MATH1061
Advanced Multivariate Calculus & Ordinary Differential Equations

Michael Kasumagic, s4430266

Semester 2, 2024

Contents

Chapter 1

Week 1

1.1 Lecture 1

In this course, we will be looking at:

- functions of several variables and calculus
- vector calculus. Rates of change of vector valued functions and applications!
- differential equations
- MATLAB - Only 6 lab sessions. MATLAB will be incorprated into assignments.

An overview of the tools of applied mathematics

- Creating and studying models of phenomena in the world:
  - physics
  - chemistry
  - biology
  - ecology
  - economics
  - engineering.
- $\boxed{\text{natural world}} \xRightarrow{\text{simplification}} \boxed{\text{mathematical model}}$.
- $\boxed{\text{mathematical model}} \xRightarrow[\text{validation}]{\text{interpretation}} \boxed{\text{natural world}}$.
- Most importantly the $\boxed{\text{mathematical model}}$ offers predictive power.
- Modelling: identify key variables and processes.
- Formulation:
  - functions of several variables
  - ordinary differential equations (involving single variable rates of change)
  - WE WILL NOT TOUCH: partial differential equations (involving functions of several variables)
  - WE WILL NOT TOUCH: statistical models

Dimensional Analysis

Definition 1.1.1: Base Quantities

There exist base quantities (or dimensions) that provide units in terms of which the units of all other physical quantities can be expressed. Conventionally, these are: mass (M), length (L), time (T) (and temperature, electric current, amount of substance, luminous intensity).

Example 1.1.1 (A falling mass)
Suppose we conduct an experiment on the time, t, it takes an object of mass m, to fall a distance of x from rest in a vaccum (near the surface of the Earth).

In Australia we find that
$$x = 4.91t^2 \text{ (metres)},$$
Our friend in the USA finds that
$$x = 16.1t^2\text{(feet)}.$$
It would be correct to write $x = ct^2$, where c is a physical quantity, depending on units, $c = \frac{1}{2}g$.

Some quantities have dimensions as a product $M^aL^bT^c$, where $a, b, c \in \mathbb{Z}$. Let $[y]$ denote the dimensions of y and $[x]$ the dimensions of x. Then $[x, y] = [x][y]$.

Example 1.1.2 (Finding dimensions of physical quantities)
Velocity $\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)$:
$$\left[\frac{\mathrm{d}x}{\mathrm{d}t}\right] = [x][t]^{-1} = LT^{-1}$$
Acceleration $\left(\frac{\mathrm{d}x}{\mathrm{d}t^2}\right)$:
$$\left(\frac{\mathrm{d}x}{\mathrm{d}t^2}\right) = [x][t]^{-2} = LT^{-2}$$
Force $m\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)$
$$[F] = [m][t]^{-1}[x][t]^{-1} = MLT^{-2}$$

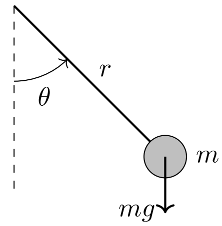
We call a quantity with dimensions $M^0L^0T^0$ **dimensionless**.
An equation that is true regardless of units is said to be **dimensionally homogeneous**. In such an equation, the dimensions of all terms must be the same.

Claim 1.1.1 Equations representing physical laws are dimensionally homogeneous.
To achieve this in our mathematical model we seek *all possible* dimensionless products among the variables. Such a collection is called **complete set**.

1.2 Lecture 2

A Simple Pendulum

Consider the simple pendulum, with mass m, length r released from angle of displacement θ, and acted upon by gravity g.



We want to find the relationship between the period, τ and θ, m, r and g. I.e. we want to find

$$\tau = f(r, m, \theta, g)$$

First, lets find the dimensions of the system.

$[\tau]$	$[r]$	$[m]$	$[\theta]$	$[g]$
T	L	M	1	LT^{-2}

Note:-

g is a dimensionless variable! In other words, its dimensions are $M^0L^0T^0$.

The product of these varaibles takes the form:

$$\tau^a r^b m^c \theta^d g^e$$

and its dimensions are

$$[\tau^a r^b m^c \theta^d g^e] = [\tau^a][r^b][m^c][\theta^d][g^e] = T^a \cdot L^b \cdot M^c \cdot 1^d \cdot L^e T^{-2e} = M^c L^{b+e} T^{a-2e}.$$

Since we desire to solve represent a physical law, by claim 1.1.1, we know we're looking for a dimensionally homogeneous system.

$$\begin{aligned} &\text{Let } M^c L^{b+e} T^{a-2e} = M^0 L^0 T^0 \\ &\text{Then} \quad \left.\begin{array}{rcl} a-2e &=& 0 \\ b+e &=& 0 \\ c &=& 0 \end{array}\right\} e = \frac{a}{2}, b = -\frac{a}{2}, d \text{ is free.} \end{aligned}$$

In other words, solve the linear system $A\underset{\sim}{x} = \underset{\sim}{0}$ (In other OTHER words, solve for the nullspace of $\mathcal{N}(A)$.)

$$\begin{aligned} \tau^a r^{-\frac{a}{2}} m^0 \theta^d g^{\frac{a}{2}} &= \left(\tau r^{-\frac{1}{2}} g^{\frac{1}{2}}\right)^a \theta^d \\ &= \left(\tau \sqrt{\frac{g}{r}}\right)^a \theta^d \end{aligned}$$

By setting (a, d) to $(1, 0)$, and $(0, 1)$, we obtain independent dimensionless products, Π_1 and Π_2.

$$\begin{aligned} \Pi_1 &= \tau\sqrt{\frac{g}{r}} \\ \Pi_2 &= \theta \end{aligned}$$

The Buckingham Π-Theorem

Theorem 1.2.1 Buckingham Π-Theorem

An equation is dimensionally homogeneous if and only if it can be written in the form

$$f(\Pi_1, \Pi_2, \ldots, \Pi_n) = 0,$$

where f is some function satisfying certain conditions (outside out scope) and $\{\Pi_1, \Pi_2, \ldots, \Pi_n\}$ is a complete set of dimensionless products corresponding to some mathematical model.

Note:-

Π_k is dimensionless, i.e. $\forall \Pi_k, [\Pi_k] = 1$

The set (Π_k) can be obtained by giving all solutions to a linear system of exponents for the model. We've found the complete set of dimensionless products for the pendulum problem, $\left\{\tau\sqrt{g/r}, \theta\right\}$. Applying the Buckingham Π-theorem now, our mathematical model is of the form:

$$f(\Pi_1, \Pi_2) = 0 \implies f(\tau\sqrt{g/r}, \theta) = 0$$

We further assume that from f we can deduce

$$\tau\sqrt{\frac{g}{r}} = h(\theta) \qquad \text{by Implict Function Theorem.}$$

We'll describe implicit function theorem in detail later.

Note:-

If $\Pi = \{\Pi_1, \Pi_2, \Pi_3\}$, then $\Pi_1 = h(\Pi_1, \Pi_2)$. More generally, if $\Pi = \{\Pi_k \mid k \in \mathbb{N}, k \leq n\}$, then $\Pi_1 = h\left(\Pi_2, \Pi_3, \ldots, \Pi_n\right)$.

Archimedes' Law

The famous "Eureka!" leaping from the bathtub one ;P

Archimedes' law applies to bodies immersed in a fluid. Consider a box of mass m, which displaces V fluid, with constant density ρ. Suppose your class mate claims that this phenomenon is described by the equation

$$m\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = mg - mVg$$

Let's verify this...

$$\begin{aligned}
[mVg] &= [m][V][g] \\
&= M \cdot L^3 \cdot LT^{-2} \\
&= ML^4T^{-2} \\
\left[m\frac{\mathrm{d}^2x}{\mathrm{d}t^2}\right] &= [m]\left[\frac{\mathrm{d}^2x}{\mathrm{d}t^2}\right] \\
&= MLT^{-2} \\
&\neq [mVg]
\end{aligned}$$

So we can conclude that this model is not dimensionally consistent.
Another classmate suggests the equation

$$m\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = mb - \rho Vg.$$

We'll similarlly analyse this like

$$\begin{aligned}
[\rho Vg] &= [\rho][V][g] \\
&= ML^{-3} \cdot L^3 \cdot LT^{-2} \\
&= ML^1T^{-2} \\
&= MLT^{-2} \\
\left[m\frac{\mathrm{d}^2x}{\mathrm{d}t^2}\right] &= [m]\left[\frac{\mathrm{d}^2x}{\mathrm{d}t^2}\right] \\
&= MLT^{-2} \\
&= [\rho Vg]
\end{aligned}$$

Which is dimensionally consistent!

Let's use Buckingham Π-theorem to establish the general form any correct model must take:

Consider the product $F^a\rho^b g^c V^d m^e$

$$\begin{aligned}
\left[F^a\rho^b g^c V^d m^e\right] &= (MLT^{-2})^a(ML^{-3})^b(LT^{-2})^c(L^3)^d(M)^e \\
&= M^aL^aT^{-2a}\cdot M^bL^{-3b}\cdot L^cT^{-2c}\cdot L^{3d}\cdot M^e \\
&= M^{a+b+e}L^{a-3b+c+3d}T^{-2a-2c} \\
\text{Let } M^0L^0T^0 &= M^{a+b+e}L^{a-3b+c+3d}T^{-2a-2c} \\
\Rightarrow \left.\begin{array}{rcl} a+b+e &=& 0 \\ a-3b+c+3d &=& 0 \\ -2a-2c &=& 0 \end{array}\right\} & \begin{array}{rcl} c &=& -a \\ b &=& -a-e \\ d &=& -a-e \end{array} \\
\text{So } F^a\rho^b g^c V^d m^e &= F^a\rho^{-a-e}g^{-a}V^{-a-e}m^e \\
&= F^a\rho^{-a}g^{-a}V^{-a}\rho^{-e}V^{-e}m^e \\
&= \left(F\rho^{-1}g^{-1}V^{-1}\right)^a\left(\rho^{-1}V^{-1}m\right)^e \\
&= \left(\frac{F}{\rho g V}\right)^a\left(\frac{m}{\rho V}\right)^e \\
\therefore \Pi &= \left\{\frac{F}{\rho g V}, \frac{m}{\rho V}\right\}.
\end{aligned}$$

Now that we've deduced Π, we know that any valid physcial law must take the form:

$$\frac{F}{\rho g V} = h\left(\frac{m}{\rho V}\right)$$

$$\Longrightarrow F = \rho g V h\left(\frac{m}{\rho V}\right)$$

> **Note:-**
>
> Generally, when proceeding with the linear algebra portion of this procedure, keep in mind the power of your desired dependent variable, and try to express all other powers in terms of it. For example, in the pendulum example, and in the Archimedes' law exmaple too, we expressed the other variables in terms of a, because this was the power of the desired dependent variables τ and F, respectively.

1.3 Lecture 3

Drag Force

Consider a sphere of radius r, moving through a viscous fluidd. We wish to model the drag force, F, dependent on the relevant variables η, the viscosity, v, the velocity, and r the radius.

Consider the product $F^a\eta^b v^c r^d$

$$\begin{aligned}
[F]^a[\eta]^b[v]^c[r]^d &= \left(MLT^{-2}\right)^a\left(ML^{-1}T^{-1}\right)^b\left(LT^{-1}\right)^c(L)^d \\
&= \left(M^aL^aT^{-2a}\right)\left(M^bL^{-b}T^{-b}\right)\left(L^cT^{-c}\right)\left(L^d\right) \\
&= M^{a+b}L^{a-b+c+d}T^{-2a-b-c} \\
\text{Let } M^0L^0T^0 &= M^{a+b}L^{a-b+c+d}T^{-2a-b-c} \\
\Rightarrow \left.\begin{array}{rcl} a+b &=& 0 \\ a-b+c+d &=& 0 \\ -2a-b-c &=& 0 \end{array}\right\} & \begin{array}{rcl} b &=& -a \\ c &=& -a \\ d &=& -a \end{array}
\end{aligned}$$

$$
\begin{aligned}
\text{So } F^a\eta^b v^c r^d &= F^a\eta^{-a}v^{-a}r^{-a} \\
&= \left(F\eta^{-1}v^{-1}r^{-1}\right)^a \\
&= \left(\frac{F}{\eta v r}\right)^a \\
\therefore \Pi &= \left\{\frac{F}{\eta v r}\right\}
\end{aligned}
$$

So from the Buckingham Π-theorem, we know that many valid physcial law must take the form:

$$
f\left(\frac{F}{\eta v r}\right) = 0
$$

$$
\implies \frac{F}{\eta v r} = k
$$

In other words, $F = k\eta v r$, where k is some dimensionless constant.

Note:-

This result, that drag force is proportional to velocity is known as Stokes' Law, and in fact, the constant $k = 6\pi$.

A Mixing Model

Initially, a tank of water with v_0 litres has m_0 grams of salt dissolved in it. Brine with n grams of salt per litre runs into the tank at a rate of x litres per minute. The contents are constantly stirred (so you can assume that the concentration of salt is always uniform throughout the tank) and water runs out of the tank at the rate of y litres per minute. Let $s(t)$ denote the amount of salt in the tank at time t (measured in minutes). Determine the equation which governs the net rate of change of salt, checking that all terms are dimensionally consistent.

Let $c(t)$ denote the density of salt at time t.
Let $v(t)$ denote the volume of water in the tank at time t.

$$
\begin{aligned}
\text{We have } c(t) &= \frac{s(t)}{v(t)} \\
\text{And } v(t) &= v_0 + (x-y)t \\
\implies c(t) &= \frac{s(t)}{v_0 + (x-y)t}.
\end{aligned}
$$

Salt enters at a rate of nx grams/minute and leaves at a rate of $yc(t)$ per minute.

$$
yc(t) = \frac{ys(t)}{v_0 + (x-y)t}.
$$

Therefore, the net rate of change is

$$
\frac{\mathrm{d}s}{\mathrm{d}t} = nx - \frac{ys}{v_0 + (x-y)t}, s(0) = m_0.
$$

Now lets check the model for dimensional homogeneity!

$$\begin{aligned}
\left[\frac{\mathrm{d}s}{\mathrm{d}t}\right] &= MT^{-1} \\
[nx] &= ML^{-3}\cdot L^3T^{-1} \\
&= MT^{-1} \\
[v_0] &= L^3 \\
[(x-y)t] &= L^3T^{-1}T \\
&= L^3 \\
\Rightarrow \left[\frac{ys}{v_0+(x-y)t}\right] &= \frac{ML^3T^{-1}}{L^3} \\
&= \frac{ML^3T^{-1}}{L^3} \\
&= MT^{-1} \\
\therefore \left[\frac{\mathrm{d}s}{\mathrm{d}t}\right] = [nx] &= \left[\frac{ys}{v_0+(x-y)t}\right]
\end{aligned}$$

So, the system is dimensionally homogeneous.

Scaling

Question, can we scale experiments in a laboratory to ensure that the observed effects are consistent?
We can use dimensionless variables, and try to preserve their values. Some examples of dimensionless variables in fluid dynamics include

$$\begin{aligned}
\text{Reynold's Number,} \qquad Re &= \frac{\rho l v}{\eta} \\
\text{Froude's Number,} \qquad Fr &= \frac{v}{\sqrt{gl}} \\
\text{Mach Number,} \qquad M &= \frac{v}{c}
\end{aligned}$$

where, ρ is the fluid denisty, l is the length of an object, v is velocty, η is the viscosity of the fluid, g is gravitational acceleration, and c is the speed of sound.

A Ship Model

Suppose our model involves Fr, which we seek to keep fixed.
The true boat has hull length $l = 20\text{m}$, at speed $v = 10\text{ms}^{-1}$
We can model this with a boat of hull length $l^* = 0.2m$, i.e. $l^* = l/100$.
What is v^*?

$$\begin{aligned}
Fr &= \frac{v}{\sqrt{gl}} = \frac{v^*}{\sqrt{gl^*}} \\
\therefore v^* &= v\sqrt{\frac{gl^*}{gl}} \\
&= v\sqrt{\frac{l^*}{l}} \\
&= 10\sqrt{\frac{0.2}{20}} \\
\therefore v^* &= 1\text{ms}^{-1}
\end{aligned}$$

Chapter 2

Week 2

2.1 Lecture 4

Multivariate Limits

Review of one-variable case

Let $f : \mathrm{D} \to \mathbb{R}$ be a function with domain D, an open subset of $\mathbb{R}$. For $a \in \mathrm{D}$ we say that the limit $\lim\limits_{x\to a} f(x)$ exists if and only if (i) the left-sided limit exist, (ii) the right-sided limit exists, and (iii) these limits equal each other, namely

$$\lim_{x\to a^-} f(x) = \lim_{x\to a^+} f(x).$$

Furthermore, if the limit exists, and is equal to the value of the function at f, namely,

$$\lim_{x\to a^-} f(x) = \lim_{x\to a^+} f(x) = f(a),$$

we say that f is continuous at a. The precise defintion of the one variable limit its

$$\lim_{x\to a} f(x) = L \iff \forall \varepsilon > 0, \exists \delta > 0 : 0 < |x - a| < \delta \implies |f(x) - L| < \varepsilon.$$

The two-variable case

When f is multivariate, finding the limit is more subtle. There are more then two ways to approach a given point. Consider

$$f(x, y) = \frac{x^2}{x^2 + y^2},$$

with domain $\mathbb{R}^2 \setminus (0, 0)$. We could approach this limit along the line $y = 0$. If $x \neq 0$,

$$f(x, 0) = \frac{x^2}{x^2 + 0} = 1,$$
$$\text{Then } \lim_{x\to 0} f(x, 0) = \lim_{x\to 0} 1 = 1$$

Suppose next we approach the limit along the line $x = 0$, if $y \neq 0$,

$$f(0, y) = \frac{0^2}{0^2 + y^2} = 0,$$
$$\text{Then } \lim_{y\to 0} f(0, y) = \lim_{y\to 0} 0 = 0.$$

Since we've approached the same point with two different paths, but found different limits,

$$\lim_{x\to 0} f(x, 0) \neq \lim_{y\to 0} f(0, y),$$

we can conclude that

$$\lim_{(x,y)\to(0,0)} f(x,y) \text{ does no exist.}$$

For the limit $\lim_{(x,y)\to(a,b)} f(x,y)$ to exist, it is necessary that *every* path in D approaching (a,b) gives the same limiting value. This presents us with a simple test for determining if a limit does not exist.

