Chapter 1: Systems of Linear Equations

NoBS Linear Algebra (Jeffrey Wang)

May 1, 2018

Section 1

Introduction

Linear equations

In linear algebra, we start with **linear equations**. In high school algebra, we learned about the linear equation $y = mx + b$. However, we will be considering linear equations with multiple variables that are ordered, which can be written in the following form:

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = b$$

The coefficient may be any real number ($a \in \mathbb{R}$). However, the variable must be to the 1 power.

Examples of linear equations

The following are linear equations:

- $x_1 = 5x_2 + 7x_3 - 12$ – Nothing looks out of the ordinary here.
- $x_2 = \sqrt{5}(9 - x_3) + \pi x_1$ – The coefficients can be non-integers and irrational.

Counterexamples of linear equations

However, the following are NOT linear equations:

- $3(x_1)^2 + 5(x_2) = 9$ – We can't have a quadratic.
- $9x_1 + 7x_2x_3 = -45$ – We also can't have variables multiplying by each other.
- $3\sqrt{x_1} + x_2 = 2$ – We furthermore can't have roots of variables, nor any other non-polynomial function.
- $\frac{x_3}{x_1} + x_2 = 3$ – Inverses are not allowed either.
- $\sin x_1 + \cos^3 x_2 - \ln x_3 + e^{-x_4} = -9$ – Obviously no transcendental functions are allowed.

System of linear equations

A **system of linear equations** (or a linear system) is a collection of linear equations that share the same variables.
If an ordered list of values are substituted into their respective variables in the system of equations and each equation of the system holds true, then we call this collection of values a **solution**.

Example of system of linear equations

Say we have the ordered list of values $(1, 3, 2)$. This would be a solution of the following system of linear equations:

$$\begin{aligned} x_1 + x_2 + x_3 &= 6 \\ x_2 + x_3 &= 5 \\ x_3 &= 2 \end{aligned}$$

because if you substituted the solution for the respective variables, you get

$$\begin{aligned} 1 + 3 + 2 &= 6 \\ 3 + 2 &= 5 \\ 2 &= 2 \end{aligned}$$

which are all valid, true statements.

Consistency

A system can have more than one solution. We call the collection of all solutions the **solution set**. If two systems have exactly the same solution set, they are actually the same system.

In high school algebra, you learned that the solution of two linear equations was the point at which they intersected; this still holds true, but in linear algebra, we'll be dealing with more generalized cases that might be too complicated to solve graphically.

There are three possibilities (and only three) to the number of solutions a system can have: 0, 1, and ∞.

A system is **inconsistent** if it has no solutions and it is **consistent** if it has at least one solution.

- No solutions - inconsistent
- Exactly one solution - consistent
- Infinitely many solutions - consistent

Section 2

Introduction to matrices

Matrix notation

An **m × n matrix** is a rectangular array (or, ordered listing) of numbers with m rows and n columns, where m and n are both natural numbers.
We can use matrices to represent systems in a concise manner.
Given the system of linear equations:

$$\begin{aligned} 3x_1 + 5x_2 &= 9 \\ 7x_2 &= 56 \end{aligned}$$

We can rewrite this in matrix notation:

$$\left[\begin{array}{cc|c} 3 & 5 & 9 \\ 0 & 7 & 56 \end{array}\right]$$

Notes about matrix notation

$$\left[\begin{array}{cc|c} 3 & 5 & 9 \\ 0 & 7 & 56 \end{array}\right]$$

Notice we have kept only the coefficients of the variables. Each row represents one equation and each column represents one variable. The last column is not a column of variables, but instead of the constants. We usually put a vertical line between the second-to-last and last column on the matrix to denote this.

Definitions of matrix notation

$$\left[\begin{array}{cc|c} 3 & 5 & 9 \\ 0 & 7 & 56 \end{array}\right]$$

The matrix we've just created contains both the coefficients and the constants. We call this kind of matrix an **augmented matrix**. If we only had the coefficients:

$$\begin{bmatrix} 3 & 5 \\ 0 & 7 \end{bmatrix}$$

then we call this the **coefficient matrix**.
You can apply many of the terminologies and the concepts we've established for systems onto matrices. For instance, a consistent system means a consistent matrix.

Solving matrices

We're going to apply the same operations we used to solve systems of linear equations back in high school algebra, except this time in matrices.

Elementary row operations

The three **elementary row operations** for matrices are:

1. **Scaling** – multiplying all of the values on a row by a constant
2. **Interchange** – swapping the positions of two rows
3. **Replacement** – adding two rows together: the row that is being replaced, and a scaled version of another row in the matrix.

Properties of elementary row operations

These row operations are reversible. In other words, you can undo the effects of them later in the process.

We consider two augmented matrices to be **row equivalent** if there is a sequence of elementary row operations that can transform one matrix into another. Two matrices A and B may appear different, but if you can go from matrix A to B using these elementary row operations, then $A \sim B$.

In fact, there are infinitely many row equivalent matrices.

Linking back to the previous section: if two augmented matrices are row equivalent, then they also have the same solution set, which again confirms that they are the same matrix/system.

Existence and uniqueness (part 1)

Two fundamental questions establish the status of a system/matrix:

1. **Existence** - does a solution exist? (i.e. Is this matrix consistent?)
2. **Uniqueness** - if a solution exists, is this solution unique?

