## Contents

# 1 Chapter 1

The entire course is about the equation

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

But we first have to begin with the basic example, a system of equations:

$$\begin{cases} x - y = -1 \\ -x + 3y = 3 \end{cases}$$

Each of these describes a line, and in this case we can see that they intersect at a point, the solution to the system of equations. In this case we only have 1 solution, but we can have systems that have 0 solutions as well as systems that have infinite solutions. To obtain an infinite number of solutions, we can have the same line twice:

$$\begin{cases} x - 2y = -1 \\ -x + 2y = 1 \end{cases}$$

These two lines meet in infinitely many points (because they're the same line). In 2d, the only other case is where the lines are parallel, and there are no solutions/points of intersections:

$$\begin{cases} x - 2y = -1 \\ x - 2y = -3 \end{cases}$$

Two systems of equations are said to be **equivalent** if they have the same solution set. Note that a system of linear equations will always have either 1 unique solution, no solutions, or infinitely many solutions.

**Definition 1.1.** A system of linear equations is **consistent** if it has at least one solution, and **inconsistent** if it has no solutions.

The **coefficient matrix** of a system of equations is a matrix that represents the coefficients of the variables in the system. This provides a method of easily writing the system of equations. The **augmented matrix** is the coefficient matrix with an extra column, in which the constants are placed. The augmented matrix can completely describe a system of equations. The idea is to solve a linear system by conducting algebraic operations on the equations involved. These operations are easier to keep track of if we do computations on the augmented matrix itself. To help in this, we have the elementary row operations:

1. Replacement: Replace one row by the sum of itself and a multiple of another row.
2. Interchange: Swap two rows.
3. Scaling: Multiply all entries in a row by a nonzero constant.

**Definition 1.2.** If there exists a sequence of row operations taking an augmented matrix $A$ to a matrix $B$, we say that $A$ and $B$ are **row equivalent**.

If the augmented matrices of two linear systems are row equivalent, then the two systems have the same solution set (they're equivalent). This is the entire reason we do row operations.

**Definition 1.3.** A rectangular matrix is in **row echelon form** if it has the following three properties:

1. All nonzero rows are above any rows of all zeros.
2. Each leading entry of a row is in a column to the right of the leading entry of the row above it.
3. All entries in a column below a leading entry are zeros.

If a matrix is in row echelon form, and additionally also satisfies the following two properties, it is in **reduced row echelon form**:

1. The leading entry of each nonzero row is 1.
2. Each leading 1 is the only nonzero entry in its column.

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

**Definition 1.4.** A **pivot position** in a matrix $A$ is a location in $A$ that corresponds to a leading 1 in the rref of $A$. A **pivot column** is a column of $A$ that contains a pivot position.

*Remark.* Almost every question in this course that can be asked of you is solvable by putting some matrix into rref.

**Theorem 1.2.** *A linear system is consistent iff the rightmost column of the augmented matrix is not a pivot column. i.e iff a row echelon form of the matrix does not have a row of the form* $[0\ 0\ 0\ 0 \dots\ 0\ b]$ *where* $b \neq 0$.

*If a linear system is consistent, then the solution set contains either a unique solution, or infinitely many solutions. We get a unique solution when there are no free variables, and we get infinitely many solutions if there are free variables.*

**Definition 1.5.** If $\vec{v_1}, \dots, \vec{v_p}$ are in $\mathbb{R}^n$, then the set of all linear combinations of $\vec{v_1}, \dots, \vec{v_p}$ is denoted by $\text{Span}\{\vec{v_1}, \dots, \vec{v_p}\}$, and is called the subset of $\mathbb{R}^n$ spanned by $\vec{v_1}, \dots, \vec{v_p}$.

We can now ask whether $\vec{b}$ is in the span of the column vectors of $A$ when talking about a vector equation.