Example 2.1.1 (Test for showing no limit exists)

$$\text{If} \begin{cases} f(x,y) \to L_1 & \text{as } (x,y) \to (a,b) \text{ along the path } C_1 \in \mathrm{D} \\ f(x,y) \to L_2 & \text{as } (x,y) \to (a,b) \text{ along the path } C_2 \in \mathrm{D} \end{cases}$$

such that $L_1 \neq L_2$ then the limit $\lim_{(x,y)\to(0,0)} f(x)$ does not exist.

Note:-

The above notation is somewhat deficient and perhaps one shoule write

$$\lim_{(x,y)\to_{\mathrm{D}}(a,b)} f(x,y)$$

to indicate that only paths in D terminating at (a,b) (which itself may or may not be in D) are considered.

Question 1

Let $\mathrm{D} = \mathbb{R}^2 \setminus (0,0)$ and $f : \mathrm{D} \to \mathbb{R}$ be given by $f(x,y)\dfrac{x^2-y^2}{x^2+y^2}$. Show that the limit $\lim_{(x,y)\to(0,0)} f(x)$ does not exist.

Solution: Consider the path $C_1 \in \mathrm{D}, y = 0$
If $x \neq 0$,

$$f(x,0) = \frac{x^2}{x^2} = 1.$$

$$\lim_{x\to 0} f(x,0) = \lim_{x\to 0} 1 = 1.$$

Consider the path $C_2 \in \mathrm{D}, x = 0$
If $y \neq 0$,

$$f(0,y) = \frac{-y^2}{y^2} = -1.$$

$$\lim_{y\to 0} f(0,y) = \lim_{y\to 0} -1 = -1 \neq \lim_{x\to 0} f(x,0).$$

Therefore, the limit $\lim_{(x,y)\to(0,0)} f(x,y)$ does not exist.

Question 2

With the same D as the previous question, consider $f(x,y) = \dfrac{xy}{x^2+y^2}$. Show that the limit $\lim_{(x,y)\to(0,0)} f(x,y)$ does not exist.

Solution: Consider the path $C_1 \in \mathrm{D}, y = 0$
If $x \neq 0$,

$$f(x,0) = 0 \implies \lim_{x\to 0} f(x,0) = 0$$

Consider the path $c_2 \in \mathrm{D}, x = 0$
If $y \neq 0$,

$$f(0,y) = 0 \implies \lim_{y\to 0} f(0,y) = 0$$

Huh. . . I really thought that would work. . .
Well, ok, Let's consider the path $C_3 \in \mathrm{D}, x = y$
If $x \neq 0$,

$$f(x,x) = \frac{x^2}{x^2+x^2} = \frac{1}{2} \implies \lim_{x\to 0} f(x,x) = \frac{1}{2}.$$

Since this limit is different from the other two, we can conclude that the limit $\lim\limits_{(x,y)\to(0,0)} f(x,y)$ does not exist.

Note:-

We can look at finitely many paths, and easily show that a limit doesn't exisit, if two paths terminating at the same point have different limiting values. But, there are infinitely many paths in $\mathbb{R}^2$ which terminate at the point (a,b). This raises the question, how can we prove a limit exists?

Consider the example

$$\lim_{(x,y)\to(0,0)} f(x,y) = 0, \text{ where } f(x,y) = \frac{3x^2y}{x^2+y^2} \text{ and } \mathrm{D} = \mathbb{R} \setminus (0,0).$$

We can argue that the above limit is correct and true by utilising a change of variables, namely

$$x = r\cos\theta \qquad y = r\sin\theta$$

Solution:

$$\begin{aligned}
f(x,y) &= f(r\cos\theta, r\sin\theta) \\
&= \frac{3r^2\cos^2\theta r\sin\theta}{r^2\cos^2\theta + r^2\sin^2\theta} \\
&= \frac{3r^3\cos^2\theta\sin\theta}{r^2\left(\cos^2\theta+\sin^2\theta\right)} \\
&= 3r\cos^2\theta\sin\theta\,|\cos\theta| \quad \leq 1 \\
|\sin\theta| &\leq 1 \\
\implies \left|\cos^2\theta\sin\theta\right| &\leq 1 \\
\implies \left|3r\cos^2\theta\sin\theta\right| &\leq 3r \\
\implies |f(x,y)| &\leq 3r \\
\therefore \text{ as } r\to 0, f(x,y) &\to 0
\end{aligned}$$

No matter what path of form (r,θ) we take, as that path approaches the point (a,b), the limit of $f(x,y)$ approaches 0. So we can argue that $\lim\limits_{(x,y)\to(0,0)} f(x,y)$ exists and is equal to 0. With everything we've learned now, let's finally, formally, define the limit of a multivariate function.

Definition 2.1.1: The Limit of a Multivariate Function

Let f be a function of two variables, whose domain D includes points arbitrarily close to (a,b). Then we say that the limit of $f(x,y)$ as (x,y) approaches (a,b) is L, written

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

if

$$\forall\varepsilon > 0, \exists\delta > 0 : 0 < \sqrt{(x-a)^2+(y-b)^2} < \delta \implies |f(x,y) - L| < \varepsilon.$$

2.2 Lecture 5

Multivariate Continuity

Definition 2.2.1: Multivariate Continuity

Given a function $f : \mathrm{D} \to \mathbb{R}$, where D is an open subset of $\mathbb{R}^2$. Let $(a, b) \in \mathrm{D}$. Then $f(x, y)$ is continuous at (a, b) If

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

ie, the limit $(x, y) \to (a, b)$ of $f(x, y)$ exists and is equal to $f(a, b)$.

Equivalently, f is continuous if

$$\forall \varepsilon > 0, \exists \delta > 0 : (x, y) \in \mathrm{D} \wedge \sqrt{(x-a)^2 + (y-b)^2} < \delta \implies |f(x, y)| < \varepsilon.$$

Note:-

An open subset is a subset which does not have boundary points. As opposed to a closed subset which has boundary points. More formally, "open" = for every point $p \in \mathrm{D}$, there is a disc with centre p that lies entirely in D. So, if you took p on the bounary, you can never find a disc which is entirely inside D, you'll always have a little bit peaking out of it.

So, is

$$f(x, y) = \frac{x^2 - y^2}{x^2 + y^2}, \mathrm{D} = \mathbb{R}^2 \setminus (0, 0)$$

continuous?
Well, even though we already found that this limit $(x, y) \to (0, 0)$ does not exist, we note that $(0, 0) \notin \mathrm{D}$. Therefore, it is continuous.
If we amend the domain to include $(0, 0)$, the function is no longer continous, because the limit $(x, y) \to (0, 0)$ does not exist.

Limit Laws

Suppose $\lim\limits_{(x,y)\to(a,b)} f(x, y)$ and $\lim\limits_{(x,y)\to(a,b)} g(x, y)$ exist.
Below, for ease of notation, lim denotes $\lim\limits_{(x,y)\to(a,b)}$.

1. $\lim(f(x, y) \pm g(x, y)) = \lim f(x, y) \pm \lim g(x, y)$.
2. $\lim cf(x, y) = c \lim f(x, y)$, where c is some constant.
3. $\lim(f(x, y) \cdot g(x, y)) = \lim f(x, y) \cdot \lim g(x, y)$.
4. $\lim \dfrac{f(x, y)}{g(x, y)} = \dfrac{\lim f(x, y)}{\lim g(x, y)}$, if $\lim g(x, y) \neq 0$.

2.3 Lecture 6

Partial Derivatives

Slope in the x-direction

Consider the surface $z = f(x, y) = 1 - x^2 - y^2$, and the point $P = (1, -1, -1)$. Holding y constant, we can we can find a cross section of the surface at that point.

$$\text{Let } g(x) = f(x, -1) = 1 - x^2 - 1 = -x^2$$

We can find the slope tangent of the surface z at the point P, in the x-direction, by differentiating $g(x)$,

$$g'(x) = -2x \implies g'(1) = -2$$

Slope in the y-direction

We can follow a similar procedure by holding x constant, and generating the y-direction cross section of the surface z

$$\text{Let } h(y) = f(1, y) = 1 - 1 - y^2 = -y^2$$

Now, we can evaluate the y-direction slope of the tanget of the surface z at the point P by taking the derivative of $h(x)$, namely

$$h'(x) = -2y \implies h'(-1) = 2.$$

Another example

Given $f(x, y) = xy^3 + x^2$, find $f_x(1, 2)$ and $f_y(1, 2)$.

$$\begin{aligned}
f(x, y) &= xy^3 + x^2 \\
\therefore f_x(x, y) &= y^3 + 2x \\
\implies f_x(1, 2) &= 8 + 2 = 10 \\
\therefore f_y(x, y) &= 3xy^2 + 0 = 3xy^2 \\
\implies f_y(1, 2) &= 3 \cdot 1 \cdot 4 = 12
\end{aligned}$$

Nothing scary :)

Partial derivative for $f(x, y, z)$

Consider the volume of a box, $V(x, y, z) = xyz$. If x changes by an an amount, say Δx, the volume will change by some amount, ΔV. One can visualise the change in volume, by imagine a box being concatonated to the original box, with width Δx, height y, and depth z. Therefore $\Delta V = yz\Delta x$. This leads us to see that

$$\frac{\Delta V}{\Delta x} = yz.$$

Now, letting $\Delta x \to 0$, we can find $\dfrac{\partial V}{\partial x} = yz$.
With partial derivatives, only one of the independent variables is allowed to change; all other variables are held constant.

$$\frac{\partial V}{\partial x} = \lim_{\Delta x \to 0} \frac{V(x + \Delta x, y, z) - V(x, y, z)}{\Delta x}$$

Higher order derivatives

$$\begin{aligned}
f_{xx} &= \frac{\partial^2 f}{\partial x^2} \\
f_{yy} &= \frac{\partial^2 f}{\partial y^2}
\end{aligned}
\qquad\qquad
\begin{aligned}
f_{xy} &= \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) \\
f_{yx} &= \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)
\end{aligned}$$

Theorem 2.3.1 Clairaut's Theorem

Suppose f is defined on a disk D that contains the point (a, b). If the functions f_{xy} and f_{yx} are both continuous, then $f_{xy}(a, b) = f_{yx}(a, b)$.

By extension, if f_{xy} and f_{yx} are both continuous everywhere, then $f_{xy} = f_{yx}$.

Question 3

Find all the first and second order derivatives of $f(x, y) = x \sin y + y \cos x$ and show that $f_{xy} = f_{yx}$.

Solution:

$$\begin{aligned}
f_x(x,y) &= \sin y - y\sin x \\
f_y(x,y) &= x\cos y + \cos x \\
f_{xx}(x,y) &= -y\cos x \\
f_{yy}(x,y) &= -x\sin y \\
f_{xy}(x,y) &= \frac{\partial f_x}{\partial y} \\
&= \cos y - \sin x \\
f_{yx}(x,y) &= \frac{\partial f_y}{\partial x} \\
&= \cos y - \sin x \\
\therefore f_{xy} &= f_{yx}
\end{aligned}$$

An unexpected feature of multivariate calculus is the possiblity that the for some point P, the partial derivatives of a function may be well defined, and yet, the function is not continuous at P. This is opposed to the one dimensional case, where we (naïvely), assert that continuity $\implies$ differentiability. Consider

$$f(x,y) = \begin{cases} \dfrac{xy}{x^2+y^2} & \text{for } (x,y) \neq (0,0) \\ 0 & \text{for } (x,y) = (0,0) \end{cases}$$

The derivatives at $(0,0)$ are well defined:

$$\frac{\partial f}{\partial x} = \frac{y^3 - x^2 y}{(x^2+y^2)^2} \implies \frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0} \frac{f(h,0) - f(0,0)}{h} = \lim_{h\to 0} \frac{0-0}{h} = 0$$

($\dfrac{\partial f}{\partial y}$ is defined and equal to $\dfrac{\partial f}{\partial x}$, because the function is symmetric, note that x and y can be simply interchanged in the function definition.)

Here, we have a function, whose derivatives are defined at a point P, and yet, is discontinuous at P.

When dealing with multivariate functions, existent derivatives are insufficient for determining differentiability.

Chapter 3

Week 3

3.1 Lecture 7

The Tangent Plane

Recall the one dimensional case, for a function $y = f(x)$, and a point $(a, f(a))$, the tangent line is given by

$$y_\top = f(a) + f'(a)(x - a)$$

and can be derived by considering

$$\begin{gathered} y_\top = mx + c \\ x = a \qquad y = f(a) \qquad m = f'(a) \\ \therefore f(a) = f'(a)a + c \\ \therefore c = f(a) - f'(a)a \\ \therefore y_\top = f'(a)x + f(a) - f'(a)a \\ \therefore y_\top = f(a) + f'(a)(x - a) \end{gathered}$$

Let's derive the vector equation for of a tangent line, consider the vector

$$\underset{\sim}{r}(x) = (x, f(x)),$$

and another vector which is some Δx away from $\underset{\sim}{r}$,

$$\underset{\sim}{r}(x + \Delta x) = (x + \Delta x, f(x + \Delta x)).$$

Now consider $\underset{\sim}{r}(x+\Delta x) - \underset{\sim}{r}(x)/\Delta x$. As we allow Δx to go to 0, $\underset{\sim}{r}(x+\Delta x) - \underset{\sim}{r}(x)$ is going to approach the tangent line, of $f(x)$ at the point $(x, f(x))$.

$$\lim_{\Delta x \to 0} \frac{\underset{\sim}{r}(x + \Delta x) - \underset{\sim}{r}(x)}{\Delta x} = \lim_{\Delta x \to 0} \begin{pmatrix} \dfrac{x + \Delta x - x}{\Delta x} \\ \dfrac{f(x + \Delta x) - f(x)}{\Delta x} \end{pmatrix} = \begin{pmatrix} 1 \\ f'(x) \end{pmatrix}$$

Therefore, the tangent vector, $\underset{\sim}{r}_\top$ is $(1, f(x))$. Points on lying on the tangent, are

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ f'(x_0) \end{pmatrix},$$

where (x, y) is the point on the line, (x_0, y_0) are the inital conditions, namely, the point lying on $f(x)$, and λ is some real number. We went to all this effort, because the tangential plane of a 3D surface, is analogous to the tangential lin on a 2D curve.

Equation for the tagent plane

On a surface $z = f(x,y)$, at the point $(a, b, f(a,b))$, the tangent plane is given by

$$\boxed{z_\top = f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b)}$$

This is analogous to the 2D case. We follow this by defining the plane as a vector

$$\boxed{\underset{\sim}{r}_\top = (r_1, r_2, r_3) = (a, b, f(a,b)) + \lambda(1, 0, f_x(a,b)) + \mu(0, 1, f_y(a,b)),}$$

where $\lambda = x - a$ and $\mu = y - b$.

I like to think of it like this:

$$\underset{\sim}{r}_x = (1, 0, f_x(a,b)) \qquad \underset{\sim}{r}_y = (1, 0, f_y(a,b))$$

are linearly independent vectors. The span $\{\underset{\sim}{r}_x, \underset{\sim}{r}_y\}$ *is* the tangent plane. $\lambda, \mu \in \mathbb{R}$ lets us take any linear combination of these planar basis vectors. This new vector, therefore, will lie on the tangent plane.

Example 3.1.1 (Tangent Plane)

Find the tangent plane to the surface $z = 1 - x^2 - y^2$ at the point $P = (1, -1, -1)$.

$$\begin{aligned} f(x,y) &= 1 - x^2 - y^2 \implies f(1,-1) = -1 \\ f_x(x,y) &= -2x \implies f_x(1,-1) = -2 \\ f_y(x,y) &= -2y \implies f_y(1,-1) = 2 \\ \therefore z_\top &= -1 - 2(x-1) + 2(y+1) \\ &= -1 - 2x + 2 + 2y + 2 \\ &= 3 - 2x + 2y \end{aligned}$$

Alternatively...

$$\underset{\sim}{r}_\top = \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 + x - 1 + 0 \\ -1 + 0 + y - 1 \\ -1 - 2x + 2 + 2y + 2 \end{pmatrix} = \begin{pmatrix} x \\ y \\ 3 - 2x + 2y \end{pmatrix}$$

Example 3.1.2 (Tangent Plane)

What is the plane tangent to the surface $f(x,y) = 4 - x^2 + 4x - y^2$, at $(1,1)$?

$$\begin{aligned} f(1,1) &= 4 - 1 + 4 - 1 = 6 \\ f_x(x,y) &= -2x + 4 \implies f_x(1,1) = 2 \\ f_y(x,y) &= -2y \implies f_y(1,1) = -2 \\ \therefore z_\top(x,y) &= 6 + 2(x-1) - 2(y-1) \\ &= 6 + 2x - 2 - 2y + 2 \\ &= 6 + 2x - 2y \end{aligned}$$

Alternatively...

$$\underset{\sim}{r}_\top = \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 6 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix} = \begin{pmatrix} 1 + x - 1 + 0 \\ 1 + 0 + y - 1 \\ 6 + 2x - 2 - 2y + 2 \end{pmatrix} = \begin{pmatrix} x \\ y \\ 6 + 2x - 2y \end{pmatrix}$$

Smoothness

A surface, $z = f(x,y)$ is smooth at a point (a,b) if f, f_x and f_y are all continuous at (a,b). Informally, as you zoom more and more into (a,b), it looks more and more like a flat plane.

