Linear Algebra

Lecture 9 - Diagonalizable Matrices and Application

Oskar Kędzierski

4 December 2023

Diagonal Matrix

Definition

The matrix $A = [a_{ij}] \in M(n \times n; \mathbb{R})$ is called **diagonal** if $a_{ij} = 0$ for any $i \neq j$, i.e.

$$A = \begin{bmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{bmatrix}.$$

Diagonal Matrix

Definition

The matrix $A = [a_{ij}] \in M(n \times n; \mathbb{R})$ is called **diagonal** if $a_{ij} = 0$ for any $i \neq j$, i.e.

$$A = \begin{bmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{bmatrix}.$$

Example

The matrices

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}$$

are diagonal.

Diagonal Matrix of Linear Endomorphism

Proposition

Let $\varphi\colon V \longrightarrow V$ *be an endomorphism of vector space* V *and let* $\mathcal{A} = (v_1, \ldots, v_n)$ *be an ordered basis of* V. *Then* $M(\varphi)_{\mathcal{A}} = [a_{ij}]$ *is diagonal if and only if* v_i *is an eigenvector of* φ. *Moreover, in such case eigenvector* v_i *is associated to the eigenvalue* a_{ii}, *i.e.* $\varphi(v_i) = a_{ii} v_i$.

Diagonal Matrix of Linear Endomorphism

Proposition

Let $\varphi\colon V \longrightarrow V$ be an endomorphism of vector space V and let $\mathcal{A} = (v_1, \ldots, v_n)$ be an ordered basis of V. Then $M(\varphi)_{\mathcal{A}} = [a_{ij}]$ is diagonal if and only if v_i is an eigenvector of φ. Moreover, in such case eigenvector v_i is associated to the eigenvalue a_{ii}, i.e. $\varphi(v_i) = a_{ii}v_i$.

Proof.

$(\Leftarrow)$ Assume each v_i is an eigenvector of φ associated to eigenvalue α_i. Then

$$\varphi(v_i) = \alpha_i v_i = 0v_1 + 0v_2 + \ldots + 0v_{i-1} + \alpha_i v_i + 0v_{i+1} + \ldots + 0v_n,$$

i.e. in the i-th column of the matrix $M(\varphi)_{\mathcal{A}}$ there is α_i in the i-th row and $0's$ elsewhere.

Diagonal Matrix of Linear Endomorphism

Proposition

Let $\varphi\colon V \longrightarrow V$ be an endomorphism of vector space V and let $\mathcal{A} = (v_1, \ldots, v_n)$ be an ordered basis of V. Then $M(\varphi)_{\mathcal{A}} = [a_{ij}]$ is diagonal if and only if v_i is an eigenvector of φ. Moreover, in such case eigenvector v_i is associated to the eigenvalue a_{ii}, i.e. $\varphi(v_i) = a_{ii} v_i$.

Proof.

($\Leftarrow$) Assume each v_i is an eigenvector of φ associated to eigenvalue α_i. Then

$$\varphi(v_i) = \alpha_i v_i = 0v_1 + 0v_2 + \ldots + 0v_{i-1} + \alpha_i v_i + 0v_{i+1} + \ldots + 0v_n,$$

i.e. in the i-th column of the matrix $M(\varphi)_{\mathcal{A}}$ there is α_i in the i-th row and $0's$ elsewhere.

($\Rightarrow$) similar to the above $\square$

Example

Let $\varphi\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be given by $\varphi((x_1, x_2)) = (8x_1 + 10x_2, -3x_1 - 3x_2)$. Then

$$M(\varphi)_{st} = \begin{bmatrix} 8 & 10 \\ -3 & -3 \end{bmatrix}, \quad w_\varphi(\lambda) = \det \begin{bmatrix} 8-\lambda & 10 \\ -3 & -3-\lambda \end{bmatrix},$$

The characteristic polynomial is $w_\varphi(\lambda) = (8-\lambda)(-3-\lambda) + 30 = \lambda^2 - 5\lambda + 6 = (\lambda - 2)(\lambda - 3)$. There are two eigenvalues $\lambda_1 = 2, \lambda_2 = 3$.

Example

Let $\varphi\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be given by
$\varphi((x_1, x_2)) = (8x_1 + 10x_2, -3x_1 - 3x_2)$. Then

$$M(\varphi)_{st} = \begin{bmatrix} 8 & 10 \\ -3 & -3 \end{bmatrix}, \quad w_\varphi(\lambda) = \det \begin{bmatrix} 8-\lambda & 10 \\ -3 & -3-\lambda \end{bmatrix},$$

The characteristic polynomial is
$w_\varphi(\lambda) = (8-\lambda)(-3-\lambda) + 30 = \lambda^2 - 5\lambda + 6 = (\lambda - 2)(\lambda - 3)$.
There are two eigenvalues $\lambda_1 = 2, \lambda_2 = 3$. In order to get corresponding eigenspaces solve

$$V_{(2)}: \begin{bmatrix} 6 & 10 \\ -3 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \iff x_1 = -\frac{5}{3}x_2,$$

i.e. $V_{(2)} = \{(-\frac{5}{3}x_2, x_2) \in \mathbb{R}^2 \mid x_2 \in \mathbb{R}\} = \mathrm{lin}((-5, 3))$

$$V_{(3)}: \begin{bmatrix} 5 & 10 \\ -3 & -6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \iff x_1 = -2x_2,$$

i.e. $V_{(3)} = \{(-2x_2, x_2) \in \mathbb{R}^2 \mid x_2 \in \mathbb{R}\} = \mathrm{lin}((-2, 1))$

Example (continued)

Recall, $\varphi((x_1, x_2)) = (8x_1 + 10x_2, -3x_1 - 3x_2)$.
The basis $\mathcal{A} = ((-5, 3), (-2, 1))$ of $\mathbb{R}^2$ consists of eigenvectors and

$$M(\varphi)_{\mathcal{A}} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix},$$

since

$$\varphi((-5, 3)) = 2(-5, 3) + 0(-2, 1),$$

$$\varphi((-2, 1)) = 0(-5, 3) + 3(-2, 1).$$

Eigenvectors for Different Eigenvalues

Theorem

Let $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ *be pairwise distinct eigenvalues of the linear endomorphism* $\varphi \colon V \longrightarrow V$. *Let* $\mathcal{A}_i \subset V_{(\alpha_i)}$ *be a finite set of linearly independent eigenvectors of* φ *associated to* α_i *for* $i = 1, \ldots, k$. *Then* $\mathcal{A} = \mathcal{A}_1 \cup \ldots \cup \mathcal{A}_k$ *is a set of linearly independent vectors.*

Eigenvectors for Different Eigenvalues

Theorem

Let $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ be pairwise distinct eigenvalues of the linear endomorphism $\varphi\colon V \longrightarrow V$. Let $\mathcal{A}_i \subset V_{(\alpha_i)}$ be a finite set of linearly independent eigenvectors of φ associated to α_i for $i = 1, \ldots, k$. Then $\mathcal{A} = \mathcal{A}_1 \cup \ldots \cup \mathcal{A}_k$ is a set of linearly independent vectors.

Proof.

For simplicity we assume that $\mathcal{A}_i = \{v_i\}$, i.e. each set $\mathcal{A}_i$ contains one vector. Assume $\gamma_1 v_1 + \gamma_2 v_2 + \ldots + \gamma_k v_k = 0$. By applying φ to both sides we get $\alpha_1\gamma_1 v_1 + \alpha_2\gamma_2 v_2 + \ldots + \alpha_k\gamma_k v_k = 0$. Repeating this procedure we get a system of linear equations:

$$U\colon \left\{ \begin{array}{rcrcccrl} \gamma_1 v_1 & + & \gamma_2 v_2 & + & \ldots & + & \gamma_k v_k & = 0 \\ \alpha_1\gamma_1 v_1 & + & \alpha_2\gamma_2 v_2 & + & \ldots & + & \alpha_k\gamma_k v_k & = 0 \\ \alpha_1^2\gamma_1 v_1 & + & \alpha_2^2\gamma_2 v_2 & + & \ldots & + & \alpha_k^2\gamma_k v_k & = 0 \\ \vdots & & \vdots & & \ddots & & \vdots & \vdots \\ \alpha_1^{k-1}\gamma_1 v_1 & + & \alpha_2^{k-1}\gamma_2 v_2 & + & \ldots & + & \alpha_k^{k-1}\gamma_k v_k & = 0 \end{array} \right.$$

$\square$

Vandermonde Determinant

One can check that the **Vandermonde determinant**

$$\det \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ \alpha_1 & \alpha_2 & \alpha_3 & \dots & \alpha_k \\ \alpha_1^2 & \alpha_2^2 & \alpha_3^2 & \dots & \alpha_k^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{k-1} & \alpha_2^{k-1} & \alpha_3^{k-1} & \dots & \alpha_k^{k-1} \end{bmatrix} = \prod_{1 \leqslant i < j \leqslant k} (\alpha_j - \alpha_i)$$

is non-zero and hence the system U can be brought by elementary row operations to a reduced echelon form

$$U: \begin{cases} \gamma_1 v_1 & & & & = 0 \\ & \gamma_2 v_2 & & & = 0 \\ & & \ddots & & \vdots \\ & & & \gamma_k v_k & = 0 \end{cases}$$

Vandermonde Determinant (continued)

Which implies that $\gamma_1 = \gamma_2 = \ldots = \gamma_k = 0$ since all vectors v_i are non-zero.

Vandermonde Determinant (continued)

Which implies that $\gamma_1 = \gamma_2 = \ldots = \gamma_k = 0$ since all vectors v_i are non-zero. In the general case one can argue in a similar way, replacing $\gamma_i v_i$ with $\gamma_i^{(1)} v_i^{(1)} + \ldots + \gamma_i^{(m_i)} v_i^{(m_i)}$, where $v_i^{(1)}, \ldots, v_i^{(m_i)} \in V_{(\alpha_i)}$ are linearly independent for $i = 1, \ldots, k$.

Vandermonde Determinant (continued)

Which implies that $\gamma_1 = \gamma_2 = \ldots = \gamma_k = 0$ since all vectors v_i are non-zero. In the general case one can argue in a similar way, replacing $\gamma_i v_i$ with $\gamma_i^{(1)} v_i^{(1)} + \ldots + \gamma_i^{(m_i)} v_i^{(m_i)}$, where $v_i^{(1)}, \ldots, v_i^{(m_i)} \in V_{(\alpha_i)}$ are linearly independent for $i = 1, \ldots, k$. Alternatively, assume that $v_1, \ldots, v_k$ are linearly dependent and the linear combination

$$\gamma_1 v_1 + \ldots + \gamma_m v_m = 0,$$

where $\gamma_i \neq 0$ involves the least number of vectors (perhaps after rearranging them). Then, by applying φ to both sides of the equation

$$\gamma_1 \alpha_1 v_1 + \ldots + \gamma_m \alpha_m v_m = 0.$$

By multiplying the first equation by α_m and subtracting it from the latter

$$\gamma_1(\alpha_1 - \alpha_m) v_1 + \ldots + \gamma_{m-1}(\alpha_{m-1} - \alpha_m) v_{m-1} = 0,$$

we get a linear combination involving $m - 1$ vectors, which leads to a contradiction.

Basis Consisting of Eigenvectors

Corollary

Let V be a finite dimensional vector space. Let $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ be pairwise distinct eigenvalues of the linear endomorphism $\varphi\colon V \longrightarrow V$. Then

Basis Consisting of Eigenvectors

Corollary

Let V be a finite dimensional vector space. Let $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ be pairwise distinct eigenvalues of the linear endomorphism $\varphi\colon V \longrightarrow V$. Then

Basis Consisting of Eigenvectors

Corollary

Let V *be a finite dimensional vector space. Let* $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ *be pairwise distinct eigenvalues of the linear endomorphism* $\varphi\colon V \longrightarrow V$. *Then*

Basis Consisting of Eigenvectors

Corollary

Let V be a finite dimensional vector space. Let $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ be pairwise distinct eigenvalues of the linear endomorphism $\varphi\colon V \longrightarrow V$. Then

i) *if $v_1, \dots, v_k \in V$ and $\varphi(v_i) = \alpha_i v_i$, $v_i \neq 0$ for $i = 1, \dots, k$ then the vectors $v_1, \dots, v_k$ are linearly independent,*

Basis Consisting of Eigenvectors

Corollary

Let V be a finite dimensional vector space. Let $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ be pairwise distinct eigenvalues of the linear endomorphism $\varphi \colon V \longrightarrow V$. Then

i) *if $v_1, \dots, v_k \in V$ and $\varphi(v_i) = \alpha_i v_i$, $v_i \neq 0$ for $i = 1, \dots, k$ then the vectors $v_1, \dots, v_k$ are linearly independent,*

ii) $\dim V_{(\alpha_1)} + \dim V_{(\alpha_2)} + \ldots + \dim V_{(\alpha_k)} \leqslant \dim V$,

Basis Consisting of Eigenvectors

Corollary

Let V be a finite dimensional vector space. Let $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ be pairwise distinct eigenvalues of the linear endomorphism $\varphi \colon V \longrightarrow V$. Then

i) *if $v_1, \ldots, v_k \in V$ and $\varphi(v_i) = \alpha_i v_i,\ v_i \neq 0$ for $i = 1, \ldots, k$ then the vectors $v_1, \ldots, v_k$ are linearly independent,*

ii) $\dim V_{(\alpha_1)} + \dim V_{(\alpha_2)} + \ldots + \dim V_{(\alpha_k)} \leqslant \dim V$,

iii) $\dim V_{(\alpha_1)} + \dim V_{(\alpha_2)} + \ldots + \dim V_{(\alpha_k)} = \dim V \iff$ *there exist a basis of V consisting of eigenvectors of $\varphi \iff$ the matrix of φ relative to some basis of V is diagonal.*

Basis Consisting of Eigenvectors

Corollary

Let V be a finite dimensional vector space. Let $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ be pairwise distinct eigenvalues of the linear endomorphism $\varphi \colon V \longrightarrow V$. Then

i) *if $v_1, \dots, v_k \in V$ and $\varphi(v_i) = \alpha_i v_i$, $v_i \neq 0$ for $i = 1, \dots, k$ then the vectors $v_1, \dots, v_k$ are linearly independent,*

ii) $\dim V_{(\alpha_1)} + \dim V_{(\alpha_2)} + \dots + \dim V_{(\alpha_k)} \leqslant \dim V$,

iii) $\dim V_{(\alpha_1)} + \dim V_{(\alpha_2)} + \dots + \dim V_{(\alpha_k)} = \dim V \iff$ *there exist a basis of V consisting of eigenvectors of φ $\iff$ the matrix of φ relative to some basis of V is diagonal.*

In the part iii) of the corollary the basis of V consists of the union of bases of $V_{(\alpha_i)}$ for $i = 1, \dots, k$.

