, q_k and r_k be the number of people arrive the web pages P, Q and R respectively after k iteration. Then

$$\begin{pmatrix} p_k \\ q_k \\ r_k \end{pmatrix} = \begin{pmatrix} 0.5 & 0 & 0.25 \\ 0.25 & 0.5 & 0.25 \\ 0.25 & 0.5 & 0.5 \end{pmatrix} \begin{pmatrix} p_{k-1} \\ q_{k-1} \\ r_{k-1} \end{pmatrix}.$$

Thus

$$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} = \cdots = \mathbf{A}^k\mathbf{x}_0,$$

where

$$\mathbf{A} = \begin{pmatrix} 0.5 & 0 & 0.25 \\ 0.25 & 0.5 & 0.25 \\ 0.25 & 0.5 & 0.5 \end{pmatrix} \text{ and } \mathbf{x}_0 = \begin{pmatrix} p_0 \\ q_0 \\ r_0 \end{pmatrix}$$

are the linkage matrix and the initial state.

Solving the characteristic equation of $\mathbf{A}$, we have

$$\begin{aligned} \begin{vmatrix} 0.5-\lambda & 0 & 0.25 \\ 0.25 & 0.5-\lambda & 0.25 \\ 0.25 & 0.5 & 0.5-\lambda \end{vmatrix} &= 0 \\ \lambda^3 - 1.5\lambda^2 + 0.5625\lambda - 0.0625 &= 0 \\ \lambda &= 1 \text{ or } 0.25 \end{aligned}$$

For $\lambda_1 = 1$, we solve

$$\begin{pmatrix} -0.5 & 0 & 0.25 \\ 0.25 & -0.5 & 0.25 \\ 0.25 & 0.5 & -0.5 \end{pmatrix} \mathbf{v} = \mathbf{0}$$

and $\mathbf{v}_1 = (2, 3, 4)^T$ is an eigenvector of $\mathbf{A}$ associated with $\lambda_1 = 1$.
For $\lambda_2 = \lambda_3 = 0.25$, we solve

$$\begin{pmatrix} 0.25 & 0 & 0.25 \\ 0.25 & 0.25 & 0.25 \\ 0.25 & 0.5 & 0.25 \end{pmatrix} \mathbf{v} = \mathbf{0}$$

and there associates only one linearly independent eigenvector $\mathbf{v}_2 = (1, 0, -1)^T$.

Thus **A** is not diagonalizable. However we have

$$\mathbf{P} = \begin{pmatrix} 2 & 1 & 4 \\ 3 & 0 & -4 \\ 4 & -1 & 0 \end{pmatrix} \text{ where } \mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{J} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.25 & 1 \\ 0 & 0 & 0.25 \end{pmatrix}.$$

(**J** is called the Jordan normal form of **A**.) When k is sufficiently large, we have

$$\begin{aligned}
\mathbf{A}^k &= \mathbf{P}\mathbf{J}^k\mathbf{P}^{-1} \\
&= \begin{pmatrix} 2 & 1 & 4 \\ 3 & 0 & -4 \\ 4 & -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.25 & 1 \\ 0 & 0 & 0.25 \end{pmatrix}^k \begin{pmatrix} 2 & 1 & 4 \\ 3 & 0 & -4 \\ 4 & -1 & 0 \end{pmatrix}^{-1} \\
&\simeq \begin{pmatrix} 2 & 1 & 4 \\ 3 & 0 & -4 \\ 4 & -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1/9 & 1/9 & 1/9 \\ 4/9 & 4/9 & -5/9 \\ 1/12 & -1/6 & 1/12 \end{pmatrix} \\
&= \begin{pmatrix} 2/9 & 2/9 & 2/9 \\ 3/9 & 3/9 & 3/9 \\ 4/9 & 4/9 & 4/9 \end{pmatrix}
\end{aligned}$$

Thus after sufficiently many iteration, the number of people arrive the web pages are given by

$$\mathbf{A}^k\mathbf{x}_0 \simeq \begin{pmatrix} 2/9 & 2/9 & 2/9 \\ 3/9 & 3/9 & 3/9 \\ 4/9 & 4/9 & 4/9 \end{pmatrix} \begin{pmatrix} p_0 \\ q_0 \\ r_0 \end{pmatrix} = (p_0 + q_0 + r_0) \begin{pmatrix} 2/9 \\ 3/9 \\ 4/9 \end{pmatrix}.$$