One way to observe this is looking at contours. As you zoom into a smooth function, the contours will straighten out. This means that, close to (a, b), the surface can be approximated with a plane; in fact, approximated with the tangent plane.
Examples of non-smooth functions include things like Brownian motion. No matter how much you zoom into Brownian motion, it always looks rough. In fact, Brownian motion is fractal.

3.2 Lecture 8

We first note, that the first order approximation of a single variable function, $f(x)$, around some point $(a, f(x))$, is a straight line tanget to that point.

$$f(a) \approx f(a) + f'(a)(x-a)$$

Example 3.2.1

Given $f(x) = \exp(x)$, estimate $e^{0.1}$.

$$\begin{gathered} f(x) = \exp(x) \implies f'(x) = \exp(x) \\ f(0) = \exp(0) = 1 \implies f'(0) = \exp(0) = 1 \\ \therefore e^{0.1} \approx f(0) + f'(0)(x-0) \\ e^{0.1} \approx 1 + x \\ \therefore e^{0.1} \approx 1.1 \end{gathered}$$

Linear Approximations for $f(x, y)$

The corresponding linear, or first order approximation for a function f, of two variables, near a known point (a, b), is the tangent plane centred on that point, given that f is a smooth. Therefore, the approximation of f at (a, b) is

$$f(x,y) \approx f(a,b) + f_x(x,y)(x-a) + f_y(x,y)(y-b)$$

Example 3.2.2

The temperature in a region is given by $T(x) = 100 - x^2 - y^2$. Find the linear approximation to $T(x, y)$ near $(0, 5)$.

$$\begin{gathered} T(x,y) = 100 - x^2 - y^2 \implies T(0,5) = 75 \\ T_x(x,y) = -2x \implies T_x(0,5) = 0 \\ T_y(x,y) = -2y \implies T_y(0,5) = -10 \\ \therefore T(x,y) \approx T(0,5) + T_x(0,5)(x-0) + T_y(0,5)(y-5) \\ T(x,y) \approx 75 + 0 + (-10)(y-5) \\ T(x,y) \approx 125 - 10y \end{gathered}$$

Example 3.2.3

Find the tangent plane to $z = \exp(-x^2)\sin y$ at $(1, \pi/2)$. Use it to approximate $e^{-(0.9)^2}\sin(1.5)$.

Note: 0.9 is close to 1, and 1.5 is close to $\pi/2$.

$$\begin{gathered} f(x,y) = \exp(-x^2)\sin y \implies f(1,\pi/2) = \exp(-1) \\ f_x(x,y) = -2x\exp(-x^2)\sin y \implies f_x(1,\pi/2) = -2\exp(-1) \\ f_y(x,y) = \exp(-x^2)\cos y \implies f_y(1,\pi/2) = 0 \\ z \approx f(1,\pi/2) + f_x(1,\pi/2)(x-1) + f_y(1,\pi/2)(y-\pi/2) \\ \implies z \approx \exp(-1)(3-2x) \\ \therefore e^{-0.9^2}\sin(1.5) \approx e^{-1}(3 - 2\cdot 0.9) \approx 0.4415 \end{gathered}$$

Differentials

If we reduce small changes Δx and Δy down to an infinitesimal changes, dx, dy, the linear approximation as

$$\mathrm{d}z = \mathrm{d}f = f_x(a,b)dx + f_y(a,b)dy$$

This infinitesimal change dz is called the total differential, and is definied by this equation (note the "=", instead of an "≈").

Consider the work done by a, $W = PV$, where P is pressure and V is volume.

$$W(P,V) = PV$$
$$W(P+\Delta P, V+\Delta V) = (P+\Delta P)(V+\Delta V) = PV + P\Delta V + V\Delta V + \Delta P\Delta V$$
$$\therefore \Delta W = P\Delta V + V\Delta V + \Delta P\Delta V$$
$$\text{Let } \mathrm{d}P = \Delta P, \text{ as } \Delta P \to 0, \mathrm{d}V = \Delta V, \text{ as } \Delta V \to 0$$
$$\therefore \mathrm{d}W = P\mathrm{d}V + V\mathrm{d}V (= W_P(P,V)\mathrm{d}P + W_V(P,V)\mathrm{d}V)\text{Note: } \Delta P\Delta V = 0$$

Estimating Error

If the error in x and y is, at mostE_x, and E_y, respectively, then a resaonable estimate for the worst-case error in the linear approximation of f at (a,b) is

$$|E| \approx |f_x(a,b)E_x| + |f_y(a,b)E_y|$$

Example 3.2.4

Suppose when making a metal barrel of base radius 1m, and height 2m, we allow an error of 5% in radius and height. Estimate the worst-case error in volume.

$$V(r,h) = \pi r h^2 \implies V_r(r,h) = 2\pi r h, V_h(r,h) = \pi r^2$$
$$\therefore V(1,2) = 2\pi, V_r(1,2) = 4\pi, V_h(1,2) = \pi$$
$$E_r = \frac{5}{100}r = \frac{5}{100}(1) = 0.05$$
$$E_h = \frac{5}{100}h = \frac{5}{100}(2) = 0.1$$
$$\implies |E| = |V_r(1,2)E_r| + |V_h(1,2)E_h| = \left|4\pi \cdot \frac{5}{100}\right| + \left|\pi \cdot \frac{10}{100}\right| = \frac{3\pi}{10}$$
$$\implies \text{pct\% error: } \frac{|E|}{V(1,2)} \cdot 100 = \frac{3\pi}{10 \cdot 2\pi} \cdot 100 = 15\%$$

Therefore, an estimate for the worst-case error in volume is 15%

Note:-

$\mathrm{d}V = V_r(1,2)\mathrm{d}r + v_h(1,2)\mathrm{d}h$, and $\mathrm{d}r = \Delta r = E_r$, $\mathrm{d}h = \Delta h = E_h$.
$|E| = |V_r(1,2)E_r| + |V_h(1,2)E_h$.
So the error in V is just the differential of V.

3.3 Lecture 9

Differentiability

Recall that in single variable calculus, a function f is differentiable at x_0 if there exists a number $f'(x_0)$ such that

$$f'(x_0) = \lim_{h\to 0} \frac{f(x_0+h) - f(x_0)}{h}$$

which can be rewritten as

$$0 = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0) - f'(x_0)h}{h}$$

Informally, we say f is differentiable at x_0 if, after enough zooming, we see a stright line $y(x) = f(x_0) + f'(x_0)(x - x_0)$, such that the difference $f(x_0 + h) - y(x_0 + h)$ goes to 0, faster then linearly in the limit as h goes to 0.

Extending to the multivariate case, let $V \subseteq \mathbb{R}^n$. We define a function $f : V \to \mathbb{R}$ to be differentiable at $\utilde{a}$ if there exists $\nabla f(\utilde{a}) \in \mathbb{R}^n$ such that

$$\lim_{\utilde{h} \to \utilde{0}} \frac{f(\utilde{a} + \utilde{h}) - f(\utilde{a}) - \nabla f(\utilde{a}) \cdot \utilde{h}}{||\utilde{h}||} = 0$$

Informally, we say that f is differentiable at $\utilde{a}$, if there is a plane $z(\utilde{a} + \utilde{h}) = f(\utilde{a}) + \nabla f(\utilde{a}) \cdot \utilde{h}$ such that $f(\utilde{a} + \utilde{h}) - z(\utilde{a} + \utilde{h})$ goes to 0 faster then linearly as $\utilde{h}$ goes to $\utilde{0}$.

As we shall see, if it is possible for a tangent plane, z to exisit at the point $\utilde{a}$, but for the function $f : V \to \mathbb{R}$ to not be differentiable at that point.

The lecturer proceeded to give us a very nice geometric interpretation of this, but I'm not going to write everything he handwrote.

Theorem 3.3.1

If a function is differentiable at a point, then it is continuous at that point. The contrapositive of this is, if a function is not continuous at a point, then it is not differentiable at that point.

Theorem 3.3.2

If all first-order partial derivatives of a function exist and are continuous at a point, then the function is differentiable at that point.

Chapter 4

Week 4

4.1 Lecture 10

Gradients and Directional Derivatives

The partial derivative f_x corresponds to the slope of $f(x, y)$ in the x-direction. We now turn out attention to the question of slopes in arbitrary directions, such as $\hat{\imath} + 2\hat{\jmath}$ or $-\hat{\jmath}$.

Let $\underset{\sim}{u} = (u_1, u_2, \ldots, u_n) \in \mathbb{R}^n : ||\underset{\sim}{u}|| = 1$. The directional derivative of f at the point $\underset{\sim}{a}$ in the direction $\underset{\sim}{u}$ be defined as:

$$f_{\underset{\sim}{u}}(\underset{\sim}{a}) \equiv \lim_{t\to 0} \frac{f(\underset{\sim}{a} + t\underset{\sim}{u}) - f(\underset{\sim}{a})}{t}.$$

It is also known as the slope of f at the point $\underset{\sim}{a}$ in the direction of $\underset{\sim}{u}$.

In practice, this is quite cumbersome though. Instead, if f is differentiable, we can observe that

$$\lim_{h\to 0} \frac{f(\underset{\sim}{a} + \underset{\sim}{h}) - f(\underset{\sim}{a}) - \nabla f(\underset{\sim}{a}) \cdot \underset{\sim}{h}}{||\underset{\sim}{h}||} = \lim_{t\to 0} \frac{f(\underset{\sim}{a} + t\underset{\sim}{u}) - f(\underset{\sim}{a}) - \nabla f(\underset{\sim}{a}) \cdot t\underset{\sim}{a}}{t} = 0$$

We can then write

$$\begin{aligned}
f_{\underset{\sim}{u}} &= \lim_{t\to 0} \frac{f(\underset{\sim}{a} + t\underset{\sim}{u}) - f(\underset{\sim}{a})}{t} + \nabla f(\underset{\sim}{a}) \cdot \underset{\sim}{u} - \nabla f(\underset{\sim}{a}) \cdot \underset{\sim}{u} \\
&= \lim_{t\to 0} \frac{f(\underset{\sim}{a} + t\underset{\sim}{u}) - f(\underset{\sim}{a}) - \nabla f(\underset{\sim}{a}) \cdot t\underset{\sim}{u}}{t} + \nabla f(\underset{\sim}{a}) \cdot \underset{\sim}{u} \\
&= \lim_{t\to 0} \frac{f(\underset{\sim}{a} + t\underset{\sim}{u}) - f(\underset{\sim}{a}) - \nabla f(\underset{\sim}{a}) \cdot t\underset{\sim}{u}}{t} + \lim_{t\to 0} \nabla f(\underset{\sim}{a}) \cdot \underset{\sim}{u} \\
&= 0 + \lim_{t\to 0} \nabla f(\underset{\sim}{a}) \cdot \underset{\sim}{u} \\
&\boxed{f_{\underset{\sim}{u}} = \nabla f(\underset{\sim}{a}) \cdot \underset{\sim}{u}}
\end{aligned}$$

We can drop our assumption that $\underset{\sim}{u}$ is a unit vector, and instead write

$$f_{\underset{\sim}{u}} = \nabla f(\underset{\sim}{a}) \cdot \frac{\underset{\sim}{u}}{||\underset{\sim}{u}||}$$

The gradient vector ∇f

The gradient vector of f, or the gradient of f is the simply the vector whose components are partial derivatives.

If we let $\{\underset{\sim}{b}_1, \underset{\sim}{b}_2, \dots, \underset{\sim}{b}_n\}$ form the canonical basis of $\mathbb{R}^n$ then,

$$\nabla \in \mathbb{R}^n : \nabla = \sum_{i=1}^{n} \underset{\sim}{b}_i \frac{\partial}{\partial x_i},$$

where x_i is the ith Cartesian dimension, or ith component of a vector in $\mathbb{R}^n$.

For example in two dimensions

$$\operatorname{grad} f = \nabla f = (f_x, f_y) = f_x \hat{\imath} + f_y \hat{\jmath}.$$

But this totally generalises into higher dimensions, for example, for a four variable function, $f(x, y, z, w)$,

$$\operatorname{grad} f = \nabla f = (f_x, f_y, f_z, f_w) = f_x \hat{\imath} + f_y \hat{\jmath} + f_z \hat{k} + f_w \hat{l}$$

Example 4.1.1

If $f(x, y) = x^2 - 3y^2 + 6y$, find the slope at $(1, 0)$ in the direction $\hat{\imath} - 4\hat{\jmath}$.

$$\nabla f = \begin{pmatrix} f_x \\ f_y \end{pmatrix} = \begin{pmatrix} 2x \\ 6 - 6y \end{pmatrix} \implies \nabla f(1, 0) = \begin{pmatrix} 2 \\ 6 \end{pmatrix}$$

$$\underset{\sim}{u} = \begin{pmatrix} 1 \\ -4 \end{pmatrix} \implies \underset{\sim}{\hat{u}} = \frac{1}{\sqrt{1^2 + 4^2}} \begin{pmatrix} 1 \\ -4 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{17} \\ -4/\sqrt{17} \end{pmatrix}$$

$$\implies f_{(1,-4)}(1, 0) = \nabla f(1, 0) \cdot \underset{\sim}{\hat{u}} = \begin{pmatrix} 1 \\ -4 \end{pmatrix} \cdot \begin{pmatrix} 1/\sqrt{17} \\ -4/\sqrt{17} \end{pmatrix} = \frac{2}{\sqrt{17}} - \frac{24}{\sqrt{17}} = \frac{-22}{\sqrt{17}}$$

Example 4.1.2

Find $f_{(1,-1)}(0, 1)$ for $f(x, y) = x - x^2y^2 + y$

$$\nabla f = \begin{pmatrix} f_x \\ f_y \end{pmatrix} = \begin{pmatrix} 1 - 2xy^2 \\ 1 - 2x^2y \end{pmatrix} \implies \nabla f(0, 1) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\underset{\sim}{u} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \implies \underset{\sim}{\hat{u}} = \frac{1}{\sqrt{1^2 + 1^2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$$

$$\implies f_{(1,-1)}(0, 1) = \nabla f(0, 1) \cdot \underset{\sim}{\hat{u}} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix} = \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}} = 0$$

Properties of the gradient vector

Two important properties of the gradient vector are:

The gradient $\nabla f(a, b)$ is perpendicular to the contour line through (a,b) and points in the direction of the greatest increase of f. In fact, the direction and magnitude of the steepest slope at (a, b) are given by $\nabla f(a, b)$ and $||\nabla f(a, b)||$.

We can understand these two facts by considering the value of $\cos\theta$ in

$$f_{\underset{\sim}{u}} = \nabla f \cdot \frac{\underset{\sim}{u}}{||\underset{\sim}{u}||} = ||\nabla f|| \cos\theta$$

Consider $\underset{\sim}{A} \cdot \underset{\sim}{B} = ||\underset{\sim}{A}||\ ||\underset{\sim}{B}|| \cos\theta$

- Case 1. $\theta = 0$
  ∇f and $\underset{\sim}{u}$ are in the same direction.
  $f_{\underset{\sim}{u}}$ is maximised, because $\cos 0 = 1$.

- Case 2. $\theta = \pi$
  ∇f and $\underset{\sim}{u}$ are in opposite directions.
  $f_{\underset{\sim}{u}}$ is minimised, because $\cos \pi = -1$.

- Case 3. $\theta = \frac{k\pi}{2}$
  ∇f and $\underset{\sim}{u}$ are perpendicular to each other.
  $f_{\underset{\sim}{u}}$ is 0. We can conclude $f(x, y) = C$, the value of f doesn't change, as we move along the contour.

