MA2001 LINEAR ALGEBRA

Diagonalization

National University of Singapore
Department of Mathematics

Motivations

- Let $\boldsymbol{A}$ be a square matrix. Then
  - $\boldsymbol{A}^m = \underbrace{\boldsymbol{A}\boldsymbol{A}\cdots\boldsymbol{A}\boldsymbol{A}}_{m \text{ times}}.$

  In general, the matrix multiplication is complicated.
  - Is there a shortcut?
- Suppose $\boldsymbol{A}$ and $\boldsymbol{B}$ are **diagonal matrices** of order n.
  - $\boldsymbol{A} = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}, \boldsymbol{B} = \begin{pmatrix} b_{11} & 0 & \cdots & 0 \\ 0 & b_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_{nn} \end{pmatrix}$
  - $\boldsymbol{A}\boldsymbol{B} = \begin{pmatrix} a_{11}b_{11} & 0 & \cdots & 0 \\ 0 & a_{22}b_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn}b_{nn} \end{pmatrix}.$

Motivations

- Let $\boldsymbol{A}$ be a square matrix. Then
  - $\boldsymbol{A}^m = \underbrace{\boldsymbol{A}\boldsymbol{A}\cdots\boldsymbol{A}\boldsymbol{A}}_{m \text{ times}}.$

  In general, the matrix multiplication is complicated.
  - Is there a shortcut?
- Suppose $\boldsymbol{A}$ is a **diagonal matrix** of order n.
  - $\boldsymbol{A} = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$
  - $\boldsymbol{A}^m = \begin{pmatrix} a_{11}^m & 0 & \cdots & 0 \\ 0 & a_{22}^m & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn}^m \end{pmatrix}.$

Motivations

- Let $\boldsymbol{A}$ be a square matrix.
  - Suppose there exists an invertible matrix $\boldsymbol{P}$ such that
    - $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{D}$ is a diagonal matrix.

    Then $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1}$.

$$
\begin{aligned}
\boldsymbol{A}^m &= (\boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1})^m \\
&= \underbrace{(\boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1})(\boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1})\cdots(\boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1})}_{m \text{ times}} \\
&= \boldsymbol{P}\boldsymbol{D}(\boldsymbol{P}^{-1}\boldsymbol{P})\boldsymbol{D}(\boldsymbol{P}^{-1}\boldsymbol{P})\cdots(\boldsymbol{P}^{-1}\boldsymbol{P})\boldsymbol{D}\boldsymbol{P}^{-1} \\
&= \boldsymbol{P}\underbrace{\boldsymbol{D}\boldsymbol{D}\cdots\boldsymbol{D}\boldsymbol{D}}_{m \text{ times}}\boldsymbol{P}^{-1} \\
&= \boldsymbol{P}\boldsymbol{D}^m\boldsymbol{P}^{-1}.
\end{aligned}
$$

Motivations

- **Example.** Suppose that each year
  - 4% of the rural population moves to the urban district.
  - 1% of the urban populations moves to the rural district.

  After n years,
  - Let a_n be the rural population;
  - Let b_n be the urban population.

  $a_n = 0.96a_{n-1} + 0.01b_{n-1}$, $b_n = 0.04a_{n-1} + 0.99b_{n-1}$.
  - $\begin{pmatrix} a_n \\ b_n \end{pmatrix} = \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix} \begin{pmatrix} a_{n-1} \\ b_{n-1} \end{pmatrix}$.

  Let $\boldsymbol{x}_n = \begin{pmatrix} a_n \\ b_n \end{pmatrix}$ and $\boldsymbol{A} = \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix}$.
  - $\boldsymbol{x}_n = \boldsymbol{A}\boldsymbol{x}_{n-1} = \boldsymbol{A}^2\boldsymbol{x}_{n-2} = \cdots = \boldsymbol{A}^n\boldsymbol{x}_0$.

Motivations

- Let $\boldsymbol{x}_n = \begin{pmatrix} a_n \\ b_n \end{pmatrix}$ and $\boldsymbol{A} = \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix}$.
  - $\boldsymbol{x}_n = \boldsymbol{A}\boldsymbol{x}_{n-1} = \boldsymbol{A}^2\boldsymbol{x}_{n-2} = \cdots = \boldsymbol{A}^n\boldsymbol{x}_0$.

  Let $\boldsymbol{P} = \begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix}$. Then $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{D} = \begin{pmatrix} 1 & 0 \\ 0 & 0.95 \end{pmatrix}$.
  - $\boldsymbol{A}^n = \boldsymbol{P}\boldsymbol{D}^n\boldsymbol{P}^{-1}$
  - $\boldsymbol{A}^n = \begin{pmatrix} 0.2 + 0.8 \cdot 0.95^n & 0.2 - 0.2 \cdot 0.95^n \\ 0.8 - 0.8 \cdot 0.95^n & 0.8 + 0.2 \cdot 0.95^n \end{pmatrix}$
    - $\begin{pmatrix} a_n \\ b_n \end{pmatrix} = \boldsymbol{x}_n = \boldsymbol{A}^n\boldsymbol{x}_0 = \boldsymbol{A}^n \begin{pmatrix} a_0 \\ b_0 \end{pmatrix}$.
    - $\begin{pmatrix} a_n \\ b_n \end{pmatrix} = \begin{pmatrix} 0.2a_0 + 0.2b_0 + (0.8a - 0.2b) \cdot 0.95^n \\ 0.8a_0 + 0.8b_0 - (0.8a - 0.2b) \cdot 0.95^n \end{pmatrix}$.

    In particular, $\begin{pmatrix} a_n \\ b_n \end{pmatrix} \xrightarrow{n\to\infty} \begin{pmatrix} 0.2(a_0 + b_0) \\ 0.8(a_0 + b_0) \end{pmatrix}$.

Motivations

- Let $\boldsymbol{A}$ be a square matrix of order 3.
  - Suppose $\boldsymbol{P} = \begin{pmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_3 \end{pmatrix}$ is invertible such that
    - $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{D} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$.

    Then $\boldsymbol{A}\boldsymbol{P} = \boldsymbol{P}\boldsymbol{D}$.
    - $\boldsymbol{A}\begin{pmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_3 \end{pmatrix} = \begin{pmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_3 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$
    - $\begin{pmatrix} \boldsymbol{A}\boldsymbol{v}_1 & \boldsymbol{A}\boldsymbol{v}_2 & \boldsymbol{A}\boldsymbol{v}_3 \end{pmatrix} = \begin{pmatrix} \lambda_1\boldsymbol{v}_1 & \lambda_2\boldsymbol{v}_2 & \lambda_3\boldsymbol{v}_3 \end{pmatrix}$.
  - Hence, $\boldsymbol{A}\boldsymbol{v}_1 = \lambda_1\boldsymbol{v}_1$, $\boldsymbol{A}\boldsymbol{v}_2 = \lambda_2\boldsymbol{v}_2$, $\boldsymbol{A}\boldsymbol{v}_3 = \lambda_3\boldsymbol{v}_3$.

Definitions

- **Definition.** Let $\boldsymbol{A}$ be a square matrix of order n.
  - Suppose that for some $\lambda \in \mathbb{R}$ and **nonzero** $\boldsymbol{v} \in \mathbb{R}^n$
    - $\boxed{\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v}}$

    λ is called an **eigenvalue** of $\boldsymbol{A}$.
    $\boldsymbol{v}$ is called an **eigenvector** of $\boldsymbol{A}$ associated with λ.
- **Example.** Let $\boldsymbol{A} = \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix}$.
  - Let $\boldsymbol{u} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$. Then $\boldsymbol{A}\boldsymbol{u} = \begin{pmatrix} 1 \\ 4 \end{pmatrix} = 1\boldsymbol{u}$.
    - $\boldsymbol{u}$ is an eigenvector of $\boldsymbol{A}$ associated to eigenvalue 1.
  - Let $\boldsymbol{v} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. $\boldsymbol{A}\boldsymbol{v} = \begin{pmatrix} 0.95 \\ -0.95 \end{pmatrix} = 0.95\boldsymbol{v}$.
    - $\boldsymbol{v}$ is an eigenvector associated to eigenvalue 0.95.

Example

- Let $\boldsymbol{B} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.
  - Let $\boldsymbol{u} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. Then $\boldsymbol{B}\boldsymbol{u} = \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix} = 3\boldsymbol{u}$.
    - $\boldsymbol{u}$ is an eigenvector of $\boldsymbol{B}$ associated to eigenvalue 3.
  - Let $\boldsymbol{v} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$. Then $\boldsymbol{B}\boldsymbol{v} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = 0\boldsymbol{v}$.
    - $\boldsymbol{v}$ is an eigenvector of $\boldsymbol{B}$ associated to eigenvalue 0.
  - Let $\boldsymbol{w} = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$. Then $\boldsymbol{B}\boldsymbol{w} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = 0\boldsymbol{w}$.
    - $\boldsymbol{w}$ is an eigenvector of $\boldsymbol{B}$ associated to eigenvalue 0.

Characteristic Equation

- Let $\boldsymbol{A}$ be a square matrix. How to find its eigenvalues?
  - $\lambda \in \mathbb{R}$ is an eigenvalue of $\boldsymbol{A}$

$\Leftrightarrow$ $\boldsymbol{A}\boldsymbol{v} = \lambda \boldsymbol{v}$ for some nonzero column vector $\boldsymbol{v}$
$\Leftrightarrow$ $\lambda \boldsymbol{v} - \boldsymbol{A}\boldsymbol{v} = \boldsymbol{0}$ for some nonzero column vector $\boldsymbol{v}$
$\Leftrightarrow$ $(\lambda \boldsymbol{I} - \boldsymbol{A})\boldsymbol{v} = \boldsymbol{0}$ for some nonzero column vector $\boldsymbol{v}$
$\Leftrightarrow$ $(\lambda \boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$ has non-trivial solution
$\Leftrightarrow$ $\lambda \boldsymbol{I} - \boldsymbol{A}$ a singular matrix
$\Leftrightarrow$ $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = 0$.

If $\boldsymbol{A}$ is of order n, then $\det(\lambda \boldsymbol{I} - \boldsymbol{A})$ is a monic polynomial in λ of degree n:
$\lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0$.

- **Definition.** Let $\boldsymbol{A}$ be a square matrix.
  - $\det(\lambda \boldsymbol{I} - \boldsymbol{A})$ is the **characteristic polynomial** of $\boldsymbol{A}$.
  - $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = 0$ is the **characteristic equation** of $\boldsymbol{A}$.

Characteristic Equation

- **Theorem.** Let $\boldsymbol{A}$ be a square matrix.
  - Then the eigenvalues of $\boldsymbol{A}$ are precisely all the roots to the characteristic equation $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = 0$.
- **Examples.**
  - Let $\boldsymbol{A} = \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix}$. Characteristic polynomial is

$$\begin{aligned}
\det(\lambda \boldsymbol{I} - \boldsymbol{A}) &= \det\left(\lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.09 \end{pmatrix}\right) \\
&= \det \begin{pmatrix} \lambda - 0.96 & -0.01 \\ -0.04 & \lambda - 0.99 \end{pmatrix} \\
&= (\lambda - 0.96)(\lambda - 0.99) - (-0.01)(-0.04) \\
&= \lambda^2 - 1.95\lambda + 0.95 \\
&= (\lambda - 0.95)(\lambda - 1).
\end{aligned}$$

Hence, $\boldsymbol{A}$ has two eigenvalues 0.95 and 1.

