
Review:

Linear algebra and vector spaces

Keywords: vector, matrix, eigenvalue, eigenvector, diagonalization, linear transformation, vector space

Compare: Vorlesung Mathematik 2

Contents

1	Matrices and vectors	2
1.1	Real Vectors	2
1.2	Real Matrices	2
1.3	Linear transformations and matrices	3
1.4	Complex matrices and vectors	4
1.5	Matrix calculus	4
2	Eigenvalues and eigenvectors	5
2.1	Definition and determination	5
2.2	*Generalized Eigenvectors*	6
3	Diagonalization	8
4	Vector spaces	9

1 Matrices and vectors

1.1 Real Vectors

- n -dimensional space \mathbb{R}^n
- elements $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ are called n -vectors

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = (x_1 \ x_2 \ \dots \ x_n)^T \quad \text{and} \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

- scalar product and norm:

$$\begin{aligned} \mathbf{x} \bullet \mathbf{y} &= x_1 y_1 + x_2 y_2 + \dots + x_n y_n \\ \|\mathbf{x}\| &= \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} \end{aligned}$$

$$\mathbf{x} \bullet \mathbf{y} = \|\mathbf{x}\| \cdot \|\mathbf{y}\| \cdot \cos \angle(\mathbf{x}, \mathbf{y})$$

- $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k \in \mathbb{R}^n$
 - If $a_1, a_2, \dots, a_k \in \mathbb{R}$, then $\mathbf{z} = a_1 \mathbf{x}_1 + a_2 \mathbf{x}_2 + \dots + a_k \mathbf{x}_k$ is called a linear combination of $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$.
 - $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ are called linearly dependent, if there exist $b_1, b_2, \dots, b_k \in \mathbb{R}$ such that $b_1 \mathbf{x}_1 + b_2 \mathbf{x}_2 + \dots + b_k \mathbf{x}_k = \mathbf{0}$ and not all $b_j = 0$.
 - $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ are called linearly independent, if a linear combination of the zero vector

$$b_1 \mathbf{x}_1 + b_2 \mathbf{x}_2 + \dots + b_k \mathbf{x}_k = \mathbf{0}$$

is possible only with $b_1 = b_2 = \dots = b_k = 0$.

1.2 Real Matrices

$\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m \in \mathbb{R}^n$

$$\mathbf{a}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}, \mathbf{a}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix}, \dots, \mathbf{a}_m = \begin{pmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{nm} \end{pmatrix} \quad \rightarrow \quad \mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}$$

is called an $n \times m$ matrix.

Notation: $\mathbf{A} \in \mathbb{R}^{n \times m}$

- The inverse matrix \mathbf{A}^{-1} of the $n \times n$ matrix $\mathbf{A} = (a_{ij})$ is defined by

$$\mathbf{A}^{-1} \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{A}^{-1} = \mathbf{I}_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

- For the $n \times n$ matrix \mathbf{A} let \mathbf{A}_{ij} denote the $(n-1) \times (n-1)$ submatrix of \mathbf{A} generated by cancelling the i -th row and the j -th column of \mathbf{A} . Then the determinant $\det \mathbf{A}$ is given (recursively) by

$$\det(\mathbf{A}) = |\mathbf{A}| = a_{11} \det \mathbf{A}_{11} - a_{12} \det \mathbf{A}_{12} + \cdots + (-1)^{n+1} a_{1n} \det \mathbf{A}_{1n}$$

