Vector Spaces and Systems of Equations

Lecture 12

Termeh Shafie

what is a vector space?

A set V is called a vector space, if it is equipped with the operations of addition and scalar multiplication in such a way that the 'usual rules' of arithmetic hold. The elements of V are generally regarded as vectors.

The 'rules' (axioms) to hold true for all vectors $\vec{u},\ \vec{v},\ \vec{w} \in V$ and scalars a, b are the following

- Closure: $\vec{u} + \vec{v} \in V, a\vec{v} \in V$
- Commutative: $\vec{u} + \vec{v} = \vec{v} + \vec{u}$
- Associativity: $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$
- Identity for Addition: There is a zero vector $\vec{0} \in V$ such that $\vec{u} + \vec{0} = \vec{u}$
- Distributive Property: $a(\vec{u} + \vec{v}) = a\vec{u} + a\vec{v}$ and $(a + b)\vec{u} = a\vec{u} + b\vec{u}$
- Associativity of Scalars: $(ab)\vec{u} = a(b\vec{u})$
- Multiplicative Identity: $1\vec{u} = \vec{u}$

what is a vector space?

Vector spaces are the foundation for:

- Linear transformations and matrices
- Eigenvalues, eigenvectors, and diagonalization
- Applications in
  - physics (e.g., quantum mechanics)
  - engineering (signal processing)
  - computer science (machine learning, graphics)

examples of vector spaces

- Real n-dimensional space $\mathbb{R}^n$:
  The set of all n-tuples of real numbers (e.g., $\mathbb{R}^2, \mathbb{R}^3$) with standard addition and scalar multiplication
  - Example: $\vec{u} = [1,2,3], \vec{v} = [4,5,6]$ and $2\vec{u} = [2,4,6]$
- Polynomials of degree n or less P_n:
  The set of all polynomials of degree $\leq n$ with real coefficients
  - Example: $P_2 = \{a_0 + a_1x + a_2x^2 \mid a_0, a_1, a_2 \in \mathbb{R}\}$
- Matrices of fixed size $M_{m\times n}$:
  The set of all $m \times n$ matrices with real (or complex) entries.
  - Example:The set of 2×2 matrices $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, B = \begin{bmatrix} 0 & 5 \\ 1 & -2 \end{bmatrix}$.

examples that are not vector spaces

- Vectors without zero vector:
  The set of all n-tuples of real numbers (e.g., $\mathbb{R}^2, \mathbb{R}^3$) with standard addition and scalar multiplication
  - Example: the set $V = \{\vec{u} \in \mathbb{R}^2 \mid u_1 + u_2 = 1\}$ because $\vec{0} \notin V$, so it is not a vector space.
- Subset of $\mathbb{R}^n$ closed under addition but not scalar multiplication
  - Example: $W = \{\vec{u} \in \mathbb{R}^2 \mid u_1 \geq 0, u_2 \geq 0\}$ because
    $\vec{u} = [1,1] \in W$ but $-1\vec{u} = [-1,-1] \notin W$
- Finite set of vectors
  - Example: $F = \{[1,0], [0,1]\}$ because
    finite sets of vectors are generally not closed under addition and scalar multiplication
- Set of matrices without closure
  - Example: $H = \{A \in M_{2\times 2} \mid \det(A) = 1\}$ because
    adding two matrices in H does not necessarily result in another matrix with $\det(A) = 1$

examples that are not vector spaces

exercise 1

Why is the set of polynomials of degree n not a vector space?

linear combinations

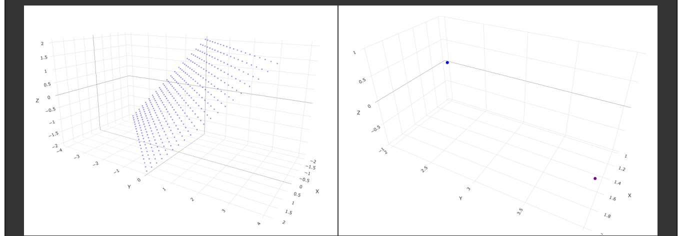
We say that $\vec{v}$ is a linear combination of $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n}$, if there exist scalars $x_1, x_2, \ldots, x_n$ such that $\vec{v} = x_1\vec{v_1} + x_2\vec{v_2} + \ldots + x_n\vec{v_n}$.

