Multivariable Calculus Notes

Jayden Li

2025-2026 Academic Year

Contents

1 Vectors

1.1 Basics

Let u, v be vectors in $\mathbb{R}^n$. They will have n components, and let them be as follows, where $u_i, v_i \in \mathbb{R}$ (scalars):

$$u = (u_1, u_2, \ldots, u_n)$$
$$v = (v_1, v_2, \ldots, v_n)$$

We can define the addition operation and the magnitude of each vector as:

$$u + v = (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n)$$
$$\|u\| = \sqrt{u_1^2 + u_2^2 + \ldots + u_n^2} = \sqrt{\sum_{i=1}^{n} u_i}$$

Since scalar addition is commutative, so is vector addition. Also, the subtraction operation is defined as:

$$u - v = u + (-v) = (u_1 - v_1, u_2 - v_2, \ldots, u_n - v_n)$$

If we have two points A and B, the vector AB from A to B is given by $AB = B - A$. If we placed the tail of the vector at A, then the head will point to B.

Two vectors are parallel if and only if they are scalar multiples of each other; that is, there exists a scalar $a \in \mathbb{R}$ such that $u = av$.

1.2 Polar Form

For a vector $u \in \mathbb{R}^2$, we can write u in polar form. Any vector can be specified with an angle from the x axis and a magnitude. Let the angle from the x axis be θ and the magnitude, the distance to the origin, be r. To convert from Cartesian to polar:

$$u = \begin{bmatrix} u_x \\ u_y \end{bmatrix} \implies \begin{cases} \theta = \arctan\left(\dfrac{u_y}{u_x}\right) \\ r = \|u\| = \sqrt{u_x^2 + u_y^2} \end{cases}$$

To convert from polar to Cartesian:

$$u = \begin{bmatrix} r\cos\theta \\ r\sin\theta \end{bmatrix} = r\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$$

Note that since $\sin^2\theta + \cos^2\theta = 1$ by the Pythagorean identity, the magnitude of u is r.

1.3 Unit Vectors

A unit vector is a vector of magnitude 1. For a vector $u \in \mathbb{R}^n$, a unit vector with the same direction is $u/\|u\|$.

1.4 Dot Product

For vectors $u = (u_1, \ldots, u_n) \in \mathbb{R}^n$ and $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$, the cross product of the two vectors $u \cdot v$ is:

$$u \cdot v = u_1v_1 + u_2v_2 + \ldots + u_nv_n = \sum_{i=1}^{n} u_iv_i$$

Since scalar multiplication is commutative, so is the dot product: $u \cdot v = v \cdot u$.

Alternatively, let θ be the angle between vectors u and v. Then:

$$u \cdot v = \|u\|\|v\|\cos\theta$$

1.5 Orthogonality

Two vectors u and v are orthogonal if and only if $u \cdot v = 0$.

1.6 Standard Bases

The standard basis vectors are as follows:

$$\begin{aligned} \hat{\imath} &= (1, 0, 0) \\ \hat{\jmath} &= (0, 1, 0) \\ \hat{k} &= (0, 0, 1) \end{aligned}$$

These vectors span the $\mathbb{R}^3$ space.

1.7 Angle Between Vectors

We can calculate the angle between vectors $u, v \in \mathbb{R}^n$ using the dot product:

$$u \cdot v = \|u\|\|v\|\cos\theta \implies \cos\theta = \frac{u \cdot v}{\|u\|\|v\|}$$

1.8 Cross Product

The cross product only exists between vectors in three dimensions. Let $u = (u_x, u_y, u_z) \in \mathbb{R}^3$ and $v = (v_x, v_y, v_z) \in \mathbb{R}^3$. The cross product $u \times v$ has the following properties:

- $u \times v$ is a vector in $\mathbb{R}^3$.
- $u \times v$ is orthogonal to both u and v. If u and v are linearly independent and span a plane, $u \times v$ is orthogonal to that plane as well.
- $u \times v = 0$ if and only if u and v are parallel/linearly dependent.
- The magnitude of $u \times v$ equals the area of the parallelogram formed by u and v: $\|u \times v\| = \|u\|\|v\| \sin\theta$

The formula for $u \times v$ can be written concisely with a determinant, and expanding by cofactors:

$$u \times v = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ u_x & u_y & u_z \\ v_x & v_y & v_z \end{vmatrix} = \hat{i} \begin{vmatrix} u_y & u_z \\ v_y & v_z \end{vmatrix} - \hat{j} \begin{vmatrix} u_x & u_z \\ v_x & v_z \end{vmatrix} + \hat{k} \begin{vmatrix} u_x & u_y \\ v_x & v_y \end{vmatrix}$$

The operation has the following properties:

- Anticommutativity: $u \times v = -(v \times u)$
- Distributive over addition: $(u + v) \times w = u \times w + v \times w$ and $u \times (v + w) = u \times v + u \times w$.
- Multiplication by a scalar: $(cu) \times v = c(u \times v) = u \times (cv)$ where $c \in \mathbb{R}$.

