Mathematics 2

Ph. D. Vitalii Konarovskyi

IPSP Summer 2019

April 1 – July 10

Contents

1 Systems of Linear Equations (Lecture Notes)

1.1 Definitions

We consider the problem of finding n scalars $x_1, \ldots, x_n \in \mathbb{F}$ which satisfy:

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \qquad\qquad\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{1.1}$$

where a_{ij}, b_i $(i = 1, \ldots, m \quad j = 1, \ldots, n)$ are given numbers from $\mathbb{F}$.

Definition 1.1 *We call (1.1) a system of linear equations with n unknowns. Any set of elements $x_1, \ldots, x_n \in \mathbb{F}$ is called a solution if it satisfies the system. The system is said to be homogeneous if $b_1 = b_2 = \cdots = b_n = 0$.*

Definition 1.2 *If we multiply the j^{th} equation by a scalar $c_j \in \mathbb{F}$, $\forall\, j = 1, \ldots, m$ and then add them, we get a new equation which is called a linear combination of equations in (1.1).*

Definition 1.3 *Two systems are equivalent if each equation in each system is a linear combination of the equations in the other system.*

Theorem 1.1 *Equivalent systems have the same solutions.*

Definition 1.4 *A system is consistent if it has at least one solution, otherwise it is inconsistent.*

1.2 Matrices and Elementary Row Operations

Definition 1.5 *Given $m, n \in \mathbb{N}$, a rectangular array of numbers $a_{ij} \in \mathbb{F}$*

$$A = (a_{ij})_{i,j=1}^{m,n} = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}$$

is called an $m \times n$ matrix. The numbers a_{ij} are called the entries of A, where i indexes the rows of A and j indexes the columns of A. We also say that A has size $m \times n$.

Definition 1.6 *The set of all $m \times n$ matrices with entries from $\mathbb{F}$ is denoted $\mathbb{F}^{m \times n}$.*

Definition 1.7 *If $A, B \in \mathbb{F}^{m \times n}$, then B is row-equivalent to A if B can be obtained from A by a finite number of elementary row-operations.*

Theorem 1.2 *If A and B are row-equivalent augmented matrices of systems of linear equations, then those systems have the same solutions.*

1.3 Row-Reduced Echelon Matrices

Let A^i be the i^{th} row vector of A and A^j be the j^{th} column vector of A.

Definition 1.8 *A is in row-echelon form (REF) if the rows of A satisfy:*

1. *either A^i is the zero vector or the first non-zero entry is 1 when read from left to right*
2. *for $i = 1, \ldots, m$, if $A^i = 0$, then $A^{i+1} = A^{i+2} = \cdots = A^m = 0$*
3. *for $i = 2, \ldots, m$, if some A^i is not the zero vector, then the first non-zero entry is 1 and occurs to the right of the initial 1 in A^{i-1}*

Definition 1.9 *The initial leading 1 is called the pivot.*

Definition 1.10 *A is in reduced row-echelon form (RREF) if A is in REF and if a column A^j containing a pivot implies that the pivot is the only non-zero entry in that column.*

Example

$$\begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad REF \quad \begin{pmatrix} 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad RREF$$

Theorem 1.3 *Every $m \times n$ matrix is row-equivalent to a matrix in RREF.*

2 Vector Spaces (Lecture Notes)

2.1 Vector Spaces

Definition 2.1 *A vector space over a field $\mathbb{F}$ is a set V, along with the operations addition and multiplication, which satisfies the following conditions:*

1. $u + v = v + u \quad \forall\, u, v \in V$
2. $(v + u) + w = v + (u + w) \quad \forall\, u, v, w \in V$
3. $\exists\, 0 \in V : 0 + v = v \quad \forall\, v \in V$
4. $\forall\, v \in V \quad \exists\, w \in V : v + w = 0, \quad w := -v$
5. $1 \cdot v = v \quad \forall\, v \in V$
6. $a(u + v) = au + av$, $(a + b)v = av + bv \quad \forall\, u, v \in V$, $a, b \in \mathbb{F}$

Definition 2.2 *$U \subseteq V$ is called a subspace of V if U is a vector space over $\mathbb{F}$ under the same operations.*

Lemma 2.1 *$U \subseteq V$ is a subspace of V if and only if:*

1. $0 \in U$
2. $\forall\, u, v \in U$, $u + v \in U$
3. $\forall\, a \in \mathbb{F}, u \in U$, $au \in U$

2.2 Bases

Definition 2.3 *Vectors $v_1, \ldots, v_n \in V$ are linearly independent if the equation*

$$a_1 v_1 + \cdots + a_n v_n = 0$$

only has the solution $a_1 = a_2 = \cdots = a_n = 0$. The set

$$\mathrm{span}\,(v_1, \ldots, v_n) = \{a_1 v_1 + \cdots + a_n v_n : a_i \in \mathbb{F},\, i = 1, \ldots, n\}$$

is the linear span of vectors $v_1, \ldots, v_n$.

Definition 2.4 *Vectors $v_1, \ldots, v_n \in V$ form a basis of V if they are linearly independent and if $V = \mathrm{span}(v_1, \ldots, v_n)$.*

Theorem 2.1 *Let $v_1, \ldots, v_n$ be a basis of a vector space V. Then for each $v \in V$ there exist unique numbers $a_1, \ldots, a_n$ such that $v = a_1 v_1 + \cdots + a_n v_n$.*

Definition 2.5 *The numbers $a_1, \ldots, a_n$ are the coordinates of v relative to the basis $v_1, \ldots, v_n$.*

Definition 2.6 *The number of basis elements of a vector space V is the dimension of the vector space V and is denoted* $\dim V$*.*

Example

$$\dim \mathbb{R}^n = n$$

$$\dim \mathbb{C}^n = n$$

$$\dim \mathbb{F}^{m\times n} = m \cdot n$$

$$\dim \mathbb{F}_n[z] = n + 1$$

2.3 Linear Maps

Definition 2.7 *Let V and W be vector spaces over $\mathbb{F}$. A function $T : V \mapsto W$ is called a linear transformation if:*

1. $T(u + v) = T\,u + T\,v \quad \forall\, u, v \in V$
2. $T(av) = aT\,v \quad \forall\, a \in \mathbb{F}, v \in V$

The set of all linear transformations from V to W is denoted by $\mathcal{L}(V, W)$. If $W = V$, then $\mathcal{L}(V) := \mathcal{L}(V, V)$

Remark *The set $\mathcal{L}(V, W)$ is a vector space over $\mathbb{F}$ under the usual operations of additions of functions and multiplication of functions by a scalar.*

3 Invertible Matrices (Lecture Notes)

3.1 Matrix of a Linear Map

Definition 3.1 *Let V be a vector space with basis $v_1, \dots, v_n$ and let W be a vector space with basis $w_1, \dots, w_m$. Given a map $T \in \mathcal{L}(V, W)$, we can write the coordinates of $T v_j$ relative to the basis $w_1, \dots, w_m$ in a matrix $M_{T v_j}$:*

$$\begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

that is,

$$T v_j = a_{1j}w_1 + \cdots + a_{mj}w_m$$

We then form the following matrix:

$$M_T := \begin{pmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \dots & a_{mj} & \dots & a_{mn} \end{pmatrix}$$

The matrix M_T is called the matrix of T relative to the pair of bases $v_1, \dots, v_n$ and $w_1, \dots, w_m$.

Theorem 3.1 *Let U, V, W be vector spaces with bases $u_1, \dots, u_l$, $v_1, \dots, v_n$, $w_1, \dots, w_m$ respectively. Let T_1, T_2 be linear maps from V to W and let S be a linear map from U to V. Let M_v be a matrix representing a vector $v \in V$. Then:*

1. $M_{T v} = M_T \cdot M_v$
2. $M_{TS} = M_T \cdot M_S$
3. $M_{cT} = cM_T$, $c \in \mathbb{F}$
4. $M_{T_1+T_2} = M_{T_1} + M_{T_2}$

3.2 Isomorphism

1. A linear map T is injective if $u \neq v \Rightarrow T u \neq T v$ ($\Leftrightarrow \ker T = \{v : T v = 0\} = \{0\}$).
2. A linear map T is surjective if $\operatorname{range} T = \{T v : v \in V\} = W$.
3. A linear map T is bijective if it is both injective and surjective; for all $w \in W$ there exists a unique $v \in V$ such that $T v = w$. T is then said to be invertible and there exists T^{-1} which is the inverse of T. The vector v is then defined as $v := T^{-1}w$.
4. $\dim(\ker T) + \dim(\operatorname{range} T) = \dim V$

Definition 3.2 *Two vector spaces V, W are called isomorphic if there exists $T : V \mapsto W$ such that T is an invertible linear map.*

Theorem 3.2 *Every n-dimensional vector space over $\mathbb{F}$ is isomorphic to $\mathbb{F}^n$.*

Theorem 3.3 *Let V be an n-dimensional vector space over $\mathbb{F}$ and let W be an m-dimensional vector space over $\mathbb{F}$. Then the set of all linear maps from V to W is isomorphic to $\mathbb{F}^{m\times n}$.*

Theorem 3.4 *Vector spaces V, W are isomorphic if and only if $\dim V = \dim W$.*

3.3 Invertible Linear Maps

Theorem 3.5 *If T is an invertible linear map from V to W, then T^{-1} is a linear map from W to V.*

Definition 3.3 *The matrix $A \in \mathbb{F}^{n\times n}$ is called invertible if there exists $B \in \mathbb{F}^{n\times n}$ such that $A \cdot B = B \cdot A = I$, where I is the identity matrix*

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and we then define $B := A^{-1}$ as the inverse of A.

Theorem 3.6 *Let $A, B \in \mathbb{F}^{n\times n}$ and let V, W be vector spaces over $\mathbb{F}$.*

1. *If A is invertible, then A^{-1} is invertible and $(A^{-1})^{-1} = A$.*
2. *If A, B are invertible, then AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$.*
3. *If A is a matrix of a linear map $T \in \mathcal{L}(V, W)$, then A is invertible if and only if T is invertible. Moreover A^{-1} is the matrix of T^{-1}.*

Theorem 3.7 *If $A \in \mathbb{F}^{n\times n}$, the following conditions are equivalent:*

1. *A is invertible.*
2. *A is row-equivalent to the $n \times n$ identity matrix. Moreover if a sequence of elementary row operations reduces A to the identity matrix, then the same sequence of operations reduces I to A^{-1}.*

Theorem 3.8 *For $A \in \mathbb{F}^{n\times n}$ the following conditions are equivalent:*

1. *A is invertible.*
2. *$Ax = 0$ only has the trivial solution $x = 0$.*
3. *$Ax = b$ has a unique solution $x = A^{-1}b$.*

4 Rank of Matrices (Lecture Notes)

Definition 4.1 *The dimension of* $\operatorname{range} T = \{T\, v : v \in V\}$ *is called the rank of a linear map* T *from* V *to* W*, which are both vector spaces over* $\mathbb{F}$*.*

$$\operatorname{rank} T = \dim\left(\operatorname{range} T\right)$$

$$\dim V = \dim\left(\ker T\right) + \dim\left(\operatorname{range} T\right) = \dim\left(\ker T\right) + \operatorname{rank} T$$

Theorem 4.1 *A linear map* $T : V \mapsto W$ *is invertible if and only if*

$$\dim V = \dim W = \operatorname{rank} T$$

Proof: T is invertible if it is bijective. Then $\ker T = \{0\}$ and $\operatorname{range} T = W$, therefore

$$\dim V = \dim W = \operatorname{rank} T$$

If $\dim V = \dim W = \operatorname{rank} T$, then $\dim\left(\ker T\right) = 0$, thus T is injective. Since $T : V \mapsto W$ and $\operatorname{range} T \subseteq W$, we have that $\dim(\operatorname{range} T) = \dim W$, thus T is surjective. Therefore, T is bijective and thus invertible. □

Definition 4.2 *For* $A \in \mathbb{F}^{m\times n}$*, the maximal number of linearly independent columns is called the rank of the matrix* A *and is denoted* $\operatorname{rank} A$*.*

Theorem 4.2 *The rank of a linear map* $T \in \mathcal{L}(V, W)$ *is equal to the rank of its matrix* M_T*, that is* $\operatorname{rank} T = \operatorname{rank} M_T$*.*

Corollary 4.1 *A matrix* $A \in \mathbb{F}^{n\times n}$ *is invertible if and only if* $\operatorname{rank} A = n$*.*

Theorem 4.3 *The rank of* $A \in \mathbb{F}^{m\times n}$ *is the maximal number of linearly independent rows.*

Theorem 4.4 *The rank of a matrix is preserved under elementary row and column transformations.*

Definition 4.3 *For a given matrix* $A = (a_{ij})_{i,j=1}^{m,n} \in \mathbb{F}^{m\times n}$*, the matrix* $A^T = (a_{ji})_{j,i=1}^{n,m}$ *is the transposed matrix of* A*.*

Corollary 4.2 $\operatorname{rank} A = \operatorname{rank} A^T$

Theorem 4.5 (Rouché-Capelli Theorem) *A system of linear equations with a matrix of coefficients* A *and augmented matrix* A' *is consistent if and only if* $\operatorname{rank} A = \operatorname{rank} A'$*.*

5 Fundamental Systems of Solutions (Lecture Notes)

5.1 Rank

Theorem 5.1 *Given the matrices $A, B \in \mathbb{F}^{n\times n}$:*

$$\operatorname{rank}(AB) \leq \min\{\operatorname{rank} A,\ \operatorname{rank} B\}$$

Corollary 5.1 *$A \in \mathbb{F}^{n\times n}$ is invertible if and only if there exists $B \in \mathbb{F}^{n\times n}$ such that*

$$AB = I \quad (\text{or } BA = I)$$

5.2 General Solutions to Systems of Linear Equations

Lemma 5.1 *Solutions x and x' to a system of linear equations have a difference $x - x'$ that is a solution to the corresponding homogeneous system of linear equations $Ay = 0$.*

Corollary 5.2 *Let x' be a solution to $Ax = b$. Then the set of all solutions is the set $\{x' + y\}$, where y is a solution to the homogeneous system $Ay = 0$.*

Definition 5.1 *Let U denote the set of all solutions to the homogeneous system of linear equations. A basis of $U = \{y : Ay = 0\}$ is called a fundamental system of solutions to the equation $Ay = 0$. Let $y^1, \dots, y^k$ be a fundamental system of solutions to $Ay = 0$. Then the general solution to the equation $Ax = b$ is:*

$$x = x' + a_1 y^1 + \cdots + a_k y^k$$

where $a_1, \dots, a_k \in \mathbb{F}$ and x' is a partial solution to $Ax = b$.

Example (Finding the Fundamental System of Solutions) *Let A be given:*

$$A = \begin{pmatrix} 1 & -2 & 1 & 0 \\ 1 & -1 & 1 & -1 \\ 0 & 1 & 0 & -1 \end{pmatrix}$$

1. *Transpose A:*

$$A^T = \begin{pmatrix} 1 & 1 & 0 \\ -2 & -1 & 1 \\ 1 & 1 & 0 \\ 0 & -1 & -1 \end{pmatrix}$$

2. *Augment the identity matrix I_n to the right of A^T:*

$$\left(\begin{array}{ccc|cccc} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ -2 & -1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & -1 & 0 & 0 & 0 & 1 \end{array}\right)$$

3. *Reduce A^T to REF:*

$$\left(\begin{array}{ccc|cccc} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 & 1 & 0 & 1 \end{array}\right)$$

The fundamental system of solutions is then:

$$y^1 = \begin{pmatrix} -1 & 0 & 1 & 0 \end{pmatrix}$$

$$y^2 = \begin{pmatrix} 2 & 1 & 0 & 1 \end{pmatrix}$$

Example (Method of Solving a System of Linear Equations) *Let the following system be given:*

$$\begin{cases} x_1 + 2x_2 - x_3 - x_4 = 1 \\ x_1 + x_3 + 2x_4 = -1 \\ 2x_1 + 2x_2 + x_4 = 0 \end{cases} \tag{5.1}$$

We rearrange the system as follows:

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n - b_1x_{n+1} = 0 \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n - b_mx_{n+1} = 0 \end{cases} \longrightarrow \begin{cases} x_1 + 2x_2 - x_3 - x_4 - x_5 = 0 \\ x_1 + x_3 + 2x_4 + x_5 = 0 \\ 2x_1 + 2x_2 + x_4 = 0 \end{cases}$$

Find the fundamental system of solutions of this new system as described in the previous example and reduce it to the following form:

$$\left(\begin{array}{ccc|ccccc} 1 & 1 & 2 & 1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 2 & 0 & 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ -1 & 2 & 1 & 0 & 0 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccccc} 1 & 1 & 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -4 & 3 & 0 & 2 & 0 \\ 0 & 0 & 0 & -1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & -1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 & 1 \end{array}\right)$$

We find for this system the fundamental system of solutions:

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

$$y^2 = \begin{pmatrix} -1 & 1 & 1 & 0 \end{pmatrix}$$

and the partial solution:

$$y^0 = \begin{pmatrix} -1 & 1 & 0 & 0 \end{pmatrix}$$

We write the general solution to (5.1) in the form:

$$y = y^0 + l_1y^1 + \cdots + l_{k-1}y^{k-1}$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + l_1 \begin{pmatrix} -4 \\ 3 \\ 0 \\ 2 \end{pmatrix} + l_2 \begin{pmatrix} -1 \\ 1 \\ 1 \\ 0 \end{pmatrix}$$

6 Determinants (Lecture Notes)

6.1 Permutations

Definition 6.1 *A permutation π of n elements is a bijective map from $\{1,\ldots,n\}$ to $\{1,\ldots,n\}$:*

$$\begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ \pi_1 & \pi_2 & \pi_3 & \ldots & \pi_n \end{pmatrix}$$

The set of all permutations of n elements is denoted S_n.

Theorem 6.1 *The number of all permutations of n elements $|S_n| = n!$.*

Example (Composition)

$$\pi \circ \sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

$$\pi^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

Definition 6.2 *An inversion pair (i,j) of $\pi \in S_n$ is a pair $i,j \in \{1,\ldots,n\}$ for which $i < j$ but $\pi_i > \pi_j$.*

Example

$$\pi = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$$

$(1,3)$ *and* $(2,3)$ *are inversion pairs.*

Definition 6.3 *The sign of $\pi \in S_n$ is defined as:*

$$\operatorname{sign} \pi = (-1)^m = \begin{cases} 1, \text{ even } m \\ -1, \text{ odd } m \end{cases}$$

where m is the number of inversion pairs. π is called an even or odd permutation, depending on the value of $\operatorname{sign} \pi$.

