MATH 22A: Vector Calculus and Linear Algebra

Final Study Guide

Denny Cao

> **Remark.** Final will focus on material in Chapters 5.1-5.7, Chapters 6.1-6.7, and Chapters 7.1-7.3, but all previous material is used in these chapters. Notes will only be for areas I am not familiar with.

Contents

§2 Matrix Algebra

§2.1 Matrix Operations

Definition 2.1. If A is an $m \times n$ matrix, and if B is an $n \times p$ matrix with columns $\boldsymbol{b_1}, \ldots, \boldsymbol{b_p}$, then the product AB is the $m \times p$ matrix whose columns are $A\boldsymbol{b_1}, \ldots, A\boldsymbol{b_p}$.

- Each column of AB is a linear combination of the columns of A using weights from the corresponding column of B.

Theorem 2.2 (Facts About Matrix Multiplication)
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. $A(AB) = (AB)C$ – Associative Law of Multiplication

b. $A(B + C) = AB + AC$ – Left Distributive Law

c. $(B + C)A = BA + CA$ – Right Distributive Law

d. $\forall r(r(AB) = (rA)B = A(rB))$

e. $I_m A = A = AI_n$ – Identity for Matrix Multiplicatoin

Theorem 2.3 (Facts About Tranpose Matrices)
Let A and B denote matrices whose sizes are appropriate for the following sums and products.

a. $(A^T)^T = A$

b. $(A + B)^T = A^T + B^T$

c. $\forall r((rA)^T = rA^T)$

d. $(AB)^T = B^T A^T$

- The transpose of a product of matrices is the product of their transposes in the **reverse** order.

§2.2 The Inverse of a Matrix

Definition 2.4. An $n \times n$ matrix is **invertible** if $\exists$ an $n \times n$ matrix C s.t.

$$CA = I \quad \text{and} \quad AC = I$$

In this case, C is an **inverse** of A, denoted A^{-1}. C is **uniquely determined** by A because if $\exists$ another matrix B that was an inverse of A, then $B = BI = B(AC) = (BA)C = IC = C$.
A matrix that is **not invertible** is called a **singular matrix** and an invertible matrix is called a **nonsingular matrix**.

Theorem 2.5 (Inverse of a 2×2 Matrix)

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. If $\det A \neq 0 \to A$ is invertible and

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

If $\det A = 0 \to A$ is not invertible.

Theorem 2.6

If A is an invertible $n \times n$ matrix, then $\forall \boldsymbol{b} \in \mathbb{R}^n$, the equation $A\boldsymbol{x} = \boldsymbol{b}$ has the unique solution $\boldsymbol{x} = A^{-1}\boldsymbol{b}$.

Proof. $\boldsymbol{x} = A^{-1}\boldsymbol{b}$ is a solution, as if we substitute, we obtain $A(A^{-1}\boldsymbol{b}) = I\boldsymbol{b} = \boldsymbol{b}$. We show there is a unique solution by showing that, if $\forall \boldsymbol{u}$, if $\boldsymbol{u}$ is a solution, then $\boldsymbol{u} = A^{-1}\boldsymbol{b}$. $A\boldsymbol{u} = \boldsymbol{b}$. We take the left inverse. $A^{-1}A\boldsymbol{u} = A^{-1}\boldsymbol{b}$, and thus $\boldsymbol{u} = A^{-1}\boldsymbol{b}$. □

Theorem 2.7 (Facts About Invertible Matrices)

a. If A is an invertible matrix, then A^{-1} is invertible and

$$(A^{-1})^{-1} = A$$

b. If $A \wedge B$ are $n \times n$ invertible matrices, then so is AB, and the inverse of AB is the product of the inverse of $A \wedge B$ in the reverse order:

$$(AB)^{-1} = B^{-1}A^{-1}$$

c. If A is an invertible matrix, then so is A^T and the inverse of A^T is the transpose of A^{-1} :

$$(A^T)^{-1} = (A^{-1})^T$$

Remark 2.8. If an elementary row operation is performed on an $m \times n$ matrix A, the resulting matrix can be written as EA, where the $m \times m$ matrix E is created by performing the same row operation on I_m.

Remark 2.9. Each elementary matrix E is invertible. The inverse of E is the elementary matrix of the same type that transforms E back into I.

Theorem 2.10

An $n \times n$ matrix A is invertible $\iff$ A is row equivalent to I_n, and in this case, any sequence of elementary row operations that reduces A to I_n also transforms I_n into A^{-1}.

§2.3 Characterizations of Invertible Matrices

Theorem 2.11 (The Invertible Matrix Theorem)

Let A be a square $n \times n$ matrix. Then the following statements are equivalent. That is, for a given A, the statements are either all true or all false.

a. A is an invertible matrix.

b. A is row equivalent to the $n \times n$ identity matrix.

c. A has n pivot positions.

d. The equation $A\boldsymbol{x} = \mathbf{0}$ has only the trivial solution.

e. The columns of A form a linearly independent set.

f. The linear transformation $\boldsymbol{x} \mapsto A\boldsymbol{x}$ is one-to-one.

g. The equation $A\boldsymbol{x} = \boldsymbol{b}$ has at least one solution for each $\boldsymbol{b} \in \mathbb{R}^n$.

h. The columns of A span $\mathbb{R}^n$.

i. The linear transformation $\boldsymbol{x} \mapsto A\boldsymbol{x}$ maps $\mathbb{R}^n$ onto $\mathbb{R}^n$.

j. There is an $n \times n$ matrix C s.t. $CA = I$.

k. There is an $n \times n$ matrix D s.t. $AD = I$.

l. A^T is an invertible matrix.

