SYLLABUS

1 Systems of linear equations

- Consistent and inconsistent systems. Equivalent systems. Elementary operations with systems. Matrix and augmented matrix of the system. Elementary row operations with matrices. Row echelon form and reduced row echelon form. Gaussian elimination. Homogeneous system. Closedness of the solutions set to a homogeneous system under addition and multiplication by scalars.

2 Matrices

- Rank of matrix as maximal number of linearly independent rows and maximal number of linearly independent columns. Rank is preserved under elementary row and column transformations. Computation of the rank of a matrix using Gaussian elimination.

  Theorem: Rank of a linear map equals to rank of its matrix (in some bases), i.e., $\mathrm{rank}(T) = \mathrm{rank}(A_T)$. Corollary: A matrix $A \in M_n(F)$ is invertible if and only if its rank is n.

  Computation of inverse matrix using Gaussian elimination.

  Theorem: For all $A, B \in M_n(F)$, $\mathrm{rank}(AB) \leq \min(\mathrm{rank}(A), \mathrm{rank}(B))$. Corollary: Two matrices $A, B \in M_n(F)$ are both invertible if and only if their product AB is invertible. Example of $\mathrm{rank}(AB) \neq \mathrm{rank}(BA)$.

  Dimension of solution set to homogeneous system of equation as the number of unknowns minus the rank of the system matrix.

- Permutations. Row notation. Cycle decomposition. Composition of permutations.

  Transposition. Adjacent transposition. Properties: (a) each permutation can be written as a composition of transpositions, $(i_1 \dots i_k) = (i_1 i_k)(i_1 i_{k-1}) \dots (i_1 i_3)(i_1 i_2)$, (b) each transposition can be written as a composition of an odd number of adjacent transpositions, $(ij) = (j-1\,j) \dots (i+1\,i+2)(i\,i+1) \dots (j-2\,j-1)(j-1\,j)$.

  Signature (or sign) of a permutation. Inversion pair.

  Theorem: $\mathrm{sgn}(\sigma) = (-1)^{\text{number of inversion pairs for } \sigma}$.

- Determinant of a matrix. Multilinear function. Alternating function.

  Proposition: If $f : V^n \to F$ is multilinear and alternating, then for any $\sigma \in S_n$, $f(v_{\sigma(1)}, \dots, v_{\sigma(n)}) = \mathrm{sgn}(\sigma) f(v_1, \dots, v_n)$.

  Theorem: Let $v_1, \dots, v_n$ be a basis of V. There exists a unique multilinear alternating function $f : V^n \to F$ with $f(v_1, \dots, v_n) = 1$.

  Theorem: Let $f : V^n \to F$ be multilinear and alternating, $v_1, \dots, v_n$ a basis of V, and $f(v_1, \dots, v_n) \neq 0$. Then any $u_1, \dots, u_n \in V$ are linearly independent if and only if $f(u_1, \dots, u_n) \neq 0$.

Determinant of a matrix as the unique multilinear alternating function of rows $\det : (F^n)^n \to F$ with $\det(e_1, \ldots, e_n) = 1$. Formulas for determinants for $n = 2, 3$.

Properties: (a) A is invertible iff $\det(A) \neq 0$, (b) $\det(AB) = \det(A) \cdot \det(B)$, (c) $\det(A)$ is the unique alternating multilinear function of columns of A such that $\det(I_n) = 1$, (d) determinant does not change if a row multiplied by a scalar is added to another row, (e) $\det(\alpha A) = \alpha^n \det(A)$.

Theorem (Laplace expansion): For any $i, j \in \{1, \ldots, n\}$, (a) $\det(A) = \sum_{k=1}^{n} (-1)^{k+j} \alpha_{kj} \det A(k|j)$ (expansion along j-th column), (b) $\det(A) = \sum_{k=1}^{n} (-1)^{i+k} \alpha_{ik} \det A(i|k)$ (expansion along i-th row).

Inverse matrix. Cramer's rule for solving systems of linear equations.

3 Invariant vector subspaces

- Change of basis. Proposition: Let $T : U \to V$ be a linear map. If $(e'_1, \ldots, e'_n) = (e_1, \ldots, e_n) \cdot P$ are two bases of U, $(f'_1, \ldots, f'_m) = (f_1, \ldots, f_m) \cdot Q$ two bases of V, then $A_T^{(e',f')} = Q^{-1} \cdot A_T^{(e,f)} \cdot P$. Corollary: For any $A \in M_{m,n}(F)$ with $\operatorname{rank}(A) = r$, there exist invertible matrices $P \in M_n(F)$ and $Q \in M_m(F)$ such that $Q^{-1} \cdot A \cdot Q = \begin{pmatrix} I_r & 0_{n-r,r} \\ 0_{m-r,r} & 0_{m-r,n-r} \end{pmatrix}$.

  Invariant subspace. Examples. Restriction of linear map to invariant subspace and its matrix.

  Eigenvectors and eigenvalues of linear operators. Spectrum of linear operator. Characteristic polynomial. Independence of characteristic polynomial of basis. Similar matrices. Theorem: $\lambda_0 \in \operatorname{Spec} T$ iff $P_T(\lambda_0) = 0$, i.e., eigenvalues are roots of the characteristic polynomial.

  Theorem: Eigenvectors corresponding to different eigenvalues are linearly independent.

  Simple spectrum. Theorem: If $T : U \to U$ has a simple spectrum, then there exists a basis $e_1, \ldots, e_n \in U$ such that $A_T = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. One says that T is diagonalizable.

  Examples of non-diagonalizable operators: (a) characteristic polynomial can be resolved not in every field, algebraically closed fields, (b) $A_T = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$.

  Algebraic and geometric multiplicities of eigenvalues. Eigenspace. Geometric multiplicity is at most algebraic multiplicity. Theorem: An operator (resp., its matrix) is diagonalizable if and only if for every eigenvalue of T, its geometric and algebraic multiplicities coincide.

4 Inner product spaces

- Definition of the inner product space. Euclidean and unitary spaces. Examples: F^n, $C([0,1], \mathbb{C})$. Norm of a vector. Cauchy-Schwarz inequality. Cauchy-Schwarz inequality in F^n and $C([0,1],\mathbb{C})$. Angle between two vectors in a Euclidean space. Triangle inequality for norms. Distance between two vectors. Properties of the distance.

- Orthogonality of two vectors, $u \perp v$. Properties of $\perp$. Orthogonal system of vectors. Orthonormal system of vectors.

  Theorem: Every orthogonal system is linearly independent. Corollary: If $\dim(V) = n$, then every orthogonal system contains at most n vectors.

  Gram-Schmidt orthogonalization. Theorem (Gram-Schmidt): If $u_1, \ldots, u_n \in V$ are linearly independent, then there exist $v_1, \ldots, v_n \in V$ orthonormal such that for all $k \leq n$, $\text{Span}\{v_1, \ldots, v_k\} = \text{Span}\{u_1, \ldots, u_k\}$. Corollary: Every inner product space has an orthonormal basis.

  Properties of ONBs. If $e_1, \ldots, e_n$ is an ONB of V, then (a) $\forall u \in V\ \exists! u_1, \ldots, u_n \in F$ such that $u = u_1 e_1 + \ldots + u_n e_n$, and $u_i = \langle u, e_i \rangle$, (b) for all $u = u_1 e_1 + \ldots + u_n e_n, v = v_1 e_1 + \ldots + v_n e_n \in V$, $\langle u, v \rangle = u_1 \overline{v}_1 + \ldots u_n \overline{v}_n$, (c) $\|u\|^2 = |u_1|^2 + \ldots + |u_n|^2$.

- Isomorphism of inner product spaces. Theorem: Any two finite dimensional inner product spaces with equal dimensions are isomorphic.

- Orthogonal complement of a set $\emptyset \neq S \subseteq V$, $S^\perp$. Properties of $\perp$: (a) $\{\overrightarrow{0}\}^\perp = V$, $V^\perp = \{\overrightarrow{0}\}$, (b) $S^\perp$ is a vector subspace of V (even if S is not), (c) $S^\perp = (\text{Span}S)^\perp$, (d) $(S^\perp)^\perp = \text{Span}S$.

