MATH 118, Spring 2020, Linear Algebra
Condensed Lecture Notes

Taken in part from
Introduction to Linear Algebra, 4e,
Gilbert Strang
and *Linear Algebra and its Applications 5e,*
David C. Lay, Steven R. Lay, Judy J. McDonald

Notes compiled by Bobby McDonald
Yale University, 2020

NOTE: I will update these notes as often as I can with the topics and examples (which will be worked out by hand in a separate document) we cover in class.

Contents

Chapter 1

Introduction to Vectors

1.1 Vectors and linear combinations

1.1. Key Ideas

- A vector $\mathbf{v}$ in two-dimensional space has two components v_1 and v_2.
- $\mathbf{v}+\mathbf{w} = \langle v_1 + w_1, v_2 + w_2 \rangle$ and $c\mathbf{v} = \langle cv_1, cv_2 \rangle$ are found a component at a time.
- A linear combination of three vectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ is $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$.
- In three dimensions, *all* linear combinations of $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ typically fill a line, then a plane, then the whole space $\mathbb{R}^3$.

Definition 1.1.1 (Vectors in $\mathbb{R}^2$)**.** *A array with only one column is called a* ***column vector****, or just a* ***vector****. Examples of vectors with two entries are*

$$\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} \sqrt{2} \\ \pi \end{bmatrix} \qquad \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

where w_1, w_2 are real numbers. The set of all vectors with two entries is called $\mathbb{R}^2$. Two vectors are ***equal*** *if and only if their corresponding entries are equal.*

Definition 1.1.2. *Given two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^2$, their* ***sum*** *is the vector $\mathbf{u}+\mathbf{v}$ obtained by adding the corresponding entries of $\mathbf{u}$ and $\mathbf{v}$. For example,*

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1+2 \\ 2+3 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$$

Given a vector $\mathbf{v}$ and a real number c, the ***scalar multiple*** *of $\mathbf{u}$ is the vector $c\mathbf{u}$ obtained by multiplying each entry of $\mathbf{u}$ by c. For example if*

$$c = 2 \text{ and } \mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \text{ then } c\mathbf{u} = 2\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}.$$

Example 1.1.3. *Given vectors* $\mathbf{u} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ *and* $\mathbf{v} = \begin{bmatrix} -3 \\ 4 \end{bmatrix}$, *find* $(-2)\mathbf{u}$, $(-2)\mathbf{v}$, *and* $\mathbf{u} + (-3)\mathbf{v}$.

Observation 1.1.4 (Vectors in $\mathbb{R}^2$)**.** We can identify the column vector $\left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right]$ with *the point* (a, b) in the plain, so we can consider $\mathbb{R}^2$ as the set of all points in the plain. We usually visualize a vector by including an arrow from the origin.

Partners 1.1.5. *Let* $\mathbf{u} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ *and* $\mathbf{v} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. *Graph* $\mathbf{u}$, $\mathbf{v}$ *and* $\mathbf{u} + \mathbf{v}$ *on the plane.*

Proposition 1.1.6 (Parallelogram Rule)**.** *Geometrically,* $\mathbf{u} + \mathbf{v}$ *is the last vertex of the parallelogram with vertices are* $\mathbf{u}$, $\mathbf{v}$ *and* $\mathbf{0}$.

Partners 1.1.7. *Let* $\mathbf{u} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. *Graph* $\mathbf{u}$, $(-2)\mathbf{u}$, *and* $3\mathbf{u}$. *What's special about* $c\mathbf{u}$ *for any* c*?*

Groups 1.1.8. *If* $\mathbf{u} = (a, b)$, *can you find an equation for the line that contains all multiples of* $\mathbf{u}$*?*