Note that the ratio does not depend on the initial state. The PageRank of the web pages P, Q and R are $2/9$, $3/9$ and $4/9$ respectively. $\square$

Fibonacci sequence

The Fibonacci sequence is defined by

$$\begin{cases} F_{k+2} = F_{k+1} + F_k, & \text{for } k \geq 0 \\ F_0 = 0, F_1 = 1. \end{cases}$$

The recursive equation can be written as

$$\begin{pmatrix} F_{k+2} \\ F_{k+1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} F_{k+1} \\ F_k \end{pmatrix}.$$

If we let

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \text{ and } \mathbf{x}_k = \begin{pmatrix} F_{k+1} \\ F_k \end{pmatrix},$$

then

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k, & \text{for } k \geq 0 \\ \mathbf{x}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \end{cases}$$

It follows that

$$\mathbf{x}_k = A^k \mathbf{x}_0.$$

To find A^k, we can diagonalize A and obtain

$$\begin{pmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix}$$

Hence

$$\begin{aligned} A^k &= \begin{pmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^k & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^k \end{pmatrix} \begin{pmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{pmatrix}^{-1} \\ &= \frac{1}{\sqrt{5}} \begin{pmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^{k+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{k+1} & \left(\frac{1+\sqrt{5}}{2}\right)^{k} - \left(\frac{1-\sqrt{5}}{2}\right)^{k} \\ \left(\frac{1+\sqrt{5}}{2}\right)^{k} - \left(\frac{1-\sqrt{5}}{2}\right)^{k} & \left(\frac{1+\sqrt{5}}{2}\right)^{k-1} - \left(\frac{1-\sqrt{5}}{2}\right)^{k-1} \end{pmatrix} \end{aligned}$$

Now

$$\mathbf{x}_k = A^k \mathbf{x}_0 = \frac{1}{\sqrt{5}} \begin{pmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^{k+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{k+1} \\ \left(\frac{1+\sqrt{5}}{2}\right)^{k} - \left(\frac{1-\sqrt{5}}{2}\right)^{k} \end{pmatrix},$$

we have

$$F_k = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2}\right)^k - \left(\frac{1-\sqrt{5}}{2}\right)^k \right).$$

We also have

$$\lim_{k\to\infty} \frac{F_{k+1}}{F_k} = \lim_{k\to\infty} \frac{((1+\sqrt{5})/2)^{k+1} - ((1-\sqrt{5})/2)^{k+1}}{((1+\sqrt{5})/2)^{k} - ((1-\sqrt{5})/2)^{k}} = \frac{1+\sqrt{5}}{2}$$

which links the Fibonacci sequence with the celebrated golden ratio. $\square$

One of the most notable and important theorems in linear algebra is the Cayley-Hamilton Theorem.

Definition

Let $\mathbf{A}$ *be an* $n \times n$ *matrix. The* **minimal polynomial** *of* $\mathbf{A}$ *is a nonzero polynomial* $m(x)$ *of minimum degree with leading coefficient 1 satisfying* $m(\mathbf{A}) = \mathbf{0}$.

Theorem (Cayley-Hamilton Theorem)

Let $\mathbf{A}$ *be an* $n \times n$ *matrix and* $p(x) = \det(\mathbf{A} - x\mathbf{I})$ *be its characteristic polynomial. Then* $p(\mathbf{A}) = \mathbf{0}$. *In other words, the minimal polynomial of* $\mathbf{A}$ *always divides the characteristic polynomial of* $\mathbf{A}$.

Proof

Let $\mathbf{B} = \mathbf{A} - x\mathbf{I}$ and

$$p(x) = \det(\mathbf{B}) = c_n x^n + c_{n-1}x^{n-1} + \cdots + c_1 x + c_0$$

be the characteristic polynomial of $\mathbf{A}$. Consider $\mathbf{B} = \mathbf{A} - x\mathbf{I}$ as an $n \times n$ matrix whose entries are polynomial in x. Write the adjoint $\mathrm{adj}(\mathbf{B})$ of $\mathbf{B}$ as a polynomial of degree $n-1$ in x with matrix coefficients

$$\mathrm{adj}(\mathbf{B}) = \mathbf{B}_{n-1}x^{n-1} + \cdots + \mathbf{B}_1 x + \mathbf{B}_0$$

where the coefficients $\mathbf{B}_i$ are $n \times n$ constant matrices. On one hand, we have

$$\begin{aligned} \det(\mathbf{B})\mathbf{I} &= (c_n x^n + c_{n-1}x^{n-1} + \cdots + c_1 x + c_0)\mathbf{I} \\ &= (c_n\mathbf{I})x^n + (c_{n-1}\mathbf{I})x^{n-1} + \cdots + (c_1\mathbf{I})x + c_0\mathbf{I} \end{aligned}$$