The rest of this chapter is dedicated to the methods used to finding whether a solution exists and is unique. The first half focuses on concepts and techniques that give us existence, and the second half focuses on uniqueness. Then, we'll combine them together in transformations.

Triangular form

For now, these questions are answered by using row operations to change the matrix into something we'll call **triangular** form. (This will be one of many ways to show existence.) This means that all values of the matrix below the "diagonal line" are zero. We can make them zero through row operations.

Case 1: not triangular form

This matrix is not in triangular form:

$$\begin{bmatrix} 2 & 3 & 4 & 2 \\ 7 & 2 & 3 & 4 \\ 4 & 19 & 9 & 9 \end{bmatrix}$$

Case 2: triangular form and consistent

This matrix is in triangular form and consistent:

$$\begin{bmatrix} 3 & 5 & 7 & 23 \\ 0 & 2 & 3 & 45 \\ 0 & 0 & 9 & 34 \end{bmatrix}$$

Case 3: triangular form but not consistent

This matrix, while in triangular form, is NOT consistent:

$$\begin{bmatrix} 7 & 8 & 9 & 10 \\ 0 & 11 & 8 & 5 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

This is because the last row says $0 = 3$, which is false. Therefore, there are no solutions for that matrix, so therefore it is inconsistent. Consistency = existence. They are the same thing.

Row reduction

Now that we're beginning to delve into solving matrices using row operations, let's develop an algorithm and identify a pattern that will let us find the solution(s) of a matrix.

Echelon form

This pattern was identified in the previous section. We called it by an informal name ("triangular form") but this is actually not its formal name. The pattern is called **echelon form** (or **row echelon form**, pronounced "esh-uh-lohn"). The official definition of echelon form is a matrix that has the following properties:

1. Any rows of all zeros must be below any rows with nonzero numbers.
2. Each "leading entry" of a row (which is the first nonzero entry, officially called a **pivot**) is in a column to the right of the leading entry/pivot of the row above it.
3. All entries in a column below a pivot must be zero.

Reduced echelon form

While echelon form is helpful, we actually want to reduce a matrix down to **reduced echelon form** (or reduced *row* echelon form) to obtain the solutions of the matrix. There are two additional requirements for a matrix to be in reduced echelon form:

4. The pivot in each nonzero row is 1. (All pivots in the matrix must be 1.)
5. Each pivot is the only nonzero entry of the column.

While there are infinitely many echelon forms for a certain matrix, there is only one unique reduced echelon form for a certain matrix.

Note that sometimes, reduced echelon form is not necessary to solve a problem in linear algebra. Ask yourself if you really need reduced echelon form. Usually, the answer is no.

Echelon form terminologies

A matrix in echelon form is an **echelon matrix**, and if in reduced echelon form, then it's a **reduced echelon matrix**.
Pivots are the leading nonzero entries of the row. A **pivot position** is a position of a certain pivot. A **pivot column** is a column that contains a pivot position.
For instance, in the following matrix:

$$\begin{bmatrix} \boxed{3} & 5 & 8 & 7 & 23 \\ 0 & \boxed{2} & 10 & 3 & 45 \\ 0 & 0 & 0 & \boxed{9} & 34 \end{bmatrix}$$

3 (located at row 1, column 1), 2 (at row 2, column 2), and 9 (at row 3, column 4) are the pivots (with their respective pivot positions).

Gaussian elimination (the row reduction algorithm)

This is the process through which you can reduce matrices. It always works, so you should use it!

Forward phase (echelon form):

1. Start with the leftmost nonzero column. This is a pivot column. The pivot position is the topmost entry of the matrix.
2. Use row operations to make all entries below the pivot position 0.
3. Move to the next column. From the previous pivot, go right one and down one position. Is this number zero? If so, move to the next column, but only move right one (don't move down one) and repeat the process. Otherwise, don't move. This is the next pivot. Repeat the steps above until all entries below pivots are zeros.

Backward phase (reduced echelon form):

1. Begin with the rightmost pivot.
2. Use row operations to make all entries above the pivot position 0.
3. Work your way to the left.
4. After every entry above the pivots are 0, see if any pivots are not 1. If so, use scaling to make it 1.

Evaluating solutions of a linear system

Reduced echelon form gives us the exact solutions of a linear system. We also found solutions by reducing our matrices to regular echelon form, plugging in values for the variables, and obtaining a solution. These are both valid ways to solve for a linear system. Now, let's interpret what these results mean.

Each column of a matrix corresponds to a variable. In our current nomenclature, we call these variables x_1, x_2, x_3, and so on. Accordingly, the first column of the matrix is associated with variable x_1 and so on. If there is a pivot in a column (i.e. a column is a pivot column), then this variable has a value assigned to it. In an augmented matrix in reduced echelon form, the value is in the corresponding row of the last column. This variable is called a **basic variable**.

What if a column does not have a pivot within it? This indicates that the value of this column's variable does not affect the linear system's solution. Therefore, it is called a **free variable**.

