Real analysis, variational calculus and functionals

Adam Boult (www.bou.lt)

April 30, 2025

Contents

Preface

This is a live document, and is full of gaps, mistakes, typos etc.

Part I

Pre-calculus

Chapter 1

Ordering of infinite sets

Chapter 2

Introduction

2.0.1 Ordered sets

Totally ordered sets

A totally ordered set is one where the relation is defined on all pairs:

$\forall a \forall b (a \leq b) \vee (b \leq a)$

Note that totality implies reflexivity.

Partially ordered sets (poset)

A partially ordered set, or poset, is one where the relation is defined between each element and itself.

$\forall a (a \leq a)$

That is, every element is related to itself.

These are also called posets.

Well-ordering

A well-ordering on a set is a total order on the set where the set contains a minimum number. For example the relation $\leq$ on the natural numbers is a well-ordering because 0 is the minimum.

The relation $\leq$ on the integers however is not a well-ordering, as there is no minimum number in the set.

2.0.2 Intervals

For a totally ordered set we can define a subset as being all elements with a relationship to a number. For example:

$[a, b] = \{x : a \leq x \wedge x \leq b\}$

This denotes a closed interval. Using the definition above we can also define an open interval:

$(a, b) = \{x : a < x \wedge x < b\}$

2.0.3 Infinitum and supremum

Infinitum

Consider a subset S of a partially ordered set T.

The infinitum of S is the greatest element in T that is less than or equal to all elements in S.

For example:

$\inf[0, 1] = 0$

$\inf(0, 1) = 0$

Supremum

The supremum is the opposite: the smallest element in T which is greater than or equal to all elements in S.

$\sup[0, 1] = 1$

$\sup(0, 1) = 1$

Max and min

If the infinitum of a set S is in S, then the infinimum is the minimum of set S. Otherwise, the minimum is not defined.

$\min[0, 1] = 0$

$\min(0, 1)$ isn't defined.

Similarly:

$\max[0, 1] = 1$

$\max(0, 1)$ isn't defined.

Chapter 3

Limits of infinite sequences

3.1 Introduction

Chapter 4

Properties of functions

4.1 Real functions

4.1.1 Real functions

Consider a function

$y = f(x)$

$f(x)$ is a real function if:

$\forall x \in \mathbb{R} f(x) \in \mathbb{R}$

4.1.2 Support

$f X \to R$

Support of f is $x \in X$ where $f(x) \neq 0$

4.1.3 Monotonic functions

Calculus stationary points finding and monotonic functions

4.1.4 Even and odd functions

Defining odd and even functions

An even function is one where:

$f(x) = f(-x)$

An odd function is one where:

$f(x) = -f(-x)$

Functions which are even and odd

If a function is even and odd:

$f(x) = f(-x) = -f(-x)$

$f(x) = -f(x)$

Then $f(x) = 0$.

Scaling odd and even functions

Scaling an even function provides an even function.

$h(x) = c.f(x)$

$h(-x) = c.f(-x)$

$h(-x) = c.f(x)$

$h(-x) = h(x)$

Scaling an odd function provides an odd function.

$h(x) = c.f(x)$

$-h(-x) = -c.f(-x)$

$-h(-x) = c.f(x)$

$-h(-x) = h(x)$

Adding odd and even functions

Note than 2 even functions added together makes an even function.

$h(x) = f(x) + g(x)$

$h(x) = f(-x) + g(-x)$

$h(-x) = f(x) + g(x)$

$h(x) = h(-x)$

And adding 2 odd functions together makes an odd function.

$h(x) = f(x) + g(x)$

$h(x) = -f(-x) - g(-x)$

$-h(-x) = f(x) + g(x)$

$-h(-x) = h(x)$

Multiplying odd and even functions

Multiplying 2 even functions together makes an even function.

$h(x) = f(x)g(x)$

$h(-x) = f(-x)g(-x)$

$h(-x) = f(x)g(x)$

$h(-x) = h(x)$

Multiplying 2 odd functions together makes an even function.

$h(x) = f(x)g(x)$

$h(-x) = f(-x)g(-x)$

$h(-x) = (-1).(-1.)f(x)g(x)$

$h(-x) = h(x)$

4.1.5 Concave and convex functions

Convex functions

A convex function is one where:

$\forall x_1, x_2 \in \mathbb{R} \forall t \in [0, 1][f(tx_1 + (1-t)x_2 \leq tf(x_1) + (1-t)f(x_2)]$

That is, for any two points of a function, a line between the two points is above the curve.

A function is strictly convex if the line between two points is strictly above the curve:

$\forall x_1, x_2 \in \mathbb{R} \forall t \in (0, 1)[f(tx_1 + (1-t)x_2 < tf(x_1) + (1-t)f(x_2)]$

An example is $y = x^2$.

Concave functions

A concave function is an upside down convex function. The line between two points is below the curve.

$\forall x_1, x_2 \in \mathbb{R} \forall t \in [0, 1][f(tx_1 + (1-t)x_2 \geq tf(x_1) + (1-t)f(x_2)]$

A function is strictly concave if the line between two points is strictly below the curve:

$\forall x_1, x_2 \in \mathbb{R} \forall t \in (0, 1)[f(tx_1 + (1-t)x_2 > tf(x_1) + (1-t)f(x_2)]$

An example is $y = -x^2$.

Affine functions

If a function is both concave and convex, then the line between two points must be the function itself. This means the function is an affine function.

$y = cx$

4.1.6 Subadditive and superadditive functions

4.2 O

4.2.1 Big O and little-o notation

Big O notation

In big O notation we are interested in t he size of a function as it getes larger. We ignore constant multiples.

$cx \in O(x)$

And addition of constants.

$cx + b \in O(x)$

If there are two terms and one is larger, we keep the largest.

$x + x^2 \in O(x^2)$

More generally we write:

$f(x) \in O(g(x))$

Little-o notation

Chapter 5

Limits and continuous functions

5.1 Limits

5.1.1 Limits of real functions

Limit operator

For a function $f(x)$,

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

We can say that L is the limit if:

$\forall \epsilon > 0 \exists \delta > 0 \forall x [0 < |x - p| < \delta \to |f(x) - L| < \epsilon]$

5.1.2 Limit superior and limit inferior

If a sequence does not converge, but stays between two points, then lim sup is upper bound, lim inf is lower bound.

5.2 Continuous functions

5.2.1 Continous functions

A function is continuous if:

$\lim_{x \to c} f(x) = f(c)$

For example a function $\frac{1}{x}$ is not continuous as the limit towards 0 is negative infinity. A function like $y = x$ is continous.

More strictly, for any $\epsilon > 0$ there exists

$\delta > 0$

$c - \delta < x < c + \delta$

Such that

$f(c) - \epsilon < f(x) < f(c) + \epsilon$

This means that our function is continuous at our limit c, if for any tiny range around $f(c)$, that is $f(c) - \epsilon$ and $f(c) + \epsilon$, there is a range around c, that is $c - \delta$ and $c + \delta$ such that all the value of $f(x)$ at all of these points is within the other range.

Limits

Why can't we use rationals for analysis?

If discontinous at not rational number, it can still be continous for all rationals.

Eg $f(x) = -1$ unless $x^2 > 2$, where $f(x) = 1$.

Continous for all rationals, because rationals dense in reals.

But can't be differentiated.

5.2.2 Reals or rationals for analysis

Why can't we use rationals for analysis?

If discontinous at not rational number, it can still be continous for all rationals.

eg $f(x) = -1$ unless $x^2 > 2$, where $f(x) = 1$.

Continous for all rationals, because rationals dense in reals

But can't be differentiated

5.2.3 Boundedness theorem

If $f(x)$ is closed and continuous in $[a, b]$ then $f(x)$ is bounded by m and M. That is:

$\exists m \in \mathbb{R} \exists M \in \mathbb{R} \forall x \in [a, b](m < f(x) < M$

5.2.4 Intermediate value theorem

Take a real function $f(x)$ on closed interval $[a, b]$, continuous on $[a, b,]$.

IVT says that for all numbers u between $f(a)$ and $f(b)$, there is a corresponding value c in $[a, b]$ such that $f(c) = u$.

That is:

$\forall u \in [min(f(a), f(b)), max(f(a), f(b))]\exists c \in [a, b](f(c) = u)$

5.2.5 Extreme value theorem

We can expand the boundedness theorem such that m and M are functions of $f(x)$ in the bound $[a, b]$. That is:

$\exists m \in \mathbb{R}\exists M \in \mathbb{R}\forall x \in [a, b](m < f(x) < M)$

Chapter 6

Transcendental and real numbers

6.1 Constructing the real numbers

6.1.1 Cauchy sequences

Cauchy sequence

A cauchy sequence is a sequence such that for an any arbitrarily small number ϵ, there is a point in the sequence where all possible pairs after this are even closer together.

$(\forall \epsilon > 0)(\exists N \in \mathbb{N} : \forall m, n \in \mathbb{N} > N)(|a_m - a_n| < \epsilon)$

This last term gives a distance between two entries. In addition to the number line, this could be used on vectors, where distances are defined.

As a example, $\frac{1}{n}$ is a cauchy sequence, $\sum_i \frac{1}{n}$ is not.

Completeness

Cauchy sequences can be defined on some given set. For example given all the numbers between 0 and 1 there are any number of different cauchy sequences converging at some point.

If it is possible to define a cauchy sequence on a set where the limit is not in the set, then the set is incomplete.

For example, the numbers between 0 and 1 but not including 0 and 1 are not complete. It is possible to define sequences which converge to these missing points.

More abstractly, you could have all vectors where $x^2+y^2 < 1$. This is incomplete (or open) as sequences on these vectors can converge to limits not in the set.

Cauchy sequences are important when considering real numbers. We could define a sequence converging on $\sqrt{2}$, but as this number is not in the set, it is incomplete.

6.1.2 Incompleteness of the rational numbers

The square root of 2 is not a rational number

Let's prove there are numbers which are not rational. Consider $\sqrt{2}$ and let's show that it being rational leads to a contradiction.

$\sqrt{2} = \frac{x}{y}$

$2 = \frac{x^2}{y^2}$

$2y^2 = x^2$

So we know that x^2 is even, and can be shown as $x = 2n$.

$2y^2 = (2n)^2$

$y^2 = 2n^2$

So y is even. But if both x and y are even, then the fraction was not reduced.

This presents a contraction so the original statement must have been false.

