Vv417 Lecture 23

Jing Liu

UM-SJTU Joint Institute

November 21, 2019

- Many applications involve applying a linear transformation repeatedly. e.g.
- Consider a small population of 800 married and 200 single female. Suppose

  30 percent of the married women get divorced each year
  and 20 percent of the single women get married each year,

  and we assume the total population remains constant.
- Imagine you want to investigate, based on the given divorce rate, the long run prospects of everlasting love before getting married.

Q: Is 30% the figure that you shall use?

- For simplicity we assume the pattern continue indefinitely into the future.
- Let $\mathbf{w}_0$ denote the vector at time zero, that is, $\mathbf{w}_0 = \begin{bmatrix} 800 \\ 200 \end{bmatrix}$.

Q How to find $\mathbf{w}_1$, the number of married and single women after one year?

- If we set up the matrix containing the percentages,

$$\mathbf{A} = \begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix}$$

- then the number of married and single women after one year is given by

$$\mathbf{w}_1 = \mathbf{A}\mathbf{w}_0 = \begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix} \begin{bmatrix} 800 \\ 200 \end{bmatrix} = \begin{bmatrix} 600 \\ 400 \end{bmatrix}$$

- To determine the number of married and single women after two years,

$$\mathbf{w}_2 = \mathbf{A}\mathbf{w}_1 = \mathbf{A}^2\mathbf{w}_0$$

  and, in general, for n years, we must compute $\mathbf{w}_n = \mathbf{A}^n\mathbf{w}_0$.

- Let us compute $\mathbf{w}_4$, $\mathbf{w}_8$, $\mathbf{w}_{16}$ in this way:

$$\mathbf{w}_4 = \begin{bmatrix} 425 \\ 575 \end{bmatrix}, \qquad \mathbf{w}_8 = \begin{bmatrix} 402 \\ 599 \end{bmatrix}, \qquad \mathbf{w}_{16} = \begin{bmatrix} 400 \\ 600 \end{bmatrix},$$

  where the entries of each have rounded to the nearest integer.

- After a certain point, we seem to always get the same answer. In fact,

$$\mathbf{w}_{20} = \begin{bmatrix} 400 \\ 600 \end{bmatrix} \quad \text{and since} \quad \mathbf{A}\mathbf{w}_{20} = \begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix} \begin{bmatrix} 400 \\ 600 \end{bmatrix} = \begin{bmatrix} 400 \\ 600 \end{bmatrix}$$

- It follows that all the succeeding vectors in the sequence remain unchanged.
- The vector $\mathbf{w} = 400\mathbf{e}_1 + 600\mathbf{e}_2$ is said to be the steady-state for the process.

Q: Do you think this steady-state depends on the initial point $\mathbf{w}_0$?

- For example, if we had started with

$$\text{1000 married women and 0 single women, then } \mathbf{w}_0 = \begin{bmatrix} 1000 \\ 0 \end{bmatrix}$$

- We can compute $\mathbf{w}_n$ as before by multiplying $\mathbf{w}_0$ by $\mathbf{A}^n$. In this case, it turns out that $\mathbf{w}_{21} = 400\mathbf{e}_1 + 600\mathbf{e}_2$, and so somehow we still end up with

the same steady-state vector.

Q: Why does this process converge, and why do we seem to get the same steady-state vector even when we change the initial vector?

- These questions are not difficult to answer if we choose a basis consisting of vectors for which the effect of the linear operator $\mathbf{A}$ is easily determined.
- In particular, if we choose a multiple of the steady-state vector, $\mathbf{x}_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, as our first basis vector, then

$$\mathbf{A}\mathbf{x}_1 = \begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

- However, if we choose $\mathbf{x}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, then the effect of $\mathbf{A}$ on $\mathbf{x}_2$ is also simple:

$$\mathbf{A}\mathbf{x}_2 = \begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix} = \frac{1}{2}\mathbf{x}_2$$

- Let us now analyse the process, using $\mathbf{x}_1$ and $\mathbf{x}_2$ as our basis vectors.

- If we express $\mathbf{w}_0 = 8000\mathbf{e}_1 + 2000\mathbf{e}_2$ as a linear combination of

$$\mathbf{x}_1 \text{ and } \mathbf{x}_2, \text{ then } \mathbf{w}_0 = 200\begin{bmatrix}2\\3\end{bmatrix} - 400\begin{bmatrix}-1\\1\end{bmatrix} = 200\mathbf{x}_1 - 400\mathbf{x}_2$$

- and it follows that

$$\mathbf{w}_1 = \mathbf{A}\mathbf{w}_0 = 200\mathbf{A}\mathbf{x}_1 - 400\mathbf{A}\mathbf{x}_2 = 200\mathbf{x}_1 - 400\left(\frac{1}{2}\right)\mathbf{x}_2$$

$$\mathbf{w}_2 = \mathbf{A}\mathbf{w}_1 = 200\mathbf{x}_1 - 400\left(\frac{1}{2}\right)^2\mathbf{x}_2$$

- In general,

$$\mathbf{w}_n = \mathbf{A}^n\mathbf{w}_0 = 200\mathbf{x}_1 - 400\left(\frac{1}{2}\right)^n\mathbf{x}_2$$

- The first component of this sum is the steady-state vector
- The second component converges to the zero as $n \to 0$.

