Linear Algebra Cheat Sheet

HKPFS Math PhD Interview Prep

1 1. System of Linear Equations

1.1 Matrix Equation Form

System: $a_{11}x_1 + \cdots + a_{1m}x_m = b_1$, ..., $a_{n1}x_1 + \cdots + a_{nm}x_m = b_n$
Matrix form: $Ax = b$ where A is $n \times m$ matrix
Augmented matrix: $(A|b)$

1.2 Solution Methods

Matrix inversion (square invertible): $x = A^{-1}b$
Cramer's rule: For invertible $n \times n$ matrix A:

$$x_i = \frac{\det(A_i)}{\det(A)}$$

where A_i is A with i-th column replaced by b

1.3 Elementary Row Operations

Type I: Interchange two rows
Type II: Multiply row by nonzero scalar
Type III: Add scalar multiple of one row to another
Elementary matrix: Apply operation to I_n
EA = result of applying operation E to A
All elementary matrices are invertible

1.4 Reduced Row Echelon Form (RREF)

Leading entry: First nonzero entry in row
Leading one: Leading entry equals 1
RREF conditions:
1. Leading one is only nonzero in its column
2. All-zero rows at bottom
3. Leading ones shift right as rows descend

Gaussian elimination: Transform to RREF via elementary operations
Leading variable: Corresponds to leading entry
Free variable: Not a leading variable

1.5 Consistency

Consistent: Has solution (solution set nonempty)
Inconsistent: No solution
$Ax = b$ inconsistent $\Leftrightarrow$ RREF of $(A|b)$ has row $[0 \cdots 0|c]$ with $c \neq 0$

1.6 Homogeneous Systems

$Ax = 0$ always has trivial solution $x = 0$
If $m > n$ (more variables than equations), infinite solutions exist
General solution structure: If x_p satisfies $Ax_p = b$, then

$$\text{Solution set} = \{x_p + y : Ay = 0\}$$

1.7 Invertibility and Solutions

For $n \times n$ matrix A:
$Ax = 0$ has nontrivial solution $\Leftrightarrow$ A not invertible
A invertible $\Leftrightarrow$ RREF of A is I_n

2 2. Determinants

2.1 Definition

2×2: $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$
$n \times n$ **(cofactor expansion):** Fix row i or column j:

$$\det(A) = \sum_{k=1}^{n} (-1)^{i+k} a_{ik} \det(\tilde{A}_{ik})$$

where $\tilde{A}_{ij}$ is (i,j)-minor (delete row i, column j)

2.2 Properties

$\det(I_n) = 1$, $\det(0) = 0$
Transpose: $\det(A^T) = \det(A)$
Switching rows/columns: Changes sign
Two identical rows/columns: $\det(A) = 0$
Scalar multiplication: $\det(cA) = c^n \det(A)$
Multiplicative: $\det(AB) = \det(A)\det(B)$
Row operation: Adding scalar multiple of row to another doesn't change determinant
Invertibility: A invertible $\Leftrightarrow \det(A) \neq 0$
If invertible: $\det(A^{-1}) = \frac{1}{\det(A)}$

2.3 Adjugate and Inverse

Adjugate: $\operatorname{adj}(A) = ((-1)^{i+j}\det(\tilde{A}_{ij}))^T$
Inverse formula: $A^{-1} = \frac{1}{\det(A)}\operatorname{adj}(A)$

2.4 Finding Inverse via Row Operations

Form $[A|I_n]$, row reduce to $[I_n|A^{-1}]$
Every invertible matrix is product of elementary matrices

3 3. Vector Spaces

3.1 Definition

Vector space V over $\mathbb{R}$ with operations + and scalar multiplication satisfying:
1. $x + y = y + x$ (commutative)
2. $(x + y) + z = x + (y + z)$ (associative)
3. $\exists 0 : x + 0 = x$ (zero element)
4. $\forall x, \exists y : x + y = 0$ (additive inverse)
5. $1x = x$
6. $a(bx) = (ab)x$
7. $(a + b)x = ax + bx$
8. $a(x + y) = ax + ay$

3.2 Examples

$\mathbb{R}^n$ with standard operations
$M_{n \times m}$ (all $n \times m$ matrices)
$P_n(\mathbb{R})$ (polynomials degree $\leq n$)
Functions $\mathbb{R} \to \mathbb{R}$ with pointwise operations

3.3 Vector Subspaces

Nonempty $S \subseteq V$ is subspace if:
1. $v_1, v_2 \in S \Rightarrow v_1 + v_2 \in S$ (closed under addition)
2. $v \in S, c \in \mathbb{R} \Rightarrow cv \in S$ (closed under scalar mult)

Examples: Solution set of $Ax = 0$; $\{0\}$; V itself
In $\mathbb{R}^2$: Only $\{0\}$, lines through origin, $\mathbb{R}^2$
In $\mathbb{R}^3$: Only $\{0\}$, lines through origin, planes through origin, $\mathbb{R}^3$

3.4 Linear Combinations

$v = a_1u_1 + \cdots + a_nu_n$ for scalars a_i and vectors u_i
Span: $\operatorname{span}(S) = \{\text{all linear combinations of vectors in } S\}$
$\operatorname{span}(S)$ is a subspace (smallest subspace containing S)
Column space: $\operatorname{col}(A) = \operatorname{span}\{\text{columns of } A\} = \{Ax : x \in \mathbb{R}^m\}$

3.5 Linear Independence

$S = \{v_1, \ldots, v_r\}$ is **linearly independent** if

$$a_1v_1 + \cdots + a_rv_r = 0 \Rightarrow a_1 = \cdots = a_r = 0$$

Otherwise **linearly dependent**
Test: Solve $a_1v_1 + \cdots + a_rv_r = 0$ (system of linear equations)

3.6 Basis and Dimension

Basis: Linearly independent set that spans V
Every vector in V has unique representation as linear combination of basis vectors
Standard basis for $\mathbb{R}^n$: $\{e_1, \ldots, e_n\}$ where e_i has 1 in position i
Dimension: $\dim(V)$ = number of vectors in any basis
All bases of V have same number of vectors
If $\dim(V) = n$ and S has n linearly independent vectors, then S is basis
If $W \subseteq V$ subspace, $\dim(W) \leq \dim(V)$

4 4. Linear Transformations

4.1 Definition

$T : V \to W$ is **linear transformation** if:
1. $T(x + y) = T(x) + T(y)$
2. $T(cx) = cT(x)$

Consequences: $T(0) = 0$;
$T(a_1x_1 + \cdots + a_rx_r) = a_1T(x_1) + \cdots + a_rT(x_r)$

4.2 Examples

$T(x) = Ax$ for matrix A (most important!)
Differentiation on $C^1(\mathbb{R})$
Integration on continuous functions
Zero transformation $T_0(v) = 0$
Identity $\mathrm{Id}_V(v) = v$

4.3 Construction

Given basis $\{v_1, \ldots, v_n\}$ of V and any $w_1, \ldots, w_n \in W$:
$\exists!$ linear $T : V \to W$ with $T(v_i) = w_i$

4.4 Null Space and Range

Null space: $N(T) = \{v \in V : T(v) = 0\}$ (kernel)
Range: $R(T) = \{T(v) : v \in V\}$ (image)
Both are subspaces
Computing range: If $\{v_1, \ldots, v_m\}$ basis for V:

$$R(T) = \operatorname{span}\{T(v_1), \ldots, T(v_m)\}$$

Nullity: $\operatorname{nullity}(T) = \dim(N(T))$
Rank: $\operatorname{rank}(T) = \dim(R(T))$

4.5 Dimension Formula

$$\operatorname{nullity}(T) + \operatorname{rank}(T) = \dim(V)$$

For matrix A ($n \times m$): - $\operatorname{rank}(A)$ = number of leading ones in RREF - $\operatorname{nullity}(A)$ = number of free variables - $\operatorname{rank}(A) + \operatorname{nullity}(A) = m$

4.6 Injectivity and Surjectivity

Injective: $T(x) = T(y) \Rightarrow x = y$
T injective $\Leftrightarrow N(T) = \{0\} \Leftrightarrow \operatorname{rank}(T) = \dim(V)$
If $\dim(V) > \dim(W)$, then T not injective
Surjective: $R(T) = W$
Isomorphism: Bijective (injective and surjective) linear map
T isomorphism $\Rightarrow \dim(V) = \dim(W)$
Inverse: If T invertible, T^{-1} also linear

5 5. Matrix Representations

5.1 Coordinate Vectors

Ordered basis $\beta = \{v_1, \ldots, v_m\}$ for V
For $v = a_1v_1 + \cdots + a_mv_m$:

$$[v]_\beta = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix}$$

5.2 Matrix of Linear Transformation

$T : V \to W$, $\beta = \{v_1, \ldots, v_m\}$ basis for V, $\gamma = \{w_1, \ldots, w_n\}$ basis for W
$[T]_\gamma^\beta$ is $n \times m$ matrix with j-th column $= [T(v_j)]_\gamma$
Key property: $[T(v)]_\gamma = [T]_\gamma^\beta [v]_\beta$
Standard matrix: For $T : \mathbb{R}^m \to \mathbb{R}^n$, use standard bases
$\text{rank}(T) = \text{rank}([T]_\gamma^\beta)$

5.3 Operations on Transformations

Addition: $(T + T')(v) = T(v) + T'(v)$
$[T + T']_\gamma^\beta = [T]_\gamma^\beta + [T']_\gamma^\beta$
Composition: $(T' \circ T)(v) = T'(T(v))$
$[T' \circ T]_\alpha^\beta = [T']_\alpha^\gamma [T]_\gamma^\beta$
Inverse: $[T^{-1}]_\beta^\gamma = ([T]_\gamma^\beta)^{-1}$
T invertible $\Leftrightarrow$ $[T]_\gamma^\beta$ invertible