- A linear combination means we add (or subtract) scalar multiples of vectors to get a new vector
- Because of the rules of vector addition, any such linear combination will be in the vector space
- Geometrically, the linear combinations of a nonzero vector form a line. The linear combinations of two nonzero vectors form a plane, unless the two vectors are collinear, in which case they form a line.

The set of all linear combinations of vectors $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n}$ is denoted by $span(\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n})$ and called the linear span of these vectors.

linear combinations: planes and lines

- Geometrically, the linear combinations of a nonzero vector form a line. The linear combinations of two nonzero vectors form a plane, unless the two vectors are collinear, in which case they form a line.



linear combinations

Let A denote the matrix whose columns are the vectors $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n}$. Expressing $\vec{v} = x_1\vec{v_1} + x_2\vec{v_2} + \ldots + x_n\vec{v_n}$ as a linear combination of the given vectors is then equivalent to solving the linear system $Ax = \vec{v}$.

- Determining whether a given vector is in the linear span of a given set of vectors, and finding coefficients for linear combinations essentially means solving a system of linear equations (*we'll return to this later*)
- Because of the rules of vector addition, any such linear combination will be in the vector space.

the span

- Given some vectors we now ask: what is the set of all vectors that you can get from a linear combination of these specific vectors?
- The answer is "*the span of these vectors*". Because of the rules of vector addition, any such linear combination will be in the vector space.

> The set of all linear combinations of vectors $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n}$ is denoted by $span(\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n})$ and called the linear span of these vectors.

example

What is the span of the vectors $\vec{v_1} = \begin{bmatrix}1\\0\end{bmatrix}, \vec{v_2} = \begin{bmatrix}1\\1\end{bmatrix}$?

We can actually get to any other vector in the plane, by taking a linear combination of $\vec{v_1}$ and $\vec{v_2}$.
In other words, the span of $\vec{v_1}$ and $\vec{v_2}$ is $\mathbb{R}^2$, i.e. the whole plane.

the span

Question 1: Geometrically, what is the span of a single vector?
Answer 1: It's a line.
Question 2: What is the span of two vectors?
Answer 2: It could be a plane, but in special cases, it could actually be a line. This will happen if one vector is just the scalar multiplication of the other: $\vec{v_2} = k\vec{v_1}$. In that case, we can never "leave" the line.

example

You expect two vectors $\vec{v_1} = \begin{bmatrix}1\\2\end{bmatrix}, \vec{v_2} = \begin{bmatrix}2\\4\end{bmatrix}$ to span the entire 2D plane ($\mathbb{R}^2$), but in fact, they only span a subspace of the plane – a single line. Can you see why?

This leads us to a very important concept in linear algebra: linear dependence and independence

linear independence

- Two equivalent (but informal) definitions of linear dependency for a set of vectors:
  1. One of more of the vectors can be expressed as a linear combination of the remaining vectors
  2. One or more of the vectors is inside the span of the remaining vectors
- Note: Linear dependence is a property of the set! (i.e. an entire set of vectors can be linearly dependent simply because two of its vectors are dependent on each other)

> A set K of n vectors is called linearly independent iff:
> $$\sum_{i=1}^{n} a_i\vec{v_i} = 0 \quad iff \quad a_1 = a_2 = \cdots = a_n = 0$$

- We call a set of linearly independent vectors that span a vector space a basis of that vector space
- The dimension of a vector space is equal to the number of vectors in its basis

linear independence

example

Let $\vec{v_1} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$, $\vec{v_2} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$, $\vec{v_3} = \begin{bmatrix} -2 \\ 1 \\ -4 \end{bmatrix}$. Are they linearly dependent?

