MATH2001
Calculus & Linear Algebra II

Michael Kasumagic, s4430266

Summer Semester, 2024-25

Contents

Chapter 1

Week 1

1.1 Lecture 1

In this course we will cover four major topics:

- Ordinary Differential Equations
- Linear Algebra
- Vector Calculus
- Integral Calculus

1.1.1 Solutions to First Order ODEs

We are comfortable solving three types of first order ODEs by now:

- Directly integrable: $\frac{\mathrm{d}y}{\mathrm{d}x} = f(x)$

$$y(x) = \int f(x)\mathrm{d}x = F(x) + C$$

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

$$\frac{1}{g(y)}\frac{\mathrm{d}y}{\mathrm{d}x} = f(x) \iff \int \frac{1}{g(y)}\frac{\mathrm{d}y}{\cancel{\mathrm{d}x}}\cancel{\mathrm{d}x} = \int f(x)\mathrm{d}x \iff G(y(x)) = F(x) + C$$

If G is invertible, then $y(x) = G^{-1}(F(x) + C)$

- Linear: $\frac{\mathrm{d}y}{\mathrm{d}x} = q(x) - p(x)y$

$$\text{Let } \mu = \exp\left(\int p(x)\mathrm{d}x\right) \implies \mu\frac{\mathrm{d}y}{\mathrm{d}x} + \mu p(x)y = \mu q(x) \iff \frac{\mathrm{d}}{\mathrm{d}x}(\mu y) = \mu q(x) \iff y(x) = \frac{1}{\mu(x)}\int \mu q(x)\mathrm{d}x$$

In many applications, we need to solve an IVP. In general this is an equation of form,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(x, y), \quad y(x_0) = y_0.$$

In other words, we seek to find solutions to the ODE which pass through the point (x_0, y_0) in the $x-y$ plane.

Example 1.1.1

$$\frac{\mathrm{d}y}{\mathrm{d}x} = x, \quad y(0) = 1 \text{ has a unique solution:}$$

$$\begin{aligned}
\frac{\mathrm{d}y}{\mathrm{d}x} &= x \\
y(x) &= \frac{1}{2}x^2 + C \\
\text{Impose } y(0) &= 1 \\
\therefore 1 &= \frac{1}{2}(0)^2 + C \\
\therefore C &= 1 \\
\therefore y(x) &= \frac{1}{2}x^2 + 1
\end{aligned}$$

Example 1.1.2

$$\frac{\mathrm{d}y}{\mathrm{d}x} = 3xy^{1/3}, \quad y(0) = 0 \text{ has more than one solution:}$$

$$\begin{aligned}
y^{-1/3}\frac{\mathrm{d}y}{\mathrm{d}x} &= 3x \\
\int y^{-1/3}\frac{\mathrm{d}y}{\mathrm{d}x}\mathrm{d}x &= \int 3x\mathrm{d}x \\
\int y^{-1/3}\mathrm{d}y &= \int 3x\mathrm{d}x \\
\frac{3}{2}y^{2/3} + C_1 &= \frac{3}{2}x^2 + C_2 \\
y^{2/3} &= x^2 + C \\
\text{Impose } y(0) &= 0 \\
0^{2/3} &= 0^2 + C \\
\implies C &= 0 \\
\therefore y^{2/3} &= x^2 \\
\therefore y &= \pm x^3
\end{aligned}$$

This is problematic. Our inital value constraint hasn't allowed us to pick one particular solution.

Note:-

The previous IVP has multiple solutions because $f(x, y) = 3xy^{1/3}$ is not differentiable at $y = 0$.

Example 1.1.3

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{x - y}{x}, \quad y(0) = 1 \text{ has no solutions:}$$

$$\begin{aligned}
\frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{x}{x} - \frac{1}{x}y \\
&= q(x) - p(x)y \\
\frac{\mathrm{d}y}{\mathrm{d}x} + p(x)y = 1
\end{aligned}$$

$$
\begin{aligned}
\mu &= \exp\left(\int p(x)\mathrm{d}x\right) \\
&= \exp\left(\int \frac{1}{x}\mathrm{d}x\right) \\
&= \exp(\ln(x)) \\
&= x
\end{aligned}
$$

$$
\begin{aligned}
\mu\frac{\mathrm{d}y}{\mathrm{d}x} + \mu p(x)y &= \mu \\
x\frac{\mathrm{d}y}{\mathrm{d}x} + y &= x \\
\frac{\mathrm{d}}{\mathrm{d}x}(xy) &= x \\
xy &= \int x\mathrm{d}x \\
&= \frac{1}{2}x^2 + C
\end{aligned}
$$

Impose $y(0) = 1$

$$
\begin{aligned}
\therefore 0 \cdot 1 &= \frac{1}{2}(0)^2 + C \\
C &= 0 \\
\therefore y(x) &= \frac{1}{2}x
\end{aligned}
$$

However, our general solution **does not** satisfy our inital value constraint, $y(0) = \frac{1}{2}(0) = 0 \neq 1$.

Note:-

Our IVP doesn't have a solution because $f(x, y) = \frac{x-y}{x}$ is not differentiable or continuous around $x = 0$.

We're kind of loosley referring to "existence and uniqueness" theorems, or Picard-Lindelöf Theorem, which generally states:

$$\text{The IVP } \frac{\mathrm{d}y}{\mathrm{d}x} = f(x, y) \quad y(x_0) = y_0$$

has a unique solution around x_0 if:

1. $f(x, y)$ is continuous around (x_0, y_0)
2. $f(x, y)$ is differentiable with respect to y around (x_0, y_0), ie $\frac{\partial f}{\partial y}$ is continuous aorund (x_0, y_0).

1.1.2 Existence and Uniqueness

Theorem 1.1.1

Consider the IVP

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(x, y), \quad y(x_0) = y_0$$

We are concerned with the conditions under which a solution exists and is unique.

1. (existence, Peano's Theorem) If $f(x, y)$ is continuous in some rectangle

$$R = \{(x, y) \mid |x - x_0| < a,\ |y - y_0| < b\}$$

then the IVP has at least one solution.

2. (uniqueness, Picard's Theorem) If $f_y(x, y) := \frac{\partial f}{\partial y}$ is also continuous in R then there is some interval $|x - x_0| \le h \le a$ which contains at least one solution.

This result only tells us that a solution exists or is unique locally. Beyond R, we simply don't know.

Example 1.1.4

$$\frac{\mathrm{d}y}{\mathrm{d}x} = x, \quad y(0) = 1$$

$$f(x, y) = x$$
$$f_y(x, y) = 0$$

These functions are both continuous over $\mathbb{R}^2$. Therefore there exists a unique solution

$$\frac{\mathrm{d}y}{\mathrm{d}x} = 3xy^{1/3}, \quad y(0) = 0 \text{ has more than one solution:}$$

$$f(x, y) = 3xy^{1/3}$$
$$f_y(x, y) = xy^{-2/3}$$

$f(x, y)$ is continuous over $\mathbb{R}^2$ so there exists at least one solution. However, f_y has a discontinuity at $y = 0$, so there may or may not be unique solutions (remember, its not an iff).

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{x - y}{x}, \quad y(0) = 1 \text{ has no solutions:}$$

$$f(x, y) = \frac{x - y}{x}$$
$$f_y(x, y) = -\frac{1}{x}$$

$f(x, y)$ and f_y both have discontinuities when $x = 0$, so we don't know from this test if there are solutions, or if the solution is unique.

Note:-

These theorems are not if and only if's. They can fail. For example, take the IVP

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{3}x^{-2/3}, \quad y(0) = 1/$$

We see

$$f(x, y) = \frac{1}{3}x^{-2/3}, \qquad f_y(x, y) = 0$$

f has a discontinuity when $x = 0$, so the theorem's fail to identify if this IVP has solutions. However, this

IVP **does** have a unique solution,

$$y(x) = x^{1/3} + 1,$$

so we need to be careful we're using these theorems correctly. **If** f and f_y are continuous in some reigon **then** there exists a unique solution in that reigon.

Example 1.1.5

Solve these:

1. $y' = y^{2/3}, \quad y(0) = 1$

$$f(x,y) = y^{2/3}$$
$$f_y(x,y) = \frac{2}{3}y^{-1/3}$$

Therefore there exist at least one solution to the IVP.

$$\begin{aligned} y^{-2/3}y' &= 1 \\ \int y^{-2/3}\mathrm{d}y &= \int 1\mathrm{d}x \\ 3y^{1/3} &= x + C \\ y^{1/3} &= \frac{1}{3}(x+C) \\ y &= \frac{1}{27}(x+C)^3 \\ \text{Impose } y(0) = 1 \end{aligned}$$

Imposing the IVP and expanding the cubic expression, will reveal 3 values for C, the nicest of which is 3. The one which satisfies our IVP is

$$y(x) = \frac{1}{27}(x+3)^3$$

Even though those other solutions exist, only one satisfies the IVP, hence this solution is unique.

2. $y' = (3x^2 + 4x + 2)/(2y - 2), \quad y(0) = 1$

$$f(x,y) = \frac{3x^2 + 4x + 2}{2(y-1)}$$

Because of the discontinuity at $y = 1$, our exisitence theorem fails to identify if solutions exist.

$$\begin{aligned} y' &= \frac{3x^2 + 4x + 2}{2(y-1)} \\ 2(y-1)y' &= 3x^2 + 4x + 2 \\ 2\int y - 1\mathrm{d}y &= \int 3x^2 + 4x + 2 \,\mathrm{d}x \\ y^2 - 2y &= x^3 + 2x^2 + 2x + C \\ y^2 - 2y + 1 &= x^3 + 2x^2 + 2x + C + 1 \\ (y-1)^2 &= x^3 + 2x^2 + 2x + C + 1 \\ \text{Impose } y(0) = 1 \\ ((1) - 1)^2 &= (0)^3 + 2(0)^2 + 2(0) + C + 1 \\ \iff C &= -1 \\ \therefore (y-1)^2 &= x^3 + 2x^2 + 2x \\ \therefore y(x) &= 1 \pm \sqrt{x^3 + 2x^2 + 2x} \end{aligned}$$

The IVP has two solutions.

1.1.3 Method of Succesive Approximations

To start, we note that it is always possible to apply a variable shift and so that the IVP is expressed:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(x,y), \qquad y(0) = 0$$

Example 1.1.6

$$y' = 2(x-1)(y-1), y(1) = 2$$

$$\begin{aligned} \text{Let } \bar{x} &= x - 1 \\ \text{Let } \bar{y} &= y - 2 \\ \text{So } \frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{\mathrm{d}\bar{y}}{\mathrm{d}\bar{x}} \\ \implies \frac{\mathrm{d}\bar{y}}{\mathrm{d}\bar{x}} &= 2\bar{x}(\bar{y}+1),\ \bar{y}(0) = 0 \end{aligned}$$

Without loss of generality we will consider this problem where the inital point is at the origin. We can restate the previous theorem 1.1.1 as follows

Theorem 1.1.2

If f and f_y are continuous in some rectangle

$$R = \{(x,y) \mid |x| \leq a,\ |y| \leq b\},$$

then there is some interval $|x| \leq h \leq a$ which contains a uniuqe solution $y = \phi(x)$ of the IVP

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(x,y), \qquad y(0) = 0$$

Equivilance with integral equation

Let $y = \phi(x)$ be the solution to the IVP

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(x,y), \qquad y(0) = 0, \tag{1}$$

and note that the function $F(x) = f(x, \phi(x))$ is a continuous function of x only. We then have

$$\phi(x) = \int_0^x F(t)\mathrm{d}t = \int_0^x f(t, \phi(t))\mathrm{d}t. \tag{2}$$

Note that $\phi(0) = 0$. This is an example of an *integral equation*. Conversely, let $\phi(x)$ satisfy the integral equation (2). By the Fundamental Theorem of Integral Calculus, $\phi'(x) = f(x, \phi(x))$, which implies that $y = \phi(x)$ is a solution of the IVP (1). In other words, the IVP (1) and the integral equation (2) are equivalent, meaning that a solution of one is a solution of the other. Herein we work with (2).

Method of succesive approximations

The goal of the approach is to generate a sequence of functions $\phi_0, \phi_1, \ldots, \phi_n, \ldots$. Starting with the initial function $\phi_0(x) = 0$ (satisfying the initial condition of (1)), the sequence is generated iteratively by

$$\phi_{n+1}(x) = \int_0^x f(t, \phi_n(t))\mathrm{d}t. \tag{3}$$

Note that each ϕ_n satisfies $\phi_n(0) = 0$, but generally not the integral equation (2) itself. However, if there is a k, such that $\phi_{k+1}(x) = \phi_k(x)$, then $\phi_k(x)$ is a solution of the integral equation (2) and hence the IVP (1). Generally this does not occur, but we may instead consider *limit functions*.

There are 4 key points to consider:

1. Do all members of the sequence exist?
2. Does the sequence converge to a limit function ϕ?
3. What are the properties of ϕ?
4. If ϕ satisfies the IVP (1), are there other solutions?

Example 1.1.7

$$y' = 2x(y+1), \qquad y(0) = 0$$

$$\phi_0(x) = 0, \qquad f(x,y) = 2x(y+1)$$
$$\phi_1(x) = \int_0^x f(t, \phi_0(t))\mathrm{d}t = \int_0^x f(t,0)\mathrm{d}t = \int_0^x 2t(0+1)\mathrm{d}t = t^2\Big|_0^x = x^2$$
$$\phi_2(x) = \int_0^x f(t, \phi_1(t))\mathrm{d}t = \int_0^x f(t,t^2)\mathrm{d}t = \int_0^x 2t(t^2+1)\mathrm{d}t = \int_0^x 2t^3 + 2t \ \mathrm{d}t = \frac{1}{2}t^4 + t^2\Big|_0^x = \frac{1}{2}x^4 + x^2$$

Similarly,

$$\phi_3(x) = x^2 + \frac{1}{2}x^4 + \frac{1}{6}x^6$$
$$\phi_4(x) = x^2 + \frac{1}{2}x^4 + \frac{1}{6}x^6 + \frac{1}{24}x^8$$

Proposition.

$$\phi_n(x) = \sum_{i=1}^{n} \frac{1}{i!}x^{2i}$$

proof. By induction: True for $n = 1$. Suppose True for $n = k$.