So we know there isn't a rational solution to $\sqrt{2}$.

6.1.3 Density of rationals and reals

Rationals are dense in reals

Reals are dense in reals

Reals are dense in rationals

6.1.4 σ-algebra

Review of algebra on a set

An algebra, Σ, on set s is a set of subsets of s such that:

- Closed under intersection: If a and b are in Σ then $a \wedge b$ must also be in Σ
- $\forall ab[(a \in \Sigma \wedge b \in \Sigma) \rightarrow (a \wedge b \in \Sigma)]$
- Closed under union: If a and b are in Σ then $a \vee b$ must also be in Σ.
- $\forall ab[(a \in \Sigma \wedge b \in \Sigma) \rightarrow (a \vee b \in \Sigma)]$

If both of these are true, then the following is also true:

- Closed under complement: If a is in Σ then $s\backslash a$ must also be in Σ

We also require that the null set (and therefore the original set, null's complement) is part of the algebra.

σ-algebra

A σ-algebra is an algebra with an additional condition:

All countable unions of sets in A are also in A.

This adds a constraint. Consider the real numbers with an algebra of all finite sets.

This contains all finite subsets, and their complements. It does not contain $\mathbb{N}$.

However a σ-algebra requires all countable unions to be including, and the natural numbers are a countable union.

The power set is a σ-algebra. All other σ-algebras are subsets of the power set.

Part II

Univariate real differentiation

Chapter 7

Univariate differentiation

7.1 Partial differentiation

7.1.1 The partial differential operator

Differential

When we change the value of an input to a function, we also change the output. We can examine these changes.

Consider the value of a function $f(x)$ at points x_1 and x_2.

$y_1 = f(x_1)$

$y_2 = f(x_2)$

$y_2 - y_1 = f(x_2) - f(x_1)$

$\frac{y_2 - y_1}{x_2 - x_1} = \frac{f(x_2) - f(x_1)}{x_2 - x_1}$

Let's define x_2 in terms of its distance from x_1:

$x_2 = x_1 + \epsilon$

$\frac{y_2 - y_1}{\epsilon} = \frac{f(x_1 + \epsilon) - f(x_1)}{\epsilon}$

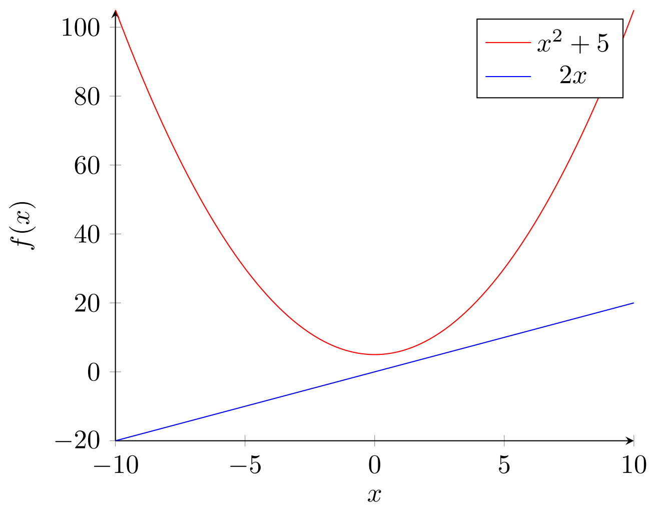
We define the differential of a function as:

$\frac{\delta y}{\delta x} = \lim_{\epsilon \to 0^+} \frac{f(x + \epsilon) - f(x)}{\epsilon}$

If this is defined, then we say the function is differentiable at that point.

Differential operator

Graph test



7.1.2 Differentiating constants, the identity function, and linear functions

Constants

$f(x) = c$

$\frac{\delta y}{\delta x} = \lim_{\epsilon \to 0^+} \frac{f(x+\epsilon) - f(x)}{\epsilon}$

$\frac{\delta y}{\delta x} = \lim_{\epsilon \to 0^+} \frac{c - c}{\epsilon} = 0$

x

$f(x) = x$

$\frac{\delta y}{\delta x} = \lim_{\epsilon \to 0^+} \frac{f(x+\epsilon) - f(x)}{\epsilon}$

$\frac{\delta y}{\delta x} = \lim_{\epsilon \to 0^+} \frac{x + \epsilon - x}{\epsilon} = 1$

Addition

$f(x) = g(x) + h(g)$

$\frac{\delta y}{\delta x} = \lim_{\epsilon \to 0^+} \frac{g(x+\epsilon) + h(x+\epsilon) - g(x) - h(x)}{\epsilon}$

$\frac{\delta y}{\delta x} = \lim_{\epsilon \to 0^+} \frac{g(x+\epsilon)-g(x)}{\epsilon} + \lim_{\epsilon \to 0^+} \frac{h(x+\epsilon)-h(x)}{\epsilon}$

$\frac{\delta y}{\delta x} = \frac{\delta g}{\delta x} + \frac{\delta h}{\delta x}$

7.1.3 Partial differentiation is a linear operator

Intro

7.1.4 The chain rule, the product rule and the quotient rule

Chain rule

$f(x) = f(g(x))$

$\frac{\delta f}{\delta x} = \lim_{\epsilon \to 0^+} \frac{f(g(x+\epsilon)) - f(g(x))}{\epsilon}$

$\frac{\delta f}{\delta x} = \lim_{\epsilon \to 0^+} \frac{g(x+\epsilon)-g(x)}{g(x+\epsilon)-g(x)} \frac{f(g(x+\epsilon)) - f(g(x))}{\epsilon}$

$\frac{\delta f}{\delta x} = \lim_{\epsilon \to 0^+} \frac{g(x+\epsilon)-g(x)}{\epsilon} \frac{f(g(x+\epsilon)) - f(g(x))}{g(x+\epsilon)-g(x)}$

$\frac{\delta f}{\delta x} = \lim_{\epsilon \to 0^+} [\frac{g(x+\epsilon)-g(x)}{\epsilon}] \lim_{\epsilon \to 0^+} [\frac{f(g(x+\epsilon)) - f(g(x))}{g(x+\epsilon)-g(x)}]$

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

Product rule

$y = f(x)g(x)$

$\frac{\delta y}{\delta x} = \lim_{\epsilon \to 0^+} \frac{f(x+\epsilon)g(x+\epsilon) - f(x)g(x)}{\epsilon}$

$\frac{\delta y}{\delta x} = \lim_{\epsilon \to 0^+} \frac{f(x+\epsilon)g(x+\epsilon) - f(x)g(x+\epsilon) + f(x)g(x+\epsilon) - f(x)g(x)}{\epsilon}$

$\frac{\delta y}{\delta x} = \lim_{\epsilon \to 0^+} \frac{f(x+\epsilon)g(x+\epsilon) - f(x)g(x+\epsilon)}{\epsilon} + \lim_{\epsilon \to 0^+} \frac{f(x)g(x+\epsilon) - f(x)g(x)}{\epsilon}$

$\frac{\delta y}{\delta x} = \lim_{\epsilon \to 0^+} g(x+\epsilon) \frac{f(x+\epsilon) - f(x)}{\epsilon} + \lim_{\epsilon \to 0^+} f(x) \frac{g(x+\epsilon) - g(x)}{\epsilon}$

$\frac{\delta y}{\delta x} = g(x) \frac{\delta f}{\delta x} + f(x) \frac{\delta g}{\delta x}$

Quotient rule

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

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

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

$$\frac{\delta}{\delta x}y=\frac{\delta f}{\delta x}\frac{1}{g(x)}-\frac{\delta g}{\delta x}\frac{f(x)}{g(x)^2}$$

$$\frac{\delta}{\delta x}y=\frac{\frac{\delta f}{\delta x}g(x)-\frac{\delta g}{\delta x}f(x)}{g(x)^2}$$

7.1.5 Differentiating natural number power functions

Other

$$\frac{\delta}{\delta x}x^n=\lim_{\delta\to 0}\frac{(x+\delta)^n-x^n}{\delta}$$

$$\frac{\delta}{\delta x}x^n=\lim_{\delta\to 0}\frac{(\sum_{i=0}^n x^i\delta^{n-i}\frac{n!}{i!(n-i)!})-x^n}{\delta}$$

$$\frac{\delta}{\delta x}x^n=\lim_{\delta\to 0}\sum_{i=0}^{n-1}x^i\delta^{n-i-1}\frac{n!}{i!(n-i)!}$$

$$\frac{\delta}{\delta x}x^n=\lim_{\delta\to 0}x^{n-1}\frac{n!}{(n-1)!(n-n+1)!}+\sum_{i=0}^{n-2}x^i\delta^{n-i-1}\frac{n!}{i!(n-i)!}$$

$$\frac{\delta}{\delta x}x^n=nx^{n-1}$$

7.1.6 L'Hôpital's rule

L'Hôpital's rule

If there are two functions which are both tend to 0 at a limit, calculating the limit of their divisor is hard. We can use L'Hopital's rule.

We want to calculate:

$$\lim_{x\to c}\frac{f(x)}{g(x)}$$

This is:

$$\lim_{x\to c}\frac{f(x)}{g(x)}=\lim_{x\to c}\frac{\frac{f(x)-0}{\delta}}{\frac{g(x)-0}{\delta}}$$

If:

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

Then

$\lim_{x\to c}\dfrac{f(x)}{g(x)}=\lim_{x\to c}\dfrac{\dfrac{f(x)-f(c)}{\delta}}{\dfrac{g(x)-f(c)}{\delta}}$

$\lim_{x\to c}\dfrac{f(x)}{g(x)}=\dfrac{f'(x)}{g'(x)}$

7.1.7 Rolle's theorem

Rolle's theorem

Take a real function $f(x)$ on closed interval $[a,b]$, differentiable on $(a,b,)$, and $f(a)=f(b)$.

Rolle's theorem states that:

$\exists c\in(a,b)(f'(c)=0)$

Generalised Rolle's theorem states that:

Generalised Rolle's theorem implies Rolle's theorem, so we only need to prove the generalised theorem.

7.1.8 Mean value theorem

Mean value theorem

Take a real function $f(x)$ on closed interval $[a,b]$, differentiable on $(a,b,)$.