Let's express each vector as the linear combination of the other two $\vec{v_3} = a\vec{v_2} + b\vec{v_1}$

$\implies \begin{bmatrix} -2 \\ 1 \\ -4 \end{bmatrix} = a \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + b \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$. This gives us three equations, one for each entry:

$$\begin{aligned} -2 &= 0a + 1b \rightarrow b = -2 \\ 1 &= 1a + 0b \rightarrow a = 1 \\ -4 &= 1a + 2b \rightarrow -3 = 1 - 4 \end{aligned}$$

Is this enough to say that they are linearly independent? No, because we have to show that any vector in this set can be expressed as a linear combination of the rest of the vectors.

linear independence

- Given 3 vectors $\vec{v_1}, \vec{v_2}, \vec{v_3}$, they are called linearly independent if and only if none of them is a linear combination of the others:

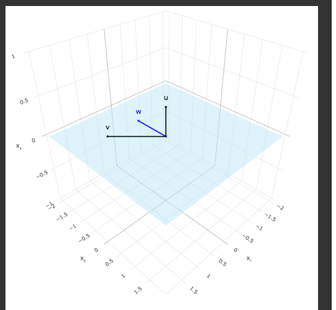
$$\vec{v_1} \neq a\vec{v_2} + b\vec{v_3} \text{ for any } a, b \in \mathbb{R}$$
$$\vec{v_2} \neq a\vec{v_1} + b\vec{v_3} \text{ for any } a, b \in \mathbb{R}$$
$$\vec{v_3} \neq a\vec{v_1} + b\vec{v_2} \text{ for any } a, b \in \mathbb{R}$$

- This is equivalent to saying that:

$$a\vec{v_1} + b\vec{v_2} + c\vec{v_3} = 0 \quad \textit{iff} \quad a = b = c = 0$$

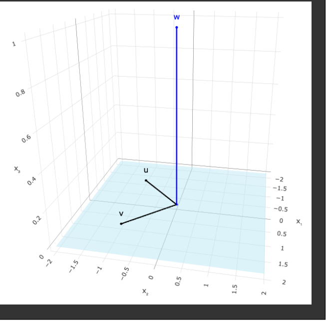
linear independence and spanning vectors

- $\vec{w}$ is in $span(\vec{u}, \vec{v})$or the plane spanned by $(\vec{u}, \vec{v})$
- $\vec{w}$ is a linear combination of $(\vec{u}, \vec{v})$, so $(\vec{u}, \vec{v}, \vec{w})$ is not linear independent.



linear independence and spanning vectors

- Since $\vec{w}$ is not in $span(\vec{u}, \vec{v})$, $(\vec{u}, \vec{v}, \vec{w})$ is linear independent.



rank

The rank of a matrix is defined as the number of linearly independent columns (or rows) of a matrix. If all of the columns are independent, we say that the matrix is of full rank. We denote the rank of matrix A by as $rank(A)$.

- Full rank $\Longleftrightarrow$ nonsingular $\Longleftrightarrow$ invertible. All of these imply that A^{-1} exists.
- If a matrix is not of full rank, it is not invertible; i.e., it is singular.

example

Let $A = \begin{bmatrix} 2 & -4 \\ 3 & -6 \end{bmatrix}$. The second column is twice the first column so the rank of the matrix is 1.

It is not full rank.

system of linear equations

- A system of linear equations is when we have two or more linear equations working together
- For the equations to "work together" they share one or more variables

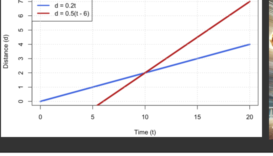
example

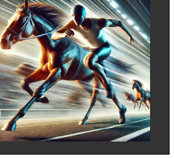
It's a race between you and a horse. You can run 0.2 km every minute. The horse can run 0.5 km every minute. But it takes 6 minutes to saddle the horse. Let d denote distance and t denote time.

We have a system of equations (that are linear):

$$d = 0.2t$$
$$d = 0.5(t - 6)$$

Where do you get caught?





system of linear equations

- Matrices are particularly useful when solving systems of equations
- Consider following system:

$$\begin{aligned} 2x - y + 3z &= 9 \\ x + 4y - 5z &= -6 \\ x - y + z &= 2 \end{aligned}$$

- This system of equations can be represented in matrix form as follows:

$$A = \begin{bmatrix} 2 & -1 & 3 \\ 1 & 4 & -5 \\ 1 & -1 & 1 \end{bmatrix} \quad \vec{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad \vec{b} = \begin{bmatrix} 9 \\ -6 \\ 2 \end{bmatrix} \implies A\vec{x} = \vec{b}$$

