MAT 307, Multivariable Calculus with Linear Algebra – Fall 2024

1 Matrices and Vector Spaces

1.1 Matrices

An $m \times n$ ***matrix*** A is a table of numbers

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

We will write $A = [a_{ij}]$, where a_{ij} is the entry in the i^{th} row and j^{th} column.

Matrices are essential to all engineering and applied sciences problems. They can be used to represent a system of linear equations; provide connectivity information in graphs/networks; represent operators on spaces, functions, etc. For example, the numbers a_{ij} could represent

- the color of the (ij)-th pixel on a screen
- the temperature in city i on day j
- the value of feature i (e.g. rating for a film on Netflix) for user j

We can also write A in terms of rows and columns as

$$A = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \end{bmatrix} = \begin{bmatrix} \vec{r}_1 \\ \vec{r}_2 \\ \vdots \\ \vec{r}_m \end{bmatrix}.$$

Here, $\vec{v}_j$, with $j = 1, \ldots, n$, are the ***column vectors*** of A and $\vec{r}_i$, with $i = 1, \ldots, m$, are the ***row vectors*** of A. Notice that each $\vec{v}_j$ is an $m \times 1$ matrix and each $\vec{r}_i$ is an $1 \times n$ matrix. For us, a ***vector*** is an $m \times 1$ matrix.

The ***transpose*** of A is an $n \times m$ matrix, denoted by A^t, obtained by flipping A along the diagonal. For example

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix}^t = \begin{bmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \\ 4 & 8 \end{bmatrix}.$$

The set of all $m \times n$ matrices is denoted by $M(m, n)$ and the set of all $m \times 1$ matrices (i.e. vectors) is denoted by $\mathbb{R}^m$. $M(m, n)$ and $\mathbb{R}^m$ are more than just sets, they have the structure of ***vector spaces***.

1.2 Vector spaces

Definition 1. A real ***vector space*** $\mathcal{V}$ is a set along with the operations of ***addition*** and ***scalar multiplication*** satisfying:

1. For all $X, Y \in \mathcal{V}$, there is an element $X + Y \in \mathcal{V}$,
2. For $X \in \mathcal{V}$ and $k \in \mathbb{R}$, there is an element $k \cdot X \in \mathcal{V}$.

Furthermore, the following conditions hold for all $k, l \in \mathbb{R}$ and $X, Y, Z \in \mathcal{V}$.

- Commutativity: $X + Y = Y + X$
- Associativity of Addition: $X + (Y + Z) = (X + Y) + Z$
- Additive Identity: there exists $\mathfrak{o} \in \mathcal{V}$ such that $X + \mathfrak{o} = X$
- Additive Inverse: there exists $-X \in \mathcal{V}$ such that $X + (-X) = \mathfrak{o}$
- Multiplicative Identity: $1 \cdot X = X$
- Associativity of Scalar Multiplications: $(kl)X = k(lX)$
- Distributativity I: $k(X + Y) = kX + kY$
- Distributativity II: $(k + l)X = kX + lY$

Remark. The number of properties listed is long but reasonable, basically they tell us that the laws of algebra work. Verifying all the vector space axioms would be rather tedious, later we will see that in most cases we don't need to verify all of them.

Example 1. Some examples of vector spaces:

1. $\mathbb{R}^n$ for $n \geq 0$.
2. A line in $\mathbb{R}^2$ through the origin.
3. A plane in $\mathbb{R}^3$ through the origin.
4. $M(m, n)$ — the set of $m \times n$ matrices.

Example 2. The following are not vector spaces. Why?

1. $\{\vec{v} = [v_1, \dots, v_n]^t \in \mathbb{R}^n \mid v_1 \geq 0\}$ — no additive inverse.
2. $\left\{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M(2,2) \mid ad - bc \neq 0\right\}$ — no additive identity.
3. $\left\{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M(2,2) \mid ad - bc = 0\right\}$ — $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ has $ad - bc = 1$, so the sum is not in the set — one says *not closed under addition.*

1.3 Linear Transformations

Given $A \in M(m,n)$, we can think of $A\vec{x}$ as A "applied" to $\vec{x}$. That is, for every $\vec{x} \in \mathbb{R}^n$, we obtain a vector $A\vec{x} \in \mathbb{R}^m$. Thus, A ***maps/transforms*** $\mathbb{R}^n$ to $\mathbb{R}^m$. We can then define a function $T_A : \mathbb{R}^n \to \mathbb{R}^m$ given by $T_A(\vec{x}) = A\vec{x}$.