Example 4.1.3

$T(x, y) = 20 - 4x^2 - y^2$ describes the temperature on the surface of a metal plate. In which direction away from the point $(2, -3)$ does the temperature change most rapidly? In which directions away from the point $(2, -3)$ is the temperature not change?

$$T(2, -3) = 20 - 4(2)^2 - (-3)^2 = 20 - 16 - 9 = -5$$

The point $(2, 3)$ lies on a contour line, lets find that line,

$$-5 = T(2, -3) = 20 - 4x^2 - y^2 \implies 4x^2 + y^2 = 25 \text{ an ellipses.}$$

$$\implies \nabla T = \begin{pmatrix} -8x \\ -2y \end{pmatrix} \implies \nabla T(2, -3) = \begin{pmatrix} -16 \\ 6 \end{pmatrix}$$

$\implies$ At $(2, 3)$, temperature is most rapidly increasing in the direction of $(-16, 6)$.

$$\text{Direction of no change} = \underset{\sim}{u} : \nabla T \cdot \underset{\sim}{u} = 0$$

$$\nabla T(2, -3) \cdot \underset{\sim}{u} = \begin{pmatrix} -16 \\ 6 \end{pmatrix} \cdot \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = -16u_1 + 6u_2 = 0$$

$$\therefore \underset{\sim}{u} \in \left\{ \begin{pmatrix} 6 \\ 16 \end{pmatrix}, \begin{pmatrix} -6 \\ -16 \end{pmatrix} \right\}$$

4.2 Lecture 11

Let's go over the one variable case. If $y = f(u)$ and $u = g(x)$, then

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}x} \qquad \text{or} \qquad y'(x) = f'(u)g'(x)$$

Example 4.2.1

A simple one first, what is $\frac{\mathrm{d}y}{\mathrm{d}x}$ given $y = (x^2 + 5)^5$

$$y = f(u) = u^5,\ u = g(x) = x^2 + 5, \qquad \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}f}{\mathrm{d}u} \cdot \frac{\mathrm{d}u}{\mathrm{d}x} = 5u^4 \cdot 2x = 10x(x^2 + 5)^4.$$

Example 4.2.2

Suppose the radius of a cylinder decreases at a rate of $r'(t) = -2\text{cm/s}$. How fast is the volume decreasing?

$$V = \pi r^2 h,\ \text{fix } h = 2,\ \implies V(r) = 2\pi r^2$$

$$\frac{\mathrm{d}r}{\mathrm{d}t} = -2,\ \frac{\mathrm{d}V}{\mathrm{d}r} = 4\pi r$$

$$\frac{\mathrm{d}V}{\mathrm{d}t} = \frac{\mathrm{d}V}{\mathrm{d}r} \cdot \frac{\mathrm{d}r}{\mathrm{d}t} = 4\pi r \cdot -2 = -8\pi r$$

$$\therefore \left.\frac{\mathrm{d}V}{\mathrm{d}t}\right|_{r=1} = -8\pi \text{cm}^3/\text{sec}.$$

The Multivariate Chain Rule

Let $U \subseteq \mathbb{R}^n$ and $V \subseteq \mathbb{R}$, and $f : U \to \mathbb{R}$, $\underset{\sim}{x} = (x_1, x_2, \ldots, x_n)$ such that $x_i : V \to \mathbb{R}$, then, provided that f and all x_i are differentiable,

$$\frac{\mathrm{d}y}{\mathrm{d}t} f(\underset{\sim}{x}(t))\bigg|_{t=t_0} = \sum_{i=1}^{n} \frac{\partial f(\underset{\sim}{x}(t_0))}{\partial x_i} x_i'(t_0)$$

Daunting as it looks, it's really quite simple, let's take $n = 2$ for example. Suppose $z = f(x, y)$ and x and y are both functions with respect to t. Then

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

If we take $n = 4$, $z = f(a, b, c, d)$,

$$\frac{\mathrm{d}z}{\mathrm{d}t} = \frac{\mathrm{d}z}{\mathrm{d}a}\frac{\mathrm{d}a}{\mathrm{d}t} + \frac{\mathrm{d}z}{\mathrm{d}b}\frac{\mathrm{d}b}{\mathrm{d}t} + \frac{\mathrm{d}z}{\mathrm{d}c}\frac{\mathrm{d}c}{\mathrm{d}t} + \frac{\mathrm{d}d}{\mathrm{d}a}\frac{\mathrm{d}d}{\mathrm{d}t}$$

Example 4.2.3

Suppose $z = f(x, y) = 4x^2 + 3y^2$ and $x(t) = \sin t$, $y(t) = \cos t$. Then

$$\begin{aligned} \frac{\mathrm{d}z}{\mathrm{d}x} &= 8x, \qquad \frac{\mathrm{d}x}{\mathrm{d}t} = \cos t \\ \frac{\mathrm{d}z}{\mathrm{d}y} &= 6y, \qquad \frac{\mathrm{d}y}{\mathrm{d}t} = -\sin t \end{aligned}$$

$$\therefore \frac{\mathrm{d}z}{\mathrm{d}t} = \frac{\mathrm{d}z}{\mathrm{d}x} \cdot \frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\mathrm{d}z}{\mathrm{d}y} \cdot \frac{\mathrm{d}y}{\mathrm{d}t} = 8x\cos t - 6y\sin t = 8\sin t\cos t - 6\cos t\sin t = 2\cos t\sin t$$

Example 4.2.4

Suppose the radius from the previous example, also has a change in height of -1cm/sec, what is the new rate of change in volume?

$$V(r, h) = \pi r^2 h, \qquad \frac{\mathrm{d}r}{\mathrm{d}t} = -2, \qquad \frac{\mathrm{d}h}{\mathrm{d}r} = -1$$

$$V_r(r, h) = 2\pi r h, \qquad V_h(r, h) = \pi r^2$$

$$\therefore \frac{\mathrm{d}V}{\mathrm{d}t} = \frac{\mathrm{d}V}{\mathrm{d}r} \cdot \frac{\mathrm{d}r}{\mathrm{d}t} + \frac{\mathrm{d}V}{\mathrm{d}h} \cdot \frac{\mathrm{d}h}{\mathrm{d}t} = -4\pi r h - \pi r^2 = -\pi r(4h + r)$$

$$\therefore \frac{\mathrm{d}V}{\mathrm{d}t}\bigg|_{\substack{r=1,\\ h=2}} = -\pi(1)(4(2) + 1) = -\pi(8 + 1) = -9\pi \text{ cm}^3/\text{sec}.$$

Extended Chain Rule with Tree Diagrams

We can represent $z = f(x(t), y(t))$ using a tree diagram:

$$\begin{aligned} z\ &-\frac{\partial z}{\partial x}\!\to x - \frac{\mathrm{d}x}{\mathrm{d}t}\!\to t \\ &-\frac{\partial z}{\partial y}\!\to y - \frac{\mathrm{d}y}{\mathrm{d}t}\!\to t \end{aligned}$$

From this, we can observe that $\frac{\mathrm{d}z}{\mathrm{d}t} = \frac{\partial z}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial z}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t}$.

To represent $z = f(u(x(t), y(t)), v(x(t), y(x)))$.

$$
\begin{aligned}
z\ -\frac{\partial z}{\partial u}{\rightarrow} u - \frac{\partial u}{\partial x}{\rightarrow} x - \frac{\mathrm{d}x}{\mathrm{d}t}{\rightarrow} t \\
-\frac{\partial u}{\partial y}{\rightarrow} y - \frac{\mathrm{d}y}{\mathrm{d}t}{\rightarrow} t \\
-\frac{\partial z}{\partial v}{\rightarrow} v - \frac{\partial v}{\partial x}{\rightarrow} x - \frac{\mathrm{d}x}{\mathrm{d}t}{\rightarrow} t \\
-\frac{\partial v}{\partial y}{\rightarrow} y - \frac{\mathrm{d}y}{\mathrm{d}t}{\rightarrow} t
\end{aligned}
$$

From this, we can gather that

$$\frac{\mathrm{d}z}{\mathrm{d}t} = \frac{\partial z}{\partial u}\left(\frac{\partial u}{\partial x}\cdot\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial u}{\partial y}\cdot\frac{\mathrm{d}y}{\mathrm{d}t}\right) + \frac{\partial z}{\partial v}\left(\frac{\partial v}{\partial x}\cdot\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial v}{\partial y}\cdot\frac{\mathrm{d}y}{\mathrm{d}t}\right).$$

Implict Differentiation

Both $y + \sin(x + y) = 0$ and $x^{11} - y_x^7 y^2 + 1 = 0$ correspond to curves in $\mathbb{R}^2$. In neither case can we explicitly solve for y as a function of x. This raises the question: can we compute $y'(x)$? The answer is yes, and we the method of computation is named implicit differentiation. In calculus I, we might have used a "brute force" approach:

Example 4.2.5

$$
\begin{aligned}
y + \sin(x+y) &= 0 \\
\frac{\mathrm{d}}{\mathrm{d}x}(y) + \frac{\mathrm{d}}{\mathrm{d}x}(\sin(x+y)) &= \frac{\mathrm{d}}{\mathrm{d}x}(0) \\
\frac{\mathrm{d}y}{\mathrm{d}x} + \cos(x+y)\frac{\mathrm{d}}{\mathrm{d}x}(x+y) &= 0 \\
\frac{\mathrm{d}y}{\mathrm{d}x} + \cos(x+y)\left(1 + \frac{\mathrm{d}y}{\mathrm{d}x}\right) &= 0 \\
\frac{\mathrm{d}y}{\mathrm{d}x} + \cos(x+y)\frac{\mathrm{d}y}{\mathrm{d}x} &= -\cos(x+y) \\
\frac{\mathrm{d}y}{\mathrm{d}x}(1 + \cos(x+y)) &= -\cos(x+y) \\
\therefore \frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{-\cos(x+y)}{1+\cos(x+y)}
\end{aligned}
$$

Alternatively, we can view the given curve C in $\mathbb{R}^2$ (such as $y + \sin(x+y) = 0$) as a contour $z = 0$ of a function $f(x, y)$. That is to say, in this case, we'd define the functoin $f(x, y) = y + \sin(x+y)$. We know that, for the point (x, y) on C, $\nabla f(x, y)$ corresponds to a vector perpendicular to the C at the point; in other words, the derivative, defined implicitly. In other words,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{f_x}{f_y}.$$

Example 4.2.6

$$
\begin{aligned}
\text{Let } f(x,y) &= y + \sin(x+y) \\
\text{Then } f_x(x,y) &= \cos(x+y)\frac{\mathrm{d}}{\mathrm{d}x}(x+y) = \cos(x+y) \\
\text{And } f_y(x,y) &= 1 + \cos(x+y)\frac{\mathrm{d}}{\mathrm{d}x}(x+y) = 1 + \cos(x+y) \\
\therefore \frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{-\cos(x+y)}{1+\cos(x+y)}
\end{aligned}
$$

Chapter 5

Week 5

5.1 Lecture 12

In a previous section, we looked at the first-order, or linear apporximation to a function. In this section, we introduce a second-order, or quadratic approximation.

One Dimensional Review

Given a function $f(x)$, we can approximate it near $x = a$ as

$$f(x) \approx f(a) + f'(a)(x-a) + f''(a)(x-a)^2.$$

This corresponds to a parabola through $(a, f(a))$, such that the derivative and second derivatives at that point match the "real" function.

$$\begin{gathered} Q(x) = f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 \implies Q(a) = f(a) \\ Q'(x) = f'(a) + f''(a)(x-a) \implies Q'(a) = f'(a) \\ Q''(x) = f''(a) \implies Q''(a) = f''(a) \end{gathered}$$

When we move to the multivariate case, we should keep this in mind.

Quadratic Approximation of $f(x, y)$

The quadratic or second-order approximation of $f(x, y)$ around the point (a, b) is a function of the form

$$Q(x, y) = c + mx + ny + Ax^2 + Bxy + Cy^2$$

such that $Q(a, b) = f(a, b)$ and all first and second order partial derivatives of Q and f are the same. Considering this, we land on the function

$$Q(x, y) = f(a, b) + f_x(a, b)(x-a) + f_y(a, b)(y-b) + \frac{1}{2}f_{xx}(a, b)(x-a)^2 + f_{xy}(a, b)(x-a)(y-b) + \frac{1}{2}f_{yy}(a, b)(y-b)$$

We should verify this function meets our criteria

$$\begin{gathered} \text{Trivially, } Q(a, b) = f(a, b) \\ Q_x(x, y) = f_x(a, b) + f_{xx}(a, b)(x-a) + f_{xy}(a, b)(y-b) \implies Q_x(a, b) = f_x(a, b) \\ Q_y(x, y) = f_y(a, b) + f_{yy}(a, b)(y-b) + f_{xy}(a, b)(x-a) \implies Q_y(a, b) = f_y(a, b) \\ Q_{xx}(x, y) = f_{xx}(a, b) \implies Q_{xx}(a, b) = f_{xx}(a, b) \\ Q_{yy}(x, y) = f_{yy}(a, b) \implies Q_{yy}(a, b) = f_{yy}(a, b) \\ Q_{xy}(x, y) = f_{xy}(a, b) \implies Q_{xy}(a, b) = f_{xy}(a, b) \end{gathered}$$

Amazing!

Example 5.1.1

Find the quadratic approximation around $(0,0)$ of

$$f(x,y) = 1 - x^2 + y^2 + xy + x^3 + x^2y^2.$$

$$\begin{aligned}
&& f(0,0) &= 1 \\
f_x(x,y) &= -2x + y + 3x^2 + 2xy^2 & \implies f_x(0,0) &= 0 \\
f_y(x,y) &= 2y + x + 2x^2y & \implies f_y(0,0) &= 0 \\
f_{xx}(x,y) &= -2 + 6x + 2y^2 & \implies f_{xx}(0,0) &= -2 \\
f_{yy}(x,y) &= 2 + 2x^2 & \implies f_{yy}(0,0) &= 2 \\
f_{xy}(x,y) &= 1 + 4xy & \implies f_{xy}(0,0) &= 1
\end{aligned}$$

$$\therefore Q(x,y) = 1 - x^2 + xy + y^2$$

Example 5.1.2

Find the linear and the quadratic approximation around $(0,0)$ of

$$f(x,y) = \exp(-x^2 - y^2)$$

$$\begin{aligned}
&& f(0,0) &= 1 \\
f_x(x,y) &= -2x\exp(-x^2 - y^2) & \implies f_x(0,0) &= 0 \\
f_y(x,y) &= -2y\exp(-x^2 - y^2) & \implies f_y(0,0) &= 0 \\
f_xx(x,y) &= (4x^2 - 2)\exp(-x^2 - y^2) & \implies f_{xx}(0,0) &= -2 \\
f_{yy}(x,y) &= (4y^2 - 2)\exp(-x^2 - y^2) & \implies f_{yy}(0,0) &= -2 \\
f_{xy}(x,y) &= 4xy\exp(-x^2 - y^2) & \implies f_{xy}(0,0) &= 0
\end{aligned}$$

$$\therefore L(x,y) = 1$$
$$\therefore Q(x,y) = 1 - x^2 - y^2$$

5.2 Lecture 13

Critical Points

A smooth, one variable function f has local maxima and minima where $f(x)$ has zero slope, i.e. where

$$\frac{\mathrm{d}f}{\mathrm{d}x} = 0.$$

The second derivative test allows us to determine the nature of $f(x)$ at that extrema,

$$\frac{\mathrm{d}^2 f}{\mathrm{d}x^2} > 0 \implies \text{local minima}$$
$$\frac{\mathrm{d}^2 f}{\mathrm{d}x^2} < 0 \implies \text{local maxima}$$
$$\frac{\mathrm{d}^2 f}{\mathrm{d}x^2} = 0 \implies \text{inconclusive}$$

Note, however, that these are not global extrema. If we define the function in a domain, then we can determine the global extrema.

A global extrema of a continuous function occurs either at the local extrema or on the boundary of its domain.

In some circumstances, there may be a local extrema in places where the derivatives aren't defined, at cusps, for example. Take, $f(x) = |x - 3|$ for instance. $f'(3)$ is not defined, but $f(3)$ is a local minima.

A more general defintion of local extrema might then be:

- A local minimum occurs at a point a where $f(x) \geq f(a) \forall x$ sufficently close to a.
- A local maximum occurs at a point a where $f(x) \leq f(a) \forall x$ sufficently close to a.

If f is continuous, a local extrema occurs either when $f'(x) = 0$ or $f'(x)$ is undefined.

Critical Points of $f(x, y)$

First we define local extrema as:

- $f(x, y)$ has a local maximum at (a, b) if $f(a, b) \geq f(x, y),\ \forall (x, y)$ sufficently close to (a, b).
- $f(x, y)$ has a local minimum at (a, b) if $f(a, b) \leq f(x, y),\ \forall (x, y)$ sufficently close to (a, b).

If f is differentiable, local extrema can only occur in points where the corresponding tangent plane is horizontal, in other words

$$\nabla f = \underset{\sim}{0}.$$

Note though, that ∇f is undefined if any partial derivatives are undeinfed.

Definition 5.2.1: Critical Point

A critical point of $f(x, y)$ occurs when $\nabla f = (0, 0)$ or ∇f is undefined.

Note:-

j The vague description "sufficently close" can be formally and precisely defined. In the simplest case $f : \mathbb{R}^2 \to \mathbb{R}$, the function f has a local maximum at (a, b) if $\exists \varepsilon > 0 : f(a, b) \geq f(x, y)\ \forall (x, y) : ||(x, y) - (a, b)|| < \varepsilon$. Analogous definitions apply for local minima and saddle points.

Definition 5.2.2: Saddle Point

A function f has a saddle point at P_0 if P_0 is a critical point of f and, if within any distance of P_0, there are points P_1 and P_2 with

$$f(P_1) > f(P_0) \qquad \text{and} \qquad f(P_2) < f(P_0)$$

Theorem 5.2.1 Second-Derivative Test

Assume that f and its first and second derivatives are all continuous at (a, b) and $\nabla f = (0, 0)$. Let

$$D = \begin{vmatrix} f_{xx}(a, b) & f_{xy}(a, b) \\ f_{yx}(a, b) & f_{yy} \end{vmatrix}.$$

Then, if

- $f_{xx}(a, b) > 0$ and $D > 0 \implies (a, b)$ is a local minimum.
- $f_{xx}(a, b) < 0$ and $D > 0 \implies (a, b)$ is a local maximum.
- $D < 0 \implies (a, b)$ is a saddle point.
- $D = 0 \implies$ second-derivative test is inconclusive.

Remark. the matrix

$$\begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}$$

is known as the Hessian matrix, and its determinant is known as the discriminant.