Parametric description of a solution set

The **parametric description** of a solution set leverages the above definitions. It is another way to represent the solution set.
For the following reduced echelon form augmented matrix:

$$\begin{bmatrix} 1 & 0 & 8 & 0 & 5 \\ 0 & 1 & 10 & 0 & 7 \\ 0 & 0 & 0 & 1 & 3 \end{bmatrix}$$

its parametric description is

$$\begin{cases} x_1 = 5 \\ x_2 = 7 \\ x_3 \text{ is free} \\ x_4 = 3 \end{cases}$$

Existence and uniqueness (part 2): of matrices

If the bottom row of an augmented matrix is all zero EXCEPT for the rightmost entry, then the matrix will be inconsistent.
Visually, if the bottom row of an augmented matrix looks like this:

$$\begin{bmatrix} 0 & \dots & 0 & b \end{bmatrix}$$

then this system is inconsistent.
Furthermore, if a system is consistent and does not contain any free variables, it has exactly one solution. If it contains one or more free variables, it has infinitely many solutions.

Section 3

Vectors and vector equations

Vectors

Back when we discussed the different kinds of solutions of a linear system, we assigned each variable of the system $x_1, x_2, \ldots, x_n$ its own column. Whatever number is its subscript is whichever column it represented in the matrix.
In fact, we'll now define these individual columns to be called **vectors**. Specifically, we call them **column vectors**: they have m number of rows but only one column. They are matrices of their own right.
Row vectors also exist, and have n columns but just one row. They are rarely used, and usually, the term vector refers to column vectors. (From now on, we will refer to column vectors simply as vectors unless otherwise specified.)

Vectors (continued)

We represent the set of all vectors with a certain dimension n by $\mathbb{R}^n$. Note that n is the number of rows a vector has. For instance, the vector $\vec{v} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ has two rows, so it is a vector within the set $\mathbb{R}^2$.

Two vectors are equal if and only if:

- they have the same dimensions
- the corresponding entries on each vector are equal

Vector operations and properties

The two fundamental vector operations are vector addition and scalar multiplication.

Vector addition: $\begin{bmatrix}1\\2\end{bmatrix} + \begin{bmatrix}3\\7\end{bmatrix} = \begin{bmatrix}1+3\\2+7\end{bmatrix} = \begin{bmatrix}4\\9\end{bmatrix}$

Scalar multiplication: $5\begin{bmatrix}3\\2\end{bmatrix} = \begin{bmatrix}5\cdot 3\\5\cdot 2\end{bmatrix} = \begin{bmatrix}15\\10\end{bmatrix}$

Vectors are commutative, associative, and distributive. This means we can use these operations together as well.

Vectors and matrices

A set of vectors can themselves form a matrix by being columns of a matrix. For instance, if we are given $\vec{v} = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}$, $\vec{u} = \begin{bmatrix} 9 \\ 5 \\ 2 \end{bmatrix}$, $\vec{w} = \begin{bmatrix} 7 \\ 8 \\ 6 \end{bmatrix}$, and they are all within the set A, then they can be represented as a matrix:

$$A = \begin{bmatrix} 1 & 9 & 7 \\ 3 & 5 & 8 \\ 4 & 2 & 6 \end{bmatrix}$$

Combining vectors

We can use vector operations to form new vectors. In fact, given a set of vectors, we are usually able to use these vector operations to represent infinitely many vectors. For instance, given the vector $\vec{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\vec{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, we can use just $\vec{v}$ and $\vec{u}$ to create any vector in $\mathbb{R}^2$, the set of all vectors with two real number entries.

Combining vectors: example

For instance, if we want to make $\vec{b} = \begin{bmatrix} 8 \\ 17 \end{bmatrix}$, we can write it as

$-9 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 17 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = -9\vec{v} + 17\vec{u} = \begin{bmatrix} 8 \\ 17 \end{bmatrix}$.

We can also put these vectors into matrices in the form $\left[\begin{array}{cc|c} \vec{v} & \vec{u} & \vec{b} \end{array}\right]$:

$$\begin{bmatrix} 1 & 1 & 8 \\ 0 & 1 & 17 \end{bmatrix}$$

We can then derive from this matrix these equations and conclusions:

$$x_2 = 17$$

$$x_1 + x_2 = 8$$

$$x_1 = 8 - 17 = -9$$

Ramifications of combining vectors

Therefore, we have just determined that the solution of this matrix will give us the coefficients (or weights) of the vectors in order to give us this vector $\vec{b}$.

Using vector operations to represent another vector is called a linear combination, which we will explore in the next section.

Section 4

Linear combinations

Linear combinations

A **linear combination** is the use of vector operations to equate a vector $\vec{y}$ to the vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p \in \mathbb{R}^n$ with coefficients, or **weights** $c_1, c_2, \ldots, c_p$.

$$\vec{y} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_p\vec{v}_p$$

An example of a linear combination is $3\vec{v}_1 - 2\vec{v}_2$.

Span

The possible linear combinations that can result from a set of vectors $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ is called the **span**.

The span is simply where this linear combination can reach. The span of $\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is simply the line parallel to $\vec{v}_1$. Notice you can't make much else out of that except scaling it. However, if you add $\vec{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ to the mix, you can now create a linear combination to get ANY value in $\mathbb{R}^2$ or the 2D plane. That means these two vectors' possibilities *spans* $\mathbb{R}^2$. For instance, $\begin{bmatrix} 9 \\ 20 \end{bmatrix} = 9\vec{v}_1 + 2\vec{v}_2$.