The scalar triple product is the determinant of the square matrix with those vectors as its rows (or columns):

$$u \cdot (v \times w) = (v \times w) \cdot u = \begin{vmatrix} u_x & u_y & u_z \\ v_x & v_y & v_z \\ w_x & w_y & w_z \end{vmatrix}$$

From the determinant, it follows that the scalar triple product of u, v, w is the volume of the parallelepiped with u, v, w as its edges.

1.9 Projection

The orthogonal projection of u onto v, written $\text{proj}_v u$, is the vector parallel to v such that the error between the projection and u is minimized.

Let p be the orthogonal projection $\text{proj}_v u$. The error of the project is $e = u - p$. The error is minimized when e is orthogonal to v. The formula for orthogonal projection is:

$$\text{proj}_v u = \frac{u \cdot v}{v \cdot v} v = \frac{u \cdot v}{\|v\|^2} v$$

The proof is left as an exercise to the reader.

Also, for an $n \times m$ matrix A with independent columns (and therefore $m \leq n$), the orthogonal projection of a vector $b \in \mathbb{R}^n$ onto the column space of A (vectors in $C(A)$ are n-dimensional, because it has n rows) is:

$$p = A(A^T A)^{-1} A^T b \in \mathbb{R}^n$$

Since A has independent columns, $A^T A$ is a full rank, square matrix, which has an inverse. This fact may occasionally be useful for projecting onto planes.

2 Lines

A line in 3D space can be represented in three ways. In each subsection below, we specify a line passing through (x_0, y_0, z_0) with slopes a, b, c.

2.1 Parametric

This parametric system with parameter t describes the line we want:

$$\begin{cases} x(t) = at + x_0 \\ y(t) = bt + y_0 \\ z(t) = ct + z_0 \end{cases}$$

2.2 Vector

A vector valued function $f : \mathbb{R} \to \mathbb{R}^n$. It takes a scalar and returns a vector: $f(t) \in \mathbb{R}^n, t \in \mathbb{R}$.

From the parametric equation of a line, we notice that for parameter t, this point lies on the line. We can define a vector valued function r with parameter t. This is the vector equation for a line.

$$r(t) = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} at + x_0 \\ bt + y_0 \\ ct + z_0 \end{bmatrix}$$

2.3 Symmetric

Recall the parametric system for a line. We solve for the parameter t in terms of x, y, z values:

$$\begin{cases} x(t) = at + x_0 \\ y(t) = bt + y_0 \\ z(t) = ct + z_0 \end{cases} \implies \begin{cases} t = \dfrac{x - x_0}{a} \\ t = \dfrac{y - y_0}{b} \\ t = \dfrac{z - z_0}{c} \end{cases}$$

For each equation in the above system, the parameter t must be equal, yielding the symmetric equation of a line:

$$\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}$$

2.4 Finding Intersection

Suppose we have two lines with the following parametric equations:

$$\begin{cases} x_1(t_1) = a_1 t_1 + x_{10} \\ y_1(t_1) = b_1 t_1 + y_{10} \\ z_1(t_1) = c_1 t_1 + z_{10} \end{cases} \qquad \begin{cases} x_2(t_2) = a_2 t_2 + x_{20} \\ y_2(t_2) = b_2 t_2 + y_{20} \\ z_2(t_2) = c_2 t_2 + z_{20} \end{cases}$$

These two lines intersect exactly where $(x_1, y_1, z_1) = (x_2, y_2, z_2)$.

$$\begin{cases} a_1 t_1 + x_{10} = a_2 t_2 + x_{20} \\ b_1 t_1 + y_{10} = b_2 t_2 + y_{20} \\ c_1 t_1 + z_{10} = c_2 t_2 + z_{20} \end{cases}$$

We know the value of a, b, c, x_0, y_0, z_0 for both equations. There are two variables to solve: t_1 and t_2. But this is an overdetermined system, with three linear equations for two variables, so there is only a solution (and intersection) if these equations are not independent. If there is, then we can use t_1 and t_2 to find the point (or infinity points) of intersection.

2.5 Distance between Lines

First find two parallel planes which contains each line, and find the distance between the parallel planes. This is the distance betwene lines.

The normal vector to the parallel planes is the cross product of the direction vectors of the two lines. Then translate the planes so that they contain the lines.

3 Planes

The equation of a plane in 3D passing through (x_0, y_0, z_0) is:

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

$$\begin{bmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0$$

Alternative, expanding the first equation, there exists a scalar value d such that the plane is described by:

$$ax + by + cz + d = 0$$

Subtracting by x_0, y_0, z_0 shifts the plane. The plane parallel to the original plane, but which passes through the origin, is given by $(x, y, z) \cdot (a, b, c) = 0$. Which means that for all points on the plane, (a, b, c) is orthogonal to it. Hence $n = (a, b, c)$ is orthogonal to the plane.

3.1 Dihedral Angle

The dihedral angle is the angle formed between two planes. It equals the angle formed between the normal vectors of each plane, which can be found using the dot product.

