Math 20C

February 26, 2021

1 The Geometry of Euclidean Space

1.1 Vectors in Two- and Three-Dimensional Space

Points are represented by ordered pairs or triplets, and they are called **Cartesian coordinates**. In cartesian coordinates, points are found by intersecting the coordinates at their respective axis, and are usually represented in the form (x,y) or (x,y,z). Each axis represents a number line called $\mathbb{R}$.

$$2D = (x, y) = \mathbb{R}^2$$

$$3D = (x, y, z) = \mathbb{R}^2$$

1.1.1 Vector Addition and Scalar Multiplication

Adding two vectors together adds up their respective components together. In other words, the x components, the y components, and the z components are added together.

$$(a, b, c) + (d, e, f) = (a + d, b + e, c + f)$$

You can multiply a vector by a scalar multiple, which is a number with just magnitude (one number).

$$\alpha(a, b, c) = (\alpha a, \alpha b, \alpha c)$$

Other definitions:

1. Zero element - $(0, 0, 0)$

2. Additive inverse - negative vector of the specified vector, travels in opposite direction but with same magnitude.

1.1.2 Geometry of Vector Operations

Components of a vector can simply also be called the coordinates of the vector, where the linear line connects the origin to the coordinate.

Geometrically, vector addition involves creating a parallelogram by drawing the other vector (we'll call it v_2), but starting at the endpoint of the first vector (v_1). This action is also applied to the other vector (v_2), where v_1 would be drawn at the endpoint of p¿ v_2.

A **bound vector** is a vector that starts at an origin, while **free vectors or vectors** do not necessarily have to start at the origin. Geometrically, multiplying a scalar to a vector will create a vector that is a certain length of the original vector.

1.1.3 The Standard Basis Vectors

There are three standard vectors that help represent a vector with multiple components: $\vec{i}, \vec{j}, \vec{k}$

$$\vec{i} = (1, 0, 0),$$
$$\vec{j} = (0, 1, 0),$$
$$\vec{k} = (0, 0, 1),$$

Vectors can then be represented in standard basis vectors:

$$\vec{a} = (a_1, a_2, a_3) = a_1 i + a_2 j + a_3 k$$

1.1.4 The Vector Joining Two Points

Vector subtraction can be explained through vector addition with a additive inverse of one of the vectors, or through geometrical method of drawing a line from the vector you wish to subtract from to the vector that is being subtracted from.

To subtract geometrically, (example: $\vec{a} - \vec{b}$) draw a vector with origin at $\vec{b}$ to $\vec{a}$, which will be the same length and direction as the vector if you were to subtract it algebraically.

1.1.5 Equations of Lines

Lines can be created using a point (we'll call it point P) and a vector ($\vec{v}$). Through parallel lines, addition at point P parallel to vector $\vec{v}$ will produce a line that is exactly parallel to $\vec{v}$. Any point on the line can be represented as $P + t\vec{v}$, where t is any real number. To quote the book: The equation of lien l through the tip of $P = (a, b, c)$ and pointing in the direction of the vector $\vec{v} = (d, e, f)$ is $l(t) = P + t\vec{v}$, where the parameter t is any real number.

Coordinate form equation:

$$x = a + td$$
$$y = b + te$$

$$z = c + tf$$

You can represent this line as an equation:

$$l(t) = (a + td, b + te, c + tf)$$

Note that you can manipulate the vector section of the equation to modify the equation's intercept form, similar to how in 2d functions, you usually have the equation $y = m(x - h) + k$), with (x,y) = (h,k). As long as (d,e,f) are the same and the point P is on the line, P can be changed.

1.1.6   Parametric Equation of a Line: Point Point Form

The parametric equations of the line l through points (x_1, y_1, z_1) and (x_2, y_2, z_2) are

$$x = x_1 + (x_2 - x_1)t,$$
$$y = y_1 + (y_2 - y_1)t,$$
$$z = z_1 + (z_2 - z_1)t,$$

Note that the places of the point P can be replaced with point x_2 as well.

1.2   Inner Product, Length, and Distance

1.2.1   Inner Product

The Inner product of two vectors is equal to the product of their individual components.
Example:

$$\vec{a} = (a_1, a_2, a_3), \vec{b} = (b_1, b_2, b_3)$$
$$\vec{a} \cdot \vec{b} = a_1 b_1 + a_2 b_2 + a_3 b_3$$

The dot product between the same vector $(\vec{a} * \vec{a})$ will always be ≥ 0 because we square the components.

Inner dot product also adheres to basic algebraic equations, such as distribution of a scalar after inner dot product.
The length of a vector is calculated through the pythagorean theorem, where each component is squared, added up, and rooted.

$$||\vec{a}|| = \sqrt{a_1^2 + a_2^2 + a_3^2} == \sqrt{\vec{a} \cdot \vec{a}} = \textbf{Length of } \vec{a}$$

1.2.2   Unit vectors

Unit vectors are vectors that have a length of one. In other words, $||\vec{v}|| = 1$. To convert any vector to a unit vector, divide the vector by its length. The process of turning a vector into a unit vector is called **normalization**.

$$\textbf{Normalized a} = \frac{\vec{a}}{||a||}$$

3

Proof has been left as an exercise to the reader (just kidding):

$$\left|\left|\frac{\vec{a}}{||\vec{a}||}\right|\right| = \left|\left|\frac{(a_1, a_2, a_3)}{||\vec{a}||}\right|\right| = \sqrt{\frac{(a_1)^2 + (a_2)^2 + (a_3)^2}{||\vec{a}||^2}} = \sqrt{\frac{||\vec{a}||^2}{||\vec{a}||^2}} = 1$$

Example) $\vec{u} = (1, 3, 4)$

1. $Length = \sqrt{1^2 + 3^2 + 4^2} = \sqrt{26}$

2. Normalize $\vec{u}$: $\frac{\vec{u}}{||\vec{u}||} = \frac{1}{\sqrt{26}} + \frac{3}{\sqrt{26}} + \frac{4}{\sqrt{26}}$

1.2.3   Distance

Distance $= \left|\left|(\vec{b} - \vec{a})\right|\right|$. Pretty self explanatory.