Span (continued)

We say that a vector $\vec{y}$ spans $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$, etc. through the notation $\vec{y} \in \text{Span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$.
The formal definition of span: if $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p \in \mathbb{R}^n$, then the set of all linear combinations of $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p$ is denoted by $\text{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ and is called the **subset of $\mathbb{R}^n$ spanned by** $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p$:

$$\text{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\} = \{c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_p\vec{v}_p : c_1, c_2, \ldots, c_p \text{ are scalars.}\}$$

Existence and uniqueness (part 3): of linear combinations

If there is a solution to the vector equation $x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_p\vec{v}_p = \vec{b}$, then there is a solution to its equivalent augmented matrix: $\begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_p & \vec{b} \end{bmatrix}$. The solution to the equivalent augmented matrix are in fact the weights of the vector equation. That means we can get $x_1, x_2, \dots, x_p$ from reducing the equivalent augmented matrix.

Example (part 1)

Let's say we have two vectors $\vec{a}_1 = \begin{bmatrix} 2 \\ 6 \\ 7 \end{bmatrix}$ and $\vec{a}_2 = \begin{bmatrix} -1 \\ -5 \\ -4 \end{bmatrix}$. Can the vector $\vec{b} = \begin{bmatrix} 8 \\ 32 \\ 24 \end{bmatrix}$ be generated as a a linear combination of $\vec{a}_1, \vec{a}_2$?

Let's consider what this means. Basically, is there a way we can find the following to be true?

$$x_1 \begin{bmatrix} 2 \\ 6 \\ 7 \end{bmatrix} + x_2 \begin{bmatrix} -1 \\ -5 \\ -4 \end{bmatrix} = \begin{bmatrix} 8 \\ 32 \\ 24 \end{bmatrix}$$

We have to add these vectors $\vec{a}_1$ and $\vec{a}_2$ together. Obviously you can't just add them together. So, can we multiply the vectors and then add them together in some form to get $\vec{b} = \begin{bmatrix} 8 \\ 32 \\ 24 \end{bmatrix}$? Maybe.

Example (part 2)

The question is, what are the coefficients needed for this to happen? We can solve for the coefficients by putting $\vec{a}_1$ and $\vec{a}_2$ in a matrix as columns and then $\vec{b}$ as the final column. Sound familiar? That's because this is an augmented matrix in the form $\begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \vec{b} \end{bmatrix}$:

$$\begin{array}{c} \begin{array}{ccc} \vec{a}_1 & \vec{a}_2 & \vec{b} \end{array} \\ \begin{bmatrix} 2 & -1 & 8 \\ 6 & -5 & 32 \\ 7 & -4 & 24 \end{bmatrix} \end{array}$$

We reduce this matrix and find the solutions from there.

Example (part 3)

Here's the reduced echelon form of the above augmented matrix:

$$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -4 \\ 0 & 0 & 0 \end{bmatrix}$$

The last column contains the answers. The first entry of the last row is $x_1 = 2$ and the second entry of the last row is $x_2 = -4$. Now, let's insert these values back into the linear combination:

$$x_1 \begin{bmatrix} 2 \\ 6 \\ 7 \end{bmatrix} + x_2 \begin{bmatrix} -1 \\ -5 \\ -4 \end{bmatrix} = \begin{bmatrix} 8 \\ 32 \\ 24 \end{bmatrix}$$

$$2 \begin{bmatrix} 2 \\ 6 \\ 7 \end{bmatrix} - 4 \begin{bmatrix} -1 \\ -5 \\ -4 \end{bmatrix} = \begin{bmatrix} 8 \\ 32 \\ 24 \end{bmatrix}$$

Example (part 4)

$$\begin{bmatrix} 4 \\ 12 \\ 14 \end{bmatrix} + \begin{bmatrix} 4 \\ 20 \\ 16 \end{bmatrix} = \begin{bmatrix} 8 \\ 32 \\ 24 \end{bmatrix}$$

$$\begin{bmatrix} 4+4 \\ 12+20 \\ 14+16 \end{bmatrix} = \begin{bmatrix} 8 \\ 32 \\ 24 \end{bmatrix}$$

$$\begin{bmatrix} 8 \\ 32 \\ 24 \end{bmatrix} = \begin{bmatrix} 8 \\ 32 \\ 24 \end{bmatrix}$$

Hey, these values $x_1 = 2$ and $x_2 = -4$ indeed hold true. That means yes, we can generate a linear combination from $\vec{a}_1$ and $\vec{a}_2$ for $\vec{b}$.

Example (part 5)

However, the question is whether $\vec{b}$ can be generated as a linear combination of $\vec{a}_1$ and $\vec{a}_2$.
Because the augmented matrix we made from the vectors was consistent, we can conclude that $\vec{b}$ can be generated as a linear combination of $\vec{a}_1$ and $\vec{a}_2$.
We did not need to do anything past the reduced echelon form. We just needed to see whether the augmented matrix was consistent.