5.4 Change of Basis

β, β' two bases for V
Change of coordinate matrix: $Q = [\text{Id}_V]_{\beta'}^\beta$
Q is invertible; $[v]_\beta = Q[v]_{\beta'}$
Similarity: Matrices A, B are similar if $\exists$ invertible Q: $B = Q^{-1}AQ$
For $T : V \to V$ with bases β, β':

$$[T]_{\beta'} = Q^{-1}[T]_\beta Q$$

where $Q = [\text{Id}_V]_{\beta'}^\beta$
Similar matrices have same determinant and rank

6 6. Inner Products

6.1 Definition

Inner product $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ satisfying:
1. $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$
2. $\langle cx, y \rangle = c\langle x, y \rangle$
3. $\langle x, y \rangle = \langle y, x \rangle$ (symmetric)
4. $\langle x, x \rangle > 0$ if $x \neq 0$ (positive definite)

Standard inner product on $\mathbb{R}^n$:

$$\langle x, y \rangle = x_1y_1 + \cdots + x_ny_n = x^Ty$$

For functions on $[a, b]$: $\langle f, g \rangle = \int_a^b f(x)g(x)dx$
For matrices: $\langle A, B \rangle = \text{tr}(A^TB)$
Weighted inner product: $\langle x, y \rangle = a_1x_1y_1 + \cdots + a_nx_ny_n$ where $a_i > 0$
Key property: $\langle Ax, y \rangle = \langle x, A^Ty \rangle$

6.2 Norm and Distance

Norm (length): $\|x\| = \sqrt{\langle x, x \rangle}$
$\|cx\| = |c|\|x\|$; $\|x\| = 0 \Leftrightarrow x = 0$
Cauchy-Schwarz: $|\langle x, y \rangle| \leq \|x\|\|y\|$
Triangle inequality: $\|x + y\| \leq \|x\| + \|y\|$

6.3 Orthogonality

Orthogonal: $x \perp y$ if $\langle x, y \rangle = 0$
Unit vector: $\|v\| = 1$
Normalize: $\frac{v}{\|v\|}$ is unit vector
Orthogonal set: Pairwise orthogonal; always linearly independent (if nonzero)
Orthonormal set: Orthogonal and all unit vectors

6.4 Orthonormal Basis

Standard basis $\{e_1, \ldots, e_n\}$ is orthonormal for $\mathbb{R}^n$
If $\{v_1, \ldots, v_r\}$ orthonormal basis and $y = a_1v_1 + \cdots + a_rv_r$:

$$a_i = \langle v_i, y \rangle$$

If orthogonal (not normalized): $a_i = \frac{\langle v_i, y \rangle}{\|v_i\|^2}$

6.5 Gram-Schmidt Process

Given basis $\{w_1, \ldots, w_n\}$, construct orthogonal basis $\{v_1, \ldots, v_n\}$:
$v_1 = w_1$
$v_i = w_i - \sum_{j=1}^{i-1} \frac{\langle w_i, v_j \rangle}{\|v_j\|^2} v_j$ for $i \geq 2$
Then normalize each v_i to get orthonormal basis: $\left\{ \frac{v_1}{\|v_1\|}, \ldots, \frac{v_n}{\|v_n\|} \right\}$
Every inner product space has orthonormal basis

6.6 Orthogonal Complement

For subset $S \subseteq V$:

$$S^\perp = \{x \in V : \langle x, y \rangle = 0 \text{ for all } y \in S\}$$

$S^\perp$ is subspace; $S^\perp = \text{span}(S)^\perp$
Dimension formula: $\dim(W) + \dim(W^\perp) = \dim(V)$
$(W^\perp)^\perp = W$ for subspace W

6.7 Orthogonal Projection

For subspace W with orthonormal basis $\{v_1, \ldots, v_r\}$:
Every $v \in V$ uniquely: $v = x + y$ where $x \in W$, $y \in W^\perp$
Projection onto W:

$$\text{proj}_W(v) = \langle v, v_1 \rangle v_1 + \cdots + \langle v, v_r \rangle v_r$$

$\text{proj}_W(v)$ minimizes $\|v - w\|$ over all $w \in W$
Finding orthonormal basis for $W^\perp$:
1. Start with orthonormal basis $\{v_1, \ldots, v_r\}$ for W
2. Extend to basis $\{v_1, \ldots, v_r, w_{r+1}, \ldots, w_n\}$ for V
3. Apply Gram-Schmidt to get $\{v_1, \ldots, v_r, v_{r+1}, \ldots, v_n\}$
4. Then $\{v_{r+1}, \ldots, v_n\}$ is orthonormal basis for $W^\perp$

7 7. Least Squares Approximation

7.1 Problem Setup

Given data points $(t_1, y_1), \ldots, (t_k, y_k)$
Find line $y = ct + d$ that best approximates data
Equivalently: solve $Ax \approx y$ where

$$A = \begin{pmatrix} t_1 & 1 \\ t_2 & 1 \\ \vdots & \vdots \\ t_k & 1 \end{pmatrix}, \quad x = \begin{pmatrix} c \\ d \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{pmatrix}$$

7.2 Least Squares Solution

Find x_0 that minimizes $\|Ax - y\|$ (distance from y to $\text{col}(A)$)
Solution: x_0 satisfies the **normal equation**:

$$A^TAx_0 = A^Ty$$

Interpretation: Ax_0 is orthogonal projection of y onto $\text{col}(A)$
$Ax_0 - y \in (\text{col}(A))^\perp = N(A^T)$

7.3 Finding Least Squares Solution

Method 1 (Normal equations):
1. Compute A^TA (always square, symmetric)
2. Compute A^Ty
3. Solve $(A^TA)x_0 = A^Ty$
4. If A^TA invertible: $x_0 = (A^TA)^{-1}A^Ty$

Method 2 (Orthogonal projection):
1. Find orthonormal basis $\{v_1, \ldots, v_m\}$ for $\text{col}(A)$ (Gram-Schmidt)
2. Compute $\text{proj}_{\text{col}(A)}(y) = \sum_{i=1}^m \langle y, v_i \rangle v_i$
3. Solve $Ax_0 = \text{proj}_{\text{col}(A)}(y)$

7.4 Properties

A^TA is invertible $\Leftrightarrow$ columns of A linearly independent
$\text{rank}(A^TA) = \text{rank}(A)$ (always!)
$\text{nullity}(A^TA) = \text{nullity}(A)$
Least squares solution always exists; unique if columns of A linearly independent

7.5 General Polynomial Fitting

For polynomial $y = a_0 + a_1t + \cdots + a_mt^m$:

$$A = \begin{pmatrix} 1 & t_1 & t_1^2 & \cdots & t_1^m \\ 1 & t_2 & t_2^2 & \cdots & t_2^m \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & t_k & t_k^2 & \cdots & t_k^m \end{pmatrix}, \quad x = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{pmatrix}$$

Then solve normal equation $(A^TA)x = A^Ty$

8 8. Eigenvalues & Eigenvectors

8.1 Definitions

For $n \times n$ matrix A:
Eigenvector: Nonzero v such that $Av = \lambda v$ for some scalar λ
Eigenvalue: Scalar λ such that $Av = \lambda v$ for some nonzero v
Eigenspace: $E_\lambda = \{v : Av = \lambda v\} = N(A - \lambda I)$

8.2 Finding Eigenvalues

Characteristic polynomial: $p_A(t) = \det(A - tI)$
Degree n polynomial with leading coefficient $(-1)^n$
λ is eigenvalue $\Leftrightarrow$ $\det(A - \lambda I) = 0$
At most n eigenvalues

8.3 Finding Eigenvectors

For eigenvalue λ: solve $(A - \lambda I)v = 0$
Solution space is E_λ (always contains nonzero vectors)

8.4 Multiplicities

Algebraic multiplicity: Largest k such that $(t - \lambda)^k | p_A(t)$
Geometric multiplicity: $\dim(E_\lambda) = \text{nullity}(A - \lambda I)$
Always: $1 \leq$ geom mult $\leq$ alg mult

8.5 Diagonalization

A is **diagonalizable** if $\exists$ invertible Q: $Q^{-1}AQ = D$ diagonal
Criteria: A diagonalizable $\Leftrightarrow$ $\exists$ basis of $\mathbb{R}^n$ of eigenvectors
$\Leftrightarrow$ For each eigenvalue: geom mult = alg mult
If A has n distinct eigenvalues, A is diagonalizable
How to diagonalize: Find eigenvectors $v_1, \ldots, v_n$ (basis)
Set $Q = [v_1 \cdots v_n]$
Then $Q^{-1}AQ = \text{diag}(\lambda_1, \ldots, \lambda_n)$ where $Av_i = \lambda_i v_i$
Eigenvectors with distinct eigenvalues are linearly independent

8.6 Applications

Powers: $A = QDQ^{-1} \Rightarrow A^n = QD^nQ^{-1}$
$D^n = \mathrm{diag}(\lambda_1^n, \ldots, \lambda_n^n)$
Matrix exponential: $e^A = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots$
If $A = QDQ^{-1}$: $e^A = Qe^DQ^{-1}$ where $e^D = \mathrm{diag}(e^{\lambda_1}, \ldots, e^{\lambda_n})$

9 9. Quick Reference

9.1 Matrix Properties

$(AB)^T = B^TA^T$
$(AB)^{-1} = B^{-1}A^{-1}$
$(A^T)^{-1} = (A^{-1})^T$
$\mathrm{rank}(A^TA) = \mathrm{rank}(A)$
$\mathrm{nullity}(A^TA) = \mathrm{nullity}(A)$

9.2 Dimension Results

$\dim(\mathbb{R}^n) = n$
$\dim(M_{n\times m}) = nm$
$\dim(P_n(\mathbb{R})) = n + 1$

9.3 Invertibility Equivalences

For $n \times n$ matrix A, TFAE:
- A invertible - $\det(A) \neq 0$ - RREF of A is I_n - $\mathrm{rank}(A) = n$ - Columns of A linearly independent - Columns of A span $\mathbb{R}^n$ - $Ax = 0$ only has trivial solution - $Ax = b$ has unique solution for any b - A is product of elementary matrices

