MIT 18.06 Linear Algebra

Complete Course Notes

Professor Gilbert Strang
Massachusetts Institute of Technology

Comprehensive Lecture Notes

Contents

Preface

These comprehensive notes cover Professor Gilbert Strang's legendary MIT 18.06 Linear Algebra course. The material emphasizes geometric intuition alongside algebraic techniques, making linear algebra accessible and applicable to numerous fields including physics, engineering, computer science, and economics.

Course Structure:

- **Unit I:** $A\mathbf{x} = \mathbf{b}$ and the Four Subspaces
- **Unit II:** Least Squares, Determinants, and Eigenvalues
- **Unit III:** Positive Definite Matrices and Applications

Part I

Unit I: $A\mathrm{x} = \mathrm{b}$ and the Four Subspaces

1 Lecture 1: The Geometry of Linear Equations

1.1 Three Perspectives on Linear Systems

Central Question

Given a system of linear equations, how do we visualize and solve it?

1.1.1 Row Picture

Each equation represents a geometric object (line in $\mathbb{R}^2$, plane in $\mathbb{R}^3$, hyperplane in $\mathbb{R}^n$). The solution is the intersection point(s).

Example 1.1. Consider the system:

$$\begin{aligned} 2x - y &= 0 \\ -x + 2y &= 3 \end{aligned}$$

In the row picture, these are two lines in the xy-plane intersecting at $(1, 2)$.

1.1.2 Column Picture

This is the most important view in linear algebra!

Rewrite the system as a linear combination of column vectors:

$$x \begin{bmatrix} 2 \\ -1 \end{bmatrix} + y \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix} \tag{1}$$

Definition 1.2 (Linear Combination)**.** A **linear combination** of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ is any vector of the form:

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$$

where $c_1, c_2, \ldots, c_n$ are scalars.

Key Question: Can we reach the right-hand side vector $\mathbf{b}$ by taking linear combinations of the column vectors?

1.1.3 Matrix Form

The compact notation:

$$A\mathbf{x} = \mathbf{b} \quad \text{where} \quad A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0 \\ 3 \end{bmatrix} \tag{2}$$

1.2 Matrix-Vector Multiplication

Definition

The product $A\mathbf{x}$ is a **linear combination of the columns of** A:

$$A\mathbf{x} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n$$

1.3 Key Insights

1. Linear algebra studies **linear combinations** of vectors
2. The fundamental question: Can we combine column vectors to produce any vector $\mathbf{b}$?
3. This leads to concepts of span, independence, and the four fundamental subspaces

2 Lecture 2: Elimination with Matrices

2.1 Gaussian Elimination

Definition 2.1 (Elimination)**.** A systematic method to solve $A\mathbf{x} = \mathbf{b}$ by converting A to upper triangular form U through elementary row operations.

2.1.1 The Algorithm

Step 1: Use the first equation to eliminate x_1 from equations $2, 3, \ldots, m$
Step 2: Use the second equation to eliminate x_2 from equations $3, 4, \ldots, m$
Step 3: Continue until upper triangular

Example 2.2. Solve:

$$\begin{aligned} x + 2y + z &= 2 \\ 3x + 8y + z &= 12 \\ 4y + z &= 2 \end{aligned}$$

Matrix form: $\begin{bmatrix} 1 & 2 & 1 \\ 3 & 8 & 1 \\ 0 & 4 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 12 \\ 2 \end{bmatrix}$

After elimination: $\begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -2 \\ 0 & 0 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 6 \\ -2 \end{bmatrix}$

2.2 Pivots

Definition 2.3 (Pivot)**.** The **pivot** is the diagonal entry used for elimination. For the elimination to proceed without row exchanges, all pivots must be non-zero.

For an $n \times n$ matrix:

- First pivot: a_{11}
- Second pivot: the $(2, 2)$ entry after first elimination
- k-th pivot: the (k, k) entry after eliminating above it

2.3 Elementary Matrices

Each elimination step can be represented as multiplication by an **elementary matrix**.