- The vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ were natural vectors to use in analyzing the process, for the effect of the matrix $\mathbf{A}$ on each of these vectors was so simple:

$$\mathbf{A}\mathbf{x}_1 = \mathbf{x}_1 = 1\mathbf{x}_1 \quad \text{and} \quad \mathbf{A}\mathbf{x}_2 = \frac{1}{2}\mathbf{x}_2$$

- For each of the two vectors, multiplying $\mathbf{A}$ is a simple scalar multiplication.
- In general, if a linear transformation is represented by an $n \times n$ matrix $\mathbf{A}$ and we can find a nonzero vector $\mathbf{x}$ so that $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$, for some scalar λ, then, for this transformation, $\mathbf{x}$ is a natural choice to use as a basis vector for $\mathbb{R}^n$.
- More precisely, we use the following terminology to refer to $\mathbf{x}$ and λ:

Definition

Suppose $\mathbf{A}$ is an $n \times n$ matrix. A scalar λ is said to be an eigenvalue of $\mathbf{A}$ if there exists a nonzero vector $\mathbf{x}$ such that

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$$

The vector $\mathbf{x}$ is said to be an eigenvector corresponding to λ.

- For example,

$$\mathbf{A} = \begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

- Since

$$\mathbf{A}\mathbf{x} = \begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \end{bmatrix} = 3 \begin{bmatrix} 2 \\ 1 \end{bmatrix} = 3\mathbf{x}$$

- It follows that $\lambda = 3$ is an eigenvalue of $\mathbf{A}$, and $\mathbf{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ is an eigenvector to

$$\lambda = 3$$

- Actually, any nonzero multiple of $\mathbf{x}$ will be a corresponding eigenvector, since

$$\mathbf{A}(\alpha\mathbf{x}) = \alpha\mathbf{A}\mathbf{x} = \alpha\lambda\mathbf{x} = \lambda(\alpha\mathbf{x})$$

- For example, $\begin{bmatrix} 4 \\ 2 \end{bmatrix}$ is also an eigenvector belonging to $\lambda = 3$:

$$\begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 12 \\ 6 \end{bmatrix} = 3 \begin{bmatrix} 4 \\ 2 \end{bmatrix}$$

- An eigenvalue problem is about finding the eigenvalue λ and the corresponding eigenvectors that satisfy the equation

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$$

- The equation can be written in the form

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

- Since we are looking for nonzero $\mathbf{x}$, λ is an eigenvalue of $\mathbf{A}$ if and only if

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

  has a nontrivial solution, that is

$$\mathrm{null}(\mathbf{A} - \lambda\mathbf{I}) \neq \{\mathbf{0}\}$$

- Any nonzero vector in $\mathrm{null}(\mathbf{A} - \lambda\mathbf{I})$ is an eigenvector belonging to λ.
- The subspace $\mathrm{null}(\mathbf{A} - \lambda\mathbf{I})$ is called the eigenspace corresponding to λ.

Q: How can we be sure that the equation

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

has a nontrivial solution?

- It is nontrivial solution if and only if

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

Definition

Suppose $\mathbf{A}$ is $n \times n$ matrix. The nth degree polynomial of λ

$$p(\lambda) = \det(\mathbf{A} - \lambda \mathbf{I})$$

is called the characteristic polynomial for the matrix $\mathbf{A}$. The equation

$$p(\lambda) = 0$$

is known as the characteristic equation for the matrix $\mathbf{A}$.

Exercise

Find the eigenvalues and the corresponding eigenvectors of the matrix

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

Solution

1. Find the characteristic equation, and solve it to find all eigenvalues,

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0 \implies \det\begin{bmatrix} 3-\lambda & 2 \\ 3 & -2-\lambda \end{bmatrix} \implies \lambda^2 - \lambda - 12 = 0$$

   Thus the eigenvalues of $\mathbf{A}$ are $\lambda_1 = 4$ and $\lambda_2 = -3$.

2. To find the eigenvectors, we must solve the two linear systems

$$(\mathbf{A} - \lambda_1\mathbf{I})\,\mathbf{x} = \mathbf{0}, \qquad \text{and} \qquad (\mathbf{A} - \lambda_2\mathbf{I})\,\mathbf{x} = \mathbf{0}$$

Solution

- For $\lambda_1 = 4$,

$$\mathbf{A} - 4\mathbf{I} = \begin{bmatrix} -1 & 2 \\ 3 & -6 \end{bmatrix}$$

- Find a basis for the eigenspace

$$\begin{aligned} -x_1 + 2x_2 &= 0 \\ 3x_1 - 6x_2 &= 0 \end{aligned} \implies x_1 = 2x_2$$

- The eigenspace is spanned by

$$\mathbf{x}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix},$$

- So any nonzero multiple of $\mathbf{x}_1$ is an eigenvector corresponding to λ_1.

- For $\lambda_2 = -3$,

$$\mathbf{A} + 3\mathbf{I} = \begin{bmatrix} 6 & 2 \\ 3 & 1 \end{bmatrix}$$

- Find a basis for the eigenspace

$$\begin{aligned} 6x_1 + 2x_2 &= 0 \\ 3x_1 + 1x_2 &= 0 \end{aligned} \implies -3x_1 = x_2$$

- The eigenspace is spanned by

$$\mathbf{x}_2 = \begin{bmatrix} -1 \\ 3 \end{bmatrix},$$

- So any nonzero multiple of $\mathbf{x}_2$ is an eigenvector corresponding to λ_2.

Exercise

Find the eigenspaces of $\mathbf{A} = \begin{bmatrix} 2 & -3 & 1 \\ 1 & -2 & 1 \\ 1 & -3 & 2 \end{bmatrix}$.