**Theorem 1.3.** *Let* $A_{m\times n}$. *Then the following are equivalent:*

1. *For each* $\vec{b}$ *in* $\mathbb{R}^m$, *the equation* $A\vec{x} = \vec{b}$ *has a solution.*
2. *Each* $\vec{b} \in \mathbb{R}^m$ *is a linear combination of the columns of* $A$.
3. *The columns of* $A$ *span* $\mathbb{R}^m$ *(span*$\{\vec{a_1}, \dots, \vec{a_n}\} = \mathbb{R}^m$*)*
4. $A$ *has a pivot position in every row.*

**Theorem 1.4.** *If* $A_{m\times n}$, $\vec{u}, \vec{v} \in \mathbb{R}^n$, $c \in \mathbb{R}$, *then*

1. $A(\vec{u}+\vec{v}) = A\vec{u} + A\vec{v}$
2. $Ac\vec{u} = A(c\vec{u})$

**Definition 1.6.** A system of linear equations is said to be **homogeneous** if it can be weitten as $A\vec{x} = \vec{0}$. Note that a homogeneous system always has at least 1 solution ($\vec{x} = \vec{0}$, which is known as the trivial solution), and we typically are looking for nontrivial solutions.

Looking at the equation $A\vec{x} = \vec{0}$, we can say that this has a nontrivial solution iff the equation has at least one free variable. Here's an example. Lets determine if the following homogeneous system has a nontrivial solution:

$$\begin{bmatrix} 3 & 5 & -4 & 0 \\ -3 & -2 & 4 & 0 \\ 6 & 1 & -8 & 0 \end{bmatrix}$$

Row reducing:

$$\begin{bmatrix} 3 & 5 & -4 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & -9 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 3 & 5 & -4 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 3 & 0 & -4 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -4/3 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Lets look at the solution to this:

$$x_1 - \frac{4}{3}x_3 = 0 \qquad x_2 = 0$$

We can write this in a different form (parametric vector form):

$$\vec{x} = \begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix} x_3$$

Any solution to $A\vec{x} = \vec{0}$ is a scalar multiple of this vector. Looking at parametric vector form in general:

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

$$A = \begin{bmatrix} 3 & 5 & -4 & 0 \\ -3 & -2 & 4 & 0 \\ 6 & 1 & -8 & 0 \end{bmatrix} \qquad \vec{b} = \begin{bmatrix} 7 \\ -1 \\ -4 \end{bmatrix}$$

Making the augmented matrix and row reducing, we get the following solutions

$$x_1 - \frac{4}{3}x_3 = -1 \qquad x_2 = 2$$

This gives us:

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} + \begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix} x_3$$

In general, the vector of scalars is known as $\vec{p}$, and $x_3$ is known as a parameter.

**Theorem 1.5.** *Suppose $A\vec{x} = \vec{b}$ is consistent for some given $\vec{b}$, and let $\vec{p}$ be a solution. Then the solution set of $A\vec{x} = \vec{b}$ is the set of all vectors of the form $\vec{w} = \vec{p} + \vec{v_h}$, where $\vec{v_h}$ is any solution to the homogenous equation $A\vec{x} = \vec{0}$.*

*Proof.*

$$\begin{aligned} A\vec{w} = A(\vec{p} + \vec{v_h}) &= A\vec{p} + A\vec{v_h} \\ &= \vec{b} + \vec{0} = \vec{b} \end{aligned}$$

This tells us that $\vec{w}$ is a solution. This isn't the full proof, as this only goes in one direction, but the other direction isn't so bad either. □

Lets talk about the notion of linear independence, something that is very important in linalg.

**Definition 1.7.** A set of vectors $\{\vec{v_1}, \ldots, \vec{v_n}\}$ in $\mathbb{R}^n$ are said to be **linearly independent** if the equation

$$x_1\vec{v_1} + x_2\vec{v_2} + \cdots + x_n\vec{v_n} = \vec{0}$$

has only the trivial solution. A set of vectors are said to be **linearly dependent** if there exists weights $c_1, \ldots, c_n$ (not all 0), s.t.

$$c_1\vec{v_1} + c_2\vec{v_2} + \cdots + c_n\vec{v_n} = \vec{0}$$

(Basically if there's a nontrivial solution to the equation). This is called the linear dependence relation.

This gives us another way to express things that we've been talking about. If we want to check whether 3 vectors are linearly independent, we can set up the homogenous equation $A\vec{x} = \vec{0}$, where $A$ is made up of the column vectors that we're given, and we can solve this. If we can only see the trivial solution, then we know that they are linearly inependent. To find a linear dependence relation, we can just find a specific solution.