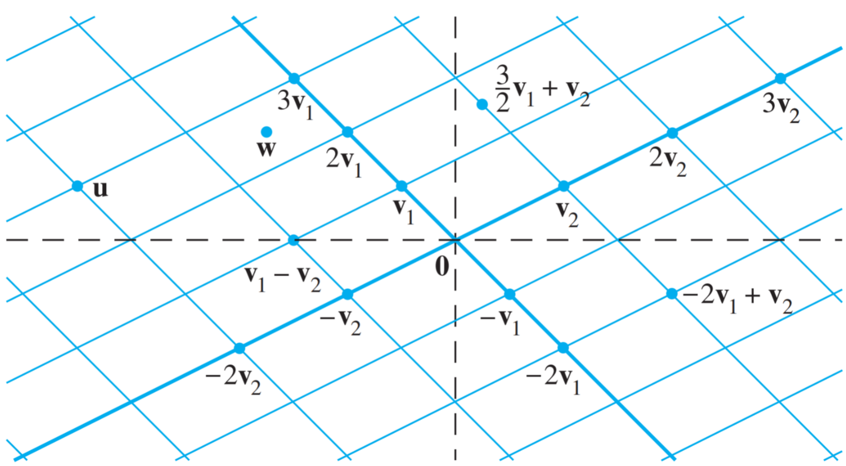
Question 1.1.9. *What properties do addition of vectors and multiplication by scalars enjoy?*

Definition 1.1.10. *Given vectors* $\mathbf{u}$ *and* $\mathbf{v}$ *in* $\mathbb{R}^n$, *and real numbers* c, d. *The vector*

$$c\mathbf{u} + d\mathbf{v}$$

is called a ***linear combination*** *of* $\mathbf{u}$ *and* $\mathbf{v}$ *with* ***weights*** c *and* d.

Example 1.1.11. *The figure below shows linear combinations of* $\mathbf{v}_1 = \left[\begin{smallmatrix} -1 \\ 1 \end{smallmatrix}\right]$ *and* $\mathbf{v}_2 = \left[\begin{smallmatrix} 2 \\ 1 \end{smallmatrix}\right]$ *where with integer weights. Estimate the linear combinations of* $\mathbf{v}_1$ *and* $\mathbf{v}_2$ *that produce* $\mathbf{u}$ *and* $\mathbf{w}$.



Definition 1.1.12. *If $\mathbf{u}, \mathbf{v}, \mathbf{w}$ are in $\mathbb{R}^3$, then the set of all linear combinations of is denoted by Span$\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ and is called the* ***span of*** $\mathbf{u}, \mathbf{v}$, ***and*** $\mathbf{w}$. *It's the collection of all vectors that look like*

$$a\mathbf{u} + b\mathbf{v} + c\mathbf{w}, \text{ with } a, b, c \text{ real numbers.}$$

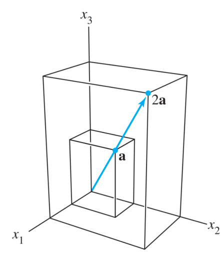
Remark 1.1.13. This definition can be extended to linear combinations of three or more vectors

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_n\mathbf{v}_n.$$

Think, Pair, Share 1.1.14. *Let* $\mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ *and* $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. *If we plotted all linear combinations of* $\mathbf{v}_1$ *and* $\mathbf{v}_2$ *what would we get?*

Observation 1.1.15 (Vectors in $\mathbb{R}^3$)**.** Vectors in $\mathbb{R}^3$ are 3×1 matrices. Like above, we can represent them geometrically in three-dimensional coordinate space. For example,

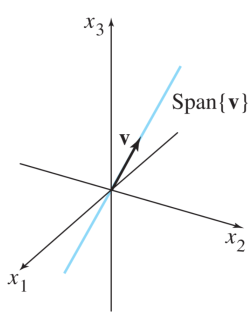
$$\mathbf{a} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$$

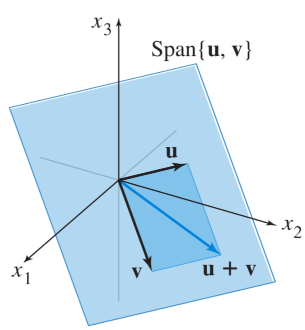


Question 1.1.16. *Let* $\mathbf{u}$ *and* $\mathbf{v}$ *be vectors in* $\mathbb{R}^3$. *What do linear combinations of* $\mathbf{u}$ *and* $\mathbf{v}$ *look like?*

Observation 1.1.17 (Geometric Descriptions of Span$\{\mathbf{u}\}$ and Span$\{\mathbf{u}, \mathbf{v}\}$)**.** Let $\mathbf{u}$ and $\mathbf{v}$ be nonzero vectors in $\mathbb{R}^3$, with $\mathbf{u}$ not a multiple of $\mathbf{v}$.

- Span$\{\mathbf{v}\}$ is the line through $\mathbf{0}$ and $\mathbf{v}$,
- Span$\{\mathbf{u}, \mathbf{v}\}$ is the plane containing $\mathbf{0}$, $\mathbf{u}$, and $\mathbf{v}$.