The mean value theorem states that:

$\exists c\in(a,b)(f'(c)=\dfrac{f(b)-f(a)}{b-a})$

7.1.9 Elasticity

Introduction

We have $f(x)$

$Ef(x)=\dfrac{x}{f(x)}\dfrac{\delta f(x)}{\delta x}$

This is the same as:

$Ef(x)=\dfrac{\delta\ln f(x)}{\delta\ln x}$

7.1.10 Smooth functions

7.1.11 Analytic function

Introduction

7.2 Higher-order differentials

7.2.1 Differentiable functions

Introduction

A differentiable function is one where the differential is defined at all points on the real line.

All differentiable functions are continuous. Not all continuous functions are differentiable.

Differentiability class

We can describe a function with its differentiability class. If a function can be differentiated n times and these differentials are all continous, then the function is class C^n.

Smooth functions

If a function can be differentiated infinitely many times to produce continous functions, it is C^∞, or smooth.

7.2.2 Critial points

Critical points

Where partial derivative are 0.

Chapter 8

Identifying and evaluating e

8.1 Exponentials

8.1.1 Defining e as a binomial

Lemma

$f(n,i) = \frac{n!}{n^i(n-i)!}$

$f(n,i) = \frac{(n-i)!\prod_{j=n-i+1}^{n} j}{n^i(n-i)!}$

$f(n,i) = \frac{\prod_{j=n-i+1}^{n} j}{n^i}$

$f(n,i) = \frac{\prod_{j=1}^{i}(j+n-i)}{n^i}$

$f(n,i) = \prod_{j=1}^{i} \frac{j+n-i}{n}$

$f(n,i) = \prod_{j=1}^{i}(\frac{n}{n} + \frac{j-i}{n})$

$f(n,i) = \prod_{j=1}^{i}(1 + \frac{j-i}{n})$

$\lim_{n\to\infty} f(n,i) = \lim_{n\to\infty} \prod_{j=1}^{i}(1 + \frac{j-i}{n})$

$\lim_{n\to\infty} f(n,i) = \prod_{j=1}^{i} 1$

$\lim_{n\to\infty} f(n,i) = 1$

Defining e

We know that:

$(a+b)^n = \sum_{i=0}^{n} a^i b^{n-i} \dfrac{n!}{i!(n-i)!}$

Let's set $b = 1$

$(a+1)^n = \sum_{i=0}^{n} a^i \dfrac{n!}{i!(n-i)!}$

Let's set $a = \dfrac{1}{n}$

$(1+\dfrac{1}{n})^n = \sum_{i=0}^{n} \dfrac{1}{n^i} \dfrac{n!}{i!(n-i)!}$

$(1+\dfrac{1}{n})^n = \sum_{i=0}^{n} \dfrac{1}{i!} \dfrac{n!}{n^i(n-i)!}$

$\lim_{n\to\infty}(1+\dfrac{1}{n})^n = \lim_{n\to\infty} \sum_{i=0}^{n} \dfrac{1}{i!} \dfrac{n!}{n^i(n-i)!}$

From the lemma above:

$\lim_{n\to\infty}(1+\dfrac{1}{n})^n = \sum_{i=0}^{\infty} \dfrac{1}{i!}$

$e = \sum_{i=0}^{\infty} \dfrac{1}{i!}$

Defining e^x

$e = \lim_{n\to\infty}(1+\dfrac{1}{n})^n$

$e^x = \lim_{n\to\infty}(1+\dfrac{1}{n})^{nx}$

$e^x = \lim_{n\to\infty} \sum_{i=0}^{nx} \dfrac{1}{n^i} \dfrac{(nx)!}{i!(nx-i)!}$

$e^x = \lim_{n\to\infty} \sum_{i=0}^{nx} \dfrac{x^i}{i!} \dfrac{(nx)!}{(nx)^i(nx-i)!}$

From the lemma:

$e^x = \sum_{i=0}^{\infty} \dfrac{x^i}{i!}$

8.1.2 Differentiating e^x

Intro

We have $e^x = \sum_{i=0}^{\infty} \dfrac{x^i}{i!}$

$$\frac{\delta}{\delta x}e^x = \frac{\delta}{\delta x}\sum_{i=0}^{\infty}\frac{x^i}{i!}$$
$$\frac{\delta}{\delta x}e^x = \sum_{i=0}^{\infty}\frac{\delta}{\delta x}\frac{x^i}{i!}$$
$$\frac{\delta}{\delta x}e^x = \sum_{i=1}^{\infty}\frac{\delta}{\delta x}\frac{x^i}{i!}$$
$$\frac{\delta}{\delta x}e^x = \sum_{i=1}^{\infty}\frac{x^{i-1}}{(i-1)!}$$
$$\frac{\delta}{\delta x}e^x = \sum_{i=0}^{\infty}\frac{x^i}{i!}$$
$$\frac{\delta}{\delta x}e^x = e^x$$

8.1.3 Differentiating exponents, logarithms and power functions

Differentiating the natural logarithm

$$\frac{\delta}{\delta x}\ln(x) = \lim_{\delta\to 0}\frac{\ln(x+\delta) - ln(x)}{\delta}$$
$$\frac{\delta}{\delta x}\ln(x) = \lim_{\delta\to 0}\frac{\ln\frac{x+\delta}{x}}{\delta}$$
$$\frac{\delta}{\delta x}\ln(x) = \lim_{\delta\to 0}\frac{\ln(1+\frac{\delta}{x})}{\delta}$$
$$\frac{\delta}{\delta x}\ln(x) = \frac{1}{x}\lim_{\delta\to 0}\frac{x}{\delta}\ln(1+\frac{\delta}{x})$$
$$\frac{\delta}{\delta x}\ln(x) = \frac{1}{x}\ln(\lim_{\delta\to 0}(1+\frac{\delta}{x})^{\frac{x}{\delta}})$$
$$\frac{\delta}{\delta x}\ln(x) = \frac{1}{x}\ln(e)$$
$$\frac{\delta}{\delta x}\ln(x) = \frac{1}{x}$$

Differentiating logarithms of other bases

$$log_a(x) = \frac{\log_b(x)}{\log_b(a)}$$
$$log_a(x) = \frac{\ln(x)}{\ln(a)}$$

$$\frac{\delta}{\delta x}\log_a(x) = \frac{\delta}{\delta x}\frac{\ln(x)}{ln(a)}$$

$$\frac{\delta}{\delta x}\log_a(x) = \frac{1}{xln(a)}$$

Exponents

$$y = a^x$$

$$\ln(y) = x\ln(a)$$

$$\frac{\delta}{\delta x}\ln(y) = \frac{\delta}{\delta x}x\ln(a)$$

$$\frac{\delta}{\delta x}\ln(y) = \ln(a)$$

$$\frac{1}{y}\frac{\delta}{\delta x}y = \ln(a)$$

$$\frac{\delta}{\delta x}a^x = a^x\ln(a)$$

Power functions

$$y = x^n$$

$$\frac{\delta}{\delta x}y = \frac{\delta}{\delta x}x^n$$

$$\frac{\delta}{\delta x}y = \frac{\delta}{\delta x}e^{n\ln(x)}$$

$$\frac{\delta}{\delta x}y = \frac{n}{x}e^{n\ln(x)}$$

$$\frac{\delta}{\delta x}y = nx^{n-1}$$

Chapter 9

The sine and cosine functions, and identifying π

9.1 Sine and cosine

9.1.1 Defing sine and cosine using Euler's formula

Euler's formula

Previously we showed that:

$e^x = \sum_{i=0}^{\infty} \dfrac{x^i}{i!}$

Consider:

$e^{i\theta}$

$e^{i\theta} = \sum_{j=0}^{\infty} \dfrac{(i\theta)^j}{j!}$

$e^{i\theta} = [\sum_{j=0}^{\infty} \dfrac{(\theta)^{4j}}{(4j)!} - \sum_{j=0}^{\infty} \dfrac{(\theta)^{4j+2}}{(4j+2)!}] + i[\sum_{j=0}^{\infty} \dfrac{(\theta)^{4j+1}}{(4j+1)!} - \sum_{j=0}^{\infty} \dfrac{(\theta)^{4j+3}}{(4j+3)!}]$

We then use this to define sin and cos functions.

$\cos(\theta) := \sum_{j=0}^{\infty} \dfrac{(\theta)^{4j}}{(4j)!} - \sum_{j=0}^{\infty} \dfrac{(\theta)^{4j+2}}{(4j+2)!}$

$\sin(\theta) := \sum_{j=0}^{\infty} \dfrac{(\theta)^{4j+1}}{(4j+1)!} - \sum_{j=0}^{\infty} \dfrac{(\theta)^{4j+3}}{(4j+3)!}$

So:

$e^{i\theta} = \cos(\theta) + i\sin(\theta)$

Alternative formulae for sine and cosine

We know

$e^{i\theta} = \cos(\theta) + i\sin(\theta)$

$e^{-i\theta} = \cos(\theta) - i\sin(\theta)$

So

$e^{i\theta} + e^{-i\theta} = cos(\theta) + i\sin(\theta) + \cos(\theta) - i\sin(\theta)$

$\cos(\theta) = \dfrac{e^{i\theta} + e^{-i\theta}}{2}$

And

$e^{i\theta} - e^{-i\theta} = cos(\theta) + i\sin(\theta) - \cos(\theta) + i\sin(\theta)$

$\sin(\theta) = \dfrac{e^{i\theta} - e^{-i\theta}}{2i}$

Sine and cosine are odd and even functions

Sine is an odd function.

$\sin(-\theta) = -\sin(\theta)$

Cosine is an even function.