Definition 2.12. A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is said to be **invertible** if $\exists$ function $S : \mathbb{R}^n \to \mathbb{R}^n$ s.t.

$$S(T(\boldsymbol{x})) = \boldsymbol{x} \quad \forall \boldsymbol{x} \in \mathbb{R}^n \tag{1}$$
$$T(S(\boldsymbol{x})) = \boldsymbol{x} \quad \forall \boldsymbol{x} \in \mathbb{R}^n \tag{2}$$

Theorem 2.13

Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation and let A be the standard matrix for T. Then T is invertible $\iff$ A is an invertible matrix. In that case, the linear transformation S given by $S(\boldsymbol{x}) = A^{-1}\boldsymbol{x}$ is the unique function satisfying equations (1) and (2).

§3 Determinants

§3.1 Introduction to Determinants

Theorem 3.1 (Cofactor Expansion)

The determinant of an $n \times n$ matrix A can be computed by cofactor expansion across any row or down any column. The expansion across the i-th row using cofactors:

$$\det A = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}$$

The cofactor expansion down the j-th columns is

$$\det A = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$$

Where $C_{ij} = (-1)^{i+j} \det A_{ij}$.

Theorem 3.2 (Determinant of Triangular Matrix)

If A is a triangular matrix, then $\det A$ is the product of the entries on the main diagonal of A.

§3.2 Properties of Determinants

Theorem 3.3 (Row Operations)

Let A be a square matrix.

a. If a multiple of one row of A is added to another row to produce a matrix B, then $\det B = \det A$

b. If two rows of A are interchanged to produce B, then $\det B = -\det A$.

c. If one row of A is multiplied by k to produce B, then $\det B = k \cdot \det A$.

Remark 3.4. As it is always possible to reduce to an echelon form U, the determinant of A is as follows:

$$\det A = \begin{cases} (-1)^r \cdot (\text{product of pivots in } U) & \text{when A is invertible} \\ 0 & \text{when A is not invertible} \end{cases}$$

Theorem 3.5

A square matrix A is invertible $\iff \det A \neq 0$.

Corollary 3.6

$\det A = 0$ when the columns of A are linearly dependent. Also, $\det A = 0$ when the **rows** of A are linearly dependent (Rows of A are columns of A^T, and linearly dependent columns of A^T make A^T singular. When A^T is singular, so is A by the Invertible Matrix Theorem.)

Theorem 3.7

If A is an $n \times n$ matrix, then $\det A^T = \det A$.

Proof Idea. Cofactor expansion of $\det A$ along the first **row** equals the cofactor expansion of $\det A^T$ along the first **column**, and then use induction to prove this. $\square$

Remark 3.8. It follows that we can perform operations on the columns of a matrix in a way that is analogous to the row operations we have considered and have the same effects on the determinants as row operations.

Theorem 3.9 (Multiplicative Property)

If A and B are $n \times n$ matrices, then $\det AB = (\det A)(\det B)$.

§3.3 Cramer's Rule, Volume, and Linear Transformations

Theorem 3.10 (Cramer's Rule)

Let A be an invertible $n \times n$ matrix. $\forall \boldsymbol{b} \in \mathbb{R}^n$, the unique solution $\boldsymbol{x}$ of $A\boldsymbol{x} = \boldsymbol{b}$ has entries given by

$$x_i = \frac{\det A_i(\boldsymbol{b})}{\det A}, \quad i = 1, 2, \ldots, n$$

where $A_i(\boldsymbol{b})$ is the matrix obtained from A by replacing column i by the vector $\boldsymbol{b}$.

Example 3.11

Use Cramer's rule to solve the system

$$\begin{aligned} 3x_1 - 2x_2 &= 6 \\ -5x_1 + 4x_2 &= 8 \end{aligned}$$

Solution. View the system as $A\boldsymbol{x} = \boldsymbol{b}$. Then:

$$A = \begin{bmatrix} 3 & -2 \\ -5 & 4 \end{bmatrix}, \quad A_1(\boldsymbol{b}) = \begin{bmatrix} 6 & -2 \\ 8 & 4 \end{bmatrix}, \quad A_2(\boldsymbol{b}) = \begin{bmatrix} 3 & 6 \\ -5 & 8 \end{bmatrix}$$

Since $\det A = 2$, the system has a unique solution. By Cramer's rule,

$$\begin{aligned} x_1 &= \frac{\det A_1(\boldsymbol{b})}{\det A} = \frac{24 + 16}{2} = 20 \\ x_2 &= \frac{\det A_2(\boldsymbol{b})}{\det A} = \frac{24 + 30}{2} = 27 \end{aligned}$$