Definition 1.1.18 (Vectors in $\mathbb{R}^n$)**.** *If n is a positive integer, $\mathbb{R}^n$ denotes the collection of ordered n-tuples of n real numbers, usually written as $n \times 1$ column matrices, such as*

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix},$$

we we sometimes denote $\langle a_1, a_2, \ldots, a_n \rangle$. The ***zero vector****, denoted* ***0*** *is the vector whose entries are all zero. We also denote $(-1)\mathbf{u} = -\mathbf{u}$.*

1.2 Lengths and dot products

1.2. Key Ideas

- The dot product $\mathbf{v} \cdot \mathbf{w}$ multiplies each component v_i by w_i and adds all $v_i w_i$.
- The length $\|\mathbf{v}\|$ of a vector $\mathbf{v}$ is the square root of $\mathbf{v} \cdot \mathbf{v}$.
- $\mathbf{u} = \mathbf{v}/\|\mathbf{v}\|$ is a **unit vector.** Its length is 1
- The dot product $\mathbf{v} \cdot \mathbf{w} = 0$ when the vectors $\mathbf{v}$ and $\mathbf{w}$ are perpendicular
- If θ is the angle between $\mathbf{v}$ and $\mathbf{w}$, then $\cos\theta = \dfrac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\|\|\mathbf{w}\|}$

Definition 1.2.1. *If $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$ are vectors in $\mathbb{R}^n$, then the* ***dot product*** *or* ***inner product*** *of $\mathbf{u}$ and $\mathbf{v}$ is $\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2$*

Question 1.2.2. *What is $\mathbf{u} \cdot \mathbf{v}$? Is it a vector?*

Example 1.2.3. *Suppose we are buying and selling candy. Gum costs \$1.00 for a pack, chocolate is \$0.75 a bar, and hard candies are \$1.50 for a roll. If we sell 10 packs of gum and 20 chocolate bars, and buy 10 rolls of hard candy, what is our total income? Okay. We know how to do this. What does it have to do with dot products?*

Groups 1.2.4. *Compute $\mathbf{u} \cdot \mathbf{v}$ for $\mathbf{u} = \begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 3 \\ 2 \\ -3 \end{bmatrix}$.*
Without any calculation, can you decide what $\mathbf{v} \cdot \mathbf{u}$ is?

Theorem 1.2.5. *Let $\mathbf{u}, \mathbf{v}$ and $\mathbf{w}$ be vectors in $\mathbb{R}^n$, and let c be a scalar. Then $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$.*

Definition 1.2.6. *The **length** (or **norm**) of $\mathbf{v}$ is the square root of dotted with itself:*

$$\|\mathbf{v}\| = \sqrt{\mathbf{v}\cdot\mathbf{v}} = \sqrt{v_1^2+v_2^2+v_3^2}, \quad \text{and} \quad \|\mathbf{v}\|^2 = \mathbf{v}\cdot\mathbf{v}$$

Example 1.2.7. *Find the norm of $\mathbf{v} = (1, -2, 2)$.*

Think, Pair, Share 1.2.8. *Let $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$. What does the norm of $\mathbf{v}$ represent geometrically?*

Definition 1.2.9. *A vector whose length is 1 is called a **unit vector**.*

Example 1.2.10. *Let $\mathbf{v} = (1, -2, 2)$. Find a unit vector $\mathbf{u}$ in the same direction as $\mathbf{v}$.*

Definition 1.2.11. *If we divide a vector by its length, we obtain a unit vector in the same direction as $\mathbf{v}$. This is called **normalizing** $\mathbf{v}$.*

Groups 1.2.12. *Show that $\mathbf{u} = (2, 3)$ and $\mathbf{v} = (-3, 2)$ meet at right angles. Hint: we've already seen that (a, b) lives on the line $ay = bx$.*

Definition 1.2.13. *Two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ are **orthogonal** (or perpendicular) if $\mathbf{u}\cdot\mathbf{v} = 0$.*

Groups 1.2.14. *Find a nonzero vector in $\mathbb{R}^3$ that is orthogonal to $\mathbf{u} = (1, 2, 3)$.*