Example (Transpositions) t_{ij} *is a transposition:*

$$t_{23} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 2 & 4 \end{pmatrix} \quad t_{24} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix}$$

Example

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 5 & 1 & 6 & 4 \end{pmatrix} = (1,\, 2,\, 3,\, 5,\, 6,\, 4)$$

Theorem 6.2 *Each permutation can be written as a composition of transpositions, where each cycle* $(i, \dots, i_k) = (i_1, i_k) \circ \dots \circ (i_1, i_2)$.

Example

$$(1,\ 2,\ 3,\ 5,\ 6,\ 4) = (1,4) \circ (1,6) \circ (1,5) \circ (1,3) \circ (1,2)$$

Theorem 6.3

$$\operatorname{sign}(\pi \circ \sigma) = \operatorname{sign}\pi \cdot \operatorname{sign}\sigma \quad \forall\, \pi, \sigma \in S_n$$

Remark *A permutation is even or odd if the number of transpositions from the decomposition in Th. 6.2 is even or odd respectively.*

6.2 Determinants

Definition 6.4 *Given* $A = (a_{ij}) \in \mathbb{F}^{n\times n}$*, the number*

$$\det A = \sum_{\pi \in S_n} \operatorname{sign}\pi\,(a_{1,\pi_1} \dots a_{n,\pi_n}) = \sum_{\pi \in S_n} \left(\operatorname{sign}\pi \prod_{i=1}^{n} a_{i,\pi_i} \right)$$

is called the determinant of A*.*

Theorem 6.4 *The determinant is a linear function of each row of the matrix:*

$$\begin{vmatrix} 1 & 2 \\ 1+2 & 2+3 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ 1 & 2 \end{vmatrix} + \begin{vmatrix} 1 & 2 \\ 2 & 3 \end{vmatrix}$$

$$\begin{vmatrix} 1c & 2c \\ 1 & 2 \end{vmatrix} = c \begin{vmatrix} 1 & 2 \\ 1 & 2 \end{vmatrix}$$

If two rows of a matrix A *are the same, then* $\det A = 0$*:*

$$\begin{vmatrix} 1 & 2 \\ 1 & 2 \end{vmatrix} = 0$$

The determinant of the identity matrix of any size is 1.

Remark *The determinant is the only map from* $\mathbb{F}^{n\times n} \mapsto \mathbb{F}$ *that satisfies the properties listed in Th. 6.4.*

Theorem 6.5 *Given $A \in \mathbb{F}^{n \times n}$:*

1. $\det A = \det A_T$
2. *If B is obtained from A by adding a multiple of one row of A to another (or a multiple of one column of A to another), then* $\det A = \det B$.
3. *Interchanging two rows or two columns introduces a factor of -1 to the determinant.*
4. $\det AB = \det A \cdot \det B$
5. *A matrix with zeros to the left of the diagonal will have a determinant equal to the product of the entries along the diagonal:*

$$\begin{vmatrix} a_{11} & \neq 0 & \neq 0 \\ 0 & \ddots & \neq 0 \\ 0 & 0 & a_{nn} \end{vmatrix} = a_{11} \cdot \dots \cdot a_{nn}$$

6. *Given matrices A, B, C of size $p \times p$, $p \times m$, $m \times m$ respectively, and the zero matrix:*

$$\begin{vmatrix} A & B \\ 0 & C \end{vmatrix} = \det A \cdot \det C$$

6.3 Computing Determinants with Cofactor Expansions

Definition 6.5 *For $i, j = 1, \dots, n$, the $i - j$ minor of A, denoted by M_{ij}, is defined to be the determinant of the matrix obtained by removing the i^{th} row and j^{th} column from A. The $i - j$ cofactor of A is $A_{ij} = (-1)^{i+j} M_{ij}$.*

Theorem 6.6 (Cofactor Expansion) *For each i^{th} row (j^{th} column, respectively) the determinant is*

$$\det A = \sum_{j=1}^{n} a_{ij} A_{ij} = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} M_{ij}$$

$$\left(\det A = \sum_{i=1}^{n} a_{ij} A_{ij} = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} M_{ij} \right)$$

7 Change of Basis (Lecture Notes)

7.1 Inverse Matrices

Definition 7.1 *The matrix*

$$\text{adj}\, A = \begin{pmatrix} A_{11} & \dots & A_{n1} \\ \vdots & & \vdots \\ A_{1n} & \dots & A_{nn} \end{pmatrix}$$

where A_{ij} is the $i-j$ minor of A, is called the classical adjoint matrix of A. Note that A_{ij} is written in the j^{th} row and i^{th} column.

Theorem 7.1 *The matrix $A \in \mathbb{F}^{n\times n}$ is invertible if and only if* $\det A \neq 0$*. If A is invertible, then*

$$A^{-1} = \frac{1}{\det A} \cdot \text{adj}\, A$$

Example

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

$$\det A = -2, \quad \text{adj}\, A = \begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix}$$

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

7.2 Cramer's Rule

Consider a system of linear equations written in the form

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

where A is the matrix of coefficients of the system. By Th. 7.1

$$\det A \cdot x = \text{adj}\, A \cdot Ax = \text{adj}\, A \cdot b$$

$$\det A\, x_j = \sum_{i=1}^{n} (\text{adj}\, A)_{ji}\, y_i = \sum_{i=1}^{n} A_{ij} y_i = \begin{vmatrix} a_{11} & \dots & b_1 & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & b_n & \dots & a_{nn} \end{vmatrix} = \det B_j$$

where B_j is the $n\times n$ matrix obtained from A by replacing the j^{th} column of A by $b = (b_1, \dots, b_n)$. Thus, the system has a unique solution given by

$$x_j = \frac{\det B_j}{\det A}$$

if and only if $\det A \neq 0$.

7.3 Change of Basis

Let V be an n-dimensional vector space over $\mathbb{F}$ with a basis $e_1, \dots, e_n$. Any vector $v \in V$ can then be written as $v = a_1 e_1 + \cdots + a_n v_n$, where $a_1, \dots, a_n \in \mathbb{F}$ are called the coordinates of v. We will denote them by

$$M_v^e = \begin{pmatrix} a_1 \\ \dots \\ a_n \end{pmatrix}_e$$

with an index e emphasizing that they are coordinates in the basis $e_1, \dots, e_n$. Let $e'_1, \dots, e'_n$ be another basis of V. Then we can write $e'_j = \sum_{i=1}^{n} \tau_{ij} e_i$.

Definition 7.2 *The matrix*

$$Q_{ee'} = Q = \begin{pmatrix} \tau_{11} & \dots & \tau_{1n} \\ \vdots & & \vdots \\ \tau_{n1} & \dots & \tau_{nn} \end{pmatrix}$$

whose columns are the columns of the coordinates of the vectors $e'_1, \dots, e'_n$ in the basis $e_1, \dots, e_n$ is called the change-of-basis matrix from the basis $e_1, \dots, e_n$ to the basis $e'_1, \dots, e'_n$. Taking $e' = (e'_1, \dots, e'_n)$, $e = (e_1, \dots, e_n)$, then

$$e' = e\, Q_{ee'}$$

Theorem 7.2 *The change-of-basis matrix $Q_{ee'}$ is invertible and $Q_{ee'}^{-1}$ is the change-of-basis matrix from e' to e.*

Now we consider the transformation of vector coordinates. Let

$$v = a_1 e_1 + \cdots + a_n e_n = a'_1 e'_1 + \cdots + a'_n e'_n$$

that is,

$$M_v^e = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}_e \qquad M_v^{e'} = \begin{pmatrix} a'_1 \\ \vdots \\ a'_n \end{pmatrix}_{e'}$$

We can then compute

$$v = \sum_{j=1}^{n} a'_j \sum_{i=1}^{n} \tau_{ij} e_i = \sum_{i=1}^{n} \sum_{j=1}^{n} a'_j \tau_{ij} e_i \Rightarrow a_i = \sum_{j=1}^{n} a'_j \tau_{ij}$$

and, in matrix form, we have

$$\begin{pmatrix} a'_1 \\ \vdots \\ a'_n \end{pmatrix}_{e'} = Q_{ee'}^{-1} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}_e = Q_{e'e} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}_e$$

7.4 Matrices of Linear Maps in Different Bases

Let V and W be vector spaces over $\mathbb{F}$. Let $e = (e_1, \ldots, e_n)$ and $e' = (e'_1, \ldots, e'_n)$ be bases in V, and let $\epsilon = (\epsilon_1, \ldots, \epsilon_m)$ and $\epsilon' = (\epsilon'_1, \ldots, \epsilon'_m)$ be bases in W. Let $Q_{ee'}$ be a change-of-basis matrix from e to e', and let $Q_{\epsilon\epsilon'}$ be a change-of-basis matrix from ϵ to ϵ', that is,

$$e' = e\, Q_{ee'} \quad \epsilon' = \epsilon\, Q_{\epsilon\epsilon'}$$

We consider a linear map $T : V \mapsto W$ and its matrix $M = M_T^{e\epsilon}$ relative to the bases e, ϵ, as well as its matrix $M' = M_T^{e'\epsilon'}$ relative to the bases e', ϵ'. Since the coordinates of $T\, e_j$ in ϵ are written in the j^{th} column of M, we have that

$$T\, e = (T\, e_1, \ldots, T\, e_n) = \epsilon\, M$$

and similarly

$$T\, e' = \epsilon'\, M'$$

We then have

$$T\, e' = T(e\, Q_{ee'}) = \epsilon'\, M' = (\epsilon\, Q_{\epsilon\epsilon'})M'$$

$$T(e\, Q_{ee'}) = (\epsilon\, Q_{\epsilon\epsilon'})M' = \epsilon(Q_{\epsilon\epsilon'} M')$$

$$\epsilon\, M\, Q_{ee'} = \epsilon\, Q_{\epsilon\epsilon'}\, M'$$

By the linear independence of ϵ, we obtain

$$M' = Q_{\epsilon\epsilon'}^{-1}\, M\, Q_{ee'} = Q_{\epsilon'\epsilon}\, M\, Q_{ee'}$$

If $W = V$ and $\epsilon = e$, $\epsilon' = e'$, then we obtain

$$M^{e'e'} = Q_{ee'}^{-1}\, M_T^{ee}\, Q_{ee'}$$

Definition 7.3 *Square matrices A and B are called similar if there exists a matrix Q, such that Q is invertible and $A = Q^{-1}BQ$.*

Remark *Two matrices are similar if and only if they represent one and the same linear map in different bases.*

Corollary 7.1 *Let A and B be similar matrices. Then* $\det A = \det B$.

8 Eigenvalues and Eigenvectors (Lecture Notes)

8.1 Definitions

Definition 8.1 *A linear operator $T : V \mapsto V$ is called diagonalizable if there exists a basis $v_1, \ldots, v_n \in V$ such that*

$$M_T = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & \ldots & \lambda_n \end{pmatrix}$$

Remark *If M_T is of the form described in Def 8.1, then $T v_i = \lambda_i v_i$, $i = 1, \ldots, n$.*

Definition 8.2 *A number $\lambda \in \mathbb{F}$ is called an eigenvalue of a linear operator T if there exists $v \neq 0$ such that $T v = \lambda v$. The vector v is called an eigenvector of T corresponding to the eigenvalue λ.*

Example *Consider $T : \mathbb{R}^3 \mapsto \mathbb{R}^3 : T(x, y, z) = (x, y, 0)$. It has eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 0$. The corresponding eigenvectors for $\lambda_1 = 1$ are $v_1 = (1, 0, 0)$ and $v_2 = (0, 1, 0)$, as well as any linear combination of the two. The corresponding eigenvector for $\lambda_2 = 0$ is $v_3 = (0, 0, 1)$ and any scalar multiple of v_3.*

Definition 8.3 *The set of all eigenvalues of a linear map $T : V \mapsto V$ is called the spectrum of T and is denoted $\operatorname{Spec} T$.*

Definition 8.4 *The set $V_\lambda = \{v : T v = \lambda v\} = \ker (T - \lambda I)$ is called the eigenspace of the linear map T corresponding to the eigenvalue λ.*

Proposition *The following statements are equivalent:*

1. *λ is an eigenvalue of T.*
2. *$T - \lambda I$ is not injective $(\ker (T - \lambda I) \neq \{0\})$.*
3. *$T - \lambda I$ is not surjective $(\operatorname{rank} (T - \lambda I) \leq n - 1,\ n = \dim V)$.*
4. *$T - \lambda I$ is not invertible.*

8.2 Characteristic Polynomials

Definition 8.5 *The matrix*

$$A - \lambda I = \begin{pmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{pmatrix}$$

is called the characteristic matrix of A and

$$\det(A - \lambda I)$$

is called the characteristic polynomial of A.

Theorem 8.1 *A number λ is an eigenvalue of a linear map T if and only if it is a root of the characteristic polynomial of M_T.*

$$\lambda \in \operatorname{Spec} T \Leftrightarrow \det(M_T - \lambda I) = 0$$

Example *Consider $T : \mathbb{R}^2 \mapsto \mathbb{R}^2 : T(x, y) = (-y, x)$. The matrix of T is*

$$M_T = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Setting its characteristic polynomial to 0:

$$\begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1 = 0$$

we find that it has no roots in $\mathbb{R}$ and therefore T has no eigenvalues or eigenvectors.

Theorem 8.2 *Matrices M_T^e and $M_T^{e'}$ of a linear map T in bases $e_1, \dots, e_n$ and $e'_1, \dots, e'_n$ respectively, have the same characteristic polynomial, i.e.* $\det(M_T^e - \lambda I) = \det(M_T^{e'} - \lambda I)$.

Corollary 8.1 *The characteristic polynomials of similar matrices coincide.*

Theorem 8.3 *If $\lambda_1, \dots, \lambda_n$ are distinct eigenvalues of a linear map T with corresponding eigenvectors $v_1, \dots, v_n$ respectively, then $v_1, \dots, v_n$ are linearly independent.*

8.3 Diagonalization

Theorem 8.4 *A linear map $T : V \mapsto V$ is diagonalizable if and only if there exists a basis $v_1, \ldots, v_n \in V$ consisting of eigenvectors of T. Moreover, if T is diagonalizable, then the matrix M_T, which is the matrix of T in the basis $v_1, \ldots, v_n$, is*

$$M_T = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & \ldots & \lambda_n \end{pmatrix}$$

where λ_i is the eigenvalue of T corresponding to v_i.

Example

$$T : \mathbb{R}^2 \mapsto \mathbb{R}^2 : M_T = A = \begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}_e$$

$$\begin{vmatrix} 1-\lambda & 3 \\ 4 & 2-\lambda \end{vmatrix} = (1-\lambda)(2-\lambda) - 12 = 0 \Rightarrow \lambda_1 = -2,\ \lambda_2 = 5$$

For λ_1 we have:

$$\begin{cases} x + 3y = -2x \\ 4x + 2y = -2y \end{cases} \Rightarrow \begin{cases} 3x + 3y = 0 \\ 4x + 4y = 0 \end{cases} \Rightarrow x = 1,\ y = -1 \Rightarrow v_1 = (1, -1)$$

For λ_2 we have:

$$\begin{cases} x + 3y = 5x \\ 4x + 2y = 5y \end{cases} \Rightarrow \begin{cases} -4x + 3y = 0 \\ 4x + -3y = 0 \end{cases} \Rightarrow x = 3,\ y = 4 \Rightarrow v_2 = (3, 4)$$

The vectors v_1 and v_2 are the eigenvectors of λ_1 and λ_2 respectively. We then have

$$M_T = \begin{pmatrix} -2 & 0 \\ 0 & 5 \end{pmatrix}$$

which is the diagonal matrix of T in the basis v_1, v_2.

9 Inner Products (Lecture Notes)

9.1 Diagonalization of Linear Maps

Theorem 9.1 *If* $\dim V = n$ *and if* $\lambda_1, \ldots, \lambda_n$ *are distinct eigenvalues of the linear map* T *with corresponding eigenvectors* $v_1, \ldots, v_n$*, then* T *is diagonalizable.*

Theorem 9.2 *Given* $T : V \mapsto V$, $\dim V = n$ *and distinct eigenvalues* $\lambda_1, \ldots, \lambda_m$ *of* T *with corresponding eigenspaces* $V_{\lambda_i} = \{v : Tv = \lambda_i v\} = \ker(M_T - \lambda I)$*, the following statements are equivalent:*

1. *T is diagonalizable.*

2. $\det(M_T - \lambda I) = (\lambda_1 - \lambda)^{n_1} \ldots (\lambda_m - \lambda)^{n_m}$ *and* $\dim V_{\lambda_i} = n_i$

3. $\dim V_{\lambda_i} + \cdots + \dim V_{\lambda_m} = n$

9.2 Scalar Products in $\mathbb{R}^3$ or $\mathbb{R}^2$

Definition 9.1 *The scalar or dot product of vectors* u *and* v *is defined as*

$$(u, v) = u \cdot v = |u||v| \cos \varphi$$

where φ *is the angle between* u *and* v*.*

Theorem 9.3 *Given* $u, v, w \in \mathbb{R}^3\,(\mathbb{R}^2)$, $a \in \mathbb{R}$*:*

1. $(u + v) \cdot w = u \cdot w + v \cdot w, \quad au \cdot v = a(u \cdot v)$

2. $u \cdot u = |u|^2 \geq 0$

3. $u \cdot u = 0 \Leftrightarrow u = 0$

4. $u \cdot v = v \cdot u$

5. $u \cdot v = 0 \Leftrightarrow u \perp v$

6. $u \cdot v = (a_1, a_2, a_3) \cdot (b_1, b_2, b_3) = a_1 b_1 + a_2 b_2 + a_3 b_3$

7. $\cos \varphi = \dfrac{u \cdot v}{|u||v|} = \dfrac{a_1 b_1 + a_2 b_2 + a_3 b_3}{\sqrt{(a_1^2 + a_2^2 + a_3^2)(b_1^2 + b_2^2 + b_3^2)}}$

Definition 9.2 *Vectors* $v_1, v_2, v_3 \in \mathbb{R}^3$ *are called an orthonormal basis in* $\mathbb{R}^3$ *if they are orthogonal and have unit length, that is*

$$v_i \cdot v_j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} = \delta_{ij}$$

9.3 Vector Products in $\mathbb{R}^3$

Definition 9.3 *The vector or cross product*

$$w = u \times v = |u||v| \sin \varphi$$

is defined as a vector w that is orthogonal to u and v with a direction given by the right hand rule. The length of w is given by the area of the parallelogram that the vectors u and v span. Note that $u \times v = -v \times u$.