  Theorem: If U is a vector subspace of V (V is finite dimensional), then $V = U \oplus U^\perp$. Corollary: Any vector $v \in V$ can be decomposed uniquely as $v = v_U + v_{U^\perp}$, where $v_U \in U$ and $v_{U^\perp} \in U^\perp$.

  Orthogonal projection, P_U. Properties: (a) P_U is linear, (b) $P_U^2 = P_U$, (c) $P_{U^\perp} = \text{id} - P_U$, (d) for all $v_1, v_2 \in V$, $\langle P_U(v_1), v_2 \rangle = \langle v_1, P_U(v_2) \rangle$, (e) $\text{Ker}P_U = U^\perp$, $\text{Im}P_U = U$. Example (projection on the line spanned by vector u): If $U = \text{Span}\{u\}$, then $P_U(v) = \frac{\langle v,u \rangle}{\|u\|^2} u$.

  Distance between sets $S_1, S_2 \subseteq V$, $\rho(S_1, S_2)$. Proposition: If U is a vector subspace of V and $v \in V$, then $\rho(v, U) = \|v_{U^\perp}\| = \|P_{U^\perp}(v)\|$.

- Dual space or the space of linear functionals on V, $V^* = \mathcal{L}(V, F)$. Example: for any $u \in V$, the map $f_u : V \to F$ defined by $f_u(v) = \langle v, u \rangle$ is in V^*. Theorem (Riesz representation theorem): If V is a finite dimensional inner product space, then for any $f \in V^*$ there exists unique $u \in V$ such that $f = f_u$, i.e., for all $v \in V$, $f(v) = \langle v, u \rangle$.

  Theorem: For any linear map $T : V \to V$ there exists unique linear map $T^* : V \to V$ such that for all $u, v \in V$, $\langle T(u), v \rangle = \langle u, T^*(v) \rangle$. T^* is the adjoint operator of T. Properties: (a) $(T^*)^* = T$, (b) $(\alpha T)^* = \overline{\alpha} T^*$, (c) $(T_1 + T_2)^* = T_1^* + T_2^*$, (d) $(T_1 T_2)^* = T_2^* T_1^*$, (e) if $e_1, \ldots, e_n$ is an ONB of V, then $A_{T^*} = (\overline{A}_T)^t$.

- Normal operator, $T^*T = TT^*$. Proposition: T is normal iff the matrix A_T of T in any ONB of V satisfies $A_T \overline{A}_T^t = \overline{A}_T^t A_T$.

Theorem: If T is normal and v is an eigenvector of T corresponding to an eigenvalue λ, then v is also an eigenvector of T^* corresponding to the eigenvalue $\overline{\lambda}$.

Canonical form of normal operator in unitary space. Theorem: If V is a finite dimensional unitary space and T is a linear operator on V, then T is normal iff there exists an ONB $e_1, \dots, e_n$ of V such that all e_i's are eigenvectors of T. (In particular, A_T is diagonal in this basis.)

Canonical form of normal operator in Euclidean space. Theorem: If V is a finite dimensional Euclidean space and T is a linear operator on V, then T is normal iff there exists an ONB of V such that the matrix of T in this basis has the block diagonal form with blocks either of size 1, $\lambda_i \in \mathbb{R}$, or of size 2, $r_i \begin{pmatrix} \cos\varphi_i & \sin\varphi_i \\ -\sin\varphi_i & \cos\varphi_i \end{pmatrix}$ with $r_i > 0$. (The blocks of size 1 correspond to deformation along corresponding basis vectors, the block of size 2 correspond to rotations of planes spanned by pairs of corresponding basis vectors by angle φ_i combined with deformation r_i.) Example: $A_T = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ - matrix of a normal operator corresponding to rotation clockwise by $\frac{\pi}{2}$.

- Self-adjoint operator, $T^* = T$. Example: P_U. Self-adjoint operators are normal. Hermitian matrix, $\overline{A}^t = A$, symmetric matrix, $A^t = A$. Proposition: T is self-adjoint iff its matrix is Hermitian in any ONB. Properties: If T_1, T_2 are self-adjoint then T_1^2, $T_1 + T_2$, αT_1, $T_1T_2 + T_2T_1$ are also self-adjoint, but T_1T_2 is not necessarily self-adjoint.

  Theorem: All eigenvalues of a self-adjoint operator are real. In particular, characteristic polynomial is real.

  Theorem: T is self-adjoint iff there exists an ONB in which A_T is diagonal and real.

- Skew-adjoint operator, $T^* = -T$. Skew-adjoint operators are normal.

  Theorem: All eigenvalues of a skew-adjoint operator are purely imaginary.

  Theorem: T is skew-adjoint iff there exists an ONB in which A_T is diagonal with purely imaginary entries.

  Proposition: T is skew-adjoint iff iT is self-adjoint.

  Theorem: For any $T \in \mathcal{L}(V, V)$, there exist unique T_1, T_2 self-adjoint such that $T = T_1 + iT_2$. Moreover, T is normal iff $T_1T_2 = T_2T_1$. Remark: Note the analogy with complex numbers: every complex number can be written in a unique way as the sum of a real number an a purely imaginary number.

- Unitary operator, $T^* = T^{-1}$. Unitary operators in Euclidean spaces are called orthogonal. Unitary operators are normal.

  Proposition: T is unitary iff for all $u, v \in V$, $\langle T(u), T(v) \rangle = \langle u, v \rangle$. Remark: The proposition states that unitary operators are isomorphisms of unitary spaces on themselves.

  Theorem: T is unitary iff T maps some (or all) ONB to ONB.

Unitary matrix, $\overline{A}^t = A^{-1}$, orthogonal matrix, $A^t = A$. Theorem: T is unitary iff in some ONB A_T is unitary. Proposition: If A is unitary, then its rows form an ONB in the space of rows $\mathbb{C}^n$, and its columns form an ONB in $\mathbb{C}^n$.

Theorem: If T is unitary, then all its eigenvalues have modulus 1, and $|\mathrm{det} A_T| = 1$.

Theorem: If T is unitary, then there exists an ONB such that A_T is diagonal with all entries on the diagonal being of modulus 1. If T is orthogonal, then there exists an ONB such that A_T is block diagonal with blocks of size 1 being either 1 or -1 and blocks of size 2 being $\begin{pmatrix} \cos\varphi_i & \sin\varphi_i \\ -\sin\varphi_i & \cos\varphi_i \end{pmatrix}$. (Canonical form of unitary/orthogonal operator.) Corollary: Every unitary (resp., orthogonal) matrix is unitary (resp., orthogonally) equivalent to a unitary (resp., orthogonal) matrix in the canonical form. Example: List of all 6 canonical forms of orthogonal operators in $\mathbb{R}^3$. (These are the only transformations of $\mathbb{R}^3$ which preserve lengths and angles between vectors.)

- Positive or positive-definite operator, $T > 0$. Non-negative or non-negative definite operator, $T \geq 0$. Properties: (a) if $T_1, T_2 \geq 0$, $\alpha_1, \alpha_2 \geq 0$ reals, then $\alpha_1 T_1 + \alpha_2 T_2 \geq 0$, (b) for any $T \in \mathcal{L}(V, V)$, $TT^* \geq 0$, (c) if $T = T^*$ then $T^2 \geq 0$, (d) if $T \geq 0$ then all its eigenvalues are non-negative ($\mathrm{spec}(T) \subseteq [0, +\infty)$), (e) if $T = T^*$ and all eigenvalues of T are non-negative, then $T \geq 0$, (f) if $T \geq 0$, then $T > 0$ iff $\mathrm{ker} T = \{\overrightarrow{0}\}$.

  Theorem (Square root of a non-negative operator): For any $T \geq 0$ there exists unique $S \geq 0$ such that $S^2 = T$, and $S > 0$ iff $T > 0$.

  Theorem (Polar decomposition): For any $T \in \mathcal{L}(V, V)$, there exist $R \geq 0$ and unitary operator S such that $T = RS$. Remark: R is uniquely defined. S is uniquely defined if T is non-degenerate, i.e., $\mathrm{ker} T = \{\overrightarrow{0}\}$. Remark: Note the analogy with trigonometric form of complex numbers $z = re^{i\varphi}$.