Example

Let $\varphi \colon \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ be given by $\varphi((x_1, x_2, x_3)) = (2x_1 - 2x_2 + x_3, 2x_2 + x_3, 4x_3)$. Then

$$M(\varphi)_{st} = \begin{bmatrix} 2 & -2 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 4 \end{bmatrix}, \qquad w_\varphi(\lambda) = (2-\lambda)^2(4-\lambda).$$

Example

Let $\varphi \colon \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ be given by $\varphi((x_1, x_2, x_3)) = (2x_1 - 2x_2 + x_3, 2x_2 + x_3, 4x_3)$. Then

$$M(\varphi)_{st} = \begin{bmatrix} 2 & -2 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 4 \end{bmatrix}, \quad w_\varphi(\lambda) = (2-\lambda)^2(4-\lambda).$$

The eigenvalues of φ are 2 and 4.

Example

Let $\varphi\colon \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ be given by $\varphi((x_1, x_2, x_3)) = (2x_1 - 2x_2 + x_3, 2x_2 + x_3, 4x_3)$. Then

$$M(\varphi)_{st} = \begin{bmatrix} 2 & -2 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 4 \end{bmatrix}, \quad w_\varphi(\lambda) = (2-\lambda)^2(4-\lambda).$$

The eigenvalues of φ are 2 and 4.

$$V_{(2)}: \begin{bmatrix} 0 & -2 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \iff x_2 = x_3 = 0,$$

$$V_{(2)} = \{(x_1, 0, 0) \in \mathbb{R}^3 \mid x_1 \in \mathbb{R}\} = \operatorname{lin}((1,0,0))$$

Example

Let $\varphi \colon \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ be given by
$\varphi((x_1, x_2, x_3)) = (2x_1 - 2x_2 + x_3, 2x_2 + x_3, 4x_3)$. Then

$$M(\varphi)_{st} = \begin{bmatrix} 2 & -2 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 4 \end{bmatrix}, \qquad w_\varphi(\lambda) = (2-\lambda)^2(4-\lambda).$$

The eigenvalues of φ are 2 and 4.

$$V_{(2)} : \begin{bmatrix} 0 & -2 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \iff x_2 = x_3 = 0,$$

$$V_{(2)} = \{(x_1, 0, 0) \in \mathbb{R}^3 \mid x_1 \in \mathbb{R}\} = \mathrm{lin}((1, 0, 0))$$

$$V_{(4)} : \begin{bmatrix} -2 & -2 & 1 \\ 0 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \iff x_1 = 0 \text{ and } x_3 = 2x_2,$$

$$V_{(4)} = \{(0, x_2, 2x_2) \in \mathbb{R}^3 \mid x_2 \in \mathbb{R}\} = \mathrm{lin}((0, 1, 2))$$

Example (continued)

$$V_{(2)} = \{(x_1, 0, 0) \in \mathbb{R}^3 \mid x_1 \in \mathbb{R}\} = \text{lin}((1, 0, 0))$$

$$V_{(4)} = \{(0, x_2, 2x_2) \in \mathbb{R}^3 \mid x_2 \in \mathbb{R}\} = \text{lin}((0, 1, 2))$$

$\dim V_{(2)} + \dim V_{(4)} = 1 + 1 < 3 = \dim \mathbb{R}^3$, therefore there is no basis of $\mathbb{R}^3$ such that matrix of φ relative to it is diagonal.

Diagonalizable Matrix

Corollary

Let V be a finite dimensional vector space and let $\dim V = n$. If the endomorphism $\varphi \colon V \longrightarrow V$ has n pairwise distinct eigenvalues then there exists a basis of V consisting of eigenvectors.

Diagonalizable Matrix

Corollary

Let V be a finite dimensional vector space and let $\dim V = n$. If the endomorphism $\varphi\colon V \longrightarrow V$ has n pairwise distinct eigenvalues then there exists a basis of V consisting of eigenvectors.

Definition

Let $A \in M(n \times n; \mathbb{R})$. We say the matrix A is **diagonalizable** if it is similar to a diagonal matrix, that is there exists an invertible matrix $C \in M(n \times n; \mathbb{R})$ such that the matrix $C^{-1}AC$ is diagonal.

Diagonalizable Matrix

Corollary

Let V be a finite dimensional vector space and let $\dim V = n$. If the endomorphism $\varphi\colon V \longrightarrow V$ has n pairwise distinct eigenvalues then there exists a basis of V consisting of eigenvectors.

Definition

Let $A \in M(n \times n; \mathbb{R})$. We say the matrix A is **diagonalizable** if it is similar to a diagonal matrix, that is there exists an invertible matrix $C \in M(n \times n; \mathbb{R})$ such that the matrix $C^{-1}AC$ is diagonal.

Proposition

Matrix $A \in M(n \times n; \mathbb{R})$ is diagonalizable $\iff$ there exists a basis of $\mathbb{R}^n$ consisting of eigenvectors of the endomorphism $\varphi\colon \mathbb{R}^n \longrightarrow \mathbb{R}^n$ given by the condition $M(\varphi)_{st} = A$.

Diagonalizable Matrix

Corollary

Let V be a finite dimensional vector space and let $\dim V = n$. If the endomorphism $\varphi\colon V \longrightarrow V$ has n pairwise distinct eigenvalues then there exists a basis of V consisting of eigenvectors.

Definition

Let $A \in M(n \times n; \mathbb{R})$. We say the matrix A is **diagonalizable** if it is similar to a diagonal matrix, that is there exists an invertible matrix $C \in M(n \times n; \mathbb{R})$ such that the matrix $C^{-1}AC$ is diagonal.

Proposition

Matrix $A \in M(n \times n; \mathbb{R})$ is diagonalizable $\iff$ there exists a basis of $\mathbb{R}^n$ consisting of eigenvectors of the endomorphism $\varphi\colon \mathbb{R}^n \longrightarrow \mathbb{R}^n$ given by the condition $M(\varphi)_{st} = A$.
Moreover, if $\mathcal{A}$ is such basis and $C = M(id)^{st}_{\mathcal{A}}$ then the matrix $C^{-1}AC$ is diagonal.

Example

Matrix $A = \begin{bmatrix} 8 & 10 \\ -3 & -3 \end{bmatrix}$ is diagonalizable. Endomorphism $\varphi((x_1, x_2)) = (8x_1 + 10x_2, -3x_1 - 3x_2)$ has two eigenvalues 2 and 3. We have computed $V_{(2)} = \mathrm{lin}((-5, 3))$ and $V_{(3)} = \mathrm{lin}((-2, 1))$. Set $\mathcal{A} = ((-5, 3), (-2, 1))$ and $C = M(id)^{st}_{\mathcal{A}}$.

Example

Matrix $A = \begin{bmatrix} 8 & 10 \\ -3 & -3 \end{bmatrix}$ is diagonalizable. Endomorphism $\varphi((x_1, x_2)) = (8x_1 + 10x_2, -3x_1 - 3x_2)$ has two eigenvalues 2 and 3. We have computed $V_{(2)} = \mathrm{lin}((-5, 3))$ and $V_{(3)} = \mathrm{lin}((-2, 1))$. Set $\mathcal{A} = ((-5, 3), (-2, 1))$ and $C = M(id)^{st}_{\mathcal{A}}$.

$$D = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} = M(\varphi)_{\mathcal{A}} = M(id)^{\mathcal{A}}_{st} M(\varphi)^{st}_{st} M(id)^{st}_{\mathcal{A}}$$

Example

Matrix $A = \begin{bmatrix} 8 & 10 \\ -3 & -3 \end{bmatrix}$ is diagonalizable. Endomorphism $\varphi((x_1, x_2)) = (8x_1 + 10x_2, -3x_1 - 3x_2)$ has two eigenvalues 2 and 3. We have computed $V_{(2)} = \text{lin}((-5, 3))$ and $V_{(3)} = \text{lin}((-2, 1))$. Set $\mathcal{A} = ((-5, 3), (-2, 1))$ and $C = M(id)^{st}_{\mathcal{A}}$.

$$D = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} = M(\varphi)_{\mathcal{A}} = M(id)^{\mathcal{A}}_{st} M(\varphi)^{st}_{st} M(id)^{st}_{\mathcal{A}}$$

$$C = \begin{bmatrix} -5 & -2 \\ 3 & 1 \end{bmatrix}, \quad C^{-1} = \begin{bmatrix} 1 & 2 \\ -3 & -5 \end{bmatrix}$$

Example (continued)

Matrix $A = \begin{bmatrix} 2 & -2 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 4 \end{bmatrix}$ is not diagonalizable. There is no basis of $\mathbb{R}^3$ consisting of eigenvalues of the endomorphism $\varphi((x_1, x_2, x_3)) = (2x_1 - 2x_2 + x_3, 2x_2 + x_3, 4x_3)$.

Example (continued)

Matrix $A = \begin{bmatrix} 2 & -2 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 4 \end{bmatrix}$ is not diagonalizable. There is no basis of $\mathbb{R}^3$ consisting of eigenvalues of the endomorphism $\varphi((x_1, x_2, x_3)) = (2x_1 - 2x_2 + x_3, 2x_2 + x_3, 4x_3)$. Matrix $B = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ is not diagonalizable (over $\mathbb{R}$). It has no (real) eigenvalues.

Application

Proposition

Let $A = \begin{bmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{bmatrix}$ *be a diagonal matrix. Then*

$$A^m = \begin{bmatrix} a_{11}^m & & 0 \\ & \ddots & \\ 0 & & a_{nn}^m \end{bmatrix} \text{ for any } m \in \mathbb{N}.$$

Application

Proposition

Let $A = \begin{bmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{bmatrix}$ *be a diagonal matrix. Then*

$$A^m = \begin{bmatrix} a_{11}^m & & 0 \\ & \ddots & \\ 0 & & a_{nn}^m \end{bmatrix} \text{ for any } m \in \mathbb{N}.$$

Remark

Note that this, in general, **does not** *hold for non–diagonal matrices, for example* $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^2 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ *and* $1^2 \neq 2$.

Application (continued)

Let $A = \begin{bmatrix} 8 & 10 \\ -3 & -3 \end{bmatrix}$. Compute A^n.

Application (continued)

Let $A = \begin{bmatrix} 8 & 10 \\ -3 & -3 \end{bmatrix}$. Compute A^n. Recall $D = C^{-1}AC$ hence $A = CDC^{-1}$. Therefore $A^n = CD^nC^{-1}$.

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

$$= \begin{bmatrix} -5 \cdot 2^n & + & 2 \cdot 3^{n+1} & -5 \cdot 2^{n+1} & + & 10 \cdot 3^n \\ 3 \cdot 2^n & - & 3^{n+1} & 3 \cdot 2^{n+1} & - & 5 \cdot 3^n \end{bmatrix} =$$

$$= 2^n \begin{bmatrix} -5 & -10 \\ 3 & 6 \end{bmatrix} + 3^n \begin{bmatrix} 6 & 10 \\ -3 & -5 \end{bmatrix} =$$

Application (continued)

Let $A = \begin{bmatrix} 8 & 10 \\ -3 & -3 \end{bmatrix}$. Compute A^n. Recall $D = C^{-1}AC$ hence $A = CDC^{-1}$. Therefore $A^n = CD^nC^{-1}$.

