Linear Algebra Study Guide

Thomas Cohn

Class Taught by Dr. David Fernández-Bretón

January 26, 2017

Part I

Start of Class to the First Exam

Chapter 1

Linear Equations

1.1 Introduction to Linear Systems

Nothing important to study or review.

1.2 Matrices, Vectors, and Gauss-Jordan Elimination

Definition: Vectors and vector spaces
A matrix with only one column is a column vector. The set of all column vectors with n components is the *vector space* $\mathbb{R}^n$

Definition: Standard representation of vectors
The standard representation of $\vec{v} = \begin{bmatrix} x \\ y \end{bmatrix}$ in the Cartesian coordinate plane is an arrow from the origin to the point.

Reduced Row-Echelon Form (rref)

(a) If a row has nonzero entries, then the first nonzero entry is a 1.

(b) If a column contains a leading 1, then all other entries in that column are 0.

(c) If a row contains a leading 1, then each row above it contains a leading 1 further to the left.

Condition **(c)** implies that rows of 0's, if any, are at the bottom of the matrix.

Types of elementary row operations

- Divide a row by a nonzero scalar.
- Subtract a multiple of one row from another row.
- Swap two rows.

1.3 On the Solutions of Linear Systems; Matrix Algebra

The Number of Solutions of a Linear System

Theorem 1.3.1

Number of solutions of a linear system
A system of equations is *consistent* if there is at least one solution; it is *inconsistent* if there are no solutions. A linear system is inconsistent if in rref, there's the row $\left[\begin{array}{cccc|c} 0 & 0 & \cdots & 0 & 0 \end{array}\right]$, representing the equation $0 = 1$.
If a linear system is consistent, then it has either *infinitely many solutions* if there is at least one free variable, or *exactly one solution* if there are no free variables.

Definition 1.3.2

Definition: Rank of a Matrix
The rank of a matrix A is the number of leading 1's in $\text{rref}(A)$, denoted $\text{rank}(A)$.

Theorem 1.3.3

Number of equations vs. number of unknowns

(a) If a linear system has exactly one solution, then the number of equations must be greater than or equal to the number of variables.

Matrix Algebra

Definition 1.3.5

Sums of matrices
The sum of two matrices of the same size is defined entry by entry:

$$\begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} + \begin{bmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nm} \end{bmatrix} = \begin{bmatrix} a_{11}+b_{11} & \cdots & a_{1m}+b_{1m} \\ \vdots & & \vdots \\ a_{n1}+b_{n1} & \cdots & a_{nm}+b_{nm} \end{bmatrix}$$

Scalar multiples of matrices
The product of a scalar with a matrix is defined entry by entry:

$$k \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} = \begin{bmatrix} ka_{11} & \cdots & ka_{1m} \\ \vdots & & \vdots \\ ka_{n1} & \cdots & ka_{nm} \end{bmatrix}$$

Definition 1.3.6

Dot product of vectors
Consider two vectors $\vec{v}$ and $\vec{w}$ with components $v_1, \ldots, v_n$ and $w_1, \ldots, w_n$. Then the dot product of $\vec{v}$ and $\vec{w}$ is $\vec{v} \cdot \vec{w} = v_1 w_1 + \cdots + v_n w_n$.

Definition 1.3.7

The product $A\vec{x}$
If A is an $n \times m$ matrix with row vectors $\vec{w_1} \ldots \vec{w_n}$, and $\vec{x}$ is a vector in $\mathbb{R}^m$, then

$$A\vec{x} = \begin{bmatrix} - & \vec{w_1} & - \\ & \vdots & \\ - & \vec{w_n} & - \end{bmatrix} \vec{x} = \begin{bmatrix} \vec{w_1} \cdot \vec{x} \\ \vdots \\ \vec{w_n} \cdot \vec{x} \end{bmatrix}$$

Theorem 1.3.8

The product $A\vec{x}$ in terms of the columns of A
If the column vectors of an $n \times m$ matrix A are $\vec{v_1}, \ldots, \vec{v_m}$ and $\vec{x}$ is a vector in $\mathbb{R}^m$ with components $x_1, \ldots, x_m$, then

$$A\vec{x} = \begin{bmatrix} | & & | \\ \vec{v_1} & \cdots & \vec{v_m} \\ | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} = x_1\vec{v_1} + \cdots + x_m\vec{v_m}$$

Definition 1.3.9

Linear combinations
A vector $\vec{b}$ in $\mathbb{R}^n$ is called a linear combination of the vectors $\vec{v_1}, \ldots, \vec{v_m}$ in $\mathbb{R}^n$ if there exist scalars $x_1, \ldots, x_m$ such that

$$\vec{b} = x_1\vec{v_1} + \cdots + x_m\vec{v_m}$$

Theorem 1.3.10

Algebraic rules for $A\vec{x}$
If A is an $n \times m$ matrix, $\vec{x}$ and $\vec{u}$ are vectors in $\mathbb{R}^m$, and k is a scalar, then

a. $A(\vec{x} + \vec{y}) = A\vec{x} + A\vec{y}$, and

b. $A(k\vec{x}) = k(A\vec{x})$.

Theorem 1.3.11

Matrix form of a linear system
We can write the linear system with augmented matrix $\left[\begin{array}{c|c} A & \vec{b} \end{array}\right]$ in matrix form as

$$A\vec{x} = \vec{b}$$

The Joy of Sets

Set relations: Equality

Definition 1.

Two sets are defined to be equal when they have precisely the same elements. When the sets A and B are equal, we write $A = B$.

Set relations: Subset

Definition 2.

If A and B are sets, then we say that A is a subset of B (or A is contained in B, or B contains A, or A is included in B, or B includes A), and write $A \subset B$ or $A \subseteq B$, provided that every element of A is an element of B.

Set operations: Complement, union, and intersection

Definition 3.

The union of sets A and B, written $A \cup B$, is the set

$$\{x|(x \in A)\text{or}(x \in B)\}$$

Definition 4.

The intersection of sets A and B, written $A \cap B$, is the set

Remark 5.

For S and T sets, $S \cap T \subset S \subset S \cup T$ and $S \cap T \subset T \subset S \cup T$

$$\{x|(x \in A)\text{and}(x \in B)\}$$

Definition 6.

Suppose A and B are sets. The difference of B and A, denoted $B \setminus A$ or $B - A$ is the set

$$\{b \in B|b \notin A\}$$

Definition 7.

Let U denote a set that contains a subset A. The complement of A (with respect to U), often written A^c, $A\complement$, $\overline{A}$, or A', is the set $U \setminus A$.

Random Notes

$\mathbb{N}$ is the set of natural numbers, $1, 2, 3, 4, \ldots$.
$\mathbb{Z}$ is the set of integers.
$\mathbb{Q}$ is the set of rational numbers.
$\mathbb{R}$ is the set of real numbers.
The cardinality of set A is $|A|$, and is defined as the number of elements the set contains.
$\emptyset$ has cardinality zero, but $\{\emptyset\}$ has cardinality one.
For any set A, $\emptyset \subset A \subset A$.
For any sets A, B, $A = B$ if and only if $A \subset B$ and $B \subset A$.
For any set S, $S \cup \emptyset = S$.
For any set S, $S \cap \emptyset = \emptyset$.
For any sets A, U such that $A \subset U$, $A\complement \cup A = U$ and $A\complement \cap A = \emptyset$.
Two sets with empty intersection are said to be disjoint.
$(A \cup B)\complement = A\complement \cap B\complement$ and $(A \cap B)\complement = A\complement \cup B\complement$

Mathematical Hygiene

Statements

Definition 1.

A statement, also called a proposition, is a sentence that is either true or false, but not both.

Negation and truth tables

The truth table for negation:

p	$\neg$
T	F
F	T

Equivalent Statements

The truth table for double negation:

p	$\neg p$	$\neg(\neg p)$
T	F	T
F	T	F

Compound statements: Conjunctions and Disjunctions

The truth table for conjunction (and):

p	q	$p \wedge q$
T	T	T
T	F	F
F	T	F
F	F	F

The truth table for disjunction (or):

p	q	$p \vee q$
T	T	T
T	F	T
F	T	T
F	F	F

Conditional Statements

The truth table for the conditional statement p implies q or if p then q:

p	q	$p \Rightarrow q$
T	T	T
T	F	F
F	T	T
F	F	T

The truth table for the negation of $p \Rightarrow q$:

p	q	$p \wedge \neg q$	$\neg(p \Rightarrow q)$
T	T	F	F
T	F	T	T
F	T	F	F
F	F	F	F

Predicates

A predicate $p(x)$ is a statement that may either be true or false, such as "$y > 4$".

Quantifiers

$\forall$ means "for all".
$\exists$ means "there exists".

Chapter 2

Linear Transformations

2.1 Introduction to Linear Transformations and Their Inverses

Definition 2.1.1

Linear Transformations
A function T from $\mathbb{R}^m$ to $\mathbb{R}^n$ is called a *linear transformation* if there exists an $n \times m$ matrix A such that $T(\vec{x} = A\vec{x}$ for all $\vec{x}$ in the vector space $\mathbb{R}^m$.

Theorem 2.1.2

The columns of the matrix of a linear transformation
Consider a linear transformation T from $\mathbb{R}^m$ to $\mathbb{R}^n$. Then the matrix of T is

$$A = \begin{bmatrix} | & | & & | \\ T(\vec{e_1}) & T(\vec{e_2}) & \cdots & T(\vec{e_m}) \\ | & | & & | \end{bmatrix}, \text{ where } \vec{e_1} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \begin{array}{l} \\ \\ \\ \leftarrow i\text{th} \\ \\ \\ \end{array}$$

Theorem 2.1.3

Linear transformations
A transformation T from $\mathbb{R}^m$ to $\mathbb{R}^n$ is linear if (and only if)

a. $T(\vec{v} + \vec{w}) = T(\vec{v}) + T(\vec{w})$, for all vectors $\vec{v}$ and $\vec{w}$ in $\mathbb{R}^m$, and

b. $T(k\vec{v}) = kT(\vec{v})$, for all vectors $\vec{v}$ in $\mathbb{R}^m$ and all scalars k.