3.2 Angle between Vector and Plane

Let v be the vector. Find the normal vector to the plane n, and find the angle θ between the vector and n. The angle between the plane and v is $90° - \theta = \pi/2 - \theta$.

3.3 Finding Intersection

Two planes intersect if and only if the vectors normal to the plane point in different direction and are not parallel.

$$a_1x + b_1y + c_1z + d_1 = 0$$
$$a_2x + b_2y + c_2z + d_2 = 0$$

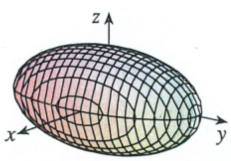
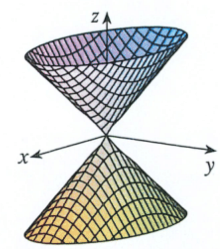
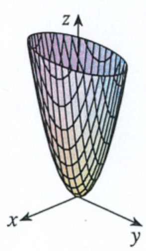
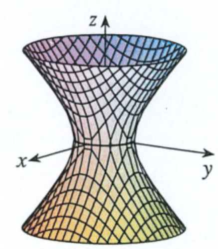
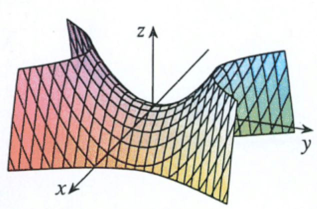
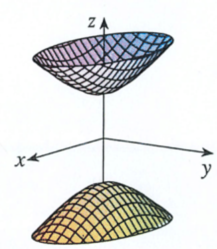
A point on the intersection of these two planes satisfy both equations. From here, set any variables (x, y, z) to a parameter t and substitute. Then use elimination to solve for the two remaining variables in terms of t (DO NOT ELIMINATE t). Now we have a parametric system describing the line of intersection.

3.4 Distance between Parallel Planes

The two lines will have the same unit normal vector n. They have the form $(x, y, z) \cdot n = a$ and $(x, y, z) \cdot n = b$. Choose any point on one plane P. We need to calculate a scalar c such that $P + cn$ is on the other plane. The distance between the planes is $\|cn\| = c$ since n is a unit vector.

4 Cylinders and Quadric Surfaces

Table 1 Graphs of Quadric Surfaces

Surface	Equation	Surface	Equation
Ellipsoid  	$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ All traces are ellipses. If $a = b = c$, the ellipsoid is a sphere.	**Cone**  	$\frac{z^2}{c^2} = \frac{x^2}{a^2} + \frac{y^2}{b^2}$ Horizontal traces are ellipses. Vertical traces in the planes $x = k$ and $y = k$ are hyperbolas if $k \neq 0$ but are pairs of lines if $k = 0$.
Elliptic Paraboloid  	$\frac{z}{c} = \frac{x^2}{a^2} + \frac{y^2}{b^2}$ Horizontal traces are ellipses. Vertical traces are parabolas. The variable raised to the first power indicates the axis of the paraboloid.	**Hyperboloid of One Sheet**  	$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$ Horizontal traces are ellipses. Vertical traces are hyperbolas. The axis of symmetry corresponds to the variable whose coefficient is negative.
Hyperbolic Paraboloid  	$\frac{z}{c} = \frac{x^2}{a^2} - \frac{y^2}{b^2}$ Horizontal traces are hyperbolas. Vertical traces are parabolas. The case where $c < 0$ is illustrated.	**Hyperboloid of Two Sheets**  	$-\frac{x^2}{a^2} - \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ Horizontal traces in $z = k$ are ellipses if $k > c$ or $k < -c$. Vertical traces are hyperbolas. The two minus signs indicate two sheets.

4.1 Trace

The intersection of a quadric surface with a plane normal to a standard basis vector is a trace. We can use the trace produced by a equation in each direction to determine which quadric surface the equation is.

4.2 Cylinders

A cylinder is a surface that consists of all lines that are parallel to a given line, and which passes through a plane curve. For example, the surface given by $y = x^2$ is a cylinder because the z component of each point can be anything, hence every line is parallel to $\hat{k}$.

5 Vector Functions

A vector a function $r : \mathbb{R} \to \mathbb{R}^n$ in n dimensions takes a scalar parameter and returns a vector in $\mathbb{R}^n$.

Let $r_1(t), r_2(t), \ldots, r_n(t)$ be the components of $r(t)$, such that $r_i : \mathbb{R} \to \mathbb{R}$ and $r(t) = (r_1(t), r_2(t), \ldots, r_t(t))$.