Example (part 6)

For this same example, is $\vec{b}$ in the span of $\{\vec{a}_1, \vec{a}_2\}$?
The span refers to the possible combinations of a linear combination.
Since $\vec{b}$ is a possible linear combination of $\{\vec{a}_1, \vec{a}_2\}$, then yes, $\vec{b}$ is in the span of $\{\vec{a}_1, \vec{a}_2\}$.
To recap, $\vec{b}$ generated as a linear combination of $\vec{a}_1, \vec{a}_2$ is the same question as whether $\vec{b}$ is in the span of $\vec{a}_1, \vec{a}_2$. They are both proven by seeing whether the augmented matrix of $\begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \vec{b} \end{bmatrix}$ is consistent.

Section 5

Basic matrix multiplication

Basic matrix multiplication

To prepare for our next section, we will first jump to a limited scope of matrix multiplication where we are multiplying a matrix by a vector. We can multiply matrices by each other. This operation multiplies the entries in a certain pattern. We can only multiply two matrices by each other if the first matrix's number of columns is equal to the second matrix's number of rows.

Basic matrix multiplication (continued)

To be clear, let's say we have the first matrix and second matrix:

- The number of columns of the first matrix must equal the number of rows of the second matrix.
- The resulting matrix will have the number of rows that the first matrix has and the number of columns that the second matrix has.

For instance, we can multiply:

$$\begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$$

but not

$$\begin{bmatrix} 1 & 8 \end{bmatrix} \begin{bmatrix} 4 \\ 2 \\ 3 \end{bmatrix}$$

Multiplying matrices

So how do we multiply matrices? Let's start with the simple example that worked.

$$\begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 1 \cdot 4 + 2 \cdot 5 + 3 \cdot 6 \end{bmatrix} = \begin{bmatrix} 32 \end{bmatrix}$$

So we took the first row's entry and multiplied it by the first column's entry, then add to it the second row's entry by the second column's entry, etc.

Multiplying matrices (continued)

Now what if we had two rows on the first matrix?

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}$$

We do what we did the first time, except this time we put the second row's products on another row.

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} = \begin{bmatrix} 1 \cdot 7 + 2 \cdot 8 + 3 \cdot 9 \\ 4 \cdot 7 + 5 \cdot 8 + 6 \cdot 9 \end{bmatrix} = \begin{bmatrix} 50 \\ 122 \end{bmatrix}$$

For now, we will not worry about cases where there are more columns on the second matrix.

Section 6

The matrix equation $A\vec{x} = \vec{b}$

The matrix equation: recalling linear combinations

Recall that a linear combination is the use of some vectors manipulated with vector operations and coefficients (called weights) to equate to another vector.

$$x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_p\vec{a}_p = \vec{b}$$

The matrix equation: recalling our previous example

In the previous section's example, we set the vectors $\vec{a}_1, \vec{a}_2$ to be the following and $\vec{b}$ to be the following:

$$x_1 \begin{bmatrix} 2 \\ 6 \\ 7 \end{bmatrix} + x_2 \begin{bmatrix} -1 \\ -5 \\ 4 \end{bmatrix} = \begin{bmatrix} 8 \\ 32 \\ 24 \end{bmatrix}$$

Instead of writing x_1, x_2 as coefficients, we can actually put them in their own vector and use matrix multiplication to achieve the same thing.

$$\begin{bmatrix} 2 & -1 \\ 6 & -5 \\ 7 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 8 \\ 32 \\ 24 \end{bmatrix}$$

Defining the matrix equation

The first matrix is A, the coefficient matrix. The vector in the middle is the collection of all weights, and it's called $\vec{x}$. The vector on the right, the linear combination of $A\vec{x}$, is called $\vec{b}$.
This equation's general form is $A\vec{x} = \vec{b}$ and is the **matrix equation**.
In general:

$$A\vec{x} = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \dots & \vec{a}_p \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} = x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_p\vec{a}_p = \vec{b}$$

This is a great way to represent linear combinations in a simple, precise manner.

Existence and uniqueness (part 4): Four ways to represent a consistent linear system

So far, we have learned four ways to represent a consistent system (where a solution exists):

1. **Pivots**: The coefficient matrix A has a pivot position in each row.
2. **Linear combination**: $\vec{b}$ is a linear combination of the columns of A.
3. **Span**: $\vec{b}$ is in the span of the columns of A.
4. **Matrix equation**: If $A\vec{x} = \vec{b}$ has a solution.

Section 7

Homogeneous linear systems

Introduction to homogeneous linear systems

A **homogeneous linear system** is a special kind of linear system in the form $A\vec{x} = \vec{b}$ where $\vec{b} = \vec{0}$.
In other words, a homogeneous linear system is in the form $A\vec{x} = \vec{0}$.
What gives about them? They have some special properties.

Properties of homogenous linear systems

- All homogeneous systems have at least one solution. This solution is called the **trivial solution** and it's when $\vec{x} = \vec{0}$. Of course $A \cdot \vec{0} = \vec{0}$ is true; that's why we call it trivial!
- But the real question is, is there a case where a homogeneous system $A\vec{x} = \vec{0}$ when $\vec{x} \neq \vec{0}$? If such a solution exists, it's called a **nontrivial solution**.
- How do we know if there is a nontrivial solution? This is only possible when $A\vec{x} = \vec{0}$ has a **free variable**. When a free variable is in the system, it allows $\vec{x} \neq 0$ while $A\vec{x} = 0$.
- Therefore, we can say whenever a **nontrivial solution** exists for a homogeneous system, it has **infinitely many solutions**. When only the **trivial solution** exists in a homogeneous system, the system has a **unique solution** (uniqueness).