Solution

1. Construct the characteristic equation, and solve it to find all eigenvalues,

$$\det \begin{bmatrix} 2-\lambda & -3 & 1 \\ 1 & -2-\lambda & 1 \\ 1 & -3 & 2-\lambda \end{bmatrix} = -\lambda(\lambda-1)^2 \implies \lambda_1 = 0, \quad \lambda_2 = \lambda_3 = 1$$

2. To find the eigenvectors, we must solve the two linear systems

$$(\mathbf{A} - \lambda_1 \mathbf{I})\,\mathbf{x} = \mathbf{0}, \qquad \text{and} \qquad (\mathbf{A} - \lambda_2 \mathbf{I})\,\mathbf{x} = \mathbf{0}$$

- The eigenspace corresponding to $\lambda = 0$ is given by $x_1 = x_2 = x_3$
- So the eigenspace corresponding to $\lambda_1 = 0$ is the span of $\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3$.

Solution

- To find the eigenspace corresponding to $\lambda = 1$, we must solve $(\mathbf{A} - \mathbf{I})\mathbf{x} = \mathbf{0}$,

$$\begin{bmatrix} 1 & -3 & 1 & 0 \\ 1 & -3 & 1 & 0 \\ 1 & -3 & 1 & 0 \end{bmatrix} \mathbf{x} = \mathbf{0} \implies \begin{aligned} x_1 - 3x_2 + x_3 &= 0 \\ x_1 - 3x_2 + x_3 &= 0 \\ x_1 - 3x_2 + x_3 &= 0 \end{aligned} \implies \begin{aligned} x_1 - 3x_2 + x_3 &= 0 \\ 0 &= 0 \\ 0 &= 0 \end{aligned}$$

- Let $x_2 = \alpha$ and $x_3 = \beta$, $x_1 = 3\alpha - \beta$, then $\begin{bmatrix} 3\alpha - \beta \\ \alpha \\ \beta \end{bmatrix} = \alpha \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$.
- Thus, the eigenspace corresponding to $\lambda = 1$ is

$$\text{span}\{ \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \}$$

- Hence, we have three eigenvalues and three eigenvectors for a 3×3 matrix.

Definition

The degree of a root (eigenvalue) of the characteristic polynomial of a matrix, that is, the number of times the root is repeated, is called the

algebraic multiplicity of the eigenvalue.

The dimension of the eigenspace corresponding to a given λ, that is the number of linearly independent eigenvectors corresponding to the eigenvalue, is called the

geometric multiplicity of the eigenvalue.

- Consider the matrix $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, which has the characteristic polynomial of

$$(1-\lambda)^2 \implies \lambda_1 = \lambda_2 = 1$$

which means we need a basis for $\operatorname{null}(\mathbf{A} - 1\mathbf{I})$

$$\implies \begin{aligned} 0x_1 + x_2 &= 0 \\ 0x_1 + 0x_2 &= 0 \end{aligned} \implies x_2 = 0, \quad \text{hence eigenspace is } \operatorname{span}\{\mathbf{e}_1\}$$

- So the geometric multiplicity is 1 but the algebraic multiplicity is 2.

Exercise

Find the eigenvalues of the matrix $\mathbf{A} = \begin{bmatrix} -2 & -1 \\ 5 & 2 \end{bmatrix}$.

Solution

- Find the characteristic equation, and solve it to find all eigenvalues,

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0 \implies \det \begin{bmatrix} -2-\lambda & -1 \\ 5 & 2-\lambda \end{bmatrix} = 0$$

$$\implies \lambda^2 + 1 = 0 \implies \lambda_1 = \mathrm{i}, \quad \text{and} \quad \lambda_2 = -\mathrm{i}$$

- It shall be not a surprise to you that the characteristic equation of a matrix

$$\mathbf{A} \quad \text{with real entries}$$

can have complex solutions.
- To allow complex eigenvalues, it is necessary to allow our scalars to be complex, and consider complex entries for vectors and matrices.

Defintiion

A complex vector space is one in which the scalars are complex numbers.

- We will be concerned only with the complex generalization of $\mathbb{R}^n$.

$$\mathbb{C}^n$$

Q: Is $\mathbb{R}^n$ a subspace of $\mathbb{C}^n$?

- Every vector $\mathbf{v}$ in $\mathbb{C}^n$ can be split into real and imaginary parts as

$$\mathbf{v} = \mathrm{Re}(\mathbf{v}) + \mathrm{i}\,\mathrm{Im}(\mathbf{v})$$

  where $\mathrm{Re}(\mathbf{v})$ and $\mathrm{Im}(\mathbf{v})$ are vectors in $\mathbb{R}^n$.

- The vector

$$\overline{\mathbf{v}} = \mathrm{Re}(\mathbf{v}) - \mathrm{i}\,\mathrm{Im}(\mathbf{v})$$

  is called the complex conjugate of $\mathbf{v}$.

- It follows that the vectors in $\mathbb{R}^n$ can be viewed as those vectors in $\mathbb{C}^n$ whose imaginary part is zero; a vector $\mathbf{v}$ in $\mathbb{C}^n$ is in $\mathbb{R}^n$ if and only if

$$\overline{\mathbf{v}} = \mathbf{v}$$

- For this part, we will need to distinguish between matrices whose entries must be real numbers, called real matrices, and matrices whose entries may be either real numbers or complex numbers, called complex matrices.
- If $\mathbf{A}$ is a complex matrix, then

$$\mathrm{Re}(\mathbf{A}) \quad \text{and} \quad \mathrm{Im}(\mathbf{A})$$

are the matrices formed from the real and imaginary parts of the entries of $\mathbf{A}$
- $\overline{\mathbf{A}}$ is the matrix formed by taking the complex conjugate of each entry in $\mathbf{A}$,

$$\overline{\mathbf{A}} = \mathrm{Re}(\mathbf{A}) - \mathrm{i}\,\mathrm{Im}(\mathbf{A})$$

- The standard operations on real matrices carry over without change to complex matrices, and all of the familiar properties continue to be true.

