23MAT112
Class Notes

Adithya Nair

Contents

Chapter 1

Eigenvalues And Eigenvectors

1.1 Introduction

Reference - **Learning from Data, Gilbert Strang as well as Chapter 6 of Introduction To Linear Algebra**

$$A\vec{x} = \lambda \vec{x}$$

This equation is a mathematical way of expressing the idea that there's some vector $\vec{x}$ that does not change in direction, but only changes in size or magnitude by some factor λ.

> **Definition 1.1.1: Eigenvector**
>
> A vector that does not change in direction after a linear transformation.

> **Definition 1.1.2: Eigenvalue**
>
> The scalar factor by which an eigenvector changes after a linear transformation

1.2 Example Matrices

1.2.1 Projection Matrices

A reminder, projection matrices are used to bring vectors outside the column space of a given matrix A to the column space of A.

$$Ax = b$$

3 Cases

When $\vec{x}$ is in the column space,

$$Px = \lambda x, where \lambda = 1$$

When $\vec{x}$ is out of the column space,

$$It\ is\ not\ an\ eigen\ vector$$

When $\vec{x}$ is orthogonal to the column space,

$$Px = \lambda x, where \lambda = 0$$

1.2.2 Permutation Matrix

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix}$$

The eigenvectors for this matrix is:

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} and \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

1.2.3 Rotation matrix

$$R = \begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix}$$

For $\theta = \frac{\pi}{2}$

$$R = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

The calculated eigenvalues for this matrix is $\pm i$

> **Note:-**
> A matrix is the representation of a linear transformation in a given basis
> $$L(\alpha a + \beta b) = \alpha L(a) + \beta L(b)$$

Question 1

B is 3×3 matrix with eigenvalues $\lambda = 0, 1, 2$.

1. What is the rank of B?

2. What is the determinant of B?

Ans. 1. - The rank of B is 2. There is one distinct eigenvalue which is 0. This means that the nullspace is one-dimensional. The nullspace is (n-r) dimensions, therefore the rank of the matrix is 2.

Ans. 2. - The determinant is 0, since there exists an eigenvalue which is 0, which means that

$$det(A - \lambda I) = 0, \text{where } \lambda \text{ is } 0$$

1.3 Diagonalizable Matrices

> **Definition 1.3.1: Diagonalizable Matrices**
>
> Let A and B be two square matrices of size $n \times n$. We say that A and B are similar if there is an invertible matrix of the same size P such that:
> $$A = PBP^{-1}$$
> Then we can say that A is **diagonalizable** if A is similar to a diagonal matrix D

Lemma 1.3.1. Suppose that A and B are two $n \times n$ matrices and P is an invertible matrix, such that $A = PBP^{-1}$. Then, $A^n = PB^nP^{-1}$

Proof. Using the principle of mathematical induction.
We are given, $A = PBP^{-1}$ to show, $A^n = PB^nP^{-1}$ **Base step -** $n = 1, A^1 = P^1B^1P^{-1}$, which is true

Induction step - Supppose $A^n = PB^n P^{-1}$. We need to show that $A^{n+1} = PB^{n+1} P^{-1}$

$$A^{n+1} = A\dot{A}^n$$

$$= (PBP^{-1})(PB^n P^{-1})$$

$$= PBB^n P^{-1} = PB^{n+1} P^{-1}$$

$\square$

Theorem 1.3.1. Let A be an $n \times n$ matrix and let $v_1, v_2, \ldots, v_k$ be eigenvectors of A with distinct eigenvalues $\lambda_1, \lambda_2, \ldots \lambda_k$. Then $v_1, v_2 \ldots v_k$ are independent. In particular, if k = n, then $v_1, v_2 \ldots v_k$ are a basis of eigenvectors for $\mathbb{R}^n$

Proof. Suppose $v_1, v_2 \ldots v_n$ are dependent such that $\exists r_i$ such that,

$$\sum_{i=1}^k r_i v_i = 0 \tag{1.1}$$

Assume that k is minimal with this property and r_k is all non-zero,

$$A.0 = r_1 A v_1 + r_2 A v_2 \cdots + r_k A v_k$$

$$0 = r_1 \lambda_1 v_1 + r_2 \lambda_2 v_2 \ldots r_k \lambda_k v_k = 0 \tag{1.2}$$

$$\lambda_k \times ?? : r_1 \lambda_k v_1 + \cdots + r_k \lambda_k v_k = 0 \tag{1.3}$$

?? - ??

$$r_1(\lambda_k - \lambda_1)v_1 + \cdots + r_{k-1}(\lambda_k - \lambda_{k-1})v_{k-1} = 0$$

$\lambda_k - \lambda_i$ is non-zero(distinct eigenvalues)
This forms a linear combination of vectors $v_1, v_2 \ldots v_{k-1}$ that equate to zero.
This contradicts the assumption that $v_1, v_2 \ldots v_{k-1}$ is the minimal set which is dependent.
Therefore the eigenvectors are independent. $\square$

Theorem 1.3.2. Let A be a n $\times$ n matrix, Then A is diagonalizable if and only if we can find a basis $v_1, \cdots, v_n$ of eigen vectors for R^n. In this case,

$$A = PDP^1 \tag{1.4}$$

where P is the matrix whose eigenvectors $v_1, \cdots, v_n$ and D is the diagonal matrix whose diagonal entries are the corresponding eigenvalues $\lambda_1, \cdots, \lambda_n$

Proof.

$$Av_i = (PDP^{-1}v_i)$$

where P is the matrix with column vectors $v_1, v_2 \cdots v_n$ and D is a diagonal matrix with the entries $\lambda_1, \cdots, \lambda_n$

$$Av_i = (PDP^{-1})P\hat{e}_i$$

$$Av_i = PD\hat{e}_i$$

$$Av_i = P\lambda_i \hat{e}_i = \lambda_i P \hat{e}_i = \lambda_i v_i$$

This proves that v_i is the eigenvector of A, and that λ_i is the corresponding eigenvalue. Because P^{-1} exists, $v_1, \cdots, v_n$ are independent, based on the theorem proven earlier, this is the basis for R^n
Part b Suppose $v_1, \cdots, v_n$ are an eigenvector basis with corresponding eigenvalues $\lambda_1, \cdots, \lambda_n$ Suppose that P is the matrix with column vectors $v_1, \cdots, v_n$

$$\text{Let} D = P^{-1}AP$$

$$D\hat{e}_i = (P^{-1}AP)\hat{e}_i$$
$$= P^{-1}Av_i = \lambda_i P^{-1}v_i = \lambda_i \hat{e}_i$$

Thus D is the diagonal matrix with the diagonal entries λ_i ☐

1.4   Symmetric Matrices

$$A = A^T$$

A symmetric matrix has real eigenvalues and orthogonal eigenvectors.

Example.

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$$

The eigenvalues for the matrix is 2 and 4 The eigenvectors are:

$$\vec{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} and \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

These eigenvectors are orthogonal.

Definition 1.4.1: Hermitian Matrices

A complex matrix in which all the entries in the given matrix are equal to their corresponding conjugate transposes.

Theorem 1.4.1. The eigenvalues for a symmetric matrix are real.

