Math 18 Notes

Khang Tong

Contents

Preface

This set of notes is a synthesis of my notes from when I took MATH 18 at UCSD and notes I prepared for discussion sections when I TA'd for the course. They are only meant to supplement lecture and in its current state make no claims to being comprehensive. Topics covered may also vary slightly between quarters.

I'd recommend before diving into the review portion of the notes to take a look at the *Tips for Studying* chapter in the appendix to plan your study.

Happy reviewing and good luck on the final!

Khang Tong
27 February 2023

Note: For convenience, all theorems have been extracted to a handout in the appendix and are referred to throughout the set of notes without duplication (unless paraphrasing them facilitates the flow of the notes).

i.e. means "that is" and is the Latin abbreviation for "id est".

e.g. means "for example" and is the Latin abbreviation for "exempli gratia".

List of Figures

Chapter 1

Linear Equations

1.1 Our First Row Reduction

Linear Equation: An equation that can be written in the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b \quad \text{where } a_i\text{'s are constants}$$

For example:

- $4x_1 + 3x_2 = \sqrt{5}x_3 + 10$ linear
- $4x_1x_2 = \sin(x_1)$ not linear

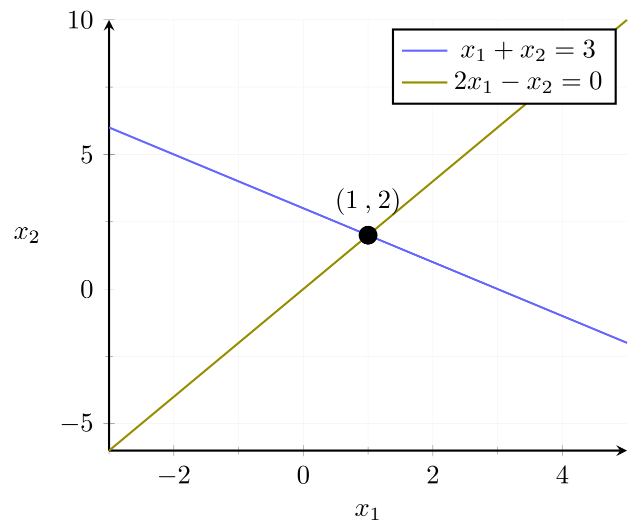
Essentially, to be linear in x_i, the only operation you can apply to x_i is multiplication by a constant.

Linear System: A collection of one or more linear equations.

Solution Set: The solution set of a linear system in $x_1, \ldots, x_n$ is the collection of all points $(s_1, \ldots, s_n)$ that satisfy all the equations.

For example:

$$\begin{aligned} x_1 + x_2 &= 3 \\ 2x_1 - x_2 &= 0 \end{aligned}$$



The intersection $(1, 2)$ comprises the *solution set* of the system.

Question: What can a solution set of a system of linear equations look like?

Answer: A point, ∅ (the empty set), or infinitely many solutions (a line, plane, etc).

An **inconsistent system** has no solution whereas a **consistent system** has 1 or infinitely many solutions.

A shorthand way of solving systems of linear equations is to write the system as an **augmented matrix**. We can rewrite the system from before as:

$$\begin{array}{ccc} \text{coef of } x_1 & \text{coef of } x_2 & \text{constant} \end{array}$$
$$\left[\begin{array}{cc|c} 1 & 1 & 3 \\ 2 & -1 & 0 \end{array}\right]$$

Row Echlon Form (REF):

1. All leading entries (the first nonzero term in a row, aka, a **pivot position**) are always to the right of the leading entry of the row above it.
2. All rows consisting of only zeroes at the bottom.
3. (1) and (2) imply that all entries below the leading entries are 0.

Reduced Row Echlon Form (RREF):

1. Is in REF.
2. The leading entry in each nonzero row is 1 (called *leading* 1).
3. Each column containing a leading 1 has zeros in all its other entries.

Ex: RREF? Pivot Columns?

(A) "Pivot positions"

$$\begin{bmatrix} 1 & 2 & 0 & 0 & 6 \\ 0 & 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

RREF, pivot cols: 1, 3, 4

(B) Pivot cols: 1, 3

$$\begin{bmatrix} 1 & -7 & 10 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Should be 0 for RREF

Not RREF, but is REF

(C)

$$\begin{bmatrix} 1 & 0 & 5 \\ 0 & 2 & 3 \\ 0 & 0 & 1 \end{bmatrix}$$

Should be 0 for RREF

Should be 1 for RREF

REF, pivot cols: 1, 2, 3

(D)

$$\begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 3 & 5 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

not REF, we have 2 leading entries in the same col!

Example 1.1

Write the system of questions as an augmented matrix and find its RREF.

$$\begin{aligned} x_2 - x_3 &= 4 + 2x_1 \\ x_1 + 2x_2 + 3x_3 &= 13 \\ 3x_1 + x_3 + 1 &= 0 \end{aligned}$$

Game plan:

1. Start at the left-most leading entry.
2. Either make the leading entry a 1 or switch with a row that already has a 1.
3. Make this the top row if not already.
4. Annihilate all entries below the 1, turning them into 0, via row operations. This is now a pivot column.
5. Ignoring the rows you already worked on, repeat for rows $2, 3, \ldots$ moving rightwards.
6. Once you arrive at the last pivot column, not work on making each entry above the 1 zero.
7. Repeat moving leftward for each pivot column.
8. Now that the matrix is in RREF, read off the solutions and celebrate with cake.

Solution:
The augmented matrix is

$$\left[\begin{array}{ccc|c} -2 & 1 & -1 & 4 \\ 1 & 2 & 3 & 13 \\ 3 & 0 & 1 & -1 \end{array}\right].$$

$$\begin{aligned} R_1 \leftrightarrow R_2 \left[\begin{array}{ccc|c} 1 & 2 & 3 & 13 \\ -2 & 1 & -1 & 4 \\ 3 & 0 & 1 & -1 \end{array}\right] &\xrightarrow[R_3 \mapsto R_3 + (-3R_1)]{R_2 \mapsto R_2 + 2R_1} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 13 \\ 0 & 5 & 5 & 30 \\ 0 & -6 & -8 & -40 \end{array}\right] \\ &\xrightarrow{R_2 \mapsto \frac{1}{5} R_2} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 13 \\ 0 & 1 & 1 & 6 \\ 0 & -6 & -8 & -40 \end{array}\right] \\ &\xrightarrow{R_3 \mapsto R_3 + 6R_2} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 13 \\ 0 & 1 & 1 & 6 \\ 0 & 0 & -2 & -4 \end{array}\right] \\ &\xrightarrow{R_3 \mapsto -\frac{1}{2} R_3} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 13 \\ 0 & 1 & 1 & 6 \\ 0 & 0 & 1 & 2 \end{array}\right] \\ &\xrightarrow[R_2 \mapsto R_2 - R_3]{R_1 \mapsto R_1 - 3R_3} \left[\begin{array}{ccc|c} 1 & 2 & 0 & 7 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 2 \end{array}\right] \\ &\xrightarrow{R_1 \mapsto R_1 - 2R_2} \left[\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 2 \end{array}\right] \quad \text{RREF!} \end{aligned}$$

The solution is the point $(-1, 4, 2)$. ■

Basic variables are variables associated with pivot columns and **free variables** are variables associated with non-pivot columns.

Example 1.2

Find the solution of a system of linear equations whose augmented matrix has given RREF:

$$\left[\begin{array}{ccccc|c} 1 & 0 & 0 & 6 & 0 & 5 \\ 0 & 0 & 1 & 3 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{array}\right]$$

Solution:

$$\left[\begin{array}{ccccc|c} 1 & 0 & 0 & 6 & 0 & 5 \\ 0 & 0 & 1 & 3 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{array}\right] \longleftrightarrow \begin{cases} x_1 + 6x_4 & = 5 \\ x_3 + 3x_4 & = -1 \\ x_5 & = 2 \end{cases} \implies \begin{cases} x_1 & = 5 - 6x_4 \\ x_3 & = -1 - 3x_4 \\ x_5 & = 2 \\ x_2 & \text{is free} \\ x_4 & \text{is free} \end{cases}$$

This is known as the **parametric description** of the solution set (i.e., basic variables written in terms of free variables) ■

To wrap up this section, we define the **identity matrix** as the $n \times n$ matrix with ones along the diagonal and zeros elsewhere $\left(\text{i.e., } I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \right)$.

1.2 The Ballad of Inspector Vector

A **vector** is an element of a *vector space* (more on that later). For now, we can think of them as a matrix with one column.

For example: $\boldsymbol{u} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}, \quad \boldsymbol{v} = \begin{bmatrix} 4 \\ 2 \\ 0 \end{bmatrix}$.

Notation:

- $\in$ – "element of", "in".
- $\notin$ – "not an element of", "not in".
- $\mathbb{R}^n$ – the set of all vectors with n entries.
- $:$ – "such that".

- $\mathbb{R}^n = \left\{ \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} : x_1, \ldots, x_n \in \mathbb{R} \right\}$ – "the set of all vectors of the form $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ such that $x_1, \ldots, x_n$ are real numbers."
- $\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} \notin \mathbb{R}^3, \quad \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \in \mathbb{R}^3.$

The **linear combination** of the set of vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ is the vector $\boldsymbol{b} = c_1\boldsymbol{v}_1 + \cdots + c_n\boldsymbol{v}_n$ with weights $c_1, \ldots, c_n$ where the c_i's are real numbers.

We know four different ways of representing a linear system and you should be comfortable translating between them:

$$\begin{cases} x_1 + 2x_2 &= 3 \\ 2x_1 + 3x_2 &= 4 \end{cases} \longleftrightarrow \quad x_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \quad \longleftrightarrow \quad \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \quad \longleftrightarrow \quad \left[\begin{array}{cc|c} 1 & 2 & 3 \\ 2 & 3 & 4 \end{array}\right]$$

"system" "vector equation" "matrix equation" "augmented matrix"

The **span** of a set of vectors, $\boldsymbol{v}_i,\ i = 1, \ldots, n$, is the set of all linear combinations of those vectors.

For example: $\text{span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\} = \mathbb{R}^3$

To determine whether a vector $\boldsymbol{b}$ is in the span of a set of vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$, build the augmented matrix $[\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n | \boldsymbol{b}]$ and determine if the system has a solution.

Example 1.3

Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$, $\boldsymbol{b} = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$. Then $\left[\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right] \longleftrightarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$.

Since the system is *inconsistent*, there does *not* exist a *linear combination* of the columns of A that gives you $\boldsymbol{b}$. I.e., $\boldsymbol{b}$ is not in the span of the columns of A.

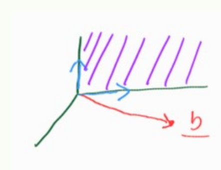


Figure 1.1: $\boldsymbol{b}$ is outside the span of the columns of A, shown in purple.

Solving for the system $A\boldsymbol{x} = \boldsymbol{b}$ is of particular interest in linear algebra, which we will spend some time on.

Notation:

- A is $m \times n$ ($A_{m \times n}$).
- $A = \big[\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n\big]$, $\boldsymbol{a}_i \in \mathbb{R}^m$.
- $\boldsymbol{x} \in \mathbb{R}^n, \quad \boldsymbol{b} \in \mathbb{R}^m$.

1. We can think of $A\boldsymbol{x}$ as representing a linear combination of the columns of A with weights x_i, ie, $A\boldsymbol{x} = x_1\boldsymbol{a}_1 + \cdots + x_n\boldsymbol{a}_n = \boldsymbol{b}$.
2. We can think of A as a *transformation matrix* or a *mapping* of $\boldsymbol{x} \in \mathbb{R}^n \mapsto \boldsymbol{b} \in \mathbb{R}^m$.

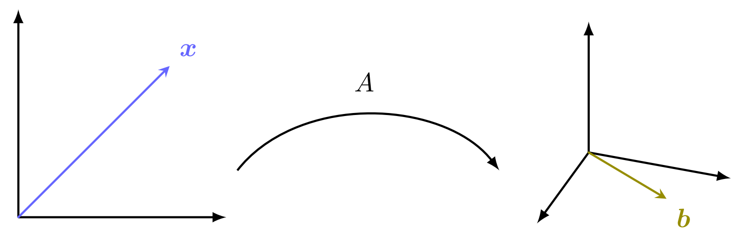


Figure 1.2: Here, we can think of A as transforming a vector from $\boldsymbol{x} \in \mathbb{R}^2$ to $\boldsymbol{b} \in \mathbb{R}^3$ via left multiplication.

Example 1.4

Let $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 4 \\ 1 & 0 & 1 \end{bmatrix}$.

a.) Describe the set of all vectors B for which $A\boldsymbol{x} = \boldsymbol{b}$ has a solution.

b.) Do the columns of A span $\mathbb{R}^3$?

Solution: Let $\boldsymbol{b} = (a,\, b,\, c)$.

a.)

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 4 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \implies \left[\begin{array}{ccc|c} 1 & 0 & 1 & a \\ 0 & 2 & 4 & b \\ 1 & 0 & 1 & c \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 1 & a \\ 0 & 1 & 2 & b/2 \\ 0 & 0 & 0 & c-a \end{array}\right]$$

$$\begin{cases} x_1 + x_3 = a \\ x_2 + 2x_3 = b/2 \\ 0 = c - a \end{cases} \implies \begin{cases} a = b_1 \\ b = b_2 \\ a = c \end{cases}$$

Thus, we can solve $A\boldsymbol{x} = \boldsymbol{b}$ exactly when $\boldsymbol{b}$ is of the form $\begin{bmatrix} b_1 \\ b_2 \\ b_1 \end{bmatrix} \left(e.g. : \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \right)$.

$$B = \left\{ \begin{bmatrix} b_1 \\ b_2 \\ b_1 \end{bmatrix} : b_1,\ b_2 \in \mathbb{R} \right\}$$

b.) No, because there are restrictions on what $\boldsymbol{b}$ can be for there to be a solution. For example, $\boldsymbol{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \in \mathbb{R}^3$ *cannot* be a solution.

■

Example 1.5

Suppose A is a matrix which has been reduced to $\begin{bmatrix} 1 & 10 & 0 & 2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. Write the solution set of $A\boldsymbol{x} = \boldsymbol{0}$ in **parametric vector form**.

Solution:

$$\begin{cases} x_1 = -10x_2 - 2x_4 \\ x_3 = -4x_4 \\ x_2 \text{ free} \\ x_4 \text{ free} \end{cases}$$

$$\implies \boldsymbol{x} = \begin{bmatrix} -10x_2 - 2x_4 \\ x_2 \\ -4x_4 \\ x_4 \end{bmatrix} = x_2 \begin{bmatrix} -10 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -2 \\ 0 \\ -4 \\ 1 \end{bmatrix}$$

Note the difference between parametric description and parametric vector form. ■

Notation:

- $\iff$ - if and only if.
- $\implies$ - implies.

A **homogeneous system** is a system of linear equations of the form $A\boldsymbol{x} = \boldsymbol{0}$.

The **trivial solution** of $A\boldsymbol{x} = \boldsymbol{0}$ is $\boldsymbol{x} = \boldsymbol{0}$. We call it trivial because something multiplied by 0 will always give you 0. Non-trivial solutions are when $\boldsymbol{x} \neq 0$.

We say that a set of column vectors of a matrix A are **linearly independent** if and only if the equation $A\boldsymbol{x} = \boldsymbol{0}$ has *only* the trivial solution. Else, they are are **linearly dependent**.

Let's try out some true or false questions:

Example 1.6

True or false? Explain.

a. A set of six vectors in $\mathbb{R}^5$ must span $\mathbb{R}^5$.

b. A set of six vectors in $\mathbb{R}^5$ cannot span $\mathbb{R}^5$.

c. Any set of three vectors in $\mathbb{R}^4$ is linearly independent.

Solution:

Form a matrix A whose columns are the 6 vectors $A = \begin{array}{c c} & \begin{array}{ccc} 1 & \dots & 6 \end{array} \\ \begin{array}{c} 1 \\ \vdots \\ 5 \end{array} & \begin{bmatrix} & & \\ & & \\ & & \end{bmatrix} \end{array}$. So A is a 5×6 matrix.

By Thm 1.4, the columns of A span $\mathbb{R}^5$ $\iff$ A has a pivot in every row. But the maximum number of pivots $= \min\{\# \text{ of rows}, \# \text{ of columns}\}$. Since A is 5×6, it has at most 5 pivots $\implies$ a. + b. are *false* (since A could have 5 pivots, or it could have less).

In order for a set of 3 vectors in $\mathbb{R}^4$ to be linearly independent, putting them into a matrix and row reducing should yield 3 pivot columns. But, since we can be given *any* 3 vectors, then we can have less than 3 pivots and c. is false.

For example: The set $\left\{ \begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\end{bmatrix} \right\}$ is linearly independent since $\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\\0&0&1\end{bmatrix}$ has a pivot in every column. But, $\left\{ \begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}1\\1\\0\\0\end{bmatrix} \right\}$ is linearly dependent since $\begin{bmatrix}1&0&1\\0&1&1\\0&0&0\\0&0&0\end{bmatrix}$ has only 2 pivots. ■

The definition of linear independence yields these three results which are handy for determining linear independence:

- A set of vectors is linearly dependent $\iff$ at least one vector in the set can be written as a linear combination of the others.
- The columns of a matrix A are linearly independent $\iff$ A has a pivot in every column.
- The columns of A are linearly dependent $\iff$ A does not have a pivot in every column.

Example 1.7

Let $A = \begin{bmatrix}1&3&-1\\-1&-5&5\\4&7&h\end{bmatrix}$. Find all values of h for which the columns of A form a linearly dependent set.

Solution: The columns of A are dependent $\iff$ A does not have a pivot in every column.

$$\begin{aligned}\begin{bmatrix}1&3&-1\\-1&-5&5\\4&7&h\end{bmatrix} &\sim \begin{bmatrix}1&3&-1\\0&-2&4\\0&-5&h+4\end{bmatrix}\\ &\sim \begin{bmatrix}1&3&-1\\0&1&-2\\0&0&h-6\end{bmatrix}\\ &\implies h=6.\end{aligned}$$

Note: If $h=6$, we can see that $A\boldsymbol{x}=\mathbf{0}$ has a nontrivial solution.

$$\begin{cases} x_1 = -5x_3 \\ x_2 = 2x_3 \\ x_3 \text{ free} \end{cases}$$

1 such solution: $-5\begin{bmatrix}1\\-1\\4\end{bmatrix} + 2\begin{bmatrix}3\\-5\\7\end{bmatrix} + \begin{bmatrix}-1\\5\\6\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}$. ■

1.3 Linear Algebra Transformed me into a Math Major

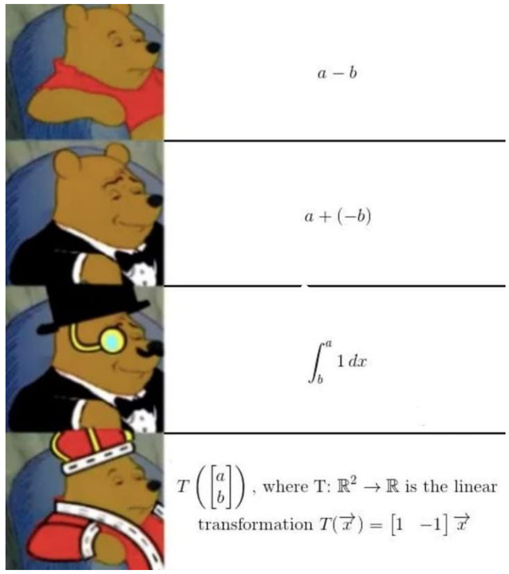


Figure 1.3: A sophisticated use of linear transformations. Credit: Shemp from the UCSD undergrad math discord

1.3.1 Function Preliminaries

Given an $m \times n$ matrix A, we can define a function T by the formula $T(\boldsymbol{x}) = A\boldsymbol{x}$.