- $\det(\mathbf{A} \cdot \mathbf{B}) = \det(\mathbf{A}) \cdot \det(\mathbf{B})$

Example 1.1

$$\begin{vmatrix} 1 & 1 & 3 & 3 \\ 1 & 2 & 1 & 2 \\ 1 & -2 & 1 & -2 \\ 0 & 1 & -2 & -1 \end{vmatrix} \\ = 1 \cdot \begin{vmatrix} 2 & 1 & 2 \\ -2 & 1 & -2 \\ 1 & -2 & -1 \end{vmatrix} - 1 \cdot \begin{vmatrix} 1 & 1 & 2 \\ 1 & 1 & -2 \\ 0 & -2 & -1 \end{vmatrix} + 3 \cdot \begin{vmatrix} 1 & 2 & 2 \\ 1 & -2 & -2 \\ 0 & 1 & -1 \end{vmatrix} - 3 \cdot \begin{vmatrix} 1 & 2 & 1 \\ 1 & -2 & 1 \\ 0 & 1 & -2 \end{vmatrix}.$$

1.3 Linear transformations and matrices

Definition 1.1 A linear transformation is a map $T : \mathbb{R}^m \rightarrow \mathbb{R}^n$ such that for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$ and all $\lambda, \mu \in \mathbb{R}$ we have:

$$T(\lambda \cdot \mathbf{x} + \mu \cdot \mathbf{y}) = \lambda \cdot T(\mathbf{x}) + \mu \cdot T(\mathbf{y})$$

Each $n \times m$ matrix \mathbf{A} defines a linear transformation by matrix multiplication

$$T_{\mathbf{A}}(\mathbf{x}) = \mathbf{A} \cdot \mathbf{x} = x_1 \mathbf{a}_1 + \cdots + x_m \mathbf{a}_m.$$

The image of the vector $\mathbf{x} \in \mathbb{R}^m$ is a linear combination of the column vectors of the matrix \mathbf{A} .

1.4 Complex matrices and vectors

Sometimes it is helpful to allow complex matrices and vectors (matrices whose elements are complex numbers). A complex matrix can be viewed as a combination of two real matrices:

$$\begin{aligned} \mathbf{A} &= \begin{pmatrix} a_{11} + ib_{11} & a_{12} + ib_{12} & \dots & a_{1m} + ib_{1m} \\ a_{21} + ib_{21} & a_{22} + ib_{22} & \dots & a_{2m} + ib_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} + ib_{n1} & a_{n2} + ib_{n2} & \dots & a_{nm} + ib_{nm} \end{pmatrix} \\ &= \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} + i \cdot \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nm} \end{pmatrix} \end{aligned}$$

1.5 Matrix calculus

- 1a. $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$
- 1b. $\mathbf{AB} \neq \mathbf{BA}$
- 2a. $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$
- 2b. $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$
- 3a. $\mathbf{A} + \mathbf{0} = \mathbf{A}$
- 3b. $\mathbf{AI} = \mathbf{IA} = \mathbf{A}$, (\mathbf{A} square)
4. $\mathbf{AB} = \mathbf{0} \not\Rightarrow \mathbf{A} = \mathbf{0}$ or $\mathbf{B} = \mathbf{0}$
5. $\mathbf{AB} = \mathbf{AC} \not\Rightarrow \mathbf{B} = \mathbf{C}$
6. $\lambda(\mathbf{A} + \mathbf{B}) = \lambda\mathbf{A} + \lambda\mathbf{B} \quad \lambda \in \mathbb{R}$
7. $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$
8. $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$
9. $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$
10. $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$
11. $(\mathbf{A}^T)^T = \mathbf{A}$
12. $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$
13. $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$
14. $(\mathbf{A}^{-1})^T = (\mathbf{A}^T)^{-1}$

For $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with $ad - bc \neq 0$ is $\mathbf{A}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

All these definitions and results can be generalized to vectors and matrices with complex entries.

2 Eigenvalues and eigenvectors

2.1 Definition and determination

Definition 2.1 If \mathbf{A} is a real (or complex) $n \times n$ matrix, then a (complex) number λ is an eigenvalue of \mathbf{A} if there is a nonzero (complex) vector $\mathbf{x} \in \mathbb{C}^n$ such that

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

Then \mathbf{x} is an eigenvector of \mathbf{A} (associated with λ).