Theorem 1.2.15. *If $\mathbf{u}$ and $\mathbf{v}$ are nonzero vectors, then*

$$\|\mathbf{u}\|\|\mathbf{v}\|\cos(\theta) = \mathbf{u}\cdot\mathbf{v}$$

Question 1.2.16. *What does this tell us about the sign of the dot product $\mathbf{u}\cdot\mathbf{v}$?*

Theorem 1.2.17 (Pythagorean Theorem)**.** *Two vectors $\mathbf{u}$ and $\mathbf{v}$ are orthogonal if and only if*

$$\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$$

1.3 Matrices

1.3. Key Ideas

- linear equations and vector equations
- solving simple systems of equations
- matrices
- **Matrix times vector:** $A\mathbf{x}$ = linear combination of the columns of A with x_i as weights.

1.3.1 Linear equations

Example 1.3.1. *Suppose we are buying and selling candy, again. Remember, gum costs \$1.00 for a pack, chocolate is \$0.75 a bar, and hard candies are \$1.50 for a roll. Suppose*

- *Monday, we sell 10 packs of gum and 20 chocolate bars and buy 10 rolls of hard candy,*
- *Tuesday, we buy 10 packs of gum, sell 10 chocolate bars, and buy/sell no hard candies,*
- *Wednesday, we buy/sell no packs of gum, buy 4 chocolate bars, and buy/sell no hard candies.*

What is our net profit?

Observation 1.3.2. We can represent the previous example using three separate equations,

$$\begin{aligned} 1.00 \times 10 \ + \ 0.75 \times 20 \ + \ 1.50 \times 10 &= \text{Monday Profit} \\ 1.00 \times 10 \ + \ 0.75 \times 10 \ + \ 1.50 \times \ \ 0 &= \text{Tuesday Profit} \\ 1.00 \times 10 \ + \ 0.75 \times 20 \ + \ 1.50 \times 10 &= \text{Wednesday Profit} \end{aligned}$$

or as a **vector equation**

$$1.00 \times \begin{bmatrix} 10 \\ 10 \\ 0 \end{bmatrix} + 0.75 \times \begin{bmatrix} 20 \\ 10 \\ 4 \end{bmatrix} + 1.50 \times \begin{bmatrix} 10 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \text{Mon. P.} \\ \text{Tues. P.} \\ \text{Wed. P.} \end{bmatrix}$$

Definition 1.3.3. *A **linear equation** in the variables $x_1, \ldots, x_n$ is an equation that can be written in the form*

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

*where b and the **coefficients** $a_1, \ldots, a_n$ are real or complex numbers.*

Example 1.3.4. *Which of the following are linear equations?*

1. $4x_1 - 5x_2 + 2 = x_1$
2. $x_2 = 2(\sqrt{6} - x_1) + x_3$
3. $4x_1 - 5x_2 = x_1x_2$
4. $x_2 = 2\sqrt{x_1} - 6$

Definition 1.3.5. *A* ***system of linear equations*** *(or a* ***linear system****) is a collection of one or more linear equations involving the same variables, say* $x_1, \ldots, x_n$*.*

Example 1.3.6. *Is* $(5, 6.5, 3)$ *in the solution set (the set of all solutions) of the system*

$$\begin{aligned} 2x_1 - x_2 + 1.5x_3 &= 8 \\ x_1 \qquad - \quad 4x_3 &= -7 \end{aligned}$$

Definition 1.3.7. *The collection of all solutions to a system is called the* ***solution set****. Two linear systems are called* ***equivalent*** *if they have the same solution set.*

Remark 1.3.8. You already know how to find the solution set to a system of two linear equations in two unknowns! Just find the intersection of the two lines!

Example 1.3.9. *What are the solution sets of the following systems?*

$$(a) \quad \begin{aligned} x_1 - 2x_2 &= -1 \\ -x_1 + 3x_2 &= 3 \end{aligned} \qquad (b) \quad \begin{aligned} x_1 - 2x_2 &= -1 \\ -x_1 + 2x_2 &= 3 \end{aligned} \qquad (c) \quad \begin{aligned} x_1 - 2x_2 &= -1 \\ 2x_1 - 4x_2 &= -2 \end{aligned}$$

Example 1.3.10. *What is the solution set of the following system? If we fix* b_1, b_2, b_3*, how many solutions will it have?*

$$\begin{aligned} x_1 \qquad\qquad &= b_1 \\ -x_1 + x_2 \qquad &= b_2 \\ -x_2 + x_3 &= b_3 \end{aligned}$$

