0. Contents

0. List of Tables

0. List of Figures

1. Algebra

Logarithms

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

Exponentials

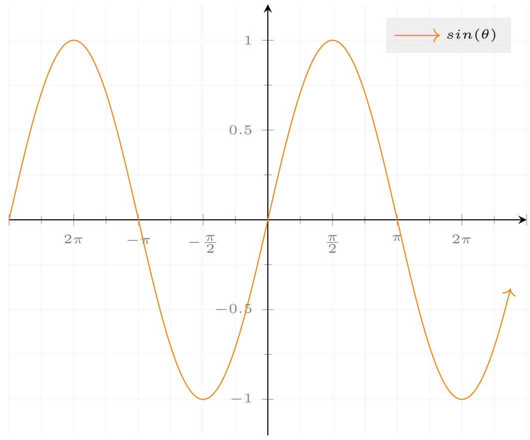
- $a^m \cdot a^n = a^{m+n}$
- $a^0 = 1 \quad (a \neq 0)$
- $a^{-n} = \frac{1}{n} \quad (a \neq 0)$
- $\frac{a^m}{a^n} = a^{m-n} \quad (a \neq 0)$
- $(a^m)^n = a^{mn}$
- $(ab)^m = a^m b^m$
- $\left(\frac{a}{b}\right)^m = \frac{a^m}{b^m} \quad (b \neq 0)$

Trigonometery

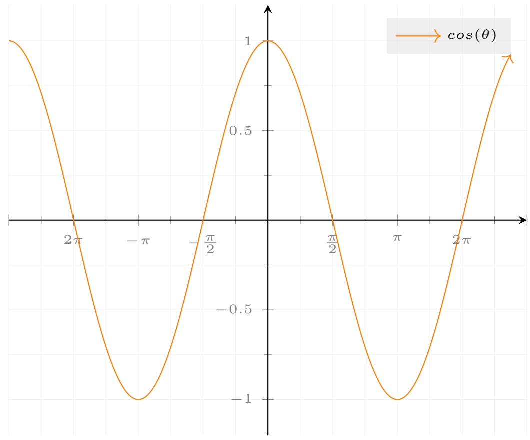
deg	0°	30°	45°	60°	90°	180°	270°	360°
rad	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$	π	$\frac{3\pi}{2}$	2π
$\sin(\theta)$	0	$\frac{1}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{\sqrt{3}}{2}$	1	0	-1	0
$\cos(\theta)$	1	$\frac{\sqrt{3}}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{1}{2}$	0	-1	0	1
$\tan(\theta)$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