5 Bilinear forms

- Definition of bilinear form, $B(\cdot, \cdot)$. Examples: (a) $B(u, v) = \sum_{i,j=1}^n a_{i,j} u_i v_j$, $u, v \in F^n$, (b) $B(f, g) = \int_0^1 K(t) f(t) g(t) dt$, $K, f, g \in C[0, 1]$, (c) $B(u, v) = f(u) g(v)$, $u, v \in V$, $f, g \in V^*$, (d) inner product in Euclidean space. Coordinate representation: If $e_1, \ldots, e_n$ is a basis of V, then $B(u, v) = \sum_{i,j=1}^n u_i v_j B(e_i, e_j)$. Gram matrix of B: $A_B = (B(e_i, e_j))_{i,j=1}^n$. Change of basis: if $(e'_1 \ldots e'_n) = (e_1 \ldots e_n) \cdot C$ then $A'_B = C^t A_B C$. Congruent matrices. Prop: Two matrices are congruent iff they represent the same bilinear form in different bases. Example: second order curves in the plane.

- Symmetric bilinear form, $B(u, v) = B(v, u)$. Skew-symmetric bilinear form, $B(u, v) = -B(v, u)$. Example: in $\mathbb{R}^2$, $u_1 v_1 + u_2 v_2$ and $u_1 v_2 - u_2 v_1$. Proposition: If B is symmetric then its matrix in any basis is symmetric, $A_B = A_B^t$.

  Theorem (Sylvester's law of intertia): (1) If B is a symmetric bilinear form in a unitary space V, then there exists a basis of V in which the matrix of B is diagonal with only 1's and 0's on the diagonal, namely, there exists $e_1, \ldots, e_n$ basis of V and $r \leq n$ such that $B(e_i, e_j) = \begin{cases} 1 & i = j \leq r \\ 0 & i = j > r \\ 0 & i \neq j \end{cases}$. In this basis $B(u, v) = u_1 v_1 + \ldots + u_r v_r$.

(2) If B is a symmetric bilinear form in a Euclidean space V, then there exists a basis of V in which the matrix of B is diagonal with only 1's, -1's, and 0's on the diagonal, namely, there exists $e_1, \ldots, e_n$ basis of V and $r < s \leq n$ such that $B(e_i, e_j) = \begin{cases} 1 & i = j \leq r \\ -1 & r < i = j \leq s \\ 0 & i = j > s \\ 0 & i \neq j \end{cases}$. In this basis $B(u, v) = u_1 v_1 + \ldots + u_r v_r - u_{r+1} v_{r+1} - \ldots - u_s v_s$.
Moreover, the number of 1's, -1's, and 0's is independent of the basis in which the matrix is of the above form. Number of 1's minus number of -1's is the signature of B.

- Quadratic form associated to a symmetric bilinear form B, $B(v, v)$. Theorem: Any quadratic form is associated to a unique symmetric bilinear form. (Here V is either Euclidean or unitary.) Polarization of the quadratic form: $B(u, v) = \frac{1}{2}\left(B(u + v, u + v) - B(u, u) - B(v, v)\right)$.

- Definition of sesquilinear form. Example: inner product in unitary space. Hermitian form (or symmetric sesquilinear form), $B(u, v) = \overline{B(v, u)}$. Coordinate representation as for bilinear forms, A_B. Remark: B is Hermitian iff A_B is Hermitian (in any basis). Change of basis: if $(e'_1 \ldots e'_n) = (e_1 \ldots e_n) \cdot C$ then $A'_B = C^t A_B \overline{C}$.

  Theorem: If B is a Hermitian form in a unitary space V, then there exists a basis of V in which the matrix of B is diagonal with only 1's, -1's, and 0's on the diagonal, namely, there exists $e_1, \ldots, e_n$ basis of V and $r < s \leq n$ such that $B(e_i, e_j) = \begin{cases} 1 & i = j \leq r \\ -1 & r < i = j \leq s \\ 0 & i = j > s \\ 0 & i \neq j \end{cases}$. In this basis $B(u, v) = u_1 \overline{v}_1 + \ldots + u_r \overline{v}_r - u_{r+1} \overline{v}_{r+1} - \ldots - u_s \overline{v}_s$.
  Moreover, the number of 1's, -1's, and 0's is independent of the basis in which the matrix is of the above form.

- Quadratic form associated to a Hermitian form B, $B(v, v)$. Any quadratic form is associated to a unique Hermitian form B. Polarization of the quadratic form: $B(u, v) = \frac{1}{2}\left(B(u + v, u + v) - B(u, u) - B(v, v)\right) + \frac{i}{2}\left(B(u + iv, u + iv) - B(u, u) - B(v, v)\right)$.

- Positive definite quadratic form, $B(v, v) > 0$ for $v \neq 0$. Remark: B is positive definite iff there exists a basis in which A_B (matrix of the associated bilinear form) is the identity matrix.

  Proposition: B is positive definite quadratic form on V iff its polarization B (symmetric bilinear form if V is Euclidean, resp., Hermitian if V is unitary) defines an inner product on V. Thus, bilinear and sesquilinear forms are generalizations of inner products in Euclidean and, resp., unitary spaces. Example: Minkowski space (over $\mathbb{R}$), $B(v,v) = v_1^2 + v_2^2 + v_3^2 - v_4^2$.

6 Other algebraic structures

- Semigroup. Examples: (a) $(\mathbb{N}, +)$, (b) all maps from a set X to itself, $M(X)$, with the operation of composition of maps $\circ$, (c) set of all subsets of a set X with operation of intersection (or union), $(P(X), \cap)$, $(P(X), \cup)$.
- Monoid, $(S, \cdot, e)$. Examples: (a) $(M(X), \circ, \mathrm{id})$, (b) $(P(X), \cap, X)$ and $(P(X), \cup, \emptyset)$.
- Group. Uniqueness of the inverse element, a^{-1}. Commutative (or Abelian) group. Examples: (a) $(\mathbb{Z}, +, 0)$ is Abelian, (b) $(\mathbb{Z}, \cdot)$ is not a group, (c) $(\mathbb{Q}, \cdot)$ is not a group, but $(\mathbb{Q} \setminus \{0\}, \cdot, 1)$ is Abelian group, (d) the group of integers modulo $n \in \mathbb{N}$, $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$, (e) cyclic group, $a^n, n \in \mathbb{Z}$, (f) general linear group, $GL(n, \mathbb{R})$ (all invertible matrices with operation of multiplication), (g) special linear group, $SL(n, \mathbb{R})$ (all matrices with determinant 1 with operation of multiplication), (h) orthogonal group, $O(n)$ (all distance preserving linear transformations of $\mathbb{R}^n$ with operation of composition, or, equivalently, all orthogonal $n \times n$ matrices with real entries with operation of multiplication), (i) special orthogonal group or rotation group, $SO(n)$ (all orthogonal matrices with determinant 1 with operation of multiplication), (j) symmetric group of degree n, S_n (group of permutations of $\{1, \ldots, n\}$).
- Subgroup. Example: $SL(n, \mathbb{R})$ is a subgroup of $GL(n, \mathbb{R})$. Isomorphism of two groups, $(G, *) \cong (G', \circ)$. Proposition: If φ is an isomorphism from $(G, *, e)$ to $(G', \circ, e')$, then (a) $\varphi(e) = e'$, (b) for all $a \in G$, $\varphi(a^{-1}) = \varphi(a)^{-1}$, (c) φ^{-1} is an isomorphism from $(G', \circ, e')$ to $(G, *, e)$.
- Ring. Examples of rings: (a) $(\mathbb{Q}, +, \cdot)$ and $(\mathbb{Z}, +, \cdot)$, (b) $(M_{n,n}(\mathbb{R}), +, \cdot)$, (c) if $(\mathbb{R}, +, *)$, where $*$ is defined as $x * y = 0$ for all $x, y \in \mathbb{R}$.

7 Functions of several variables

- Definition of $y = f(x_1, \ldots, x_n)$. Other notation: $z = f(x, y)$, $w = f(x, y, z)$. $\|x\| = \sqrt{x_1^2 + \ldots + x_n^2}$. Open ball $B(x, r)$. Limit of f at $a \in \mathbb{R}^n$, $\lim_{x \to a} f(x)$. (Mind: $f(a)$ may be undefined.) Properties of the limit.

