

VU LEARNING HUB

vulearninghub.xyz

HANDOUT

MTH5206

Linear Algebra

MTH5206 by VU Learning HUB

Prepared for VU students • Free study resources

<https://vulearninghub.xyz>

COMPREHENSIVE STUDY HANDOUT

MTH5206

Linear Algebra

Lecture-wise study notes prepared for Virtual University of Pakistan students.
Covers all major topics with objectives, key definitions, summaries and exam pointers.

9 study units

Table of Contents

Placeholder for table of contents

0

Unit 1: Systems of Linear Equations

Learning Objectives

- Objective 1:** Solve linear systems by elimination

Objective 2: Use matrices and row reduction

Objective 3: Classify systems: unique, infinite, or no solution

A linear system is a set of equations linear in the unknowns; its solution set is what linear algebra exists to find. Gaussian elimination, systematically eliminating variables, reduces any system to row-echelon form, from which back-substitution yields solutions. The augmented matrix packages the system, letting us operate on coefficients alone.

Row reduction (Gauss-Jordan) to reduced row-echelon form exposes the solution directly. Pivots mark basic variables; non-pivot columns give free variables. The classification follows: a pivot in every variable column means a unique solution; free variables mean infinitely many; a row $[0 \dots 0 \mid b]$ with b nonzero means no solution (inconsistent).

Applications abound: network flows, economic models, curve fitting. Very important: row operations (swap, scale, add multiples) are the course's fundamental skill; every later topic, from inverses to eigenvalues, rests on fluent elimination. Practice until it is automatic.

Key Definitions

Row-echelon form is defined as a matrix with leading entries stairstepping right, zeros below each.

Pivot is defined as the leading nonzero entry of a row in echelon form.

Consistent system is defined as a linear system possessing at least one solution.

Key Points to Remember

- Gaussian elimination $\rightarrow$ echelon form $\rightarrow$ back-substitution.
- Augmented matrix; row ops: swap, scale, add multiples.
- Pivots in all variable columns: unique; free variables: infinite; $[0 \dots 0 \mid b]$: none.
- Elimination fluency underlies the entire course.

Unit Summary

Gaussian elimination classifies and solves linear systems through row reduction, with pivots revealing the solution structure.

Exam Pointer

Very important: solve systems by elimination and classify them; reduced row-echelon fluency is assumed everywhere later.

Unit 2: Matrix Algebra

Learning Objectives

Objective 1: Perform matrix operations

Objective 2: Find inverses and use them

Objective 3: Work with transposes and special matrices

Matrices add and scale entrywise; multiplication is defined by rows-times-columns, requiring compatible dimensions, and is not commutative: AB differs from BA in general. Matrix multiplication encodes composition of linear transformations, which is why order matters. Partitioned and block multiplication extend the rules.

The inverse A^{-1} satisfies $AA^{-1} = I$; it exists exactly when A is square with full pivot rank (nonzero determinant). Compute it by row-reducing $[A \mid I]$ to $[I \mid A^{-1}]$. Inverses solve systems ($x = A^{-1}b$), though elimination is usually cheaper. Properties: $(AB)^{-1} = B^{-1}A^{-1}$, $(A^{-1})^{-1} = A$.

The transpose A^T flips rows and columns; symmetric matrices ($A^T = A$) recur in optimization and statistics. Special forms, diagonal, triangular, orthogonal, have computational advantages. Very important: non-commutativity is the classic trap; never assume $AB = BA$.

Key Definitions

Matrix inverse is defined as the matrix A^{-1} with $AA^{-1} = A^{-1}A = I$, existing for nonsingular square A .

Transpose is defined as the matrix A^T with rows and columns interchanged.

Symmetric matrix is defined as a square matrix equal to its transpose.

Key Points to Remember

- Multiplication: rows-by-columns, order matters ($AB \neq BA$ generally).
- Inverse via $[A \mid I] \rightarrow [I \mid A^{-1}]$; exists iff square and full rank.
- $(AB)^{-1} = B^{-1}A^{-1}$; $(A^T)^T = A$.
- Special forms (diagonal, triangular, symmetric, orthogonal) simplify computation.

Unit Summary

Matrix algebra's non-commutative multiplication, inverses and transposes provide the computational language of linear systems.

Exam Pointer

Very important: compute inverses by row reduction; prove or use $(AB)^{-1} = B^{-1}A^{-1}$.