Definition 2.4 (Elementary Matrix)**.** An elementary matrix E_{ij} performs the row operation: Row i gets Row i minus ℓ_{ij} times Row j

Example 2.5. To eliminate below the first pivot:

$$E_{21} = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

where -3 is chosen to make the $(2, 1)$ entry zero.

Key Property: $E_{32}(E_{21}A) = U$

2.4 Permutation Matrices

When a pivot is zero, we must exchange rows using a **permutation matrix** P.

Example 2.6.

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{exchanges rows 1 and 2}$$

3 Lecture 3: Multiplication and Inverse Matrices

3.1 Five Ways to View Matrix Multiplication

Let $C = AB$ where A is $m \times n$ and B is $n \times p$.

3.1.1 Method 1: Entry by Entry

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = (\text{row } i \text{ of } A) \cdot (\text{column } j \text{ of } B)$$

3.1.2 Method 2: Columns of C

$$\text{Column } j \text{ of } C = A \times (\text{column } j \text{ of } B)$$

Each column of C is a **linear combination** of the columns of A.

3.1.3 Method 3: Rows of C

$$\text{Row } i \text{ of } C = (\text{row } i \text{ of } A) \times B$$

Each row of C is a linear combination of the rows of B.

3.1.4 Method 4: Column times Row (Most Important!)

$$AB = \sum_{k=1}^{n} (\text{column } k \text{ of } A)(\text{row } k \text{ of } B)$$

This expresses AB as a sum of rank-1 matrices!

Example 3.1.

$$\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \begin{bmatrix} 1 & 6 \end{bmatrix} = \begin{bmatrix} 2 & 12 \\ 3 & 18 \\ 4 & 24 \end{bmatrix}$$

This is a rank-1 matrix (all rows are multiples of each other).

3.1.5 Method 5: Block Multiplication

Partition matrices into blocks and multiply the blocks.

3.2 Matrix Inverses

Definition 3.2 (Inverse Matrix)**.** Matrix A is **invertible** (or **non-singular**) if there exists a matrix A^{-1} such that:

$$A^{-1}A = I \quad \text{and} \quad AA^{-1} = I$$

Theorem 3.3. *If A^{-1} exists, it is unique.*

3.2.1 When Does A^{-1} Exist?

- A must be square
- All columns must be independent
- All rows must be independent
- Determinant $\det(A) \neq 0$
- All pivots must be non-zero

3.3 Computing A^{-1}

Gauss-Jordan Method:

$$[A \mid I] \xrightarrow{\text{row operations}} [I \mid A^{-1}] \tag{3}$$

Example 3.4 (2 by 2 Matrix)**.** For $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$:

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

3.4 Properties of Inverses

1. $(AB)^{-1} = B^{-1}A^{-1}$ (reversal!)
2. $(A^T)^{-1} = (A^{-1})^T$
3. $(A^{-1})^{-1} = A$

4 Lecture 4: LU Decomposition

4.1 The LU Factorization

Theorem 4.1 (LU Decomposition)**.** *If elimination on A requires no row exchanges, then:*

$$A = LU$$

where L is lower triangular with ones on the diagonal, and U is upper triangular.

- U is the result of elimination (upper triangular)
- L contains the multipliers from elimination
- L has ones on the diagonal (by convention)

Example 4.2.

$$A = \begin{bmatrix} 2 & 1 \\ 8 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} = LU$$

The multiplier 4 appears in L.

4.2 Why LU is Useful

To solve $A\mathbf{x} = \mathbf{b}$:

$$\begin{aligned} A\mathbf{x} &= \mathbf{b} \\ LU\mathbf{x} &= \mathbf{b} \\ L\mathbf{c} &= \mathbf{b} \quad \text{(forward substitution, find } \mathbf{c}\text{)} \\ U\mathbf{x} &= \mathbf{c} \quad \text{(back substitution, find } \mathbf{x}\text{)} \end{aligned}$$

Both steps use triangular systems, which are fast to solve!

4.3 Computational Cost

- Computing LU: approximately $\frac{n^3}{3}$ operations
- Solving $L\mathbf{c} = \mathbf{b}$ and $U\mathbf{x} = \mathbf{c}$: approximately n^2 operations each
- **Advantage:** Factor once, solve many systems quickly

4.4 Permutation: PA = LU

If row exchanges are needed:

$$PA = LU$$

where P is a permutation matrix.