$\square$

Theorem 3.12 (Inverse Formula)

Let A be an invertible $n \times n$ matrix. Then

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

Proof. This is derived from Cramer's rule, as the j-th column of A^{-1} is a vector $\boldsymbol{x}$ that satisfies $A\boldsymbol{x} = \boldsymbol{e}_j$, where $\boldsymbol{e}_j$ is the j-th column of the identity matrix, and the i-th entry of $\boldsymbol{x}$ is the (i, j)-entry of A^{-1}. By Cramer's rule:

$$\left\{(i, j) - \text{entry of } A^{-1}\right\} = x_i = \frac{\det A_i(\boldsymbol{e}_j)}{\det A}$$

A cofactor expansion down column i of $A_i(\boldsymbol{e_j})$ shows that

$$\det A_i(\boldsymbol{e}_j) = (-1)^{i+j} \det A_{ji} = C_{ji}$$

Thus

$$A^{-1} = \frac{1}{\det A} \begin{bmatrix} C_{11} & C_{21} & \dots & Cn1 \\ C_{12} & C_{22} & \dots & Cn2 \\ \vdots & \vdots & & \vdots \\ C_{1n} & C_{2n} & \dots & C_{nn} \end{bmatrix}$$

The matrix of the cofactors on the right side is called the **adjugate** of A. □

Remark 3.13. The adjugate matrix is the **transpose** of the matrix of cofactors.

Theorem 3.14 (Geometric Interpretation of Determinants)

If A is a 2×2 matrix, the area of the parallelogram determined by the columns of A is $|\det A|$. If A is a 3×3 matrix, the volume of the parallelepiped determined by the columns of A is $|\det A|$.

Theorem 3.15

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation determined by a 2×2 matrix A. If S is a parallelogram in $\mathbb{R}^2$, then

$$\{\text{area of } T(S)\} = |\det A| \cdot \{\text{area of } S\}$$

If T is determined by a 3×3 matrix A, and if S is a parallelepiped in $\mathbb{R}^3$, then

$$\{\text{volume of } T(S)\} = |\det A| \cdot \{\text{volume of } S\}$$

§4 Vector Spaces

§4.1 Vector Spaces and Subspaces

Definition 4.1. To show a space is a **vector space**, must show that:

- $\forall a, b \in V(a+b \in V)$. (Closed under vector addition)
- $\forall c, \forall a \in V(ca \in V)$. (Closed under scalar multiplication)
- $\exists 0 \in V$.

Definition 4.2. A **subspace** of a vector V is a subset H of V that has three properties:

- Zero vector of V is in H
- H is closed under vector addition.
- H is closed under scalar multiplication.

Theorem 4.3

$\boldsymbol{v_1}, \dots, \boldsymbol{v_p} \subseteq V \to \text{Span}\{\boldsymbol{v_1}, \dots, \boldsymbol{v_p}\}$ is a subspace of V.

Remark 4.4. We call $\text{Span}\{\boldsymbol{v_1}, \dots, \boldsymbol{v_p}\}$ **the subspace spanned** (or **generated**) by $\{\boldsymbol{v_1}, \dots, \boldsymbol{v_p}\}$. Given any subspace H of V, a **spanning** (or **generating**) **set** for H is a set $\{\boldsymbol{v_1}, \dots, \boldsymbol{v_p}\}$ in H such that $H = \text{Span}\{\boldsymbol{v_1}, \dots, \boldsymbol{v_p}\}$.

§4.2 Null Spaces, Column Spaces, and Linear Transformations

Definition 4.5. The **null space** of an $m \times n$ matrix A, written as $\text{Nul}A$, is the set of all solutions of the homogeneous equation $A\boldsymbol{x} = \boldsymbol{0}$. In set notation:

$$\text{Nul}A = \{\boldsymbol{x} \mid \boldsymbol{x} \in \mathbb{R}^n \wedge A\boldsymbol{x} = \boldsymbol{0}\}$$

Remark 4.6. A better description of $\text{Nul}A$ is the set of all $\boldsymbol{x} \in \mathbb{R}^n$ that are mapped into the zero vector of $\mathbb{R}^m$ via the linear transformation $\boldsymbol{x} \mapsto A\boldsymbol{x}$.

Theorem 4.7

The null space of an $m \times n$ matrix A is a subspace of $\mathbb{R}^n$. Equivalently, the set of all solutions to a system $A\boldsymbol{x} = \boldsymbol{0}$ of m homogeneous linear equations in n unknowns is a subspace of $\mathbb{R}^n$.

Definition 4.8. The **column space** of an $m \times n$ matrix A, written as $\text{Col}A$, is the set of all linear combinations of the columns of A. If $A = \begin{bmatrix} \boldsymbol{a_1} & \cdots & \boldsymbol{a_n} \end{bmatrix}$, then

$$\text{Col}A = \text{Span}\{\boldsymbol{a_1}, \dots, \boldsymbol{a_n}\}$$

We can write this as

$$\text{Col}A = \{\boldsymbol{b} \mid \exists \boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{b} = A\boldsymbol{x}\}$$

Theorem 4.9

The column space of an $m \times n$ matrix A is a subspace of $\mathbb{R}^m$.