Note:

- The **domain** of T is $\mathbb{R}^n$.
- The **range** of $T = \{\boldsymbol{b} \in \mathbb{R}^m : A\boldsymbol{x} = \boldsymbol{b} \text{ is consistent}\}$.

Notation:

- $\underbrace{T}_{\substack{\text{Name of} \\ \text{function}}} : \underbrace{\mathbb{R}^n}_{\text{domain}} \longrightarrow \underbrace{\mathbb{R}^m}_{\substack{\textbf{codomain} \\ \text{“where the} \\ \text{outputs live”}}}$
- $\underbrace{\boldsymbol{x} \mapsto A\boldsymbol{x}}_{\text{formula for } T}$ – “$\boldsymbol{x}$ maps to $A\boldsymbol{x}$”
- $\operatorname{Im} T$ – The **image** of T, aka the range of T, aka $T(\mathbb{R}^n)$.
- $\subseteq$ – subset of or equal to.

For example: Take the function $f(x) = x^2$.

- $f : \mathbb{R} \to \mathbb{R}$
- $x \mapsto x^2$

The function $f(x)$ is a real-valued function that takes real numbers as its input and outputs real numbers. Even though the *range* of f is nonnegative, we say that the *codomain* is $\mathbb{R}$ since the output is taken from $\mathbb{R}$. In other words, the range may equal the codomain, or it may be a subset of it. I.e., range $\subseteq$ codomain.

Example 1.8

If A is a 7×11 matrix, what do p and q have to be in order to define a function $T : \mathbb{R}^p \to \mathbb{R}^q$ by the formula $T(\boldsymbol{x}) = A\boldsymbol{x}$?

Solution: Two matrices A and B, where A is $m \times n$ and B is $r \times k$, are **conformable** if their dimensions are suitable for defining some operation. The product AB is defined if $n = r$ and the product $AB = C$ has dimension $m \times k$.

In other words, in order to multiply two matrices, the number of columns of the first must match the number of rows of the second. The resulting product has the number of rows of the first matrix and the number of columns of the second matrix.

Therefore, the product $A\boldsymbol{x}$ makes sense $\iff$ $\boldsymbol{x} \in \mathbb{R}^{11}$ and we have that $A\boldsymbol{x} \in \mathbb{R}^7 \implies p = 11,\ q = 7.$ ■

Check out this page to review computing products of matrices.

Definition 1.9

A function T is a **linear transformation** if

a. $T(\mathbf{0}) = \mathbf{0}$

b. $T(c_1\boldsymbol{u} + c_2\boldsymbol{v}) = c_1T(\boldsymbol{u}) + c_2T(\boldsymbol{v})$

Example 1.10

Suppose $\operatorname{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2\} = \mathbb{R}^2$. Suppose $T : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transformation. Show that if $T(\boldsymbol{u}_1) = \mathbf{0}$ and $T(\boldsymbol{u}_2) = \mathbf{0}$, then for any $\boldsymbol{x} \in \mathbb{R}^2,\ T(\boldsymbol{x}) = \mathbf{0}$.

Solution: Let $\boldsymbol{x}$ be any vector in $\mathbb{R}^2$. We want to calculate $T(\boldsymbol{x})$. Since $\operatorname{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2\} = \mathbb{R}^2$, we can find $x_1,\ x_2$ so that $\boldsymbol{x} = x_1\boldsymbol{u}_1 + x_2\boldsymbol{u}_2$. So then

$$\begin{aligned} T(\boldsymbol{x}) &= T(x_1\boldsymbol{u}_1 + x_2\boldsymbol{u}_2) \\ &= x_1T(\boldsymbol{u}_1) + x_2T(\boldsymbol{u}_2) && \text{(since } T \text{ is linear)} \\ &= x_1 \cdot \mathbf{0} + x_2 \cdot \mathbf{0} \\ &= \mathbf{0}. \end{aligned}$$

■

Example 1.11

Show that the transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(x_1, x_2) = (4x_1 - 3x_2, |x_2|)$ is not linear. *Hint:* Find a specific example to show that at least one of the linearity properties fail.

Solution: Let $\boldsymbol{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $c = -2$. Then

$$\begin{aligned} T(c\boldsymbol{u}) &= T\left(\begin{bmatrix} -2 \\ -2 \end{bmatrix}\right) = \begin{bmatrix} -2 \\ 2 \end{bmatrix} \\ cT(\boldsymbol{u}) &= -2T\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = -2\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ -2 \end{bmatrix} \\ &\implies T(c\boldsymbol{u}) \neq cT(\boldsymbol{u}) \\ &\implies T \text{ is not linear.} \end{aligned}$$

■

1.3.2 The Standard Matrix of a Linear Transformation

Problem: Given a linear transformation T, find a matrix A so that $T(\boldsymbol{x}) = A\boldsymbol{x}$.

Idea: Given $T : \mathbb{R}^n \to \mathbb{R}^m$ and T linear,

$$\begin{aligned} T\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right) &= T\left(\begin{bmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}\right) + T\left(\begin{bmatrix} 0 \\ x_2 \\ \vdots \\ 0 \end{bmatrix}\right) + \cdots + T\left(\begin{bmatrix} 0 \\ 0 \\ \vdots \\ x_n \end{bmatrix}\right) \\ &= x_1 T\left(\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}\right) + x_2 T\left(\begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}\right) \cdots + x_n T\left(\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}\right) \\ &= x_1 T(\boldsymbol{e}_1) + x_2 T(\boldsymbol{e}_2) + \cdots + x_n T(\boldsymbol{e}_n) \\ &= \begin{bmatrix} T(\boldsymbol{e}_1) \; T(\boldsymbol{e}_2) \; \cdots \; T(\boldsymbol{e}_n) \end{bmatrix} \boldsymbol{x} \end{aligned}$$

Where $\boldsymbol{e}_i$ is a **standard unit vector** with $\boldsymbol{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, $\boldsymbol{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \boldsymbol{e}_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$ In other words, $\boldsymbol{e}_i$ is the ith column of the identity matrix I_n.

$A = \begin{bmatrix} T(\boldsymbol{e}_1) \; T(\boldsymbol{e}_2) \; \cdots \; T(\boldsymbol{e}_n) \end{bmatrix}$ is called the **standard matrix for the linear transformation** T. Thus, we can see that by using the linearity of T and decomposing $\boldsymbol{x}$ into a linear combination of standard unit vectors, we can find A by figuring out what happens to $\boldsymbol{e}_i$ under T.

Example 1.12

Find the standard matrix A for the linear transformation

a.) $T : \mathbb{R}^2 \to \mathbb{R}^2$ which reflects vectors over the line $x_1 = x_2$ and then stretches the first coordinate by 3.

b.) $T : \mathbb{R}^3 \to \mathbb{R}^4$ with $T(x_1, x_2, x_3) = (x_1 + x_2,\, x_2 + x_3,\, x_1 + x_3,\, 0)$.

Solution:

a.) By the previous discussion, $A = \begin{bmatrix} T\left(\begin{bmatrix}1\\0\end{bmatrix}\right) & T\left(\begin{bmatrix}0\\1\end{bmatrix}\right)\end{bmatrix}$. A reflection over the line $x_1 = x_2$ results in the vector coordinates switching places and stretching the first coordinate by 3 results in multiplying the first coordinate by 3. Thus,

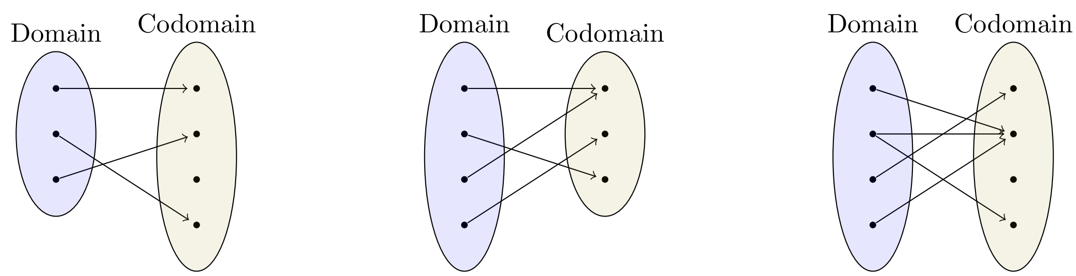
$$\begin{aligned} T(\boldsymbol{e}_1) &= \begin{bmatrix}0\\1\end{bmatrix} \\ T(\boldsymbol{e}_2) &= \begin{bmatrix}3\\0\end{bmatrix} \\ \implies A &= \begin{bmatrix}0 & 3\\1 & 0\end{bmatrix} \end{aligned}$$

b.)

$$\begin{aligned} T(\boldsymbol{x}) &= \begin{bmatrix} x_1 + x_2 \\ x_2 + x_3 \\ x_1 + x_3 \\ 0 \end{bmatrix} = x_1 \begin{bmatrix}1\\0\\1\\0\end{bmatrix} + x_2 \begin{bmatrix}1\\1\\0\\0\end{bmatrix} + x_3 \begin{bmatrix}0\\1\\1\\0\end{bmatrix} \\ \implies A &= \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

■

A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is **onto** if $\operatorname{Im} T = \mathbb{R}^m$, and **1-to-1** if $T(\boldsymbol{x}_1) = T(\boldsymbol{x}_2) \implies \boldsymbol{x}_1 = \boldsymbol{x}_2$.



(a) This function is *1-to-1* because each input in the domain maps to exactly one unique output in the codomain. It is not onto because there are points in the codomain that do not get mapped to.

(b) This function is *onto* because for each output in the codomain, there exists an input in the domain that maps to it. It is not 1-to-1 because there exists an output that gets mapped to by two distinct inputs.

(c) This is not even a function because there exists an input that maps to two distinct outputs.

Figure 1.4: Visualization of *1-to-1* and *onto* functions.

In other words, imagine a blank wall that you throw a bucket of paint at. The paint is your domain and the wall is your codomain. You are *onto* if the paint covers the *entire* wall, i.e., all the paint gets onto the wall, although perhaps multiple drops of paint cover the same part of the wall. You are *1-to-1* if each drop of paint gets assigned to a unique point on the wall (no piece of the wall gets two drops of paint), although you may not cover the entire wall.

If $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation with standard matrix A, then

- T is onto $\iff$ $A\boldsymbol{x} = \boldsymbol{b}$ is consistent for all $\boldsymbol{b} \in \mathbb{R}^m$.
  $\iff$ A has a pivot in every row.
- T is 1-to-1 $\iff$ $A\boldsymbol{x} = \mathbf{0}$ has only the trivial solution
  $\iff$ A has a pivot in every column.

Example 1.13

Let $T : \mathbb{R}^3 \to \mathbb{R}^4$ with $T(x_1, x_2, x_3) = (x_1 + x_2,\, x_2 + x_3,\, x_1 + x_3,\, 0)$. Is T onto? If T 1-to-1?

Solution:

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

So, T is 1-to-1 but not onto. ∎

Chapter 2

mAtrIX aLgeBrA

2.1 Matrix Operations, Inverses, IMT, Oh My!

This chapter is where the holy grail theorem of this class makes its introduction – The Invertible Matrix Theorem (IMT). It's big and it has a lot of equivalent statements. Make sure you spend some time looking at this theorem and convince yourself why one statement would imply that another is true. Thorough understanding of all the statements in this theorem will demonstrate a thorough understanding of a lot of the concepts in this class.

Warning: The IMT can only be applied to square matrices!

If A is $n \times n$, then the **inverse** of A, if it exists, denoted A^{-1}, satisfies $A^{-1}A = AA^{-1} = I_n$.

A is **invertible** if and only if you can show that one of the statements in the IMT is true for A.

From the IMT, an $n \times n$ matrix A is invertible $\iff$ the RREF of A is I_n. Thus, to find A^{-1}, assuming that A is invertible:

1. Form the augmented matrix $[A \mid I_n]$.
2. Row reduce this matrix to RREF.
3. A will row reduce to I_n and the augmented side will become A^{-1} via your row operations. In other words,
$$\left[A \mid I_n\right] \sim \left[I_n \mid A^{-1}\right]$$

If A does not row reduce to the identity matrix, then A is *not* invertible.

Example 2.1

Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$. Find A^{-1}.

Solution:

$$\begin{aligned}\left[A \mid I_3\right] &= \left[\begin{array}{ccc|ccc}1 & 2 & 3 & 1 & 0 & 0\\ 2 & 4 & 5 & 0 & 1 & 0\\ 3 & 5 & 6 & 0 & 0 & 1\end{array}\right]\\ &\sim \left[\begin{array}{ccc|ccc}1 & 0 & 0 & 1 & -3 & 2\\ 0 & 1 & 0 & -3 & 3 & -1\\ 0 & 0 & 1 & 2 & -1 & 0\end{array}\right]\end{aligned}$$

Thus, $A^{-1} = \begin{bmatrix}1 & -3 & 2\\ -3 & 3 & -1\\ 2 & -1 & 0\end{bmatrix}$. ■

Note: If A is invertible and you just want to find the jth column of A^{-1}, row reduce $\left[A \mid \boldsymbol{e}_j\right]$.

For example: If B is 5×5 and invertible, to find the 4th column of B^{-1}, row reduce $\left[B \mid \boldsymbol{e}_4\right]$, i.e., $\left[\begin{array}{c|c} & \\ B & \begin{bmatrix}0\\0\\0\\1\\0\end{bmatrix} \\ & \end{array}\right]$.

Let $A = \begin{bmatrix}a & b\\ c & d\end{bmatrix}$. For 2×2 matrices, the formula for its inverse, if it exists, is

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

Example 2.2

Find the inverse of $A = \begin{bmatrix}6 & -1\\ 2 & 1\end{bmatrix}$ and $B = \begin{bmatrix}3 & -1\\ 6 & -2\end{bmatrix}$ if it exists.

Solution:

$$A^{-1} = \frac{1}{8}\begin{bmatrix}1 & 1\\ -2 & 6\end{bmatrix} = \begin{bmatrix}1/8 & 1/8\\ -1/4 & 3/4\end{bmatrix}.$$

For B, note that $ad-bc = (3)(-2)-(-1)(6) = 0 \implies B$ has no inverse. ■

$T:\mathbb{R}^n \to \mathbb{R}^n$ is **invertible** if there exist $S:\mathbb{R}^n\to\mathbb{R}^n$ such that $S(T(\boldsymbol{x})) = T(S(\boldsymbol{x})) = \boldsymbol{x}$.

Example 2.3

Define $T:\mathbb{R}^2\to\mathbb{R}^2$ by $T(\boldsymbol{x}) = \begin{bmatrix}6x_1 - x_2\\ 2x_1 + x_2\end{bmatrix}$. Show that T is invertible and find a formula for T^{-1}.

Solution: Recall that $T(\boldsymbol{x}) = A\boldsymbol{x}$ where $A = \left[T(\boldsymbol{e}_1)\; T(\boldsymbol{e}_2)\right] = \begin{bmatrix}6 & -1\\ 2 & 1\end{bmatrix}$.

From example 2.2, $A^{-1} = \begin{bmatrix} 1/8 & 1/8 \\ -1/4 & 3/4 \end{bmatrix}$. Hence, we have that $T^{-1}(\boldsymbol{x}) = A^{-1}\boldsymbol{x} = \begin{bmatrix} \frac{1}{8}x_1 + \frac{1}{8}x_2 \\ -\frac{1}{4}x_1 + \frac{3}{4}x_2 \end{bmatrix}$. ■

Example 2.4

Is $A = \begin{bmatrix} 1 & 0 & 0 & 3 \\ 2 & 5 & 10 & 4 \\ -1 & 3 & 6 & 9 \\ 2 & 1 & 2 & -1 \end{bmatrix}$ invertible? *Hint:* Do not do any row reducing!

Solution: Nope. Since column $3 = 2 \cdot$ column 2, we have that the columns of A are not linearly independent. Hence, A is not invertible by the IMT. ■

Example 2.5

Suppose C is an 8×8 matrix whose columns span $\mathbb{R}^8$. What is the span of the rows of C?

Solution: By the IMT, since the columns of C span $\mathbb{R}^8$, then C^T is also invertible. By the IMT, the columns of C^T span $\mathbb{R}^8$. But the columns of C^T are the rows of C. Thus, the rows of C span $\mathbb{R}^8$. ■

Example 2.6

Suppose B is a 20×20 matrix with 18 pivots. How many solutions are there to $B\boldsymbol{x} = \mathbf{0}$?

Solution: Since B has less than 20 pivots, by the IMT, $B\boldsymbol{x} = \mathbf{0}$ must have more than the trivial solution. So $B\boldsymbol{x} = \mathbf{0}$ has infinitely many solutions. ■

Chapter 3

Determinants

Our goal in this chapter is the define the *determinant* of an $n \times n$ matrix and study its properties. Determinants have a nice geometric interpretation which 3Blue1Brown does an amazing job showcasing in his Essence of Linear Algebra series. In particular, his videos on determinants and spaces are helpful for visualizing the concepts in this section.

Note: In class, we went over vector spaces before determinants and we will be assuming knowledge of chapter 4 throughout this chapter. I only numbered this chapter 3 to align with the chapters in the textbook (I might change this in the future if I decide to make this a more standalone set of notes).

3.1 You are Determinated

Definition 3.1

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. The **determinant** of a 2×2 matrix is $ad - bc$.

Recall that the formula for the inverse of a 2×2 matrix A is given by $A^{-1} = \dfrac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. Note that this inverse exists $\iff$ $\det A \neq 0$ which is one of the statements of The Invertible Matrix Theorem.

Notation:

- $\det A$ – The determinant of A.
- $|A|$ – The determinant of A.
- a_{ij} – The ith row and jth column entry of A.

For $n > 2$, one way determinants are found is by iteratively summing up determinants of smaller subsets of the matrix A. This process is known as **cofactor expansion**.

We define the *cofactor expansion along the 1st row* of A to be:

$$\det A = a_{11} \det A_{11} - a_{12} \det A_{12} + a_{13} \det A_{13} + \cdots + (-1)^{1+n} a_{1n} \det A_{1n}$$

where A_{ij} is the $(n-1) \times (n-1)$ submatrix obtained by deleting the ith row and jth column of A.