5.1 Limits

$$\lim_{t\to a} r(t) = \left(\lim_{t\to a} r_1(t), \lim_{t\to a} r_2(t), \ldots, \lim_{t\to a} r_n(t)\right)$$

5.2 Derivatives

$$r'(t) = \lim_{h\to 0} \frac{r(t+h) - r(t)}{h} = \left(\ldots, \lim_{h\to 0} \frac{r_i(t+h) - r_i(t)}{h}, \ldots\right) = (r_1'(t), r_2'(t), \ldots, r_n'(t))$$

Suppose $f : \mathbb{R} \to \mathbb{R}$ is a scalar function. Derivative rules for vector functions $(u, v : \mathbb{R} \to \mathbb{R}^n)$:

$$\frac{\mathrm{d}}{\mathrm{d}t}(u(t) + v(t)) = u'(t) + v'(t)$$
$$\frac{\mathrm{d}}{\mathrm{d}t}(u(t) \cdot v(t)) = u'(t) \cdot v(t) + u(t) \cdot v'(t)$$
$$\frac{\mathrm{d}}{\mathrm{d}t} u(f(t)) = u'(f(t))f'(t)$$

5.3 Integrals

$$\int_a^b r(t)\,\mathrm{d}t = \left(\int_a^b r_1(t)\,\mathrm{d}t, \int_a^b r_2(t)\,\mathrm{d}t, \ldots, \int_a^b r_n(t)\,\mathrm{d}t\right)$$

5.4 Arc Length

The arc length of a vector function between two parameters is given by:

$$S = \int_{t_1}^{t_2} \sqrt{\sum_{i=1}^{n} (r_i'(t))^2}\,\mathrm{d}t$$

In 2 and 3 dimensions, where $r(t) = (x(t), y(t))$ and $r(t) = (x(t), y(t), z(t))$:

$$S = \int_{t_1}^{t_2} \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2}\,\mathrm{d}t$$
$$S = \int_{t_1}^{t_2} \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}z}{\mathrm{d}t}\right)^2}\,\mathrm{d}t$$

5.4.1 Arc Length Parameterization

For a vector function between t_1 and t_2, the arc length between t_1 and an arbitrary parameter $t \in [t_1, t_2]$ is:

$$s(t) = \int_{t_1}^{t} \sqrt{\sum_{i=1}^{n} (r_i'(u))^2}\, \mathrm{d}u = \int_{t_1}^{t} \|r'(u)\|\, \mathrm{d}u$$

The arc length function is always increasing, because $s'(t) = \|r(t)\| > 0$.

We can reparameterize (change parameter) of the original function r to arrive at the arc length parameterization. From $s(t) = \int_{t_1}^{t} \|r(u)\|\, \mathrm{d}u$, solve for the original parameter t in terms of s, and change the parameter of $r(t)$ to $r(s)$. Change the bounds of the parameter accordingly.

In the arc length parameterization, the distance traveled (arc length) on two intervals of equal length are always the same. **The arc length on the interval $a \leq t \leq b$ is equal to $b - a$.** That is:

$$\int_{a}^{b} \|r'(s)\|\, \mathrm{d}s = b - a$$

which is true if $\|r'(s)\| = 1$ ($\int_a^b 1\, \mathrm{d}s = b - a$).

5.5 Curvature

For a vector valued function $r : \mathbb{R} \to \mathbb{R}^n$, we find its arc length parameterization $r(s)$. Because the magnitude of the derivative of the arc length parameter $\|r'(s)\| = 1$, the derivative of $r(s)$ equals the unit tangent vector function $\hat{T}$: $\hat{T}(s) = r'(s)$.

The curvature of the curve is how quickly the unit tangent vector $\hat{T}(s)$ changes with respect to the arc length. But $\|\hat{T}(s)\| = 1$ (unit vector), acceleration only comes from changes in direction. The curvature $\kappa(s) = \|T'(s)\|$ measures how curved the path is.

For some vector valued function $r(t) = (r_1(t), r_2(t), \ldots, r_n(t))$ where $r_i : \mathbb{R} \to \mathbb{R}$ on the interval $a \leq t \leq b$:

$$s(t) = \int_{a}^{t} \|r'(u)\|\, \mathrm{d}u \qquad \text{(Arc length parameter)}$$

Then solve for t in terms of the arc length parameter s and substitute t for this function of s: $t = t(s)$, for the arc length parameterization $r(s)$.

$$\hat{T}(s) = \frac{r'(s)}{\|r'(s)\|} = r'(s) \qquad \text{(Unit tangent vector)}$$

$$\hat{T}'(s) = \frac{\mathrm{d}\hat{T}}{\mathrm{d}s} \qquad \text{(Acceleration)}$$

$$\boxed{\kappa(s) = \|\hat{T}'(s)\| = \|r''(s)\|} \qquad \text{(Curvature)}$$

The curvature in terms of t can be found by the chain rule, because s is a function of t:

$$\hat{T}'(t) = \frac{\mathrm{d}}{\mathrm{d}t}\hat{T}(s(t)) = \hat{T}'(s(t))s'(t) = \hat{T}'(s)\|r'(t)\|$$

$$\implies \hat{T}'(s) = \frac{\hat{T}'(t)}{\|r'(t)\|} \implies \|\hat{T}'(s)\| = \boxed{\kappa(t) = \frac{\|\hat{T}'(t)\|}{\|r'(t)\|}}$$

(Since $s(t) = \int_a^t \|r'(u)\|\, \mathrm{d}u$, by the Fundamental Theorem of Calculus $s'(t) = \|r'(t)\|$. $\kappa(s) = \kappa(t)$ because the arc length s equals the parameter t by definition.)