Proof. Since $\text{Span}\{\boldsymbol{a_1}, \dots, \boldsymbol{a_n}\}$ is a subspace of $\mathbb{R}^m$ by Theorem 4.3 (the column vectors of A are in $\mathbb{R}^m$), it follows that $\text{Col}A$ is a subspace of $\mathbb{R}^m$. $\square$

Remark 4.10. The column space of an $m \times n$ matrix A is all of $\mathbb{R}^m \iff$ The equation $A\boldsymbol{x} = \boldsymbol{b}$ has a solution $\forall \boldsymbol{b} \in \mathbb{R}^m$

Definition 4.11. The **kernel** (or **null space**) of a **linear transformation** T from a vector space V into a vector space W is the set of all $\boldsymbol{u} \in V$ s.t. $T(\boldsymbol{u}) = \boldsymbol{0}$ (the zero vector in W). The **range** of T is the set of all vectors in W of the form $T(\boldsymbol{x})$ for some $\boldsymbol{x} \in V$.

Remark 4.12. If T is a matrix transformation $T(\boldsymbol{x}) = A\boldsymbol{x}$, then the kernel and the range of T is just the null space and the column space of A.

§4.3 Linearly Independent Sets; Bases

Theorem 4.13

An indexed set $\{\boldsymbol{v_1}, \dots, \boldsymbol{v_p}\}$ of two or more vectors, with $\boldsymbol{v_1} \neq 0$, is linearly dependent if and only if some $\boldsymbol{v_j}$ (with $j > 1$) is a linear combination of the preceding vectors, $\boldsymbol{v_1}, \dots, \boldsymbol{v_{j-1}}$.

Definition 4.14. Let H be a subspace of a vector space V. An indexed set of vectors $\mathcal{B} = \{\boldsymbol{b_1}, \dots, \boldsymbol{b_p}\}$ in V is a **basis** for H if

(i) $\mathcal{B}$ is a linearly independent set, and

(ii) the subspace spanned by $\mathcal{B}$ coincides with H:

$$H = \text{Span}\{\boldsymbol{b_1}, \dots, \boldsymbol{b_p}\}$$

Theorem 4.15 (The Spanning Set Theorem)

Let $S = \{\boldsymbol{v_1}, \dots, \boldsymbol{v_p}\}$ be a set in V, and let $H = \text{Span}\{\boldsymbol{v_1}, \dots, \boldsymbol{v_p}\}$.

a. If one of the vectors in S, $\boldsymbol{v_k}$, is a linear combination of the remaining vectors in S, then the set formed from S by removing $\boldsymbol{v_k}$ still spans H.

b. If $H \neq \{\boldsymbol{0}\}$, some subset of S is a basis for H.

Theorem 4.16

The pivot columns of a matrix A form a basis for $\text{Col}A$.

Remark 4.17. Be careful to use the pivot columns of A itself for the basis of $\text{Col}A$!

- We can view basis as a spanning set that is as small as possible, or
- a linearly independent set that is as large as possible

§4.4 Coordinate Systems

Theorem 4.18 (The Unique Representation Theorem)

Let $\mathcal{B} = \{\boldsymbol{b_1}, \ldots, \boldsymbol{b_n}\}$ be a basis for a vector space V. Then $\forall \boldsymbol{x} \in V \exists$ a unique set of scalars $c_1, \ldots, c_n$ s.t.

$$\boldsymbol{x} = c_1\boldsymbol{b_1} + \cdots + c_n\boldsymbol{b_n}$$

Definition 4.19. Suppose $\mathcal{B} = \{\boldsymbol{b_1}, \ldots, \boldsymbol{b_n}\}$ is a basis for V and $\boldsymbol{x} \in V$. The **coordinates of $\boldsymbol{x}$ relative to the basis $\mathcal{B}$** (or the **$\mathcal{B}$-coordinates of $\boldsymbol{x}$**) are the weights $c_1, \ldots, c_n$ s.t. $\boldsymbol{x} = c_1\boldsymbol{b_1} + \cdots + c_n\boldsymbol{b_n}$.

- If $c_1, \ldots, c_n$ are the $\mathcal{B}$-coordinates of $\boldsymbol{x}$, then the vector in $\mathbb{R}^n$

$$[\boldsymbol{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

  is the **coordinate vector of $\boldsymbol{x}$ (relative to $\mathcal{B}$)**, or the **$\mathcal{B}$-coordinate vector of $\boldsymbol{x}$**. The mapping $\boldsymbol{x} \mapsto [\boldsymbol{x}]_{\mathcal{B}}$ is the **coordinate mapping (determined by $\mathcal{B}$)**.