Note that if we have a set of two vectors, linear independence is just if they are scalar multiples of each other, but this only works for sets of 2 vectors. For higher numbers:

**Theorem 1.6.** *A set $S = \{\vec{v_1}, \ldots, \vec{v_p}\}$ of two or more vectors is linearly dependent iff at least one of the vectors in $S$ is a linear combination of the others. In fact, if $S$ is linearly dependent and $\vec{v_1} \neq \vec{0}m$ then some $\vec{v_j}$ with $j > 1$ is a linear combination of the preceding vectors $\vec{v}_1, \ldots, \vec{v}_{j-1}$.*

*Note that this does not mean that every vector in $S$ is a linear combination of the others. It just means that at least one of them is.*

**Theorem 1.7.** *If a set contains more vectors than there are entries in each vector, then the set is linearly dependent. That is, any set of vectors $\{\vec{v_1}, \ldots, \vec{v_p}\} \in \mathbb{R}^n$ is linearly dependent if $p > n$.*

**Theorem 1.8.** *If a set contains the zero vector, then the set is linearly dependent.*

*Proof.* Suppose $\vec{v}_1 = \vec{0}$. Then $1\vec{v}_1 + 0\vec{v_2} + \cdots + 0\vec{v_p} = \vec{0}$. This is a nontrivial linear dependence relation, so we have linear dependence. □

## 1.1 Linear Transformations

**Definition 1.8.** $T : \mathbb{R}^n \to \mathbb{R}^m$ is a transformation (or function, mapping, map) that sends a vector in $\mathbb{R}^n$ to a vector in $\mathbb{R}^m$ ($\vec{x} \mapsto T(\vec{x})$).

A transformation $T$ is linear if it satisfies two conditions:

1. $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v})\ \forall \vec{u}, \vec{v}$ in the domain of $T$
2. $T(c\vec{u}) = cT(\vec{u})\ \forall \vec{u}$ in the domain of $T$.

Essentially, a transformation is linear if it respects linear combinations.

What types of linear transformations do we have? It turns out that matrix transformations are the only finite dimensional linear transforms (integrals and derivatives act in this way but they act on infinite dimensional spaces). In this course, we'll be looking at matrix transformations.

Let's do an example. Let $A$ be

$$A = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}$$

and let's define some vectors

$$\vec{u} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \qquad \vec{b} = \begin{bmatrix} 3 \\ 2 \\ -5 \end{bmatrix}, \qquad \vec{c} = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}$$

Let's define a transformation $T : \mathbb{R}^2 \to \mathbb{R}^3$:

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

If we enact this on a general vector, we get

$$T(\vec{x}) = \begin{bmatrix} x_1 - 3x_2 \\ 3x_1 + 5x_2 \\ -x_1 + 7x_2 \end{bmatrix}$$

If we enact this transformation on $\vec{u}$:

$$T(\vec{u}) = A\vec{u} = \begin{bmatrix} 5 \\ 1 \\ -9 \end{bmatrix}$$

One way of saying this is that the result is the image of $\vec{u}$ under the transformation $T$. We can also ask to find an $\vec{x}$ in $\mathbb{R}^2$ whose image under $T$ is $\vec{b}$ (solve $T(\vec{x}) = \vec{b}$ for $\vec{x}$). To do this, we can just solve the matrix equation $A\vec{x} = \vec{b}$. We can also check to see if there is more than one $\vec{x}$ such that $T(\vec{x}) = \vec{b}$. We see that if we solve the system of equations, we see that there is only one unique solution, which means that there is only one $\vec{x}$. Had we found any free variables, we could have multiple vectors that would satisfy the equation. The last question we want to ask is whether $\vec{c}$ is in the range of $T$. We want to check whether there exists an $\vec{x}$ such that $A\vec{x} = \vec{c}$, which is the same as asking whether $T(\vec{x}) = \vec{c}$ has a solution, or whether $A\vec{x} = \vec{c}$.

If $T$ is a linear transformation, then

$$T(c\vec{u} + d\vec{v}) = cT(\vec{u}) + dT(\vec{v})$$

If $T$ is linear, then

$$T(\vec{0}) = T(\vec{u} - \vec{u}) = T(\vec{u}) + -1T(\vec{u}) = T(\vec{u}) - T(\vec{u}) = \vec{0}$$

Note that the input zero vector and the output zero vector could have different dimensions. Also note that the converse of this isn't true, so if the transformation sends the zero vector to the zero vector, the transformation is not necessarily linear.