Theorem 9.4 *Given $u, v, w \in \mathbb{R}^3$ and $a \in \mathbb{R}$:*

1. $u \times v = 0 \Leftrightarrow u \parallel v$
2. $(u + v) \times w = u \times w + v \times w$
3. $(au) \times v = a(u \times v)$
4. *If i, j, k form an orthonormal basis and $i \times j = k$, then*

$$u \times v = (a_1 i + a_2 j + a_3 k) \times (b_1 i + b_2 j + b_3 k) = \begin{vmatrix} i & j & k \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

5. $u \cdot (v \times w) = v \cdot (w \times u) = w \cdot (u \times v)$
6. $u \times (v \times w) = v(u \cdot w) - w(u \cdot v)$

9.4 Inner Products

Definition 9.4 *An inner product on a vector space V over $\mathbb{F}$ is a map $\langle \cdot, \cdot \rangle : V \times V \mapsto \mathbb{F}$ such that we have:*

1. *Linearity in the first slot:*

$$\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle \quad \forall\, u, v, w \in V$$

$$\langle au, v \rangle = a\langle u, v \rangle \quad \forall\, u, v \in V, a \in \mathbb{F}$$

2. *Positivity:*

$$\langle u, u \rangle \geq 0 \quad \forall\, u \in V$$

3. *Positive definiteness:*

$$\langle u, u \rangle = 0 \Leftrightarrow u = 0$$

4. *Conjugate symmetry:*

$$\langle u, v \rangle = \overline{\langle v, u \rangle} \quad \forall\, u, v \in V$$

$$(\text{if } \mathbb{F} = \mathbb{R},\ \langle u, v \rangle = \langle v, u \rangle)$$

Remark

$$\langle u, av \rangle = \overline{a} \langle u, v \rangle$$

Definition 9.5 *An inner product space is a vector space over $\mathbb{F}$ together with an inner product $\langle \cdot, \cdot \rangle$.*

Example

1. $V = \mathbb{R}^3$

$$\langle u, v \rangle = (u, v) = u \cdot v = a_1 b_1 + a_2 b_2 + a_3 b_3$$

2. $V = \mathbb{F}^n$

$$\langle u, v \rangle = \sum_{i=1}^{n} a_i \overline{b_i} = a_1 \overline{b_1} + \cdots + a_n \overline{b_n}$$

$$(V = \mathbb{R}^n \Rightarrow \langle u, v \rangle = a_1 b_1 + \cdots + a_n b_n)$$

3. $V = \mathbb{F}[z]$ *or* $V = C([0, 1])$

$$\langle f, g \rangle = \int\limits_0^1 f(z) \overline{g(z)} \, dz$$

Definition 9.6 *A map $\| \cdot \| : V \mapsto [0, \infty)$ is a norm on V if we have:*

1. *Positive homogeneity:*

$$\|av\| = |a| \|v\| \quad \forall \, a \in \mathbb{F}, \, v \in V$$

2. *Positive definiteness:*

$$\|v\| = 0 \Leftrightarrow v = 0$$

3. *Triangle inequality:*

$$\|u + v\| \leq \|u\| + \|v\| \quad \forall \, u, v \in \mathbb{F}$$

Theorem 9.5 (Cauchy-Schwarz Inequality) *Let $\|v\| = \sqrt{\langle v, v \rangle}$. Then for all $u, v \in V$*

$$|\langle u, v \rangle| \leq \|u\| \|v\|$$

Moreover, we have equality if and only if u and v are linearly dependent.

10 Orthonormal Bases (Lecture Notes)

10.1 Inner Product and Norm Maps

Definition 10.1 *A map $\langle \cdot, \cdot \rangle : V \times V \mapsto \mathbb{F}$ satisfying:*

1. $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle \quad \forall\, u, v, w \in V$
   $\langle au, v \rangle = a\langle u, v \rangle \quad \forall\, u, v \in V, a \in \mathbb{F}$
2. $\langle u, u \rangle \geq 0 \quad \forall\, u \in V$
3. $\langle u, u \rangle = 0 \Leftrightarrow u = 0$
4. $\langle u, v \rangle = \overline{\langle v, u \rangle} \quad \forall\, u, v \in V$

is called an inner product on the vector space V over $\mathbb{F}$.

Definition 10.2 *A map $\|\cdot\| : V \mapsto [0, \infty)$ satisfying*

1. $\|av\| = |a|\|v\| \quad \forall\, a \in \mathbb{F},\, v \in V$
2. $\|v\| = 0 \Leftrightarrow v = 0$
3. $\|u + v\| \leq \|u\| + \|v\| \quad \forall\, u, v \in \mathbb{F}$

is called a norm on the vector space V over $\mathbb{F}$.

Example *For $V = \mathbb{C}^n$, $u = (a_1, \ldots, a_n)$, $v = (b_1, \ldots, b_n)$, $w = (d_1, \ldots, d_n)$ we check that $\langle \cdot, \cdot \rangle$ satisfies the properties of the inner product:*

1. $\langle u + v, w \rangle = (a_1 + b_1)\overline{d_1} + \cdots + (a_n + b_n)\overline{d_n} = a_1\overline{d_1} + \cdots + a_n\overline{d_n} + \cdots + b_1\overline{d_1} + \cdots + b_n\overline{d_n}$
   $= \langle u, w \rangle + \langle v, w \rangle$
2. $\langle u, u \rangle = a_1\overline{a_1} + \cdots + a_n\overline{a_n} = |a_1|^2 + \cdots + |a_n|^2 \geq 0$
3. $\langle u, u \rangle = a_1\overline{a_1} + \cdots + a_n\overline{a_n} = |a_1|^2 + \cdots + |a_n|^2 = 0 \Rightarrow u = 0$
4. $\langle u, v \rangle = \overline{\overline{a_1\overline{b_1} + \cdots + a_n\overline{b_n}}} = \overline{\overline{a_1}\overline{\overline{b_1}} + \cdots + \overline{a_n}\overline{\overline{b_n}}} = \overline{b_1\overline{a_1} + \cdots + b_n\overline{a_n}} = \overline{\langle v, u \rangle}$

Definition 10.3 *A vector space V over $\mathbb{F}$ with an inner product $\langle \cdot, \cdot \rangle$ is called an inner product space.*

Theorem 10.1 (Triangle Inequality) *Let $\|u\| := \sqrt{\langle u, u \rangle}$. Then for all $u, v \in V$*

$$\|u + v\| \leq \|u\| + \|v\|$$

Remark *For all $u, v, w \in V$, $a \in \mathbb{F}$*

$$\langle u, v + w \rangle = \overline{\langle v + w, u \rangle} = \overline{\langle v, u \rangle} + \overline{\langle w, u \rangle} = \langle u, v \rangle + \langle u, w \rangle$$

$$\langle u, av \rangle = \overline{a}\langle u, v \rangle$$

Example *For $V = \mathbb{R}^n$, $u = (a_1, \dots, a_n)$, $v = (b_1, \dots, b_n)$, we have:*

$$\langle u, v \rangle = a_1 b_1 + \dots + a_n b_n$$

$$\|u\| = \sqrt{\langle u, u \rangle} = \sqrt{a_1^2 + \dots + a_n^2}$$

Cauchy-Schwarz Inequality:

$$|a_1 b_1 + \dots + a_n b_n| \leq \sqrt{(a_1^2 + \dots + a_n^2)(b_1^2 + \dots + b_n^2)}$$

Triangle Inequality:

$$\sqrt{(a_1 + b_1)^2 + \dots + (a_n + b_n)^2} \leq \sqrt{(a_1^2 + \dots + a_n^2)(b_1^2 + \dots + b_n^2)}$$

Corollary 10.1 *The function $\|u\| = \sqrt{\langle u, u \rangle}$ is a norm on the inner product space V.*

Example *We check that $\|u\| = \sqrt{\langle u, u \rangle}$ is a norm:*

1. $\|av\| = \sqrt{\langle av, av \rangle} = \sqrt{a\bar{a}\langle v, v \rangle} = \sqrt{|a|^2 \langle v, v \rangle} = |a|\|v\|$
2. $\|v\| = \sqrt{\langle v, v \rangle} = 0 \Rightarrow v = 0$
3. $\|u + v\| \leq \|u\| + \|v\|$

10.2 Orthonormal Bases

Definition 10.4 *Vectors u and v are orthogonal if $\langle u, v \rangle = 0$*

Theorem 10.2 (Pythagorean Theorem) *If $u, v \in V$ are orthogonal, then*

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2$$

Definition 10.5 *Vectors $e_1, \dots, e_n \in V$ are called orthogonal if*

$$\langle e_i, e_j \rangle = 0,\ i \neq j$$

They are orthonormal if

$$\langle e_i, e_j \rangle = \delta_{ij}$$

Theorem 10.3 *Every list of non-zero orthogonal vectors in V is linearly independent.*

Definition 10.6 *An orthonormal basis of V is a list of orthonormal vectors that form a basis in V.*

Theorem 10.4 *If $e_1, \dots, e_n$ form an orthonormal basis in V, then*

$$v = \langle v, e_1 \rangle e_1 + \dots + \langle v, e_n \rangle e_n, \quad \forall\, v \in V$$

$$\|v\|^2 = \langle v, v \rangle = |\langle v, e_1 \rangle|^2 + \dots + |\langle v, e_n \rangle|^2$$

10.3 The Gram-Schmidt Orthogonalization Procedure

Theorem 10.5 *If $v_1, \ldots, v_n$ is a list of linearly independent vectors in an inner product space V, then there exist vectors $e_1, \ldots, e_n$ that are orthonormal such that*

$$\operatorname{span}\{v_1, \ldots, v_k\} = \operatorname{span}\{e_1, \ldots, e_k\}, \quad k = 1, \ldots, m$$

Proof: First we set $e_1 = \dfrac{v_1}{\|v_1\|}$. Then $e_2 = \dfrac{v_2 - \langle v_2, e_1\rangle e_1}{\|v_2 - \langle v_2, e_1\rangle e_1\|}$. In general:

$$e_k = \frac{v_k - \langle v_k, e_1\rangle e_1 - \cdots - \langle v_k, e_{k-1}\rangle e_{k-1}}{\|v_k - \langle v_k, e_1\rangle e_1 - \cdots - \langle v_k, e_{k-1}\rangle e_{k-1}\|}$$

□

Example *Take $v_1 = (1, 1, 0)$, $v_2 = (2, 1, 1) \in \mathbb{R}^3$.*

$$e_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{2}}(1, 1, 0)$$

$$\langle v_2, e_1\rangle = \frac{1}{\sqrt{2}}(2, 1, 1) \cdot (1, 1, 0) = \frac{3}{\sqrt{2}}$$

$$e_2 = \frac{v_2 - \langle v_2, e_1\rangle e_1}{\|v_2 - \langle v_2, e_1\rangle e_1\|} = \frac{(2, 1, 1) - \frac{3}{2}(1, 1, 0)}{\|(2, 1, 1) - \frac{3}{2}(1, 1, 0)\|} = \frac{\frac{1}{2}(1, -1, 2)}{\|\frac{1}{2}(1, -1, 2)\|} = \frac{1}{\sqrt{6}}(1, -1, 2)$$

Corollary 10.2 *Every finite-dimensional inner product space has an orthonormal basis.*

Corollary 10.3 *Every orthonormal list of vectors in V can be extended to an orthonormal basis of V.*

11 Orthogonal Projections (Lecture Notes)

Definition 11.1 *Let V be a vector space over $\mathbb{F}$. Let U be a subset, but not necessarily a subspace, of V. The set*

$$U^{\perp} = \{v \in V : \langle v, u\rangle = 0 \quad \forall\, u \in U\}$$

is the orthogonal complement of U.

Lemma 11.1 *If U is a subset of V, then $U^{\perp} = (\operatorname{span} U)^{\perp}$.*

Proposition (from Mathematics 1) *Let U_1 and U_2 be vector subspaces of V. Then $V = U_1 \oplus U_2$ if and only if:*

1. $\forall\, v \in V$, $v = u_1 + u_2$, $u_1 \in U_1$, $u_2 \in U_2$
2. $U_1 \cap U_2 = \{0\}$

Theorem 11.1 *If U is a vector subspace of V, then $V = U \oplus U^{\perp}$.*

Proof: Let $e_1, \dots, e_m$ be an orthonormal basis of U. Then

$$u_1 = \langle v, e_1\rangle e_1 + \cdots + \langle v, e_m\rangle e_m \in U$$

Checking that $u_2 = v - u_1 \in U^{\perp}$:

$$\langle u_2, e_1\rangle = \langle v, e_1\rangle - \langle v, e_1\rangle\langle e_1, e_1\rangle - \cdots - \langle v, e_m\rangle\langle e_m, e_1\rangle = 0 \Rightarrow \langle u_2, e_j\rangle = 0$$

We have

$$u_2 \in \{e_1, \dots, e_m\}^{\perp} = \operatorname{span}\{e_1, \dots, e_m\}^{\perp} = U^{\perp}$$

Thus the first condition $v = u_1 + u_2$ is fulfilled. Now let $u \in U \cap U^{\perp}$. Then

$$\langle u \in U, u \in U^{\perp}\rangle = 0 = \|u\|^2 \Rightarrow u = 0$$

Thus the second condition $U \cap U^{\perp} = \{0\}$ is fulfilled and $V = U \oplus U^{\perp}$. $\square$

Theorem 11.2 *If U is a subset of V, then $(U^{\perp})^{\perp} = \operatorname{span} U$. In particular, if U is a subspace of V, then $(U^{\perp})^{\perp} = U$.*

Definition 11.2 *The map $P_U : V \mapsto V$ defined as $P_U(v) = u_1$, where $u_1 \in U$ is such that $v = u_1 + u_2$, $u_2 \in U^{\perp}$, is called an orthogonal projection. If $e_1, \dots, e_m$ form an orthonormal basis in U, then*

$$P_U(v) = \langle v, e_1\rangle e_1 + \cdots + \langle v, e_m\rangle e_m$$

Theorem 11.3 *Let U be a subspace of V and $v \in V$. Then*

$$\|v - P_U(v)\| \leq \|v - u\| \quad \forall\, u \in U$$

Moreover, the equality holds if and only if $u = P_U(v)$.

12 Adjoint Operators (Lecture Notes)

12.1 Dual Space

Let V be a finite-dimensional inner product space over $\mathbb{F}$. Take $u \in V$ and consider the map

$$f_u(v) = \langle v, u \rangle \quad \forall\, v \in V$$

f_u is a linear map from V to $\mathbb{F}$ and is called a functional on V.

Theorem 12.1 (Riesz Representation Theorem) *Given a finite-dimensional inner product space V and $f \in \mathcal{L}(V, \mathbb{R})$, there exists a unique $u \in V$ such that*

$$f(v) = \langle v, u \rangle \quad \forall\, v \in V$$

Proof: Given an orthonormal basis $e_1, \ldots, e_n$, any vector can be written as

$$v = \langle v, e_1 \rangle e_1 + \cdots + \langle v, e_n \rangle e_n$$

$$\begin{aligned} f(v) = f\big(\langle v, e_1 \rangle e_1 + \cdots + \langle v, e_n \rangle e_n\big) &= \langle v, e_1 \rangle f(e_1) + \cdots + \langle v, e_n \rangle f(e_n) \\ &= \big\langle v, \overline{f(e_1)} e_1 + \cdots + \overline{f(e_n)} e_n \big\rangle = \langle v, u \rangle \end{aligned}$$

We find that there exists $u \in V$ such that $f(v) = \langle v, u \rangle$. Additionally, taking $v = u_1 - u_2$:

$$f(v) = \langle v, u_1 \rangle = \langle v, u_2 \rangle$$

$$0 = f(v) - f(v) = \langle v, u_1 - u_2 \rangle = \langle u_1 - u_2, u_1 - u_2 \rangle = \|u_1 - u_2\|^2 \Rightarrow u_1 = u_2$$

we find that u is unique. $\square$

Definition 12.1 *The set $V^* = \mathcal{L}(V, \mathbb{F})$ is the set of all linear functionals on V and is called the dual space of V.*

Remark

$$\langle v, u_1 \rangle = \langle v, u_2 \rangle \quad \forall\, v \in V \Rightarrow u_1 = u_2$$

Remark *By Th. 12.1, V^* only contains functionals of the form $f(v) = \langle v, u \rangle \quad \forall\, v \in V$.*

12.2 Adjoint Operators

Theorem 12.2 *For $T \in \mathcal{L}(V)$ there exists a unique linear map $T^* \in \mathcal{L}(V)$ such that*

$$\langle T\, v, u \rangle = \langle v, T^*\, u \rangle \quad u, v \in V$$

Proof: Consider $f^u(v) = \langle T\, v, u \rangle$. Then there exists $u' \in V$ such that

$$f^u(v) = \langle T\, v, u \rangle = \langle v, u' \rangle$$

Simply take $T^*\, u = u'$, then

$$\langle T\, v, u \rangle = \langle v, T^*\, u \rangle \quad \forall\, v, u \in V$$

and T^* is a unique map from V to V. Now take $u_1, u_2 \in V$.

$$\langle v, T^*(u_1 + u_2) \rangle = \langle T\, v, u_1 + u_2 \rangle = \langle T\, v, u_1 \rangle + \langle T\, v, u_2 \rangle = \langle v, T^*\, u_1 \rangle + \langle v, T^*\, u_2 \rangle$$

$$= \langle v, T^*\, u_1 + T^*\, u_2 \rangle \quad \forall\, v \in V$$

$$\Rightarrow T^*(u_1 + u_2) = T^*\, u_1 + T^*\, u_2$$

The same can be shown for $T^*(au) = a\, T^*\, u$, $a \in \mathbb{F}$. Thus T^* is linear. □

Definition 12.2 *The operator T^* is called the adjoint of T.*

Theorem 12.3 *Let $e_1, \ldots, e_n$ form an orthonormal basis in V and $T \in \mathcal{L}(V)$. Then the entries of the matrix M_T of T in the basis $e_1, \ldots, e_n$ are given by*

$$a_{ij} = \langle T\, e_j, e_i \rangle$$

Theorem 12.4 *The matrix M_{T^*} in any orthonormal basis of V is the complex conjugate, transposed matrix of the matrix M_T of $T \in \mathcal{L}(V)$.*

Proof:

$$M_{T^*} = (b_{ij})$$

$$b_{ij} = \langle T^*\, e_j, e_i \rangle = \overline{\langle e_i, T^*\, e_j \rangle} = \overline{\langle T\, e_i, e_j \rangle} = \overline{a_{ji}}$$

□

Theorem 12.5 *For $T, S \in \mathcal{L}(V)$, $a \in \mathbb{F}$:*

1. $(T + S)^* = T^* + S^*$
2. $(a\, T)^* = a\, T^*$
3. $(TS)^* = S^* T^*$
4. $(T^*)^* = T$

12.3 Self-Adjoint Operators

Definition 12.3 *If $T \in \mathcal{L}(V)$ satisfies $T = T^*$, then T is called self-adjoint. If a matrix A satisfies $A = A^*$, then A is self-adjoint.*

Remark *T is self-adjoint if its matrix M_T, in any orthonormal basis, is adjoint.*

Remark *If each entry of A is real and A is self-adjoint, then A is symmetric.*

Theorem 12.6 *Let $T, S \in \mathcal{L}(V)$ be self-adjoint and $a \in \mathbb{F}$. Then T^2, $T + S$, $a\,T$, $TS + ST$ are also self-adjoint.*

Theorem 12.7 *If $T \in \mathcal{L}(V)$ is self-adjoint, then each eigenvalue of T is real and eigenvectors corresponding to distinct eigenvalues are orthogonal.*

Proof:

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle T\,v, v\rangle = \langle v, T^*\,v\rangle = \langle v, T\,v\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\langle v, v\rangle$$

$$\Rightarrow \lambda = \overline{\lambda} \Rightarrow \lambda \in \mathbb{R}$$

$$\lambda_1\langle v_1, v_2\rangle = \langle T\,v_1, v_2\rangle = \langle v_1, T\,v_2\rangle = \lambda_2\langle v_1, v_2\rangle$$

$$\Rightarrow \langle v_1, v_2\rangle = 0 \Rightarrow v_1 \perp v_2$$

□

Theorem 12.8 *Given that $T \in \mathcal{L}(V)$ is self-adjoint:*

1. *There exists an orthogonal basis of V, each vector of which is an eigenvector of T corresponding to a real eigenvalue of T.*
2. *There exists an orthonormal basis in which M_T has a diagonal form, where each entry is real.*

13 Unitary Operators (Lecture Notes)

13.1 Definitions

Let V and W be inner product vectors spaces over the same field $\mathbb{F}$.