2.2 Linear Transformations in Geometry

Definition 2.2.1

Orthogonal Projections
Consider a line L in the coordinate plane running through the origin. Any vector $\vec{x}$ in $\mathbb{R}^2$ can be written uniquely as

$$\vec{x} = \vec{x}^{\parallel} + \vec{x}^{\perp}$$

where $\vec{x}^{\parallel}$ is parallel to line L and $\vec{x}^{\perp}$ is perpendicular to L.
The transformation $T(\vec{x}) = \vec{x}^{\parallel}$ from $\mathbb{R}^2$ to $\mathbb{R}^2$ is called the *orthogonal projection of $\vec{x}$ onto L*, often denoted by $\text{proj}_L(\vec{x})$. If $\vec{w}$ is a nonzero vector parallel to L, then

$$\text{proj}_L(\vec{x}) = \left(\frac{\vec{x} \cdot \vec{w}}{\vec{w} \cdot \vec{w}}\right) \vec{w}.$$

In particular, if $\vec{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ is a *unit* vector parallel to L, then

$$\text{proj}_L(\vec{x}) = (\vec{x} \cdot \vec{u})\vec{u}.$$

The transformation $T(\vec{x}) = \text{proj}_L(\vec{x})$ is linear, with matrix

$$P = \frac{1}{w_1^2 + w_2^2} \begin{bmatrix} w_1^2 & w_1 w_2 \\ w_1 w_2 & w_2^2 \end{bmatrix} = \begin{bmatrix} u_1^2 & u_1 u_2 \\ u_1 u_2 & u_2^2 \end{bmatrix}.$$

Definition 2.2.2

Reflections
Consider a line L in the coordinate plane, running through the origin, and let $\vec{x} = \vec{x}^{\parallel} + \vec{x}^{\perp}$ be a vector in $\mathbb{R}^2$. The linear transformation $T(\vec{x}) = \vec{x}^{\parallel} - \vec{x}^{\perp}$ is called the *reflection of $\vec{x}$ about L*, often denoted by $\text{ref}_L(\vec{x})$:

$$\text{ref}_L(\vec{x}) = \vec{x}^{\parallel} - \vec{x}^{\perp}.$$

We have a formula relating $\text{ref}_L(\vec{x})$ to $\text{proj}_L(\vec{x})$:

$$\text{ref}_L(\vec{x}) = 2\text{proj}_L(\vec{x}) - \vec{x} = 2(\vec{x} \cdot \vec{u})\vec{u} - \vec{x}.$$

The matrix of T is of the form $\begin{bmatrix} a & b \\ b & -a \end{bmatrix}$, where $a^2 + b^2 = 1$. Conversely, any matrix of this form represents reflection about a line.

Theorem 2.2.3

Rotations
The matrix of a counterclockwise rotation in $\mathbb{R}^2$ through an angle θ is

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

Note that this matrix is of the from $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, where $a^2 + b^2 = 1$. Conversely, any matrix of this form represents a rotation.

Theorem 3.2.4

Rotations combined with a scaling
A matrix of the form $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ represents a rotation combined with a scaling.

More precisely, if r and θ are the polar coordinates of the vector $\begin{bmatrix} a \\ b \end{bmatrix}$, then $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ represents a rotation through θ combined with a scaling by r.

Theorem 2.2.5

Horizontal and vertical shears

The matrix of a *horizontal Shear* is of the from $\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$, and the matrix of a *vertical shear* is of the form $\begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$, where k is an arbitrary constant.

Summary

Transformation	**Matrix**
Scaling by k	$kI_2 = \begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix}$
Orthogonal **projection** onto line L	$\begin{bmatrix} u_1^2 & u_1u_2 \\ u_1u_2 & u_2^2 \end{bmatrix}$, where $\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ is a unit vector parallel t
Reflection about a line	$\begin{bmatrix} a & b \\ b & -a \end{bmatrix}$, where $a^2 + b^2 = 1$.
Rotation through angle θ	$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ or $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, where $a^2 + b^2 = 1$.
Rotation through angle θ combined with **scaling** by r	$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} = r \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$
Horizontal shear	$\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$
Vertical shear	$\begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$

2.3 Matrix Products

Definition 2.3.1

Matrix Multiplication

- **a.** Let B be an $n \times p$ matrix and A a $q \times m$ matrix. The product BA is defined if (and only if) $p = q$.
- **b.** If B is an $n \times q$ matrix and A a $p \times m$ matrix, then the product BA is defined as the matrix of the linear transformation $T(\vec{x}) = B(A\vec{x})$. This means that $T(\vec{x}) = B(A\vec{x}) = (BA)\vec{x}$, for all $\vec{x}$ in the vector space $\mathbb{R}^m$. The product BA is an $n \times m$ matrix.

Theorem 2.3.2

The columns of the matrix product
Let B be an $n \times p$ matrix and A a $p \times m$ matrix with columns $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_m}$. Then, the product BA is

$$BA = B \begin{bmatrix} | & | & & | \\ \vec{v_1} & \vec{v_2} & \cdots & \vec{v_m} \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ B\vec{v_1} & B\vec{v_2} & \cdots & B\vec{v_m} \\ | & | & & | \end{bmatrix}.$$

To find BA, we can multiply B by the columns of A and combine the resulting vectors.

Theorem 2.3.3

Matrix multiplication is noncommutative
$AB \neq BA$, in general. However, at times it does happen that $AB = BA$; then we say that the matrices A and B *commute*.

Theorem 2.3.4

The entries of the matrix product
Let B be an $n \times p$ matrix and A a $p \times m$ matrix. The ijth entry of BA is the dot products of the ith row of B with the jth column of A.

$$BA = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & & \vdots \\ b_{i1} & b_{i2} & \cdots & b_{ip} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pj} & \cdots & a_{pm} \end{bmatrix}$$

is the $n \times m$ matrix whose ijth entry is

$$b_{i1}a_{1j} + b_{i2}a_{2j} + \cdots + b_{ip}a_{pj} = \sum_{k=1}^{p} b_{ik}a_{kj}.$$

Theorem 2.3.5

Multiplying with the identity matrix
For an $n \times m$ matrix A, $AI_m = I_nA = A$.

Theorem 2.3.6

Matrix multiplication is associative

$$(AB)C = (A(BC)$$

We can simply write ABC for the product $(AB)C = A(BC)$.

Theorem 2.3.7

Distributive property for matrices
If A and B are $n \times p$ matrices, and C and D are $p \times m$ matrices, then

$$A(C + D) = AC + AD, \text{ and}$$

$$(A + B)C = AC + BC.$$

Theorem 2.3.8

If A is an $n \times p$ matrix, B is a $p \times m$ matrix, and k is a scalar, then

$$(kA)B = A(kB) = k(AB)$$

2.4 The inverse of a Linear Transformation

Definition 2.4.1

Invertible Functions
A function T from X to Y is called invertible if the equation $T(x) = y$ as a unique solution x in X for each y in Y.
In this case, the inverse T^{-1} from Y to X is defined by

$$T^{-1}(y) = (\text{the unique } x \text{ in } X \text{ such that } T(x) = y)\,.$$

To put it differently, the equation

$$x = T^{-1}(y) \quad \text{means that} \quad y = T(x)$$

parNote that

$$T^{-1}(T(x)) = x \quad \text{and} \quad T(T^{-1}(y)) = y$$

for all x in X and for all y in Y.
Conversely, if L is a function from Y to X such that

$$L(T(x)) = x \quad \text{and} \quad T(L(y)) = y$$

for all x in X and for all y in Y, then T is invertible and $T^{-1} = L$.
If a function is invertible, then so is T^{-1} and $(T^{-1})^{-1} = T$.

Definition 2.4.2

Invertible matrices
A square matrix A is said to be *invertible* if the linear transformation $\vec{y} = T(\vec{x}) = A\vec{x}$ is invertible. In this case, the matrix of T^{-1} is denoted by A^{-1}. If the linear transformation $\vec{y} = T(\vec{x}) = A\vec{x}$ is invertible, then its inverse is $\vec{x} = T^{-1}(\vec{y}) = A^{-1}\vec{y}$.

Theorem 2.4.3

Invertibility
An $n \times n$ matrix A is invertible if (and only if) $\text{rref}(A) = I_n$ or, equivalently, if $\text{rank}(A) = n$.

Theorem 2.4.4

Invertibility and linear systems
Let A be an $n \times n$ matrix.

a. Consider a vector $\vec{b}$ in $\mathbb{R}^n$. If A is invertible, then the system $A\vec{x} = \vec{b}$ has the unique solution $\vec{x} = A^{-1}\vec{b}$. If A is noninvertible, then the system $A\vec{x} = \vec{b}$ has infinitely many solutions or none.

b. Consider the special case when $\vec{b} = \vec{0}$. The system $A\vec{x} = \vec{0}$ has $\vec{x} = \vec{0}$ as a solution. If A is invertible, then this is the only solution. If A is noninvertible, then the system $A\vec{x} = \vec{0}$ has infinitely many solutions.

Theorem 2.4.5

Finding the inverse of a matrix
To find the *inverse* of an $n \times n$ matrix A, form the $n \times (2n)$ matrix $[A|I_n]$ and compute $\text{rref}\,[A|I_n]$

- If $\text{rref}\,[A|I_n]$ is of the form $[I_n|B]$, then A is invertible, and $A^{-1} = B$.
- If $\text{rref}\,[A|I_n]$ is of another form (i.e., its left half fails to be I_n), then A is not invertible. Note that the left half of $\text{rref}\,[A|I_n]$ is $\text{rref}(A)$.

Theorem 2.4.6

Multiplying with the inverse
For an invertible $n \times n$ matrix A, $A^{-1}A = I_n$ and $AA^{-1} = I_n$.

Theorem 2.4.7

The inverse of a product of matrices
If A and B are invertible $n \times n$ matrices, then BA is invertible as well, and $(BA)^{-1} = A^{-1}B^{-1}$.

Theorem 2.4.8

A criterion for invertibility
Let A and B be two $n \times n$ matrices such that $BA = I_n$. Then

a. A and B are both invertible,

b. $A^{-1} = B$ and $B^{-1} = A$, and

c. $AB = I_n$.

Theorem 2.4.9

Inverse and determinant of a 2×2 matrix

a. The 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is invertible if (and only if) $ad - bc \neq 0$. Quantity $ad - bc$ is called the *determinant* of A, written $\det(A)$:

$$\det(A) = \det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc.$$

b. If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is invertible, then

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

Theorem 2.4.10

Geometrical interpretation of the determinant of a 2×2 matrix
If $A = \begin{bmatrix} \vec{v} & \vec{w} \end{bmatrix}$ is a 2×2 matrix with nonzero columns $\vec{v}$ and $\vec{w}$, then

$$\det(A) = \det \begin{bmatrix} \vec{v} & \vec{w} \end{bmatrix} = ||\vec{v}|| \sin\theta ||\vec{w}||,$$

where θ is the oriented angle from $\vec{v}$ to $\vec{w}$, with $-\pi \leq \theta \leq \pi$. It follows that

- $|\det(A)| = ||\vec{v}|| |\sin\theta| ||\vec{w}||$ is the *area of the parallelogram* spanned by $\vec{v}$ and $\vec{w}$.
- $\det(A) = 0$ if $\vec{v}$ and $\vec{w}$ are *parallel*, meaning that $\theta = 0$ or $\theta = \pi$.
- $\det(A) > 0$ if $0 < \theta < \pi$, and
- $\det(A) < 0$ if $-\pi < \theta < 0$.

Chapter 3

Subspaces of $\mathbb{R}^n$ and Their Dimensions

3.1 Image and Kernel of a Linear Transformation

Definition 3.1.1

Image of a function
The *image* of a function consists of all the values the function takes in its target space. If f is a function from X to Y, then $\text{image}(f) = \{f(x) : x \text{ in } X\} = \{b \text{ in } Y : b = f(x), \text{ for some } x \text{ in } X\}$

Definition 3.1.2

Span
Consider the vectors $\vec{v_1}, \ldots, \vec{v_m}$ in $\mathbb{R}^n$. The set of all linear combinations $c_1\vec{v_1} + \cdots + c_m\vec{v_m}$ of the vectors $\vec{v_1}, \ldots, \vec{v_m}$ is called their *span*:

$$\text{span}(\vec{v_1}, \ldots, \vec{v_m}) = \{c_1\vec{v_1} + \cdots + c_m\vec{v_m} : c_1, \ldots, c_m \text{ in } \mathbb{R}^n\}$$

Theorem 3.1.3

Image of a linear transformation
The image of a linear transformation $T(\vec{x}) = A\vec{x}$ is the span of the column vectors of A.* We denote the image of T by $\text{im}(T)$ or $\text{im}(A)$.
*The image of T is also called the *column space* of A.