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

$$= \begin{bmatrix} -5 \cdot 2^n + 2 \cdot 3^{n+1} & -5 \cdot 2^{n+1} + 10 \cdot 3^n \\ 3 \cdot 2^n - 3^{n+1} & 3 \cdot 2^{n+1} - 5 \cdot 3^n \end{bmatrix} =$$

$$= 2^n \begin{bmatrix} -5 & -10 \\ 3 & 6 \end{bmatrix} + 3^n \begin{bmatrix} 6 & 10 \\ -3 & -5 \end{bmatrix} =$$

$$= 2^n \left(3 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 8 & 10 \\ -3 & -3 \end{bmatrix} \right) + 3^n \left(-2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 8 & 10 \\ -3 & -3 \end{bmatrix} \right) =$$

$$= (3 \cdot 2^n - 2 \cdot 3^n) I_2 + (-2^n + 3^n) A.$$

Application (continued)

$$A = \begin{bmatrix} 8 & 10 \\ -3 & -3 \end{bmatrix},$$

$$A^n = (3 \cdot 2^n - 2 \cdot 3^n)I + (-2^n + 3^n)A.$$

Note that for $n = 2$

$$A^2 = 5A - 6I,$$

we recover the characteristic polynomial $w_A(\lambda) = \lambda^2 - 5\lambda + 6$.

Application (continued)

$$A = \begin{bmatrix} 8 & 10 \\ -3 & -3 \end{bmatrix},$$

$$A^n = (3 \cdot 2^n - 2 \cdot 3^n)I + (-2^n + 3^n)A.$$

Note that for $n = 2$

$$A^2 = 5A - 6I,$$

we recover the characteristic polynomial $w_A(\lambda) = \lambda^2 - 5\lambda + 6$. Since A and I are linearly independent it follows that

$$\lambda^n \equiv (-2^n + 3^n)\lambda + (3 \cdot 2^n - 2 \cdot 3^n) \pmod{w_A(\lambda)},$$

i.e. the polynomial

$$\lambda^n - \left[(-2^n + 3^n)\lambda + (3 \cdot 2^n - 2 \cdot 3^n)\right],$$

is divisible by the polynomial $w_A(\lambda)$.

Determinant of a Diagonalizable Matrix

Proposition

Let $A \in M(n \times n; \mathbb{R})$ *be a diagonalizable matrix and let* $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ *denote the eigenvalues of* A*. Then*

$$\det A = \lambda_1 \cdot \ldots \cdot \lambda_n.$$

Determinant of a Diagonalizable Matrix

Proposition

Let $A \in M(n \times n; \mathbb{R})$ be a diagonalizable matrix and let $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ denote the eigenvalues of A. Then

$$\det A = \lambda_1 \cdot \ldots \cdot \lambda_n.$$

Proof.

Let

$$D = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}.$$

Then

$$\det A = w_A(0) = w_D(0) = \lambda_1 \cdot \ldots \cdot \lambda_n.$$

□

Coefficients of Characteristic Polynomial

Remark

In general, for any matrix $A \in M(n \times n; \mathbb{R})$

$$w_A(\lambda) = \sum_{i=0}^{n} (-1)^i \left(\sum_{\substack{J \subset \{1,\dots,n\} \\ \#J = n-i}} \det A_{J;J} \right) \lambda^i,$$

where if $J = \{j_1, \dots, j_{n-i}\}$ and $1 \leqslant j_1 < \dots < j_{n-i} \leqslant n$

$$\det A_{J;J} = \det A_{j_1,\dots,j_{n-i};j_1,\dots,j_{n-i}},$$

denotes a minor of order $(n-i)$ (so called **principal minor***).*

Coefficients of Characteristic Polynomial

Remark

In general, for any matrix $A \in M(n \times n; \mathbb{R})$

$$w_A(\lambda) = \sum_{i=0}^{n} (-1)^i \left(\sum_{\substack{J \subset \{1,\ldots,n\} \\ \#J = n-i}} \det A_{J;J} \right) \lambda^i,$$

where if $J = \{j_1, \ldots, j_{n-i}\}$ and $1 \leqslant j_1 < \ldots < j_{n-i} \leqslant n$

$$\det A_{J;J} = \det A_{j_1,\ldots,j_{n-i};j_1,\ldots,j_{n-i}},$$

denotes a minor of order $(n-i)$ (so called **principal minor***).*

In other words, the coefficient of λ^i is equal to $(-1)^i$ times the sum of all prinicipal minors of order $(n-i)$.

Coefficients of Characteristic Polynomial (continued)

Proof.

If $A(\lambda) = [a_{ij}(\lambda)]$ where $a_{ij}(\lambda)$ are differentiable functions of variable λ, the **Jacobi formula** holds

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} \det A(\lambda) = \mathrm{Tr}(\mathrm{adj}(A(\lambda))\frac{\mathrm{d}}{\mathrm{d}\lambda}A(\lambda)),$$

where $\frac{\mathrm{d}}{\mathrm{d}\lambda}A(\lambda) = [\frac{\mathrm{d}}{\mathrm{d}\lambda}a_{ij}(\lambda)]$ and for $B = [b_{ij}]$ the **trace** of matrix $B \in M(n \times n; \mathbb{R})$ is equal to $\mathrm{Tr}(B) = \sum_{i=1}^{n} b_{ii}$.

Coefficients of Characteristic Polynomial (continued)

Proof.

If $A(\lambda) = [a_{ij}(\lambda)]$ where $a_{ij}(\lambda)$ are differentiable functions of variable λ, the **Jacobi formula** holds

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} \det A(\lambda) = \mathrm{Tr}(\mathrm{adj}(A(\lambda)) \frac{\mathrm{d}}{\mathrm{d}\lambda} A(\lambda)),$$

where $\frac{\mathrm{d}}{\mathrm{d}\lambda} A(\lambda) = [\frac{\mathrm{d}}{\mathrm{d}\lambda} a_{ij}(\lambda)]$ and for $B = [b_{ij}]$ the **trace** of matrix $B \in M(n \times n; \mathbb{R})$ is equal to $\mathrm{Tr}(B) = \sum_{i=1}^{n} b_{ii}$.

If $A = [a_{ij}] \in M(n \times n; \mathbb{R})$ is a square matrix, by the Jacobi formula

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} w_A(\lambda) = \mathrm{Tr}(\mathrm{adj}(A - \lambda I)(-I)) = - \sum_{\substack{J \subset \{1,\dots,n\} \\ \#J = n-1}} \det(A - \lambda I)_{J;J}.$$

Coefficients of Characteristic Polynomial (continued)

Proof.

Using induction one can show that

$$\frac{\mathrm{d}^i}{\mathrm{d}\lambda^i} w_A(\lambda) = (-1)^i i! \sum_{\substack{J \subset \{1,\dots,n\} \\ \#J = n-i}} \det(A - \lambda I)_{J;J}.$$

Coefficients of Characteristic Polynomial (continued)

Proof.

Using induction one can show that

$$\frac{\mathrm{d}^i}{\mathrm{d}\lambda^i} w_A(\lambda) = (-1)^i i! \sum_{\substack{J \subset \{1,\dots,n\} \\ \#J = n-i}} \det(A - \lambda I)_{J;J}.$$

The claim follows from the Taylor formula, i.e.

$$w_A(\lambda) = \sum_{i=0}^{n} \frac{1}{i!} \frac{\mathrm{d}^i}{\mathrm{d}\lambda^i} w_A(0) \lambda^i.$$

$\square$

Coefficients of Characteristic Polynomial (continued)

Proof.

Using induction one can show that

$$\frac{\mathrm{d}^i}{\mathrm{d}\lambda^i} w_A(\lambda) = (-1)^i i! \sum_{\substack{J \subset \{1,\dots,n\} \\ \#J = n-i}} \det(A - \lambda I)_{J;J}.$$

The claim follows from the Taylor formula, i.e.

$$w_A(\lambda) = \sum_{i=0}^{n} \frac{1}{i!} \frac{\mathrm{d}^i}{\mathrm{d}\lambda^i} w_A(0) \lambda^i.$$

□

Remark

The Jacobi formula follows directly form the chain rule for total derivatives (note that $\frac{\partial}{\partial a_{ij}} \det A = (-1)^{i+j} \det A_{ij}$ *hence* $d(\det)_A = \operatorname{adj}(A)$ *).*

Coefficients of Characteristic Polynomial (continued)

The coefficients of characteristic polynomial are also symmetric functions of eigenvalues (permuting, i.e. changing the order of factors does not change the coefficients).

$$(\lambda-\lambda_1)(\lambda-\lambda_2)=\lambda^2-(\lambda_1+\lambda_2)\lambda+\lambda_1\lambda_2,$$

$$(\lambda-\lambda_1)(\lambda-\lambda_2)(\lambda-\lambda_3)=\lambda^3-(\lambda_1+\lambda_2+\lambda_3)\lambda^2+(\lambda_1\lambda_2+\lambda_1\lambda_3+\lambda_2\lambda_3)\lambda-$$
$$-\lambda_1\lambda_2\lambda_3,$$

$$(\lambda-\lambda_1)(\lambda-\lambda_2)(\lambda-\lambda_3)(\lambda-\lambda_4)=\lambda^4-(\lambda_1+\lambda_2+\lambda_3+\lambda_4)\lambda^3+$$
$$+(\lambda_1\lambda_2+\lambda_1\lambda_3+\lambda_1\lambda_4+\lambda_2\lambda_3+\lambda_2\lambda_4+\lambda_3\lambda_4)\lambda^2-$$
$$-(\lambda_1\lambda_2\lambda_3+\lambda_1\lambda_2\lambda_4+\lambda_1\lambda_3\lambda_4+\lambda_2\lambda_3\lambda_4)\lambda+\lambda_1\lambda_2\lambda_3\lambda_4.$$

$$\vdots$$

Elementary Symmetric Polynomials

Definition

The $m-$th symmetric polynomial in variables $x_1, \ldots, x_n$, where $m \geqslant 0$ (assume $e_0 = 1$) is,

$$e_m = e_m(x_1, \ldots, x_n) = \sum_{1 \leqslant i_1 < i_2 < \ldots < i_m \leqslant n} x_{i_1} x_{i_2} \cdot \ldots \cdot x_{i_m}.$$

Elementary Symmetric Polynomials (continued)

Proposition

The coefficients of the characteristic polynomial w_A of matrix $A \in M(n \times n; \mathbb{C})$ are (up to a sign) elementary symmetric polynomials of the (complex) eigenvalues of A, i.e.

$$w_A(\lambda) = \sum_{j=0}^{n} (-1)^{n-j} e_j(\lambda_1, \ldots, \lambda_n) \lambda^{n-j}.$$

Proof.

Omitted (use induction). $\square$

[0] cf. I. G. Macdonald *Symmetric Functions and Hall Polynomials*, Oxford 2015

Partitions

Definition

A partition μ of a natural number $n \in \mathbb{N}$ is any sequence of natural numbers $\mu_1, \mu_2, \mu_3 \ldots$ such that

$$|\mu| = \mu_1 + \mu_2 + \mu_3 + \ldots = n,$$

and

$$\mu_1 \geqslant \mu_2 \geqslant \mu_3 \geqslant \ldots .$$

The numbers $\mu_1, \mu_2, \mu_3 \ldots$ are called **parts** of μ. The number of non–zero parts $l(\mu)$ of μ is called the **length** of μ.

[0]alternatively cf. R. P. Stanley *Enumerative Combinatorics vol. 2*, Cambridge 2001

Partitions

Definition

A partition μ of a natural number $n \in \mathbb{N}$ is any sequence of natural numbers $\mu_1, \mu_2, \mu_3 \ldots$ such that

$$|\mu| = \mu_1 + \mu_2 + \mu_3 + \ldots = n,$$

and

$$\mu_1 \geqslant \mu_2 \geqslant \mu_3 \geqslant \ldots .$$

The numbers $\mu_1, \mu_2, \mu_3 \ldots$ are called **parts** of μ. The number of non–zero parts $l(\mu)$ of μ is called the **length** of μ.

Example

$(2, 2, 1, 0, 0)$ is a partition of the number 5 of length 3, i.e. $|\mu| = 5$ and $l(\mu) = 3$. It is denoted alternatively as $\mu = (1^1 2^2 3^0 \ldots)$.

[0] alternatively cf. R. P. Stanley *Enumerative Combinatorics vol. 2*, Cambridge 2001

Monomial Symmetric Polynomials

Definition

For any partition $\mu = (1^{k_1} 2^{k_2} 3^{k_3} \dots)$ such that $|\mu| = m$ and $l(\mu) \leqslant n$ the $m-$th monomial symmetric polynomial m_μ is given by the formula

$$m_\mu = m_\mu(x_1, \dots, x_n) = \frac{1}{k_1!k_2! \dots k_n!} \sum_{\sigma \in S_n} x_1^{\mu(\sigma(1))} x_2^{\mu(\sigma(2))} \cdot \dots \cdot x_n^{\mu(\sigma(n))},$$

where S_n denotes the $n-$th symmetric group (i.e. the group of all permutations of the set $\{1, \dots, n\}$).