Then $\phi_{k+1}(x) = \int_0^x f(t, \phi_k(t))\mathrm{d}t = \int_0^x 2t\left(1 + \sum_{i=1}^{k} \frac{1}{i!}t^{2i}\right)\mathrm{d}t = \int_0^x \left(2t + \sum_{i=1}^{k} \frac{2}{i!}t^{2i+1}\right)\mathrm{d}t = t^2 + \sum_{i=2}^{k+1} \frac{1}{i!}t^{2i}\Big|_0^x$

$$\therefore \phi_{k+1} = x^2 + \sum_{i=2}^{k+1} \frac{1}{i!}x^{2i} = \sum_{i=1}^{k+1} \frac{1}{i!}x^{2i}$$

So the proposition is true $\forall n \in \mathbb{N}$. □

$$\lim_{n\to\infty} \phi_n(x) = \lim_{n\to\infty} \sum_{i=1}^{n} \frac{1}{i!}x^{2i} \text{ exists} \iff \text{the series converges}$$

Applying the ratio test between two successive terms, j and $j+1$, as j goes to infinty,

$$\lim_{j\to\infty} \left| \frac{\frac{x^{2j+2}}{(j+1)!}}{\frac{x^{2j}}{j!}} \right| = \lim_{j\to\infty} \left| \frac{x^{2j+2}}{(j+1)!} \cdot \frac{j!}{x^{2j}} \right| = \lim_{j\to\infty} \left| \frac{x^2}{j+1} \right| = 0$$

Therefore, the series converges!
Therefore, the limit, as $n \to \infty$ of ϕ_n exists.

1.1.4 Exact First Order ODEs

Definition 1.1.1: Exact First Order ODE

Recall that if $z = f(x, y)$ is a differentiable function of x and y, where $x = g(t)$ and $y = h(t)$ are both differentiable functions of t, then z is a differentiable function of t, whose derivative is given by the chain rule:

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

Now suppose the equation

$$f(x, y) = C$$

defines y implicitly as a function of x. Then $y = y(x)$ can be show to satisfy a first order ODE obtained by using the chain rule above. In this case, $z = f(x, y(x)) = C$, so,

$$\frac{\mathrm{d}z}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x}C = 0 = \frac{\partial f}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}x} + \frac{\partial f}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}x}$$
$$\implies f_x + f_y y' = 0$$

A first order ODE of form

$$P(x, y) + Q(x, y)\frac{\mathrm{d}y}{\mathrm{d}x} = 0$$

is called exact if there is a function $f(x, y)$ such that

$$f_x(x, y) = P(x, y) \quad \text{and} \quad f_y(x, y) = Q(x, y).$$

The solution is then given implicitly by the equation

$$f(x, y) = C,$$

where C can usually be determined by some intial condition.

Theorem 1.1.3 Test for Exactness

Let $P, Q, \dfrac{\partial P}{\partial y}, \dfrac{\partial Q}{\partial x}$ be continuous over some reigon of interest. Then

$$P(x, y) + Q(x, y)\frac{\mathrm{d}y}{\mathrm{d}x} = 0$$

is an exact ODE if and only if

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$$

everywhere in the reigon

Proof. 1. Prove: ODE is exact $\implies \dfrac{\partial P}{\partial y} = \dfrac{\partial Q}{\partial x}$.

Recall Clairout's Theorem,

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} \text{ if both } f_{xy} \text{ and } f_{yx} \text{ are continuous in the reigon.}$$

$$\text{Suppose ODE is exact} \implies \exists f(x, y) : \frac{\partial f}{\partial x} = P(x, y), \frac{\partial f}{\partial y} = Q(x, y)$$

$$\frac{\partial P}{\partial y} = \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial Q}{\partial x}, \text{ by Clairout's Theorem.}$$

2. Prove: $\dfrac{\partial P}{\partial y} = \dfrac{\partial Q}{\partial x} \implies$ ODE is exact.

Suppose $\dfrac{\partial P}{\partial y} = \dfrac{\partial Q}{\partial x}$. We seek a function f such that $f_x = P, f_y = Q$.

$$\text{Take } f(x,y) = \int_{x_0}^{x} P(x',y)\mathrm{d}x' + \int_{y_0}^{y} Q(x_0,y') + C$$
$$f_x(x,y) = \frac{\partial}{\partial x}\left(\int_{x_0}^{x} P(x',y)\mathrm{d}x' + \cancel{\int_{y_0}^{y} Q(x_0,y')\mathrm{d}y'}\right) = P(x,y)$$
$$f_y(x,y) = \frac{\partial}{\partial y}\left(\cancel{\int_{x_0}^{x} P(x',y)\mathrm{d}x'} + \int_{y_0}^{y} Q(x_0,y')\mathrm{d}y'\right) = Q(x,y)$$

Therefore $P(x,y) + Q(x,y)\dfrac{\mathrm{d}y}{\mathrm{d}x} = 0$ is an exact ODE $\iff \dfrac{\partial P}{\partial y} = \dfrac{\partial Q}{\partial x}$ everywhere in the reigon. $\square$

Example 1.1.8

Solve the ODE $2x + e^y + xe^y y' = 0$

$$\begin{aligned}
P(x,y) &= 2x + e^y \\
\frac{\partial P}{\partial y} &= e^y \\
Q(x,y) &= xe^y \\
\frac{\partial Q}{\partial x} &= e^y \\
\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x} &\Rightarrow \text{ODE is exact} \\
\therefore \exists f(x,y) : f_x(x,y) &= P = 2x + e^y \\
\text{and } f_y(x,y) &= Q = xe^y \\
\implies f = \int P\mathrm{d}x &= \int 2x + e^y \mathrm{d}x \\
&= x^2 + xe^y + g(y) \\
\implies f_y(x,y) = xe^y &= \frac{\partial}{\partial y}\left(x^2 + xe^y + g(y)\right) \\
xe^y &= xe^y + \frac{\mathrm{d}g}{\mathrm{d}y} \\
\implies \frac{\mathrm{d}g}{\mathrm{d}y} &= 0 \\
\therefore f(x,y) &= x^2 + xe^y + C
\end{aligned}$$

All solutions to ODE: $f(x,y) = k$.

$$\iff x^2 + xe^y = k' \qquad (k' = k - C)$$
$$\iff y = \ln\left(\frac{k' - x^2}{x}\right)$$

1.2 Lecture 2

1.2.1 Almost Exact ODEs and Integrating Factors

Suppose we have

$$P(x,y) + Q(x,y)\frac{\mathrm{d}y}{\mathrm{d}x} = 0$$

and

$$\frac{\partial P}{\partial y} \neq \frac{\partial Q}{\partial x}.$$

This is not an exact ODE, but can we do anything with it anyway? Let's consider an "integrating factor" (not to be confused with integrating factors used when solving linear ODEs).

The general idea though, is to multiple the ODE by some function, $h(x, y)$ such that the resulting ODE

$$h(x,y)P(x,y) + h(x,y)Q(x,y)\frac{\mathrm{d}y}{\mathrm{d}x} = 0$$

is exact. We know that this new ODE is exact if and only if

$$\frac{\partial}{\partial y}(hP) = \frac{\partial}{\partial x}(hQ)$$

Let's find a h which accomplishes this:

$$\frac{\partial}{\partial y}(hP) = \frac{\partial}{\partial x}(hQ)$$
$$h_y P + hP_y = h_x Q + hQ_x \iff h_y P - h_x Q + h(P_y - Q_x) = 0$$

Solving for h in general requires us to solve this first order partial differential equation, which is very nasty and also outside the scope of this course.

However, we can consider a simplier case, where h is a function of one of the varaibles x or y. Let's try $h = h(x)$:

$$\frac{\mathrm{d}h}{\mathrm{d}x} = h(x)\frac{P_y - Q_x}{Q} = h\hat{f}$$

Suppose $\hat{f}$ is a function of one variable, x. Then we are left with a first order seperable ODE which we can solve. Once we've solved for h, we can find $f(x, y)$ that solves the exact ODE we wanted to solve.

This is not a great technique, it often doesn't work and requires a lot of trial an error. For example, if $h = h(x)$ didn't yield an approriate $f(x, y)$, we could try $h = h(y)$ or $h = h(x) + h(y)$, which would also give us a seperable ODE to solve. If this technique does work, it hints to some underlying symmetry in the differential system we're solving.

Example 1.2.1

$$\text{Solve } (3xy + y^2) + (x^2 + xy)\frac{\mathrm{d}y}{\mathrm{d}x} = 0$$

$$P(x,y) = 3xy + y^2 \qquad Q(x,y) = x^2 + xy$$
$$\frac{\partial P}{\partial y} = 3x + 2y \neq 2x + y = \frac{\partial Q}{\partial x}$$

So, this ODE is not exact. Can we multiply through by some integrating factor, h?

$$\text{Take } h = h(x) \neq 0$$
$$h(3xy + y^2) + h(x^2 + xy)\frac{\mathrm{d}y}{\mathrm{d}x} = \hat{p} + \hat{q}\frac{\mathrm{d}y}{\mathrm{d}x} = 0$$

is exact

$$\iff \frac{\partial \hat{p}}{\partial y} = \frac{\partial \hat{q}}{\partial x} \iff h(3x + 2y) = h_x(x^2 + xy) + h(2x + y) \iff h(x + y) = h_x x(x + y)$$

Supposing that $x + y \neq 0$, we can simplify and find

$$h = h'x$$

Supposing that $x \neq 0$, we can see

$$h' = \frac{1}{x}h$$

This is a seperable first order ODE we can simply solve,

$$\int \frac{1}{h}\frac{\mathrm{d}h}{\mathrm{d}x}\mathrm{d}x = \int \frac{1}{x}\mathrm{d}x \iff \int \frac{1}{h}\mathrm{d}h = \int \frac{1}{x}\mathrm{d}x \iff \ln|h| = \ln|x| + \hat{\alpha} \iff h(x) = \alpha x,\ \alpha = \exp(\hat{\alpha}).$$

We're free to choose $\alpha > 0$, so we'll take $\alpha = 1$ for simplicity, and then multiple our original ODE by our integrating factor $h = h(x) = x$. Check:

$$h(3xy + y^2) + h(x^2 + xy)\frac{\mathrm{d}y}{\mathrm{d}x} = 0 \iff x(3xy + y^2) + x(x^2 + xy)\frac{\mathrm{d}y}{\mathrm{d}x} = 0$$

$$\iff (3x^2y + xy^2) + (x^3 + x^2y)\frac{\mathrm{d}y}{\mathrm{d}x} = 0,\ P(x,y) = 3x^2y + xy^2,\ Q(x,y) = x^3 + x^2y$$

$$\frac{\partial P}{\partial y} = 3x^2 + 2xy = 3x^2 + 2xy = \frac{\partial Q}{\partial x}$$

So this ODE is exact. Therefore, there exists some functoin, $f(x,y)$ such that $f_x = P$ and $f_y = Q$

$$\text{Take } f(x,y) = x^3y + \frac{1}{2}x^2y^2 \implies f_x = 3x^2y + xy^2 = P,\ f_y = x^3 + x^2y$$

Therefore, the solution to our ODE is

$$f(x,y) = K \iff x^3y + \frac{1}{2}x^2y^2 = K \iff \frac{1}{2}x^2y^2 + x^3y - K = 0 \iff y = \frac{-x^2 \pm \sqrt{x^3 + 2K}}{x}$$

Purely for fun, we're going to apply an inital condition, $y(1) = 0$

$$\text{Then } 0 = \frac{-1 \pm \sqrt{1 + 2K}}{1} \iff 0 = K \text{ and we choose the positive branch}$$

So our final solution is

$$y(x) = \frac{x^2 + \sqrt{x^3}}{x} = \sqrt{x} - x$$

1.2.2 Hyperbolic Functions

Definition 1.2.1: Hyperbolic Functions

$$\cosh(x) = \frac{e^x + e^{-x}}{2},$$
$$\sinh(x) = \frac{e^x - e^{-x}}{2},$$
$$\tanh(x) = \frac{\sinh(x)}{\cosh(x)} = \frac{1 - e^{-2x}}{1 + e^{-2x}},$$

Corollary 1.2.1 Hyperbolic-Pythagorean Identity

$$\cosh^2(x) - \sinh^2(x) = 1$$

This follows from direct calculation.

Note that the Pythagorean identity $\cos^2 x + \sin^2 x = 1$ allows us to paramaterise the unit circle, namely by setting $x(t) = \cos t,\ y(t) = \sin t$, which gives us the equation of a unit circle, $\cos^2 t + \sin^2 t = x^2 + y^2 = 1$.

If instead, we set $x(t) = \cosh t,\ y(t) = \sinh t$, we can see

$$\cosh^2 t - \sinh^2 t = x^2 - y^2 = 1$$

which is the equation for a hyperbola.
Also following from direct calculation, similar to their trigonometric counterparts, the hyperbolic functions satisfy

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x} \cosh x &= \frac{e^x - e^{-x}}{2} = \sinh x, \\ \frac{\mathrm{d}}{\mathrm{d}x} \sinh x &= \frac{e^x + e^{-x}}{2} = \cosh x \end{aligned}$$

Note that $\cosh(0) = 1,\ \cosh(x) \geq 1$ and $\cosh(x)$ is an even function ($\cosh(-x) = \cosh(x)$); $\sinh(0) = 0,\ \sinh(x)$ is an odd function $\sinh(-x) = -\sinh(x)$.

