21-241 Matrices and Linear Transformations Notes

Compiled by Salman Hajizada

Exam 2 Edition

Chapter 1

Linear Equations in Linear Algebra

1.1 Systems of linear equations

Definition: A system of linear equations is a collection of one or more linear equations involving the same set of variables. A system has either:

- No solution
- Infinite solutions
- A unique solution

The info about a linear system can be recorded in a matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

The one with just the coefficients is called **coefficient matrix**
The one with the last col with the product is called the **augmented matrix**

There are 3 elementary row operations to carry out on the matrix:

- Replacement: by a sum of itself and a multiple of another row
- Interchange: swap rows
- Scaling: scale a row by a scalar

1.2 Row Reduction and Echelon Forms

Definition: A matrix is in **echelon form** if

- All nonzero rows are above any rows of zeros
- Each leading entry of a row is in a column ro the right of the leading entry of the row above it
- All entries in a column below a leading entry are zeros

A matrix is in **reduced echelon form** if

- It is in echelon form
- The leading entry in each nonzero row is 1
- Each leading 1 is the only nonzero entry in its column

Theorem 1 (Uniqueness of the Reduced Echelon Form)**.** *Each matrix is row equivalent to one and only one reduced echelon matrix*

Definition 1 (Pivot Positon)**.** *A **pivot position** in A is a location in A that corresponds to a leading 1 in the reduced echelon form of A*

Row Reduction Algorithm

1. Start with left non-zero column. This is a pivot column. Pivot position is at the top
2. Choose a nonzero entry in the pivot column as a pivot. Interchange to move it into pivot position
3. Use row replacement to make zeros below pivot position
4. Ignore the row with the pivot and the ones above. Apply steps 1-3 to the remaining matrix until no more rows to modify
5. Starting with rightmost pivot and working up and left, make zeros above each pivot. Make every pivot a 1 by scaling

Variables corresponding to pivot columns are **basic**, the other ones are **free**

Theorem 2 (Existence and Uniqueness theorem)**.**

A linear system is consistent iff the right column of the augmented matrix is NOT a pivot column i.e. iff echelon form has NO row of form:

$$\begin{bmatrix} 0 & \dots & 0 & b \end{bmatrix} \text{ with } b \text{ nonzero}$$

If a system is consistent then solution set contain either a unique solution if there are 0 free variables or infinite solutions if there is at least 1 free variable.

1.3 Vector Equations

Algebraic Properties of Vectors in $\mathbb{R}^n$

$\forall$ $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ and scalars c and d:

1. **Commutativity of addition:** $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
2. **Associativity of addition:** $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$
3. **Additive identity:** $\mathbf{u} + \mathbf{0} = \mathbf{u}$
4. **Additive inverse:** $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$
5. **Distributivity of scalar multiplication over addition (vectors):** $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$
6. **Distributivity of scalar multiplication over addition (scalars):** $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$
7. **Associativity of scalar multiplication:** $(cd)\mathbf{u} = c(d\mathbf{u})$
8. **Scalar identity:** $1\mathbf{u} = \mathbf{u}$

A linear combination is a vector $\mathbf{y} = c_1\mathbf{v}_1 + \ldots + c_p\mathbf{v}_p$

Definition 2. *Span*$\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ *is the set of all linear combinations of* $\mathbf{v_1}, \ldots, \mathbf{v_p}$
It is the set of all vectors that can be written in the form $c_1\mathbf{v}_1 + \ldots + c_p\mathbf{v}_p$

1.4 The Matrix Equation Ax = b

Definition 3.

If A is an $m \times n$ matrix, with columns $\mathbf{a_1}, \ldots, \mathbf{a_n}$ and if $\mathbf{x} \in \mathbb{R}^n$,

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

This is defined only if the number of columns of A equals number of entires in $\mathbf{x}$

Theorem 3.

If A is an $m \times n$ matrix and $\mathbf{b} \in \mathbb{R}^n$, then $A\mathbf{x} = \mathbf{b}$ has the same solution set as

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = b$$

which has the same solution set as the system of linear equations with matrix

$$\begin{bmatrix} \mathbf{a}_1 & \ldots & \mathbf{a}_n & \mathbf{b} \end{bmatrix}$$

Theorem 4.

Let A be a $m \times n$ matrix. The following statements are equivalent

a. $\forall\ \mathbf{b} \in \mathbb{R}^m, A\mathbf{x} = \mathbf{b}$ has a solution

b. $\forall\ \mathbf{b}$ is a linear combination of the columns of A

c. The columns of A span $\mathbb{R}^m$

d. A has a pivot in every row

Warning: A is a **coefficient matrix** in this theorem

Theorem 5.