Notice that

$$T_A(\vec{x}+\vec{y}) = A(\vec{x}+\vec{y}) = A\vec{x} + A\vec{y} = T_A(\vec{x}) + T_A(\vec{y}) \quad \text{and}$$

$$T_A(c\vec{x}) = A(c\vec{x}) = cA\vec{x} = cT_A(\vec{x}).$$

Functions between vectors spaces that satisfy these two properties will be the focus of this course going forward.

Definition 2. A map $T : \mathcal{V} \to \mathcal{W}$ where both $\mathcal{V}$ and $\mathcal{W}$ are vector spaces is called ***a linear transformation or linear map*** if $T(X+Y) = T(X) + T(Y)$ and $T(cX) = cT(X)$ for all $X, Y \in \mathcal{V}$ and for all scalars c.

Example 3. The map $f : \mathbb{R}^2 \to \mathbb{R}^2$ given by $f(x,y) = (2x, 5y)$ is a linear transformation that corresponds to stretching the x-direction by 2 and the y-direction by 5. Another example is $g : \mathbb{R}^2 \to \mathbb{R}^2$ given by $g(x,y) = \left(\frac{x-\sqrt{3}\,y}{2}, \frac{\sqrt{3}\,x+y}{2}\right)$, which corresponds to rotating the plane counter-clockwise around the origin by angle $\pi/3$.

Our first observation is that for any $A \in M(m,n)$, the map $T_A : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation.

Our second observation is that $T(\mathbf{o}) = \mathbf{o}$ because $T(X) = T(\mathbf{o} + X) = T(\mathbf{o}) + T(X)$

For linear transformations between Euclidean spaces, we have the following result.

Theorem 1. *Any linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ can be defined as multiplication by an $m \times n$ matrix A. That is, $T = T_A$ for some $A \in M(m,n)$.*

Proof. We will first build the matrix $A \in M(m,n)$ and then verify that $T = T_A$. Let A be the matrix whose i^{th} column is $\vec{v}_i = T(\hat{e}_i)$ for $1 \le i \le n$. Any $\vec{x} \in \mathbb{R}^n$ can be written uniquely (this is due to Proposition 3 below) as $\sum_{i=1}^n x_i \hat{e}_i$. Since T is a linear transformation,

$$T(\vec{x}) = T\left(\sum_{i=1}^n x_i \hat{e}_i\right) = \sum_{i=1}^n x_i T(\hat{e}_i) = \sum_{i=1}^n x_i \vec{v}_i = A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = A\vec{x} = T_A(\vec{x}).$$

□

1.4 Linear dependence

Definition 3. Let $S = \{A_1, A_2, \dots, A_k\}$ be a set of elements of vector space $\mathcal{V}$. We say $C \in V$ is ***linearly dependent on*** S if there exists $b_i \in \mathbb{R}$, such that

$$C = \sum_{i=1}^{k} b_i A_i = b_1 A_1 + b_2 A_2 + \cdots + b_k A_k.$$

In this case, we say C is a ***linear combination*** of the A_i's

Example 4. (Computing grades) Let $\vec{h}, \vec{q}, \vec{m}, \vec{f}$ be column vectors such that h_i is the homework average, q_i is the quiz average, m_i is the midterm grade, and m_i is the file exam grade of the i^{th} student in this class. The final grades $\vec{g}$ for all the students can be computed as

$$\vec{g} = 0.25 \cdot \vec{h} + 0.20 \cdot \vec{q} + 0.25 \cdot \vec{m} + 0.3 \cdot \vec{f}$$

That is, $\vec{g}$ is linearly dependent on $\{\vec{h}, \vec{q}, \vec{m}, \vec{g}\}$.

Definition 4. Let $S = \{A_1, A_2, \dots, A_k\} \subset \mathcal{V}$, then

- S is **linearly dependent** if $S = \{\mathfrak{o}\}$ or if there is at least one A_i that is a linear combination of the ***other*** elements of S. This is equivalent to saying that at least one A_i is linearly dependent on $S \setminus \{A_i\}$.
- S is **linearly independent** if it is not linearly dependent, i.e. no A_i is a linear combination of the other elements of S.