Theorem 3.1.4

Some properties of the image
The image of a linear transformation T (from $\mathbf{R}^m$ to $\mathbf{R}^n$) has the following properties:

a. The zero vector $\vec{0}$ in $\mathbb{R}^n$ is in the image of T.

b. The image of T is *closed under addition*: If $\vec{v_1}$ and $\vec{v_2}$ are in the image of T, then so is $\vec{v_1} + \vec{v_2}$.

c. The image of T is *closed under scalar multiplication*: If $\vec{v}$ is in the image of T and k is an arbitrary scalar, then $k\vec{v}$ is in the image of T as well.

Definition 3.1.5

Kernel
The *kernel** of a linear transformation $T(\vec{x}) = A\vec{x}$ from $\mathbb{R}^m$ to $\mathbb{R}^n$ consists of all zeros of the transformation, that is, the solutions of the equation $T(\vec{x}) = A\vec{x} = \vec{0}$. In other words, the kernel of T is the solution set of the linear system

$$A\vec{x} = \vec{0}.$$

We denote the kernel of T by $\ker(T)$ or $\ker(A)$.
*The kernel of T is also called the *null space* of A.

Theorem 3.1.6

Some properties of the kernel
Consider a linear transformation T from $\mathbb{R}^m$ to $\mathbb{R}^n$.

- **a.** The zero vector $\vec{0}$ in $\mathbb{R}^m$ is in the kernel of T.
- **b.** The kernel is closed under addition.
- **c.** The kernel is closed under scalar multiplication.

Theorem 3.1.7

When is ker$(A) = \{\vec{0}\}$**?**

- **a.** Consider an $n \times n$ matrix A. Then $\ker(A) = \{\vec{0}\}$ if (and only if) $\operatorname{rank}(A) = n$.
- **b.** Consider an $n \times m$ matrix A. If $\ker(A) = \{\vec{0}\}$, then $m \leq n$. Equivalently, if $m > n$, then there are nonzero vectors in the kernel of A.
- **c.** For a *square* matrix A, we have $\ker(A) = \{\vec{0}\}$ if (and only if) A is invertible.

Summary 3.1.8

Various characterizations of invertible matrices
For an $n \times n$ matrix A, the following statements are equivalent; that is, for a given A, they are either all true or all false.

- **i.** A is invertible
- **ii.** The linear system $A\vec{x} = \vec{b}$ has a unique solution $\vec{x}$, for all $\vec{b}$ in $\mathbb{R}^n$.
- **iii.** $\operatorname{rref}(A) = I_n$.
- **iv.** $\operatorname{rank}(A) = n$.
- **v.** $\operatorname{im}(A) = \mathbb{R}^n$.
- **vi.** $\ker(A) = \{\vec{0}\}$.

3.2 Subspaces of $\mathbb{R}^n$; Bases and Linear Independence

Definition 3.2.1

Subspaces of $\mathbb{R}^n$
A subset W of the vector space $\mathbb{R}^n$ is called a (linear) *subspace* of $\mathbb{R}^n$ if it has the following three properties:

- **a.** W contains the zero vector in $\mathbb{R}^n$.
- **b.** W is closed under addition: If $\vec{w_1}$ and $\vec{w_2}$ are both in W, then so is $\vec{w_1} + \vec{w_2}$.
- **c.** W is closed under scalar multiplication: If $\vec{w}$ is in W and k is an arbitrary scalar, then $k\vec{w}$ is in W.

Theorem 3.2.2

Image and kernel are subspaces
If $T(\vec{X}) = A\vec{x}$ is a linear transformation from $\mathbb{R}^m$ to $\mathbb{R}^n$, then

- $\ker(T) = \ker(A)$ is a subspace of $\mathbb{R}^m$, and
- $\text{image}(T) = \text{im}(A)$ is a subspace of $\mathbb{R}^n$.

Definition 3.2.3

Redundant vectors; linear independence; basis
Consider vectors $\vec{v_1}, \ldots, \vec{v_m}$ in $\mathbb{R}^n$.

- **a.** We say that a vector $\vec{v_i}$ in the list $\vec{v_1}, \ldots, \vec{v_m}$ is *redundant* if $\vec{v_i}$ is a linear combination of the preceding vectors $\vec{v_1}, \ldots, \vec{v_{i-1}}$.*
- **b.** The vectors $\vec{v_1}, \ldots, \vec{v_m}$ are called *linearly independent* if none of them is redundant. Otherwise, the vectors are called *linearly dependent* (meaning that at least one of them is redundant).
- **c.** We say that the vectors $\vec{v_1}, \ldots, \vec{v_m}$ in a subspace V of $\mathbb{R}^n$ form a *basis* of V if they span V and are linearly independent.

*We call the first vector, $\vec{v_1}$ redundant if it is the zero vector. This agrees with the convention that the *empty linear combination* of vectors is the zero vector.

Theorem 3.2.4

Basis of the image
To construct a basis of the image of a matrix A, list all the column vectors of A, and omit the redundant vectors from this list.

Theorem 3.2.5

Linear independence and zero components
Consider vectors $\vec{v_1}, \ldots, \vec{v_m}$ in $\mathbb{R}^n$. If $\vec{v_1}$ is nonzero, and if each of the vectors $\vec{v_i}$ has a nonzero entry in a component where all the preceding vectors $\vec{v_1}, \ldots, \vec{v_{i-1}}$ have a 0, then the vectors $\vec{v_1}, \ldots, \vec{v_m}$ are linearly independent.

Definition 3.2.6

Linear Relations
Consider the vectors $\vec{v_1}, \ldots, \vec{v_m}$ in $\mathbb{R}^n$. An equation of the form

$$c_1\vec{v_1} + \cdots + c_m\vec{v_m} = \vec{0}$$

is called a (linear) *relation* among the vectors $\vec{v_1}, \ldots, \vec{v_m}$. There is always the *trivial relation*, with $c_1 = \cdots = c_m = 0$. *Nontrivial relations* (where at least one coefficient c_i is nonzero) may or may not exist among the vectors $\vec{v_1}, \ldots, \vec{v_m}$.

Theorem 3.2.7

Relations and linear dependence
The vectors $\vec{v_1}, \ldots, \vec{v_m}$ in $\mathbb{R}^n$ are linearly dependent if (and only if) there are nontrivial relations among them.

Theorem 3.2.8

Kernel and relations
The vectors in the kernel of an $n \times m$ matrix A correspond to the linear relations among the column vectors $\vec{v_1}, \ldots, \vec{v_m}$ of A: The equation

$$A\vec{x} = \vec{0} \quad \text{means that} \quad x_1\vec{v_1} + \cdots + x_m\vec{v_m} = \vec{0}.$$

In particular, the column vectors of A are linearly independent if (and only if) $\ker(A) = \{\vec{0}\}$, or, equivalently, if $\mathrm{rank}(A) = m$. This condition implies that $m \leq n$.
Thus, we can find at most n linearly independent vectors in $\mathbb{R}^n$.

Summary 3.2.9

Various characterizations of linear independence
For a list $\vec{v_1}, \ldots, \vec{v_m}$ of vectors in $\mathbb{R}^n$, the following statements are equivalent:

- **i.** Vectors $\vec{v_1}, \ldots, \vec{v_m}$ are linearly independent.
- **ii.** None of the vectors $\vec{v_1}, \ldots, \vec{v_m}$ is redundant, meaning that none of them is a linear combination of the preceding vectors.
- **iii.** None of the vectors $\vec{v_i}$ is a linear combination of the other vectors $\vec{v_1}, \ldots, \vec{v_{i-1}}, \vec{v_{i+1}}, \ldots, \vec{v_m}$ in the list.
- **iv.** There is only the trivial relation among the vectors $\vec{v_1}, \ldots, \vec{v_m}$, meaning that the equation $c_1\vec{v_1} + \cdots + c_m\vec{v_m} = \vec{0}$ has only the solution $c_1 = \cdots = c_m = 0$.
- **v.** $\ker \begin{bmatrix} | & & | \\ \vec{v_1} & \cdots & \vec{v_m} \\ | & & | \end{bmatrix} = \{\vec{0}\}$.
- **vi.** $\mathrm{rank} \begin{bmatrix} | & & | \\ \vec{v_1} & \cdots & \vec{v_m} \\ | & & | \end{bmatrix} = m$.

Theorem 3.2.10

Basis and unique representation
Consider the vectors $\vec{v_1}, \ldots, \vec{v_m}$ in a subspace V of $\mathbb{R}^n$.
The vectors $\vec{v_1}, \ldots, \vec{v_m}$ form a basis of V if (and only if) every vector $\vec{v}$ in V can be expressed *uniquely* as a linear combination

$$\vec{v} = c_1\vec{v_1} + \cdots + c_m\vec{v_m}.$$

(In Section 3.4, we will call the coefficients $c_1, \ldots, c_m$ the *coordinates* of $\vec{v}$ with respect to the basis $\vec{v_1}, \ldots, \vec{v_m}$.

3.3 The Dimension of a Subspace of $\mathbb{R}^n$

Theorem 3.3.1

Consider vectors $\vec{v_1}, \ldots, \vec{v_p}$ and $\vec{}_1, \ldots, \vec{w_q}$ in a subspace V of $\mathbb{R}^n$. If the vectors $\vec{v_1}, \ldots, \vec{v_p}$ are linearly independent, and the vectors $\vec{w_1}, \ldots, \vec{w_q}$ span V, then $q \geq p$.

Theorem 3.3.2

Number of vectors in a basis
All bases of a subspace V of $\mathbb{R}^n$ consist of the same number of vectors.

Definition 3.3.3

Dimension
Consider a subspace V of $\mathbb{R}^n$. The number of vectors in a basis of V is called the *dimension* of V, denoted by $\dim(V)$.

Theorem 3.3.4

Finding Bases of Kernel and Image

Independent vectors and spanning vectors in a subspace of $\mathbb{R}^n$
Consider a subspace V of $\mathbb{R}^n$ with $\dim(V) = m$.

- **a.** We can find *at most* m linearly independent vectors in V.
- **b.** We need *at least* m vectors to span V.
- **c.** If m vectors in V are linearly independent, then they form a basis of V.
- **d.** If m vectors in V span V, then they form a basis of V.

Theorem 3.3.5

Using rref to construct a basis of the image
To construct a basis of the image of A, pick the column vectors of A that correspond to the columns of $\text{rref}(A)$ containing the leading 1's.

Theorem 3.3.6

Dimension of the image
For any matrix A, $\dim(\text{im } A) = \text{rank}(A)$.

Theorem 3.3.7

Rank-nullity theorem
For any $n \times m$ matrix A, the equation

$$\dim(\ker A) + \dim(\text{im } A) = m$$

holds. The dimension of $\ker(A)$ is called the *nullity* of A, and in Theorem 3.3.6 we observed that $\dim(\text{im } A) = \text{rank}(A)$. Thos, we can write the preceeding equation alternatively as

$$(\text{nullity of } A) + (\text{rank of } A) = m$$

Some authors go as far as to call this the *fundamental theorem of linear algebra*.