Parametric vector form

When we have free variables, we describe basic variables with relation to the free variables. We group the weights of the free variables to the basic variables through parametric vector form.

If I have:

$x_1 = 3x_2 + 4x_3$

$x_2 = 2x_3$

x_3 free

Then we will do necessary replacements to arrive at:

$x_1 = 10x_3$

$x_2 = 2x_3$

$x_3 = x_3$

Parametric vector form (continued)

If we combine them in $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ then we get $\vec{x} = \begin{bmatrix} 10x_3 \\ 2x_3 \\ x_3 \end{bmatrix}$. Factor out the x_3 and we get $\vec{x} = x_3 \begin{bmatrix} 10 \\ 2 \\ 1 \end{bmatrix}$. Now, we have represented the solution $\vec{x}$ to the matrix equation $A\vec{x} = \vec{b}$ in parametric vector form $x_3 \begin{bmatrix} 10 \\ 2 \\ 1 \end{bmatrix}$.

Non-homogeneous linear systems in terms of homogeneous linear systems

What linear system is a non-homogeneous linear system? Basically, it's anything where $A\vec{x} \neq \vec{0}$. So, basically most matrix equations.

What's fascinating is that non-homogeneous linear systems with infinitely many solutions can actually be represented through homogeneous linear systems that have nontrivial solutions. This is because non-homogeneous linear systems are in fact homogeneous linear systems with a *translation* by a certain constant vector $\vec{p}$, where $\vec{x} = \vec{p} + t\vec{v}$. $t\vec{v}$ is where all of the free variables are located.

And guess what? Since this system has infinitely many solutions, the translation is taken out when $A\vec{x} = \vec{0}$. That means, for a system with infinitely many solutions (i.e. one that can be written in parametric vector form), the nonconstant vector $\vec{v}$ **(the one with the free variables), without the constant vector $\vec{p}$, is a possible solution**!

Section 8

Linear independence

Introduction to linear independence

Now, within our linear combination, could we have duplicate vectors? Yes. Graphically speaking, if two vectors are multiples of each other, they would be parallel. The whole point of linear combinations is to make new vectors. What use is two vectors that are multiples of each other? That's redundant.

This is where the concept of **linear independence** comes in. Are there any redundant variables in the linear combination? Or, are there any redundant rows in a reduced matrix? If so, there are redundancies and therefore we say the combination is linearly dependent. If nothing is redundant, then it's linearly independent.

Formal definition of linear independence

Mathematically, however, we have a formal way of defining linear independence. If the linear combination $x_1\vec{v}_1 + x_2\vec{v}_2 + \ldots x_p\vec{v}_p = \vec{0}$ has *only the trivial solution*, then it's linearly independent. Otherwise, if there exists a nontrivial solution, it must be linearly dependent.
Why is that? If there are redundant vectors, then there exists weights for them to cancel each other out. Let's say we have two vectors
$\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}$. Clearly, $\vec{v}_2$ is a multiple of $\vec{v}_1$. So we can say
$2\vec{v}_1 = \vec{v}_2$. We can rearrange this to get $2\vec{v}_1 - 1\vec{v}_2 = \vec{0}$. Therefore, per our formal definition, $\{\vec{v}_1, \vec{v}_2\}$ is linearly dependent.

Linear independence and uniqueness

If a set of vectors are linearly dependent, then there are infinitely many ways to make a combination. Therefore, there would be infinitely many solutions for $A\vec{x} = \vec{b}$. However, if a set of vectors are linearly independent, then there would only be at most one way to form the solution for $A\vec{x} = \vec{b}$. This leads us to a conceptual connection that is very important: **if a set of vectors are linearly independent, at most one solution can be formed**. This means linear independence implies uniqueness of solution.

Linear independence of a set of one vector

A set with one vector contained within is linearly independent unless this vector is the zero vector. The zero vector is linearly dependent but has only the trivial solution.

$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ is linearly independent. $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ is linearly dependent and has only the trivial solution.

Linear independence of a set of two vectors

Given two vectors $\vec{v}, \vec{u}$, then this set is linearly independent as long as $\vec{v}$ is not in the span of $\vec{u}$. It could also be the other way around, but the same result still holds true.

Linear independence of a set of multiple vectors

We have several rules for determining linear dependence of a set of multiple vectors. If one rule holds true, then the set is linearly dependent.

- If at least one of the vectors in the set is a linear combination of some other vectors in the set, then the set is linearly dependent. (Note that not all have to be a linear combination of each other for this to hold true. Just one needs to be a linear combination of others.)
- If there are more columns than rows in a coefficient matrix, that system is linearly dependent.
- If a set contains the zero vector $\vec{0}$, then the set is linearly dependent. It's easy to create the zero vector with any vector; just make the weight 0, lol.

Why is this true?

"If there are more columns than rows in a coefficient matrix, that system is linearly dependent."