Example 5.2.1

Find and classify the critical points of $f(x, y) = \exp(-x^2 + y^2)$.

$$\begin{aligned}
\nabla f &= (-2x\exp(-x^2+y^2), 2y\exp(-x^2+y^2)) \\
(-2x\exp(-x^2+y^2),\ 2y\exp(-x^2+y^2)) &= (0,0) \implies (0,0) \text{ is the only critical point.} \\
D &= \begin{vmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{vmatrix} \\
&= \begin{vmatrix} 4x^2\exp(-x^2+y^2) & -4xy\exp(-x^2+y^2) \\ -4xy\exp(-x^2+y^2) & 4y^2\exp(-x^2+y^2) \end{vmatrix} \\
D(0,0) &= \begin{vmatrix} -2 & 0 \\ 0 & 2 \end{vmatrix} \\
&= -4 < 0 \implies (0,0) \text{ is a saddle point.}
\end{aligned}$$

Global Maxima and Minima

To find the global extrema values of f on a closed and bounded domain, D, we need to compare both.

- the values of f at the critical points, where $\nabla f = \underset{\sim}{0}$, and
- the extrema values on the boundary of D.

5.3 Lecture 14

Constrained Optimisation and Lagrange Multipliers

Practically, we often need to optimise a function subject to certain constraints.

Suppose $f(x, y)$ is a function we seek to maximise, which is constrained by some constraint euqation (or condition) $g(x, y)$, then the extrema occurs when the contour $g(x, y) = 0$ is tangent to a countour of $f(x, y)$, ie,

$$\nabla f = \lambda \nabla g$$

for some λ, known as the Lagrange multiplier.

Example 5.3.1

Find the minimum value of $x^2 + y^2$ subject to the constraint $x + y = 1$.

$$\begin{aligned}
f(x,y) = x^2 + y^2, \qquad g(x,y) &= x + y - 1,\ g(x,y) = 0 \\
\nabla f &= \lambda \nabla g \\
(2x, 2y) &= \lambda(1,1) \\
\nabla f = \lambda \nabla g &\Rightarrow 2x = \lambda 1,\ 2y = \lambda 1 \\
&\Rightarrow x = \frac{\lambda}{2} = y \\
&\Rightarrow x + y - 1 \Rightarrow 2x = 1 \Rightarrow x = \frac{1}{2} \Rightarrow y = \frac{1}{2} \\
\therefore f\left(\frac{1}{2}, \frac{1}{2}\right) &= \frac{1}{2} \text{ is a local minimum.}
\end{aligned}$$

Check:

$$f(\frac{1}{2} + \varepsilon, \frac{1}{2} - \varepsilon),\ \text{satisfying } x + y = 1$$

$$f(\frac{1}{2}+\varepsilon, \frac{1}{2}-\varepsilon) = \left(\frac{1}{2}+\varepsilon\right)^2 + \left(\frac{1}{2}-\varepsilon\right)^2 = \frac{1}{4}+\frac{1}{4}+\varepsilon^2 = \frac{1}{2}+\varepsilon^2$$

Example 5.3.2

Let $A = 4xy$ describe the area of a rectangle centred on the origin, with width $2x$ and height $2y$. What is the maximum area of this rectangle, inside a bounding elipse, described by $x^2/a^2 + y^2/b^2 = 1$?

$$\begin{gathered}
f(x,y) = 4xy, \qquad g(x,y) = \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1,\ g(x,y) = 0 \\
\nabla f = \lambda \nabla g \Rightarrow (4y, 4x) = \lambda \left(\frac{2x}{a^2}, \frac{y}{b^2}\right) \\
\Rightarrow 4y = \lambda \frac{2x}{a^2}, \qquad 4x = \lambda \frac{2y}{b^2} \\
\Rightarrow \frac{4y}{4x} = \frac{2x\lambda}{a^2} \cdot \frac{b^2}{2y\lambda} \Rightarrow y^2 = \frac{x^2 b^2}{a^2}
\end{gathered}$$

Note that $x = 0, y = 0 \Rightarrow A = 0$. Since our goal is to maximise area, $x \neq 0$ and $y \neq 0$.

$$\begin{gathered}
\Rightarrow \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0 \Rightarrow \frac{x^2}{a^2} + \frac{x^2 b^2}{b^2 a^2} = 1 \Rightarrow x^2 = \frac{a^2}{2} \Rightarrow y^2 = \frac{b^2}{2} \\
\therefore \text{max area at } (x,y) = \left(\frac{a}{\sqrt{2}}, \frac{b}{\sqrt{2}}\right) \\
\therefore \text{max area} = f\left(\frac{a}{\sqrt{2}}, \frac{b}{\sqrt{2}}\right) = 4 \cdot \frac{a}{\sqrt{2}} \cdot \frac{b}{\sqrt{2}} = 2ab
\end{gathered}$$

Lagrange multipliers can be extended to any number of dimensions.

Example 5.3.3

Let $V(x,y,z) = xyz$ describe the volume of a box, where $x, y, z > 0$. Find the maximum volume of the box subject to the constraint that the vertex (x,y,z) lies on the plane $z + 2x + 3y = 6$.

$$\begin{gathered}
f(x,y,z) = xyz, \qquad g(x,y,z) = 2x + 3y + z - 6,\ g(x,y,z) = 0. \\
\nabla f = \lambda \nabla g \Rightarrow (yz, xz, xy) = \lambda(2,3,1) \\
\left.\begin{aligned} yz &= 2\lambda \\ xz &= 3\lambda \\ xy &= 1\lambda \end{aligned}\right\} \Rightarrow \begin{aligned} &\frac{yz}{xy} = \frac{2\lambda}{3\lambda} \Rightarrow \frac{y}{x} = \frac{2}{3} \Rightarrow y = \frac{2}{3}x \\ &\frac{yz}{xy} = \frac{2\lambda}{1\lambda} \Rightarrow \frac{z}{x} = 2 \Rightarrow z = 2x \end{aligned} \\
\Rightarrow 2x + 3y + z - 6 = 0 \Rightarrow 2x + 2x + 2x = 6 \Rightarrow x = 1 \Rightarrow y = \frac{2}{3},\ z = 2 \\
\therefore \text{Max volume at } \left(1, \frac{2}{3}, 2\right) \\
\therefore \text{Max volume is } f\left(1, \frac{2}{3}, 2\right) = \frac{4}{3}
\end{gathered}$$

If we consider the square of the distance function,

$$D^2 = (x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2 =: f(x,y,z)$$

we can treat problems of finding the distance (ie minimum distance) between a point and a plane, as optimisation problems using Lagrange multipliers. The point is encoded into the distance functoin, as above, and the plane of interest is encoded into the constraint, $g(x,y,z)$.

Example 5.3.4

Use Lagrange multipliers to find the distance between the point $(1,2,3)$ and the plane $x+y+z=1$

$$\text{We seek to minimise } D^2 = f(x,y,z) = (x-1)^2 + (y-2)^2 + (z-3)^2$$
$$\text{constrained by } g(x,y,z) = x+y+z-1,\ g(x,y,z) = 0$$
$$\nabla f = \lambda \nabla g = \begin{pmatrix} 2(x-1) \\ 2(y-2) \\ 2(z-3) \end{pmatrix} = \begin{pmatrix} \lambda \\ \lambda \\ \lambda \end{pmatrix}$$
$$\left.\begin{array}{l} 2x-2=\lambda \\ 2y-4=\lambda \\ 2y-6=\lambda \end{array}\right\} \Rightarrow \begin{array}{l} x = 1+\frac{\lambda}{2} \\ y = 2+\frac{\lambda}{2} \\ z = 3+\frac{\lambda}{2} \end{array}$$
$$x+y+z-1=0 \Rightarrow 1+\frac{\lambda}{2}+2+\frac{\lambda}{2}+3+\frac{\lambda}{2} \Rightarrow \frac{3\lambda}{2} = -5 \Rightarrow \lambda = \frac{-10}{3}$$
$$\Rightarrow x = \frac{-2}{3}, \qquad y = \frac{1}{3}, \qquad z = \frac{4}{3}\text{Optimised } f\left(\frac{-2}{3},\frac{1}{3},\frac{4}{3}\right) = \frac{25}{9}+\frac{25}{9}+\frac{25}{9} = \frac{25}{3}$$
$$\text{Distance : } D = \sqrt{f\left(\frac{-2}{3},\frac{1}{3},\frac{4}{3}\right)} = \sqrt{\frac{25}{3}} = \frac{5}{\sqrt{3}} = \frac{5\sqrt{3}}{3}$$

Chapter 6

Week 6

6.1 Lecture 15

Extedning Lagrange Multipliers

Suppose we want to have arbitrary constraints, and variables. To optimise the function $f(x_1, x_2, \ldots, x_L)$ with constraints $g_i(x_1, x_2, \ldots, x_L)$, for $i = 1, \ldots, N$,

$$\bar{f}(x_1, \ldots, x_L, \lambda_1, \ldots \lambda_N) = f(x_1, \ldots, x_L) - \sum_{i=1}^{N} \lambda_i g_i(x_1, \ldots, x_L)$$

The solution is obtained by solving the system of equations,

$$\begin{aligned}
\frac{\partial \bar{f}}{\partial x_k} &= \frac{\partial f}{\partial x_k} - \sum_{i=0}^{N} \lambda_i \frac{\partial g_i}{\partial x_k} = 0, \quad && k = 1, \ldots, L \\
\frac{\partial \bar{f}}{\partial \lambda_i} &= g_i = 0, && i = 1, \ldots, N
\end{aligned}$$

We call $\bar{f}$ the Lagrangian function, or the Lagrangian of f.

Note:-

This method of Lagrange multipliers can be generalised to take into accoutn inequality constraints of the form $g(x_1, \ldots, x_L) \leq c$. These are known as Karush-Kuhn-Tucker conditions, and will be studied in further coruses.

The Envelope of a Family of Functions

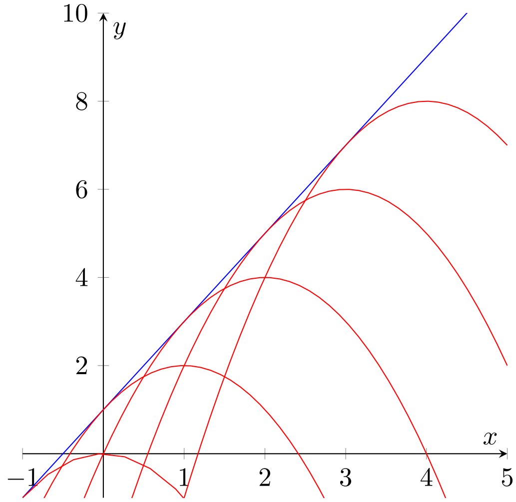
Consider a function of two variables, $f(x, t)$. We could interpret this as a "family" of 1 varaible functions,

$$y^{(t)}(x) = f(x, t).$$

That is to say that, for each value of t, there is an associated one variable function, which can then be interpretted as a cross section.

Let's consider the example

$$f(x, t) = 2t - (x - t)^2$$



We can see that the curve $y = 2x + 1$ is a tangent line to each of the cross=sections. In this case, $y = 2x + 1$ is an envelope for a family of one-variable functions.

The envelope, in general, is a curve which has the property that it is tangential to each cross-section at some point. The derivative of $y^{(t)}(x)$ is simply

$$\frac{\mathrm{d}y^{(t)}}{\mathrm{d}x} = f_x(x, t).$$

Let $t = g(x)$ be a solution (if it exists) of $f_t(x, g(x)) = 0$. Then define the evelope $E(x)$ to be

$$E(x) = f(x, g(x))$$

The curve $E(x)$ intersects the cruve $y^{(t*)}(x)$ at the point x_0 whenever $t* = g(x_0)$. Moreover, these curves have the same derivative at x_0.

Example 6.1.1

Show that the enveloping function of $f(x, t) = 2t - (x - t)^2$ is $y = 2x + 1$.

$$\begin{aligned}
y^{(t)}(x) &= f(x, t) \\
\frac{\partial f}{\partial t} &= 2 + 2(x - t) = 2 + 2x - 2t \\
\text{Let } f_t(x, t) &= 0 \\
\text{Then } 2 + 2x - 2t &= 0 \\
1 + x - t &= 0 \\
1 + x &= t \\
\text{Substitute } t = g(x) = x + 1 &\to f(x, t) \\
f(x, g(x)) &= 2(x + 1) - (x - x - 1)^2 \\
&= 2x + 2 - 1 \\
&= 2x + 1 \\
\therefore E(x) &= f(x, g(x)) = 2x + 1
\end{aligned}$$

6.2 Lecture 16

Line Integrals

Say we have a curve in $\mathbb{R}^3$ parametrised by

$$x = x(t), \qquad y = y(t), \qquad z = z(t), \qquad a \le t \le b.$$

In other words, $\underset{\sim}{r}(t) = (x(t), y(t), z(t)),\ t \in [a, b]$. We can work out the arc length of such a curve. Each individual segment of the arc length can be approximated,

$$\Delta s_i \approx \left|\left|\frac{\mathrm{d}\underset{\sim}{r}}{\mathrm{d}t}\right|\right|_{t=t_i} \Delta t_i$$

The arc length, S will be the sum of all these segments, namely,

$$S = \sum_{i=1}^{n} \Delta s_i \approx \sum_{i=1}^{n} \Delta t_i \left|\left|\frac{\mathrm{d}\underset{\sim}{t}}{\mathrm{d}t}\right|\right|_{t=t_i}$$

If we take the limit, as $n \to \infty$ of the expression, $\Delta t \to 0$ and $\Delta s \to 0$, and we see

$$S = \int_a^b \left|\left|\frac{\mathrm{d}\underset{\sim}{r}}{\mathrm{d}t}\right|\right| \mathrm{d}t = \int_a^b \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}z}{\mathrm{d}t}\right)^2}\,\mathrm{d}t$$

We use new notation to describe this,

$$S = \int_C \mathrm{d}s \implies \mathrm{d}s = \left|\left|\frac{\mathrm{d}\underset{\sim}{r}}{\mathrm{d}t}\right|\right| \mathrm{d}t$$

The line integral over C. In general, if $f(x, y)$ is defined on a smooth curve $C \in \mathbb{R}^2$ then the line integral of f over C is

$$\int_C f(x, y)\mathrm{d}s$$

If C is parametrised by $\underset{\sim}{r}(t),\ t \in [a, b]$ and $f(x, y) > 0$, then this line integral gives the ribbon area under f along C. Line integrals are generaltisations of one-dimensional definite integrals.

Example 6.2.1

Find the length of the helix described by

$$\begin{cases} x = \cos 10t \\ y = \sin 10t \\ z = t \end{cases}, \qquad t \in [0, \pi]$$

$$\underset{\sim}{r}(t) = (\cos(10t), \sin(10t), t) \Rightarrow \frac{\mathrm{d}\underset{\sim}{r}}{\mathrm{d}t} = (-10\sin(10t), 10\cos(10t), 1)$$

$$\frac{\mathrm{d}s}{\mathrm{d}t} = \left|\left|\frac{\mathrm{d}\underset{\sim}{r}}{\mathrm{d}t}\right|\right| = \sqrt{100\sin^2(10t) + 100\cos^2(10t) + 1} = \sqrt{100(\sin^2(10t) + \cos^2(10t)) + 1} = \sqrt{101}$$

$$\therefore S = \int_C \mathrm{d}s = \int_0^{\pi} \sqrt{101}\mathrm{d}t = \sqrt{101}t\Big|_0^{\pi} = \pi\sqrt{101} - 0\sqrt{101} = \pi\sqrt{101}$$

Example 6.2.2

Find the arc length of a semicircle of radius 2.

$$\underset{\sim}{r} = (2\cos t, 2\sin t),\ t \in [0, \pi]$$

$$\frac{\mathrm{d}\underset{\sim}{r}}{\mathrm{d}t} = (-2\sin t, 2\cos t)$$

$$\left|\left|\frac{\mathrm{d}\underset{\sim}{r}}{\mathrm{d}t}\right|\right| = \sqrt{4\sin^2 t + 4\cos^2 t} = 2$$

$$\therefore S = \int_C \mathrm{d}s - \int_0^{\pi} 2\mathrm{d}t = 2t|_0^{\pi} = 2\pi$$

6.3 Lecture 17

Polar Curves

A polar curve is the set of points whose polar coordinates satisfy

$$r = f(\theta)$$

Here, r is the distance from the origin, and θ is the angular displacement from the positive x axis. Any polar curve can be parametrised as

$$r = f(\theta) \implies \begin{matrix} x(\theta) = f(\theta)\cos\theta \\ y(\theta) = f(\theta)\sin\theta \end{matrix} \qquad \text{or} \qquad \underset{\sim}{r}(\theta) = f(\theta)\cos(\theta)\hat{\imath} + f(\theta)\sin(\theta)\hat{\jmath},\ \theta \in [\alpha, \beta]$$

For a polar curve, the arc length is expressed

$$S = \int_\alpha^\beta \sqrt{f'(\theta)^2 + f(\theta)^2}\ \mathrm{d}\theta.$$

We find this by differentiating $\underset{\sim}{r}$ with respect to θ and finding the square of the magnitude. We end up canceling a bunch of the trigonometric functions, which square and sum to give 1. All thats left over is the root of the sum of the function squared and its derivative squared.

Word Done by a Constant Force

In one dimension, the work done by a constant force, F, along a straight line of length d is $W = Fd$.

In two dimensions, the work done by a constant force in moving a particle along a straight line from P to Q is $W = \underset{\sim}{F} \cdot \overrightarrow{PQ}$. This is because only the component of F which is perpendicular to $\overrightarrow{PQ}$ is contributing to work done. This, in essence, transforms our 2 dimesnionsal problem, back into a one dimensional problem. Where θ is the angle from $\overrightarrow{PQ}$ to $\underset{\sim}{F}$,

$$W = \left|\left|\underset{\sim}{F}\right|\right| \cos\theta \cdot \left|\left|\overrightarrow{PQ}\right|\right| = \underset{\sim}{F} \cdot \overrightarrow{PQ}$$

Word Done Over a Curve

Lets now consider a more general case of work done by an object, through a force field.

$$\underset{\sim}{F}(x, y, z) = F_1(x, y, z)\hat{\imath} + F_2(x, y, z)\hat{\jmath} + F_3(x, y, z)\hat{k}$$

where F_1, F_2, F_3 are continuous functions, and the object moves along a curve C.