  Proposition: Let $u_1 = u_1(t), \ldots, u_n = u_n(t)$ be real functions defined on an interval containing $t_0 \in \mathbb{R}$. If for all $i \in \{1, \ldots, n\}$, $\lim_{t \to t_0} u_i(t) = a_i$ and $\lim_{x \to a} f(x) = L$, then $\lim_{t \to t_0} f(u_1(t), \ldots, u_n(t)) = L$. (This is useful in proving non-existence of the limit: if for two different choices of u_i's, the corresponding limits $\lim_{t \to t_0} f(u_1(t), \ldots, u_n(t))$ are different, then the limit of f at a does not exist.)

  Examples: (a) $f(x, y) = \frac{x^2 y}{x^4 + y^2}$, for $x^2 + y^2 > 0$ (limit at 0 along every line is 0, but the limit at 0 does not exist), (b) $f(x, y) = \frac{xy}{x^2 + y^2}$, for $x^2 + y^2 > 0$.

Proposition: If $f(x_1, \ldots, x_n) = g(\|x\|)$ for some $g : \mathbb{R} \to \mathbb{R}$, then $\lim_{x\to 0} f(x) = \lim_{r\to 0+} g(r)$. Such f's are called isotropic. Example: $\lim_{(x,y)\to(0,0)}(x^2+y^2)\ln(x^2+y^2) = \lim_{r\to 0+} r\ln r = 0$.

- Continuous function at $a \in \mathbb{R}^n$. Continuous function on D. Examples: (a) polynomials, (b) rational functions (in their domains). Bounded set. Closed set. Examples: (a) $B(x,r)$ bounded but not closed, (b) $\overline{B}(x,r)$ bounded and closed, (c) $S(x,r)$ bounded and closed.

  Theorem: If a function is continuous on a bounded and closed set D, then it is bounded on D and attains its maximal and minimal values at some points of D.

  Connected set. Theorem: If f is continuous on a connected set D, then for all $x, y \in D$, f takes any value between $f(x)$ and $f(y)$ at some points of D.

8 Differential calculus of functions of several variables

- Partial derivative of f with respect to x_i at $a \in \mathbb{R}^n$, $\frac{\partial f}{\partial x_i}(a) = f_{x_i}(a) = f'_{x_i}(a) = \frac{\partial}{\partial x_i} f(a)$. Special notation in 2 and 3 dimensions, e.g., f_x, f_y, w_z. Properties of partial derivative with respect to x_i.

- Geometric interpretation of f_x in two dimensions. Tangent plane to $z = f(x,y)$ at (a,b), $z = f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b)$. Examples: (a) $f(x,y) = \begin{cases} 1 & xy \neq 0 \\ 0 & xy = 0 \end{cases}$ (the tangent plane at $(0,0)$ is $z = 0$), (b) $f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & x^2+y^2 > 0 \\ 0 & x = y = 0 \end{cases}$ ($f_x(a,b)$ and $f_y(a,b)$ exist for all $(a,b) \in \mathbb{R}^2$, but f is discontinuous at 0).

- Differentiable function at $a \in \mathbb{R}^n$. Theorem: If all $f_{x_i}(a)$ are continuous at a, then f is differentiable at a. Continuously differentiable function. Example: $f(x,y) = \begin{cases} (x^2+y^2)\sin\frac{1}{x^2+y^2} & x^2+y^2 > 0 \\ 0 & x = y = 0 \end{cases}$ is differentiable at $(0,0)$, but f_x, f_y are not continuous. Remark: If f is differentiable at a, then f is continuous at a.

- Chain rule. Theorem: If f is a differentiable function of $x_1, \ldots, x_n$, where $x_i = x_i(t_1, \ldots, t_m)$ are functions of $t_1, \ldots, t_m$ such that $\frac{\partial x_i}{\partial t_j}$ exist for all i, j, then there exist $\frac{\partial f}{\partial t_j}$ for all j, and $\frac{\partial f}{\partial t_j} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial t_j} + \ldots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial t_j}$. Remark: Differentiability of f is crucial, although only existence of $\frac{\partial f}{\partial x_i}$ appears in the chain rule.

- Directional derivative, $\frac{\partial f}{\partial \ell}(a) = D_\ell f(a)$. Geometric interpretation.

  Theorem: If f is differentiable at a, then for any direction ℓ, there exists $D_\ell f(a)$, and $D_\ell f(a) = (f_{x_1}(a), \ldots, f_{x_n}(a)) \cdot \ell$. Gradient of f, $\nabla f(x)$.

Theorem: If f is differentiable at a, then the maximum of $D_\ell f(a)$ over all unit vectors ℓ equals $\|\nabla f(a)\|$ and it is attained by a vector with the same direction as $\nabla f(a)$. In particular, the gradient of f does not depend on the choice of coordinates.

- Higher order derivatives, $f_{x_i x_j} = (f_{x_i})_{x_j} = \frac{\partial^2}{\partial x_j \partial x_i} f = \frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial}{\partial x_j}(\frac{\partial f}{\partial x_i})$, $f_{x_i x_i} = \frac{\partial^2 f}{\partial x_i^2}$, etc. Special notation in 2 and 3 dimensions.

  Theorem: If f is defined in a ball around a, and $f_{x_i x_j}$, $f_{x_j x_i}$ are continuous at a, then $f_{x_i x_j}(a) = f_{x_j x_i}(a)$. Example: $f(x,y) = \begin{cases} xy\frac{x^2-y^2}{x^2+y^2} & x^2+y^2>0 \\ 0 & x=y=0 \end{cases}$ $(f_{xy}(0,0) = -1 \neq 1 = f_{yx}(0,0))$.

- Implicit differentiation. If $F(x_1, \dots, x_n, y) = 0$, then for all i, $\frac{\partial F}{\partial x_i} + \frac{\partial F}{\partial y}\frac{\partial y}{\partial x_i} = 0$, i.e., $\frac{\partial y}{\partial x_i} = -\frac{F_{x_i}}{F_y}$.

- Taylor's formula. Multiindex $k = (k_1, \dots, k_n)$. $|k| = k_1 + \ldots + k_n$, $k! = k_1! \cdot \ldots \cdot k_n!$, $f^{(k)} = \frac{\partial^{|k|} f}{\partial x_1^{k_1} \ldots \partial x_n^{k_n}}$, $(x-a)^k = (x_1-a_1)^{k_1} \cdot \ldots \cdot (x_n - a_n)^{k_n}$.

  Theorem: If f is continuously differentiable to order m in a ball around a, then $f(x) = \sum_{k:|k|<m} \frac{f^{(k)}(a)}{k!}(x-a)^k + \sum_{k:|k|=m} \frac{f^{(k)}(a+\theta(x-a))}{k!}(x-a)^k$, where $\theta \in (0,1)$ is a parameter dependent on a and x. (Lagrange remainder.) Example: Taylor's formula in 2 and 3 dimensions.

  Open set. Convex set.

  Theorem (Lagrange's mean value theorem): If f is differentiable in an open convex set G, then for all $x, y \in G$ there exists $\theta \in (0,1)$ such that $f(y) - f(x) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x + \theta(y-x))(y-x) = \nabla f(x + \theta(y-x)) \cdot (y-x)$.

  Corollary: If all partial derivatives of f equal to 0 everywhere in G, then f is constant in G.

9 Extrema of functions of several variables

- Local extrema. Local minimum and local maximum. Strict local minimum and local maximum.

  Theorem (necessary condition): If f has local extremum at $a \in \mathbb{R}^n$ and f_{x_i} exists for some i, then $f_{x_i}(a) = 0$.

  Critical point. Remark: If f has local extremum at a, then a is critical, but not vice versa. Saddle point. Example: $f(x,y) = y^2 - x^2$.

- Positive and negative definite symmetric quadratic forms, $Q(z) = \sum_{i,j=1}^n a_{ij} z_i z_j$.

  Theorem (sufficient condition): If f is twice continuously differentiable in a ball around a and $f_{x_i}(a) = 0$ for all i, then f attains a local maximum (resp., minumum) at a if the (symmetric) quadratic form $Q(z) = \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(a) z_i z_j$ is negative (resp., positive) definite. If Q takes positive and negative values, then f does not have local extremum at a.