For example:

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

where we deleted the entries in pink.

Definition 3.2

Let $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$. The **determinant** of A is defined to be the *cofactor expansion* of A along *any* row or column.

Restating Thm 3.1 in slightly different words:

We can compute $\det A$ using cofactor expansion along any row or column and we always get the same answer. I.e.,

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij} \quad \longleftarrow \text{ along } j\text{th column}$$

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij} \quad \longleftarrow \text{ along } i\text{th row}$$

An immediate application of this theorem is that we can choose to compute $\det A$ in the easiest way possible.

Example 3.3

Find $\begin{vmatrix} 2 & 1 & 3 & 4 \\ 0 & -1 & 6 & 5 \\ 0 & 0 & 3 & 2 \\ 0 & 0 & 0 & 7 \end{vmatrix}$.

Solution: We notice that the 4th row has many zeroes, so a lot of terms in its cofactor expansion will be zero.

$$\begin{aligned} \begin{vmatrix} 2 & 1 & 3 & 4 \\ 0 & -1 & 6 & 5 \\ 0 & 0 & 3 & 2 \\ 0 & 0 & 0 & 7 \end{vmatrix} &= (-1)^{4+1} \cdot 0 + (-1)^{4+2} \cdot 0 + (-1)^{4+3} \cdot 0 + (-1)^{4+4} \cdot 7 \cdot \begin{vmatrix} 2 & 1 & 3 \\ 0 & -1 & 6 \\ 0 & 0 & 3 \end{vmatrix} \\ &= 7 \begin{vmatrix} 2 & 1 & 3 \\ 0 & -1 & 6 \\ 0 & 0 & 3 \end{vmatrix} \\ &= 7 \left((-1)^{3+3} \cdot 3 \begin{vmatrix} 2 & 1 \\ 0 & -1 \end{vmatrix} \right) \\ &= 7 \cdot 3 \cdot (-2) \\ &= -42. \end{aligned}$$

■

A **triangular matrix** is a square matrix whose entries above or below the main diagonal are 0. An **upper triangular** matrix is a triangular matrix with the form $U = \begin{bmatrix} a_{11} & & & \\ & \ddots & * & \\ & 0 & \ddots & \\ & & & a_{nn} \end{bmatrix}$ and a **lower triangular** matrix is a triangular matrix with the form $L = \begin{bmatrix} a_{11} & & & \\ & \ddots & 0 & \\ & * & \ddots & \\ & & & a_{nn} \end{bmatrix}$ where the $*$ and diagonal entries can be any real number.

So then Thm 3.2 states that the determinant of a triangular matrix is the product of its diagonal entries:

- $\det\left(\begin{bmatrix} a_{11} & & & \\ & \ddots & 0 & \\ & * & \ddots & \\ & & & a_{nn} \end{bmatrix}\right) = a_{11} \cdot a_{22} \cdot \cdots \cdot a_{nn}.$
- $\det\left(\begin{bmatrix} a_{11} & & & \\ & \ddots & * & \\ & 0 & \ddots & \\ & & & a_{nn} \end{bmatrix}\right) = a_{11} \cdot a_{22} \cdot \cdots \cdot a_{nn}.$

Just one more helpful theorem before we get into an example:

From Thm 3.3, we have that if A is an $n \times n$ matrix:

- Interchanging 2 rows of A multiplies $\det A$ by -1.
- Adding a multiple of one row of A to another row of A does not change $\det A$.
- Multiplying a row of A by k multiples $\det A$ by k.

Example 3.4

Suppose $\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = -3$. What is $\begin{vmatrix} g & h & i \\ 8d+2g & 8e+2h & 8f+2i \\ a & b & c \end{vmatrix}$?

Solution: Let the given matrix be A and the second be B. The operations applied to get from A to B are:

1. Interchange row 1 and 3 $\implies$ $(-1)\det A$.
2. $8\cdot$ row 2 $\implies$ $8 \cdot \det A$.
3. row 2 + $2\cdot$ row 3 $\implies$ no change.

Thus, $\det B = (-1)(8)\det A = 24$. ■

Properties of determinants: Let A and B be $n \times n$ matrices,

a.) $\det(AB) = \det A \cdot \det B$.

b.) $\det\left(A^T\right) = \det A$.

c.) $\det\left(A^{-1}\right) = \dfrac{1}{\det A}$, (if A is invertible).

Note: $\det(A + B) \neq \det A + \det B$.

Example 3.5

Let $B = \begin{bmatrix} 3 & -7 & 1 \\ 2 & 0 & 5 \\ 1 & 4 & 1 \end{bmatrix}$ so then $\det B = -73$.

Find:

a.) $\det\left(B^T\right)$

b.) $\det\left(B^5\right)$

c.) $\det\left(B^{-1}\right)$

Solution:

a.) $\det\left(B^T\right) = -73$.

b.) $\det\left(B^5\right) = (-73)^5$.

c.) $\det\left(B^{-1}\right) = -\dfrac{1}{73}$.

■

3.2 Applications of Determinants

3.2.1 Cramer's Rule

Cramer's Rule: If A is an invertible $n \times n$ matrix, then for every $\boldsymbol{b} \in \mathbb{R}^n$, $A\boldsymbol{x} = \boldsymbol{b}$ has the unique solution $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ where $x_j = \dfrac{\det(A_j(\boldsymbol{b}))}{\det A}$ where $A_j(\boldsymbol{b})$ is the matrix A except that the jth column is replaced by $\boldsymbol{b}$.

Example 3.6

Use Cramer's Rule to find all α such that the system has a unique solution and describe the solution.

$$\begin{aligned} 3\alpha x_1 - 2x_2 &= 4 \\ -6x_1 + \alpha x_2 &= 1 \end{aligned}$$

Solution: Let $A = \begin{bmatrix} 3\alpha & -2 \\ -6 & \alpha \end{bmatrix}$. Then

$$\begin{aligned} \det A &= 3\alpha^2 - 12 \\ &= 3(\alpha - 2)(\alpha + 2) \\ &\implies \quad \text{So then the system has a unique solution whenever } \alpha \neq \pm 2. \end{aligned}$$

Let $A_1(\boldsymbol{b}) = \begin{bmatrix} 4 & -2 \\ 1 & \alpha \end{bmatrix}$, $A_2(\boldsymbol{b}) = \begin{bmatrix} 3\alpha & 4 \\ -6 & 1 \end{bmatrix}$

So then

$$x_1 = \frac{4\alpha + 2}{3(\alpha - 2)(\alpha + 2)}$$
$$x_2 = \frac{3\alpha + 24}{3(\alpha - 2)(\alpha + 2)}$$

when $\alpha \neq \pm 2$. ■

3.2.2 Geometric Intuition

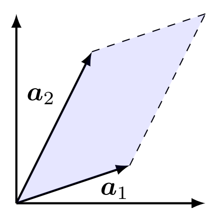


Figure 3.1: Area of a parallelogram.

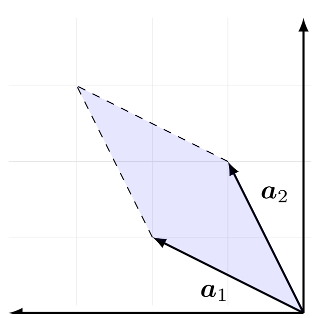
If $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ are vectors in $\mathbb{R}^2$, then the area of the parallelogram determined by $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ equals

$$\Big| \det([\boldsymbol{a}_1 \ \boldsymbol{a}_2]) \Big|.$$

Example 3.7

Find the area of the parallelogram with vertices $(0, 0)$, $(-1, 2)$, $(-2, 1)$, and $(-3, 3)$.

Solution:



$$\begin{vmatrix} -2 & -1 \\ 1 & 2 \end{vmatrix} = -4 + 1 = -3.$$

Thus, the area of the parallelogram is the absolute value of of this determinant, which is 3.

Note that due to the geometric properties of parallelograms, we would've still gotten the same answer if we had $\boldsymbol{a}_2 = (-3, 3)$ instead. ■

Similarly, in $\mathbb{R}^3$, the volume of the parallelepiped determined by the columns of a 3×3 matrix A equals

$$\Big| \det A \Big|.$$

Example 3.8

Find the volume of the parallelepiped with one vertex at the origin and adjacent vertices $(-1, 1, 5)$, $(2, 0, 3)$, and $(4, 4, -2)$.

Solution: We have that $A = \begin{bmatrix} -1 & 2 & 4 \\ 1 & 0 & 4 \\ 5 & 3 & -2 \end{bmatrix}$. Since there's a zero in the second row, lets do cofactor expansion along the second row:

$$\begin{aligned} \begin{vmatrix} -1 & 2 & 4 \\ 1 & 0 & 4 \\ 5 & 3 & -2 \end{vmatrix} &= -1 \begin{vmatrix} 2 & 4 \\ 3 & -2 \end{vmatrix} + 0 - 4 \begin{vmatrix} -1 & 2 \\ 5 & 3 \end{vmatrix} \\ &= -1(-4 - 12) - 4(-3 - 10) \\ &= 68 \\ \implies \text{volume } &= 68. \end{aligned}$$

■

The determinant of a transformation matrix A is how areas are scaled after the transformation.

Example 3.9

Suppose a transformation doubles the area of a unit square. What is $\det A$, where A is the transformation matrix for said transformation?

Solution:

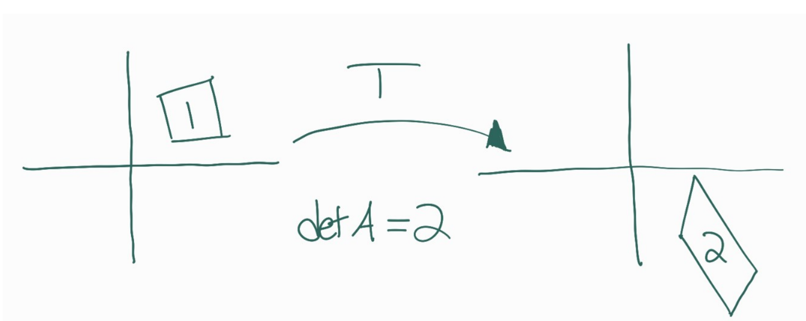


Figure 3.2: The determinant of a transformation tells you how areas are scaled after a transformation.

■

Note: If $\det A = 0$, then areas get "squished" after the transformation.

Example 3.10

Suppose $\det A = 0$ for a 3×3 transformation matrix A. What are the possible values for $\dim \operatorname{Nul} A$ and describe what's going on geometrically during the transformation.

Solution: Since A is a 3×3 matrix, then the transformation transforms 3D space into 3D space. Since $\det A = 0$, then there exists some areas that get "squished" or mapped to $\mathbf{0}$ after the transformation.

Thus, $1 \leq \dim \operatorname{Nul} A \leq 3$ since there must exist some nonzero vectors that get mapped to zero.

If $\dim \operatorname{Nul} A = 1$, then there exists a line in $\mathbb{R}^3$ such that all vectors on that line get mapped to $\mathbf{0}$ after the transformation. By Rank-Nullity, $\operatorname{rank} A = 2$; so $\mathbb{R}^3$ space gets squished down into a plane (in $\mathbb{R}^3$) since the column space of A spans a plane.

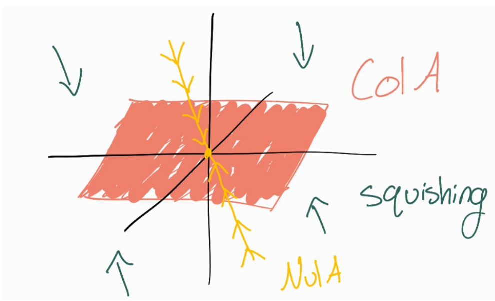


Figure 3.3: Under the transformation, $\mathbb{R}^3$ gets "squished" down into a plane, as shown by the column space of A. Since the nullity is 1, there exists a line such that all vectors on that line get mapped to $\mathbf{0}$ under the transformation.

If $\dim \operatorname{Nul} A = 2$, then there exists a plane in $\mathbb{R}^3$ such that all vectors on that plane gets mapped to $\mathbf{0}$ after the transformation. So then $\mathbb{R}^3$ space gets squished into a line (since $\operatorname{rank} A = 1$).

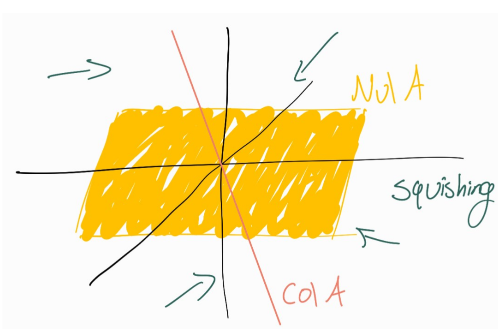


Figure 3.4: Under the transformation, $\mathbb{R}^3$ gets "squished" down into a line, as shown by the column space of A. Since the nullity is 2, there exists a plane such that all vectors on that plane get mapped to $\mathbf{0}$ under the transformation.

■

Chapter 4

Vector Spaces

4.1 Are you a Sub.......space?

Idea: $\mathbb{R}^1$, $\mathbb{R}^2$, $\mathbb{R}^3$, ... are all examples of a more abstract object called a **vector space**: a collection of objects called *vectors*. The more formal definition of a vector space is not as important as the next definition, but I'll include it here for completeness, but I won't dignify it with its own box.

A **vector space** is a nonempty set V of objects, called *vectors*, coupled with two operations which we call *addition* and *multiplication by scalars* (real numbers), subject to the ten axioms below. The axioms must hold for all vectors $\boldsymbol{u}$, $\boldsymbol{v}$, and $\boldsymbol{w}$ in V and for all scalars c and d.

1. $\boldsymbol{u} + \boldsymbol{v} \in V$. (closed under addition)
2. $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$. (commutativity)
3. $(\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w} = \boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w})$ (associativity)
4. There exists an *additive identity* element called $\mathbf{0} \in V$ such that $\boldsymbol{u} + \mathbf{0} = \boldsymbol{u}$.
5. For each $\boldsymbol{u} \in V$, there exists an *additive inverse* element called $-\boldsymbol{u} \in V$ such that $\boldsymbol{u} + (-\boldsymbol{u}) = \mathbf{0}$.
6. $c\boldsymbol{u} \in V$. (closed under scalar multiplication)
7. $c(\boldsymbol{u} + \boldsymbol{v}) = c\boldsymbol{u} + c\boldsymbol{v}$. (distributive property for scalars)
8. $(c + d)\boldsymbol{u} = c\boldsymbol{u} + d\boldsymbol{u}$ (distributive property for vectors)
9. $c(d\boldsymbol{u}) = (cd)\boldsymbol{u}$
10. There exists a *multiplicative identity* element called 1 such that $1\boldsymbol{u} = \boldsymbol{u}$.

More often, we will be more concerned about determining whether something is a *subspace* of a *vector space*. This is because we then only need to check three conditions instead of ten, and conveniently, a subspace of a vector space is itself a vector space so it's a much quicker way to find vector spaces.

Definition 4.1

A **subspace** of a vector space V is a subset W of V that satisfies these three conditions:

1. $\mathbf{0} \in W$.
2. W is closed under addition ($\boldsymbol{u}, \boldsymbol{v} \in W \implies \boldsymbol{u} + \boldsymbol{v} \in W$).
3. W is closed under scalar multiplication ($c \in \mathbb{R}, \boldsymbol{u} \in W \implies c\boldsymbol{u} \in W$).

Note that 3 $\implies$ 1 so we really only need to check 2 and 3.

Examples of nonvector spaces:

- $\mathbb{Z}$ = the set of integers $\{\dots, -2, -1, 0, 1, 2, \dots\}$.
  Reason: not closed under scalar multiplication ($3 \in \mathbb{Z}$ but $\frac{1}{2} \cdot 3 = \frac{3}{2} \notin \mathbb{Z}$).
- V = the set of polynomials of degree 3.
  Reasons: 0 is not a degree 3 polynomial. Also, $x^3 + (-x^3) = 0 \notin V$ (not closed under addition).
- The 3rd quadrant in $\mathbb{R}^2$, i.e., $\left\{ \begin{bmatrix} x \\ y \end{bmatrix} : x, y \leq 0 \right\} = W$.
  Reason: e.g. $\begin{bmatrix} -3 \\ -2 \end{bmatrix} \in W$ but $-2 \begin{bmatrix} -3 \\ -2 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix} \notin W$.

Example 4.2

Let $M_{4\times4}$ be the set of 4×4 matrices. Determine if the given subset is a *subspace* of $M_{4\times4}$:

a. W = the set of 4×4 invertible matrices.

b. S = the set of 4×4 diagonal matrices.

Solution:

a. No, since $\mathbf{0} \notin W$. $\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ is not invertible.

b. A **diagonal matrix** is a square matrix whose only nonzero entries are along the **main diagonal** (the diagonal that starts at the top-left element and goes down towards the bottom-right).

   Yes. The set of matrices in S take on the form $\begin{bmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & d \end{bmatrix}$, $a, b, c, d \in \mathbb{R}$.

   1 $\mathbf{0} \in S$ because $\mathbf{0} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ with $a = b = c = d = 0$.

   2 Let $A, B \in S$. We need to show that $A + B \in S$.

   Let $A = \begin{bmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & d \end{bmatrix}$ and $B = \begin{bmatrix} e & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & g & 0 \\ 0 & 0 & 0 & h \end{bmatrix}$. Then $A + B = \begin{bmatrix} a+e & 0 & 0 & 0 \\ 0 & b+f & 0 & 0 \\ 0 & 0 & c+g & 0 \\ 0 & 0 & 0 & d+h \end{bmatrix} \in S$.

   3 Let $A \in S$, $k \in \mathbb{R}$. We need to show $kA \in S$.

$$kA = \begin{bmatrix} ka & 0 & 0 & 0 \\ 0 & kb & 0 & 0 \\ 0 & 0 & kc & 0 \\ 0 & 0 & 0 & kd \end{bmatrix} \in S.$$

■

4.2 All About that Space

Let A be an $m \times n$ matrix.

- The **null space** of A is the set of solutions to $A\boldsymbol{x} = \mathbf{0}$.
  Notation: $\operatorname{Nul} A = \{\boldsymbol{x} \in \mathbb{R}^n : A\boldsymbol{x} = \mathbf{0}\}$.
- The **column space** of A is the span of its columns.
  Notation: $\operatorname{Col} A = \{\boldsymbol{b} \in \mathbb{R}^m : A\boldsymbol{x} = \boldsymbol{b} \text{ is consistent}\}$.
- The **row space** of A is the span of its rows.
  Notation: $\operatorname{Row} A = \{\boldsymbol{b} \in \mathbb{R}^n : A^T\boldsymbol{x} = \boldsymbol{b} \text{ is consistent}\}$.
- Let V be a vector space. Let $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_p \in V$. Let $S = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_p\}$. S is called a **basis** for V if
  (i.) S is linearly independent.
  (ii.) $\operatorname{Span} S = V$ (i.e., S spans V).