Theorem 3.3.8

Finding bases of the kernel and image by inspection
Suppose you are able to spot the redundant columns of matrix A.
Express each redundant column as a linear combination of the preceding columns, $\vec{v_i} = c_1\vec{v_1} + \cdots + c_{i-1}\vec{v_{i-1}}$; write a corresponding relation, $-c_1\vec{v_1} - \cdots - c_{i-1}\vec{v_{i-1}} + \vec{v_i} = \vec{0}$; and generate the vector

$$\begin{bmatrix} -c1 \\ \vdots \\ -c_{i-1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

in the kernel of A. The vectors so constructed form a basis of the kernel of A.
The nonredundant columns form a basis of the image of A.

Bases of $\mathbb{R}^n$

Theorem 3.3.9

Bases of $\mathbb{R}^n$
The vectors $\vec{v_1}, \ldots, \vec{v_n}$ in $\mathbb{R}^n$ form a basis of $\mathbb{R}^n$ if (and only if) the matrix

$$\begin{bmatrix} | & & | \\ \vec{v_1} & \cdots & \vec{v_n} \\ | & & | \end{bmatrix}$$

is invertible.

Summary 3.3.10

Various characterizations of invertible matrices
For an $n \times n$ matrix A, the following statements are equivalent.

- **i.** A is invertible
- **ii.** The linear system $A\vec{x} = \vec{b}$ has a unique solution $\vec{x}$, for all $\vec{b}$ in $\mathbb{R}^n$.
- **iii.** $\text{rref}(A) = I_n$.
- **iv.** $\text{rank}(A) = n$.
- **v.** $\text{im}(A) = \mathbb{R}^n$.
- **vi.** $\ker(A) = \{\vec{0}\}$.
- **vii.** The column vectors of A form a basis of $\mathbb{R}^n$.
- **viii.** The column vectors of A span $\mathbb{R}^n$.
- **ix.** The column vectors of A are linearly independent.

3.4 Coordinates

Definition 3.4.1

Coordinates in a subspace of $\mathbb{R}^n$
Consider a basis $\mathfrak{B} = (\vec{v_1}, \vec{v_2}, \ldots, \vec{v_m})$ of a subspace V of $\mathbb{R}^n$. By Theorem 3.2.10, any vector $\vec{x}$ in V can be written uniquely as

$$\vec{x} = c_1\vec{v_1} + c_2\vec{v_2} + \cdots + c_m\vec{v_m}.$$

The scalars $c_1, c_2, \ldots, c_m$ are called the $\mathfrak{B}$-*coordinates* of $\vec{x}$, and the vector

$$\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix}$$

is the $\mathfrak{B}$-*coordinate vector* of $\vec{x}$, denoted by $[\vec{x}]_{\mathfrak{B}}$. Thus,

$$[\vec{x}]_{\mathfrak{B}} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix} \text{ means that } \vec{x} = c_1\vec{v_1} + c_2\vec{v_2} + \cdots + c_m\vec{v_m}.$$

Note that

$$\vec{x} = S[\vec{x}]_{\mathfrak{B}}, \text{ where } S = \begin{bmatrix} | & | & & | \\ \vec{v_1} & \vec{v_2} & \cdots & \vec{v_m} \\ | & | & & | \end{bmatrix}, \text{ an } n \times m \text{ matrix.}$$

Theorem 3.4.2

Linearity of Coordinates
If $\mathfrak{B}$ is a basis of subspace V of $\mathbb{R}^n$, then

- **a.** $[\vec{x} + \vec{y}]_{\mathfrak{B}} = [\vec{x}]_{\mathfrak{B}} + [\vec{y}]_{\mathfrak{B}}$, for all vectors $\vec{x}$ and $\vec{y}$ in V, and
- **b.** $[k\vec{x}]_{\mathfrak{B}} = k[\vec{x}]_{\mathfrak{B}}$, for all $\vec{x}$ in V and for all scalars k.

Theorem 3.4.3

The matrix of a linear transformation
Consider a linear transformation T from $\mathbb{R}^n$ to $\mathbb{R}^n$ and a basis $\mathfrak{B} = (\vec{v_1}, \ldots, \vec{v_n})$ of $\mathbb{R}^n$. Then there exists a unique $n \times n$ matrix B that transforms $[\vec{x}]_{\mathfrak{B}}$ into $[T(\vec{x})]_{\mathfrak{B}}$:

$$[T(\vec{x})]_{\mathfrak{B}} = B[\vec{x}]_{\mathfrak{B}},$$

for all $\vec{x}$ in $\mathbb{R}^n$. This matrix B is called the $\mathfrak{B}$-*matrix of* T. We can construct B column by column, as follows:

$$B = \begin{bmatrix} | & & | \\ [T(\vec{v_1})]_{\mathfrak{B}} & \cdots & [T(\vec{v_n})]_{\mathfrak{B}}. \\ | & & | \end{bmatrix}.$$

Theorem 3.4.4

Standard matrix versus $\mathfrak{B}$-matrix
Consider a linear transformation T from $\mathbb{R}^n$ to $\mathbb{R}^n$ and a basis $\mathfrak{B} = (\vec{v_1}, \ldots, \vec{v_n})$ of $\mathbb{R}^n$. Let B be the $\mathfrak{B}$-matrix of T, and let A be the standard matrix of T [such that $T(\vec{x}) = A\vec{x}$ for all $\vec{x}$ in $\mathbb{R}^n$]. Then

$$AS = SB,\ B = S^{-1}AS, \text{ and } A = SBS^{-1}, \text{ where } S = \begin{bmatrix} | & & | \\ \vec{v_1} & \cdots & \vec{v_n} \\ | & & | \end{bmatrix}.$$

Definition 3.4.5

Similar matrices
Consider two $n \times n$ matrices A and B. We say that A is similar to B if there exists an invertible matrix S such that

$$AS = SB, \text{ or } B = S^{-1}AS$$

Theorem 3.4.6

Similarity is an equivalence relation

a. An $n \times n$ matrix A is similar to A itself (*reflexivity*).

b. If A is similar to B, then B is similar to A (*symmetry*).

c. If A is similar to B and B is similar to C, then A is similar to C (*transitivity*).

Theorem 3.4.7

When is the $\mathfrak{B}$-matrix of T diagonal?
Consider a linear transformation $T(\vec{x}) = A\vec{x}$ from $\mathbb{R}^n$ to $\mathbb{R}^n$. Let $\mathfrak{B} = (\vec{v_1}, \ldots, \vec{v_n})$ be a basis of $\mathbb{R}^n$.
Then the $\mathfrak{B}$-matrix B of T is *diagonal* if and only if $T(\vec{v_1}) = c_1\vec{v_1}, \ldots, T(\vec{v_n}) = c_n\vec{v_n}$ for some scalars $c_1, \ldots, c_n$.
From a geometrical point of view, this means that $T(\vec{v_j})$ is *parallel* to $\vec{v_j}$ for all $j = 1, \ldots, n$.

Part II

First Exam to the Second Exam

Chapter 4

Linear Spaces

4.1 Introduction to Linear Spaces

Definition 4.1.1

Linear spaces (or vector spaces)
A *linear space* V is a set endowed with a rule for addition (if f and g are in V, then so is $f + g$) and a rule for scalar multiplication (if f is in V and k in $\mathbb{R}$, then kf is in V) such that the operations satisfy the following eight rules (for all f, g, h in V and all c, k in $\mathbb{R}$):

1. $(f + g) + h = f + (g + h)$.

2. $f + g = g + f$.

3. There exists a *neutral element* n in V such that $f + n = f$, for all f in V. This n is unique and denoted by 0.

4. For each f in V, there exists a g in V such that $f + g = 0$. This g is unique and denoted by $(-f)$.

5. $k(f + g) = kf + kg$.

6. $(c + k)f = cf + ck$.

7. $c(kf) = (ck)f$.

8. $1f = f$.

Definition 4.1.2

Subspaces
A subset W of a linear space V is called a *subspace* of V if

a. W contains the neutral element 0 of V.

b. W is closed under addition (if f and g are in W, then so is $f + g$).

c. W is closed under scalar multiplication (if f is in W and k is a scalar, then kf is in W).

We can summarize parts b and c by saying that W is closed under linear combinations.

Definition 4.1.3

Span, linear independence, basis, coordinates
Consider the elements $f_1, \ldots, f_n$ in a linear space V.

a. We say that $f_1, \ldots, f_n$ *span* V if every f in V can be expressed as a linear combination of $f_1, \ldots, f_n$.

b. We say that f_i is *redundant* if it is a linear combination of $f_1, \ldots, f_{i-1}$. The elements $f_1, \ldots, f_n$ are called *linearly independent* if none of them is redundant. This is the case if the equation

$$c_1 f_1 + \cdots + c_n f_n = 0$$

has only the trivial solution

$$c_1 = \cdots = c_n = 0.$$

c. We say that elements $f_1, \ldots, f_n$ are a *basis* of V if they span V and are linearly independent. This means that every f in V can be written uniquely as a linear combination $f = c_1 f_1 + \cdots + c_n f_n$. The coefficients $c_1, \ldots, c_n$ are called the *coordinates* of f with respect to the basis $\mathfrak{B} = (f_1, \ldots, f_n)$. The vector

$$\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

in $\mathbb{R}^n$ is called the $\mathfrak{B}$-*coordinate vector* of f, denoted by $[f]_{\mathfrak{B}}$.

The transformation

$$L(f) = [f]_{\mathfrak{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \text{ from } V \text{ to } \mathbb{R}^n$$

is called the $\mathfrak{B}$-*coordinate transformation*, sometimes denoted by $L_{\mathfrak{B}}$.

Theorem 4.1.4

Linearity of the coordinate transformation $L_{\mathfrak{B}}$
If $\mathfrak{B}$ is a basis of a linear space V, then

a. $[f + g]_{\mathfrak{B}} = [f]_{\mathfrak{B}} + [g]_{\mathfrak{B}}$, for all elements f and g of V, and

b. $[kf]_{\mathfrak{B}} = k[f]_{\mathfrak{B}}$, for all f in V and for all scalars k.

Theorem 4.1.5

Dimension
If a linear space V has a basis with n elements, then all other bases of V consist of n elements as well. We say that n is the *dimension* of V:

$$\dim(V) = n.$$

Summary 4.1.6

Finding a basis of a linear space V

a. Write down a typical element of V, in terms of some arbitrary constants.

b. Using the arbitrary constants as coefficients, express your typical element as a linear combination of some elements of V.

c. Verify that the elements of V in this linear combination are linearly independent; then they will form a basis of V.

Theorem 4.1.7

Linear differential equations
The solutions of the differential equation

$$f''(x) + af'(x) + bf(x) = 0 \text{ (where } a \text{ and } b \text{ are constants)}$$

form a two-dimensional subspace of the space C^∞ of smooth functions.
More generally, the solutions of the differential equation

$$f^{(n)}(x) + a_{n-1}f^{n-1}(x) + \cdots + a_1f'(x) + a_0f(x) = 0$$

(where $a_0, \ldots, a_{n-1}$ are constants) form an n-dimensional subspace of C^∞. A differential equation of this form is called an nth-*order linear differential equation with constant coefficients*.

Definition 4.1.8

Finite dimensional linear spaces
A linear space V is called *finite dimensional* if it has a (finite) basis $f_1, \ldots, f_n$, so that we can define its dimension $\dim(V) = n$. See Definition 4.1.5. Otherwise, the space is called *infinite dimensional*.

