

## 0. Contents

6.1 Sequences . . . . .	10
6.2 Series . . . . .	10

List of Tables . . . . .	1
List of Figures . . . . .	1
1 Algebra . . . . .	2
2 Graphs . . . . .	3
3 Linear Algebra . . . . .	5
3.1 Basics . . . . .	5
3.1.1 Matrix Multiplication . . . . .	5
3.2 Basis . . . . .	5
3.2.1 Converting Basis . . . . .	6
3.2.2 Rotating in a different Basis . . . . .	6
3.3 Products . . . . .	6
3.3.1 Inner Product . . . . .	6
3.3.2 Outer Product . . . . .	6
3.4 Operations . . . . .	6
3.4.1 Projections . . . . .	6
3.4.2 Rotations . . . . .	6
3.4.3 Scale . . . . .	6
3.4.4 Shear . . . . .	6
3.4.5 Pivots . . . . .	6
3.4.6 Transpose . . . . .	6
3.4.7 Conjugate . . . . .	6
3.4.8 Complex Conjugate . . . . .	6
3.5 Properties . . . . .	7
3.5.1 Orthogonoal . . . . .	7
3.5.2 Positive . . . . .	7
3.6 Spaces . . . . .	7
3.7 Eigenvectors . . . . .	7
3.7.1 Eigenbasis . . . . .	8
3.8 Factorization . . . . .	8
3.8.1 Gram-Schmidt . . . . .	8
3.8.2 Singular Value Decomposition . . . . .	8
4 Calculus . . . . .	8
5 Multivariable Calculus . . . . .	9
6 Sequences , Series . . . . .	10

## 0. List of Tables

## 0. List of Figures

# 1. Algebra

## Logarithms

- $\log_e = \ln$
- $\log_a a = 1$
- $\log_a (1) = 0$
- $\log_a \left(\frac{1}{a}\right) = -1$
- $\log_a (n) + \log_a (m) = \log_a (mn)$
- $\log_a (n) - \log_a (m) = \log_a \left(\frac{n}{m}\right)$
- $\log_a (x^n) = n \cdot \log_a (x)$
- $\log_a \left(\frac{1}{y}\right) = -\log_a (y)$