Example 1.2.2

Prove that:

1. $\cosh^2 x = \frac{1}{2}(\cosh(2x) + 1)$

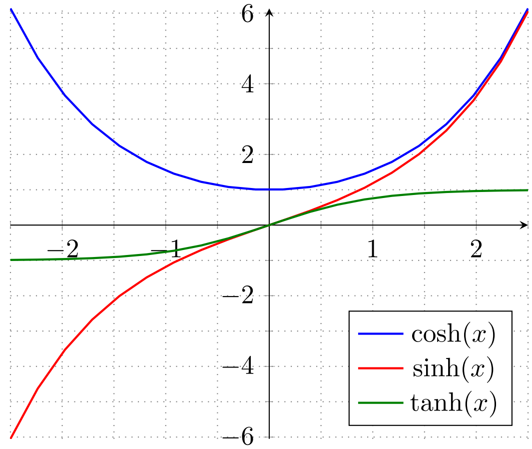
$$\begin{aligned} \cosh^2 x &= \left(\frac{e^x + e^{-x}}{2}\right)^2 \\ &= \frac{e^{2x} + 2e^0 + e^{-2x}}{4} \\ &= \frac{1}{2} \cdot \frac{e^{2x} + e^{-2x} + 2}{2} \\ &= \frac{1}{2}(\cosh 2x + 1) \end{aligned}$$

2. $\sinh^2 x = \frac{1}{2}(\cosh(2x) - 1)$

$$\begin{aligned} \sinh^2 x &= \left(\frac{e^x - e^{-x}}{2}\right)^2 \\ &= \frac{e^{2x} - 2e^0 + e^{-2x}}{4} \\ &= \frac{1}{2} \cdot \frac{e^{2x} + e^{-2x} - 2}{2} \\ &= \frac{1}{2}(\cosh 2x - 1) \end{aligned}$$

3. $\sinh(2x) = 2\cosh x \sinh x$

$$\sinh 2x = \frac{e^{2x} - e^{-2x}}{2} = \frac{(e^x + e^{-x})(e^x - e^{-x})}{2} \cdot \frac{2}{2} = 2 \cdot \frac{e^x + e^{-x}}{2} \cdot \frac{e^x - e^{-x}}{2} = 2\cosh x \sinh x$$



Looking at the plots of the functions, we can deduce that

$$\operatorname{dom} \cosh x = \mathbb{R} \qquad \operatorname{dom} \sinh x = \mathbb{R} \qquad \operatorname{dom} \tanh x = \mathbb{R}$$
$$\operatorname{ran} \cosh x = [1, \infty) \qquad \operatorname{ran} \sinh x = \mathbb{R} \qquad \operatorname{ran} \tanh x = (-1, 1)$$

Definition 1.2.2: Reciprocal Hyperbolic Functions

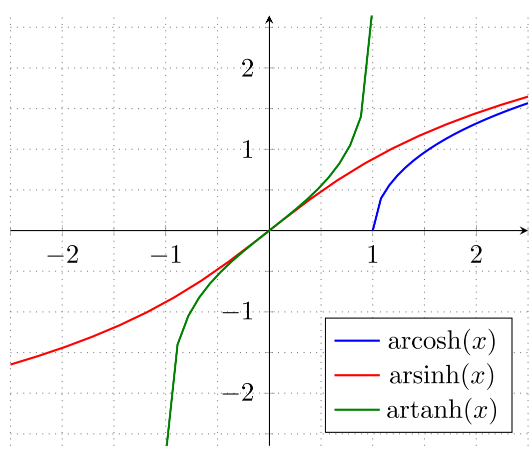
$$\coth(x) = \frac{1}{\tanh(x)} = \frac{\cosh(x)}{\sinh(x)} = \frac{1 + e^{-2x}}{1 - e^{-2x}}$$
$$\operatorname{sech}(x) = \frac{1}{\cosh(x)} = \frac{2}{e^x - e^{-x}}$$
$$\operatorname{csch}(x) = \frac{1}{\operatorname{sech}(x)} = \frac{2}{e^x + e^{-x}}$$

1.2.3 Inverse Hyperbolic Functions

Definition 1.2.3: Inverse Hyperbolic Functions

If f is ... f^{-1} is denoted ...:

$$\cosh(x) \qquad \operatorname{arcosh}(x)$$
$$\sinh(x) \qquad \operatorname{arsinh}(x)$$
$$\tanh(x) \qquad \operatorname{artanh}(x)$$



$$\begin{aligned} \operatorname{dom}\operatorname{arcosh}x &= [1,\infty) & \operatorname{dom}\operatorname{arsinh}x &= \mathbb{R} & \operatorname{dom}\operatorname{artanh}x &= (-1,1) \\ \operatorname{ran}\operatorname{arcosh}x &= [0,\infty) & \operatorname{ran}\operatorname{arsinh}x &= \mathbb{R} & \operatorname{ran}\operatorname{artanh}x &= \mathbb{R} \end{aligned}$$

We have the following:

$$\int \frac{\mathrm{d}x}{\sqrt{1+x^2}} = \operatorname{arsinh}x + C$$
$$\int \frac{\mathrm{d}x}{\sqrt{1-x^2}} = \operatorname{arcosh}x + C,\ \ x > 1$$

Example 1.2.3

Show $\dfrac{\mathrm{d}}{\mathrm{d}x}(\operatorname{arsinh}x) = \dfrac{1}{\sqrt{1+x^2}}$.

$$\begin{aligned} \operatorname{arsinh}x &= y(x) \\ x &= \sinh y \\ \iff \frac{\mathrm{d}}{\mathrm{d}x}(x) &= \frac{\mathrm{d}}{\mathrm{d}x}(\sinh y) \\ \iff 1 &= \frac{\mathrm{d}y}{\mathrm{d}x}\cdot\cosh y \\ \iff \frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{1}{\cosh y} \\ &= \frac{1}{\cosh(\operatorname{arsinh}x)} \\ &= \frac{1}{\sqrt{\cosh^2(\operatorname{arsinh}x)}} \\ &= \frac{1}{\sqrt{1+\sinh^2(\operatorname{arsinh}x)}} \\ &= \frac{1}{\sqrt{1+\sinh(\operatorname{arsinh}x)\sinh(\operatorname{arsinh}x)}} \\ &= \frac{1}{\sqrt{1+x\cdot x}} \\ &= \frac{1}{\sqrt{1+x^2}} \end{aligned}$$

Show $\dfrac{\mathrm{d}}{\mathrm{d}x}(\operatorname{arcosh}x) = \dfrac{1}{\sqrt{x^2-1}}$.

$$\begin{aligned} \operatorname{arcosh}x &= y(x) \\ x &= \cosh y \\ \iff \frac{\mathrm{d}}{\mathrm{d}x}(x) &= \frac{\mathrm{d}}{\mathrm{d}x}(\cosh y) \\ \iff 1 &= \frac{\mathrm{d}y}{\mathrm{d}x}\cdot\sinh y \\ \iff \frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{1}{\sinh y} \\ &= \frac{1}{\sinh(\operatorname{arcosh}x)} \\ &= \frac{1}{\sqrt{\sinh^2(\operatorname{arcosh}x)}} \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{\sqrt{\cosh^2(\operatorname{arcosh} x) - 1}} \\
&= \frac{1}{\sqrt{\cosh(\operatorname{arcosh} x)\cosh(\operatorname{arcosh} x) - 1}} \\
&= \frac{1}{\sqrt{x \cdot x - 1}} \\
&= \frac{1}{\sqrt{x^2 - 1}}
\end{aligned}
$$

Example 1.2.4

Evaluate $\displaystyle\int \frac{\mathrm{d}x}{\sqrt{1+x^2}}$

$$
\begin{aligned}
1 + \sinh^2 t &= \cosh^2 t \\
\text{Let } x &= \sinh t \\
\implies \frac{\mathrm{d}x}{\mathrm{d}t} &= \cosh t \Rightarrow \mathrm{d}x = \cosh t \ \mathrm{d}t \\
\therefore \int \frac{\mathrm{d}x}{\sqrt{1+x^2}} &= \int \frac{\cosh t}{\sqrt{1+\sinh^2 t}} \mathrm{d}t \\
&= \int \frac{\cosh t}{\sqrt{\cosh^2 t}} \mathrm{d}t \\
&= \int \frac{\cosh t}{\cosh t} \mathrm{d}t \\
&= \int \mathrm{d}t \\
&= t + C \\
&= \operatorname{arsinh} x + C
\end{aligned}
$$

Evaluate $\displaystyle\int \frac{\mathrm{d}x}{\sqrt{x^2-1}}$

$$
\begin{aligned}
\cosh^2 t - 1 &= \sinh^2 t \\
\text{Let } x &= \cosh t \\
\implies \frac{\mathrm{d}x}{\mathrm{d}t} &= \sinh t \Rightarrow \mathrm{d}x = \sinh t \ \mathrm{d}t \\
\therefore \int \frac{\mathrm{d}x}{\sqrt{x^2-1}} &= \int \frac{\sinh t}{\sqrt{\cosh^2 t - 1}} \mathrm{d}t \\
&= \int \frac{\sinh t}{\sqrt{\sinh^2 t}} \mathrm{d}t \\
&= \int \frac{\sinh t}{\sinh t} \mathrm{d}t \\
&= \int \mathrm{d}t \\
&= t + C \\
&= \operatorname{arcosh} x + C, \ x \geq 1
\end{aligned}
$$

Example 1.2.5

Show that $\dfrac{\mathrm{d}}{\mathrm{d}x}(\operatorname{artanh} x) = \dfrac{1}{1-x^2}$

$$\begin{aligned}
y &= \operatorname{artanh} x \\
\tanh y &= \tanh \operatorname{artanh} x \\
\frac{\mathrm{d}}{\mathrm{d}x}(\tanh y) &= \frac{\mathrm{d}}{\mathrm{d}x}(x) \\
\frac{\mathrm{d}y}{\mathrm{d}x}(1-\tanh^2 y) &= 1 \\
\frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{1}{1-\tanh^2 y} \\
\frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{1}{1-\tanh^2 \operatorname{artanh} x} \\
&= \frac{1}{1-x^2}
\end{aligned}$$

Using partial fractions, we also find that

$$\int \frac{\mathrm{d}x}{1-x^2} = \frac{1}{2}\ln\left(\frac{1+x}{1-x}\right) + C$$

In fact, we have the following identities

$$\begin{aligned}
\operatorname{artanh} x &= \frac{1}{2}\ln\left(\frac{1+x}{1-x}\right) \\
\operatorname{arsinh} x &= \ln\left(x + \sqrt{x^2+1}\right) \\
\operatorname{arcosh} x &= \ln\left(x + \sqrt{x^2-1}\right)
\end{aligned}$$

Example 1.2.6

Show that $\operatorname{arsinh} x = \ln\left(x + \sqrt{x^2+1}\right)$

$$\begin{aligned}
y &= \operatorname{arsinh} x \\
\sinh y &= x \\
&= \frac{e^y - e^{-y}}{2} \\
2x &= e^y - e^{-y}
\end{aligned}$$

Let $z = e^y > 0$

$$\begin{aligned}
2x &= z - \frac{1}{z} \\
0 &= z^2 - 2xz - 1 \\
\therefore z &= \frac{2x \pm \sqrt{4x^2 - 4(1)(-1)}}{2(1)} \\
&= \frac{2x \pm \sqrt{4x^2+4}}{2} \\
&= \frac{2x \pm \sqrt{4(x^2+1)}}{2} \\
&= \frac{2x \pm 2\sqrt{x^2+1}}{2}
\end{aligned}$$

$$= x \pm \sqrt{x^2+1}$$

Since $z > 0$, we'll pick the positive branch

$$e^y = x + \sqrt{x^2+1}$$
$$y = \ln\left(x + \sqrt{x^2+1}\right)$$
$$\therefore \operatorname{arsinh} x = \ln\left(x + \sqrt{x^2+1}\right)$$

1.2.4 The Cateary Problem

One of the most famous problems where hyperbolic functions are used is in determining the profile of a heavy chain (of constant density ρ) suspended from two points of equal height (known as a catenary curve). To derive the differential equation satisfied by the profile $y(x)$, we look at the forces acting on a small element of arc.
Let $T(x)$ be the tensile force in the chain with constant horizontal component H (since the load is not a function of x) and vertical component $V(x)$. The vertical components of the tensile force at either end of the arc are V and $V + \delta V$. The mass of the arc will be $\rho(\delta s)$, so that the force due to gravity is $\rho g(\delta s)$ The horizontal equilibrium is the trivial relation $H = H$, whereas the vertical equilibrium is the more informative

$$(V + \delta V) = V + \rho g(\delta s).$$

Divind both sides by δx gives

$$\frac{\delta V}{\delta x} = \rho g \frac{\delta s}{\delta x}.$$

From geometry, we also have the approximation

$$\frac{\delta y}{\delta x} \approx \frac{V}{H}.$$

We also have the approximation to the arclength δs

$$(\delta x)^2 \approx (\delta x)^2 + (\delta y)^2 \implies \frac{\delta s}{\delta x} \approx \sqrt{1 + \left(\frac{\delta y}{\delta x}\right)^2}.$$

Finally we take the limit $\delta x \to 0$ so that $\delta y \to 0$ and $\delta s \to 0$ similtaneously. We then have the following equations

$$\frac{\mathrm{d}V}{\mathrm{d}x} = \rho g \frac{\mathrm{d}s}{\mathrm{d}x},$$
$$V = H\frac{\mathrm{d}y}{\mathrm{d}x},$$
$$\frac{\mathrm{d}s}{\mathrm{d}x} = \sqrt{1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}.$$

Putting these equations together, yields the ODE satisfied by the profile of the chain, $y(x)$,

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} = \frac{\rho g}{H}\sqrt{1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}.$$

1.2.5 Linear Second-Order Non-Homogenous ODEs and the Wronskian

1.3 Lecture 3

1.3.1 Variation of Parameters

We've seen that for a linear second-order, non-homogenous IVP,

$$y'' + p(x)y' + q(x)y = r(x), \quad y(x_0) = y_0$$

if p, q, r are continous on an open interval I, and the inital condition, $x_0 \in I$, then there exisits a solution to the IVP. The solution will be a linear combination of the solution in the homogenous case and the particular case, $y(x) = y_H(x) + y_P(x)$. Assuming the homogenous case is a linear combination of linearly independent ys, ie $W(y_1, y_2) \neq 0$, we can apply variation of parameters. The process is as follows:

1. Solve $y'' + p(x)y' + q(x)y = 0$, and obtain a fundamental set of solutions, y_1, y_2. Calculate the Wronskian, $W(y_1, y_2)(x) = W$.
2. Set $y_P = u(x)y_1(x) + v(x)y_2(x)$ and substitute into the ODE. We also impose the condition, $u'y_1 + v'y_2 = 0$. We can freely impose this condition because we have two functions, u, v, and only one equation they must satisfy, the ODE.
3. We obtain
$$u(x) = -\int \frac{y_2 r}{W} \mathrm{d}x, \qquad v(x) = \int \frac{y_1 r}{W} \mathrm{d}x.$$

This approach is a variation of the reduction of order, which prescribes taking some solution, y, of the associated ODE, and using it to find a particular solution.