Characteristic Equation

- **Theorem.** Let $\boldsymbol{A}$ be a square matrix.
  - Then the eigenvalues of $\boldsymbol{A}$ are precisely all the roots to the characteristic equation $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = 0$.
- **Examples.**
  - Let $\boldsymbol{B} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$. Characteristic polynomial:

$$\begin{aligned} \det(\lambda \boldsymbol{I} - \boldsymbol{B}) &= \begin{vmatrix} \lambda - 1 & -1 & -1 \\ -1 & \lambda - 1 & -1 \\ -1 & -1 & \lambda - 1 \end{vmatrix} \\ &= \lambda^3 - 3\lambda^2 \\ &= \lambda^2(\lambda - 3). \end{aligned}$$

Hence, $\boldsymbol{B}$ has two eigenvalues 0 and 3.

Characteristic Equation

- **Theorem.** Let $\boldsymbol{A}$ be a square matrix.
  - Then the eigenvalues of $\boldsymbol{A}$ are precisely all the roots to the characteristic equation $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = 0$.
- **Examples.**
  - Let $\boldsymbol{C} = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 2 \\ 1 & 1 & 1 \end{pmatrix}$. Characteristic polynomial:

$$\begin{aligned} \det(\lambda \boldsymbol{I} - \boldsymbol{C}) &= \begin{vmatrix} \lambda & 1 & 0 \\ 0 & \lambda & -2 \\ -1 & -1 & \lambda - 1 \end{vmatrix} \\ &= \lambda^3 - \lambda^2 - 2\lambda + 2 \\ &= (\lambda - 1)(\lambda - \sqrt{2})(\lambda + \sqrt{2}). \end{aligned}$$

Hence, $\boldsymbol{C}$ has three eigenvalues 1, $\sqrt{2}$ and $-\sqrt{2}$.

Main Theorem for Invertible Matrices

- **Theorem.** Let $\boldsymbol{A}$ be a square matrix of order n. Then the following are equivalent:
  1. $\boldsymbol{A}$ is invertible.
  2. The reduced row-echelon form of $\boldsymbol{A}$ is $\boldsymbol{I}_n$.
  3. The homogeneous linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ has only the trivial solution.
  4. The linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ has exactly one solution.
  5. $\boldsymbol{A}$ is the product of elementary matrices.
  6. $\det(\boldsymbol{A}) \neq 0$.
  7. The rows of $\boldsymbol{A}$ form a basis for $\mathbb{R}^n$.
  8. The columns of $\boldsymbol{A}$ form a basis for $\mathbb{R}^n$.
  9. $\operatorname{rank}(\boldsymbol{A}) = n$.
  10. 0 is not an eigenvalue of $\boldsymbol{A}$.

Main Theorem for Invertible Matrices

- **Proof.** It remains to show that “10” is equivalent to “6”:
  - 0 is not an eigenvalue of $\boldsymbol{A}$

$$
\begin{aligned}
&\Leftrightarrow 0 \text{ is not a root to } \det(\lambda \boldsymbol{I} - \boldsymbol{A}) = 0 \\
&\Leftrightarrow \det(0\boldsymbol{I} - \boldsymbol{A}) \neq 0 \\
&\Leftrightarrow \det(-\boldsymbol{A}) \neq 0 \\
&\Leftrightarrow (-1)^n \det(\boldsymbol{A}) \neq 0 \\
&\Leftrightarrow \det(\boldsymbol{A}) \neq 0.
\end{aligned}
$$

Upper Triangular Matrices

- Let $\boldsymbol{A}$ be an **upper triangular** matrix of order n:
  - $\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$

  Its characteristic polynomial is $\det(\lambda \boldsymbol{I} - \boldsymbol{A})$:

$$\begin{aligned} \det(\lambda \boldsymbol{I} - \boldsymbol{A}) &= \begin{vmatrix} \lambda - a_{11} & -a_{12} & -a_{13} & \cdots & a_{1n} \\ 0 & \lambda - a_{22} & -a_{23} & \cdots & a_{2n} \\ 0 & 0 & \lambda - a_{33} & \cdots & -a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda - a_{nn} \end{vmatrix} \\ &= (\lambda - a_{11})(\lambda - a_{22})(\lambda - a_{33}) \cdots (\lambda - a_{nn}). \end{aligned}$$

Upper Triangular Matrices

- **Theorem.** Let $\boldsymbol{A}$ be an upper (or lower) triangular matrix. Then its eigenvalues are all the diagonal entries of $\boldsymbol{A}$.
  - More precisely, if $\boldsymbol{A} = (a_{ij})_{n \times n}$ is upper triangular $(a_{ij} = 0$ if $i > j)$ or lower triangular $(a_{ij} = 0$ if $i < j)$,
    - then the eigenvalues of $\boldsymbol{A}$ are $a_{11}, a_{22}, \ldots, a_{nn}$.
- **Examples.**
  - $\begin{pmatrix} -1 & 3.5 & 14 \\ 0 & 5 & -26 \\ 0 & 0 & 2 \end{pmatrix}$. Eigenvalues: -1, 5 and 2.
  - $\begin{pmatrix} -2 & 0 & 0 \\ 99 & 0 & 0 \\ 10 & -4.5 & 10 \end{pmatrix}$. Eigenvalues: -2, 0 and 10.

Eigenspace

- Let $\boldsymbol{A}$ be a square matrix of order n.
  - Let λ be an eigenvalue of $\boldsymbol{A}$.

  Let $\mathbf{0} \neq \boldsymbol{v} \in \mathbb{R}^n$. Then
  - $\boldsymbol{v}$ is an eigenvector of $\boldsymbol{A}$ associated to λ

    $\Leftrightarrow \boldsymbol{A}\boldsymbol{v} = \lambda \boldsymbol{v}$
    $\Leftrightarrow (\lambda \boldsymbol{I} - \boldsymbol{A})\boldsymbol{v} = \mathbf{0}$
    $\Leftrightarrow \boldsymbol{v}$ is a nonzero vector in the nullspace of $\lambda \boldsymbol{I} - \boldsymbol{A}$.
- **Definition.** Let $\boldsymbol{A}$ be a square matrix and λ an eigenvalue of $\boldsymbol{A}$. (Then $\lambda \boldsymbol{I} - \boldsymbol{A}$ is singular.)
  - The **eigenspace** of $\boldsymbol{A}$ associated to λ is the nullspace of $\lambda \boldsymbol{I} - \boldsymbol{A}$, denoted by E_λ (or $E_{\boldsymbol{A},\lambda}$).
    - E_λ consists of all the eigenvectors of $\boldsymbol{A}$ associated to λ, and the zero vector $\mathbf{0}$. Note that $\dim E_\lambda \geq 1$.

Examples

- $\boldsymbol{A} = \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix}$ has eigenvalues 1 and 0.95.
  - The eigenspace E_1 is the nullspace of $1\boldsymbol{I} - \boldsymbol{A}$.
    - $1\boldsymbol{I} - \boldsymbol{A} = \begin{pmatrix} 0.04 & -0.01 \\ -0.04 & 0.01 \end{pmatrix}$.
    - $(1\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \mathbf{0} \Leftrightarrow \boldsymbol{x} = t \begin{pmatrix} 0.25 \\ 1 \end{pmatrix}$, $t \in \mathbb{R}$.

    Then $E_1 = \operatorname{span}\left\{ \begin{pmatrix} 0.25 \\ 1 \end{pmatrix} \right\}$, and $\dim(E_1) = 1$.
  - The eigenspace $E_{0.95}$ is the nullspace of $0.95\boldsymbol{I} - \boldsymbol{A}$.
    - $0.95\boldsymbol{I} - \boldsymbol{A} = \begin{pmatrix} -0.01 & -0.01 \\ -0.04 & -0.04 \end{pmatrix}$.
    - $(0.95\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \mathbf{0} \Leftrightarrow \boldsymbol{x} = t \begin{pmatrix} -1 \\ 1 \end{pmatrix}$, $t \in \mathbb{R}$.

    Then $E_{0.95} = \operatorname{span}\left\{ \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\}$, and $\dim(E_{0.95}) = 1$.

Examples

- $\boldsymbol{B} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ has eigenvalues 3 and 0.
  - The eigenspace E_3 is the nullspace of $3\boldsymbol{I} - \boldsymbol{B}$.
    - $3\boldsymbol{I} - \boldsymbol{B} = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$.
    - $(3\boldsymbol{I} - \boldsymbol{B})\boldsymbol{x} = \boldsymbol{0} \Leftrightarrow \boldsymbol{x} = t\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, t \in \mathbb{R}$.

    Then $E_3 = \operatorname{span}\left\{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right\}$, and $\dim(E_3) = 1$.

Examples

- $\boldsymbol{B} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ has eigenvalues 3 and 0.
  - The eigenspace E_0 is the nullspace of $0\boldsymbol{I} - \boldsymbol{B}$.
    - $0\boldsymbol{I} - \boldsymbol{B} = -\boldsymbol{B} = \begin{pmatrix} -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \end{pmatrix}$.
    - $(0\boldsymbol{I} - \boldsymbol{B})\boldsymbol{x} = \boldsymbol{0} \Leftrightarrow \boldsymbol{x} = s\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, s, t \in \mathbb{R}$.

    $E_0 = \operatorname{span}\left\{\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}\right\}$, and $\dim(E_0) = 2$.
- **Note:** If $\boldsymbol{A}$ is singular, then 0 is an eigenvalue of $\boldsymbol{A}$.
  - The eigenspace E_0 is the nullspace of $\boldsymbol{A}$.

Examples

- $\boldsymbol{C} = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 2 \\ 1 & 1 & 1 \end{pmatrix}$ has eigenvalues 1, $\sqrt{2}$ and $-\sqrt{2}$.
  - The eigenspace E_1 is the nullspace of $1\boldsymbol{I} - \boldsymbol{C}$.
    - $1\boldsymbol{I} - \boldsymbol{C} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -2 \\ -1 & -1 & 0 \end{pmatrix}$.
    - $(1\boldsymbol{I} - \boldsymbol{C})\boldsymbol{x} = \boldsymbol{0} \Leftrightarrow \boldsymbol{x} = t \begin{pmatrix} -2 \\ 2 \\ 1 \end{pmatrix}, t \in \mathbb{R}$.

    $E_1 = \operatorname{span}\left\{ \begin{pmatrix} -2 \\ 2 \\ 1 \end{pmatrix} \right\}$, and $\dim(E_1) = 1$.