Monomial Symmetric Polynomials

Definition

For any partition $\mu = (1^{k_1} 2^{k_2} 3^{k_3} \dots)$ such that $|\mu| = m$ and $l(\mu) \leqslant n$ the $m-$th monomial symmetric polynomial m_μ is given by the formula

$$m_\mu = m_\mu(x_1, \dots, x_n) = \frac{1}{k_1! k_2! \dots k_n!} \sum_{\sigma \in S_n} x_1^{\mu(\sigma(1))} x_2^{\mu(\sigma(2))} \cdot \dots \cdot x_n^{\mu(\sigma(n))},$$

where S_n denotes the $n-$th symmetric group (i.e. the group of all permutations of the set $\{1, \dots, n\}$).

Example

Let $\mu = (2, 1, 0)$ and $n = 3$, then

$$m_\mu(x_1, x_2, x_3) = x_1^2 x_2 + x_1^2 x_3 + x_1 x_2^2 + x_2^2 x_3 + x_1 x_3^2 + x_2 x_3^2.$$

Monomial Symmetric Polynomials (continued)

Remark

The constant $\frac{1}{k_1!k_2!\ldots k_n!}$ *is chosen to make coefficients of all monomials in* m_μ *equal to* 1. *For example, let* $\mu = (1,1,0)$ *and* $n = 3$, *then*

$$m_\mu(x_1, x_2, x_3) = x_1x_2 + x_1x_3 + x_2x_3.$$

For example, let $\mu = (1,1,1)$ *and* $n = 3$, *then*

$$m_\mu(x_1, x_2, x_3) = x_1x_2x_3.$$

Analogously for $\mu = (2,0,0)$ *and* $n = 3$

$$m_\mu(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2.$$

Complete Symmetric Polynomials

Definition

For any $m \geqslant 0$ the $m-$th complete symmetric polynomial h_m in variables $x_1, \dots, x_n$ is given by the formula

$$h_m = h_m(x_1, \dots, x_n) = \sum_{|\mu|=m} m_\mu(x_1, \dots, x_n).$$

We set $h_0 = 1$ and $h_m = 0$ for any $m > n$.

Complete Symmetric Polynomials

Definition

For any $m \geqslant 0$ the $m-$th complete symmetric polynomial h_m in variables $x_1, \ldots, x_n$ is given by the formula

$$h_m = h_m(x_1, \ldots, x_n) = \sum_{|\mu|=m} m_\mu(x_1, \ldots, x_n).$$

We set $h_0 = 1$ and $h_m = 0$ for any $m > n$.

Example

Let $n = 3$, then

$$h_1(x_1, x_2, x_3) = x_1 + x_2 + x_3,$$

$$h_2(x_1, x_2, x_3) = \mu_{(2,0,0)} + \mu_{(1,1,0)} = x_1^2 + x_2^2 + x_3^2 + x_1 x_2 + x_1 x_3 + x_2 x_3.$$

The polynomial h_m is sum of all monomials in variables $x_1, \ldots, x_n$ of degree m.

Power Symmetric Polynomials

Definition

For any $m \geqslant 1$ the $m-$th power symmetric polynomial p_m in variables $x_1, \ldots, x_n$ is given by the formula

$$p_m = p_m(x_1, \ldots, x_n) = m_{(1^m)} = x_1^m + \ldots + x_n^m.$$

We set $p_0 = n$.

Power Symmetric Polynomials

Definition

For any $m \geqslant 1$ the $m-$th power symmetric polynomial p_m in variables $x_1, \ldots, x_n$ is given by the formula

$$p_m = p_m(x_1, \ldots, x_n) = m_{(1^m)} = x_1^m + \ldots + x_n^m.$$

We set $p_0 = n$.

Example

For $m = 2$ and $n = 3$

$$p_2(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2.$$

Symmetric Polynomials

Definition

Polynomial $P(x_1, \ldots, x_n) \in \mathbb{C}[x_1, \ldots, x_n]$ is **symmetric**, if for any $\sigma \in S_n$

$$P(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) = P(x_1, \ldots, x_n).$$

Proposition

Any symmetric polynomial in n variables is a polynomial of $h_1, \ldots, h_n$ (resp. of $p_1, \ldots, p_n$, resp. of $e_1, \ldots, e_n$).

Proof.

Omitted. $\square$

Newton Identities

Let

$$E(t) = (1+x_1 t)(1+x_2 t)\cdot\ldots\cdot(1+x_n t) = 1+e_1 t+e_2 t^2+\ldots+e_n t^n =$$

$$= \sum_{m=0}^{n} e_m(x_1,\ldots,x_n)t^m,$$

be the generating function for the elementary symmetric polynomials. Similarly, let

$$H(t) = \frac{1}{1-x_1 t}\cdot\frac{1}{1-x_2 t}\cdot\ldots\cdot\frac{1}{1-x_n t} =$$

$$= (1+x_1 t+x_1^2 t^2+\ldots)(1+x_2 t+x_2^2 t^2+\ldots)\cdot\ldots\cdot(1+x_n t+x_n^2 t^2+\ldots) =$$

$$= 1 + h_1 t + h_2 t^2 + \ldots = \sum_{m=0}^{\infty} h_m(x_1,\ldots,x_n)t^m.$$

$$P(t) = \frac{x_1}{1-tx_1}+\frac{x_2}{1-tx_2}+\ldots+\frac{x_n}{1-tx_n} = p_1+p_2 t+p_3 t+\ldots =$$

$$= \sum_{m=0}^{\infty} p_{m+1}(x_1,\ldots,x_n)t^m.$$

Newton Identities (continued)

The following (easy to check) equations hold

$$H(t)E(-t) = 1,$$

$$P(t) = \frac{H'(t)}{H(t)},$$

$$P(-t) = \frac{E'(t)}{E(t)},$$

giving raise (by the uniqueness of the Taylor expansion, comparing the coefficients at t^k) to the following identities, respectively

$$\sum_{m=0}^{k} (-1)^m e_m h_{m-k} = 0, \quad \text{for} \quad k \geqslant 1$$

$$\sum_{m=0}^{k} h_m p_{k-m+1} = (k+1)h_{k+1}, \quad \text{for} \quad k \geqslant 0,$$

$$\sum_{m=0}^{k} (-1)^{k-m} e_m p_{k-m+1} = (k+1)e_{k+1}, \quad \text{for} \quad k \geqslant 0.$$

Newton Identities (continued)

Usually, those identities are written in a slightly different but equivalent (simple exercise) way

$$\sum_{m=0}^{k} (-1)^m e_m h_{m-k} = 0, \quad \text{for } k \geqslant 1$$

$$\sum_{m=1}^{k} h_{k-m} p_m = k h_k, \quad \text{for } k \geqslant 1,$$

$$\sum_{m=1}^{k} (-1)^{m-1} e_{k-m} p_m = k e_k, \quad \text{for } k \geqslant 1.$$

Moreover, it is possible to express e_m's and h_m's solely in terms of p_m's which lead for example to formulas for the coefficients of the characteristic polynomial w_A in terms of $\mathrm{tr}(A), \mathrm{tr}\left(A^2\right), \ldots, \mathrm{tr}(A^n)$ (see Faddeev–LeVerrier algorithm).

Newton Identities (continued)

Proposition

The following formulas hold

$$h_m(x_1, \ldots, x_n) = \sum_{\substack{|\mu|=m \\ \mu=(1^{k_1}2^{k_2}\ldots)}} \frac{1}{\prod_j j^{k_j} k_j!} p_\mu,$$

$$e_m(x_1, \ldots, x_n) = \sum_{\substack{|\mu|=m \\ \mu=(1^{k_1}2^{k_2}\ldots)}} (-1)^{m-l(\mu)} \frac{1}{\prod_j j^{k_j} k_j!} p_{\mu_1} p_{\mu_2} \cdot \ldots,$$

where

$$p_\mu = p_{\mu_1} p_{\mu_2} \cdot \ldots.$$

Proof.

Omitted. Observe that the identities do not depend on n (i.e. the number of variables). $\square$

Newton Identities – Example

$$h_1 = p_{(1)} = p_1,$$

$$h_2 = p_{(2)} + p_{(1,1)} = p_{(1^0 2^1 \dots)} + p_{(1^2 2^0 \dots)} =$$

$$= \frac{1}{1^0 \cdot 0! \cdot 2^1 \cdot 1!} p_2 + \frac{1}{1^2 \cdot 2!} p_1 p_1 = \frac{1}{2}(p_1^2 + p_2),$$

$$h_3 = p_{(3)} + p_{(2,1)} + p_{(1,1,1)} = p_{(1^0 2^0 3^1 \dots)} + p_{(1^1 2^1 \dots)} + p_{(1^3 \dots)} =$$

$$= \frac{1}{3^1 1!} p_3 + \frac{1}{1^1 \cdot 1! \cdot 2^1 \cdot 1!} p_2 p_1 + \frac{1}{1^3 3!} p_1 p_1 p_1 = \frac{1}{3} p_3 + \frac{1}{2} p_1 p_2 + \frac{1}{6} p_1^3,$$

$$\vdots$$

Newton Identities – Example

$$e_1 = p_{(1)} = p_1,$$

$$e_2 = -p_{(2)} + p_{(1,1)} = -p_{(1^0 2^1 \ldots)} + p_{(1^2 2^0 \ldots)} =$$

$$= \frac{1}{1^0 \cdot 0! \cdot 2^1 \cdot 1!} p_2 + \frac{1}{1^2 \cdot 2!} p_1 p_1 = \frac{1}{2}(p_1^2 - p_2),$$

$$e_3 = p_{(3)} - p_{(2,1)} + p_{(1,1,1)} = p_{(1^0 2^0 3^1 \ldots)} - p_{(1^1 2^1 \ldots)} + p_{(1^3 \ldots)} =$$

$$= \frac{1}{3^1 1!} p_3 - \frac{1}{1^1 \cdot 1! \cdot 2^1 \cdot 1!} p_2 p_1 + \frac{1}{1^3 3!} p_1 p_1 p_1 = \frac{1}{3} p_3 - \frac{1}{2} p_1 p_2 + \frac{1}{6} p_1^3,$$

$$\vdots$$

Newton Identities – Example (continued)

Three numbers $x, y, z \in \mathbb{R}$ satisfy the following system of equations

$$\begin{cases} x + y + z = 2, \\ x^2 + y^2 + z^2 = 6, \\ x^3 + y^3 + z^3 = 8. \end{cases}$$

Determine xyz. The problem can be solved using the identity

$$e_3 = \frac{1}{3}p_3 - \frac{1}{2}p_1 p_2 + \frac{1}{6}p_1^3,$$

that is

$$xyz = \frac{1}{3} \cdot 8 - \frac{1}{2} \cdot 2 \cdot 6 + \frac{1}{6} \cdot 2^3 =$$

$$= \frac{8}{3} - 6 + \frac{4}{3} = -2.$$

In fact, $x = 1, y = 2, z = -1$ (up to a permutation).

Schur Polynomials

For any monomial $x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdot \ldots \cdot x_n^{\alpha_n}$ define the antisymmetric (or skew–symmetric) function

$$a_\alpha(x_1, \ldots, x_n) = \sum_{\sigma \in S_n} \operatorname{sgn} \sigma(\sigma . x^\alpha),$$

where

$$\sigma . x^\alpha = x_{\sigma(1)}^{\alpha_1} x_{\sigma(2)}^{\alpha_2} \cdot \ldots x_{\sigma(n)}^{\alpha_n}.$$

For example, if $\alpha = (1, 2, 0)$ and $n = 3$ then

$$a_\alpha(x_1, x_2, x_3) = x_1 x_2^2 + x_2^2 x_3 + x_1 x_3^2 - x_1^2 x_2 - x_2^2 x_3 - x_1 x_3^2.$$

Schur Polynomials (continued)

The alternative definition of the determinant implies that

$$a_\alpha(x_1, \ldots, x_n) = \det \begin{bmatrix} x_1^{\alpha_1} & x_1^{\alpha_2} & x_1^{\alpha_3} & \cdots & x_1^{\alpha_n} \\ x_2^{\alpha_1} & x_2^{\alpha_2} & x_2^{\alpha_3} & \cdots & x_2^{\alpha_n} \\ x_3^{\alpha_1} & x_3^{\alpha_2} & x_3^{\alpha_3} & \cdots & x_3^{\alpha_n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_n^{\alpha_1} & x_n^{\alpha_2} & x_n^{\alpha_3} & \cdots & x_n^{\alpha_n} \end{bmatrix}$$

Schur Polynomials (continued)

From the properties of the determinant it follows that

$$a_\alpha(x_1, \dots, x_i, \dots, x_j, \dots, x_n) = -a_\alpha(x_1, \dots, x_j, \dots, x_i, \dots, x_n),$$

(i.e. a_α is alternating and hence antisymmetric) and that

$$a_\alpha = 0,$$

if some $\alpha_i = \alpha_j$ for $i \neq j$. It follows that the polynomial a_α is divisible by the polynomial $x_i - x_j$ in the ring $\mathbb{Z}[x_1, \dots, x_n]$. For example for $\alpha = (1, 2, 0)$ and $n = 3$

$$a_\alpha(x_1, x_2, x_3) = x_1 x_2^2 + x_2^2 x_3 + x_1 x_3^2 - x_1^2 x_2 - x_2^2 x_3 - x_1 x_3^2 =$$

$$= (x_1 - x_2)(-x_1 x_2 + x_1 x_3 + x_2 x_3 - x_3^2) =$$

$$= (x_2 - x_1)(x_3 - x_1)(x_3 - x_2).$$

Schur Polynomials (continued)

Without the loss of generality one can assume that

$$\alpha_1 > \alpha_2 > \ldots > \alpha_n \geqslant 0.$$

This implies that $\alpha_1 \geqslant n-1, \alpha_2 \geqslant n-2, \ldots$ therefore if $\delta = (n-1, n-2, n-3, \ldots, 2, 1, 0)$ then

$$\mu = \alpha - \delta,$$

has non–negative components. Moreover

$$\mu_1 - \mu_2 = (\alpha_1 - (n-1)) - (\alpha_2 - (n-2)) = \alpha_1 - \alpha_2 - 1 \geqslant 0,$$

$$\mu_2 - \mu_3 = (\alpha_2 - (n-2)) - (\alpha_3 - (n-3)) = \alpha_2 - \alpha_3 - 1 \geqslant 0,$$

$$\vdots$$

that is μ is a partition. This can be reversed, that is for any partition μ, the $\alpha = \mu + \delta$ gives a non–zero function a_α. Observe that $a_\delta(x_1, \ldots, x_n)$ is the Vandermonde determinant.