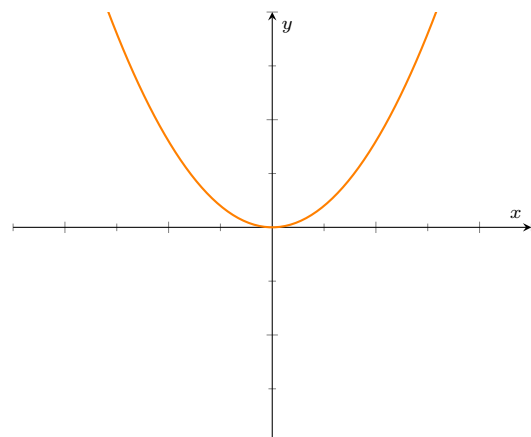
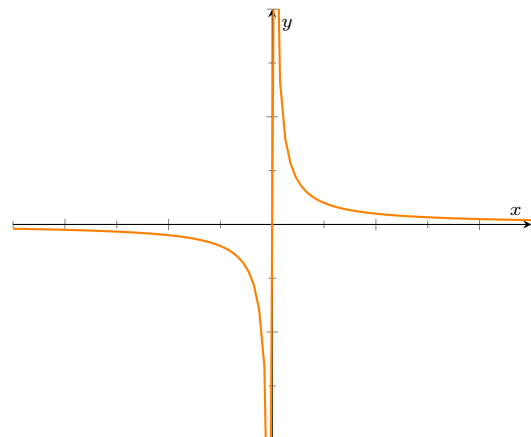
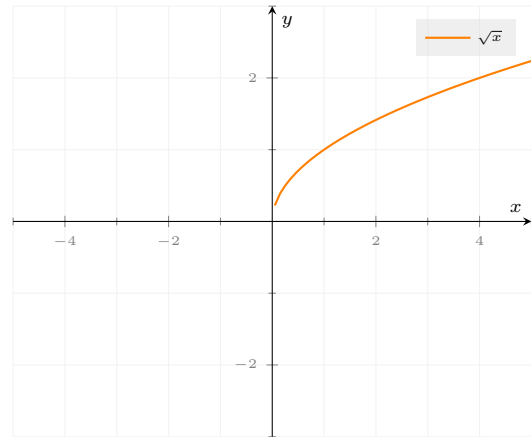
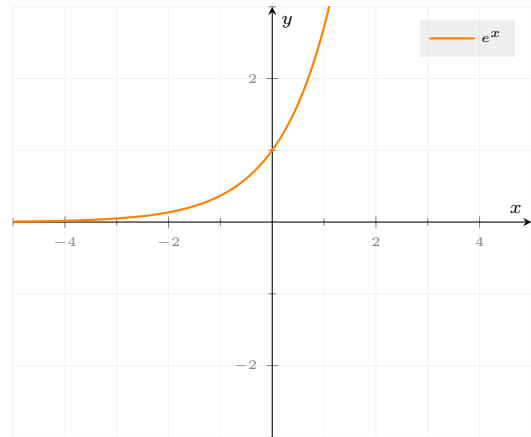
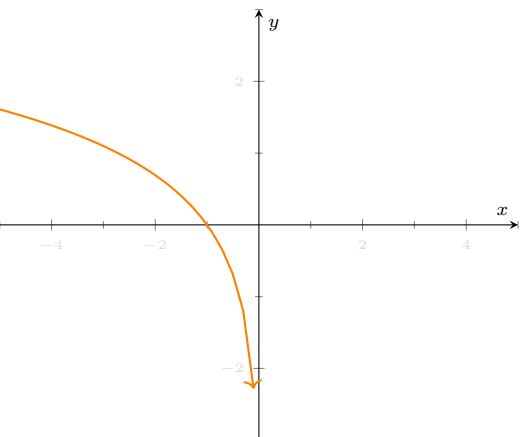
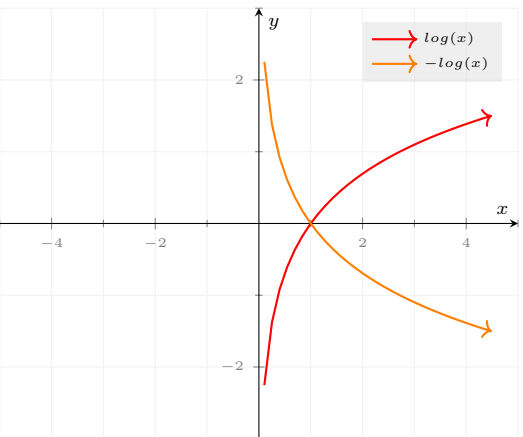
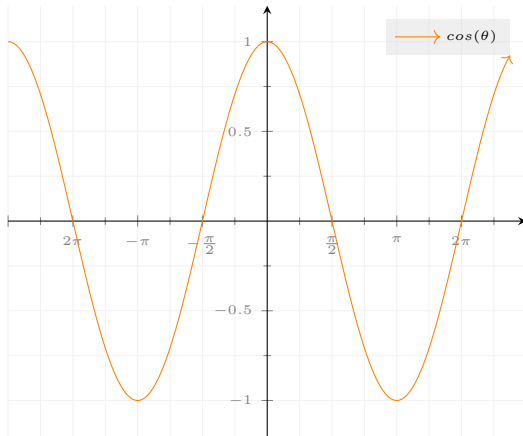
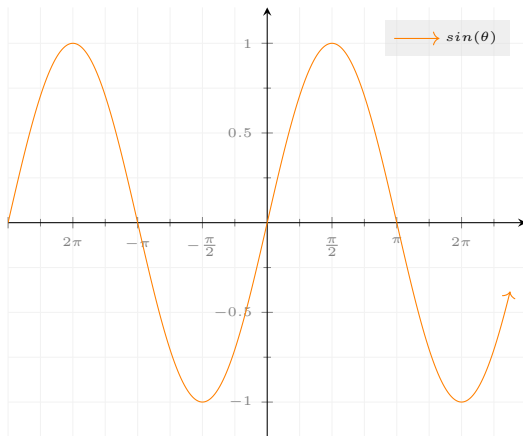
## Exponentials