Examples: (a) if $n=1$, then $Q(z)=\frac{\partial^2 f}{\partial x_1^2}z_1^2$, and we recover the known sufficient condition for real functions, (b) if $n=2$, then Q is positive definite iff $\frac{\partial^2 f}{\partial x_1^2}\frac{\partial^2 f}{\partial x_2^2}-(\frac{\partial^2 f}{\partial x_1\partial x_2})^2>0$ and $\frac{\partial^2 f}{\partial x_1^2}>0$, and negative definite iff $\frac{\partial^2 f}{\partial x_1^2}\frac{\partial^2 f}{\partial x_2^2}-(\frac{\partial^2 f}{\partial x_1\partial x_2})^2>0$ and $\frac{\partial^2 f}{\partial x_1^2}<0$.

Hessian matrix of f, $(\frac{\partial^2 f}{\partial x_1\partial x_2})_{i,j=1}^n$.

- Sylvester's criterion: A symmetric quadratic form $Q(z)=\sum_{i,j=1}^n a_{ij}z_iz_j$ is positive definite iff $a_{11}>0$, $\begin{vmatrix} a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}>0$, $\begin{vmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix}>0$, …, $\det A>0$.

  Remark: Q is negative definite iff $-Q$ is positive definite. Thus, Q is negative definite iff $a_{11}<0$, $\begin{vmatrix} a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}>0$, $\begin{vmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix}<0$, …

  Examples: (a) $f(x,y)=x^4+y^4-4xy+1$, (b) $f(x,y)=x^2+2xy+y^2$, (c) $f(x,y)=xy^3$. (In (b) and (c) $f_{xx}f_{yy}-(f_{xy})^2=0$.)

- Finding extrema of a function on a bounded closed set. Example: Find global maximum and minimum of $f(x,y)=x^2-2xy+2y$ on $\{(x,y):0\le x\le 3, 0\le y\le 2\}$.

- Jacobi matrix of $g_1(x_1,\dots,x_n),\dots,g_m(x_1,\dots,x_n)$, $(\frac{\partial g_i}{\partial x_j})_{i,j}$. Jacobian, $\frac{\partial(g_1,\dots,g_n)}{\partial(x_1,\dots,x_n)}$.

- Method of Lagrange multipliers.

  Theorem: Let $f,g_1,\dots,g_m$ be continuously differentiable on an open set G and the rank of the Jacobi matrix $\begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \dots & \frac{\partial g_1}{\partial x_n}\\ & \dots & \\ \frac{\partial g_m}{\partial x_1} & \dots & \frac{\partial g_m}{\partial x_n}\end{pmatrix}$ equals m for all $x\in G$. Let $F(x)=f(x)-\lambda_1 g_1(x)-\dots\lambda_m g_m(x)$. If f has local extremum subject to constraints $g_1(x)=\dots=g_m(x)=0$ at point a, then there exist $\lambda_1,\dots,\lambda_m\in\mathbb{R}$, such that $\frac{\partial F}{\partial x_i}(a)=0$ for all i.

  Remark: condition on the Jacobi matrix means that none of the constraints $g_i(x)=0$ follows from the others. $\lambda_1,\dots,\lambda_m$ are Lagrange multipliers. F is Lagrangian.

  Strategy: (a) check the rank condition, (b) write down the Lagrangian, (c) find all solutions $x_1,\dots,x_n,\lambda_1,\dots,\lambda_m$ of the system of equations $\begin{cases}\frac{\partial F}{\partial x_i}(x)=0 & i\in\{1,\dots,n\}\\ g_j(x)=0 & j\in\{1,\dots,m\}\end{cases}$, (d) evaluate f at all found $(x_1,\dots,x_n)$.

  Examples: (a) maximum of $f(x,y,z)=x+2y+3z$ on the curve $\begin{cases} x-y+z=1\\ x^2+y^2=1\end{cases}$, (b) maximum of $f(x,y)=xy$ given $x+y-p=0$, (c) maximum and minimum of $f(x_1,\dots,x_n)=\sum_{i,j=1}^n a_{ij}x_ix_j$ with $a_{ij}=a_{ji}$ given $x_1^2+\dots+x_n^2=1$.

- Examples of change of variables:

(a) $A = (\frac{\partial f}{\partial x_1})^2 + \ldots + (\frac{\partial f}{\partial x_n})^2$ and $B = \frac{\partial^2 f}{\partial x_1^2} + \ldots + \frac{\partial^2 f}{\partial x_n^2}$ do not depend on the choice of ONB,

(b) For $z = f(x,y)$, expressions for $A = (z_x)^2 + (z_y)^2 = \|\nabla f\|^2$ and $B = z_{xx} + z_{yy}$ (Laplacian of f) in polar coordinates: $A = (z_r)^2 + \frac{1}{r^2}(z_\varphi)^2$, $B = z_{rr} + \frac{1}{r^2} z_{\varphi\varphi} + \frac{1}{r} z_\varphi$,

(c) solve $yz_x - xz_y = 0$ using polar coordinates,

(d) linear transport equation, $u_t + cu_x = 0$, substitute $\xi = t$, $\eta = x - ct$,

(e) wave equation, $u_{tt} = a^2 u_{xx}$, substitute $\xi = x - at$, $\eta = x + at$.

10 Uniform convergence

- Interchange of limits, integrals, and derivatives is not always allowed:

  (a) $f_n(x) = \begin{cases} 1 & x > \frac{1}{n} \\ nx & x \in [0, \frac{1}{n}] \\ 0 & x \le 0 \end{cases}$ are continuous, but the limit is not,

  (b) $f_n(x) = \begin{cases} 1 & x \in [0,1], x = \frac{p}{q}, q \le n \\ 0 & \text{else} \end{cases}$ are integrable, but the limit is not,

  (c) $f_n(x) = \begin{cases} nx & x \in [0, \frac{1}{n}] \\ n(\frac{2}{n} - x) & x \in [\frac{1}{n}, \frac{2}{n}] \\ 0 & \text{else} \end{cases}$ converge to 0, but their integrals do not converge to 0,

  (d) $f_n(x) = |x|^{1+\frac{1}{n}}, x \in [-1,1]$, differentiable at 0, but the limit is not,

  (e) $f_n(x) = \frac{1}{n} \sin nx, x \in [0, \pi]$ converge to 0, but their derivatives do not converge to 0.

- Uniform convergence of $f_n(x)$ to f on E, $f_n(x) \overset{E}{\rightrightarrows} f(x)$ as $n \to \infty$. Example: $f_n(x) = x^n$ converge uniformly on $[0,q]$ for $q \in (0,1)$, but not on $[0,1]$.

  Properties: (a) if f_n are continuous on E and $f_n \overset{E}{\rightrightarrows} f$ as $n \to \infty$, then f is continuous,

  (b) if f_n are integrable on $[a,b]$ and $f_n \overset{[a,b]}{\rightrightarrows} f$ as $n \to \infty$, then f is integrable on $[a,b]$ and $\int_a^b f(x)dx = \lim_{n\to\infty} \int_a^b f_n(x)dx$,

  (c) if f_n are differentiable on $[a,b]$, $f_n \to \varphi$ and $f_n' \rightrightarrows \psi$ on $[a,b]$ as $n \to \infty$, then $\varphi' = \psi$.

  Cauchy criterion of uniform convergence: $f_n \rightrightarrows f$ on E as $n \to \infty$ iff for any $\varepsilon > 0$ there exists N such that for all $n \ge N$ and $m > 0$, $\sup_{x\in E} |f_{n+m}(x) - f_n(x)| < \varepsilon$.

  Theorem (Weierstrass, sufficient condition): If there exists a sequence a_n such that $\lim_{n\to\infty} a_n = 0$ and $\sup_{x\in E} |f_n(x) - f(x)| \le a_n$, then $f_n \rightrightarrows f$ on E as $n \to \infty$.

- Uniform convergence of series. Example: $\sum_{k=1}^\infty \frac{\sin kx}{k^2}$.