Theorem

- Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in $\mathbb{C}^n$, and if α is a scalar, then:

$$1. \quad \overline{\overline{\mathbf{u}}} = \mathbf{u} \qquad 2. \quad \overline{\alpha\mathbf{u}} = \overline{\alpha}\,\overline{\mathbf{u}} \qquad 3. \quad \overline{\mathbf{u} \pm \mathbf{v}} = \overline{\mathbf{u}} \pm \overline{\mathbf{v}}$$

- Let $\mathbf{A}$ be an $m \times k$ complex matrix and $\mathbf{B}$ be a $k \times n$ complex matrix, then:

$$1. \quad \overline{\overline{\mathbf{A}}} = \mathbf{A} \qquad 2. \quad \overline{(\mathbf{A}^{\mathrm{T}})} = (\overline{\mathbf{A}})^{\mathrm{T}} \qquad 3. \quad \overline{\mathbf{AB}} = \overline{\mathbf{A}}\,\overline{\mathbf{B}}$$

Proof

- To show $\overline{\mathbf{AB}} = \overline{\mathbf{A}}\,\overline{\mathbf{B}}$, we consider the ijth element of the left-hand side

$$\left[\overline{\mathbf{AB}}\right]_{ij} = \overline{[\mathbf{AB}]_{ij}} = \overline{\sum_{p}^{k} [\mathbf{A}]_{ip} [\mathbf{B}]_{pj}} = \sum_{p}^{k} \overline{[\mathbf{A}]_{ip}}\,\overline{[\mathbf{B}]_{pj}}$$

$$= \sum_{p}^{k} \left[\overline{\mathbf{A}}\right]_{ip} \left[\overline{\mathbf{B}}\right]_{pj} = \left[\overline{\mathbf{A}}\,\overline{\mathbf{B}}\right]_{ij}$$

- Except for the use of complex scalars, the notions of linear combination, linear independence, subspace, spanning, basis, dimension, determinant, inverse, can be extended to complex vectors and matrices without much modification.

Exercise

Suppose

$$\mathbf{v} = \begin{bmatrix} 3+\mathrm{i} \\ -2\mathrm{i} \end{bmatrix} \quad \text{and} \quad \mathbf{A} = \begin{bmatrix} 1+\mathrm{i} & -\mathrm{i} \\ 4 & 6-2\mathrm{i} \end{bmatrix}$$

Find $\det(\mathbf{A})$ and determine whether $\mathbf{Av}$ and $\mathbf{v}$ are linearly independent.

Solution

- Determinant

$$\det(\mathbf{A}) = (1+\mathrm{i})(6-2\mathrm{i}) + 4\mathrm{i} = 8 + 8\mathrm{i}$$

- Compute $\mathbf{Av} = (3+\mathrm{i})\begin{bmatrix} 1+\mathrm{i} \\ 4 \end{bmatrix} + (-2\mathrm{i})\begin{bmatrix} -\mathrm{i} \\ 6-2\mathrm{i} \end{bmatrix} = \begin{bmatrix} 4\mathrm{i} \\ 8-8\mathrm{i} \end{bmatrix}$

Solution

- As before, we need to determine whether there is non-zero scalar α_1 and α_2

$$\alpha_1 \begin{bmatrix} 3+\mathrm{i} \\ -2\mathrm{i} \end{bmatrix} + \alpha_2 \begin{bmatrix} 4\mathrm{i} \\ 8-8\mathrm{i} \end{bmatrix} = \mathbf{0}$$

- Note the scalar field is $\mathbb{C}$, so α_1 and α_2 need not be real, and $\mathbf{0} \in \mathbb{C}^2$.
- However, the equivalence theorem is applicable, so we can simply consider

$$\underbrace{\begin{bmatrix} 3+\mathrm{i} & 4\mathrm{i} \\ -2\mathrm{i} & 8-8\mathrm{i} \end{bmatrix}}_{\mathbf{B}} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

  and conclude $\mathbf{Av}$ and $\mathbf{v}$ are linearly independent if and only if $\det(\mathbf{B}) \neq 0$

$$\det(\mathbf{B}) = (3+\mathrm{i})(8-8\mathrm{i}) - 8 = 24 - 16\mathrm{i} \neq 0$$

- Therefore $\mathbf{v}$ and $\mathbf{Av}$ are linearly independent.

- As in the real case, λ is a complex eigenvalue of $\mathbf{A}$ if and only if

  there exists a nonzero vector $\mathbf{x}$ in $\mathbb{C}^n$ such that $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$.

- Each such $\mathbf{x}$ in $\mathbb{C}^n$ is called a complex eigenvector of $\mathbf{A}$ corresponding to λ.

Exercise

Find the eigenspace of the matrix corresponding to the eigenvalue $\lambda = \mathrm{i}$.

$$\mathbf{A} = \begin{bmatrix} -2 & -1 \\ 5 & 2 \end{bmatrix}$$

Solution

- Compute the matrix

$$\mathbf{A} - \lambda\mathbf{I} = \begin{bmatrix} -2 - \mathrm{i} & -1 \\ 5 & 2 - \mathrm{i} \end{bmatrix}$$

- We know $\mathbf{A} - \lambda\mathbf{I}$ must be singular, thus we only need one of the two rows.