Examples

- $\boldsymbol{C} = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 2 \\ 1 & 1 & 1 \end{pmatrix}$ has eigenvalues 1, $\sqrt{2}$ and $-\sqrt{2}$.
  - The eigenspace $E_{\sqrt{2}}$ is the nullspace of $\sqrt{2}\boldsymbol{I} - \boldsymbol{C}$.
    - $\sqrt{2}\boldsymbol{I} - \boldsymbol{C} = \begin{pmatrix} \sqrt{2} & 1 & 0 \\ 0 & \sqrt{2} & -2 \\ -1 & -1 & \sqrt{2} - 1 \end{pmatrix}$.
    - $(\sqrt{2}\boldsymbol{I} - \boldsymbol{C})\boldsymbol{x} = \boldsymbol{0} \Leftrightarrow \boldsymbol{x} = t \begin{pmatrix} -1 \\ \sqrt{2} \\ 1 \end{pmatrix}, t \in \mathbb{R}$.

    $E_{\sqrt{2}} = \operatorname{span}\left\{ \begin{pmatrix} -1 \\ \sqrt{2} \\ 1 \end{pmatrix} \right\}$, and $\dim(E_{\sqrt{2}}) = 1$.

Examples

- $C = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 2 \\ 1 & 1 & 1 \end{pmatrix}$ has eigenvalues 1, $\sqrt{2}$ and $-\sqrt{2}$.
  - The eigenspace $E_{-\sqrt{2}}$ is the nullspace of $-\sqrt{2}\boldsymbol{I} - \boldsymbol{C}$.
    - $-\sqrt{2}\boldsymbol{I} - \boldsymbol{C} = \begin{pmatrix} -\sqrt{2} & 1 & 0 \\ 0 & -\sqrt{2} & -2 \\ -1 & -1 & -\sqrt{2}-1 \end{pmatrix}$.
    - $(-\sqrt{2}\boldsymbol{I} - \boldsymbol{C})\boldsymbol{x} = \boldsymbol{0} \Leftrightarrow \boldsymbol{x} = t\begin{pmatrix} -1 \\ -\sqrt{2} \\ 1 \end{pmatrix}$, $t \in \mathbb{R}$.

    $E_{-\sqrt{2}} = \mathrm{span}\left\{\begin{pmatrix} -1 \\ -\sqrt{2} \\ 1 \end{pmatrix}\right\}$, and $\dim(E_{-\sqrt{2}}) = 1$.

Diagonalization

Diagonalizable Matrices

- **Definition.** Let $\boldsymbol{A}$ be a square matrix.
  - $\boldsymbol{A}$ is called **diagonalizable** if there exists an **invertible** matrix $\boldsymbol{P}$ such that $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ is a **diagonal** matrix.
- **Examples**.
  - $\boldsymbol{A} = \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix}$ and $\boldsymbol{P} = \begin{pmatrix} 0.25 & -1 \\ 1 & 1 \end{pmatrix}$
    - Then $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{D} = \begin{pmatrix} 1 & 0 \\ 0 & 0.95 \end{pmatrix}$.

    Then $\boldsymbol{A}$ is diagonalizable.
  - Note that the diagonal entries of $\boldsymbol{D}$ are the eigenvalues of $\boldsymbol{A}$.
    - The columns of $\boldsymbol{P}$ are eigenvectors of $\boldsymbol{A}$ associated to these eigenvalues.

Diagonalizable Matrices

- **Definition.** Let $\boldsymbol{A}$ be a square matrix.
  - $\boldsymbol{A}$ is called **diagonalizable** if there exists an **invertible** matrix $\boldsymbol{P}$ such that $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ is a **diagonal** matrix.
- **Examples**.
  - $\boldsymbol{B} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ and $\boldsymbol{P} = \begin{pmatrix} 1 & -1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$.
    - $\boldsymbol{P}^{-1}\boldsymbol{B}\boldsymbol{P} = \boldsymbol{D} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. So $\boldsymbol{B}$ is diagonalizable.
  - Note that the diagonal entries of $\boldsymbol{D}$ are the eigenvalues of $\boldsymbol{B}$.
    - The columns of $\boldsymbol{P}$ are eigenvectors of $\boldsymbol{B}$ associated to these eigenvalues.

Examples

- Prove that $\boldsymbol{M} = \begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix}$ is not diagonalizable.
  - Suppose there exists invertible $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ such that
    - $\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} \begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$.
      i.e., $\begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$.
    - $\begin{pmatrix} 2a & 2b \\ a+2c & b+2d \end{pmatrix} = \begin{pmatrix} \lambda a & \mu b \\ \lambda c & \mu d \end{pmatrix}$.
    If $a \neq 0$, then $\lambda = 2$, and $a + 2c = 2c \Rightarrow a = 0$; so $a = 0$.
    If $b \neq 0$, then $\mu = 2$, and $b + 2d = 2d \Rightarrow b = 0$; so $b = 0$.
    - Then $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ c & d \end{pmatrix}$ is singular.
  - Therefore, $\boldsymbol{M}$ is not diagonalizable.

Criterion of Diagonalizability

- Let $\boldsymbol{A}$ be a square matrix of order n.
  - Suppose that $\boldsymbol{A}$ is diagonalizable.
    - There exist invertible matrices $\boldsymbol{P}$ such that $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ is a diagonal matrix $\boldsymbol{D}$, i.e., $\boldsymbol{A}\boldsymbol{P} = \boldsymbol{P}\boldsymbol{D}$.

    Let $\boldsymbol{P} = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{pmatrix}$ and $\boldsymbol{D} = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$.

    - $\boldsymbol{A}\begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{pmatrix} \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$.
    - $\begin{pmatrix} \boldsymbol{A}\boldsymbol{v}_1 & \cdots & \boldsymbol{A}\boldsymbol{v}_n \end{pmatrix} = \begin{pmatrix} \lambda_1\boldsymbol{v}_1 & \cdots & \lambda_n\boldsymbol{v}_n \end{pmatrix}$.
  - Then $\boldsymbol{A}\boldsymbol{v}_i = \lambda_i \boldsymbol{v}_i, \quad i = 1, \ldots, n$.

  λ_i is an eigenvalue of $\boldsymbol{A}$, $\boldsymbol{v}_i$ is an eigenvector associated to λ_i.
  - $\boldsymbol{P}$ is invertible $\Rightarrow \boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ are linearly independent.

Criterion of Diagonalizability

- Let $\boldsymbol{A}$ be a square matrix of order n.
  - Suppose $\boldsymbol{A}$ has n linearly independent eigenvectors.
    - $\boldsymbol{A}\boldsymbol{v}_1 = \lambda_1\boldsymbol{v}_1, \boldsymbol{A}\boldsymbol{v}_2 = \lambda_2\boldsymbol{v}_2, \ldots, \boldsymbol{A}\boldsymbol{v}_n = \lambda_n\boldsymbol{v}_n,$
      - where $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ are linearly independent.

    Let $\boldsymbol{P} = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{pmatrix}$. Then $\boldsymbol{P}$ is invertible.
    - Let $\boldsymbol{D} = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$.

$$\begin{aligned} \boldsymbol{A}\boldsymbol{P} &= \boldsymbol{A}\begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{A}\boldsymbol{v}_1 & \cdots & \boldsymbol{A}\boldsymbol{v}_n \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1\boldsymbol{v}_1 & \cdots & \lambda_n\boldsymbol{v}_n \end{pmatrix} \\ &= \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{pmatrix} \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} = \boldsymbol{P}\boldsymbol{D}. \end{aligned}$$

  - $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{D}$; so $\boldsymbol{A}$ is diagonalizable.

Criterion of Diagonalizability

- **Theorem.** Let $\boldsymbol{A}$ be a square matrix of order n.
  - $\boldsymbol{A}$ is diagonalizable

    $\Leftrightarrow$ $\boldsymbol{A}$ has n linearly independent eigenvectors.
- **Remark.** Suppose that $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{D}$ is diagonal.
  - The diagonal entries of $\boldsymbol{D}$ are eigenvalues of $\boldsymbol{A}$:
    - $\lambda_1, \ldots, \lambda_n$, which may be repeated.

    $\boldsymbol{D}$ is not unique unless $\boldsymbol{A}$ has only one eigenvalue.
  - The columns of $\boldsymbol{P}$ are eigenvectors of $\boldsymbol{A}$:
    - $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$, which are linearly independent.
    - $\boldsymbol{v}_i$ is an eigenvector of $\boldsymbol{A}$ associated to λ_i.

    $\boldsymbol{P}$ is not unique. For instance,
    - $\boldsymbol{v}_i$ can be replaced by a nonzero multiple of $\boldsymbol{v}_i$.

Diagonalization

- **Algorithm of Diagonalization**
  - Let $\boldsymbol{A}$ be a square matrix of order n.
    1. Solve $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = 0$ to find eigenvalues of $\boldsymbol{A}$.
    2. For each eigenvalue λ_i of $\boldsymbol{A}$,
       - find a basis S_i for the eigenspace E_{λ_i}.

$$\boldsymbol{A} \text{ is diagonalizable} \Leftrightarrow |S_1| + \cdots + |S_k| = n,$$
$$\boldsymbol{A} \text{ is not diagonalizable} \Leftrightarrow |S_1| + \cdots + |S_k| < n.$$

Suppose $\boldsymbol{A}$ is diagonalizable. Then

- $S_1 \cup \cdots \cup S_k = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is a basis for $\mathbb{R}^n$.
- $\boldsymbol{A}$ is diagonalized by $\boldsymbol{P} = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{pmatrix}$.

Remarks

- $\det(\lambda \boldsymbol{I} - \boldsymbol{A})$ is a polynomial of λ in degree n.
  - It can be completely factorized as
    - $(\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n)$, $\lambda_i \in \mathbb{C}$.

    But $\lambda_1, \ldots, \lambda_n$ are not necessarily real numbers.
  - If some λ_i is not real,
    - then $\boldsymbol{A}$ is not diagonalizable (over $\mathbb{R}$).
- **Example.** Let $\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.
  - $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = \lambda^2 + 1 = (\lambda - i)(\lambda + i)$.
    - $\boldsymbol{A}$ is not diagonalizable over $\mathbb{R}$.
    - $\begin{pmatrix} -i & i \\ 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} -i & i \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$.

Remarks

- Suppose that $\det(\lambda \boldsymbol{I} - \boldsymbol{A})$ can be completely factorized:
  - $(\lambda - \lambda_1)^{r_1}(\lambda - \lambda_2)^{r_2} \cdots (\lambda - \lambda_k)^{r_k}$,
    - where $\lambda_1, \lambda_2, \ldots, \lambda_k$ are all distinct.

    Then r_i is the **algebraic multiplicity** $a(\lambda_i)$ of λ_i.
  - Let E_i be the eigenspace of $\boldsymbol{A}$ associated to λ_i.
    - $\dim E_i$ is the **geometric multiplicity** $g(\lambda_i)$ of λ_i.
  - One can prove (MA2101) that $g(\lambda_i) \leq a(\lambda_i)$.