- $a^m \cdot a^n = a^{m+n}$
- $a^0 = 1 \quad (a \neq 0)$
- $a^{-n} = \frac{1}{a^n} \quad (a \neq 0)$
- $\frac{a^m}{a^n} = a^{m-n} \quad (a \neq 0)$
- $(a^m)^n = a^{mn}$
- $(ab)^m = a^m b^m$
- $\left(\frac{a}{b}\right)^m = \frac{a^m}{b^m} \quad (b \neq 0)$

## Trigonometry

deg	0°	30°	45°	60°	90°	180°	270°	360°
rad	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$	$\pi$	$\frac{3\pi}{2}$	$2\pi$
$\sin(\theta)$	0	$\frac{1}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{\sqrt{3}}{2}$	1	0	-1	0
$\cos(\theta)$	1	$\frac{\sqrt{3}}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{1}{2}$	0	-1	0	1
$\tan(\theta)$	0	$\frac{1}{\sqrt{3}}$	1	$\sqrt{3}$	N/A	0	N/A	1

## 2. Graphs

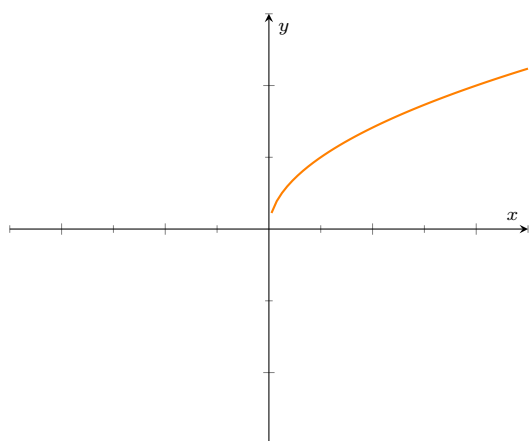


24

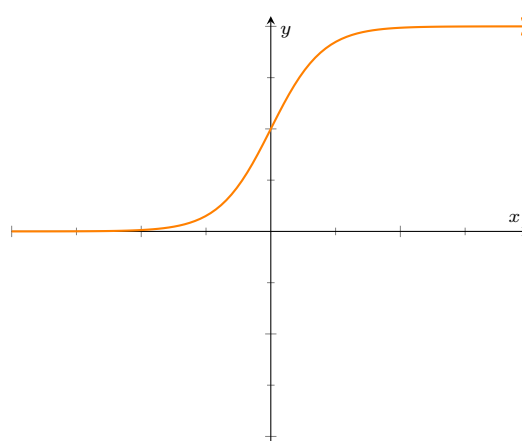
25

26

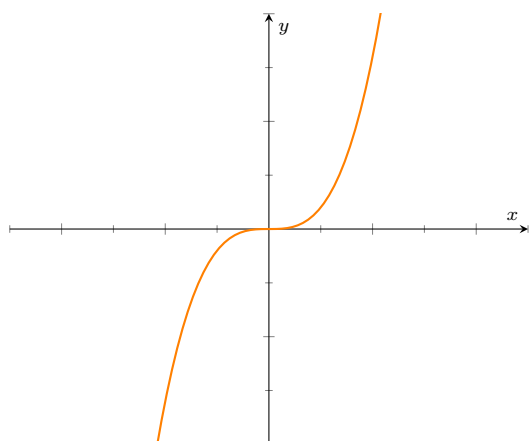
27



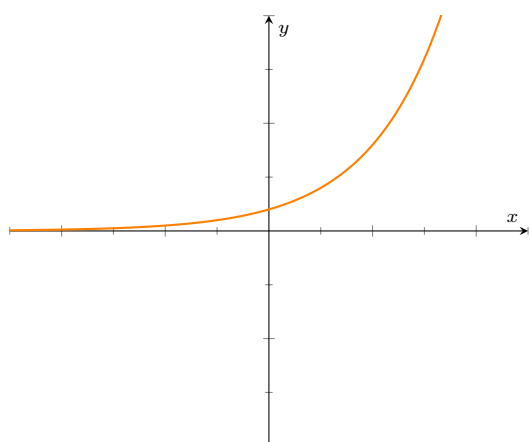
28



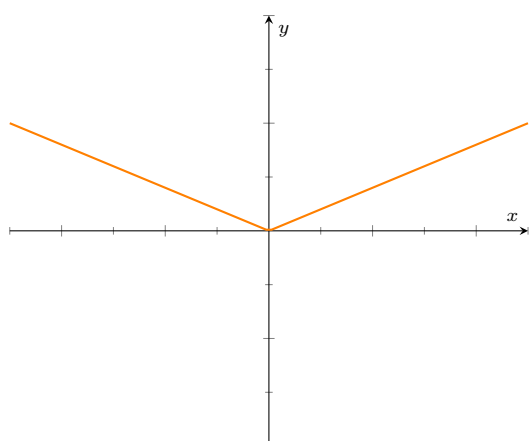
32



29



30



31

-§-

## 3. Linear Algebra

### 3.1. Basics

$$\begin{aligned}\vec{a} + \vec{b} &= \vec{b} + \vec{a} \\ 2\vec{a} &= \vec{a} + \vec{a} \\ \|\vec{a}\|^2 &= \sum_i a_i^2 \\ A_{ij}^T &= A_{ji}\end{aligned}$$

#### 3.1.1. Matrix Multiplication

Right way to think about it

33

Why is it called Linear Algebra ?

34

Linear Transformations

35

### 3.2. Basis

#### Basis Vector

##### DEFINITION 3.1.

The set of vectors that are used to define the co-ordinate system that we are operating in , are called basis vectors.

Every time that we define two vectors , we are also simultaneously making a choice for what basis we are operating in. It just so happens that most often , this basis is on the x , y plane with  $\hat{i}$  , and  $\hat{j}$  being the basis vectors.

36

37

38

39

40