1.2.4   Angle Between Two Vectors

Angles are in radians, and range from $0 \leq \theta \leq \pi$. Using the law of cosines and substituting the sides for the distance of each vector (including the subtraction of the two vectors), we get the equation

$$\vec{ab} = ||\vec{a}|| * ||\vec{b}|| \cos\theta$$

Proof: $a^2 = b^2 + c^2 - 2bc\cos\theta$, where $a = ||\vec{a} - \vec{b}||$, $b = ||\vec{b}||$, $c = \vec{a}$.

$$||\vec{a} - \vec{b}||^2 = ||\vec{a}||^2 + ||\vec{b}||^2 - 2||\vec{a}||||\vec{b}|| \cos\theta$$

$$(\vec{a} - \vec{b}) * (\vec{a} - \vec{b}) = ||\vec{a}||^2 + ||\vec{b}||^2 - 2||\vec{a}||||\vec{b}|| \cos\theta$$

$$(\vec{a})^2 + (\vec{b})^2 - 2\vec{ab} = ||\vec{a}||^2 + ||\vec{b}||^2 - 2||\vec{a}||||\vec{b}|| \cos\theta$$

$$||\vec{a}||^2 + ||\vec{b}||^2 - 2\vec{ab} = ||\vec{a}||^2 + ||\vec{b}||^2 - 2||\vec{a}||||\vec{b}|| \cos\theta$$

$$-2\vec{ab} = -2||\vec{a}|| * ||\vec{b}|| \cos\theta$$

$$\vec{ab} = ||\vec{a}|| * ||\vec{b}|| \cos\theta$$

isolating θ gives us the equation for finding the angle between two vectors:

$$\theta = \arccos\left(\frac{\vec{a} \cdot \vec{b}}{||\vec{a}|| \; ||\vec{b}||}\right)$$

1.2.5 Cauchy-Schwarz Inequality

$$|\vec{a} \cdot \vec{b}| \leq ||a|| \; ||b||$$

if and only if $\vec{a}$ is a scalar multiple of $\vec{b}$, or one of them is 0. Both of the proofs that were given in this class were kinda lacking, so I formed this proof from a Khan Academy Video.

We know that for any vector,

$$||\vec{v}|| \geq 0$$

We'll represent this as a function (equation of distance from point on a line to origin):

$$||t\vec{y} - \vec{x}|| \geq 0 \text{ squaring both sides} \rightarrow p(t) = ||t\vec{y} - \vec{x}||^2 \geq 0$$

t is a scalar, while y and x are vectors.

$$||\vec{a}||^2 = \sqrt{\vec{a} \cdot \vec{a}}^2 = \vec{a} \cdot \vec{a}$$

$$p(t) = (t\vec{y} - \vec{x}) \cdot (t\vec{y} - \vec{x}) = (\vec{y} \cdot \vec{y})t^2 - (2\vec{x} \cdot \vec{y})t + (\vec{x} \cdot \vec{x})1 \geq 0$$

We can consider the vectors to be constants (because they are and won't change), substituting the following:

$$A = \vec{y} \cdot \vec{y}$$

$$B = 2\vec{x} \cdot \vec{y}$$

$$C = \vec{x} \cdot \vec{x}$$

We make the equation

$$p(t) = At^2 - Bt + C \geq 0$$

Plugging in $t = \frac{b}{2a}$:

$$p(\frac{b}{2a}) = A\frac{B^2}{4A^2} - B\frac{B}{2A} + C \rightarrow \frac{B^2}{4A} - \frac{2B^2}{4A} + C \rightarrow C - \frac{B^2}{4A} \geq 0 \rightarrow$$

$$4AC \geq B^2$$

Plugging back in the vectors:

$$4 * \vec{x} \cdot \vec{x} * \vec{y} \cdot \vec{y} = 4(\vec{x} \cdot \vec{y})^2 \rightarrow \sqrt{||\vec{x}||^2||\vec{y}||^2} = (\vec{x} \cdot \vec{y})^2 \rightarrow ||\vec{x}|| \; ||\vec{y}|| \geq \vec{x} \cdot \vec{y}$$

We can absolute value both sides to finally get the formula:

$$||\vec{x}|| \; ||\vec{y}|| \geq |\vec{x} \cdot \vec{y}|$$

This was the simplest proof I could find.