  Note that $a(\lambda_1) + a(\lambda_2) + \cdots + a(\lambda_k) = n$.
  - If $\dim E_i < a(\lambda_i)$ for some i,
    - then $\dim E_1 + \dim E_2 + \cdots + \dim E_k < n$;

    consequently, $\boldsymbol{A}$ is not diagonalizable.

Remarks

- Let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be distinct eigenvalues of $\boldsymbol{A}$.
  - and $\boldsymbol{v}_i$ be an eigenvector of $\boldsymbol{A}$ associated to λ_i.

  Then $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ are linearly independent.
- **Proof.** Let $k = 2$. Suppose $c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 = \boldsymbol{0}$.

$$\begin{aligned}
\boldsymbol{0} = \boldsymbol{A}\boldsymbol{0} &= \boldsymbol{A}(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2) \\
&= c_1(\boldsymbol{A}\boldsymbol{v}_1) + c_2(\boldsymbol{A}\boldsymbol{v}_2) \\
&= (c_1\lambda_1)\boldsymbol{v}_1 + (c_2\lambda_2)\boldsymbol{v}_2, \\
\boldsymbol{0} = \lambda_1\boldsymbol{0} &= \lambda_1(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2) \\
&= (c_1\lambda_1)\boldsymbol{v}_1 + (c_2\lambda_1)\boldsymbol{v}_2.
\end{aligned}$$

  - Then $c_2\lambda_2\boldsymbol{v}_2 = c_2\lambda_1\boldsymbol{v}_2$, i.e., $c_2(\lambda_2 - \lambda_1)\boldsymbol{v}_2 = \boldsymbol{0}$.
    - $\boldsymbol{v}_2 \neq 0$, $\lambda_1 \neq \lambda_2$; so $c_2 = 0$ & $c_1 = 0$.

  The general case can be proved by mathematical induction. (Exercise.)

Diagonalization

- **Algorithm of Diagonalization**
  - Let $\boldsymbol{A}$ be a square matrix of order n.

Case 1. If $\det(\lambda\boldsymbol{I} - \boldsymbol{A})$ cannot be completely factorized,
- then $\boldsymbol{A}$ is not diagonalizable.

Case 2. If $\det(\lambda\boldsymbol{I} - \boldsymbol{A})$ can be completely factorized,
- for each λ_i, find a basis S_i for its eigenspace.

2a. If $|S_i| < a(\lambda_i)$ for some i,
- then $\boldsymbol{A}$ is not diagonalizable.

2b. If $|S_i| = a(\lambda_i)$ for all i,
- then $\boldsymbol{A}$ is diagonalizable.
- $S_1 \cup \cdots \cup S_k = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is a basis for $\mathbb{R}^n$.
- $\boldsymbol{P} = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{pmatrix}$ diagonalizes $\boldsymbol{A}$.

Examples

- Let $\boldsymbol{B} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.

Step 1. $\det(\lambda \boldsymbol{I} - \boldsymbol{B}) = (\lambda - 3)\lambda^2$.

- ○ $\boldsymbol{B}$ has eigenvalues $\lambda = 3$ and $\lambda = 0$.

Step 2. Find bases for eigenspaces:

- ○ E_3: $\{(1,1,1)^{\mathrm{T}}\}$.
- ○ E_0: $\{(-1,1,0)^{\mathrm{T}}, (-1,0,1)^{\mathrm{T}}\}$.

Step 3. $\boldsymbol{P} = \begin{pmatrix} 1 & -1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$. Then $\boldsymbol{P}^{-1}\boldsymbol{B}\boldsymbol{P} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

38 / 96

Examples

- Let $\boldsymbol{B} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.

Step 1. $\det(\lambda \boldsymbol{I} - \boldsymbol{B}) = (\lambda - 3)\lambda^2$.

- ○ $\boldsymbol{B}$ has eigenvalues $\lambda = 3$ and $\lambda = 0$.

Step 2. Find bases for eigenspaces:

- ○ E_3: $\{(1,1,1)^{\mathrm{T}}\}$.
- ○ E_0: $\{(-1,1,0)^{\mathrm{T}}, (-1,0,1)^{\mathrm{T}}\}$.

Step 3. $\boldsymbol{P} = \begin{pmatrix} -1 & 1 & -1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$. Then $\boldsymbol{P}^{-1}\boldsymbol{B}\boldsymbol{P} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

- The ith column of $\boldsymbol{P}$ is an eigenvector of $\boldsymbol{B}$ associated to the ith diagonal entry (eigenvalue) of $\boldsymbol{P}^{-1}\boldsymbol{B}\boldsymbol{P}$.

39 / 96

Examples

- Let $\boldsymbol{C} = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 2 \\ 1 & 1 & 1 \end{pmatrix}$.

Step 1. $\det(\lambda \boldsymbol{I} - \boldsymbol{C}) = (\lambda - 1)(\lambda - \sqrt{2})(\lambda + \sqrt{2})$.

- $\boldsymbol{C}$ has eigenvalues $\lambda = 1$, $\sqrt{2}$ and $-\sqrt{2}$.

Step 2. Find bases for eigenspaces:

- E_1: $\{(-2, 2, 1)^{\mathrm{T}}\}$.
- $E_{\sqrt{2}}$: $\{(-1, \sqrt{2}, 1)^{\mathrm{T}}\}$.
- $E_{-\sqrt{2}}$: $\{(-1, -\sqrt{2}, 1)^{\mathrm{T}}\}$.

Step 3. $\boldsymbol{P} = \begin{pmatrix} -2 & -1 & -1 \\ 2 & \sqrt{2} & -\sqrt{2} \\ 1 & 1 & 1 \end{pmatrix}$, $\boldsymbol{P}^{-1}\boldsymbol{C}\boldsymbol{P} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & -\sqrt{2} \end{pmatrix}$.

Examples

- Let $\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ -3 & 5 & 2 \end{pmatrix}$.

Step 1. $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = (\lambda - 1)(\lambda - 2)^2$.

- $\boldsymbol{A}$ has eigenvalues $\lambda = 1$ and 2.

Step 2. Find bases for eigenspaces:

- $1\boldsymbol{I} - \boldsymbol{A} = \begin{pmatrix} 0 & 0 & 0 \\ -1 & -1 & 0 \\ 3 & -5 & -1 \end{pmatrix}$.
- $(1\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0} \Leftrightarrow \boldsymbol{x} = t \begin{pmatrix} 1 \\ -1 \\ 8 \end{pmatrix}, t \in \mathbb{R}$.
- $E_1 = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ -1 \\ 8 \end{pmatrix} \right\}$.

Examples

- Let $\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ -3 & 5 & 2 \end{pmatrix}$.

Step 1. $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = (\lambda - 1)(\lambda - 2)^2$.

- ○ $\boldsymbol{A}$ has eigenvalues $\lambda = 1$ and 2.

Step 2. Find bases for eigenspaces:

- ○ $2\boldsymbol{I} - \boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ 3 & -5 & 0 \end{pmatrix}$.
- ○ $(2\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0} \Leftrightarrow \boldsymbol{x} = t \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, t \in \mathbb{R}$.
- ○ $E_2 = \operatorname{span}\left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$.

Examples

- Let $\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ -3 & 5 & 2 \end{pmatrix}$.

Step 1. $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = (\lambda - 1)(\lambda - 2)^2$.

- ○ $\boldsymbol{A}$ has eigenvalues $\lambda = 1$ and 2.

Step 2. Find bases for eigenspaces:

- ○ $E_1 = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ -1 \\ 8 \end{pmatrix} \right\}$,
- ○ $E_2 = \operatorname{span}\left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$.

Step 3. Since there are only two linearly independent eigenvectors, $\boldsymbol{A}$ is not diagonalizable.

Examples

- Let $\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$.

Step 1. $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = \lambda^2 - \lambda - 1$.

- $\boldsymbol{A}$ has eigenvalues $\frac{1+\sqrt{5}}{2}$ and $\frac{1-\sqrt{5}}{2}$.

Step 2. Find eigenspaces:

- $\left(\frac{1+\sqrt{5}}{2}\boldsymbol{I} - \boldsymbol{A}\right)\boldsymbol{x} = \boldsymbol{0} \Leftrightarrow \boldsymbol{x} = t\begin{pmatrix} 1 \\ \frac{1+\sqrt{5}}{2} \end{pmatrix}, t \in \mathbb{R}$.
  - $E_{\frac{1+\sqrt{5}}{2}} = \operatorname{span}\left\{\begin{pmatrix} 1 \\ \frac{1+\sqrt{5}}{2} \end{pmatrix}\right\}$.
- $\left(\frac{1-\sqrt{5}}{2}\boldsymbol{I} - \boldsymbol{A}\right)\boldsymbol{x} = \boldsymbol{0} \Leftrightarrow \boldsymbol{x} = t\begin{pmatrix} 1 \\ \frac{1-\sqrt{5}}{2} \end{pmatrix}, t \in \mathbb{R}$.
  - $E_{\frac{1-\sqrt{5}}{2}} = \operatorname{span}\left\{\begin{pmatrix} 1 \\ \frac{1-\sqrt{5}}{2} \end{pmatrix}\right\}$.

44 / 96

Examples

- Let $\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$.

Step 1. $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = \lambda^2 - \lambda - 1$.

- $\boldsymbol{A}$ has eigenvalues $\frac{1+\sqrt{5}}{2}$ and $\frac{1-\sqrt{5}}{2}$.

Step 2. Find eigenspaces:

- $E_{\frac{1+\sqrt{5}}{2}} = \operatorname{span}\left\{\begin{pmatrix} 1 \\ \frac{1+\sqrt{5}}{2} \end{pmatrix}\right\}$.
- $E_{\frac{1-\sqrt{5}}{2}} = \operatorname{span}\left\{\begin{pmatrix} 1 \\ \frac{1-\sqrt{5}}{2} \end{pmatrix}\right\}$.

Step 3. $\boldsymbol{P} = \begin{pmatrix} 1 & 1 \\ \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \end{pmatrix}$. $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix}$.

45 / 96

Examples

- **Theorem.** Let $\boldsymbol{A}$ be a square matrix of order n.
  - If $\boldsymbol{A}$ has n distinct eigenvalues,
    - then $\boldsymbol{A}$ is diagonalizable.

  Proof. Suppose $\boldsymbol{A}$ has distinct eigenvalues $\lambda_1, \dots, \lambda_n$.
  - Let $\boldsymbol{v}_i$ be an eigenvector of $\boldsymbol{A}$ associated to λ_i.
  - It is known that $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ are linearly independent.

  Therefore, $\boldsymbol{A}$ is diagonalizable.
- **Example.** Let $\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 2 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 4 \end{pmatrix}$.
  - $\boldsymbol{A}$ has eigenvalues $1, 2, 3, 4$; so $\boldsymbol{A}$ is diagonalizable.
  - Can you diagonalize it? (Exercise.)

Application

- Suppose that $\boldsymbol{A}$ is diagonalizable.
  - There exists an invertible matrix $\boldsymbol{P}$ such that
    - $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$ is diagonal.
    - $\boldsymbol{A} = \boldsymbol{P} \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} \boldsymbol{P}^{-1}$.
  - Let m be a nonnegative integer. Then
    - $\boldsymbol{A}^m = \boldsymbol{P} \begin{pmatrix} \lambda_1^m & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n^m \end{pmatrix} \boldsymbol{P}^{-1}$.