On the other hand, we have

$$\begin{aligned} \mathbf{B}\mathrm{adj}(\mathbf{B}) &= (\mathbf{A} - x\mathbf{I})(\mathbf{B}_{n-1}x^{n-1} + \cdots + \mathbf{B}_1 x + \mathbf{B}_0) \\ &= -\mathbf{B}_{n-1}x^n + (\mathbf{A}\mathbf{B}_{n-1} - \mathbf{B}_{n-2})x^{n-1} + \cdots + (\mathbf{A}\mathbf{B}_1 - \mathbf{B}_0)x + \mathbf{A}\mathbf{B}_0 \end{aligned}$$

Proof.

Comparing the coefficients of

$$\det(\mathbf{B})\mathbf{I} = \mathbf{B}\,\mathrm{adj}(\mathbf{B}),$$

we get

$$\begin{aligned} c_n\mathbf{I} &= -\mathbf{B}_{n-1} \\ c_{n-1}\mathbf{I} &= \mathbf{A}\mathbf{B}_{n-1} - \mathbf{B}_{n-2} \\ &\vdots \\ c_1\mathbf{I} &= \mathbf{A}\mathbf{B}_1 - \mathbf{B}_0 \\ c_0\mathbf{I} &= \mathbf{A}\mathbf{B}_0 \end{aligned}$$

If we multiply the first equation by $\mathbf{A}^n$, second by $\mathbf{A}^{n-1}, \cdots$, the last by $\mathbf{I}$ and add up the resulting equations, we obtain

$$\begin{aligned} p(\mathbf{A}) &= c_n\mathbf{A}^n + c_{n-1}\mathbf{A}^{n-1} + \cdots + c_1\mathbf{A} + c_0\mathbf{I} \\ &= -\mathbf{A}^n\mathbf{B}_{n-1} + (\mathbf{A}^n\mathbf{B}_{n-1} - \mathbf{A}^{n-1}\mathbf{B}_{n-2}) + \cdots + (\mathbf{A}^2\mathbf{B}_1 - \mathbf{A}\mathbf{B}_0) + \mathbf{A}\mathbf{B}_0 \\ &= \mathbf{0} \end{aligned}$$

Example

Let

$$\mathbf{A} = \begin{pmatrix} 4 & -2 & 1 \\ 2 & -2 & 3 \\ 2 & -2 & 3 \end{pmatrix}.$$

The characteristic equation is

$$p(x) = -x^3 + 7x^2 - 16x + 12.$$

So

$$-\mathbf{A}^3 + 7\mathbf{A}^2 - 16\mathbf{A} + 12\mathbf{I} = \mathbf{0}.$$

Example

Now

$$\mathbf{A}^2 = \begin{pmatrix} 14 & -10 & 5 \\ 10 & -6 & 5 \\ 10 & -10 & 9 \end{pmatrix}.$$

We have

$$\begin{aligned} \mathbf{A}^3 &= 7\mathbf{A}^2 - 16\mathbf{A} + 12\mathbf{I} \\ &= 7\begin{pmatrix} 14 & -10 & 5 \\ 10 & -6 & 5 \\ 10 & -10 & 9 \end{pmatrix} - 16\begin{pmatrix} 4 & -2 & 1 \\ 2 & -2 & 3 \\ 2 & -2 & 3 \end{pmatrix} + 12\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 46 & -38 & 19 \\ 38 & -30 & 19 \\ 38 & -38 & 27 \end{pmatrix} \end{aligned}$$