Example 4.3

Given an $m \times n$ matrix A:

1. $\operatorname{Nul} A$ is a subspace of $\mathbb{R}^n$.
2. $\operatorname{Col} A$ is a subspace of $\mathbb{R}^m$.
3. $\operatorname{Row} A$ is a subspace of $\mathbb{R}^n$.

Example 4.4

Let $A = \begin{bmatrix} 1 & 2 & 3 & 0 \\ 0 & 1 & 1 & -1 \end{bmatrix}$

a. $\operatorname{Nul} A$ is a subspace of $\mathbb{R}^4$.

b. $\operatorname{Col} A$ is a subspace of $\mathbb{R}^2$.

c. Find a set of vectors that span $\operatorname{Nul} A$.

d. Find a nonzero vector in $\operatorname{Col} A$.

Solution:

c. $\left[A \mid \mathbf{0}\right] \sim \left[\begin{array}{cccc|c} 1 & 0 & 1 & 2 & 0 \\ 0 & 1 & 1 & -1 & 0 \end{array}\right]$. So $A\boldsymbol{x} = \mathbf{0}$ has solutions of the form:

$$\begin{cases} x_1 = -x_3 - 2x_4 \\ x_2 = x_4 - x_3 \\ x_3,\ x_4 \text{ free} \end{cases}$$

$$\implies \boldsymbol{x} = x_3 \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

So $\operatorname{Nul} A = \operatorname{Span}\left\{ \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 0 \\ 1 \end{bmatrix} \right\}$.

d. $\begin{bmatrix} 1 \\ 0 \end{bmatrix} \in \operatorname{Col} A$. ■

Example 4.5

Let $A = \begin{bmatrix} 1 & 3 \\ 0 & 1 \\ 2 & -1 \end{bmatrix}$.

a. Is $\boldsymbol{b} = \begin{bmatrix} 2 \\ 1 \\ 5 \end{bmatrix}$ in $\operatorname{Col} A$?

b. Is $\boldsymbol{c} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$ in $\operatorname{Nul} A$?

Solution:

a.) By definition, $\boldsymbol{b} \in \operatorname{Col} A \iff A\boldsymbol{x} = \boldsymbol{b}$ is consistent. But,

$$\begin{bmatrix} A \mid \boldsymbol{b} \end{bmatrix} \sim \left[\begin{array}{cc|c} 1 & 3 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 8 \end{array}\right] \text{ is inconsistent.}$$

Thus, $\boldsymbol{b} \notin \operatorname{Col} A$.

b.) By Definition, $\boldsymbol{c} \in \operatorname{Nul} A \iff A\boldsymbol{c} = \boldsymbol{0}$.

$$\begin{aligned} A\boldsymbol{c} &= 3 \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} + -1 \begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ -1 \\ 7 \end{bmatrix} \\ &\neq \boldsymbol{0}. \end{aligned}$$

Thus, $\boldsymbol{c} \notin \operatorname{Nul} A$.

■

Example 4.6

Are the following a basis for $\mathbb{R}^2$?

a.) $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$

b.) $\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ -2 \end{bmatrix} \right\}$

c.) $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$

Solution: Recall that we need to check that the set of vectors 1. Is linearly independent and 2. Spans $\mathbb{R}^2$.

a.) Yes.

b.) No, the vectors are linearly dependent.

c.) No, even though the set spans $\mathbb{R}^2$, the set is linearly dependent.

■

Note: $\left\{ \boldsymbol{e}_1, \ldots, \boldsymbol{e}_n \right\}$ is called the **standard basis** for $\mathbb{R}^n$.

Example 4.7

Let $V = \text{Span} \left\{ \overset{\boldsymbol{u}}{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}}, \overset{\boldsymbol{v}}{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}}, \overset{\boldsymbol{w}}{\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}} \right\}$ Find a basis for V.

Solution:

Let $A = \begin{bmatrix} \boldsymbol{u}\ \boldsymbol{v}\ \boldsymbol{w} \end{bmatrix}$. The RREF of A shows us the dependence relation between the columns of A.

$$A \sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

We see that col 1 + col 2 = col 3. To find a basis for V, since $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{w}$ but $\boldsymbol{u}$, $\boldsymbol{v}$ are linearly independent, we can take out $\boldsymbol{w}$ and take the basis to be $\left\{ \boldsymbol{u}, \boldsymbol{v} \right\} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\}$. ■

Essentially, to find a basis for a set of vectors, take their span and start removing "redundant" vectors – vectors that are linear combinations of the other vectors which when removed, result in the remaining vectors still spanning the same space.

Another way to put this is to *keep only the pivot columns.*

Example 4.8

Let $A = \begin{bmatrix} 3 & -1 & 0 & 3 \\ 6 & -2 & 3 & 9 \\ 3 & -1 & 2 & 5 \end{bmatrix}$.

a.) Find a basis for $\operatorname{Nul} A$.

b.) Find a basis for $\operatorname{Col} A$.

c.) Find a basis for $\operatorname{Row} A$.

Solution: A has RREF $= \begin{bmatrix} 1 & -1/3 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

a.) $A\boldsymbol{x} = \mathbf{0}$ has solution

$$\begin{aligned} \boldsymbol{x} &= \begin{bmatrix} \frac{1}{3}x_2 - x_4 \\ x_2 \\ -x_4 \\ x_4 \end{bmatrix} \\ &= x_2 \begin{bmatrix} 1/3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ 0 \\ -1 \\ 1 \end{bmatrix}. \end{aligned}$$

Thus, a basis for $\operatorname{Nul} A$ is $\left\{ \begin{bmatrix} 1/3 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ -1 \\ 1 \end{bmatrix} \right\}$.

Note: In general, to find a basis for $\operatorname{Nul} A$, write down the solution of $A\boldsymbol{x} = \mathbf{0}$ in *parametric vector form (PVF)*. The vectors that appear in the PVF are *always* a basis for $\operatorname{Nul} A$.

b.) We notice that col 1 + col 3 = col 4 and that col 2 = $-1/3$ col 1.

The RREF of A shows us that the columns of A are linearly dependent. I.e., we can write the 2nd and 4th columns as linear combinations of the 1st and 3rd columns. So, a basis for $\operatorname{Col} A$ is $\left\{ \begin{bmatrix} 3 \\ 6 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \\ 2 \end{bmatrix} \right\}$.

Note: In general, to find a basis for $\operatorname{Col} A$, write A in RREF form to see which columns are pivot columns. Then, take those columns from the original matrix A as your basis. While row operations preserve pivot columns, they do not preserve column space so you *must* use the original matrix A.

c.) *Approach:* A basis for $\operatorname{Row} A$ is *always* given by the set of nonzero rows in the RREF form of A (⚠ CONTRAST THIS WITH $\operatorname{Col} A$).

Thus, a basis for $\operatorname{Row} A$ is rows 1 and 2, i.e., $\left\{ \begin{bmatrix} 1 \\ -1/3 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \right\}$.

■

4.3 Dimension

The **dimension** of a vector space is essentially how many vectors are required to form a basis for that vector space.

Notation:

- $\dim \mathbb{R}^n = n \quad \Big(\text{since } \Big\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\Big\} \text{ is a basis for } R^n\Big)$.
- $\dim V = 2$ for $V = \text{Span}\left\{\begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}2\\1\\1\end{bmatrix}\right\}$ (since col 3 = col 1 + col 2).

General Summary: If A is an $m \times n$ matrix:

- $\dim \text{Nul}\, A = \#$ of free variables.
  $= \textbf{Nullity}\, A$
- $\dim \text{Col}\, A = \#$ of pivot columns of A
  $= \textbf{Rank}\, A$
  $= \dim \text{Row}\, A$
- $\dim \text{Col}\, A + \dim \text{Nul}\, A = \#$ of pivot columns $+ \#$ of free variables.
  $= \#$ of columns of A
  (this follows from and is equivalent to the Rank-Nullity theorem).

Example 4.9

What is the maximum possible rank of a 10×14 matrix?

Solution: Since max # of pivots $= \min \Big\{\#$ of rows, # of cols $\Big\}$, this implies that max rank $= 10$ (rank ≤ 10). ■

Example 4.10

Suppose the solutions of a homogeneous system of 5 linear equations in 6 unknowns are all multiplies of one nonzero solution. Will the system have a solution for every possible choice of constants on the right hand side of the equations?

Solution: $A\boldsymbol{x} = \boldsymbol{0}$ has a nontrivial solution of the form $c\boldsymbol{v}$. So $\dim \text{Nul}\, A = 1$. But then by the Rank-Nullity theorem, $\dim \text{Nul}\, A + \dim \text{Col}\, A = 6 \implies \dim \text{Col}\, A = 5 \implies \text{Col}\, A = \mathbb{R}^5$. So that columns of A span $\mathbb{R}^5$ (since we have 5 pivot rows). So yes, $A\boldsymbol{x} = \boldsymbol{b}$ is consistent for every $\boldsymbol{b} \in \mathbb{R}^5$. ■

Example 4.11

Let $A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$. Find:

a.) $\operatorname{rank} A$

b.) $\dim \operatorname{Nul} A$

c.) $\dim \operatorname{Row} A$

Solution: $A \sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$.

a.) From the RREF of A, we see that there are two pivot columns. Thus $\operatorname{rank} A = 2$. Also note that a basis for $\operatorname{Col} A$ has 2 vectors: $\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\}$ and so $\dim \operatorname{Col} A = \operatorname{rank} A$.

b.) $A\boldsymbol{x} = \mathbf{0} \implies \boldsymbol{x} = \begin{bmatrix} -x_3 \\ -x_3 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$ We know from this that a basis for $\operatorname{Nul} A$ is $\left\{ \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \right\}$. Thus, $\dim \operatorname{Nul} A = 1$.

c.) From Rank-Nullity, since $\dim \operatorname{Row} A + \dim \operatorname{Nul} A = 3$, then $\dim \operatorname{Row} A = 2$.

■

4.4 Coordinates

Now we get to see why we learned about basis vectors.

Recall (The Unique Representation Theorem): Let $B = \left\{ \boldsymbol{v}_1, \ldots, \boldsymbol{v}_p \right\}$ be a basis for some vector space V. Then if $\boldsymbol{x}$ is in V, then there is *exactly one way* to write $\boldsymbol{x}$ as a linear combination of $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$ (otherwise, B would not be a linearly independent set).

Also recall that $\left\{ \boldsymbol{e}_1, \ldots, \boldsymbol{e}_n \right\}$ is the *standard basis* for $\mathbb{R}^n$.

Vocab/Notation:

1) Given $\boldsymbol{x} \in \mathbb{R}^n$ and if $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, then $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n$. The numbers $x_1, \ldots, x_n$ are called the **coordinates of $\boldsymbol{x}$ relative to the standard basis.**

2) Suppose $B = \left\{ \boldsymbol{b}_1, \ldots, \boldsymbol{b}_n \right\}$ is another basis for $\mathbb{R}^n$. Then by The Unique Representation Theorem, given any $\boldsymbol{x} \in \mathbb{R}^n$, we have $\boldsymbol{x} = c_1\boldsymbol{b}_1 + \cdots + c_n\boldsymbol{b}_n$ for some numbers $c_1, \ldots, c_n$. The numbers $c_1, \ldots, c_n$ are called the **coordinates of $\boldsymbol{x}$ relative to the basis** B.

Notation: $[\boldsymbol{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$

3) If $B = \left\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\right\}$ is a basis for $\mathbb{R}^n$, the matrix $P_B = \begin{bmatrix}\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\end{bmatrix}$ is called the **change of coordinate matrix from B to the standard basis**. The reason for the name is because $P_B[\boldsymbol{x}]_{\mathcal{B}} = \boldsymbol{x}$, so left multiplication by P_B tells you how to write a vector (written in coordinates relative to the basis B) with respect to the standard basis.

Ever since we first learned about graphing, we've thought in terms of the standard basis. The point $(1, 2)$ we generally think of as going across 1 (1 times $\boldsymbol{e}_1$) and up 2 (2 times $\boldsymbol{e}_2$). However, there are times and applications where it is more convenient to think in terms of a different basis.

Take for example, the basis $B = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ for $\mathbb{R}^2$ and the point $\boldsymbol{x} = (1, 2)$.

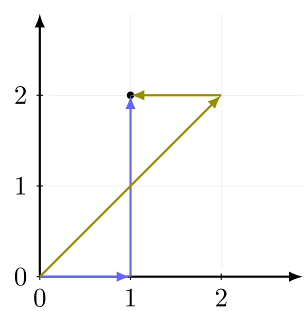


Figure 4.1: In blue, we take the standard basis to get to the point $(1, 2)$, which is how we usually think about it. In olive, we take a different route using the basis vectors from B. Note that we are still talking about the same point, but we are taking a different path to get there. Our worldview and way of navigating has shifted. To get to $(1, 2)$ using the basis vectors of B, note that we have to go 2 in the direction of the first basis vector and -1 in the direction of the second basis vector. Thus, $[\boldsymbol{x}]_B = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$.

Example 4.12

Let $B = \left\{ \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 5 \\ -6 \end{bmatrix} \right\}$. B is a basis for $\mathbb{R}^2$.

Let $\boldsymbol{x} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$. Find $[\boldsymbol{x}]_{\mathcal{B}}$, the coordinates of $\boldsymbol{x}$ relative to B.

Solution: We know from our definition above that $P_B[\boldsymbol{x}]_{\mathcal{B}} = \boldsymbol{x}$ where $P_B = \begin{bmatrix}\boldsymbol{b}_1, \boldsymbol{b}_2\end{bmatrix} = \begin{bmatrix} 1 & 5 \\ -2 & -6 \end{bmatrix}$. We have two approaches:

a.) Row reduce $\begin{bmatrix}\boldsymbol{b}_1 \ \boldsymbol{b}_2 \mid \boldsymbol{x}\end{bmatrix}$.

b.) $P_B[\boldsymbol{x}]_{\mathcal{B}} = \boldsymbol{x} \implies P_B^{-1}\boldsymbol{x} = [\boldsymbol{x}]_{\mathcal{B}}$.

Using our inverse formula for 2×2 matrices, we have that $P_B^{-1} = \frac{1}{4}\begin{bmatrix} -6 & -5 \\ 2 & 1 \end{bmatrix} \implies [\boldsymbol{x}]_{\mathcal{B}} = \begin{bmatrix} -29/4 \\ 9/4 \end{bmatrix}$.

■

Example 4.13

Let $B = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} \right\}$. B is a basis for $\mathbb{R}^3$.

a.) Find P_B, the change of coordinate matrix from B to the standard basis.

b.) If $[\boldsymbol{x}]_{\mathcal{B}} = \begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix}$, find $\boldsymbol{x}$.

Solution:

a.) $P_B = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 4 \\ 0 & 0 & 6 \end{bmatrix}$.

b.) $\boldsymbol{x} = P_B[\boldsymbol{x}]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 8 \\ 6 \end{bmatrix}$.

■

Remark: P_B^{-1} is the **change of coordinate matrix from the standard matrix to** B.

Problem: Let B and D be bases for $\mathbb{R}^n$. How do we find the change of coordinate matrix from B to D?

Let's work through the standard basis:

$$\begin{array}{ccccc} B & & \text{Standard basis} & & D \\ [\boldsymbol{x}]_{\mathcal{B}} & \xmapsto[\substack{\text{Multiply on} \\ \text{left by } P_B}]{} & \boldsymbol{x} & \xmapsto[\substack{\text{Multiply on} \\ \text{left by } P_D^{-1}}]{} & [\boldsymbol{x}]_{\mathcal{D}} \end{array}$$

Thus, the **change of coordinate matrix from** B **to** D is $(P_D)^{-1}P_B$.

Notation: $\mathcal{D} \overset{P}{\leftarrow} \mathcal{B}$

Example 4.14

Let $B = \left\{ \begin{bmatrix} 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 6 \\ 3 \end{bmatrix} \right\}$ and $D = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}$.

a.) Find $\mathcal{D} \overset{P}{\leftarrow} \mathcal{B}$.

b.) If $[\boldsymbol{x}]_{\mathcal{B}} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, find $[\boldsymbol{x}]_{\mathcal{D}}$.

Solution: From the discussion above:

$$\begin{aligned}\mathcal{D} \overset{P}{\leftarrow} \mathcal{B} &= P_D^{-1} P_B \\ &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 6 \\ 4 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 6 & 9 \\ 4 & 3 \end{bmatrix}.\end{aligned}$$

Thus, $[\boldsymbol{x}]_{\mathcal{D}} = \mathcal{D} \overset{P}{\leftarrow} \mathcal{B}[\boldsymbol{x}]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$.

Let's check our answers:

Both $P_B[\boldsymbol{x}]_{\mathcal{B}}$ and $P_D[\boldsymbol{x}]_{\mathcal{D}}$ should give us $\boldsymbol{x}$, the coordinates of $\boldsymbol{x}$ relative to the standard basis.

- $P_B[\boldsymbol{x}]_{\mathcal{B}} = \begin{bmatrix} 2 & 6 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} -2 \\ 5 \end{bmatrix} = \boldsymbol{x}$.
- $P_D[\boldsymbol{x}]_{\mathcal{D}} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 5 \end{bmatrix} = \begin{bmatrix} -2 \\ 5 \end{bmatrix} = \boldsymbol{x}$.

Additionally,

- $[\boldsymbol{x}]_{\mathcal{B}} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ means that $\boldsymbol{x} = 2\boldsymbol{b}_1 - \boldsymbol{b}_2 = \begin{bmatrix} -2 \\ 5 \end{bmatrix}$.
- $[\boldsymbol{x}]_{\mathcal{D}} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$ means that $\boldsymbol{x} = 3\boldsymbol{d}_1 + 5\boldsymbol{d}_2 = \begin{bmatrix} -2 \\ 5 \end{bmatrix}$.

■

Note: If V is a vector space with bases $B = \left\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\right\}$ and $D = \left\{\boldsymbol{d}_1, \ldots, \boldsymbol{d}_n\right\}$, then the first column of $\mathcal{D} \overset{P}{\leftarrow} \mathcal{B}$ tells us how to write $\boldsymbol{b}_1$ as a linear combination of $\boldsymbol{d}_1, \ldots, \boldsymbol{d}_n$, the second column of $\mathcal{D} \overset{P}{\leftarrow} \mathcal{B}$ tells us how to write $\boldsymbol{b}_2$ as a linear combination of $\boldsymbol{d}_1, \ldots, \boldsymbol{d}_n$, and so on.

Example 4.15

Let $B = \left\{\boldsymbol{b}_1, \boldsymbol{b}_2\right\}$ and $F = \left\{\boldsymbol{c}_1, \boldsymbol{c}_2\right\}$ be bases for V. If $\boldsymbol{b}_1 = 3\boldsymbol{c}_1 - 6\boldsymbol{c}_2$ and $\boldsymbol{b}_2 = 11\boldsymbol{c}_1 + 5\boldsymbol{c}_2$, find $\mathcal{F} \overset{P}{\leftarrow} \mathcal{B}$.