9.4 Common Mistakes to Avoid

$\det(A + B) \neq \det(A) + \det(B)$ in general

Eigenvectors must be nonzero by definition

Similar matrices have same eigenvalues but not necessarily same eigenvectors

Change of basis matrix depends on order

In least squares: A^TA may not be invertible

Gram-Schmidt must maintain order of vectors

Mathematical Analysis Cheat Sheet

HKPFS Math PhD Interview Prep

1 1. Real Numbers

1.1 Field Axioms

Field F with operations $+$ and $\cdot$:
Addition: (A1) Commutative, (A2) Associative, (A3) Zero exists, (A4) Additive inverse exists
Multiplication: (M1) Commutative, (M2) Associative, (M3) Unity $1 \neq 0$ exists, (M4) Multiplicative inverse exists (for $a \neq 0$)
Distributive: (D) $a(b+c) = ab + ac$
Key facts: $0 \cdot a = 0$; $(-1)(-1) = 1$; $ab = 0 \Rightarrow a = 0$ or $b = 0$

1.2 Ordered Field

Relation $\leq$ satisfying:
(a) Reflexive: $a \leq a$
(b) Antisymmetric: $a \leq b$ and $b \leq a \Rightarrow a = b$
(c) Transitive: $a \leq b$ and $b \leq c \Rightarrow a \leq c$
(d) Compatible with $+$: $a \leq b \Rightarrow a + c \leq b + c$
(e) Compatible with $\cdot$: $a \leq b, 0 \leq c \Rightarrow ac \leq bc$
Properties: $a > 0 \Leftrightarrow -a < 0$; $1 > 0$; $a > b > 0 \Rightarrow b^{-1} > a^{-1} > 0$
Absolute value: $|a| = \max\{a, -a\}$
$|a| \geq 0$; $|ab| = |a||b|$; $|a+b| \leq |a| + |b|$ (triangle inequality); $||a| - |b|| \leq |a - b|$

1.3 Completeness (LUB Property)

Supremum: For non-empty $S \subseteq \mathbb{R}$ bounded above, $\sup(S)$ exists and is the *least* upper bound
Infimum: For non-empty S bounded below, $\inf(S) = -\sup(-S)$
Archimedean Property: For any $a \in \mathbb{R}$, $\exists N \in \mathbb{N} : N > a$
Corollary: $\forall \epsilon > 0, \exists N : 1/N < \epsilon$
Density of $\mathbb{Q}$: For $a < b$, $\exists p/q \in \mathbb{Q} : a < p/q < b$
Nested Intervals: If $I_n = [a_n, b_n]$ with $I_{n+1} \subseteq I_n$, then $\bigcap_{n=1}^{\infty} I_n \neq \emptyset$
$\mathbb{R}$ is uncountable (Cantor's diagonal argument)

2 2. Sequences

2.1 Convergence

$(a_n) \to L$ if $\forall \epsilon > 0, \exists N : |a_n - L| < \epsilon$ for all $n \geq N$
Uniqueness: Limit is unique if it exists
Boundedness: Convergent sequences are bounded
Negation: $(a_n) \not\to L$ if $\exists \epsilon > 0$ and infinitely many n with $|a_n - L| \geq \epsilon$

2.2 Arithmetic of Limits

If $(a_n) \to L_1$ and $(b_n) \to L_2$:
- $(ca_n) \to cL_1$ - $(a_n + b_n) \to L_1 + L_2$ - $(a_n b_n) \to L_1 L_2$ - $(a_n/b_n) \to L_1/L_2$ if $b_n \neq 0, L_2 \neq 0$ - If $a_n \leq b_n$, then $L_1 \leq L_2$
Sandwich Theorem: If $a_n \leq c_n \leq b_n$ and $(a_n), (b_n) \to L$, then $(c_n) \to L$

2.3 Subsequences and Monotone Sequences

Subsequence: (a_{n_k}) where $n_1 < n_2 < \cdots$
If $(a_n) \to L$, then every subsequence $\to L$
Monotone Convergence Theorem: Bounded monotone sequence converges
- Increasing bounded above: $(a_n) \to \sup\{a_n\}$ - Decreasing bounded below: $(a_n) \to \inf\{a_n\}$

2.4 Limit Superior and Inferior

For bounded (a_n):
$\limsup_{n\to\infty} a_n = \lim_{n\to\infty} \sup\{a_k : k \geq n\}$
$\liminf_{n\to\infty} a_n = \lim_{n\to\infty} \inf\{a_k : k \geq n\}$
Always: $\liminf a_n \leq \limsup a_n$
(a_n) converges $\Leftrightarrow \liminf a_n = \limsup a_n$
Bolzano-Weierstrass: Every bounded sequence has convergent subsequence

2.5 Cauchy Sequences

(a_n) is Cauchy if $\forall \epsilon > 0, \exists N : |a_m - a_n| < \epsilon$ for all $m, n \geq N$
Cauchy Criterion: (a_n) converges $\Leftrightarrow$ (a_n) is Cauchy
Cauchy sequences are bounded

3 3. Series

3.1 Convergence of Series

Series $\sum_{i=1}^{\infty} a_i$ converges if partial sums $s_n = \sum_{i=1}^{n} a_i$ converge
If $\sum a_i$ converges, then $a_n \to 0$
Cauchy Criterion: $\sum a_i$ converges $\Leftrightarrow$ $\forall \epsilon > 0, \exists N : |a_{n+1} + \cdots + a_m| < \epsilon$ for all $m > n \geq N$

3.2 Absolute Convergence

$\sum a_i$ is **absolutely convergent** if $\sum |a_i| < \infty$
Absolute convergence $\Rightarrow$ convergence
Rearrangement: Absolutely convergent series can be rearranged without changing sum

3.3 Convergence Tests

Comparison Test: If $0 \leq a_n \leq b_n$ and $\sum b_n$ converges, then $\sum a_n$ converges
Root Test: Let $\alpha = \limsup |a_n|^{1/n}$ - If $\alpha < 1$: absolutely convergent - If $\alpha > 1$: divergent
Ratio Test: - If $\limsup |a_{n+1}/a_n| < 1$: absolutely convergent - If $\liminf |a_{n+1}/a_n| > 1$: divergent
Alternating Series Test: If $a_1 \geq a_2 \geq \cdots \geq 0$ and $a_n \to 0$, then $\sum (-1)^{n+1} a_n$ converges
Integral Test: If $f : [1, \infty) \to \mathbb{R}_{\geq 0}$ decreasing, then $\sum f(n)$ converges $\Leftrightarrow \int_1^{\infty} f(x)dx < \infty$

4 4. Limits of Functions

4.1 Definition

$\lim_{x\to\alpha} f(x) = L$ if $\forall \epsilon > 0, \exists \delta > 0 : |f(x) - L| < \epsilon$ whenever $0 < |x - \alpha| < \delta$
One-sided limits: - Right: $\lim_{x\to\alpha^+} f(x) = L$ - Left: $\lim_{x\to\alpha^-} f(x) = L$
$\lim_{x\to\alpha} f = L \Leftrightarrow \lim_{x\to\alpha^-} f = L = \lim_{x\to\alpha^+} f$

4.2 Sequential Criterion

$\lim_{x\to\alpha} f(x) = L \Leftrightarrow$ for all sequences $(a_n) \to \alpha$ (with $a_n \neq \alpha$), $f(a_n) \to L$

4.3 Arithmetic of Limits

If $\lim_{x\to\alpha} f = L_1$ and $\lim_{x\to\alpha} g = L_2$:
- $\lim(cf) = cL_1$ - $\lim(f+g) = L_1 + L_2$ - $\lim(fg) = L_1 L_2$ - $\lim(f/g) = L_1/L_2$ if $L_2 \neq 0$
Sandwich Theorem: If $f \leq h \leq g$ and $\lim f = \lim g = L$, then $\lim h = L$

4.4 Limits at Infinity

$\lim_{x\to\infty} f(x) = L$ if $\forall \epsilon > 0, \exists M : |f(x) - L| < \epsilon$ for all $x > M$
$\lim_{x\to\alpha} f(x) = \infty$ if $\forall C > 0, \exists \delta > 0 : f(x) > C$ whenever $0 < |x - \alpha| < \delta$

5 5. Continuity

5.1 Definition

$f : I \to \mathbb{R}$ is **continuous at** $\alpha \in I$ if $\forall \epsilon > 0, \exists \delta > 0 : |f(x) - f(\alpha)| < \epsilon$ whenever $|x - \alpha| < \delta$
Equivalently: $\lim_{x\to\alpha} f(x) = f(\alpha)$
Sequential Criterion: f continuous at $\alpha \Leftrightarrow$ for all $(a_n) \to \alpha$, $f(a_n) \to f(\alpha)$

5.2 Properties

If f, g continuous at α: - $cf, f \pm g, fg$ continuous at α - f/g continuous at α if $g(\alpha) \neq 0$
Composition: If f continuous at α and g continuous at $f(\alpha)$, then $g \circ f$ continuous at α

5.3 Key Theorems

Extreme Value Theorem: If $f : [a, b] \to \mathbb{R}$ continuous, then f attains max and min on $[a, b]$
Intermediate Value Theorem: If $f : [a, b] \to \mathbb{R}$ continuous, then f takes every value between $f(a)$ and $f(b)$
Image of Interval: Continuous image of interval is interval
Inverse Function: If $f : I \to \mathbb{R}$ continuous and injective, then: - f strictly monotone - $f^{-1} : f(I) \to I$ is continuous