Proof. The claim:
$$For\, Ax = \lambda x, A = A^T$$
$$\lambda \in R$$

Take:
$$Ax = \lambda x$$

And the complex conjugate:
$$\overline{A}\overline{x} = \overline{\lambda}\overline{x} \quad (\overline{A} = A)(real\, matrix)$$
$$(\overline{A}\overline{x})^T = (\overline{\lambda}\overline{x})^T \Leftrightarrow \overline{x}^T A^T = \overline{\lambda}\overline{x}^T \tag{1.5}$$

?? × x
$$\overline{x}^T A^T x = \lambda \overline{x^T} x \Leftrightarrow \overline{x}^T Ax = \overline{\lambda}\overline{x}^T x \Leftrightarrow \overline{x}^T \lambda x = \overline{\lambda}\overline{x}^T x \Leftrightarrow \lambda = \overline{\lambda}$$

Then λ is real, when $\overline{x}^T x \neq 0$ ☐

Lemma 1.4.1. Let A be a symmetric matrix. If v and w are eigenvectors with distinct eigenvalues $\lambda \& \mu$ then v & w are orthogonal.

Proof.
$$Av.w = (Av)^T w = v^T A^T w = v^T Aw = v.Aw$$
$$\Leftrightarrow \lambda v.w = \mu v.w \Leftrightarrow (\lambda - \mu)v.w = 0$$
$$\Leftrightarrow v.w = 0$$

What this means is that, when A is symmetric

$$A = PDP^{-1} = PDP^T \quad (P^{-1} = P^T \text{ for orthogonal matrices})$$

Theorem 1.4.2. Let A be a symmetric matrix. Then we can find a diagonal matrix D and an orthogonal matrix P such that,

$$A = QDQ^T$$

In particular, every symmetric matrix is diagonalisable.

1.5 Positive Definite Matrices

Definition 1.5.1: Positive Definite Matrices

Any matrix A for which the following is true:

$$\vec{x}^T A \vec{x} > 0, \forall \vec{x} \neq \vec{0}$$

is called a **positive definite matrix**

For such a matrix:

- The matrix is always a full-rank matrix.

- The eigenvalues are all positive

- $A + B$ is also a positive definite matrix, when A and B are positive definite.

- For a rectangular $m \times n$ matrix, we check for $A^T A$ or AA^T,

$$x^T(A^T A)x = (x^T A^T)(Ax) - (Ax)^T(Ax) = \|Ax\|^2$$

This means that $A^T A$ or AA^T is always a positive definite matrix, unless x lies in the nullspace of A, in which case $Ax = 0$

1.6 Similar Matrices

Refer to the definition in section **??**. A is any square matrix. When A is real & symmetric

$$A = PDP^T = QDQ^T$$

A and D are similar matrices

Example.

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, D = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$$

Note:-

Similar matrices have the same eigenvalues

Proof.

$$A\vec{x} - \lambda x$$

$$AI\vec{x} = APP^{-1}x = \lambda \vec{x}$$

$$\Rightarrow P^{-1}APP^{-1}\vec{x} = \lambda P^{-1}x$$

$$B(P^{-1}x) = \lambda(P^{-1}\vec{x})$$

□

1.7   Singular Value Decomposition

$$a = u\sigma v^t$$

where, u,v are orthogonal matrices, $\sum$ is a diagonal matrix. we know that $r(a) \subseteq r^n$ and the $c(a) \subseteq r^m$, all matrices are linear transformations. take two orthogonal vectors u_1, u_2 such that:

$$av_1 = u_1, av_2 = u_2 \text{ where r is the rank of the matrix.}$$

$$u = \begin{bmatrix} u_1 & u_2 & u_3 & \ldots & u_r & \ldots u_m \end{bmatrix}$$

$$v = \begin{bmatrix} v_1 & v_2 & v_3 \ldots v_r & \ldots & v_n \end{bmatrix}$$

$$av = u\sigma \text{ (where } \sigma \text{ are the factors of } v_i)$$

$$a = u\sigma v^{-1} = u\sigma v^t \text{ (since v is orthogonal)}$$

$$a^t a = (u\sigma v^t)^t (u\sigma v^t)$$

$$= v\sigma^2 v^t$$

1.8   Procedure For The Factorization

1. v is the eigenvectors of $a^t a$

2. the eigenvalues are the same for u and v. use $av = u\sigma$

Chapter 2

Fourier Series

2.1   A Crash Course On Infinite Series

There was this paradox, formulated by Zeno. He proclaimed that, "Suppose that someone is running, towards a particular place. He cannot reach there, unless he crosses half the distance. If I have to reach the desination, I have to cross the halfway mark of the halfway mark, and so on and so forth....". The proposition is that he will never be able to reach the destination, given this logic.

   Let us assume uniform speed, of the runner. Suppose the distance to be covered is 1 unit. The midway mark is $\frac{1}{2}$ units, then $\frac{1}{4}$, then $\frac{1}{8}$ and so on so forth. If he has a uniform speed, it will take T units for the first half to be covered. Then $\frac{T}{2}$ then $\frac{T}{4}$ and so on and so forth.

   Let's add these times.

$$T + \frac{T}{2} + \frac{T}{4} + \cdots + \frac{T}{2^{n-1}} \approx 2T$$

   One of the ways to deal with this is a concept known as **partial sum.**. What should be done is to add the first n terms.

   Taking this case,

$$s_1 = T, s_2 = \frac{3T}{2}, s_3 = \frac{7T}{4}, s_n = \frac{(2^n - 1)T}{2^{n-1}} = T(2 - \frac{1}{2^{n-1}})$$

When $n \to \infty$, $\frac{1}{2^{n-1}} \to 0$ then the total time taken is 2T.

   Let's stop assuming that the runner is moving at uniform speed.

   Let's say that the first half is covered in T time, then the next half is covered in $\frac{T}{2}$ time and $\frac{T}{3}$ time.

   The total time is:

$$T + \frac{T}{2} + \frac{T}{3} + \cdots + \frac{T}{n}$$

Taking the partial sum,

$$s_1 = T, s_2 = \frac{3T}{2}, s_3 = \frac{11T}{6}$$

   It is difficult to arrive at the general term of s_n with the traditional methods.

   Let's try a graphical analysis,

   Take the graph of y = $\frac{1}{x}$ GRAPH REQUIRED HERE.

in the limit, $\sum_{k=1}^{n} A_k > \int \frac{1}{x}$

$$\Rightarrow 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} > \int_1^{n+1} \frac{1}{x} dx = ln(n+1)$$

$$\Rightarrow s_n = T + \frac{T}{2} + \frac{T}{3} + \ldots \frac{T}{n} > Tln(n+1)$$

But this means, that Zeno was correct, the runner will never reach his destination. The logarithmic curve is not finite.

The case in which the runner was running at uniform speed is called a convergent series, which means it arrives at a finite sum.

The case in which the runner was running at a non-uniform speed, the series we obtained is called a divergent series, which means it tends towards a non-finite sum.

Earlier, this was solved using sums and approximations, until bright mathematicians such as Euler came along.

2.2 Fourier Series

The main reference is going to be - Olver, Introduction To Partial Differential Equations. Olver was a mathematician, just like his father before him. This textbook is structured in a way where the underlying frameworks behind the mathematics is covered well. This section of Fourier Series is covered in Chapter 3 and Chapter 7 of the textbook.

Calling back to the previous section

$$A\vec{x} = \lambda\vec{x}$$
$$L\vec{f}] = \lambda\vec{f}$$

This is an analogue to the way differential equations work

When certain differential problems can be solved using the eigenvalue methods, these problems are known as **Sturm-Liouville**

It is important that we learn to solve these equations by hand rather than solve them computationally.

Let's begin,

$$\vec{v}.\vec{w} = 0, \text{that implies} v \perp \vec{w}$$

Recall, when sets have two operations with certain properties they are known as vector spaces or linear spaces.

There is a subcategory of them known as inner product spaces, e.g. the dot product. These spaces have a third operation known as the inner product.