Application

- Suppose that $\boldsymbol{A}$ is diagonalizable.
  - There exists an invertible matrix $\boldsymbol{P}$ such that
    - $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$ is diagonal.
    - $\boldsymbol{A} = \boldsymbol{P} \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} \boldsymbol{P}^{-1}$.
  - Suppose that $\boldsymbol{A}$ is invertible. Then for any integer m,
    - $\boldsymbol{A}^m = \boldsymbol{P} \begin{pmatrix} \lambda_1^m & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n^m \end{pmatrix} \boldsymbol{P}^{-1}$.

Examples

- Let $\boldsymbol{A} = \begin{pmatrix} -4 & 0 & -6 \\ 2 & 1 & 2 \\ 3 & 0 & 5 \end{pmatrix}$.
  - $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = (\lambda + 1)(\lambda - 1)(\lambda - 2)$.
    - $(-1\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0} \Leftrightarrow \boldsymbol{x} = t \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}, t \in \mathbb{R}$.
    - $(1\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0} \Leftrightarrow \boldsymbol{x} = t \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, t \in \mathbb{R}$.
    - $(2\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0} \Leftrightarrow \boldsymbol{x} = t \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, t \in \mathbb{R}$.
  - $\boldsymbol{P} = \begin{pmatrix} -2 & 0 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$. $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$.

Examples

- Let $\boldsymbol{A} = \begin{pmatrix} -4 & 0 & -6 \\ 2 & 1 & 2 \\ 3 & 0 & 5 \end{pmatrix}$.
  - $\boldsymbol{P} = \begin{pmatrix} -2 & 0 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$. $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$.
    - $\boldsymbol{A}^m = \boldsymbol{P} \begin{pmatrix} (-1)^m & 0 & 0 \\ 0 & 1^m & 0 \\ 0 & 0 & 2^m \end{pmatrix} \boldsymbol{P}^{-1}$.

$$\begin{aligned} \boldsymbol{A}^{10} &= \boldsymbol{P} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1024 \end{pmatrix} \boldsymbol{P}^{-1} \\ &= \cdots = \begin{pmatrix} -1022 & 0 & -2046 \\ 0 & 1 & 0 \\ 1023 & 0 & 2047 \end{pmatrix}. \end{aligned}$$

Examples

- The **Fibonacci numbers** a_n are defined by
  - $a_0 = 0$, $a_1 = 1$ and $a_n = a_{n-1} + a_{n-2}$ for $n \geq 2$.

  Note that $a_{n+1} = a_{n-1} + a_n$ for $n \geq 1$.
  - $\begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix} = \begin{pmatrix} a_n \\ a_{n-1} + a_n \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} a_{n-1} \\ a_n \end{pmatrix}$.

  Let $\boldsymbol{x}_n = \begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix}$ and $\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$.
  - $\boldsymbol{x}_n = \boldsymbol{A}\boldsymbol{x}_{n-1} = \boldsymbol{A}^2\boldsymbol{x}_{n-2} = \cdots = \boldsymbol{A}^n\boldsymbol{x}_0$, $\boldsymbol{x}_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

  We have diagonalized $\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$.
  - $\boldsymbol{P} = \begin{pmatrix} 1 & 1 \\ \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \end{pmatrix}$. $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix}$.

Examples

- The **Fibonacci numbers** F_n are defined by
  - $a_0 = 0$, $a_1 = 1$ and $a_n = a_{n-1} + a_{n-2}$ for $n \geq 2$.

  Let $\boldsymbol{x}_n = \begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix}$ and $\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$.

$$\begin{aligned}
\boldsymbol{x}_n = \boldsymbol{A}^n \boldsymbol{x}_0 &= \boldsymbol{P} \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix}^n \boldsymbol{P}^{-1} \boldsymbol{x}_0 \\
&= \boldsymbol{P} \begin{pmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^n & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^n \end{pmatrix} \boldsymbol{P}^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{\sqrt{5}} \left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}} \left(\frac{1-\sqrt{5}}{2}\right)^n \\ \frac{1}{\sqrt{5}} \left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \frac{1}{\sqrt{5}} \left(\frac{1-\sqrt{5}}{2}\right)^{n+1} \end{pmatrix}
\end{aligned}$$

  Therefore, $a_n = \dfrac{1}{\sqrt{5}} \left(\dfrac{1+\sqrt{5}}{2}\right)^n - \dfrac{1}{\sqrt{5}} \left(\dfrac{1-\sqrt{5}}{2}\right)^n$.

Orthogonal Diagonalization

Introduction

- Recall that an $n \times n$ matrix $\boldsymbol{A}$ is **diagonalizable**

  $\Leftrightarrow$ $\boldsymbol{A}$ has n **linearly independent eigenvectors**

  $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ (associated to $\lambda_1, \ldots, \lambda_n$).

  Then $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{D}$, where

  - $\boldsymbol{P} = (\boldsymbol{v}_1 \quad \boldsymbol{v}_2 \quad \cdots \quad \boldsymbol{v}_n)$, $\boldsymbol{D} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$

- In order to find $\boldsymbol{P}^{-1}$, we may need:
  - Gauss-Jordan elimination: $(\boldsymbol{P} \mid \boldsymbol{I}) \dashrightarrow (\boldsymbol{I} \mid \boldsymbol{P}^{-1})$.
  - Adjoint matrix: $\boldsymbol{P}^{-1} = \dfrac{1}{\det(\boldsymbol{P})} \mathbf{adj}(\boldsymbol{P})$.
- **Note**: If $\boldsymbol{P}$ is **orthogonal**, then $\boldsymbol{P}^{-1} = \boldsymbol{P}^{\mathrm{T}}$.

Introduction

- Let $\boldsymbol{B} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$. Then it can be diagonalized by
  - $\boldsymbol{P} = \begin{pmatrix} 1 & -1 & -1 \\ 1 & 1 & -1 \\ 1 & 0 & 2 \end{pmatrix}$.

  We can verify that the columns of $\boldsymbol{P}$, which are eigenvectors of $\boldsymbol{B}$, form an **orthogonal** basis for $\mathbb{R}^3$.
  - Normalizing:
    - $\boldsymbol{R} = \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix}$.
  - $\boldsymbol{R}$ is an orthogonal matrix, which also diagonalizes $\boldsymbol{B}$.

55 / 96

Definition

- **Definition.** A square matrix $\boldsymbol{A}$ is called **orthogonally diagonalizable** if it can be diagonalized by an **orthogonal** matrix. That is,
  - there exists an **orthogonal** matrix $\boldsymbol{P}$ such that
    - $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}$ $(= \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})$ is a **diagonal** matrix.

  $\boldsymbol{P}$ is said to **orthogonally diagonalize** $\boldsymbol{A}$.
- **Remarks.** For any eigenvalue λ of $\boldsymbol{A}$, we can always choose an orthonormal basis for the associated eigenspace E_λ.

  Suppose further that $\boldsymbol{A}$ is **orthogonally diagonalizable**.
  - Then $\boldsymbol{A}$ is diagonalizable, and $\boldsymbol{A}$ has n linearly independent eigenvectors.
  - For distinct eigenvalues $\lambda \neq \mu$,
    - Every eigenvector of λ is orthogonal to that of μ.

56 / 96

Classification

- **Theorem**. A square matrix is **orthogonally diagonalizable**

  $\Leftrightarrow$ it is a **symmetric** matrix.
- **Proof**. $(\Rightarrow)$ Suppose $\boldsymbol{A}$ is orthogonally diagonalizable.
  - There is an orthogonal matrix $\boldsymbol{P}$ & a diagonal matrix $\boldsymbol{D}$
    - such that $\boldsymbol{D} = \boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}$.

    Since $\boldsymbol{D}$ is diagonal, it is also symmetric.
    - $\boldsymbol{D} = \boldsymbol{D}^{\mathrm{T}} = (\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P})^{\mathrm{T}} = \boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}$.

    Therefore, $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}$.
  - Note that both $\boldsymbol{P}$ and $\boldsymbol{P}^{\mathrm{T}}$ are invertible.
    - By Cancellation Law: $\boldsymbol{A} = \boldsymbol{A}^{\mathrm{T}}$.

    That is, $\boldsymbol{A}$ is symmetric.

  $(\Leftarrow)$ is left in MA2101 Linear Algebra II.

Algorithm

- **Algorithm**. (Orthogonally diagonalize symmetric matrix).

  Let $\boldsymbol{A}$ be a **symmetric** matrix of order n.
  1. Find all distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$.
  2. For each eigenvalue λ_i, find an **orthonormal** basis for the eigenspace E_{λ_i}.
     (i) Find a basis S_{λ_i} for E_{λ_i}.
     (ii) Use Gram-Schmidt process to transfer S_{λ_i} to an orthonormal basis T_{λ_i} for E_{λ_i}.
  3. Let $T = T_{\lambda_1} \cup T_{\lambda_2} \cup \cdots \cup T_{\lambda_k}$,
     - $T = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is an orthonormal basis for $\mathbb{R}^n$.

     $\boldsymbol{P} = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{pmatrix}$ orthogonally diagonalizes $\boldsymbol{A}$.

Algorithm

- Compare with the **algorithm for diagonalization**:

  Let $\boldsymbol{A}$ be a square matrix of order n.

  1. Find all distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$.
  2. For each eigenvalue λ_i, find a basis for the eigenspace E_{λ_i}.
  3. Let $S = S_{\lambda_1} \cup S_{\lambda_2} \cup \cdots \cup S_{\lambda_k}$.
     (i) If $|S| < n$, then $\boldsymbol{A}$ is not diagonalizable.
     (ii) If $|S| = n$, say $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$,
       - $\boldsymbol{P} = \begin{pmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n \end{pmatrix}$ diagonalizes $\boldsymbol{A}$.

Algorithm

- **Remarks.** Let $\boldsymbol{A}$ be a **symmetric** matrix of order n.

  1. Every eigenvalue of $\boldsymbol{A}$ is a real number.
  2. Write the characteristic polynomial
     - $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = (\lambda - \lambda_1)^{r_1} \cdots (\lambda - \lambda_k)^{r_k}$,
       - where $\lambda_1, \ldots, \lambda_k$ are distinct eigenvalues.

       Then $\dim E_{\lambda_1} = r_1, \ldots, \dim E_{\lambda_k} = r_k$.

     $\therefore$ $\dim E_{\lambda_1} + \cdots + \dim E_{\lambda_k} = r_1 + \cdots + r_k = n$.
  3. If each basis S_{λ_i} for E_{λ_i} is orthonormal, then
     - $T = S_{\lambda_1} \cup \cdots \cup S_{\lambda_k} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is an orthonormal set. (Exercise.)
     - $\boldsymbol{P} = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{pmatrix}$ is an orthogonal matrix.