Similarly to how we defined the line integral initally, we start by chopping up C into N finite pieces, and observing that the work done by the ith line segment is

$$W_i \approx \underset{\sim}{F}(P_i) \cdot \underset{\sim}{T}(P_i)\Delta s_i,$$

where P_i is the ith sample point, $\underset{\sim}{T}$ is a vector functoin, which is the unit vector tangent to C, and $\underset{\sim}{F}$ is a vector function which is the value of the vector field at the sample point. The dot product represents the $\underset{\sim}{F}$ component tangent to C at the sample point, and Δs_i is a scale factor, the distance from the start to the end of the segment. The total work done, is the sum of all these segments,

$$W \approx \sum W_i = \sum_{\forall i} \underset{\sim}{F}(P_i) \cdot \underset{\sim}{T}(P_i)\Delta s_i.$$

Just as before, we take the limit as the number of segments goes to infinty, or as the segment distance goes to 0, and observe that

$$W = \int_C \underset{\sim}{F} \cdot \underset{\sim}{T}\mathrm{d}s$$

To evaluate the line integral, we use a parameterisation of the curve C. Let C be parameterised by

$$\underset{\sim}{r}(t) = (x(t), y(t), z(t)), \; t \in [a, b]$$

We can write $\underset{\sim}{T}$ in terms o this parameterisation,

$$\underset{\sim}{T}(P_i) = \frac{\underset{\sim}{r}'(t_i)}{||\underset{\sim}{r}'(t_i)||}.$$

Hence the work done over the ith arc is approximately

$$W_i \approx \underset{\sim}{F}(\underset{\sim}{r}(t_i)) \cdot \underset{\sim}{r}'(t_i)\Delta t$$

Summing up all these arcs gives us

$$W \approx \sum_{\forall i} W_i = \sum_{\forall i} \underset{\sim}{F}(\underset{\sim}{r}'t_i)) \cdot \underset{\sim}{r}'(t_i)\Delta t.$$

Finally, we take the limit as $\Delta t \to 0$,

$$W = \int_a^b \underset{\sim}{F}(\underset{\sim}{r}(t)) \cdot \underset{\sim}{r}'(t) \; \mathrm{d}t.$$

This line integral is commonly expressed

$$\int_C \underset{\sim}{F} \cdot \mathrm{d}\underset{\sim}{r}$$

but this is merely a notational convenience, you could consider

$$\mathrm{d}\underset{\sim}{r} = \mathrm{d}x\hat{\imath} + \mathrm{d}y\hat{\jmath} + \mathrm{d}z\hat{k}$$

but again, this is just a notational convenience, and only really serves to remind us where this formula came from.

Example 6.3.1

Evaluate $\int_C \underset{\sim}{F} \cdot \mathrm{d}\underset{\sim}{r}$ where $\underset{\sim}{F} = (xy, yz, zx)$ and C is parameterised $x = t$, $y = t^2$, and $z = t^3$, $t \in [0, 1]$.

$$\begin{aligned}
\underset{\sim}{r}(t) &= (x(t), y(t), z(t)) = (t, t^2, t^3) \\
\therefore \underset{\sim}{r}(t) &= (1, 2t, 3t^2) \\
\underset{\sim}{F}(\underset{\sim}{r}(t)) &= (x(t)y(t), y(t)z(t), z(t)x(t)) = (t^3, t^5, t^4) \\
\therefore \underset{\sim}{F}(\underset{\sim}{r}(t)) \cdot \underset{\sim}{r}'(t) &= (1 \cdot t^3) + (2t \cdot t^5) + (3t^2 \cdot t^4) = t^3 + 2t^6 + 3t^6 = t^3 + 5t^6 \\
\therefore \int_C \underset{\sim}{F} \cdot \mathrm{d}\underset{\sim}{r} &= \int_0^1 t^3 + 5t^6 \; \mathrm{d}t = \frac{1}{4}t^4 + \frac{5}{7}t^7 \bigg|_0^1 = \frac{27}{28}
\end{aligned}$$

Chapter 7

Week 7

7.1 Lecture 18

Conservative Fields

Gradient fields are a type of force field with a special property of being *conservative*. There is a relatively easy way to determine if a field is conservative.

Let A and B be points, then if $\int_A^B \utilde{F} \cdot \mathrm{d}\utilde{r}$ is independent of the path taken (that is, forall paths connected the points A and B, the result is the same), then $\utilde{F}$ is called a conservative field.

That is, if $\utilde{F}$ is conservative, then all paths connecting A and B will give the same result.

If there exists a function $f(x, y)$ such that $\utilde{F} = \nabla f$, then $\utilde{F}$ is a gradient field, with a potential function f. All gradient fields are conservative, and if you can find a potential function, then

$$\int_A^B \utilde{F} \cdot \mathrm{d}\utilde{r} = f(B) - f(A)$$

Note:-
This equation is really important!!

Example 7.1.1
Show that $\utilde{F}(x, y) = (x + y, x + 1)$ is a gradient field.

$$\begin{aligned}
\nabla f &= \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) \\
\frac{\partial f}{\partial x} &= x + y \\
\Rightarrow f(x, y) &= \int \frac{\partial f}{\partial x} \mathrm{d}x \\
&= \frac{1}{2}x^2 + xy + g(y) \\
\frac{\partial f}{\partial y} &= x + 1 \\
&= \frac{\partial}{\partial y}\left(\frac{1}{2}x^2 + xy + g(y)\right) \\
&= x + g'(y) \\
\Rightarrow g'(y) &= 1 \\
\Rightarrow g(y) &= y
\end{aligned}$$

$$\therefore \left(f(x,y) : \underset{\sim}{F} = \nabla f\right) = \frac{1}{2}x^2 + xy + y$$

Example 7.1.2

Show that $\underset{\sim}{F} = \frac{x+y}{2}\hat{\imath} + \frac{1}{2}y\hat{\jmath}$ is not a gradient field.

$$\begin{aligned}
\text{Suppose } \exists f(x,y) &: \underset{\sim}{F} = \nabla f \\
\text{Then } \frac{\partial f}{\partial x} &= \frac{1}{2}xy \\
\Rightarrow f &= \int \frac{\partial f}{\partial x} \mathrm{d}x \\
&= \int \frac{1}{2}x + \frac{1}{2}y \ \mathrm{d}x \\
&= \frac{1}{4}x^2 + \frac{1}{2}xy + g(y) \\
\frac{\partial f}{\partial y} &= \frac{\partial}{\partial y}\left(\frac{1}{4}x^2 + \frac{1}{2}xy + g(y)\right) \\
&= \frac{1}{2}x + g'(y) \\
&= \frac{1}{2}y & ※ \\
\because g'(y) &= \frac{1}{2}y - \frac{1}{2}x & ※ \\
\therefore \nexists f(x,y) &: \underset{\sim}{F} = \nabla f
\end{aligned}$$

It turns out that you don't need to find a potential function to check if a force field $\underset{\sim}{F} = (F_1, F_2)$ is a gradient field. If $\underset{\sim}{F}$ is a gardient field,

$$\frac{\partial F_1}{\partial y} = \frac{\partial F_2}{\partial x}$$

7.2 Lecture 19

Conservation of Energy in Three-Dimensional Space

We've seen that if $\underset{\sim}{F}$ is conservative, $\underset{\sim}{F} = \nabla f$. Take $V(\underset{\sim}{r}) = -f(\underset{\sim}{r})$. By Newton's Second Law,

$$\underset{\sim}{F} = -\nabla V = m\ddot{\underset{\sim}{r}}$$

where $\ddot{\underset{\sim}{r}}$ is the second derivative of position with respect to time, ie, acceleration. We can then apply the dot product of $\dot{\underset{\sim}{r}}$ to both sides:

$$-\nabla V \cdot \dot{\underset{\sim}{r}} = m\ddot{\underset{\sim}{r}} \cdot \dot{\underset{\sim}{r}}$$

Let's now analyse the LHS

$$\begin{aligned}
LHS &= -\left(\frac{\partial V}{\partial x}, \frac{\partial V}{\partial y}, \frac{\partial V}{\partial z}\right) \cdot \left(\frac{\mathrm{d}x}{\mathrm{d}t}, \frac{\mathrm{d}y}{\mathrm{d}t}, \frac{\mathrm{d}z}{\mathrm{d}t}\right) \\
&= -\left(\frac{\partial V}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial V}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t} + \frac{\partial V}{\partial z}\frac{\mathrm{d}z}{\mathrm{d}t}\right) \\
&= -\frac{\mathrm{d}V}{\mathrm{d}t}
\end{aligned}$$

Now the RHS

$$RHS = m\ddot{\underset{\sim}{r}} \cdot \underset{\sim}{r}$$

$$
\begin{aligned}
&= m\left(\frac{\mathrm{d}^2x}{\mathrm{d}t^2}, \frac{\mathrm{d}^y}{\mathrm{d}t^2}, \frac{\mathrm{d}^z}{\mathrm{d}t^2}\right) \cdot \left(\frac{\mathrm{d}x}{\mathrm{d}t}, \frac{\mathrm{d}y}{\mathrm{d}t}, \frac{\mathrm{d}z}{\mathrm{d}t}\right) \\
&= m\left(\frac{\mathrm{d}^2x}{\mathrm{d}t^2}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\mathrm{d}^y}{\mathrm{d}t^2}\frac{\mathrm{d}y}{\mathrm{d}t} + \frac{\mathrm{d}^z}{\mathrm{d}t^2}\frac{\mathrm{d}z}{\mathrm{d}t}\right) \\
&= \frac{m}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}z}{\mathrm{d}t}\right)^2\right) \\
&= \frac{m}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left|\left|\dot{\underset{\sim}{r}}(t)\right|\right|^2
\end{aligned}
$$

Hence the second law of motion becomes

$$
\begin{aligned}
-\frac{\mathrm{d}V}{\mathrm{d}t} &= \frac{m}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left|\left|\dot{\underset{\sim}{r}}(t)\right|\right|^2 \\
0 &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{m}{2}\left|\left|\dot{\underset{\sim}{r}}(t)\right|\right|^2 + V(\underset{\sim}{r}(t))\right) \\
E &= \frac{m}{2}\left|\left|\dot{\underset{\sim}{r}}(t)\right|\right|^2 + V(\underset{\sim}{r}(t))
\end{aligned}
$$

where E is constant, introduced in the integration in the final step.

This is the energy equation. The first term is a function of velocity. It is the contribution to the total energy due to the object's motion, and is called kinetic energy (K.E.). Note that $\left|\left|\frac{\mathrm{d}\underset{\sim}{r}}{\mathrm{d}t}\right|\right|^2 \geq 0$, so kinetic energy is non-negative.

The second term is a function of poistion, called potential energy (P.E.), because it has the potential to be converted to kinetic energy.

Central Forces

A force is called central if it has the form:

$$
\underset{\sim}{F} = F(r)\hat{\underset{\sim}{r}} = \frac{F(r)}{r}\underset{\sim}{r}
$$

where $\underset{\sim}{r}(t) = x\hat{\imath} + y\hat{\jmath} + z\hat{k}$ and $r = \left|\left|\underset{\sim}{r}\right|\right|$.

The magnitude of a central force is dependent only on the distance of the object from the origin. Such a force is called attractive if it acts towards the origin ($F(r) < 0$), and repulsive if it acts away from the origin ($F(r) > 0$).

Theorem 7.2.1

If $F(r)$ is continuous over some domain D, then the central field $\underset{\sim}{F} = F(r)\hat{\underset{\sim}{r}}$ is conservative throughout D.

Proof.

Consider the gradient of the radial distance

$$
\nabla\underset{\sim}{r} = \frac{\mathrm{d}r}{\mathrm{d}x}\hat{\imath} + \frac{\mathrm{d}r}{\mathrm{d}y}\hat{\jmath} + \frac{\mathrm{d}r}{\mathrm{d}z}\hat{k} = \frac{x}{r}\hat{\imath} + \frac{y}{r}\hat{\jmath} + \frac{z}{r}\hat{k} = \frac{1}{r}\left(x\hat{\imath} + y\hat{\jmath} + z\hat{k}\right) = \frac{\underset{\sim}{r}}{r} = \hat{\underset{\sim}{r}}
$$

Let $f(r)$ be the anti derivative of $F(r)$. ie, $F(r) = f'(r)$. Since F is continuous, f must exist.

$$
\nabla f(r) = \frac{\partial f}{\partial x}\hat{\imath} + \frac{\partial f}{\partial y}\hat{\jmath} + \frac{\partial f}{\partial z}\hat{k} = \frac{\mathrm{d}f}{\mathrm{d}r}\frac{\partial r}{\partial x}\hat{\imath} + \frac{\mathrm{d}f}{\mathrm{d}r}\frac{\partial r}{\partial y}\hat{\jmath} + \frac{\mathrm{d}f}{\mathrm{d}r}\frac{\partial r}{\partial h}\hat{k} = \frac{\mathrm{d}f}{\mathrm{d}r}\left(\frac{\partial r}{\partial x}\hat{\imath} + \frac{\partial r}{\partial y}\hat{\jmath} + \frac{\partial r}{\partial z}\hat{k}\right) = f'(r)\nabla\underset{\sim}{r} = F(r)\hat{\underset{\sim}{r}}
$$

Therefore, the central force $\underset{\sim}{F}$ is a gradient field (hence conservative) with potential function $f(r)$.
Now, we set

$$
V(r) = -f(r), \qquad \implies \underset{\sim}{F}(\underset{\sim}{r}) = -\nabla V(r)
$$

Angular Momentum

Consider a particle with mass m, velocity $\underset{\sim}{v}$ and position $\underset{\sim}{r}$, we say its angular momentum is

$$\underset{\sim}{L} = m(\underset{\sim}{r} \times \underset{\sim}{v})$$

If force is exerted on a particle and we consider Newton's Second Law, $\underset{\sim}{F} = m\underset{\sim}{a}$,

$$\dot{\underset{\sim}{L}} = \frac{\mathrm{d}}{\mathrm{d}t}\left(m(\underset{\sim}{r} \times \underset{\sim}{v})\right) = m\left(\dot{\underset{\sim}{r}} \times \underset{\sim}{v} + \underset{\sim}{r} \times \dot{\underset{\sim}{v}}\right) = m\left(\underset{\sim}{v} \times \underset{\sim}{v} + \underset{\sim}{r} \times \underset{\sim}{a}\right) = \underset{\sim}{r} \times m\underset{\sim}{a} = \underset{\sim}{r} \times \underset{\sim}{F}$$

We call $\underset{\sim}{r} \times \underset{\sim}{F}$ torque.

Theorem 7.2.2

Under all central forces, the angular momentum $\underset{\sim}{L}$ is conserved.

Proof.

Recall that for a central force

$$\underset{\sim}{F} = \frac{F(r)}{r}\underset{\sim}{r}$$

$$\dot{\underset{\sim}{L}} = \underset{\sim}{r} \times \underset{\sim}{F} = \underset{\sim}{r} \times \left(\frac{F(r)}{r}\underset{\sim}{r}\right) = \frac{F(r)}{r}\left(\underset{\sim}{r} \times \underset{\sim}{r}\right) = \underset{\sim}{0}$$

This is to say, that at all times, angular momentum, $\underset{\sim}{L}$ does not change, it is conserved. Thus, central forces conserve both energy, and anuglar momentum.

7.3 Lecture 20

Theorem 7.3.1

If a central force acts on an object in three dimensions, its motion is restricted to a plane.

Consider the plan P upon which the inital position vector, $\underset{\sim}{r}_0 = \underset{\sim}{r}(t_0)$, and inital velocity vector, $\underset{\sim}{v}_0 = \underset{\sim}{v}(t_0)$, lie. Since $\underset{\sim}{L} = m(\underset{\sim}{r_0} \times \underset{\sim}{v_0})$, $\underset{\sim}{L}$ is perpendicular to P.

Since $\underset{\sim}{L}$ is conserved, consider the vectors $\underset{\sim}{r}_1 = \underset{\sim}{r}(t_1)$ and $\underset{\sim}{v}_1 = \underset{\sim}{v}(t_1)$,

$$\underset{\sim}{L} = m(\underset{\sim}{r}_0 \times \underset{\sim}{v}_0) = m(\underset{\sim}{r}_1 \times \underset{\sim}{v}_1)$$

Since $\underset{\sim}{r}_1$ and $\underset{\sim}{v}_1$ are orthognal to $\underset{\sim}{L}$ they must therefore also lie in P.