Remark 4.20. Let $P_{\mathcal{B}} = \begin{bmatrix} \boldsymbol{b_1} & \boldsymbol{b_2} & \ldots & \boldsymbol{b_n} \end{bmatrix}$. The vector equation $\boldsymbol{x} = c_1\boldsymbol{b_1} + c_2\boldsymbol{b_2} + \cdots + c_n\boldsymbol{b_n}$ is equivalent to

$$\boldsymbol{x} = P_{\mathcal{B}}[\boldsymbol{x}]_{\mathcal{B}}$$

$P_{\mathcal{B}}$ is the **change-of-coordinates matrix** from $\mathcal{B}$ to the standard basis of $\mathbb{R}^n$. Left-multiplication by $P_{\mathcal{B}}$ transforms the coordinate vector $[\boldsymbol{x}]_{\mathcal{B}}$ into $\boldsymbol{x}$.

Since the columns of $P_{\mathcal{B}}$ form a basis for $\mathbb{R}^n$, $P_{\mathcal{B}}$ is invertible (by the Invertible Matrix Theorem). We can thus multiply by $P_{\mathcal{B}}^{-1}$ on the left to convert $\boldsymbol{x}$ into its $\mathcal{B}$-coordinate vector:

$$P_{\mathcal{B}}^{-1}\boldsymbol{x} = [\boldsymbol{x}]_{\mathcal{B}}$$

The correspondence $\boldsymbol{x} \mapsto [\boldsymbol{x}]_{\mathcal{B}}$, produced here by $P_{\mathcal{B}}^{-1}$, is the coordinate mapping mentioned earlier. Since $P_{\mathcal{B}}^{-1}$ is an invertible matrix, the coordinate mapping ins a one-to-one linear transformation from $\mathbb{R}^n$ onto $\mathbb{R}^n$, by the Invertible Matrix Theorem.

Theorem 4.21

Let $\mathcal{B} = \{\boldsymbol{b_1}, \ldots, \boldsymbol{b_n}\}$ be a basis for a vector space V. Then the coordinate mapping $\boldsymbol{x} \mapsto [\boldsymbol{x}]_{\mathcal{B}}$ is a one-to-one transformation from V onto $\mathbb{R}^n$.

§4.5 Dimension of a Vector Space

Theorem 4.22

If a vector space V has a basis $\mathcal{B} = \{\boldsymbol{b_1}, \ldots, \boldsymbol{b_n}\}$, then any set in V containing more than n vectors must be linearly dependent.

Theorem 4.23

If a vector space V has a basis of n vectors, then every basis of V must consist of exactly n vectors.

Definition 4.24. If V is spanned by a finite set, then V is said to be **finite-dimensional**, and the **dimension** of V, written as $\dim V$, is the number of vectors in a basis for V. The dimension of the zero vector space $\{\mathbf{0}\}$ is defined to be zero. If V is not spanned by a finite set, then V is said to be **infinite-dimensional**.

Theorem 4.25

Let H be a subspace of a finite-dimensional vector space V. Any linearly independent set in H can be expanded, if necessary, to a basis for H. Also, H is finite-dimensional and

$$\dim H \leq \dim V$$

Theorem 4.26 (The Basis Theorem)

Let V be a p-dimensional vector space, $p \geq 1$. Any linear independent set of exactly p elements in V is automatically a basis for V. Any set of exactly p elements that spans V is automatically a basis for V.

> **Remark 4.27.** The dimension of $\mathrm{Nul}A$ is the number of free variables in the equation $A\boldsymbol{x} = \mathbf{0}$, and the dimension of $\mathrm{Col}A$ is the number of pivot columns in A. Thus, $\dim \mathrm{Nul}A + \dim \mathrm{Col}A = \#$ of free variables $+ \#$ of pivot columns $= \#$ of columns.

§4.6 Rank

Definition 4.28. If A is an $m \times n$ matrix, each row of A has n entries and thus can be identified with a vector in $\mathbb{R}^n$. The set of all linear combinations oft he row vectors is called the **row space** of A and is denoted by $\mathrm{Row}A$. Each row has n entries, so $\mathrm{Row}A$ is a subspace of $\mathbb{R}^n$. Since the rows of A are identified with the columns of A^T, we could also write that $\mathrm{Col}A^T = \mathrm{Row}A$.

Theorem 4.29

If two matrices A and B are row equivalent, then their row spaces are the same. If B is in echelon form, the nonzero rows of B form a basis for the row space of A as well as for that of B.

Definition 4.30. The **rank** of A is the dimension of the column space of A.

Theorem 4.31 (The Rank Theorem)

The dimensions of the column space and the row space of an $m \times n$ matrix A are equal. This common dimension, the rank of A, equals the number of pivot positions in A and satisfies the equation

$$\operatorname{rank} A + \dim \operatorname{Nul} A = n$$

Theorem 4.32 (The Invertible Matrix Theorem (continued))

Let A be an $n \times n$ matrix. Then the following statements are each equivalent to the statement that A is an invertible matrix.

m. The columns of A form a basis of $\mathbb{R}^n$.

n. $\operatorname{Col} A = \mathbb{R}^n$.

o. $\dim \operatorname{Col} A = n$.