Solution: From our note above, $\mathcal{F} \overset{P}{\leftarrow} \mathcal{B} = \begin{bmatrix} 3 & 11 \\ -6 & 5 \end{bmatrix}$. ■

4.5 Additional Examples in $\mathbb{P}_n$ and $M_{m\times n}$

Notation:

- $\mathbb{P}_n = \left\{a_0 + a_1x + \cdots + a_nx^n : a_0, \ldots, a_n \in \mathbb{R}\right\}$ – the set of all polynomials of degree $\leq n$.

- $M_{m\times n}$, $\mathbb{R}^{m\times n}$ – the set of all $m \times n$ matrices.

Example 4.16

Is $B = \left\{1,\, 1-t,\, 2-4t+t^2,\, 6-18t+4t^2-t^3\right\}$ a basis for $\mathbb{P}_3$?

Polynomial	Coordinate vector in $\mathbb{R}^4$
1	$\begin{bmatrix}1\\0\\0\\0\end{bmatrix}$
$1-t$	$\begin{bmatrix}1\\-1\\0\\0\end{bmatrix}$
$2-4t+t^2$	$\begin{bmatrix}2\\-4\\1\\0\end{bmatrix}$
$6-18t+4t^2-t^3$	$\begin{bmatrix}6\\-18\\4\\-1\end{bmatrix}$

A consequence of Thm 4.8 is that B is a basis for

$\mathbb{P}_3 \iff \left\{\begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}1\\-1\\0\\0\end{bmatrix}, \begin{bmatrix}2\\-4\\1\\0\end{bmatrix}, \begin{bmatrix}6\\-18\\4\\-1\end{bmatrix}\right\}$ is a basis for $\mathbb{R}^4$.

Since $\begin{bmatrix}1&1&2&6\\0&-1&-4&-18\\0&0&1&4\\0&0&0&-1\end{bmatrix}$ has 4 pivots, then *yes*, B is a basis for $\mathbb{P}_3$.

Example 4.17

a.) Take $S =$ the set of 4×4 diagonal matrices from example 4.2. Find a basis for S. What is the dimension of S?

b.) A matrix M is called *skew-symmetric* if $M^T = -M$. Note that a skew-symmetric matrix is necessarily square. Show that the set of all skew-symmetric matrices, denoted W, is a subspace of $M_{n\times n}$. What is a basis for W? What is the dimension of W?

Solution:

a.) The general form for a vector in S is $\begin{bmatrix}a&0&0&0\\0&b&0&0\\0&0&c&0\\0&0&0&d\end{bmatrix}$ where a, b, c, d are real numbers. Then

$$\begin{bmatrix}a&0&0&0\\0&b&0&0\\0&0&c&0\\0&0&0&d\end{bmatrix} = a\begin{bmatrix}1&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} + b\begin{bmatrix}0&0&0&0\\0&1&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} + c\begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&1&0\\0&0&0&0\end{bmatrix} + d\begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&1\end{bmatrix}.$$

So, a basis for S is $\left\{\begin{bmatrix}1&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix}, \begin{bmatrix}0&0&0&0\\0&1&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix}, \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&1&0\\0&0&0&0\end{bmatrix}, \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&1\end{bmatrix}\right\}$ and S has dimension 4.

b.) (i.) $\mathbf{0}^T = -\mathbf{0}$ so $\mathbf{0} \in W$.

(ii.) Let $A, B \in W$. Note that for A and B to be in W, A and B must satisfy $A^T = -A$ and $B^T = -B$. Then,

$$\begin{aligned}(A+B)^T &= A^T + B^T \\ &= -A - B \\ &= -(A+B).\end{aligned}$$

Thus, $A + B \in W$.

(iii.) Let $k \in \mathbb{R}$. Then,

$$\begin{aligned}(kA)^T &= kA^T \\ &= -(kA).\end{aligned}$$

Thus, $kA \in W$.

Hence, W is a subspace of $M_{n\times n}$.

We will show an example basis for $W \subseteq M_{3\times 3}$ but it can easily be generalized for $W \subseteq M_{n\times n}$. Let $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$. For A to be skew-symmetric,

$$\begin{aligned} & A^T = -A \\ \implies & \begin{bmatrix} a & d & g \\ b & e & h \\ c & f & i \end{bmatrix} = \begin{bmatrix} -a & -b & -c \\ -d & -e & -f \\ -g & -h & -i \end{bmatrix}.\end{aligned}$$

For the diagonal entries, $a = -a \implies a = 0$. Similarly, $e = i = 0$. For the nondiagonal entries, we see that $a_{ij} = -a_{ji}$. For example, $b = -d$. Thus, we have a general form for A:

$$\begin{bmatrix} 0 & b & c \\ -b & 0 & f \\ -c & -f & 0 \end{bmatrix} = b\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} + f\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$$

Thus, a basis for W is $\left\{ \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} \right\}$ and W has dimension 3.

(This last bit is probably beyond the scope of this class in terms of difficulty, in the sense that you're not expected to know how to calculate the number of triangular entries).

Note that we only needed to know what the entries were on the upper triangular part of the matrix to determine an entire matrix in W, since the diagonals are all 0 and the lower triangular entries are symmetrical (and negative) with the upper triangular entries.

The number of entries in the upper triangular part with the diagonal is $\dfrac{n(n+1)}{2}$ so then the number of entries in the upper triangular part without the diagonal is $\dfrac{n(n+1)}{2} - n = \dfrac{n(n-1)}{2}$. Thus, the dimension of W when W is a subspace of $M_{n\times n}$ is $\dfrac{n(n-1)}{2}$.

■

Chapter 5

Eigenvalues and Eigenvectors

To motivate our study of eigenvalues and eigenvectors, let's look an an example:

Example 5.1

Let $T(\boldsymbol{x}) = \begin{bmatrix} 6 & 0 \\ 0 & -2 \end{bmatrix} \boldsymbol{x}$. Give a geometric description of the action of T on vectors in $\mathbb{R}^2$.

Is it possible to give a simple geometric description of $S : \mathbb{R}^2 \to \mathbb{R}^2$ where $S(\boldsymbol{x}) = \begin{bmatrix} 6 & 16 \\ -1 & -4 \end{bmatrix} \boldsymbol{x}$?

Solution: We have that $T(\boldsymbol{e}_1) = \begin{bmatrix} 6 \\ 0 \end{bmatrix}$, $T(\boldsymbol{e}_2) = \begin{bmatrix} 0 \\ -2 \end{bmatrix}$. So T stretches the $\boldsymbol{e}_1$ component of $\boldsymbol{x}$ by 6 and reflects and stretches the $\boldsymbol{e}_2$ component of $\boldsymbol{x}$ by 2.

For S, we can't explain with the standard basis. ■

Problem: Given a linear transformation T from $\mathbb{R}^n \to \mathbb{R}^n$, we want to find a basis where the action of T on those basis vectors is a simple stretch or reflection.

5.1 Eigenstuff

Definition 5.2

Let A be an $n \times n$ matrix. A *nonzero* vector $\boldsymbol{v}$ is called an **eigenvector** for A with **eigenvalue** λ if

$$A\boldsymbol{v} = \lambda \boldsymbol{v}.$$

So *eigenvectors* are vectors such that, when multiplied by A, merely undergo a scalar multiplication after the transformation, and the amount that the eigenvector is scaled by is called the *eigenvalue* associated with that eigenvector.

Example 5.3

Let $A = \begin{bmatrix} 3 & 2 \\ 3 & 8 \end{bmatrix}$. Is $\boldsymbol{u} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ or $\boldsymbol{v} = \begin{bmatrix} -4 \\ 2 \end{bmatrix}$ an eigenvector of A?

Solution:

- $A\boldsymbol{u} = \begin{bmatrix} 3 & 2 \\ 3 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -5 \end{bmatrix}$. $\boldsymbol{u}$ is not an eigenvector of A since $A\boldsymbol{u}$ is not a simple scaling of $\boldsymbol{u}$.
- $A\boldsymbol{v} = \begin{bmatrix} 3 & 2 \\ 3 & 8 \end{bmatrix} \begin{bmatrix} -4 \\ 2 \end{bmatrix} = \begin{bmatrix} -8 \\ 4 \end{bmatrix} = 2\boldsymbol{v}$. So $\boldsymbol{v}$ is an eigenvector of A with eigenvalue 2.

■

Okay, so given an $n \times n$ matrix A, how do we find its eigenvectors?

Well,

$$\begin{aligned} A\boldsymbol{v} = \lambda\boldsymbol{v} &\iff A\boldsymbol{v} - \lambda\boldsymbol{v} = \mathbf{0} \\ &\iff (A - \lambda I_n)\boldsymbol{v} = \mathbf{0}. \end{aligned}$$

(Note that we need to include the identity matrix I_n when we factor out $\boldsymbol{v}$ so that the matrices remain conformable). But, by definition, $\boldsymbol{v} \neq \mathbf{0}$, which implies that $A - \lambda I_n$ is *not invertible*. Thus,

$$\det(A - \lambda I_n) = 0.$$

So, to find the eigenvectors of A:

1. Set $\det(A - \lambda I_n) = 0$ and solve for λ.
2. For each eigenvalue from 1, solve $(A - \lambda I_n)\boldsymbol{v} = \mathbf{0}$ to find the eigenvectors for λ.

Note: The null space of $A - \lambda I_n$ is called the **eigenspace** corresponding to λ.

Example 5.4

Let $A = \begin{bmatrix} 6 & 16 \\ -1 & -4 \end{bmatrix}$.

a.) Find the eigenvalues of A.

b.) Find the bases for the eigenspaces of A.

Solution:

a.)

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} 6-\lambda & 16 \\ -1 & -4-\lambda \end{vmatrix} \\ &= (6-\lambda)(-1-\lambda) + 16 \\ &= \lambda^2 - 2\lambda - 8 \\ &= (\lambda - 4)(\lambda + 2) \end{aligned}$$

So, $\det(A - \lambda I) = 0 \iff (\lambda - 4)(\lambda + 2) = 0 \iff \lambda = 4, -2$. These are our eigenvalues.

b.) <u>For $\lambda_1 = 4$:</u>

We wish to solve $(A - 4I)\boldsymbol{v} = \mathbf{0}$, i.e., we need to solve $\begin{bmatrix} 2 & 16 \\ -1 & -8 \end{bmatrix} \boldsymbol{v} = \mathbf{0}$.

We have that $\begin{bmatrix} 2 & 16 \\ -1 & -8 \end{bmatrix} \sim \begin{bmatrix} 1 & 8 \\ 0 & 0 \end{bmatrix}$. So then the general solution is $\boldsymbol{v} = v_2 \begin{bmatrix} -8 \\ 1 \end{bmatrix}$.

So eigenvectors for $\lambda_1 = 4$ are all multiples of $\begin{bmatrix} -8 \\ 1 \end{bmatrix}$.

Thus, a basis for our eigenspace is $\left\{ \begin{bmatrix} -8 \\ 1 \end{bmatrix} \right\}$.

We can check our answer: $A \begin{bmatrix} -8 \\ 1 \end{bmatrix} = \begin{bmatrix} -32 \\ 4 \end{bmatrix} = 4 \begin{bmatrix} -8 \\ 1 \end{bmatrix}$, which is exactly what we'd expect from our definition of eigenvalues and eigenvectors.

<u>For $\lambda_2 = -2$:</u>

$$\begin{aligned} \text{Solve } (A + 2I)\boldsymbol{v} = \mathbf{0} &\implies \begin{bmatrix} 8 & 16 \\ -1 & -2 \end{bmatrix} \boldsymbol{v} = \mathbf{0} \\ &\implies \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \boldsymbol{v} = \mathbf{0}. \end{aligned}$$

Thus, our general solution is $\boldsymbol{v} = v_2 \begin{bmatrix} -2 \\ 1 \end{bmatrix}$ and a basis for our eigenspace is $\left\{ \begin{bmatrix} -2 \\ 1 \end{bmatrix} \right\}$.

■

Remark: This example combines a lot of concepts from previous chapters. Please make sure you take extra time to digest this example and justify every step we took in our solution.

Some definitions:

- $\det(A - \lambda I_n)$ is called the **characteristic polynomial** of A.
- $\det(A - \lambda I_n) = 0$ is called the **characteristic equation**.
- The **multiplicity** of an eigenvalue of A is the number of times that λ is a root of the characteristic polynomial.

Also note that 5.1 states that if A is triangular, then the eigenvalues of A are the entries on its main diagonal.

Example 5.5

Let $C = \begin{bmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 4 & 1 & 3 & 0 \\ 10 & 5 & 6 & 4 & 2 \end{bmatrix}$.

a.) What is the characteristic polynomial of C?

b.) List its eigenvalues and their multiplicities.

Solution:

a.) Recall that the determinant of a triangular matrix is the product of its entries along the main diagonal. Thus, the characteristic polynomial is

$$(3-\lambda)(-1-\lambda)(-\lambda)(3-\lambda)(2-\lambda).$$

b.)

Eigenvalue	Multiplicity
3	2
-1	1
0	1
2	1

■

Example 5.6

Let $A = \begin{bmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{bmatrix}$. Analyze the long term behavior of $x_{k+1} = Ax_k$ for $k = 0, 1, 2, \ldots$ with $x_0 = \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix}$.

Solution: The idea here is that since eigenvectors are vectors that experience a simple stretching when transformed by A, we can examine the long term behavior by taking the limit along these paths.

$$\begin{aligned} \det(A-\lambda I) &= \lambda^2 - 1.92\lambda + 0.92 \\ &= (\lambda-1)(\lambda-0.92) \\ &\implies \lambda = 1,\ 0.92 \end{aligned}$$

These eigenvalues correspond to $\boldsymbol{v}_1 = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$ and $\boldsymbol{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ respectively. Since our eigenvectors form a basis for $\mathbb{R}^2$, we can write x_0 in terms of eigenvectors.

So then $\boldsymbol{x}_0 = \begin{bmatrix} \boldsymbol{v}_1 \, \boldsymbol{v}_2 \end{bmatrix} \boldsymbol{c}$ for weights c_i, to be determined:

$$\begin{aligned} \boldsymbol{c} &= \begin{bmatrix} \boldsymbol{v}_1 \, \boldsymbol{v}_2 \end{bmatrix}^{-1} \boldsymbol{x}_0 \\ &= \begin{bmatrix} 0.125 \\ 0.225 \end{bmatrix}. \end{aligned}$$

So then we have

$$\begin{aligned} \boldsymbol{x}_1 &= A\boldsymbol{x}_0 \\ &= A\left(\begin{bmatrix} \boldsymbol{v}_1 \, \boldsymbol{v}_2 \end{bmatrix} \boldsymbol{c}\right) \\ &= c_1 A\boldsymbol{v}_1 + c_2 A\boldsymbol{v}_2 \\ &= c_1 \boldsymbol{v}_1 + c_2 0.92 \boldsymbol{v}_2 && \text{(since } v_i\text{'s are eigenvectors of } A\text{)} \end{aligned}$$

and

$$\begin{aligned} \boldsymbol{x}_2 &= A\boldsymbol{x}_1 \\ &= c_1 A\boldsymbol{v}_1 + c_2(0.92)A\boldsymbol{v}_2 \\ &= c_1\boldsymbol{v}_1 + c_2(0.92)^2\boldsymbol{v}_2 \end{aligned}$$

In general,

$$\boldsymbol{x}_k = c_1\boldsymbol{v}_1 + c_2(0.92)^k\boldsymbol{v}_2.$$

So the $\boldsymbol{v}_2$ term goes to 0 as $k \to \infty$.

Hence, $\boldsymbol{x}_k$ tends to $c\boldsymbol{v}_1 = \begin{bmatrix} 0.375 \\ 0.625 \end{bmatrix}$. ■

This next example is cool:

Example 5.7

Let A be a 2×2 matrix.
Let $\tau = \operatorname{trace}(A) =$ the sum of the diagonal components of A.
Let $\Delta = \det A$.
Find a nice clean formula for the characteristic equation and then use that to find the eigenvalues.

Solution: Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. So then $\tau = a + d$ and $\Delta = ad - bc$.

$$\begin{aligned}
\det(A - \lambda I) &= \begin{vmatrix} a-\lambda & b \\ c & d-\lambda \end{vmatrix} \\
&= (a-\lambda)(d-\lambda) - bc \\
&= \lambda^2 - a\lambda - d\lambda + ad - bc \\
&= \lambda^2 - (a+d)\lambda + (ad - bc) \\
&\boxed{0 = \lambda^2 - \tau\lambda + \Delta}
\end{aligned}$$

Solving for λ,

$$\boxed{\lambda = \frac{\tau \pm \sqrt{\tau^2 - 4\Delta}}{2}}$$

■

5.2 Diagonalizable Alley

Some definitions and theorems to get this section started:

Two matrices A and B are called **similar** if there is some invertible matrix P so that $A = PBP^{-1}$.

Thm 5.4: If A and B are similar, then A and B have the same eigenvalues.

Definition 5.8

A square matrix is called **diagonalizable** if it is *similar* to a diagonal matrix, i.e., A is diagonalizable $\iff A = PDP^{-1}$ for some diagonal matrix D.

Theorem 5.9: (Numbering doesn't align with the handout in the appendix)

An $n \times n$ matrix A is diagonalizable $\iff$ A has a set of n linearly independent eigenvectors $\left\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\right\}$. In this case, $A = PDP^{-1}$ where $P = \begin{bmatrix}\boldsymbol{v}_1 \, \boldsymbol{v}_2 \, \ldots \, \boldsymbol{v}_n\end{bmatrix}$ and $D = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}$ where $A\boldsymbol{v}_i = \lambda_i \boldsymbol{v}_i$.

So to diagonalize A, build P by placing all the eigenvectors of A into a matrix so that D contains the corresponding eigenvalues to those eigenvectors in the same order.

Note:

- P and D are *not* unique because we can reorder columns of P, multiply it by scalars, etc. But once you fix P, then D is determined.
- A is diagonalizable $\iff$ $\mathbb{R}^n$ has a basis consisting of eigenvectors of A.

  A is diagonalizable $\iff$ the dimensions of the eigenspaces of A add up to n.

Example 5.10

Are the following matrices diagonalizable?

$A = \begin{bmatrix} 3 & -9 \\ -2 & 6 \end{bmatrix}$ $\qquad B = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ $\qquad C = \begin{bmatrix} 5 & 1 \\ 0 & 5 \end{bmatrix}$

Solution: Let's do A and C first (confirm on your own that you get the same eigenvalues/eigenvectors):

a.)

Eigenvalue	Eigenspace	Dimension of eigenspace
0	Span $\left\{\begin{bmatrix} 3 \\ 1 \end{bmatrix}\right\}$	1
9	Span $\left\{\begin{bmatrix} -3 \\ 2 \end{bmatrix}\right\}$	1

Since the sum of the dimensions of the eigenspaces is 2, which is the number of columns of A, then A is diagonalizable.