Definition 13.1 *$T \in \mathcal{L}(V,W)$ preserves inner products if*

$$\langle T\,v, T\,u\rangle_W = \langle v, u\rangle_V \quad \forall\, u, v \in V$$

T is then called an isomorphism of V onto W.

Theorem 13.1 *Let $\dim V = \dim W$ be finite. Given $T \in \mathcal{L}(V,W)$, the following statements are equivalent:*

1. *T preserves inner products.*
2. *T is an inner product vector space isomorphism, i.e. it is invertible and preserves inner products.*
3. *T maps some (then every) orthonormal basis in V to an orthonormal basis in W.*

Corollary 13.1 *Inner product spaces V and W are isomorphic if and only if $\dim V = \dim W$.*

Theorem 13.2 *$T \in \mathcal{L}(V,W)$ preserves inner products if and only if*

$$\|T\,v\| = \|v\| \quad \forall\, v \in V$$

Definition 13.2 *If the operator $T \in \mathcal{L}(V)$ preserves inner products, then T is called a unitary operator.*

Theorem 13.3 *$T \in \mathcal{L}(V)$ is unitary if and only if*

$$T^*T = TT^* = I \Leftrightarrow T^* = T^{-1}$$

Definition 13.3 *A matrix $A \in \mathbb{F}^{n\times n}$ is called unitary if*

$$A^*A = AA^* = I \Leftrightarrow A^* = A^{-1}$$

If $A \in \mathbb{R}^{n\times n}$ is unitary, i.e. $A^TA = I$, then A is called an orthogonal matrix.

Theorem 13.4 *$T \in \mathcal{L}(V)$ is unitary if and only if its matrix in some (then every) orthonormal basis is unitary.*

Theorem 13.5

1. *Given a self-adjoint matrix $A \in \mathbb{F}^{n\times n}$, there exists a unique matrix P such that $P^{-1}AP$ is diagonal.*
2. *If A is a real symmetric matrix, then there exists a real orthogonal matrix P such that $P^{-1}AP$ is diagonal.*

Theorem 13.6 *$A \in \mathbb{F}^{n\times n}$ is unitary if and only if its rows (columns) form an orthonormal basis in $\mathbb{F}^n$.*

13.2 Normal Operators

Definition 13.4 *A linear map $T \in \mathcal{L}(V)$ is called normal if $TT^* = T^*T$. A matrix A is called normal if $AA^* = A^*A$.*

Theorem 13.7 (Spectral Theorem) *If V is a finite-dimensional inner product space over $\mathbb{C}$ and $T \in \mathcal{L}(V)$, then T is normal if and only if there exists an orthonormal basis in V consisting of eigenvectors of T.*

Corollary 13.2 *Let $T \in \mathcal{L}(V)$ be a normal operator and let V be a finite-dimensional inner product space over $\mathbb{C}$. Let $\lambda_1, \ldots, \lambda_m$ be distinct eigenvalues of T. Then*

1. $V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_m}$, $V_{\lambda_i} = \{v : Tv = \lambda_i v\} = \ker(T - \lambda_i I)$
2. $i \neq j \Rightarrow V_{\lambda_i} \perp V_{\lambda_j}$ *i.e.* $\forall\, v \in V_{\lambda_i},\, u \in V_{\lambda_j},\, \langle v, u\rangle = 0$

Remark *Th. 13.7 tells us that $T \in \mathcal{L}(V)$ is normal if and only if M_T is diagonal with respect to an orthonormal basis $e_1, \ldots, e_n \in V$, i.e. there exists a unitary matrix U such that*

$$UM_T U^* = UM_T U^{-1} = \begin{pmatrix} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n \end{pmatrix}$$

Remark *Diagonal decomposition allows us to easily compute powers and functions of matrices. Let*

$$A = UDU^{-1} \quad D = \begin{pmatrix} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n \end{pmatrix}$$

Then

$$A^n = (UDU^{-1})^n = UD^nU^{-1} = U \begin{pmatrix} \lambda_1^n & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n^n \end{pmatrix} U^{-1}$$

Thus we can define

$$f(A) = U \begin{pmatrix} f(\lambda_1) & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & f(\lambda_n) \end{pmatrix} U^{-1}$$

Take the example of calculating e^A:

$$e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k = U \left(\sum_{k=0}^{\infty} \frac{1}{k!} D^k \right) U^{-1}$$

$$= U \begin{pmatrix} \sum_{k=0}^{\infty} \frac{1}{k!}\lambda_1^k & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \sum_{k=0}^{\infty} \frac{1}{k!}\lambda_n^k \end{pmatrix} U^{-1} = U \begin{pmatrix} e^{\lambda_1^n} & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & e^{\lambda_n^n} \end{pmatrix} U^{-1}$$

14 Bilinear Forms (Lecture Notes)

14.1 Definitions

Definition 14.1 *A function $B : V \times V \mapsto \mathbb{F}$, where V is a vector space over $\mathbb{F}$, is called a bilinear form if*

1. $B(au + bv, w) = aB(u, w) + bB(v, w)$
2. $B(w, au + bv) = aB(w, u) + bB(w, v)$

Example

1. *For a vector space V over $\mathbb{R}$, the inner product is a bilinear form: $B(u, v) = \langle u, v\rangle$.*
2. *For vectors $u = (x_1, \ldots, x_n)$ and $v = (y_1, \ldots, y_n)$ in $\mathbb{F}$, a bilinear form can be defined as the following:*

$$B(u, v) = \sum_{i,j=1}^{n} a_{ij}\, x_i\, y_j$$

   If $n = 2$, this is, for example

$$B(u, v) = \sum_{i,j=1}^{2} a_{ij}\, x_i\, y_j = x_1y_1 + 2x_2y_1 + 3x_1y_2 + 7x_2y_2$$

   where a_{ij} are arbitrary scalars.

3. *For functions $f, g \in C([0, 1])$, we can define the bilinear form as*

$$B(f, g) = \int_0^1 k(x)f(x)g(x)\, dx \quad k(x) \in \mathrm{C}([0, 1])$$

Definition 14.2 *The matrix $A = (a_{ij})_{i,j=1}^{n}$, where $a_{ij} = B(e_i, e_j)$, is called the matrix of B in the basis $e_1, \ldots, e_n$, also called the Gram matrix of B.*

$$B(v, u) = \begin{pmatrix} x_1 & \ldots & x_n \end{pmatrix} \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = M_v^T A M_u$$

14.2 Change of Basis

When changing bases, a bilinear form B can be written as follows:

$$B(v,u) = \left(M_v^e\right)^T A^e\, M_u^e = \left(QM_v^{e'}\right)^T A^e\left(QM_u^{e'}\right) = \left(M_v^{e'}\right)^T\left(Q^T A^e Q\right) M_u^{e'}$$

where $Q = Q_{ee'}$ is the change-of-basis matrix from the basis e to the basis e'.

Theorem 14.1 *Let A^e be the matrix of B in the basis e. Then*

$$A^{e'} = Q_{ee'}^T A^e Q_{ee'} = Q^T A^e Q$$

where Q is the change-of-basis matrix from e to e', is the matrix of B in the basis e'.

Definition 14.3 *The rank of the matrix of a bilinear form B is called the rank of B and is denoted* $\operatorname{rank} B$.

Theorem 14.2 *If* $\dim V = n$*, then the following conditions are equivalent:*

1. $\operatorname{rank} B = n$
2. $\forall\, v \neq 0 \quad \exists\, u : B(v,u) \neq 0$
3. $\forall\, u \neq 0 \quad \exists\, v : B(v,u) \neq 0$

Definition 14.4 *If B satisfies one of the conditions in Th. 14.2, then it is non-degenerate.*

Theorem 14.3 *Let B be a symmetric bilinear form in a real vector space V. Then there exists a basis $e_1, \ldots, e_n \in V$ in which the matrix of B is diagonal with only 1's, -1's, and 0's on the diagonal, that is*

$$\exists\, e_1, \ldots, e_n \in V,\ s \leq r \leq n : a_{ij} = B(e_i, e_j) = \begin{cases} 1, & i = j \leq s \\ -1, & s < i = j \leq r \\ 0, & i = j > r \\ 0, & i \neq j \end{cases}$$

In this case, $B(v,u) = x_1y_1 + \cdots + x_sy_s - x_{s+1}\,y_{s+1} - \cdots - x_ry_r$.

Theorem 14.4 (Sylvester's Law of Inertia) *The number of 1's, -1's, and 0's in Th. 14.3 is independent of the basis in which the matrix of B is written.*

14.3 Quadratic Forms

Let B be a symmetric bilinear form in the vector space V.

Definition 14.5 *The map from V to $\mathbb{F}$ given by $v \mapsto B(v,v)$ is called the quadratic form associated with B.*

Remark *Any quadratic form in a real vector space uniquely determines the associated bilinear form:*

$$B(v,u) = \frac{1}{2}\big[B(v+u, v+u) - B(v,v) - B(u,u)\big]$$

Theorem 14.5 *For every quadratic form $B(v,v) \in \mathbb{R}^n$, there exists a basis such that*

$$B(v,v) = x_1^2 + \cdots + x_s^2 - x_{s+1}^2 - \cdots - x_r^2$$

Example *Consider the quadratic form $B(v,v) = x_1^2 + 2x_1x_2 - 4x_1x_3 + 4x_3^2 - 6x_2x_3$. We can rewrite it as*

$$B(v,v) = (x_1^2 + 2x_1x_2 - 2x_12x_3 - 2x_22x_3 + x_2^2 + 4x_3^2) + 2x_22x_3 - x_2^2 - 6x_2x_3$$

$$= (x_1 + x_2 - 2x_3)^2 - 2x_2x_3 - x_2^2$$

$$= (x_1 + x_2 - 2x_3)^2 - (x_2^2 + 2x_2x_3 + x_3^2) + x_3^2$$

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

We can then simply assign a new basis:

$$z_1 = x_1 + x_2 - 2x_3 \qquad z_2 = x_3 \qquad z_3 = x_2 + x_3$$

such that $B(v,v) = z_1^2 + z_2^2 - z_3^2$.

Definition 14.6 *A quadratic form B is positive-definite if $B(v,v) > 0 \quad \forall\, v \neq 0$. It is negative-definite if $B(v,v) < 0 \quad \forall\, v \neq 0$.*

Theorem 14.6 *A quadratic form B is positive-definite if and only if $M_1 > 0, \ldots, M_n > 0$. It is negative-definite if and only if $M_1 < 0$, $M_2 > 0$, $M_3 < 0$, $M_4 > 0 \ldots$. The numbers M_i are the principal minors of the matrix of B given by*

$$M_1 = a_{11}, \quad M_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad M_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \quad \ldots$$

15 Topology in $\mathbb{R}^d$ (Lecture Notes)

15.1 Norms in $\mathbb{R}^d$

The distance between $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ is given by

$$\|x - y\| = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}$$

Properties of norms and distance:

1. $\|x\| \geq 0$
2. $\|x\| = 0 \Leftrightarrow x = (0, \ldots, 0)$
3. $\|ax\| = |a|\|x\|$
4. $\|x + y\| \leq \|x\| + \|y\|$
5. $\big|\|x\| - \|y\|\big| \leq \|x - y\|$
6. $\|x - y\| \geq 0$
7. $\|x - y\| = 0 \Leftrightarrow x = y$
8. $\|x - y\| = \|y - x\|$
9. $\|x - y\| \leq \|x - z\| + \|z - y\|$
10. $\big|\|x - z\| - \|y - z\|\big| \leq \|x - y\|$

15.2 Limits in $\mathbb{R}^d$

Definition 15.1 *A sequence $\left(x^{(n)}\right)_{n\geq 1}$ of elements in $\mathbb{R}^d$ converges if there exists $x \in \mathbb{R}^d$ such that*

$$\|x^{(n)} - x\| \to 0,\ n \to \infty$$

That is

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} : \forall n \geq N \quad \|x^{(n)} - x\| < \varepsilon$$

Theorem 15.1 *If $x^{(n)} \to x$, $n \to \infty$ and $x^{(n)} \to y$, $n \to \infty$, then $x = y$.*

Theorem 15.2 *If $x^{(n)} \to x$, $n \to \infty$, then*

$$\forall y \in \mathbb{R}^d \quad \left\|x^{(n)} - y\right\| \to \|x - y\|,\ n \to \infty$$

15.3 Limit Points in $\mathbb{R}^d$

Definition 15.2 *The set*

$$B_r(x) = \{y \in \mathbb{R}^d : \|x - y\| < r\}$$

is called an open ball of radius $r > 0$ and center $x \in \mathbb{R}^d$.

Definition 15.3 *The set*

$$\overline{B_r}(x) = \{y \in \mathbb{R}^d : \|x - y\| \leq r\}$$

is called a closed ball of radius $r > 0$ and center $x \in \mathbb{R}^d$.

Definition 15.4 *A set $A \subseteq \mathbb{R}^d$ is bounded if $\exists\, r > 0 : A \subseteq B_r(0)$.*

Definition 15.5 *A point $x_0 \in \mathbb{R}^d$ is a limit point of $A \subseteq \mathbb{R}^d$ if*

$$\forall r > 0 \quad \exists x \in A,\ x \neq x_0 : x \in B_r(x_0)$$

Theorem 15.3 *A point $x_0 \in \mathbb{R}^d$ is a limit point of $A \subseteq \mathbb{R}^d$ if and only if there exists a sequence $\left(x^{(n)}\right)_{n\geq 1}$ such that*

1. $x_0 \neq x^{(n)} \in A \quad \forall n$
2. $x^{(n)} \to x_0$, $n \to \infty$

15.4 Open Sets

Definition 15.6 *A point $x_0 \in A$ is called an inner point of A if $\exists\, r > 0 : B_r(x_0) \subseteq A$.*

Definition 15.7 *A set $A \subseteq \mathbb{R}^d$ is open if each point of A is an inner point of A, that is*

$$\forall\, x \in A \quad \exists\, r > 0 : B_r(x) \subseteq A$$

Example *The set $A = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 > 0\}$ is open. If $x = (x_1, x_2) \in A$, we can take $r = x_1 > 0$. Then $B_r(x) \subseteq A$.*

Remark *The set $A = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \geq 0\}$ is not open, since the points $(0, x_2)$, where $x_2 \in \mathbb{R}$, are not inner points of A.*

Theorem 15.4 *The union of any number of open sets is open.*

Theorem 15.5 *The intersection of a finite number of open sets is open.*

Remark *The intersection of any number of open sets is not open in general.*

Example *Consider the open set $A_r = B_r(x)$. Then consider*

$$\bigcap_{r>0} A_r = \{x : x \in A_r = B_r(x) \quad \forall\, r > 0\} = \{x\}$$

Note that the set $\{x\}$ is not open.

15.5 Closed Sets

Definition 15.8 *A set A is closed if it contains all its limit points.*

Theorem 15.6 *A set $A \subseteq \mathbb{R}^d$ is closed if and only if the set*

$$\mathbb{R}^d \setminus A = \{x \in \mathbb{R}^d : x \notin A\}$$

is open.