Schur Polynomials (continued)

Definition

For any partition μ and $\delta = (n-1, n-2, \ldots, 2, 1, 0)$ the Schur polynomial (in variables $x_1, \ldots, x_n$) is the symmetric polynomial in $\mathbb{Z}[x_1, \ldots, x_n]$ given by the formula

$$s_\mu = s_\mu(x_1, \ldots, x_n) = \frac{a_{\mu+\delta}}{a_\delta}.$$

Remark

Schur polynomials for μ such that $|\mu| = m$ form a $\mathbb{Z}-$basis of the homogeneous symmetric polynomials of degree m. Schur polynomials play an important role in combinatorics, algebraic geometry, representation theory of the symmetric group, general linear group and the unitary group.

Schur Polynomials (continued)

$$s_\mu = \frac{\det \begin{bmatrix} x_1^{\mu_1+(n-1)} & x_1^{\mu_2+(n-2)} & x_1^{\mu_3+(n-3)} & \cdots & x_1^{\mu_n+0} \\ x_2^{\mu_1+(n-1)} & x_2^{\mu_2+(n-2)} & x_1^{\mu_3+(n-3)} & \cdots & x_2^{\mu_n+0} \\ x_3^{\mu_1+(n-1)} & x_3^{\mu_2+(n-2)} & x_1^{\mu_3+(n-3)} & \cdots & x_3^{\mu_n+0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_n^{\mu_1+(n-1)} & x_n^{\mu_2+(n-2)} & x_1^{\mu_3+(n-3)} & \cdots & x_n^{\mu_n+0} \end{bmatrix}}{\det \begin{bmatrix} x_1^{n-1} & x_1^{n-2} & x_1^{n-3} & \cdots & x_1^{0} \\ x_2^{n-1} & x_2^{n-2} & x_2^{n-3} & \cdots & x_2^{0} \\ x_3^{n-1} & x_3^{n-2} & x_3^{n-3} & \cdots & x_3^{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_n^{n-1} & x_n^{n-2} & x_n^{n-3} & \cdots & x_n^{0} \end{bmatrix}}.$$

Schur Polynomials – Example

$$\delta = (2,1,0)$$

$$s_\delta = -\left(x_2 - x_1\right)\left(x_3 - x_1\right)\left(x_3 - x_2\right)$$

$$s_{(2,0,0)}(x_1, x_2, x_3) = \frac{1}{s_\delta} \det \begin{bmatrix} x_1^4 & x_1^1 & x_1^0 \\ x_2^4 & x_2^1 & x_2^0 \\ x_3^4 & x_3^1 & x_3^0 \end{bmatrix} =$$

$$= \frac{1}{s_\delta}\left(-\left(x_2 - x_1\right)\cdot\right.$$

$$\left.\cdot\left(x_3 - x_1\right)\left(x_3 - x_2\right)\left({x_3}^2 + x_2x_3 + x_1x_3 + {x_2}^2 + x_1x_2 + {x_1}^2\right)\right) =$$

$$= x_1^2 + x_2^2 + x_3^2 + x_1x_2 + x_1x_3 + x_2x_3.$$

Schur Polynomials – Another Example

$$\delta = (2,1,0)$$

$$s_\delta = -\,(x_2 - x_1)\,(x_3 - x_1)\,(x_3 - x_2)$$

$$s_{(1,1,0)}(x_1, x_2, x_3) = \frac{1}{s_\delta} \det \begin{bmatrix} x_1^3 & x_1^2 & x_1^0 \\ x_2^3 & x_2^2 & x_2^0 \\ x_3^3 & x_3^2 & x_3^0 \end{bmatrix} =$$

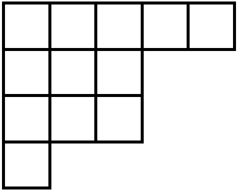
$$= \frac{1}{s_\delta}\left(-\,(x_2 - x_1)\,(x_3 - x_1)\,(x_3 - x_2)\,(x_2 x_3 + x_1 x_3 + x_1 x_2)\right) =$$

$$= x_1 x_2 + x_1 x_3 + x_2 x_3 .$$

Semistandard Young Tableau

Definition

For any partition μ a semistandard Young tableau T of shape μ is a way of placing numbers into the diagram (μ_1 boxes in the first row, μ_2 in the second, etc.)



such that

i) numbers in rows are weakly increasing (from left to right),

ii) numbers in columns are strictly increasing (top to down).

Semistandard Young Tableau

Definition

The set $SSYT_\mu$ is the set of all semistandard Young tableaux and $SSYT_\mu(n)$ is the set of all semistandard Young tableuax with entries not greater than n. For any $T \in SSYT_\mu(n)$

$$x^T = x_1^{\#1's} x_2^{\#2's} \cdot \ldots \cdot x_n^{\#n's},$$

that x_j is raised to the number of occurrence of j in T.

---

[1] cf. B. E. Sagan, *The Symmetric Group*, Springer 2001

Semistandard Young Tableau

Definition

The set $SSYT_\mu$ is the set of all semistandard Young tableaux and $SSYT_\mu(n)$ is the set of all semistandard Young tableuax with entries not greater than n. For any $T \in SSYT_\mu(n)$

$$x^T = x_1^{\#1's} x_2^{\#2's} \cdot \ldots \cdot x_n^{\#n's},$$

that x_j is raised to the number of occurrence of j in T.

Proposition

$$s_\mu(x_1, \ldots, x_n) = \sum_{T \in SSYT_\mu(n)} x^T.$$

Proof.

Omitted.[1] □

[1] cf. B. E. Sagan, *The Symmetric Group*, Springer 2001

Semistandard Young Tableau - Example

$$\mu = (2,0,0), \quad n = 3,$$

$$SSYT_\mu(n) = \left\{ \boxed{1}\boxed{1}, \boxed{2}\boxed{2}, \boxed{3}\boxed{3}, \boxed{1}\boxed{2}, \boxed{1}\boxed{3}, \boxed{2}\boxed{3} \right\},$$

$$s_\mu = x_1^2 + x_2^2 + x_3^2 + x_1x_2 + x_1x_3 + x_2x_3.$$

$$\mu = (1,1,0), \quad n = 3,$$

$$SSYT_\mu(n) = \left\{ \begin{array}{|c|}\hline 1 \\ \hline 2 \\ \hline \end{array}, \begin{array}{|c|}\hline 1 \\ \hline 3 \\ \hline \end{array}, \begin{array}{|c|}\hline 2 \\ \hline 3 \\ \hline \end{array} \right\},$$

$$s_\mu = x_1x_2 + x_1x_3 + x_2x_3.$$

Pieri's Formula

Proposition

$$s_\mu s_{(m)} = \sum_\nu s_\nu,$$

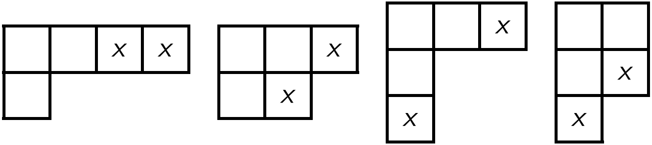
where the sum is over all paritions ν obtained from μ by adding m boxes but no two in a single column.

Proof.

Omitted. $\square$

Pieri's Formula – Example

$$s_{(2,1)}s_{(2)} = s_{(4,1)} + s_{(3,2)} + s_{(3,1,1)} + s_{(2,2,1)},$$



Symmetric Matrix – Spectral Theorem

Definition

Matrix $A \in M(n \times n; \mathbb{R})$ is called **symmetric** if $A^\intercal = A$.

Proposition

Let $A \in M(n \times n; \mathbb{R})$ be a symmetric matrix. Then A is diagonalizable.

Symmetric Matrix – Spectral Theorem

Definition

Matrix $A \in M(n \times n; \mathbb{R})$ is called **symmetric** if $A^{\mathsf{T}} = A$.

Proposition

Let $A \in M(n \times n; \mathbb{R})$ be a symmetric matrix. Then A is diagonalizable.

Moreover there exists an **orthogonal** basis of $\mathbb{R}^n$ consisting of eigenvectors of the endomorphism $M(\varphi)_{st} = A$, i.e. vectors of that basis are pairwise perpendicular.

Example

Let

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

Then

$$w_A(\lambda) = -(\lambda + 3)[(1 - \lambda)(-2 - \lambda) - 4] = -(\lambda + 3)^2(\lambda - 2),$$

$$V_{(-3)} = \text{lin}((-1, 2, 0), (0, 0, 1)),$$

$$V_{(2)} = \text{lin}((2, 1, 0)),$$

and the eigenvectors are pairwise perpendicular.

Minimal Polynomial

Definition

Let $A \in M(n \times n; \mathbb{R})$. The minimal polynomial μ_A of the matrix A is a non-zero monic polynomial with real coefficients of the least degree such that $\mu_A(A) = 0$.

Equivalently, the minimal polynomial of A is the non-zero monic polynomial of the least degree which image under the map

$$\mathbb{R}[x] \ni P(x) \mapsto P(A) \in M(n \times n; \mathbb{R}),$$

is the zero matrix (or which divides each $P(x) \in \mathbb{R}[x]$ with $P(A) = 0$).

Minimal Polynomial

Definition

Let $A \in M(n \times n; \mathbb{R})$. The minimal polynomial μ_A of the matrix A is a non-zero monic polynomial with real coefficients of the least degree such that $\mu_A(A) = 0$.

Equivalently, the minimal polynomial of A is the non-zero monic polynomial of the least degree which image under the map

$$\mathbb{R}[x] \ni P(x) \mapsto P(A) \in M(n \times n; \mathbb{R}),$$

is the zero matrix (or which divides each $P(x) \in \mathbb{R}[x]$ with $P(A) = 0$).

By the Cayley–Hamilton Theorem the minimal polynomial of A divides the characteristic polynomial of A, i.e. $\mu_A \mid w_A$.

Minimal Polynomial

Remark

The degree of the minimal polynomial μ_A is equal to the smallest number $m \geqslant 1$ such that

$$A^m \in \mathrm{lin}(A^{m-1}, \ldots, A^1, A^0),$$

and if

$$A^m = \alpha_{m-1} A^{m-1} + \ldots + \alpha_1 A^1 + \alpha_0 A^0,$$

for some $\alpha_i \in \mathbb{R}$, then

$$\mu_A(\lambda) = \lambda^m - (\alpha_{m+1} \lambda^{m-1} + \ldots + \alpha_1 \lambda + \alpha_0).$$

Example

Let $A = \begin{bmatrix} 8 & 10 \\ -3 & -3 \end{bmatrix}$. Then $w_A(\lambda) = (\lambda - 2)(\lambda - 3)$ and the only monic divisors of w_A are $w_A, \lambda - 2, \lambda - 3$ and 1. Since A is not a diagonal matrix then $\mu_A = w_A$.

Example

Let $A = \begin{bmatrix} 8 & 10 \\ -3 & -3 \end{bmatrix}$. Then $w_A(\lambda) = (\lambda - 2)(\lambda - 3)$ and the only monic divisors of w_A are $w_A, \lambda - 2, \lambda - 3$ and 1. Since A is not a diagonal matrix then $\mu_A = w_A$.

Let $B = \begin{bmatrix} 2 & -2 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 4 \end{bmatrix}$. Then $w_B(\lambda) = (2 - \lambda)^2(4 - \lambda)$. Then only monic divisors of w_B are $-w_B, (\lambda - 2)^2, \lambda - 2, (\lambda - 2)(\lambda - 4)$, $\lambda - 4$ and 1. It can be checked that $\mu_B = -w_B$.

Example

Let $A = \begin{bmatrix} 8 & 10 \\ -3 & -3 \end{bmatrix}$. Then $w_A(\lambda) = (\lambda - 2)(\lambda - 3)$ and the only monic divisors of w_A are $w_A, \lambda - 2, \lambda - 3$ and 1. Since A is not a diagonal matrix then $\mu_A = w_A$.