-§-

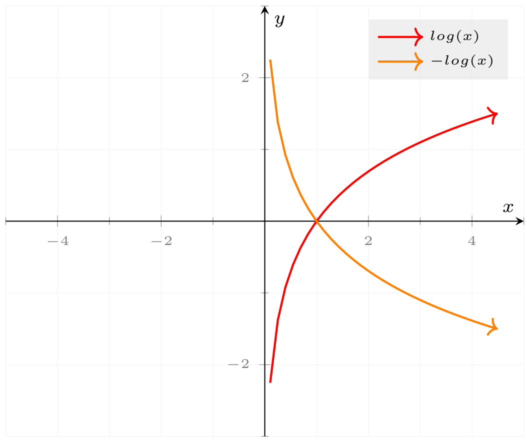
2. Graphs



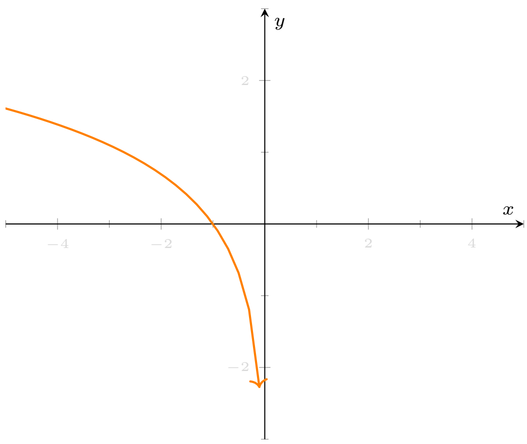
20



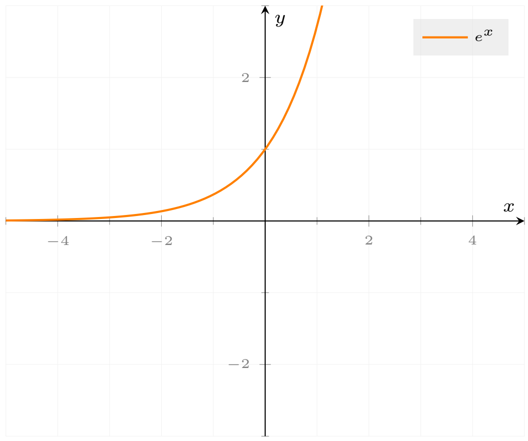
21



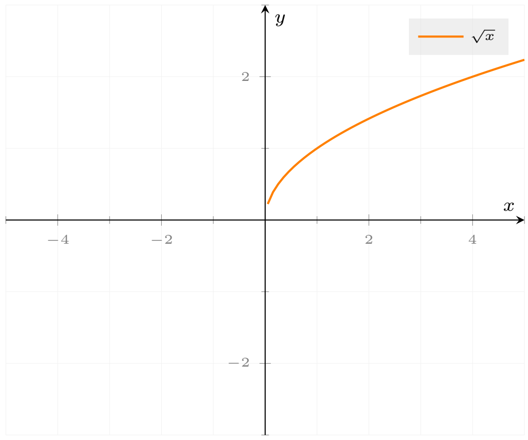
22



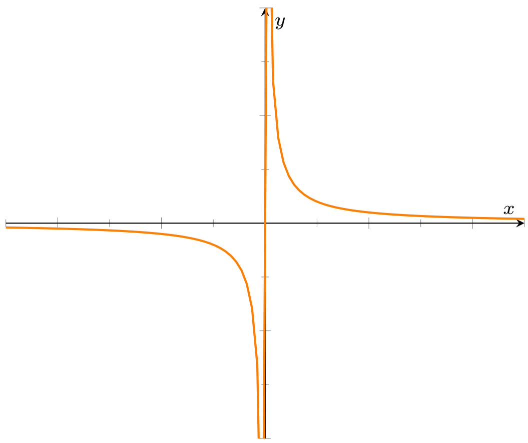
23



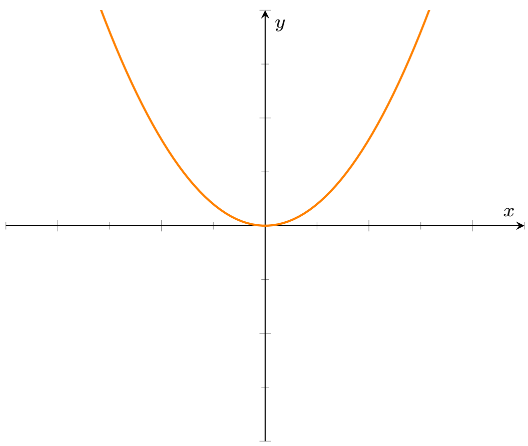
24



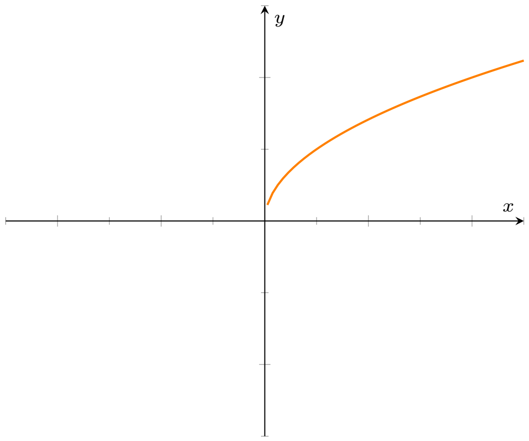
25



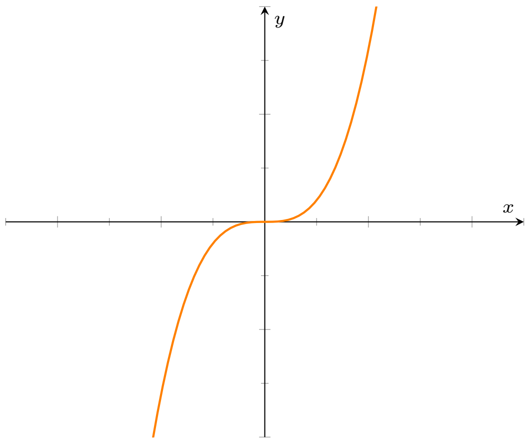
26



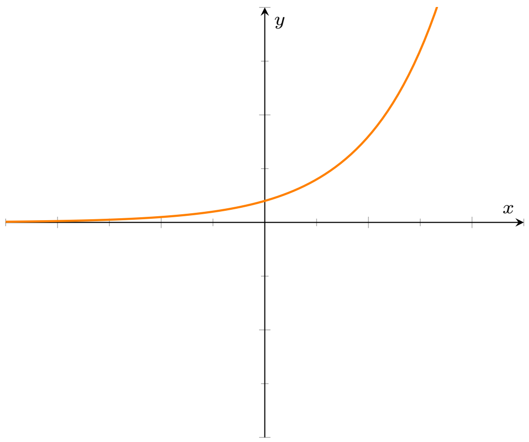
27



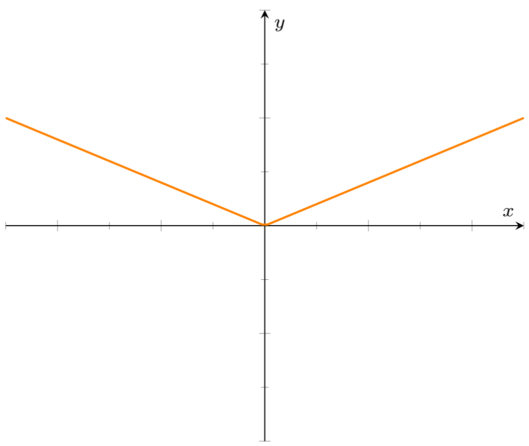
28



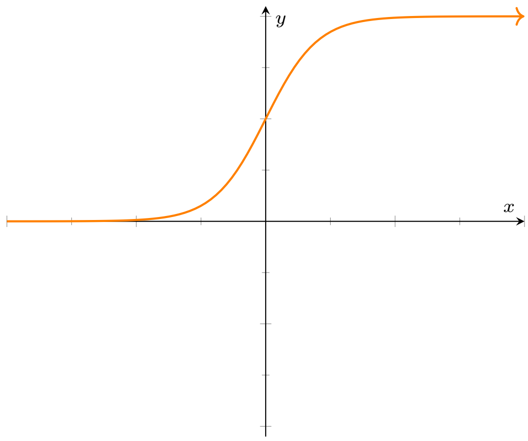
29



30



31



32

-§-

3. Linear Algebra

3.1. Basics

$$\overrightarrow{a} + \overrightarrow{b} = \overrightarrow{b} + \overrightarrow{a}$$
$$2\overrightarrow{a} = \overrightarrow{a} + \overrightarrow{a}$$
$$\|\overrightarrow{a}\|^2 = \sum_i a_i^2$$
$$A_{ij}^T = A_{ji}$$

3.1.1. Matrix Multiplication

Right way to think about it

Why is it called Linear Algebra ?

Linear Transformations

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 happends that most often , this basis is on the x , y plane with $\hat{i}$, and $\hat{j}$ being the basis vectors.