$\cos(-\theta) = \cos(\theta)$

9.1.2 De Moive's formula

$e^{i\theta} = \cos(\theta) + i\sin(\theta)$

Let $\theta = nx$:

$e^{inx} = \cos(nx) + i\sin(nx)$

$(e^{ix})^n = \cos(nx) + i\sin(nx)$

$(\cos(x) + i\sin(x))^n = \cos(nx) + i\sin(nx)$

9.1.3 Expanding sine and cosine

Expansion

$\sin(\alpha + \beta) = \sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta)$

$\cos(\alpha + \beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta)$

9.1.4 Addition of sine and cosine

Adding waves with same frequency

We know that:

$a\sin(bx+c) = a\sin(bx)\cos(c) + a\sin(c)\cos(bx)$

So:

$a\sin(bx+c)+d\sin(bx+e) = a\sin(bx)\cos(c)+a\sin(c)\cos(bx)+d\sin(bx)\cos(e)+d\sin(e)\cos(bx)$

We know that:

$\sin(\theta) = \frac{e^{i\theta} - e^{-i\theta}}{2i}$

So:

$a\sin(bx+c) + d\sin(bx+f) = a\frac{e^{i(bx+c)} - e^{-i(bx+c)}}{2i} + d\frac{e^{i(bx+f)} - e^{-i(bx+f)}}{2i}$

$a\sin(bx+c) + d\sin(bx+f) = \frac{a(e^{i(bx+c)} - e^{-i(bx+c)}) + d(e^{i(bx+f)} - e^{-i(bx+f)})}{2i}$

$a\sin(bx+c) + d\sin(bx+f) = \frac{a(e^{ibx}e^{ic} - e^{-ibx}e^{-ic}) + d(e^{ibx}e^{if} - e^{-ibx}e^{-if})}{2i}$

$a\sin(bx+c) + d\sin(bx+f) = \frac{(e^{ibx}(ae^{ic} + de^{if}) - e^{-ibx}(ae^{-c} + d^{-if})}{2i}$

$a_i\sin(b_ix+c_i) + a_j\sin(b_jx+c_j) = a_i\sin(b_ix+c_i) + a_j\sin(b_ix+c_j)$

$a_i\sin(b_ix+c_i) + a_j\sin(b_jx+c_j) = a_i\sin(b_ix)\cos(c_i) + a_i\sin(c_i)\cos(b_ix) + a_j\sin(b_ix)\cos(c_j) + a_j\sin(c_j)\cos(b_ix)$

9.1.5 Calculus of sine and cosine

Unity

Note that with imaginary numbers we can reverse all is. So:

$e^{i\theta} = \cos(\theta) + i\sin(\theta)$

$e^{-i\theta} = \cos(\theta) - i\sin(\theta)$

$e^{i\theta}e^{-i\theta} = (\cos(\theta) + i\sin(\theta))(\cos(\theta) - i\sin(\theta))$

$e^{i\theta}e^{-i\theta} = \cos(\theta)^2 + \sin(\theta)^2$

$e^{i\theta}e^{-i\theta} = e^{i\theta-i\theta} = e^0 = 1$

So:

$\cos(\theta)^2 + \sin(\theta)^2 = 1$

Note that if $\cos(\theta)^2 = 0$, then $\sin(\theta)^2 = \pm 1$

That is, if the real part of $e^{i\theta}$ is 0, the imaginary part is ± 1. And visa versa.

Similarly if the derivative of the real part of $e^{i\theta}$ is 0, the imaginary part is ± 1. And visa versa.

Sine and cosine are linked by their derivatives

Note that these functions are linked in their derivatives.

$\frac{\delta}{\delta\theta}\cos(\theta) = \sum_{j=0}^{\infty} \frac{(\theta)^{(4j+3)}}{(4j+3)!} - \sum_{j=0}^{\infty} \frac{(\theta)^{4j+1}}{(4j+1)!}$

$\frac{\delta}{\delta\theta}\cos(\theta) = -\sin(\theta)$

Similarly:

$\frac{\delta}{\delta\theta}\sin(\theta) = cos(\theta)$

Both sine and cosine oscillate

$\frac{\delta^2}{\delta\theta^2}\sin(\theta) = -\sin(\theta)$

$\frac{\delta^2}{\delta\theta^2}\cos(\theta) = -\cos(\theta)$

So for either of:

$y = \cos(\theta)$

$y = \sin(\theta)$

We know that

$\frac{\delta^2}{\delta\theta^2}y(\theta) = -y(\theta)$

Consider $\theta = 0$.

$e^{i.0} = \cos(0) + i\sin(0)$

$1 = \cos(0) + i\sin(0)$

$\sin(0) = 0$

$\cos(0) = 1$

Similarly we know that the derivative:

$\sin'(0) = \cos(0) = 1$

$\cos'(0) = -\sin(0) = 0$

Consider $\cos(\theta)$.

As $\cos(0)$ is static at $\theta = 0$, and is positive, it will fall until $\cos(\theta) = 0$.

While this is happening, $\sin(\theta)$ is increasing. As:

$\cos(\theta)^2 + \sin(\theta)^2 = 1$

$\sin(\theta)$ will equal 1 where $\cos(\theta) = 0$.

Due to symmetry this will repeat 4 times.

Let's call the length of this period τ.

Where $\theta = \tau * 0$

- $\cos(\theta) = 1$
- $\sin(\theta) = 0$

Where $\theta = \tau * \frac{1}{4}$

- $\cos(\theta) = 0$
- $\sin(\theta) = 1$

Where $\theta = \tau * \frac{2}{4}$

- $\cos(\theta) = -1$
- $\sin(\theta) = 0$

Where $\theta = \tau * \frac{3}{4}$

- $\cos(\theta) = 0$
- $\sin(\theta) = -1$

Relationship between$\cos(\theta)$ and $\sin(\theta)$

Note that $\sin(\theta + \frac{\tau}{4}) = \cos(\theta)$

Note that $\sin(\theta) = \cos(\theta)$ at

- $\tau * \frac{1}{8}$
- $\tau * \frac{5}{8}$

And that all these answers loop. That is, add any integer multiple of τ to θ and the results hold.

$e^{i\theta} = e^{i\theta + n\tau}$

$n \in \mathbb{N}$

$e^{i\theta} = \cos(\theta) + i\sin(\theta)$

$e^{i\theta} = \cos(\theta + n\tau) + i\sin(\theta + n\tau)$

$e^{i\theta} = e^{i(\theta+n\tau)}$

Calculus of trig

Relationship between cos and sine

$\sin(x + \frac{\pi}{2}) = \cos(x)$

$\cos(x + \frac{\pi}{2}) = -\sin(x)$

$\sin(x + \pi) = -\sin(x)$

$\cos(x + \pi) = -\cos(x)$

$\sin(x + \tau) = \sin(x)$

$\cos(x + \tau) = \cos(x)$

Chapter 10

Polar coordinates

10.1 Polar coordinates

10.1.1 Polar co-ordinates

All complex numbers can be shown in polar form

Consider a complex number

$z = a + bi$

We can write this as:

$z = r\cos(\theta) + ir\sin(\theta)$

Polar forms are not unique

Because the functions loop:

$ae^{i\theta} = a(\cos(\theta) + i\sin(\theta))$

$ae^{i\theta} = a(\cos(\theta + n\tau) + i\sin(\theta + n\tau))$

$ae^{i\theta} = ae^{i\theta + n\tau}$

Additionally:

$ae^{i\theta} = a(\cos(\theta) + i\sin(\theta))$

$ae^{i\theta} = a(\cos(\theta) + i\sin(\theta))$

$ae^{i\theta} = -a(\cos(\theta) - i\sin(\theta))$

$ae^{i\theta} = -a(\cos(\theta + \frac{\pi}{2}) + i\sin(\theta + \frac{\pi}{2}))$

Real and imaginary parts of a complex number in polar form

We can extract the real and imaginary parts of this number.

$Re(z) := r\cos(\theta)$

$Im(z) := r\sin(\theta)$

Alternatively:

$Re(z) = r\frac{e^{i\theta}+e^{-i\theta}}{2}$

$Im(z) = r\frac{e^{i\theta}-e^{-i\theta}}{2i}$

10.1.2 Moving between polar and cartesian coordinates

All polar numbers can be shown as Cartesian

$ae^{i\theta} = a(\cos(\theta) + i\sin(\theta))$

$ae^{i\theta} = a\cos(\theta) + ia\sin(\theta)$

$z = a + bi$

$e^{i\theta} =$

$e^{x} = \sum_{i=0}^{\infty}\frac{x^i}{i!}$

10.1.3 Arithmetic of polar coordinates

Addition

$z_3 = z_1 + z_2$

$z_3 = a_1e^{i\theta_1} + a_2e^{i\theta_2}$

$z_3 = a_1[\cos(\theta_1) + i\sin(\theta_1)] + a_2[\cos(\theta_2) + i\sin(\theta_2)]$

$z_3 = [a_1\cos(\theta_1) + a_2\cos(\theta_2)] + i[a_1\sin(\theta_1) + a_2\sin(\theta_2)]$

Multiplication

$z_3 = z_1.z_2$

$z_3 = a_1e^{i\theta_1}a_2e^{i\theta_2}$

$z_3 = a_1a_2e^{i(\theta_1+\theta_2)}$

$a_3 = a_1a_2$

$\theta_3 = \theta_1 + \theta_2$

Chapter 11

Power series, Taylor series and Maclaurin series

11.1 Power series

11.1.1 Power series

of the form:

$\sum_{n=0} a_n (x-c)^n$

Smoothness of power series

Power series are all smooth. That is, they are infinitely differentiable.

11.2 Taylor series

11.2.1 Taylor series

$f(x)$ can be estimated at point c by identifying its repeated differentials at point c.

The coefficients of an infinate number of polynomials at point c allow this.

$f(x) = \sum_{i=0}^{\infty} a_i (x-c)^i$

$f'(x) = \sum_{i=1}^{\infty} a_i (x-c)^{i-1} i$

$f''(x) = \sum_{i=2}^{\infty} a_i (x-c)^{i-2} i(i-1)$

$f^j(x) = \sum_{i=j}^{\infty} a_i (x-c)^{i-j} \dfrac{i!}{(i-j)!}$

For $x = c$ only the first term in the series is non-zero.

$f^j(c) = \sum_{i=j}^{\infty} a_i (c-c)^{i-j} \frac{i!}{(i-j)!}$

$f^j(c) = a_i j!$

So:

$a_j = \frac{f^j(c)}{j!}$

So:

$f(x) = \sum_{i=0}^{\infty} (x-c)^i \frac{f^i(c)}{i!}$

11.2.2 Convergence

If $x = c$ then the power series will be equal to a_0.

For other values the power series may not converge.

11.2.3 Cauchy-Hadamard theorem

Radius of convergence:

$\frac{1}{R} = \limsup_{n\to\infty}(|a_n|^{\frac{1}{n}})$

11.2.4 Maclaurin series

A Taylor series around $c = 0$.

$f(x) = \sum_{i=0}^{\infty} (x-c)^i \frac{f^i(c)}{i!}$

$f(x) = \sum_{i=0}^{\infty} (x)^i \frac{f^i(0)}{i!}$

For example, for:

$f(x) = (1-x)^{-1}$

$f^i(0) = i!$

So, around $x = 0$:

$f(x) = \sum_{i=0}^{\infty} (x)^i$

11.2.5 Analytic functions

(root test, direct comparison test, rate of convergence, radius of convergence)

Chapter 12

Matrix exponents and Taylor series of matrices

12.1 Taylor series of matrices

12.1.1 Taylor series of matrices

We can also use Taylor series to evaluate functions of matrices.