5.4 Uniform Continuity

$f : I \to \mathbb{R}$ is **uniformly continuous** if $\forall \epsilon > 0, \exists \delta > 0 : |f(x) - f(y)| < \epsilon$ whenever $|x - y| < \delta$ (for all $x, y \in I$)
Key difference: δ depends only on ϵ, not on point
Uniform continuity $\Rightarrow$ continuity
Theorem: Continuous on $[a, b] \Rightarrow$ uniformly continuous
Lipschitz: $|f(x) - f(y)| \leq C|x - y|$ for all x, y ($C > 0$)
Lipschitz $\Rightarrow$ uniformly continuous

6 6. Sequences of Functions

6.1 Pointwise Convergence

$(f_n) \to f$ pointwise if $\forall x \in I$, $f_n(x) \to f(x)$
Pointwise limit of continuous functions may not be continuous

6.2 Uniform Convergence

$(f_n) \to f$ uniformly (written $(f_n) \rightrightarrows f$) if $\forall \epsilon > 0, \exists N : |f_n(x) - f(x)| < \epsilon$ for all $n \geq N$ and all $x \in I$
Key: N depends only on ϵ, not on x
Cauchy Criterion: $(f_n) \rightrightarrows f \Leftrightarrow (f_n)$ uniformly Cauchy: $\forall \epsilon > 0, \exists N : |f_m(x) - f_n(x)| < \epsilon$ for all $m > n \geq N$ and all x

6.3 Properties

Continuity Preservation: If f_n continuous and $(f_n) \rightrightarrows f$, then f continuous
Uniform convergence $\Rightarrow$ pointwise convergence (not converse!)
Weierstrass M-Test: If $|f_n(x)| \leq M_n$ for all x and $\sum M_n < \infty$, then $\sum f_n$ converges uniformly

7 7. Power Series

7.1 Radius of Convergence

Power series: $f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n$
Let $\beta = \limsup |a_n|^{1/n}$ and $R = 1/\beta$ (radius of convergence)
Convergence: - Absolutely for $|x - x_0| < R$ - Diverges for $|x - x_0| > R$ - Uniformly on $[x_0 - R_0, x_0 + R_0]$ for any $R_0 < R$

7.2 Differentiation and Integration

$f(x) = \sum a_n(x-x_0)^n$ with radius R
Term-by-term differentiation: $f'(x) = \sum na_n(x-x_0)^{n-1}$ has same radius R
f is infinitely differentiable on (x_0-R, x_0+R)
$f^{(n)}(x_0) = n!a_n$
Term-by-term integration: $\int_{x_0}^x f(t)dt = \sum \frac{a_n}{n+1}(x-x_0)^{n+1}$ has same radius R

7.3 Important Series

$e^x = \sum_{n=0}^\infty \frac{x^n}{n!}$ (all x)
$\sin x = \sum_{n=0}^\infty \frac{(-1)^n x^{2n+1}}{(2n+1)!}$ (all x)
$\cos x = \sum_{n=0}^\infty \frac{(-1)^n x^{2n}}{(2n)!}$ (all x)
$\frac{1}{1-x} = \sum_{n=0}^\infty x^n$ ($|x|<1$)
$\ln(1+x) = \sum_{n=1}^\infty \frac{(-1)^{n+1}x^n}{n}$ ($|x|<1$)

8 8. Differentiation

8.1 Definition

$f'(\alpha) = \lim_{x\to\alpha} \frac{f(x)-f(\alpha)}{x-\alpha}$ if limit exists
Differentiable at $\alpha \Rightarrow$ continuous at α

8.2 Rules

$(cf)' = cf'$; $(f \pm g)' = f' \pm g'$
Product: $(fg)' = f'g + fg'$
Quotient: $(f/g)' = \frac{f'g - fg'}{g^2}$ if $g \neq 0$
Chain Rule: $(g \circ f)'(\alpha) = g'(f(\alpha)) \cdot f'(\alpha)$
Inverse: If f differentiable, bijective, $f' \neq 0$, then $(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}$

8.3 Mean Value Theorems

Rolle's Theorem: If f continuous on $[a,b]$, differentiable on (a,b), and $f(a)=f(b)$, then $\exists\zeta \in (a,b) : f'(\zeta) = 0$
Mean Value Theorem: If f continuous on $[a,b]$, differentiable on (a,b), then $\exists\zeta \in (a,b)$:

$$f'(\zeta) = \frac{f(b)-f(a)}{b-a}$$

Cauchy MVT: $\exists\zeta \in (a,b)$:

$$f'(\zeta)(g(b)-g(a)) = g'(\zeta)(f(b)-f(a))$$

8.4 Applications of MVT

$f' \equiv 0$ on interval $\Rightarrow f$ constant
$f' > 0 \Rightarrow f$ strictly increasing
$f' \geq 0 \Rightarrow f$ increasing
f' bounded $\Rightarrow f$ Lipschitz

8.5 L'Hôpital's Rules

Type $0/0$**:** If $\lim_{x\to\alpha} f(x) = 0 = \lim_{x\to\alpha} g(x)$ and $\lim_{x\to\alpha} \frac{f'(x)}{g'(x)} = L$, then

$$\lim_{x\to\alpha} \frac{f(x)}{g(x)} = L$$

Type ∞/∞**:** If $\lim_{x\to\alpha} f(x) = \pm\infty$, $\lim_{x\to\alpha} g(x) = \pm\infty$, and $\lim_{x\to\alpha} \frac{f'(x)}{g'(x)} = L$, then

$$\lim_{x\to\alpha} \frac{f(x)}{g(x)} = L$$

Works for $\alpha = \pm\infty$ and $L = \pm\infty$

8.6 Taylor's Theorem

If f is n times differentiable:

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + R_n(x)$$

Lagrange Remainder: $\exists\zeta$ between x_0 and x:

$$R_n(x) = \frac{f^{(n)}(\zeta)}{n!}(x-x_0)^n$$

Cauchy Remainder:

$$R_n(x) = \int_{x_0}^x \frac{(x-t)^{n-1}}{(n-1)!} f^{(n)}(t)dt$$

Taylor Series: If $\lim_{n\to\infty} R_n(x) = 0$:

$$f(x) = \sum_{n=0}^\infty \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n$$

9 9. Riemann Integration

9.1 Definitions

Partition: $P = \{a = x_0 < x_1 < \cdots < x_n = b\}$
Norm: $\|P\| = \max(x_i - x_{i-1})$
Tagged partition: Choose $t_i \in [x_{i-1}, x_i]$
Riemann sum: $S(f; \dot{P}) = \sum_{i=1}^n f(t_i)(x_i - x_{i-1})$
f is **Riemann integrable** if $\exists L : \forall\epsilon > 0, \exists\delta > 0$ such that $|S(f;\dot{P}) - L| < \epsilon$ whenever $\|\dot{P}\| < \delta$
Write $L = \int_a^b f(x)dx$

9.2 Darboux Sums

Upper sum: $U(f;P) = \sum \sup_{[x_{i-1},x_i]} f \cdot (x_i - x_{i-1})$
Lower sum: $L(f;P) = \sum \inf_{[x_{i-1},x_i]} f \cdot (x_i - x_{i-1})$
$U(f) = \inf_P U(f;P)$; $L(f) = \sup_P L(f;P)$
Always: $L(f;P) \leq L(f) \leq U(f) \leq U(f;P)$
Criterion: f Riemann integrable $\Leftrightarrow U(f) = L(f)$
$\Leftrightarrow \forall\epsilon > 0, \exists P : U(f;P) - L(f;P) < \epsilon$

9.3 Integrability

Riemann integrable functions are bounded
Monotone $\Rightarrow$ integrable
Continuous $\Rightarrow$ integrable
Lebesgue Criterion: Bounded f integrable $\Leftrightarrow$ discontinuity set has measure zero

9.4 Properties

If f, g integrable on $[a,b]$:
$\int_a^b (cf) = c\int_a^b f$
$\int_a^b (f+g) = \int_a^b f + \int_a^b g$
$f \leq g \Rightarrow \int_a^b f \leq \int_a^b g$
$|f|$ integrable and $|\int_a^b f| \leq \int_a^b |f|$
fg integrable
$\int_a^b f = \int_a^c f + \int_c^b f$ for $c \in (a,b)$

9.5 Fundamental Theorems of Calculus

FTC I: If F continuous on $[a,b]$, differentiable on (a,b), and F' integrable:

$$\int_a^b F'(x)dx = F(b) - F(a)$$

FTC II: If f integrable on $[a,b]$, define $F(x) = \int_c^x f(t)dt$: - F is continuous on $[a,b]$ - If f continuous at x_0, then $F'(x_0) = f(x_0)$

9.6 Techniques

Integration by parts: If u, v continuous, u', v' integrable:

$$\int_a^b uv' + \int_a^b u'v = u(b)v(b) - u(a)v(a)$$

Substitution: If f continuous, u continuously differentiable:

$$\int_a^b f(u(x))u'(x)dx = \int_{u(a)}^{u(b)} f(t)dt$$

MVT for Integrals: If f continuous on $[a,b]$, $\exists\zeta \in [a,b]$:

$$f(\zeta) = \frac{1}{b-a}\int_a^b f(x)dx$$

9.7 Integration of Series

If $(f_n) \rightrightarrows f$ uniformly on $[a,b]$ and each f_n integrable:
f is integrable and $\lim_{n\to\infty} \int_a^b f_n = \int_a^b f$
For power series: can integrate term-by-term within radius of convergence

10 10. Quick Reference

10.1 Important Limits

$\lim_{n\to\infty} \frac{1}{n} = 0$
$\lim_{n\to\infty} n^{1/n} = 1$
$\lim_{n\to\infty} c^{1/n} = 1$ for $c > 0$
$\lim_{n\to\infty} \frac{c^n}{n!} = 0$ for any c
$\lim_{x\to 0} \frac{\sin x}{x} = 1$
$\lim_{x\to\infty} (1 + 1/x)^x = e$

10.2 Common Series

Geometric: $\sum_{n=0}^\infty r^n = \frac{1}{1-r}$ if $|r| < 1$
Harmonic: $\sum_{n=1}^\infty \frac{1}{n}$ diverges
p-series: $\sum_{n=1}^\infty \frac{1}{n^p}$ converges iff $p > 1$

10.3 Key Inequalities

Bernoulli: $(1+x)^n \geq 1 + nx$ for $x > -1, n \in \mathbb{N}$
AM-GM: $\frac{a_1 + \cdots + a_n}{n} \geq \sqrt[n]{a_1 \cdots a_n}$
Cauchy-Schwarz: $|\sum a_i b_i| \leq \sqrt{\sum a_i^2}\sqrt{\sum b_i^2}$

10.4 Common Mistakes

1. Pointwise convergence $\not\Rightarrow$ uniform convergence. Example: $f_n(x) = x^n$ on $[0,1]$. Then $f_n(x) \to f(x) = \begin{cases} 0, & x < 1 \\ 1, & x = 1 \end{cases}$. Limit f is discontinuous $\Rightarrow$ convergence not uniform.
2. Continuous $\not\Rightarrow$ differentiable. Example: $f(x) = |x|$. Continuous everywhere, not differentiable at $x = 0$. **3. Differentiable $\not\Rightarrow f'$ continuous.** Example: $f(x) = \begin{cases} x^2\sin(1/x), & x \neq 0 \\ 0, & x = 0 \end{cases}$. Then $f'(x) = 2x\sin(1/x) - \cos(1/x)$ for $x \neq 0$, $f'(0) = 0$. f' exists but oscillates wildly near 0 (not continuous).