Examples

- Let $\boldsymbol{A} = \begin{pmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix}$.
  1. Find eigenvalues: For 2×2 matrix,
     - $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = \lambda^2 - \mathrm{tr}(\boldsymbol{A})\lambda + \det(\boldsymbol{A})$.
     - $\lambda^2 - 2\lambda + \frac{3}{4} = (\lambda - \frac{1}{2})(\lambda - \frac{3}{2})$.

     $\therefore$ $\lambda = \frac{1}{2}$ and $\lambda = \frac{3}{2}$.
  2. Find eigenvectors. For $\lambda = \frac{1}{2}$,
     - Solve $(\lambda \boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$:
       - $\begin{pmatrix} -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.
     - $\begin{pmatrix} x \\ y \end{pmatrix} = t \begin{pmatrix} 1 \\ 1 \end{pmatrix} \overset{\text{normalizing}}{\Longrightarrow} \boldsymbol{v}_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$.

61 / 96

Examples

- Let $\boldsymbol{A} = \begin{pmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix}$.
  1. Find eigenvalues: For 2×2 matrix,
     - $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = \lambda^2 - \mathrm{tr}(\boldsymbol{A})\lambda + \det(\boldsymbol{A})$.
     - $\lambda^2 - 2\lambda + \frac{3}{4} = (\lambda - \frac{1}{2})(\lambda - \frac{3}{2})$.

     $\therefore$ $\lambda = \frac{1}{2}$ and $\lambda = \frac{3}{2}$.
  2. Find eigenvectors. For $\lambda = \frac{3}{2}$,
     - Solve $(\lambda \boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$:
       - $\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.
     - $\begin{pmatrix} x \\ y \end{pmatrix} = t \begin{pmatrix} -1 \\ 1 \end{pmatrix} \overset{\text{normalizing}}{\Longrightarrow} \boldsymbol{v}_2 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$.

62 / 96

Examples

- Let $\boldsymbol{A} = \begin{pmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix}$.

1. Find eigenvalues: For 2×2 matrix,
   - $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = \lambda^2 - \operatorname{tr}(\boldsymbol{A})\lambda + \det(\boldsymbol{A})$.
   - $\lambda^2 - 2\lambda + \frac{3}{4} = (\lambda - \frac{1}{2})(\lambda - \frac{3}{2})$.

   $\therefore$ $\lambda = \frac{1}{2}$ and $\lambda = \frac{3}{2}$.

3. Let $\boldsymbol{P} = \begin{pmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$. Then
   - $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{3}{2} \end{pmatrix}$.

Examples

- Let $\boldsymbol{B} = \begin{pmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 3 & 1 \\ -1 & 1 & 1 & 3 \end{pmatrix}$.

1. Find the eigenvalues. The characteristic polynomial

$$\begin{aligned} & \det(\lambda \boldsymbol{I} - \boldsymbol{B}) \\ = & \det \begin{pmatrix} \lambda - 1 & 1 & -1 & 1 \\ 1 & \lambda - 1 & 1 & -1 \\ -1 & 1 & \lambda - 3 & -1 \\ 1 & -1 & -1 & \lambda - 3 \end{pmatrix} \\ = & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ = & \lambda^4 - 8\lambda^3 + 16\lambda^2 = \lambda^2(\lambda - 4)^2. \end{aligned}$$

   - The eigenvalues are $\lambda = 0$ and $\lambda = 4$.

Examples

- Let $\boldsymbol{B} = \begin{pmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 3 & 1 \\ -1 & 1 & 1 & 3 \end{pmatrix}$.

  2. Find the eigenvectors. Let $\lambda = 0$. Solve

     ○ $(\lambda \boldsymbol{I} - \boldsymbol{B})\boldsymbol{x} = \boldsymbol{0}$.

     $$\begin{pmatrix} -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & -3 & -1 \\ 1 & -1 & -1 & -3 \end{pmatrix} \begin{pmatrix} w \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

     - $\begin{pmatrix} w \\ x \\ y \\ z \end{pmatrix} = s \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 2 \\ 0 \\ -1 \\ 1 \end{pmatrix}, s, t \in \mathbb{R}$.

Examples

- Let $\boldsymbol{B} = \begin{pmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 3 & 1 \\ -1 & 1 & 1 & 3 \end{pmatrix}$.

  2. Find the eigenvectors. Let $\lambda = 0$. Set

     ○ $\boldsymbol{u}_1 = (1, 1, 0, 0)$ and $\boldsymbol{u}_2 = (2, 0, -1, 1)$.

     $$\begin{aligned} \boldsymbol{v}_1 &= \boldsymbol{u}_1 = (1, 1, 0, 0) \\ \boldsymbol{v}_2 &= \boldsymbol{u}_2 - \frac{\boldsymbol{u}_2 \cdot \boldsymbol{v}_1}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_2}\boldsymbol{v}_1 = (1, -1, -1, 1). \end{aligned}$$

     ○ Normalizing:

     $$\begin{aligned} \boldsymbol{w}_1 &= \frac{\boldsymbol{v}_1}{\|\boldsymbol{v}_1\|} = (\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}, 0, 0) \\ \boldsymbol{w}_2 &= \frac{\boldsymbol{v}_2}{\|\boldsymbol{v}_2\|} = (\tfrac{1}{2}, -\tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2}). \end{aligned}$$

Examples

- Let $\boldsymbol{B} = \begin{pmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 3 & 1 \\ -1 & 1 & 1 & 3 \end{pmatrix}$.

  2. Find the eigenvectors. Let $\lambda = 4$. Solve
     - $(\lambda \boldsymbol{I} - \boldsymbol{B})\boldsymbol{x} = \boldsymbol{0}$.
       $\begin{pmatrix} 3 & 1 & -1 & 1 \\ 1 & 3 & 1 & -1 \\ -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} w \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$
       - $\begin{pmatrix} w \\ x \\ y \\ z \end{pmatrix} = s \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 0 \\ 1 \end{pmatrix}$, $s, t \in \mathbb{R}$.

Examples

- Let $\boldsymbol{B} = \begin{pmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 3 & 1 \\ -1 & 1 & 1 & 3 \end{pmatrix}$.

  2. Find the eigenvectors. Let $\lambda = 4$. Set
     - $\boldsymbol{u}_3 = (\frac{1}{2}, -\frac{1}{2}, 1, 0)$ and $\boldsymbol{u}_4 = (-\frac{1}{2}, \frac{1}{2}, 0, 1)$.

$$\begin{aligned} \boldsymbol{v}_3 &= \boldsymbol{u}_3 = (\tfrac{1}{2}, -\tfrac{1}{2}, 1, 0) \\ \boldsymbol{v}_4 &= \boldsymbol{u}_4 - \frac{\boldsymbol{u}_4 \cdot \boldsymbol{v}_3}{\boldsymbol{v}_3 \cdot \boldsymbol{v}_3}\boldsymbol{v}_3 = (-\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}, 1). \end{aligned}$$

     - Normalizing:

$$\begin{aligned} \boldsymbol{w}_3 &= \frac{\boldsymbol{v}_3}{\|\boldsymbol{v}_3\|} = (\tfrac{1}{\sqrt{6}}, -\tfrac{1}{\sqrt{6}}, \tfrac{2}{\sqrt{6}}, 0) \\ \boldsymbol{w}_4 &= \frac{\boldsymbol{v}_2}{\|\boldsymbol{v}_2\|} = (-\tfrac{1}{\sqrt{12}}, \tfrac{1}{\sqrt{12}}, \tfrac{1}{\sqrt{12}}, \tfrac{3}{\sqrt{12}}). \end{aligned}$$

Examples

- Let $\boldsymbol{B} = \begin{pmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 3 & 1 \\ -1 & 1 & 1 & 3 \end{pmatrix}$.

  3. Let $\boldsymbol{P} = \begin{pmatrix} \boldsymbol{w}_1 & \boldsymbol{w}_2 & \boldsymbol{w}_3 & \boldsymbol{w}_4 \end{pmatrix}$.

     ◦ $\boldsymbol{P} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{2} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{12}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{2} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} \\ 0 & -\frac{1}{2} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{12}} \\ 0 & \frac{1}{2} & 0 & \frac{3}{\sqrt{12}} \end{pmatrix}$.

     Then $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$ is diagonal.

Quadratic Forms and Conic Sections

Quadratic Form

- A **homogeneous** polynomial in degree 2 in variables x, y:
  - $f(x, y) = ax^2 + bxy + cy^2$, $\quad a, b, c$ are real constants.

  It is known as a **quadratic form** in variables x, y.

- **Definition.** A **quadratic form** in n variables $x_1, \ldots, x_n$ is
  - $Q(x_1, \ldots, x_n) = \sum_{i=1}^{n} q_{ii}x_i^2 + \sum_{i<j} q_{ij}x_ix_j$.

- **Examples.**
  - $Q(x, y) = x^2 + y^2 - xy$.
  - $Q(x, y, z) = x^2 + 2y^2 + 3z^2 + 4xy + 5xz + 6yz$.
  - $Q(x_1, \ldots, x_n) = x_1^2 + x_2^2 + \cdots + x_n^2$.

Quadratic Form

- Let $\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$, $\boldsymbol{A} = (a_{ij})_{n\times n}$ a symmetric matrix.
  - $\boldsymbol{A}\boldsymbol{x} = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n a_{1j}x_j \\ \vdots \\ \sum_{j=1}^n a_{nj}x_j \end{pmatrix}$.

$$\begin{aligned} \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} &= (x_1, \ldots, x_n) \begin{pmatrix} \sum_{j=1}^n a_{1j}x_j \\ \vdots \\ \sum_{j=1}^n a_{nj}x_j \end{pmatrix} \\ &= \sum_{i=1}^n x_i \left(\sum_{j=1}^n a_{ij}x_j \right) \\ &= \sum_{i=1}^n \sum_{j=1}^n a_{ij}x_ix_j \\ &= \sum_{i=1}^n a_{ii}x_i^2 + \sum_{i<j} 2a_{ij}x_ix_j. \end{aligned}$$

Quadratic Form

- $Q(x_1, \ldots, x_n) = \sum_{i=1}^{n} q_{ii}x_i^2 + \sum_{i<j} q_{ij}x_ix_j$ is a quadratic form.
  - Let $\boldsymbol{x} = (x_1, \ldots, x_n)^{\mathrm{T}}$ and $\boldsymbol{A} = (a_{ij})_{n\times n}$ be defined by
    - $a_{ii} = q_{ii}$ and $a_{ij} = a_{ji} = \frac{1}{2}q_{ij}$ for $i < j$.

    Then $Q(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$, $\boldsymbol{x} \in \mathbb{R}^n$.
- **Examples.**
  - $Q(x, y) = 2x^2 + 3y^2$ is a quadratic form in x and y.
    - Let $\boldsymbol{x} = \begin{pmatrix} x \\ y \end{pmatrix}$ and $\boldsymbol{A} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$.
    - Then $Q(x, y) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$.
  - $Q(x, y) = x^2 + y^2 - xy$ is a quadratic form in x and y.
    - Let $\boldsymbol{x} = \begin{pmatrix} x \\ y \end{pmatrix}$ and $\boldsymbol{A} = \begin{pmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix}$.
    - Then $Q(x, y) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$.