#### Span

##### DEFINITION 3.2.

All the points that you can reach in a plane with the basis vectors that you chose is called the span.

The span of  $\vec{a}$  and  $\vec{b}$  is the set of all of their linear combinations , i.e. ,  $x\vec{a} + y\vec{b}$  , where  $x, y \in \mathbb{R}$ .

Usually most basis vectors will you give you the entire plane , except if they line up , i.e, if the angle between them is 0 , or if both vectors are just 0.

41

42

43

A vector is called **linearly dependent** on another if it is in the span of the other. If a vector adds another dimension then it is **linearly independent**.

44

45

46

The scalar projection of your basis vector onto another is how you find out what the scalar value of that

47

48

vector is , in the span defined by your chosen basis vector.

Basically ,  $\vec{v}$  in basis  $\vec{b}_1, \vec{b}_2$ , where  $\vec{b}_1, \vec{b}_2$  are orthogonal to each other , is given by

$$\begin{bmatrix} \frac{\vec{v} \cdot \vec{b}_1}{\|\vec{b}_1\|^2} \\ \frac{\vec{v} \cdot \vec{b}_2}{\|\vec{b}_2\|^2} \end{bmatrix}$$

Maps of images in CNNs can be thought of as basis changes and vector projections to more feature rich spaces.

### 3.2.1. Converting Basis

### 3.2.2. Rotating in a different Basis

Figure out matrix  $A$  that transforms from basis  $b_1$  to  $b_2$ .

Figure out the rotation as a transformation by matrix multiplication in basis  $b_1$ , i.e. find the matrix that you need to multiply by for that rotation.

Transform the vector that the transformation needs to be applied to , from basis  $b_2$  into basis  $b_1$ .

Apply the rotation by matrix multiplication in basis  $b_1$ .

Transformed the rotated matrix back into basis  $b_2$  by multiplying by  $A^{-1}$ .