4. $\sum a_n$ converges $\not\Rightarrow \sum a_n^2$ converges. Example: $a_n = \frac{(-1)^n}{\sqrt{n}}$. $\sum a_n$ converges (Alternating Series Test), but $\sum a_n^2 = \sum \frac{1}{n}$ diverges (harmonic series).
5. Ratio test inconclusive when limit $= 1$**.** Examples: (a) $a_n = \frac{1}{n}$: $\frac{a_{n+1}}{a_n} \to 1$, yet $\sum a_n$ diverges. (b) $a_n = \frac{1}{n^2}$: $\frac{a_{n+1}}{a_n} \to 1$, yet $\sum a_n$ converges.
$\Rightarrow$ Ratio test gives no information when the limit equals 1.

Multivariable Calculus

HKPFS Math PhD Interview Prep

1 1. Vectors & Geometry

1.1 Dot Product

$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n$
$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\|\|\mathbf{v}\|\cos\theta$
$\|\mathbf{v}\| = \sqrt{\mathbf{v}\cdot\mathbf{v}}$
Orthogonal: $\mathbf{u} \perp \mathbf{v} \Leftrightarrow \mathbf{u}\cdot\mathbf{v} = 0$
Projection of u onto v:
$\mathrm{proj}_{\mathbf{v}}\mathbf{u} = \left(\frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{v}\|^2}\right)\mathbf{v}$

1.2 Cross Product (in $\mathbb{R}^3$)

$\mathbf{u}\times\mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}$
$= (u_2v_3 - u_3v_2, u_3v_1 - u_1v_3, u_1v_2 - u_2v_1)$
$\|\mathbf{u}\times\mathbf{v}\| = \|\mathbf{u}\|\|\mathbf{v}\|\sin\theta$ (area of parallelogram)
Properties: $\mathbf{u}\times\mathbf{v} = -\mathbf{v}\times\mathbf{u}$, $\mathbf{u}\times\mathbf{u} = \mathbf{0}$
Scalar triple product:
$(\mathbf{u}\times\mathbf{v})\cdot\mathbf{w} = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$ (volume)

1.3 Lines & Planes

Line through point P parallel to v:
$\mathbf{r}(t) = \mathbf{p} + t\mathbf{v}$
Plane with normal $\mathbf{n} = (A, B, C)$ through (a, b, c):
$A(x-a) + B(y-b) + C(z-c) = 0$
General form: $Ax + By + Cz = D$

1.4 Coordinate Systems

Polar: $x = r\cos\theta$, $y = r\sin\theta$
$r = \sqrt{x^2+y^2}$, $\tan\theta = \frac{y}{x}$
Cylindrical: (r, θ, z) where (r, θ) is polar in xy-plane
Spherical: (ρ, ϕ, θ)

$$x = \rho\sin\phi\cos\theta, \quad y = \rho\sin\phi\sin\theta$$
$$z = \rho\cos\phi$$
$$\rho = \sqrt{x^2+y^2+z^2}, \quad \tan\phi = \frac{\sqrt{x^2+y^2}}{z}$$

2 2. Functions & Limits

2.1 Multivariable Functions

Domain: Set $X \subseteq \mathbb{R}^n$ where f is defined
Range: $\{f(\mathbf{x}) : \mathbf{x} \in X\}$
Level curve at height c: $\{(x, y) : f(x, y) = c\}$
Graph: $\{(x, y, f(x, y)) : (x, y) \in X\} \subseteq \mathbb{R}^3$

2.2 Quadric Surfaces

Ellipsoid: $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$
Elliptic paraboloid: $\frac{z}{c} = \frac{x^2}{a^2} + \frac{y^2}{b^2}$
Hyperbolic paraboloid: $\frac{z}{c} = \frac{y^2}{b^2} - \frac{x^2}{a^2}$
Hyperboloid (1 sheet): $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$
Hyperboloid (2 sheets): $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1$

2.3 Topology

Open ball: $B(\mathbf{a}, r) = \{\mathbf{x} : \|\mathbf{x}-\mathbf{a}\| < r\}$
Closed ball: $\overline{B}(\mathbf{a}, r) = \{\mathbf{x} : \|\mathbf{x}-\mathbf{a}\| \leq r\}$
Interior point: $\mathbf{a} \in X$ is interior if $\exists r > 0 : B(\mathbf{a}, r) \subseteq X$
Boundary point: Every ball around **a** contains points in X and not in X
Open set: Every point is an interior point
Closed set: Complement is open (equivalently: contains all boundary points)
Bounded: $\exists M > 0 : \|\mathbf{x}\| \leq M$ for all $\mathbf{x} \in X$
Compact: Closed and bounded

2.4 Limits & Continuity

Limit: $\lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x}) = L$ if $\forall\varepsilon > 0, \exists\delta > 0 :$
$0 < \|\mathbf{x}-\mathbf{a}\| < \delta \Rightarrow \|f(\mathbf{x}) - L\| < \varepsilon$
Continuous at a: $\lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x}) = f(\mathbf{a})$
Showing limit DNE: Approach along different paths and get different limits
Sandwich theorem: If $f \leq g \leq h$ and $\lim f = \lim h = L$, then $\lim g = L$

3 3. Differentiation

3.1 Partial Derivatives

$\frac{\partial f}{\partial x_j}(\mathbf{a}) = \lim_{h\to 0}\frac{f(\mathbf{a}+h\mathbf{e}_j) - f(\mathbf{a})}{h}$
Notation: $f_{x_j}(\mathbf{a})$, $D_{x_j}f(\mathbf{a})$, $\frac{\partial f}{\partial x_j}$
Higher order: $f_{x_ix_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$
Clairaut's theorem: If $f \in C^k$, then mixed partials commute: $f_{x_ix_j} = f_{x_jx_i}$

3.2 Gradient & Derivative

Gradient: $\nabla f(\mathbf{a}) = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right)$
Derivative (Jacobian): For $f : \mathbb{R}^n \to \mathbb{R}^m$, $f = (f_1, \ldots, f_m)$:
$Df = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$
For scalar f: $Df = (\nabla f)^T$ (row vector)

3.3 Differentiability

f is **differentiable at a** if $\nabla f(\mathbf{a})$ exists and
$\lim_{\mathbf{x}\to\mathbf{a}} \frac{f(\mathbf{x}) - [f(\mathbf{a}) + \nabla f(\mathbf{a})\cdot(\mathbf{x}-\mathbf{a})]}{\|\mathbf{x}-\mathbf{a}\|} = 0$
Sufficient condition: If all partial derivatives exist and are continuous in a neighborhood of **a**, then f is differentiable at **a**
Class C^k: All partial derivatives up to order k exist and are continuous
Differentiable $\Rightarrow$ continuous (but not conversely)

3.4 Tangent Plane

Tangent plane to $z = f(x, y)$ at $(a, b, f(a, b))$:
$z = f(a, b) + f_x(a, b)(x-a) + f_y(a, b)(y-b)$
For surface $F(x, y, z) = c$: Normal vector is $\nabla F = (F_x, F_y, F_z)$

3.5 Directional Derivative

$D_{\mathbf{u}}f(\mathbf{a}) = \lim_{h\to 0}\frac{f(\mathbf{a}+h\mathbf{u}) - f(\mathbf{a})}{h}$
where **u** is a unit vector.
If f is differentiable: $D_{\mathbf{u}}f(\mathbf{a}) = \nabla f(\mathbf{a})\cdot\mathbf{u}$
Maximum rate of increase: Direction of ∇f, rate $= \|\nabla f\|$
Minimum rate: Direction of $-\nabla f$, rate $= -\|\nabla f\|$
∇f is perpendicular to level curves/surfaces

3.6 Chain Rule

For $h = f \circ \mathbf{g}$ where $f : \mathbb{R}^n \to \mathbb{R}^k$, $\mathbf{g} : \mathbb{R}^m \to \mathbb{R}^n$:
$Dh(\mathbf{t}) = Df(\mathbf{g}(\mathbf{t}))\cdot D\mathbf{g}(\mathbf{t})$
Special case: $z = f(x, y)$, $x = x(t)$, $y = y(t)$:
$\frac{dz}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}$
Polar coordinates conversion:
$\frac{\partial f}{\partial r} = \cos\theta\frac{\partial f}{\partial x} + \sin\theta\frac{\partial f}{\partial y}$
$\frac{\partial f}{\partial \theta} = -r\sin\theta\frac{\partial f}{\partial x} + r\cos\theta\frac{\partial f}{\partial y}$