4.2 Linear Transformations and Isomorphisms

Definition 4.2.1

Linear transformations, image, kernel, rank, nullity
Consider two linear spaces V and W. A function T from V to W is called a *linear transformation* if

$$T(f + g) = T(f) + T(g) \quad \text{and} \quad T(kf) = kT(f)$$

for all elements f and g of V and for all scalars k. These two rules are referred to as the *sum rule* and the *constant-multiple rule*, respectively.
For a linear transformation T from V to W, we let

$$\text{im}(T) = \{T(f) : f \text{ in } V\}$$

and

$$\ker(T) = \{f \text{ in } V : T(f) = 0\}.$$

Note that $\text{im}(T)$ is a subspace of target space W and that $\ker(T)$ is a subspace of domain V.
If the image of T is finite dimensional, then $\dim(\text{im}T)$ is called the *rank* of T, and if the kernel of T is finite dimensional, then $\dim(\ker T)$ is the *nullity* of T.
If V is finite dimensional, then the rank-nullity theorem holds. See Theorem 3.3.7:

$$\dim(V) = \text{rank}(T) + \text{nullity}(T) = \dim(\text{im}T) + \dim(\ker(T)).$$

Definition 4.2.2

Isomorhpisms and isomorphic spaces
An invertible linear transformation T is called an *isomorphism*. We say that the linear space V is isomorphic to the linear space W if there exists an isomorphism T from V to W.

Theorem 4.2.3

Coordinate transformations are isomorphisms
If $\mathfrak{B} = (f_1, f_2, \ldots, f_n)$ is a basis of a linear space V, then the *coordinate transformation* $L_{\mathfrak{B}}(f) = [f]_{\mathfrak{B}}$ from V to $\mathbb{R}^n$ is an isomorphism. Thus, V is isomorphic to $\mathbb{R}^n$; the linear spaces V and $\mathbb{R}^n$ have the same structure.

Theorem 4.2.4

Properties of isomorphisms

a. A linear transformation T from V to W is an isomorphism if (and only if) $\ker(T) = \{0\}$ and $\mathrm{im}(T) = W$.

In parts (b) through (d), the linear spaces V and W are assumed to be finite dimensional.

b. The linear space V is isomorphic to W if (and only if) $\dim(V) = \dim(W)$.

c. Suppose T is a linear transformation from V to W with $\ker(T) = \{0\}$. If $\dim(V) = \dim(W)$, then T is an isomorphism.

d. Suppose T is a linear transformation from V to W with $\mathrm{im}(T) = W$. If $\dim(V) = \dim(W)$, then T is an isomorphism.

4.3 The Matrix of a Linear Transformation

Definition 4.3.1

The $\mathfrak{B}$-matrix of a linear transformation
Consider a linear transformation T from V to V, where V is an n-dimensional linear space. Let $\mathfrak{B}$ be a basis of V. Consider the linear transformation $L_{\mathfrak{B}} \circ T \circ L_{\mathfrak{B}}^{-1}$ from $\mathbb{R}^n$ to $\mathbb{R}^n$, with standard matrix B, meaning that $B\vec{x} = L_{\mathfrak{B}}(T(L_{\mathfrak{B}}^{-1}(\vec{x})))$. This matrix B is called the $\mathfrak{B}$-matrix of transformation T. Letting $f = L_{\mathfrak{B}}^{-1}(\vec{x})$ and $\vec{x} = [f]_{\mathfrak{B}}$, we find that

$$[T(f)]_{\mathfrak{B}} = B\,[f]_{\mathfrak{B}}\,, \qquad \text{for all } f \text{ in } V.$$

Theorem 4.3.2

The columns of the $\mathfrak{B}$-matrix of a linear transformation
Consider a linear transformation T from V to V, and let B be the matrix of T with respect to a basis $\mathfrak{B} = (f_1, f_2, \ldots, f_n)$ of V. Then

$$B = \begin{bmatrix} | & & | \\ [T(f_1)]_{\mathfrak{B}} & \cdots & [T(f_n)]_{\mathfrak{B}} \\ | & & | \end{bmatrix}.$$

The columns of B are the $\mathfrak{B}$-coordinate vectors of the transforms of the basis elements $f_1, \ldots, f_n$ of V.

Definition 4.3.3

Change of basis matrix
Consider two bases $\mathfrak{A}$ and $\mathfrak{B}$ of an n-dimensional linear space V. Consider the linear transformation $L_{\mathfrak{A}} \circ L_{\mathfrak{B}}^{-1}$ from $\mathbb{R}^n$ to $\mathbb{R}^n$, with standard matrix S, meaning that $S\vec{x} = L_{\mathfrak{A}}(L_{\mathfrak{B}}^{-1}(\vec{x}))$ for all $\vec{x}$ in $\mathbb{R}^n$. This invertible matrix S is called the *change of basis matrix* from $\mathfrak{B}$ to $\mathfrak{A}$, sometimes denoted by $S_{\mathfrak{B}\to\mathfrak{A}}$. Letting $f = L_{\mathfrak{B}}^{-1}(\vec{x})$ and $\vec{x} = [f]_{\mathfrak{B}}$, we find that

$$[f]_{\mathfrak{A}} = S[f]_{\mathfrak{B}}, \quad \text{for all } f \text{ in } V.$$

If $\mathfrak{B} = (b_1, \ldots, b_i, \ldots, b_n)$, then

$$[b_i]_{\mathfrak{A}} = S[b_i]_{\mathfrak{B}} = S\vec{e}_i = (i\text{th column of } S),$$

so that

$$S_{\mathfrak{B}\to\mathfrak{A}} = \begin{bmatrix} | & & | \\ [b_1]_{\mathfrak{A}} & \cdots & [b_n]_{\mathfrak{A}} \\ | & & | \end{bmatrix}.$$

Theorem 4.3.4

Change of basis in a subspace of $\mathbb{R}^n$
Consider a subspace V of $\mathbb{R}^n$ with two bases $\mathfrak{A} = (\vec{a_1}, \ldots, \vec{a_m})$ and $\mathfrak{B} = (\vec{b_1}, \ldots, \vec{b_m})$. Let S be the change of basis matrix from $\mathfrak{B}$ to $\mathfrak{A}$. Then the following equation holds:

$$\begin{bmatrix} | & & | \\ \vec{b_1} & \cdots & \vec{b_m} \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ \vec{a_1} & \cdots & \vec{a_m} \\ | & & | \end{bmatrix} S.$$

Theorem 4.3.5

Change of basis for the matrix of a linear transformation
Let V be a linear space with two given bases $\mathfrak{A}$ and $\mathfrak{B}$. Consider a linear transformation T from V to V, and let A and B be the $\mathfrak{A}$-matrix and $\mathfrak{B}$-matrix of T, respectively. Let S be the change of basis matrix from $\mathfrak{B}$ to $\mathfrak{A}$. Then A is similar to B, and

$$AS = SB \quad \text{or} \quad A = sBS^{-1} \quad \text{or} \quad B = S^{-1}AS.$$

Chapter 5

Orthogonality and Least Squares

5.1 Orthogonal Projections and Orthonormal Bases

Definition 5.1.1

Orthogonality, length, unit vectors

a. Two vectors $\vec{v}$ and $\vec{w}$ in $\mathbb{R}^n$ are called *perpendicular* or *orthogonal* if $\vec{v} \cdot \vec{w} = 0$.

b. The *length* (or magnitude or norm) of a vector $\vec{v}$ in $\mathbb{R}^n$ is $||\vec{v}|| = \sqrt{\vec{v} \cdot \vec{v}}$.

c. A vector $\vec{u}$ in $\mathbb{R}^n$ is called a *unit vector* if its length is 1, (i.e., $||\vec{u}|| = 1$, or $\vec{u} \cdot \vec{u} = 1$).

Definition 5.1.2

Orthonormal vectors
The vectors $\vec{u_1}, \vec{u_2}, \ldots, \vec{u_m}$ in $\mathbb{R}^n$ are called *orthonormal* if they are all unit vectors and orthogonal to one another:

$$\vec{u_i} \cdot \vec{u_j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Theorem 5.1.3

Properties of orthonormal vectors

a. Orthonormal vectors are linearly independent.

b. Orthonormal vectors $\vec{u_1}, \ldots, \vec{u_n}$ in $\mathbb{R}^n$ form a basis of $\mathbb{R}^n$.

Theorem 5.1.4

Orthogonal Projection
Consider a vector $\vec{x}$ in $\mathbb{R}^n$ and a subspace V of $\mathbb{R}^n$. Then we can write

$$\vec{x} = \vec{x}^{||} + \vec{x}^{\perp},$$

where $\vec{x}^{||}$ is in V and $\vec{x}^{\perp}$ is perpendicular to V, and this representation is unique.
The vector $\vec{x}^{||}$ is called the *orthogonal projection* of $\vec{x}$ onto V, denoted by $\text{proj}_V \vec{x}$.
The transformation $T(\vec{x}) = \text{proj}_V \vec{x} = \vec{x}^{||}$ from $\mathbb{R}^n$ to $\mathbb{R}^n$ is linear.

Theorem 5.1.5

Formula for the orthogonal projection
If V is a subspace of $\mathbb{R}^n$ with an orthonormal basis $\vec{u_1}, \ldots, \vec{u_m}$, then

$$\text{proj}_V \vec{x} = \vec{x}^{\parallel} = (\vec{u_1} \cdot \vec{x})\vec{u_1} + \cdots + (\vec{u_m} \cdot \vec{x})\vec{u_m},$$

for all $\vec{x}$ in $\mathbb{R}^n$.

Theorem 5.1.6

Consider an orthonormal basis $\vec{u_1}, \ldots, \vec{u_n}$ of $\mathbb{R}^n$. Then

$$\vec{x} = (\vec{u_1} \cdot \vec{x})\vec{u_1} + \cdots + (\vec{u_n} \cdot \vec{x})\vec{u_n},$$

for all $\vec{x}$ in $\mathbb{R}^n$.

Definition 5.1.7

Orthogonal Complement
Consider a subspace V of $\mathbb{R}^n$. The *orthogonal complement* $V^\perp$ of V is the set of those vectors $\vec{x}$ in $\mathbb{R}^n$ that are orthogonal to all vectors in V:

$$V^\perp = \{\vec{x} \text{ in } \mathbb{R}^n : \vec{v} \cdot \vec{x} = 0, \text{ for all } \vec{v} \text{ in } V\}.$$

Note that $V^\perp$ is the kernel of the orthogonal projection onto V.

Theorem 5.1.8

Properties of the orthogonal complement
Consider a subspace V of $\mathbb{R}^n$.

- **a.** The orthogonal complement $V^\perp$ of V is a subspace of $\mathbb{R}^n$.
- **b.** The intersection of V and $V^\perp$ consists of the zero vector: $V \cap V^\perp = \{\vec{0}\}$.
- **c.** $\dim(V) + \dim(V^\perp) = n$.
- **d.** $(V^\perp)^\perp = V$.

Theorem 5.1.9

Pythagorean theorem
Consider two vectors $\vec{x}$ and $\vec{y}$ in $\mathbb{R}^n$. The equation

$$||\vec{x} + \vec{y}||^2 = ||\vec{x}||^2 + ||\vec{y}||^2$$

holds if (and only if) $\vec{x}$ and $\vec{y}$ are orthogonal.

Theorem 5.1.10

An inequality for the magnitude of proj$_V(\vec{x})$
Consider a subspace V of $\mathbb{R}^n$ and a vector $\vec{x}$ in $\mathbb{R}^n$. Then

$$||\text{proj}_V \vec{x}|| \leq ||\vec{x}||.$$

The statement is an equality if (and only if) $\vec{x}$ is in V.