These exist for vectors. We are going to take the vector space for all functions.

$$L^P : \int fg dx = <f,g>, \text{Where P} = 2$$

Here, L^2 is the defined norm we are going to take, and $\langle f,g \rangle$ is the notation for the inner product.

When we take the inner product for a function on itself,

$$\int ff dx = \langle f,f \rangle dx = \|f\|^2$$

Lemma 2.2.1. Under the rescaled L^2 inner product, the trigonometric functions: 1, $\cos x$, $\sin x$, $\cos 2x$, $\sin 2x$, ... satisfy the following orthogonality conditions:

- $\langle cos kx, \cos lx \rangle = \langle \sin kx, \sin lx \rangle = 0 \forall k \neq l$

- $\langle \cos kx, \sin lx \rangle = 0 \ \forall \ k, l$

- $\|1\| = \sqrt{2}, \| \cos kx \| = \| \sin kx \| = 1 \forall k \neq 0$

Proof. Take the equation,

$$\frac{1}{\pi} \int_{-\pi}^{\pi} fg dx = \langle f,g \rangle$$

For the first orthogonality condition,

$$\langle \cos kx, \cos lx \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} \cos kx \cos lx dx$$

$$\int \cos(Ax)\cos(Bx)dx$$

$$= \int \cos(A+B)x + \cos(A-B)x$$

$$= \frac{\sin(A+B)x}{A+B} + \frac{\sin(A-B)x}{A-B} + C$$

Taking this to the first equation.

$$\langle \cos kx, \cos lx \rangle = \frac{1}{\pi}\left[\frac{\sin(k+l)x}{k+l} + \frac{\sin(k-l)x}{k-l}\right]^{\pi}_{-\pi}$$

Which is equal to 0

Proving,

$$\langle \cos kx \sin lx \rangle = 0 \ \forall \ k, l \tag{2.1}$$

$$\frac{1}{\pi}\int_{-\pi}^{\pi} \cos kx \cos lx dx$$

$$\cos kx \sin lx = \frac{1}{2}(\sin(k+l)x - \sin(k-l)x)$$

$$= \frac{1}{2\pi}\left[-\frac{\cos k+l}{k+l} + \frac{\cos k-l}{k-l}\right]^{\pi}_{-\pi}$$

Which equals to 0.

Proving,

$$\|1\|^2 = \sqrt{2}, \|\cos kx\| = \|\sin kx\| = 1 \forall k \neq 0$$

$$\|1\| = \frac{1}{\pi}\int_{\pi}^{\pi} 1 dx$$

$$= \frac{1}{\pi}[x]^{\pi}_{-\pi}$$

$$\frac{1}{\pi}(\pi + \pi)$$

$$\|1\|^2 = 2$$

$$\|1\| = \sqrt{2}$$

For $\|cos x\|$

$$\frac{1}{\pi}\int_{\pi}^{-\pi} (cos^2 x)$$

$$\frac{1}{\pi}\int_{\pi}^{-\pi} \frac{1 + \cos 2x}{2}$$

$\square$

> **Definition 2.2.1: Fourier Series**
>
> The Fourier series of a function $f(x)$ is defined on $-\pi \leq x \leq \pi$ is f(x) where
>
> $$f(x) \cong \frac{a_0}{2} + \sum_{k=1}^{\infty}[a_k \cos kx + b_k \sin kx]$$
>
> where coefficients are given by the inner product formulae:
>
> $$a_k = \langle f, \cos kx \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos (kx) dx \ \forall k = 0, 1, 2 \ldots$$
>
> $$b_k = \langle f, \sin kx \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin (kx) dx \ \forall k = 0, 1, 2 \ldots$$

The Fourier series is an expression of a function using an infinite series of sin and cos functions, the previous lemma proved that the functions are orthogonal to each other. Thus we can use these functions as a basis to express all functions.

> **Question 2**
>
> Find the Fourier Series for $f(x) = \{m, x > 0 \text{ and } -m, x < 0\}$

> **Solution:-**
>
> $$a_0 = \langle f, 1 \rangle$$
>
> $$\langle f, 1 \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)$$
>
> $$= \int_{-\pi}^{\pi} f(x) dx = \int_{-\pi}^{0} f(x) + \int_{0}^{\pi} f(x) dx$$
>
> $$= [-mx]_{\pi}^{0} + [mx]_{0}^{\pi}$$
>
> $$= [-\pi k] + \pi k$$
>
> $$= 0$$
>
> $$a_k = \langle f, \cos x \rangle$$
>
> $$\frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos (kx)$$

$$a_k = 0$$

$$b_k = \frac{2}{\pi}[1 - \cos (k\pi)]$$

2.3 Change Of Scale

Take a function f(x) such that the period p is some $2L$, where L is some natural number. This would not naturally fit within the interval of 2π. What we do is, set $x = \frac{p}{2\pi}v$ then $f(\frac{p}{2\pi}v + p) = f(\frac{p}{2\pi}(v + 2\pi))$

$$x = \frac{p}{2\pi}v$$

$$f(\frac{p}{2\pi}v + p) = f(\frac{p}{2\pi}(v + 2\pi))$$

$$x = \frac{2L}{2\pi}$$

$$x = \frac{L}{\pi}v$$

To simplify,

$$g(v + 2\pi)$$

$$g(v) = f(\frac{p}{2\pi}v)$$

$$g(v) = \frac{a_0}{2} + \sum_{k=1}^{\infty}[a_k \cos kv + b_k \sin kv]$$

g(v) is periodic with p $= 2\pi$

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty}[a_k cos(\frac{k\pi x}{L}) + b_k sin(\frac{k\pi x}{L})]$$

Finding a_0

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} g(v)dv$$

$$= \frac{1}{\pi} \int_{-L}^{L} g(\frac{\pi}{L}x)(\frac{\pi}{L}x)dx$$

$$\frac{1}{L} \int_{-L}^{L} f(x)dx = a_0$$

Finding a_k

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} g(v) \cos kvdv$$

$$\frac{1}{L} \int_{-L}^{L} g(\frac{\pi x}{L}) \cos \frac{k\pi x}{L}(\frac{\pi}{L})dx$$

$$= \frac{1}{L} \int_{-L}^{L} f(x) \cos \frac{k\pi x}{L}dx = a_k$$

2.4   Odd And Even Functions

Definition 2.4.1: Even Function

A function such that $f(x) = f(-x)$

Definition 2.4.2: Odd Function

A function such that $f(-x) = -f(x)$

Theorem 2.4.1. The Fourier coefficients of the functions $f_1 + f_2$ are the sum of the corresponding coefficients of f_1 and f_2, The Fourier coefficients of cf is c times the Fourier coefficients of f.

$$f(x) = x + \pi, -\pi < x < \pi$$

Question 3

Find the Fourier series of $f(x) = |x|, -\pi < x < \pi$

2.5   Integrating Fourier Series

There is an underlying assumption that all the functions we are dealing with are **convergent**.

When integrating a function's fourier series. The problem arises that $\frac{a_0}{2}$ is not being integrated. We are trading summation for integration and $\frac{a_0}{2}$ is not contained within the summation.

2.6   Differentiation And Integration Of Fourier Series

2.6.1   Background Work

Half Range Expansions

Sometimes, we're given a function f(x) on some interval [0,L]. We can only find the Fourier Series of periodic functions. To convert non-periodic functions, we limit them to the range $[-\pi, \pi]$ and then extend it periodically.

When we're given a range for a given function, what we do is assume both "even" and "odd" extensions(refer to **??** and **??**.