Unit 3: Determinants

Learning Objectives

Objective 1: Compute determinants by expansion and row reduction

Objective 2: Apply determinant properties

Objective 3: Use Cramer's rule and adjugates

The determinant is a scalar encoding of a square matrix: nonzero exactly when the matrix is invertible, its magnitude measuring volume scaling. For 2×2 , $ad - bc$; for larger, cofactor expansion along any row or column (Laplace), with the checkerboard sign pattern. Row reduction computes determinants efficiently: swaps flip the sign, scaling scales it, adding multiples leaves it unchanged.

Properties make theory work: $\det(AB) = \det(A)\det(B)$, $\det(A^T) = \det(A)$, and triangular matrices give the product of diagonal entries. A zero row or duplicate rows force $\det = 0$. These properties, more than raw computation, drive proofs.

Cramer's rule expresses each solution component as a ratio of determinants; elegant but computationally inferior to elimination beyond tiny systems. The adjugate formula $A^{-1} = \text{adj}(A)/\det(A)$ similarly illuminates structure. Very important: use row-reduction properties for computation; reserve expansion for small or sparse matrices.

Key Definitions

Determinant is defined as the scalar $\det(A)$ nonzero exactly for invertible matrices, measuring volume scaling.

Cofactor is defined as the signed minor $(-1)^{i+j}$ times the minor determinant.

Adjugate is defined as the transpose of the cofactor matrix; $A^{-1} = \text{adj}(A)/\det(A)$.

Key Points to Remember

- 2×2 : $ad - bc$; larger: cofactor expansion or row reduction with sign tracking.
- Row ops: swap flips sign, scale scales, replacement preserves.
- $\det(AB) = \det(A)\det(B)$; $\det(A^T) = \det(A)$; triangular $\rightarrow$ diagonal product.
- Cramer/adjugate: theoretically illuminating, computationally poor.

Unit Summary

Determinants test invertibility and scale volumes, computed via expansion or row operations and governed by multiplicative properties.

Exam Pointer

Very important: compute determinants efficiently and apply $\det(AB) = \det(A)\det(B)$ in proofs.

Unit 4: Vector Spaces

Learning Objectives

Objective 1: Verify vector space axioms

Objective 2: Test subspaces

Objective 3: Work with linear combinations and span

A vector space is a set closed under addition and scalar multiplication satisfying ten axioms (commutativity, associativity, identities, inverses, distributivity). Examples: $\mathbb{R}^n$, matrices $M_{\{m,n\}}$, polynomials P_n , function spaces. The axioms guarantee that familiar algebraic manipulations work; verifying them is the entry ticket for any new example.

A subspace is a subset closed under the operations (equivalently, containing zero and closed under linear combinations). Solution sets of homogeneous systems, spans, and eigenspaces are the canonical subspaces. The subspace test (nonempty + closed under addition and scaling) is the working tool.

Linear combinations $a_1 v_1 + \dots + a_k v_k$ and their span (all such combinations) describe reachability: the column space is the span of columns, i.e. all b for which $Ax = b$ is solvable. Very important: the subspace test and span arguments are proof staples; master the 'closed under linear combinations' phrasing.

Key Definitions

Vector space is defined as a set with addition and scalar multiplication satisfying the ten axioms.

Subspace is defined as a subset closed under addition and scalar multiplication.

Span is defined as the set of all linear combinations of given vectors.

Key Points to Remember

- Ten axioms; examples: $\mathbb{R}^n$, matrices, polynomials, functions.
- Subspace test: contains zero (or nonempty) + closed under addition and scaling.
- Span = all linear combinations; column space = solvability set of $Ax=b$.
- Homogeneous solution sets are always subspaces.

Unit Summary

Vector spaces abstract linear structure through axioms; subspaces and spans capture closure and reachability, the language of solvability.

Exam Pointer

Very important: verify subspaces with the closure test; express solution sets as spans.

Unit 5: Linear Independence, Basis and Dimension

Learning Objectives

Objective 1: Test linear independence

Objective 2: Find bases for subspaces

Objective 3: Use dimension in rank-nullity arguments

Vectors are linearly independent if only the trivial combination gives zero; dependence means redundancy, one vector is a combination of others. The pivot columns of a matrix are independent; free-variable columns depend on them. Testing independence is row reduction on the vectors as columns.

A basis is an independent spanning set: minimal enough to avoid redundancy, large enough to reach everything. Every vector then has unique coordinates relative to the basis. The standard basis $e_1, \dots, e_n$ is the default, but bases adapted to a problem (eigenbases!) simplify enormously.