Solution

- Let us use the second equation,

$$(\mathbf{A} - \lambda \mathbf{I})\,\mathbf{x} = \begin{bmatrix} -2-\mathrm{i} & -1 \\ 5 & 2-\mathrm{i} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\implies 5x_1 + (2-i)x_2 = 0$$

$$\implies \mathbf{x} = \alpha \begin{bmatrix} 2-\mathrm{i} \\ -5 \end{bmatrix} \qquad \text{for any } \alpha \neq 0.$$

- Thus the eigenspace corresponding to $\lambda = \mathrm{i}$ is the span of

$$\begin{bmatrix} 2-\mathrm{i} \\ -5 \end{bmatrix}$$

Theorem

If λ is an eigenvalue of a real $n \times n$ matrix $\mathbf{A}$, and if $\mathbf{x}$ is a eigenvector belonging to λ, then $\overline{\lambda}$ is also an eigenvalue of $\mathbf{A}$, and $\overline{\mathbf{x}}$ is a corresponding eigenvector.

Proof

- Since λ is an eigenvalue of $\mathbf{A}$ and $\mathbf{x}$ is a corresponding eigenvector, we have

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x} \implies \overline{\mathbf{A}\mathbf{x}} = \overline{\lambda\mathbf{x}} \implies \overline{\mathbf{A}}\overline{\mathbf{x}} = \overline{\lambda}\overline{\mathbf{x}}$$

- However, $\mathbf{A} = \overline{\mathbf{A}}$, since $\mathbf{A}$ is real, we have

$$\mathbf{A}\overline{\mathbf{x}} = \overline{\lambda}\overline{\mathbf{x}}$$

which shows $\overline{\lambda}$ is an eigenvalue of $\mathbf{A}$, and $\overline{\mathbf{x}}$ is a corresponding eigenvector.

- The correct choice of the basis often can simplify a particular problem.
- For a linear operator $L\colon \mathcal{V} \to \mathcal{V}$, the goal is to find a basis $\mathcal{B}_{\mathcal{V}}$ such that

$$[L]_{\mathcal{B}_{\mathcal{V}}}$$

the transformation matrix with respect to $\mathcal{B}_{\mathcal{V}}$ is as simple as possible.

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$$

- Recall the relationship between coordinate vectors with respect to two bases

$$[\mathbf{v}]_{\mathcal{A}} = \mathbf{P}_{\mathcal{B}\to\mathcal{A}}\,[\mathbf{v}]_{\mathcal{B}}$$

where $\mathbf{P}_{\mathcal{B}\to\mathcal{A}}$ is the transition matrix from basis $\mathcal{B}$ to basis $\mathcal{A}$.

Theorem

Let $\mathcal{V}$ be a vector space, and $L\colon \mathcal{V} \to \mathcal{V}$, and $\mathcal{A}$ and $\mathcal{B}$ be bases of $\mathcal{V}$, then

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

where $\mathbf{P} = \mathbf{P}_{\mathcal{B}\to\mathcal{A}}$, $\mathbf{A} = [L]_{\mathcal{A}}$ and $\mathbf{B} = [L]_{\mathcal{B}}$

Proof

- Let $\mathbf{x} = [\mathbf{v}]_{\mathcal{B}}$ be the coordinate vector of $\mathbf{v} \in \mathcal{V}$ with respect to $\mathcal{B}$ and

$$\mathbf{y} = \mathbf{P}\mathbf{x}, \qquad \text{where } \mathbf{P}_{\mathcal{B}\to\mathcal{A}}$$

  then $\mathbf{y} = [\mathbf{v}]_{\mathcal{A}}$. Now suppose $\mathbf{t} = \mathbf{A}\mathbf{y}$ and $\mathbf{z} = \mathbf{B}\mathbf{x}$, then

$$\mathbf{t} = [L(\mathbf{v})]_{\mathcal{A}}, \qquad \text{and} \qquad \mathbf{z} = [L(\mathbf{v})]_{\mathcal{B}}$$

- The transition matrix from $\mathcal{A}$ to $\mathcal{B}$ is $\mathbf{P}^{-1}$, so $\mathbf{P}^{-1}\mathbf{t} = \mathbf{z}$, which implies

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{y} = \mathbf{B}\mathbf{x} \implies \mathbf{P}^{-1}\mathbf{A}\mathbf{P}\mathbf{x} = \mathbf{B}\mathbf{x} \implies (\mathbf{P}^{-1}\mathbf{A}\mathbf{P} - \mathbf{B})\mathbf{x} = \mathbf{0}$$

- This is true for every $\mathbf{x} \in \mathbb{R}^n$, so the column space must be trivial, thus

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

$$\begin{tikzcd}
{[\mathbf{v}]_{\mathcal{A}}} \arrow[r, "\mathbf{A}"] & {[L(\mathbf{v})]_{\mathcal{A}}} \arrow[d, "\mathbf{P}^{-1}"] \\
{[\mathbf{v}]_{\mathcal{B}}} \arrow[u, "\mathbf{P}"] \arrow[r, "\mathbf{B}"'] & {[L(\mathbf{v})]_{\mathcal{B}}}
\end{tikzcd}$$

Definition

Suppose $\mathbf{A}$ and $\mathbf{B}$ are $n \times n$ matrices. $\mathbf{B}$ is said to be similar to $\mathbf{A}$ if there exists a nonsingular matrix $\mathbf{P}$ such that $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. The transformation

$$T\colon \mathbf{A} \mapsto \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$$

is called a similarity transformation on $\mathbf{A}$.

- Note that if $\mathbf{B}$ is similar to $\mathbf{A}$, then $\mathbf{A}$ is similar to $\mathbf{B}$. Thus, we may simply say that $\mathbf{A}$ and $\mathbf{B}$ are similar matrices.

Exercise

Let D be the derivative operator on $\mathcal{P}_2$. Find the matrix representation of D w.r.t

$$\mathcal{S} = \{1, x, x^2\}$$

and the matrix representation of D with respect to

$$\mathcal{B} = \{1, 2x, 4x^2 - 2\}$$

Find the eigenvalues and eigenvectors of the two matrices.