5 Lecture 5: Transposes, Permutations, Vector Spaces

5.1 Transpose

Definition 5.1 (Transpose)**.** The **transpose** of A, denoted A^T, has entries $(A^T)_{ij} = A_{ji}$.

Properties:

1. $(A^T)^T = A$
2. $(AB)^T = B^T A^T$ (reversal!)
3. $(A + B)^T = A^T + B^T$
4. $(A^{-1})^T = (A^T)^{-1}$

5.2 Symmetric Matrices

Definition 5.2 (Symmetric Matrix)**.** Matrix A is **symmetric** if $A^T = A$.

Theorem 5.3. *For any matrix R, the product $R^T R$ is always symmetric.*

Symmetric matrices are fundamental in applications (quadratic forms, optimization, physics).

5.3 Permutation Matrices

Definition 5.4 (Permutation Matrix)**.** A **permutation matrix** P is the identity matrix with reordered rows.

Key Property: $P^{-1} = P^T$ (permutations are orthogonal!)
There are $n!$ permutation matrices of size $n \times n$.

5.4 Vector Spaces

Definition 5.5 (Vector Space)**.** A **vector space** is a set of vectors closed under:

1. Addition: $\mathbf{v} + \mathbf{w}$ is in the space
2. Scalar multiplication: $c\mathbf{v}$ is in the space

Example 5.6. $\mathbb{R}^n$ is the space of all column vectors with n real components.

5.5 Subspaces

Definition 5.7 (Subspace)**.** A **subspace** of $\mathbb{R}^n$ is a subset that is itself a vector space.

Requirements:

1. Contains the zero vector
2. Closed under addition
3. Closed under scalar multiplication

Example 5.8 (Subspaces of $\mathbb{R}^3$)**.**

- All of $\mathbb{R}^3$
- Any plane through the origin
- Any line through the origin
- The zero vector alone: $\{\mathbf{0}\}$

6 Lecture 6: Column Space and Nullspace

6.1 Column Space

Definition 6.1 (Column Space)**.** The **column space** $C(A)$ consists of all linear combinations of the columns of A:

$$C(A) = \{A\mathbf{x} : \mathbf{x} \in \mathbb{R}^n\}$$

Key Question: For which vectors $\mathbf{b}$ does $A\mathbf{x} = \mathbf{b}$ have a solution?
Answer: When $\mathbf{b} \in C(A)$.

Theorem 6.2. *$C(A)$ is a subspace of $\mathbb{R}^m$ (where A is $m \times n$).*

6.2 Nullspace

Definition 6.3 (Nullspace)**.** The **nullspace** $N(A)$ consists of all solutions to $A\mathbf{x} = \mathbf{0}$:

$$N(A) = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}\}$$

Theorem 6.4. *$N(A)$ is a subspace of $\mathbb{R}^n$ (where A is $m \times n$).*

6.3 Key Examples

Example 6.5 (Invertible Matrix)**.** If A is invertible, then:

- $N(A) = \{\mathbf{0}\}$ (only the trivial solution)
- $C(A) = \mathbb{R}^n$ (all of $\mathbb{R}^n$)

7 Lecture 7: Solving $A\mathrm{x} = \mathbf{0}$

7.1 Reduced Row Echelon Form (RREF)

Definition 7.1 (RREF)**.** The **reduced row echelon form** $R = \mathrm{rref}(A)$ is obtained by:

1. Producing echelon form (forward elimination)
2. Making all pivots equal to 1
3. Eliminating above pivots (backward elimination)

7.2 Pivot and Free Variables

Definition 7.2.

- **Pivot columns**: Columns containing pivots
- **Free columns**: Columns without pivots
- **Pivot variables**: Variables corresponding to pivot columns
- **Free variables**: Variables corresponding to free columns (can be anything!)

Number of free variables $= n - r$ where r is the rank.