Example 1.3.1

Derivation of $u(x)$ and $v(x)$ of the variation of parameters.

$$y'' + p(x)y' + q(x)y = r(x) \tag{1}$$

Suppose we solved the homogenous case, $y'' + py' + qy = 0$.

$$\begin{gathered}
\implies \exists y_1(x), y_2(x) : W(y_1, y_2)(x) \neq 0,\ y_H(x) = Ay_1(x) + By_2(x) \\
y_P(x) = u(x)y_1(x) + v(x)y_2(x) \qquad (2) \\
\therefore y_P' = u'y_1 + uy_1' + v'y_2 + vy_2
\end{gathered}$$

Impose that $u'y_1 + v'y_2 = 0$, then

$$\begin{gathered}
y_P' = uy_1' + vy_2' \\
\therefore y_P'' = u'y_1' + uy_1'' + v'y_2' + vy_2''
\end{gathered}$$

We'll now substitute (2)'s derivatives back into (1), and find

$$(u'y_1' + uy_1'' + v'y_2' + vy_2'') + p(uy_1' + vy_2') + q(uy_1 + vy_2) = r$$

Consider $uy_1'' + puy_1' + quy_1$ and $vy_2'' + pvy_2' + qvy_2$, and note that they are solutions to the homogenous case, and are therefore equal to 0. So we can simply cancel them out, and are left with:

$$u'y_1' + v'y_2' = r$$

In fact, the entire system has been reduced to the system of equations

$$\begin{gathered}
\begin{cases} u'y_1 + v'y_2 = 0 \\ u'y_1' + v'y_2' = r \end{cases} \\
\iff \begin{pmatrix} y_1(x) & y_2(x) \\ y_1'(x) & y_2'(x) \end{pmatrix} \begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} 0 \\ r(x) \end{pmatrix} \\
\begin{pmatrix} y_1(x) & y_2(x) \\ y_1'(x) & y_2'(x) \end{pmatrix} = \hat{W},\ \det \hat{W} = \det W(y_1, y_2)(x) = W \neq 0 \implies \hat{W} \text{ is invertible.} \\
\hat{W}^{-1} = \frac{1}{\det \hat{W}} \begin{pmatrix} y_2'(x) & -y_2(x) \\ -y_1' & y_1(x) \end{pmatrix} \\
\therefore \begin{pmatrix} u' \\ v' \end{pmatrix} = \hat{W}^{-1} \begin{pmatrix} 0 \\ r(x) \end{pmatrix} = \frac{1}{\det \hat{W}} \begin{pmatrix} y_2'(x) & -y_2(x) \\ -y_1' & y_1(x) \end{pmatrix} \begin{pmatrix} 0 \\ r(x) \end{pmatrix} \\
\begin{pmatrix} u' \\ v' \end{pmatrix} = \frac{1}{W} \begin{pmatrix} -y_2 r \\ y_1 r \end{pmatrix}
\end{gathered}$$

$$\iff \begin{cases} u' = \dfrac{-y_2 r}{W} \\ v' = \dfrac{y_1 r}{W} \end{cases} \iff \begin{cases} u = -\displaystyle\int \frac{y_2 r}{W} \mathrm{d}x \\ v = \displaystyle\int \frac{y_1 r}{W} \mathrm{d}x \end{cases}$$

Example 1.3.2
Solve

$$y'' - 4y' + 5y = \frac{2e^{2x}}{\sin x}$$

using variation of parameters.

$$y = y_H + y_P$$

Let's ansatz that $y_H = e^{\lambda x}$

$$\iff \lambda^2 - 4 + 5 = 0 \iff \lambda_{1,2} = 2 \pm i \iff y_H = Ae^{2x}\cos x + Be^{2x}\sin x$$

$$W = \det W(y_1, y_2)(x) = \det \begin{pmatrix} e^{2x}\cos x & e^{2x}\sin x \\ 2e^{2x}\cos x - e^{2x}\sin x & 2e^{2x}\sin x + e^{2x}\cos x \end{pmatrix} = e^{4x} \neq 0$$

Find $y_P = uy_1 + vy_2$

$$u(x) = -\int \frac{y_2 r}{W} \mathrm{d}x = -\int \frac{e^{2x}\sin x \frac{2e^{2x}}{\sin x}}{e^{4x}} \mathrm{d}x = -2\int 1 \ \mathrm{d}x = -2x$$

$$v(x) = \int \frac{y_1 r}{W} \mathrm{d}x = \int \frac{e^{2x}\cos x \frac{2e^{2x}}{\sin x}}{e^{4x}} \mathrm{d}x = 2\int \cot x \ \mathrm{d}x = 2\ln|\sin x|$$

$$\implies y_P = 2\ln|\sin x| \, e^{2x}\sin x - 2xe^{2x}\cos x$$

$$\implies y = Ae^{2x}\cos x + Be^{2x}\sin x + 2\ln|\sin x| \, e^{2x}\sin x - 2xe^{2x}\cos x$$

Example 1.3.3
Solve for y_P, given

$$y'' - 4y' + 5y = \frac{2e^{2x}}{\sin x}$$

using reduction of order.

$$y = y_H + y_P y_P = U(x) y_H = U(x)e^{2x}\sin x$$

$$y_P' = U'e^{2x}\sin x + 2Ue^{2x}\sin x + Ue^{2x}\cos x$$

$$y_P'' = e^{2x}\left(U''\sin x + 2U'(2\sin x + \cos x) + U(3\sin x + 4\cos x)\right)$$

Plug y_P and its derivatives back into the ODE

$$\begin{gathered} e^{2x}\left(U''\sin x + 2U'(2\sin x + \cos x) + U(3\sin x + 4\cos x)\right) \\ -4e^{2x}(U'\sin x + 2U\sin x + U\cos x) \\ +5e^{2x}(U(x)\sin x) \\ = \frac{2e^{2x}}{\sin x} \\ e^{2x}\left(U''\sin x + 4U'\sin x + 2U'\cos x + 3U\sin x + 4U\cos x\right) \\ +e^{2x}(-4U'\sin x - 8U\sin x - 4U\cos x) \\ +e^{2x}(5U(x)\sin x) \\ = \frac{2e^{2x}}{\sin x} \end{gathered}$$

$$\begin{aligned}
U''e^{2x}\sin x + 2U'e^{2x}\cos x &= \frac{2e^{2x}}{\sin x} \\
U''\sin^2 x + 2U'\cos x\sin x &= 2 \\
\frac{\mathrm{d}}{\mathrm{d}x}\left(U'\sin^2 x\right) &= 2 \\
\int \frac{\mathrm{d}}{\mathrm{d}x}\left(U'\sin^2 x\right)\mathrm{d}x &= \int 2\mathrm{d}x \\
U'\sin^2 x &= 2x \\
\therefore U' &= \frac{2x}{\sin^2 x} \\
\therefore U &= 2\ln|\sin x| - 2x\frac{\cos x}{\sin x} \\
\implies y_P &= e^{2x}\sin x\left(2\ln|\sin x| - 2x\frac{\cos x}{\sin x}\right) \\
&= 2e^{2x}\sin x\ln|\sin x| - 2xe^{2x}\cos x
\end{aligned}$$

Which is the same answer we got when solving this using variation of parameters.

1.3.2 Vector Spaces

Note:-

$\mathbb{F}$ stands for $\mathbb{R}$ or $\mathbb{C}$.
Thus, if a statement holds for both $\mathbb{R}$ and $\mathbb{C}$, we say it holds for $\mathbb{F}$.
Elements of $\mathbb{F}$ called scalars.

Definition 1.3.1: Vector Space

Let V be a nonempty set on which the operations addition ('+') and scalar mulitplication ('$\cdot$') are defined. V is called a vector space over $\mathbb{F}$ if the following hold for all $\underset{\sim}{u}, \underset{\sim}{v}, \underset{\sim}{w} \in V$ and $k, l \in \mathbb{F}$:

(V1) Closure: $\underset{\sim}{u} + \underset{\sim}{v} \in V$

(V2) Additive Communitativity: $\underset{\sim}{u} + \underset{\sim}{v} = \underset{\sim}{v} + \underset{\sim}{u}$

(V3) Additive Associativity: $\underset{\sim}{u} + (\underset{\sim}{v} + \underset{\sim}{w}) = (\underset{\sim}{u} + \underset{\sim}{v}) + \underset{\sim}{w}$

(V4) Additive Identity: $\exists \underset{\sim}{0} : \underset{\sim}{u} + \underset{\sim}{0} = \underset{\sim}{u}$

(V5) Additive Inverse: $\forall \underset{\sim}{u}, \exists(-\underset{\sim}{u}) : \underset{\sim}{u} + (-\underset{\sim}{u}) = \underset{\sim}{0}$

(V6) Closure Under Scalar Multiplication: $k \cdot \underset{\sim}{u} \in V$

(V7) Multiplicative-Additive Distributivity: $k \cdot (\underset{\sim}{u} + \underset{\sim}{v}) = k \cdot \underset{\sim}{u} + k \cdot \underset{\sim}{v}$

(V8) Additive-Multiplicative Distributivity: $(k + l) \cdot \underset{\sim}{u} = k \cdot \underset{\sim}{u} + l \cdot \underset{\sim}{u}$

(V9) Multiplicative-Multiplicative Distributivity: $k \cdot (l \cdot \underset{\sim}{u}) = (kl) \cdot \underset{\sim}{u}$

(V10) Multiplicative Identity: $1 \cdot \underset{\sim}{u} = \underset{\sim}{u}$

Elements of a vector space are called *vectors*

To decide if a given nonempty set is a vector space, we suggest following

1. Identify what V is, what are its elements?
2. Identity what $+$ and $\cdot$ are.
3. Verify closure (V1, V6)

4. Identity identities and inverses (V4, V5, V10)

5. Verify communitativity, associativity and distributivity axioms (V2, V3, V7, V8, V9)

Example 1.3.4

Consider the set of $n-$tupples, $\mathbb{F}^n$, where $n \in \mathbb{N}$. Is $\mathbb{F}^n$ a vector space?

1. $V = \left\{ \underset{\sim}{u} = (u_1, u_2, \ldots, u_n) \,\middle|\, \forall i \in \mathbb{N},\ u_i \in \mathbb{F} \right\}$

2. $\underset{\sim}{u} + \underset{\sim}{v} = (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n)$
   $k \cdot \underset{\sim}{u} = (ku_1, ku_2, \ldots, ku_n)$

3. $\underset{\sim}{u} + \underset{\sim}{v} = (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n)$. $\forall i \in \mathbb{N},\ u_i + v_i \in \mathbb{F}$. Therefore, $\forall \underset{\sim}{u}, \underset{\sim}{v} \in \mathbb{F}^n,\ \underset{\sim}{u} + \underset{\sim}{v} \in \mathbb{F}^n$.
   $k \cdot \underset{\sim}{u} = (ku_1, ku_2, \ldots, ku_n)$. $\forall i \in \mathbb{N},\ k\underset{\sim}{u}_i \in \mathbb{F}$. Therefore, $\forall \underset{\sim}{u} \in \mathbb{F}^n,\ k \in \mathbb{F},\ k \cdot \underset{\sim}{u} \in \mathbb{F}^n$.

4. $\underset{\sim}{0} = (0, 0, \ldots, 0)$, with n entries.
   $\forall \underset{\sim}{u} \in \mathbb{F}^n,\ \exists(-\underset{\sim}{u}) = (-u_1, -u_2, \ldots, -u_n) \in \mathbb{F}^n$
   $1 \in \mathbb{F},\ 1 \cdot \underset{\sim}{u} = 1$

5. $\underset{\sim}{u} + \underset{\sim}{v} = (u_1 + v_1, \ldots, u_n + v_n) = (v_1 + u_1, \ldots, v_n + u_n) = \underset{\sim}{v} + \underset{\sim}{u}$
   $\underset{\sim}{u} + (\underset{\sim}{v} + \underset{\sim}{w}) = (u_1 + (v_1 + w_1), \ldots, u_n + (v_n + w_n)) = ((u_1 + v_1) + w_1, \ldots, (u_n + v_n) + w_n) = (\underset{\sim}{u} + \underset{\sim}{v}) + \underset{\sim}{w}$
   $k \cdot (\underset{\sim}{u} + \underset{\sim}{v}) = (k(u_1 + v_1), \ldots, k(u_1 + v_1)) = (ku_1 + ku_1, \ldots, ku_n + kv_n) = (ku_1, \ldots, ku_n) + (kv_1, \ldots, kv_n) = k \cdot \underset{\sim}{u} + k \cdot \underset{\sim}{v}$
   $(k+l) \cdot \underset{\sim}{u} = ((k + l)u_1, \ldots, (k + l)u_n) = (ku_1 + lu_1, \ldots, ku_n + lu_n) = (ku_1, \ldots, ku_n) + (lu_1, \ldots, lu_n) = k \cdot \underset{\sim}{u} + l \cdot \underset{\sim}{u}$
   $k \cdot (l \cdot \underset{\sim}{u}) = k \cdot (lu_1, \ldots, lu_n) = (klu_1, \ldots, klu_n) = kl \cdot (u_1, \ldots, u_n) = (kl) \cdot \underset{\sim}{u}$

Therefore $\mathbb{F}^n$ is a vector space.

Example 1.3.5

Consider the set of $m \times n$ matricies with scalar entries, $M_{m \times n}(\mathbb{F})$. Is this a vector space?

1. $V = \left\{ \begin{pmatrix} a_{1,1} & a_{1,2} & \ldots & a_{1,n} \\ a_{2,1} & a_{2,2} & \ldots & a_{2,n} \\ \vdots & & \ddots & \\ a_{m,1} & a_{m,2} & \ldots & a_{m,n} \end{pmatrix} \middle| \forall i, j \in \mathbb{N},\ a_{i,j} \in \mathbb{F} \right\}$

2. '+' is matrix addition; entry-wise addition. $\forall A, B \in M_{m \times n}(\mathbb{F}),\ \exists C :\ \forall i, j \in \mathbb{N},\ c_{i,j} = a_{i,j} + b_{i,j}$.
   '·' is scalar multiplication. $kA = C$ where each entry of C, $c_{i,j} = ka_{i,j}$.