*Remark.* Under some mild assumptions, every linear transformation is a matrix transformation.

What this means, is that associated to every linear transformation, there is a matrix.

**Theorem 1.9.** *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. then there exists a unique matrix $A$ such that*

$$T(\vec{x}) = A\vec{x} \qquad \forall\, \vec{x} \in \mathbb{R}^n$$

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

$$A = [T(\vec{e}_1) \;\; T(\vec{e}_2) \;\; \dots \;\; T(\vec{e}_n)]$$

*You can think of the $\vec{e}$ vectors as the standard basis vectors for the space (in the case of $\mathbb{R}^3$, they're $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$). This is called the **standard matrix** for the transformation $T$.*

Let's do an example of this. Find the standard matrix $A$ for the dilation transformation $T(\vec{x}) = 3\vec{x}$ ($\vec{x} \in \mathbb{R}^2$). To do this, we take the column vectors of the two by two identity matrix, enact the transformation on them, and then splice the outputs together into one matrix:

$$A = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$$

Let's do another one. Let $T$ be the transformation that rotates every point in $\mathbb{R}^2$ about the origin counterclockwise by some angle $\phi$. Looking at $T(\begin{bmatrix} 1 \\ 0 \end{bmatrix})$, we see that this rotates to $\begin{bmatrix} \cos\phi \\ \sin\phi \end{bmatrix}$, and $T(\begin{bmatrix} 0 \\ 1 \end{bmatrix})$ rotates to $\begin{bmatrix} -\sin\phi \\ \cos\phi \end{bmatrix}$, leaving us with the matrix

$$A = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}$$

In particular, if $\phi = \frac{\pi}{2}$, we see that

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

**Definition 1.9.** A map $T : \mathbb{R}^n \to \mathbb{R}^m$ is **onto** $\mathbb{R}^m$ (or **surjective**) if each $\vec{b} \in \mathbb{R}^m$ is the image of at least one $\vec{x} \in \mathbb{R}^n$.

**Definition 1.10.** A map $T : \mathbb{R}^n \to \mathbb{R}^m$ is **one to one** (or **injective**) of each $\vec{b} \in \mathbb{R}^m$ is the image of at most one $\vec{x} \in \mathbb{R}^n$.

Note that these are not mutually exclusive, we can have all combinations of surjective and injective for a transformation.

**Theorem 1.10.** *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then $T$ is 1-1 iff $T(\vec{x}) = \vec{0}$ has only the trivial solution.*

**Theorem 1.11.** *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Let $A$ be its standard matrix. Then*

1. *$T$ maps $\mathbb{R}^n$ onto $\mathbb{R}^m$ iff the columns of $A$ span $\mathbb{R}^m$.*
2. *$T$ is 1-1 iff the columns of $A$ are linearly independent.*

Moving back to the TFAE we talked about previously, we see that the confition for being onto is equivalent to some pretty easy things to do, like checking whether $A$ has a pivot position in every row.

# 2 Chapter 2

The **zero matrix** ($0_{m\times n}$) is the $m$ by $n$ matrix with all entries 0. The $n \times n$ identity matrix is the matrix of all zeros and 1s along the diagonal. Square matrices are matrices where $m = n$. What is really important in matrix algebra is matrix multiplication. If we have $A_{m\times n}$ and $B_{n\times p}$, then

$$AB = A[\vec{b}_1\ \vec{b}_2 \dots\ \vec{b}_p]$$

That is, $AB$ is the $m \times p$ matrix whose columns are $A\vec{b_1}, \dots, A\vec{b}_p$.

The inverse of a $2 \times 2$ matrix can be given via

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

Where the quantity $ad - bc$ is known as the determinant, and is denoted $det(A)$. There is no simple rule for the higher dimensions of matrices.