Theorem 15.7 *The intersection of any number of closed sets is a closed set, and the union of a finite number of closed sets is also a closed set.*

Definition 15.9 *The set $\bar{A}$, which consists of all points of A and all limit points of A, is called the closure of A.*

Example

$$A = \{x \in \mathbb{R}^2 : x_1 > 0\} \Rightarrow \bar{A} = \{x \in \mathbb{R}^2 : x_1 \geq 0\}$$

$$\overline{B_r(x)} = \{y : \|x - y\| \leq r\} = \overline{B_r}(x)$$

16 Functions of Several Variables (Lecture Notes/Slides)

16.1 Compact Sets in $\mathbb{R}^d$

Definition 16.1 *An open cover of a set $K \subseteq \mathbb{R}^d$ is a collection G_α, $\alpha \in T$ of open subsets of $\mathbb{R}$ such that $K \subseteq \bigcup_{\alpha \in T} G_\alpha$.*

Definition 16.2 *A subset $K \subseteq \mathbb{R}^d$ is called a compact set if every open cover G_α, $\alpha \in T$ of K contains a finite subcover, that is*

$$\exists\, \alpha_1, \ldots, \alpha_n : K \subseteq \bigcup_{i=1}^{n} G_{\alpha_i}$$

Theorem 16.1 *If $K \subseteq \mathbb{R}^d$, then the following statements are equivalent:*

1. *K is compact.*
2. *K is closed and bounded.*
3. *If $\left(x^{(n)}\right)_{n \geq 1}$ is a sequence of elements from K, then there always exists a subsequence $\left(x^{(n_k)}\right)_{n \geq 1}$ such that $x^{n_k} \to x_0$ and $x_0 \in K$.*

16.2 Examples of Functions of Several Variables

1. Real-valued functions of one variable $f : D \mapsto \mathbb{R}$, $D \subseteq \mathbb{R}$:
   (a) $f(x) = 2x$
   (b) $f(x) = \sin x$
   (c) $f(x) = \sqrt{1 - x^2}$, $x \in [-1, 1]$
2. Real-valued functions of several variables $f : D \mapsto \mathbb{R}$, $D \subseteq \mathbb{R}^d$:
   (a) $f(x, y) = 3x + 2y$
   (b) $f(x, y) = x^2 + y^2$
   (c) $f(x, y) = \sin x \sin y$

   The set $D_a = \{x \in D : f(x) = a\}$ is called a level set of f.
3. Vector-valued functions of several variables $f : D \mapsto \mathbb{R}^m$, $D \subseteq \mathbb{R}^d$:
   (a) $f(x, y) = (\cos x \sin y, \sin x \cos y) = \nabla \sin x \sin y$
4. Vector-valued functions of one variable $f : D \mapsto \mathbb{R}^d$, $D \subseteq \mathbb{R}$:
   (a) $f(t) = (1 + 2t, t, 3 - t)$
   (b) $f(t) = (\cos t, \sin t)$

16.3 Limits of Functions

Definition 16.3 *Given $f : D \mapsto \mathbb{R}^m$, $D \subseteq \mathbb{R}^d$ and a limit point x_0 of D, the point $P \in \mathbb{R}^m$ is called the limit of f at the point x_0 if*

$$\forall \left(x^{(n)}\right)_{n\geq 1} \in D : x^{(n)} \neq x_0,\, x^{(n)} \to x_0$$

one has $f\left(x^{(n)}\right) \to P$. We write $p = \lim\limits_{x\to x_0} f(x)$.

Theorem 16.2 *Given $f : D \mapsto \mathbb{R}^m$ and a limit point x_0 of D, we have $p = \lim\limits_{x\to x_0} f(x)$ if and only if*

$$\forall\, \varepsilon > 0 \quad \exists\, \delta > 0 : \forall\, x \in D,\, x \neq x_0,\, \|x - x_0\| < \delta \Rightarrow \|f(x) - p\| < \varepsilon$$

Remark *If $f : D \mapsto \mathbb{R}^m$ and $f = (f_1, \ldots, f_m)$, then $p = \lim\limits_{x\to x_0} f(x)$ if and only if*

$$\forall\, i \quad p_i = \lim_{x\to x_0} f_i(x), \quad p = (p_1, \ldots, p_m)$$

Example

1. For $f(x,y) = \dfrac{x^2 y}{x^2+y^2}$, $\quad (x,y) \in \mathbb{R} \setminus \{0\}$ we have

$$0 \leq \left|\frac{x^2 y}{x^2+y^2}\right| = \frac{x^2|y|}{x^2+y^2} \leq \frac{(x^2+y^2)|y|}{x^2+y^2} = |y| \Rightarrow \lim_{(x,y)\to(0,0)} \frac{x^2 y}{x^2+y^2} = 0$$

2. For $f(x,y) = \dfrac{xy}{x^2+y^2}$, $\quad (x,y) \in \mathbb{R} \setminus \{0\}$, consider $x = y$, then $x = -y$:

$$f(x,x) = \frac{x^2}{x^2+x^2} = \frac{1}{2} \qquad f(x,-x) = \frac{-x^2}{x^2+x^2} = -\frac{1}{2}$$

This implies that $\lim\limits_{(x,y)\to(0,0)} \dfrac{xy}{x^2+y^2}$ does not exist.

Theorem 16.3 *Given $f : D \mapsto \mathbb{R}^m$, $D \subseteq \mathbb{R}^d$ and a limit point x_0 of D, then $p = \lim\limits_{x\to x_0} f(x)$ if and only if for any map $\alpha : (0,\varepsilon) \mapsto D$ such that*

1. $\lim\limits_{t\to 0} \alpha(t) = x_0$

2. $\alpha(t) \neq x_0 \quad \forall\, t \in (0,\varepsilon)$

one has $\lim\limits_{t\to 0} f(\alpha(t)) = p$.

Example $\lim_{(x,y)\to(0,0)} \dfrac{x^2 y}{x^4+y^2}$ *exists along every line, but does not exist in general. Consider the following two cases:*

$$x := at,\, y := bt \Rightarrow f(x,y) = \frac{a^2t^2bt}{a^4t^4b^2t^2} = \frac{a^2bt}{a^4t^2+b^2} \Rightarrow f(x,y) = f(at,bt) \to 0,\, t \to 0$$

$$x := t,\, y := t^2 \Rightarrow f(x,y) = \frac{t^4}{t^4+t^4} = \frac{1}{2} \Rightarrow f(x,y) = f(t,t^2) \to \frac{1}{2},\, t \to 0$$

This implies that the limit does not exist.

17 Continuous Functions (Lecture Notes)

17.1 Definitions and Basic Properties

Definition 17.1 *A function $f : D \mapsto \mathbb{R}^m$, $D \subseteq \mathbb{R}^d$ is continuous at a limit point $x_0 \in D$ if $\lim\limits_{x \to x_0} f(x) = f(x_0)$. This definition is equivalent to the following:*

$$\forall\, \varepsilon > 0 \quad \exists\, \delta > 0 : \forall\, x \in D \quad \|x - x_0\| < \delta \Rightarrow \|f(x) - f(x_0)\| < \varepsilon$$

Definition 17.2 *A point $x_0 \in D$ is called an isolated point of D if there exists $r > 0$ such that $B_r(x_0) \cap D = \{x_0\}$. We assume that any function is continuous at any isolated point.*

Definition 17.3 *A function $f : D \mapsto \mathbb{R}^m$ is called continuous on D if it is continuous at each point of D. In this case $f \in \mathrm{C}(D, \mathbb{R}^m)$. Additionally, if $m = 1$, we write $\mathrm{C}(D, \mathbb{R}) := \mathrm{C}(D)$*

Remark *If x_0 is an inner point of D, then f is continuous at x_0 if and only if*

$$\forall\, \varepsilon > 0 \quad \exists\, \delta > 0 : f(B_\delta(x_0)) \subseteq B_\varepsilon(f(x_0))$$

Theorem 17.1 *For functions $f : D \mapsto \mathbb{R}$, $g : D \mapsto \mathbb{R}$ that are continuous at x_0, the following are also continuous at x_0:*

1. $cf \quad \forall\, c \in \mathbb{R}$
2. $f + g$
3. fg
4. $\dfrac{f}{g}$, $g(x_0) \neq 0$

Theorem 17.2 *A function $f = (f_1, \ldots, f_m) : D \mapsto \mathbb{R}^m$ is continuous at x_0 if and only if $f_i : D \mapsto \mathbb{R}$ is continuous at x_0 for each $i = 1, \ldots, m$.*

Theorem 17.3 *If a function $f : D \mapsto M$, $D \subseteq \mathbb{R}^d$, $M \subseteq \mathbb{R}^m$ is continuous at $x_0 \in D$ and $g : M \mapsto \mathbb{R}^n$ is continuous at $y_0 = f(x_0) \in M$, then $h(x) = g(f(x)) = (g \circ f)(x)$ is continuous at x_0.*

17.2 Examples of Continuous Functions

1. Constant Functions
   $f(x) = c, \quad x \in \mathbb{R}^d$, $c \in \mathbb{R}$

2. Coordinate Functions
   $\pi_k : \mathbb{R}^d \mapsto R, \quad k = 1, \ldots, d$
   $\pi_k(x_1, \ldots, x_d) = x_k, \quad x = (x_1, \ldots, x_d) \in \mathbb{R}^d$

   Proof of continuity: Let $x_0 = (x_1^0, \ldots, x_d^0)$. Let $\varepsilon > 0$ be given. We have

   $$|\pi_k(x) - \pi_k(x_0)| = |x_k - x_k^0|$$

   $$= \sqrt{(x_k - x_k^0)^2} \leq \sqrt{(x_1 - x_1^0)^2 + \cdots + (x_d - x_d^0)^2} = \|x - x_0\|$$

   Taking $\delta = \varepsilon$, we have $\|x - x_0\| < \delta \Rightarrow |\pi_k(x) - \pi_k(x_0)| < \varepsilon$. $\square$

3. Polynomials
$P(x_1, \ldots, x_d) = \sum_{k_1=0}^{n_1} \cdots \sum_{k_d=0}^{n_d} a_{k_1 \ldots k_d} \, x_1^{k_1} \ldots x_d^{k_d}, \quad x \in \mathbb{R}^d$

4. Rational Functions
$R(x) = \frac{P(x)}{Q(x)}, \quad x \in D$
for polynomials P, Q on $\mathbb{R}^d$ and $D = \{x \in \mathbb{R}^d : Q(x) \neq 0\}$

5. Other Functions
$f(x_1, x_2) = e^{\sqrt{x_1^2 + x_2^2}} \sin(x_1 + x_1 x_2^2), \quad (x_1, x_2) \in \mathbb{R}^2$

17.3 Characterization of Continuous Functions

Given $f : D \mapsto \mathbb{R}^m$ and $B \subseteq \mathbb{R}^m$, we set the preimage of B:

$$f^{-1}(B) = \{x \in D : f(x) \in B\}$$

If $A \subseteq D$, then we say that A is open in D if

$$\forall \, x \in A \quad \exists \, r > 0 : D \cap B_r(x_0) \subseteq A$$

Remark *If D is open, then A is open in D if and only if A is open in $\mathbb{R}^d$.*

Theorem 17.4 *If $f : D \mapsto \mathbb{R}^m$, $D \subseteq \mathbb{R}^d$, then f is continuous on D if and only if, for any open set $G \in \mathbb{R}^m$, the set $f^{-1}(G) = \{x \in D : f(x) \in G\}$ is open in D.*

17.4 Continuous Functions on Compact Sets

Theorem 17.5 *Given a compact set D in $\mathbb{R}^d$ and $f \in \mathrm{C}(D, \mathbb{R}^m)$, the set*

$$f(D) = \{f(x) : x \in D\}$$

is compact in $\mathbb{R}^m$.

Theorem 17.6 *Given a compact set K in $\mathbb{R}^d$ and a continuous function $f : K \mapsto \mathbb{R}$:*

1. *f is bounded on K, i.e. $\exists \, C > 0 : |f(x)| \leq C \quad \forall \, x \in K$*

2. *$\exists \, x_*, x^* \in K : f(x_*) = \min_{x \in K} f(x)$, $f(x^*) = \max_{x \in K} f(x)$*

Theorem 17.7 *If K is a compact set and $f : K \mapsto \mathbb{R}^m$ is continuous, then f is uniformly continuous, that is,*

$$\forall \, \varepsilon > 0 \quad \exists \, \delta < 0 : \forall \, x', x'' \in K \quad \|x' - x''\| < \delta \Rightarrow \|f(x') - f(x'')\| < \varepsilon$$

18 Differentiation of Functions of Several Variables (Part I)

(**Lecture Notes**/**Slides**)

18.1 Functions of One Variable

Consider the function $f : (a, b) \mapsto \mathbb{R}$, which is differentiable at $x_0 \in (a, b)$. Then there exists

$$\lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x_0)}{\Delta x} = f'(x_0)$$

We take any line

$$g(x) = f(x_0) + m(x - x_0)$$

through the point $(x_0, f(x_0))$, and consider the approximation

$$f(x) - g(x) = f(x) - f(x_0) - m(x - x_0)$$

Then

$$\frac{f(x) - g(x)}{x - x_0} = \frac{f(x) - f(x_0)}{x - x_0} - m \to 0,\ x \to x_0 \Leftrightarrow m = f'(x_0)$$

Thus $f(x) - g(x) = f(x) - f(x_0) - m(x - x_0) = o(x - x_0)$ if and only if g is the tangent line to f at x_0, i.e. $m = f'(x_0)$.

18.2 Definitions

The functional $L : \mathbb{R}^d \mapsto \mathbb{R}$ is called linear if $\forall\, x, y \in \mathbb{R}^d,\ a \in \mathbb{R}$

1. $L(x + y) = L(x) + L(y)$
2. $L(ax) = a\, L(x)$

Additionally, by Th. 12.1 (Riesz Representation Theorem)

$$\exists\, v = (v_1, \ldots, v_d) \in \mathbb{R}^d : L(x) = \langle v, x\rangle \quad \forall\, x \in \mathbb{R}^d$$

As before, we will approximate a function $f : \mathbb{R}^d \mapsto \mathbb{R}$ by another function

$$g(x) = a + L(x - x_0) = a + \langle v, x - x_0\rangle$$

Definition 18.1 *Let x_0 be an inner point of $D \subseteq \mathbb{R}^d$. The function $f : D \mapsto \mathbb{R}$ is called differentiable at x_0 if there exists a linear function $L(x) = \langle v, x\rangle$ such that*

$$f(x) - f(x_0) - L(x - x_0) = o(\|x - x_0\|),\ x \to x_0 \Leftrightarrow \lim_{x \to x_0} \frac{f(x) - f(x_0) - L(x - x_0)}{\|x - x_0\|} = 0$$

The function $g(x) = f(x_0) + L(x - x_0)$, $x \in \mathbb{R}^d$ is the tangent plane to f through the point $(x_0, f(x_0))$.

Definition 18.2 *The function L is called the differential of f at x_0 and is denoted $df(x_0) = L$. Alternatively, $df(x_0) = v_1\, dx_1^0 + \cdots + v_d\, dx_d^0$.*

18.3 Partial Derivatives

Let $e_1, \ldots, e_d$ be the standard basis in $\mathbb{R}^d$.

Definition 18.3 *The limit*

$$\frac{\partial f}{\partial x_k}(x_0) = f'_{x_k}(x_0) = \lim_{\Delta x_k \to 0} \frac{f(x_1^0, \ldots, x_k^0 + \Delta x_k, \ldots, x_d^0) - f(x_1^0, \ldots, x_k^0, \ldots, x_d^0)}{\Delta x_k}$$

$$= \lim_{t \to 0} \frac{f(x_0 + t\, e_k) - f(x_0)}{t}$$

if it exists, is called the partial derivative of f at x_0 with respect to x_k.

Definition 18.4 *The vector*

$$\nabla f(x_0) = \left(\frac{\partial f}{\partial x_1}(x_0), \ldots, \frac{\partial f}{\partial x_d}(x_0) \right)$$

is called the gradient of f at x_0.

Theorem 18.1 *If f is differentiable at x_0, then for each $k = 1, \ldots, d$, there exists*

$$\frac{\partial f}{\partial x_k}(x_0) = \lim_{\Delta x_k \to 0} \frac{f(x_1^0, \ldots, x_k^0 + \Delta x_k, \ldots, x_d^0) - f(x_1^0, \ldots, x_d^0)}{\Delta x_k}$$

Moreover, the differential of f is defined as

$$df(x_0) = \frac{\partial f}{\partial x_1}(x_0)\, dx_1 + \cdots + \frac{\partial f}{\partial x_d}(x_0)\, dx_d$$

The linear map in Def. 18.1 then has the following form:

$$L\, x = \langle \nabla f(x_0), x \rangle$$

Proof: Assume f is differentiable. Then

$$\lim_{x \to x_0} \frac{f(x) - f(x_0) - L(x - x_0)}{\|x - x_0\|} = 0$$

This also means that

$$\lim_{t \to 0} \frac{f(x_0 + t\, e_k) - f(x_0) - L(t\, e_k)}{\|t\, e_k\|} = \lim_{t \to 0} \frac{f(x_0 + t\, e_k) - f(x_0)}{t} - L\, e_k = 0$$

Then $\frac{\partial f}{\partial x_k}(x_0) = L\, e_k = v_k$. □

Remark *$g(x) = f(x_0) + \nabla f(x_0)(x - x_0)$ is the tangent plane to the graph of f through the point $(x_0, f(x_0))$.*

Theorem 18.2 *If f is differentiable at x_0, then f is continuous at x_0.*

Example *The function $f(x, y) = \frac{xy}{x^2+y^2}$ is discontinuous at 0, so f is not differentiable at 0, but the partial derivatives of f exist at each point of $\mathbb{R}^2$. Therefore, the inverse statement to Th. 18.1 is not true.*

Theorem 18.3 *Let x_0 be an inner point of D and let $f : D \mapsto \mathbb{R}$ be given. If*

1. $\exists\, \varepsilon > 0 : \forall\, z \in B_\varepsilon(x_0) \quad \exists\, \dfrac{\partial f}{\partial x_k}(z) \quad \forall\, k = 1, \ldots, d$

2. $\dfrac{\partial f}{\partial x_k}$ *is continuous at x_0 for all $k = 1, \ldots, d$*

then f is differentiable at x_0.

Corollary 18.1 *If f has continuous partial derivatives on D, where D is open, then f is differentiable at each point of D. The set $\mathrm{C}^1(D)$ is the set of all differentiable functions on D. Thus*

$$f \in \mathrm{C}^1(D) \Leftrightarrow \frac{\partial f}{\partial x_k} \in \mathrm{C}(A) \quad \forall\, k$$

Theorem 18.4 *If $f : D \mapsto \mathbb{R}$ and $g : D \mapsto \mathbb{R}$ are differentiable at $x_0 \in D$, then the following are also differentiable at x_0:*

1. $c\,f$
2. $f + g$
3. fg
4. $\dfrac{f}{g}$, $g(x_0) \neq 0$

Theorem 18.5 *Let $f : D \mapsto \mathbb{R}$ be differentiable at x^0. Let $x_k = x_k(t_1, \ldots, t_m)$ be such that $x_k^0 = x_k(t_1^0, \ldots, t_m^0)$ and $\exists\, \frac{\partial x_k}{\partial t_j}(t^0), \forall\, j, k$. Then for the function $h(t) = f(x(t))$, there exist partial derivatives*

$$\frac{\partial h}{\partial t_j}(t^0) = \sum_{k=1}^{d} \frac{\partial f}{\partial x_k}(x^0) \frac{\partial x_k}{\partial t_j}(t^0)$$

19 Differentiation of Functions of Several Variables (Part II)

(**Lecture Notes**)

19.1 Derivatives of Real-Valued Functions

Theorem 19.1 (Chain Rule I) *Let the function $f : D \mapsto \mathbb{R}$ be differentiable at x_0 and let $x_k = x_k(t_1, \dots, t_m)$ be such that the partial derivatives $\frac{\partial x_k}{\partial t_j}(t^0)$ exist for all j, k. We take $x_k^0 = x_k(t_1^0, \dots, t_m^0)$. Then for the function*

$$h(t_1, \dots, t_m) = f\big(x_1(t_1, \dots, t_m), \dots, x_d(t_1, \dots, t_m)\big)$$

there exists

$$\frac{\partial h}{\partial t_j}(t_0) = \sum_{k=1}^{d} \frac{\partial f}{\partial x_k}(x_0) \frac{\partial x_k}{\partial t_j}(t_0), \quad x(t_0) = x_0$$

Definition 19.1 *The limit*

$$\frac{\partial f}{\partial l}(x_0) = \lim_{t \to 0} \frac{f(x_0 + tl) - f(x_0)}{t}$$

if it exists, is the directional derivative of $f : D \mapsto \mathbb{R}$ at an inner point x_0 of D in the direction of the vector $l = (l_1, \dots, l_d) \in \mathbb{R}^d$.