3. $\forall i, j \in \mathbb{N}, i \leq m, j \leq n,\ a_{i,j}, b_{i,j} \in \mathbb{F} \implies a_{i,j} + b_{i,j} \iff A + B \in M_{m \times n}(\mathbb{F})$
   $\forall k \in \mathbb{F}, ka_{i,j} \in \mathbb{F} \iff k \cdot A \in M_{m \times n}(\mathbb{F})$

4. There exists a zero matrix, with all 0 entries.
   For all matricies, A, there exisits a matrix $-A$, such that $-A = \begin{pmatrix} -a_{1,1} & -a_{1,2} & \ldots & -a_{1,n} \\ -a_{2,1} & -a_{2,2} & \ldots & -a_{2,n} \\ \vdots & & \ddots & \\ -a_{m,1} & -a_{m,2} & \ldots & -a_{m,n} \end{pmatrix}$
   $1 \in \mathbb{F}, \forall A \in M_{m \times n}(\mathbb{F}),\ 1 \cdot A = A$.

5. These axioms are essentially extensions of the $n-$tupple proofs we just gave, and I am not going to write them all out right now, but rest assured: they hold.

Therefore, $M_{m \times n}(\mathbb{F})$ is a vector space.

Other examples of vector spaces

- The set of continous real-valued functions on $[a, b]$, $C[a, b]$.

$$[a,b] = D \subset \mathbb{R}.\ C[a,b] = V = \{f : D \to \mathbb{R} \mid f \text{ is continuous on } D.\}$$
$$\forall f, g \in C[a,b],\ f(x) + g(x) = (f+g)(x)$$
$$\forall k \in \mathbb{R}, f \in C[a,b],\ k \cdot f(x) = (k \cdot f)(x)$$
$$\underset{\sim}{0}(x) = 0;\ (-f)(x) = -f(x)$$

- The set of polynomials of degree at most n, $P_n(\mathbb{F})$

$$P_n(\mathbb{F}) = V = \left\{a_0 + a_1x + a_2x^2 + \cdots + a_nx^n \,\middle|\, \forall i \in \mathbb{Z}_{\geq 0},\ a_i \in \mathbb{F}\right\}$$
$$\forall p, q \in P_n(\mathbb{F}),\ p + q = (p_0 + q_0) + (p_1 + q_1)x + (p_2 + q_2)x^2 + \cdots + (p_n + q_n)x^n$$
$$\forall k \in \mathbb{F},\ p \in P_n(\mathbb{F}),\ k \cdot p = (kp_0) + (kp_1)x + (kp_2)x^2 + \cdots + (kp_n)x^n$$
$$\exists \underset{\sim}{0} \in P_n(\mathbb{F}),\ \underset{\sim}{0} = 0; \forall p \in P_n(\mathbb{F}), \exists(-p) : p + (-p) = \underset{\sim}{0}, (-p) = -p$$

- The set of solutions to a homogenous linear ODE. Consider $y'' + p(x)y' + q(x)y = 0, y = y(x)$

$$V = \{y \mid y \text{ is a solution to the linear homogenous ODE under consideration}\}$$
$$\forall y_1, y_2 \in V,\ y_1 + y_2 = (y_1 + y_2)(x) \in V$$
$$\forall k \in \mathbb{F}, \forall y \in V,\ k \cdot y = (k \cdot y)(x) \in V$$
$$\underset{\sim}{0} = 0 \in V; \forall y \in V,\ -y = (-1 \cdot y)(x) \in V$$

1.3.3 Linear Algebra Concepts

Linear combination

For $\underset{\sim}{v_1}, \underset{\sim}{v_2}, \ldots, \underset{\sim}{v_n} \in V$ and $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{F}$, we call

$$\alpha_1\underset{\sim}{v_1} + \alpha_2\underset{\sim}{v_2} + \cdots + \alpha_n\underset{\sim}{v_n}$$

a linear combination of the vectors $\underset{\sim}{v_1}, \underset{\sim}{v_2}, \ldots, \underset{\sim}{v_n}$.

Linear independence

A non-empty set of vectors $S = \left\{\underset{\sim}{v_1}, \underset{\sim}{v_2}, \ldots, \underset{\sim}{v_n}\right\} \subseteq V$ is said to be linearly dependent if there exist scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$ not all zero such that

$$\alpha_1\underset{\sim}{v_1} + \alpha_2\underset{\sim}{v_2} + \cdots + \alpha_n\underset{\sim}{v_n} = \underset{\sim}{0}.$$

Otherwise, S is called linearly independent, i.e. S is linearly independent if

$$\alpha_1\underset{\sim}{v_1} + \alpha_2\underset{\sim}{v_2} + \cdots + \alpha_n\underset{\sim}{v_n} = \underset{\sim}{0} \implies \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0.$$

Subspace

A subset $W \subseteq V$ is called a subspace of V if W is also a vector space with the same addition and scalar multiplication as V. In particular, W is required to close under addition and scalar multiplication.

Span

The span of a non-empty set of vectors $S = \left\{\underset{\sim}{v_1}, \underset{\sim}{v_2}, \ldots, \underset{\sim}{v_n}\right\} \subseteq V$ is the set of all linear combinations of vectors in S, denoted $\operatorname{span} S$. The set $\operatorname{span} S$ is a subspace of V.

$$\operatorname{span} S = \operatorname{span}\left\{\underset{\sim}{v_1}, \underset{\sim}{v_2}, \ldots, \underset{\sim}{v_n}\right\} = \left\{\alpha_1\underset{\sim}{v_1} + \alpha_2\underset{\sim}{v_2} + \cdots + \alpha_n\underset{\sim}{v_n} \,\middle|\, \forall i \in \mathbb{N},\ \alpha_i \in \mathbb{F}\right\}$$

If the span of S is equal to the vector space V, then S is said to span V.

1.3.4 Basis

Let $\beta = \left\{\underset{\sim}{v_1}1, \underset{\sim}{v_2}, \ldots, \underset{\sim}{v_n}\right\} \subseteq V$ be a set of vectors in V. β is a basis for V If

(B1) β is linearly independent

(B2) β spans V. ($\operatorname{span}\beta = V$)

Note, however, that the notion of basis is only defined for finite sets. A nonzero vector space is **finite-dimensional** if it contains a finite number of vectors that form a basis; n is a finite number. If no such set exists, the vector space is called **infinite-dimensional**
Basis enables us to concretely define a sense of dimensionality for a vector space, namely

$$n = \dim V = |\operatorname{span}\beta|.$$

It is said that V is $n-$dimensional, or has n dimensions.

An **ordered basis** for a vector space is a basis endowed with a specific order. For some vector spaces, there is a canonical ordered basis, called a standard basis. For example,

$$\text{Standard basis of } \mathbb{R}^3 : \beta = \left\{\hat{\imath}, \hat{\jmath}, \hat{k}\right\} = \left\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}\right\} \implies \dim\mathbb{R}^3 = 3$$

$$\text{Standard basis of } P_3(\mathbb{R}) : \beta = \left\{1, x, x^2, x^3\right\} \implies \dim P_3(\mathbb{R}) = 4$$

$$\text{Standard basis of } M_{m\times n}(\mathbb{R}) : \beta = \left\{\begin{matrix}\begin{pmatrix}1&0&0\\0&0&0\end{pmatrix}, & \begin{pmatrix}0&1&0\\0&0&0\end{pmatrix}, & \begin{pmatrix}0&0&1\\0&0&0\end{pmatrix}, \\ \begin{pmatrix}0&0&0\\1&0&0\end{pmatrix}, & \begin{pmatrix}0&0&0\\0&1&0\end{pmatrix}, & \begin{pmatrix}0&0&0\\0&0&1\end{pmatrix}\end{matrix}\right\} \implies \dim M_{m\times n}(\mathbb{R}) = 6$$

Chapter 2

Week 2

2.1 Lecture 4

2.1.1 Decomposition Theory

Let β be a set of vectors in the vector space V. Then, β is a basis for V if and only if each each vector in V can be expressed as a unique linear combination of vectors in β.

Lenma 2.1.1 β is a basis of $V \implies$ all $\underset{\sim}{w} \in V$ is a unique linear combination of $v \in \beta$.

Assume β is a basis for V.

$\implies \forall \underset{\sim}{w} \in V,\ \underset{\sim}{w} = \displaystyle\sum_{i=1}^{n} \alpha_i \underset{\sim}{v}_i$, where each α_i is unique.

Suppose $\underset{\sim}{w}$ is not a unique linear combination, then it could also be expressed as $\displaystyle\sum_{i=1}^{n} \beta_i \underset{\sim}{v}_i$.

Then, $\underset{\sim}{0} = \underset{\sim}{w} + (-\underset{\sim}{w}) = \displaystyle\sum_{i=1}^{n} \alpha_i \underset{\sim}{v}_i - \sum_{i=1}^{n} \beta_i \underset{\sim}{v}_i = \sum_{i=1}^{n} (\alpha_i - \beta_i) \underset{\sim}{v}_i$.
$\implies \alpha_i - \beta_i = 0,\ \forall i \iff \alpha_i = \beta_1,\ \forall i$
Which is a contradition, because we assumed that $\underset{\sim}{w}$ did not have a unique linear combination.
Therefore, $\underset{\sim}{w}$ has a unique linear combination.

Lenma 2.1.2 All $\underset{\sim}{w} \in V$ is a unique linear combination of $v \in \beta \implies \beta$ is a basis of V.

Assume $\forall \underset{\sim}{w} \in V$ is a unique linear combination of vectors in $\underset{\sim}{v}_i \in \beta$.
$\implies \beta \setminus V = \operatorname{span}\{\beta\}$, by axiom **B2**.
Since $\underset{\sim}{w}$ is assumed to be unique, we only have one choice of coeficcents, α_i.
$\implies \beta$ satisfies **B1**
$\implies \beta$ is a basis for V.

Theorem 2.1.1 Decomposition Theory

$$\beta \text{ forms a basis of } V \iff \forall \underset{\sim}{w} \in V,\ \underset{\sim}{w} = \sum_{i=1}^{n} \alpha_i \underset{\sim}{v}_i,\ \forall i, \underset{\sim}{v}_i \in \beta,\ \text{ is unique.}$$

Note:-
Not all bases of V are unique!

2.1.2 Transition Matrix

Let β be an ordered basis for the vector space V. For $\underset{\sim}{u} \in V$ let $a_1, a_2, \ldots, a_n$ be the unique scalars such that

$$\underset{\sim}{u} = \sum_{i=1}^{n} a_i \underset{\sim}{v}_i.$$

Definition (Coordinate Vector). The coordinate vector of $\underset{\sim}{u}$ reletive to β is given by

$$\left[\underset{\sim}{u}\right]_\beta = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

We can denote the $i-$th component as $\left[\underset{\sim}{u}\right]_\beta^i$

Definition 2.1.1: Transition Matrix

Let β' be another ordered basis of V. The transition matrix From β to β', denoted by $P_{\beta \to \beta'}$ relates the two coordinate vectors of $\underset{\sim}{u}$ by

$$\left[\underset{\sim}{u}\right]_{\beta'} = P_{\beta \to \beta'} \left[\underset{\sim}{u}\right]_\beta$$

If β'' is yet another ordered basis of V then

$$P_{\beta' \to \beta''} P_{\beta \to \beta'} = P_{\beta \to \beta''} \implies P_{\beta \to \beta'} P_{\beta' \to \beta} = P_{\beta \to \beta} = I$$

Example 2.1.1

Consider two ordered bases $\beta = \{1, x\}$ and $\beta' = \{1 + x, 2x\}$ of the vector space $P_1(\mathbb{F})$. $\underset{\sim}{u} = a + bx$ can be rewritten as

$$a(1 + x) + \frac{1}{2}(b - a)(2x),$$

so we have

$$\left[\underset{\sim}{u}\right]_\beta = \begin{pmatrix} a \\ b \end{pmatrix} \qquad \text{and} \qquad \left[\underset{\sim}{u}\right]_{\beta'} = \begin{pmatrix} a \\ \frac{1}{2}(b - a) \end{pmatrix}$$

Therefore the transtion matrix $P_{\beta \to \beta'}$ is given

$$\begin{pmatrix} 1 & 0 \\ \frac{-1}{2} & \frac{1}{2} \end{pmatrix}$$

which satisifies $\left[\underset{\sim}{u}\right]_{\beta'} = P_{\beta \to \beta'} \left[\underset{\sim}{u}\right]_\beta$

In general, with vector space V, with $\dim V = n$, and two bases $\beta = \{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \ldots, \underset{\sim}{u}_n\}$ and $\beta' = \{\underset{\sim}{u}'_1, \underset{\sim}{u}'_2, \ldots, \underset{\sim}{u}'_n\}$ of V,

$$P_{\beta \to \beta'} = \left(\left[\underset{\sim}{u_1}\right]_{\beta'} \middle| \left[\underset{\sim}{u_2}\right]_{\beta'} \middle| \ldots \middle| \left[\underset{\sim}{u_n}\right]_{\beta'} \right)$$

2.1.3 Real Inner Product Spaces

Dot Product

Sometimes called the Euclidean inner product, for two vectors $\underset{\sim}{u}, \underset{\sim}{v} \in \mathbb{R}^n$, the dot product is given by

$$\underset{\sim}{u} \cdot \underset{\sim}{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n = \sum_{i=1}^{n} u_i v_i = \underset{\sim}{u}^T \underset{\sim}{v}.$$

This is a map from $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, and the following key properties:

(i) Symmetric: $\underset{\sim}{u} \cdot \underset{\sim}{v} = \underset{\sim}{v} \cdot \underset{\sim}{v}$

(ii) Linear 1: $\big(\underset{\sim}{u} + \underset{\sim}{v}\big) + \underset{\sim}{w} = (\underset{\sim}{u} \cdot \underset{\sim}{w}) + (\underset{\sim}{v} \cdot \underset{\sim}{w})$

(iii) Linear 2: $(k\underset{\sim}{u}) \cdot \underset{\sim}{v} = k(\underset{\sim}{u} \cdot \underset{\sim}{v})$

(iv) Positive Definite 1: $\forall \underset{\sim}{u} \in \mathbb{R},\ \underset{\sim}{u} \cdot \underset{\sim}{u} = \big|\big|\underset{\sim}{u}\big|\big|^2 \geq 0$

(v) Positive Definite 2: $\underset{\sim}{u} \cdot \underset{\sim}{u} = 0 \iff \underset{\sim}{u} = \underset{\sim}{0}$.