For e.g.

$$f(x) = \sin(x), 0 \le x \le \pi,$$

Find the Fourier Series for: a) Even periodic extension b) Odd periodic extension

Ans. For even periodic extension...

$$a_0 = 2 \int_0^\pi \sin(x)dx$$

$$a_0 = 2[-\cos(x)]_0^\pi$$

$$a_0 = 2[2]$$

$$a_0 = 4$$

$$a_k = \frac{2}{\pi} \int_0^\pi \sin x \cos kx dx$$

$$a_k = \frac{2}{\pi} \int_0^\pi \frac{\sin(kx + x) + \sin(kx - x)}{2}$$

$$a_k = \frac{2}{\pi}[-\frac{1}{2(k+1)} \cos(kx + x) - \frac{1}{2(k-1)} \cos(kx - x)]_0^\pi$$

$$b_k = 0$$

$$= \frac{4}{(1 - k^2)\pi}, \text{for even k}$$

$$0, \text{for odd k}$$

Question 4

$$f(x) = \frac{2k}{L}x, 0 < x < \frac{L}{2}, \frac{2k}{L}(L - x), \frac{L}{2} < x < L$$

Find the fourier series for even and odd periodic extensions.

Solution:-

For even periodic extension

$$a_k = \frac{2}{\pi} \int_0^{\frac{L}{2}} \frac{2k}{L} x + 2 \int_{\frac{L}{2}}^{L} \frac{2k}{L} (L-x) dx$$

2.7 Piecewise Continuous Functions

The hallmark of the Fourier series is that you can have expansions for even **discontinuous** functions.

Definition 2.7.1: Piecewise Continuous Functions

A function f(x) is said to be piecewise continuous on an interval [a,b] if it is defined and continuous ecxept possibly at a finite number of points $a \leq x_1 \leq x_2 \leq \cdots \leq x_n \leq b$ Furthermore, at each point of discontinuity, we require that the left and right hand limits exists.

$$f(x_k^-) = \lim_{x \to x_k^-} f(x); f(x_k^+) = \lim_{x \to x_k^+} f(x)$$

At the ends of the domain, the left hand is ignored at a and the right hand limit is ignored at b.

We take the basis,
$$\sigma(x) = 1, x > 0 \ and \ 0, x < 0$$

Then
$$h(x) = \beta\sigma(x - \xi) = \beta, x > \xi \ and \ 0, x < \xi$$

Definition 2.7.2: Piecewise C^1

A function is called piecewise $C^1(U \subset R)$ (continuous and continuously differentiable) on the interval [a,b] except at a finite number of points $a \leq x_1 \leq x_2 \leq \cdots \leq x_n \leq b$. At each exceptional point, the left and the right hand limits of both the function and its derivative exists.

2.8 Complex Fourier Series

Refer to 3Blue1Brown's video on it

$$f(x) \cong \frac{a_0}{2} + \sum_{k=1}^{\infty}[a_k \cos kx + b_k \sin kx]$$

The complex form of this is

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{-ikx}$$

Here the vector space is complex functions. The inner product is

$$\langle f, g \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\overline{g(x)}dx$$

$$C_k = \langle f, e^{i\pi x} \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)e^{-i\pi x}$$

Note:-
Prove fundamental theorems pertaining to the project. Some projects can be on discrete mathematics. Numerical methods for solving Singular Vector Decomposition and implementing eigenvalues and eigenvectors.

$$e^{ikx} = \cos(kx) + i\sin(kx)$$
$$e^{-ikx} = \cos(kx) - i\sin(kx)$$
$$\cos(kx) = \frac{e^{ikx} + e^{-ikx}}{2}; \sin(kx) = \frac{e^{ikx} - e^{-ikx}}{2i}$$
$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{ikx}$$

$$c_k = \langle f, e^{ikx} \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-ikx} dx$$

$$\langle e^{ikx}, e^{ilx} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ikx} e^{-lx} dx$$

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(k-l)x}$$

$$\frac{1}{2\pi i} \left[\frac{e^{i(k-l)x}}{(k-l)} \right]$$

When k = l, the inner product yields 1, when k $\neq$ l then the inner product yields 0

2.9 Differentiation And Integration

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} [a_k cos(kx) + b_k sin(kx)]$$

The discussion was surrounding whether or not we can differentiate or integrate functions that have been expressed with their Fourier series... Only when these functions converge. Looking at integration first, the integral of sin x is -cos x, and the cos term gets converted to sin.

The problem arises with the constant term $\frac{a_0}{2}$, whiich is not periodic.

$$\frac{a_0}{2} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) dx$$

This term is known as the mean̈of the function. A Fourier Series is only integrable when the mean is zero.

Lemma 2.9.1. If f(x) is 2π- periodic, then its integral $g(x) = \int_0^x f(y) dy$ is 2π-periodic if and only if $\int_{-\pi}^{\pi} f(x) dx = 0$, so that f(x) has zero mean on the interval $[-\pi, \pi]$.

Theorem 2.9.1. If f is piecewise continuous and has mean zero on the interval $[-\pi, \pi]$ then its Fourier Series

$$f(x) \cong \sum_{k=1}^{\infty} [a_k cos(kx) + b_k sin(kx)]$$

can be integrated term by term to produce the Fourier Series.

$$g(x) = \int_0^x f(y) dy \cong m + \sum_{k=1}^{\infty} -\frac{b_k}{k} coskx + \frac{a_k}{k} sinkx$$

where $m = \int_{-\pi}^{\pi} g(x) \, dx$

Note:-

m is the inner product of g(x) with itself. The remarkable thing about this is it is equal to $\sum_{k=1}^{\infty} \frac{b_k}{k}$

Chapter 3

Taylor Series

This concept marks the start of multi-variable calculus.

Refer to 3Blue1Brown's videos on multi-varable calculus taken on Khan Academy, as well as MIT 18.02 SC.

Earlier, we were trying to expand a function into orthogonal functions.

Converting a function into a series, allows us to approximate them. The Taylor Series is important because it gives us the means to find the derivative of polynomials in higher dimensions.

3.1   Taylor Polynomial

> **Definition 3.1.1: Polynomials**
>
> A polynomial is an expression of x such that:
>
> $$p(x) = a_0 + a_1 x + \cdots + a_n x^n, n \geq 0$$

The goal of the Taylor Series is to approximate a function into a polynomial, known as the Taylor Polynomial.

At the functions at 0 should equal to the polynomial at 0 $x = 0$, $f(0) = p(0) = a_0$, so should the derivatives, $f'(0) = p'(0) = a_1$ and $f''(0) = p''(0) = 2a_2$ and so on so forth.

Thus for,

$$f^{(n)} x = n! a_n$$

Where

$$f^{(1)} = f' = \frac{df}{dx} \ , \quad f^{(2)} = f'' = \frac{d^2 f}{dx^2}$$

The number of terms is determined by the amount of derivatives that can be done on f(x)

> **Theorem 3.1.1.** Let f be a function with derivatives of order n, at the point x = 0. Then there exists one and only one polynomial P of degree $\leq$ n, which satisfies the (n+1) conditions.
>
> $$P(0) = f(0), P^{(1)}(0) = f^{(1)}(0), \ldots P^{(n)}(0) = f^{(n)}(0)$$
>
> $$P(x) = \sum_{k=0}^{n} \frac{f^{(n)}(0)}{n!} x^k = f(0) + f^{(1)}(0) + x + \frac{f^{(2)}(0)}{2} x^2 + \ldots \frac{f^{(n)}(0)}{n!} x^n$$

For a Taylor polynomial at x = a, we take $(x - a)$ instead to get the coefficients again, since $x - a$ is equal to zero.

$$P(x) = a_0 + a_1(x - a) + a_2(x - a)^2 + \cdots + a_n(x - a)^n$$

$$P(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!} (x - a)^k$$

18

3.1.1 Notation

$$P = T_n f(x)$$

Where, P is the Taylor polynomial and T_n is the Taylor operator and x is the point at which the polynomial is evaluated.