Example 1.3.11. *How can we interpret solutions to systems of equations with three variables geometrically?*

1.3.2 Matrices

Observation 1.3.12. We can also represent our first example more concisely using something called a *matrix* and multiplying it by a vector.

$$1.00 \times \begin{bmatrix} 10 \\ 10 \\ 0 \end{bmatrix} + 0.75 \times \begin{bmatrix} 20 \\ 10 \\ 4 \end{bmatrix} + 1.50 \times \begin{bmatrix} 10 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 10 & 20 & 10 \\ 10 & 10 & 0 \\ 0 & 4 & 0 \end{bmatrix} \begin{bmatrix} 1.00 \\ 0.75 \\ 1.50 \end{bmatrix}$$

Definition 1.3.13. *A matrix is an array of numbers. If A is a matrix with m rows and n columns, then A is called an $m \times n$ matrix. The entry in the ith row and jth column of A is called the (i, j) entry of A.*

Definition 1.3.14 (Matrix times a vector)**.** *If A has columns $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$, and $\mathbf{x}$ is the vector (c, d, e), then*

$$A\mathbf{x} = \begin{bmatrix} \mathbf{u} & \mathbf{v} & \mathbf{w} \end{bmatrix} \begin{bmatrix} c \\ d \\ e \end{bmatrix} = c\mathbf{u} + d\mathbf{v} + e\mathbf{w}.$$

Example 1.3.15. *Compute the product $A\mathbf{x}$ where*

$$A = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \qquad \mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Procedure 1.3.16. *If $A\mathbf{x} = \mathbf{b}$, then the ith entry of $\mathbf{b}$ is the dot product of the ith row of A with $\mathbf{x}$.*

1.3.3 Linear equations and matrices

Example 1.3.17. *Compute the product $A\mathbf{x}$ where*

$$A = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Example 1.3.18. *What if $A\mathbf{x} = \mathbf{b}$ where A and $\mathbf{b}$ are given, but $\mathbf{x}$ is unknown? How could we find $\mathbf{x}$ if we're told*

$$A = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

Example 1.3.19. *What is $\mathbf{x}$ if*

$$A = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Chapter 2

Solving Linear Equations

2.1 Vectors and linear equations

2.1. Key Ideas

- systems of equations can have no, one, or many solutions
- a system of equations with at least one solution is called consistent
- systems can be solved using back substitution

Example 2.1.1. *How many solutions do each of the following systems have?*

$$(a)\quad \begin{aligned} x_1 - 2x_2 &= -1 \\ -x_1 + 3x_2 &= 3 \end{aligned} \qquad (b)\quad \begin{aligned} x_1 - 2x_2 &= -1 \\ -x_1 + 2x_2 &= 3 \end{aligned} \qquad (c)\quad \begin{aligned} x_1 - 2x_2 &= -1 \\ 2x_1 - 4x_2 &= -2 \end{aligned}$$

Proposition 2.1.2. *A system of linear equations has*

1. *no solution, or*
2. *exactly one solution, or*
3. *infinitely many solutions.*

A system is called **consistent** if it has either one or infinitely many solutions, and **inconsistent** if it has no solution.

Example 2.1.3. *Are the following system consistent?*

$$(a)\quad \begin{aligned} x_1 - 2x_2 &= -1 \\ -x_1 + 3x_2 &= 3 \end{aligned} \qquad (b)\quad \begin{aligned} x_1 - 2x_2 &= -1 \\ -x_1 + 2x_2 &= 3 \end{aligned} \qquad (c)\quad \begin{aligned} x_1 - 2x_2 &= -1 \\ 2x_1 - 4x_2 &= -2 \end{aligned}$$

Remark 2.1.4. We will be interested in two fundamental questions about linear systems:

1. Is the system consistent?
2. If a system is consistent, is the solution unique?

Example 2.1.5. *Determine if the following system of equations is consistent.*

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ 5x_1 \phantom{{}-2x_2} - 5x_3 &= 10 \end{aligned}$$

2.2 The idea of elimination

2.2. Key Ideas

- A linear system becomes upper triangular after elimination.
- We subtract ℓ_{ij} times equation j from equation i to make the (i,j) entry zero, where

$$\ell_{ij} = \frac{(i,j) \text{ entry}}{\text{pivot in row } j}.$$

- The upper triangular system is solved by back substitution.