Let's axiom-itize this!

Inner Product

An inner product is a function on a vector space, V that maps two vectors in V, say $(\underset{\sim}{u}, \underset{\sim}{v})$, to a real number, denoted $\big\langle \underset{\sim}{u}, \underset{\sim}{v} \big\rangle$, such that, for all $\underset{\sim}{u}, \underset{\sim}{v}, \underset{\sim}{w} \in V$:

(**I1**) Symmetric: $\big\langle \underset{\sim}{u}, \underset{\sim}{v} \big\rangle = \big\langle \underset{\sim}{v}, \underset{\sim}{v} \big\rangle$

(**I2**) Linear 1: $\big\langle \underset{\sim}{u} + \underset{\sim}{v}, \underset{\sim}{w} \big\rangle = \big\langle \underset{\sim}{u}, \underset{\sim}{w} \big\rangle + \big\langle \underset{\sim}{v}, \underset{\sim}{w} \big\rangle$

(**I3**) Linear 2: $\big\langle k\underset{\sim}{u}, \underset{\sim}{v} \big\rangle = k \big\langle \underset{\sim}{u}, \underset{\sim}{v} \big\rangle$

(**I4**) Positive Definite 1: $\big\langle \underset{\sim}{u}, \underset{\sim}{u} \big\rangle \geq 0$

(**I5**) Positive Definite 2: $\big\langle \underset{\sim}{u}, \underset{\sim}{u} \big\rangle = 0 \iff \underset{\sim}{u} = \underset{\sim}{0}$.

Note:-

Complex inner product spaces are beyond the scope of this course

Example 2.1.2 (Weighted Dot Product)

$\forall \underset{\sim}{u}, \underset{\sim}{v} \in \mathbb{R}^n$, with scalars $\gamma_i > 0, \forall i$,

$$\big\langle \underset{\sim}{u}, \underset{\sim}{v} \big\rangle = \gamma_1 u_1 v_1 + \gamma_2 u_2 v_2 + \cdots + \gamma_i u_i v_i = \sum_{i=1}^{n} \gamma_i u_i v_i$$

It's a dot product, except we can "weight" certain components higher then others. A special case of the weighted dot product is $\gamma_i = 1$ for all i, and this is the familar dot product.
STRESS: $\gamma_i > 0 \iff$ (**I5**)

Example 2.1.3 (Inner Product Generated by a Matrix)

Let $\underset{\sim}{u}, \underset{\sim}{v} \in \mathbb{R}^n,\ A \in M_{n\times n}(\mathbb{R}),\ A$ is invertible (therefore $\det A \neq 0$). Define an inner product,

$$\big\langle \underset{\sim}{u}, \underset{\sim}{v} \big\rangle = (A\underset{\sim}{u}) \cdot (A\underset{\sim}{v})$$

Note that: $(A\underset{\sim}{u}) \cdot (A\underset{\sim}{v}) = (A\underset{\sim}{u})^T A\underset{\sim}{v} = \underset{\sim}{u}^T A^T A\underset{\sim}{v}$.
Interestingly, the weifhted dot product is just a special case of this kind of inner product, only where $A^T A = A^2 = \operatorname{diag}(\gamma_1, \gamma_2, \ldots, \gamma_n) \iff A = \operatorname{diag}\big(\sqrt{\gamma_1}, \sqrt{\gamma_2}, \ldots, \sqrt{\gamma_n}\big)$.
STRESS: We assumed $\det A \neq 0 \implies \det(A^T A) = (\det A)^2 \neq 0$. Again $\det A \neq 0 \iff$ (**I5**)

Example 2.1.4 (Inner Product on $M_{n\times n}(\mathbb{R})$)

Note that $\dim M_{n\times n}(\mathbb{R}) = n^2$. We start by noting that the trace of a matrix $\underset{\sim}{u} \in M_{n\times n}(\mathbb{R})$ is

$$\operatorname{Tr} \underset{\sim}{u} = \sum_{i=1}^{n} \underset{\sim}{u}_{i,i}.$$

The inner product is defined as

$$\langle \underset{\sim}{u}, \underset{\sim}{v} \rangle = \text{Tr}\left\{\underset{\sim}{u}^T \underset{\sim}{v}\right\}$$

This is a more interesting example, so We'll prove that it is an inner product.
Lemma. I1: $\langle \underset{\sim}{u}, \underset{\sim}{v} \rangle = \text{Tr}\left(\underset{\sim}{u}^T \underset{\sim}{v}\right) = \text{Tr}\left(\left(\underset{\sim}{u}^T \underset{\sim}{v}\right)^T\right) = \text{Tr}\left(\underset{\sim}{v}^T \underset{\sim}{u}\right) = \langle \underset{\sim}{v}, \underset{\sim}{u} \rangle$ □
Lemma. I2, I3: $\langle \alpha \underset{\sim}{u} + \beta \underset{\sim}{v}, \underset{\sim}{w} \rangle = \text{Tr}\left((\alpha \underset{\sim}{u}^T + \beta \underset{\sim}{v}^T)\underset{\sim}{w}\right) = \text{Tr}\left(\alpha \underset{\sim}{u}^T \underset{\sim}{w} + \beta \underset{\sim}{v}^T \underset{\sim}{w}\right) = \text{Tr}\left(\alpha \underset{\sim}{u}^T \underset{\sim}{w}\right) + \text{Tr}\left(\beta \underset{\sim}{v}^T \underset{\sim}{w}\right) = \alpha \langle \underset{\sim}{u}, \underset{\sim}{w} \rangle + \beta \langle \underset{\sim}{v}, \underset{\sim}{w} \rangle$ □
Lemma. I4, I5: $\langle \underset{\sim}{u}, \underset{\sim}{u} \rangle = \text{Tr}\left(\underset{\sim}{u}^T \underset{\sim}{u}\right) = \sum_{i=1}^n \sum_{j=1}^n (u_{i,j})^2 \geq 0$. It follows that this expression can only equal 0 if and only if $\underset{\sim}{u} = \underset{\sim}{0}$.
Theorem. Therefore, this is an inner product □

Example 2.1.5 (Standard Inner Product on $P_n(\mathbb{R})$)

$$\langle \underset{\sim}{u}, \underset{\sim}{v} \rangle = p_0 q_0 + p_1 q_1 + \cdots + p_n q_n$$

Example 2.1.6 (Evaluation Inner Product on $P_n(\mathbb{R})$)
Let $x_0, x_1, \ldots, x_n \in \mathbb{R}$ all be distinct ($x_i \neq x_j \iff i \neq j$). Then

$$\left\langle \underset{\sim}{p}, \underset{\sim}{q} \right\rangle = p(x_0)q(x_0) + p(x_1)q(x_1) + \cdots + p(x_n)q(x_n) = \sum_{i=0}^n p(x_i)q(x_i)$$

Example 2.1.7 (Innter Product on $C[a,b]$)
Let $\underset{\sim}{f}, \underset{\sim}{g} \in C[a,b]$.

$$\left\langle \underset{\sim}{f}, \underset{\sim}{g} \right\rangle = \int_a^b f(x)g(x)\mathrm{d}x$$

2.2 Lecture 5

2.2.1 Orthognality

Norm

The norm (or magnitude or length) of an element $\underset{\sim}{v} = (v_1, \ldots, v_n) \in \mathbb{R}^n$ is given by the familar expression

$$\|\underset{\sim}{v}\| = \sqrt{\underset{\sim}{v} \cdot \underset{\sim}{v}} = \sqrt{v_1^2 + \cdots + v_n^2}$$

This can be rewritten using the real inner product space on V notation we've just developed,

$$\|\underset{\sim}{v}\| = \sqrt{\langle \underset{\sim}{u}, \underset{\sim}{u} \rangle}$$

We could use any inner product space, like the examples we presented in the previous chapter.
Definition (Unit Vector). A vector $\underset{\sim}{v} \in V$, with $\|\underset{\sim}{v}\| = 1$ is called a unit vector.

Definition 2.2.1: Distance Function

Inspured by our intution on $\mathbb{R}^n$, we can define the distance $d\left(\underset{\sim}{u}, \underset{\sim}{v}\right)$ between two vectors

$$d\left(\underset{\sim}{u}, \underset{\sim}{v}\right) = \|\underset{\sim}{u} - \underset{\sim}{v}\|$$

This distance is symmetric, following from

$$\left\langle \underset{\sim}{u} - \underset{\sim}{v}, \underset{\sim}{v} - \underset{\sim}{u} \right\rangle = \left\langle \underset{\sim}{u}, \underset{\sim}{u} \right\rangle - \left\langle \underset{\sim}{u}, \underset{\sim}{v} \right\rangle - \left\langle \underset{\sim}{v}, \underset{\sim}{u} \right\rangle + \left\langle \underset{\sim}{v}, \underset{\sim}{v} \right\rangle = \left\langle \underset{\sim}{v} - \underset{\sim}{u}, \underset{\sim}{u} - \underset{\sim}{v} \right\rangle$$

Critically, the notions of norm and distance are relative to the inner product itself.

Definition (Orthognal). Two vectors $\underset{\sim}{u}, \underset{\sim}{v} \in V$ are orthognol iff $\left\langle \underset{\sim}{u}, \underset{\sim}{v} \right\rangle = 0$.
This generalises to all inner product spaces, and gives us a basis to talk about a general "angle" between vectors.

Example 2.2.1

Consider $V = P_2(\mathbb{R})$ and $\underset{\sim}{u} = x,\ \underset{\sim}{v} = x^2 \in V$. Let

$$\left\langle \underset{\sim}{u}, \underset{\sim}{v} \right\rangle = \int_{-1}^{1} u(x)v(x)\mathrm{d}x = \int_{-1}^{1} x^3 \mathrm{d}x = 0$$

Therefore, in this inner product space, the vectors are orthognol. But if we changed our inner product space

$$\left\langle \underset{\sim}{u}, \underset{\sim}{v} \right\rangle = \int_{0}^{1} u(x)v(x)\mathrm{d}x = \int_{0}^{1} x^3 \mathrm{d}x = \frac{1}{4}$$

In this inner product space, the vectors are not orthognol.

Pythagorean Theorem

Let V be a real product space and let $\underset{\sim}{u}, \underset{\sim}{v} \in V$ then

$$\|\underset{\sim}{u} + \underset{\sim}{v}\|^2 = \|\underset{\sim}{u}\|^2 + \|\underset{\sim}{v}\|^2 \iff \left\langle \underset{\sim}{u}, \underset{\sim}{v} \right\rangle = 0$$

This very well known facto genralises to all inner product spaces.

Proof.

$$\|\underset{\sim}{u} + \underset{\sim}{v}\|^2 = \left\langle \underset{\sim}{u} + \underset{\sim}{v}, \underset{\sim}{u} + \underset{\sim}{v} \right\rangle = \left\langle \underset{\sim}{u}, \underset{\sim}{u} \right\rangle + 2\left\langle \underset{\sim}{u}, \underset{\sim}{v} \right\rangle + \left\langle \underset{\sim}{v}, \underset{\sim}{v} \right\rangle = \|\underset{\sim}{u}\|^2 + \|\underset{\sim}{v}\|^2 + 2\left\langle \underset{\sim}{u}, \underset{\sim}{v} \right\rangle$$

Then

$$\|\underset{\sim}{u} + \underset{\sim}{v}\|^2 = 0 \iff \left\langle \underset{\sim}{u}, \underset{\sim}{v} \right\rangle = 0$$

□

Cauchy-Schwarz Inequality

Let V be a real inner product space, and let $\underset{\sim}{u}, \underset{\sim}{v} \in V$ then

$$\left|\left\langle \underset{\sim}{u}, \underset{\sim}{v} \right\rangle\right| \leq \|\underset{\sim}{u}\|\ \|\underset{\sim}{v}\|$$

This inequality is an equality iff $\underset{\sim}{u}$ or $\underset{\sim}{v}$ is a scalar multiple of the other vector

Proof. First we'll consider the trivial case. Without loss of generality, suppose that $\underset{\sim}{u} = \underset{\sim}{0}$.
Then $0 = \left\langle \underset{\sim}{u}, \underset{\sim}{v} \right\rangle = \|\underset{\sim}{u}\|\ \|\underset{\sim}{v}\| = 0$

For the non-trivial case, take $\underset{\sim}{u}, \underset{\sim}{v} \in V : \underset{\sim}{u}, \underset{\sim}{v} \neq 0$.
Let $a = \|\underset{\sim}{u}\|^2 > 0,\ b = \left\langle \underset{\sim}{u}, \underset{\sim}{v} \right\rangle,\ c = \|\underset{\sim}{v}\|^2 > 0$.
Consider $t \in \mathbb{R}$ and $(t\underset{\sim}{u}, \underset{\sim}{v}) \in V$.

$$\implies 0 \leq \|t\underset{\sim}{u} + \underset{\sim}{v}\|^2 = \left\langle t\underset{\sim}{u} + \underset{\sim}{v}, t\underset{\sim}{u} + \underset{\sim}{v} \right\rangle$$

$$
\begin{aligned}
&= \|u\|^2 t^2 + 2\left\langle \underset{\sim}{u}, \underset{\sim}{v} \right\rangle t + \|v\|^2 \\
&= at^2 + 2bt + c
\end{aligned}
$$

Consider this as a degree-2 polynomial in t, and consider the conditions under which it has real solutions

$$0 \leq at^2 + 2bt + c \iff b^2 - 4ac \leq 0 \iff b^2 \leq 4ac \iff \langle \underset{\sim}{u}, \underset{\sim}{v} \rangle^2 \leq \|\underset{\sim}{u}\|^2 \ \|\underset{\sim}{v}\|^2 \iff \langle \underset{\sim}{u}, \underset{\sim}{v} \rangle \leq \|\underset{\sim}{u}\| \ \|\underset{\sim}{v}\|$$

which is the Cauchy-Schwarz inequality we sought to prove. □

Triangle Inequality

Let V be a real inner product space, and let $\underset{\sim}{u}, \underset{\sim}{v} \in V$. Then

$$\|\underset{\sim}{u} + \underset{\sim}{v}\| \leq \|\underset{\sim}{u}\| + \|\underset{\sim}{v}\|$$

Proof.

$$
\begin{aligned}
\|\underset{\sim}{u} + \underset{\sim}{v}\|^2 &= \|\underset{\sim}{u}\|^2 + 2\left\langle \underset{\sim}{u}, \underset{\sim}{v} \right\rangle + \|\underset{\sim}{v}\|^2 \\
&\leq \|\underset{\sim}{u}\|^2 + 2\left\langle \underset{\sim}{u}, \underset{\sim}{v} \right\rangle + \|\underset{\sim}{v}\|^2 \\
&\leq \|\underset{\sim}{u}\|^2 + 2\|\underset{\sim}{u}\|\|\underset{\sim}{v}\| + \|\underset{\sim}{v}\|^2 && \text{(C-S Inequality)} \\
\iff \|\underset{\sim}{u} + \underset{\sim}{v}\| &\leq \|\underset{\sim}{u}\| + \|\underset{\sim}{v}\|
\end{aligned}
$$

□

Angle Between Two Vectors

In $\mathbb{R}^n$, given a $\cdot$ product, θ between $\underset{\sim}{u}, \underset{\sim}{v} \in \mathbb{R}^n$ is given by

$$\theta = \arccos\left(\frac{\underset{\sim}{u} \cdot \underset{\sim}{v}}{\|\underset{\sim}{u}\| \ \|\underset{\sim}{v}\|}\right), \ \theta \in [0, \pi]$$

Using this intution, consider an inner product space V, with $\langle \, , \, \rangle$

$$\implies \theta = \arccos\left(\frac{\langle \underset{\sim}{u}, \underset{\sim}{v} \rangle}{\|\underset{\sim}{u}\| \ \|\underset{\sim}{v}\|}\right)$$

Note that

$$\frac{\langle \underset{\sim}{u}, \underset{\sim}{v} \rangle}{\|\underset{\sim}{u}\| \ \|\underset{\sim}{v}\|} \leq 1 \iff -1 \leq \frac{\langle \underset{\sim}{u}, \underset{\sim}{v} \rangle}{\|\underset{\sim}{u}\| \ \|\underset{\sim}{v}\|} \leq 1,$$

which aligns perfectly the cos, whose range is $[-1, 1]$ and arccos whose domain is $[-1, 1]$.