Solution

- Find the images of $\mathcal{S}$ under D with respect to $\mathcal{S}$.

$$\begin{aligned} D(1) &= 0 \\ &= 0\cdot 1 + 0\cdot x + 0\cdot x^2 \\ D(x) &= 1 \\ &= 1\cdot 1 + 0\cdot x + 0\cdot x^2 \\ D(x^2) &= 2x \\ &= 0\cdot 1 + 2\cdot x + 0\cdot x^2 \end{aligned} \implies [D]_{\mathcal{S}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \implies \lambda_{1,2,3} = 0$$

- Similarly, we can find $[D]_{\mathcal{B}} = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{bmatrix}$, and the eigenvalues are again

$$\lambda_1 = \lambda_2 = \lambda_3 = 0$$

- Since all eigenvalues are zero, the eigenspace is the null space, so $\mathrm{span}(\mathbf{e}_1)$.

Q: Are the eigenvalues and eigenspaces of similar matrices always same?

Theorem

Let $\mathbf{A}$ and $\mathbf{B}$ be $n \times n$ matrices. If $\mathbf{B}$ is similar to $\mathbf{A}$, then the two matrices have the same characteristic polynomial and, consequently, the same eigenvalues.

Proof

- Let $P_{\mathbf{A}}(\lambda)$ and $P_{\mathbf{B}}(\lambda)$ denote the characteristic polynomial of $\mathbf{A}$ and $\mathbf{B}$.
- If $\mathbf{A}$ and $\mathbf{B}$ are similar, then there exists a nonsingular matrix $\mathbf{P}$ such that

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

$$\begin{aligned} P_{\mathbf{B}}(\lambda) = \det(\mathbf{B} - \lambda\mathbf{I}) &= \det(\mathbf{P}^{-1}\mathbf{A}\mathbf{P} - \lambda\mathbf{P}^{-1}\mathbf{P}) \\ &= \det(\mathbf{P}^{-1}(\mathbf{A} - \lambda\mathbf{I})\mathbf{P}) \\ &= \det(\mathbf{P}^{-1})\det(\mathbf{A} - \lambda\mathbf{I})\det(\mathbf{P}) = P_{\mathbf{A}}(\lambda) \end{aligned}$$

- Given

$$\mathbf{T} = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} \qquad \text{and} \qquad \mathbf{S} = \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix}$$

- It is easily seen that the eigenvalues of $\mathbf{T}$ are $\lambda_1 = 2$ and $\lambda_2 = 3$. Why?
- Since it is triangular, the determinant is the product of its diagonal elements,

$$\begin{bmatrix} a_{11} - \lambda & & \\ & \ddots & \\ & & a_{nn} - \lambda \end{bmatrix}$$

- If $\mathbf{A} = \mathbf{S}^{-1}\mathbf{TS}$, then the eigenvalues of $\mathbf{A}$ should be the same as those of $\mathbf{T}$

$$\mathbf{A} = \begin{bmatrix} 2 & -3 \\ -3 & 5 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} -1 & -2 \\ 6 & 6 \end{bmatrix}$$

You can check the eigenvalues are the same easily.

Q: How about the eigenvectors?

- There are various ways to think about products of the form

$\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ in which $\mathbf{A}$ and $\mathbf{P}$ are $n \times n$ matrices and $\mathbf{P}$ is invertible,

- If they are viewed as similar transformations

$$T\colon \mathbf{A} \mapsto \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$$

many properties of the matrix $\mathbf{A}$ are preserved under such transformations

- For example,

$$\begin{aligned}
\mathbf{B} &= \mathbf{P}^{-1}\mathbf{A}\mathbf{P} \\
\det(\mathbf{B}) &= \det(\mathbf{P}^{-1}\mathbf{A}\mathbf{P}) \\
\det(\mathbf{B}) &= \frac{1}{\det(\mathbf{P})}\det(\mathbf{A})\det(\mathbf{P}) \implies \det(\mathbf{B}) = \det(\mathbf{A})
\end{aligned}$$

- In general, any property that is preserved by a similarity transformation is called a similarity invariant and is said to be invariant under similarity.

Similarity Invariants

1. Eigenvalues — $\mathbf{A}$ and $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ have the same eigenvalues.
2. Determinant — $\mathbf{A}$ and $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ have the same determinant.
3. Trace — $\mathbf{A}$ and $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ have the same trace.
4. Rank — $\mathbf{A}$ and $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ have the same rank.
5. Nullity — $\mathbf{A}$ and $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ have the same nullity.
6. Invertibility — $\mathbf{A}$ is invertible if and only if $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ is invertible
7. Eigenspace dimension — If λ is an eigenvalue of $\mathbf{A}$ (and hence of $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$) then the eigenspace of $\mathbf{A}$ corresponding to λ and the eigenspace of $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ corresponding to λ have the same dimension.

Theorem

If λ_1, λ_2, ..., λ_k are distinct eigenvalues of an $n \times n$ matrix $\mathbf{A}$ with corresponding eigenvectors $\mathbf{x}_1$, $\mathbf{x}_2$, ..., $\mathbf{x}_k$, then $\mathbf{x}_1$, $\mathbf{x}_2$, ..., $\mathbf{x}_k$ are linearly independent.