1.2.6 Perpendicular(orthogonal) vectors

If two vectors are perpendicular ($\theta = 90$ and $\cos(\theta) = 0$), the vectors are considered **orthogonal**. If there is a system of mutually orthogonal unit vectors (all vectors are orthogonal to each other and are unit vectors), the whole system is considered **orthonormal**.
You can figure out if two vectors are orthogonal through the angle function, in which the dot products of the two vectors are 0.

$$\vec{a} \cdot \vec{b} = 0$$

1.2.7 Ortogonal Projection

Orthogonal Projection is when two vectors form a right angle by using one of the vectors as a hypotenuse, and the other vector as a leg of the triangle.
The **orthogonal projection** of vector $\vec{v}$ on $\vec{a}$ is the vector $\alpha\vec{a}$ that creates a right triangle with $\vec{v}$. A projection along a vector is the projected vector required to form a right triangle with another vector (which is said to be "along" the vector). The vector (we'll call it $\vec{v}$) that projects on another vector ($\vec{a}$) forms a vector by the equation

$$\vec{v} = \alpha\vec{a} + \vec{q}$$

Where α is a constant, and $\vec{q}$ is always orthogonal to a.
We solve for α by multiplying both sides by $\vec{a}$, where we get $\alpha = \frac{\vec{a}\cdot\vec{v}}{||a||^2}$.

$$\vec{p} = \alpha\vec{a} = \frac{|\vec{a} \cdot \vec{v}|}{||a||^2}\vec{a}$$

Examples:

1. $\vec{i} + \vec{j}$ on $\vec{i} - 2\vec{j}$:
$$\vec{p} = \frac{-1}{5}(1, -2, 0) = (\frac{-1}{5}, \frac{2}{5}, 0)$$

2. $(1, 0)$ on $(0, 1)$
$$\vec{p} = 0$$

Dot product of two vectors are zero, so the projecting vector is also zero. Note that if $\vec{a}$ and $\vec{v}$ are orthagonal, the projection will be zero.

3. $(-1, 1)$ on $(1, 0)$
$$\vec{p} = \frac{-1}{1}(1, 0) = (-1, 0)$$

1.2.8 Triangle Inequality

The distance of two vectors added together is always equal to or less than the two vectors' distance added together.

$$||a + b|| \leq ||a|| + ||b||$$

This can be visualized through drawing any triangle at all, in which the addition of the two legs are always greater than the actual hypotenuse.

1.3 Matrices, Determinants, and the Cross Product

$\vec{a} \times \vec{b}$ represents the dot product of vector a and vector b. In the dot product, the ordering matters.

1.3.1 2 x 2 Matrices

In a 2 x 2 matrix, there are 4 scalars that are arranged in a 2 x 2 matrix, with a matrix having the form $m \times n$, where m is the number of columns, and n is the number of rows. The vectors in the example takes the form row then column.
Example:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

The **determinant** of a 2x2 matrix is the product of diagonal from top left to bottom right, which is subtracted by the diagonal starting top right to bottom left. Determinants are marked through the use of vertical lines in contrast to square brackets which denotes a matrix.
Example:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

1.3.2 3 x 3 Matrices

A 3 x 3 matrix is a matrix with 3 rows and 3 columns.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

The **determinant** of a 3x3 matrix is calculated through the following steps:

1. Cross out the first row. It will be used in the final calculation after the 2x2 determinant of each column

2. For each column, cross it out, and calculate the determinant with the 2x2 matrix that is left over after crossing the row and column out. This will be multiplied by the item in the first row with the same column.

3. Add the first column determinant, subtract the second column determinant, and add the third column determinant.

Example:

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = 1 \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} - 2 \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} + 3 \begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} = 1(5*9-6*8) - 2(4*9-6*7) + 3(4*8-5*7)$$

$$= 1(-3) - 2(-6) + 3(-3) = -3 + 12 - 9 = 0$$

1.3.3   Properties of Determinants

1. If you interchange 2 columns or rows, the determinant changes sign (+ to -). This applies to both 2×2 and 3×3 matrices. This allows for any row to be the first row in 3×3 matrices.

2. If two rows or columns are identical, the determinant will be 0.

3. Scalars can be factored out of any row or column, and redistrubted to other rows/columns. They can also be taken out of the determinant, laying outside as a coefficient.

4. Adding to a row or column to another row or column will not change the result of the determinant.

1.3.4   The Cross Product

The cross product is an application of the determinant between 2 vectors. It usually assumes the 3 column x 2 row form, in which we would only cross out columns to create 2x2 matrices which we find determinants of.
Example:
$$\vec{a} = (a_1, a_2, a_3), \ \vec{b} = (b_1, b_2, b_3)$$

$$\vec{a} \times \vec{b} = \begin{vmatrix} i & j & k \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \begin{vmatrix} a_2 & b_3 \\ a_2 & b_3 \end{vmatrix} \vec{i} - \begin{vmatrix} a_1 & b_3 \\ a_1 & b_3 \end{vmatrix} \vec{j} + \begin{vmatrix} a_1 & b_2 \\ a_1 & b_2 \end{vmatrix} \vec{k}$$

Remember that the vector columns and rows aren't interchanged, or else it will change the sign of the determinant. In addition, $\vec{a} \times \vec{b} = 0$ when a and b are parallel, or one of the vectors are 0. The cross product of two vectors will always create a vector orthogonal to any vector that is spanned (that can be created) by the two original vectors.
The triple product is when we calculate the cross product of two vectors, and then use that cross product to calculate an inner dot product.

$$\textbf{Triple Product} = (\vec{a} \times \vec{b}) \cdot \vec{c}$$

The c vector simply becomes the coefficient of the cross product between the a and b vector.

$$(\vec{a} \times \vec{b}) \cdot \vec{c} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = - \begin{vmatrix} c_1 & c_2 & c_3 \\ b_1 & b_2 & b_3 \\ a_1 & a_2 & a_3 \end{vmatrix} = \begin{vmatrix} c_1 & c_2 & c_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

Using the cross product between two vectors and a trig identity, we get the equation:

$$||a \times b||^2 = ||a||^2 ||b||^2 - (a \cdot b)^2 = ||a||^2 ||b||^2 - ||a||^2 ||b||^2 \cos^2 \theta = ||a||^2 ||b||^2 \sin^2 \theta$$

Which gives
$$||a \times b|| = ||a|| \ ||b|| \sin \theta$$

Geometrically, determinants represent the area of a parallelogram formed with the adjacent sides of two angles. In 3d space, its the volume of a shape called a parallelepiped, with the vectors being the adjacent sides of the shape. Remember that the triple product is the same thing as a 3×3 determinant.