Quadratic Form

- $Q(x_1, \ldots, x_n) = \sum_{i=1}^{n} q_{ii} x_i^2 + \sum_{i<j} q_{ij} x_i x_j$ is a quadratic form.
  - Let $\boldsymbol{x} = (x_1, \ldots, x_n)^{\mathrm{T}}$ and $\boldsymbol{A} = (a_{ij})_{n\times n}$ be defined by
    - $a_{ii} = q_{ii}$ and $a_{ij} = a_{ji} = \frac{1}{2} q_{ij}$ for $i < j$.

    Then $Q(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x}$, $\boldsymbol{x} \in \mathbb{R}^n$.
- **Examples.**
  - $Q(x, y, z) = x^2 + 2y^2 + 3z^2 + 4xy + 5xz + 6yz.$
    - It is a quadratic form in variables x, y, z.
    - Let $\boldsymbol{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ and $\boldsymbol{A} = \begin{pmatrix} 1 & 2 & \frac{5}{2} \\ 2 & 2 & 3 \\ \frac{5}{2} & 3 & 3 \end{pmatrix}$.
    - Then $Q(x, y, z) = \boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x}$.

Simplification

- Suppose the quadratic form is presented as
  - $Q(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x}, \quad \boldsymbol{x} = (x_1, \ldots, x_n)^{\mathrm{T}} \in \mathbb{R}^n,$

  where $\boldsymbol{A}$ is a **symmetric** matrix of order n.
- Recall that $\boldsymbol{A}$ is **orthogonally diagonalizable**.
  - There exists an orthogonal matrix $\boldsymbol{P}$ such that
    - $\boldsymbol{P}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{P} = \boldsymbol{D} = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$.

    Let $\boldsymbol{y} = \boldsymbol{P}^{\mathrm{T}} \boldsymbol{x} = (y_1, \ldots, y_n)^{\mathrm{T}} \in \mathbb{R}^n$. Then $\boldsymbol{x} = \boldsymbol{P} \boldsymbol{y}$.

$$\begin{aligned} Q(\boldsymbol{x}) &= (\boldsymbol{P}\boldsymbol{y})^{\mathrm{T}} \boldsymbol{A} (\boldsymbol{P}\boldsymbol{y}) = \boldsymbol{y}^{\mathrm{T}} (\boldsymbol{P}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{P}) \boldsymbol{y} \\ &= \begin{pmatrix} y_1 & \cdots & y_n \end{pmatrix} \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \\ &= \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2. \end{aligned}$$

Examples

- Let $Q(x,y) = x^2 - xy + y^2$.
  - $Q(x,y) = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$.
  - Orthogonally diagonalize $\begin{pmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix}$.
    - $\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{3}{2} \end{pmatrix}$.
  - Let $\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}}(x+y) \\ \frac{1}{\sqrt{2}}(-x+y) \end{pmatrix}$.

$$\begin{aligned} Q(x,y) &= \tfrac{1}{2}(x')^2 + \tfrac{3}{2}(y')^2 \\ &= \tfrac{1}{4}(x+y)^2 + \tfrac{3}{4}(-x+y)^2. \end{aligned}$$

Examples

- Let $Q(x,y,z) = x^2 + 2y^2 + z^2 + 2xz$.
  - $Q(x,y,z) = \begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$.
  - Orthogonally diagonalize $\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$.
    - $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{D} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \boldsymbol{P} = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}$.
  - Let $\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \boldsymbol{P}^{\mathrm{T}}\boldsymbol{x} = \begin{pmatrix} \frac{1}{\sqrt{2}}(x+z) \\ y \\ \frac{1}{\sqrt{2}}(-x+z) \end{pmatrix}$.
    - $Q(x,y,z) = 2(x')^2 + 2(y')^2 + 0(z')^2 = (x+z)^2 + 2y^2$.

Quadratic Equation

- A **quadratic equation** in variable x is of the form
  - $ax^2 + bx = c$.
- **Definition.** A **quadratic equation** in variables x and y is
  - $ax^2 + bxy + cy^2 + dx + ey = f$.

  The graph of a quadratic equation is a **conic section**.
- **Note.** Let $\boldsymbol{x} = \begin{pmatrix} x \\ y \end{pmatrix}$, $\boldsymbol{A} = \begin{pmatrix} a & \frac{1}{2}b \\ \frac{1}{2}b & c \end{pmatrix}$ and $\boldsymbol{b} = \begin{pmatrix} d \\ e \end{pmatrix}$.
  - $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{x} = f$.
- **Definition.** $ax^2 + bxy + cy^2 = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ is the quadratic form **associated** with the quadratic equation.
  - $ax^2 + bxy + cy^2 + dx + ey = f$.

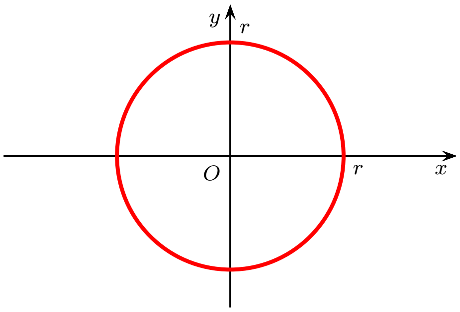
Classification of Conics

- Classification of **conic sections**.
  - **Degenerated** conic sections.
    - The whole plane $\mathbb{R}^2$: $0 = 0$.
    - Empty set: $x^2 + y^2 = -1$.
    - A point: $x^2 + y^2 = 0$.
    - A line: $x = 0$ or $x^2 = 0$.
    - A pair of distinct lines: $x^2 - y^2 = 0$.
  - **Non-degenerated** conic sections.
    - Circle: $x^2 + y^2 = 1$.
    - Ellipse: $x^2 + 2y^2 = 1$.
    - Hyperbola: $x^2 - y^2 = 1$.
    - Parabola: $x^2 - y = 0$.

Standard Forms

- Standard form of **circle** or **ellipse**:
  - $\dfrac{x^2}{\alpha^2} + \dfrac{y^2}{\beta^2} = 1, \quad \alpha > 0, \beta > 0.$
    - $\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} \frac{1}{\alpha^2} & 0 \\ 0 & \frac{1}{\beta^2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 1.$

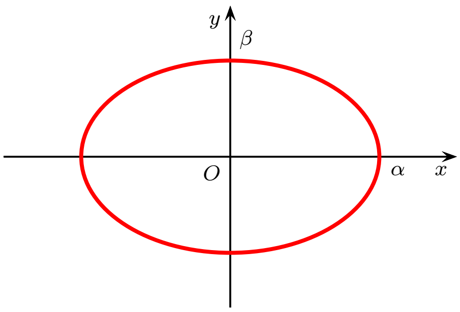
    If $\alpha = \beta$, it is a circle of radius $r = \alpha = \beta$.



Standard Forms

- Standard form of **circle** or **ellipse**:
  - $\dfrac{x^2}{\alpha^2} + \dfrac{y^2}{\beta^2} = 1, \quad \alpha > 0, \beta > 0.$
    - $\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} \frac{1}{\alpha^2} & 0 \\ 0 & \frac{1}{\beta^2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 1.$

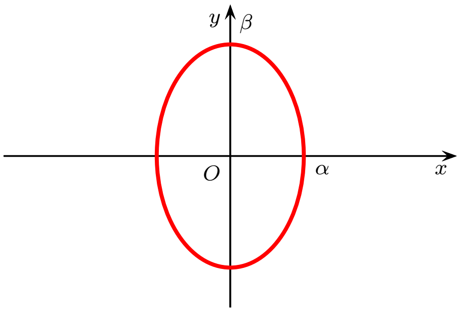
    If $\alpha > \beta$, ellipse of major radius α, minor radius β:



Standard Forms

- Standard form of **circle** or **ellipse**:
  - $\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} = 1, \quad \alpha > 0, \beta > 0.$
    - $\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} \frac{1}{\alpha^2} & 0 \\ 0 & \frac{1}{\beta^2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 1.$

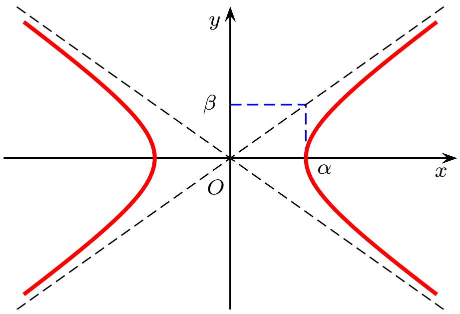
    If $\alpha < \beta$, ellipse of major radius β, minor radius α:



Standard Forms

- Standard form of **hyperbola**:
  - Case 1: $\frac{x^2}{\alpha^2} - \frac{y^2}{\beta^2} = 1, \quad \alpha > 0, \beta > 0.$
    - $\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} \frac{1}{\alpha^2} & 0 \\ 0 & -\frac{1}{\beta^2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 1.$

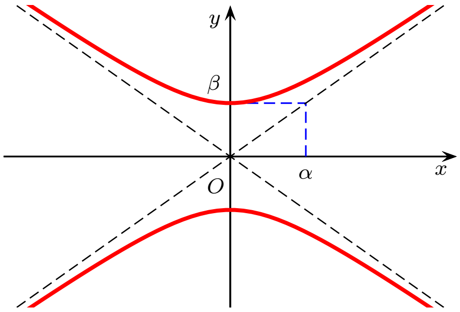
    Semi-major axis α and semi-minor axis β.



Standard Forms

- Standard form of **hyperbola**:
  - Case 2: $-\dfrac{x^2}{\alpha^2}+\dfrac{y^2}{\beta^2}=1, \quad \alpha>0, \beta>0.$
    - $\begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} -\frac{1}{\alpha^2} & 0 \\ 0 & \frac{1}{\beta^2} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}=1.$

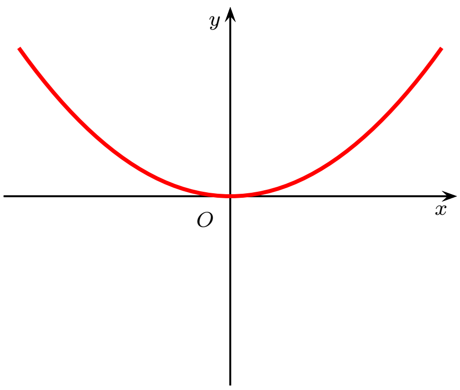
    Semi-major axis β and semi-minor axis α.



Standard Forms

- Standard form of **parabola**:
  - Case 1: $x^2=\alpha y, \quad |\alpha|/4\neq 0$ is the focal length.
    - $\begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}+\begin{pmatrix} 0 & -\alpha \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}=0.$

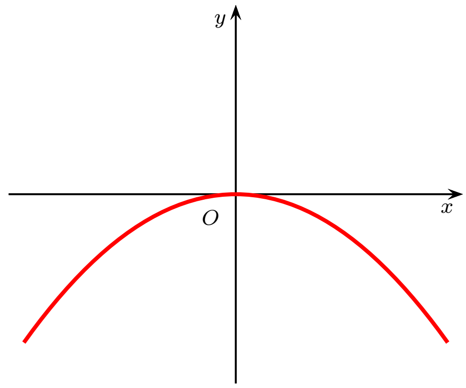
    Suppose that $\alpha>0$.