Example 2.2.1. *In Section 1.3 we determined whether the following systems were consistent using geometric and substitution arguments. Is there an algebraic way to do this without substitution?*

$$(a) \quad \begin{aligned} x_1 - 2x_2 &= -1 \\ -x_1 + 3x_2 &= 3 \end{aligned} \qquad (b) \quad \begin{aligned} x_1 - 2x_2 &= -1 \\ -x_1 + 2x_2 &= 3 \end{aligned} \qquad (c) \quad \begin{aligned} x_1 - 2x_2 &= -1 \\ 2x_1 - 4x_2 &= -2 \end{aligned}$$

Remark 2.2.2 (The idea of elimination)**.** We can solve a system by replacing it with an equivalent system that's easier to solve. We can do this by replacing one equation by adding multiples of equations, interchanging equations, or multiplying an equation by a nonzero constant.

Example 2.2.3. *Determine if the following system of equations is consistent without substitution*

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ 5x_1 - 5x_3 &= 10 \end{aligned}$$

Example 2.2.4. *Determine if the following system is consistent:*

$$\begin{aligned} x_2 - 4x_3 &= 8 \\ 2x_1 - 3x_2 + 2x_3 &= 1 \\ 4x_1 - 6x_2 + 4x_3 &= 1 \end{aligned}$$

2.3 Elimination using matrices

2.3. Key Ideas

- we can record information about a system in a matrix
- we can use elementary row operations to reduce matrices
- row reduction algorithm and how to use it to solve a system of equations

Definition 2.3.1. *The essential information in a linear system can be recorded into a rectangular array called a* ***matrix****. For example, given the system*

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ 5x_1 - 5x_3 &= 10 \end{aligned}$$

with the coefficients of each variable aligned in columns, the matrix

$$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ 5 & 0 & -5 \end{bmatrix}$$

is called the ***coefficient matrix*** *of the system, and*

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 5 & 0 & -5 & 10 \end{array}\right]$$

is called the ***augmented matrix*** *of the system. The* ***size*** *of a matrix tells us how many rows and columns it has. An* $\boldsymbol{m \times n}$ ***matrix*** *has m rows and n columns.*

Remark 2.3.2. Matrices will make our lives much easier when solving systems of linear equations!

Example 2.3.3. *Solve the system*

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ 5x_1 - 5x_3 &= 10 \end{aligned}$$

using a matrix.

Definition 2.3.4 (Elementary Row Operations)**.**

1. *(Replacement) replace one row by the sum of itself and a multiple of another row.*
2. *(Interchanging) Interchange two rows.*
3. *(Scaling) Multiply all entries in a row by a nonzero constant.*

Two matrices are called ***row equivalent*** *if there is a sequence of elementary row operations that transforms one matrix into the other.*

Remark 2.3.5. If the augmented matrices of two linear systems are row equivalent, then the two systems have the same solution set.

Example 2.3.6. *Determine if the following system is consistent:*

$$\begin{aligned} x_2 - 4x_3 &= 8 \\ 2x_1 - 3x_2 + 2x_3 &= 1 \\ 4x_1 - 8x_2 + 12x_3 &= 1 \end{aligned}$$

Definition 2.3.7. *The leftmost nonzero entry in a row is called the* ***leading entry****. A matrix is in* ***echelon form*** *(or* ***row echelon form****) if it has the following three properties:*

1. *all zero row are at the bottom*
2. *each leading entry of a row to the right of the leading entry above it*
3. *all entries in a column below a leading entry are zeros*

If a matrix in echelon form satisfies the following addition conditions, then it is in ***reduced echelon form*** *(or* ***reduced row echelon form****:*

4. *the leading entry in each nonzero row is* 1*.*
5. *each leading* 1 *is the only nonzero entry in its column.*

Example 2.3.8. *Which of the following is in echelon form? Reduced echelon form?*

$$\begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 0 & 0 & 0 & 5/2 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 & 29 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 & -5 \\ 0 & 1 & 2 & 12 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Remark 2.3.9. Any nonzero matrix can be **row reduced** into infinitely many matrices in echelon form. However, *reduced* echelon form for a matrix is unique.