1.3.5 Equations of Planes

We gotta cover what **normal** means to vectors. A vector is **normal** to a plane if it is completely perpendicular to the plane. With this, we know that a point on the plane should create a dot product of 0 with a vector that is normal to the plane.

We shall define two vectors (P_0 is a point on vector P):

$$\vec{P} = (x, y, z) \tag{1}$$
$$\vec{P_0} = (x_0, y_0, z_0) \tag{2}$$
$$\overrightarrow{P_0P} = (x - x_0, y - y_0, z - z_0) \tag{3}$$
$$\vec{n} = Ai + Bj + Ck \tag{4}$$

To know if $\vec{P}$ is on the the same plane as $\vec{P_0}$, we calculate the dot product $\overrightarrow{P_0P} \cdot \vec{n} = 0$, and ensure that the dot product does indeed compute to 0.

$$\overrightarrow{P_0P} \cdot \vec{n} = A(x - x_0) + B(y - y_0) + C(z - z_0) = 0$$

which can be simplified to

$$\overrightarrow{P_0P} \cdot \vec{n} = Ax + By + Cz + D = 0 \tag{5}$$
$$D = -Ax_0 - By_0 - Cz_0 \tag{6}$$

The equation for a plane that is normal to vector $\vec{v} = (A, B, C)$ and intersects the plane at a point (x_0, y_0, z_0) satisfies the following equation:

$$P = A(x - x_0) + B(y - y_0) + C(z - z_0) = 0$$

We can usually find vectors that are orthogonal to the plane through cross product of two vectors.

We don't actually know what A, B, C, D are, and they are determined by the plane and are a scalar multiple. If you know linear algebra, a reduced echelon matrix would probably help you solve for the scalar. Else, the cross product of vectors on the plane can create a cross product which will determine the scalars of the vectors.

1.3.6 Distance: Point to Plane

The distance from (x_1, y_1, z_1) to the plane $Ax + By + Cz + D = 0$ is

$$Distance = \frac{|Ax_1 + By_1 + Cz_1 + D|}{\sqrt{A^2 + B^2 + C^2}}$$

2 Differentiation

2.1 The Geometry of Real-Valued Functions

This section discusses actual functions that would be applied to vectors, and seems to have some similarity to linear algebra again.

2.1.1 Functions and Mappings

Functions in Vector Calculus have a similarity to functions on 2d space in that a function applies certain operations on the provided input. We assume that the function has a domain of A (in which $\vec{x}$ is in A, I think it means that x must have the same amount of components as A has of columns), and applies a set of operations on the given vector. These functions are called **vector-valued functions** when the output of the function is over 1 (meaning that it creates a vector and not a scalar, and is in $\mathbb{R}^m$ with $m > 1$). When $m = 1$, it's considered a **scalar-valued function** because it produces a scalar.

Let's look at some functions that are in the book:

$$f : (x, y, z) \mapsto (x^2 + y^2 + z^2)^{-3/2}$$

In this function (which is a scalar-valued function because it produces one number, making it a scalar), given a vector (x, y, z) in $\mathbb{R}^3$. The arrow with a vertical line starting at the end means "maps to", in which the function of (x, y, z) maps to the specified scalar.

This is an example of a **function of several variables** when the function takes in more than one value in order to produce an output.

$$g(x) = g(x_1, x_2, x_3, x_4, x_5, x_6) = (x_1 x_2 x_3 x_4 x_5 x_6, \sqrt{x_1^2 + x_6^2})$$

This is a function with domain $\mathbb{R}^6$ with a range of $\mathbb{R}^2$, making it a vector-valued function, in which $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$, which means that A is a subset in (domain of) $\mathbb{R}^n$, which is then has a range of $\mathbb{R}^m$.

In real life problems, functions of several variables are very common, as there usually isn't an exact correlation between just two variables; its a combination of many variables, which would produce such an output, usually a scalar if quantifying something, or a vector if necessary (like in physics, with velocity being a vector).

Alright, let's see what functions can actually do in this class.

2.1.2 Graphs of Functions

Let's define the symbols to make you look smarter:

1. $x \in A$ means that x is in A, meaning that x has variables that are held in A.

2. $A \subset \mathbb{R}^m$ means that A is a subset of $\mathbb{R}^m$, meaning that A holds numbers that would be found in $\mathbb{R}^m$, and that A has the same dimension as $\mathbb{R}^m$.

3. $\rightarrow$ simply notes conversion usually from one dimension to another, which could be the same dimension.

Note that 3d graphs are already hard enough to see, and any dimension above that (with perhaps the exclusion of time in a 3d graph, making it 4d) would be hard to visualize graphs above 3D.