Theorem 19.2 *If $f : D \mapsto \mathbb{R}$ is differentiable at an inner point x_0 of D, then for any vector $l = (l_1, \dots, l_d) \in \mathbb{R}^d$, the directional derivative in the direction of l exists and*

$$\frac{\partial f}{\partial l}(x_0) = \langle \nabla f(x_0), l \rangle = \sum_{k=1}^{d} \frac{\partial f}{\partial x_k}(x_0) \cdot l_k$$

Theorem 19.3 *If $f : D \mapsto \mathbb{R}$ is differentiable at an inner point x_0 of D, then*

$$\max_{\|l\|=1} \frac{\partial f}{\partial l}(x_0) = \|\nabla f(x_0)\|$$

Moreover, the maximum is attained by a vector with the same direction as $\nabla f(x_0)$.

Proof: By the Cauchy-Schwarz inequality,

$$\frac{\partial f}{\partial l}(x_0) = \langle \nabla f(x_0), l \rangle \leq \|\nabla f(x_0)\| \|l\| = \|\nabla f(x_0)\|$$

Taking $\tilde{l} = \frac{\nabla f(x_0)}{\|\nabla f(x_0)\|}$,

$$\frac{\partial f}{\partial l} = \langle \nabla f(x_0), \tilde{l} \rangle = \frac{1}{\|\nabla f(x_0)\|} \langle \nabla f(x_0), \nabla f(x_0) \rangle = \|\nabla f(x_0)\|$$

$\square$

19.2 Derivatives of Vector-Valued Functions

Definition 19.2 *Let x_0 be an inner point of D. A function $f : D \mapsto \mathbb{R}^m$, $D \subseteq \mathbb{R}^d$ is called differentiable at x_0 if there exists a linear map $L : \mathbb{R}^d \mapsto \mathbb{R}^m$ such that*

$$f(x) - f(x_0) - L(x - x_0) = o(\|x - x_0\|),\ x \to x_0$$

In the standard basis, the linear map L can be given by a matrix, which is called the derivative of f at x_0:

$$f'(x_0) = \begin{pmatrix} v_{11} & \dots & v_{1d} \\ \vdots & & \vdots \\ v_{m1} & \dots & v_{md} \end{pmatrix}$$

Theorem 19.4 *A function $f : D \mapsto \mathbb{R}^m$, $f = (f_1, \dots, f_m)$ is differentiable at x_0 if and only if $f_k : D \mapsto \mathbb{R}$ is differentiable at x_0 for all $k = 1, \dots, m$. Moreover*

$$f'(x_0) = \left(\frac{\partial f_i}{\partial x_j}(x_0)\right)_{i,j=1}^{m,d} = \begin{pmatrix} \dfrac{\partial f_1}{x_1}(x_0) & \dots & \dfrac{\partial f_1}{x_d}(x_0) \\ \vdots & & \vdots \\ \dfrac{\partial f_m}{x_1}(x_0) & \dots & \dfrac{\partial f_m}{x_d}(x_0) \end{pmatrix}$$

Definition 19.3 *The matrix*

$$f'(x_0) = \left(\frac{\partial f_i}{\partial x_j}(x_0)\right)_{i,j=1}^{m,d}$$

is called the Jacobian matrix of f at x_0. If $m = d$ then the determinant

$$\frac{\partial(f_1, \dots, f_d)}{\partial(x_1, \dots, x_d)} = \det f'(x_0)$$

is called the Jacobian determinant of f at x_0. A point x_0 at which $\det f'(x_0) = 0$ is called a singular point.

Theorem 19.5 (Chain Rule II) *Let $D \subseteq \mathbb{R}^d$, $M \subseteq \mathbb{R}^m$ be open. If $f : D \mapsto M$ is differentiable at x_0 and if $g : M \mapsto \mathbb{R}^m$ is differentiable at $y_0 = f(x_0)$, then the function $h = g \circ f : D \mapsto \mathbb{R}^m$ is differentiable at x_0 and*

$$h'(x_0) = g'(f(x_0)) \cdot f'(x_0)$$

Corollary 19.1 *Under the assumptions of Th. 19.5 and if $n = m = d$, then*

$$\frac{\partial(h_1, \dots, h_d)}{\partial(x_1, \dots, x_d)} = \frac{\partial(g_1, \dots, g_d)}{\partial(f_1, \dots, f_d)} \frac{\partial(f_1, \dots, f_d)}{\partial(x_1, \dots, x_d)}$$

20 Implicit Function Theorem, Higher Order Derivatives

(**Lecture Notes**)

20.1 Implicit Function Theorem

Let D be an open set in $\mathbb{R}^d$. We denote by $\mathrm{C}^1(D, \mathbb{R}^m)$ the set of functions $f : D \mapsto \mathbb{R}^m$ which are differentiable at each point $x \in D$ and for which $f' : D \mapsto \mathbb{R}^{m+d}$ is continuous.

Theorem 20.1 *Let D be an open set in $\mathbb{R}^d$. For $f : D \mapsto \mathbb{R}^d$, $x_0 \in D$, and $y_0 = f(x_0)$, assume that the following conditions hold:*

1. $f \in \mathrm{C}^1(D, \mathbb{R}^d)$
2. $\det f'(x_0) \neq 0$

Then an open set $G \subseteq D$, which contains x_0, and a ball $B_r(y_0)$ exist such that

1. *$f : G \mapsto B_r(y_0)$ is a bijective map*
2. *the inverse map $g = f^{-1} : B_r(y_0) \mapsto G$ belongs to $\mathrm{C}^1(B_r(y_0), \mathbb{R}^d)$*
3. $g'(y) = (f^{-1})'(y) = \Big(f'\big(g(y)\big)\Big)^{-1} \quad \forall\, y \in B_r(y_0)$

Theorem 20.2 (Implicit Function Theorem) *Let G be an open set in $\mathbb{R}^{d+m}$ and take a point $(x_0, y_0) \in G$, where $x_0 \in \mathbb{R}^d$ and $y_0 \in \mathbb{R}^m$. Assume that $F : G \mapsto \mathbb{R}^m$ satisfies the following properties:*

1. $F(x_0, y_0) = 0$
2. $F \in \mathrm{C}^1(G, \mathbb{R}^m)$
3. $\det F'_y(x_0, y_0) \neq 0$, *where F'_y is the derivative of F with respect to y*

Then a ball $B_r(x_0) \subset \mathbb{R}^d$ and a unique function $h : B_r(x_0) \mapsto \mathbb{R}^m$, $h \in \mathrm{C}^1(B_r(x_0), \mathbb{R}^m)$ exist such that

1. $h(x_0) = y_0$
2. $F(x, h(x)) = 0 \quad \forall\, x \in B_r(x_0)$
3. $h'(x) = -\Big(F'_y\big(x, h(x)\big)\Big)^{-1} F'_x(x, h(x))$

Example *Consider* $F(x_1, x_2, y) = x_1^2 + x_2^2 + y^2 - 1$. *This equality defines the unit sphere in* $\mathbb{R}^3$. *Take* $x_0 = (0, 0)$ *and* $y_0 = 1$. *Then*

1. $F(x_0, y_0) = 0$
2. $F \in \mathrm{C}^1(G, \mathbb{R})$
3. $F'_y(x_0, y_0) = 2y_0 = 2 \neq 0$

So, a ball $B_r(x_0)$ *and a unique function* $y = h(x_1, x_2) : B_r(x_0) \mapsto \mathbb{R}$ *exist, in this example,* $y = h(x_1, x_2) = \sqrt{1 - x_1^2 - x_2^2}$, *such that*

1. $y_0 = h(x_1^0, x_2^0)$
2. $F(x_1, x_2, h(x_1, x_2)) = 0$
3. $F'_x(x, y) = (2x_1, 2x_2)$
   $h'(x_1, x_2) = -\frac{1}{2y}(2x_1, 2x_2) = -\left(\dfrac{x_1}{h(x_1, x_2)}, \dfrac{x_2}{h(x_1, x_2)}\right) \quad \forall\,(x_1, x_2) \in B_r(x_0)$

20.2 Higher Order Derivatives

Let D be an open set in $\mathbb{R}^d$ and consider a function $f : D \mapsto \mathbb{R}^d$.

Definition 20.1 *The second order partial derivative of* f *at* $x_0 \in D$, *if it exists, is*

$$\frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right)(x_0) = \frac{\partial^2 f}{\partial x_i\,\partial x_j}(x_0) \qquad \left(= \frac{\partial^2 f}{\partial x_i^2} \text{ if } i = j\right)$$

Theorem 20.3 (Schwarz's Theorem) *If the second order partial derivatives* $\frac{\partial^2 f}{\partial x\,\partial y}$ *and* $\frac{\partial^2 f}{\partial y\,\partial x}$ *exist and are continuous on* D, *then they are equal:*

$$\frac{\partial^2 f}{\partial x\,\partial y} = \frac{\partial^2 f}{\partial y\,\partial x}$$

Definition 20.2 *The matrix*

$$\mathrm{Hess}_{x_0} f = f''(x_0) = \left(\frac{\partial^2 f}{\partial x_i\,\partial x_j}\right)_{i,j=1}^{d}$$

is called the second order derivative of f *at* x_0, *or the Hessian matrix of* f.

Example *Consider* $f(x, y) = xe^y + y$. *Then* $\frac{\partial f}{\partial x} = e^y$ *and* $\frac{\partial f}{\partial y} = xe^y + 1$. *Thus we obtain*

$$f''(x, y) = \begin{pmatrix} 0 & e^y \\ e^y & xe^y \end{pmatrix}$$

By the same way, one can introduce the n^{th}*-order derivative of* f*:*

$$\frac{\partial^n f}{\partial x_{i_1} \ldots \partial x_{i_n}}(x_0) = \frac{\partial}{\partial x_{i_1}}\left(\frac{\partial}{\partial x_{i_2}} \cdots \left(\frac{\partial f}{\partial x_{i_n}}\right) \cdots\right)(x_0)$$

Definition 20.3 *We define* $\mathrm{C}^n(D)$ *as a class of functions* $f : D \mapsto \mathbb{R}$ *such that*

$$\frac{\partial^n f}{\partial x_{i_1} \dots \partial x_{i_n}}$$

exists and is continuous on D *for every* $i_1, \dots, i_n = 1, \dots, d$.

20.3 Taylor's Theorem

Theorem 20.4 *We assume* $f \in \mathrm{C}^n(D)$. *Let* $x_0, x \in D$ *be given and for all* $\theta \in [0,1]$ *we have* $(1-\theta)x_0 + \theta x \in D$. *Then*

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(x_0)(x-x_0)^2 + \dots$$

$$\dots + \frac{1}{(n-1)!} f^{(n-1)}(x_0)(x-x_0)^{n-1} + \frac{1}{n!} f^{(n)}\big((1-\theta)x_0 + \theta x\big)(x-x_0)^n$$

where θ *is some point from* $[0,1]$, *and*

$$f^{(k)}(x_0)(x-x_0)^k = \sum_{i_1,\dots,i_k=1}^{d} \frac{\partial^k f(x_0)}{\partial x_{i_1} \dots \partial x_{i_k}} (x_{i_1} - x^0_{i_1}) \dots (x_{i_k} - x^0_{i_k})$$

Remark *If* $n = 2$, *then*

$$f(x) = f(x_0) + \big\langle \nabla f(x_0), x - x_0 \big\rangle + \frac{1}{2} \big\langle f''(\tilde{x})(x-x_0), x - x_0 \big\rangle$$

where $\tilde{x} = (1-\theta)x_0 + \theta x$, $\theta \in [0,1]$ *and* $\big\langle f''(\tilde{x})(x-x_0), x-x_0 \big\rangle$ *is a bilinear form.*

Example *Consider* $f(x,y) = \sin x \sin y$. *Then*

$$\nabla f(x,y) = (\cos x \sin y, \sin x \cos y)$$

$$f''(x,y) = \begin{pmatrix} -\sin x \sin y & \cos x \cos y \\ \cos x \cos y & -\sin x \sin y \end{pmatrix} \left(= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ at } (0,0) \right)$$

Close to the point $x_0 = (0,0)$, $f(x,y) \approx 0 + 0 + \frac{1}{2}xy + \frac{1}{2}xy = xy$.

21 Extrema of Functions of Several Variables (Lecture Notes/Slides)

21.1 Necessary Conditions of Local Extrema

Consider $f : D \mapsto \mathbb{R}$, $D \subseteq \mathbb{R}^d$.

Definition 21.1

- *A point $x_0 \in D$ is called a local maximum (minimum) of f if there exists $r > 0$ such that*

    1. $B_r(x_0) \subseteq D$
    2. $f(x) \leq f(x_0) \quad \forall\, x \in B_r(x_0) \quad \big(f(x) \geq f(x_0) \quad \forall\, x \in B_r(x_0)\big)$

- *If $f(x_0) \geq f(x) \quad \forall\, x \in D \quad \big(f(x_0) \leq f(x) \quad \forall\, x \in D\big)$, then the point x_0 is called the global maximum (minimum).*

- *If x_0 is a local maximum or a local minimum then x_0 is called a local extremum.*

Theorem 21.1 *If x_0 is a local extremum of f, then assuming $\nabla f(x_0)$ exists, $\nabla f(x_0) = 0$.*

Definition 21.2 *If x_0 is an inner point of D for which $\nabla f(x_0) = 0$, then x_0 is called a critical point of f.*

Remark *In general, if x_0 is a critical point, it is not necessarily a local extremum.*

Example *Consider the function $f(x, y) = x^2 - y^2$ from Figure 1. The point $x_0 = (0, 0)$ is a critical point of f but it is not a local extremum. In the case of $f(x, y) = x^2 - y^2$, the point $x_0 = (0, 0)$ is called a saddle point.*

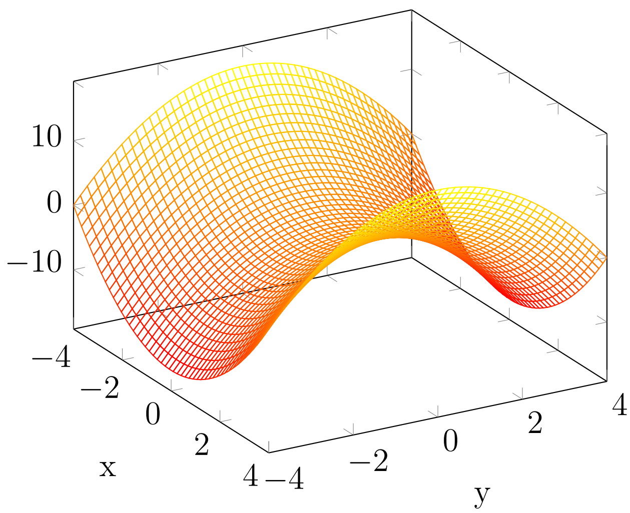


Figure 1: $f(x, y) = x^2 - y^2$

21.2 Sufficient Conditions of Local Extrema

Theorem 21.2 *Let D be an open set in $\mathbb{R}^d$. Consider $f \in \mathrm{C}^2(D)$ and assume $x_0 \in D$ is a critical point of f.*

1. *If $f''(x_0)$ is positive-definite, then x_0 is a local minimum of f.*
2. *If $f''(x_0)$ is negative-definite, then x_0 is a local maximum of f.*
3. *If $f''(x_0)$ is indefinite, i.e. $\langle f''(x_0)\, u, u\rangle > 0$ and $\langle f''(x_0)\, v, v\rangle < 0$ for some u and v, then x_0 is not a local extremum of f.*

(Refer to Def. 20.2 and Th. 14.6)

Corollary 21.1 *Let D be open in $\mathbb{R}^2$. Assume $\frac{\partial f}{\partial x}(x_0, y_0) = 0$ and $\frac{\partial f}{\partial y}(x_0, y_0) = 0$.*

1. *If $\dfrac{\partial^2 f}{\partial x^2}(x_0, y_0) > 0$ and $\det f''(x_0, y_0) > 0$, then (x_0, y_0) is a local minimum.*
2. *If $\dfrac{\partial^2 f}{\partial x^2}(x_0, y_0) < 0$ and $\det f''(x_0, y_0) > 0$, then (x_0, y_0) is a local maximum.*
3. *If $\det f''(x_0, y_0) < 0$, then (x_0, y_0) is not a local extremum.*

Example *Consider $f(x, y, z) = x^2 + y^2 + z^2 + 2x + 4y - 6z + xy$.*

1. *Find critical points at which $\nabla f(x) = 0$:*

$$\frac{\partial f}{\partial x} = 2x + 2 + y = 0 \qquad \frac{\partial f}{\partial y} = 2y + 4 + x = 0 \qquad \frac{\partial f}{\partial z} = 2z - 6 = 0$$

$$\begin{cases} 2x + y = -2 \\ x + 2y = -4 \\ 2z = 6 \end{cases} \Rightarrow x = 0,\ y = -2,\ z = 3$$

We find $x_0 = (0, 2, -3)$ is a critical point.

2. *Check if x_0 is a local extremum:*

$$f''(x, y, z) = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \Rightarrow M_1 = 2 > 0,\ M_2 = 3 > 0,\ M_3 = 6 > 0$$

$f''(x, y, z)$ is positive-definite, therefore x_0 is a local minimum.