- Techniques for solving:
  - substitution
  - elimination
  - matrix inversion
  - Cramer's rule

we use this example in the following to illustrate each

substitution

- Choose the easiest variable to solve for and plug this expression for the variable into the other two equations you did not yet use
- Three things can happen:

1. There is the same number of equations as unknowns $\implies$ uniquely determined
2. There are more unknowns than equations i.e. infinite number of solutions $\implies$ underdetermined
3. There are more equations than unknowns (equations are contradictory) $\implies$ overdetermined

substitution

$$\begin{aligned} 2x - y + 3z &= 9 & (1) \\ x + 4y - 5z &= -6 & (2) \\ x - y + z &= 2 & (3) \end{aligned}$$

$$A = \begin{bmatrix} 2 & -1 & 3 \\ 1 & 4 & -5 \\ 1 & -1 & 1 \end{bmatrix} \quad \vec{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad \vec{b} = \begin{bmatrix} 9 \\ -6 \\ 2 \end{bmatrix} \implies A\vec{x} = \vec{b}$$

- We solve equation (3) and substitute it into equation (2) and (1): $x = y - z + 2$

$\implies 2(y - z + 2) - y + 3z = 9 \implies 2y - 2z + 4 - y + 3z = 9 \implies y + z = 5 \quad (4)$

$\implies (y - z + 2) + 4y - 5z = -6 \implies y - z + 2 + 4y - 5z = -6 \implies 5y - 6z = -8 \quad (5)$

- Now solve equations (4) and (5):

From (4) we have $y = 5 - z$ which is substituted into equation (5)

$5(5 - z) - 6z = -8 \implies 25 - 5z - 6z = -8 \implies -11z = -33 \implies \boxed{z = 3}$

Substitute back into y: $\boxed{y = 5 - 3 = 2}$

and finally from $x = y - z + 2$ we get $\boxed{x = 2 - 3 + 2 = 1}$

Gaussian elimination

- We eliminate one variable by combining equations

$$\begin{aligned} 2x - y + 3z &= 9 \quad (1) \\ x + 4y - 5z &= -6 \quad (2) \\ x - y + z &= 2 \quad (3) \end{aligned} \qquad A = \begin{bmatrix} 2 & -1 & 3 \\ 1 & 4 & -5 \\ 1 & -1 & 1 \end{bmatrix} \quad \vec{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad \vec{b} = \begin{bmatrix} 9 \\ -6 \\ 2 \end{bmatrix} \implies A\vec{x} = \vec{b}$$

- Eliminate x:
  - subtracting equation (3) from (2): $(x + 4y - 5z) - (x - y + z) = -6 - 2 \implies 5y - 6z = -8 \quad (4)$
  - subtract $2\times$ equation (3) from equation (1): $(2x - y + 3z) - 2(x - y + z) = 9 - 2(2) \implies y + z = 5 \quad (5)$
- Solve for y and z:
  - from equation (5) we get a new equation: $y = 5 - z$ (6) which is substituted into (4):
    $5(5 - z) - 6z = -8 \implies 25 - 5z - 6z = -8 \implies -11z = -33 \implies \boxed{z = 3}$
  - from (6) we get $\boxed{y = 5 - 3 = 2}$ and from (3) we get $x = y - z + 2 \implies \boxed{x = 2 - 3 + 2 = 1}$

matrix inversion

- $A\vec{x} = \vec{b}$ can be used to create an **augmented matrix** by taking our A and adding the column $\vec{b}$ to it as a new column on the right
- We can then attempt to get the identity matrix in all but the last column
- If we can do this, then we get what is called reduced **row echelon form**
- We can do this, then we can read off the answers to the system by returning to equation form, since each row will provide the value of one of the variables
- If the rank of A is equal to its number of rows, then the system has at least one solution (i.e. each equation is linearly independent, and so cannot produce a contradiction)
- If the rank of A is equal to its number of columns, then the system has at most one solution (i.e. the number of independent rows, corresponding to equations, is at least as great as the number of variables)
- When the matrix A is square and nonsingular, we can invert the matrix to figure out what the unique solution is to the equation: $\vec{x} = A^{-1}\vec{b}$

matrix inversion

$$\begin{aligned} 2x - y + 3z &= 9 \quad (1) \\ x + 4y - 5z &= -6 \quad (2) \\ x - y + z &= 2 \quad (3) \end{aligned} \qquad A = \begin{bmatrix} 2 & -1 & 3 \\ 1 & 4 & -5 \\ 1 & -1 & 1 \end{bmatrix} \quad \vec{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad \vec{b} = \begin{bmatrix} 9 \\ -6 \\ 2 \end{bmatrix} \implies A\vec{x} = \vec{b}$$