p. $\operatorname{rank} A = n$

q. $\operatorname{Nul} A = \{\mathbf{0}\}$

r. $\dim \operatorname{Nul} A = 0$

§4.7 Change of Basis

Theorem 4.33

Let $\mathcal{B} = \{\boldsymbol{b_1}, \ldots, \boldsymbol{b_n}\}$ and $\mathcal{C} = \{\boldsymbol{c_1}, \ldots, \boldsymbol{c_n}\}$ be bases of a vector space V. Then there is a unique $n \times n$ matrix $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$ such that

$$[\boldsymbol{x}]_{\mathcal{C}} = \underset{\mathcal{C} \leftarrow \mathcal{B}}{P} [\boldsymbol{x}]_{\mathcal{B}}$$

The columns of $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P} [\boldsymbol{x}]_{\mathcal{B}}$ are the $\mathcal{C}$-coordinate vectors of the vectors in the basis $\mathcal{B}$. That is,

$$\underset{\mathcal{C} \leftarrow \mathcal{B}}{P} [\boldsymbol{x}]_{\mathcal{B}} = \begin{bmatrix} [\boldsymbol{b_1}]_{\mathcal{C}} & [\boldsymbol{b_2}]_{\mathcal{C}} & \cdots & [\boldsymbol{b_n}]_{\mathcal{C}} \end{bmatrix}$$

Definition 4.34. $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$ is called the **change-of-coordinates matrix from $\mathcal{B}$ to $\mathcal{C}$**. Multiplication by this matrix converts $\mathcal{B}$-coordinates to $\mathcal{C}$-coordinates.

Remark 4.35. The columns of the change-of-coordinates matrix are linearly independent because they are coordinate vectors of the linearly independent set $\mathcal{B}$. As the change-of-coordinates matrix is square, it follows that i must be invertible, by the Invertible Matrix Theorem. Left-multiplying both sides of the equation by the inverse yields

$$(\underset{\mathcal{C} \leftarrow \mathcal{B}}{P})^{-1} [\boldsymbol{x}]_{\mathcal{C}} = [\boldsymbol{x}]_{\mathcal{B}}$$

> Thus, $(\underset{\mathcal{C}\leftarrow\mathcal{B}}{P})^{-1}$ is the matrix that converts $\mathcal{C}$-coordinates into $\mathcal{B}$-coordinates. That is,
>
> $$(\underset{\mathcal{C}\leftarrow\mathcal{B}}{P})^{-1} = \underset{\mathcal{B}\leftarrow\mathcal{C}}{P}$$

We obtain the change-of-coordinate matrix by the following:

$$\left[\boldsymbol{c_1} \quad \boldsymbol{c_2} \mid \boldsymbol{b_1} \quad \boldsymbol{b_2}\right] \sim \left[I \mid \underset{\mathcal{C}\leftarrow\mathcal{B}}{P}\right]$$

§5 Eigenvalues and Eigenvectors

Definition 5.1. An **eigenvector** of an $n \times n$ matrix A is a nonzero vector $\boldsymbol{x}$ such that $A\boldsymbol{x} = \lambda\boldsymbol{x}$ for some scalar λ. A scalar λ is called an **eigenvalue** of A if there is a nontrivial solution $\boldsymbol{x}$ of $A\boldsymbol{x} = \lambda\boldsymbol{x}$; such an $\boldsymbol{x}$ is called an *eigenvector corresponding to* λ.

§5.1 Eigenvectors and Eigenvalues

Theorem 5.2

The eigenvalues of a triangular matrix are the entries on its main diagonal.

Definition 5.3. The set of *all* solutions of $(A - \lambda I)\boldsymbol{x} = \boldsymbol{0}$ is the null space of the matrix $A - \lambda I$. Thus, this set is a *subspace* of $\mathbb{R}^n$ and is called the **eigenspace** of A corresponding to λ. The eigenspace consists of the zero vector nd all the eigenvectors corresponding to λ.

Theorem 5.4

If $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r$ are eigenvectors that correspond to distinct eigenvalues $\lambda_1, \ldots, \lambda_r$ of an $n \times n$ matrix A, then the set $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r\}$ is linearly independent.

§5.2 The Characteristic Equation

Definition 5.5. Finding the eigenvalues of A is, by the Inverse Matrix Theorem, equivalent to finding all λ such that $A - \lambda I$ is *not* invertible. This matrix fails to be invertible when the determinant is zero, and thus the eigenvalues of A are the solutions of the **characteristic equation**

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

Definition 5.6. If A is an $n \times n$ matrix, then $\det(A - \lambda I)$ is a polynomial of degree n, the **characteristic polynomial**.

Definition 5.7. The **algebraic multiplicity** of an eigenvalue λ is its multiplicity as a root of the characteristic equation.

Theorem 5.8 (The Iverible Matrix Theorem (continued))

Let A be an $n \times n$ matrix. Then A is invertible if and only if:

s. The number 0 is *not* an eigenvalue of A.

t. The determinant of A is *not* zero.

Theorem 5.9 (Properties of Determinants)

Let A and B be $n \times n$ matrices.

a. A is invertible if and only if $\det A \neq 0$.

b. $\det AB = (\det A)(\det B)$.

c. $\det A^T = \det A$.

d. If A is triangular, then $\det A$ is the product of the entries on the main diagonal of A.

e. A row replacement operation on A does not change determinant. A row interchange changes the sign of determinant. A row scaling also scales determinant by same scale factor.