7.3 Computing the Nullspace

Algorithm:

1. Reduce A to $R = \mathrm{rref}(A)$
2. Identify pivot and free columns
3. For each free variable:
   - Set it to 1 (set other free variables to 0)
   - Solve for pivot variables
   - This gives a **special solution**
4. $N(A)$ = span of special solutions

Example 7.3.

$$A = \begin{bmatrix} 1 & 2 & 2 & 2 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 8 & 10 \end{bmatrix} \to R = \begin{bmatrix} 1 & 2 & 0 & -2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Pivot columns: 1, 3. Free columns: 2, 4.
Special solution 1 (set $x_2 = 1, x_4 = 0$):

$$\mathbf{s}_1 = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

Special solution 2 (set $x_2 = 0, x_4 = 1$):

$$\mathbf{s}_2 = \begin{bmatrix} 2 \\ 0 \\ -2 \\ 1 \end{bmatrix}$$

$$N(A) = \text{span}\{\mathbf{s}_1, \mathbf{s}_2\}$$

Theorem 7.4.

$$\dim(N(A)) = n - r$$

where r is the rank of A.

8 Lecture 8: Solving $A\mathbf{x} = \mathbf{b}$

8.1 Solvability Condition

Theorem 8.1 (Solvability)**.** *$A\mathbf{x} = \mathbf{b}$ is solvable if and only if $\mathbf{b} \in C(A)$.*

Practical Test: If a row of A reduces to zeros, the corresponding entry in $\mathbf{b}$ (after the same row operations) must also be zero.

8.2 Complete Solution

When $A\mathbf{x} = \mathbf{b}$ is solvable, the **complete solution** is:

$$\mathbf{x} = \mathbf{x}_p + \mathbf{x}_n$$

where:

- $\mathbf{x}_p$ = particular solution (set all free variables to 0)
- $\mathbf{x}_n$ = any vector in $N(A)$

8.3 Rank and Solutions

For $m \times n$ matrix A with rank r:

Case	Condition	Solutions
Full column rank	$r = n < m$	0 or 1 solution
Full row rank	$r = m < n$	infinitely many for every $\mathbf{b}$
Full rank (square)	$r = m = n$	Unique solution
Not full rank	$r < \min(m, n)$	0 or infinitely many

9 Lecture 9: Independence, Basis, Dimension

9.1 Linear Independence

Definition 9.1 (Linear Independence)**.** Vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are **linearly independent** if:

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = \mathbf{0} \implies c_1 = c_2 = \cdots = c_n = 0$$

Otherwise, they are **linearly dependent**.

Column perspective: Columns of A are independent if and only if $N(A) = \{\mathbf{0}\}$ if and only if rank $= n$.

9.2 Spanning

Definition 9.2 (Span)**.** Vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ **span** a space if every vector in that space can be written as their linear combination.

9.3 Basis

Definition 9.3 (Basis)**.** A **basis** for a vector space is a set of vectors that:

1. Are linearly independent
2. Span the space

Key Property: Every vector in the space has a **unique** representation as a linear combination of basis vectors.

Example 9.4 (Standard Basis for $\mathbb{R}^n$)**.**

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

9.4 Dimension

Definition 9.5 (Dimension)**.** The **dimension** of a space is the number of vectors in any basis for that space.

Theorem 9.6. *All bases for a given space have the same number of vectors.*

Key Dimensions:

- $\dim(C(A)) = r$ (rank of A)
- $\dim(N(A)) = n - r$
- Basis for $C(A)$: pivot columns of A
- Basis for $N(A)$: special solutions

10 Lecture 10: The Four Fundamental Subspaces

10.1 The Big Picture

For an $m \times n$ matrix A of rank r:

The Four Fundamental Subspaces

1. **Column Space** $C(A)$ in $\mathbb{R}^m$, dimension r
2. **Nullspace** $N(A)$ in $\mathbb{R}^n$, dimension $n - r$
3. **Row Space** $C(A^T)$ in $\mathbb{R}^n$, dimension r
4. **Left Nullspace** $N(A^T)$ in $\mathbb{R}^m$, dimension $m - r$