Remark. The zero element $\mathfrak{o} \in \mathcal{V}$ is linearly dependent on any non-empty subset of $\mathcal{V}$ because for any $A \in \mathcal{V}$ one has $\mathfrak{o} = 0 \cdot A$. Therefore, any set S containing $\mathfrak{o}$ is linearly dependent. The nullset $\{\}$ is linearly independent, since it is not linearly dependent as it contains no elements.

Exercise 1. Understand why the following are true.

- $\{\cos^2(x), \sin^2(x), 1\}$ is linearly dependent.
- $\{\cos(x), \sin(x), 1\}$ is linearly independent.
- $\{[1, 0, 0]^t, [1, 0, 1]^t, [0, 0, 1]^t\}$ is linearly dependent.
- $\{[1, 1, 0]^t, [2, 0, 1]^t, [1, 0, 1]^t\}$ is linearly independent.

Proposition 2. *A set* $S = \{X_1, \dots, X_n\} \subset \mathcal{V}$ *is linearly dependent if and only if there are* $c_1, \dots, c_n \in \mathbb{R}$ *such that*

$$\mathfrak{o} = c_1 X_1 + \cdots + c_n X_n \quad \textbf{and} \quad c_j \neq 0 \textit{ for some } j.$$

Proof. (⇒) Assume that $\{X_1, \ldots, X_n\}$ is linearly dependent. Then, either $S = \{\mathfrak{o}\}$ or for some j, $X_j = \sum_{i \neq j} c_i X_i$. Let $c_j = -1$, then $\mathfrak{o} = \sum_{i=1}^n c_i X_i$ and $c_j = -1 \neq 0$.
(⇐) Conversely, assume that there are $c_1, \ldots, c_n \in \mathbb{R}$ with $\mathfrak{o} = \sum_{i=1}^n c_i X_i$ and $c_j \neq 0$ for some j. Then, either $S = \{\mathfrak{o}\}$ or $X_j = \sum_{i \neq j} \left(-\frac{c_i}{c_j}\right) X_i$ is a (nonempty) linear combination, which means that S is linearly dependent. □

The ***span*** of S is the set of ***all*** linear combinations of elements of S. That is

$$\text{span}(S) = \left\{ C \in \mathcal{V} \mid C = \sum_{i=1}^{k} b_i A_i \text{ for } b_i \in \mathbb{R} \right\}.$$

Exercise 2. Draw pictures to explain why these are true.

- $\text{span}\{[1,0,0]^t, [1,0,1]^t, [0,0,1]^t\}$ is a plane in $\mathbb{R}^3$.
- $\text{span}\{[1,1,0]^t, [2,0,1]^t, [1,0,1]^t\}$ is all of $\mathbb{R}^3$.

Proposition 3. *The standard basis vectors $\hat{e}_i$ of $\mathbb{R}^n$ where the i-th coordinate is* 1 *while the rest are* 0 *for $i = 1, 2, \cdots, n$ are linearly independent and span $\mathbb{R}^n$.*

Can you see why? Note that $\mathbb{R}^n$ is also the span of the $\{\hat{e}_i\}_{i=1}^n$ since any $\vec{v} \in \mathbb{R}^n$ can be written as $\vec{v} = \sum_{i=1}^n v_i \hat{e}_i$.

Exercise 3. Can you show that $\{ \begin{bmatrix} 1 \\ a \\ b \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ c \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \}$ is linearly independent and spans $\mathbb{R}^3$? Here a, b, c are some fixed real numbers. You need to prove both parts of the statement under the assumption that a, b, c are ***any*** real numbers.

We will now focus our attention on linear dependence/independence for vectors and matrices. Our goal is to find an algorithm to determine whether a set of vectors is linearly dependent or not.

2 Linear Systems of Equations

2.1 Linear systems

A ***linear equation*** in a variable x has the form $ax = b$ where a, b are scalars. Equations such as $\sin(x) = e^x$, $x^2 = 1/x$ are ***not*** linear and are not considered. A general **linear equations** in n variables has the form

$$c_1 x_1 + c_2 x_2 + \cdots c_n x_n = c_0,$$

where x_i are the variables/unknowns and c_i are scalars.