Cauchy criterion: $\sum_{n=1}^{\infty} u_n(x)$ converges uniformly on E iff $\forall\varepsilon > 0\ \exists N$ such that $\forall n \geq N, m > 0$, $\sup_{x\in E} |\sum_{k=n+1}^{n+m} u_k(x)| < \varepsilon$.

Weierstrass: if $\exists a_n$ such that $\sup_{x\in E} |u_n(x)| \leq a_n$ and $\sum_{n=1}^{\infty} a_n < \infty$ then $\sum_{n=1}^{\infty} u_n(x)$ converges uniformly.

If u_n are continuous and $\sum_{n=1}^{\infty} u_n(x)$ converges uniformly, then $s(x) = \sum_{n=1}^{\infty} u_n(x)$ is also continuous.

If u_n are integrable on $[a, b]$ and $\sum_{n=1}^{\infty} u_n(x)$ converges uniformly, then $\int_a^b s(x)dx = \sum_{n=1}^{\infty} \int_a^b u_n(x)dx$.

If u_n are differentiable on $[a, b]$, $\sum_{n=1}^{\infty} u_n(x)$ converges, and $\sum_{n=1}^{\infty} u_n'(x)$ converges uniformly, then $s'(x) = \sum_{n=1}^{\infty} u_n'(x)$.

Dirichlet test: If $u_n(x) = a_n(x)b_n(x)$, where (a) $a_n(x) \rightrightarrows 0$ on E and for each $x \in E$, $a_n(x)$ is a non-increasing sequence, (b) there exists C such that $\sup_{x\in E} |\sum_{k=1}^{n} b_k(x)| \leq C$ for all n, then $\sum_{n=1}^{\infty} u_n(x)$ converges uniformly.

Examples: (a) $\sum_{n=1}^{\infty} \frac{x^n}{n!}$ converges uniformly on $[-R, R]$, but not on $\mathbb{R}$, (b) $\sum_{n=1}^{\infty} e^{-n^5x^2} \sin nx$ converges uniformly on $\mathbb{R}$, (c) Dirichlet series $\sum_{n=1}^{\infty} \frac{\sin nx}{n^\alpha}$ converges uniformly on $\mathbb{R}$ for $\alpha > 1$, converges uniformly on any $[a, b]$ which does not contain $2\pi k$, converges on $\mathbb{R}$ but not uniformly.

Remark: Properties of uniform convergence of series are used in explicit calculations. Examples: $\sum_{n=1}^{\infty} nq^n$, $\sum_{n=1}^{\infty} \frac{q^n}{n}$, Taylor series for $\ln(1+x)$ and $\arctan x$. Approximate calculations of integrals using Taylor series expansion, $\int_0^T \frac{\sin x}{x} dx$.

- Uniform convergence of functions of several variables, $f(x, y) \rightrightarrows \varphi(x)$ on $X \subseteq \mathbb{R}^n$ as $y \to y_0 \in \mathbb{R}^m$.

  Cauchy: $f(x, y) \rightrightarrows \varphi(x)$ on X as $y \to y_0$ iff $\forall\varepsilon > 0\ \exists\delta > 0$ such that for all y_1, y_2 with $0 < \|y_1 - y_0\|, \|y_2 - y_0\| < \delta$, $\sup_{x\in X} |f(x, y_1) - f(x, y_2)| < \varepsilon$.

  Weierstrass: If there exists $g(y)$ such that $\lim_{y\to y_0} g(y) = 0$ and $\sup_{x\in S} |f(x, y) - \varphi(x)| \leq g(y)$, then $f(x, y) \rightrightarrows \varphi(x)$ on X as $y \to y_0$.

  Conditions for interchange of limits, limit and integral, limit and derivative can be formulated similarly as in previous situations, but it is often easier to use joint continuity:

  Jointly continuous function $f(x, y)$, $x \in X \subset \mathbb{R}^n$, $y \in Y \subset \mathbb{R}^m$. Example: $f(x, y) = \begin{cases} 1 & xy \neq 0 \\ 0 & xy = 0 \end{cases}$ is continuous in x and in y at 0, but not jointly continuous at 0.

  Theorem: Let $f(x, y)$ be defined on $X \times Y$, where X is closed and bounded set in $\mathbb{R}^n$ (e.g., $X = [a_1, b_1] \times \ldots \times [a_n, b_n]$) and Y is closed and bounded set in $\mathbb{R}^m$ (e.g., $Y = [c_1, d_1] \times \ldots \times [c_m, d_m]$). If f is jointly continuous on $X \times Y$, then for any $y_0 \in Y$, $f(x, y) \overset{X}{\rightrightarrows} f(x, y_0)$ as $y \to y_0$.

  Corollaries: (a) If $f(x, y)$ is jointly continuous on $X \times Y$, where $X = [a, b]$ and Y is closed and bounded in $\mathbb{R}^m$, then $I(y) = \int_a^b f(x, y)dx$ is continuous on Y, i.e., for all $y_0 \in Y$, $\lim_{y\to y_0} \int_a^b f(x, y)dx = \int_a^b f(x, y_0)dx = \int_a^b (\lim_{y\to y_0} f(x, y))dx$.

(b) If $f_y(x,y)$ is jointly continuous on $[a,b]\times[c,d]$, then $I'(y)=\int_a^b f_y(x,y)dx$.

(c) If $f(x,y)$ is jointly continuous on $[a,b]\times[c,d]$, then $\int_c^d I(y)dy = \int_a^b(\int_c^d f(x,y)dy)dx$.

(d) Let $J(y)=\int_{a(y)}^{b(y)} f(x,y)dx$, where $a\le a(y), b(y)\le b$, there exist $a'(y), b'(y)$ for all $c\le y\le d$. If $f_y(x,y)$ is jointly continuous on $[a,b]\times[c,d]$, then $J'(y) = \int_{a(y)}^{b(y)} f_y(x,y)dx + f(b(y),y)b'(y) - f(a(y),y)a'(y)$.

Examples: (a) $I(y)=\int_0^1 \ln(x^2+y^2)dx$ is continuous for all $y>0$ and $I'(y) = 2\arctan\frac{1}{y}$. However, $I'(0)=\pi\neq 0=\int_0^1 \frac{d}{dy}\ln(x^2+y^2)|_{y=0}dx$.

(b) $f(x,y)=x^y$ on $[0,1]\times[a,b]$, $0<a<b$. $\int_0^1 \frac{x^b-x^a}{\ln x}dx=\ln\frac{1+b}{1+a}$.

(c) for $f(x,y) = \begin{cases} \frac{y^2-x^2}{(x^2+y^2)^2} & x,y\in[0,1], x^2+y^2>0 \\ 0 & x=y=0 \end{cases}$ (f is discontinuous at 0), $\int_0^1(\int_0^1 f(x,y)dx)dy=\frac{\pi}{4}\neq-\frac{\pi}{4}=\int_0^1(\int_0^1 f(x,y)dy)dx$.

(d) $J(y)=\int_0^y \frac{(y-x)^{n-1}}{(n-1)!}f(x)dx$, $J^{(n)}(y)=f(y)$.

11 Metric, normed, and Hilbert spaces

- Metric space, (X,ρ). Examples: (a) $X=\mathbb{R}$, $\rho(x,y)=|x-y|$, (b) $X=\mathbb{C}$, $\rho(x,y)=|x-y|$, (c) $X=\mathbb{R}^n$, $\rho(x,y)=\sqrt{(x_1-y_1)^2+\ldots+(x_n-y_n)^2}$, (d) $X=\mathbb{R}^n$, $\rho(x,y)=\sum_{i=1}^n|x_i-y_i|$, (e) $X=\mathbb{R}^n$, $\rho(x,y)=\max_{1\le i\le n}|x_i-y_i|$, (f) $X=C_{[a,b]}$, $\rho(f,g)=\max_{x\in[a,b]}|f(x)-g(x)|$, (g) $X=C^k_{[a,b]}$, $\rho(f,g)=\sum_{i=0}^k\max_{x\in[a,b]}|f^{(i)}(x)-g^{(i)}(x)|$, (h) $X=\ell_2$, $\rho(x,y)=\sqrt{\sum_{i=1}^\infty(x_i-y_i)^2}$, (i) $X=\ell_1$, $\rho(x,y)=\sum_{i=1}^\infty|x_i-y_i|$, (j) $X=\ell_\infty$, $\rho(x,y)=\sup_{i\ge1}|x_i-y_i|$.