4 4. Optimization

4.1 Taylor's Theorem

First order: $f(\mathbf{x}) = f(\mathbf{a}) + Df(\mathbf{a})(\mathbf{x}-\mathbf{a}) + R_1(\mathbf{x}, \mathbf{a})$
where $\lim_{\mathbf{x}\to\mathbf{a}}\frac{R_1(\mathbf{x},\mathbf{a})}{\|\mathbf{x}-\mathbf{a}\|} = 0$
Second order:
$f(\mathbf{x}) = f(\mathbf{a}) + Df(\mathbf{a})(\mathbf{x}-\mathbf{a}) + \frac{1}{2}(\mathbf{x}-\mathbf{a})^T H_f(\mathbf{a})(\mathbf{x}-\mathbf{a}) + R_2$
Hessian matrix:
$H_f = \begin{pmatrix} f_{x_1x_1} & \cdots & f_{x_1x_n} \\ \vdots & \ddots & \vdots \\ f_{x_nx_1} & \cdots & f_{x_nx_n} \end{pmatrix}$
Total differential: $df(\mathbf{a}, \mathbf{h}) = Df(\mathbf{a})\mathbf{h}$
Incremental change:
$\Delta f = f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) \approx df(\mathbf{a}, \mathbf{h})$

4.2 Extrema

Critical point: $\nabla f(\mathbf{a}) = \mathbf{0}$ or $\nabla f(\mathbf{a})$ DNE
Necessary condition: If **a** is interior local extremum and $\nabla f(\mathbf{a})$ exists, then $\nabla f(\mathbf{a}) = \mathbf{0}$
Saddle point: Critical point that is not a local extremum

4.3 Second Derivative Test

Let **a** be a critical point of $f \in C^2$.

- If $H_f(\mathbf{a})$ is positive definite $\Rightarrow$ local min
- If $H_f(\mathbf{a})$ is negative definite $\Rightarrow$ local max
- If $\det H_f(\mathbf{a}) \neq 0$ and $H_f(\mathbf{a})$ is indefinite $\Rightarrow$ saddle point

Sylvester's criterion: Let $d_k = \det(H_k)$ (leading principal minors)

- Positive definite: $d_k > 0$ for all $k = 1, \ldots, n$
- Negative definite: $d_k < 0$ for odd k, $d_k > 0$ for even k

For $n = 2$: Let $D = f_{xx}f_{yy} - (f_{xy})^2$

- $D > 0, f_{xx} > 0 \Rightarrow$ local min
- $D > 0, f_{xx} < 0 \Rightarrow$ local max
- $D < 0 \Rightarrow$ saddle point
- $D = 0 \Rightarrow$ inconclusive

4.4 Extreme Value Theorem

If $f : X \to \mathbb{R}$ is continuous and X is compact, then f attains global max and min on X
Strategy: Find critical points in interior, then check boundary

4.5 Lagrange Multipliers

To optimize $f(\mathbf{x})$ subject to $g(\mathbf{x}) = c$:
At extremum **a** (if $\nabla g(\mathbf{a}) \neq \mathbf{0}$):
$\exists\lambda : \nabla f(\mathbf{a}) = \lambda\nabla g(\mathbf{a})$
Multiple constraints $g_1 = c_1, \ldots, g_k = c_k$ (with $\{\nabla g_j\}$ linearly indep.):
$\nabla f = \lambda_1\nabla g_1 + \cdots + \lambda_k\nabla g_k$
Procedure:

1. Solve $\nabla f = \lambda\nabla g$ and $g = c$ simultaneously
2. Evaluate f at all solutions
3. Compare to find max/min

5 5. Integration

5.1 Double Integrals

Definition:
$\iint_R f\,dA = \lim_{\Delta x,\Delta y\to 0}\sum_{i,j} f(\mathbf{x}_{ij})\Delta x_i \Delta y_j$
Fubini's theorem (rectangle): $R = [a,b]\times[c,d]$
$\iint_R f\,dA = \int_a^b\int_c^d f(x,y)\,dy\,dx = \int_c^d\int_a^b f(x,y)\,dx\,dy$
Type 1 region: $D = \{(x,y) : a \le x \le b, g(x) \le y \le h(x)\}$
$\iint_D f\,dA = \int_a^b\int_{g(x)}^{h(x)} f(x,y)\,dy\,dx$
Type 2 region: $D = \{(x,y) : c \le y \le d, g(y) \le x \le h(y)\}$
$\iint_D f\,dA = \int_c^d\int_{g(y)}^{h(y)} f(x,y)\,dx\,dy$
Volume under graph: $V = \iint_D f(x,y)\,dA$ (if $f \ge 0$)
Area of region: $A = \iint_D 1\,dA$

5.2 Triple Integrals

$\iiint_B f\,dV = \int_a^b\int_c^d\int_p^q f(x,y,z)\,dz\,dy\,dx$
For elementary region:
$\iiint_D f\,dV = \int_a^b\int_{g(x)}^{h(x)}\int_{\varphi(x,y)}^{\psi(x,y)} f\,dz\,dy\,dx$
Volume: $V = \iiint_D 1\,dV$

5.3 Change of Variables

Jacobian: For $x = x(u,v)$, $y = y(u,v)$:
$\frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix} = x_u y_v - x_v y_u$
Change of variables formula:
$\iint_D f(x,y)\,dx\,dy = \iint_{D'} f(x(u,v),y(u,v))\left|\frac{\partial(x,y)}{\partial(u,v)}\right|du\,dv$
Polar coordinates: $x = r\cos\theta$, $y = r\sin\theta$
$dA = dx\,dy = r\,dr\,d\theta$
Cylindrical: $x = r\cos\theta$, $y = r\sin\theta$, $z = z$
$dV = r\,dr\,d\theta\,dz$
Spherical: $x = \rho\sin\phi\cos\theta$, $y = \rho\sin\phi\sin\theta$, $z = \rho\cos\phi$
$dV = \rho^2\sin\phi\,d\rho\,d\phi\,d\theta$
For $\mathbb{R}^3$:
$\frac{\partial(x,y,z)}{\partial(u,v,w)} = \begin{vmatrix} x_u & x_v & x_w \\ y_u & y_v & y_w \\ z_u & z_v & z_w \end{vmatrix}$

6 6. Vector Calculus

6.1 Curves & Paths

Path: $\boldsymbol{\gamma} : I \to \mathbb{R}^n$ continuous, $I \subseteq \mathbb{R}$ interval
Curve: Image $\boldsymbol{\gamma}(I)$
Velocity: $\mathbf{v}(t) = \boldsymbol{\gamma}'(t)$
Speed: $\|\mathbf{v}(t)\| = \|\boldsymbol{\gamma}'(t)\|$
Arc length: $L = \int_a^b \|\boldsymbol{\gamma}'(t)\|\,dt$
Parametrization: Injective C^1 path with image C
Reparametrization: $\boldsymbol{\gamma}_2 = \boldsymbol{\gamma}_1 \circ \phi$
where $\phi : [c,d] \to [a,b]$ bijective C^1
Orientation-preserving if $\phi(c) = a$, $\phi(d) = b$
Orientation-reversing if $\phi(c) = b$, $\phi(d) = a$

6.2 Differential Operators

Del operator: $\nabla = \frac{\partial}{\partial x_1}\mathbf{e}_1 + \cdots + \frac{\partial}{\partial x_n}\mathbf{e}_n$
Gradient: $\nabla f = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right)$
(scalar $\to$ vector)
Divergence:
$\nabla\cdot\mathbf{F} = \text{div}\,\mathbf{F} = \frac{\partial F_1}{\partial x_1} + \cdots + \frac{\partial F_n}{\partial x_n}$
(vector $\to$ scalar)
Curl (in $\mathbb{R}^3$):
$\nabla\times\mathbf{F} = \text{curl}\,\mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_1 & F_2 & F_3 \end{vmatrix}$
$= \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}\right)\mathbf{i} + \left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}\right)\mathbf{j} + \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)\mathbf{k}$
Key identities:
- $\nabla\times(\nabla f) = \mathbf{0}$ (curl of gradient is zero)
- $\nabla\cdot(\nabla\times\mathbf{F}) = 0$ (divergence of curl is zero)

Conservative field: $\mathbf{F} = \nabla f$ for some scalar f

6.3 Line Integrals

Scalar line integral:
$\int_{\boldsymbol{\gamma}} f\,ds = \int_a^b f(\boldsymbol{\gamma}(t))\|\boldsymbol{\gamma}'(t)\|\,dt$
Independent of orientation
Vector line integral:
$\int_{\boldsymbol{\gamma}} \mathbf{F}\cdot d\mathbf{s} = \int_a^b \mathbf{F}(\boldsymbol{\gamma}(t))\cdot\boldsymbol{\gamma}'(t)\,dt$
Also written: $\int_{\boldsymbol{\gamma}} F_1\,dx_1 + \cdots + F_n\,dx_n$
Orientation: Reversing orientation changes sign of vector line integral but not scalar
Closed curve notation: $\oint_C f\,ds$, $\oint_C \mathbf{F}\cdot d\mathbf{s}$

6.4 Surface Integrals

Parametrized surface: $\boldsymbol{\Phi} : D \subseteq \mathbb{R}^2 \to \mathbb{R}^3$, D open connected
Tangent vectors: $\mathbf{T}_s = \frac{\partial\boldsymbol{\Phi}}{\partial s}$, $\mathbf{T}_t = \frac{\partial\boldsymbol{\Phi}}{\partial t}$
Normal vector: $\mathbf{N}(s,t) = \mathbf{T}_s \times \mathbf{T}_t$
Smooth surface: $\mathbf{N} \neq \mathbf{0}$ everywhere
Surface area: $A = \iint_D \|\mathbf{N}(s,t)\|\,ds\,dt$
Scalar surface integral:
$\iint_{\boldsymbol{\Phi}} f\,dS = \iint_D f(\boldsymbol{\Phi}(s,t))\|\mathbf{N}(s,t)\|\,ds\,dt$
Vector surface integral:
$\iint_{\boldsymbol{\Phi}} \mathbf{F}\cdot d\mathbf{S} = \iint_D \mathbf{F}(\boldsymbol{\Phi}(s,t))\cdot\mathbf{N}(s,t)\,ds\,dt$
Orientable surface: Can define continuous unit normal everywhere
Closed surface notation: $\oint_S f\,dS$, $\oint_S \mathbf{F}\cdot d\mathbf{S}$