The curvature at any point of a circle is the reciprocal of its radius.

The osculating circle is a circle that best approximates the curvature of the function at a given point. The radius of the osculating circle is the reciprocal of the curvature (the circle and the function have the same curvature). The center of the osculating circle is the center of curvature.

5.6 Normal Vector

The unit normal vector $\hat{N}(t)$ is the unit vector normal to the curve and the tangent vector.

$$\hat{N}(t) = \frac{\hat{T}'(t)}{\|\hat{T}'(t)\|}$$

By this definition, $\hat{N}(t)$ points towards the center of curvature.

The center of the osculating circle and the center of curvature is $\hat{N}(t)/\kappa$ from the point $r(t)$.

5.7 Binormal Vector

$\hat{T}(t)$ and $\hat{N}(t)$ form a plane, known as the osculating plane. The unit binormal vector is orthogonal to the osculating plane, and is given by:

$$\hat{B}(t) = \hat{T}(t) \times \hat{N}(t)$$

$\hat{B}(t)$ is already a unit vector, because ($\hat{T}(t)$ and $\hat{N}(t)$ are orthogonal so the angle $\theta = \pi/2$):

$$\|\hat{B}(t)\| = \|\hat{T}(t) \times \hat{N}(t)\| = \|\hat{T}(t)\|\|\hat{N}(t)\| \sin\theta = (1)(1)\sin\left(\frac{\pi}{2}\right) = 1$$

The unit tangent, unit normal and unit binormal vectors span and form a basis for the $\mathbb{R}^3$ space. They are also orthogonal to each other. This is called the TNB basis.

Since the magnitude of $\hat{B}(t)$ is 1:

$$\|\hat{B}(t)\|^2 = 1 \implies \frac{\mathrm{d}}{\mathrm{d}t}(\hat{B}(t) \cdot \hat{B}(t)) = \frac{\mathrm{d}}{\mathrm{d}t}1 = 0 \implies 2\hat{B}'(t) \cdot \hat{B}(t) = 0 \implies \hat{B}'(t) \cdot \hat{B}(t) = 0$$

So $\hat{B}'(t)$ is orthogonal to $\hat{B}(t)$. By definition $\hat{B}(t) = \hat{T}(t) \cdot \hat{N}(t)$:

$$\hat{B}'(t) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\hat{T}(t) \times \hat{N}(t)\right) = \hat{T}'(t) \times \hat{N}(t) + \hat{T}(t) \times \hat{N}'(t) = \cancelto{0}{\hat{N}(t) \times \hat{N}(t)} + \hat{T}(t) \times \hat{N}'(t) = \hat{T}(t) \times \hat{N}'(t)$$

So $\hat{B}'(t)$ is also orthogonal to $\hat{T}(t)$. Since $\hat{T}(t), \hat{N}(t), \hat{B}(t)$ form an orthogonal basis, $\hat{B}'(t)$ is parallel to $\hat{N}(t)$.

5.8 Torsion

From above, the unit binormal vector is parallel to the unit normal vector. Using the arc length parameterization, let τ be the value such that:

$$\hat{B}'(s) = -\tau\hat{N}(s)$$

The scalar value τ is torsion. It measures how much the curve "twists," or how much it deviates or lifts from the osculating plane at any given point. Taking the dot product of $\hat{N}(s)$ on both sides:

$$\begin{aligned} \hat{B}'(s) \cdot \hat{N}(s) &= -\tau\left(\hat{N}(s) \cdot \hat{N}(s)\right) \\ &= -\tau\|N(s)\|^2 \\ &\boxed{\tau = -\hat{B}'(s) \cdot \hat{N}(s)} \end{aligned}$$

It is not convenient to calculate τ in terms of the arc length parameter, we can apply the chain rule. Recall that $\mathrm{d}s/\mathrm{d}t = \|r'(t)\|$:

$$\tau = -\frac{\mathrm{d}\hat{B}}{\mathrm{d}s} \cdot \hat{N} = -\frac{\mathrm{d}\hat{B}/\mathrm{d}t}{\mathrm{d}s/\mathrm{d}t} \cdot \hat{N} = \boxed{-\frac{\hat{B}(t) \cdot \hat{N}(t)}{\|r'(t)\|}}$$

By convention, positive torsion means the curve is "twisting up" from the osculating plane.

5.9 Tangential and Normal Acceleration

We can break acceleration down into components: one parallel to the direction of motion (velocity) and one orthogonal to it. In physics this would be the linear acceleration and tangential acceleration.