  Open ball, $B(a,r)$. Examples: open ball in $\mathbb{R}^2$ with metric from (c)-(e).

  Closure of S in (X,ρ), $\overline{S}$. Closed ball. Properties: (a) $S\subseteq\overline{S}$, (b) $\overline{\overline{S}}=\overline{S}$, (c) if $S_1\subseteq S_2$, then $\overline{S}_1\subseteq\overline{S}_2$, (d) $\overline{S_1\cup S_2}=\overline{S}_1\cup\overline{S}_2$.

  Closed set. Open set. Remarks: (a) S is closed in (X,ρ) iff $X\setminus S$ is open, (b) X and $\emptyset$ are both open and closed in (X,ρ), (c) intersections and unions of open sets (closed sets).

  Limit of sequence $x_n\in X$. Examples: (a) $\mathbb{R}$, (b) $C_{[a,b]}$. Continuous function on (X,ρ). Cauchy sequence.

  Complete metric space. Examples: (a) $\mathbb{R}$, (b) $\mathbb{R}^n$ (with any of (c-e) metrics), (c) $C_{[a,b]}$, (d) ℓ_2.

  Compact metric space. Remark: if X is compact, then it is closed and bounded. Proposition: $X\subseteq\mathbb{R}^d$ is compact iff X is closed and bounded in $\mathbb{R}^d$. Examples: (a) $\overline{B}(a,r)$ and $S(a,r)$ are compacts in $\mathbb{R}^d$, (b) $\overline{B}(0,1)$ is not a compact in ℓ_2 (although it is closed and bounded).

  Theorem: If f is a continuous function on a compact S, then it is bounded and attains its maximal and minimal values at some points of S.

- Normed (vector) space. Every normed space is a metric space, $\rho(x,y) = \|x-y\|$. Examples: (a) $\mathbb{R}^n$, $\|x\| = \sqrt{x_1^2+\ldots+x_n^2}$, (b) $\mathbb{R}^n$, $\|x\| = \sum_{i=1}^n |x_i|$, (c) $\mathbb{R}^n$, $\|x\| = \max_{1\le i\le n}|x_i|$, (d) $C_{[a,b]}$, $\|f\| = \max_{x\in[a,b]}|f(x)|$, (e) ℓ_p, $\|x\| = (\sum_{i=1}^\infty |x_i|^p)^{\frac{1}{p}}$, $p\in[1,\infty]$.

  Equivalence of norms. Theorem: If X is a finite dimensional normed space, then any two norms on it are equivalent. Example: In $\mathbb{R}^n$ all the three norms (a-c) are equivalent. Remark: If two norms are equivalent, then (a) x_n converges to x in one norm iff x_n converges to x in the other norm, (b) if f is continuous with respect to one norm, then it is continuous with respect to the other norm.

  Banach space, complete normed space. Examples: (a) $\mathbb{R}^n$, (b) $C_{[a,b]}$, (c) ℓ_2, (d) any finite dimensional closed space.

- Hilbert space, complete vector space with an inner product. Hilbert space is a normed space, $\|x\| = \sqrt{\langle x,x\rangle}$.

  Theorem: A normed vector space X over $\mathbb{R}$ admits inner product iff the parallelogramm identity holds, i.e., for all $x,y\in X$, $\|x+y\|^2+\|x-y\|^2 = 2(\|x\|^2+\|y\|^2)$. In this case, one defines the inner product as $\langle x,y\rangle = \frac{1}{4}(\|x+y\|^2-\|x-y\|^2)$. Example: ℓ_p, $p\in[1,\infty]$ is Hilbert iff $p=2$.

  Separable Hilbert space. Example: $\{e^{i\alpha x}\}_{\alpha\in\mathbb{R}}$ is an uncountable basis of $L = \{y(x) = \sum_{j=1}^n a_j e^{i\alpha_j x}, n\in\mathbb{N}\}$. Completion of L is a non-separable Hilbert space.

  Theorem (Gram-Schmidt): Any separable Hilbert space has an ONB.

  Properties of ONB: If $\{e_i\}_{i\ge1}$ is an ONB, then for any $x\in X$, $x = \sum_{i=1}^\infty x_i e_i$ (i.e., $\|x-\sum_{i=1}^n x_i e_i\|\to 0$ as $n\to\infty$), where $x_i = \langle x, e_i\rangle$ (Fourier coefficients). Fourier series of x. Parseval's identity, $\|x\|^2 = \sum_{i=1}^\infty |x_i|^2$.

12 Ordinary differential equations

- Ordinary differential equation (ODE), $F(x,y,y',\ldots,y^{(n)}) = 0$ or $y^{(n)} = f(x,y,y',\ldots,y^{(n-1)})$. Order of the ODE. Classical solution. Examples: $y^{(n)} = f(x)$.

  General first order ODE, $y' = f(x,y)$. Geometric interpretation of solutions. Direction field. Integral curve. Isoclines. Example: $y' = x^2+y^2$, isoclines are circles centred at 0.

  Examples to various behaviors of solutions: (a) normal reproduction, $y' = \lambda y$, (b) logistic curve, $y' = y(1-y)$, (c) harvesting, $y' = y(1-y)-q$, (d) explosion, $y' = \lambda y^2$.

  General solution to n-th order ODE, $y = \varphi(x, C_1,\ldots,C_n)$, $C_1,\ldots,C_n$ constants. Particular solution. Initial value problem.

  Inverse problem: given relation $y = \varphi(x, C_1,\ldots,C_n)$, $C_1,\ldots,C_n$ constants, find (n-th order) ODE for y. Example: (a) ODE for all circles in the plane, (b) ODE for a family of parabolas in the plane.

- Separable first order ODEs. (1) $y' = f(x)$ with $y(x_0) = y_0$ admits solution $y(x) = y_0 + \int_{x_0}^{x} f(x)dx$. In differentials: $dy = f(x)dx$. Integral curves (vertical shifts). Example: $y' = \frac{1}{x}$.

  (2) $y' = f(y)$ or $dy = f(y)dx$. Solutions are given by $x = x_0 + \int_{y_0}^{y} \frac{dy}{f(y)}$ and $y = y_0$ if $f(y_0) = 0$. Integral curves (horizontal shifts). Remark: In this case, several integral curves may pass through the same point (x_0, y_0). Examples: (a) $y' = y^2$, (b) $y' = 3\sqrt[3]{y^2}$ (here 2 integral curves pass through $(2, 0)$).

  (3) $y' = f(x)g(y)$. If $g(y) \neq 0$, rewrite $\frac{dy}{g(y)} = f(x)dx$ and integrate. If $g(y_0) = 0$ for some y_0, then $y = y_0$ is also a solution to the ODE.

  (4) $M(x)N(y)dx + P(x)Q(y)dy = 0$. First solve the case $P(x) \neq 0$, $N(y) \neq 0$ using integration. Then examine zeros of P and N. Examples: (a) $x(y^2 - 1)dx + y(x^2 - 1)dy = 0$, (b) $xydx + (x + 1)dy = 0$, (c) $x^2y^2dy = (y - 1)dx$.

  Remark: pay attention when divide by a function which takes value 0 at some point, some solution may be lost.

  (5) $y' = f(ax + by)$. Change variables (x, y) to (x, z), where $z = ax + by$. Then $dz = adx + bdy = (a + bf(z))dx$, which is a separable equation.

- Homogeneous first order ODEs. (1) Each homogeneous first order ODE can be reduced to a separable ODE by changing the variables (x, y) to (x, t), where $y = tx$ (and $dy = xdt + tdx$). Examples: (a) $y' = \frac{2xy}{x^2 - y^2}$, (b) $xdy = (x + y)dx$.

  (2) A first order ODE can sometimes be reduced to a homogeneous ODE by changing variables (x, y) to (x, z), where $y = z^m$, and m is tuned appropriately. Example: $2x^4yy' + y^4 = 4x^6$ is reduced to a homogeneous equation by substituting $y = z^{\frac{3}{2}}$.