If A is an $m \times n$ matrix and $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, and c is a scalar

a. $A\ (\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v}$;

b. $A\ (c\mathbf{u}) = c(A\mathbf{u})$;

1.5 Solution sets of Linear Systems

Homogenous Linear Systems: if it can be written in form $A\mathbf{x} = \mathbf{0}$
This system always has at least one solution ($\mathbf{x} = \mathbf{0}$): **trivial solution**
The equation has a **non-trivial** solution iff it has at least one free variable

Non-homogenous Linear Systems: if it can be written in form $A\mathbf{x} = \mathbf{b}$, $\mathbf{b}$ non-zero

Theorem 6.

Let $A\mathbf{x} = \mathbf{b}$ be consistent for some $\mathbf{b}$, let $\mathbf{p}$ be a solution. Then the solution set is the set of all vectors in form $\mathbf{w} = \mathbf{p} + \mathbf{v}_h$, where $\mathbf{v}_h$ is any solution of $A\mathbf{x} = \mathbf{0}$

Writing a solution set in parametric vector form

1. Row reduce to reduced echelon form
2. Express each basic variable in terms of any free variables
3. Write a typical solution as a vector whose entries depend on free variables
4. Decompose it into a linear combination of vectors (with numbers only inside) using free variables as parameters

1.7 Linear Independence

Definition 4.

$\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ is **linearly independent** if

$$x_1\mathbf{v}_1 + \ldots + x_p\mathbf{v}_p = \mathbf{0}$$

has only the trivial solution. It is said to be **linearly dependent** if $\exists c_1, \dots, c_p$, not all zero, such that

$$c_1\mathbf{v}_1 + \ldots + c_p\mathbf{v}_p = \mathbf{0}$$

A set of two vectors is linearly dependent if at least one is a multiple of the other

Theorem 7 (Characterization of Linearly independent Sets)**.**

A set of two or more vectors is linearly dependent iff at least one of them is a linear combination the others.

Theorem 8.

If a set contains more vectors than entries in each vector, then it is linearly dependant

Theorem 9.

If a set contains the zero vector, then it is linearly dependant

1.8 Introduction to Linear Transformations

A **transformation** T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a rule that assigns each vector $\mathbf{x}$ in $\mathbb{R}^n$ to a vector $T(\mathbf{x})$ in $\mathbb{R}^m$

- Domain: $\mathbb{R}^n$
- Co-domain: $\mathbb{R}^m$
- Range: set of all images $T(\mathbf{x})$

Definition 5. *T is linear iff*

- $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$
- $T(c\mathbf{u}) = cT(\mathbf{u})$

If T is linear then:

- $T(\mathbf{0}) = \mathbf{0}$
- $T(c\mathbf{u} + d\mathbf{v}) = cT(\mathbf{u}) + dT(\mathbf{v})$
- $T(c_1\mathbf{v}_1 + \ldots + c_p\mathbf{v}_p) = T(c_1\mathbf{v}_1) + \ldots + T(c_p\mathbf{v}_p)$

1.9 The Matrix of a Linear Transformation

Theorem 10. *For a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$, there exists a unique matrix A such that*

$$T(\mathbf{x}) = A\mathbf{x} \quad \forall\, \mathbf{x} \in \mathbb{R}^n.$$

This matrix A is given by

$$A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \dots & T(\mathbf{e}_n) \end{bmatrix},$$

where $\mathbf{e}_j$ denotes the j-th column of the identity matrix in $\mathbb{R}^n$.

Definition 6.

T is **onto** (surjective) if

- Range of T = Co-domain

- $\forall\ \mathbf{v} \in codomain(B), \exists\ \mathbf{u} \in domain(A)$ such that $T(\mathbf{u}) = \mathbf{v}$
- Find standard matrix
$$A = \begin{bmatrix} T(c_1) & \dots & T(c_n) \end{bmatrix}$$
Onto if $A\mathbf{u} = \mathbf{v}$ has a solution for each $\mathbf{v}$ in range

Definition 7.