## 3.3. Products

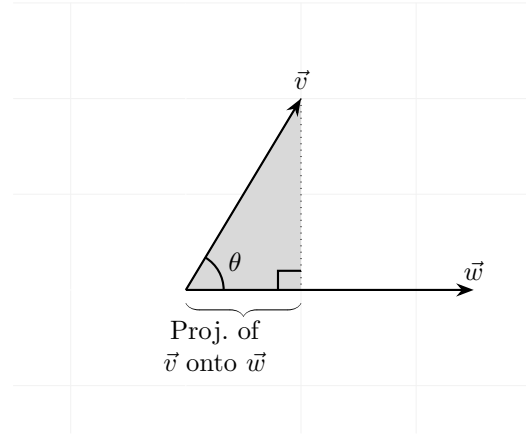
### 3.3.1. Inner Product

$$\begin{aligned} \vec{a} \cdot \vec{b} &= a_1b_1 + a_2b_2 + \dots + a_nb_n \\ &= \sum_i a_ib_i \\ \vec{a} \cdot \vec{a} &= \|\vec{a}\|^2 \\ \vec{a} \cdot \vec{b} &= \vec{b} \cdot \vec{a} \\ \vec{a} \cdot (\vec{b} + \vec{c}) &= \vec{a} \cdot \vec{b} + \vec{a} \cdot \vec{c} \\ \vec{a} \cdot (x\vec{b}) &= x(\vec{a} \cdot \vec{b}) \\ \vec{a} \cdot \vec{b} &= \vec{a} \vec{b}^T \end{aligned}$$

### 3.3.2. Outer Product

## 3.4. Operations

### 3.4.1. Projections



## Scalar Projection

It is the length (scalar) of the 'shadow' cast by one vector onto another.

It is basically how much  $\vec{a}$  goes in direction of  $\vec{b}$  as a scalar.

$$\frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\|} = \|\vec{b}\| \cos(\theta)$$

## Vector Projection

It is the a vector of the scalar projection , i.e. , how much  $\vec{a}$  goes in direction of  $\vec{b}$ , created by multiplying the scalar projection with a unit vector in direction of  $\vec{b}$ .

Vector in direction  $\vec{a}$  with length 1 :

$$\frac{\vec{a}}{\|\vec{a}\|}$$

Vector Projection :

$$\frac{\vec{a}}{\|\vec{a}\|} \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\|} = \vec{a} \frac{\vec{a} \cdot \vec{b}}{\vec{a} \cdot \vec{a}}$$

### 3.4.2. Rotations

### 3.4.3. Scale

### 3.4.4. Shear

### 3.4.5. Pivots

### 3.4.6. Transpose

### 3.4.7. Conjugate

### 3.4.8. Complex Conjugate

## 3.5. Properties

Zero Matrix Row Matrix Column Matrix Square Matrix Diagonal Matrix Symmetric Matrix Scalar Matrix Null Matrix Singular Matrix Invertible Matrix Transpose Permutation Matrix Orthogonal Upper Triangular Lower Triangular Positive Definite Matrix Positive Semi-Definite Matrix Bilinearity Cofactor Matrix Orthogonal Matrix Eigenvalue Matrix

### 3.5.1. Orthogonal

Orthogonal Matrices are a special kind of matrix that combine both inverse matrices and transpose matrices. Specifically they combine them in a way, such that for any matrix  $Q$  if  $Q^{-1} = Q^T$ , then  $Q$  is an Orthogonal Matrix. It is more commonly written as the following :

$$QQ^T = I = Q^T Q$$

Properties :

- Orthogonal Matrix preserves norms / lengths :  
 $(Q\vec{x})^T(Q\vec{x}) = \|Q\vec{x}\|^2 = \vec{x}^T Q^T Q \vec{x} = \vec{x}^T \vec{x} = \|\vec{x}\|^2$