  (3) $y' = f(\frac{a_1x + b_1y + c_1}{a_2x + b_2y + c_2})$ can be reduced to a homogeneous ODE by changing variables (x, y) to (u, v), where $x = u + \xi$, $y = v + \eta$, and (ξ, η) is a solution to the system of linear equations $\begin{cases} a_1\xi + b_1\eta + c_1 = 0 \\ a_2\xi + b_2\eta + c_2 = 0 \end{cases}$ (i.e., (ξ, η) is the point of intersection of the lines $a_1x + b_1y + c_1 = 0$ and $a_2x + b_2y + c_2 = 0$). If the above system does not have solutions (lines are parallel), then the ODE can be rewritten in the form $y' = F(a_2x + b_2y)$, for some F, and hence can be reduced to a separable ODE (see (5) above). Example: $(2x - 4y + 6)dx + (x + y - 3)dy = 0$.

- Applications: solving physical problems using ODEs.

- Linear first order ODE, $y' + Py = Q$. Homogeneous equation, $y' + Py = 0$. General solution to homogeneous equation, $y = Ce^{-\int P(x)dx}$. General solution to inhomogeneous equation. Strategy: (a) find general solution to the homogeneous equation, (b) replace constant C by a (unknown) function of x, $C(x)$, and substitute in the inhomogeneous equation to find $C(x)$. Example: $xy' - 2y = 2x^4$.

  Reductions to linear first order ODEs: (1) change the role of the variables, $x = x(y)$. Example: $y = (2x + y^3)y'$.

  (2) Bernoulli equation, $y' + a(x)y = b(x)y^n$, $n \neq 1$. Substitute $z = \frac{1}{y^{n-1}}$.

(3) Riccati equation, $y' + a(x)y + b(x)y^2 = c(x)$. General solution is not always possible to find. If, however, $y_1 = y_1(x)$ is a particular solution, then the substitution $y(x) = z(x) + y_1(x)$ reduces Riccati to Bernoulli.

- Existence and uniqueness of solutions.

  Banach fixed-point theorem: If (X, ρ) is a metric space and $T : X \to X$ a contraction, then there exists a unique $x_* \in X$ such that $T(x_*) = x_*$. Furthermore, for any $x_0 \in X$, if $x_{n+1} = T(x_n)$, then $x_n \to x_*$ in (X, ρ).

  Picard-Lindelöf theorem: Let $f = f(x, y)$ be (a) uniformly Lipschitz continuous in y and (b) continuous in x, then there exists $\varepsilon > 0$ such that the initial value problem $y' = f(x, y), y(x_0) = y_0$ has a unique solution on $[x_0 - \varepsilon, x_0 + \varepsilon]$. Moreover, the functions $y_0(x) = y_0$, $y_k(x) = y_0 + \int_{x_0}^{x} f(s, y_{k-1}(s))ds$, $k \geq 1$, converge uniformly to the unique solution. (Picard iteration.) Example: $y' = x - y^2$, $y(0) = 0$, $y_0(x) = 0$, $y_1(x) = \frac{x^2}{2}$, $y_2(x) = \frac{x^2}{2} - \frac{x^4}{20}$.

  Example: $y' = 3y^{\frac{2}{3}}$, $y(0) = 0$, at least two solutions (a) $y(x) = 0$ for all x, (b) $y(x) = x^3$ for $x < 0$ and $y(x) = 0$ for $x \geq 0$.

- Higher order ODEs. Reduction of the order: (a) $F(x, y^{(k)}, \dots, y^{(n)}) = 0$, (b) $F(y, y', \dots, y^{(n)}) = 0$. Example: $2yy'' = (y')^2 + 1$.

- Linear ODE of nth order, $a_0 y^{(n)} + a_1 y^{(n-1)} + \ldots + a_n y = F(x)$, where $a_i(x)$ are continuous functions. Let (a, b) be an interval such that $a_0(x) \neq 0$ for all $x \in (a, b)$. On the interval (a, b) rewrite the equation as $L_n y = y^{(n)} + p_1 y^{(n-1)} + \ldots + p_n y = f(x)$, where $p_i(x)$ are continuous functions on (a, b). Homogeneous linear ODE, $L_n y = 0$.

  Theorem: If y_1, y_2 are particular solutions to $L_n y = 0$ and C is a constant, then $(y_1 + y_2)$ and Cy_1 are also solutions to $L_n y = 0$.

  Linearly independent functions. Examples: (a) $1, x, x^2, \dots, x^n$ are linearly independent, (b) $1, \sin^2 x, \cos^2 x$ are linearly dependent. Wronski determinant, $W(x) = W(y_1, \dots, y_n)$.

  Theorem: If $y_1, \dots, y_n$ are solutions to $L_n y = 0$ on (a, b), then either $W(x) \neq 0$ for all $x \in (a, b)$ ($y_1, \dots, y_n$ are linearly independent), or $W(x) = 0$ for all $x \in (a, b)$ ($y_1, \dots, y_n$ are linearly dependent).

  Fundamental system of solutions to $L_n y = 0$. Theorem: If $y_1, \dots, y_n$ is a fundamental system of solutions to $L_n y = 0$, then $y = C_1 y_1 + \ldots C_n y_n$ is the general solution to $L_n y = 0$. Example: $y'' - y = 0$.

  Inverse problem: Given a linearly independent system of functions $y_1, \dots, y_n$, find the nth order linear ODE $L_n y = 0$ which they solve. Example: x, x^2 is a fundamental system for $y'' - \frac{2}{x}y' + \frac{2}{x^2}y = 0$ on intervals $(0, +\infty)$ and $(-\infty, 0)$.

- Inhomogeneous linear ODEs, $L_n y = f$. Theorem: If y_1 is a particular solution to $L_n y = f$ and y_2 is the general solution to $L_n y = 0$, then $(y_1 + y_2)$ is the general solution to $L_n y = f$. Example: $y'' + y = 3x$. Strategy to find general solution to $L_n y = f$: (a) general solution to $L_n y = 0$, $y = C_1 y_1 + \ldots C_n y_n$, (b) replace C_i by $C_i(x)$ and substitute into $L_n y = f$, write system of linear equations for $\frac{dC_i}{dx}$ with

Wronski matrix of $y_1, \ldots, y_n$, $\begin{pmatrix} y_1 & \ldots & y_n \\ y_1' & \ldots & y_n' \\ & \ldots & \\ y_1^{(n-1)} & \ldots & y_n^{(n-1)} \end{pmatrix} \begin{pmatrix} C_1'(x) \\ C_2'(x) \\ \ldots \\ C_n'(x) \end{pmatrix} = \begin{pmatrix} 0 \\ \ldots \\ 0 \\ f(x) \end{pmatrix}$.

Example: $xy'' - y' = x^2$.

- Linear ODE with constant coefficients, $a_0 y^{(n)} + \ldots + a_n y = f(x)$, where $a_i \in \mathbb{R}$, $a_0 \neq 0$. Characteristic equation, $a_0\lambda^n + \ldots + a_n = 0$. Fundamental system of solutions of homogeneous equation: (a) simple real root, (b) simple complex root, (c) real root of multiplicity k, (d) complex root of multiplicity k. Examples: (a) $y'' - y = 0$, (b) $y'' + a^2 y = 0$, (c) $y''' - y'' - y' + y = 0$, (d) $y'''' + 8y'' + 16y = 0$.

  Particular solutions to inhomogeneous equation $Ly = f$: (a) $f(x) = P_m(x)e^{\gamma x}$, (b) $f(x) = e^{\alpha x}(P_m(x)\cos\beta x + Q_m(x)\sin\beta x)$, (c) $f(x) = f_1(x) + \ldots + f_k(x)$. Example: $y''' - 6y'' + 9y' = xe^{3x} + e^{3x}\cos 2x$.

13 Literature

1. A. Schüler. Calculus. Lecture notes, Leipzig University 2006.

2. K. Hoffman and R. Kunze. Linear algebra.

3. I. Lankham, B. Nachtergaele, A. Schilling. Linear algebra (as an introduction to abstract mathematics).

4. G. Fichtenholz. Differential- und Integralrechnung.

5. S. Mac Lane and G. Birkhoff. Algebra.

6. T.L. Chow. Mathematical methods for physicists: A concise introduction.

7. G. Knieper. Mathematik für Physiker.