Orthogonal Complement

Let U be a subset of the real inner product space V. The orthogonal complement of U, denoted $U^\perp$, is the set of all vectors in V that are orthogonal to every vector in U. That is

$$U^\perp = \left\{ \underset{\sim}{v} \in V \ \middle| \ \langle \underset{\sim}{v}, \underset{\sim}{u} \rangle = 0, \ \forall \underset{\sim}{u} \in U \right\}$$

This is a vector space with addition and sclar multiplication inherited from V.

Example 2.2.2

For $A \in M_{m\times n}(\mathbb{R})$, $\text{Row}(A)^\perp = \mathcal{N}(A)$ with respect to the Euclidean inner product.

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix} = \begin{pmatrix} \underset{\sim}{r}_1 \\ \underset{\sim}{r}_2 \\ \vdots \\ \underset{\sim}{r}_m \end{pmatrix} = \begin{pmatrix} \underset{\sim}{c}_1 & \underset{\sim}{c}_2 & \dots & \underset{\sim}{c}_n \end{pmatrix}$$

$$\text{Col}(A) = \text{span}\left\{\underset{\sim}{c}_1, \underset{\sim}{c}_2, \dots, \underset{\sim}{c}_m\right\}$$
$$\text{Row}(A) = \text{span}\left\{\underset{\sim}{r}_1^T, \underset{\sim}{r}_2^T, \dots, \underset{\sim}{r}_n^T\right\}$$

$$\text{Given } \underset{\sim}{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \text{ Take } A\underset{\sim}{x} = \begin{pmatrix} \underset{\sim}{r}_1 \\ \underset{\sim}{r}_2 \\ vdots \\ \underset{\sim}{r}_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \underset{\sim}{r}_1\underset{\sim}{x} \\ \underset{\sim}{r}_2\underset{\sim}{x} \\ \vdots \\ \underset{\sim}{r}_n\underset{\sim}{x} \end{pmatrix} = \begin{pmatrix} \underset{\sim}{r}_1^T \cdot \underset{\sim}{x} \\ \underset{\sim}{r}_2^T \cdot \underset{\sim}{x} \\ \vdots \\ \underset{\sim}{r}_n^T \cdot \underset{\sim}{x} \end{pmatrix}$$

This tells us that

$$A\underset{\sim}{x} = 0 \iff \underset{\sim}{r}_i^t \cdot \underset{\sim}{x} = 0, \ \forall i = 1, 2, \dots, n$$

which brings us to the conclusion that

$$\underset{\sim}{x} \in \mathcal{N}(A) \iff x \in \text{Row}(A)^\perp \iff \text{Row}(A)^\perp = \mathcal{N}(A)$$

$U^\perp$ is an example of a subspace

A nonempty set W of a vector space V is a subspace of V is it si a vector space with the same addition and scalar multiplication as V. To verify that a subset is a subspace, one checks the following

(1) $\underset{\sim}{0} \in W$

(2) $\underset{\sim}{u} + \underset{\sim}{u} \in W, \ \forall \underset{\sim}{u} \in W$

(3) $k\underset{\sim}{u} \in W, \ \forall \underset{\sim}{u} \in W, \ \forall k \in \mathbb{F}$

Now We can prove that $U^\perp$ is a subspace

2.2.2 Setting Up the Gram-Schmidt Process

Orthogonal Set

Let V be a real inner product space, a nonempty set of vectors in V is orthogonal if each vector in the set is orthogonal to all the other vectors in the set. That is, the set $\{\underset{\sim}{v}_1, \underset{\sim}{v}_1, \dots, \underset{\sim}{v}_1\} \subseteq V$ is orthogonal if

$$\left\langle \underset{\sim}{v}_i, \underset{\sim}{v}_j \right\rangle = 0, \quad i \neq j$$

Let S be a finite set of vectors in V such that $\underset{\sim}{0} \notin S$ and $|S| < \dim V$. Then

$$S \text{ orthogonal} \implies S \text{ linearly independent}$$

Proof. Let $S = \{\underset{\sim}{v}_1, \underset{\sim}{v}_2, \dots, \underset{\sim}{v}_n\} : \langle \underset{\sim}{v}_i, \underset{\sim}{v}_1 \rangle = 0, \ i \neq j, \ \|\underset{\sim}{v}_i\|^2 > 0.$

$$\implies \underset{\sim}{0} = k_1\underset{\sim}{v}_1 + \dots + k_n\underset{\sim}{v}_n \implies \forall \underset{\sim}{v}_i \in S, \ \left\langle \underset{\sim}{0}, \underset{\sim}{v}_i \right\rangle = 0 = \left\langle k_1\underset{\sim}{v}_1 + \dots + k_n\underset{\sim}{v}_n, \underset{\sim}{v}_i \right\rangle$$
$$\iff \underset{\sim}{0} = k_1 \left\langle \underset{\sim}{v}_1, \underset{\sim}{v}_i \right\rangle + k_2 \left\langle \underset{\sim}{v}_2, \underset{\sim}{v}_i \right\rangle + \dots + k_i \left\langle \underset{\sim}{v}_i, \underset{\sim}{v}_i \right\rangle$$
$$\iff \underset{\sim}{0} = k_i \|\underset{\sim}{v}_i\|^2 \iff k_i = 0 \iff S \text{ linearly independent}$$

$\square$

Orthonormal Basis

An orthogonal set of vectors in V is called orthonormal if all the vectors in the set are unit vectors, that is a set,

$$\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \ldots, \underset{\sim}{e}_n\} \subset V : \langle \underset{\sim}{e}_i, \underset{\sim}{e}_j \rangle = \delta_{i,j}$$

where the Kronecker delta is defined by

$$\delta_{i,j} = \begin{cases} 0, & i \neq j, \\ 1, & i = j \end{cases}$$

Note:-

$$A = \begin{pmatrix} \delta_{1,1} & \delta_{1,2} & \ldots & \delta_{1,n} \\ \delta_{2,1} & \delta_{2,2} & \ldots & \delta_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{m,1} & \delta_{m,2} & \ldots & \delta_{m,n} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix} = I$$

Example 2.2.3

Given $\mathbb{R}^n$ endowed with the dot product, the set

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\} \subset \mathbb{R}^3$$

is an orthonormal basis, according to $\cdot$.

The set

$$S = \left\{ \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}, \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix} \right\} \subset \mathbb{R}^3$$

is orthonormal but not a basis.

The set

$$S \cup \left\{ \begin{pmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{6}} \end{pmatrix} \right\} \subset \mathbb{R}^3$$

is both orthonormal and forms a basis in $\mathbb{R}^3$.

Definition (Orthonormal Basis). An orthonormal basis for V is a basis or V that is also an orthonormal set.

Decomposition Theorem

Let $S = \{\underset{\sim}{e}_1, \ldots, \underset{\sim}{e}_n\}$ be an orthonormal basis for V and let $\underset{\sim}{u} \in V$. Then

$$\underset{\sim}{u} = \langle \underset{\sim}{u}, \underset{\sim}{e}_1 \rangle^2 \underset{\sim}{e}_1 + \cdots + \langle \underset{\sim}{u}, \underset{\sim}{e}_n \rangle^2 \underset{\sim}{e}_n$$

and

$$\|\underset{\sim}{u}\|^2 = \langle \underset{\sim}{u}, \underset{\sim}{e}_1 \rangle^2 + \cdots + \langle \underset{\sim}{u}, \underset{\sim}{e}_n \rangle^2$$

Note:-

This is, basically, what we do when we write a vector $\underset{\sim}{u} = (1, 2) \in \mathbb{R}^2$ as $\underset{\sim}{u} = 1\hat{\imath} + 2\hat{\jmath}$ and $\|\underset{\sim}{u}\|^2 = a^2 + b^2$.

We've decomposed the vector $\underset{\sim}{u}$ into scalars multiplied by basis vectors.

Proof. Suppose S is a basis
(1)

$$\implies \forall \underset{\sim}{u} \in V,\ \underset{\sim}{u} = \sum_{i=1}^{n} a_i \underset{\sim}{e}_i,\ a_i \in \mathbb{F} \text{ is unique.}$$

$$\begin{aligned}
\implies \left\langle \underset{\sim}{u}, \underset{\sim}{e_i} \right\rangle &= \left\langle \sum_{j=1}^{n} a_j \underset{\sim}{e_i}, \underset{\sim}{e_i} \right\rangle \\
&= \sum_{j=1}^{n} a_j \left\langle \underset{\sim}{e_i}, \underset{\sim}{e_i} \right\rangle \\
&= \sum_{j=1}^{n} a_j \delta_{j,i} \\
&= a_j \left\langle \underset{\sim}{e_j}, \underset{\sim}{e_j} \right\rangle \\
&= a_j \delta_{j,j} \\
&= a_j \\
\implies a_i = \left\langle \underset{\sim}{u}, \underset{\sim}{e_i} \right\rangle
\end{aligned}$$

(2)

$$\|\underset{\sim}{u}\|^2 = \left\langle \underset{\sim}{u}, \underset{\sim}{u} \right\rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} \left\langle \underset{\sim}{u}, \underset{\sim}{e}_j \right\rangle \left\langle \underset{\sim}{u}, \underset{\sim}{e}_i \right\rangle \left\langle \underset{\sim}{e}_i, \underset{\sim}{e}_j \right\rangle = \sum_{i=1}^{n} \left\langle \underset{\sim}{u}, \underset{\sim}{e}_i \right\rangle^2$$

□

Orthogonal Projection

Let U be a finite-dimensional subspace of the real inner product space V. Then each $\underset{\sim}{v} \in V$ can be written in a unique way as

$$\underset{\sim}{v} = \underset{\sim}{u} + \underset{\sim}{w},\ \underset{\sim}{u}, \underset{\sim}{w} \in U^{\perp}$$

In the proof, we will assume that U has an orthonormal basis $S = \left\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \dots, \underset{\sim}{e}_k\right\}$

Proof. Let $\underset{\sim}{v} \in V$. $\underset{\sim}{v}$ can be expressed

$$\underset{\sim}{v} = \underset{\sim}{u} + (\underset{\sim}{v} - \underset{\sim}{u}) - \underset{\sim}{u} + \underset{\sim}{w},\ \forall \underset{\sim}{u} \in U,\ \underset{\sim}{w} = \underset{\sim}{v} - \underset{\sim}{u}$$

We want to find $\underset{\sim}{u} : \left\langle \underset{\sim}{u}, \underset{\sim}{w} \right\rangle = 0 = \left\langle \underset{\sim}{u}, \underset{\sim}{v} - \underset{\sim}{u} \right\rangle$.
$\exists S$, an orthonormal basis for U.

$$\implies \underset{\sim}{u} = a_1 \underset{\sim}{e}_1 + a_2 \underset{\sim}{e}_2 + \cdots + a_k \underset{\sim}{e}_k$$

We impose $\underset{\sim}{w} : \left\langle \underset{\sim}{e}_i, \underset{\sim}{w} \right\rangle = 0$.

$$\implies \underset{\sim}{w} = \underset{\sim}{u} - \underset{\sim}{v} = \underset{\sim}{v} - \sum_{i=1}^{n} \left\langle \underset{\sim}{u}, \underset{\sim}{e}_i \right\rangle \underset{\sim}{e}_i,\ \left\langle \underset{\sim}{u}, \underset{\sim}{w} \right\rangle = 0 \implies \underset{\sim}{w} \in U^{\perp}$$

□

Note:-

This orthonormal projection is unique.

Let $\underset{\sim}{v} = \underset{\sim}{u} + \underset{\sim}{w}$ and $\underset{\sim}{v} = \underset{\sim}{u}' + \underset{\sim}{w}'$, where $\underset{\sim}{u}, \underset{\sim}{u}' \in U$ and $\underset{\sim}{w}, \underset{\sim}{w}' \in U^{\perp}$.
$\implies \underset{\sim}{u} + \underset{\sim}{w} = \underset{\sim}{u}' + \underset{\sim}{w}' \iff U \ni \underset{\sim}{u} - \underset{\sim}{u}' = \underset{\sim}{w}' - \underset{\sim}{w} \in U^{\perp}$※$\because U \cap U^{\perp} = \left\{\underset{\sim}{0}\right\}$.