> **Example.** $T_n[e^x]$ at $x = 0$
> $$P = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!}$$

> **Example.** $T_n[\sin x]$ at $x = 0$
> $$x - \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!}$$

> **Example.** $T_n[\cos x]$ at $x = 0$
> $$1 - \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots + (-1)^n \frac{x^{2n}}{2n!}$$

3.2 Calculus Of Taylor Polynomials

> **Theorem 3.2.1.** The Taylor operator T_n has the following properties:
>
> 1. If c_1 and c_2 are constants, then $T_n(c_1 f + c_2 g) = c_1 T_n(f) + c_2 T_n(g)$
>
> 2. Differentiation: The derivative of a Taylor polynomial of f is a Taylor polynomial of f', in fact, we have $(T_n f)' = T_{n-1}(f')$
>
> 3. Integration: An indefinite integral of a Taylor polynomial of f is a Taylor polynomial of an indefinite integral of f. More precisely, if $g(x) : \int_a^x f(t)dt$ then we have $T_{n+1}[g(x)] = \int T_n f(x)dx$

> **Theorem 3.2.2** (The substitution property)**.** Let $g(x) = f(cx)$, where c is a constant. Then we have,
>
> $$T_n[g(x; a)] = T_n[f(cx; a)]$$
>
> In particular, when a = 0,
> $$T_n[g(x)] = T_n[f(cx)]$$
>
> **Proof.**
> $$g(x) = f(cx); g^k(c^k)f^k(cx)$$
> $$\sum_{k=0}^{n} \frac{f^k(ca)}{k!}(cx - ca)^k$$
> $$= T_n[f(cx; ca)]$$
>
> $\square$

> **Theorem 3.2.3.** Let P_n be a polynomial of degree $n \geq 1$. Let f and g be two functions with derivatives of order n at 0 and assume that
> $$f(x) = P_n(x) + x^n g(x),$$
> where $g(x) \to 0$ as $x \to 0$. Then P_n is the Taylor Polynomial generated by f at 0

$$\frac{1}{1-x} = 1 + x + x^2 + \cdots + x^n + \frac{x^{n+1}}{1-x}$$

$$T_n[\frac{1}{1-x}] = 1 + x + x^2 + \cdots + x^n$$

The integral of this function,

$$T_{n+1}[-ln(1-x)] = x + \frac{x^2}{2} + \frac{x^3}{3} + \cdots + \frac{x^{n+1}}{N+1}$$

3.3 Error

Since the Taylor polynomial is nothing but an approximation, we must understand just how much the approximation loses in information. That is denoted by $E(x)$ or the error.

$$f(x) - \sum_{k=0}^{n} \frac{f^{(k)}a}{k!}(x-a)^k = E(x)$$

Theorem 3.3.1. Assume $f(x)$ has a continuous second derivative $f''(x)(f \in C^2)$ in some neighbourhood of a constant a. Then, for every x in this neighbourhood, we have $f(x) = f(a) + f'(a)(x-a) + E_1(x)$, where,

$$E_1 = \int_a^x (x-t)f''(t)dt$$

Proof.

$$f(x) - f(a) - f'(a)(x-a) = E_1(x)$$

$$\int_a^x f'(t)dt - f'a(x-a)$$

$$\int_a^x f'(t)dt - f'a\int_a^x dt[\text{Conversion into definite integral form}]$$

$$\int_a^x [f'(t) - f'(a)]dt = E_1(x)[\text{Integration is a linear operator}]$$

$$\int_a^x udv = [f'(t) - f'(a)](t-x) - \int_a^x (t-x)f''(t)dt$$

$$= \int_a^x (x-t)f''(t)dt[\text{Applying } \int uv \text{ rule }]$$

$\square$

Theorem 3.3.2. Assume $f(x)$ has a continuous $(n+1)^{th}$ derivative $f^{(n+1)}(x)(f \in C^{n+1})$ in some neighbourhood of a constant a. Then, for every x in this neighbourhood, we have $f(x) = \sum_{k=0}^{n} \frac{f^{(k)}}{k!}(x-a)^k + E_n(x)$, where,

$$E_n = \frac{1}{n!} \int_a^x (x-t)^n f^{(n+1)}(t)dt$$

Proof. Using mathematical induction.
i) Base case, n = 1 (Proven in previous theorem)
ii) Inductive step, Assume E_n is true, we have to show E_{n+1} is true.

$$E_{n+1}(x) - E_n(x) = -\frac{f^{(n+1)(a)}}{(n+1)!}(x-a)^{n+1}$$

$$E_{n+1}(x) = \frac{1}{n!} \int_a^x (x-t)^n f^{(n+1)}(x)dt - \left(-\frac{f^{(n+1)(a)}}{n+1!}(x-a)^{n+1}\right)$$

$$= \frac{1}{n!} \int_a^x (x-t)^n f^{(n+1)}(t)dt - -\frac{f^{(n+1)}(a)}{n!} \int_a^x (x-t)^n dt$$

$$\frac{1}{n!} \int_a^x (x-t)^n [f^{(n+1)}(t) - f^{(n+1)}(a)dt]$$

$$u = f^{(n+1)}(t) - f^{(n+1)}(a)$$

$$Taking \frac{dv}{dt} = -(x-t)^n$$

Applying uv rule,

$$E_{n+1}(x) = \frac{1}{n!} \int_a^x [f^{n+1}(x) - f^{n+1}(a)]\frac{(x-t)^{n+1}}{n+1} + \frac{1}{n!} \int_a^x \frac{(x-t)^{n+1}}{n+1} f^{(n+2)}(t)dt$$

We get,

$$E_n = \frac{1}{n!} \int_a^x (x-t)^n f^{(n+1)}(t)dt$$

$\square$

Note:-

The word neighbourhood here means it's just an open interval.

$$(a - \varepsilon, a + \varepsilon)$$

3.4   Error Estimation

If we know the bounds on the derivative of a function, we can estimate the boundaries on the error of the function.

Theorem 3.4.1. If the $(n+1)^{th}$ derivative of f satisfies the inequalities,

$$m \leq f^{(n+1)}(t) \leq M$$

for all t in some interval containing a, then for every x in this interval, we have the following inequalities/estimates

$$\frac{m(x-a)^{n+1}}{(n+1)!} \leq E_n(x) \leq \frac{M(x-a)^{n+1}}{(n+1)!} \text{ if } x > a$$

$$\frac{m(a-x)}{(n+1)!} \leq (-1)^{n+1} E_n(x) \leq \frac{M(a-x)^{n+1}}{(n+1)!} \text{ if } x < a$$

Proof. i) $x > a, t \in [a,x], then (x-a)^n > 0$ We know,

$$m \leq f^{(n+1)}(t) \leq M$$

Performing an integral operation and multiplying by $\frac{(x-a)^{n+1}}{(n+1)!)}$

$$\int_a^x m \frac{(x-t)^n}{n!} dt \leq \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)} dt \leq \int_a^x M \frac{(x-t)^n}{n!} dt$$

Upon integration, and substituting $E_n(x)$

$$\frac{m(x-a)^{n+1}}{(n+1)!} \leq E_n(x) \leq \frac{M(x-a)^{n+1}}{(n+1)!}$$

ii) $x < a, t \in [x,a], (-1)^n (x-t)^n \geq 0$

$$m \int_x^a (-1)^n \frac{(x-t)^n}{n!} \leq (-1)^n \int_x^a f^{(n+1)} \frac{(x-t)^n}{n!} \leq (-1)^n \int_x^a M \frac{(x-t)^{n+1}}{n!} dt$$