Graph of a Function: Let f: $U \subset \mathbb{R}^n \rightarrow \mathbb{R}$, meaning that given variables in $\mathbb{R}^m$, produce a scalar. The graph of f would be $\mathbb{R}^n + 1$, in which upon graphing all the variables, graph an additional point denoting the output of f.

$$\text{graph } f = (x_1, \ldots, x_n, f(x_1, \ldots, x_n)) \in \mathbb{R}^n + 1 | (x_1, \ldots, x_n) \in U$$

Where U is the domain over $(x_1, \ldots, x_n)$.

2.1.3 Level Sets, Curves, and Surfaces

A **level set** is a set where the function is set to a constant, such as 1. If you were to apply this to a full function, you would get subsections of the function at points where the certain constant is equal to the original function (in a 2d graph, this would be a flat line at a specified z or constant). This graph of a level set is then called a level curve or level contour, because it does indeed outline a certain z level for inputs of $\mathbb{R}^2$.
Here's a mathematical interpretation of Level Curves:

$$\{x \in U | f(x) = c\} \subset \mathbb{R}^n$$

The graph of the function only spans $\mathbb{R}^n$ because the function does not actually appear to be graphed, and it is only a constant. One example of level curves are circles, graphed by $x^2 + y^2 = c^2$, in which the variables must equal a certain constant.

2.1.4   The Method of Sections

In a section of graph f, a part of graph f intersects with a given plane. I don't think this section is that important.

2.2   Limits and Continuity

Like 2d graphs ($\mathbb{R}^2$, graphs in $\mathbb{R}^n$ also have limits and definitions for continuity, which are similar to its 2d graph definitions

2.2.1   Open Sets

An Open disk or ball of radius r is a disk/sphere based around a point (called $\vec{x_0}$) where the equation ($\|\vec{x} - \vec{x_0}\| < r$) holds true. We know that this makes a circle because it's max distance is from the center to the edge of the disk or ball, which is equal to the radius. This open set is represented usually as $D_r(\vec{x_0})$, where r is the radius of the open set and the whole function is in a set in $\mathbb{R}^n$.

$$D_r(\vec{x_0}) \subset U \subset \mathbb{R}^n \tag{7}$$
$$D_r(\vec{x_0}) = \|\vec{x} - \vec{x_0}\| < r \tag{8}$$

This also has a theorem, where for each $\vec{x_0} \in \mathbb{R}^n$ and $r > 0$, $D_r(\vec{x_0})$ is an open set. U is used to define values which make sense for the function and make the function true.

2.2.2   Boundary

Boundary points parallel boundary points in real life. In math, a **boundary point** is where for a point x where $x \in \mathbb{R}^n$, and every 'neighborhood' of x contains at least one point in A and at least one point not in A, where $A \subset \mathbb{R}^n$. A neighborhood is pretty much defined as a subset of a function's output, or possible solutions to a function. Note that the point x can be in either A or not in A. As long the neighborhood it has a point that is/isn't in A, then it is considered to be a boundary point.

2.2.3   Limits

If you're in math 20c, you probably will know what limits are already. It's probably better to use a previous explanation rather than using the mathematical explanation that they give here. To redefine an informal definition of limits (given by Prof. Hammock), a limit is when a function approaches a certain value, (such as $\lim_{(x,y)\to(a,b)} f(x,y) = L$), which is sufficiently close enough to (a,b), but not equal to (a,b).
Limits attempt to generalize where a function goes as it closes in on a certain point.

This is the formal definition:

Defining variables:

1. f: $A \subset \mathbb{R}^n \to \mathbb{R}^m$, where A is an open set.

2. x is assumed to be within open set A, or a boundary point of A.

3. $x \neq x_0$

Limit is defined as the following: a function that approaches a value called b as x approaches x_0. This means that as x gets closer to x_0, function f should reach closer to the value b.

These limits have the same types of properties as limits you've learned of before (or i hope you learned them). The only real new property is that for limits with a vector (b is not $\mathbb{R}$), the b will have multiple components and will be the same dimension as the function's output.

For multiple variable limits, we must keep some things in mind:

1. Any path that approaches (a,b) must equal the same thing. Similar to how the left and right side must equal the same value/point in order for the limit to potentially exist, any side must equal the same value/point in order to exist.

2. Because of this, if the limit as f (x,y) approaches (a,b) yields different points/numbers, then the limit does not exist.

2.2.4   Continuity

Continuity is a feature of a graph where at a certain point, the point is continuous if the limit approaching that point is equal to the point actually at that point.
$$\lim_{x \to x_0} f(x) = f(x_0)$$

2.2.5   Substitution

Given that a point is continuous, you can substitute the limit for the actual value of x, and this property applies to all functions, even multi-variable.

You can also substitute multivariables for one variable if applicable, say in the example
$$\lim_{(x,y) \to (0,0)} \frac{\sin(xy)}{xy}$$
You can substitute u for (x,y), which would result in an easier limit to solve. Remember to evaluate the expression from one side, especially if the limit approaches an endpoint of a graph when not substituting.

Polar coordinates can also be used to help simplify limits:

$$x = r\cos\theta \tag{9}$$
$$y = r\sin\theta \tag{10}$$
$$r = \sqrt{x^2 + y^2} \tag{11}$$

2.2.6 Methods for Limits

These are some methods to try out and combine when solving limits:

1. Convert the function to one-variable, solve using past math class lessons

2. Convert to polar coordinates, and use trig identities to simplify the limit

3. Use squeeze theorem to squeeze a limit between two other limits that equal the same thing

4. Use continuity when possible

5. Show that limit doesn't exist when using different paths

2.3 Differentiation

Graphs are hard in 3D space. That's why we use derivatives to help find rates of change, minima, maxima, and other features of a graph.