One possible way to diagonalize A is $A = PDP^{-1}$ where $P = \begin{bmatrix} 3 & -3 \\ 1 & 2 \end{bmatrix}$ and $D = \begin{bmatrix} 0 & 0 \\ 0 & 9 \end{bmatrix}$.

c.)

Eigenvalue	Eigenspace	Dimension of eigenspace
5	Span $\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\}$	1

C is not diagonalizable since the sum of the dimensions of the eigenspaces is less than the number of columns of C.

b.) First, we need to find the eigenvalues and eigenvectors of B.

$$\begin{aligned}
\det(B-\lambda I) &= \begin{vmatrix} -\lambda & 1 & 1 \\ 1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{vmatrix} \\
&= -\lambda \begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} - \begin{vmatrix} 1 & 1 \\ 1 & -\lambda \end{vmatrix} + \begin{vmatrix} 1 & -\lambda \\ 1 & 1 \end{vmatrix} \\
&= -\lambda(\lambda^2-1) - (-\lambda-1) + (1+\lambda) \\
&\vdots \\
&= -(\lambda+1)(\lambda-2)(\lambda+1) \\
\implies \lambda &= -1,\, 2.
\end{aligned}$$

<u>For $\lambda_1 = -1$:</u>

$$\begin{aligned}
(B-(-1)I)\boldsymbol{v} = \mathbf{0} &\implies \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \boldsymbol{v} = \mathbf{0} \\
&\implies \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \boldsymbol{v} = \mathbf{0} \\
&\implies \boldsymbol{v} = v_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + v_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.
\end{aligned}$$

So then the eigenspace is Span $\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$.

<u>Similarly, for $\lambda_2 = 2$,</u> $\boldsymbol{v} = v_3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ so the eigenspace is Span $\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$.

Eigenvalue	Eigenspace	Dimension of eigenspace
-1	Span $\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$	2
2	Span $\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$	1

B is diagonalizable because the sum of the dimensions of the eigenspaces is 3, which is the number of columns of B.

Hence, we could have $B = PDP^{-1}$ where $P = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$ and $D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$.

∎

Tying up some loose ends with some theorems:

- Eigenvectors corresponding to distinct eigenvalues are linearly independent.
- If an $n \times n$ matrix A has n distinct eigenvalues, then A is diagonalizable.

  Note: The converse is *not* true (see B in the above example).

- If A is $n \times n$, then $\det A$ is the product of the eigenvalues of A with their multiplicities.
- The sum of the diagonal entries of A is the sum of the eigenvalues of A with their multiplicities.

Example 5.11

$B = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$. Then $\lambda = -1, 2$ where -1 has multiplicity 2.
So then the sum of the diagonal entries of $B = 0 = -1(2) + 2(1)$.
And $\det B = (-1)(-1)(2) = 2$.

Example 5.12

True or false?

a.) If A is diagonalizable, then it is invertible.

b.) If A is invertible, then it is diagonalizable.

Solution:

a.) False. $A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ is diagonalizable because it is already a diagonal matrix but A is not invertible.

b.) False. $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ is invertible since $\det A = 1 \neq 0$ but has only 1 linearly independent eigenvector.

■

Big picture idea: If $T : \mathbb{R}^n \to \mathbb{R}^n$, with $T(\boldsymbol{x}) = A(\boldsymbol{x})$, then T has a very simple geometric action on $\mathbb{R}^n \iff A$ is diagonalizable.

Chapter 6

Orthogonality and Least Squares

6.1 I think that issue is entirely orthogonal to the issue here because the Commonwealth is acknowledging...

To kick off this section let's define some terms. Let $\boldsymbol{u}$, $\boldsymbol{v} \in \mathbb{R}^n$.

- $\boldsymbol{u}$ and $\boldsymbol{v}$ are called **orthogonal** if $\boldsymbol{u} \cdot \boldsymbol{v} = 0$.

  $\boldsymbol{u}$, $\boldsymbol{v}$ orthogonal means that the vectors form a right angle.

- The **magnitude** of $\boldsymbol{u}$ is its length and we write $||\boldsymbol{u}||$ to mean magnitude. We have that $||\boldsymbol{u}|| = \sqrt{\boldsymbol{u} \cdot \boldsymbol{u}}$.

- $\operatorname{dist}(\boldsymbol{u}, \boldsymbol{v})$ is the **distance** between $\boldsymbol{u}$ and $\boldsymbol{v}$ and is equal to $||\boldsymbol{u} - \boldsymbol{v}||$ or $||\boldsymbol{v} - \boldsymbol{u}||$.

Example 6.1

Let $\boldsymbol{u} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ and $\boldsymbol{v} = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$.

a.) Are $\boldsymbol{u}$ and $\boldsymbol{v}$ orthogonal?

b.) What is the distance between $\boldsymbol{u}$ and $\boldsymbol{v}$?

Solution:

a.) $\boldsymbol{u} \cdot \boldsymbol{v} = -4 + 12 = 8 \neq 0$. So, $\boldsymbol{u}$ and $\boldsymbol{v}$ are *not* orthogonal.

b.) $||\boldsymbol{u} - \boldsymbol{v}|| = \left|\left| \begin{bmatrix} 4 \\ -1 \end{bmatrix} \right|\right| = \boxed{\sqrt{17}.}$

■

So why do we care about orthogonality? Because we can use it to find projections of vectors. The projection of a vector onto a line, plane, etc,... is the point on the line, plane, etc,... that is closest to the vector.

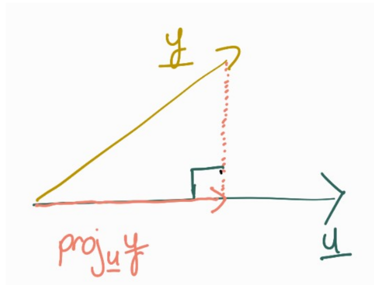


Figure 6.1: The projection of $\boldsymbol{y}$ onto the vector $\boldsymbol{u}$ can be seen as the portion of $\boldsymbol{y}$ in the direction of $\boldsymbol{u}$ and is given by what its shadow would be if a light source were directly above $\boldsymbol{y}$.

Definition 6.2

For two vectors $\boldsymbol{y}$, $\boldsymbol{u} \in \mathbb{R}^n$,

$$\text{proj}_{\boldsymbol{u}}\boldsymbol{y} = \frac{\boldsymbol{y}\cdot\boldsymbol{u}}{||\boldsymbol{u}||^2}\boldsymbol{u} = \frac{\boldsymbol{y}\cdot\boldsymbol{u}}{\boldsymbol{u}\cdot\boldsymbol{u}}\boldsymbol{u}$$

Read as the projection of the vector $\boldsymbol{y}$ onto the vector $\boldsymbol{u}$. In other words, the component (or the amount of $\boldsymbol{y}$) that is in the direction of $\boldsymbol{u}$.

Example 6.3

Let $\boldsymbol{y} = \begin{bmatrix}1\\2\end{bmatrix}$ and $\boldsymbol{u} = \begin{bmatrix}-6\\2\end{bmatrix}$. Find a point on Span $\{\boldsymbol{u}\}$ closest to $\boldsymbol{y}$.

Solution: We are looking for the projection of $\boldsymbol{y}$ onto the line that $\boldsymbol{u}$ spans. Since $\boldsymbol{y}\cdot\boldsymbol{u} = -2$ and $||\boldsymbol{u}||^2 = 40$, then $\text{proj}_{\boldsymbol{u}}\boldsymbol{y} = -\frac{1}{20}\begin{bmatrix}-6\\2\end{bmatrix} = \begin{bmatrix}3/10\\-1/10\end{bmatrix}$ is the point on Span $\{\boldsymbol{u}\}$ closest to $\boldsymbol{y}$. ■

We can find also projections onto planes (and higher dimensions). Let $S = \left\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\right\}$ where $\boldsymbol{v}_i \in \mathbb{R}^n$. Then S is **orthogonal** if $\boldsymbol{v}_i \cdot \boldsymbol{v}_j = 0$ for $i \neq j$.

If a basis for a vector space is orthogonal, then we call it an **orthogonal basis**.

Suppose W is a subspace of $\mathbb{R}^n$. Then the **orthogonal complement** of W, denoted $W^\perp$, is the set of all vectors which are orthogonal to W. I.e., $W^\perp = \left\{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{w}\cdot\boldsymbol{x} = 0 \text{ for all } \boldsymbol{w} \in W\right\}$.

An **orthonormal basis** is an orthogonal basis where each vector is a unit vector (i.e., $||\boldsymbol{v}_i|| = 1$ for all $i \leq p$).

Definition 6.4

Let $W \subseteq \mathbb{R}^n$ and $\left\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p\right\}$ be an orthogonal basis for W. Then

$$\text{proj}_W\boldsymbol{y} = \frac{\boldsymbol{y}\cdot\boldsymbol{u}_1}{\boldsymbol{u}_1\cdot\boldsymbol{u}_1}\boldsymbol{u}_1 + \cdots + \frac{\boldsymbol{y}\cdot\boldsymbol{u}_p}{\boldsymbol{u}_p\cdot\boldsymbol{u}_p}\boldsymbol{u}_p$$

Also,

$$\boldsymbol{y} = \text{proj}_W\boldsymbol{y} + \text{proj}_{W^\perp}\boldsymbol{y}$$

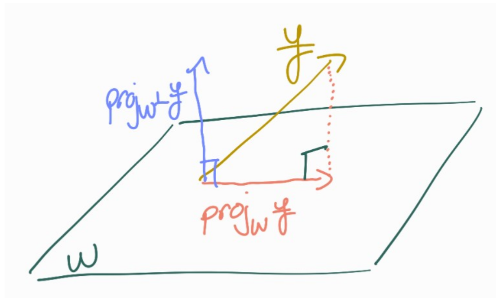


Figure 6.2: The projection of $\boldsymbol{y}$ onto W. Note that $\boldsymbol{y}$ can be decomposed into its **orthogonal components** such that $\boldsymbol{y} = \text{proj}_W \boldsymbol{y} + \text{proj}_{W^\perp} \boldsymbol{y}$.

Example 6.5

Let $W = \text{Span}\left\{ \begin{array}{c} \boldsymbol{u}_1 \\ \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \end{array}, \begin{array}{c} \boldsymbol{u}_2 \\ \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \end{array} \right\}$. Let $\boldsymbol{y} = \begin{bmatrix} 5 \\ 1 \\ 3 \end{bmatrix}$.

Find $\text{proj}_W \boldsymbol{y}$, i.e., find the orthogonal projection of $\boldsymbol{y}$ onto W.

Solution: First, we should check that $\boldsymbol{u}_1$, $\boldsymbol{u}_2$ are orthogonal:

$$\boldsymbol{u}_1 \cdot \boldsymbol{u}_2 = 1 + 1 - 2 = 0.$$

We have, $\boldsymbol{y} \cdot \boldsymbol{u}_1 = 3$, $\boldsymbol{y} \cdot \boldsymbol{u}_2 = 12$, $||\boldsymbol{u}_1||^2 = 3$, $||\boldsymbol{u}_2||^2 = 6$.

So then $\text{proj}_W \boldsymbol{y} = \frac{3}{3} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} + \frac{12}{6} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}$. ∎

In the previous example, we relied on the fact that we knew an orthogonal basis for W. But what if W isn't orthogonal?

We can make it orthogonal! (Gram-Schmidt)

Example 6.6: A slight detour

Let $\boldsymbol{v}_1 = \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}$, $\boldsymbol{v}_2 = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$, and $\boldsymbol{v}_3 = \begin{bmatrix} 2 \\ 0 \\ 4 \end{bmatrix}$.

a.) Show that the vectors form an orthogonal set but not an orthonormal set.

b.) Make them an orthonormal set.

Solution:

$$\text{a.)} \quad \begin{array}{r|l} \boldsymbol{v}_1 \cdot \boldsymbol{v}_2 = 0 & ||\boldsymbol{v}_1|| = \sqrt{5} \\ \boldsymbol{v}_1 \cdot \boldsymbol{v}_3 = 0 & ||\boldsymbol{v}_2|| = 1 \\ \boldsymbol{v}_2 \cdot \boldsymbol{v}_3 = 0 & ||\boldsymbol{v}_3|| = 2\sqrt{5} \end{array}$$

All pairwise dot products are zero, so the vectors form an orthogonal set. But, since the magnitudes of the vectors are not one, then they do not form an orthonormal set.

b.) To make the set orthonormal, we just need all magnitudes to be one.

Let $\boldsymbol{u}_1 = \dfrac{1}{||\boldsymbol{v}_1||}\boldsymbol{v}_1 = \dfrac{1}{\sqrt{5}}\begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}$

$\boldsymbol{u}_2 = \dfrac{1}{||\boldsymbol{v}_2||}\boldsymbol{v}_2 = \boldsymbol{v}_2$

$\boldsymbol{u}_3 = \dfrac{1}{||\boldsymbol{v}_3||}\boldsymbol{v}_3 = \dfrac{1}{2\sqrt{5}}\begin{bmatrix} 2 \\ 0 \\ 4 \end{bmatrix}.$

So $\left\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\right\}$ forms an orthonormal set.

■

6.2 Gram-Schmidt

Problem: Given a set of vectors $\left\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\right\}$ which is a basis for a vector space V but which do not form an orthogonal set, how do we construct an orthogonal basis for V?

Answer: Use Gram-Schmidt.

The Gram-Schmidt Process: Suppose V is a finite dimensional inner product space (don't need to worry about what that means for this class) and that $\left\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\right\}$ is a basis for V. Then an orthogonal basis for V, $\left\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\right\}$, can be found using the following process:

1.) $\boldsymbol{u}_1 = \boldsymbol{v}_1$.

2.) $\boldsymbol{u}_2 = \boldsymbol{v}_2 - \dfrac{\boldsymbol{v}_2 \cdot \boldsymbol{u}_1}{||\boldsymbol{u}_1||^2}\boldsymbol{u}_1$.

3.) $\boldsymbol{u}_3 = \boldsymbol{v}_3 - \dfrac{\boldsymbol{v}_3 \cdot \boldsymbol{u}_1}{||\boldsymbol{u}_1||^2}\boldsymbol{u}_1 - \dfrac{\boldsymbol{v}_3 \cdot \boldsymbol{u}_2}{||\boldsymbol{u}_2||^2}\boldsymbol{u}_2$.

$\vdots$

p.) $\boldsymbol{u}_p = \boldsymbol{v}_p - \dfrac{\boldsymbol{v}_p \cdot \boldsymbol{u}_1}{||\boldsymbol{u}_1||^2}\boldsymbol{u}_1 - \cdots - \dfrac{\boldsymbol{v}_p \cdot \boldsymbol{u}_{p-1}}{||\boldsymbol{u}_{p-1}||^2}\boldsymbol{u}_{p-1}$.

Then $\left\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p\right\}$ forms an orthogonal basis for V.

Example 6.7

Let $\boldsymbol{v}_1 = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}$, $\boldsymbol{v}_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$, $\boldsymbol{v}_3 = \begin{bmatrix} 3 \\ 7 \\ -1 \end{bmatrix}$. Note that $\left\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\right\}$ is a basis for $\mathbb{R}^3$. Construct an orthogonal basis for $\mathbb{R}^3$.

Solution: We were told that the v_i's form a basis for $\mathbb{R}^3$, but if we weren't then we should verify that they do form a basis for $\mathbb{R}^3$ before proceeding.

Step 1: $\boldsymbol{u}_1 = \boldsymbol{v}_1 = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}$.

Step 2:
$$\begin{aligned}
\boldsymbol{u}_2 &= \boldsymbol{v}_2 - \frac{\boldsymbol{v}_2 \cdot \boldsymbol{u}_1}{||\boldsymbol{u}_1||^2}\boldsymbol{u}_1 \\
&= \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} - \frac{2}{5}\begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} 1/5 \\ 2/5 \\ -1 \end{bmatrix}.
\end{aligned}$$

Step 3:
$$\begin{aligned}
\boldsymbol{u}_3 &= \boldsymbol{v}_3 - \frac{\boldsymbol{v}_3 \cdot \boldsymbol{u}_1}{||\boldsymbol{u}_1||^2}\boldsymbol{u}_1 - \frac{\boldsymbol{v}_3 \cdot \boldsymbol{u}_2}{||\boldsymbol{v}_2||^2}\boldsymbol{u}_2 \\
&= \begin{bmatrix} 3 \\ 7 \\ -1 \end{bmatrix} - \left(-\frac{1}{5}\right)\begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} - \frac{22/5}{6/5}\begin{bmatrix} 1/5 \\ 2/5 \\ -1 \end{bmatrix} \\
&= \begin{bmatrix} 8/3 \\ 16/3 \\ 8/3 \end{bmatrix}.
\end{aligned}$$

So then the orthogonal basis we've constructed is

$$\left\{ \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1/5 \\ 2/5 \\ -1 \end{bmatrix}, \begin{bmatrix} 8/3 \\ 16/3 \\ 8/3 \end{bmatrix} \right\}$$

Note: You can check your work by confirming that the above is mutually orthogonal.

We can make this an orthonormal basis by dividing each vector by its magnitude, i.e., let $\boldsymbol{w}_i = \dfrac{\boldsymbol{u}_i}{||\boldsymbol{u}_i||^2}$ for $i = 1, 2, 3$. ■

Example 6.8

Expand $\boldsymbol{v}_1 = \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}$, $\boldsymbol{v}_2 = \begin{bmatrix} 2 \\ 0 \\ 4 \end{bmatrix}$ into an orthogonal basis for $\mathbb{R}^3$.

Solution: Note that $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are already orthogonal (if they weren't then you could Gram-Schmidt the shit outta that bitch). So we just need to find a vector that is orthogonal to $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$.

Notice that the second entry of each vector is 0. So any vector with a nonzero 2nd entry should be outside the span of $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$.

Although it might be easy to see what an orthogonal vector would be, for funzies, say you can't easily find it upon inspection.

Let $\boldsymbol{v}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$. This is our initial "guess" which we will make orthogonal via Gram-Schmidt. Since $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are already orthogonal, then Gram-Schmidt won't change them. So let $\boldsymbol{u}_1 = \boldsymbol{v}_1$ and $\boldsymbol{u}_2 = \boldsymbol{v}_2$.