6.5 Fundamental Theorems

Green's theorem: Let $C = \partial D$ positively oriented, $\mathbf{F} = (F_1, F_2)$:
$\oint_C \mathbf{F}\cdot d\mathbf{s} = \oint_C F_1\,dx + F_2\,dy$
$= \iint_D \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)dx\,dy$
Equivalently: $\oint_C \mathbf{F}\cdot d\mathbf{s} = \iint_D (\nabla\times\mathbf{F})\cdot\mathbf{k}\,dA$
Divergence theorem ($\mathbb{R}^2$): Let $C = \partial D$, $\mathbf{n}$ outward normal:
$\oint_C \mathbf{F}\cdot\mathbf{n}\,ds = \iint_D \nabla\cdot\mathbf{F}\,dA$
Stokes' theorem: Let S be orientable surface, ∂S oriented consistently:
$\iint_S (\nabla\times\mathbf{F})\cdot d\mathbf{S} = \oint_{\partial S} \mathbf{F}\cdot d\mathbf{s}$
Gauss/Divergence theorem: Let D be solid region, ∂D oriented outward:
$\oint_{\partial D} \mathbf{F}\cdot d\mathbf{S} = \iiint_D \nabla\cdot\mathbf{F}\,dV$

6.6 Important Facts

- Positively oriented boundary: D on left when traversing C
- Right-hand rule for consistent orientation on surface
- Green's is 2D Stokes
- For conservative field $\mathbf{F} = \nabla f$:
  $\int_{\boldsymbol{\gamma}} \mathbf{F}\cdot d\mathbf{s} = f(\text{end}) - f(\text{start})$ (path-independent)

Mathematics Interview Preparation

1 Linear Algebra

1.1 Systems and Matrices

1. Define the **Rank** and **Nullity** of a matrix. State the **Dimension Formula** (Rank-Nullity Theorem) for a matrix $A \in M_{n\times m}$.

2. What are the conditions for a system $Ax = b$ to be **inconsistent**? Relate this to the RREF of the augmented matrix $(A|b)$.

3. List at least 5 equivalent conditions for an $n \times n$ matrix A to be **Invertible** (The Invertible Matrix Theorem).

4. Explain **Cramer's Rule**. When is it applicable?

1.2 Vector Spaces and Linear Maps

1. Define a **Subspace**. Why is the union of two subspaces generally not a subspace?

2. Define **Linear Independence**. How do you test if a set of vectors is linearly independent using a homogeneous system?

3. Let $T : V \to W$ be a linear transformation. Define the **Kernel** (Null Space) and **Image** (Range).

4. What does it mean for a linear map to be an **Isomorphism**? What does this imply about the dimensions of V and W?

5. Explain the relationship between the matrix of a transformation $[T]_{\mathcal{B}}^{\mathcal{C}}$ and a change of basis matrix Q. How does Q relate $[v]_{\mathcal{B}}$ and $[v]_{\mathcal{B}'}$?

   Quick note: The change-of-basis matrix Q converts coordinates between bases: $[v]_{\mathcal{B}} = Q[v]_{\mathcal{B}'}$, and $[v]_{\mathcal{B}'} = Q^{-1}[v]_{\mathcal{B}}$. Matrices of the same transformation in different bases are related by
   $$[T]_{\mathcal{B}'} = Q^{-1}[T]_{\mathcal{B}}Q.$$

6. Define the **Dual Space** V^*. If V is finite-dimensional, is $V \cong V^*$?

   Quick note: The dual space V^* is the set of all linear functionals $f : V \to \mathbb{F}$, forming a vector space itself. If $\dim(V) = n$, then $\dim(V^*) = n$, so V and V^* are isomorphic as vector spaces. However, this isomorphism is *not canonical*—it depends on the choice of basis (or an extra structure such as an inner product). The dual basis $\{e^1, \ldots, e^n\}$ is defined by $e^i(e_j) = \delta_{ij}$.

1.3 Inner Products and Orthogonality

1. State the properties of an **Inner Product**. What is the relationship between the inner product and the norm?

2. State the **Cauchy-Schwarz Inequality**. When does equality hold?

3. Describe the **Gram-Schmidt Process**. What is its geometric interpretation?

   Quick note: The **Gram-Schmidt process** takes a linearly independent set of vectors $\{v_1, v_2, \ldots, v_n\}$ in an inner product space and converts it into an **orthonormal set** $\{q_1, q_2, \ldots, q_n\}$ that spans the same subspace. Each new vector is made orthogonal to the previous ones by subtracting its projection components:

   $$q_1 = \frac{v_1}{\|v_1\|}, \quad q_k = \frac{v_k - \sum_{j=1}^{k-1} \langle v_k, q_j \rangle q_j}{\|v_k - \sum_{j=1}^{k-1} \langle v_k, q_j \rangle q_j\|}.$$

   Geometric interpretation: Gram–Schmidt can be seen as the process of constructing a sequence of perpendicular directions from arbitrary basis vectors — like turning skewed axes into orthogonal ones. It provides an orthonormal basis suitable for projections and for QR decomposition ($A = QR$).

4. Define the **Orthogonal Complement** $S^\perp$. What is the dimension of $W^\perp$ if W is a subspace of a finite-dimensional space V?

5. Explain the **Least Squares Problem**. Deriving from the geometry of projections, why is the solution given by the normal equation $A^T A x = A^T y$?

   Quick note: The least squares solution x_{LS} minimizes $\|Ax - y\|^2$ by projecting y onto the column space of A. The residual $r = y - Ax_{\text{LS}}$ is orthogonal to $\text{Col}(A)$, giving the condition $A^T r = 0$, which leads to the **normal equation:**

   $$A^T A x_{\text{LS}} = A^T y.$$

   When the columns of A are linearly independent, $A^T A$ is invertible and the unique solution is

   $$x_{\text{LS}} = (A^T A)^{-1} A^T y.$$

6. Under what condition is the matrix $A^T A$ invertible?

7. What is a **Positive Definite Matrix**? Give three equivalent characterizations (e.g., eigenvalues, pivots, energy).

   Quick note: A symmetric matrix $A \in \mathbb{R}^{n \times n}$ is **positive definite** if

   $$x^T A x > 0 \quad \text{for all } x \neq 0.$$

   This means it defines a strictly positive quadratic form. Several equivalent characterizations exist:

   - **Energy test:** $x^T A x > 0$ for all nonzero x (definition).
   - **Eigenvalue test:** All eigenvalues of A are positive.
   - **Pivot (principal minors) test:** All leading principal minors $\det(A_k)$ are positive.

1.4 Eigenvalues and Diagonalization

1. Define **Algebraic Multiplicity** and **Geometric Multiplicity**. What is the inequality relationship between them?

   Quick note: Algebraic multiplicity $m_a(\lambda)$ = number of times λ is a root of the characteristic polynomial $\det(A-\lambda I) = 0$. Geometric multiplicity $m_g(\lambda) = \dim(\mathrm{Ker}(A-\lambda I))$. Inequality: $1 \le m_g(\lambda) \le m_a(\lambda)$, with equality $\Rightarrow$ the eigenvalue's eigenspace is "complete."

2. What is the precise condition for a matrix A to be **Diagonalizable**?

   Quick note: A matrix A is diagonalizable $\iff$ it has n linearly independent eigenvectors $\iff$ $m_g(\lambda) = m_a(\lambda)$ for each eigenvalue. If A has n distinct eigenvalues $\Rightarrow$ automatically diagonalizable.

3. If a matrix A is symmetric ($A^T = A$), what can you say about its eigenvalues and eigenvectors?

   Quick note: Real symmetric $\Rightarrow$ all eigenvalues are real, eigenvectors for distinct eigenvalues are orthogonal, and A is orthogonally diagonalizable:

   $$A = QDQ^T, \quad Q^TQ = I.$$

   (Spectral Theorem.)

4. Explain how to compute the matrix exponential e^A using diagonalization.

   Quick note: If $A = PDP^{-1}$ with $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, then

   $$e^A = Pe^DP^{-1}, \quad \text{where } e^D = \mathrm{diag}(e^{\lambda_1}, \ldots, e^{\lambda_n}).$$

   This works since $A^k = PD^kP^{-1}$ and exponentials preserve similarity transformations.

2 Mathematical Analysis

2.1 Real Numbers and Topology

1. Define the **Supremum** (LUB). How does the **Completeness Axiom** distinguish $\mathbb{R}$ from $\mathbb{Q}$?

   Quick note: The **supremum** $\sup S$ of a nonempty, bounded-above set $S \subseteq \mathbb{R}$ is the smallest real number u with $s \le u$ for all $s \in S$. The **Completeness Axiom** states every nonempty, bounded-above subset of $\mathbb{R}$ has a supremum in $\mathbb{R}$. $\mathbb{R}$ is **complete**; $\mathbb{Q}$ is not — e.g. $\{q \in \mathbb{Q} : q^2 < 2\}$ has no supremum in $\mathbb{Q}$.

2. State the **Archimedean Property** and the **Density of** $\mathbb{Q}$.

3. What is a **Cauchy Sequence**? State the Cauchy Criterion for convergence in $\mathbb{R}$.

4. State the **Bolzano-Weierstrass Theorem**.

   Quick note: Every bounded sequence in $\mathbb{R}$ has a **convergent subsequence**. This expresses the compactness of closed, bounded intervals in $\mathbb{R}$: bounded $\Rightarrow$ at least one limit (accumulation) point exists.