Why? Recall that the number of entries a vector has is equivalent to the number of dimensions it contains. If there are more vectors than dimensions, the vectors will be redundant within these dimensions. For instance, $\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ only spans a line in $\mathbb{R}^2$. If we add $\vec{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ to the set, we can now span all of $\mathbb{R}^2$ because $\vec{v}_1$ is not parallel to $\vec{v}_2$. But if we add $\vec{v}_3 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$, it doesn't add anything new to the span. We're still in $\mathbb{R}^2$ but with a redundant vector. Furthermore, $\vec{v}_1 = 2\vec{v}_2 - 1\vec{v}_3$, so clearly that would be a linearly dependent set of vectors.

Section 9

Linear transformations

Background of linear transformations

We have studied the matrix equation $A\vec{x} = \vec{b}$. This equation is kind of like $y = mx + b$, which is just a relation. These relations can be represented as functions in regular algebra as $f(x) = mx + b$. Now, we will transition to the equivalent of functions in linear algebra, called linear transformations.

Introduction to linear transformations

A **linear transformation**, also called a **function** or **mapping**, is where we take the vector $\vec{x}$ and manipulate it to get the result $T(\vec{x})$. This is the same concept as turning x into $f(x)$. How do we manipulate it? We have several options, but for now, we will stick with the coefficient matrix A. Using the coefficient matrix for manipulation means we are using a special form of linear transformation called a **matrix transformation**. (Note of clarification: If we talk about A, then we are referring to a matrix transformation. Otherwise, we are referring to a general linear transformation.)

Linear transformations: definitions

For the coefficient matrix A with size $m \times n$ (where m is the number of rows and n is the number of columns), the number of columns n it has is the **domain** ($\mathbb{R}^n$) while the number of rows m it has is the **codomain** ($\mathbb{R}^m$). The codomain is different from the range.
Specifically, the transformed vector $\vec{x}$ itself, $T(\vec{x})$, is called the **image** of $\vec{x}$ under T. Note that the image is like the $\vec{b}$ in $A\vec{x} = \vec{b}$, and in fact, $T(\vec{x}) = \vec{b}$.

Linear transformation: things to think about

Something to keep in mind: matrix equations can have no solutions, one solution, or infinitely many solutions. If no $\vec{x}$ can form an image, it means there's no solutions. If only one $\vec{x}$ forms an image, then there is only one solution. But if we can find more than one $\vec{x}$ (i.e. there's probably a basic variable) then we have infinitely many solutions and not a unique solution. Why isn't the codomain the range? Well, the range is the places where the solution to this matrix exists. The image is the actual vector $\vec{x}$ transformed, while the codomain is just the new domain where the image resides after this transformation has shifted. Not all of the codomain is in the range of the transformation. Remember, this is linear algebra, and we're shifting the dimensions here with these transformations, so we have to differentiate between which dimension we're in before and after. If $A\vec{x} = \vec{b}$, then $\vec{b}$ is in the range of the transformation $\vec{x} \mapsto A\vec{x}$.

Linear transformation notation

By the way, the notation to represent a transformation's domain to codomain is:

$$T : \mathbb{R}^n \mapsto \mathbb{R}^m$$

Note that $\mapsto$ is read as "maps to", so that's why we'll call use the terms transformations and mappings interchangeably. It's just natural!

Properties of linear transformations

It's important to note that matrix transformations are a specific kind of linear transformation, so you can apply these properties to matrix transformations, but they don't solely apply to matrix transformations. Also, don't forget later on that these properties hold true for other kinds of linear transformations. With that in mind, the following properties apply to all linear transformations:

- $T(\vec{u}+\vec{v}) = T(\vec{u}) + T(\vec{v})$, for $\vec{u}, \vec{v} \in \text{Domain}\{T\}$
- $T(c\vec{u}) = cT(\vec{u})$, for $c \in \mathbb{R}, \vec{u} \in \text{Domain}\{T\}$
- $T(\vec{0}) = \vec{0}$ - oh yeah, that's right. Those "linear functions" $y = mx + b$ with $b \neq 0$ you learned about in Algebra I? They're not actually "linear functions" in linear algebra.

So why do these properties even matter? In problems, you are given $T(\vec{u})$ but not $\vec{u}$ itself (i.e. $\vec{b}$ and not $\vec{x}$ in $A\vec{x} = \vec{b}$), so you must isolate $T(\vec{u})$ from these operations to do anything meaningful with it.

Section 10

The matrix of a linear transformation

Background of the matrix of a linear transformation

In the previous section, we mainly dealt with $\vec{x}$ in the matrix equation $A\vec{x} = \vec{b}$. Why is $A\vec{x} = \vec{b}$ a transformation? Because, every linear transformation that satisfies $\mathbb{R}^m \mapsto \mathbb{R}^n$ is a matrix transformation. (Note: This means only linear transformations with $\mathbb{R}^m \mapsto \mathbb{R}^n$ is a matrix transformation. What was said before about not all linear transformations being matrix transformations *still holds true*.) And because they are matrix transformations, we can use the form $\vec{x} \mapsto A\vec{x}$ to represent our transformations. In this section, we will shift the focus from $\vec{x}$ to A.

Now, when we map $\vec{x} \mapsto A\vec{x}$, we have technically do have a coefficient in the front of the lone $\vec{x}$. So really, we should say $I\vec{x} \mapsto A\vec{x}$. What matrix I always satisfies $\vec{x} = I\vec{x}$? The **identity matrix** I_n, which is a matrix size $n \times n$ with all zeroes except for ones down the diagonal.