Position-Dependent Unit Vectors for Central Force Fields

The acceleration of an object in a central force field depends only on the radial distance r. For central orce problems, we use polarcoordinates in $\mathbb{R}^2$

$$x = r\cos\theta$$
$$y = r\sin\theta$$

We can express the position vector in a different way. θ is the angle between $\underset{\sim}{r}$ and the x-axis.

$$\begin{aligned} \underset{\sim}{r} &= x\hat{\imath} + y\hat{\jmath} \\ &= r\cos\theta\hat{\jmath} + r\sin\theta\hat{\jmath} \\ &= r(\cos\theta\hat{\jmath} + \sin\theta\hat{\imath}) \\ &= r\hat{\underset{\sim}{r}} \end{aligned}$$

Another unit vector is given by

$$\hat{\underset{\sim}{\theta}} = -\sin\theta\hat{\imath} + \cos\theta\hat{\jmath}$$

This $\underset{\sim}{\hat{\theta}}$ is perpendicular to $\underset{\sim}{\hat{r}}$

$$\underset{\sim}{\hat{r}} \cdot \underset{\sim}{\hat{\theta}} = (\cos\theta \cdot -\sin\theta + \sin\theta \cdot \cos\theta) = 0$$

and

$$\underset{\sim}{\hat{r}} \times \underset{\sim}{\hat{\theta}} = \begin{vmatrix} \hat{\imath} & -\hat{\jmath} & \hat{k} \\ \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \end{vmatrix} = \hat{k}$$

So, the vectors $\underset{\sim}{\hat{r}}$ and $\underset{\sim}{\hat{\theta}}$ form a basis for $\mathbb{R}^2$. Because these vectors are both orthognal and normal, we call them "orthonormal," which is kinda cute. Unlike the x-y Cartesian system we're used to, this co-ordinate system varies with the angle the particle makes with the x-axis. We can also expand these coordinates into concepts of position, velocity and anular momentum.

$$\begin{aligned}
\underset{\sim}{r} &= r\underset{\sim}{\hat{r}} \\
&= r(t)\underset{\sim}{\hat{r}}(\theta(t)) \\
\underset{\sim}{v} = \underset{\sim}{\dot{r}} &= \dot{r}\underset{\sim}{\hat{r}} + r\frac{\mathrm{d}}{\mathrm{d}t}\underset{\sim}{\hat{r}}\dot{\theta} \\
&= \dot{r}\underset{\sim}{\hat{r}} + r\dot{\theta}\underset{\sim}{\hat{\theta}} \\
\underset{\sim}{L} &= m(\underset{\sim}{r} \times \underset{\sim}{v}) \\
&= m(r\underset{\sim}{\hat{r}}) \times (\dot{r}\underset{\sim}{\hat{r}} + r\dot{\theta}\underset{\sim}{\hat{\theta}}) \\
&= mr^2\dot{\theta}\underset{\sim}{\hat{\theta}}
\end{aligned}$$

Motion in a Gravitational Field

It can be shown that the trajectory of an object of mass m in the gravitational system is given by the polar curve

$$r = \frac{l}{1 + e\cos\theta}$$

This is an equation for a conic section, symmetric about the x-axis. Thus we conclude

The orbits of all objects acting under such a gravitational field must be conic sections, either circle, ellipse, parabola, or branch-hyperbola, with a focus at the origin.

This is called Kepler's First Law of Planetary Motion.

Note:-

This won't be on the final exam, so we'll leave it at that. We could show that all gravitational orbits are conic sections, Stewart, ed.8 p916-917 for additional reading. For now, though, we'll just move on to ODEs ;)

Chapter 8

Week 8

8.1 Lecture 21

Ordinary Differential Equations

Differential equations (DEs) is an equation which consists one or more derivatives of an unknown function. There are two types of DEs:

- (ODEs) Ordinary DEs, where the unknown function is a function of only one variable.
- (PDEs) Partial DEs, where the unknown function is a function of more than one variable. We will not consider PDEs at all in this class.

In ODEs, one often takes time, t, as the independent variable, instead of x. Also derivatives such as $x'(t)$ and $x''(t)$ with respect to time, are often denoted $\dot{x}$ and $\ddot{x}$.

Some examples of ODEs:

- Unbounded population growth: $P'(t) = kP$
- Motion due to gravity: $mx''(t) = -kx$
- Spring system: $mx''(t) = -kx$
- Interaction between electric charges: $r'' = Kr^{-2}$
- General form of linear ODEs: $y^{(n)} = f(x, y', y'', \ldots y^{(n-1)})$

We will only study linear ODEs in this course.

Solution to an ODE

Suppose we are given an ODE for y which is an unknown function of x. Then $y = f(x)$ is said to be a solution to the ODE if the ODE is satisfied when $y = f(x)$ and its derivatives are substituted into the equation.

When asked to *solve* an ODE, you are expected to find all possible solutions. This means that you need to find the general solution to the ODE. For an ODE that involves only the unknown function of y and its first derivative, the general solution will involve one arbitrary constant.

Example 8.1.1

Show that $y = A\exp(x^2/2)$ is a general solution to $y' = xy$.

$$\begin{aligned} y' &= xy \\ &= xA\exp(x^2/2) \\ xy &= xA\exp(x^2/2) \end{aligned}$$

Since LHS = RHS, $y = A\exp(x^2/2)$ is a general solution to $y' = xy$.

Example 8.1.2

Find the general solutionn to the DE $y' = x^2$

$$\begin{aligned} y' &= x^2 \\ \Rightarrow y &= \int y' \mathrm{d}x \\ &= \int x^2 dx \\ &= \frac{1}{3}x^3 + C, \ C \in \mathbb{R} \end{aligned}$$

Order of an ODE

The order of an ODE is defined by the highest order derivative present in the ODE

- $\frac{\mathrm{d}P}{\mathrm{d}t} = kP$: First Order
- $m\ddot{x} = -kx$: Second Order
- $x(y'')^2 + y'y''' + 4y^5 = yy'$: Third Order
- $y' = xy$: First Order

Inital Value Problem

An IVP is the problem of solving and ODE subject to some inital conditions of the form $y(t_0) = a, \ y'(t_0) = b$, etc. The solution to an IVP no longer contains any arbitrary constants, because they've been determined by the inital conditions.

Example 8.1.3

A flow-meter in a pipline measures flow-through as $2 + \sin t$L/sec. How much fluid passes through the pipline from time 0 to time t.

$$\frac{\mathrm{d}V}{\mathrm{d}t} = 2 + \sin t, \quad \left[\frac{\mathrm{d}V}{\mathrm{d}t}\right] = L^3T^{-1}, \quad [t] = T^1$$

$$\text{Let's introduce 2 varaibles: } [\omega] = T^{-1}, \quad [\alpha] = L^3T^{-1}$$

These varaibles enable the sin to be dimensionless, and for the RHS to match the dimensions of the LHS. With these variables, let us suggest a new model:

$$\frac{\mathrm{d}V}{\mathrm{d}t} = \alpha\left(2 + \sin(\omega t)\right) \Rightarrow V = \int \frac{\mathrm{d}V}{\mathrm{d}t}\mathrm{d}t = 2\alpha t - \frac{\alpha}{\omega}\cos(\omega t) + C, \ C \in \mathbb{R}$$

$$V(0) = 0 = 0 - \frac{\alpha}{\omega}\cos 0 + C \iff 0 = C - \frac{\alpha}{\omega} \iff C = \frac{\alpha}{\omega}$$

Therefore, the final solution is

$$V(t) = 2\alpha t - \frac{\alpha}{\omega}\cos(\omega t) + \frac{\alpha}{\omega}$$

Solution Curves

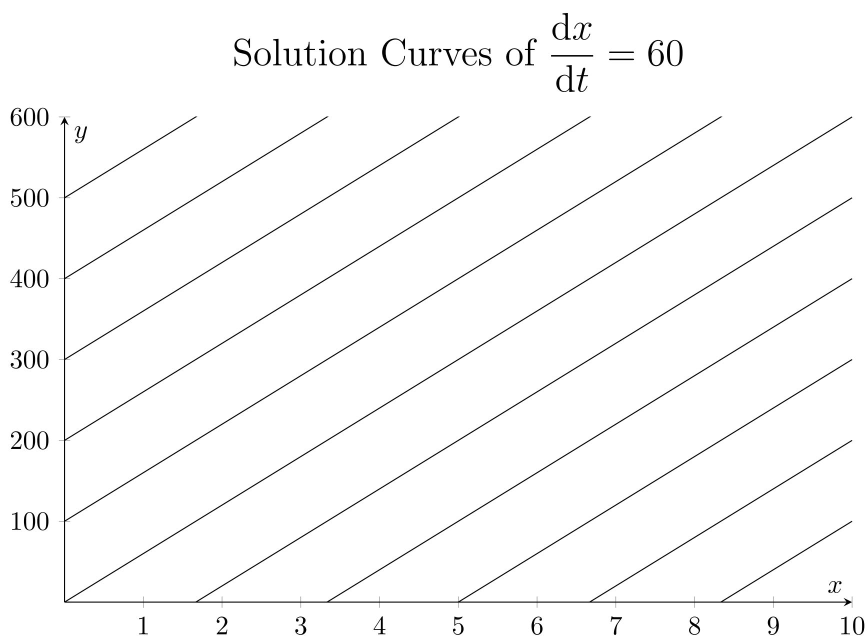
We'll intoduce this with an example. Suppose we want to find the position of a bike, traveling at a constant 60km/hr, along a perfectly straight road, at time t. If $x = x(t)$ is the distance travelled at time t, the ODE is

$$\frac{\mathrm{d}x}{\mathrm{d}t} = 60$$

and we can find the solution by directly integrating,

$$x = \int \frac{\mathrm{d}x}{\mathrm{d}t}\mathrm{d}t = \int 60\mathrm{d}t = 60t + C,\ C \in \mathbb{R}$$

To determine that constant C, we need more information, like the inital position at time 0. In such a case, we'd be solving the IVP. For different values of C, we'd find different solutions, suppose we graphed some of these solutions



These curves, $y = 60t + C$ are called the solution curves to the ODE. Note that in this particular case, the solutions curves are all straight lines of slope 60.

Analytical and Numerical Solutions

To solve an ODE (or IVP) analytically means to give a solution curve in terms of continuous functions defined over a specified interval, where the solution is obtained exactly by analytic means, for example, integration. The solution satisfies the ODE and inital conditions on direct substitution.

To solve an ODE or IVP numerically means to use an algorithm to generate a sequence of points which approximate a solution curve.

8.2 Lecture 22

Slope Fields

We have seen that in order to solve

$$\frac{\mathrm{d}y}{\mathrm{d}t} = f(t)$$

we only need to integrate. However, for the more general first-order ODE,

$$\frac{\mathrm{d}y}{\mathrm{d}t} = f(t, y)$$

this doesn't work anymore. Nonetheless, the ODE gives a qualitative picture of the solution by noting at $(t, y) = (a, b)$ the slope of $y(t)$ is $f(a, b)$. So, as follows, we can

- In the ty-plane at $(t, y) = (a, b)$ draw a small straight line with slope $f(a, b)$.
- Repeat this process for many different values of (a, b).
- The resulting diagram is called the slope field.

Note that the slope field can be generated without having to solve the ODE.

Equilibrium Solutions

An equilibrium solution is a constant solution $y(t) = c$ to the ODE

$$\frac{\mathrm{d}y}{\mathrm{d}t} = f(t, y)$$

The graph of an equilibrium solution is a horizontal line. Such a line has a slope of 0, $y' = 0$, and only happens if $f(t, y) = 0$ has a solution $y = c$ for some $c \in \mathbb{R}$.

Example 8.2.1

Find the equilibrium solutions of $y'(t) = -3(y - 1)$

$$\begin{gathered} y'(t) = 0,\ \forall t \implies y(t) = c,\ \forall t \\ y(t) = 1 \implies y'(0),\ \forall t \end{gathered}$$

Consider the long term behaviour

$$\begin{gathered} y > 1 \implies y'(t) < 0 \implies \text{decreasing solution} \\ y < 1 \implies y'(t) > 0 \implies \text{increasing solution} \\ \implies \lim_{t\to\infty} y(t) = 1 \end{gathered}$$

No matter what the initial conditions are, in the long term, the system will settle at the equilibrium $y(t) = 1$. This can be obsrved nicely on the slope field

Example 8.2.2

Find the equilibrium solutions of $y'(t) = 2t + 1$

$$\begin{gathered} \nexists c \in \mathbb{R} : \forall t,\ y'(t) = 0 \implies \text{no equilibrium} \\ \int y'(t)\mathrm{d}t = \int 2t + 1\mathrm{d}t = t^2 + t + c \end{gathered}$$

Consider the long term behaviour

$$\lim_{t\to\infty} y(t) = +\infty$$

No matter what the inital conditions are, in the long term, the system will diverge positively. No equilibrium point.

Example 8.2.3

Find the equilibrium solutions of $y'(t) = y(1 - y)$

$$\begin{gathered} \forall y'(t) = 0 \implies \forall t, y(t) = c,\ c \in \mathbb{R} \\ y = 1,\ y = 0,\ \text{make equilibrium solutions} \\ y'(t) = 0(1 - 0) = 1(1 - 1) = 0 \end{gathered}$$

Consider the long term behaviour

$$\begin{aligned} y > 1 &\implies y' < 0 \implies \text{decreasing solution} \\ 0 < y < 1 &\implies y' > 0 \implies \text{increasing solution} \\ y < 0 &\implies y' < 0 \implies \text{decreasing solution} \end{aligned}$$

Long term behaviour depends on the inital conditions. If, for example, $y(0) = 2$, then the system will settle at the equilibrium 1. If $y(0) = -2$, in the long term, the system will diverge negatively.

Stability of Equilibrium Solutions

A pencil sitting balanced vertically is in an equilibrium state. But make one small pertubrbation and it will topple. This is an unstable equilibrium. On the other hand, a pendulum hanging vertically is also in an equilibrium state. But if you perturb it slightly, it will eventually settle back down at its equilibrium point; this is a stable equilibrium.

From a slope field you can identify stable and unstable equilibriums by looking at if curves tend towards the equilibrium or away from it as time increases.

Formally, an equilibrium solution $y(t) = y_0$ to the DE $y' = f(t, y)$ is stable if the inital value problem:

$$\frac{\mathrm{d}y}{\mathrm{d}t} = f(t, y), \qquad y(0) = y_0 \pm \varepsilon$$

has a solution which satisfies

$$\lim_{t\to\infty} y(t) = y_0.$$

In other words, if you start from sufficently close to a stable equilibrium, the you will approach that equilibrium solution.

- Stable: All slopes sufficently close to the equilibrium converge to the equilibrium.
- Unstable: All slopes sufficently close to the equilibrium diverge from the the equilibrium.
- Semistable: Some slopes sufficently close to the equilibrium converge to the equilibrium.

Existence and Uniqueness of Solutions

Consider the IVP

$$\frac{\mathrm{d}y}{\mathrm{d}t} = f(t, y), \qquad f(t_0) = y_0$$

There exists a theorem asserting that, if $f(t, y)$ is smooth (that is continuous and first deriviatve is continuous) in some rectangle about (t_0, y_0) there exists a unique solution $y = y_1(t)$ in some small neighbourhood of (t_0, y_0).

Theorem 8.2.1

Let f and $\frac{\mathrm{d}f}{\mathrm{d}y}$ be continuous in some rectangle $\alpha \leq t_0 \leq \beta$ and $a \leq y \leq b$, which contains (t_0, y_0). Then, in some interval $t_0 - h \leq t_0 \leq t_0 + h \ \in [\alpha, \beta]$, there is a unique solution $y(t)$ to the inital value problem.

An important consequence of this is that equilibrium solutions cannot be crossed by other solution curves. In fact, no solutions curves can cross each other because this would mean that in some point(s), $y'(t)$ has more than one value. Equilibrium solutions therefore partition the solution space.

8.3 Lecture 23

Euler's Method for Numerical Solutions to ODEs

The method gives a simple approximation solution to an ODE and is closely related to the notion of a slope field.

Euler's Method using Tangent Lines

The equation of the straight line with slope m passes through the point $(t, y) = (a, b)$ is

$$y = b + m(t - a)$$

Euler's method uses tangent lines as an approximation to solution curves. The tangent line to a solution curve $y' = f(t, y)$ at (t_0, y_0) is

$$y = y_0 + f(t_0, y_0)(t - t_0)$$

This approximates the cruve when t is close to t_0. Imagine then, a family of solution curves to the differential equation. We can calculate an approximate value for y at some later time by taking lots of small steps in time. At each step we will use the tangent line to the solution through our current point. This is Euler's Method.

Using Δt as the step size for the algorithm,

$$t_1 = t_0 + \Delta t, \quad t_2 = t_1 + \Delta t, \quad \ldots, \quad t_n = t_{n-1} + \Delta t$$

To then approximate the y values,

$$y_1 = y_0 + f(t_0, y_0)\Delta t, \quad y_2 = y_1 + f(t_1, y_1)\Delta t, \quad \ldots, \quad y_n = y_{n-1} + f(t_{n-1}, y_{n-1})\Delta t,$$

Example 8.3.1

Use Euler's Method with $\Delta t = 0.2$ to approximate a solution to $y(0.6)$ for the IVP

$$\frac{\mathrm{d}y}{\mathrm{d}t} = 2t, \qquad y(0) = 0.$$

Compare your approximation to the real value

$$t_0 = 0, \quad t_1 = t_0 + \Delta t = 0 + 0.2 = 0.2, \quad t_2 = 0.4, \quad t_3 = 0.6$$

Use Euler's Method to find a numerical solution

$$\begin{gathered}
y_0 = y(0) = 0, \\
y_1 = y_0 + f(t_0)\Delta t = 0 + (0.0)0.2 = 0, \\
y_2 = y_1 + f(t_1)\Delta t = 0 + (0.4)0.2 = 0.08, \\
y_3 = y_2 + f(t_2)\Delta t = 0.08 + (0.8)0.2 = 0.24 \\
\therefore y(0.6) \approx 0.24
\end{gathered}$$

Now let's find the analytical solution

$$\begin{gathered}
\frac{\mathrm{d}y}{\mathrm{d}t} = 2t \implies y = \int \frac{\mathrm{d}y}{\mathrm{d}t}\mathrm{d}t = \int 2t\mathrm{d}t = t^2 + c \\
y(0) = 0 = (0)^2 + c \iff c = 0 \\
\therefore y(t) = t^2 \implies t(0.6) = 0.36 \\
\implies \varepsilon = |y_3 - y(0.6)| = 0.12
\end{gathered}$$

This method becomes more accurate as Δt becomes smaller, but this means that a larger number of steps are required to find the numerical solution. We could use MATLAB, for instance, to solve the previous example, but with $\Delta = 0.05$,

```
t = 0:0.05:0.6;
y(1) = 0;
for i = 1:12
    y(i+1) = y(i) + 2*t(i)*0.05;
end
disp(y(13));
plot(t, y, '-');
```

Output: 0.3300 $\implies \varepsilon = |y_{13} - y(0.6)| = 0.03$ which is clearly a more accurate approximation!