### Orthonormal

#### 3.5.2. Positive

### 3.6. Spaces

Vector Spaces

Null Space

Column Space

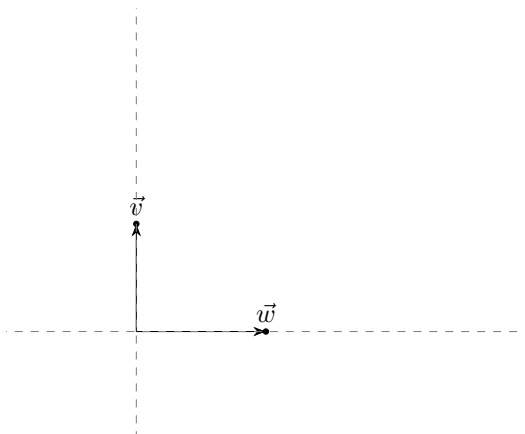
Row Space

Eigenspace

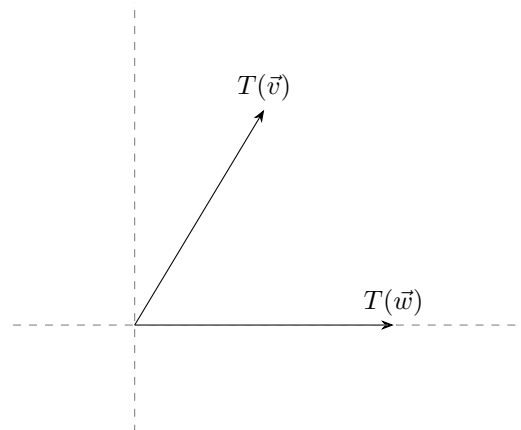
Fundamental Theorem of Linear Algebra

### 3.7. Eigenvectors

All vectors that stay on their span post a transformation.



transformation



After the above transformation we can see that  $\vec{v}$  is an eigenvector of the transformation  $T$ , because it stays on its span (dotted line). Usually we scale the eigenvectors down to unit size by dividing by the magnitude  $\frac{\vec{v}}{\|\vec{v}\|}$ .

So for any transformation  $T(\vec{x})$  in linear algebra, we can get a good picture of what  $T(\vec{x})$  has done to our original space by just looking at the eigenvalues.

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

Where  $T$  is some transformation matrix,  $\lambda$  is the eigenvalue and  $\vec{x}$  is the eigenvector. Since LHS is vector multiple, and RHS is a scalar we need to fix that by multiplying with the identity matrix :

$$\begin{aligned} T\vec{x} &= \lambda I \vec{x} \\ (T - \lambda I)\vec{x} &= \vec{0} \end{aligned}$$

So either  $T - \lambda I = 0$  or  $\vec{x} = 0$ . The 2nd case is trivial.

We know that for any matrix transformation to be 0, the transformation is reducing the determinant, i.e. area between the basis vectors, to be 0. This means  $\det(T - \lambda I) = 0$ . As an example for a 2D matrix :

$$T = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$\begin{aligned} \det(T - \lambda I) &= \det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right) \\ &= \det\left(\begin{bmatrix} a - \lambda & b \\ c & d - \lambda \end{bmatrix}\right) \end{aligned}$$

$$(a - \lambda)(d - \lambda) - bc = 0$$

$$\lambda^2 - (a + d)\lambda + (ad - bc) = 0$$

Eigenvalues are solutions to the above equation.

- Not all transformation will have eigenvectors
- Eigenbasis:** When we choose our basis vectors



such that they are eigenvectors, this is called an eigenbasis.

- All eigenvectors + eigenvalues is called the **Spectrum** of the space.
- The reason we care about eigenthings is because they tell us everything about a linear transformation by reducing it to it's bare essentials, i.e. the eigenvectors can communicate all relevant information as succinctly as possible.
- How many can we have? Eigenvectors with different eigenvalues are linearly independent. Since there cannot be more linearly independent vectors than the dimension of the space. Therefore there cannot be more eigenvectors than the dimension of the space.