Theorem 5.1.11

Cauchy-Schwarz inequality
If $\vec{x}$ and $\vec{y}$ are vectors in $\mathbb{R}^n$, then

$$|\vec{x} \cdot \vec{y}| \leq ||\vec{x}||\,||\vec{y}||.$$

This statement is an equality if (and only if) $\vec{x}$ and $\vec{y}$ are parallel.

Definition 5.1.12 Angle between two vectors
Consider two nonzero vectors $\vec{x}$ and $\vec{y}$ in $\mathbb{R}^n$. The angle θ between these vectors is defined as

$$\theta = \arccos \frac{\vec{x} \cdot \vec{y}}{||\vec{x}||\,||\vec{y}||}$$

5.2 Gram-Schmidt Process and QR Factorization

Theorem 5.2.1

The Gram-Schmidt process
Consider a basis $\vec{v_1}, \dots, \vec{v_m}$ of a subspace V of $\mathbb{R}^n$. For $j = 2, \dots, m$, we resolve the vector $\vec{v_j}$ into its components parallel and perpendicular to the span of the preceeding vectors, $\vec{v_1}, \dots, \vec{v_{j-1}}$:

$$\vec{v_j} = \vec{v_j}^{\parallel} + \vec{v_j}^{\perp}, \quad \text{with respect to span}(\vec{v_1}, \dots, \vec{v_{j-1}}).$$

Then

$$\vec{u_1} = \frac{1}{||\vec{v_1}||}\vec{v_1}, \;\; \vec{u_2} = \frac{1}{||\vec{v_2}^{\perp}||}\vec{v_2}^{\perp}, \dots, \;\; \vec{u_j} = \frac{1}{||\vec{v_j}^{\perp}||}\vec{v_j}^{\perp}, \dots, \vec{u_m} = \frac{1}{||\vec{v_m}^{\perp}||}\vec{v_m}^{\perp}$$

is an orthonormal basis of V. By Theorem 5.1.7, we have

$$\vec{v_j}^{\perp} = \vec{v_j} - \vec{v_j}^{\parallel} = \vec{v_j} - (\vec{u_1} \cdot \vec{v_j})\vec{u_1} - \cdots - (\vec{u_{j-1}} \cdot \vec{v_j})\vec{u_{j-1}}$$

The QR Factorization

Theorem 5.2.2

QR **factorization**
Consider an $n \times m$ matrix M with linearly independent columns $\vec{v_1}, \dots, \vec{v_m}$. Then there exists an $n \times m$ matrix Q whose columns $\vec{u_1}, \dots, \vec{u_m}$ are orthonormal and an upper triangular matrix R with positive diagonal entries such that

$$M = QR$$

This representation is unique. Furthermore, $r_{11} = ||\vec{v_1}||$, $r_{jj} = ||\vec{v_j}^{\perp}||$ (for $j = 2, \dots, m$) and $r_{ij} = \vec{u_i} \cdot \vec{v_j}$ (for $i \leq j$).

Theorem 5.2.3

QR **factorization**
Consider an $n \times m$ matrix M with linearly independent columns $\vec{v_1}, \dots \vec{v_m}$. Then the columns $\vec{u_1}, \dots, \vec{u_m}$ of Q and the entries r_{ij} of R can be computed in the following order:

first column of R, first column of Q;

second column of R, second column of Q;

third column of R, third column of Q;

and so on.

More specifically,

$$r_{11} = ||\vec{v_1}||, \quad \vec{u_1} = \frac{1}{r_{11}}\vec{v_1};$$

$$r_{12} = \vec{u_1} \cdot \vec{v_2}, \quad \vec{v_2}^{\perp} = \vec{v_2} - r_{12}\vec{u_1} \quad r_{22} = ||\vec{v_2}^{\perp}||, \quad \vec{u_2} = \frac{1}{r_{22}}\vec{v_2}^{\perp};$$

$$r_{13} = \vec{u_1} \cdot \vec{v_3}, \quad r_{23} = \vec{u_2} \cdot \vec{v_3}, \quad \vec{v_3}^{\perp} = \vec{v_3} - r_{13}\vec{u_1} - r_{23}\vec{u_2}, \quad r_{33} = ||\vec{v_3}^{\perp}||, \quad \vec{u_3} = \frac{1}{r_{33}}\vec{v_3}^{\perp};$$

and so on.

5.3 Orthogonal Transformations and Orthogonal Matrices

Definition 5.3.1

Orthogonal Transformations and orthogonal matrices
A linear transformation T from $\mathbb{R}^n$ to $\mathbb{R}^n$ is called *orthogonal* if it preserves the length of vectors:

$$||T(\vec{x})|| = ||\vec{x}||, \quad \text{for all } \vec{x} \text{ in } \mathbb{R}^n.$$

If $T(\vec{x}) = A\vec{x}$ is an orthogonal transformation, we say that A is an *orthogonal matrix*.

Theorem 5.3.2

Orthogonal transformations preserve orthogonality
Consider an orthogonal transformation T from $\mathbb{R}^n$ to $\mathbb{R}^n$. If the vectors $\vec{v}$ and $\vec{w}$ in $\mathbb{R}^n$ are orthogonal, then so are $T(\vec{v})$ and $T(\vec{w})$.

Theorem 5.3.3

Orthogonal transformations and orthonormal bases

- **a.** A linear transformation T from $\mathbb{R}^n$ to $\mathbb{R}^n$ is orthogonal if (and only if) the vectors $T(\vec{e_1}), T(\vec{e_2}), \ldots, T(\vec{e_n})$ form an orthonormal basis of $\mathbb{R}^n$.
- **b.** An $n \times n$ matrix A is orthogonal if (and only if) its columns form an orthonormal basis of $\mathbb{R}^n$.

Theorem 5.3.4

Products and inverses of orthogonal matrices

- **a.** The product AB of two orthogonal $n \times n$ matrices A and B is orthogonal.
- **b.** The inverse A^{-1} of an orthogonal $n \times n$ matrix A is orthogonal.

Definition 5.3.5

The transpose of a matrix; symmetric and skew-symmetric matrices
Consider an $m \times n$ matrix A.
The *transpose* A^T of A is the $n \times m$ matrix whose ijth entry is the jith entry of A: The roles of rows and columns are reversed.
We say that a square matrix A is *symmetric* if $A^T = A$, and A is called *skew-symmetric* if $A^T = -A$.

Theorem 5.3.6

If $\vec{v}$ and $\vec{w}$ are two (column) vectors in $\mathbb{R}^n$, then $\vec{v} \cdot \vec{w} = \vec{v}^T\vec{w}$.

Theorem 5.3.7

Consider an $n \times n$ matrix A. The matrix A is orthogonal if (and only if) $A^T A = I_n$, or, equivalently, if $A^{-1} = A^T$.

Summary 5.3.8

Orthogonal Matrices
Consider an $n \times n$ matrix A. Then the following statements are equivalent:

i. A is an orthogonal matrix.

ii. The transformation $L(\vec{x}) = A\vec{x}$ preserves lengths; that is, $||A\vec{x}|| = ||\vec{x}||$ for all $\vec{x}$ in $\mathbb{R}^n$.

iii. The columns of A form an orthonormal basis of $\mathbb{R}^n$.

iv. $A^T A = I_n$.

v. $A^{-1} = A^T$.

vi. A preserves the dot product, meaning that $(A\vec{x}) \cdot (A\vec{y}) = \vec{x} \cdot \vec{y}$ for all $\vec{x}$ and $\vec{y}$ in $\mathbb{R}^n$.

Theorem 5.3.9

Properties of the transpose

a.	$(A+B)^T = A^T + B^T$	for all $m \times n$ matrices A and B.
b.	$(kA)^T = kA^T$	for all $m \times n$ matrices A and for all scalars k.
c.	$(AB)^T = B^T A^T$	for all $m \times p$ matrices A and for all $p \times n$ matrices B.
d.	$\operatorname{rank}(A^T) = \operatorname{rank}(A)$	for all matrices A.
e.	$(A^T)^{-1} = (A^{-1})^T$	for all invertible $n \times n$ matrices A.

Theorem 5.3.10

The matrix of an orthogonal projection
Consider a subspace V of $\mathbb{R}^n$ with orthonormal basis $\vec{u_1}, \vec{u_2}, \ldots, \vec{u_m}$. The matrix P of the orthogonal projection onto V is

$$P = QQ^T, \quad \text{where} \quad Q = \begin{bmatrix} | & | & & | \\ \vec{u_1} & \vec{u_2} & \cdots & \vec{u_m} \\ | & | & & | \end{bmatrix}.$$

Pay attention to the order of the factors (QQ^T as opposed to $Q^T Q$). Note that matrix P is symmetric, since $P^T = (QQ^T)^T = (Q^T)^T Q^T = QQ^T = P$.

5.4 Least Squares and Data Fitting

Theorem 5.4.1

For any matrix A, $(\text{im}A)^\perp = \ker(A^\perp)$.

Theorem 5.4.2

a. If A is an $n \times m$ matrix, then $\ker(A) = \ker(A^T A)$.

b. If A is an $n \times m$ matrix with $\ker(A) = \{\vec{0}\}$, then $A^T A$ is invertible.

An Alternative Characterization of Orthogonal Projections

Theorem 5.4.3

Consider a vector $\vec{x}$ in $\mathbb{R}^n$ and a subspace V of $\mathbb{R}^n$. Then the orthogonal projection $\text{proj}_V\vec{x}$ is the vector in V *closest* to $\vec{x}$, in that

$$||\vec{x} - \text{proj}_V\vec{x}|| < ||\vec{x} - \vec{v}||,$$

for all $\vec{v}$ in V different from $\text{proj}_V\vec{x}$.

Definition 5.4.4

Least-squares solution
Consider a linear system $A\vec{x} = B$, where A is an $n \times m$ matrix. A vector $\vec{x}^*$ in $\mathbb{R}^m$ is called a *least-squares solution* of this system if $||\vec{b} - A\vec{x}^*|| \leq ||\vec{b} - A\vec{x}||$ for all $\vec{x}$ in $\mathbb{R}^m$.

Theorem 5.4.5

The normal equation
The least-squares solutions of the system $A\vec{x} = \vec{b}$ are the exact solutions of the (consistent) system $A^TA\vec{x} = A^T\vec{b}$. The system $A^TA\vec{x} = A^T\vec{b}$ is called the *normal equation* of $A\vec{x} = \vec{b}$.

Theorem 5.4.6

If $\ker(A) = \vec{0}$, then the linear system $A\vec{x} = \vec{b}$ has the unique least squares solution $\vec{x}^* = (A^TA)^{-1}A^T\vec{b}$.

Theorem 5.4.7

The matrix of an orthogonal projection
Consider a subspace V of $\mathbb{R}^n$ with basis $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_m}$. Let

$$A = \begin{bmatrix} | & | & & | \\ \vec{v_1} & \vec{v_2} & \cdots & \vec{v_m} \\ | & | & & | \end{bmatrix}$$

Then the matrix of the orthogonal projection onto V is $A(A^TA)^{-1}A^T$.

5.5 Inner Product Spaces

Definition 5.5.1

Inner products and inner product spaces
An *inner product* in a linear space V is a rule that assigns a real scalar (denoted by $\langle f, g\rangle$) to any pair f, g elements of V, such that the following properties hold for all f, g, h in V, and all c in $\mathbb{R}$.