Definition 2.3.10. *A* ***pivot position*** *in a matrix* A *is a location in* A *that corresponds to a leading* 1 *in the reduced echelon form of* A*. A* ***pivot column*** *is a column of* A *that contains a pivot position. A* ***pivot*** *is a nonzero entry in a pivot position.*

Example 2.3.11. *Label the pivot positions and pivot columns of the matrices above.*

Example 2.3.12. *Row reduce the matrix A to echelon form and locate pivot columns.*

$$A = \begin{bmatrix} 0 & -3 & -6 & 4 & 9 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 1 & 4 & 5 & -9 & -7 \end{bmatrix}$$

Procedure 2.3.13 (Row Reduction Algorithm)**.** *To transform a matrix into echelon form:*

1. *begin with the leftmost nonzero column*
2. *interchange rows if necessary to move a nonzero entry in this column to the top*
3. *use row replacement to create zeros in all positions below the new pivot.*
4. *ignore the row containing the pivot positions and all rows, if any, above it and apply steps 1-3 to the submatrix that remains*

If you want reduced echelon form, add one more step

5. *Beginning with the rightmost pivot and working upward and to the left, create zeros above each pivot. If a pivot is not* 1*, make it* 1 *by scaling.*

Example 2.3.14. *Apply elementary row operations to transform the following matrix into echelon form, and then reduced echelon form.*

$$\begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix}$$

Definition 2.3.15. *Steps 1-4 above are called the* ***forward phase*** *of the row reduction algorithm. Step 5 is called the* ***backward phase.***

Example 2.3.16. *Find the general solution of a linear system whose augmented matrix can be reduced to the matrix below.*

$$\left[\begin{array}{ccc:c} 1 & 0 & -5 & 1 \\ 0 & 1 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Definition 2.3.17. *The variables corresponding to pivot columns of a matrix are called* ***basic variables****, the other variables are called* ***free variables****.*

Remark 2.3.18. Whenever a system is consistent, the solution set can be described explicitly by solving the *reduced* system of equations for the basic variables in terms of the free variables.

Example 2.3.19. *Find the general solution of a system whose augmented matrix is reduced to*

$$\left[\begin{array}{ccccc:c} 1 & 6 & 2 & -5 & -2 & -4 \\ 0 & 0 & 2 & -8 & -1 & 3 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{array}\right]$$

Example 2.3.20. *Determine the existence and uniqueness of the solutions to the system*

$$\begin{aligned} 3x_2 - 6x_3 + \ \ 6x_4 + 4x_5 &= -5 \\ 3x_1 - 7x_2 + \ \ 8x_3 - 5x_4 + 8x_5 &= 9 \\ 3x_1 - 9x_2 + 12x_3 - 9x_4 + 6x_5 &= 15 \end{aligned}$$

Theorem 2.3.21. *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, where there are no free variables, or

(ii) infinitely many solutions, when there is at least one free variable.

Using the theorem, and the rest of this section, we have the following procedure to find and describe all the solutions of a linear system.

Procedure 2.3.22 (Using Row Reduction to Solve a Linear System)**.**

1. *Write the augmented matrix of the system.*
2. *Use the row reduction algorithm to write the matrix in echelon form. If the system is inconsistent, stop, there are no solutions; otherwise, go to the next step.*
3. *Use the row reduction algorithm to write the matrix in reduced echelon form.*
4. *Write the system of equations corresponding to the reduced matrix.*
5. *Solve each basic variable in terms of any free variables.*

2.4 Rules for matrix operations

2.4. Key Ideas

- The (i, j) entry of AB is the dot product of row i of A with column j of B.
- An $m \times n$ matrix times an $n \times p$ matrix gives an $m \times p$ matrix, and uses mnp separate multiplications.
- $A(BC) = (AB)C$, but $AB \neq BA$ in general

Definition 2.4.1. *If A is an $m \times n$ matrix (m rows and n columns), then the entry in the ith row and jth column of A, typically denoted a_{ij}, is called the (i, j)-**entry** of A. We write $A = [a_{ij}]$ using this notation. Columns of A are vectors in $\mathbb{R}^m$, usually denoted $\mathbf{a}_1, \ldots \mathbf{a}_n$. We often write:*

$$A = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix}.$$

*The **main diagonal** of $A = [a_{ij}]$ is the entries $a_{11}, a_{22}, a_{33}, \ldots$. A **zero matrix** is one whose entries are all zero. The **identity matrix** is a $n \times n$ square matrix with ones on the main diagonal and zeros everywhere else, usually denoted I_n*

Definition 2.4.2. *Two matrices are **equal** if they have the same size and their corresponding entires are equal. If A and B are matrices of the same size, then the **sum** $A+B$ is the matrix whose entries are the sums of the corresponding entries in A and B.*

Example 2.4.3. *Let* $A = \begin{bmatrix} 1 & 2 & 3 \\ -4 & 5 & -6 \end{bmatrix}$, $B = \begin{bmatrix} 4 & 5 & 6 \\ 7 & -8 & 9 \end{bmatrix}$, *and* $C = \begin{bmatrix} 1 & 3 \\ 5 & -6 \end{bmatrix}$.
Find $A + B$, $B + A$, and $A + C$.