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 $\overrightarrow{a}$ and $\overrightarrow{b}$ is the set of all of their linear combinations , i.e. , $x\overrightarrow{a} + y\overrightarrow{b}$, 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.

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**.

The scalar projection of your basis vector onto another is how you find out what the scalar value of that vector is , in the span defined by your chosen basis vector.

Basically , $\overrightarrow{v}$ in basis $\overrightarrow{b_1}$, $\overrightarrow{b_2}$, where $\overrightarrow{b_1}$, $\overrightarrow{b_2}$ are orthogonal to each other , is given by

$$\begin{bmatrix} \frac{\overrightarrow{v} \bullet \overrightarrow{b_1}}{\left\| \overrightarrow{b_1} \right\|} \\ \frac{\overrightarrow{v} \bullet \overrightarrow{b_2}}{\left\| \overrightarrow{b_2} \right\|} \end{bmatrix}$$

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

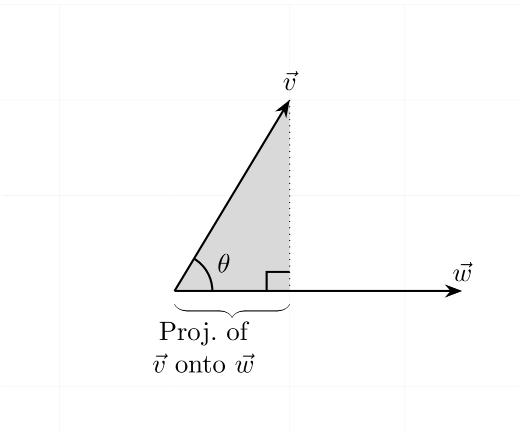
3.3.1. Inner Product

$$\begin{aligned}
\overrightarrow{a} \bullet \overrightarrow{b} &= a_1 b_1 + a_2 b_2 + \ldots + a_n b_n \\
&= \sum_i a_i b_i \\
\overrightarrow{a} \bullet \overrightarrow{a} &= \left\| \overrightarrow{a} \right\|^2 \\
\overrightarrow{a} \bullet \overrightarrow{b} &= \overrightarrow{b} \bullet \overrightarrow{a} \\
\overrightarrow{a} \bullet \left(\overrightarrow{b} + \overrightarrow{c} \right) &= \overrightarrow{a} \bullet \overrightarrow{b} + \overrightarrow{a} \bullet \overrightarrow{c} \\
\overrightarrow{a} \bullet \left(x \overrightarrow{b} \right) &= x \left(\overrightarrow{a} \bullet \overrightarrow{b} \right) \\
\overrightarrow{a} \bullet \overrightarrow{b} &= \overrightarrow{a} \, \overrightarrow{b}^T
\end{aligned}$$

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 $\overrightarrow{a}$ goes in direction of $\overrightarrow{b}$ as a scalar.

$$\frac{\overrightarrow{a} \bullet \overrightarrow{b}}{\left\| \overrightarrow{a} \right\|} = \left\| \overrightarrow{b} \right\| \, cos(\theta)$$

Vector Projection

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

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

$$\frac{\overrightarrow{a}}{\left\| \overrightarrow{a} \right\|}$$

Vector Projection :

$$\frac{\overrightarrow{a}}{\left\| \overrightarrow{a} \right\|} \frac{\overrightarrow{a} \bullet \overrightarrow{b}}{\left\| \overrightarrow{a} \right\|} = \overrightarrow{a} \frac{\overrightarrow{a} \bullet \overrightarrow{b}}{\overrightarrow{a} \bullet \overrightarrow{a}}$$

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 Symmetrix 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 Ciculant Matrix Orthogonal Matrix Eigenvalue Matrix

3.5.1. Orthogonoal

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 :

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

Properties :