Example

Multiplying by **A** *gives*

$$
\begin{aligned}
\mathbf{A}^4 &= 7\mathbf{A}^3 - 16\mathbf{A}^2 + 12\mathbf{A} \\
&= 7(7\mathbf{A}^2 - 16\mathbf{A} + 12) - 16\mathbf{A}^2 + 12\mathbf{A} \\
&= 33\mathbf{A}^2 - 100\mathbf{A} + 84\mathbf{I} \\
&= 33\begin{pmatrix} 14 & -10 & 5 \\ 10 & -6 & 5 \\ 10 & -10 & 9 \end{pmatrix} - 100\begin{pmatrix} 4 & -2 & 1 \\ 2 & -2 & 3 \\ 2 & -2 & 3 \end{pmatrix} + 84\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 146 & -130 & 65 \\ 130 & -114 & 65 \\ 130 & -130 & 81 \end{pmatrix}
\end{aligned}
$$

Example

Thus we can use Cayley-Hamilton theorem to express power of $\mathbf{A}$ *in terms of* $\mathbf{A}$ *and* $\mathbf{A}^2$*. We can also use Cayley-Hamilton theorem to find* $\mathbf{A}^{-1}$ *as follow*

$$\begin{aligned}
12\mathbf{I} &= \mathbf{A}^3 - 7\mathbf{A}^2 + 16\mathbf{A} \\
\mathbf{A}^{-1} &= \frac{1}{12}(\mathbf{A}^2 - 7\mathbf{A} + 16\mathbf{I}) \\
&= \frac{1}{12}\left(\begin{pmatrix} 14 & -10 & 5 \\ 10 & -6 & 5 \\ 10 & -10 & 9 \end{pmatrix} - 7\begin{pmatrix} 4 & -2 & 1 \\ 2 & -2 & 3 \\ 2 & -2 & 3 \end{pmatrix} + 16\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right) \\
&= \frac{1}{6}\begin{pmatrix} 1 & 2 & -1 \\ -2 & 5 & -1 \\ -2 & 2 & 2 \end{pmatrix}.
\end{aligned}$$

□

Example

Find the minimal polynomial of $\mathbf{A} = \begin{pmatrix} 2 & -3 & 1 \\ 1 & -2 & 1 \\ 1 & -3 & 2 \end{pmatrix}$.

Solution: The characteristic polynomial of $\mathbf{A}$ is

$$p(x) = -x(x-1)^2.$$

It can be proved that every root of the characteristic polynomial must also be a root of the minimal polynomial. Thus the minimal polynomial must be one of the following polynomials

$$x(x-1) \;\text{ or }\; x(x-1)^2.$$

By direct calculation

$$\mathbf{A}(\mathbf{A}-\mathbf{I}) = \begin{pmatrix} 2 & -3 & 1 \\ 1 & -2 & 1 \\ 1 & -3 & 2 \end{pmatrix} \begin{pmatrix} 1 & -3 & 1 \\ 1 & -3 & 1 \\ 1 & -3 & 1 \end{pmatrix} = \mathbf{0}.$$

Hence the minimal polynomial of $\mathbf{A}$ is $x(x-1)$.

Definition

Let $\mathbf{A}$ *be an* $n \times n$ *matrix. The* **matrix exponential** *of* $\mathbf{A}$ *is defined as*

$$\exp(\mathbf{A}) = \sum_{k=0}^{\infty} \frac{\mathbf{A}^k}{k!}.$$

Theorem (Properties of matrix exponential)

Let $\mathbf{A}$ *and* $\mathbf{B}$ *be two* $n \times n$ *matrix and* a, b *be any scalars. Then*

1. $\exp(\mathbf{0}) = \mathbf{I}$;
2. $\exp(-\mathbf{A}) = \exp(\mathbf{A})^{-1}$;
3. $\exp((a+b)\mathbf{A}) = \exp(a\mathbf{A})\exp(b\mathbf{A})$;
4. *If* $\mathbf{AB} = \mathbf{BA}$, *then* $\exp(\mathbf{A}+\mathbf{B}) = \exp(\mathbf{A})\exp(\mathbf{B})$;
5. *If* $\mathbf{A}$ *is nonsingular, then* $\exp(\mathbf{A}^{-1}\mathbf{BA}) = \mathbf{A}^{-1}\exp(\mathbf{B})\mathbf{A}$;
6. $\det(\exp(\mathbf{A})) = e^{\mathrm{tr}(\mathbf{A})}$. *(*$\mathrm{tr}(\mathbf{A}) = a_{11} + a_{22} + \cdots + a_{nn}$ *is the trace of* $\mathbf{A}$.*)*

Theorem

If $\mathbf{D} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$ *is a diagonal matrix, then*

$$\exp(\mathbf{D}) = \begin{pmatrix} e^{\lambda_1} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2} & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & e^{\lambda_n} \end{pmatrix}.$$

Moreover, for any $n \times n$ *matrix* $\mathbf{A}$*, if there exists nonsingular matrix* $\mathbf{P}$ *such that* $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D}$ *is a diagonal matrix. Then*

$$\exp(\mathbf{A}) = \mathbf{P}\exp(\mathbf{D})\mathbf{P}^{-1}.$$

Example

Find $\exp(\mathbf{A}t)$ *where* $\mathbf{A} = \begin{pmatrix} 4 & 2 \\ 3 & -1 \end{pmatrix}$.