Let $B = \begin{bmatrix} 2 & -2 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 4 \end{bmatrix}$. Then $w_B(\lambda) = (2 - \lambda)^2(4 - \lambda)$. Then only monic divisors of w_B are $-w_B, (\lambda - 2)^2, \lambda - 2, (\lambda - 2)(\lambda - 4)$, $\lambda - 4$ and 1. It can be checked that $\mu_B = -w_B$. Equivalently, the matrix

$$B^2 = \begin{bmatrix} 4 & -8 & 4 \\ 0 & 4 & 6 \\ 0 & 0 & 16 \end{bmatrix},$$

is not a linear combination of matrices B and I_3.

Minimal Polynomials of Similar Matrices

Proposition

Let $A, B \in M(n \times n; \mathbb{R})$ *be similar matrices. Then* $\mu_A = \mu_B$.

Minimal Polynomials of Similar Matrices

Proposition

Let $A, B \in M(n \times n; \mathbb{R})$ be similar matrices. Then $\mu_A = \mu_B$.

Proof.

If $A = C^{-1}BC$ then $0 = \mu_A(A) = C^{-1}\mu_A(B)C$ therefore $\mu_A(B) = 0$. By definition $\mu_B \mid \mu_A$ and analogously $\mu_A \mid \mu_B$. Since both polynomials are monic $\mu_A = \mu_B$. □

Minimal Polynomials of Similar Matrices

Proposition

Let $A, B \in M(n \times n; \mathbb{R})$ be similar matrices. Then $\mu_A = \mu_B$.

Proof.

If $A = C^{-1}BC$ then $0 = \mu_A(A) = C^{-1}\mu_A(B)C$ therefore $\mu_A(B) = 0$. By definition $\mu_B \mid \mu_A$ and analogously $\mu_A \mid \mu_B$. Since both polynomials are monic $\mu_A = \mu_B$. $\square$

Remark

Non-similar matrices can have the same minimal polynomials. For example

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix},$$

have the same minimal polynomial

$$\mu_A(\lambda) = \mu_B(\lambda) = (\lambda - 1)(\lambda - 2)$$

Criterion for Diagonalizability

Theorem

Let $A \in M(n \times n; \mathbb{R})$. Matrix A is diagonalizable if and only if the minimal polynomial of A factors as follows

$$\mu_A(\lambda) = (\lambda - \alpha_1)(\lambda - \alpha_2)\ldots(\lambda - \alpha_k),$$

where $\alpha_i \in \mathbb{R}$ and $\alpha_i \neq \alpha_j$, i.e. α_i are pairwise distinct numbers.

Criterion for Diagonalizability

Theorem

Let $A \in M(n \times n; \mathbb{R})$. Matrix A is diagonalizable if and only if the minimal polynomial of A factors as follows

$$\mu_A(\lambda) = (\lambda - \alpha_1)(\lambda - \alpha_2)\dots(\lambda - \alpha_k),$$

where $\alpha_i \in \mathbb{R}$ and $\alpha_i \neq \alpha_j$, i.e. α_i are pairwise distinct numbers.

Example

$$A = \begin{bmatrix} 8 & 10 \\ -3 & -3 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & -2 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 4 \end{bmatrix}$$

$$\mu_A(\lambda) = (\lambda - 2)(\lambda - 3),$$

$$\mu_B(\lambda) = (\lambda - 2)^2(\lambda - 4).$$

Matrix A is diagonalizable and matrix B is not diagonalizable.

Criterion for Diagonalizability (continued)

Example

The minimal polynomial of matrix

$$C = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

is equal to its characteristic polynomial. The minimal polynomial has pairwise different complex roots so the matrix C diagonalizes over $\mathbb{C}$ but not over $\mathbb{R}$.

Criterion for Diagonalizability (continued)

Corollary

Matrix $A \in M(n \times n; \mathbb{C})$ *of finite order (i.e.,* $A^m = I$ *for some* $m \geqslant 1$*) is diagonalizable (over* $\mathbb{C}$*).*

Criterion for Diagonalizability (continued)

Corollary

Matrix $A \in M(n \times n; \mathbb{C})$ of finite order (i.e., $A^m = I$ for some $m \geqslant 1$) is diagonalizable (over $\mathbb{C}$).

Proof.

The minimal polyomial of A divides the polynomial $x^m - 1$ which has only simple roots. $\square$

Criterion for Diagonalizability (continued)

Corollary

Matrix $A \in M(n \times n; \mathbb{C})$ of finite order (i.e., $A^m = I$ for some $m \geqslant 1$) is diagonalizable (over $\mathbb{C}$).

Proof.

The minimal polyomial of A divides the polynomial $x^m - 1$ which has only simple roots. $\square$

Warning

This theorem fails in positive characteristic, take say
$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \in M(2 \times 2; \mathbb{F}_2)$.

Criterion for Diagonalizability (continued)

Proof.

($\Rightarrow$) Let $D = C^{-1}AC$, by the previous proposition $\mu_A = \mu_D$. Let $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ be all pairwise distinct eigenvalues of matrix D. For any $i = 1, \dots, k$, $v_i \in V_{(\alpha_i)}$

$$(D - \alpha_j I)v_i = (\alpha_i - \alpha_j)v_i \quad \text{for} \quad j = 1, \dots, k.$$

Criterion for Diagonalizability (continued)

Proof.

$(\Rightarrow)$ Let $D = C^{-1}AC$, by the previous proposition $\mu_A = \mu_D$. Let $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ be all pairwise distinct eigenvalues of matrix D. For any $i = 1, \ldots, k$, $v_i \in V_{(\alpha_i)}$

$$(D - \alpha_j I)v_i = (\alpha_i - \alpha_j)v_i \quad \text{for} \quad j = 1, \ldots, k.$$

It follows that for any $i = 1, \ldots, k$, $v_i \in V_{(\alpha_i)}$ and any $m_1, \ldots, m_k \geqslant 0$

$$[(D - \alpha_1 I)^{m_1} \ldots (D - \alpha_k I)^{m_k}]v_i = (\alpha_i - \alpha_1)^{m_1} \cdot \ldots \cdot (\alpha_i - \alpha_k)^{m_k} v_i.$$

Criterion for Diagonalizability (continued)

Proof.

($\Rightarrow$) Let $D = C^{-1}AC$, by the previous proposition $\mu_A = \mu_D$. Let $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ be all pairwise distinct eigenvalues of matrix D. For any $i = 1, \ldots, k$, $v_i \in V_{(\alpha_i)}$

$$(D - \alpha_j I)v_i = (\alpha_i - \alpha_j)v_i \quad \text{for} \quad j = 1, \ldots, k.$$

It follows that for any $i = 1, \ldots, k$, $v_i \in V_{(\alpha_i)}$ and any $m_1, \ldots, m_k \geqslant 0$

$$[(D - \alpha_1 I)^{m_1} \ldots (D - \alpha_k I)^{m_k}]v_i = (\alpha_i - \alpha_1)^{m_1} \cdot \ldots \cdot (\alpha_i - \alpha_k)^{m_k} v_i.$$

Since for any $P(x) \in \mathbb{R}[x]$

$$P(D) = 0 \iff P(D)v_i = 0 \quad \text{for any} \quad i = 1, \ldots, k, \ v_i \in V_{(\alpha_i)},$$

it follows that the minimal polynomial $\mu_A(\lambda)$ is equal to

$$\mu_D(\lambda) = (\lambda - \alpha_1) \cdot \ldots \cdot (\lambda - \alpha_k).$$

Criterion for Diagonalizability (continued)

Proof.

($\Leftarrow$) Let

$$Q_i(\lambda) = \frac{\mu_A(\lambda)}{\lambda - \alpha_i} \quad \text{for} \quad i = 1, \ldots, k.$$

Since

$$\mathrm{GCD}(Q_1(\lambda), \ldots, Q_k(\lambda)) = 1,$$

there exist polynomials $P_1, \ldots, P_k \in \mathbb{R}[x]$ such that

$$P_1(\lambda)Q_1(\lambda) + \ldots + P_k(\lambda)Q_k(\lambda) = 1. \qquad (*)$$

Criterion for Diagonalizability (continued)

Proof.

($\Leftarrow$) Let

$$Q_i(\lambda) = \frac{\mu_A(\lambda)}{\lambda - \alpha_i} \quad \text{for} \quad i = 1, \dots, k.$$

Since

$$\mathrm{GCD}(Q_1(\lambda), \dots, Q_k(\lambda)) = 1,$$

there exist polynomials $P_1, \dots, P_k \in \mathbb{R}[x]$ such that

$$P_1(\lambda)Q_1(\lambda) + \ldots + P_k(\lambda)Q_k(\lambda) = 1. \qquad (*)$$

Since $\mu_A \mid w_A$, the numbers $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ are eigenvalues of matrix A.

Criterion for Diagonalizability (continued)

Proof.

$(\Leftarrow)$ Let

$$Q_i(\lambda) = \frac{\mu_A(\lambda)}{\lambda - \alpha_i} \quad \text{for} \quad i = 1, \ldots, k.$$

Since

$$\mathrm{GCD}(Q_1(\lambda), \ldots, Q_k(\lambda)) = 1,$$

there exist polynomials $P_1, \ldots, P_k \in \mathbb{R}[x]$ such that

$$P_1(\lambda)Q_1(\lambda) + \ldots + P_k(\lambda)Q_k(\lambda) = 1. \qquad (*)$$

Since $\mu_A \mid w_A$, the numbers $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ are eigenvalues of matrix A. The union of bases of the eigenspaces $V_{(\alpha_1)}, \ldots, V_{(\alpha_k)}$ is a linearly independent set.

Criterion for Diagonalizability (continued)

Proof.

$(\Leftarrow)$ Let

$$Q_i(\lambda) = \frac{\mu_A(\lambda)}{\lambda - \alpha_i} \quad \text{for} \quad i = 1, \ldots, k.$$

Since

$$\mathrm{GCD}(Q_1(\lambda), \ldots, Q_k(\lambda)) = 1,$$

there exist polynomials $P_1, \ldots, P_k \in \mathbb{R}[x]$ such that

$$P_1(\lambda)Q_1(\lambda) + \ldots + P_k(\lambda)Q_k(\lambda) = 1. \qquad (*)$$

Since $\mu_A \mid w_A$, the numbers $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ are eigenvalues of matrix A. The union of bases of the eigenspaces $V_{(\alpha_1)}, \ldots, V_{(\alpha_k)}$ is a linearly independent set. It is enough to show that it spans $\mathbb{R}^n$.

Criterion for Diagonalizability (continued)

Proof.

For any (column) vector $v \in \mathbb{R}^n$ and any $i = 1, \ldots, k$

$$Q_i(A)v \in V_{(\alpha_i)},$$

because

$$\ker(A - \alpha_i I) = V_{(\alpha_i)}$$

and

$$(A - \alpha_i I)(Q_i(A))v = \mu_A(A)v = 0.$$

Criterion for Diagonalizability (continued)

Proof.

For any (column) vector $v \in \mathbb{R}^n$ and any $i = 1, \ldots, k$

$$Q_i(A)v \in V_{(\alpha_i)},$$

because

$$\ker(A - \alpha_i I) = V_{(\alpha_i)}$$

and

$$(A - \alpha_i I)(Q_i(A))v = \mu_A(A)v = 0.$$

Let $v \in \mathbb{R}^n$ be any (column) vector. Substituting matrix A to the equation (**??**) and multiplying it by vector v on the right

$$v = Q_1(A)\left(P_1(A)v\right) + \ldots + Q_k(A)\left(P_k(A)v\right),$$

where

$$Q_i(A)\left(P_i(A)v\right) \in V_{(\alpha_i)} \quad \textit{for} \quad i = 1, \ldots, k.$$

Diagonalizability on Invariant Subspace

Definition

Let $\varphi\colon V \to V$ be an endomorphism and let $W \subset V$ be subspace. Then W is an **invariant subspace** of φ if

$$\varphi(W) \subset W.$$

Diagonalizability on Invariant Subspace

Definition

Let $\varphi \colon V \to V$ be an endomorphism and let $W \subset V$ be subspace. Then W is an **invariant subspace** of φ if

$$\varphi(W) \subset W.$$

Proposition

Let W be an invariant subspace of a diagonalizable endomorphism $\varphi \colon V \to V$.

Diagonalizability on Invariant Subspace

Definition

Let $\varphi \colon V \to V$ be an endomorphism and let $W \subset V$ be subspace. Then W is an **invariant subspace** of φ if

$$\varphi(W) \subset W.$$

Proposition

Let W be an invariant subspace of a diagonalizable endomorphism $\varphi \colon V \to V$. Then $\varphi|_W \colon W \to W$ is diagonalizable.

Diagonalizability on Invariant Subspace

Definition

Let $\varphi \colon V \to V$ be an endomorphism and let $W \subset V$ be subspace. Then W is an **invariant subspace** of φ if

$$\varphi(W) \subset W.$$

Proposition

Let W be an invariant subspace of a diagonalizable endomorphism $\varphi \colon V \to V$. Then $\varphi|_W \colon W \to W$ is diagonalizable.

Proof.