Consider e^M

We can evaluate this as:

$e^M = \sum_{k=0}^\infty \frac{1}{k!}M^k$

Part III

Univariate real integration

Chapter 13

The Riemann integral, definite and indefinite integrals, and anti-derivatives

13.1 The Riemann integral

13.1.1 Riemann sums

Given a function $f(x)$ and an interval $[a, b]$, we can divide $[a, b]$ into n sections and calculate:

$\sum_{j=0}^{n(b-a)} f(a + \frac{j}{n})$

This is the Riemann sum.

13.1.2 Riemann integral

We take the limit of the Riemann sum as $n \to \infty$

$\int_a^b f(x)dx := \lim_{n\to\infty} \sum_{j=0}^{n(b-a)} f(a + \frac{j}{n})$

13.1.3 Linearity

$\int_a^b f(x) + g(x)dx = \lim_{n\to\infty} \sum_{j=0}^{n(b-a)} f(a + \frac{j}{n}) + g(a + \frac{j}{n})$

$\int_a^b f(x) + g(x)dx = \lim_{n\to\infty} \sum_{j=0}^{n(b-a)} f(a + \frac{j}{n}) + \lim_{n\to\infty} \sum_{j=0}^{n(b-a)} g(a + \frac{j}{n})$

$\int_a^b f(x) + g(x)dx = \int_a^b f(x)dx + \int_a^b g(x)dx$

13.1.4 Continuation

$\int_a^b f(x)dx + \int_b^c f(x)dx = \lim_{n\to\infty} \sum_{j=0}^{n(b-a)} f(a + \frac{j}{n}) + \lim_{n\to\infty} \sum_{j=0}^{n(c-b)} f(b + \frac{j}{n})$

$\int_a^b f(x)dx + \int_b^c f(x)dx = \lim_{n\to\infty} [\sum_{j=0}^{n(b-a)} f(a + \frac{j}{n}) + \sum_{j=0}^{n(c-b)} f(b + \frac{j}{n})]$

$\int_a^b f(x)dx + \int_b^c f(x)dx = \lim_{n\to\infty} [\sum_{j=0}^{n(b-a)} f(a + \frac{j}{n}) + \sum_{j=n(b-a)}^{n(c-b)+n(b-a)} f(b + \frac{j - n(b-a)}{n})]$

$\int_a^b f(x)dx + \int_b^c f(x)dx = \lim_{n\to\infty} [\sum_{j=0}^{n(b-a)} f(a + \frac{j}{n}) + \sum_{j=n(b-a)}^{n(c-a)} f(a + \frac{j}{n})]$

$\int_a^b f(x)dx + \int_b^c f(x)dx = \lim_{n\to\infty} [\sum_{j=0}^{n(c-a)} f(a + \frac{j}{n})]$

$\int_a^b f(x)dx + \int_b^c f(x)dx = \int_a^c f(x)dx$

13.2 Definite and indefinite integrals

13.2.1 Definite integrals

Definite integrals are between two points.

$\int_0^1 f(x)dx$

13.2.2 Indefinite integrals

Indefinite integrals are not. Eg +c at end. The antiderivative.

$\int f(x)dx$

13.2.3 Unsigned definite integral

$\int_{[0,1]} f(x)dx$

13.3 Anti-derivatives

13.3.1 Anti-derivative

Taking the derivative of a function provides another function. The anti-derivative of a function is a function which, when differentiated, provides the original function.

As this function can include any additive constant, there are an infinite number of anti-derivatives for any function.

Chapter 14

Integration by parts

14.1 Integration by parts

14.1.1 Integration by parts

We have:

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

We want that in terms of y.

We know from the product rule of differentiation:

$y = a(x)b(x)$

Means that:

$\frac{\delta y}{\delta x} = a'(x)b(x) + a(x)b'(x)$

So let's relabel $f(x)$ as $h'(x)$

δ

$\frac{\delta y}{\delta x} = h'(x)g(x)$

$\frac{\delta y}{\delta x} + h(x)g'(x) = h'(x)g(x) + h(x)g'(x)$

$y + \int h(x)g'(x) = \int h'(x)g(x) + h(x)g'(x)$

$y + \int h(x)g'(x) = h(x)g(x)$

$y = h(x)g(x) - \int h(x)g'(x)$

For example:

$\frac{\delta y}{\delta x} = x.\cos(x)$

$f(x) = \cos(x)$

$g(x) = x$

$h(x) = \sin(x)$

$g'(x) = 1$

So:

$y = x \int \cos(x)dx - \int \sin(x)dx$

$y = x\sin(x) - \cos(x) + c$

Chapter 15

The fundamental theorem of calculus

15.1 The fundamental theorem of calculus

15.1.1 Mean value theorem for integration

Take function $f(x)$. From the extreme value theorem we know that:

$\exists m \in \mathbb{R} \exists M \in \mathbb{R} \forall x \in [a, b](m < f(x) < M)$

15.1.2 Fundamental theorem of calculus

From continuation we know that:

$\int_a^{x_1} f(x)dx + \int_{x_1}^{x_1+\delta x} f(x)dx = \int_a^{x_1+\delta x} f(x)dx$

$\int_x^{x_1+\delta x} f(x)dx = \int_a^{x_1+\delta} f(x)dx - \int_a^{x_1} f(x)dx$

Indefinite integrals

Chapter 16

Lebesque integrals

16.1 Lebesque integrals

16.1.1 Lebesque integrals

Chapter 17

The tangent function, and evaluating π

17.1 Tangent

17.1.1 Tan

The $\tan(\theta)$ function is defined as:

$\tan(\theta) := \frac{\sin(\theta)}{\cos(\theta)}$

Behaviour around 0

$\sin(0) = 0$

$\cos(0) = 1$

$\tan(0) := \frac{\sin(0)}{\cos(0)}$

$\tan(0) = \frac{0}{1}$

$\tan(0) = 0$

Behaviour around $\cos(\theta) = 0$

$\tan(\theta) = \frac{\sin(\theta)}{\cos(\theta)}$

So $\tan(\theta)$ is undefined where $\cos(\theta) = 0$.

This happens where:

$\theta = \frac{\tau}{4} + \frac{1}{2}n\tau$

$\theta = \frac{1}{4}\tau(1+2n)$

Where $n \in \mathbb{Z}$.

Derivatives

$\tan(\theta) = \frac{\sin(\theta)}{\cos(\theta)}$

$\frac{\delta}{\delta\theta}\tan(\theta) = \frac{\delta}{\delta\theta}\frac{\sin(\theta)}{\cos(\theta)}$

$\frac{\delta}{\delta\theta}\tan(\theta) = \frac{\cos(\theta)}{\cos(\theta)} + \frac{\sin^2(\theta)}{\cos^n(\theta)}$

$\frac{\delta}{\delta\theta}\tan(\theta) = 1 + \tan^2(\theta)$

Note this is always positive. This means:

$\lim_{\cos(\theta)\to 0^+} = -\infty$

$\lim_{\cos(\theta)\to 0^-} = \infty$

17.1.2 Inverse functions

Inverse trigonometric functions

$\sin(\arcsin(\theta)) := \theta$

$\cos(\arccos(\theta)) := \theta$

$\tan(\arctan(\theta)) := \theta$

17.1.3 Integrals

Cosine and sine

$\arccos(\theta)$, $\arcsin(\theta)$ and difficulty of inversing

In order to determine τ we need inverse functions for $\cos(\theta)$ or $\sin(\theta)$.

These are the $\arccos(\theta)$ and $\arcsin(\theta)$ functions respectively.

However this is not easily calculated. Instead we look for another function.

Calculating $\arctan(\theta)$

So we want a function to inverse this. This is the $\arctan(\theta)$ function.

If $y = \tan(\theta)$, then:

$\theta = \arctan(y)$

We know the derivative for $\tan(\theta)$ is:

$$\frac{\delta}{\delta\theta}\tan(\theta) = 1 + \tan^2(\theta)$$

$$\frac{\delta y}{\delta\theta} = 1 + y^2$$

So

$$\frac{\delta\theta}{\delta y} = \frac{1}{1+y^2}$$

$$\frac{\delta}{\delta y}\arctan(y) = \frac{1}{1+y^2}$$

So the value for $\arctan(k)$ is:

$$\arctan(k) = \arctan(a) + \int_a^k \frac{\delta}{\delta y}\arctan(y)\delta y$$

$$\arctan(k) = \arctan(a) + \int_a^k \frac{1}{1+y^2}\delta y$$

What do we know about this function? We know it can map to multiple values of θ because the underlying $\sin(\theta)$ and $\cos(\theta)$ functions also loop.

We know that one of the results for $\arctan(0)$ is 0.

17.1.4 Calculating τ

As we note above, $\sin(\theta) = \cos(\theta)$ at $\theta = \tau * \frac{1}{8}$

This is also where $\tan(\theta) = 1$.

$$\arctan(k) = \arctan(a) + \int_a^k \frac{1}{1+y^2}\delta y$$

We start from $a = 0$.

$$\arctan(k) = \arctan(0) + \int_0^k \frac{1}{1+y^2}\delta y$$

We know that one of the results for $\arctan(0)$ is 0.

$$\arctan(k) = \int_0^k \frac{1}{1+y^2}\delta y$$

We want $k = 1$

$$\arctan(1) = \int_0^1 \frac{1}{1+y^2}\delta y$$

$$\frac{\tau}{8} = \int_0^1 \frac{1}{1+y^2}\delta y$$

$\tau = 8\int_0^1 \frac{1}{1+y^2}\delta y$

We know that the $\cos(\theta)$ and $\sin(\theta)$ functions cycle with period τ.