- **a.** $\langle f, g\rangle = \langle g, f\rangle$ (symmetry).
- **b.** $\langle f + h, g\rangle = \langle f, g\rangle + \langle h, g\rangle$.
- **c.** $\langle cf, g\rangle = c\langle f, g\rangle$.
- **d.** $\langle f, f\rangle > 0$, for all nonzero f in V (positive definiteness.

A linear space endowed with an inner product is called an *inner product space.*

Definition 5.5.2

Norm, orthogonality
The *norm* (or magnitude) of an element f of an inner product space is

$$||f|| = \sqrt{\langle f, f \rangle}.$$

Two elements f and g of an inner product space are called *orthogonal* (or perpendicular) if $\langle f, g \rangle = 0$.

Theorem 5.5.3

Orthogonal Projection
If $g_1, \ldots, g_m$ is an orthonormal basis of a subspace W of an inner product space V, then

$$\text{proj}_W f = \langle g_1, f \rangle g_1 + \cdots + \langle g_m, f \rangle g_m,$$

for all f in V.

Part III

Second Exam to the Final

Chapter 6

Determinants

6.1 Introduction to Determinants

The Determinant of a 3×3 Matrix

Definition 6.1.1

Determinant of a 3×3 matrix, in terms of the columns
If $A = \begin{bmatrix} \vec{u} & \vec{v} & \vec{w} \end{bmatrix}$, then $\det A = \vec{u} \cdot (\vec{v} \times \vec{w})$.
A 3×3 matrix A is invertible if (and only if) $\det(A) \neq 0$.

Theorem 6.1.2

Sarrus's rule
To find the determinant of a 3×3 matrix A, write the first two columns of A to the right of A. Then multiply the entries along the six diagonals shown in the book. Add or subtract these diagonal products, as shown in the diagram. (I don't know how to put this into LaTeX, so check the textbook.)

$$\det A = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}.$$

Linearity Properties of the Determinant

No theorems or definitions, except that the determinant is not a linear transformation, but has some linear properties.

The Determinant of an $n \times n$ Matrix

Definition 6.1.3

Patterns, inversions, and determinants
A *pattern* in an $n \times n$ matrix A is a way to choose n entries of the matrix so that there is one chosen entry in each row and in each column of A.
With a pattern P we associate the product of all its entries, denoted prod P.
Two entries in a pattern are said to be *inverted* if one of them is located to the right and above the other in the matrix.
The *signature* of a pattern P is defined as $\text{sgn } P = (-1)^{(\text{number of inversions in } P)}$.
The determinant of A is defined as

$$\det A = \sum (\text{sgn } P)(\text{prod } P),$$

where the sum is taken over all $n!$ patterns P in the matrix A. Thus, we are summing up the products associated with all patterns with an even number of inversions, and we are subtracting the products associated with the patterns with an odd number of inversions.

Theorem 6.1.4

Determinant of a triangular matrix
The determinant of an (upper or lower) triangular matrix is the product of the diagonal entries of the matrix. In particular, the determinant of a diagonal matrix is the product of its diagonal entries.

6.2 Properties of the Determinant

Theorem 6.2.1

Determinant of the transpose
If A is a square matrix, then $\det(A^T) = \det A$.

Theorem 6.2.2

Linearity of the determinant in the rows and columns
Consider the fixed row vectors $\vec{v_1}, \dots, \vec{v_{i-1}}, \vec{v_{i+1}}, \dots \vec{v_n}$ with n components. Then the function

$$T(\vec{x}) = \det \begin{bmatrix} - & \vec{v_1} & - \\ & \vdots & \\ - & \vec{v_{i-1}} & - \\ - & \vec{x} & - \\ - & \vec{v_{i+1}} & - \\ & \vdots & \\ - & \vec{v_n} & - \end{bmatrix} \quad \text{from } \mathbb{R}^{1\times n} \text{ to } \mathbb{R}$$

is a linear transformation. This property is referred to as *linearity of the determinant in the* ith *row.* Likewise, the determinant is *linear in all the columns.*

Theorem 6.2.3

Elementary row operations and determinants

- **a.** If B is obtained from A by dividing a row of A by a scalar k, then $\det B = (1/k)\det A$. (i.e. $\det A = k\det B$)
- **b.** If B is obtained from A by a row swap, then $\det B = -\det A$. We say that the determinant is *alternating* on the rows.
- **c.** If B is obtained from A by adding a multiple of a row of A to another row, then $\det B = \det A$.

Analogous results hold for elementary column operations.

Theorem 6.2.4

Invertibility and determinant
A square matrix A is invertible if and only if $\det A \neq 0$.

Theorem 6.2.6

Determinants of products and powers
If A and B are $n \times n$ matrices and m is a positive integer, then

- **a.** $\det(AB) = (\det A)(\det B)$, and
- **b.** $\det(A^m) = (\det A)^m$.

Theorem 6.2.7

Determinants of similar matrices
If matrix A is similar to B, then $\det A = \det B$.

Theorem 6.2.8

Determinant of an inverse
If A is an invertible matrix, then

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

Definition 6.2.9

Minors
For an $n \times n$ matrix A, let A_{ij} be the matrix obtained by omitting the ith row and the jth column of A. The determinant of the $(n-1) \times (n-1)$ matrix A_{ij} is called a *minor* of A.

Theorem 6.2.10

Laplace expansion (or cofactor expansion)
We can compute the determinant of an $n \times n$ matrix A by Laplace expansion down any column or along any row.

Expansion down the jth column:

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij})$$

Expansion along the ith row:

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij})$$

Theorem 6.2.11

The determinant of a linear transformation
Consider a linear transformation T from V to V, where V is a finite-dimensional linear space. If $\mathfrak{B}$ is a basis of V and B is the $\mathfrak{B}$-matrix of T, then we define $\det T = \det B$. This determinant is independent of the basis $\mathfrak{B}$ we choose.

6.3 Geometrical Interpretations of the Determinant; Cramer's Rule

Theorem 6.3.1

The determinant of an orthogonal matrix is either 1 or -1.

Definition 6.3.2

Rotation matrices
An orthogonal $n \times n$ matrix A with $\det A = 1$ is called a *rotation matrix*, and the linear transformation $T(\vec{x}) = A\vec{x}$ is called a *rotation*.

Theorem 6.3.3

The determinant in terms of the columns
If A is an $n \times n$ matrix with columns $\vec{v_1}, \vec{v_2}, \dots, \vec{v_n}$, then $|\det A| = ||\vec{v_1}||||\vec{v_2}^\perp|| \cdots ||\vec{v_n}^\perp||$, where $\vec{v_k}^\perp$ is the component of $\vec{v_k}$ perpendicular to span$(\vec{v_1}, \dots, \vec{v_{k-1}})$.

Theorem 6.3.4

Volume of a parallelepiped in $\mathbb{R}^3$
Consider a 3×3 matrix $A = \begin{bmatrix} \vec{v_1} & \vec{v_2} & \vec{v_3} \end{bmatrix}$. Then the volume of the parallelepiped defined by $\vec{v_1}$, $\vec{v_2}$, and $\vec{v_3}$ is $|\det A|$.

Definition 6.3.5

Parallelepipeds in $\mathbb{R}^n$
Consider the vectors $\vec{v_1}, \vec{v_2}, \dots, \vec{v_m}$ in $\mathbb{R}^n$. The m-paralellepiped defined by the vectors $\vec{v_1}, \dots, \vec{v_m}$ is the set of all vectors in $\mathbb{R}^n$ of the form $c_1\vec{v_1} + c_2\vec{v_2} + \cdots + c_m\vec{v_m}$, where $0 \leq c_i \leq 1$. The m-volume $V(\vec{v_1}, \dots, \vec{v_m})$ of this m-parallelepiped is defined recursively by $V(\vec{v_1}) = ||\vec{v_1}||$ and $V(\vec{v_1}, \dots, \vec{v_m}) = V(\vec{v_1}, \dots, \vec{v_{m-1}})||\vec{v_m}^\perp||$.

Theorem 6.3.6

Volume of a parallelepiped in $\mathbb{R}^n$
Consider the vectors $\vec{v_1}, \vec{v_2}, \dots, \vec{v_m}$ in $\mathbb{R}^n$. Then the m-volume of the m-parallelepiped defined by the vectors $\vec{v_1}, \dots, \vec{v_m}$ is $\sqrt{\det(A^T A)}$, where A is the $n \times m$ matrix with columns $\vec{v_1}, \vec{v_2}, \dots, \vec{v_m}$. In particular, if $m = n$, this volume is $|\det A|$.

Theorem 6.3.7

Expansion factor
Consider a linear transformation $T(\vec{x}) = A\vec{x}$ from $\mathbb{R}^2$ to $\mathbb{R}^2$. Then $|\det A|$ is the *expansion factor*

$$\frac{\text{area of } T(\Omega)}{\text{area of } \Omega}$$

of T on parallelograms Ω.
Likewise, for a linear transformation $T(\vec{x}) = A\vec{x}$ from $\mathbb{R}^n$ to $\mathbb{R}^n$, $|\det A|$ is the expansion factor of T on n-parallelepipeds:

$$V(A\vec{v_1}, \dots, A\vec{v_n}) = |\det A| V(\vec{v_1}, \dots, \vec{v_n}),$$

for all vectors $\vec{v_1}, \dots, \vec{v_n}$ in $\mathbb{R}^n$.

Theorem 6.3.8

Carmer's rule
Consider the linear system $A\vec{x} = \vec{b}$, where A is an invertible $n \times n$ matrix. The components x_i of the solution vector $\vec{x}$ are

$$x_i = \frac{\det(A_{\vec{b},i})}{\det A},$$

where $A_{\vec{b},i}$ is the matrix obtained by replacing the ith column of A by $\vec{b}$.

Theorem 6.3.9

Adjoint and inverse of a matrix
Consider an invertible $n \times n$ matrix A. The *classical adjoint* adj(A) is the $n \times n$ matrix whose ijth entry is $(-1)^{i+j}\det(A_{ij})$. Then

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

Chapter 7

Eigenvalues and Eigenvectors

7.1 Diagonalization

Definition 7.1.1

Diagonalizable matrices
Consider a linear transformation $T(\vec{x}) = A\vec{x}$ from $\mathbb{R}^n$ to $\mathbb{R}^n$. Then A (or T) is said to *diagonalizable* if the matrix B of T with respect to some basis is diagonal.

By Theorem 3.4.4 and Definition 3.4.5, matrix A is diagonalizable if (and only if) A is similar to some diagonal matrix B, meaning that there exists an invertible matrix S such that $S^{-1}AS = B$ is diagonal.

To *diagonalize* a square matrix A means to find an invertible matrix S and a diagonal matrix B such that $S^{-1}AS = B$.

Definition 7.1.2

Eigenvectors, eigenvalues, and eigenbases
Consider a linear transformation $T(\vec{x}) = A\vec{x}$ from $\mathbb{R}^n$ to $\mathbb{R}^n$.

A nonzero vector $\vec{v}$ in $\mathbb{R}^n$ is called an **eigenvector** of A (or T) if $A\vec{v} = \lambda\vec{v}$ for some scalar λ. This λ. is called the **eigenvalue** associated with the eigenvector $\vec{v}$.