2.3.1 Partial Derivatives

Similar to derivatives in one-variable calculus, the definition of a partial derivative is similar to a derivative of one function. It is most likely called a partial derivative in calculus because the definition of a partial derivative comes from the change of only one variable. Partial derivatives become easy when we assume that all other variables are said to be constant.

Assume that $U \subset \mathbb{R}^n$ (Domain spans $\mathbb{R}^n$), and f : $U \subset \mathbb{R}^n \to \mathbb{R}$. A **partial derivative with respect to the n-th variable** is defined by

$$\frac{\partial f}{\partial x_j}(x_1, \ldots, x_n) = \lim_{h \to 0} \frac{f(x_1, x_2, \ldots, x_j + h, \ldots, x_n)}{h} \tag{12}$$
$$= \lim_{h \to 0} \frac{f(\vec{x} + he_j - f(x)}{h} \tag{13}$$

In the simplified version, the variables are all within vector x, while the h is multiplied with the variable the partial derivative is with respect to. Domain of the derivative should be all real numbers, or $x \in \mathbb{R}^n$.
You should know by now the different forms that the derivative may be written, along with the main rules for differentiation.

2.3.2   The Linear or Affine Approximation

A plane is tangent to a graph if the the x and y axes at a certain point are equal to both the partial derivative with respect to x and the partial derivative with respect to y. This is also similar to slope intercept form of 2d functions. This tangent plane represents a tangent line because each partial derivative describes the instantaneous rate at which a variable (either x or y) is changed at a certain point.

Note that the graph can be represented either as equal to z (or f(x,y) which is also z) or can equal zero, with the z on the other side. Onto definition:

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be differentiable at $\vec{x_0} = (x_0, y_0)$. The tangent plane in $\mathbb{R}^3$ at (x_0, y_0) is defined by the equation

$$z \; (or f(x_0, y_0)) = \left[\frac{\partial f}{\partial x}(x_0, y_0) \right](x - x_0) + \left[\frac{\partial f}{\partial y}(x_0, y_0) \right](y - y_0) + f(x_0, y_0)$$

This is also equal to the equation

$$0 = \left[\frac{\partial f}{\partial x}(x_0, y_0) \right](x - x_0) + \left[\frac{\partial f}{\partial y}(x_0, y_0) \right](y - y_0)$$

If you subtract x from both sides. I don't know if this is actually true, but i think it is.

2.3.3   Gradients

This was on the homework, so I think it's good to put it here.

The **gradient** of a function $f : U \subset \mathbb{R}^m \to \mathbb{R}$ is equal to a 1 x n matrix, with each column in the matrix being the partial derivative of the n-th variable. A gradient is denoted as **D**f (x).

$$\mathbf{D}f(x) = \begin{bmatrix} \dfrac{\partial f}{\partial x_1} & \cdots & \dfrac{\partial f}{\partial x_n} \end{bmatrix}$$

Evaluating a gradient at a point requires that you plug in the variables into each of the derivatives to get a scalar.

2.4   Introduction to Paths and Curves

2.4.1   Paths and Curves

Graphed functions on 2d were usually called functions, but they could also be curves, even if they don't actually have a curve. In mathematical terms, a **curve** is a set of values that maps an interval of real numbers onto a plane or space (which is just graph in whatever dimension), which is also called a **path**. In

these functions however, we use only one independent parameter that changes all dimensions: t, which usually represents time. In such curves and paths, we consider the function to parameterize the path of the function, as we have all dependent variables (x, y, z) to all be dependent to the independent variable t. This is also similar to one of our first lessons, where we created an Equation of lines (1.1.5) to graph such a curve/path. Note that t doesn't exactly mean time, which allows us to go into negative t's.

The book contains a big distinction between all of these curves / paths: Assume that a path in $\mathbb{R}^n$ is a function c: [a, b] $\rightarrow \mathbb{R}^n$, where a and b is the domain of the parameter t (i think?).

1. Just like the assumption says, any graphing of the function is a path.

2. If n = 2, it is a path in **plane**.

3. If n = 3, it is a path in **space**.

4. c(a) and c(b) are the endpoints of the path.

5. c(t) traces the curve C, meaning as t varies, the curve C becomes more complete for the domain.

6. For each componenent of c, we can create **component functions**, which are the functions for each component that takes in the value of t.

Note from the video lecture: c(t) usually represents the position of a particle at time t.

2.4.2   Velocity and Tangents to Paths

Component functions are useful in differentiating paths. Similar to the differentiation of 2d and 3d graphs, a differentiable path has a velocity as a time t, equal to

$$c'(t) = \lim_{h \to 0} \frac{c(t+h) - c(t)}{h} = (x'(t), y'(t), z'(t)) = x'(t)i + y'(t)j + z'(t)k$$

Which is just the definition of derivative with respect to t. Because we have component functions, we can easily determine the velocity of a path at a certain time. Note that this is the velocity vector, and does not tell where the position of the path is. Like other derivatives, the vector is parallel to the tangent line at (x(t), y(t), z(t)).

Related to slopes are **tangent vectors**, which determine a line tangent to path c (any function that creates a path).

The speed of the particle can be determined from the norm of velocity, as the distance of the velocity should give the speed of the particle (makes sense, alright?).