Therefore $cos(n.\tau) = \cos(0)$

Chapter 18

Other trigonometric functions

18.1 Other

18.1.1 Other functions

Reciprocal trigonometric functions

Standard

$\csc(\theta) := \dfrac{1}{\sin(\theta)}$

$\sec(\theta) := \dfrac{1}{\cos(\theta)}$

$\cot(\theta) := \dfrac{1}{\tan(\theta)}$

Hyperbolic

$csch(\theta) := \dfrac{1}{\sinh(\theta)}$

$sech(\theta) := \dfrac{1}{\cosh(\theta)}$

$coth(\theta) := \dfrac{1}{\tanh(\theta)}$

Inverse trigonometric functions

Reciprocal standard

$\csc(arccsc(\theta)) := \theta$

$\sec(arcsec(\theta)) := \theta$

$\cot(arccot(\theta)) := \theta$

Reciprocal hyperbolic

$csch(arccsch(\theta)) := \theta$

$sech(arcsech(\theta)) := \theta$

$coth(arccoth(\theta)) := \theta$

18.2 Hyperbolic functions

18.2.1 Hyperbolic functions

Hyperbolic functions

$\sinh(\theta) := \sin(i\theta)$

$\cosh(\theta) := \cos(i\theta)$

$\tanh(\theta) := \tan(i\theta)$

Inverse trigonometric functions

$\sinh(arcsinh(\theta)) := \theta$

$\cosh(arccosh(\theta)) := \theta$

$\tanh(arctan(\theta)) := \theta$

Chapter 19

Fourier analysis

19.1 Fourier analysis

19.1.1 Representing wave functions

Wave function are of the form:

$\cos(ax+b)$

$\sin(ax+b)$

We can use the following identities:

- $\cos(x) = \sin(x + \frac{\tau}{8})$
- $\sin(-x) = -\sin(x)$
- $\sin(a+b) = \sin(a)\cos(b) + \sin(b)\cos(a)$

So we can write any function as:

Using e

19.1.2 Harmonics

19.1.3 Fourier series

Fourier series

Motivation: we have a function we want to display as another sort of function.

More specifically, a function can be shown as a combination of sinusoidal waves.

To frame this let's imagine a sound wave, with values $f(t)$ for all time values t. We can imagine this as a summation of sinusoidal functions. That is:

$f(t) = \sum_{n=0}^{\inf} a_n cos(nw_0 t)$

We want to get another function $F(\xi)$ for all frequencies ξ.

Combinations of wave functions

We can add sinusoidal waves to get new waves.

For example

$s_N(x) = 2\sin(x+3) + \sin(-4x) + \frac{1}{2}\cos(x)$

As a summation of series

We can simplify arbitrary series using the following identities:

$\cos(x) = \sin(x + \frac{\tau}{8})$

$\sin(-x) = -\sin(x)$

So we have:

$s(x) = 2\sin(x+3) - \sin(4x) + \frac{1}{2}\sin(x + \frac{\tau}{8})$

We can put this into the following format:

$s(x) = \sum_{i=1}^{m} a_i \sin(b_i x + c_i)$

Where:

$a = [2, -1, \frac{1}{2}]$

$b = [1, 4, 1]$

$c = [3, 0, \frac{\tau}{8}]$

Ordering by b

We can move terms around to get:

$s(x) = \sum_{i=1}^{m} a_i \sin(b_i x + c_i)$

Where:

$a = [2, \frac{1}{2}, -1]$

$b = [1, 1, 4]$

$c = [3, \frac{\tau}{8}, 0]$

Adding waves with same frequency

We know that:

$\sin(a+b) = \sin(a)\cos(b) + \sin(b)\cos(a)$

So:

$\sin(b_i x + c_i) = \sin(b_i x)\cos(c_i) + \sin(c_i)\cos(b_i x)$

If 2 terms have the same value for b_i, then:

$a_i \sin(b_i x + c_i) + a_j \sin(b_j x + c_j) = a_i \sin(b_i x + c_i) + a_j \sin(b_i x + c_j)$

$a_i \sin(b_i x + c_i) + a_j \sin(b_j x + c_j) = a_i \sin(b_i x)\cos(c_i) + a_i \sin(c_i)\cos(b_i x) + a_j \sin(b_i x)\cos(c_j) + a_j \sin(c_j)\cos(b_i x)$

So we now get for:

$s(x) = \sum_{i=1}^{m} a_i \sin(b_i x + c_i)$

$a = [, -1]$

$b = [, 4]$

$c = [, 0]$

19.1.4 Fourier transforms

Fourier transform

$\hat{f}(\Xi) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i x \Xi} dx$

Inverse Fourier transform

$f(x) = \int_{-\infty}^{\infty} \hat{f}(\Xi) e^{2\pi i x \Xi} d\Xi$

Fourier inversion theorem

Chapter 20

Generalising factorials: The gamma function

20.1 Introduction

20.1.1 Gamma function

The gamma function expands the factorial function to the real (and complex) numbers

We want:

$f(1) = 1$

$f(x+1) = xf(x)$

There are an infinite number of functions which fit this. The function could fluctuate between the natural numbers.

The function we use is:

$\Gamma(z) = \int_0^\infty x^{z-1}e^{-x}dx$

Part IV

Ordinary differential equations

Chapter 21

First-order Ordinary Differential Equations (ODEs)

21.1 Introduction

21.1.1 Order of differential equations

21.1.2 Implicit and explit differential equations

An ordinary differential equation is one with only one independent variable. For example:

$\frac{dy}{dx} = f(x)$

The order of a differential equation is the number of differentials of y included. For example one with the second derivative of y is of order 2.

Ordinary equations can can either implicit or explicit. An explicit function shows the highest order derivative as a function of other terms.

An implicit function is one which is not explicit.

A linear ODE is an explicit ODE where the derivative terms of y do not multiply together, that is, in the form:

$y^{(n)} = \sum_i a_i(x)y^{(i)} + r(x)$

First-order ODEs

We have an evolution:

$\frac{dy}{dt}=f(t,y)$

And a starting condition:

$y_0=f(t_0)$

We now discuss various ways to solve these.

21.2 First-order Ordinary Differential Equations

21.2.1 Ordinary differential equations

An ordinary differential equation is one with only one independent variable. For example:

$\frac{dy}{dx}=f(x)$

The order of a differential equation is the number of differentials of y included. For example one with the second derivative of y is of order 2.

Ordinary equations can can either implicit or explicit. An explicit function shows the highest order derivative as a function of other terms.

An implicit function is one which is not explicit.

A linear ODE is an explicit ODE where the derivative terms of y do not multiply together, that is, in the form:

$y^{(n)}=\sum_i a_i(x)y^{(i)}+r(x)$

First-order ODEs

We have an evolution:

$\frac{dy}{dt}=f(t,y)$

And a starting condition:

$y_0=f(t_0)$

We now discuss various ways to solve these.

21.2.2 Linear first-order Ordinary Differential Equations

Linear ODEs

For some we can write:

$\frac{dy}{dt}=f(t,y)$

$\frac{dy}{dt}=q(t)-p(t)y$

This can be solved by multiplying by an unknown function $\mu(t)$:

$\frac{dy}{dt} + p(t)y = q(t)$

$\mu(t)[\frac{dy}{dt} + p(t)y] = \mu(t)q(t)$

We can then set $\mu(t) = e^{\int p(t)dt}$. This means that $\frac{d\mu}{dt} = p(t)u(t)$

$\frac{d}{dt}[\mu(t)y] = \mu(t)q(t)$

$\mu(t)y = \int \mu(t)q(t)dt + C$

In some cases, this can then be solved.

Example

$\frac{\delta y}{\delta x} = cy$

$y = Ae^{c(y+a)}$

$\frac{\delta^2 y}{\delta x^2} = cy$

$y = Ae^{\sqrt{c}(y+a)}$

21.2.3 Separable first-order Ordinary Differential Equations

For some we can write:

$\frac{dy}{dt} = f(t, y)$

$\frac{dy}{dt} = \frac{g(t)}{h(y)}$

We can then do the following:

$h(y)\frac{dy}{dt} = g(t)$

$\int h(y)\frac{dy}{dt}dt = \int g(t)dt + C$

$\int h(y)dy = \int g(t)dt + C$

In some cases, these functions can then be integrated and solved.

21.3 Second-order Ordinary Differential Equations

21.3.1 Linear second-order Ordinary Differential Equations

These are of the form

$$\frac{d^2y}{dt^2} + p(t)\frac{dy}{dt} + q(t)y = g(t)$$

There are two types. Homogenous equations are where $g(t) = 0$. Otherwise they are heterogenous.

We explore the case with constants:

$$a\frac{d^2y}{dt^2} + b\frac{dy}{dt} + cy = 0$$

Chapter 22

Second-order Ordinary Differential Equations (ODEs)

Part V

Univariate optimisation

Chapter 23

Univariate optimisation

23.1 Unconstrained optimisation

23.1.1 Introduction to unconstrained optimisation

Goals

We want to identify either the maximum or the minimum.

There exist local minima and global minima.

Optimising through limits

If we are looking to minimise a function, and the limits are ∞ or $-\infty$ then we can optimise by taking large or small values.

We can examine this for each variable.

This also applies for maximising a function.

Optimisation through stationary points

Stationary points of a function are points where marginal changes do not have an impact on the value of the function. As a result they are either local maxima or minima.

Optimisation through algorithms

If we cannot identify stationary points easily, we can instead use algorithms to identify optima.

Stationary points of strictly concave and convex functions

If a function is strictly concave it will only have one stationary point, a local, and global, maxima.

If a function is strictly convex it will only have one stationary point, a local, and global, minima.

23.1.2 Local optima

23.1.3 Optimising convex functions

23.1.4 Analytic optimisation

Convex and concave functions

Convex functions only have one minimum, and concave functions have only one maximum.

If a function is not concave or convex, it may have multiple minima

If a function is convex, then there is only one critical point – the local minimum. We can identify this this by looking for critical points using first-order conditions.

Similarly, if a function is concave, then there is only one critical point – the local maximum.

We can identify whether a function is concave or convex by evaluating the Hessian matrix.

Evaluating multiple local optima

We can evaluate each of the local minima or maxima, and compare the sizes.

We can identify these by taking partial derivatives of the function in question and identifying where this function is equal to zero.

$u = f(x)$

$u_{x_i} = \frac{\delta f}{\delta x_i} = 0$

We can then solve this bundle of equations to find the stationary values of x.

After identifying the vector x for these points we can then determine whether or not the points are minima or maxima by examining the second derivative at these points. If it is positive it is a local minima, and therefore not an optimal point. Points beyond these will be higher, and may be higher than any local maxima.

23.1.5 Stationary points and first-order conditions

23.1.6 Local minima, maxima and inflection points

23.1.7 Optimising convex and non-convex differentiable functions

Part VI

Multivariate real analysis: Scalar fields

Chapter 24

Multivariate functions

24.1 Multivariate space

24.1.1 Regions

A region is a subset

Type-I regions (y-simple regions)

Type-II regions (x-simple regions)

Elementary regions

An elementary region is a region which is either a type-I region or a type-II region.