Upon applying the limits and switching the limits for the middlemost term, we get

$$\frac{m(a-x)}{(n+1)!} \leq (-1)^{n+1} E_n(x) \leq \frac{M(a-x)^{n+1}}{(n+1)!}$$

$\square$

The Taylor Polynomial is used in the approximation of a polynomial. For a practical use, we will solve the integral, this integral has no formal way of solving for.

$$\frac{2}{\sqrt{\pi}} \int_0^{\frac{1}{2}} e^{-t^2} dt = erf(x) [Error Function]$$

Taking the case of e^x We know that,

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24}$$

Taking some interval [-c,0] , $c > 0 \in R$ When $x < 0$

$$m \leq f^{(n+1)}(x) \leq M$$

$$e^{-c} \leq f^{(n+1)}(x) \leq 1$$

$$e^{-c} \leq e^x \leq 1$$

Now we can use the theorem to find the bounds of the error function.

$$\frac{e^{-c} - x^5}{5!} \le (-1)^5 E_4(x) \le \frac{(-x)^5}{5!}$$

Since c is an arbitrary number and e^{-c} forms an almost negligible value, we take that entire term to be zero, and we are more interested in the upper bound.

$$0 \le (-1)^5 E_4(x) \le \frac{(-x)^5}{5!}$$

Substituing $-t^2$ instead of x

$$E_4(-t^2) \ge -\frac{t^{10}}{5!}$$

$$0 \le \frac{t^1 0}{5!} \le \frac{1}{2^{10}} \frac{1}{5!} = 8.14x10^{-6}$$

$$0 \le \frac{t^{10}}{5!} \le 8.14 \times 10^{-6}$$

Substituting $-t^2$ in the Taylor polynomial, we get the final answer to be 0.4612548616898148

3.5   Little-Oh!

We learned that we can write a function $f(x)$ as

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + E_n(x)$$

Taking an interval in the neighbourhood of a constant a, $a - c, a + c$, if the function $f(x)$ is bounded by this interval, then the Error function is also bounded by,

$$f^{n+1}(x) \le M$$

$$|E_n(x)| \le \frac{M|(x-a)^{n+1}|}{(n+1)!}$$

$$0 \le |\frac{E_n(x)}{x-a}| \le \frac{M|x-a|^n}{(n+1)!}$$

When $x \to a$ then usually the value becomes undefined but, on the right hand side, the function is tending to zero faster.

What this means is that,

$$\frac{E_n(x)}{(x-a)^n} \to 0$$

$$E_n(x) = o((x-a)^n$$

<div>

Definition 3.5.1

Assumg $g(x) \ne 0$ for all $x \ne a$ in some interval containing a (neighbourhood of a). The notation $f(x) = o(g(x))$ as $x \to a$ means that

$$\lim_{x \to a} \frac{f(x)}{g(x)} = 0$$

</div>

$f(x)$ may be approximated near a by a polynomial in $(x-a)$ of degree n, and the error in this approximation is of smaller order than $(x-a)^n$ as $x \to a$

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + o(x-a)^n \, as \, x \to a$$

The further away you get from a, the less accurate the Taylor polynomial becomes

Example.
$$f(x) = o(1) \text{ as } x \to a$$
$$\lim_{x \to a} f(x) = 0$$

Example.
$$f(x) = o(x) \text{ as } x \to 0$$
$$\lim_{x \to 0} \frac{f(x)}{x} = 0$$

Example.
$$f(x) = h(x) + o(g(x))$$
$$\lim_{x \to a} \frac{f(x) - h(x)}{g(x)} = 0$$

Example.
$$\sin x = x + o(x) \text{ as } x \to 0$$
$$\lim_{x \to 0} \frac{\sin x - x}{x} = 0$$
$$\lim_{x \to 0} \frac{\sin x}{x} = 1$$

Theorem 3.5.1 (Algebra of Little-o)**.** As $x \to a$, we have the following,

$$a) o(g(x)) \pm ox)(g(x)) = o(g(x))$$
$$b) o(c(g(x)) = o(g(x) \text{ if } c \neq 0$$
$$c) f(x).o(g(x)) = o(f(x).g(x))$$
$$d) o(o(g(x)) = o(g(x)$$
$$e) \frac{1}{1 + g(x)} = 1 - g(x) + o(g(x)) \text{ if } g(x) \to 0 as x \to a$$

Example.
$$tanx = x + \frac{1}{3}x^3 + o(x^3)$$

1. Taylor Series will be done as a video

2. Fourier Transforms will be done as a video

3. Laplace Transforms will be done as a video

Chapter 4

Optimization

The background needed here is single variable calculus(18.01SC from MIT). What we require for the course is multi-variable calculus.

4.1 Resources To Refer To

1. Robert Ghrist - series on single variable and multivariable calculus.

2. 3Blue1Brown - a series of videos on multivariable calculus.

4.2 Differential calculus of scalar and vector fields

We're dealing with transformations.
$$T : V \to W$$
These transformations are either linear or nonlinear. Here V and W are subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively.

Taking for example, Where n and m = 1. An example would be $y = f(x)$, where $x \in \mathbb{R}, y \in \mathbb{R}$. This would be known as a real valued function of a real variable.

When n = 1, m > 1: Vector valued function of a real variable.

When n > 1, m = 1: Real valued function of a vector variable. Examples include temperature and norm. (Scalar field)

When n and m > 1: Vector valued functions of a vector variable. (Vector field)

On dealing with spaces, we use notations to represent a vector.

1. $R^2 : (x, y)$

2. $R^3 : (x, y, z)$

3. $R^n : (x_1, \ldots, x_n)$

4.3 Open ball and open sets

> **Note:-**
>
> In single variable calculus, we have a notation for representing intervals using (a,b) and [a,b]. We use these intervals when discussing differentiability over an open interval. While continuity is measured over a closed interval. **Why? Homework for today.** My answer - Differentiability can only be checked on a continuous function. If continuity is true from $[a, b]$ then we cannot say anything about the LHL and RHL of a and b respectively, extending the argument further, we cannot say anything about the LHD and RHD of a and b.

- Let $\underline{a}$ be a given point in R^n and let r be a given positive number. The set of all points $\underline{x}$ in R^n such that

$$\|\underline{x} - \underline{a}\| < r$$

is called an open n-ball of radius r and center a.

- We denote this set by $B(\underline{a})$ or $B(\underline{a}; r)$.

    - in R^1: open interval
    - in R^2: solid disc
    - in R^3: solid sphere
    - in R^n: n-ball

Definition 4.3.1: Interior Point

Let S be a subset of $\mathbb{R}^n$, and assume that $\underline{a} \in S$. Then $\underline{a}$ is called an interior point of S if there is an open n-ball with center at $\underline{a}$, all of whose points belong to S
Notation: int S(the set of all interior points of S)

Definition 4.3.2: Open Set

A set S in $\mathbb{R}^n$ is open if all its points are interior points. In other words, S is open if and only if S = int S

An open set containing point $\underline{a}$ is also called as neighborhood of $\underline{a}$

If A_1 and A_2 are subsets of R^1, their Cartesian product $A_1 \times A_2$ is the set in R^2 defined by

$$A_1 \times A_2 = \{(a_1, a_2) | a_1 \in A_1, a_2 \in A_2\}$$

Theorem 4.3.1. If A_1 and A_2 are open sets, then $A_1 \times A_2$ is also open in R^2

Proof. There is a 1-ball around a_1 in A_1 and for a_2 in A_2. These balls are represented by $B(a_1; r_1)$ and $B(a_2; r_2)$. Where r = $\min(r_1, r_2)$. The 2-ball is $x_1, x_2 \in B(a; r)$

$$\|x - a\| < r \rightarrow |x_1 - a_1| < r_1$$
$$|x_2 - a_2| < r_2$$
$$x_1 \in B(a_1; r_1) \rightarrow x_1 \in A_1$$
$$x_2 \in B(a_2; r_2) \rightarrow x_2 \in A_2$$

Since x_1 and x_2 are interior points of A_1 and A_2, the cartesian product is $A_1 \times A_2$ $\qquad\square$

We were talking about transformations on vector spaces, but on specific vector spaces with an additional structure added. These vector spaces are known as **normed linear spaces** or vector spaces with an inner product operation.