10.2 Bases for the Four Subspaces

Subspace	**In**	**Dim**	**Basis**
$C(A)$	$\mathbb{R}^m$	r	Pivot columns of A
$N(A)$	$\mathbb{R}^n$	$n-r$	Special solutions
$C(A^T)$	$\mathbb{R}^n$	r	First r rows of R
$N(A^T)$	$\mathbb{R}^m$	$m-r$	From $EA = R$

10.3 Orthogonality

Theorem 10.1 (Fundamental Theorem of Linear Algebra, Part 1)**.** *1. $C(A^T) \perp N(A)$ (orthogonal complements in $\mathbb{R}^n$)*

2. $C(A) \perp N(A^T)$ (orthogonal complements in $\mathbb{R}^m$)

10.4 Dimension Relations

$$\dim(C(A)) + \dim(N(A^T)) = m \tag{4}$$

$$\dim(C(A^T)) + \dim(N(A)) = n \tag{5}$$

Or equivalently: $r + (m - r) = m$ and $r + (n - r) = n$.

11 Lecture 11: Matrix Spaces and Rank 1 Matrices

11.1 New Vector Spaces

Example 11.1 (Space of 3 by 3 Matrices)**.** Let M = all 3×3 matrices. This is a vector space with dimension 9.

Example 11.2 (Symmetric Matrices)**.** Let S = all 3×3 symmetric matrices. This is a subspace of M with dimension 6.

11.2 Rank One Matrices

Definition 11.3 (Rank 1 Matrix)**.** Every rank 1 matrix can be written as: $A = \mathbf{u}\mathbf{v}^T =$ (column) $\times$ (row)

Theorem 11.4. *Every $m \times n$ matrix of rank r can be expressed as a sum of r rank-1 matrices.*

Part II
Unit II: Least Squares, Determinants and Eigenvalues

12 Lecture 12: Graphs and Networks

12.1 Incidence Matrix

Definition 12.1 (Incidence Matrix)**.** For a graph with n nodes and m edges, the incidence matrix A is $m \times n$.

12.2 Kirchhoff's Laws

Current Law (KCL): $A^T\mathbf{y} = \mathbf{0}$
Voltage Law (KVL): Potential differences around loops sum to zero.

13 Lecture 14: Orthogonal Vectors and Subspaces

13.1 Orthogonality of Vectors

Definition 13.1 (Orthogonal Vectors)**.** Vectors $\mathbf{x}$ and $\mathbf{y}$ are **orthogonal** if: $\mathbf{x}^T\mathbf{y} = 0$

Theorem 13.2 (Pythagorean Theorem)**.** *If* $\mathbf{x} \perp \mathbf{y}$*, then:* $\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 = \|\mathbf{x} + \mathbf{y}\|^2$

13.2 Orthogonal Subspaces

Definition 13.3 (Orthogonal Subspaces)**.** Subspaces S and T are orthogonal if every vector in S is orthogonal to every vector in T.

Theorem 13.4 (Fundamental Orthogonalities)**.**
1. *Row space perpendicular to nullspace (in* $\mathbb{R}^n$*)*
2. *Column space perpendicular to left nullspace (in* $\mathbb{R}^m$*)*

13.3 Orthogonal Complements

Definition 13.5 (Orthogonal Complement)**.** The orthogonal complement $S^\perp$ of subspace S contains all vectors perpendicular to S.

Theorem 13.6.

$$N(A) = C(A^T)^\perp \tag{6}$$
$$N(A^T) = C(A)^\perp \tag{7}$$

And: $\dim(S) + \dim(S^\perp) = n$ *(in* $\mathbb{R}^n$*).*

14 Lecture 15: Projections onto Subspaces

14.1 Projection onto a Line

Problem: Project vector $\mathbf{b}$ onto line through $\mathbf{a}$.
Key equation: Error is perpendicular to $\mathbf{a}$: $\mathbf{a}^T(\mathbf{b} - x\mathbf{a}) = 0$
Solution: $x = \frac{\mathbf{a}^T\mathbf{b}}{\mathbf{a}^T\mathbf{a}}$, $\quad \mathbf{p} = \mathbf{a}x = \mathbf{a}\frac{\mathbf{a}^T\mathbf{b}}{\mathbf{a}^T\mathbf{a}}$

14.2 Projection Matrix

Definition 14.1 (Projection Matrix onto Line)**.** $P = \frac{\mathbf{a}\mathbf{a}^T}{\mathbf{a}^T\mathbf{a}}$ Then $\mathbf{p} = P\mathbf{b}$ projects $\mathbf{b}$ onto $\mathbf{a}$.