Simple regions

A simple region is a region which is both a type-I and a type-II region.

24.1.2 Curves and closed curves

In a space we can identify a curve between two points. If the input in the real numbers then this curve is unique.

For more general scalar fields this will not be the case. Two points in $\mathbb{R}^2$ could be joined by an infinite number of paths.

A curve can be defined as a function on the real numbers. The curve itself is totally ordered, and homogenous to the real number line.

We can write the curve therefore as:

$r : [a, b] \to C$

Where a and b are the start and end points of the curve, and C is the resulting curve.

Closed curves

If the start and end point of the curve are the same then the curve is closed. We can write this as:

$\oint_C f(r)ds = \int_a^b f(r(t))|r'(t)|dt$

24.1.3 Surfaces

24.1.4 Length of a curve

We have a curve from a to b in $\mathbf{R}^n$.

$f : [a, b] \rightarrow \mathbf{R}^n$

We divide this into n segments.

The ith cut is at:

$t_i = a + \frac{i}{n}(b - a)$

So the first cut is at:

$t_0 = a$

$t_n = b$

The distance between two sequential cuts is:

$||f(t_i) - f(t_{i-1}||$

The sum of all these differences is:

$L = \sum_{i=1}^n ||f(t_i) - f(t_{i-1})||$

The limit is:

$L = \lim_{n\rightarrow\infty} \sum_{i=1}^n ||f(t_i) - f(t_{i-1})||$

Method 1

$L = \lim_{n\rightarrow\infty} \sum_{i=1}^n ||f(t_i) - f(t_{i-1})||$

$L = \lim_{n\rightarrow\infty} \sum_{i=1}^n ||\frac{f(t_i) - f(t_{i-1})}{\Delta t}||\Delta t$

$L = \lim_{n\rightarrow\infty} \sum_{i=1}^n ||f'(t)||\Delta t$

$L = \int_a^b ||f'(t)||dt$

Method 2

$L = \lim_{n\to\infty} \sum_{i=1}^{n} ||f(t_i) - f(t_{i-1})||$

$L = \lim_{n\to\infty} \sum_{i=1}^{n} \sqrt{(f(t_i) - f(t_{i-1}))^* M(f(t_i) - f(t_{i-1}))}$

$L = \int_a^b \sqrt{(dt)^T M(dt)}$

Chapter 25

Multivariate differentiation of scalar fields - partial differentiation, del and grad

25.1 Partial differentiation of scalar fields

25.1.1 Scalar fields

A scalar field is a function on an underlying input which produces a real output.

Inputs are not limited to real numbers. In this section we consider functions on vector spaces.

25.1.2 Del

$\nabla = (\sum_{i=1}^n e_i \frac{\delta}{\delta x_i})$

Where e are the basis vectors.

This on its own means nothing. It is similar to the partial differentiation function.

25.1.3 Gradient

In a scalar field we can calculate the partial derivative at any point with respect to one input.

We may wish to consider these collectively. To do that we use the gradient operator.

We previously introduced the Del operator where:

$\nabla = (\sum_{i=1}^n e_i \frac{\delta}{\delta x_i})$

Where e are the basis vectors.

This on its own means nothing. It is similar to the partial differentiation function.

We now multiply Del by the function. This gives us:

$\nabla f = (\sum_{i=1}^n e_i \frac{\delta f}{\delta x_i})$. This gives us a vector in the underlying vector space.

This is the gradient.

Chapter 26

Total differentiation of scalar fields

26.1 Total differentiation of scalar fields

26.1.1 Total differentiation

Consider a multivariate function.

$f(x)$.

We can define:

$\Delta f(x, \Delta x) := f(x + \Delta x) - f(x)$

$\Delta f(x, \Delta x) = \sum_{i=1}^n f(x + \Delta x_i + \sum_{j=0}^{i-1} \Delta x_j) - f(x + \sum_{j=0}^{i-1} \Delta x_j)$

$\Delta f(x, \Delta x) = \sum_{i=1}^n \Delta x_i \frac{f(x + \Delta x_i + \sum_{j=0}^{i-1} \Delta x_j) - f(x + \sum_{j=0}^{i-1} \Delta x_j)}{\Delta x_i}$

$\frac{\Delta f}{\Delta x_k} = \sum_{i=1}^n \frac{\Delta x_i}{\Delta x_k} \frac{f(x + \Delta x_i + \sum_{j=0}^{i-1} \Delta x_j) - f(x + \sum_{j=0}^{i-1} \Delta x_j)}{\Delta x_i}$

$\lim_{\Delta x_k \to 0} \frac{\Delta f}{\Delta x_k} = \sum_{i=1}^n \lim_{\Delta x_k \to 0} \frac{\Delta x_i}{\Delta x_k} \frac{f(x + \Delta x_i + \sum_{j=0}^{i-1} \Delta x_j) - f(x + \sum_{j=0}^{i-1} \Delta x_j)}{\Delta x_i}$

$\frac{df}{dx_k} = \sum_{i=1}^n \frac{dx_i}{dx_k} \frac{\delta f}{\delta x_i}$

26.1.2 Total differentiation of a univariate function

For a univariate function total differentiation is the same as partial differentiation.

$$\frac{df}{dx} = \frac{dx}{dx}\frac{\delta f}{\delta x}$$
$$\frac{df}{dx} = \frac{\delta f}{\delta x}$$

Chapter 27

Directional derivative of scalar fields

27.1 Directional derivative of scalar fields

27.1.1 Directional derivative

We have a function, $f(\mathbf{x})$.

Given a vector v, we can identify by how much this scalar function changes as you move in that direction.

$\nabla_v f(x) := \lim_{\delta \to 0} \dfrac{f(\mathbf{x}+\delta\mathbf{v}) - f(\mathbf{x})}{\delta}$

The directional derivative is the same dimension as underlying field.

Other

Differentiation of scalar field, df, can be defined as a vector field where grad is 0. can differ with orientation, scale

Chapter 28

Multivariate integration of scalar fields

28.1 Integration of scalar fields

28.1.1 Line integral of scalar fields

28.1.2 Double integral of scalar fields

28.1.3 Surface integral of scalar fields

28.1.4 Gradient theorem

In a scalar field, the line integral of the gradient field is the difference between the value of the scalar field at the start and end points.

This generalises the fundamental theorem of calclulus.

28.1.5 Green's theorem

We have a curve C on a plane.

Inside this is region D.

We have two functions: $L(x, y)$ and $M(x, y)$ defined on the region and curve.

$\oint_C (Ldx + Mdy) = \int \int_D (\frac{\delta M}{\delta x} - \frac{\delta L}{\delta y})dxdy$

28.1.6 Differential forms

Type-I

For type-I, we can integrate over y, then integrate over x.

Type-II

For type-II, we can integrate over x, then integrate over y.

Chapter 29

Generalising the binomial coefficient formula: The beta function

29.1 Introduction

29.1.1 Beta function

The beta function expand the binomial coefficient formula to the real (and complex) numbers.

We want to expand the binomial coefficient function.

$(\frac{n}{k}) = \frac{n!}{k!(n-k)!}$

We do this as:

$B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$

Chapter 30

Multivariate optimisation of scalar fields

30.1 Unconstrained multivariate optimisation

30.1.1 Introduction

30.2 Optimisation with linear equality constraints

30.2.1 Single equality constraint

Constrained optimisation

Rather than maximise $f(x)$, we want to maximise $f(x)$ subject to $g(x) = 0$.

We write this, the Lagrangian, as:

$\mathcal{L}(x, \lambda) = f(x) - \sum_k^m \lambda_k [g_k(x) - c_k]$

We examine the stationary points for both vector x and λ. By including the latter we ensure that these points are consistent with the constraints.

Solving the Langrangian with one constraint

Our function is:

$\mathcal{L}(x, \lambda) = f(x) - \lambda[g(x) - c]$

The first-order conditions are:

$\mathcal{L}_\lambda = -[g(x) - c]$

$\mathcal{L}_{x_i} = \frac{\delta f}{\delta x_i} - \lambda \frac{\delta g}{\delta x_i}$

The solution is stationary so:

$\mathcal{L}_{x_i} = \frac{\delta f}{\delta x_i} - \lambda \frac{\delta g}{\delta x_i} = 0$

$\lambda \frac{\delta g}{\delta x_i} = \frac{\delta f}{\delta x_i}$

$\lambda = \frac{\frac{\delta f}{\delta x_i}}{\frac{\delta g}{\delta x_i}}$

Finally, we can use the following in practical applications.

$\frac{\frac{\delta f}{\delta x_i}}{\frac{\delta g}{\delta x_i}} = \frac{\frac{\delta f}{\delta x_j}}{\frac{\delta g}{\delta x_j}}$

30.2.2 Multiple equality constraints

Solving the Langrangian with many constraints

This time we have:

$\mathcal{L}_{x_i} = \frac{\delta f}{\delta x_i} - \sum_k^m \lambda_k \frac{\delta g_k}{\delta x_i} = 0$

$\mathcal{L}_{x_j} = \frac{\delta f}{\delta x_j} - \sum_k^m \lambda_k \frac{\delta g_k}{\delta x_j} = 0$

$\frac{\delta f}{\delta x_i} - \sum_k^m \lambda_k \frac{\delta g_k}{\delta x_i} = \frac{\delta f}{\delta x_j} - \sum_k^m \lambda_k \frac{\delta g_k}{\delta x_j}$

30.3 Linear programming

30.3.1 Inequality constraints

linear programming means of the form max $c^T x$ st. $Ax <= b$ $x >= 0$ this is the canonical form

Lagrangians with inequality constraints

We can add constraints to an optimisation problem. These constraints can be equality constraints or inequality constraints. We can write constrained optimisation problem as:

Minimise $f(x)$ subject to

$g_i(x) \leq 0$ for $i = 1, ..., m$

$h_i(x) = 0$ for $i = 1, \ldots, p$

We write the Lagrangian as:

$\mathcal{L}(x,\lambda,\nu)=f(x)+\sum_{i=1}^m \lambda_i g_i(x)+\sum_{i=1}^p \nu_i h_i(x)$

If we try and solve this like a standard Lagrangian, then all of the inequality constraints will instead by equality constraints.

Affinity of the Lagrangian

The Lagrangian function is affine with respect to λ and ν.