Solution: Diagonalizing **A**, we have

$$\begin{pmatrix} 1 & 2 \\ -3 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 4 & 2 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 0 \\ 0 & 5 \end{pmatrix}.$$

Therefore

$$\begin{aligned} \exp(\mathbf{A}t) &= \begin{pmatrix} 1 & 2 \\ -3 & 1 \end{pmatrix} \begin{pmatrix} e^{-2t} & 0 \\ 0 & e^{5t} \end{pmatrix} \begin{pmatrix} 1 & 2 \\ -3 & 1 \end{pmatrix}^{-1} \\ &= \frac{1}{7} \begin{pmatrix} e^{-2t} + 6e^{5t} & -2e^{-2t} + 2e^{5t} \\ -3e^{-2t} + 3e^{5t} & 6e^{-2t} + e^{5t} \end{pmatrix} \end{aligned}$$

Example

Find $\exp(\mathbf{A}t)$ *where* $\mathbf{A} = \begin{pmatrix} 0 & 1 & 3 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$.

Solution: First compute

$$\mathbf{A}^2 = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{A}^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Therefore

$$\begin{aligned} \exp(\mathbf{A}t) &= \mathbf{I} + \mathbf{A}t + \frac{1}{2}\mathbf{A}^2 t^2 \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 3 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} t + \frac{1}{2} \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} t^2 \\ &= \begin{pmatrix} 1 & t & 3t + 2t^2 \\ 0 & 1 & 2t \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

Example

Find $\exp(\mathbf{A}t)$ *where* $\mathbf{A} = \begin{pmatrix} 4 & 1 \\ 0 & 4 \end{pmatrix}$.

Solution:

$$\begin{aligned}
\exp(\mathbf{A}t) &= \exp\begin{pmatrix} 4t & t \\ 0 & 4t \end{pmatrix} \\
&= \exp\left(\begin{pmatrix} 4t & 0 \\ 0 & 4t \end{pmatrix} + \begin{pmatrix} 0 & t \\ 0 & 0 \end{pmatrix}\right) \\
&= \exp\begin{pmatrix} 4t & 0 \\ 0 & 4t \end{pmatrix} \exp\begin{pmatrix} 0 & t \\ 0 & 0 \end{pmatrix} \\
& \quad \left(\text{since} \begin{pmatrix} 4t & 0 \\ 0 & 4t \end{pmatrix}\begin{pmatrix} 0 & t \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & t \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 4t & 0 \\ 0 & 4t \end{pmatrix}\right) \\
&= \begin{pmatrix} e^{4t} & 0 \\ 0 & e^{4t} \end{pmatrix}\left(\mathbf{I} + \begin{pmatrix} 0 & t \\ 0 & 0 \end{pmatrix}\right) \\
&= \begin{pmatrix} e^{4t} & te^{4t} \\ 0 & e^{4t} \end{pmatrix}
\end{aligned}$$

Example

Find $\exp(\mathbf{A}t)$ *where*

$$\mathbf{A} = \begin{pmatrix} 5 & -1 \\ 1 & 3 \end{pmatrix}.$$

Solution: Solving the characteristic equation

$$\begin{aligned} \begin{vmatrix} 5-\lambda & -1 \\ 1 & 3-\lambda \end{vmatrix} &= 0 \\ (5-\lambda)(3-\lambda)+1 &= 0 \\ \lambda^2 - 8\lambda + 16 &= 0 \\ (\lambda-4)^2 &= 0 \\ \lambda &= 4, 4 \end{aligned}$$

we see that $\mathbf{A}$ has only one eigenvalue $\lambda = 4$.