T is **one-to-one** (injective)if

- $T(\mathbf{u}_1) = T(\mathbf{u}_2) \implies \mathbf{u}_1 = \mathbf{u}_2$
- Find standard matrix
$$A = \begin{bmatrix} T(c_1) & \dots & T(c_n) \end{bmatrix}$$
$A\mathbf{u} = \mathbf{b}$ has a unique solution for each $\mathbf{b} \in Range(T)$

Theorem 11. *T is one-to-one iff $T(\mathbf{x}) = \mathbf{0}$ has a ONLY the trivial solution*

Theorem 12. *Let $T : \mathbb{R}^n \to \mathbb{R}^m$, let A be standard matrix for T. Then*

- *T maps $\mathbb{R}^n$ onto $\mathbb{R}^m$ iff the columns of A span $\mathbb{R}^m$*
- *T is one-to-one iff the columns of A are linearly independent*

Chapter 2

Matrix Algebra

2.1 Matrix Operations

- **Square Matrix**: row and columns are the same
- **Diagonal entries**: $a_{11}, a_{22}, \ldots$
- **Diagonal matrix**: square $n \times n$ matrix who's non-diagonal entires are zero
- **Zero matrix**: $m \times n$ matrix where all entries are 0

$[a_{ij}]_{m\times n} \pm [b_{ij}]_{m\times n} = [a_{ij} \pm b_{ij}]_{m\times n}$

Two matrices are **equal** if they have the same size and their entries are equal (duh)

Theorem 1.

Properties of Matrices

Let A, B and C be matrices of the same size, and let r and s be scalars

- $A + B = B + A$
- $(A + B) + C = A + (B + C)$
- $A + 0 = A$
- $r(A + B) = rA + rB$
- $(r + s)A = rA + sA$
- $r(sA) = (rs)A$

Definition 1. *If A is a $m \times n$ matrix and B is an $n \times p$ matrix with columns $\mathbf{b}_1, \ldots, \mathbf{b}_p$, then*

$$AB = A\begin{bmatrix}\mathbf{b}_1 \ldots \mathbf{b}_p\end{bmatrix} = \begin{bmatrix}A\mathbf{b}_1 \ldots A\mathbf{b}_p\end{bmatrix}$$

- Each column of AB is a linear combination of the columns of A using weights from columns of B
- **The number of columns of A must match the number of rows in B**
- Row column rule: $(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}$
- $row_i(AB) = row_i(A) \cdot B$

Theorem 2.

Properties of Matrices

Let A be an $m \times n$ matrix, and let B and C have sizes for which the indicated sums and products are defined.

$A(BC) = (AB)C$ (associative law of multiplication)

$A(B + C) = AB + AC$ (left distributive law)

$(B + C)A = BA + CA$ (right distributive law)

$r(AB) = (rA)B = A(rB)$ for any scalar r

$I_m A = A = AI_n$ (identity for matrix multiplication)

WARNING

Let A be an $m \times n$ matrix, and let B and C have sizes for which the indicated sums and products are defined.

- In general, $AB \neq BA$
- In general, $AB = AC$ **does not imply** $B = C$
- In general, $AB = 0$ **does not imply** $A = 0 \vee B = 0$

Power of a matrix: $A^k = \underbrace{AA \ldots A}_{k}$

The transpose of a matrix: A^T is obtained by swapping its rows and columns of A.

Theorem 3.

- $(A^T)^T = A$
- $(A + B)^T = A^T + B^T$
- $(rA)^T = rA^T$
- $(AB)^T = B^T A^T$

2.2 The Inverse of a Matrix

An $n \times n$ matrix A is said to be **invertible** if there is an $n \times n$ matrix C such that

$$CA = I_n \quad \text{and} \quad AC = I_n,$$

In this case, C is an **inverse** of A

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

Theorem 4. *Let* $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ *If* $ad - bc \neq 0$*, then* A *is invertEBIL and*

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

$detA = ad - bc$

Theorem 5. *If* A *is invertible, then* $\forall$ $\mathbf{b} \in \mathbb{R}^n$, $A\mathbf{x} = \mathbf{b}$ *has the unique solution* $\mathbf{x} = A^{-1}\mathbf{b}$

Theorem 6. *1. If* A *invertible, then* A^{-1} *is invertEbil and* $(A^{-1})^{-1} = A$

2. *If A and B invertible, $(AB)^{-1} = B^{-1}A^{-1}$*
3. *If A invertible, $(A^T)^{-1} = (A^{-1})^T$*

- An **elementary matrix** is a matrix obtained by performing a single elementary row operation on I
- If a row op is performed on A, the result can be written as EA, where E is made by performing the same row op on I
- Each E is an invertible and its inverse is the elementary matrix that transforms E back into I.