Let μ_φ be the minimal polynomial of φ. Then for any $w \in W$

$$\mu_\varphi(\varphi)(w) = 0,$$

that is the minimal polynomial of $\varphi|_W$ divides the minimal polynomial of φ hence it has simple roots. $\square$

Simultaneously Diagonalizable Endomorphisms

Proposition

Let $\varphi_i \colon V \to V$ *where* $i \in I$ *be a family of diagonalisable endomorphisms. Then endomorphisms* φ_i *commute, i.e., for any* $i, j \in I$

$$\varphi_i \circ \varphi_j = \varphi_j \circ \varphi_i,$$

if and only if there exists a basis $\mathcal{A}$ *of* V *such that matrices* $M(\varphi_i)_{\mathcal{A}}^{\mathcal{A}}$ *are diagonal for each* $i \in I$*, that is endomorphisms* φ_i *are simultaneously diagonalizable.*

Simultaneously Diagonalizable Endomorphisms

Proposition

Let $\varphi_i \colon V \to V$ where $i \in I$ be a family of diagonalisable endomorphisms. Then endomorphisms φ_i commute, i.e., for any $i, j \in I$

$$\varphi_i \circ \varphi_j = \varphi_j \circ \varphi_i,$$

if and only if there exists a basis $\mathcal{A}$ of V such that matrices $M(\varphi_i)_{\mathcal{A}}^{\mathcal{A}}$ are diagonal for each $i \in I$, that is endomorphisms φ_i are simultaneously diagonalizable.

Proof.

($\Leftarrow$) if $M(\varphi_i)_{\mathcal{A}}^{\mathcal{A}}$, $M(\varphi_j)_{\mathcal{A}}^{\mathcal{A}}$ are diagonal then

$$M(\varphi_i)_{\mathcal{A}}^{\mathcal{A}} M(\varphi_j)_{\mathcal{A}}^{\mathcal{A}} = M(\varphi_j)_{\mathcal{A}}^{\mathcal{A}} M(\varphi_i)_{\mathcal{A}}^{\mathcal{A}}.$$

Simultaneously Diagonalizable Endomorphisms

Proposition

Let $\varphi_i \colon V \to V$ where $i \in I$ be a family of diagonalisable endomorphisms. Then endomorphisms φ_i commute, i.e., for any $i, j \in I$

$$\varphi_i \circ \varphi_j = \varphi_j \circ \varphi_i,$$

if and only if there exists a basis $\mathcal{A}$ of V such that matrices $M(\varphi_i)_{\mathcal{A}}^{\mathcal{A}}$ are diagonal for each $i \in I$, that is endomorphisms φ_i are simultaneously diagonalizable.

Proof.

($\Leftarrow$) if $M(\varphi_i)_{\mathcal{A}}^{\mathcal{A}}$, $M(\varphi_j)_{\mathcal{A}}^{\mathcal{A}}$ are diagonal then

$$M(\varphi_i)_{\mathcal{A}}^{\mathcal{A}} M(\varphi_j)_{\mathcal{A}}^{\mathcal{A}} = M(\varphi_j)_{\mathcal{A}}^{\mathcal{A}} M(\varphi_i)_{\mathcal{A}}^{\mathcal{A}}.$$

($\Rightarrow$) induction of $n = \dim V$.

Simultaneously Diagonalizable Endomorphisms

Proposition

Let $\varphi_i \colon V \to V$ where $i \in I$ be a family of diagonalisable endomorphisms. Then endomorphisms φ_i commute, i.e., for any $i, j \in I$

$$\varphi_i \circ \varphi_j = \varphi_j \circ \varphi_i,$$

if and only if there exists a basis $\mathcal{A}$ of V such that matrices $M(\varphi_i)_{\mathcal{A}}^{\mathcal{A}}$ are diagonal for each $i \in I$, that is endomorphisms φ_i are simultaneously diagonalizable.

Proof.

($\Leftarrow$) if $M(\varphi_i)_{\mathcal{A}}^{\mathcal{A}},\ M(\varphi_j)_{\mathcal{A}}^{\mathcal{A}}$ are diagonal then

$$M(\varphi_i)_{\mathcal{A}}^{\mathcal{A}} M(\varphi_j)_{\mathcal{A}}^{\mathcal{A}} = M(\varphi_j)_{\mathcal{A}}^{\mathcal{A}} M(\varphi_i)_{\mathcal{A}}^{\mathcal{A}}.$$

($\Rightarrow$) induction of $n = \dim V$. If $n = 1$ the statement is obvious.

Simultaneously Diagonalizable Endomorphisms (continued)

Proof.

Assume there exists $j \in I$ such that $\dim V_{\varphi_j,(\lambda)} < \dim V$, where $\lambda \in \mathbb{R}$ is an eigenvalue of φ_j and $W = V_{\varphi_j,(\lambda)}$ is an eigenspace of φ_j (otherwise each φ_i is a uniform scalling).

Simultaneously Diagonalizable Endomorphisms (continued)

Proof.

Assume there exists $j \in I$ such that $\dim V_{\varphi_j,(\lambda)} < \dim V$, where $\lambda \in \mathbb{R}$ is an eigenvalue of φ_j and $W = V_{\varphi_j,(\lambda)}$ is an eigenspace of φ_j (otherwise each φ_i is a uniform scalling). Then for any $i \in I$ and any $v \in V_{\varphi_j,(\lambda)}$

$$\varphi_j(\varphi_i(v)) = \varphi_i(\varphi_j(v)) = \varphi_i(\lambda v) = \lambda(\varphi_i(v)),$$

i.e.,

$$\varphi_i(V_{\varphi_j,(\lambda)}) \subset V_{\varphi_j,(\lambda)}.$$

Simultaneously Diagonalizable Endomorphisms (continued)

Proof.

Assume there exists $j \in I$ such that $\dim V_{\varphi_j,(\lambda)} < \dim V$, where $\lambda \in \mathbb{R}$ is an eigenvalue of φ_j and $W = V_{\varphi_j,(\lambda)}$ is an eigenspace of φ_j (otherwise each φ_i is a uniform scalling). Then for any $i \in I$ and any $v \in V_{\varphi_j,(\lambda)}$

$$\varphi_j(\varphi_i(v)) = \varphi_i(\varphi_j(v)) = \varphi_i(\lambda v) = \lambda(\varphi_i(v)),$$

i.e.,

$$\varphi_i(V_{\varphi_j,(\lambda)}) \subset V_{\varphi_j,(\lambda)}.$$

The family $\varphi_j|_W$ commute and each $\varphi_j|_W$ is diagonalisable, therefore by the inductive assumption the family is simultaneously diagonalisable (by repeating the argument on each eigenspace of φ_j). $\square$

Vandermonde Determinant

Proposition

For any $\alpha_1, \dots, \alpha_k \in \mathbb{R}$

$$V(\alpha_1, \dots, \alpha_k) = \det \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ \alpha_1 & \alpha_2 & \alpha_3 & \dots & \alpha_k \\ \alpha_1^2 & \alpha_2^2 & \alpha_3^2 & \dots & \alpha_k^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{k-1} & \alpha_2^{k-1} & \alpha_3^{k-1} & \dots & \alpha_k^{k-1} \end{bmatrix} =$$

$$= \prod_{1 \leqslant i < j \leqslant k} (\alpha_j - \alpha_i).$$

Vandermonde Determinant

Proposition

For any $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$

$$V(\alpha_1, \ldots, \alpha_k) = \det \begin{bmatrix} 1 & 1 & 1 & \ldots & 1 \\ \alpha_1 & \alpha_2 & \alpha_3 & \ldots & \alpha_k \\ \alpha_1^2 & \alpha_2^2 & \alpha_3^2 & \ldots & \alpha_k^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{k-1} & \alpha_2^{k-1} & \alpha_3^{k-1} & \ldots & \alpha_k^{k-1} \end{bmatrix} =$$

$$= \prod_{1 \leqslant i < j \leqslant k} (\alpha_j - \alpha_i).$$

Proof.

Proof by induction on k. For $k = 2$

$$V(\alpha_1, \alpha_2) = \det \begin{bmatrix} 1 & 1 \\ \alpha_1 & \alpha_2 \end{bmatrix} = \alpha_2 - \alpha_1.$$

Vandermonde Determinant (continued)

Proof.

$$V(\alpha_1, \ldots, \alpha_k) = \det \begin{bmatrix} 1 & 1 & 1 & \ldots & 1 \\ \alpha_1 & \alpha_2 & \alpha_3 & \ldots & \alpha_k \\ \alpha_1^2 & \alpha_2^2 & \alpha_3^2 & \ldots & \alpha_k^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{k-1} & \alpha_2^{k-1} & \alpha_3^{k-1} & \ldots & \alpha_k^{k-1} \end{bmatrix} \overset{\substack{r_k - \alpha_1 r_{k-1} \\ r_{k-1} - \alpha_1 r_{k-2} \\ \vdots \\ r_3 - \alpha_1 r_2 \\ r_2 - \alpha_1 r_1}}{=}$$

Vandermonde Determinant (continued)

Proof.

$$V(\alpha_1,\dots,\alpha_k) = \det \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ \alpha_1 & \alpha_2 & \alpha_3 & \dots & \alpha_k \\ \alpha_1^2 & \alpha_2^2 & \alpha_3^2 & \dots & \alpha_k^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{k-1} & \alpha_2^{k-1} & \alpha_3^{k-1} & \dots & \alpha_k^{k-1} \end{bmatrix} \overset{\substack{r_k - \alpha_1 r_{k-1} \\ r_{k-1} - \alpha_1 r_{k-2} \\ \vdots \\ r_3 - \alpha_1 r_2 \\ r_2 - \alpha_1 r_1}}{=}$$

$$= \det \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & \alpha_2 - \alpha_1 & \alpha_3 - \alpha_1 & \dots & \alpha_k - \alpha_1 \\ 0 & \alpha_2^2 - \alpha_1\alpha_2 & \alpha_3^2 - \alpha_1\alpha_3 & \dots & \alpha_k^2 - \alpha_1\alpha_k \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \alpha_2^{k-1} - \alpha_1\alpha_2^{k-2} & \alpha_3^{k-1} - \alpha_1\alpha_3^{k-2} & \dots & \alpha_k^{k-1} - \alpha_1\alpha_k^{k-2} \end{bmatrix} =$$

Vandermonde Determinant (continued)

Proof.
(by the Laplace formula along the first column)

$$= \det \begin{bmatrix} \alpha_2 - \alpha_1 & \alpha_3 - \alpha_1 & \dots & \alpha_k - \alpha_1 \\ \alpha_2^2 - \alpha_1\alpha_2 & \alpha_3^2 - \alpha_1\alpha_3 & \dots & \alpha_k^2 - \alpha_1\alpha_k \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_2^{k-1} - \alpha_1\alpha_2^{k-2} & \alpha_3^{k-1} - \alpha_1\alpha_3^{k-2} & \dots & \alpha_k^{k-1} - \alpha_1\alpha_k^{k-2} \end{bmatrix} =$$

$$= \det \begin{bmatrix} \alpha_2 - \alpha_1 & \alpha_3 - \alpha_1 & \dots & \alpha_k - \alpha_1 \\ (\alpha_2 - \alpha_1)\alpha_2 & (\alpha_3 - \alpha_1)\alpha_3 & \dots & (\alpha_k - \alpha_1)\alpha_k \\ \vdots & \vdots & \ddots & \vdots \\ (\alpha_2 - \alpha_1)\alpha_2^{k-2} & (\alpha_3 - \alpha_1)\alpha_3^{k-2} & \dots & (\alpha_k - \alpha_1)\alpha_k^{k-2} \end{bmatrix} =$$

(by dividing the j-th column by the factor $(\alpha_{j+1} - \alpha_1)$)

Vandermonde Determinant (continued)

Proof.

$$= \prod_{1\leqslant 1<j\leqslant k} (\alpha_j - \alpha_1) \det \begin{bmatrix} 1 & 1 & \dots & 1 \\ \alpha_2 & \alpha_3 & \dots & \alpha_k \\ \alpha_2^2 & \alpha_3^2 & \dots & \alpha_k^2 \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_2^{k-2} & \alpha_3^{k-2} & \dots & \alpha_k^{k-2} \end{bmatrix} =$$

(by the inductive assumption)

$$= \prod_{1\leqslant 1<j\leqslant k} (\alpha_j - \alpha_1) \prod_{2\leqslant i<j\leqslant k} (\alpha_j - \alpha_i) = V(\alpha_1, \dots, \alpha_k).$$

Polynomial Interpolation

Proposition

Let $x_1, \ldots, x_{n+1} \in \mathbb{R}$ be pairwise distinct points, i.e. $x_i \neq x_j$ for all $1 \leqslant i < j \leqslant n+1$. For any $y_1, \ldots, y_{n+1} \in \mathbb{R}$ there exists a unique polynomial $P(x)$ of degree at most n such that

$$P(x_i) = y_i \text{ for } i = 1, \ldots, n+1.$$

Polynomial Interpolation

Proposition

Let $x_1, \dots, x_{n+1} \in \mathbb{R}$ be pairwise distinct points, i.e. $x_i \neq x_j$ for all $1 \leqslant i < j \leqslant n+1$. For any $y_1, \dots, y_{n+1} \in \mathbb{R}$ there exists a unique polynomial $P(x)$ of degree at most n such that

$$P(x_i) = y_i \text{ for } i = 1, \dots, n+1.$$

Proof.