$\mathcal{L}(x,\lambda,\nu)=f(x)+\sum_{i=1}^m \lambda_i g_i(x)+\sum_{i=1}^p \nu_i h_i(x)$

$\mathcal{L}_{\lambda_i}(x,\lambda,\nu)=g_i(x)$

$\mathcal{L}_{\nu_i}(x,\lambda,\nu)=h_i(x)$

As the partial differential is constant, the partial differential is an affine function.

30.3.2 Primal and dual problems

The primal problem

We already have this.

The dual problem

We can define the Lagrangian dual function:

$g(\lambda,\nu)=\inf_{x\in X}\mathcal{L}(x,\lambda,\nu)$

That is, we have a function which chooses the returns the value of the optimised Lagrangian, given the values of λ and ν.

This is an unconstrained function.

We can prove this function is concave (how?).

The infimum of a set of concave (and therefore also affine) functions is concave.

The supremum of a set of convex (and therefore also affine) functions is convex.

Given a function with inputs x, what values of x maximise the function?

We explore constrained and unconstrained optimisation. The former is where restrictions are placed on vector x, such as a budget constraint in economics.

The dual problem is concave

The duality gap

We refer to the optimal solution for the primary problem as p^*, and the optimal solution for the dual problem as d^*.

The duality gap is p^*-d^*.

30.3.3 Complementary slackness for linear optimisation

30.3.4 Farkas' lemma

We have matrix A and vector b.

Either:

- $Ax = b$; $x \geq 0$
- $A^T y \geq 0$; $b^T y < 0$

30.4 Quadratic optimisation

30.4.1 The quadratic optimisation problem

30.5 Constrainted non-linear optimisation

30.5.1 Weak duality theorem

The duality gap ($p^* - d^*$ is non-negative.

30.5.2 Lagrange multipliers

30.5.3 The dual problem for non-linear optimisation

30.5.4 The weak duality theorem

30.6 Constrained convex optimisation

30.6.1 Slater's condition

Strong duality

Strong duality is where the duality gap is 0.

Slater's condition

Slater's condition says that strong duality holds if there is an input where the inequality constraints are satisified strictly.

That is they are $g(x) < 0$, not $g(x) \leq 0$

This means that the conditions are slack.

This only applies if the problem is convex. That is, if Slater's condition holds, and the problem is convex, then strong duality holds.

30.6.2 The strong duality theorem

30.6.3 Karush-Kuhn-Tucker conditions

If our problem is non-convex, or if Slater's condition does not hold, how else can be find a solution?

A solution, p^* can satisify KKT conditions.

30.7 Sort

30.7.1 Unconstrained envelope theorem

Consider a function which takes two parameters:

$f(x, \alpha$

We want to choose x to maximise f, given α.

$V(\alpha) = \sup_{x \in X} f(x, \alpha)$

There is a subset of X where $f(x, \alpha) = V(\alpha)$.

$X^*(\alpha) = \{x \in X | f(x, \alpha) = V(\alpha)\}$

This means that $V(\alpha) = f(x^*, \alpha)$ for $x^* \in X^*$.

Let's assume that there is only one x^*.

$V(\alpha) = f(x^*, \alpha)$

What happens to the value function as we relax α?

$V_{\alpha_i}(\alpha) = f_{\alpha_i}(x^*(\alpha), \alpha)$.

$V_{\alpha_i}(\alpha) = f_x \dfrac{\delta x^*}{\delta \alpha} + f_{\alpha_i}$.

We know that $f_x = 0$ from first order conditions. So:

$V_{\alpha_i}(\alpha) = f_{\alpha_i}$.

That is, at the optimum, as the constant is relaxed, we can treat the x^* as fixed, as the first-order movement is 0.

30.7.2 Identifying upper and lower bounds of linear programming

In min/max problem, any feasibly solution is an upper/lower bound.

can we get a bound at the other side? yes, by doing linear combinations of inequalities eg maximise $30x + 100y$ subject to: $4x + 10y <= 40$ $x >= 3$

We can identify a lower bound by inputting something which works, for example $x = 3$ and $y = 0$. This gives us a lower bound of 90.

To get an upper bound we can manipulate the constraints: $40x + 100y <= 400$ $10x >= 30$ And then: $40x + 100y <= 370 + 30$ $40x + 100y <= 370 + 10x$ $30x + 100y <= 370$

So we have an upper bound of 370.

This lower bound is a result of doing linear combinations of the inequalities. For different combinations, we could have a lower lower bound.

This is the dual problem. How do we choose the linear combination of inequalities such that the resulting lower bound is minimised?

30.7.3 Hessian matrix

We can take a function and make a matrix of its second order partial derivatives. This is the Hessian matrix, and it describes the local curvature of the function.

If the function f has n parameters, the Hessian matrix is $n \times n$, and is defined as:

$H_{ij} = \frac{\delta^2 f}{\delta x_i \delta x_j}$

If the function is convex, then the Hessian matrix is positive semi-definite for all points, and vice versa.

If the function is concave, then the Hessian matrix is negative semi-definite for all points, and vice versa.

We can diagnose critical points by evaluating the Hessian matrix at those points.

If it is positive definite, it is a local minimum. If it is negative definite it is a local maximum. If there are both positive and negative eigenvalues it is a saddle point.

Part VII

Multivariate real analysis: Vector fields

Chapter 31

Multivariate differentiation of vector fields, including Jacobians, scalar potential, conservative vector fields, divergence, Solenoidal vector fields, Laplace operator, curl, hodge stars and hodge duals

31.1 Partial differentiation of vector fields

31.1.1 Jacobian matrix

If we have n inputs and m functions such that:

$f_i(\mathbf{x})$

The Jacobian is a matrix where:

$J_{ij} = \frac{\delta f_i}{\delta x_j}$

31.2 Scalar potential

31.2.1 Scalar potential

Given a vector field $\mathbf{F}$ we may be able to identify a scalar field P such that:

$\mathbf{F} = -\nabla P$

31.2.2 Non-uniqueness of scalar potentials

Scalar potentials are not unique.

If P is a scalar potential of $\mathbf{F}$, then so is $P + c$, where c is a constant.

31.2.3 Conservative vector fields

Not all vector fields have scalar potentials. Those that do are conservative.

For example if a vector field is the gradient of a scalar height function, then the height is a scalar potential.

If a vector field is the rotation of water, there will not be a scalar potential.

31.3 Divergence

31.3.1 Divergence

This takes a vector field and produces a scalar field.

It is the dot product of the vector field with the del operator.

$divF = \nabla.F$

Where $\nabla = (\sum_{i=1}^n e_i \frac{\delta}{\delta x_i})$

$divF = \sum_{i=1}^n e_i \frac{\delta F_i}{\delta x_i}$

31.3.2 Divergence as net flow

Divergence can be thought of as the net flow into a point.

For example, if we have a body of water, and a vector field as the velocity at any given point, then the divergence is 0 at all points.

This is because water is incompressible, and so there can be no net flows.

Areas which flow out are sources, while areas that flow inwards are sinks.

31.3.3 Solenoidal vector fields

If there is no divergence, then the vector field is called solenoidal.

31.3.4 The Laplace operator

Cross product of divergence with the gradient of the function.

$\Delta f = \nabla . \nabla f$

$\Delta f = \sum_{i=1}^n \frac{\delta^2 f}{\delta x_i^2}$

31.4 Curl

31.4.1 Curl

The curl of a vector field is defined as:

$curl\mathbf{F} = \nabla \times \mathbf{F}$

Where: $\nabla = (\sum_{i=1}^n e_i \frac{\delta}{\delta x_i})$

And: $\mathbf{x} \times \mathbf{y} = \| |\mathbf{x}| | | |\mathbf{y}| | \sin(\theta)\mathbf{n}$

The curl of a vector field is another vector field.

The curl measures the rotation about a given point. For example if a vector field is the gradient of a height map, the curl is 0 at all points, however for a rotating body of water the curl reflects the rotation at a given point.

31.4.2 Divergence of the curl

If we have a vector field $\mathbf{F}$, the divergence of its curl is 0:

$\nabla .(\nabla \times \mathbf{F}) = 0$

31.4.3 Vector potential

Given a vector field $\mathbf{F}$ we may be able to identify another vector field A such that:

$\mathbf{F} = \nabla \times \mathbf{A}$

Existence:

We know that the divergence of the curl for any vector field is 0, so this applies to A:

$\nabla .(\nabla \times \mathbf{A}) = 0$

Therefore:

$\nabla . \mathbf{F} = 0$

This means that if there is a vector potential of $\mathbf{F}$, then $\mathbf{F}$ has no divergence.

31.4.4 Non-uniqueness of vector potentials

Vector potentials are not unique.

If $\mathbf{A}$ is a vector potential of $\mathbf{F}$, then so is $\mathbf{A} + \nabla c$, where c is a scalar field and ∇c is its gradient.

31.4.5 Conservative vector fields

Not all vector fields have scalar potentials. Those that do are conservative.

For example if a vector field is the gradient of a scalar height function, then the height is a scalar potential.

If a vector field is the rotation of water, there will not be a scalar potential.

31.4.6 Hodge stars

The Hodge star operator is a generalisation of cross product. In 3d space if we have a plane, we can get a vector perpendicular and visa versa. Generally, we are in n-dimensional space and we input k vectors and get out $n - k$ vectors.

31.4.7 Hodge duals

Chapter 32

Multivariate integration of vector fields

32.1 Integration of vector fields

32.1.1 Line integral of vector fields

We may wish to integrate along a curve in a vector field.

We previously showed that we can write a curve as a function on the real line:

$r : [a, b] \to C$

The integral is therefore the sum of the function at all points, with some weighting. We write this:

$\int_C f(r)ds = \lim_{\Delta srightarrow0} \sum_{i=0}^{n} f(r(t_i))\Delta s_i$

In a vector field we use

$\int_C f(r)ds = \int_a^b f(r(t)).r'(t)dt$

32.1.2 Double integral of vector fields

32.1.3 Surface integral for vector fields

32.2 Stoke's theorem

32.2.1 The divergence theorem

32.2.2 Stoke's theorem

Part VIII

Partial differential equations (PDEs)

Chapter 33

Partial Differential Equations (PDEs)

33.1 Introduction

Part IX

Variational calculus / calculus of variations / functionals

Chapter 34

Variational calculus/functionals

34.1 Introduction

34.1.1 Introduction