- Find the determinant of A to make sure it's invertible (via cofactor expansion along first row): $\det(A) = -11$
- This is not zero, the matrix is nonsingular, and we can invert it
- To invert the matrix we compute 9 minors:
  $M_{11} = -1,\ M_{12} = 6,\ M_{13} = -5,\ M_{21} = 2,\ M_{22} = -1,\ M_{23} = -1,\ M_{31} = -7,\ M_{32} = -13,\ M_{33} = 9$
- Now use the formula for the inverse from the previous lecture $A^{-1} = \frac{1}{|A|}C^T$ where $C_{ij} = (-1)^{i+j}M_{i,j}$
- This gives: $\mathbf{A}^{-1} = \frac{1}{-11} \cdot \begin{bmatrix} -1 & -2 & -7 \\ -6 & -1 & 13 \\ -5 & 1 & 9 \end{bmatrix}$
- And finally we multiply with $\vec{b}$ to get the solution: $\vec{x} = A^{-1}\vec{b} = \frac{1}{-11} \cdot \begin{bmatrix} -1 & -2 & -7 \\ -6 & -1 & 13 \\ -5 & 1 & 9 \end{bmatrix} \cdot \begin{bmatrix} 9 \\ -6 \\ 2 \end{bmatrix} = \boxed{\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}}$

Cramer's rule

- only works when there an equal number of equations and unknowns (i.e. a square matrix A) and A is nonsingular
- this rules states that we can solve for $\vec{x}$ using the formula:

$$x_i = \frac{|B_i|}{|A|}$$

where the matrix B_i is formed by replacing the ith column of A (the column corresponding to variable x_i) with $\vec{b}$

1. take the determinant of A, to check to make sure that we can apply this rule and determine the denominator of each x_i
2. we form the B_i by replacing each of the three columns by $\vec{b}$
3. compute the determinants of the matrices B_i
4. apply Cramer's rule according to formula

Cramer's rule

$$\begin{aligned} 2x - y + 3z &= 9 & (1)\\ x + 4y - 5z &= -6 & (2)\\ x - y + z &= 2 & (3) \end{aligned}$$

$$A = \begin{bmatrix} 2 & -1 & 3 \\ 1 & 4 & -5 \\ 1 & -1 & 1 \end{bmatrix} \quad \vec{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad \vec{b} = \begin{bmatrix} 9 \\ -6 \\ 2 \end{bmatrix} \implies A\vec{x} = \vec{b}$$

- we already know the determinant of A: $\det(A) = -11$
- form the B_i by replacing each of the three columns by $\vec{b}$

$$B_1 = \begin{bmatrix} 9 & -1 & 3 \\ -6 & 4 & -5 \\ 2 & -1 & 1 \end{bmatrix}, B_2 = \begin{bmatrix} 2 & 9 & 3 \\ 1 & -6 & -5 \\ 1 & 2 & 1 \end{bmatrix}, B_3 = \begin{bmatrix} 2 & -1 & 9 \\ 1 & 4 & -6 \\ 1 & -1 & 2 \end{bmatrix}$$

- compute determinant for each B_i: $\det(B_1) = -11$, $\det(B_2) = -22$, $\det(B_3) = -33$
- apply Cramer's rule:

$$x = \frac{\det(B_1)}{\det(A)}, \quad y = \frac{\det(B_2)}{\det(A)}, \quad z = \frac{\det(B_3)}{\det(A)} \implies x = \frac{-11}{-11} = 1, \quad y = \frac{-22}{-11} = 2, \quad z = \frac{-33}{-11} = 3$$

Cramer's rule

exercise 2

Use Cramer's Rule to solve the linear system where $A\vec{x} = \vec{b}$ where

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$