The polynomial $P(x) = y$ is given by the equation

$$\det \begin{bmatrix} 1 & 1 & 1 & 1 & \dots & 1 \\ x & x_1 & x_2 & x_3 & \dots & x_{n+1} \\ x^2 & x_1^2 & x_2^2 & x_3^2 & \dots & x_{n+1}^2 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ x^n & x_1^n & x_2^n & x_3^n & \dots & x_{n+1}^n \\ y & y_1 & y_2 & y_3 & \dots & y_{n+1} \end{bmatrix} = 0.$$

Polynomial Interpolation (continued)

Remark

Note that the coefficient of y *is equal to* $(-1)^{n+1}V(x_1,\ldots,x_{n+1})$ *hence it is non–zero.*

Polynomial Interpolation (continued)

Remark

Note that the coefficient of y is equal to $(-1)^{n+1}V(x_1,\dots,x_{n+1})$ hence it is non–zero.

Remark

Equivalently,

$$P(x) = \sum_{i=1}^{n+1} y_i P_i(x),$$

where

$$P_i(x) = \frac{V(x_1,\dots,x_{i-1},x,x_{i+1},\dots,x_{n+1})}{V(x_1,\dots,x_{n+1})},$$

for $i = 1,\dots,n+1$ are polynomials of degree n such that

$$P_i(x_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Polynomial Interpolation (continued)

Example

The equation of the line passing through points $(x_1, y_1), (x_2, y_2)$ is

$$y = y_1 \frac{x_2 - x}{x_2 - x_1} + y_2 \frac{x - x_1}{x_2 - x_1}.$$

Polynomial Interpolation (continued)

Example

The equation of the line passing through points $(x_1, y_1), (x_2, y_2)$ is

$$y = y_1 \frac{x_2 - x}{x_2 - x_1} + y_2 \frac{x - x_1}{x_2 - x_1}.$$

Example

The equation of the parabola (or a line) passing through points $(x_1, y_1), (x_2, y_2), (x_3, y_3)$ is

$$y = y_1 \frac{(x_3 - x)(x_3 - x_2)(x_2 - x)}{(x_3 - x_1)(x_3 - x_2)(x_2 - x_1)} + y_2 \frac{(x_3 - x_1)(x_3 - x)(x - x_1)}{(x_3 - x_1)(x_3 - x_2)(x_2 - x_1)} +$$

$$+ y_3 \frac{(x - x_1)(x - x_2)(x_2 - x_1)}{(x_3 - x_1)(x_3 - x_2)(x_2 - x_1)} = y_1 \frac{(x_3 - x)(x_2 - x)}{(x_3 - x_1)(x_2 - x_1)} + y_2 \frac{(x_3 - x)(x - x_1)}{(x_3 - x_2)(x_2 - x_1)} +$$

$$+ y_3 \frac{(x - x_1)(x - x_2)}{(x_3 - x_1)(x_3 - x_2)}.$$

Polynomial Interpolation (continued)

Remark

The polynomial $P(x)$ *is called* **Lagrange interpolation polynomial**.

Polynomial Interpolation (continued)

Remark

The polynomial $P(x)$ is called **Lagrange interpolation polynomial**.

Proposition

If function $f \in \mathcal{C}^{n+1}([a,b])$ and polynomial $P(x)$ of degree at most n satisfy

$$P(x_i) = f(x_i) \text{ for } i = 1, \ldots, n+1$$

for pairwise distinct $x_1, \ldots, x_{n+1} \in [a,b]$ then for any $x \in [a,b]$ there exists $\min(x, x_1, \ldots, x_{n+1}) < \xi < \max(x, x_1, \ldots, x_{n+1})$ such that

$$f(x) - P(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \prod_{i=1}^{n+1} (x - x_i).$$

Example

Let $f(x) = \cos x$. Then there exists a unique polynomial $P(x)$ of degree at most 11, which attains the same values as the function $f(x)$ at 12 pairwise distinct points $x_1, \ldots, x_{12} \in [0, \pi]$ and

$$|f(x) - P(x)| \leqslant \frac{\pi^{12}}{12!} < 0.002$$

for any $x \in [0, \pi]$.

Equation of an Affine Hyperplane

Proposition

Let $p_1, \ldots, p_n \in \mathbb{R}^n$ be n points, where
$p_1 = (p_{11}, \ldots, p_{1n}), p_2 = (p_{21}, \ldots, p_{2n}), \ldots, p_{n-1} = (p_{n1}, \ldots, p_{nn})$.
An equation of an affine hyperplane passing through $p_1, \ldots, p_n$ (if it is unique up to a non–zero constant, i.e., $p_1, \ldots, p_n$ do not lie on an affine subspace of dimension $n-2$) is given by the equation

$$\det \begin{bmatrix} x_1 & x_2 & \cdots & x_n & 1 \\ p_{11} & p_{12} & \cdots & p_{1n} & 1 \\ p_{21} & p_{22} & \cdots & p_{2n} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ p_{(n-1)1} & p_{(n-1)2} & \cdots & p_{(n-1)n} & 1 \\ p_{n1} & p_{n2} & \cdots & p_{nn} & 1 \end{bmatrix} = 0.$$

Example

Equation of a line passing through points $(a_1, b_1), (a_2, b_2) \in \mathbb{R}^2$ is equal to

$$\det \begin{bmatrix} x_1 & x_2 & 1 \\ a_1 & b_1 & 1 \\ a_2 & b_2 & 1 \end{bmatrix} = 0,$$

that is, by the Laplace expansion along the first row,

$$(b_1 - b_2)x_1 - (a_1 - a_2)x_2 + a_1 b_2 - a_2 b_1 = 0.$$

For example, the line passing through $(1, 2), (2, 5)$ has equation

$$\det \begin{bmatrix} x_1 & x_2 & 1 \\ 1 & 2 & 1 \\ 2 & 5 & 1 \end{bmatrix} = 0,$$

i.e.,

$$-3x_1 + x_2 + 1 = 0.$$

Hoffman–Wieland Theorem

Proposition

Let $A, B \in M(n \times n; \mathbb{R})$ be two symmetric matrices, i.e. $A = A^\intercal, B = B^\intercal$. Let $\lambda_1(A) \leqslant \lambda_2(A) \leqslant \dots \lambda_n(A)$ and $\lambda_1(B) \leqslant \lambda_2(B) \leqslant \dots \lambda_n(B)$ be all eigenvalues of A and B. Then

$$\sum_{i=1}^{n} (\lambda_i(A) - \lambda_i(B))^2 \leqslant \|A - B\|_F^2.$$

Proof.

Let $A = QD_AQ^\intercal$ and $B = PD_BP^\intercal$ be spectral decompositions of A and B, respectively, where D_A, D_B are diagonal matrices with weakly increasing elements along the main diagonal. Then

$$\sum_{i=1}^{n} (\lambda_i(A) - \lambda_i(B))^2 = \operatorname{Tr}\left((D_A - D_B)^2\right) =$$

$$= \operatorname{Tr}\left(D_A^2\right) - 2\operatorname{Tr}(D_AD_B) + \operatorname{Tr}\left(D_B^2\right).$$

Hoffman–Wielandt Theorem

Proof.

Moreover

$$\|A - B\|_F^2 = \mathrm{Tr}\big((A-B)^2\big) = \mathrm{Tr}\left((QD_AQ^\intercal - PD_BP^\intercal)^2\right) =$$

$$\mathrm{Tr}\big(D_A^2\big) - 2\,\mathrm{Tr}(QD_AQ^\intercal PD_BP^\intercal) + \mathrm{Tr}\big(D_B^2\big).$$

Let $U = [u_{ij}] = Q^\intercal P$ be an orthogonal matrix. Then the inequality is equivalent to

$$\mathrm{Tr}(U^\intercal D_A U D_B) \leqslant \mathrm{Tr}(D_A D_B),$$

or equivalently

$$\sum_{i,j=1}^{n} \lambda_i(A)\lambda_j(B)u_{ij}^2 \leqslant \sum_{i=1}^{n} \lambda_i(A)\lambda_i(B).$$

Hoffman–Wielandt Theorem (continued)

Proof.

$$\sum_{i,j=1}^{n} \lambda_i(A)\lambda_j(B)u_{ij}^2 \leqslant \sum_{i=1}^{n} \lambda_i(A)\lambda_i(B).$$

The left–hand side is a linear function in the entries of a doubly–stochastic matrix, and, it assumes a maximum at a vertex of the polyhedron of doubly–stochastic matrices, which, by the Birkhoff–von Neumann, is a matrix of some permutation $\sigma \in S_n$. The theorem follows by the rearrangement inequality. $\square$

Hoffman–Wielandt Theorem (continued)

Remark

The inequality becomes an equation when it is possible to diagonalize A and B simultaneously keeping the order of eigenvalues.

Remark

Similar inequality holds for any complex matrices and its eigenvalues.

Rearrangement Inequality

Proposition

For any real numbers

$$x_1 \leqslant x_2 \leqslant \ldots \leqslant x_n,$$

$$y_1 \leqslant y_2 \leqslant \ldots \leqslant y_n,$$

and any permutation $\sigma \in S_n$

$$x_1 y_n + x_2 y_{n-1} + \ldots + x_n y_1 \leqslant$$

$$\leqslant x_1 y_{\sigma(1)} + x_2 y_{\sigma(2)} + \ldots + x_n y_{\sigma(n)} \leqslant$$

$$\leqslant x_1 y_1 + x_2 y_2 + \ldots + x_n y_n.$$

Proof.

(sketch, by induction) If $x_1 \geqslant x_2$ and $y_1 \geqslant y_2$ then

$$(x_1 - x_2)(y_1 - y_2) \geqslant 0,$$

$$x_1 y_1 + x_2 y_2 \geqslant x_1 y_2 + x_2 y_1.$$

Rearrangement Inequality

Proof.

Let σ be a permutation maximizing the product and assume there exists i such that $\sigma(i) = j \neq i$ and choose the biggest such i. Then there exists $k < i$ such that $\sigma(k) = i$. Consider the terms

$$x_i y_{\sigma(i)} = x_i y_j, \quad \text{and} \quad x_k y_{\sigma(k)} = x_k y_i.$$

Since $j < i$ we have that

$$x_k \geqslant x_i \quad \text{and} \quad y_j \geqslant y_i.$$

If $x_k = x_i$ or $y_j = y_i$ then modifying the permutation σ such that $\sigma'(i) = i$ and $\sigma'(k) = j$ and $\sigma'(m) = \sigma(m)$ otherwise does not change the sum. Assume that

$$x_k > x_i \quad \text{and} \quad y_j > y_i.$$

But then

$$\sum_i x_i y_{\sigma'(i)} > \sum_i x_i y_{\sigma(i)}.$$

Minimax Theorems

Proposition

Let $M \in M(n \times n; \mathbb{C})$ be any Hermitian matrix with real eigenvalues

$$\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n,$$

and an orthogonal basis of $\mathbb{C}^n$

$$\mathcal{B} = (v_1, \ldots, v_n),$$

such that

$$Mv_j = \lambda_j v_j,$$

for $j = 1, \ldots, j$. Then

$$\min_{\substack{v \in V, v \neq 0 \\ \dim V = k}} R(M, v) \leqslant \lambda_k \leqslant \max_{\substack{v \in V, v \neq 0 \\ \dim V = n-k+1}} R(M, v)$$

The inequalities are sharp (for the upper bound take for example $V = \mathrm{lin}(v_k, \ldots, v_n)$ and $v = v_k$).

Minimax Theorems (continued)

Proof.

Let $V \subset \mathbb{C}^n$ be a subspace such that $\dim V = n - k + 1$. By the dimension count there exists $w \neq 0$ such that $\|w\| = 1$ and

$$w \in V \cap \mathrm{lin}(v_1, \ldots, v_k).$$

Assume

$$w = \sum_{j=1}^{k} \alpha_j v_j.$$

Then

$$\max_{\substack{v \in V, v \neq 0 \\ \dim V = n-k+1}} R(M, v) \geqslant R(M, w) = \sum_{j=1}^{k} \lambda_j |\alpha_j|^2 \geqslant \lambda_k \sum_{j=1}^{k} |\alpha_j|^2 = \lambda_k.$$

Obviously $R(M, v_k) = \lambda_k$.

Minimax Theorems (continued)

Proof.

The second inequality follows in a similar manner by considering (exercise)

$$w \in V \cap \mathrm{lin}(v_k, \ldots, v_n).$$

$\square$

Minimax Theorems (continued)

Proof.

The second inequality follows in a similar manner by considering (exercise)

$$w \in V \cap \operatorname{lin}(v_k, \ldots, v_n).$$

$\square$

Remark

The same proof works for a real symmetric matrix $A \in M(n \times n; \mathbb{R})$.

Courant–Fischer Theorem

Corollary

Let $M \in M(n \times n; \mathbb{C})$ be any Hermitian matrix with real eigenvalues

$$\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n.$$

Then

$$\max_{\dim V = k} \min_{v \in V, v \neq 0} R(M, v) = \lambda_k = \min_{\dim V = n-k+1} \max_{v \in V, v \neq 0} R(M, v).$$