Definition 2.4.4. *If r is a scalar and A is a matrix, then the **scalar multiple** rA is the matrix whose entries are r times the corresponding entries of A. Notationally, $-A$ stands for $(-1)A$, and $A - B = A + (-1)B$.*

Example 2.4.5. *Let* $A = \begin{bmatrix} 1 & 2 & 3 \\ -4 & 5 & -6 \end{bmatrix}$ *and* $B = \begin{bmatrix} 4 & 5 & 6 \\ 7 & -8 & 9 \end{bmatrix}$. *Find $2B$ and $A - 2B$.*

Theorem 2.4.6. *Let A, B, and C be matrices of the same size, and 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. $r(sA) = (rs)A$.

Definition 2.4.7. *If A is an $m \times n$ matrix, and B is an $n \times p$ matrix with columns $\mathbf{b}_1, \ldots, \mathbf{b}_p$, then the product AB is the $m \times p$ matrix whose columns are $A\mathbf{b}_1, \ldots, A\mathbf{b}_p$. That is*

$$AB = \begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 \cdots & A\mathbf{b}_p \end{bmatrix}.$$

Remark 2.4.8. If the number of columns of A doesn't match the number of rows of B, then the product AB is *undefined.*

Example 2.4.9. *Compute AB and BA, when $A = \begin{bmatrix} 2 & 1 \\ -3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & 5 & 1 \\ 2 & -8 & 3 \end{bmatrix}$.*

Procedure 2.4.10 (Row-Column Rule for AB)**.** *If the product AB is defined, then the (i,j)-entry of AB is the sum of the products of corresponding entries from row i of A and column j of B. If $(AB)_{ij}$ denotes the (i,j)-entry in AB, and A is an $m \times n$ matrix, then*

$$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + \cdots + a_{in}b_{nj}$$

Example 2.4.11. *With A and B from Example 2.4.9, compute AB using the row-column rule.*

Theorem 2.4.12. *Let A be an $m \times n$ matrix, and let B and C have the right sizes so that the following sums and products are defined.*

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

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

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

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

e. $I_m A = A = AI_n$

Example 2.4.13. *Let $A = \begin{bmatrix} 2 & -3 \\ -4 & 6 \end{bmatrix}$, $B = \begin{bmatrix} 8 & 4 \\ 5 & 5 \end{bmatrix}$, $C = \begin{bmatrix} 5 & -2 \\ 3 & 1 \end{bmatrix}$, and $D = \begin{bmatrix} 3 & 9 \\ 2 & 6 \end{bmatrix}$.*

(a) Find AB and BA.

(b) Find AC.

(c) Find AD.

Watchout! 2.4.14. *Here are some important warnings for matrix multiplication:*

1. *In general, $AB \neq BA$.*
2. *Cancellation laws do not hold for multiplication; $CA = CB$ (or $AC = BC$) does not mean $A = B$.*
3. *If $AB = 0$, this does not mean $A = 0$ or $B = 0$.*

Definition 2.4.15. *If A is an $m \times n$ matrix, the **transpose** of A is the $n \times m$ matrix, denoted A^T, whose columns are formed from the corresponding rows of A.*

Example 2.4.16. *Let* $A = \begin{bmatrix} a & b & d \end{bmatrix}$, $B = \begin{bmatrix} 8 & 4 \\ 5 & 5 \\ 6 & 2 \end{bmatrix}$, *and* $C = \begin{bmatrix} 5 & -2 & 1 & 3 \\ 3 & 1 & 2 & -6 \end{bmatrix}$.

Find A^T, B^T, and C^T.

Theorem 2.4.17. *Let A and B be matrices who are the right size for the following operations.*

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

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

c. $(rA)^T = rA^T$ *(for any scalar r)*

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