So we just need to find $\boldsymbol{u}_3$.

$$\begin{aligned} \boldsymbol{u}_3 &= \boldsymbol{v}_3 - \frac{\boldsymbol{v}_3 \cdot \boldsymbol{u}_1}{||\boldsymbol{u}_1||^2}\boldsymbol{u}_1 - \frac{\boldsymbol{v}_3 \cdot \boldsymbol{u}_2}{||\boldsymbol{v}_2||^2}\boldsymbol{u}_2 \\ &= \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} - \left(-\frac{1}{5}\right)\begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix} - \frac{4}{20}\begin{bmatrix} 2 \\ 0 \\ 4 \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}. \end{aligned}$$

Thus, $\left\{ \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$ forms an orthogonal basis for $\mathbb{R}^3$. ■

Example 6.9

Let $W = \text{Span}\left\{ \begin{bmatrix} 2 \\ 0 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$. Let $\boldsymbol{y} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$. Find $\text{proj}_W \boldsymbol{y}$. Check your work by finding an orthogonal basis for $W^\perp$ and showing that $\boldsymbol{y} = \text{proj}_W \boldsymbol{y} + \text{proj}_{W^\perp} \boldsymbol{y}$.

Solution: Notice that $\begin{bmatrix} 2 \\ 0 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ are not orthogonal. Use Gram-Schmidt to find orthogonal basis vectors for W (otherwise, our formula for $\text{proj}_W \boldsymbol{y}$ would not apply).

$$\begin{aligned} \boldsymbol{u}_1 &= \begin{bmatrix} 2 \\ 0 \\ 4 \end{bmatrix}. \\ \boldsymbol{u}_2 &= \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \frac{4}{20}\begin{bmatrix} 2 \\ 0 \\ 4 \end{bmatrix} \\ &= \begin{bmatrix} -2/5 \\ 1 \\ 1/5 \end{bmatrix}. \end{aligned}$$

So then,

$$\begin{aligned}
\text{proj}_W \boldsymbol{y} &= \frac{\boldsymbol{y}\cdot\boldsymbol{u}_1}{||\boldsymbol{u}_1||^2}\boldsymbol{u}_1 + \frac{\boldsymbol{y}\cdot\boldsymbol{u}_2}{||\boldsymbol{u}_2||^2}\boldsymbol{u}_2 \\
&= \frac{1}{10}\begin{bmatrix}2\\0\\4\end{bmatrix} + \frac{1}{2}\begin{bmatrix}-2/5\\1\\1/5\end{bmatrix} \\
&= \begin{bmatrix}0\\1/2\\1/2\end{bmatrix}.
\end{aligned}$$

To find an orthogonal basis for $W^\perp$, we need to find all vectors $\boldsymbol{x}$ such that $\boldsymbol{w}\cdot\boldsymbol{x} = 0$ for all $\boldsymbol{w} \in W$. From ex 1.10, we saw that we only needed to check that the basis vectors map to 0:

$$\begin{aligned}
\left[\begin{array}{ccc|c}2&0&4&0\\0&1&1&0\end{array}\right] &\sim \left[\begin{array}{ccc|c}1&0&2&0\\0&1&1&0\end{array}\right] \\
\implies \boldsymbol{x} &= x_3\begin{bmatrix}-2\\-1\\1\end{bmatrix}.
\end{aligned}$$

Thus, the set of all vectors that satisfy $W^T\boldsymbol{x} = 0$ is given by $W^\perp = \text{Span}\left\{\begin{bmatrix}-2\\-1\\1\end{bmatrix}\right\}$. Therefore, a basis for $W^\perp$ is $\left\{\overset{\boldsymbol{w}}{\begin{bmatrix}-2\\-1\\1\end{bmatrix}}\right\}$. So then,

$$\begin{aligned}
\text{proj}_{W^\perp}\boldsymbol{y} &= \frac{\boldsymbol{y}\cdot\boldsymbol{w}}{\boldsymbol{w}\cdot\boldsymbol{w}}\boldsymbol{w} \\
&= -\frac{1}{2}\begin{bmatrix}-2\\-1\\1\end{bmatrix} \\
&= \begin{bmatrix}1\\1/2\\-1/2\end{bmatrix}
\end{aligned}$$

so that

$$\begin{aligned}
\text{proj}_W\boldsymbol{y} + \text{proj}_{W^\perp}\boldsymbol{y} &= \begin{bmatrix}0\\1/2\\1/2\end{bmatrix} + \begin{bmatrix}1\\1/2\\-1/2\end{bmatrix} \\
&= \begin{bmatrix}1\\1\\0\end{bmatrix} \\
&= \boldsymbol{y}
\end{aligned}$$

as desired. ■

6.3 Least Squares

Problem: What if $A\boldsymbol{x} = \boldsymbol{b}$ is inconsistent? We can try to find the next best thing; i.e., find $\hat{\boldsymbol{x}}$ so that $A\hat{\boldsymbol{x}}$ is as close as possible to $\boldsymbol{b}$. In other words, find $A\hat{\boldsymbol{x}}$ in $\text{Col}\,A$ which has a solution, $A\hat{\boldsymbol{x}} = \hat{\boldsymbol{b}}$.

Given W a subspace of $\mathbb{R}^n$ and $\boldsymbol{y} \in \mathbb{R}^n$, The Best Approximation Theorem tells us that:

- $\text{proj}_W \boldsymbol{y}$ is the closest point in W to $\boldsymbol{y}$.
- The distance from $\boldsymbol{y}$ to W is $||\boldsymbol{y} - \text{proj}_W \boldsymbol{y}||$.

Example 6.10

Let $W = \text{Span}\left\{ \begin{bmatrix} 2 \\ 0 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$ and $\boldsymbol{y} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$.

a.) Find the point in W which is closest to $\boldsymbol{y}$.

b.) Find the distance from $\boldsymbol{y}$ to W.

Solution:

a.) The point in W which is closest to $\boldsymbol{y}$ is given by $\text{proj}_W \boldsymbol{y}$ which we found in example 6.9 to be $\begin{bmatrix} 0 \\ 1/2 \\ 1/2 \end{bmatrix}$.

b.) The distance from $\boldsymbol{y}$ to W is given by,

$$\begin{aligned} ||\boldsymbol{y} - \text{proj}_W \boldsymbol{y}|| &= \left|\left| \begin{bmatrix} 1 \\ 1/2 \\ -1/2 \end{bmatrix} \right|\right| \\ &= \sqrt{3/2}. \end{aligned}$$

■

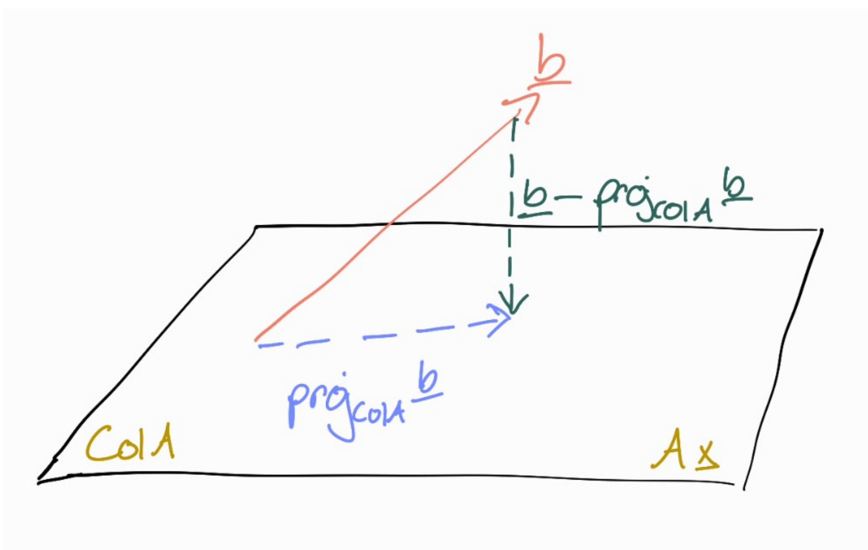


Figure 6.3: Here, we can see that $\boldsymbol{b}$ lies outside the column space of A. Thus, the system $A\boldsymbol{x} = \boldsymbol{b}$ is inconsistent. So, we want to find the point on $\text{Col}\, A$ that is as close as possible to $\boldsymbol{b}$, which we can denote $A\hat{\boldsymbol{x}}$. By The Best Approximation Theorem, we see that $\text{proj}_{\text{Col}\, A}\boldsymbol{b}$ is the point we're looking for. Also note then that, by definition 6.4, $\boldsymbol{b} - \text{proj}_{\text{Col}\, A}\boldsymbol{b} = \boldsymbol{b} - A\hat{\boldsymbol{x}}$ is *orthogonal* to $\text{Col}\, A$.

Since $\boldsymbol{b} - A\hat{\boldsymbol{x}}$ is orthogonal to $\operatorname{Col} A$, then $A^T(\boldsymbol{b} - A\hat{\boldsymbol{x}}) = \boldsymbol{0} \implies A^TA\hat{\boldsymbol{x}} = A^T\boldsymbol{b}$. This is known as the **normal equations** for $A\boldsymbol{x} = \boldsymbol{b}$.

We call $\hat{\boldsymbol{x}}$ a **least squares** solution since solving for the *normal equations* is effectively the same as minimizing the quantity $||\boldsymbol{b} - A\hat{\boldsymbol{x}}||^2$ (but you don't need to know how to do this optimization problem for this class — just make sure you understand figure 6.3).

To compute the least squares problem:

1.) Compute A^TA and $A^T\boldsymbol{b}$.

2.) Either
   (a) Form the augmented matrix for the system $A^TA\hat{\boldsymbol{x}} = A^T\boldsymbol{b}$ and row reduce or,
   (b) If A^TA is invertible, then find $(A^TA)^{-1}$ and calculate $\hat{\boldsymbol{x}} = (A^TA)^{-1}A^T\boldsymbol{b}$.

Example 6.11

Find the least squares solutions of $A\boldsymbol{x} = \boldsymbol{b}$ where $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix}$, $\boldsymbol{b} = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}$.

Solution: We have that

$$\begin{aligned} A^TA &= \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 5 & 3 \\ 3 & 3 \end{bmatrix}. \end{aligned}$$

and

$$\begin{aligned} A^T\boldsymbol{b} &= \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ 6 \end{bmatrix}. \end{aligned}$$

So then augment $A^TA\hat{\boldsymbol{x}} = A^T\boldsymbol{b}$:

$$\left[\begin{array}{cc|c} 5 & 3 & 0 \\ 3 & 3 & 6 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & -3 \\ 0 & 1 & 5 \end{array}\right]$$

Hence,

$$\hat{\boldsymbol{x}} = \begin{bmatrix} -3 \\ 5 \end{bmatrix}.$$

Note: We say least squares *solutions* because we can have infinitely many solutions if the columns of A are dependent (not full rank) which means that A^TA is not invertible. But, we are always guaranteed at least one least square solution. ■

Example 6.12

Suppose you collect some data: given the amount of math problems Khang solved that day, how many gummy bear snacks does Khang have left in his office to give to his best student? You randomly sampled 3 days and got points $(1, 0)$, $(6, 9)$, and $(4, 2)$. Find the best fit line to see if you can uncover an association.

Solution: The general equation for a line is given by $y = mx + b$. If all points fit on a single line, we would have:

$$\begin{cases} 1m + b &= 0 \\ 6m + b &= 9 \\ 4m + b &= 2 \end{cases} \longleftrightarrow \begin{bmatrix} 1 & 1 \\ 6 & 1 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 9 \\ 2 \end{bmatrix}$$

But, the system

$$\left[\begin{array}{cc|c} 1 & 1 & 0 \\ 6 & 1 & 9 \\ 4 & 1 & 2 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right]$$

is inconsistent. There does not exist m and b which gives us a line that fits all 3 points. But, let $A = \begin{bmatrix} 1 & 1 \\ 6 & 1 \\ 4 & 1 \end{bmatrix}$ and $\boldsymbol{b} = \begin{bmatrix} 0 \\ 9 \\ 2 \end{bmatrix}$. Then the least squares solution gives us:

$$\begin{aligned} \hat{\boldsymbol{x}} &= (A^T A)^{-1} A^T \boldsymbol{b} \\ &= \begin{bmatrix} 65/38 \\ -99/38 \end{bmatrix} \end{aligned}$$

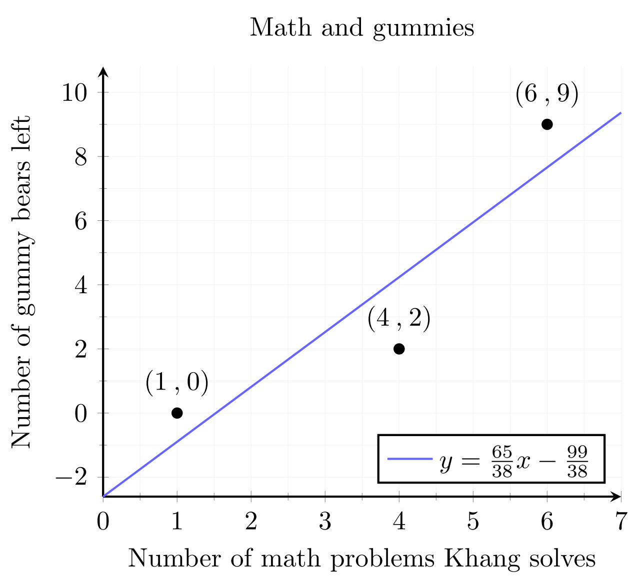


Figure 6.4: The best fit line for our data found via least squares.

■

Epilogue

Congratulations, you've made it through one of my favorite courses as an undergrad! I hope your journey was a beautiful one filled with a new-found appreciation for what math can be. Up until this point, all of your math courses have likely been geared towards just getting you ready for calculus but, there's a whole wide world of math out there and there's probably something for everyone.

If your AP calc teacher destroyed your love for math, I hope this course was the first step towards healing your soul.

In seriousness, if anything, and even if you never take another math course again, I hope that this course expanded your perspective and relationship with math. If you struggled, I hope that you improved your studenting skills to rise to the occasion. If you didn't do as well as you had hoped, I hope you learned something about how you can improve next quarter. And if nothing else, I hope you had some amount of fun and made a new friend in discussion section.

Above all, be proud of your accomplishments and be kind to yourself.

In this chapter I wanted to highlight some courses, both math and non-math you could take after MATH 18. I was an ESYS: Ecology, Behavior, and Evolution major for my first 2 years before switching into applied math. I also did a double minor in music and linguistics so I was able to take quite a few varied courses throughout my undergrad.

Whether you're continuing on your journey through math or are just looking for some interesting electives to take, here are some classes I took that were either deeply insightful or really fun.

Math

- MATH 120: Complex analysis

  Calculus but with complex numbers (learn why $e^{i\pi} + 1 = 0$ and how to plot functions with imaginary numbers).

- MATH 155: Geometric computer graphics

  The math behind how computers represent physical space, simulate lighting, draw curves, etc... The professor Sam Buss has consulted for Rockstar Games and helped create the physics engine used in GTA III. Particularly interesting if you're interested in 3D modeling (eg Blender).

  Your projects will look like professional late-80s cgi animations.

- MATH 181: Mathematical statistics

  Before this class I thought stats was boring. Now I'm in grad school for it. Highly recommend taking with David Quarfoot if possible!

Music

- MUS 137a-c: Jazz theory & improv

  Learn jazz from grammy award winning composer and producer Kamau Kenyatta. Spring quarter is entirely dedicated to recording a jazz album with the entire class.

Linguistics

- LIGN 7/8: Sign language and their cultures / languages and cultures in America

  Learning about our world and the people that inhabit it is a lifelong pursuit and these classes examine our cultural history through a linguistic lens. Also fulfills your DEI requirement.

- LIGN 119: First and second language learning: from childhood through adolescence

  Interesting class if you're interested in learning languages. A lecture towards the end includes real babies.

Misc

- POLI 160AA: Introduction to policy analysis

  Learn the many ways to divide a cake (this was very helpful for a hungry Khang at 8am).

- BIEB 166: Animal behavior & communication

  Busy bee, buzz buzz.

I'll leave you with three quotes from Dr. Larry Fleinhardt, a character from one of my favorite shows – Numb3rs:

- *"The magical element of water penetrates the impenetrable like the sleeping mind dreams the solution to a problem."*
- *"While being young is an accident of time, youth is a permanent state of mind."*
- *"You can contemplate silence, but you can never find it."*

Appendix A

Tips for Studying

I referred to David Quarfoot's Studenting 101 slides for some parts of this chapter.

Time Management

The most important asset you have is time. Whether it's time with a loved one, time in the stock market, or time invested in your education, effective use of your time is the difference between being able to balance a heavy course load with all the other social things which make college worthwhile and becoming overwhelmed and shutting down.

Figure out how many days you have left until the final and plan accordingly. It is better to spread 10 hours of studying over 10 days than it is to cram two 5 hour sessions right before.

Also equally important is how you will allocate time to reviewing various parts of the course.

Setting the Macro Picture and Planning your Study

Let's say you have 10 days before the final and plan on studying an hour each day. How do you know what to focus on for each session?

Before you begin delving deep into your review, I recommend taking a moment to create a broad list of topics for the whole course. Draw out a map if that's what works for you, mark connections, and fill in details to determine your weakest areas.

In general, for each concept, you should be able to (starting from shallow to deep knowledge):

- Define what it means and how to do any calculations related to it, if applicable.
- Understand why this concept *has* to be true and its importance/motivation. Ask yourself: why did we learn this? What can we do now with this knowledge that we couldn't before?
- Find connections across other topics in the course.

For example:

In this class we learned about null spaces. At first, we learned that the null space of a transformation is the set of all vectors that get mapped to $\mathbf{0}$. Then, we learned how to find the null space through some examples.

Diving deeper, we realized that a matrix A is invertible if and only if the dimension of the null space of A is 0. The null space of A also gives us information about whether the columns of A are linearly independent and about the value of its determinant.

An application of the null space we saw is that we used it to define what an eigenspace is so that we could find eigenvectors of A which we then used to determine if A was diagonalizable. We also found that if the null space of A includes only the zero vector, then $A^T A$ is invertible and so we can find a unique least-squares solution for $A\boldsymbol{x} = \boldsymbol{b}$.

As you can see, just going down the rabbit hole of one concept led us to find connections to every single chapter we went over. This only works if you have a solid understanding of concepts from this class and so doing this exercise will uncover areas where your conceptual foundations are weak which you can then target during your study sessions.

Once you have an idea of which topics you should spend your time on, how should you study those topics? The best way to get better at math is to do math so, keep the review portion as brief as necessary and do problems.

If your professor posts suggested practice problems, work on those first. If they think that the concepts presented in those problems are worth doing, then it's likely that they'll show up in some form on the final.

Also, rework key homework problems *without* context. Knowing which section a homework problem comes from can often give you hints on which concepts or techniques to apply but, on the final, you won't have that context.

Misc

Lastly, a few days before the final, attend to factors that affect performance:

- **Sleep:** I don't think it really needs to be said how important sleep is for your brain. Allow your subconscious mind to form connections and soak in your studying by giving it enough rest.
- **Nutrition:** A well-fed brain is a happy brain that's more willing to focus.
- **Mental health:** College is hard and so is life. Things understandably get in the way which can bring us down. Realize that one exam does not define who you are. Try to do what you can to attend to personal matters so that you can be at your best on exam day, but also be forgiving of yourself if things don't go as planned. Life is messy and inelegant and you're allowed to be imperfect.