For example, the set of points in the xy-plane satisfying $ax + by = c$ form a **line**; the set of points in $\mathbb{R}^3$ satisfying $ax + by + cz = d$ form a **plane**. The set of points in $\mathbb{R}^n$ satisfying the general linear equation above is a **hyperplane**.

A **system of linear equations** is a collection of linear equations.

$$\begin{aligned} a_{11}\,x_1 + a_{12}\,x_2 + \cdots + a_{1n}\,x_n &= b_1 \\ a_{21}\,x_1 + a_{22}\,x_2 + \cdots + a_{2n}\,x_n &= b_2 \\ &\vdots \\ a_{m1}\,x_1 + a_{m2}\,x_2 + \cdots + a_{mn}\,x_n &= b_m \end{aligned}$$

The $m \times n$ matrix $A = [a_{ij}]$ is called the **coefficient matrix** of the system. Let $A = \begin{bmatrix} \vec{v}_1 \, \vec{v}_2 \, \cdots \, \vec{v}_n \end{bmatrix}$ be the column vectors of A. Setting $\vec{b} = \begin{bmatrix} b_1 \, b_2 \, \cdots \, b_m \end{bmatrix}^t$, the system of linear equations is equivalent to saying that

$$x_1\,\vec{v}_1 + x_2\,\vec{v}_2 + \cdots + x_n\vec{v}_n = \vec{b}$$

That is, $\vec{b}$ is a linear combination of $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$. Here, $x_1, \ldots, x_n$ are the ***unknowns*** of the system. The expression on the left-hand size appears frequently enough to *define* the short-hand operation

$$A\,\vec{x} = A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \overset{def}{=\!=} x_1\vec{v}_1 + \cdots + x_n\vec{v}_n.$$

Thus, the system of linear equations is reduced to the expression $A\,\vec{x} = \vec{b}$.

We can also go backwards. For any $A \in M(m,n)$ and vector $\vec{b} \in \mathbb{R}^m$, the expression $A\,\vec{x} = \vec{b}$ defines a system of m linear equations with n unknowns given by $\vec{x} \in \mathbb{R}^n$.

The ***system has a solution*** if and only if $\vec{b}$ is a linear combination of the column vectors of A (i.e. $\vec{b}$ is in the span of $\{v_1, \ldots, v_n\}$). A system that does not have a solution is called ***inconsistent***.

The $m \times (n+1)$ matrix $[A \mid \vec{b}]$ formed by both the coefficients a_{ij} and the constants b_i is called the ***augmented matrix***.

Example 5.

$$\begin{aligned} 2x - y - z &= 0 \\ -2x + y - z &= -6 \\ x - 2y + 3z &= 3 \end{aligned}$$

This system consists of three equations for three unknowns x, y, and z. In matrix notation, the system can be written as

$$A = \begin{bmatrix} 2 & -1 & -1 \\ -2 & 1 & -1 \\ 1 & -2 & 3 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 0 \\ -6 \\ 3 \end{bmatrix}, \quad \text{and} \quad [A \mid \vec{b}\,] = \left[\begin{array}{ccc|c} 2 & -1 & -1 & 0 \\ -2 & 1 & -1 & -6 \\ 1 & -2 & 3 & 3 \end{array}\right].$$

To solve this system, we can use the method of ***elimination of variables***: the first equation gives $z = 2x + y$. Substituting this into the second and third equations,

$$-4x + 2y = -6, \qquad 7x - 5y = 3.$$

The first equation gives $y = 2x - 3$. Substituting into the second equation gives $x = 4$. Hence, $x = 4$, $y = 5$, and $z = 3$.

Proposition 4. *The column vectors of A are linearly independent if and only if the only solution to $A\vec{x} = \vec{0}$ is $\vec{x} = \vec{0}$.*

Proof. By definition $A\vec{x} = x_1\vec{v}_1 + \cdots + x_n\vec{v}_n$, where $\vec{v}_i$ are the column vectors of A. From this point of view, this proposition is the ***contrapositive*** of Proposition 2. □

Since the solution set of a system $A\vec{x} = \vec{0}$ determines the linear dependence/independence of the columns of A, we want to find a way to solve linear systems $A\vec{x} = \vec{b}$ efficiently. In the next lecture, we will develop the method of ***Gaussian elimination***, but for now, let us focus on the geometry of linear systems.