Remark: If \mathbf{x} is an eigenvector associated with the eigenvalue λ , then so is $\alpha\mathbf{x}$ for every real number $\alpha \neq 0$.

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

How to find eigenvalues? The equation can be written as

$$\begin{aligned} \mathbf{A}\mathbf{x} &= \lambda\mathbf{x} \\ \Leftrightarrow \mathbf{A}\mathbf{x} - \lambda\mathbf{I}\mathbf{x} &= \mathbf{0} \\ \Leftrightarrow (\mathbf{A} - \lambda\mathbf{I})\mathbf{x} &= \mathbf{0} \end{aligned}$$

This is a homogeneous linear system of equations. It has a solution $\mathbf{x} \neq \mathbf{0}$ if and only if the matrix $(\mathbf{A} - \lambda\mathbf{I})$ is singular which means that it has determinant equal to 0.

$$(\mathbf{A} - \lambda\mathbf{I}) \text{ singular} \Leftrightarrow \underbrace{\det(\mathbf{A} - \lambda\mathbf{I})}_{p_A(\lambda)} = 0$$

$p_A(\lambda) = 0$ is called the characteristic equation of \mathbf{A} . The function $p_A(\lambda)$ is a polynomial of degree n in λ , called the characteristic polynomial of \mathbf{A} .

Determination of the eigenvalues and eigenvectors

1. The polynomial equation $p_A(\lambda) = 0$ has always n complex solutions (counted with multiplicity) and may have no real solutions. If $\lambda_1, \dots, \lambda_r \in \mathbb{C}$ are the pairwise distinct solutions (the eigenvalues of \mathbf{A}) with the multiplicities k_1, \dots, k_r then the characteristic polynomial can be written as

$$p_A(\lambda) = (\lambda_1 - \lambda)^{k_1} (\lambda_2 - \lambda)^{k_2} \dots (\lambda_r - \lambda)^{k_r}.$$

The multiplicity k_i of the zero λ_i is called algebraic multiplicity of the eigenvalue λ_i . Generally, the determination of the (exact) zeros is impossible for $n \geq 5$ and we have to use numerical methods.

2. For each eigenvalue λ_i ($1 \leq i \leq r$) we compute the so called eigenspace for λ_i

$$V(\lambda_i) = \{ \mathbf{x} \in \mathbb{C}^n \mid (\mathbf{A} - \lambda_i\mathbf{I})\mathbf{x} = \mathbf{0} \}.$$

The dimension of the vector space $V(\lambda_i)$ is called the geometric multiplicity of the eigenvalue λ_i .

Definition 2.2 The spectral radius of a quadratic matrix A is the real number

$$\rho(A) := \max\{|\lambda_1|, \dots, |\lambda_r|\}.$$

2.2 *Generalized Eigenvectors*

To solve some interesting problems we have to generalize the notion of eigenvectors.

Definition 2.3 A vector $\mathbf{x} \in \mathbb{C}^n$ is called generalized eigenvector of degree $l \in \mathbb{N}$ associated to the eigenvalue λ of \mathbf{A} , if

$$(\mathbf{A} - \lambda \mathbf{I})^l \mathbf{x} = \mathbf{0} \quad \text{and} \quad (\mathbf{A} - \lambda \mathbf{I})^{l-1} \mathbf{x} \neq \mathbf{0}.$$

Of course, an eigenvector is a generalized eigenvector of degree 1.