Theorem 7. *A is invertible iff A is row equivalent to I, and any sequence of row ops that reduces A to I also transforms I to A^{-1}*

Finding A^{-1}: Row reduce $\begin{bmatrix} A & I \end{bmatrix}$. If A row-eq to I then $\begin{bmatrix} A & I \end{bmatrix}$is row-eq to $\begin{bmatrix} I & A^{-1} \end{bmatrix}$

2.3 Characterizations of Matrices

Theorem 8.

The following statements about a square n by n matrix A are equivalent

- A is an invertible matrix.
- A is row equivalent to the $n \times n$ identity matrix.
- A has n pivot positions.
- The equation $Ax = 0$ has only the trivial solution.
- The columns of A form a linearly independent set.
- The linear transformation $x \mapsto Ax$ is one-to-one.
- The equation $Ax = b$ has at least one solution for each $b \in \mathbb{R}^n$.
- The columns of A span $\mathbb{R}^n$.
- The linear transformation $x \mapsto Ax$ maps $\mathbb{R}^n$ onto $\mathbb{R}^n$.
- There is an $n \times n$ matrix C such that $CA = I$.
- There is an $n \times n$ matrix D such that $AD = I$.
- A^T is an invertible matrix.

If $AB = I$, then A and B invertible with $B = A^{-1}$ and $A = B^{-1}$
A linear transformation is invertible if $\exists S : \mathbb{R}^n \to \mathbb{R}^n$ such that:

$$\begin{aligned} S(T(\mathbf{x})) = x \quad \text{for all } \mathbf{x} \in \mathbb{R}^n \\ T(S(\mathbf{x})) = x \quad \text{for all } \mathbf{x} \in \mathbb{R}^n \end{aligned}$$

Theorem 9. *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ and A be the matrix. T is invertible iff A is invertible. Then $S(\mathbf{x}) = A^{-1}\mathbf{x}$*

2.4 Partitioned Matrices

Trivial stuff honestly just partition and multiply.
Ok actually there is one good theorem here

Theorem 10 (Column-Row Expansion of AB)**.**

If A is $m \times n$ and B is $n \times p$:

$$AB = \begin{bmatrix} col_1(A) & \dots & col_n(A) \end{bmatrix} \begin{bmatrix} row_1(B) \\ \vdots \\ row_n(B) \end{bmatrix} = col_1(A)row_1(B) + \dots + col_n(A)row_n(B)$$

The inverse of a block upper triangular matrix:

$$A^{-1} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ 0 & A_{22}^{-1} \end{bmatrix}$$

2.5 Matrix Factorizations

Assume A is $m \times n$ and can be reduced to Echelon Form without row interchanges.
Then $A = LU$,

- $L : m \times m$ lower triangular with 1's in diagonal
- $U : m \times n$ echelon form of A

$A\mathbf{x} = \mathbf{b}$ can be written as $L(U\mathbf{x}) = \mathbf{b}$. Writing $\mathbf{y} = U\mathbf{x}$, we get:

$$L(U\mathbf{x}) = \mathbf{b}$$

$$L\mathbf{y} = \mathbf{b}$$
$$U\mathbf{x} = \mathbf{y}$$

We can get $\mathbf{x}$ by solving the easy pair of equations:

1. Solve $L\mathbf{y} = \mathbf{b}$ for $\mathbf{y}$
2. Solve $U\mathbf{x} = \mathbf{y}$ for $\mathbf{x}$

LU Factorization Algorithm (boring)

1. Reduce A to echelon form U using a sequence of row replacement operations
2. Place entries in L such that the same sequence of row operations reduces L to I

LU Factorization Algorithm (fast)

For a matrix $A : m \times n$

1. Initialize L as the identity matrix $m \times m$ and U as A.
2. For each row $i = 1, 2, \dots, m$:
   - Use row operations to create zeros below the diagonal in U by subtracting multiples of the current row from the rows below.
   - For each row operation, divide the corresponding entries below the pivot in column i by the pivot, and store these multipliers in the corresponding positions of L.
3. Continue this process until U is in echelon form and L has recorded the multipliers such that $A = LU$.