5. Define **Limit Superior** ($\limsup$) and **Limit Inferior** ($\liminf$). Under what condition does a sequence converge?

2.2 Series and Convergence

1. Distinguish between **Absolute Convergence** and **Conditional Convergence**.
2. State the **Ratio Test** and **Root Test**. When are these tests inconclusive?
3. Does $\sum a_n$ converging imply $\sum a_n^2$ converges?
4. If $a_n \to 0$, does $\sum a_n$ necessarily converge? Give a counter-example.

2.3 Continuity and Differentiation

1. Define **Continuity** using the $\epsilon - \delta$ definition.
2. Define **Uniform Continuity**. How does the definition differ from pointwise continuity?
3. True or False: If f is continuous on a bounded interval (a, b), it is uniformly continuous. (Hint: Consider $f(x) = 1/x$).
4. State the **Intermediate Value Theorem** and the **Extreme Value Theorem**. What topological properties of the domain are required?

   Quick note: **Intermediate Value Theorem (IVT):** If f is continuous on $[a, b]$ and y lies between $f(a)$ and $f(b)$, then $\exists c \in (a, b)$ such that $f(c) = y$. Requires the domain to be **connected (interval)**.

   Extreme Value Theorem (EVT): If f is continuous on a closed, bounded interval $[a, b]$, then $\exists c, d \in [a, b]$ such that

   $$f(c) = \max_{x \in [a,b]} f(x), \qquad f(d) = \min_{x \in [a,b]} f(x).$$

   Requires the domain to be **compact (closed and bounded)**.
5. Does differentiability at a point imply continuity at that point? Does continuity imply differentiability? (Provide the standard counter-example for the latter).
6. State the **Mean Value Theorem**.

   Quick note: If f is continuous on $[a, b]$ and differentiable on (a, b), then

   $$\exists c \in (a, b) \text{ such that } f'(c) = \frac{f(b) - f(a)}{b - a}.$$

   Requires f to be **continuous on** $[a, b]$ and **differentiable on** (a, b).
7. Can a derivative f' exist everywhere but be discontinuous? (Hint: $x^2 \sin(1/x)$).

2.4 Sequences of Functions and Integration

1. Define **Pointwise Convergence** vs. **Uniform Convergence** of a sequence of functions (f_n).
2. Why is Uniform Convergence important for swapping limits with integrals or derivatives?

   Quick note: Uniform convergence allows limit operations to pass through continuous processes like integration and differentiation. If $f_n \to f$ **uniformly** and each f_n is integrable on $[a, b]$, then

   $$\lim_{n \to \infty} \int_a^b f_n(x)\, dx = \int_a^b \lim_{n \to \infty} f_n(x)\, dx.$$

   Uniform convergence also ensures f is continuous if each f_n is continuous. Pointwise convergence alone **does not** guarantee these properties.

3. State the **Weierstrass M-Test**.
4. Define the **Riemann Integral** using partitions and Darboux sums ($U(f, P)$ and $L(f, P)$).
5. State both parts of the **Fundamental Theorem of Calculus**.

   Quick note: **FTC Part I (Derivative of Integral):** If f is continuous on $[a, b]$ and $F(x) = \int_a^x f(t)\,dt$, then $F'(x) = f(x)$.

   FTC Part II (Integral of Derivative): If F is differentiable on $[a, b]$ with continuous derivative F', then
   $$\int_a^b F'(x)\,dx = F(b) - F(a).$$

3 Multivariable Calculus

3.1 Differentiation in $\mathbb{R}^n$

1. Define the **Gradient** ∇f. What is its geometric relationship to level surfaces?
2. True or False: If all partial derivatives exist at a point, the function is differentiable at that point.

   Quick note: **False.** Existence of all partials does not guarantee differentiability — they must fit together to form a linear approximation. Example: $f(x, y) = \frac{x^2 y}{x^2+y^2}$ at $(0, 0)$ has partials 0, but f isn't differentiable.
3. State the **Inverse Function Theorem**. What does the Jacobian determinant tell you about local invertibility?

   Quick note: If $f : \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable near a and $\det(Df(a)) \neq 0$, then there exist neighborhoods V of a and W of $f(a)$ such that $f : V \to W$ is a bijection with a continuously differentiable inverse f^{-1}.
   $$D(f^{-1})(f(a)) = [Df(a)]^{-1}.$$
4. State **Clairaut's Theorem** regarding mixed partial derivatives.

   Quick note: If f_{xy} and f_{yx} exist and are **continuous** near a point (a, b), then
   $$f_{xy}(a, b) = f_{yx}(a, b).$$
5. Explain the method of **Lagrange Multipliers**. Why do we solve $\nabla f = \lambda \nabla g$?
6. How do you classify critical points using the **Hessian Matrix** (Second Derivative Test)? Relate this to positive/negative definiteness.

3.2 Integration and Vector Calculus

1. State **Fubini's Theorem**. When can you swap the order of integration?

   Quick note: If $f(x, y)$ is continuous (or absolutely integrable) on a rectangular region, then
   $$\int_a^b \int_c^d f(x, y)\,dy\,dx = \int_c^d \int_a^b f(x, y)\,dx\,dy.$$

   Order of integration may be swapped when f is continuous or $|f|$ is integrable (Fubini/Tonelli).

2. Explain the **Change of Variables Formula** for multiple integrals. What is the role of the Jacobian determinant?

   Quick note: For a transformation $T(u, v) = (x, y)$,

$$\int_V f(x,y)\,dA = \int_U f(T(u,v))\,\big|\det J_T(u,v)\big|\,du\,dv.$$

   The Jacobian determinant $\big|\det J_T\big|$ gives the local area or volume scaling factor caused by the coordinate change. Examples: polar ($dA = r\,dr\,d\theta$), spherical ($dV = \rho^2 \sin\phi\,d\rho\,d\phi\,d\theta$).

3. Define the **Curl** and **Divergence** of a vector field.

   Quick note:

$$\nabla \cdot \mathbf{F} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}, \quad \nabla \times \mathbf{F} = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}, \frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}, \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right).$$

   $\nabla \cdot \mathbf{F}$ (divergence) is a scalar measuring outflow or source strength. $\nabla \times \mathbf{F}$ (curl) is a vector measuring local rotation.

4. State **Green's Theorem**.

   Quick note:

$$\oint_C (P\,dx + Q\,dy) = \iint_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dA.$$

   Relates a line integral around a closed curve to a double integral over the enclosed region (circulation = curl form in $\mathbb{R}^2$).

5. State the **Divergence Theorem** (Gauss's Theorem).

   Quick note:

$$\oint_{\partial V} \mathbf{F} \cdot \mathbf{n}\,dS = \iiint_V (\nabla \cdot \mathbf{F})\,dV.$$

6. State **Stokes' Theorem**.

   Quick note:

$$\oint_{\partial S} \mathbf{F} \cdot d\mathbf{r} = \iint_S (\nabla \times \mathbf{F}) \cdot \mathbf{n}\,dS.$$

7. What is a **Conservative Vector Field**? How is it related to path independence of line integrals?

   Quick note: A vector field $\mathbf{F}$ is conservative if there exists a scalar potential ϕ such that $\mathbf{F} = \nabla\phi$. Then $\int_C \mathbf{F} \cdot d\mathbf{r}$ is path-independent and $\oint_C \mathbf{F} \cdot d\mathbf{r} = 0$ for every closed curve C. In simply connected regions, $\mathbf{F}$ is conservative if and only if $\nabla \times \mathbf{F} = \mathbf{0}$.

4 Topology and Metric Spaces

1. Define a **Metric Space**. What are the three properties of a distance function?

2. Define **Open** and **Closed** sets. Can a set be both? Can a set be neither?

3. What is the definition of **Compactness** (using open covers)?

   Quick note: A set K in a topological space is **compact** if every open cover of K has a finite subcover.

4. State the **Heine-Borel Theorem**. In which spaces does it apply?

   Quick note: In $\mathbb{R}^n$, a set is **compact** if and only if it is **closed and bounded**. This equivalence (the Heine-Borel property) holds only in finite-dimensional Euclidean spaces. In infinite-dimensional spaces, closed and bounded sets need not be compact (for example, the unit ball in ℓ^2).

5. Define **Connectedness**.

   Rigorous definition: A topological space X is said to be **connected** if there do not exist two nonempty disjoint open sets $U, V \subseteq X$ such that

   $$X = U \cup V.$$

   Equivalently, X is connected if the only subsets of X that are both open and closed (clopen sets) are $\varnothing$ and X itself. If such a decomposition $X = U \cup V$ exists with both U and V nonempty and open (or equivalently, both closed), then X is said to be **disconnected**.

5 The "Trap" Section: Common Counter-Examples

1. **Trap:** If a sequence of continuous functions f_n converges pointwise to f on $[0,1]$, is f continuous?

   Counterexample: Let

   $$f_n(x) = x^n \quad \text{on } [0,1].$$

   Each f_n is continuous, but $f_n \to f$ pointwise where

   $$f(x) = \begin{cases} 0, & 0 \le x < 1, \\ 1, & x = 1. \end{cases}$$

   The limit function f is *not* continuous at $x = 1$. Pointwise convergence of continuous functions does not, in general, preserve continuity.

2. **Trap:** Is the union of an infinite number of closed sets always closed?

   Counterexample: Define

   $$F_n = \left[\frac{1}{n}, 1\right].$$

   Each F_n is closed, but

   $$\bigcup_{n=1}^{\infty} F_n = (0,1].$$

3. **Trap:** If a function has a local minimum at a, is the Hessian matrix $H_f(a)$ always positive definite?

   Counterexample: Let

   $$f(x,y) = x^4 + y^4.$$

   At $(0,0)$, f has a local (and global) minimum, yet

   $$H_f(0,0) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

   which is *positive semidefinite*, not positive definite.