**Theorem 2.1.** *If $A_{n\times n}$ is invertible, then for every $\vec{b} \in \mathbb{R}^n$, the equation $A\vec{x} = \vec{b}$ has the unique solution $\vec{x} = A^{-1}\vec{b}$.*

**Theorem 2.2.** *If $A$ is invertible, then $A^{-1}$ is invertible, and $(A^{-1})^{-1} = A$. If $A_{n\times n}$, $B_{n\times n}$ are invertible, then so is $AB$, and $(AB)^{-1} = B^{-1}A^{-1}$.*

*Proof.*

$$AB(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$$

□

We can keep track of the three elementary row operations through matrix multiplication:

**Definition 2.1.** An **elementary matrix** is one that is obtained from the identity matrix via a single elementary row operation.

Some examples of this are

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix} \qquad E_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

Note that if we take an arbitrary matrix $A$ and compute $E_1A$, we see that the output is the same as the elementary operation used to compute $E_1$. This lets use describe row operations in terms of matrix multiplication. Also note that each elementary matrix $E$ is invertible, with $E^{-1}$ being the elementary matrix of the same type that reverts the row operation.

**Theorem 2.3.** *$A_{n\times n}$ is invertible iff $A$ is row equivalent to the identity matrix $I_{n\times n}$, and in this case, and sequence of row operations that transforms $A$ into $I$, also transforms $I$ into $A^{-1}$.*

This gives us an algorithm for computing the inverse of $A$ of any size, if it exists:

1. set up the augmented matrix $[A_{n\times n}|I_{n\times n}]$
2. row reduce until you get $[I_{n\times n}|?]$

This ? will be $A^{-1}$. If $A$ is not invertible, we won't be able to get the identity on the left side of the augmented matrix.

**Theorem 2.4.** *The idea of this theorem is that an invertible matrix $A$ is "very nice"! Basically, we have a set of equivalent statements (TFAE): //TODO: put it in here combining f and i, we see that if $A$ is invertible, then the transformation $T : \vec{x} \to A\vec{x}$ is both 1-1 and onto.*

How do we use this theorem? When we are presented with a situation in which you have an invertible matrix $A$, we want to remember this big list of properties that are also true. This is a huge source of T/F questions.

**Definition 2.2.** If $A_{n\times n} = [a_{ij}]$, then $det(A)$ is the sum of $n$ terms of the form $a_{ij}det(A_{ij})$, with plus and minus signs alternating, where the entries $a_{11}, a_{12}, a_{13}, \ldots, a_{1n}$ are from the first row of $A$.

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

Note that this is inductively defined, so to compute the determinant of a 3 by 3, we only need to know how to compute the determinant of a 2 by 2 matrix.

We can also compute determinants via the cofactor expansion method.

**Definition 2.3.** The $(i, j)$-cofactor of $A$ is

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

Thus by definition, $detA = a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n}$. We say that this is the cofactor expansion across the first row.

*Remark.* One massive fact is that we can compute $det(A)$ via a cofactor expansion across any row or down any column.

**Theorem 2.5.** *The determinant of $A_{n\times n}$ can be computed via cofactor expansion across any row or down any column. The expansion across the $i$th row is*

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

*And the expansion down the $j$th column is*

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

**Theorem 2.6.** ***Cramer's Rule****: Let $A_{n\times n}$, $\vec{b} \int \mathbb{R}^n$ be any vector. Let $A_i(\vec{b})$ be the matrix obtained from $A$ by replacing its $i$th column with $\vec{b}$.*

*Let $A_{n\times n}$ be invertible. For any $\vec{b} \in \mathbb{R}^n$, the unique solution to $A\vec{x} = \vec{b}$ has entries given by*

$$x_i = \frac{det(A_i(\vec{b}))}{det(A)}$$

It turns out that there is a general formula for $A^{-1}$, called the adjugate (or classical adjoint):

**Theorem 2.7.** *Let $A$ be an $n \times n$ invertible matrix. Then*

$$A^{-1} = \frac{1}{det(A)}\, adj(A)$$

*Where $adj(A)$ is the adjugate matrix of $A$ and is the matrix of transpose cofactors:*

$$adj(A) = \begin{bmatrix} C_{11} & C_{21} & C_{31} & \dots & C_{n1} \\ C_{12} & C_{22} & & & \\ C_{13} & & \ddots & & \\ \vdots & & & \ddots & \\ C_{1n} & & & & C_{nn} \end{bmatrix}$$