Properties:

1. $P^T = P$ (symmetric)
2. $P^2 = P$ (projecting twice equals projecting once)
3. $\text{Rank}(P) = 1$

14.3 Projection onto Subspace

Problem: Project $\mathbf{b}$ onto column space of A.
Find $\mathbf{p} = A\hat{\mathbf{x}}$ closest to $\mathbf{b}$.
Key equation: Error $\mathbf{e} = \mathbf{b} - A\hat{\mathbf{x}}$ must be perpendicular to $C(A)$: $A^T(\mathbf{b} - A\hat{\mathbf{x}}) = \mathbf{0}$

Normal Equations

$A^T A\hat{\mathbf{x}} = A^T\mathbf{b}$

Solution: $\hat{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b}, \quad \mathbf{p} = A\hat{\mathbf{x}} = A(A^TA)^{-1}A^T\mathbf{b}$

Definition 14.2 (Projection Matrix onto $C(A)$)**.** $P = A(A^TA)^{-1}A^T$

15 Lecture 16: Least Squares

15.1 Least Squares Problem

Scenario: $A\mathbf{x} = \mathbf{b}$ has no solution (more equations than unknowns).
Goal: Find $\hat{\mathbf{x}}$ that minimizes $\|\mathbf{b} - A\mathbf{x}\|^2 = \mathbf{e}^T\mathbf{e}$.

Theorem 15.1 (Least Squares Solution)**.** $\hat{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b}$ *minimizes* $\|A\mathbf{x} - \mathbf{b}\|^2$.

15.2 Application: Fitting a Line

Data: Points $(t_1, b_1), (t_2, b_2), \ldots, (t_m, b_m)$
Model: $b = C + Dt$ (line)

System: $\begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$

Usually no exact solution! Use least squares.

16 Lecture 17: Orthogonal Matrices and Gram-Schmidt

16.1 Orthonormal Vectors

Definition 16.1 (Orthonormal Vectors)**.** Vectors $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n$ are **orthonormal** if: $\mathbf{q}_i^T\mathbf{q}_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$

16.2 Orthogonal Matrix

Definition 16.2 (Orthogonal Matrix)**.** Matrix Q with orthonormal columns satisfies: $Q^TQ = I$ If Q is square, then $Q^T = Q^{-1}$.

16.3 Gram-Schmidt Process

Input: Independent vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$
Output: Orthonormal vectors $\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3$

Gram-Schmidt Algorithm

Step 1: $\mathbf{A} = \mathbf{a}, \quad \mathbf{q}_1 = \frac{\mathbf{A}}{\|\mathbf{A}\|}$
Step 2: $\mathbf{B} = \mathbf{b} - (\mathbf{q}_1^T\mathbf{b})\mathbf{q}_1, \quad \mathbf{q}_2 = \frac{\mathbf{B}}{\|\mathbf{B}\|}$
Step 3: $\mathbf{C} = \mathbf{c} - (\mathbf{q}_1^T\mathbf{c})\mathbf{q}_1 - (\mathbf{q}_2^T\mathbf{c})\mathbf{q}_2, \quad \mathbf{q}_3 = \frac{\mathbf{C}}{\|\mathbf{C}\|}$

16.4 QR Decomposition

Theorem 16.3 (QR Factorization)**.** *Every matrix A with independent columns can be written as: $A = QR$ where Q has orthonormal columns and R is upper triangular with positive diagonal.*

17 Lecture 18: Properties of Determinants

17.1 Three Fundamental Properties

Defining Properties

Property 1: $\det(I) = 1$
Property 2: Exchanging two rows reverses the sign.
Property 3: The determinant is linear in each row separately.