Proof

Assume they are linearly dependent, that is, let $r < k$ be the dimension of

$$\operatorname{span}(\mathcal{S}) \qquad \text{where} \qquad \mathcal{S} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$$

and $\mathcal{W} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_r\}$ are linearly independent, under this assumption,

$$\mathcal{W} \cup \{\mathbf{x}_{r+1}\} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_r, \mathbf{x}_{r+1}\}$$

are linearly dependent, and there are scalar α_1, ..., α_{r+1} that are not all zero

$$\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \cdots + \alpha_r \mathbf{x}_r + \alpha_{r+1} \mathbf{x}_{r+1} = \mathbf{0}$$

$$\alpha_1 \mathbf{x}_1 + \cdots + \alpha_r \mathbf{x}_r = -\alpha_{r+1} \mathbf{x}_{r+1}$$

Proof

Since $\alpha_{r+1} \neq 0$, otherwise $\mathbf{x}_1, \dots, \mathbf{x}_r$ would be linearly dependent. Thus

$$\alpha_{r+1}\mathbf{x}_{r+1} \neq \mathbf{0} \implies \alpha_1\mathbf{x}_1 + \cdots + \alpha_r\mathbf{x}_r \neq \mathbf{0} \implies \alpha_1, \dots, \alpha_r \text{ cannot be all zero}$$

Multiply the equation by $\mathbf{A}$,

$$\begin{aligned}
\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \cdots + \alpha_r\mathbf{x}_r + \alpha_{r+1}\mathbf{x}_{r+1} &= \mathbf{0} \\
\alpha_1\mathbf{A}\mathbf{x}_1 + \alpha_2\mathbf{A}\mathbf{x}_2 + \cdots + \alpha_r\mathbf{A}\mathbf{x}_r + \alpha_{r+1}\mathbf{A}\mathbf{x}_{r+1} &= \mathbf{0} \\
\alpha_1\lambda_1\mathbf{x}_1 + \alpha_2\lambda_2\mathbf{x}_2 + \cdots + \alpha_r\lambda_r\mathbf{x}_r + \alpha_{r+1}\lambda_{r+1}\mathbf{x}_{r+1} &= \mathbf{0}
\end{aligned}$$

If subtract λ_{r+1} times of the first equation from the third equation

$$\begin{aligned}
\alpha_1(\lambda_1 - \lambda_{r+1})\mathbf{x}_1 + \cdots + \alpha_r(\lambda_r - \lambda_{r+1})\mathbf{x}_r + \alpha_{r+1}(\lambda_{r+1} - \lambda_{r+1})\mathbf{x}_{r+1} &= \mathbf{0} \\
\alpha_1(\lambda_1 - \lambda_{r+1})\mathbf{x}_1 + \cdots + \alpha_r(\lambda_r - \lambda_{r+1})\mathbf{x}_r &= \mathbf{0}
\end{aligned}$$

Since λ_i's are distinct and $\{\alpha_1, \dots, \alpha_r\}$ cannot be all zero, not all coefficients of the last equation are zero, which contradicts the independence of $\mathcal{W}$, so the extension set $\mathcal{W} \cup \{\mathbf{x}_{r+1}\}$ is linearly independent, thus so does $\mathcal{S}$.

Definition

A matrix $\mathbf{A}_{n\times n}$ is said to be diagonalizable if there is an invertible matrix $\mathbf{P}$ and a diagonal matrix $\mathbf{D}$ such that

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

Theorem

An $n \times n$ matrix $\mathbf{A}$ is diagonalizable if and only if

$$\mathbf{A} \text{ has } n \text{ linearly independent eigenvectors.}$$

Furthermore, when $\mathbf{A}$ is diagonalizable, then

1. The columns of the diagonalizing matrix $\mathbf{P}$ are the n eigenvectors.
2. The diagonal elements of $\mathbf{D}$ are the corresponding eigenvalues.

Proof

Suppose that the matrix $\mathbf{A}$ has linearly independent eigenvector $\mathbf{x}_1, \ldots, \mathbf{x}_n$, and λ_i is the eigenvalue of $\mathbf{A}$ corresponding to $\mathbf{x}_i$ for each i, some of the λ_i's may be equal, and $\mathbf{P}$ be the matrix whose jth column is $\mathbf{x}_j$.

Proof

It follows that
$$\begin{aligned}\mathbf{AP} &= \begin{bmatrix}\mathbf{Ax}_1 & \cdots & \mathbf{Ax}_n\end{bmatrix} \\ &= \begin{bmatrix}\lambda_1\mathbf{x}_1 & \cdots & \lambda_n\mathbf{x}_n\end{bmatrix} \\ &= \begin{bmatrix}\mathbf{x}_1 & \cdots & \mathbf{x}_n\end{bmatrix}\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix} = \mathbf{PD}\end{aligned}$$

The matrix $\mathbf{P}$ has n linearly independent columns, so $\mathbf{P}$ is invertible and

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

Suppose $\mathbf{A}$ is diagonalizable, then there exists a nonsingular matrix $\mathbf{P}$ such that

$$\mathbf{AP} = \mathbf{PD}$$

If $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are the columns of $\mathbf{P}$, then

$$\mathbf{Ax}_i = \lambda_i\mathbf{x}_i, \qquad \lambda_i = d_{ii} \quad \text{for each } i$$

Since $\mathbf{P}$ is invertible, thus $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are linearly independent.

Exercise

If it is possible, find a matrix $\mathbf{P}$ that diagonalizes

$$\mathbf{A} = \begin{bmatrix} 2 & -3 \\ 2 & -5 \end{bmatrix}$$

and the corresponding diagonal matrix $\mathbf{D}$ that is similar to $\mathbf{A}$, explain otherwise.