2.2 Geometry

For the system in Example 5 the solution occurs at the intersection of all three planes. The normal to each plane is understood as the row vectors $\vec{r}_i$ of A. Thus, the row vectors determine how corresponding planes intersect one another.

Consider two planes $a_{11}x + a_{12}y + a_{13}z = b_1$ and $a_{21}x + a_{22}y + a_{23}z = b_2$. Let A be the coefficient matrix and $[A \mid \vec{b}\,]$ the augmented matrix. The two planes in $\mathbb{R}^3$ will

- intersect in a line if the two planes are not parallel. Let $\vec{r}_1$ and $\vec{r}_2$ be the row vectors of A, then the two planes are not parallel if and only if $\{\vec{r}_1, \vec{r}_2\}$ is linearly independent. For example, consider the augmented matrix

$$[A \mid \vec{b}\,] = \left[\begin{array}{ccc|c} 4 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \end{array}\right]$$

- coincide in the same plane if the rows of the ***augmented*** matrix are linearly dependent. For example, consider the augmented matrix

$$[A \mid \vec{b}\,] = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 2 & 0 & 0 & 2 \end{array}\right]$$

- be parallel if $\vec{r}_1$ and $\vec{r}_2$ are linearly dependent but the two row vectors of $[A \mid \vec{b}]$ are linearly independent. For example, consider the augmented matrix

$$[A \mid \vec{b}] = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 2 \end{array}\right]$$

Two equations in three unknowns have either no common solutions (i.e. inconsistent) or infinitely many solutions.

For systems with three equations to have a solution, we have to intersect the line or plane solving two of the equations with the plane representing the third equation. Three planes can intersect in

- 0 points when the system is inconsistent because either all three planes are parallel; or the three planes intersect in two or three lines that are pairwise parallel. Here are some examples,

$$[A \mid \vec{b}] = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 2 \\ 1 & 0 & 0 & 3 \end{array}\right], \qquad \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{array}\right], \qquad \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 \end{array}\right]$$

- infinitely many points, when three planes intersect along the same line (eg: the system $x = 2$, $y = 2$, $x - y = 0$), or when they overlap in the same plane (eg: the system $x = 1$, $2x = 2$, $3x = 3$). The system is consistent and $\vec{b}$ lies in the same plane or line spanned by the three column vectors of A.

$$[A \mid \vec{b}] = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 2 \\ 1 & -1 & 0 & 0 \end{array}\right], \qquad \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 2 & 0 & 0 & 2 \\ 3 & 0 & 0 & 3 \end{array}\right]$$

- 1 point, when the three row vectors $\vec{r}_i$ of A are linearly independent. As we shall see later, $\vec{b}$ can then be expressed ***uniquely*** as a linear combination of the column vectors. An example augmented matrix is

$$[A \mid \vec{b}] = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

Definition 5. The ***rank*** r of a system of m linear equations in n unknowns is the ***maximal number of linearly independent <u>rows</u> in the augmented matrix*** $[A \mid \vec{b}]$. In general, rank(A) of a matrix A is the maximal number of linearly independent rows in A.[1]

[1] Why is this not a very good definition? We will find a better one in the next lecture.

For a system of m linear equations in n unknowns, where we have

$$A\vec{x} = \vec{b}, \qquad A \in M(m,n), \qquad \vec{x} \in \mathbb{R}^n, \qquad \vec{b} \in \mathbb{R}^m$$

Next time, we will show that if $m < n$, there are either no solution or infinitely many solutions. If $m \geq n$, then there maybe 0, 1, or infinitely many solutions. The number of solutions to a linear system depends on ***both*** $r' = \text{rank}(A)$ and $r = \text{rank}([A \mid \vec{b}])$. We will use Gauss-Jordan elimination to create an efficient algorithm to compute r and r'.

By computing r and r' for the examples above, we find ***empirically*** that

- If $r' < r$, the system is inconsistent and has no solution.
- If $r = r' < n$, then the system has infinitely many solution.
- If $r = r' = n$, the system has a unique solution.