Refresher:

$$T : V \rightarrow W, V \subseteq R^n, W \subseteq R^m$$

Definition 4.3.3: Exterior And Boundary

A point x is said to be exterior to a set S in $\mathbb{R}^n$ if there is an n-ball $B(\vec{x})$ containing no points of S. The set of all points in $\mathbb{R}^n$ exterior to S is called the exterior of S and is denoted by ext S.

$$\|x - a\| > r$$

A point which is neither exterior nor interior is said to be the boundary point of S.
The set of all boundary points of the set S is called the boundary of S. It is denoted by ∂S

$$\|x - a\| = r$$

4.4   Limits And Continuity

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

Suppose a small number ε ($\varepsilon > 0, \varepsilon \in \mathbb{R}$) and $|x - a| < \varepsilon \to |f(x) - b| < \delta$) where δ and ε are two positive small numbers.

This is what a limit means. Consider a function $f : S \to \mathbb{R}^m$ where $S \subseteq R^n$. If $\vec{a} \in \mathbb{R}^n$ and $\vec{b} \in \mathbb{R}^m$, we write

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

to mean that

$$\lim_{\|x-a\|} \|f(\vec{x} - \vec{b}) = 0$$

In $\mathbb{R}^2$:

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

In $\mathbb{R}^3$:

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

A function is said to be continuous at $\vec{a}$ if f is defined at $\vec{a}$ and if

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

Theorem 4.4.1. If $\lim_{x \to a} f(x) = b$ and $\lim_{x \to a} g(x) = c$

1. $\lim_{x \to a}[f(x) + g(x)] = b + c$

2. $\lim_{x \to a}[\lambda(x)] = \lambda b$

3. $\lim_{x \to a} f(x) \cdot g(x) = b \cdot c$

4. $\lim_{x \to a} \|f(x)\| = \|b\|$

A function f is continuous at a point, if and only if each component f_k is continuous at that point.

$$\vec{f} = (f_1(x), f_2(x), \ldots, f_m(x))$$

Theorem 4.4.2. Let f and g be the functions such that the composite function fog is defined at a, where

$$(fog)(\vec{x}) = \vec{f}[\vec{g}(x)]$$

If g is continuous at a and if f is continuous at $g(a)$, then the conmposition fog is continous at a.

Example - $sin(x^2y), \log(x^2 + y^2), \frac{e^{x+y}}{x+y}$

Take this example

$$f(x, y) = \frac{xy}{x^2 + y^2} \text{when } (x, y) \neq 0$$
$$= 0, \text{when } (x, y) = 0$$

When taking the limit on this function, $\lim_{x \to a} f(x)$ Take the x-axis. Then the function would give 0. Take the y-axis, then the function would also give 0. But, when $x = y$, the function gives $\frac{1}{2}$ until $x = y = 0$, the function gives 0.

Therefore this function is not continuous.

4.5 Differentiation(Scalar Fields)

A scalar field is a real valued function taking a vector variable.

$$T : V \to W, \text{where } V \subseteq \mathbb{R}^n, W \subseteq \mathbb{R}$$

The function,

$$f(\vec{x}) \quad f : S \to \mathbb{R}, S \subseteq \mathbb{R}^n$$

$\vec{a}$ is an interior point of S.
In the figure,

$$\|\vec{a} - hy - \vec{a}\| = \|h\vec{y}\| = |h|\|\vec{y}\|$$

Where,

$$|h|\|\vec{y}\| < r$$

Or $|h|\|y\|$ is an interior point
We know that for single-variable calculus,

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

This expression is known as the difference quotient.
Using these terms, we can arrive at a definition for the difference quotient and derivative for a scalar field.

$$\frac{f(\vec{a} + h\vec{y}) - f(\vec{a})}{h}$$

Definition 4.5.1: Derivative

Given a scalar field, $f : S \to \mathbb{R}$, where $S \subseteq \mathbb{R}^n$. Let a be an interior point of S and let $\vec{y}$ be an arbitrary point in $\mathbb{R}^n$. The derivative of f at a with respect to $\vec{y}$ is denoted by the symbol $f'(a; y)$ and is defined by the equation,

$$f'(a; y) = \lim_{h \to 0} \frac{f(\vec{a} + h\vec{y} - f(\vec{a}))}{h}$$

when the the limit exists.

---

e.g. Let's say the direction $\vec{y} = 0$

$$f'(a; 0) = \vec{0}$$

e.g. f is linear

$$f(\alpha a + \beta b) = \alpha f(a) + \beta f(b)$$

$$f'(\vec{a}; y) = f(\vec{y})$$

e.g. $f(\vec{x}) = \|x\|^2$

$$f(\vec{x} + h\vec{y}) = \|\vec{x} + h\vec{y}\|^2$$
$$= (\vec{x} + h\vec{y}) \cdot (x + hy)$$
$$= \|\vec{x}\|^2 + 2h\vec{x} \cdot \vec{y} + h^2\|\vec{y}\|^2$$
$$f(x) = \|x^2\|$$
$$f'(x; y) = \lim_{h \to 0} \frac{2h\vec{x} \cdot \vec{y} + h\|y\|^2}{h} = 2x \cdot \vec{y} + h\|y\| = 2x \cdot y$$

28

4.6 Directional Derivative

> **Definition 4.6.1**
>
> If $\vec{y}$ is a unit vector, the derivative $f'(\vec{a}; \vec{y})$ is called the directional derivative of f at $\vec{a}$. In particular, if $\vec{y} = \hat{e}^k$(the kth coordinate unit vector) The directional derivative $f'(\vec{a}; \hat{e}_k)$ is called the partial derivative with respect to $\hat{e}_k$ and is also denoted by $D_k f(\vec{a})$. Thus,
>
> $$D_k f(\vec{a}) = f'(\vec{a}; \hat{e}_k)$$
>
> $$D_1 f(\vec{a}) = \frac{\partial f(\vec{a})}{\partial x_1}$$

For derivatives of multiple variables,

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial f}{\partial y \partial x} = D_{21} f(\vec{a})$$

For the function $y = f(x)$, let's say we remove the limit from the difference quotient, we'd still need to account for the error.

$$f'(a) = \left[\frac{f(a+h) - f(a)}{h}\right] + E(a, h)$$

$$\Rightarrow f'(a) - \frac{f(a+h) - f(a)}{h}] = E(a, h)$$

As $h \to 0, E(a, h) \to 0$

$$f(a+h) - f(a) = \frac{f(a+h) - f(a)}{h}.h = (f'(a)0 = 0)$$

$$\Rightarrow f(a+h) = f(a) + hf'(a) + hE(a, h)$$

Referring back to chapter 3, this is a way to write $o(x)$ notation.

$$f(a+h) - f(a) = hf'(a) + o(h)$$

Taking this to multivariable calculus. The argument can be made,

$$f(\vec{a} + h\vec{y}) - f(\vec{a}) = \left(\frac{f(a + hy) - f(a)}{h}.h\right)$$

The argument can be made that when h tends to 0, $E(\vec{a}; h)$ also tends to zero. e.g.

$$f(x, y) = \left\{\frac{xy^2}{x^2 + y^4}; (x, y \neq 0)\right\}$$

$$= \{0; \ (x, y) = (0, y)\}$$

$$\vec{a} + h\vec{y} = (0, 0) + h(a, b) = h(a, b)$$

$$f(\vec{a} + h\vec{y}) = f(h(a, b)) = f(ha, hb) = \frac{hab^2}{a^2 + h^2 b^4}$$

$$\frac{f(a+h) - f(a)}{h} = \frac{ab^2}{a^2 + h^2 b^4};$$

$$\lim_{h \to 0} f'(a; y) = \frac{b^2}{a}$$

The problem arises when we look at the function. We have made the argument that when the derivative of a function exists, it implies continuity. But that does not seem to be true in multivariable calculus.