22 Conditional Local Extrema (Lecture Notes)

22.1 Some Exercises

1. Find the point in the plane $3x + 4y + z = 1$ closest to $(-1, 1, 1)$. For this we minimize the squared distance function

$$f(x, y, z) = (x + 1)^2 + (y - 1)^2 + (z - 1)^2$$

with the constraint $3x + 4y + z = 1$ given by the equation of the plane.

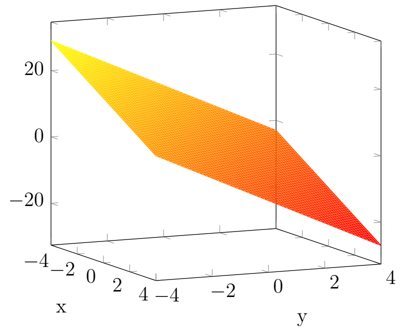


Figure 2: $3x + 4y + z = 1$

2. Find the minimum distance between two curves in $\mathbb{R}^2$ given by $x^2 + 2y^2 = 1$ (ellipse) and $x + y = 4$ (line). For this we minimize the function

$$f(x_1, y_1, x_2, y_2) = (x_1 - x_2)^2 + (y_1 - y_2)^2$$

with the constraints $x_1^2 + 2y_1^2 = 1$ and $x_2 + y_2 = 4$ given by the equations of the two curves, then take the square root of the result.

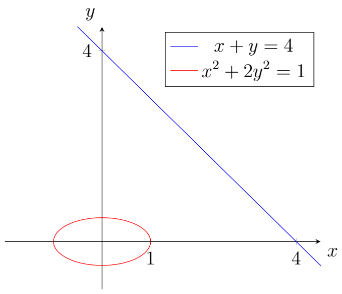


Figure 3

3. Find the size of an open rectangular bath of volume V for which its surface area is a minimum. For this we minimize the function

$$f(x, y, z) = 2xy + 2yz + xz$$

with the constraint $xyz = V$.

22.2 Method of Lagrange Multipliers

Consider functions $f : D \mapsto \mathbb{R}$, $g_i : D \mapsto \mathbb{R}$, $i = 1, \ldots, m$, where $D \subseteq \mathbb{R}^d$ is an open set. We want to find conditional local extrema of f subject to constraints $g_1 = 0, \ldots, g_m = 0$.

Definition 22.1 *Let $M = \{x \in D : g_1(x) = 0, \ldots, g_m(x) = 0\}$. A point $x_0 \in D$ is called a conditional local maximum (minimum) of f subject to the constraints $g_1 = 0, \ldots, g_m = 0$ if*

$$\exists\, r > 0 : \forall\, x \in B_r(x_0) \cap M \quad f(x_0) \geq f(x) \quad \big(f(x_0) \leq f(x)\big)$$

If x_0 is a conditional local maximum or minimum, then x_0 is called a conditional local extremum.

Theorem 22.1 *We assume that $m < d$, $f \in \mathrm{C}^1(D)$, $g_i \in \mathrm{C}^1(D)$, $i = 1, \ldots, m$ and the matrix*

$$\left(\frac{\partial g_i}{\partial x_j}(x_0) \right)_{i,j=1}^{m,d}$$

has rank m at x_0, where x_0 is a conditional local extremum of f subject to the constraints $g_1 = 0, \ldots, g_m = 0$. Then there exist real numbers $\lambda_1, \ldots, \lambda_m$ for which x_0 is a critical point of the function

$$F(x) = f(x) - \lambda_1 g_1(x) - \cdots - \lambda_m g_m(x)$$

that is

$$\frac{\partial F}{\partial x_j}(x_0) = 0 \quad \forall\, j = 1, \ldots, d$$

Method of Lagrange Multipliers:

(i) Find all solutions $x_1, \ldots, x_d$, $\lambda_1, \ldots, \lambda_m$ of the system

$$\begin{cases} \dfrac{\partial F}{\partial x_j}(x) = 0 \quad j = 1, \ldots, d \\ g_i(x) = 0 \quad i = 1, \ldots, m \end{cases}$$

(ii) Determine which of the critical points are conditional local extrema of f. This can usually be done using intuitive or physical arguments.

Example (Solving Ex. 1, 2 from 22.1 using the Method of Lagrange Multipliers)

1. *Solving Ex. 1*
   We must minimize the function

$$f(x, y, z) = (x+1)^2 + (y-1)^2 + (z-1)^2$$

   with the constraint

$$g(x, y, z) = 3x + 4y + z - 1 = 0$$

   We then have

$$F(x, y, z) = (x+1)^2 + (y-1)^2 + (z-1)^2 - \lambda(3x + 4y + z - 1)$$

$$\frac{\partial F}{\partial x} = 2(x+1) - 3\lambda = 0 \quad \frac{\partial F}{\partial y} = 2(y-1) - 4\lambda = 0 \quad \frac{\partial F}{\partial z} = 2(z-1) - \lambda = 0$$

   Solving the system of equations for λ in terms of x, y, z yields $\lambda = -\frac{1}{13}$. Thus

$$x = -\frac{29}{26} \quad y = \frac{11}{13} \quad z = \frac{25}{26}$$

2. *Solving Ex. 2*
   We must minimize the function

$$f(x_1, y_1, x_2, y_2) = (x_1 - x_2)^2 + (y_1 - y_2)^2$$

   with the constraints

$$g_1(x_1, y_1, x_2, y_2) = x_1^2 + 2y_1^2 - 1 = 0 \qquad g_2(x_1, y_1, x_2, y_2) = x_2 + y_2 - 4 = 0$$

   We then have

$$F(x_1, y_1, x_2, y_2) = (x_1 - x_2)^2 + (y_1 - y_2)^2 - \lambda_1(x_1^2 + 2y_1^2 - 1) - \lambda_2(x_2 + y_2 - 4)$$

$$\frac{\partial F}{\partial x_1} = 2(x_1 - x_2) - 2\lambda_1 x_1 = 0 \qquad \frac{\partial F}{\partial y_1} = 2(y_1 - y_2) - 4\lambda_1 y_1 = 0$$

$$\frac{\partial F}{\partial x_2} = -2(x_1 - x_2) - \lambda_2 = 0 \qquad \frac{\partial F}{\partial y_2} = -2(y_1 - y_2) - \lambda_2 = 0$$

   Solving the system of equations and ignoring unphysical solutions, we obtain

$$(x_1, y_1) = \left(\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right) \quad (x_2, y_2) = \left(2 + \frac{1}{2\sqrt{6}}, 2 - \frac{1}{2\sqrt{6}}\right)$$

   Thus the minimum distance between the two curves is $\sqrt{2}\left(2 - \frac{3}{\sqrt{6}}\right) = 2\sqrt{2} - \sqrt{3}$.

23 Basic Concepts of Differential Equations (Lecture Notes/Slides)

23.1 Models Leading to Differential Equations

1. **Population Growth and Decay**
   Let $P(t)$ be the number of members of a population, which, in order to simplify the mathematical model, we will assume can take any positive value, and let a be the rate of change of that population. This results in the differential equation

   $$P'(t) = a\,P(t)$$

2. **Spread of Epidemics**
   Let the rate of change of the infected population be proportional to the product of the number of people already infected and the number of people susceptible to infection but not already infected. This leads to the differential equation

   $$I'(t) = r\,I(t)\big(S - I(t)\big)$$

   where S is the total number of members of the population, $I(t)$ is the number of infected members at time t, and r is some positive constant. If at time $t = 0$ there were I_0 infected people, then we can add to the equation the condition

   $$I(0) = I_0$$

3. **Simple Pendulum**
   Consider a pendulum with length $l = 1$. Its motion as a function of time can be described by the differential equation

   $$\theta'' + g\sin\theta = 0$$

   where g is the gravitational constant. If the amplitude is small, we can approximate $\sin\theta \approx \theta$, resulting in the differential equation

   $$\theta'' + g\,\theta = 0$$

23.2 Basic Definitions

Definition 23.1

- *A differential equation is an equation that contains one or more derivatives of an unknown function.*
- *The order of a differential equation is the order of the highest derivative that it contains.*
- *If a differential equation involves an unknown function of only one variable, it is called an ordinary differential equation.*

We will consider differential equations of the form

$$y^{(n)} = f\left(x, y, y', \dots, y^{(n-1)}\right) \tag{23.1}$$

Example (Some Examples of Differential Equations)

1. $y' = x^2$
2. $y' = (y^2 + 1)x^2$
3. $y'' = 2x - 2y' - y$
4. $y^{(n)} = y' - \sin y + x$

Definition 23.2 *A solution to a differential equation of the form (23.1) is a function $y = y(x)$ that is defined on some open interval (a, b) and can be differentiated n times such that*

$$y^{(n)}(x) = f\left(x, y(x), y'(x), \dots, y^{(n-1)}(x)\right) \quad \forall\, x \in (a, b)$$

Definition 23.3

- *The graph of a solution of a differential equation is called a solution curve.*
- *A curve C is said to be an integral curve of a differential equation if every function $y = y(x)$ whose graph is a segment of C is a solution to the differential equation.*

Example *Consider the differential equation*

$$y' = -\frac{x}{y} \tag{23.2}$$

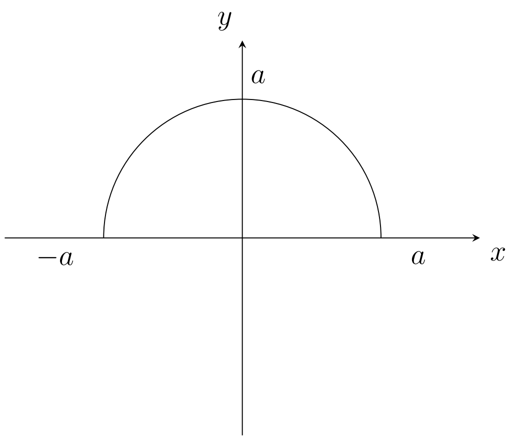
The functions

$$y_{1,2}(x) = \pm\sqrt{a^2 - x^2},\ x \in (-a, a)$$

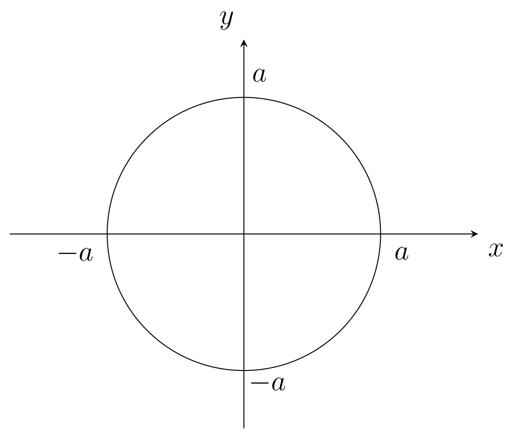
where a is a constant, are solutions to (23.2). The graphs of these functions are then solution curves of the differential equation. The integral curve of the differential equation is given by the equation of a circle

$$x^2 + y^2 = a^2$$

Note that this equation is not a solution curve because it is not a graph of some function $y = y(x)$.



(a) Solution curve $y_1(x) = \sqrt{a^2 - x^2}$



(b) Integral curve $x^2 + y^2 = a^2$

Example *Solving the differential equation*

$$y'' = e^x \tag{23.3}$$

can be done by integration:

$$y'(x) = \int y''(x)\,dx = \int e^x\,dx = e^x + c_1$$

$$y(x) = \int y'(x)\,dx = \int (e^x + c_1)\,dx = e^x + c_1 x + c_2$$

with constants c_1 and c_2.

23.3 Initial Value Problem

As seen with (23.3), there can be infinitely many solutions to a differential equation, depending on the constants involved. The problem of finding solutions to (23.1) which satisfy the initial conditions

$$y(x_0) = p_0,\ y'(x_0) = p_1, \ldots, y^{(n-1)}(x_0) = p_{n-1} \tag{23.4}$$

for some x_0 from the domain of y is called the initial value problem.

Example *To solve (23.3) with initial conditions $y(0) = 1$, $y'(0) = 0$, we first take the general solution and ensure it satisfies the initial condition $y(0) = 1$:*

$$y(0) = e^0 + c_2 = 1 \Rightarrow c_2 = 0$$

Now we differentiate the general solution and ensure it satisfies the initial condition $y'(0) = 0$:

$$y'(0) = e^0 + c_1 = 0 \Rightarrow c_1 = -1$$

We thus obtain the particular solution to this initial value problem:

$$y(x) = e^x - x$$

which, under the right conditions, is the unique solution.

Example *Consider the differential equation*

$$y' = 2\sqrt{y} \tag{23.5}$$

with the initial condition $y(0) = 0$. Then $y(x) = 0$ and $y(x) = x^2$, $x \geq 0$ are two different solutions to this problem. This shows that solutions to the initial value problem are not unique.

Remark *If f is continuous, then the differential equation $y' = f(x)$ with the initial condition $y(x_0) = y_0$ has the following solution:*

$$y(x) = y_0 + \int_{x_0}^{x} f(t)\,dt$$

23.4 Directional Fields of First Order Differential Equations

Consider the equation

$$y' = f(x, y) \tag{23.6}$$

We will discuss the graphical method of solving (23.6). Recall that $y = y(x)$ is a solution to (23.6) if $y'(x) = f(x, y(x))$ for all x from some interval. So, the slope of the integral curve of (23.6) through a point (x_0, y_0) is given by the number $f(x_0, y_0)$. If f is defined on a set R, we can construct a directional field for (23.6) in R by drawing a short line segment or vector with slope $f(x, y)$ through each point (x, y) in R. Then integral curves of (23.6) are continuous curves tangent to the vectors in the directional field.

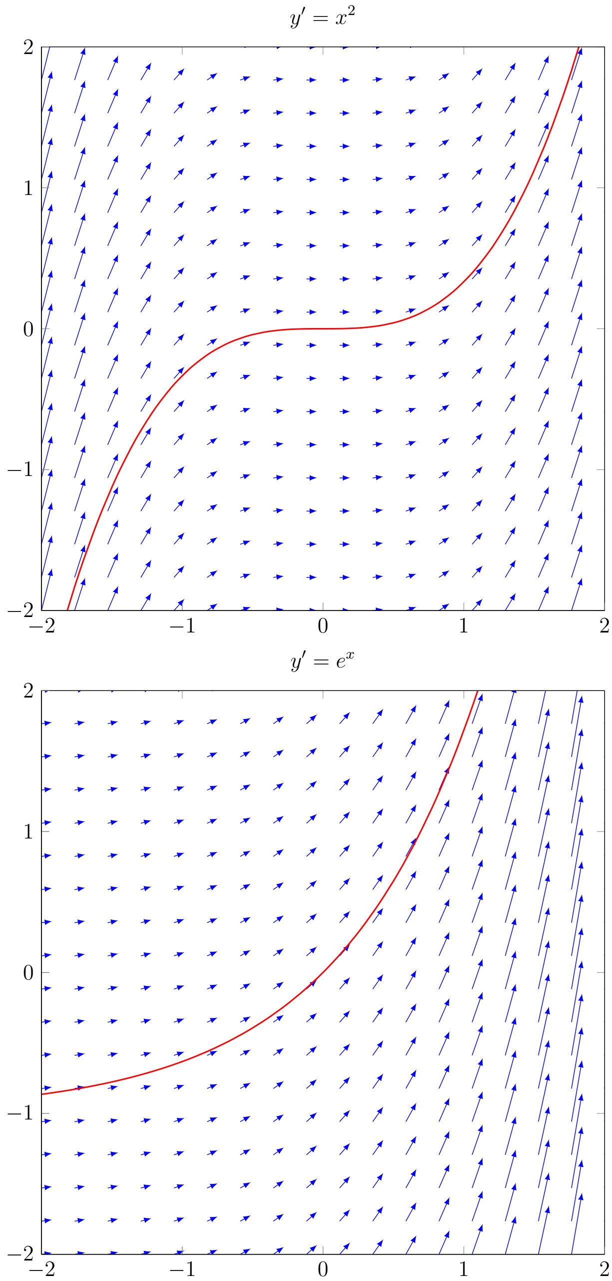


Figure 5: Examples of directional fields

24 First Order Differential Equations (Lecture Notes)

24.1 Separable Differential Equations

Definition 24.1 *A differential equation of the form*

$$y' = f(y)g(x) \tag{24.1}$$

is called a separable differential equation.

In this section, we will describe the method of solving equations of the form (24.1). First, we rewrite (24.1) in the form

$$h(y)y' = g(x)$$

where $h(y) := \frac{1}{f(y)}$. Assume that $h(y)$ and $g(x)$ have antiderivatives $H(y)$ and $G(x)$ respectively. By the chain rule we have

$$\frac{d}{dx}H\big(y(x)\big) = H'\big(y(x)\big) \cdot y'(x) = h\big(y(x)\big) \cdot y'(x)$$

Thus

$$\frac{d}{dx}H\big(y(x)\big) = \frac{d}{dx}G(x)$$

Integrating both sides of this equation will yield

$$H\big(y(x)\big) = G(x) + c \tag{24.2}$$

Consequently, any differentiable function $y = y(x)$ that satisfies (24.2) is a solution to (24.1). To find this solution, we must find the antiderivatives of h and g.

Example (Separation of Variables)

$$y' = \frac{dy}{dx} = x(1+y^2) \Rightarrow \frac{dy}{1+y^2} = x\,dx \Rightarrow \arctan y = \frac{x^2}{2} + c$$

$$y = \tan\left(\frac{x^2}{2} + c\right)$$

Example (Differential Equation with Implicit Solution)

$$y' = \frac{dy}{dx} = \frac{2x+1}{5y^4+1} \Rightarrow (5y^4+1)\,dy = (2x+1)\,dx$$

$$y^5 + y = x^2 + x + c$$

Remark *In dividing (24.1) by $f(y)$ we may lose some solutions, namely the constant solution $y(x) = y_0$, where $f(y_0) = 0$.*

Example *We will solve the initial value problem:*

$$y' = 2xy^2 \qquad y(0) = 1$$

First, we find the general solution of the differential equation by separation of variables:

$$y' = \frac{dy}{dx} = 2xy^2$$

$$\frac{dy}{y^2} = 2x\,dx,\ y \neq 0$$

$$-\frac{1}{y} = x^2 + c$$

$$y = -\frac{1}{x^2 + c} \tag{24.3}$$

Note that $y = 0$ is also a solution. Moreover, it cannot be written in the form (24.3). However, this means that it cannot satisfy the initial condition. We therefore take the former solution (24.3) to find c:

$$y(0) = -\frac{1}{0^2 + c} = 1 \Rightarrow c = -1$$

We then have the solution

$$y(x) = \frac{1}{1 - x^2},\ x \in (-1, 1)$$

24.2 Linear First Order Differential Equations

Definition 24.2 *A first order differential equation is said to be linear if it can be written in the form*

$$y' + p(x)\,y = f(x) \tag{24.4}$$

If $f = 0$, (24.4) is called homogeneous, otherwise, it is called nonhomogeneous.