Dimension is the number of vectors in any basis (well-defined by the basis theorem). Rank-nullity: $\text{rank}(A) + \text{nullity}(A) = n$ for $m \times n$ A , linking column space and null space dimensions. Very important: dimension counting proves impossibilities (no basis of $\mathbb{R}^3$ has 2 vectors) and the rank-nullity relation organizes all 'four subspaces' reasoning.

Key Definitions

Linear independence is defined as only the trivial linear combination yields the zero vector.

Basis is defined as a linearly independent spanning set; coordinates are unique.

Dimension is defined as the number of vectors in any basis of the space.

Key Points to Remember

- Independence test: row-reduce vectors as columns; pivots mark independent ones.
- Basis = independent + spanning; unique coordinates follow.
- Dimension well-defined; $\text{rank} + \text{nullity} = n$.
- Dimension arguments prove spanning/independence claims quickly.

Unit Summary

Independence, basis and dimension turn spanning sets into coordinate systems, with rank-nullity relating a matrix's fundamental subspaces.

Exam Pointer

Very important: find bases for column/null spaces and apply rank-nullity; prove dimension-based claims.

Unit 6: Linear Transformations

Learning Objectives

Objective 1: Represent transformations as matrices

Objective 2: Analyze kernel and range

Objective 3: Compose transformations and change bases

A linear transformation T preserves addition and scaling: $T(u+v) = T(u)+T(v)$, $T(cu) = cT(u)$. Every $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ is multiplication by a unique matrix A (columns are images of basis vectors); conversely every matrix defines one. Geometry becomes algebra: rotations, projections, reflections are matrices.

The kernel (null space: vectors mapping to zero) measures information lost; the range (column space) measures what is reached. T is one-to-one iff kernel is trivial, onto iff range is all of the codomain; for square matrices these coincide (invertible). The rank-nullity theorem restated: $\dim \ker + \dim \text{range} = n$.

Composition is matrix multiplication; change of basis uses $P^{-1}AP$ (similarity), previewing diagonalization. Very important: kernel/range analysis is the conceptual core; examiners ask whether a given transformation is one-to-one/onto and expect kernel-range reasoning.

Key Definitions

Linear transformation is defined as a map preserving vector addition and scalar multiplication.

Kernel is defined as the subspace of vectors mapped to zero.

Similar matrices is defined as $B = P^{-1}AP$ for invertible P ; representing one transformation in different bases.

Key Points to Remember

- T linear iff matrix multiplication; standard matrix from basis images.
- Kernel = lost information; range = reached set; 1-1 iff trivial kernel.
- Square case: 1-1, onto, invertible all equivalent.
- Composition = multiplication; change of basis = similarity $P^{-1}AP$.

Unit Summary

Linear transformations are matrices in disguise; kernel and range analyze injectivity and surjectivity, with similarity handling basis changes.

Exam Pointer

Very important: find standard matrices and test one-to-one/onto via kernel and range.

Unit 7: Eigenvalues and Eigenvectors

Learning Objectives

- Objective 1:** Compute eigenvalues and eigenvectors
- Objective 2:** Diagonalize matrices where possible
- Objective 3:** Interpret eigenvalues geometrically and dynamically

An eigenvector v of A satisfies $Av = \lambda v$: the transformation stretches v by λ , preserving its line. Eigenvalues solve the characteristic equation $\det(A - \lambda I) = 0$; each eigenvalue's eigenspace is the null space of $(A - \lambda I)$. The characteristic polynomial's degree n gives (counting multiplicity) n eigenvalues over the complex numbers.

Diagonalization writes $A = PDP^{-1}$ with D diagonal of eigenvalues and P of eigenvectors, possible exactly when n independent eigenvectors exist (guaranteed for distinct eigenvalues). Powers become trivial: $A^k = PD^kP^{-1}$. Symmetric matrices always diagonalize orthogonally (spectral theorem), vital in statistics and optimization.

Dynamically, eigenvalues govern linear recurrences and differential systems: $|\lambda| > 1$ grows, < 1 decays. Google's PageRank is an eigenvector computation. Very important: the characteristic equation and the diagonalization criterion are the unit's exam core; symmetric-matrix guarantees are bonus material.

Key Definitions

Eigenvalue/eigenvector is defined as a scalar λ and nonzero v with $Av = \lambda v$.

Characteristic equation is defined as $\det(A - \lambda I) = 0$, whose roots are the eigenvalues.

Diagonalization is defined as writing $A = PDP^{-1}$ with D diagonal; possible with n independent eigenvectors.