$$\begin{aligned} \det(\mathbf{A}-4\mathbf{I})\mathbf{v} &= \mathbf{0} \\ \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}\mathbf{v} &= \mathbf{0} \end{aligned}$$

we can find only one linearly independent eigenvector $\mathbf{v} = (1,1)^T$. Thus $\mathbf{A}$ is not diagonalizable. To find $\exp(\mathbf{A}t)$, we may take

$$\mathbf{T} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

and let ($\mathbf{J}$ is called the Jordan normal form of $\mathbf{A}$.)

$$\begin{aligned} \mathbf{J} &= \mathbf{T}^{-1}\mathbf{A}\mathbf{T} \\ &= \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 5 & -1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 5 & -1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 4 & 1 \\ 0 & 4 \end{pmatrix} \end{aligned}$$

Consider

$$\begin{aligned} \det(\mathbf{A} - 4\mathbf{I})\mathbf{v} &= \mathbf{0} \\ \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \mathbf{v} &= \mathbf{0} \end{aligned}$$

we can find only one linearly independent eigenvector $\mathbf{v} = (1, 1)^T$. Thus $\mathbf{A}$ is not diagonalizable. To find $\exp(\mathbf{A}t)$, we may use the so called generalized eigenvector. A vector $\mathbf{v}$ is called a generalized eigenvector of rank 2 associated with eigenvalue λ if it satisfies

$$\begin{cases} (\mathbf{A} - \lambda\mathbf{I})\mathbf{v} \neq \mathbf{0}, \\ (\mathbf{A} - \lambda\mathbf{I})^2\mathbf{v} = \mathbf{0}. \end{cases}$$

Now if we take

$$\begin{cases} \mathbf{v}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\ \mathbf{v}_1 = (\mathbf{A} - 4\mathbf{I})\mathbf{v}_0 = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \end{cases}$$

then

$$\begin{cases} (\mathbf{A} - 4\mathbf{I})\mathbf{v}_0 = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \neq \mathbf{0}, \\ (\mathbf{A} - 4\mathbf{I})^2\mathbf{v}_0 = (\mathbf{A} - 4\mathbf{I})\mathbf{v}_1 = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \mathbf{0}. \end{cases}$$

So $\mathbf{v}_0$ is a generalized eigenvector of rank 2 associated with eigenvalue $\lambda = 4$.

We may let

$$\mathbf{T} = [\mathbf{v}_1\ \mathbf{v}_0] = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix},$$

and

$$\begin{aligned}
\mathbf{J} &= \mathbf{T}^{-1}\mathbf{A}\mathbf{T} \\
&= \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 5 & -1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 5 & -1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 4 & 1 \\ 0 & 4 \end{pmatrix}
\end{aligned}$$

($\mathbf{J}$ is called the Jordan normal form of $\mathbf{A}$.)

Then

$$\begin{aligned}
& \exp(\mathbf{A}t) \\
= \; & \exp(\mathbf{TJT}^{-1}t) \\
= \; & \mathbf{T}\exp(\mathbf{J}t)\mathbf{T}^{-1} \\
= \; & \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} e^{4t} & te^{4t} \\ 0 & e^{4t} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} \text{ (By above Example)} \\
= \; & \begin{pmatrix} e^{4t} + te^{4t} & -te^{4t} \\ te^{4t} & e^{4t} - te^{4t} \end{pmatrix}
\end{aligned}$$

$\square$

Exercise

Find $\exp(\mathbf{A}t)$ *for each of the following matrices* $\mathbf{A}$.

1. $\begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}$
2. $\begin{pmatrix} 4 & -1 \\ 1 & 2 \end{pmatrix}$
3. $\begin{pmatrix} 0 & -4 & 1 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix}$
4. $\begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix}$

Answer

1. $\frac{1}{7}\begin{pmatrix} 3e^{5t}+4e^{-2t} & 3e^{5t}-3e^{-2t} \\ 4e^{5t}-4e^{-2t} & 4e^{5t}+3e^{-2t} \end{pmatrix}$
2. $e^{2t}\begin{pmatrix} 1+t & t \\ -t & 1-t \end{pmatrix}$
3. $\begin{pmatrix} 1 & -t & t-6t^2 \\ 0 & 1 & 3t \\ 0 & 0 & 1 \end{pmatrix}$
4. $\begin{pmatrix} e^{3t} & te^{3t} & \frac{1}{2}t^2e^{3t} \\ 0 & e^{3t} & te^{3t} \\ 0 & 0 & e^{3t} \end{pmatrix}$