Solution

The eigenvalues of $\mathbf{A}$ are $\lambda_1 = 1$ and $\lambda_2 = -4$. Corresponding to λ_1 and λ_2, we have the eigenvectors $\mathbf{x}_1 = 3\mathbf{e}_x + \mathbf{e}_y$ and $\mathbf{x}_2 = \mathbf{e}_x + 2\mathbf{e}_y$, so

$$\mathbf{P} = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} \qquad \text{and} \qquad \mathbf{D} = \begin{bmatrix} 1 & 0 \\ 0 & -4 \end{bmatrix}$$

You can easily verify that $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$.

- Q: Why diagonalizing matrices $\mathbf{P}$ are not unique?
- Q: How to determine whether a matrix $\mathbf{A}_{n\times n}$ is diagonalizable in general?

Exercise

Consider $\mathbf{A} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 1 & 0 & 2 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 2 & 0 & 0 \\ -1 & 4 & 0 \\ -3 & 6 & 2 \end{bmatrix}$, are they diagonalizable?

Solution

- Both matrices have the same eigenvalues,

$$\lambda_1 = 4, \qquad \text{and} \qquad \lambda_2 = \lambda_3 = 2$$

- However, $\mathbf{A}$ has only two linearly independent eigenvectors $\mathbf{e}_2$ and $\mathbf{e}_3$.
- On the other hand, $\mathbf{B}$ has three linearly independent eigenvectors

$$\mathbf{x}_1 = \mathbf{e}_2 + 3\mathbf{e}_3, \qquad \mathbf{x}_2 = 2\mathbf{e}_1 + \mathbf{e}_2, \qquad \mathbf{x}_3 = \mathbf{e}_3$$

Therefore $\mathbf{B}$ is diagonalizable, but $\mathbf{A}$ is not diagonalizable.

- If $\mathbf{A}$ is diagonalizable, then $\mathbf{A}$ can be factored into a product $\mathbf{PDP}^{-1}$, and

$$\mathbf{A}^2 = \mathbf{PDP}^{-1}\mathbf{PDP}^{-1} = \mathbf{PD}^2\mathbf{P}^{-1}$$

and, in general, $\mathbf{A}^k = \mathbf{PD}^k\mathbf{P}^{-1} = \mathbf{P}\begin{bmatrix} \lambda_1^k & & & \\ & \lambda_2^k & & \\ & & \ddots & \\ & & & \lambda_n^k \end{bmatrix}\mathbf{P}^{-1}$

- If we have a factorization $\mathbf{A} = \mathbf{PDP}^{-1}$, it is easy to compute powers of $\mathbf{A}$.
- Given a scalar α, the exponential e^α can be defined using a power series

$$e^\alpha = 1 + \alpha + \frac{1}{2!}\alpha^2 + \frac{1}{3!}\alpha^3 + \cdots = \lim_{m\to\infty}\sum_{k=0}^{m}\frac{1}{k!}\alpha^k$$

- Similarly, for any $n \times n$ matrix $\mathbf{A}$, we can define the matrix exponential $e^{\mathbf{A}}$ in terms of the convergent power series

$$e^{\mathbf{A}} = \mathbf{I} + \mathbf{A} + \frac{1}{2!}\mathbf{A}^2 + \frac{1}{3!}\mathbf{A}^3 + \cdots$$

- The matrix exponential occurs in many applications. For a diagonal matrix

$$\mathbf{D} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

the matrix exponential is easy to compute:

$$\begin{aligned} e^{\mathbf{D}} &= \lim_{m\to\infty} \left(\mathbf{I} + \mathbf{D} + \frac{1}{2!}\mathbf{D}^2 + \frac{1}{3!}\mathbf{D}^3 + \cdots + \frac{1}{m!}\mathbf{D}^m \right) \\ &= \lim_{m\to\infty} \begin{bmatrix} \sum_{k=0}^{m} \frac{1}{k!}\lambda_1^k & & \\ & \ddots & \\ & & \sum_{k=0}^{m} \frac{1}{k!}\lambda_n^k \end{bmatrix} = \begin{bmatrix} e^{\lambda_1} & & & \\ & e^{\lambda_2} & & \\ & & \ddots & \\ & & & e^{\lambda_n} \end{bmatrix} \end{aligned}$$

- It is expensive to compute the matrix exponential for an arbitrary matrix $\mathbf{A}$

$$e^{\mathbf{A}} = \mathbf{I} + \mathbf{A} + \frac{1}{2!}\mathbf{A}^2 + \frac{1}{3!}\mathbf{A}^3 + \cdots$$

- However, if $\mathbf{A}$ is diagonalizable, then

$$e^{\mathbf{A}} = \mathbf{P}\left(\mathbf{I} + \mathbf{D} + \frac{1}{2!}\mathbf{D}^2 + \frac{1}{3!}\mathbf{D}^3 + \cdots\right)\mathbf{P}^{-1} = \mathbf{P}e^{\mathbf{D}}\mathbf{P}^{-1}$$

Exercise

Compute $e^{\mathbf{A}}$ where $\mathbf{A} = \begin{bmatrix} -2 & -6 \\ 1 & 3 \end{bmatrix}$

Solution

The eigenvalues of $\mathbf{A}$ are $\lambda_1 = 1$ and $\lambda_2 = 0$, with the corresponding eigenvectors $\mathbf{x}_1 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$ and $\mathbf{x}_2 = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$, so $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1} = \begin{bmatrix} -2 & -3 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ -1 & -2 \end{bmatrix}$

$$\Longrightarrow e^{\mathbf{A}} = \mathbf{P}e^{\mathbf{D}}\mathbf{P}^{-1} = \begin{bmatrix} -2 & -3 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} e^1 & 0 \\ 0 & e^0 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ -1 & -2 \end{bmatrix} = \begin{bmatrix} 3-2e & 6-6e \\ e-1 & 3e-2 \end{bmatrix}$$