Standard Forms

- Standard form of **parabola**:
  - Case 1: $x^2 = \alpha y$, $|\alpha|/4 \neq 0$ is the focal length.
    - $\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 & -\alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0.$

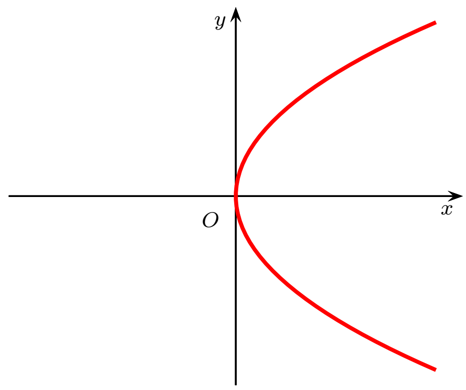
    Suppose that $\alpha < 0$.



Standard Forms

- Standard form of **parabola**:
  - Case 2: $y^2 = \alpha x$, $|\alpha|/4 \neq 0$ is the focal length.
    - $\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} -\alpha & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0.$

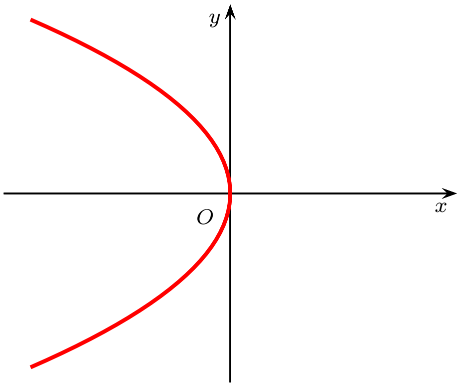
    Suppose that $\alpha > 0$.



Standard Forms

- Standard form of **parabola**:
  - Case 2: $y^2 = \alpha x$, $|\alpha|/4 \neq 0$ is the focal length.
    - $\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} -\alpha & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0.$

    Suppose that $\alpha < 0$.



Classification

- Classify $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{x} = f$, $\boldsymbol{x} \in \mathbb{R}^2$.
  1. Orthogonally diagonalize $\boldsymbol{A}$.
     - $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$, $\boldsymbol{P}$ an orthogonal matrix.
  2. Let $\boldsymbol{y} = \boldsymbol{P}^{\mathrm{T}}\boldsymbol{x}$. Then
     - $\boldsymbol{y}^{\mathrm{T}} \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} \boldsymbol{y} + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{y} = f.$
  3. Complete the squares.
- **Remark.** λ and μ are eigenvalues of $\boldsymbol{A}$; $\lambda\mu = \det(\boldsymbol{A})$.
  - Suppose the conic section is **non-degenerate**.
    - $\det(\boldsymbol{A}) > 0 \Leftrightarrow$ ellipse (or circle).
    - $\det(\boldsymbol{A}) < 0 \Leftrightarrow$ hyperbola.
    - $\det(\boldsymbol{A}) = 0 \Leftrightarrow$ parabola.

Examples

- $x^2 - xy + y^2 - x - y = 1.$

  Let $\boldsymbol{x} = \begin{pmatrix} x \\ y \end{pmatrix}$, $\boldsymbol{A} = \begin{pmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix}$ and $\boldsymbol{b} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$.
  - $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{x} = 1.$

1. Orthogonally diagonalize $\boldsymbol{A}$.
   - $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{3}{2} \end{pmatrix}$, where $\boldsymbol{P} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$.
2. Let $\begin{pmatrix} x' \\ y' \end{pmatrix} = \boldsymbol{y} = \boldsymbol{P}^{\mathrm{T}}\boldsymbol{x} = \begin{pmatrix} \frac{1}{\sqrt{2}}(x+y) \\ \frac{1}{\sqrt{2}}(-x+y) \end{pmatrix}$.
   - $\boldsymbol{y}^{\mathrm{T}} \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{3}{2} \end{pmatrix} \boldsymbol{y} + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{y} = 1.$

Examples

- $x^2 - xy + y^2 - x - y = 1.$

  Let $\boldsymbol{x} = \begin{pmatrix} x \\ y \end{pmatrix}$, $\boldsymbol{A} = \begin{pmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix}$ and $\boldsymbol{b} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$.
  - $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{x} = 1.$

3. $\frac{1}{2}(x')^2 + \frac{3}{2}(y')^2 - \sqrt{2}\,(x') = 1.$
   - $\frac{1}{2}(x' - \sqrt{2})^2 + \frac{3}{2}(y')^2 = 1 + \frac{1}{2}(\sqrt{2})^2 = 2.$
   - $\dfrac{(x' - \sqrt{2})^2}{2^2} + \dfrac{(y')^2}{(2/\sqrt{3})^2} = 1.$

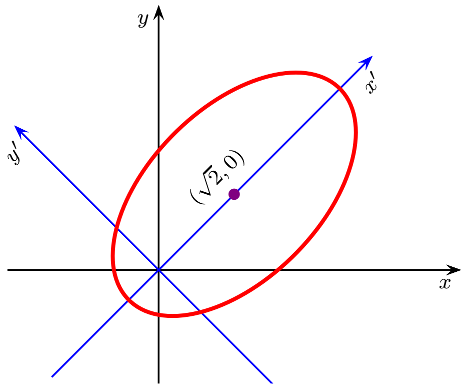
Note that $\boldsymbol{P} = \begin{pmatrix} \cos\frac{\pi}{4} & -\sin\frac{\pi}{4} \\ \sin\frac{\pi}{4} & \cos\frac{\pi}{4} \end{pmatrix}$, $\begin{pmatrix} x' \\ y' \end{pmatrix} = \boldsymbol{P}^{\mathrm{T}} \begin{pmatrix} x \\ y \end{pmatrix}$.

- The x'- and y'-axis is obtained by rotating the x- and y-axis about the origin O anticlockwise by $\pi/4$.

Examples

- $x^2 - xy + y^2 - x - y = 1.$
  - $\dfrac{(x' - \sqrt{2})^2}{2^2} + \dfrac{(y')^2}{(2/\sqrt{3})^2} = 1.$

  The x'- and y'-axis is obtained by rotating the x- and y-axis about the origin O anticlockwise by $\pi/4$.



Examples

- $2x^2 + 24xy + 9y^2 + 20x - 6y = 5.$

  Let $\boldsymbol{x} = \begin{pmatrix} x \\ y \end{pmatrix}$, $\boldsymbol{A} = \begin{pmatrix} 2 & 12 \\ 12 & 9 \end{pmatrix}$ and $\boldsymbol{b} = \begin{pmatrix} 20 \\ -6 \end{pmatrix}$.
  - $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{x} = 5.$

  1. Orthogonally diagonalize $\boldsymbol{A}$ (Exercise).
     - $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} 18 & 0 \\ 0 & -7 \end{pmatrix}$, where $\boldsymbol{P} = \begin{pmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix}$.
  2. Let $\begin{pmatrix} x' \\ y' \end{pmatrix} = \boldsymbol{y} = \boldsymbol{P}^{\mathrm{T}}\boldsymbol{x} = \begin{pmatrix} \frac{3}{5}x + \frac{4}{5}y \\ -\frac{4}{5}x + \frac{3}{5}y \end{pmatrix}$.
     - $\boldsymbol{y}^{\mathrm{T}} \begin{pmatrix} 18 & 0 \\ 0 & -7 \end{pmatrix} \boldsymbol{y} + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{y} = 5.$

Examples

- $2x^2 + 24xy + 9y^2 + 20x - 6y = 5$.

  Let $\boldsymbol{x} = \begin{pmatrix} x \\ y \end{pmatrix}$, $\boldsymbol{A} = \begin{pmatrix} 2 & 12 \\ 12 & 9 \end{pmatrix}$ and $\boldsymbol{b} = \begin{pmatrix} 20 \\ -6 \end{pmatrix}$.

  - $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{x} = 5$.

  3. $18(x')^2 - 7(y')^2 + \frac{36}{5}x' - \frac{98}{5}y' = 5$.
     - $18(x' + \frac{1}{5})^2 - 7(y' + \frac{7}{5})^2 = -8$.
     - $-\dfrac{(x' + \frac{1}{5})^2}{(2/3)^2} + \dfrac{(y' + \frac{7}{5})^2}{(\sqrt{8/7})^2} = 1$.

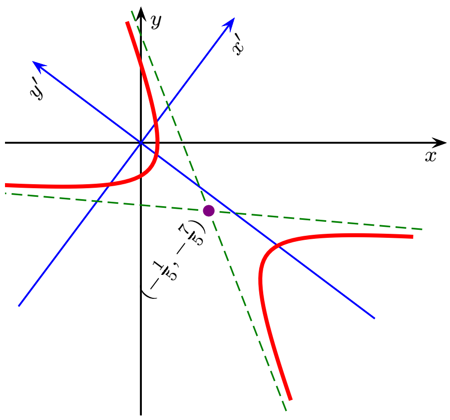
  Note that $\boldsymbol{P} = \begin{pmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix}$ and $\boldsymbol{y} = \boldsymbol{P}^{\mathrm{T}}\boldsymbol{x}$.

  - The x'- and y'-axis is obtained by rotating the x- and y-axis about the origin O anticlockwise by $\cos^{-1}(\frac{3}{5})$.

Examples

- $2x^2 + 24xy + 9y^2 + 20x - 6y = 5$.
  - $-\dfrac{(x' + \frac{1}{5})^2}{(2/3)^2} + \dfrac{(y' + \frac{7}{5})^2}{(\sqrt{8/7})^2} = 1$.

  The x'- and y'-axis is obtained by rotating the x- and y-axis about the origin O anticlockwise by $\cos^{-1}(\frac{3}{5})$.



Remark

- Let $\boldsymbol{P}$ be orthogonal of order 2. Then $\det(\boldsymbol{P}) = \pm 1$.
  - $\det(\boldsymbol{P}) = 1 \Rightarrow \boldsymbol{P} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$.
    - Let $\boldsymbol{y} = \boldsymbol{P}^{\mathrm{T}}\boldsymbol{x}$. Then the new axes are obtained by rotating the original axes about O anticlockwise by θ.
  - $\det(\boldsymbol{P}) = -1 \Rightarrow \boldsymbol{P} = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$.
    - $\boldsymbol{P} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.
    - Let $\boldsymbol{y} = \boldsymbol{P}^{\mathrm{T}}\boldsymbol{x}$. Then the new axes are obtained by first rotating the original axes about O anticlockwise by θ, then reflecting w.r.t. the x'-axis.

  By multiplying the 2nd column of $\boldsymbol{P}$ by -1 if necessary, we can always diagonalize a symmetric $\boldsymbol{A}$ by an orthogonal matrix with determinant 1.