**Theorem 2.8.** *If $A_{2\times 2}$, then the area of the parallelogram determined by the columns of $A$ is $|det(A)|$. If $A_{3\times 3}$, the volume of the parallelpiped determined by the columns of $A$ is $|det(A)|$. This scales to higher dimensions as well. We can construct the shapes by taking the vectors and adding them head to tail, which forms a full shape.*

**Theorem 2.9.** *Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation with standard matrix $A$. If $S$ is a parallelogram in $\mathbb{R}^2$, then the area of the image $T(S)$ is equal to $|det(A)|A_S$, where $A_S$ is the area of $S$. Essentially, the area of the transformed region is the area of the original region scaled by the determinant of the transformation's matrix.*

*This theorem also scales to $n$ dimensions (in $\mathbb{R}^3$, the volume of $T(S)$ is the volume of $S$ scaled by $|det(A)|$). Note that $S$ can be any region with finite area/volume.*

# 3 Chapter 4: Abstract Vector Spaces

The main idea is that we are replacing $\mathbb{R}^n$ with an abstract vector spaces, and we want to study the general properties of these spaces. There is a list of vector space axioms that all vector spaces must follow: closure, commutativity, associativity, existence of a zero vector, existence of a additive inverse, closure under scalar multiples, the distributive property for scalars (both directions), scalar associativity, and finally that scalar multiplication by 1 returns the same vector.

Note that $\mathbb{R}^n$ satisfies all of these properties, and thus is a vector space. Lets look at another one. Let the set $C$ be the set of all polynomials up to degree 2. Remember that this is the set of things of the form

$$a_2x^2 + a_1x + a_0$$

such that $a_2$, $a_1$, and $a_0 \in \mathbb{R}$. If we go through our list, we see that it satisfies all of the conditions, and it is actually $\mathbb{R}^3$. However, one interesting thing is that vector spaces can be infinite dimensional, so we could instead look at the set of all polynomials, regardless of degree.

**Definition 3.1.** A vector space $V$ (abbreviated to v.s.) is a set of objects called vectors that is equipped with two operations, vector addition (+) and scalar multiplication (in our case we say the scalars are real numbers). Both of these satisfy the vector space axioms.

**Definition 3.2.** A subspace of a vector space $V$ is a subset $H$ of $V$ that has 3 properties:

1. the zero vector of $V$ is in $H$.
2. $H$ is closed under vector addition.
3. $H$ is closed under scalar multiplication.

This is a subset of a vs that is a vs in its own right. This definition then allows us to check only these three things. This is better than checking every vs axiom.

What is the simplest possible subspace of a vector space? What about the subspace that consists of only the zero vector? This is known as the zero subspace, $H = \{\vec{0}\}$. Every vs has a zero subspace. Also note that $H = V$ is also technically a subspace of $V$.

If we look at $\mathbb{R}^3$, $\mathbb{R}^2$ is not even a subset of this, but if we choose $\vec{v} = [x\, y\, 0]$, we see that we have a vs that is isomorphic to $\mathbb{R}^2$, but is not $\mathbb{R}^2$.

**Theorem 3.1.** *If $\vec{v}_1, \ldots, \vec{v}_p$ are vectors in a vs $V$, then $Span\{\vec{v}_1, \ldots, \vec{v}_p\}$ is a subspace of $V$.*

We call $H := Span\{\vec{v}_1, \ldots \vec{v}_p\}$ the subspace spanned by the set of vectors, and the subspace is the subspace generated by the vectors.

**Definition 3.3.** The null space of a matrix $Nul(A)$ is the set of all solutions to the homogeneous equation $A\vec{x} = \vec{0}$.

**Theorem 3.2.** *The null space of $A_{m\times n}$ is a subspace of $\mathbb{R}^n$. Equivalently, the set of all solutions to $A\vec{x} = \vec{0}$ is a subpsace of $\mathbb{R}^n$.*

Note that it is very easy to check whether a given vector is in the null space of $A$, we can just look at $A\vec{x}$ and see if it equals $\vec{0}$.

How do we explicitly describe $Nul(A)$? If we are given a matrix, and asked to find a spanning set for the null space of that matrix, we can simple row reduce $[A\ \vec{0}]$, and write the solution out in parametric vector form.