17.2 Key Properties

1. Two equal rows implies $\det(A) = 0$
2. Row operations leave determinant unchanged (elimination)
3. Row of zeros implies $\det(A) = 0$
4. Upper triangular: $\det(U) = d_1 d_2 \cdots d_n$
5. $\det(A) = 0$ if and only if A is singular
6. $\det(AB) = \det(A) \cdot \det(B)$
7. $\det(A^T) = \det(A)$

18 Lecture 19: Determinant Formulas

18.1 Cofactor Expansion

Definition 18.1 (Cofactor)**.** The cofactor C_{ij} is: $C_{ij} = (-1)^{i+j}\det(M_{ij})$ where M_{ij} is obtained by deleting row i and column j.

Theorem 18.2 (Cofactor Expansion)**.** *Along row i:* $\det(A) = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}$

18.2 Cramer's Rule

Theorem 18.3 (Cramer's Rule)**.** *To solve* $A\mathbf{x} = \mathbf{b}$*:* $x_j = \frac{\det(B_j)}{\det(A)}$ *where* B_j *is* A *with column* j *replaced by* $\mathbf{b}$*.*

19 Lecture 20: Applications of Determinants

19.1 Area and Volume

Theorem 19.1.

- $|\det(A)|$ *equals area of parallelogram (in* $\mathbb{R}^2$*)*
- $|\det(A)|$ *equals volume of parallelepiped (in* $\mathbb{R}^3$*)*

20 Lecture 21: Eigenvalues and Eigenvectors

20.1 Definition

Definition 20.1 (Eigenvalue and Eigenvector)**.** λ is an eigenvalue of A with eigenvector $\mathbf{x} \neq \mathbf{0}$ if: $A\mathbf{x} = \lambda\mathbf{x}$

20.2 Finding Eigenvalues

$A\mathbf{x} = \lambda\mathbf{x} \implies (A - \lambda I)\mathbf{x} = \mathbf{0}$

For non-trivial solution:

Characteristic Equation

$\det(A - \lambda I) = 0$

20.3 Properties of Eigenvalues

Theorem 20.2.

1. *Sum of eigenvalues equals trace of* A
2. *Product of eigenvalues equals determinant of* A

21 Lecture 22: Diagonalization

21.1 Diagonalizing a Matrix

Theorem 21.1 (Diagonalization)**.** *If* A *has* n *linearly independent eigenvectors, then:* $A = S\Lambda S^{-1}$ *where* S *has eigenvectors as columns and* Λ *is diagonal with eigenvalues.*

21.2 Powers of A

Theorem 21.2. $A^k = S\Lambda^k S^{-1}$

Computing Λ^k is easy since it's diagonal!

21.3 Stability

Theorem 21.3 (Stability of difference equations)**.** *For* $\mathbf{u}_{k+1} = A\mathbf{u}_k$*:*

- $\mathbf{u}_k \to \mathbf{0}$ *if all* $|\lambda_i| < 1$
- $\mathbf{u}_k$ *explodes if any* $|\lambda_i| > 1$

22 Lecture 23: Differential Equations

22.1 System of ODEs

$$\frac{d\mathbf{u}}{dt} = A\mathbf{u}, \quad \mathbf{u}(0) = \mathbf{u}_0 \tag{8}$$

22.2 Solution via Eigenvalues

If $A\mathbf{x} = \lambda\mathbf{x}$, then $\mathbf{u}(t) = e^{\lambda t}\mathbf{x}$ solves the equation.
General solution: $\mathbf{u}(t) = c_1 e^{\lambda_1 t}\mathbf{x}_1 + c_2 e^{\lambda_2 t}\mathbf{x}_2 + \cdots + c_n e^{\lambda_n t}\mathbf{x}_n$

22.3 Matrix Exponential

Definition 22.1 (Matrix Exponential)**.** $e^{At} = I + At + \frac{(At)^2}{2!} + \frac{(At)^3}{3!} + \cdots$

Solution: $\mathbf{u}(t) = e^{At}\mathbf{u}_0$
If $A = S\Lambda S^{-1}$: $e^{At} = Se^{\Lambda t}S^{-1}$

23 Lecture 24: Markov Matrices

23.1 Markov Matrices

Definition 23.1 (Markov Matrix)**.** A matrix A is Markov if:

1. All entries $a_{ij} \geq 0$
2. Each column sums to 1

23.2 Properties

Theorem 23.2. *For a Markov matrix:*

1. $\lambda = 1$ *is always an eigenvalue*
2. *All other eigenvalues satisfy* $|\lambda_i| \leq 1$

The eigenvector for $\lambda = 1$ is the steady state.