Key Points to Remember

- Eigenvectors: invariant lines; eigenvalues from $\det(A - \lambda I) = 0$.
- Eigenspace = $\text{null}(A - \lambda I)$; diagonalize iff n independent eigenvectors.
- $A^k = PD^kP^{-1}$; symmetric matrices diagonalize orthogonally.
- $|\lambda|$ governs growth/decay in dynamical systems; PageRank is eigenvector-based.

Unit Summary

Eigenvalues reveal a transformation's invariant directions; diagonalization exploits them to simplify powers and dynamics.

Exam Pointer

Very important: compute eigenpairs and diagonalize; state when diagonalization is possible.

Unit 8: Orthogonality and Least Squares

Learning Objectives

Objective 1: Apply Gram-Schmidt orthogonalization

Objective 2: Compute orthogonal projections

Objective 3: Solve least-squares problems

Orthogonal vectors (zero dot product) generalize perpendicularity; orthonormal sets are independent automatically and give easy coordinates (dot with each basis vector). The Gram-Schmidt process orthogonalizes any independent set, producing QR factorization $A = QR$ (Q orthonormal columns, R upper triangular), the stable engine of numerical linear algebra.

The orthogonal projection of y onto a subspace W is the closest vector in W to y ; the residual is orthogonal to W . This best-approximation property makes projection the geometric heart of least squares. Orthonormal bases make projection a sum of simple terms.

Least squares solves $Ax = b$ approximately when no exact solution exists: the normal equations $A^T A x = A^T b$ give the minimizer of $\|Ax - b\|^2$. Regression lines, curve fitting and data science rest on this. Very important: least-squares numericals and projection computations are standard; understand 'closest vector' geometrically.

Key Definitions

Orthogonal is defined as vectors with zero inner product; orthonormal adds unit length.

Orthogonal projection is defined as the closest vector in a subspace to a given vector.

Least-squares solution is defined as the x minimizing $\|Ax - b\|^2$, solving $A^T A x = A^T b$.

Key Points to Remember

- Orthonormal sets: automatic independence, easy coordinates.

- Gram-Schmidt -> QR factorization; projection = closest vector, residual orthogonal.
- Normal equations $A^T A x = A^T b$; foundation of regression.
- Geometry first: projection minimizes distance, hence 'least squares'.

Unit Summary

Orthogonality yields stable coordinates and best approximations; least squares projects data onto model spaces, founding regression.

Exam Pointer

Very important: run Gram-Schmidt, compute projections, and solve least-squares numericals via normal equations.

Unit 9: Applications: Markov Chains, Geometry and Data

Learning Objectives

- Objective 1:** Analyze Markov chains with eigen-methods
- Objective 2:** Apply linear algebra to geometry
- Objective 3:** Preview data-science uses: SVD and PCA

A Markov chain's transitions form a stochastic matrix (nonnegative columns summing to 1); state distributions evolve by multiplication. The eigenvalue 1 gives the steady-state vector: long-run behavior regardless of start (for regular chains). Population movement and market-share models are textbook applications.

Geometry is linear algebra made visible: transformations rotate, reflect, project and shear; determinants measure area/volume change (and orientation via sign); cross products and planes use dot products. Computer graphics chains these transformations via matrix multiplication.

Data science extends the course: the singular value decomposition $A = U \Sigma V^T$ generalizes diagonalization to all matrices, powering compression and latent-semantic analysis; principal component analysis (PCA) projects data onto top-variance directions, i.e. eigenvectors of the covariance matrix. Very important: the steady-state computation and the SVD/PCA statements are capstone material; exams favor Markov numericals.

Key Definitions

Stochastic matrix is defined as a nonnegative matrix with columns summing to 1; transitions of a Markov chain.

Steady-state vector is defined as the probability distribution fixed by the transition matrix (eigenvalue 1).

Singular value decomposition is defined as $A = U \Sigma V^T$ generalizing diagonalization to arbitrary matrices.

Key Points to Remember

- Markov: stochastic matrix powers; steady state = eigenvector for $\lambda=1$.

- Geometry: transformations as matrices; determinant = volume scale (+ orientation).
- SVD generalizes diagonalization; PCA = top eigenvectors of covariance.
- Applications unify the course: eigenvalues, orthogonality and transformations at work.

Unit Summary

Markov chains, geometry and data science apply eigenvalues, orthogonality and transformations, with SVD and PCA as the modern capstone.

Exam Pointer

Very important: compute a Markov steady state; state the SVD and PCA eigenvalue connection.