2.4.3 Tangent Line

Tangent line is similar to the graphs of linear approximation the the previous section (2.3). Assuming that c(t_0) is not a zero vector ($c(t_0) \neq 0$), the equation of the tangent line at point c(t_0) is

$$l(t) = c(t_0) + c'(t_0)(t - t_0)$$

Which copies the form of every other tangent line in any other space, including 2d.

2.5 Properties of the Derivative

Functions can be derived using the past derivative techniques that we learned in other classes, like chain rule, product rule (i don't remember the name), and quotient rule.

Here are derivative properties, assuming that each function is differential at x:

1. Constant Mulitiple Rule: taking a differential function at x, The coefficient of the function after taking the derivative will remain.

2. Sum rule: If a function is two functions with the same domains and ranges (domains have to be in same dimension, range has to also be in same dimension but doesn't have to be same as domain), the derivative function that contains the other two functions will just be the derivative of the other two functions added together.

3. Product rule: Assuming the domain is any dimension and the range is a scalar, The chain rule applies to functions like this, meaning

$$\frac{d}{dx}(f(x)g(x)) = f(x)\frac{d}{dx}g(x) + \frac{d}{dx}f(x)g(x)$$

4. Quotient rule: Assuming the domain is any dimension and the range is a scalar, the quotient rule is the same thing as the quotient rule in single-variable calculus:

$$\frac{d}{dx}\frac{f(x)}{g(x)} = \frac{g(x)\frac{d}{dx}f(x) - f(x)\frac{d}{dx}g(x)}{g(x)^2}$$

Note that these derivative rules apply for when f and g take in the same input dimension and give the same output dimension. There are different rules for when the functions don't have th same domain and range.

2.5.1 Chain Rule

This one is pretty intuitive. Let $g : U \subset \mathbb{R}^n \to \mathbb{R}^m$ and $f : V \subset \mathbb{R}^m \to \mathbb{R}^p$. Now suppose that we want to find the derivative of $f(g(\vec{x}))$, where g is differentiable at x. Because g(x) creates an output that's in $\mathbb{R}^m$, we can use this output to serve as the input to f, which makes this expression make sense. Taking the derivative of this yields

$$\frac{d}{dx}(f \cdot g) = \frac{d}{dx}g(x)f(\text{evaluated result of } (g(x)))$$

The evaluated part in g(x) means that we consider the g(x) inside f to be just a number, similar to how x inside g is just a number.

2.5.2 First Special Case of the Chain Rule

This special case is when the domain and the final range are the same (meaning inside function takes in something in $\mathbb{R}$, and the outer function outputs $\mathbb{R}$ as well. This is a special case because the result would instead be a dot product with the gradient of the outer function with the vector of the inner function, because the two derivatives would give a vector that the dot product could be applied to in order to create a scalar.

For this example, consider $c : U \subset \mathbb{R}^1 \to \mathbb{R}^3$, $f : V \subset \mathbb{R}^3 \to \mathbb{R}^1$.

$$\frac{dh}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt} \tag{14}$$

$$\frac{dh}{dt} = \nabla f(c(t)) \cdot c'(t) \tag{15}$$

$$\nabla f(c(t)) = (\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z})_{(x,y,z)} \tag{16}$$

$$c'(t) = (\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt})_t \tag{17}$$

2.5.3 Second Special Case of the Chain Rule

This one is pretty hard to explain. Let u, v, w, and f be functions $\mathbb{R}^3 \to \mathbb{R}^1$, and g be $\mathbb{R}^3 \to \mathbb{R}^3$.

$$g(x, y, z) = (u(x, y, z), v(x, y, z), w(x, y, z))$$

$$\nabla g(x, y, z) = \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} & \frac{\partial w}{\partial z} \end{bmatrix}_{(x,y,z)}$$

$$h(x) = f(g(x, y, z))$$

$$h'(x) = \nabla f(g(x, y, z))\nabla g(x, y, z)$$

$$h'(x, y, z) = \begin{bmatrix} \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} & \frac{\partial f}{\partial w} \end{bmatrix} \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} & \frac{\partial w}{\partial z} \end{bmatrix}_{(x,y,z)}$$

2.6 Gradients and Directional Derivatives

2.6.1 Gradients in $\mathbb{R}^3$

Recall the modified definition from 2.1:

The **gradient** of a function $f : U \subset \mathbb{R}^3 \to \mathbb{R}$ is equal to a 1 x n matrix, with each column in the matrix being the partial derivative of the n-th variable. A gradient is denoted as ∇ f.

$$\nabla f = \left(\frac{\partial f}{\partial x_1} \quad \cdots \quad \frac{\partial f}{\partial x_n} \right)$$

2.6.2 Directional Derivatives

Given a function that takes in three variables, we want to identify how a function's slope changes over time. We define a starting point of $\mathbf{x}$, which is changed by the unit vector $\mathbf{v}$ and the independent variable t. The rate of change in this function is the definition of a **directional derivative**. We also assume $f : \mathbb{R}^3 \to \mathbb{R}$.

$$\frac{d}{dt} f(\mathbf{x} + t\mathbf{v}) \Big|_{t=0}$$

Note that $\mathbf{v}$ is a unit vector, which helps describe the general movement of the input. The vector is a unit vector in order to show the change in correspondence to the change in distance by 1.