Part III
Unit III: Positive Definite Matrices

24 Lecture 25: Symmetric Matrices

24.1 Spectral Theorem

Theorem 24.1 (Spectral Theorem). *If A is symmetric, then:*

1. *All eigenvalues are real*
2. *Eigenvectors can be chosen orthonormal*
3. *$A = Q\Lambda Q^T$ where $Q^T Q = I$*

24.2 Positive Definite Matrices

Definition 24.2 (Positive Definite). Symmetric matrix A is positive definite if: $\mathbf{x}^T A\mathbf{x} > 0 \quad \text{for all } \mathbf{x} \neq \mathbf{0}$

24.3 Tests for Positive Definiteness

Theorem 24.3 (Equivalent Conditions). *For symmetric A, these are equivalent:*

1. *Energy test: $\mathbf{x}^T A\mathbf{x} > 0$ for all $\mathbf{x} \neq \mathbf{0}$*
2. *Eigenvalue test: All $\lambda_i > 0$*
3. *Pivot test: All pivots greater than 0*
4. *Determinant test: All upper-left sub-determinants greater than 0*

25 Lecture 26: Complex Matrices and FFT

25.1 Complex Vectors

For complex vectors, the inner product is: $\mathbf{z}^H\mathbf{w} = \bar{z}_1 w_1 + \bar{z}_2 w_2 + \cdots + \bar{z}_n w_n$ where $\mathbf{z}^H = \bar{\mathbf{z}}^T$ (conjugate transpose).

25.2 Hermitian Matrices

Definition 25.1 (Hermitian Matrix). A is Hermitian if $A^H = A$.

This is the complex analog of symmetric matrices.

25.3 Fast Fourier Transform

The FFT computes the discrete Fourier transform in $O(n \log n)$ time instead of $O(n^2)$.

26 Lecture 29: Singular Value Decomposition

26.1 The SVD Theorem

Theorem 26.1 (SVD). *Every $m \times n$ matrix A can be factored as: $A = U\Sigma V^T$ where:*

- *U is $m \times m$ orthogonal*
- *Σ is $m \times n$ diagonal with singular values*
- *V is $n \times n$ orthogonal*

26.2 Applications of SVD

1. Image compression
2. Principal Component Analysis (PCA)
3. Pseudoinverse
4. Low-rank approximation
5. Least squares

27 Final Review

27.1 The Big Picture

Topic	**Key Ideas**
Linear equations	Elimination, LU
Vector spaces	Subspaces, basis, dimension
Four subspaces	Column, null, row, left null
Orthogonality	Projections, least squares
Determinants	Properties, cofactors
Eigenvalues	Diagonalization, powers
Positive definite	Tests, applications
SVD	Four subspaces, compression

27.2 Key Formulas

Matrix operations:

$$(AB)^T = B^T A^T \tag{9}$$
$$(AB)^{-1} = B^{-1} A^{-1} \tag{10}$$
$$\det(AB) = \det(A)\det(B) \tag{11}$$

Projections: $P = A(A^TA)^{-1}A^T$
Least squares: $\hat{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b}$
Eigenvalues: $\det(A - \lambda I) = 0$
Diagonalization: $A = S\Lambda S^{-1}$

Conclusion

This completes the comprehensive notes for MIT 18.06 Linear Algebra. The course emphasizes:

- Geometric intuition over formalism
- Understanding connections between concepts
- Applications alongside theory
- The beauty and power of linear algebra

"Linear algebra is the mathematics of the 21st century."
— Professor Gilbert Strang

Resources for Further Study:

- Textbook: Introduction to Linear Algebra by Gilbert Strang
- Video lectures: MIT OpenCourseWare 18.06
- Practice problems: MIT OCW problem sets