Take this function, and put in $(k, \sqrt{k})$. The function gives $\frac{1}{2}$ everywhere except at 0. This means that the function is discontinuous, by our understanding of continuity.

Let's try extending the definition within single variable calculus, in such a way that the existence of a derivative implies continuity.

$$y = f(x), x, y \in \mathbb{R}$$

For single-variable calculus, this works as a definition for the derivative.

$$f(x + h) - f(x) = hf'(x) + o(h)$$

What term would fit that same description for multi-variable calculus?

$$f(\vec{a} + h\vec{v}) - f(\vec{a}) = (?) + o\|v\|$$

4.7   Total derivative

Definition 4.7.1: Total Derivative

We say that f is differentiable at $\vec{a}$, if there exists a linear transformation.

$$T_{\vec{a}} : \mathbb{R}^n \to \mathbb{R}$$

from $\mathbb{R}^n$ to $\mathbb{R}$, and a scalar function $E(\vec{a}; \vec{v})$ such that

$$f(\vec{a} + \vec{v}) = f(\vec{a}) + T_{\vec{a}}(\vec{v}) + \|\vec{v}\|E(\vec{a}; \vec{v})$$

for $\|v\| < r$, where $E(\vec{a}; \vec{v}) \to 0$ as $\|\vec{v}\| \to 0$. The linear transformation $T_{\vec{a}}$ is called the total derivative of f at $\vec{a}$

$$f(\vec{a} + \vec{v}) = f(\vec{a}) + T_{\vec{a}}(\vec{v}) + \|\vec{v}\|E(\vec{a}; \vec{v}) \qquad \text{(By definition)}$$

$$\frac{f(\vec{a} + h\vec{y} - f)\vec{a}}{h} = T_a(y) + \frac{|h|\|y\|E(\vec{a}; h\vec{y}}{h} \qquad \text{(Dividing both sides by } h\text{)}$$

$$f'(\vec{a}; \vec{y}) = T_{\vec{a}}(\vec{y}) \qquad \left(\lim_{h \to 0}\right)$$

Theorem 4.7.1. Assume f is differentiable at $\vec{a}$, with total derivative $f'(a; y)$ exists for every $\vec{y}$ in $\mathbb{R}^n$ and we have,

$$T_{\vec{a}}(\vec{y}) = f'(\vec{a}; \vec{y})$$

Moreover, $f'(a; \vec{y})$ is a **linear combination** of $\vec{y}$. In fact, if $\vec{y} = (y_1, \ldots, y_n)$, we have

$$f'(\vec{a}; \vec{y}) = \sum_{k=1}^{n} D_k f(\vec{a}) y_k$$

Proof.

$$T_{\vec{a}}(\vec{y}) = y_1 T_{\vec{a}}(\hat{e}_1) + y_2 T_{\vec{a}}(\hat{e}_2) + \cdots + y_n T_{\vec{a}}(\hat{e}_n)$$

$$T_{\vec{a}}(\sum_{k=1}^{n} y_k \hat{e}_k) = \sum_{k=1}^{n} y_k T_{\vec{a}}(\hat{e}_k)$$

$$f'(\vec{a}; \hat{e}_k) = T_{\vec{a}}(\hat{e}_k)$$

$$f'(\vec{a}; \vec{y}) = \sum_{k=1}^{n} D_k f(\vec{a}) y_k$$

> ### Definition 4.7.2: Gradient Operator
>
> The gradient operator is a linear operator that helps calculate the total derivative.
>
> $$T_{\vec{a}}(\vec{y}) = \nabla f(\vec{a}) \cdot (\vec{y})$$
>
> $$\text{where, } \nabla = \frac{\partial()}{\partial x_1}\hat{e}_1 + \frac{\partial()}{\partial x_2}\hat{e}_2 \ldots \frac{\partial()}{\partial x_n}\hat{e}_n$$
>
> $$\nabla f = \frac{\partial f}{\partial x_1}\hat{e}_1 + \frac{\partial f}{\partial x_2}\hat{e}_2 + \cdots + \frac{\partial f}{\partial x_n}\hat{e}_n$$

This gradient, is a vector. The total derivative,

$$f'(\vec{a}; \vec{y}) = \nabla f \cdot \vec{\hat{y}}$$

> ### Note:-
>
> $$\nabla \times \vec{v} \text{ is known as a curl.}$$
> $$\nabla \cdot \vec{v} \text{ is known as a divergence.}$$

If the partial derivatives for a function exist and they are continuous, then the total derivative exist. If the total derivative exists, then the partial derivatives exist.

4.8   Chain Rule

For one variable,

$$g(t) = f(p(t))$$

When there are multiple functions, we apply the chain rule.

$$g'(t) = f'p'; f' = \frac{df}{dp}, p' = \frac{dp}{dt}$$

$$\frac{dg}{dt} = \frac{df}{dp}\frac{dp}{dt}$$

Taking this to vectors,

$$g(t) = f(\vec{r}(t))$$

The derivative of a vector function,

$$\vec{r}(t) - (x(t), y(t))$$
$$\vec{r'}(t) = (x'(t), y'(t))$$

> **Theorem 4.8.1.** Let f be a scalar field defined on an open set $S \subseteq R^n$, and let $\vec{r}$ be a vector-valued function which maps an interval J from $\mathbb{R}^1$ into S. Define the composite function $g = f o \vec{r}$ on J, by the equation.
>
> $$g(t) = f(\vec{r}(t)) \text{ if } t \in J$$
>
> Let t be a point in J at which $\vec{r'}(t)$ exists and assume that f is differentiable at $\vec{r}(t)$. Then $g'(t)$ exists and is equal to the dot product,
>
> $$g'(t) = \nabla f(\vec{a}) \cdot \vec{r'}(t), \text{ where } \vec{a} = \vec{r}(t)$$

Proof.

$$g(t) = f(r(t))$$

$$g(t+h) - g(t) = f(r(t+h)) - f(r(t))$$

$$= f(\vec{y} + \vec{a}) - f(\vec{a}) \qquad\qquad y = \vec{r}(t+h) - \vec{r}(t) \text{ and } \vec{a} = \vec{r}(t)$$

$$\frac{g(t+h) - g(t)}{h} = \nabla f \cdot \frac{\vec{r}(t+h) - \vec{r}(t)}{h} + \frac{\|y\|}{h} E(\vec{a}; \vec{y}) \qquad\qquad \text{Applying limit } h \to 0$$

$$g'(t) = \nabla f \cdot r'(t) + 0$$

$\square$

32