Definition 5.10. A is **similar to B** if there is an invertible matrix P such that $P^{-1}AP = B$, or equivalently $A = PBP^{-1}$. Writing Q for P^{-1}, we have $Q^{-1}BQ = A$, and thus similarity is symmetric. Changing A to $P^{-1}AP$ is called a **similarity transformation**.

Theorem 5.11

If $n \times n$ matrices A and B are similar, then they have the same characteristic polynomial and hence the same eigenvalues (with the same multiplicity).

Proof. If $B = P^{-1}AP$, then:

$$B - \lambda I = P^{-1}AP - \lambda P^{-1}P = P^{-1}(AP - \lambda P) = P^{-1}(A - \lambda I)P$$

From the multiplicative property, we compute

$$\begin{aligned}\det(B - \lambda I) &= \det(P^{-1}(A - \lambda I)P) \\ &= \det(P^{-1})\det(A - \lambda I)\det(P)\end{aligned}$$

Since $\det(P^{-1})\det(P) = \det(P^{-1}P) = \det(I) = 1$, $\det(B - \lambda I) = \det(A - \lambda I)$, as desired. $\square$

Remark 5.12.

- Same eigenvalues does not imply similarity! Must have same eigenvalues corresponding to the same eigenvectors.
- Similarity is not the same as row equivalence, as row changes usually change eigenvalues.

Example 5.13 (Dynamical Systems)

Let $A = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix}$. Analyze the long-term behavior of the dynamical system defined by $\boldsymbol{x}_{k+1} = A\boldsymbol{x}_k (k = 0, 1, 2, \dots)$, with $\boldsymbol{x}_0 = \begin{bmatrix} .6 \\ .4 \end{bmatrix}$.

Solution. First step is to find eigenvalues of A and a basis for each eigenspace. Characteristic equation is:

$$0 = \lambda^2 - 1.92\lambda + .92$$

thus, $\lambda = 1, .92$.

The corresponding eigenvectors for $\lambda = 1$ and $\lambda = .92$ are found by $A - \lambda I = 0$ and solving the homogeneous system. The eigenvectors are:

$$\boldsymbol{v}_1 = \begin{bmatrix} 3 \\ 5 \end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Every multiple of each also solves the corresponding homogeneous system.

We then write $\boldsymbol{x}_0$ in terms of $\boldsymbol{v}_1, \boldsymbol{v}_2$, which can be done as $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ is a basis for $\mathbb{R}^2$ (They are linearly independent and there are two vectors). Thus, there exists weights c_1, c_2 such that

$$\boldsymbol{x}_0 = c_1\boldsymbol{v_1} + c_2\boldsymbol{v}_2 = \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

We can rewrite:

$$\begin{aligned} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} &= \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 \end{bmatrix}^{-1} \boldsymbol{x}_0 = \begin{bmatrix} 3 & 1 \\ 5 & -1 \end{bmatrix}^{-1} \begin{bmatrix} .6 \\ .4 \end{bmatrix} \\ &= -\frac{1}{8} \begin{bmatrix} -1 & -1 \\ -5 & 3 \end{bmatrix} \begin{bmatrix} .6 \\ .4 \end{bmatrix} = \begin{bmatrix} .125 \\ .225 \end{bmatrix} \end{aligned}$$

Note that we could've solved by finding the RREF as well.

As $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are eigenvectors of A, with $A\boldsymbol{v}_1 = \boldsymbol{v}_1$ and $A\boldsymbol{v}_2 = .92\boldsymbol{v}_2$, we can easily compute each $\boldsymbol{x}_k$:

$$\begin{aligned} \boldsymbol{x}_1 = A\boldsymbol{x}_0 &= c_1 A\boldsymbol{v}_1 + c_2 A\boldsymbol{v}_2 \\ &= c_1\boldsymbol{v}_1 + c_2(.92)\boldsymbol{v}_2 \\ \boldsymbol{x}_2 = A\boldsymbol{x}_1 &= c_1 A\boldsymbol{v}_1 + c_2(.92)A\boldsymbol{v}_2 \\ &= c_1\boldsymbol{v}_1 + c_2(.92)^2\boldsymbol{v}_2 \end{aligned}$$

In general,

$$\boldsymbol{x}_k = c_1\boldsymbol{v}_1 + c_2(.92)^k\boldsymbol{v}_2 \ (k = 0, 1, 2, \dots)$$

The rest is trivial. Substitute for c_1, c_2, and then find $\lim\limits_{k\to\infty}$. □

We form a general solution:

$$\boldsymbol{x}_k = c_1\lambda_1^k\boldsymbol{v}_1 + \cdots + c_n\lambda_n^k\boldsymbol{v}_n$$

§5.3 Diagonalization

> **Remark 5.14.** We use diagonalization to find powers A^k quickly for large values of k.

Definition 5.15. A square matrix A is said to be **diagonalizable** if A is similar to a diagonal matrix, that is, if $A = PDP^{-1}$ for some invertible matrix P and some diagonal matrix D.