Velocity v is defined as the first derivative:

$$v(t) = r'(t) = \underbrace{\|r'(t)\|}_{\mathrm{d}s/\mathrm{d}t} \underbrace{\frac{r'(t)}{\|r'(t)\|}}_{\hat{T}(t)} = \left(\frac{\mathrm{d}s}{\mathrm{d}t}\right) \hat{T}(t)$$

Acceleration a is the second derivative or the derivative of velocity v:

$$a(t) = v'(t) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)\hat{T}(t)\right) = \left(\frac{\mathrm{d}^2 s}{\mathrm{d}t^2}\right)\hat{T}(t) + \left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)\hat{T}'(t)$$

Recall the definition of the unit normal vector $\hat{N}(t) = \hat{T}'(t)/\|\hat{T}'(t)\|$ and of curvature $\kappa = \|\hat{T}'(t)\|/(\mathrm{d}s/\mathrm{d}t)$:

$$= \left(\frac{\mathrm{d}^2 s}{\mathrm{d}t^2}\right)\hat{T}(t) + \left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^2 \left(\frac{\|T'(t)\|}{\frac{\mathrm{d}s}{\mathrm{d}t}}\right) \frac{\hat{T}'(t)}{\|T'(t)\|} = \left(\frac{\mathrm{d}^2 s}{\mathrm{d}t^2}\right)\hat{T}(t) + \left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^2 \kappa\hat{N}(t)$$

Arc length s and curvature κ are scalars. Acceleration can be written as a linear combination of the unit tangent vector $\hat{T}(t)$ and the unit normal vector $\hat{N}(t)$. Thus, we can write a_T, a_N, the tangential and normal components of acceleration, respectively:

$$\boxed{a_T = \frac{\mathrm{d}^2 s}{\mathrm{d}t^2} \qquad a_N = \kappa\left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^2}$$

It then follows that acceleration can be expressed as:

$$a(t) = a_T\hat{T}(t) + a_N\hat{N}(t)$$

5.10 Alternative Formula for Curvature

From before, we have:

$$r'(t) = v(t) = \left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)\hat{T}(t) \qquad r''(t) = a(t) = \left(\frac{\mathrm{d}^2 s}{\mathrm{d}t^2}\right)\hat{T}(t) + \left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)\hat{T}'(t)$$

Taking the cross product of these two functions:

$$\begin{aligned}
r'(t) \times r''(t) &= \left(\left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)\hat{T}(t)\right) \times \left(\left(\frac{\mathrm{d}^2 s}{\mathrm{d}t^2}\right)\hat{T}(t) + \left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)\hat{T}'(t)\right) \\
&= \left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)\hat{T}(t) \times \left(\frac{\mathrm{d}^2 s}{\mathrm{d}t^2}\right)\hat{T}(t) + \left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)\hat{T}(t) \times \left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)\hat{T}'(t) \\
&= \left(\frac{\mathrm{d}s}{\mathrm{d}t}\frac{\mathrm{d}^2 s}{\mathrm{d}t^2}\right)\left(\cancelto{0}{\hat{T}(t) \times \hat{T}(t)}\right) + \left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^2 \left(\hat{T}(t) \times \hat{T}'(t)\right) \\
&= \left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^2 \left(\hat{T}(t) \times \hat{T}'(t)\right)
\end{aligned}$$

Since $\hat{T}(t) \times \hat{T}(t) = 0$. Since $\hat{T}(t)$ and $\hat{T}'(t)$ are orthogonal, the magnitude of the cross product $\hat{T}(t) \times \hat{T}'(t)$ is $\|\hat{T}(t)\|\|\hat{T}'(t)\| = \|\hat{T}'(t)\|$ since $\hat{T}(t)$ is a unit vector.

$$\|r'(t) \times r''(t)\| = \left\|\left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^2 \left(\hat{T}(t) \times \hat{T}'(t)\right)\right\| = \left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^2 \|T'(t)\| = \left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^2 \frac{\|T'(t)\|}{\mathrm{d}s/\mathrm{d}t}\frac{\mathrm{d}s}{\mathrm{d}t}$$

Definition of curvature $\kappa = \|T'(t)\|/(\mathrm{d}s/\mathrm{d}t)$:

$$= \left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^2 \kappa \frac{\mathrm{d}s}{\mathrm{d}t} = \left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^3 \kappa$$

From the definition of arc length:

$$s(t) = \int_a^t \|r'(u)\|\,\mathrm{d}u \implies \frac{\mathrm{d}s}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\int_a^t \|r'(u)\|\,\mathrm{d}u = \|r'(t)\| \implies \left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^3 = \|r'(t)\|^3$$

Dividing these:

$$\frac{\left(\frac{\mathrm{d}s}{\mathrm{d}s}\right)^3 \kappa}{\left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^3} = \boxed{\kappa = \frac{\|r'(t) \times r''(t)\|}{\|r'(t)\|^3}}$$

Note that the magnitude of the cross product can be interpreted as the area of the parallelogram formed by the two vectors. When working with two dimensional functions, the actual cross product cannot be calculated, but its magnitude's interpretation as the area can be found through the determinant:

$$\kappa = \frac{1}{\|r'(t)\|^3}\begin{vmatrix} r'(t) \\ r''(t) \end{vmatrix}$$

6 Partial Derivatives

A multivariable function takes in multiple variables. The domain is a n-tuple of numbers, denoted $\mathbb{R}^n$. The codomain is the set the function maps to, often $\mathbb{R}$.