This description is very bad, please practice this on a non-square matrix

Chapter 3

Determinants

3.1 Intro to Determinants

First, we denote A_{ij} to be the sub-matrix of A formed by deleting the ith row and jth column of A

Definition 1. *For $n \geq 2$, the **determinant** of a $n \times n$:*

$$\begin{aligned} \det A \quad &= a_{11}\det(A_{11}) - a_{12}\det(A_{12}) + \ldots + (-1)^{1+n}a_{1n}\det(A_{1n}) = \\ &= \textstyle\sum_{j=1}^{n}(-1)^{1+j}a_{1j}\det(A_{1j}) \end{aligned}$$

The (i, j)-cofactor of A: $C_{ij} = (-1)^{i+j}\det(A_{ij})$

Theorem 1. *The determinant of A can be computed by cofactor expansion across any row or column. Across row i:*

$$\det(A) = a_{i1}C_{i1} + \ldots + a_{in}C_{in}$$

Across column j:

$$\det(A) = a_{1j}C_{1j} + \ldots + a_{nj}C_{nj}$$

Signs of the cofactor:

$$\begin{bmatrix} + & - & + & \cdots \\ - & + & - & \\ + & - & + & \\ \vdots & & & \ddots \end{bmatrix}$$

Theorem 2. *If A triangular, then $\det(A)$ is the product of the entires on the main diagonal of A*

3.2 Properties of Determinants

Theorem 3.

Row Operations

For a matrix $A : m \times n$

a. If a multiple of a row in A is added to another to produce B: $\det(B) = \det(A)$

b. If two rows of A are swapped to produce B: $\det(B) = -\det(A)$

c. If a row of A is multiplied by k to produce B: $\det(B) = k \cdot \det(A)$

Two main strats where this is useful:

- Row reduce a matrix to echelon form to simplify calculation (without row interchange!)
- Factor out a common factor from a row to simplify calculation

Theorem 4. *A is invertible iff $\det(A) \neq 0$ (trivial)*

Theorem 5. $\det(A) = \det(A^T)$

Theorem 6. *If A and B $n \times n$, then* $\det(AB) = \det(A) \cdot \det(B)$

Lemma 1. $\det(A^{-1}) = \frac{1}{\det(A)}$

3.3 Crammer's Rule

Let $A_i(\mathbf{b})$ be the matrix obtained from A by replacing column i by the vector $\mathbf{b}$

Theorem 7 (Cramer's Rule)**.** *Let $A : n \times n$, invertible, the solution $\mathbf{x}$ of $A\mathbf{x} = \mathbf{b}$ is given by:*

$$x_i = \frac{\det(A_i(\mathbf{b}))}{\det(A)}, \text{ for } i = 1, 2, \ldots, n$$

The matrix of cofactors is called the **adjugate or classical adjoint** of A, denoted by adj A

Theorem 8 (an Inverse Forumla)**.**

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} C_{11} & C_{21} & \ldots & C_{n1} \\ C_{12} & C_{22} & \ldots & C_{n2} \\ \vdots & \vdots & & \\ C_{1n} & C_{2n} & \ldots & C_{nn} \end{bmatrix} = \frac{adj(A)}{\det(A)}$$

Theorem 9. *If $A : 2 \times 2$, Area of parallelogram determined by columns of A is* $\det(A)$
If $A : 3 \times 3$, Volume of parallelepiped determined by columns of A is $\det(A)$

Chapter 4

Vector Spaces and Subspaces

4.1 Vector Spaces and Subspaces

Definition 2. *A **vector space** is a nonempty set V of objects called vectors, on which there are defined two operations: addition and scalar multiplication. The axioms below must hold for all* $\mathbf{u}, \mathbf{w}, \mathbf{v} \in V$

1. $\mathbf{u} + \mathbf{v} \in V$ (Closure under addition)
2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
3. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
4. $\mathbf{0} \in V$, such that $\mathbf{u} + \mathbf{0} = \mathbf{u}$
5. $\forall\ \mathbf{u} \in V$, $\exists\ -\mathbf{u} \in V$ such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$
6. $c\mathbf{u} \in V$
7. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$
8. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$
9. $c(d\mathbf{u}) = (cd)\mathbf{u}$
10. $1\mathbf{u} = \mathbf{u}$

Definition 3. *A **subspace** of a space V is a subset H of V that:*

- *The zero vector of V is in H*
- *H is closed under vector addition:* $\mathbf{u} + \mathbf{v} \in H$ *for all* $\mathbf{u}, \mathbf{v} \in H$
- *H is closed under scalar multiplication:* $c\mathbf{u} \in H$ *for all scalars c and vectors* $\mathbf{u} \in H$

Theorem 10. *If* $\mathbf{v}_1, \ldots, \mathbf{v}_p$ *are in a vector space V, then Span*$\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ *is a subspace of V*