**Theorem 3.3.** *Let $S = \{\vec{v}_1, \dots, \vec{v}_p\}$ be vectors in $V$, and let $H$ be the span of $S$. If one of the vectors in $S$ is a linear combination of the remaining vectors, then the set formed from $S$ by removing the vector still spans $H$. If $H \neq \{\vec{0}\}$, some subset of $S$ is a basis for $H$.*

This is the Spanning Set theorem. We can use this to find bases for $Nul(A)$ and $Col(A)$, by looking at the column vectors in the matrix that are linearly independent (basically get the pivot columns). Note that row-equivalent matrices don't share basis vectors.

How do we find a basis for $Nul(A)$? It turns out that our process of explicitly describing a spanning set for $Nul(A)$ automatically is a basis.

**Theorem 3.4.** *Let $B = \{\vec{b}_1, \dots, \vec{b}_n\}$ be a basis for a vs $V$. Then for each $\vec{x} \in V$, there exists a unique set of scalars $c_1, \dots, c_n$ such that $\vec{x}$ can be represented by the basis vectors of $V$ as a superposition.*

The coordinates of $\vec{x}$ relative to the basis of $B$, also known as the $B$-coordinates are the weights $c_1, \dots, c_n$ such that $\vec{x} = c_1\vec{b}_1 + \cdots + c_n\vec{b}_n$. We write this as $[\vec{x}]_B = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$. This vector is called the coordinate vector of $\vec{x}$ relative to $B$. The map $\vec{x} \to [\vec{x}]_B$ is called the coordinate mapping determined by $B$.

**Definition 3.4.** $\mathcal{E} = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ is the standard basis of $\mathbb{R}^n$.

**Theorem 3.5.** *Let $\mathcal{B} = \{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_n\}$ be a basis for a vector space $V$. Then the coordinate mapping $\vec{x} \to [\vec{x}]_{\mathcal{B}}$ is a one to one and onto linear transformation from $V$ onto $\mathbb{R}^n$.*

**Definition 3.5.** A linear transformation $T : V \to W$ is said to be an isomorphism if $T$ is both one to one and onto, and $V$ and $W$ are said to be isomorphic, written $V \cong W$.

The most important part of this is that vector spaces that are isomorphic are "identical" as vector spaces.

Lets look at an example. Take the vector space of polynomials up to degree 3 ($\mathbb{P}_3$). We have the basis

$$\mathcal{B} = \{1, t, t^2, t^3\}$$

A generic polynomial of degree 3 is

$$p(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3$$

$$[\vec{p}]_{\mathcal{B}} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

Thus by the previous theorem, the coordinate mapping

$$\vec{p} \to [\vec{p}]_{\mathcal{B}}$$

is an isomorphism from $\mathbb{P}_3$ to $\mathbb{R}^4$, so $\mathbb{P}_3$ is isomorphic to $\mathbb{R}^4$.

Another example is that $\mathbb{R}^2$ is not a subspace of $\mathbb{R}^3$, but it is isomorphic to a subspace of $\mathbb{R}^3$ (the one where the third element is always 0).

**Theorem 3.6.** *If a vector space $V$ has a basis $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_n\}$ then any set in $V$ containing more than $n$ vectors must be linearly dependent.*

This is the analog to the similar statement we had back when we first started talking about linear independence. The we immediately get:

**Theorem 3.7.** *If a vector space $V$ has a basis of $n$ vectors, then every basis of $V$ must have exactly $n$ vectors. We call $n$ the dimension of $V$.*

**Definition 3.6.** If $V$ is spanned by a finite set, then $V$ is said to be finite-dimensional, and the dimension of $V$, written as $dim(V)$ is the number of vectors in a basis for $V$. If $V$ is not spanned by a finite set, then $V$ is said to be infinite-dimensional.

Let's look at a really easy one, $\mathbb{R}^n$. This is finite dimensional, and $dim(\mathbb{R}^n) = n$. An example of an infinite dimensional vector space would be $\mathbb{P}$, the space of all polynomials. Note that a subspace of $\mathbb{R}^4$ does not necessarily have to be of dimension 4.

**Theorem 3.8.** *Let $H$ be a subspace of a finite-dimensionial vector space $V$/ Any linearly independent set in $H$ can be expanded, if necessary, to a basis for $H$. Also, $H$ is finite-dimensional and $dim(H) \leq dim(V)$.*