Let's look at what this actually looks like:

$$\lim_{h \to 0} \frac{f(x + hv) - f(x)}{h}$$

It seems to be the definition of the derivative, but with an extra v instead of just h. Perhaps it means that we can just multiply the end derivative by v, because in the above limit, we will only find variables of h after canceling out f(x+hv) with f(x). Consider the derivative of x^2 through the above limit to see such a proof. We will generalize a directional derivative in $\mathbb{R}^3$.

$$\mathbf{D}f(x)v = \nabla f(x)v = \left[\frac{\partial f}{\partial x}(x) \right] v_1 + \left[\frac{\partial f}{\partial y}(x) \right] v_2 + \left[\frac{\partial f}{\partial z}(x) \right] v_3$$

Note that each component of the vector is multiplied by the respective partial derivative. This equation also reflects the chainrule, considering that v is the velocity, which means the derivative of the inside function.

2.6.3 Directions of Fastest Increase

Assuming $\nabla f(x) \neq 0$, $\nabla f(x)$ points in the direction along which f is increasing the fastest. This is simply equal to the derivative of f(x) at point x, as long as the derivative at the point is not zero.

2.6.4 Gradients and Tangent Planes to Level Sets

Recall what level sets are: level sets are when you set a function $f : \mathbb{R}^n \to \mathbb{R}$, and set f to equal a constant real number, like 5. In this case, we'll consider n to be 3 ($\mathbb{R}^3$).

If $f(x_0, y_0, z_0) = k$, where k is a constant real number (level set/ surface), and c(t) is defined to be the path of x, y, z where $c(0) = (x_0, y_0, z_0)$ and $c'(0) = v$ then $\nabla f(x_0, y_0, z_0) \cdot v = 0$. Simplified, this looks like

$$0 = \frac{d}{dt} f(c(t)) \Big|_{t=0} = \nabla f(c(0)) \cdot v$$

Similar to tangent lines are the tangent planes to level surfaces, which is the dot product of the gradient f with the distance from (x_0, y_0, z_0). This assumes that the gradient is not 0.

3 Higher-Order Derivatives: Maxima and Minima

3.1 Iterated Partial Derivatives

Because derivatives of functions like f: $\mathbb{R}^3 \to \mathbb{R}$ exist and are continuous, we could potentially derive a partial derivative further with respect to one of the variables. For example, if $\frac{\partial f}{\partial x}$ is differentiable, we would say that the function is **twice continuously differentiable** given that all other partial derivatives are also differentiable. This extends further than twice; third continuously differentiable functions exist, and so on. Deriving $\frac{\partial f}{\partial x} * \frac{\partial}{\partial x}$ is an example of iterated partial derivatives, because we iterate over the derivative of the same variable. When combining two derivatives, like $\frac{\partial^2 f}{\partial x \partial y}$, we get a derivative called **mixed partial derivatives**.

3.1.1 The Mixed Partials Are Equal

Note that the order of derivatives do not matter when mixing partial derivatives of functions, which applies to any continuously differentiable function. This means that the following two are equivalent:

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$$

3.2 Not Covered

3.3 Extrema of Real-Valued functions

Let f: $\mathbb{R}^n \to \mathbb{R}^m$, then $f(x_0)$ is a local max if f(x) ¡ $f(x_0)$ for x near x_0. $f(x_0)$ is a local min if $f(x_0) \leq$ f(x) for all x near x_0. Both local minimums and maximums

are local extrema.

Note that $f'(x) = 0$ is a qualification for a local extremum, but may not be a local max or min due to a saddle point.

Critical points occur where $f'(x) = 0$ or $f'(x)$ does not exist.

Let f: $\mathbb{R}^n \to \mathbb{R}^m$, then $\vec{x_0} \in \mathbb{R}^n$ is a critical point if either f isn't differentiable at x_0 or the derivative at x_0 is 0. This theorem is mutual, so a local extremum can only happen at where the derivative is zero or does not exist, and a local extremum must imply thatthe slope is 0.

3.3.1 Second derivative test

The second derivative test for f(x, y): Let f be of class C^2, $(x_0, y_0) \in \mathbb{R}^n$.

$$\textbf{Discriminant} = D = \frac{\partial^2 f}{\partial x^2}(x_0, y_0) * \frac{\partial^2 f}{\partial x^2}(x_0, y_0) - \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)$$

The **discriminant** helps tell us about a point given that the first partial derivatives all equal zero. There are three results for discriminant:

1. If D ¿ 0 and $\frac{\partial f}{\partial x^2}$ ¿ 0, f has a local minimum at that point.

2. If D ¿ 0 and $\frac{\partial f}{\partial x^2}$ ¡ 0, f has a local maximum at that point.

3. If D ¡ 0, then f has a saddle point at that point.

4. If D = 0, test is inconclusive.

3.3.2 Closed and bounded critical points

We must define what closed and bounded mean.

1. **Closed** means that the function contains all of its boundary points, such as $x^2 + y^2 = 1$.

2. **Bounded** means that the function is contained in some disk, meaning it is finite, and does not expand to infinity in any direction.

The definitions for open and unbounded should be fairly intuitive.

If f is a continuous function on a closed boudned set, D, in $\mathbb{R}^2$, then f attains an absolute max and an absolute min at some points x_0, x_1 in D. This simply means there has to be some point where the maximum and minimum are achieved in a closed area.

The method to find min and max is the following:

1. Find f at critical points in the interior of D (where derivative is zero)

2. Find f at extrema on the boundary.

3. Select the most extreme values from 1 and 2, highest is absolute max, lowest is absolute min.