Theorem 5.16

An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.

In fact, $A = PDP^{-1}$ with D a diagonal matrix, if and only if the columns of P are n linearly independent eigenvectors of A. In this case, the diagonal etnries of D are eigenvalues of A that correspond, respectively, to the eigenvectors in P.

Remark 5.17. In other words, A is diagonalizable if and only if there are enough eigenvectors to form a basis of $\mathbb{R}^n$. We call such a basis an **eigenvector basis** of $\mathbb{R}^n$.

Remark 5.18. When checking if P and D work, to avoid computing P^{-1}, it is sufficient to verify that $AP = PD$, as this is equivalent to $A = PDP^{-1}$ when P is invertible.

Theorem 5.19

An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

Proof. With n distinct eigenvalues, the set of eigenvectors $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is a linearly independent set. $\square$

Theorem 5.20

Let A be an $n \times n$ matrix whose distinct eigenvalues are $\lambda_1, \ldots, \lambda_p$.

a. For $1 \leq k \leq p$, the dimension of the eigenspace for λ_k is less than or equal to the multiplicity of the eigenvalue λ_k.

b. The matrix A is diagonalizable if and only if the sum of the dimensions of the eigenspace equals n, and this happens if and only if:

   i. the characteristic polynomial factors completely into linear factors and

   ii. the dimension of the eigenspace for each λ_k equals the multiplicity of λ_k.

c. If A is diagonalizable and $\mathcal{B}_k$ is a basis for the eigenspace corresponding to λ_k for each k, then the total collection of vectors in the sets $\mathcal{B}_1, \ldots, \mathcal{B}_p$ forms an eigenvector basis for $\mathbb{R}^n$.

§5.4 Eigenvectors and Linear Transformations

Remark 5.21. Recall that any linear transformation T from $\mathbb{R}^n \to \mathbb{R}^m$ can be implemented via left-multiplication by a matrix A, called the standard matrix of T. We now attempt to understand $A = PDP^{-1}$ in terms of linear transformations.

- Let V be an n-dimensional vector space, let W be an m-dimensional vector space, let T be any linear transformation from V to W. To associate a matrix with T, choose (ordered) bases $\mathcal{B}$ and $\mathcal{C}$ for V and W respectively.

Given any $\boldsymbol{x}$ in V, the coordinate vector $[\boldsymbol{x}]_{\mathcal{B}}$ is in $\mathbb{R}^n$ and the coordinate vector of its image, $[T(\boldsymbol{x})]_{\mathcal{C}}$, is in $\mathbb{R}^m$.

- As $\boldsymbol{x} = r_1\boldsymbol{b}_1 + \cdots + r_n\boldsymbol{b}_n$, $T(x) = T(r_1\boldsymbol{b}_1 + \cdots + r_n\boldsymbol{b}_n) = r_1T(\boldsymbol{b}_1) + \cdots + r_nT(\boldsymbol{b}_n)$ since T is linear. Then, $[T(\boldsymbol{x})]_{\mathcal{C}} = r_1[T(\boldsymbol{b}_1)]_{\mathcal{C}} + \cdots + r_n[T(\boldsymbol{b}_n)]_{\mathcal{C}}$.
- We can rewrite
$$[T(\boldsymbol{x})]_{\mathcal{C}} = M[\boldsymbol{x}]_{\mathcal{B}}$$
where
$$M = \begin{bmatrix}[T(\boldsymbol{b}_1)]_{\mathcal{C}} & \cdots & T(\boldsymbol{b}_n)]_{\mathcal{C}}\end{bmatrix}$$
The matrix M is a matrix representation of T called the **matrix for $\boldsymbol{T}$ relative to the bases $\boldsymbol{\mathcal{B}}$ and $\boldsymbol{\mathcal{C}}$**.
- We just left multiply by M.

Definition 5.22. When V and W have same base, M is the **basis for $\boldsymbol{T}$ relative to $\boldsymbol{\mathcal{B}}$**, or the **$\boldsymbol{\mathcal{B}}$-matrix for $\boldsymbol{T}$**, denoted $[T]_{\mathcal{B}}$.

Theorem 5.23

Suppose $A = PDP^{-1}$ where D is a diagonal $n \times n$ matrix. If $\mathcal{B}$ is the basis for $\mathbb{R}^n$ formed from the columns of P, then D is the $\mathcal{B}$-matrix for the transformation $\boldsymbol{x} \mapsto A\boldsymbol{x}$.

§5.5 Complex Eigenvalues

Theorem 5.24

Let A be a 2×2 matrix with a complex eigenvalue $\lambda = a - bi$ $(b \neq 0)$ and an associated eigenvector $\boldsymbol{v} \in \mathbb{C}^2$. Then:

$$A = PCP^{-1}, P = \begin{bmatrix}\text{Re } \boldsymbol{v} & \text{Im } \boldsymbol{v}\end{bmatrix}, C = \begin{bmatrix} a & -b \\ b & d \end{bmatrix}$$

§5.6 Discrete Dynamical Systems