6.1 Cylindrical Coordinates

Cylindrical coordinates are (r, θ, z), where r and θ map onto x and y in the same way as polar coordinates in 2 dimensions:

$$\begin{aligned}
x &= r\cos\theta \\
y &= r\sin\theta
\end{aligned}$$

and z is just the z coordinate. As with polar coordinates:

$$r^2 = x^2 + y^2$$
$$\tan\theta = \frac{y}{x}$$

6.2 Limits

A multivariable limit is denoted:

$$\lim_{(x,y)\to(a,b)} f(x,y) = L$$

where x and y approach a and b at the same time. For the limit to exist, approaching (a,b) from all (infinitely many) paths must evaluate to the same value. In 2 dimensions, there are only 2 paths; approaching from the left and the right.

To show a limit does not exist at a given point, show that approacing the point from two different paths produces different values. This can be done by setting y to a function of x, or vice versa.

For limits to the origin, we can evaluate the limit by changing $f(x,y)$ to cylindrical/polar coordinates $g(r,\theta)$:

$$\lim_{(x,y)\to(a,b)} f(x,y) = \lim_{r\to 0} g(r,\theta)$$

The limit exists if and only if it produces the same value for all values of θ. If it does not, then approaching from different directions (different paths) produce different values, and hence the limit does not exist.

6.3 Continuity

A function f is continuous at the point (a,b) if and only if:

$$\lim_{(x,y)\to(a,b)} f(x,y) = f(a,b)$$

All polynomials in x and y are continuous. Limits of known continuous functions can be evaluated by direct substitution, like in single variable calculus.

6.4 Differentiability

A function $f : \mathbb{R}^n \to \mathbb{R}$ is continuous at a point a if there exists a linear function L such that (where $x, a \in \mathbb{R}^n$):

$$\lim_{x\to a} \frac{|f(x) - L(x)|}{\|x - a\|} = 0$$

which means the denominator $\|x - a\|$ grows slower than the numerator $|f(x) - L(x)|$, which is that the linear function L approaches f faster than x approaches a.

Like in single variable calculus, a function is not differentiable at a point if there is a sharp turn, vertical asymptotes, etc.

6.5 Partial Derivatives

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function with variables $x_1, x_2, \ldots, x_n$. The partial derivative with respect to any variable is the derivative if we hold all other variables constant, and denoted as:

$$\frac{\partial f}{\partial x_i} = f_{x_i}(\ldots)$$

f must be differentiable at a point if its partial derivatives are continuous at that point. But this is not a necessary condition (i.e. there are differentiable functions without continuous partial derivatives).

Clairaut's theorem states that if a function is defined on a point (a, b) and its second partial derivatives f_{xy} and f_{yx} are both continuous, then $f_{xy}(a, b) = f_{yx}(a, b)$.

6.6 Tangent Plane and Normal Line

The tangent plane to the curve of $f(x, y, z)$ at (a,b) is:

$$z = f_x(a, b)(x - a) + f_y(a, b)(y - b) + f(a, b)$$

The x coefficient is the partial derivative with respect to x and the y coefficient is the partial derivative with respect to y. x, y are shifted by a, b, respectively, and we add $f(a, b)$ so that it will pass through the point on the curve (a, b).

Using this formula, the tangent line to the curve $f(x)$ at a is (looks familiar):

$$y = f_x(a)(x - a) + f(a) = f'(a)(x - a) + f(a)$$

The normal line at a given point is the line orthogonal to the tangent plane and passing through said point. The normal line to a surface defined by $F(x, y, z) = 0$ at (x_0, y_0, z_0) is:

$$\begin{aligned} x &= x_0 + F_x(x_0, y_0, z_0)t \\ y &= y_0 + F_y(x_0, y_0, z_0)t \\ z &= z_0 + F_z(x_0, y_0, z_0)t \end{aligned}$$

To calculate the tangent plane to $F(x, y, z)$ at (x_0, y_0, z_0), first observe that the gradient:

$$\nabla F(x_0, y_0, z_0) = (F_x(x_0, y_0, z_0), F_y(x_0, y_0, z_0), F_z(x_0, y_0, z_0))$$

is parallel to the normal line, which is orthogonal to the tangent plane. Therefore, the tangent plane is given by:

$$(x - x_0, y - y_0, z - z_0) \cdot \nabla F = 0$$

To calculate the normal line to a surface given by $z = f(x, y)$, express it as:

$$F(x, y, z) = f(x, y) - z = 0$$

Notice that $F_z(x, y, z) = -1$. The normal line to this surface at (x_0, y_0, z_0) where $z_0 = f(x_0, y_0)$ is:

$$\begin{aligned} x &= x_0 + F_x(x_0, y_0, z_0)t = x_0 + f_x(x_0, y_0)t \\ y &= y_0 + F_y(x_0, y_0, z_0)t = y_0 + f_y(x_0, y_0)t \\ z &= z_0 + F_z(x_0, y_0, z_0)t = z_0 - t \end{aligned}$$

6.7 Directional Derivatives

The directional derivative is a measure of the rate of change of a function along a certain direction vector.