Example 8.3.2

Consider the IVP

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \sin(ty), \qquad y(0) = 1$$

Write MATLAB code to esitmate $y(2)$, using Euler's Method with step size $\Delta t = 0.01$.

```
Dt = 0.1
t = 0:Dt:2;
y(1) = 1;
for i = 1:length(t)-1
    y(i+1) = y(i) + sin(t(i)*y(i))*Dt;
end
y(end)
```

Output: 1.8243

A particular issue can occur if the Δt is too large. For instance, after a step, the approximation might have jumped over the equilibrium solution, which is not analytically correct. The whole calculation would need to be thrown out or adjusted at that point.

Chapter 9

Week 9

9.1 Lecture 24

Heun's Method - An Extension of Euler's Method

We can think about Euler's method in a different way. Instead of approximating the derivative from first principles, let us integrate the ODE from t_n to t_{n+1}

$$\text{Suppose } y - \phi(t), \text{ is a solution to the IVP} \begin{cases} y' = f(t, y) \\ y(t_0) = y_0 \end{cases}$$

$$\int_{t_n}^{t_{n+1}} \frac{\mathrm{d}\phi}{\mathrm{d}t} \mathrm{d}t = \int_{t_n}^{t_{n+1}} f(t, \phi(t)) \mathrm{d}t$$

Geometerically, this integral can be represented as the area under the curve $f(t, \phi(t)))$, and this area can be approximated with triangles of width $\Delta t = t_{n+1} - t_n$ and height $f(t_n, \phi(t_n))$,

$$\phi(t_{n+1}) = \phi(t_n) + f(t_n, \phi(t_n))\Delta t \qquad \text{or} \qquad y_{n+1} = y_n + f(t_n, y_n)\Delta t$$

Fundamentally though, we're using a rectangular approximation of the area, and better area approximating shapes are availible, trapeziums, for example. In that case, we replace the intergral with the average of the two endpoints:

$$\phi(t_{n+1}) = \phi(t_n) + \frac{\Delta t}{2} \left(f(t_n, \phi(t_n)) + f(t_{n+1}, \phi(t_{n+1}))\right)$$

As before, replacing ϕ with y leads to

$$y_{n+1} = y_n + \frac{\Delta t}{2} \left(f(t_n, y_n) + f(t_{n+1}, y_{n+1})\right)$$

But clearly, this is problematic, since our formula for y_{n+1} is dependent on itself. To get around this, we can use Euler's method to approximate y_{n+1}.

$$y_{n+1} = y_n + \frac{\Delta t}{2} \left(f(t_n, y_n) + f(t_{n+1}, \hat{y}_{n+1})\right)$$

where, $\hat{y}_{n+1} = y_n + f(t_n, y_n)\Delta t$. This is called Heun's method.

Seperable First-Order ODEs

This is one of several classes of ODEs we will study. It is very important to be skilled at identifying and solving this type of ODE. A first order ODE is seperable, if it can be written

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(x)g(y)$$

- $\frac{\mathrm{d}y}{\mathrm{d}x} = y(1-y)$: **Seperable**. $f(x) = 1,\ g(y) = y(1-y)$.

- $\frac{\mathrm{d}y}{\mathrm{d}x} = \exp(x+y)$: **Seperable**. $f(x) = \exp(x),\ g(y) = \exp(y)$.
- $y' = \exp(x+y)^2$: **Unseperable**. $y' = \exp(x^2 + 2xy + y^2)$.
- $\dot{y} = \frac{ty+y}{t^2}$. **Seperable**. $f(t) = \frac{t+1}{t},\ g(y) = y$
- $\frac{\mathrm{d}y}{\mathrm{d}t} = ty + y^2$: **Unseperable**. $\dot{y} = (t+y)y$

We can solve seperable ODEs following these steps:

0. Determine that the ODE is seperable; can be written $\dfrac{\mathrm{d}y}{\mathrm{d}x} = f(x)g(y)$.
1. Rewrite the expression as $\dfrac{1}{g(y)}\dfrac{\mathrm{d}y}{\mathrm{d}x} = f(x)$
2. Integrate both sides with respect to x: $\displaystyle\int \frac{1}{g(y)}\frac{\mathrm{d}y}{\mathrm{d}x}\mathrm{d}x = \int f(x)\mathrm{d}x$
3. Note that the integral on the LHS is a substitution and hence replace it with $\displaystyle\int \frac{\mathrm{d}y}{g(y)} = \int f(x)\mathrm{d}x$
4. If we're lucky, one or both sides can actually be evaluated.
5. If we're very lucky, we can explicitly express y as a function of x.

Singular Solutions

In step 1, we rewrite $y' = f(x)g(y)$ as

$$\frac{1}{g(y)}\frac{\mathrm{d}y}{\mathrm{d}x} = f(x)$$

We can't divide by 0, so if $g(y) = 0$ we have a problem.

If there is such an a such that $g(a) = 0$, then the ODE will have an equilibrium solution at $y(x) = a$. This is easy to check:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(x)g(y), \qquad \text{at a } \frac{\mathrm{d}y}{\mathrm{d}x} = g(0)f(x) = 0 \cdot f(x) = 0$$

Such a solution is called a singular solution, because it won't generally arise from an earlier step. We need to always check for singular/equilibrium solutions.

9.2 Lecture 25

Linear First-Order ODEs

If you cannot solve a differential equation by seperation, the next plan of attack should be to test if it is linear and use integration factor method.

If a first order DE can be written in form

$$\frac{\mathrm{d}y}{\mathrm{d}t} + p(t)y = q(t)$$

it is linear.

- $y' = x$: **Linear**. $p(x) = 0,\ q(x) = x$
- $y' = \log x$: **Linear**. $p(x) = 0,\ q(x) = \log x$
- $y' = y^2$: **Non-linear**.
- $y' = y$: **Linear**. $p(x) = 1,\ q(x) = 0$
- $y'y = 1$: **Non-linear**.

- $y' = y + 3\sin x$: **Linear**. $p(x) = -1,\ q(x) = 3\sin x$
- $y' = t^5 - t^2 y$: **Linear**. $p(x) = t^2,\ q(x) = t^5$
- $y' = \sin y + e^x$ **Non-linear**.
- $(y')^2 = y$: **Non-linear**.

In general, let $\mathfrak{L} = \dfrac{\mathrm{d}}{\mathrm{d}t} + p(t)$ be a differential operator so that applying $\mathfrak{L}[y],\ y \in C^1(\mathbb{R})$

$$\mathfrak{L}[y] = \frac{\mathrm{d}}{\mathrm{d}t} y + p(t)y$$

Definition (Linear). $\forall \alpha, \beta \in \mathbb{R},\ y_1, y_2 \in C^1(\mathbb{R}),\ \mathfrak{L}[\alpha y_1 + \beta y_2] = \alpha \mathfrak{L}[y_1] + \beta \mathfrak{L}[y_2]$.
Claim. $\mathfrak{L}$ is linear.

Proof.

$$\begin{aligned}
\mathfrak{L}[\alpha y_1 + \beta y_2] &= \frac{\mathrm{d}}{\mathrm{d}t}(\alpha y_1 + \beta y_2) + p(t)(\alpha y_1 + \beta y_2) \\
&= \alpha \frac{\mathrm{d}}{\mathrm{d}t} y_1 + \beta \frac{\mathrm{d}}{\mathrm{d}t} y_2 + \alpha p(t) y_1 + \beta p(t) y_2 \\
&= \alpha \frac{\mathrm{d}}{\mathrm{d}t} y_1 + \alpha p(t) y_1 + \beta \frac{\mathrm{d}}{\mathrm{d}t} y_2 + \beta p(t) y_2 \\
&= \alpha \left(\frac{\mathrm{d}}{\mathrm{d}t} y_1 + p(t) y_1 \right) + \beta \left(\frac{\mathrm{d}}{\mathrm{d}t} y_2 + p(t) y_2 \right) \\
&= \alpha \mathfrak{L}[y_1] + \beta \mathfrak{L}[y_2]
\end{aligned}$$

□

Now, recall the product rule for differentiation, written here in reverse

$$f \frac{\mathrm{d}g}{\mathrm{d}t} + \frac{\mathrm{d}f}{\mathrm{d}t} g = \frac{\mathrm{d}}{\mathrm{d}t}(fg)$$

This formula is the key to solving first order linear ODEs

Example 9.2.1
Solve $\dfrac{\mathrm{d}y}{\mathrm{d}t} + \left(\dfrac{2t}{t^2+1} \right) y = t^2 + 1$

$$\text{This ODE is linear, } p(t) = \frac{2t}{t^2+1},\ q(t) = t^2 + 1$$

Multiply through by $(t^2 + 1)$. This is called the intergrating factor, for future discussion.

$$(t^2+1)\frac{\mathrm{d}y}{\mathrm{d}t} + 2ty = (t^2+1)^2$$

This is just like the product rule! $f = t^2 + 1,\ \dfrac{\mathrm{d}f}{\mathrm{d}t} = 2t,\ g = y,\ \dfrac{\mathrm{d}g}{\mathrm{d}t} = \dfrac{\mathrm{d}y}{\mathrm{d}t}$. So we can replace the entire LHS with the RHS of the product rule, as we've written it above.

$$\frac{\mathrm{d}}{\mathrm{d}t}\big((t^2+1)y\big) = (t^2+1)^2$$

And now we can integrate both sides with respect to t and cancel it on the LHS

$$\begin{aligned}
\int \frac{\mathrm{d}}{\mathrm{d}t}\big((t^2+1)y\big)\,\mathrm{d}t &= \int t^4 + 2t^2 + 1\ \mathrm{d}t \\
\iff (t^2+1)y + c_2 &= \frac{t^5}{5} + \frac{2t^3}{3} + t + c_1 \\
\therefore y = \frac{\frac{t^5}{5} + \frac{2t^3}{3} + t + c}{t^2+1}&, \quad c = c_1 - c_2 \in \mathbb{R}
\end{aligned}$$

The Intergrating Factor

Can we always multiply a first-order linear ODE by some function, such that the LHS matches the product rule?

Yes.

Given the ODE

$$\frac{\mathrm{d}y}{\mathrm{d}t} + p(t)y = q(t)$$

we can multiply through by the yet-to-be-found intergrating factor $I(t)$:

$$I(t)\frac{\mathrm{d}y}{\mathrm{d}t} + I(t)p(t)y = I(t)q(t)$$

We can think of I as g, and y as f, in the product rule above. Then the first term is gf' and consequently we want the second term to be $fg' = yI'$. Since this term equals Ipy, we can infer that I must satisfy the ODE

$$I'(t) = I(t)p(t)$$

But this is a seperable ODE, which we already know how to solve!

$$I(t) = \exp\left(\int p(t)\mathrm{d}t\right)$$

9.3 Lecture 26

Applications of First-Order ODEs

Newton's Law of Cooling

Newton's law of cooling states that the rate at which a body cools is proportional to the temperature difference between the body and its surrounding medium.

If T is the temperature of the body and T_m is the temperature of the surrounding medium, then, according to Newton's law

$$\frac{\mathrm{d}T}{\mathrm{d}t} = -k(T - T_m), \qquad k > 0$$

Here, the constant k is chosen such that if $T > T_m,\ T'(t)$ is negative, describing cooling.

Electrical Engineering

Chapter 10

Week 10

10.1 Lecture 27

Single-Species Population Models

Malthusian Model

Logistic Model

Malthusian models problematically does not have “damping factor.” That is, the population will grow without bound, which doesn’t reflect real life, where effects like overpopulation and limited food supply are contributing factors to population.

$$\frac{\mathrm{d}P}{\mathrm{d}t} = rP\left(1 - \frac{P}{\theta}\right)$$

where θ is some large constant known as the carrying capacity. Note, $\theta >> 1$, otherwise

$$\frac{\mathrm{d}P}{\mathrm{d}t} \approx rP$$

which is the same Malthusian model as before. As population grows and approaches the carrying capacity, the then the growth comes to a standstill. If we introduce new varaibles $y = P/\theta$ and $\tau = rt$. Then the ODE becomes

$$\frac{\mathrm{d}y}{\mathrm{d}\tau} = y(1 - y)$$

which we can easily solve, and find

$$y(\tau) = \frac{y_0 e^{\tau}}{1 - y_0 + y_0 e^{\tau}}$$

and finally, we we let $P(0) = P_0$ and substitue back the original variables, we see

$$P(t) = \frac{\theta P_0 e^{rt}}{\theta - P_0 + P_0 e^{rt}}$$

The model will have equilibria. Assuming $r > 0$,

$$\frac{\mathrm{d}P}{\mathrm{d}t} = rP\left(1 - \frac{P}{\theta}\right) \implies \begin{cases} p = \theta \\ p = 0 \end{cases} \text{ are equilibria}$$

We can then analyse the stability,

$$\begin{aligned} P > \theta &\implies \tfrac{\mathrm{d}P}{\mathrm{d}t} < 0 \\ 0 < P < \theta &\implies \tfrac{\mathrm{d}P}{\mathrm{d}t} > 0 \\ P < 0 &\implies \tfrac{\mathrm{d}P}{\mathrm{d}t} < 0 \end{aligned}$$

So we can determine that the equilibrium at $P = \theta$ is stable, and the equilibrium at $P = 0$ is unstable.

Multi-Species Models

Predator-Prey System

Consider an isolated system where the prey (say rabbits, denoted $R(t)$) have an infinite food supply and the predators (wolves, denoted $W(t)$) feed on the prey. Assume

- In the absence of predators, $\dfrac{\mathrm{d}R}{\mathrm{d}t} = \alpha R, \quad \alpha > 0$
- In the absence of prey, $\dfrac{\mathrm{d}W}{\mathrm{d}t} = -\beta W, \quad \beta > 0$
- With both spieces present, the principal cause of death in prey is being eaten by predators, and the birth and death rates of predators depends on the prey.
- We assume that the interaction is at a rate proportinoal to both populations.

Under these assumptions, the system evolves according to

$$\begin{cases} \dfrac{\mathrm{d}R}{\mathrm{d}t} = \alpha R - \gamma RW = \gamma R\left(\dfrac{\alpha}{\gamma} - W\right) \\[2ex] \dfrac{\mathrm{d}W}{\mathrm{d}t} = -\beta R + \delta RW = -\delta W\left(\dfrac{\beta}{\delta} - R\right) = \delta W\left(R - \dfrac{\beta}{\gamma}\right) \end{cases}$$

These are called the Lotka-Volterra equations and are an example of a system of coupled ODEs (in this case, first-order, non-linear). We can evaluate the equilibria of the system:

$$\frac{\mathrm{d}R}{\mathrm{d}t} = 0, \text{ if} \qquad R = 0, \quad W = \frac{\alpha}{\gamma}$$

$$\left.\begin{aligned} \left.\frac{\mathrm{d}W}{\mathrm{d}t}\right|_{R=0} &= 0 \qquad \implies W = 0 \\ \left.\frac{\mathrm{d}W}{\mathrm{d}t}\right|_{W=\alpha/\gamma} &= 0 \quad \implies R = \frac{\beta}{\delta} \end{aligned}\right\} \quad \begin{gathered}\text{Equilibria:}\\ (R, W) = (0,0) \text{ and } \left(\tfrac{\beta}{\delta}, \tfrac{\alpha}{\gamma}\right)\end{gathered}$$

We can plot these on an R–W plane called the phase plane, and curves on the plane are called phase trajectories

$$\frac{\mathrm{d}W}{\mathrm{d}R} = \frac{\dfrac{\mathrm{d}W}{\mathrm{d}t}}{\dfrac{\mathrm{d}R}{\mathrm{d}t}} = \frac{-\beta W + \delta RW}{\alpha R - \gamma RW}$$

Competing Species

Consider two populations $P_1(t)$ and $P_2(t)$. Let's initally suppose they are kept separate from one another.

$$\begin{aligned} \frac{\mathrm{d}P_1}{\mathrm{d}t} &= P_1(\alpha_1 - \beta_1 P_1) \\ \frac{\mathrm{d}P_2}{\mathrm{d}t} &= P_2(\alpha_2 - \beta_2 P_2) \end{aligned} \qquad \begin{gathered}\text{Populations competing}\\ \text{with themselves}\\ \text{Uncoupled ODEs}\end{gathered}$$

These populations compete only with themselves, which is just the logistic model we introduced previously. Now let's suppose we bring these populations, so that they compete for shared resrouces. This is reflected by introducing a growth parameter $\gamma_{ij} > 1$ which expresses how population j affects population i. Now

$$\frac{\mathrm{d}P_1}{\mathrm{d}t} = P_1(\alpha_1 - \beta_1 P_1 - \gamma_{12} P_2)$$

$$\frac{\mathrm{d}P_2}{\mathrm{d}t} = P_2(\alpha_2 - \beta_2 P_2 - \gamma_{21} P_1)$$

is a coupled system of ODEs.

10.2 Lecture 28

Second-Order Differential Equations

Just trying to figure out why my preamble is broken.