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \text{etc.}$$

Why does this matter? Well, in $I\vec{x} \mapsto A\vec{x}$, $\vec{x}$ is on both sides. So really, the transformation represents a change from I, the identity matrix, to A. We shall name A the **standard matrix**.
The columns of the identity matrix, which we shall call $\{\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n}\}$, correlate to the columns of the standard matrix, which are just transformations of the identity matrix's columns (i.e. $A = [T(\vec{e_1})\ T(\vec{e_2})\ \ldots\ T(\vec{e_n})]$).

Another vector representation format

We can write vectors another way: $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = (x_1, x_2)$. This comes in handy when we want to write a transformation as a tuple, like we do for a function: $T(x_1, x_2) = (x_1 + x_2, 3x_2) = \begin{bmatrix} x_1 + x_2 \\ 3x_2 \end{bmatrix}$.

$$T(x_1, x_2) = (x_1 + x_2, 3x_2)$$

is the same thing as

$$T(\vec{x}) = A\vec{x} = \vec{b}$$

is the same thing as

$$T(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}) = \begin{bmatrix} 1 & 1 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ 3x_2 \end{bmatrix}$$

(Note that a tuple is the generic term for an ordered pair with more than two values. For instance, $(1, 3, 4)$ is a 3-tuple. And technically, $(2, 5)$ is an ordered pair, also called a 2-tuple.)

Existence and uniqueness (part 5): Onto and one-to-one

Two ways we can characterize transformations are whether they are **onto mappings**, **one-to-one mappings**, both, or neither. They help us determine the existence and uniqueness of solution(s) for transformations.
Formal definitions:
A mapping $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ is **onto** $\mathbb{R}^m$ if for each $\vec{b}$ in the codomain $\mathbb{R}^m$, **there exists a** $\vec{x}$ in the domain $\mathbb{R}^n$ so that $T(\vec{x}) = \vec{b}$.
A mapping $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ is **one-to-one** if each $\vec{b}$ in the codomain $\mathbb{R}^m$ is the image (or result) of **at most one** $\vec{x}$ in the domain $\mathbb{R}^n$.
These definitions connect transformations back to the matrix equation and existence and uniqueness.

One-to-one and onto meaning

If a mapping is onto, that means every value in its domain maps to something else on the new codomain. ("No Value Left Behind!") That means there shall always exist a solution $\vec{x}$ for $T(\vec{x}) = \vec{b}$, right? Well, $T(\vec{x}) = A\vec{x}$. And if $T(\vec{x})$ always has a solution, so does $A\vec{x}$. Remember what $A\vec{x}$ always having a solution means? The columns of its matrix A span the codomain $\mathbb{R}^m$.

If a mapping $T(\vec{x}) = \vec{b}$ is one-to-one, which means it has one unique solution, then similarly, $A\vec{x} = \vec{b}$ must also have a unique solution. And that means the columns of A are linearly independent, which is true because that's also a way to say a matrix's columns form a unique solution.

Now, practically speaking, we need to use a few evaluating techniques to determine whether transformations are onto, one-to-one, both, or neither.

Implications of one-to-one and onto

Onto:

- Is there a pivot in each row? If yes, it IS onto.
- Do the columns of A span the codomain $\mathbb{R}^m$? If yes, it IS onto.
- Is $m < n$ (i.e. the number of rows is less than the number of columns)? If yes, it is NOT onto.

One-to-one:

- Is there a pivot in each column? If yes, it IS onto.
- Are the columns of A linearly independent? If yes, it IS one-to-one.
- Is $m > n$ (i.e. the number of rows is greater than the number of columns)? If yes, it is NOT one-to-one.
- Are there any free variables? If yes, it is NOT one-to-one.

In summary

- **Onto** mapping (also called **surjective**): does a solution exist? If yes, then the transformation is onto. More precisely, T maps the domain $\mathbb{R}^m$ onto the codomain $\mathbb{R}^n$.
- **One-to-one** mapping (also called **injective**): are there infinitely many solutions? (i.e. are there any free variables in the equation?) If yes, then the transformation is NOT one-to-one.

If a transformation is one-to-one, then the columns of A will be linearly independent.
If there is a pivot in every **row** of a standard matrix A, then T is **onto**.
If there is a pivot in every **column** of a standard matrix A, then T is **one-to-one**.

Section 11

Summary of Chapter 1: Ways to represent existence and uniqueness

Existence

The following concepts are related to or show **existence** of a solution:

- If a pivot exists in every row of a matrix.
- If a system is consistent.
- If weights exist for a linear combination formed from the columns of the coefficient matrix to equate the solution.
- If the solution is in the span of a set of vectors, usually the set of vectors in a linear combination or the columns of the coefficient matrix.
- If there exists a solution for $A\vec{x} = \vec{b}$.
- If a transformation is onto or surjective.

Uniqueness

The following concepts are related to or show **uniqueness** of a solution:

- If a pivot exists in every column of a matrix.
- If there are no free variables (solely basic variables) in a solution.
- If a homogeneous linear system $A\vec{x} = \vec{0}$ has only the trivial solution where $\vec{x} = 0$ is the only solution to $A\vec{x} = \vec{0}$.
- If a solution can be expressed in parametric vector form.
- If a set of vectors is linearly independent.
- If a transformation is one-to-one or injective.