### 3.7.1. Eigenbasis

## 3.8. Factorization

### 3.8.1. Gram-Schmidt

### 3.8.2. Singular Value Decomposition

Inverse Distribution through parentheses Transpose  
Distribution through parentheses Moving Matrices to  
and fro from equality Moving Inverses to and fro from  
equality Moving Transposes to and fro from equality

-§-

## 4. Calculus

### Derivative

How much the value of the function  $f(x)$  changes between a point  $x$ , and another point  $\Delta x$  away from  $x$  as the distance between the points approaches 0.

$$\frac{df}{dx} = \lim_{\Delta x \rightarrow 0} \left( \frac{f(x + \Delta x) - f(x)}{\Delta x} \right)$$

### Sum Rule

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

### Power Rule

$$F(x) = (f(x))^n$$

$$F'(x) = n(f(x))^{n-1} \cdot f'(x)$$

### Product Rule

$$F(x) = f(x) \cdot g(x)$$

$$F'(x) = f(x)g'(x) + f'(x)g(x)$$

### Chain Rule

$$F'(x) = f'(g(x)) \cdot g'(x)$$

### Quotient Rule

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

### Second Derivative Test

If  $f'(c) = 0$ , and  $f''$  is continuous over the interval containing  $c$  then :

$f''(c) < 0 \Rightarrow f$  has local maximum at  $x = c$

$f''(c) > 0 \Rightarrow f$  has local minimum at  $x = c$

$f''(c) = 0 \Rightarrow f$  test is inconclusive

**Common Derivatives**

$$f(x) = \frac{1}{x} \quad , \quad f'(x) = -\frac{1}{x^2}$$

$$f(x) = e^x \quad , \quad f'(x) = e^x$$

$$f(x) = \ln(x) \quad , \quad f'(x) = \frac{1}{x}$$

$$f(x) = 2^x \quad , \quad f'(x) = 2^x \ln(x)$$

$$f(x) = \sqrt{x} \quad , \quad f'(x) = \frac{1}{2 \cdot \sqrt{x}}$$

$$f(x) = \sin(x) \quad , \quad f'(x) = \cos(x)$$

$$f(x) = \cos(x) \quad , \quad f'(x) = -\sin(x)$$

$$f(x) = \sin^{-1}(x) \quad , \quad f'(x) = \frac{1}{\sqrt{1-x^2}}$$

$$f(x) = \cos^{-1}(x) \quad , \quad f'(x) = -\frac{1}{\sqrt{1-x^2}}$$

$$f(x) = \tan^{-1}(x) \quad , \quad f'(x) = \frac{1}{1+x^2}$$

**5. Multivariable Calculus**

Assume the other variables are just constants , and differentiate that way.

**Notation**

$$\frac{\delta}{\delta x} f = f_x$$

**Multivariate Chain Rule**

$$\frac{d}{dx} f(x, y, z) = \frac{\delta f}{\delta x} \frac{dx}{dt} + \frac{\delta f}{\delta y} \frac{dy}{dt} + \frac{\delta f}{\delta z} \frac{dz}{dt}$$

**Gradient**

The derivative of a vector sent to a scalar , i.e. ,  $f : \mathbb{R}^n \rightarrow \mathbb{R}$

The direction of steepest descent in a specific direction  $\hat{v} : \nabla_x f(x) \bullet \hat{v}$

$$\nabla_x f(x_1, x_2, \dots, x_n) = \begin{bmatrix} \frac{\delta f}{\delta x_1} \\ \frac{\delta f}{\delta x_2} \\ \vdots \\ \frac{\delta f}{\delta x_n} \end{bmatrix}$$

**Directional Derivative****Jacobian**

A vector pointing in the direction of steepest uphill slope

The derivative of a vector sent to another vector , i.e. ,  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$

$$J = \begin{bmatrix} \frac{\delta f}{\delta x_1} & , & \frac{\delta f}{\delta x_2} & , & \frac{\delta f}{\delta x_3} & , & \dots \end{bmatrix}$$

**Hessian Matrix****Multivariate 2nd derivative test****Multivariate Taylor Series****Lagrangian Method****Maclaurin Series****Taylor Series**

-§-

-§-

## 6. Sequences , Series

### 6.1. Sequences

170 Geometric Sequences Arithmetic Sequences

### 6.2. Series

171 Infinite Geometric Series

$$\sum_{i=0}^{\infty} \alpha^i = 1 + \alpha + \alpha^2 + \dots = \frac{1}{1 - \alpha} \quad , \quad |\alpha| < 1$$

-§-