Example 2.1 The matrix

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

has the eigenvalue 1 of (algebraic) multiplicity 3 with $\dim V(1) = 1$ (geometric multiplicity). We have:

$$\begin{array}{lll} (\mathbf{A} - \mathbf{I}) \mathbf{e}_1 = \mathbf{0} & (\mathbf{A} - \mathbf{I}) \mathbf{e}_2 = \mathbf{e}_1 & (\mathbf{A} - \mathbf{I})^2 \mathbf{e}_2 = \mathbf{0} \\ (\mathbf{A} - \mathbf{I}) \mathbf{e}_3 = \mathbf{e}_1 + \mathbf{e}_2 & (\mathbf{A} - \mathbf{I})^2 \mathbf{e}_3 = \mathbf{e}_1 & (\mathbf{A} - \mathbf{I})^3 \mathbf{e}_3 = \mathbf{0} \end{array}$$

This means, that \mathbf{e}_1 is an eigenvector, \mathbf{e}_2 a generalized eigenvector of degree 2 and \mathbf{e}_3 a generalized eigenvector of degree 3.

Theorem 2.1 Let $\mathbf{A} \in \mathbb{C}^{n \times n}$ be a complex (or real) matrix with

$$p_{\mathbf{A}}(\lambda) = (\lambda_1 - \lambda)^{k_1} (\lambda_2 - \lambda)^{k_2} \dots (\lambda_r - \lambda)^{k_r}.$$

- Let λ be an eigenvalue of \mathbf{A} of (algebraic) multiplicity l . Then there exist l linearly independent generalized eigenvectors (of degree $\leq l$). This means:

$$\dim\{ \mathbf{x} \in \mathbb{C}^n \mid (\mathbf{A} - \lambda \mathbf{I})^l \mathbf{x} = \mathbf{0} \} = l.$$

- Generalized eigenvectors associated to pairwise different eigenvalues of \mathbf{A} are linearly independent.
- There exists a basis $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n$ of \mathbb{C}^n consisting of generalized eigenvectors of \mathbf{A} . If \mathbf{P} is the matrix with this basis as the columns, then

$$\mathbf{P}^{-1} \mathbf{A} \mathbf{P} = \begin{pmatrix} \boxed{\mathbf{A}_1} & & & \mathbf{0} \\ & \boxed{\mathbf{A}_2} & & \\ & & \ddots & \\ \mathbf{0} & & & \boxed{\mathbf{A}_r} \end{pmatrix}$$

with $\mathbf{A}_i \in \mathbb{C}^{k_i \times k_i}$ for all $i = 1, 2, \dots, r$.

Example 2.2 Let $n = 2$ and $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

1. Characteristic polynomial:

$$\begin{aligned} p_A(\lambda) &= \det \begin{pmatrix} a - \lambda & b \\ c & d - \lambda \end{pmatrix} \\ &= \lambda^2 - \underbrace{(a + d)}_{=: \text{tr}(A)} \lambda + \underbrace{ad - bc}_{=: \det(A)} = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \end{aligned}$$

$$\text{with } \lambda_{1,2} = \frac{a + d}{2} \pm \sqrt{\frac{(a + d)^2}{4} - \det(A)}.$$

2. For each λ_i ($i = 1, 2$) we solve the linear system

$$\begin{pmatrix} a - \lambda_i & b \\ c & d - \lambda_i \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

If $n = 2$, we have four different cases:

1. $\lambda_1, \lambda_2 \in \mathbb{R}$, $\lambda_1 \neq \lambda_2$

$$\text{Example: } \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

2. $\lambda = \lambda_1 = \lambda_2 \in \mathbb{R}$ with $\dim V(\lambda) = 2$

$$\text{Example: } \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

3. $\lambda = \lambda_1 = \lambda_2 \in \mathbb{R}$ with $\dim V(\lambda) = 1$

$$\text{Example: } \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$$

4. $\lambda_2 = \overline{\lambda_1} \in \mathbb{C} - \mathbb{R}$

$$\text{Example: } \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix} \text{ with } \phi \neq k\pi$$

3 Diagonalization

Let \mathbf{A} and \mathbf{P} be $n \times n$ matrices with \mathbf{P} invertible. Then \mathbf{A} and $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ have the same eigenvalues (because they have the same characteristic polynomial).