To solve equations of the form (24.4), we first consider the corresponding homogeneous equation, when $f = 0$:

$$y' + p(x)\,y = 0 \tag{24.5}$$

We can solve (24.5) using the method of separation of variables:

$$\ln|y| = -P(x) + k$$

where $P(x) := \int p(x)\,dx$ and k is a constant. We then have

$$y = e^k\, e^{-P(x)} \quad \text{and} \quad y = -e^k\, e^{-P(x)}$$

Consequently, we can write solutions to (24.5) as

$$y = Ce^{-P(x)} \tag{24.6}$$

where $C \in \mathbb{R}$. Note that for $C = 0$, $y = 0$ is also a solution to (24.5).

Definition 24.3 *(24.6) is called the general solution to the homogeneous equation (24.5).*

In order to solve (24.4), we assume that C in (24.6) depends on x, that is,

$$y = C(x)\, e^{-P(x)} \tag{24.7}$$

We then substitute (24.7) into (24.4) and obtain

$$C(x) = \int f(x)\, e^{P(x)}\, dx + C_1$$

where C_1 is some constant of integration. We then have

$$y(x) = \left(\int f(x)\, e^{P(x)}\, dx \right) e^{-P(x)} + C_1 e^{-P(x)}$$

Remark *The general solution to (24.4) can be written as the sum of a partial solution to (24.4) and the general solution to the homogeneous equation (24.5).*

Example *Take the differential equation*

$$y' - 2xy = e^{x^2}$$

First solve

$$y' - 2xy = 0$$

$$y = Ce^{-\int(-2x)\,dx} = Ce^{x^2}$$

Now substitute $y = C(x)e^{x^2}$ into the differential equation to obtain

$$C(x) = \int dx = x + C_1$$

We then have

$$y(x) = (x + c_1)e^{x^2} = xe^{x^2} + C_1 e^{x^2}$$

24.3 Transformation of Nonlinear Functions

Here we consider the Bernoulli Equation, which is of the form

$$y' + p(x)\,y = f(x)\,y^r \tag{24.8}$$

where $r \in \mathbb{R} \setminus \{0, -1\}$. Let y_1 be a nontrivial solution to

$$y_1' + p(x)\,y_1 = 0$$

Then we find a solution to (24.8) in the form

$$y = u\,y_1$$

where u is some function. So, substituting $y = u\,y_1$ into (24.8), we obtain

$$u'\,y_1 + u\big(y_1' + p(x)\,y_1\big) = u'\,y_1 = f(x)(u\,y_1)^r$$

which is a separable differential equation and can be solved.

Example *Take the differential equation*

$$y' - y = xy^2$$

First solve

$$y' - y = 0$$

$$y = Ce^x$$

Then take $y_1(x) = e^x$. We substitute $y(x) = u(x)\,y_1(x) = u(x)e^x$ into the differential equation:

$$u'e^x = u^2xe^{2x}$$

$$u' = u^2xe^x$$

and we find a separable differential equation which we can solve:

$$u = -\frac{1}{(x-1)e^x + c}$$

Thus we find

$$y = -\frac{e^x}{(x-1)e^x + c} = -\frac{1}{x - 1 + ce^{-x}}$$

25 Existence and Uniqueness of Solutions, Higher Order Linear Differential Equations (Lecture Notes)

25.1 Homogeneous Nonlinear Differential Equations

An equation which can be written as

$$y' = q\left(\frac{y}{x}\right) \tag{25.1}$$

is called a homogeneous nonlinear differential equation. To solve (25.1), we find a solution in the form $y(x) = x\,u(x)$.

Example *Consider the homogeneous nonlinear differential equation*

$$y' = \frac{y + xe^{-\frac{y}{x}}}{x} \tag{25.2}$$

We substitute $y(x) = x\,u(x)$ into (25.2) and obtain

$$u'x + u = \frac{ux + xe^{-u}}{x} \Rightarrow u'x = e^{-u} \Rightarrow u = \ln(\ln|x| + c)$$

$$y = x\ln(\ln|x| + c)$$

25.2 Existence and Uniqueness of Solutions

Theorem 25.1 *(Peano) If f is continuous on $(a,b)\times(c,d)$ and $x_0 \in (a,b)$, $y_0 \in (c,d)$, then there exists $\varepsilon > 0$ such that the initial value problem*

$$y' = f(x,y),\; y(x_0) = y_0 \tag{25.3}$$

has a solution on $[x_0-\varepsilon, x_0+\varepsilon]$.

Theorem 25.2 *(Picard-Lindelöf) If $f = f(x,y)$ is*

1. *uniformly Lipschitz continuous in y, that is*

$$\exists\, L > 0 : |f(x,y_1) - f(x,y_2)| \le L|y_1 - y_2|$$

2. *continuous in x*

then there exists $\varepsilon > 0$ such that (25.3) has a unique solution on $[x_0-\varepsilon, x_0+\varepsilon]$.

Remark *If*

$$\left|\frac{\partial f}{\partial y}\right| \le C \quad \forall\,(x,y) \in (a,b)\times(c,d)$$

then the first condition from Th. 25.2 is satisfied.

25.3 Higher Order Linear Differential Equations with Constant Coefficients

Definition 25.1 *An equation of the form*

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \cdots + a_2 y'' + a_1 y' + a_0 y = F(x) \quad a_n \neq 0 \tag{25.4}$$

is called a higher order linear differential equation with constant coefficients. If $F = 0$, then (25.4) is called homogeneous.

We substitute

$$y = Ce^{\lambda x} \tag{25.5}$$

into the corresponding homogeneous equation of (25.4)

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \cdots + a_2 y'' + a_1 y' + a_0 y = 0 \quad a_n \neq 0$$

If λ is a solution to the resulting polynomial equation

$$a_n \lambda^n + a_{n-1} \lambda^{n-1} + \cdots + a_1 \lambda + a_0 = 0 \tag{25.6}$$

then (25.5) is a solution to the homogeneous equation of (25.4).

Definition 25.2 *(25.6) is called the characteristic polynomial of*

$$L\,y := a_n y^{(n)} + a_{n-1} y^{(n-1)} + \cdots + a_2 y'' + a_1 y' + a_0 y = 0$$

Example *Consider the differential equation*

$$y''' - 6y'' + 11y' - 6y = 0$$

We obtain the characteristic polynomial

$$\lambda^3 - 6\lambda^2 + 11\lambda - 6 = 0 \Rightarrow \lambda_1 = 1,\ \lambda_2 = 2,\ \lambda_3 = 3$$

and we find the general solution

$$y = C_1 e^x + C_2 e^{2x} + C_3 e^{3x}$$

Definition 25.3 *The set $\{y_1, \ldots, y_n\}$ is a fundamental system of solutions to $L\,y = 0$ if every solution y can be written as a linear combination of $\{y_1, \ldots, y_n\}$:*

$$y = C_1 y_1 + \cdots + C_n y_n$$

for some constants $C_1, \ldots, C_n$.

Definition 25.4 *The functions $y_1, \ldots, y_n$ are linearly independent if*

$$C_1 y_1(x) + \cdots + C_n y_n(x) = 0 \quad \forall\, x \Rightarrow C_1 = C_2 = \cdots = C_n = 0$$

Theorem 25.3 *A set of n solutions $\{y_1, \dots, y_n\}$ to $L\,y = 0$ is a fundamental system of solutions if and only if $y_1, \dots, y_n$ are linearly independent.*

If the characteristic polynomial of $L\,y = 0$ can be written as

$$P(\lambda) = (\lambda - \lambda_1)^{k_1} \dots (\lambda - \lambda_m)^{k_m}$$

we have the following cases:

1. $\lambda \in \mathbb{R}$, $k = 1$
   Solution: $y = e^{\lambda x}$

2. $\lambda \in \mathbb{R}$, $k > 1$
   Solutions: $y_1 = e^{\lambda x}$, $y_2 = xe^{\lambda x}, \dots, y_k = x^{k-1}e^{\lambda x}$

3. $\lambda = a \pm bi$, $k = 1$
   Solutions: $y_1 = e^{ax}\cos bx$, $y_2 = e^{ax}\sin bx$

4. $\lambda = a \pm bi$, $k > 1$
   Solutions:
   $y_1 = e^{ax}\cos bx$, $y_2 = e^{ax}\sin bx$, $y_3 = xe^{ax}\cos bx$, $y_4 = xe^{ax}\sin bx \dots$
   $y_{2k-1} = x^{k-1}e^{ax}\cos bx$, $y_{2k} = x^{k-1}e^{ax}\sin bx$

Example *Consider the differential equation*

$$y'' + 4y = 0$$

The characteristic polynomial yields

$$\lambda^2 + 4 = 0 \Rightarrow \lambda_1 = 2i,\ \lambda_2 = -2i$$

$$P(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) = (\lambda - 2i)(\lambda + 2i)$$

and we obtain the solution

$$y = C_1 \cos 2x + C_2 \sin 2x$$

26 Systems of Linear Differential Equations (Lecture Notes)

26.1 Rewriting Scalar Differential Equations as Systems

Example *Consider the differential equation describing the motion of a pendulum of length $l = 1$ as a function of time:*

$$\theta'' + g \sin\theta = 0 \tag{26.1}$$

We define the velocity $v(t) = \theta(t)$, from which $\theta''(t) = v'(t)$ follows. We can then rewrite (26.1) as a system of differential equations:

$$\begin{cases} v' = -g\sin\theta \\ v = \theta' \end{cases}$$

Example *Consider the differential equation*

$$a_0 y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1}y' + a_n y = F(t) \quad a_0 \neq 0$$

We define

$$\begin{aligned} y_1 &:= y \\ y_2 &:= y' = y_1' \\ y_3 &:= y'' = y_2' \\ &\vdots \\ y_n &:= y^{(n-1)} = y_{n-1}' \end{aligned}$$

We can then write

$$\begin{cases} y_1' = y_2 \\ y_2' = y_3 \\ \quad\vdots \\ y_{n-1}' = y_n \\ y_n' = -\dfrac{a_1}{a_0}y_n - \dfrac{a_2}{a_0}y_{n-1} - \cdots - \dfrac{a_{n-1}}{a_0}y_2 - \dfrac{a_n}{a_0}y_1 + F(t) \end{cases}$$

Definition 26.1 *A system of the form*

$$\begin{cases} y_1' = g_1(t, y_1, \ldots, y_n) \\ \qquad\vdots \\ y_n' = g_n(t, y_1, \ldots, y_n) \end{cases}$$

is called a first order system of differential equations.

26.2 Linear Systems of Differential Equations

Definition 26.2 *A first order system of differential equations of the form*

$$\begin{cases} y_1' = a_{11}(t)y_1(t) + \cdots + a_{1n}(t)y_n(t) + f_1(t) \\ \qquad\vdots \\ y_n' = a_{n1}(t)y_1(t) + \cdots + a_{nn}(t)y_n(t) + f_n(t) \end{cases} \tag{26.2}$$

is called a linear system of differential equations. Defining

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \quad A(t) = \begin{pmatrix} a_{11}(t) & \dots & a_{1n}(t) \\ \vdots & & \vdots \\ a_{n1}(t) & \dots & a_{nn}(t) \end{pmatrix} \quad f(t) = \begin{pmatrix} f_1(t) \\ \vdots \\ f_n(t) \end{pmatrix}$$

where A is called the coefficient matrix of (26.2), we can rewrite (26.2) as

$$y' = A(t)y + f(t)$$

If $f = 0$, then (26.2) is called homogeneous.

Theorem 26.1 *If A and f are continuous functions on (a, b) and $t_0 \in (a, b)$, $y_0 \in \mathbb{R}^n$, then*

$$y' = A(t)y + f(t),\ y(t_0) = y_0$$

has a unique solution on (a, b).

26.3 Homogeneous Systems of Linear Equations

Here, we will consider

$$y' = A(t)y \tag{26.3}$$

We have the trivial solution $y = (0, \dots, 0)$. Let $y^1, \dots, y^n$ be vector-valued functions which are solutions to (26.3). Then

$$y = c_1 y^1 + \cdots + c_n y^n \tag{26.4}$$

is also a solution to (26.3) for any constants $c_1, \dots, c_n$.

Definition 26.3 *If any solution to (26.3) can be written in the form (26.4) for some constants $c_1, \dots, c_n$, then $y^1, \dots, y^n$ is called the fundamental system of solutions to (26.3).*

Definition 26.4 *$y^1, \dots, y^n$ are called linearly independent on (a, b) if the equality*

$$c_1 y^1(t) + \cdots + c_n y^n(t) = 0 \quad \forall\, t \in (a, b)$$

implies $c_1 = \cdots = c_n = 0$

Theorem 26.2 *Let the $n \times n$ matrix $A = A(t)$ be continuous on (a, b). Then a set of solutions $y^1, \dots, y^n$ to (26.3) on (a, b) is a fundamental system of solutions if and only if it is linearly independent on (a, b).*

26.4 Solving Homogeneous Systems of Linear Differential Equations with Constant Coefficients

We will find solutions to (26.3), assuming the matrix A is constant, as solutions to higher order linear differential equations.

Example *Consider the system*

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} -4 & -3 \\ 6 & 5 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \Leftrightarrow \begin{cases} y_1' = -4y_1 - 3y_2 \\ y_2' = 6y_1 + 5y_2 \end{cases} \tag{26.5}$$

We will find solutions in the form

$$y_1 = x_1 e^{\lambda t} \quad y_2 = x_2 e^{\lambda t} \tag{26.6}$$

where x_1, x_2 are some constants. We then have

$$y_1' = \lambda x_1 e^{\lambda t} \quad y_2' = \lambda x_2 e^{\lambda t} \tag{26.7}$$

Substituting (26.6) and (26.7) into (26.5), we obtain

$$\begin{cases} -4x_1 - 3x_2 = \lambda x_1 \\ 6x_1 + 5x_2 = \lambda x_2 \end{cases} \Leftrightarrow \begin{cases} (-4-\lambda)x_1 - 3x_2 = 0 \\ 6x_1 + (5-\lambda)x_2 = 0 \end{cases} \tag{26.8}$$

To find nontrivial solutions, we find λ such that $\det(A - \lambda I) = 0$. We obtain $\lambda_1 = 2$ and $\lambda_2 = -1$. For $\lambda_1 = 2$ we have

$$\begin{cases} -6x_1 - 3x_2 = 0 \\ 6x_1 + 3x_2 = 0 \end{cases} \Rightarrow \begin{cases} y_1 = -e^{2t} \\ y_2 = 2e^{2t} \end{cases}$$

For $\lambda_2 = -1$ we have

$$\begin{cases} -3x_1 - 3x_2 = 0 \\ 6x_1 + 6x_2 = 0 \end{cases} \Rightarrow \begin{cases} y_1 = -e^{-t} \\ y_2 = e^{-t} \end{cases}$$

The general solution to (26.5) is then given by

$$y_1 = -c_1 e^{2t} - c_2 e^{-t}$$

$$y_2 = 2c_1 e^{2t} + c_2 e^{-t}$$

In order to solve

$$y' = Ay \tag{26.9}$$

we first find eigenvalues of the matrix A from the equation $\det(A - \lambda I) = 0$. Let

$$\det(A - \lambda I) = (\lambda - \lambda_1)^{k_1} \dots (\lambda - \lambda_m)^{k_m}$$

1. If $\lambda = \lambda_j \in \mathbb{R}$ and the fundamental system of solutions to

$$(A - \lambda I)x = 0 \tag{26.10}$$

consists of $k = k_j$ solutions $x^1, \dots, x^k$, then

$$y^1 = x^1 e^{\lambda t}, \dots, y^k = x^k e^{\lambda t}$$

are linearly independent solutions to (26.9).

2. If $\lambda \in \mathbb{R}$ and the fundamental system of solutions consists of $m < k$ solutions, then we find solutions to (26.9) in the form

$$y = \left(x^{(1)} + x^{(2)}t + \dots + x^{(k-m)}t^{k-m}\right)e^{\lambda t}$$

3. If $\lambda \in \mathbb{C} \setminus \mathbb{R}$, then we find solutions to (26.9) as before. We obtain

$$y = y^1 + iy^2$$

Then take y^1 and y^2 as linearly independent solutions to (26.9).

Example *Consider*

$$\begin{cases} y_1' = 2y_1 + y_2 + y_3 \\ y_2' = -2y_1 - y_3 \\ y_3' = 2y_1 + y_2 + 2y_3 \end{cases} \tag{26.11}$$

Solving the characteristic equation

$$\begin{vmatrix} 2-\lambda & 1 & 1 \\ -2 & -\lambda & -1 \\ 2 & 1 & 2-\lambda \end{vmatrix} = 0$$

we obtain $\lambda_1 = 2$ and $\lambda_2 = \lambda_3 = 1$. For $\lambda_1 = 2$ we find the fundamental system of solutions to

$$\begin{cases} x_2 + x_3 = 0 \\ -2x_1 - 2x_2 - x_3 = 0 \\ 2x_1 + x_2 = 0 \end{cases}$$

and obtain

$$y_1 = e^{2t} \quad y_2 = -2e^{2t} \quad y_3 = 2e^{2t}$$

Now for $\lambda_2 = \lambda_3 = 1$, we first find the number of vectors in the fundamental system of solutions to

$$\begin{cases} x_1 + x_2 + x_3 = 0 \\ -2x_1 - x_2 - x_3 = 0 \\ 2x_1 + x_2 + x_3 = 0 \end{cases}$$

The rank of the corresponding matrix is 2, so the fundamental system of solutions contains only 1 vector. We substitute solutions in the form

$$y_1 = (x_1 + z_1 t)e^t \quad y_2 = (x_2 + z_2 t)e^t \quad (x_3 + z_3 t)e^t$$

into (26.11), find the fundamental system of solutions, then obtain

$$y = c_1 \left[\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} t \right] e^t + c_2 \left[\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} t \right] e^t$$

Combining with the solutions obtained from $\lambda_1 = 2$, we obtain the general solution to (26.11)

$$y_1 = c_2 e^t + c_3 e^{2t}$$

$$y_2 = -(c_1 + c_2)e^t - c_2 t e^t - 2c_3 e^{2t}$$

$$y_3 = c_1 e^t + c_2 t e^t + 2c_3 e^{2t}$$