  Side note: as cliché as it might sound, confidence does play a big role on exam day. Have faith that on exam day, you've done all that you could to prepare and do your best.

Appendix B

Theorem Handout

These theorems are ripped directly from our textbook for convenient referencing. *You by no means need to have these memorized!* (Many of these you'll end up internalizing anyways since they just state more explicitly concepts you should have a solid understanding of and because you use them so often, like Thm 1.4 or the IMT). As long as you are able to understand the theorems and how you might use them to answer a T/F problem or help with a computational problem then you'll be more than set.

Theorem 1.1. Each matrix is row equivalent to one and only one reduced echelon matrix.

Theorem 1.2 (Existence and Uniqueness)**.** A linear system is consistent if and only if the rightmost column of the augmented matrix is not a pivot column.

If a linear system is consistent, then the solution set contains either (i) a unique solution, when there are no free variables, or (ii) infinitely many solutions, when there is at least one free variable.

Theorem 1.3. If A is an $m \times n$ matrix, with columns $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$, and if $\boldsymbol{b}$ is in $\mathbb{R}^m$, the matrix equation

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

has the same solution set as the vector equation

$$x_1\boldsymbol{a}_1 + \cdots + x_n\boldsymbol{a}_n = \boldsymbol{b}$$

which, in turn, has the same solution set as the system of linear equations whose augmented matrix is

$$[\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n|\boldsymbol{b}].$$

Theorem 1.4. Let A be an $m \times n$ matrix. Then the following statements are logically equivalent. That is, for a particular A, either they are all true statements or they are all false.

a. For each $\boldsymbol{b} \in \mathbb{R}^m$, the equation $A\boldsymbol{x} = \boldsymbol{b}$ has a solution.

b. Each $\boldsymbol{b} \in \mathbb{R}^m$ is a linear combination of the columns of A.

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

d. A has a pivot in every row.

Theorem 1.5. If A is an $m \times n$ matrix, $\boldsymbol{u}$ and $\boldsymbol{v}$ are vectors in $\mathbb{R}^n$, and c is a scalar, then:

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

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

Theorem 1.6. Suppose the equation $A\boldsymbol{x} = \boldsymbol{b}$ is consistent for some $\boldsymbol{b}$, and let $\boldsymbol{p}$ be a solution. Then the solution set of $A\boldsymbol{x} = \boldsymbol{b}$ is the set of all vectors of the form $\boldsymbol{w} = \boldsymbol{p} + \boldsymbol{v}_h$ where $\boldsymbol{v}_h$ is any solution of the homogeneous equation $A\boldsymbol{x} = \boldsymbol{0}$.

Theorem 1.7 (Characterization of Linearly Dependent Sets)**.** An indexed set $S = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\}$ of two or more vectors is linearly dependent if and only if at least one of the vectors in S is a linear combination of the others. In fact, if S is linearly dependent and $\boldsymbol{v}_1 \neq \boldsymbol{0}$, then some $\boldsymbol{v}_j$ (with $j > 1$) is a linear combination of the preceding vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{j-1}$.

Theorem 1.8. If a set contains more vectors than there are entries in each vector, then the set is linearly dependent. That is, any set $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\}$ in $\mathbb{R}^n$ is linearly dependent if $p > n$.

Theorem 1.9. If a set $S = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\}$ in $\mathbb{R}^n$ contains the zero vector, then the set is linearly dependent.

Theorem 1.10. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then there exists a unique matrix A such that

$$T(\boldsymbol{x}) = A\boldsymbol{x} \quad \text{for all } \boldsymbol{x} \text{ in } \mathbb{R}^n.$$

In fact, A is the $m \times n$ matrix whose jth column is the vector $T(\boldsymbol{e}_j)$ where $\boldsymbol{e}_j$ is the jth column of the identity matrix in $\mathbb{R}^n$:

$$A = [T(\boldsymbol{e}_1), \ldots, T(\boldsymbol{e}_n)]$$

Theorem 1.11. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then T is one-to-one if and only if the equation $T(\boldsymbol{x}) = \boldsymbol{0}$ has only the trivial solution.

Theorem 1.12. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, and let A be the standard matrix for T. Then:

a. T maps $\mathbb{R}^n$ onto $\mathbb{R}^m$ if and only if the columns of A span $\mathbb{R}^m$.

b. T is one-to-one if and only if the columns of A are linearly independent.

Theorem 2.1. Let A, B, and C be matrices of the same size, and let r and s be scalars.

a. $A + B = B + A$

b. $(A + B) + C = A + (B + C)$

c. $A + 0 = A$

d. $r(A + B) = rA + rB$

e. $(r + s)A = rA + sA$

f. $r(sA) = (rs)A$

Theorem 2.2. 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. $A(BC) = (AB)C$ (associative law of multiplication)

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

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

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

e. $I_mA = A = AI_n$ (identity for matrix multiplication)

Theorem 2.3. 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. For any scalar r, $(rA)^T = rA^T$

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

Theorem 2.4. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. If $ad - bc \neq 0$, then A is invertible and

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

If $ad - bc = 0$, then A is not invertible.

Theorem 2.5. If A is an invertible $n \times n$ matrix, then for each $\boldsymbol{b} \in \mathbb{R}^n$, the equation $A\boldsymbol{x} = \boldsymbol{b}$ has the unique solution $\boldsymbol{x} = A^{-1}\boldsymbol{b}$.

Theorem 2.6.

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

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

b. If A and B are $n \times n$ invertible matrices, then so is AB.

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

c. If A is an invertible matrix, then so is A^T

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

Theorem 2.7. An $n \times n$ matrix A is invertible if and only if 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}.

Theorem 2.8 (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 such that $CA = I$.

k. There is an $n \times n$ matrix D such that $AD = I$.

l. A^T 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. $\dim \operatorname{Nul} A = 0$.

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

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

A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is said to be **invertible** if there exists a function $S : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$S(T(\boldsymbol{x})) = \boldsymbol{x} \quad \text{for all } \boldsymbol{x} \in \mathbb{R}^n \tag{B.1}$$

$$T(S(\boldsymbol{x})) = \boldsymbol{x} \quad \text{for all } \boldsymbol{x} \in \mathbb{R}^n \tag{B.2}$$

The next theorem shows that if such an S exists, it is unique and must be a linear transformation. We call S the **inverse** of T and write it as T^{-1}.

Theorem 2.9. 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 if and only if 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 B.1 and B.2.

Given $A = [a_{ij}]$, the (i, j)-**cofactor** of A is the number C_{ij} given by

$$C_{ij} = (-1)^{i+j} \det A_{ij} \tag{B.3}$$

Theorem 3.1. The determinant of an $n \times n$ matrix A can be computed by a cofactor expansion across any row or down any column. The expansion across the ith row using the cofactors in B.3 is

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

The cofactor expansion down the jth column is

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

Theorem 3.2. If A is a triangular matrix, then $\det A$ is the product of the entries on the main diagonal of A.

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 $\det B = k \cdot \det A$.

Theorem 3.4. A square matrix A is invertible if and only if $\det A \neq 0$.

Theorem 3.5. If A is an $n \times n$ matrix, then $\det A^T = \det A$.

Theorem 3.6 (Multiplicative Property)**.** If A and B are $n \times n$ matrices, then $\det AB = (\det A)(\det B)$.

Theorem 3.7 (Cramer's Rule)**.** Let A be an invertible $n \times n$ matrix. For any $\boldsymbol{b} \in \mathbb{R}^n$, the unique solution $\boldsymbol{x}$ of $A\boldsymbol{x} = \boldsymbol{b}$ has entries given by

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

Theorem 3.8 (An Inverse Formula)**.** Let A be an invertible $n \times n$ matrix. Then

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

Theorem 3.9. 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 3.10. 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

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

Theorem 4.1. If $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$ are in a vector space V, then $\operatorname{Span}\{\boldsymbol{v}_1, \ldots \boldsymbol{v}_p\}$ is a subspace of V.

Theorem 4.2. 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 of $A\boldsymbol{x} = \boldsymbol{0}$ of m homogeneous linear equations in n unknowns is a subspace of $\mathbb{R}^n$.

Theorem 4.3. The column space of an $m \times n$ matrix A is a subspace of $\mathbb{R}^m$.

Theorem 4.4. An indexed set $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\}$ of two or more vectors, with $\boldsymbol{v}_1 \neq \boldsymbol{0}$, is linearly dependent if and only if some $\boldsymbol{v}_j$ (with $j > 1$) is a linear combination of the preceding vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{j-1}$.

Theorem 4.5 (The Spanning Set Theorem)**.** Let $S = \{\boldsymbol{v}_1, \ldots \boldsymbol{v}_p\}$ be a set in V, and let $H = \operatorname{Span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\}$.

a. If one of the vectors in S – say, $\boldsymbol{v}_k$ – 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 4.6. The pivot columns of a matrix A form a basis for $\operatorname{Col} A$

Theorem 4.7 (The Unique Representation Theorem)**.** Let $\mathcal{B} = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$ be a basis for a vector space V. Then for each $\boldsymbol{x}$ in V, there exists a unique set of scalars $c_1, \ldots, c_n$ such that

$$\boldsymbol{x} = c_1\boldsymbol{b}_1 + \cdots + c_n\boldsymbol{b}_n$$

Theorem 4.8. 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 linear transformation from V onto $\mathbb{R}^n$.

Theorem 4.9. 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 4.10. If a vector space V has a basis of n vectors, then every basis of V must consist of exactly n vectors.

Theorem 4.11. 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 4.12 (The Basis Theorem)**.** Let V be a p-dimensional vector space, $p \geq 1$. Any linearly 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.

Theorem 4.13. 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.

Theorem 4.14 (Rank-Nullity)**.** 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, also equals the number of pivot positions in A and satisfies the equation

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

Theorem 4.15. Let $\mathcal{B} = \{\boldsymbol{b}_1, \dots, \boldsymbol{b}_n\}$ and $\mathcal{C} = \{\boldsymbol{c}_1, \dots, \boldsymbol{c}_n\}$ 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 $\underset{\mathcal{C} \leftarrow \mathcal{B}}{P}$ are the $\mathcal{C}$-coordinate vectors of the vectors in the basis $\mathcal{B}$. That is,

$$\underset{\mathcal{C} \leftarrow \mathcal{B}}{P} = \begin{bmatrix} [\boldsymbol{b}_1]_{\mathcal{C}} \cdots [\boldsymbol{b}_n]_{\mathcal{C}} \end{bmatrix}$$

Theorem 5.1. The eigenvalues of a triangular matrix are the entries on its main diagonal.

Theorem 5.2. If $\boldsymbol{v}_1, \dots, \boldsymbol{v}_r$ are the eigenvectors that correspond to distinct eigenvalues $\lambda_1, \dots, \lambda_r$ of an $n \times n$ matrix A, then the set $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_r\}$ is linearly independent.

Theorem 5.3 (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 the determinant. A row interchange changes the sign of the determinant. A row scaling also scales the determinant by the same scalar factor.

Theorem 5.4. 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 multiplicities).

Theorem 5.5 (The Diagonalization Theorem)**.** 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 entries of D are eigenvalues of A that correspond, respectively, to the eigenvectors in P.

Theorem 5.6. An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

Theorem 5.7. Let A be an $n \times n$ matrix whose distinct eigenvalues are $\lambda_1, \dots, \lambda_p$.

a. For $1 \leq k \leq p$, the dimension of the eigenspace for 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 eigenspaces equals n, and this happens if and only if (i) the characteristic 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 $\mathcal{B}_1, \dots, \mathcal{B}_p$ forms an eigenvector basis for $\mathbb{R}^n$.

Theorem 5.8. If $A = PDP^{-1}$, then $A^k = PD^kP^{-1}$.

Theorem 6.1. Let $\boldsymbol{u}, \boldsymbol{v}$, and $\boldsymbol{w}$ be vectors in $\mathbb{R}^n$, and let c be a scalar. Then

a. $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{u}$

b. $(\boldsymbol{u} + \boldsymbol{v}) \cdot \boldsymbol{w} = \boldsymbol{u} \cdot \boldsymbol{w} + \boldsymbol{v} \cdot \boldsymbol{w}$

c. $(c\boldsymbol{u}) \cdot \boldsymbol{v} = c(\boldsymbol{u} \cdot \boldsymbol{v}) = \boldsymbol{u} \cdot (c\boldsymbol{v})$

d. $\boldsymbol{u} \cdot \boldsymbol{u} \geq 0$, and $\boldsymbol{u} \cdot \boldsymbol{u} = 0$ if and only if $\boldsymbol{u} = \boldsymbol{0}$

Theorem 6.2 (The Pythagorean Theorem)**.** Two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ are orthogonal if and only if $||\boldsymbol{u} + \boldsymbol{v}||^2 = ||\boldsymbol{u}||^2 + ||\boldsymbol{v}||^2$.

Theorem 6.3. Let A be an $m \times n$ matrix. The orthogonal complement of the row space of A is the null space of A, and the orthogonal complement of the column space of A is the null space of A^T:

$$(\mathrm{Row}\, A)^{\perp} = \mathrm{Nul}\, A \text{ and } (\mathrm{Col}\, A)^{\perp} = \mathrm{Nul}\, A^T$$

Theorem 6.4. If $S = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p\}$ is an orthogonal set of nonzero vectors in $\mathbb{R}^n$, then S is linearly independent and hence is a basis for the subspace spanned by S.

Theorem 6.5. Let $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p\}$ be an orthogonal basis for a subspace W of $\mathbb{R}^n$. For each $\boldsymbol{y}$ in W, the weights of the linear combination

$$\boldsymbol{y} = c_1\boldsymbol{u}_1 + \cdots + c_p\boldsymbol{u}_p$$

are given by

$$c_j = \frac{\boldsymbol{y} \cdot \boldsymbol{u}_j}{\boldsymbol{u}_j \cdot \boldsymbol{u}_j} \quad (j = 1, \ldots, p)$$

Theorem 6.6. An $m \times n$ matrix U has orthonormal columns if and only if $U^TU = I$.

Theorem 6.7. Let U be an $m \times n$ matrix with orthonormal columns, and let $\boldsymbol{x}$ and $\boldsymbol{y}$ be in $\mathbb{R}^n$. Then

a. $||U\boldsymbol{x}|| = ||\boldsymbol{x}||$

b. $(U\boldsymbol{x}) \cdot (U\boldsymbol{y}) = \boldsymbol{x} \cdot \boldsymbol{y}$

c. $(U\boldsymbol{x}) \cdot (U\boldsymbol{y}) = 0$ if and only if $\boldsymbol{x} \cdot \boldsymbol{y} = 0$

Theorem 6.8 (The Orthogonal Decomposition Theorem)**.** Let W be a subspace of $\mathbb{R}^n$. Then each $\boldsymbol{y} \in \mathbb{R}^n$ can be written uniquely in the form

$$\boldsymbol{y} = \hat{\boldsymbol{y}} + \boldsymbol{z}$$

where $\hat{\boldsymbol{y}}$ is in W and $\boldsymbol{z}$ is in $W^{\perp}$. In fact, if $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p\}$ is any orthogonal basis of W, then

$$\hat{\boldsymbol{y}} = \frac{\boldsymbol{y} \cdot \boldsymbol{u}_1}{\boldsymbol{u}_1 \cdot \boldsymbol{u}_1}\boldsymbol{u}_1 + \cdots + \frac{\boldsymbol{y} \cdot \boldsymbol{u}_p}{\boldsymbol{u}_p \cdot \boldsymbol{u}_p}\boldsymbol{u}_p$$

and $\boldsymbol{z} = \boldsymbol{y} - \hat{\boldsymbol{y}}$.

Theorem 6.9 (The Best Approximation Theorem)**.** Let W be a subspace of $\mathbb{R}^n$, let $\boldsymbol{y}$ be any vector in $\mathbb{R}^n$, and let $\hat{\boldsymbol{y}}$ be the orthogonal projection of $\boldsymbol{y}$ onto W. Then $\hat{\boldsymbol{y}}$ is the closest point in W to $\boldsymbol{y}$, in the sense that

$$||\boldsymbol{y} - \hat{\boldsymbol{y}}|| < ||\boldsymbol{y} - \boldsymbol{v}||$$

for all $\boldsymbol{v} \in W$ distinct from $\hat{\boldsymbol{y}}$.

Theorem 6.10. If $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p\}$ is an orthonormal basis for a subspace W of $\mathbb{R}^n$, then

$$\mathrm{proj}_W \boldsymbol{y} = (\boldsymbol{y} \cdot \boldsymbol{u}_1)\boldsymbol{u}_1 + \cdots + (\boldsymbol{y} \cdot \boldsymbol{u}_p)\boldsymbol{u}_p$$

If $U = [\boldsymbol{u}_1\ \boldsymbol{u}_2\ \cdots\ \boldsymbol{u}_p]$, then

$$\mathrm{proj}_W \boldsymbol{y} = UU^T\boldsymbol{y} \text{ for all } \boldsymbol{y} \in \mathbb{R}^n.$$

Theorem 6.11 (The Gram-Schmidt Process)**.** See section 6.2.

Theorem 6.12 (The QR Factorization)**.** If A is an $m \times n$ matrix with linearly independent columns, then A can be factored as $A = QR$, where Q is an $m \times n$ matrix whose columns form an orthonormal basis for $\operatorname{Col} A$ and R is an $n \times n$ upper triangular invertible matrix with positive entries on its diagonal.

Theorem 6.13. The set of least-squares solutions of $A\boldsymbol{x} = \boldsymbol{b}$ coincides with the nonempty set of solutions of the normal equations $A^T A\boldsymbol{x} = A^T\boldsymbol{b}$.

Theorem 6.14. Let A be an $m \times n$ matrix. The following statements are logically equivalent:

a. The equation $A\boldsymbol{x} = \boldsymbol{b}$ has a unique least-squares solution for each $\boldsymbol{b} \in \mathbb{R}^m$.

b. The columns of A are linearly independent.

c. The matrix $A^T A$ is invertible.

When these statements are true, the least-squares solution $\hat{\boldsymbol{x}}$ is given by

$$\hat{\boldsymbol{x}} = (A^T A)^{-1} A^T \boldsymbol{b}.$$

Theorem 6.15. Given an $m \times n$ matrix A with linearly independent columns, let $A = QR$ be a QR factorization of A as in thm 6.12. Then, for each $\boldsymbol{b} \in \mathbb{R}^m$, the equation $A\boldsymbol{x} = \boldsymbol{b}$ has a unique least-squares solution, given by

$$\hat{\boldsymbol{x}} = R^{-1} Q^T \boldsymbol{b}.$$

Index

References

[1] D. C. Lay, S. R. Lay, and J. McDonald. *Linear algebra and its applications.* Pearson Education, 2016.

[2] D. Quarfoot. *Studenting 101: School and Life Advice I Wish I Had Heard.* University of California, San Diego.

[3] L. Stevens. *Lectures in MATH 18.* University of California, San Diego, 2018.