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

Orthonormal

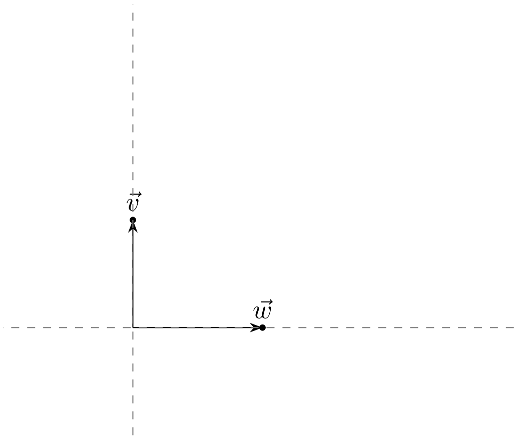
3.5.2. Positive

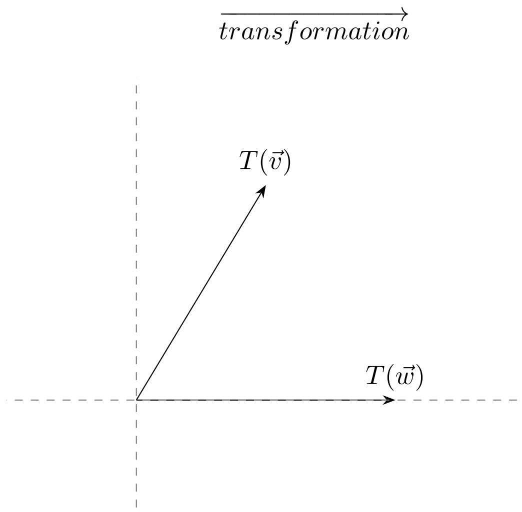
3.6. Spaces

Vector Spaces
Null Space
Column Space
Row Space
Eigenspace
Fundamental Theorem of Linear Algebra

3.7. Eigenvectors

All vectors that stay on thier span post a 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 , λ 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 :

$$T\vec{x} = \lambda I\vec{x}$$
$$(T - \lambda I)\vec{x} = \vec{0}$$

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 eigenthingies 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 thant 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 Δx away from x as the distance between the points approaches 0.

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

Sum Rule

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

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

$$f(x) = \frac{1}{x} \quad , \quad f'(x) = \frac{1}{-x^2}$$
$$f(x) = e^x \quad , \quad f'(x) = e^x$$
$$f(x) = \ln(x) \quad , \quad f'(x) = \frac{1}{x}$$
$$f(x) = 2^x \quad , \quad f'(x) = 2^x \ln(x)$$
$$f(x) = \sqrt{x} \quad , \quad f'(x) = \frac{1}{2 \cdot \sqrt{x}}$$
$$f(x) = sin(x) \quad , \quad f'(x) = cos(x)$$
$$f(x) = cos(x) \quad , \quad f'(x) = -sin(x)$$
$$f(x) = sin^{-1}(x) \quad , \quad f'(x) = \frac{1}{\sqrt{(1-x^2)}}$$
$$f(x) = cos^{-1}(x) \quad , \quad f'(x) = -\frac{1}{\sqrt{(1-x^2)}}$$
$$f(x) = tan^{-1}(x) \quad , \quad f'(x) = \frac{1}{1+x^2}$$

Maclaurin Series

Taylor Series

-§-

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. , $f : \mathbb{R}^n \to \mathbb{R}$

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

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

Directional Derivative

Jacobian
A vector pointing in the direction of steepest uphill slope

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

$$f : \mathbb{R}^n \to \mathbb{R}^m$$
$$f(x_1, x_2, x_3, \ldots)$$
$$J = \begin{bmatrix} \frac{\delta f}{\delta x_1} & , & \frac{\delta f}{\delta x_2} & , & \frac{\delta f}{\delta x_3} & , & \ldots \end{bmatrix}$$

Hessian Matrix

Multivariate 2nd derivative test

Multivariate Taylor Series

Lagrangian Method

-§-

6. Sequences , Series

6.1. Sequences

Geometric Sequences Arithmetic Sequences

6.2. Series

Infinite Geometric Series

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

-§-