The directional derivative of a function $f : \mathbb{R}^n \to \mathbb{R}$ with respect to a unit vector $\mathbf{u} \in \mathbb{R}^n$ is:

$$D_{\mathbf{u}} f(\mathbf{x}) = \lim_{h \to 0} \frac{f(\mathbf{x} + \mathbf{u}) - f(\mathbf{x})}{h}$$

If $\mathbf{u} = (u_1, u_2, \dots, u_n)$ and $\|\mathbf{u}\| = 1$, the directional derivative with respect to $\mathbf{u}$ is:

$$D_{\mathbf{u}} f(\mathbf{x}) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} u_i = \nabla f(\mathbf{x}) \cdot \mathbf{u}$$

where ∇f is the gradient. For $f : \mathbb{R}^2 \to \mathbb{R}$ and $\mathbf{u} = (a, b)$:

$$D_{\mathbf{u}} f(x, y) = a\frac{\partial f}{\partial x} + b\frac{\partial f}{\partial y} = a f_x(x, y) + b f_y(x, y)$$

6.8 Gradient

For a function $f : \mathbb{R}^n \to \mathbb{R}$, the gradient $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$ is defined as:

$$\nabla f(x_1, x_2, \dots, x_n) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n}\right)$$

In two dimensions:

$$\nabla f(x, y) = (f_x(x, y), f_y(x, y))$$

The gradient at a point is the direction of greatest change at that point. That is, if $\mathbf{u} \in \mathbb{R}^n$, $D_{\mathbf{u}} f$ is greatest when $\mathbf{u}$ has the same direction as $\nabla f(\mathbf{x})$.

For some value $c \in \mathbb{R}$, the level curve of value c is the curve defined by $f(\mathbf{x}) = c$. At some point $\mathbf{x}_1$, the gradient at that point $\nabla f(\mathbf{x}_1)$ is always perpendicular to the level curve of value $f(\mathbf{x}_1)$ when graphed on the same plane.

6.9 Relative Extrema

$\mathbf{x}_1$ is a critical point of $f : \mathbb{R}^n \to \mathbb{R}$ if $\nabla f(\mathbf{x}_1) = 0$ or it does not exist.

To determine if $\mathbf{x}_1$ is a local minimum, local maximum or a saddle point, calculate the Hessian matrix:

$$H = \begin{bmatrix} f_{xx}(\mathbf{x}_1) & f_{xy}(\mathbf{x}_1) \\ f_{yx}(\mathbf{x}_1) & f_{yy}(\mathbf{x}_1) \end{bmatrix}$$

The Second Partials Test:

- If $\det H > 0$, $\mathbf{x}_1$ is a local minimum if $f_{xx}(\mathbf{x}_1) > 0$ and a local maximum if $f_{xx}(\mathbf{x}_1) < 0$.
- If $\det H < 0$, $\mathbf{x}_1$ is a saddle point.
- If $\det H = 0$, the test is inconclusive.

A saddle point is one in which some paths passing through the point results in a local minimum, and some other paths result in a local maximum.

6.10 Chain Rule

For one independent variable: if $x = g(t), y = g(t)$ and $z = f(x, y) = f(g(t), h(t))$:

$$\frac{\mathrm{d}z}{\mathrm{d}t} = \frac{\partial z}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial z}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t}$$

For two independent variables: if $x = g(u, v), y = h(u, v)$ and $z = f(x, y) = f(g(u, v), h(u, v))$:

$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial u} \qquad \frac{\partial z}{\partial v} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial v}$$

6.11 Lagrange Multipliers

We can find the absolute/relative extrema inside a boundary using by identifying points where $\nabla f = 0$. Lagrange Multipliers will find relative extrema on the boundary.

Let $f, g : \mathbb{R}^n \to \mathbb{R}$ and $k \in \mathbb{R}$. We find the relative extrema of f on the boundary $g(\mathbf{x}) = k$. Potential optimal points $\curvearrowleft_1$ on the boundary are where the gradients of f and g are scalar multiples, that is:

$$\nabla f(\mathbf{x}_1) = \lambda \nabla g(\mathbf{x}_1)$$
$$g(\mathbf{x}_1) = k$$

where $\lambda \in \mathbb{R}$, and is called the Lagrange Multiplier. Then, absolute minimum/maximum on the boundary can be found by comparing the values of $f(\mathbf{x}_1)$ for all candidate points.

To find the absolute minimum/maximum of a function within a boundary, find candidate points on the boundary with Lagrange Multipliers, and candidate points inside the boundary where $\nabla f(\mathbf{x}_1) = 0$, and compare.

7 Integration