The vector $\underset{\sim}{u} \in U$ is called the orthognol projection of $\underset{\sim}{v}$ onto U and is given by

$$\mathrm{proj}_U(\underset{\sim}{v}) = \left\langle \underset{\sim}{v}, \underset{\sim}{e}_1 \right\rangle \underset{\sim}{e}_1 + \left\langle \underset{\sim}{v}, \underset{\sim}{e}_2 \right\rangle \underset{\sim}{e}_2 + \cdots + \left\langle \underset{\sim}{v}, \underset{\sim}{e}_k \right\rangle \underset{\sim}{e}_k$$

Likewise, the vector $\underset{\sim}{w} \in U^{\perp}$ is called the orthognol projection of $\underset{\sim}{v}$ onto $U^{\perp}$ and is given by

$$\operatorname{proj}_{U^{\perp}} = \underset{\sim}{v} - \operatorname{proj}_{U}(\underset{\sim}{v})$$

We take it for granted here, but it is possible to prove that

$$\dim V = \dim U + \dim U^{\perp}$$

which is a fact that can be helpful in determining the orthogonal complement of a subspace U. Indeed, suppose you have managed to find $\dim V - \dim U$ linearly independent vectors that are orthogonal to U. Then these vectors will in fact form a basis for $U^{\perp}$.

Example 2.2.4 (Orthogonal Projection in $\mathbb{R}^3$)

Let $\mathbb{R}^3$ be endowed with the usual dot product, and let

$$U = \operatorname{span}\left\{(0,1,0), \left(-\frac{4}{5}, 0, \frac{3}{5}\right)\right\}, \quad \underset{\sim}{v} = (1,1,1).$$

Find the orthogonal projections of $\underset{\sim}{v}$ onto U and $U^{\perp}$.

$$\begin{aligned}
\underset{\sim}{u} &= \operatorname{proj}_{U}(\underset{\sim}{v}) \\
&= \left\langle \underset{\sim}{v}, \underset{\sim}{e_1} \right\rangle \underset{\sim}{e_1} + \left\langle \underset{\sim}{v}, \underset{\sim}{e_2} \right\rangle \underset{\sim}{e_2} \\
&= 1\underset{\sim}{e_1} + \left(\frac{3}{6} - \frac{4}{5}\right) \underset{\sim}{e_2} \\
&= \left(\frac{4}{25}, 1, \frac{-3}{25}\right) \\
\underset{\sim}{w} &= \underset{\sim}{v} - \operatorname{proj}_{U}(\underset{\sim}{v}) \\
&= \underset{\sim}{v} - \underset{\sim}{u} \\
&= (1,1,1) - \left(\frac{4}{25}, 1, \frac{-3}{25}\right) \\
&= \left(\frac{21}{25}, 0, \frac{28}{25}\right)
\end{aligned}$$

Now, lets find that third basis vector

$$\begin{aligned}
\left\langle \underset{\sim}{w}, \underset{\sim}{e_1} \right\rangle = \left\langle \underset{\sim}{w}, \underset{\sim}{e_2} \right\rangle &= 0 \\
\therefore \underset{\sim}{e_3} &= \frac{\underset{\sim}{w}}{\|\underset{\sim}{w}\|} \\
&= \left(\frac{3}{5}, 0, \frac{4}{5}\right)
\end{aligned}$$

And so, we've found that

$$\dim U + \dim U^{\perp} = 2 + 1 = 3 = \dim \mathbb{R}^3$$
$$\implies \mathbb{R}^3 = \operatorname{span}\left\{(0,1,0), \left(\frac{-4}{5}, 0, \frac{3}{5}\right), \left(\frac{3}{5}, 0, \frac{4}{5}\right)\right\}$$

2.3 Lecture 6

2.3.1 Gram-Schmidt Process

Construction of Orthonormal Basis

In the previous example, we can see how we started with a pair of vectors, but were able to algorthmically find a third vector, to span the entire parent-vectorspace.
This can be generalised to turn a linearly independent set of vectors into an orthonormal set of vectors with the same span as the original set. Applying the algorithm to a basis thus turns the basis into an orthonormal basis. Hence:

$$\boxed{\text{Every finite-dimensional real inner product space has an orthonormal basis.}}$$

Let $\left\{\underset{\sim}{v}_1, \ldots, \underset{\sim}{v}_n\right\}$ be a linearly independent set of vectors in the real inner product space V. The corresponding algorithm is called the Gram-Schmidt process

---

Algorithm 1: Gram-Schmidt Process

Input: A linearly independent set of vectors $\left\{\underset{\sim}{v}_1, \ldots, \underset{\sim}{v}_n\right\} \subset V$
Output: An Orthonormal basis for V

Step 1

- $\underset{\sim}{w}_1 \leftarrow \underset{\sim}{v}_1;$
- $W_1 \leftarrow \text{span}\left\{\underset{\sim}{w}_1\right\} = \text{span}\left\{\underset{\sim}{e}_1\right\};$
- $\underset{\sim}{e}_1 \leftarrow \dfrac{\underset{\sim}{v}_1}{\|\underset{\sim}{v}_1\|}, \quad \|\underset{\sim}{e}_1\| = 1;$

Step 2

- $\underset{\sim}{w}_2 \leftarrow v_2 - \text{proj}_{W_1}(\underset{\sim}{v}_2) = \underset{\sim}{v}_2 - \left\langle \underset{\sim}{v}_2, \underset{\sim}{e}_1 \right\rangle \underset{\sim}{e}_1 = \underset{\sim}{v}_2 - \dfrac{\left\langle \underset{\sim}{v}_2, \underset{\sim}{w}_1 \right\rangle}{\|\underset{\sim}{w}_1\|^2} \underset{\sim}{w}_1;$
- $\implies \left\langle \underset{\sim}{w}_2, \underset{\sim}{w}_1 \right\rangle = 0;$
- $W_2 \leftarrow \text{span}\{W_1, W_2\} = \text{span}\left\{\underset{\sim}{e}_1, \underset{\sim}{e}_2\right\};$
- $\underset{\sim}{e}_2 \leftarrow \dfrac{\underset{\sim}{w}_2}{\|\underset{\sim}{w}_2\|};$

Step 3

- $\underset{\sim}{w}_3 = \underset{\sim}{v}_3 - \dfrac{\left\langle \underset{\sim}{v}_3, \underset{\sim}{w}_1 \right\rangle}{\|\underset{\sim}{w}_1\|^2} \underset{\sim}{w}_1 - \dfrac{\left\langle \underset{\sim}{v}_3, \underset{\sim}{w}_2 \right\rangle}{\|\underset{\sim}{w}_2\|^2} \underset{\sim}{w}_2 = \underset{\sim}{v}_3 - \left\langle \underset{\sim}{v}_3, \underset{\sim}{e}_1 \right\rangle \underset{\sim}{e}_1 - \left\langle \underset{\sim}{v}_3, \underset{\sim}{e}_2 \right\rangle \underset{\sim}{e}_2;$
- $\implies \left\langle \underset{\sim}{w}_3, \underset{\sim}{w}_i \right\rangle = 0, \ \forall i \in \{1, 2\};$
- $W_3 \leftarrow \text{span}\left\{\underset{\sim}{w}_1, \underset{\sim}{w}_2, \underset{\sim}{w}_3\right\};$
- $\underset{\sim}{e}_3 \leftarrow \dfrac{\underset{\sim}{w}_3}{\|\underset{\sim}{w}\|^2};$

Step n

- $\underset{\sim}{w}_n = \underset{\sim}{v}_n - \displaystyle\sum_{k=1}^{n-1} \dfrac{\left\langle \underset{\sim}{v}_n, \underset{\sim}{w}_k \right\rangle}{\|\underset{\sim}{w}_k\|^2} \underset{\sim}{w}_k = \underset{\sim}{v}_n - \sum_{k=1}^{n-1} \left\langle \underset{\sim}{v}_1, \underset{\sim}{e}_k \right\rangle \underset{\sim}{e}_k;$
- $\implies \left\langle \underset{\sim}{w}_n, \underset{\sim}{w}_i \right\rangle = 0, \ \forall i \in \{1, \ldots, n-1\};$
- $W_n \leftarrow \text{span}\left\{\underset{\sim}{w}_1, \ldots, \underset{\sim}{w}_n\right\};$
- $\underset{\sim}{e}_n \leftarrow \dfrac{\underset{\sim}{w}_n}{\|\underset{\sim}{w}_n\|^2};$

return $W_n = \text{span}\left\{\underset{\sim}{w}_1, \ldots, \underset{\sim}{w}_n\right\} = \text{span}\left\{\underset{\sim}{e}_1, \ldots, \underset{\sim}{e}_n\right\};$

Example 2.3.1

Construct an orthonormal basis for $P_1(\mathbb{R})$ with inner product

$$\left\langle \underset{\sim}{p}, \underset{\sim}{q} \right\rangle = \int_{-1}^{1} p(x)q(x)\mathrm{d}x,$$

where the basis is

$$\beta = \left\{ \underset{\underset{\sim}{v}_1}{1+x} \ , \ \underset{\underset{\sim}{v}_2}{1-2x} \right\}.$$

$$\begin{aligned}
\underset{\sim}{w}_1 &= \underset{\sim}{v}_1 = (1+x) \\
\|\underset{\sim}{w}_1\| &= \|1+x\| = \sqrt{\int_{-1}^{1} p(x)^2 \mathrm{d}x} \\
&= \sqrt{\int_{-1}^{1} 1+2x+x^2 \mathrm{d}x} \\
&= \sqrt{x+x^2+\frac{1}{3}x^3\Big|_{-1}^{1}} \\
&= \sqrt{(1)+(1)^2+\frac{1}{3}(1)^3-(-1)-(-1)^2-\frac{1}{3}(-1)^3} \\
&= \sqrt{1+1+\frac{1}{3}+1-1+\frac{1}{3}} \\
\therefore \|\underset{\sim}{w}_1\| &= \sqrt{\frac{8}{3}} = \frac{2\sqrt{3}}{3} \\
\underset{\sim}{w}_2 &= \underset{\sim}{v}_2 - \frac{\left\langle \underset{\sim}{v}_2, \underset{\sim}{w}_1 \right\rangle}{\|\underset{\sim}{w}_1\|^2} \underset{\sim}{w}_1 \\
&= (1-2x) - \frac{3}{8}(1+x) \int_{-1}^{1} (1-2x)(1+x) \mathrm{d}x \\
&= (1-2x) - \frac{3}{8}(1+x) \int_{-1}^{1} 1-x-2x^2 \ \mathrm{d}x \\
&= (1-2x) - \frac{3}{8}(1+x) \ x - \frac{1}{2}x^2 - \frac{2}{3}x^3\Big|_{-1}^{1} \\
&= (1-2x) - \frac{3}{8}(1+x) \left((1) - \frac{1}{2}(1)^2 - \frac{2}{3}(1)^3 - (-1) + \frac{1}{2}(-1)^2 + \frac{2}{3}(-1)^3 \right) \\
&= (1-2x) - \frac{3}{8}(1+x) \left(1 - \frac{1}{2} - \frac{2}{3} + 1 + \frac{1}{2} - \frac{2}{3} \right) \\
&= (1-2x) - \frac{3}{8}(1+x) \left(\frac{2}{3} \right) \\
&= (1-2x) - \frac{1}{4}(1+x) \\
\therefore \underset{\sim}{w}_2 &= \frac{3}{4}(1-3x) \\
\|\underset{\sim}{w}_2\| &= \sqrt{\left(\frac{3}{4}\right)^2 \int_{-1}^{1} (1-3x)^2 \ \mathrm{d}x} \\
&= \sqrt{\frac{9}{16} \int_{-1}^{1} 1-6x+9x^2 \ \mathrm{d}x} \\
&= \sqrt{\frac{9}{16} \ x-3x^2+3x^3\Big|_{-1}^{1}}
\end{aligned}$$

$$= \sqrt{\frac{9}{16}\left((1) - 3(1)^2 + 3(1)^3 - (-1) + 3(-1)^2 - 3(-1)^3\right)}$$

$$\therefore \|\underset{\sim}{w}_2\| = \frac{3\sqrt{2}}{2}$$

$$\underset{\sim}{e}_1 = \frac{\underset{\sim}{w}_1}{\|\underset{\sim}{w}_1\|} = \frac{\sqrt{3}}{2}(1+x)$$

$$\underset{\sim}{e}_2 = \frac{\underset{\sim}{w}_2}{\|\underset{\sim}{w}_2\|} = \frac{\sqrt{2}}{4}(1-3x)$$

$$\therefore P_1(\mathbb{R}) = \operatorname{span}\left\{\frac{\sqrt{3}}{2}(1+x), \frac{\sqrt{2}}{4}(1-3x)\right\}$$

2.3.2 Least Sqaures Approximation

A problem in linear algebra is the following:

Given a vector $\underset{\sim}{v}$ in a real inner product space, V,
give the best approsimation to $\underset{\sim}{v}$ in a finite-dimensional subspace U of V.

This is called the "least squares problem."
By "best approximation," we mean to find a vector in a subspace of minimal distance to a given vector in the ambient vector space. So the answer to the problem is to

find $\underset{\sim}{u} \in U : d(\underset{\sim}{u}, \underset{\sim}{v})$ is as small as possible.

That is, find $\underset{\sim}{u} \in U$ that minimises $\|\underset{\sim}{v} - \underset{\sim}{u}\|$.

Theorem 2.3.1 Best Approximation Theorem

If U is a finite-dimensional subspace of a real product space V, and $\underset{\sim}{v} \in V$, then $\text{proj}_U(\underset{\sim}{v})$ is the best approximation to $\underset{\sim}{v}$ from U, given by

$$\underset{\sim}{v} - \text{proj}_U(\underset{\sim}{v}) \leq \|\underset{\sim}{v} - \underset{\sim}{u}\|, \ \forall \underset{\sim}{u} \in U$$

Chapter 3

Week 3

3.1 Lecture 7

3.1.1 Applications of Least Squares Approximation

Fitting a curve to data

Expierments yield data (assuming x is unique and exact).

$$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$$

which can include some measurement error. Theory may predict a polynomial relation between x and y, but experimental data rarely matches theoretical predictions exactly. We seek a least sqaures polynomial function of best fit (a regression).

3.1.2 Eigenvalues and Eigenvectors

Non-singular matricies

For $n \times n$ sqaure matricies A, we have several