A basis $\vec{v_1}, \dots, \vec{v_n}$ of $\mathbb{R}^n$ is called an **eigenbasis** for A (or T) if the vectors $\vec{v_1}, \dots, \vec{v_n}$ are eigenvectors of A, meaning that $A\vec{v_1} = \lambda_1\vec{v_1}, \dots, A\vec{v_n} = \lambda_n\vec{v_n}$ for some scalars $\lambda_1, \dots, \lambda_n$.

Theorem 7.1.3

Eigenbases and diagonalization
The matrix A is diagonalizable if (and only if) there exists an eigenbasis for A. If $\vec{v_1}, \dots, \vec{v_n}$ is an eigenbasis for A, with $A\vec{v_1} = \lambda_1\vec{v_1}, \dots, A\vec{v_n} = \lambda_n\vec{v_n}$, then the matrices

$$S = \begin{bmatrix} | & | & & | \\ \vec{v_1} & \vec{v_2} & \cdots & \vec{v_n} \\ | & | & & | \end{bmatrix} \text{ and } B = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

will diagonalize A, meaning that $S^{-1}AS = B$.

Conversely, if the matrices S and B diagonalize A, then the column vectors of S will form an eigenbasis for A, and the diagonal entries of B will be the associated eigenvalues.

Theorem 7.1.4

The possible real eigenvalues of an orthogonal matrix are 1 and -1.

Summary 7.1.5

Various characterizations of invertible matrices
For an $n \times n$ matrix A, the following statements are equivalent.

i. A is invertible.

ii. The linear system $A\vec{x} = \vec{b}$ has a unique solution $\vec{x}$, for all $\vec{b}$ in $\mathbb{R}^n$.

iii. $\text{rref}A = I_n$.

iv. $\text{rank}A = n$.

v. $\text{im}A = \mathbb{R}^n$.

vi. $\ker A = \{\vec{0}\}$

vii. The column vectors of A form a basis of $\mathbb{R}^n$.

viii. The column vectors of A span $\mathbb{R}^n$.

ix. The column vectors of A are linearly independent.

x. $\det A \neq 0$.

xi. 0 fails to be an eigenvalue of A.

7.2 Finding the Eigenvalues of a Matrix

Theorem 7.2.1

Eigenvalues and determinants; characteristic equation
Consider an $n \times n$ matrix A and a scalar λ. Then λ is an eigenvalue of A if (and only if) $\det(A - \lambda I_n) = 0$. This is called the *characteristic equation* (or the *secular equation*) of matrix A.

Theorem 7.2.2

Eigenvalues of a triangular matrix
The eigenvalues of a triangular matrix are its diagonal entries.

Definition 7.2.3

Trace
The sum of the diagonal entries of a square matrix A is called the *trace* of A, denoted by $\text{tr}A$.

Theorem 7.2.4

Characteristic equation of a 2×2 matrix A

$$\det(A - \lambda I_2) = \lambda^2 - (\text{tr}A)\lambda + \det A = 0$$

Theorem 7.2.5

Characteristic polynomial
If A is an $n \times n$ matrix, then $\det(A - \lambda I_n)$ is a polynomial of degree n, of the form

$$(-\lambda)^n + (\text{tr}A)(-\lambda)^{n-1} + \cdots + \det A$$

$$= (-1)^n\lambda^n + (-1)^{n-1}(\text{tr}A)\lambda^{n-1} + \cdots + \det A.$$

This is called the *characteristic polynomial* of A, denoted by $f_A(\lambda)$.

Definition 7.2.6

Algebraic multiplicity of an eigenvalue
We say that an eigenvalue λ_0 of a square matrix A has *algebraic multiplicity* k if λ_0 is a root of multiplicity k of the characteristic polynomial $f_A(\lambda)$, meaning that we can write $f_A(\lambda) = (\lambda_0 - \lambda)^k g(\lambda)$ for some polynomial $g(\lambda)$ with $g(\lambda_0) \neq 0$. We write $\text{almu}(\lambda_0) = k$.

Theorem 7.2.7

Number of eigenvalues
An $n \times n$ matrix has *at most* n real eigenvalues, even if they are counted with their algebraic multiplicities. If n is odd, then an $n \times n$ matrix has *at least* one eigenvalue.

Theorem 7.2.8

Eigenvalues, determinant, and trace
If an $n \times n$ matrix A has the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ listed with their algebraic multiplicities, then $\det A = \lambda_1 \lambda_2 \cdots \lambda_n$, and $\text{tr} A = \lambda_1 + \lambda_2 + \cdots + \lambda_n$.

7.3 Finding the Eigenvectors of a Matrix

Theorem 7.3.1

Eigenspaces
Consider an eigenvalue λ of an $n \times n$ matrix A. Then the kernel of the matrix $A - \lambda I_n$ is called the *eigenspace* associated with λ, denoted by E_λ:

$$E_\lambda = \ker(A - \lambda I_n) = \{\vec{v} \text{ in } \mathbb{R}^n : A\vec{v} = \lambda \vec{v}\}.$$

Definition 7.3.2

Geometric multiplicity
Consider an eigenvalue λ of an $n \times n$ matrix A. The dimension of eigenspace $E_\lambda = \ker(A - \lambda I_n)$ is called the *geometric multiplicity* of eigenvalue λ, denoted by $\text{gemu}(\lambda)$. Thus,

$$\text{gemu}(\lambda) = \text{nullity}(A - \lambda I_n) = n - \text{rank}(A - \lambda I_n)$$

Theorem 7.3.3

Eigenbases and geometric multiplicities

- **a.** Consider an $n \times n$ matrix A. If we find a basis of each eigenspace of A and concatenate all these bases, then the resulting eigenvectors $\vec{v_1}, \ldots, \vec{v_s}$ will be linearly independent. (Note that s is the sum of the geometric multiplicities of the eigenvalues of A.) This result implies that $s \leq n$.
- **b.** Matrix A is diagonalizable if (and only if) the geometric multiplicities of the eigenvalues add up to n (meaning that $s = n$ in part a).

Theorem 7.3.4

An $n \times n$ matrix with n distinct eigenvalues
If an $n \times n$ matrix A has n distinct eigenvalues, then A is diagonalizable. We can construct an eigenbasis by finding an eigenvector for each eigenvalue.

Theorem 7.3.5

The eigenvalues of similar matrices
Suppose matrix A is similar to B. Then

a. Matrices A and B have the same characteristic polynomial; that is, $f_A(\lambda) = f_B(\lambda)$.

b. $\text{rank}A = \text{rank}B$ and $\text{nullity}A = \text{nullity}B$.

c. Matrices A and B have the same eigenvalues, with the same algebraic and geometric multiplicities. (However, the eigenvectors need not be the same.)

d. Matrices A and B have the same determinant and the same trace: $\det A = \det B$ and $\text{tr}A = \text{tr}B$.

Theorem 7.3.6

Algebraic versus geometric multiplicity
If λ is an eigenvalue of a square matrix A, then $\text{gemu}(\lambda) \leq \text{almu}(\lambda)$.

Theorem 7.3.7

Strategy for diagonalization
Suppose we are asked to determine whether a given $n \times n$ matrix A is diagonalizable. If so, we wish to find an invertible matrix S such that $S^{-1}AS = B$ is diagonal.

We can proceed as follows.

a. Find the eigenvalues of A by solving the characteristic equation $f_A(\lambda) = \det(A - \lambda I_n) = 0$.

b. For each eigenvalue λ, find a basis of the eigenspace $E_\lambda = \ker(A - \lambda I_n)$.

c. Matrix A is diagonalizable if (and only if) the dimensions of the eigenspaces add up to n. In this case, we find an eigenbasis $\vec{v_1}, \ldots, \vec{v_n}$ for A by concatenating the bases of the eigenspaces we found in part b.

Let

$$S = \begin{bmatrix} | & | & & | \\ \vec{v_1} & \vec{v_2} & \cdots & \vec{v_n} \\ | & | & & | \end{bmatrix}. \text{ Then } S^{-1}AS = B = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

where λ_j is the eigenvalue associated with $\vec{v_j}$.

7.4 More on Dynamical Systems

Theorem 7.4.2

Powers of a diagonalizable matrix
If

$$S^{-1}AS = B = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix},$$

then

$$A^t = SB^tS^{-1} = S \begin{bmatrix} \lambda_1^t & 0 & \cdots & 0 \\ 0 & \lambda_2^t & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^t \end{bmatrix} S^{-1}$$

Definition 7.4.3

The eigenvalues of a linear transformation
Consider a linear transformation T from V to V, where V is a linear space. A scalar λ is called an *eigenvalue* of T if there exists a nonzero element f of V such that $T(f) = \lambda f$.

Such an f is called an *eigenfunction* if V consists of functions, an *eigenmatrix* if V consists of matrices, and so on. In theoretical work, the inclusive term *eigenvector* is often used for f.

Now suppose that V is finite dimensional. Then a basis $\mathfrak{B}$ of V consisting of eigenvectors of T is called an *eigenbasis* for T. We say that transformation T is *diagonalizable* if the matrix of T with respect to some basis is diagonal. Transformation T is diagonalizable if (and only if) there exists an eigenbasis for T.

7.5 Complex Eigenvalues

Theorem 7.5.2

Fundamental theorem of algebra
Any polynomial $p(\lambda)$ with complex coefficients splits; that is, it can be written as a product of linear factors

$$p(\lambda) = k(\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n),$$

for some complex numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$, and k. (The λ_i need not be distinct.) Therefore, a polynomial $p(\lambda)$ of degree n has precisely n complex roots if they are properly counted with their multiplicities.

Theorem 7.5.4

A complex $n \times n$ matrix has n complex eigenvalues if they are counted with their algebraic multiplicities.

Theorem 7.5.5

Trace, determinant, and eigenvalues
Consider an $n \times n$ matrix A with complex eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, listed with their algebraic multiplicities. Then

$$\mathrm{tr}A = \lambda_1 + \lambda_2 + \cdots + \lambda_n$$
$$\det A = \lambda_1 \lambda_2 \cdots \lambda_n.$$

Chapter 8

Symmetric Matrices and Quadratic Forms

8.1 Symmetric Matrices

Theorem 8.1.1

Spectral theorem
A matrix A is *orthogonally diagonalizable* (i.e., there exists an orthogonal S such that $S^{-1}AS = S^T AS$ is diagonal) if and only if A is *symmetric* (i.e., $A^T = A$).

Theorem 8.1.2

Consider a symmetric matrix A. If $\vec{v_1}$ and $\vec{v_2}$ are eigenvectors of A with *distinct* eigenvalues λ_1 and λ_2, then $\vec{v_1} \cdot \vec{v_2} = 0$; that is, $\vec{v_2}$ is orthogonal to $\vec{v_1}$.

Theorem 8.1.3

A symmetric $n \times n$ matrix A has n real eigenvalues if they are counted with their algebraic multiplicities.

Theorem 8.1.4

Orthogonal diagonalization of a symmetric matrix A

a. Find the eigenvalues of A, and find a basis of each eigenspace.

b. Using the Gram-Schmidt process, find an *orthonormal* basis of each eigenspace.

c. Form an orthonormal eigenbasis $\vec{v_1}, \vec{v_2}, \dots, \vec{v_n}$ for A by concatenation the orthonormal bases you found in part b, and let

$$S = \begin{bmatrix} | & | & & | \\ \vec{v_1} & \vec{v_2} & \cdots & \vec{v_n} \\ | & | & & | \end{bmatrix}.$$

S is orthogonal, and $S^{-1}AS$ will be diagonal.

Chapter 9

Information not in the Textbook