Definition 3.1 An $n \times n$ matrix \mathbf{A} is diagonalizable if there is an invertible matrix \mathbf{P} and a diagonal matrix \mathbf{D} such that

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

Two natural questions:

1. Which square matrices are diagonalizable?
2. If \mathbf{A} is diagonalizable, how do we find the matrix \mathbf{P} ?

Theorem 3.1 An $n \times n$ matrix \mathbf{A} is diagonalizable if and only if it has a set of n linearly independent eigenvectors $\mathbf{p}_1, \dots, \mathbf{p}_n$. In this case,

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \text{diag}(\lambda_1, \dots, \lambda_n),$$

where \mathbf{P} is the matrix with $\mathbf{p}_1, \dots, \mathbf{p}_n$ as its columns, and $\lambda_1, \dots, \lambda_n$ are the corresponding eigenvalues.

Many of the matrices encountered in economics are (real) symmetric and for these matrices we have the following important result.

Theorem 3.2 (Spectral Theorem for symmetric matrices) If the real $n \times n$ matrix \mathbf{A} is symmetric ($\mathbf{A} = \mathbf{A}^T$), then:

1. All n eigenvalues $\lambda_1, \dots, \lambda_n$ are real.
2. Eigenvectors that correspond to different eigenvalues are orthogonal.
3. There exists an orthogonal and real matrix \mathbf{P} ($\mathbf{P}^{-1} = \mathbf{P}^T$) such that

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \text{diag}(\lambda_1, \dots, \lambda_n).$$

The columns $\mathbf{p}_1, \dots, \mathbf{p}_n$ of the matrix \mathbf{P} are eigenvectors of unit length corresponding to the eigenvalues $\lambda_1, \dots, \lambda_n$.

Example 3.1 The matrix $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$ has the eigenvalues and eigenvectors

$$\begin{aligned} \lambda_1 = 2 & & \mathbf{p}_1 &= \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ \lambda_2 = 3 & & \mathbf{p}_2 &= \begin{pmatrix} 1 \\ 2 \end{pmatrix} \end{aligned}$$

Hence $\mathbf{P} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$, $\mathbf{P}^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$ and:

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

4 Vector spaces

Definition 4.1 A (real) vector space is a set V together with two operations

$$\begin{aligned} + : V \times V &\rightarrow V \text{ (vector addition)} \\ \cdot : \mathbb{R} \times V &\rightarrow V \text{ (scalar multiplication)} \end{aligned}$$

such that for all $u, v, w \in V$ and all $\alpha, \beta \in \mathbb{R}$:

- $u + (v + w) = (u + v) + w$
- $u + v = v + u$
- There exists an element $0 \in V$, called the zero element, such that $v + 0 = 0 + v = v$ for all $v \in V$.
- For every $v \in V$ there exists an element $-v \in V$, called the additive inverse of v , such that $v + (-v) = 0$.
- $\alpha \cdot (\beta \cdot v) = (\alpha \cdot \beta) \cdot v$
- $1 \cdot v = v$ ($1 \in \mathbb{R}$)
- $\alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v$
- $(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$

Example 4.1 The space \mathbb{R}^n with the well-known (componentwise) vector addition and scalar multiplication is a real vector space.

Example 4.2 The space \mathbb{C}^n with the well-known (componentwise) vector addition and scalar multiplication is a real vector space.

Example 4.3 The set \mathcal{F} of functions $\mathbb{R}^n \rightarrow \mathbb{R}$ can be given the structure of a (real) vector space, where the operations are defined pointwise. For any $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ and any $\alpha \in \mathbb{R}$ define:

$$\begin{aligned} + : \mathcal{F} \times \mathcal{F} &\rightarrow \mathcal{F} & (f + g)(x) &= f(x) + g(x) \\ \cdot : \mathbb{R} \times \mathcal{F} &\rightarrow \mathcal{F} & (\alpha \cdot f)(x) &= \alpha \cdot f(x) \end{aligned}$$