

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

Mathematics Review

Dustin Pluta

Graduate Statistics Bootcamp
Department of Statistics
University of California, Irvine

2019

Goals for Lecture

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

- Familiarize everyone with functions
- Review differentiation and integration
- Review basic optimization of functions
- References: Abbott [2001], Casella and Berger [2002]

Why is Calculus Important?

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

- 1 **Derivatives** for optimizing functions, e.g. finding the maximum likelihood estimator.
- 2 **Integrals** for calculating probabilities.
- 3 **Sequences and limits** for asymptotic theory, e.g. law of large numbers and central limit theorem.
- 4 **Taylor series** for many, many things. When in doubt, try a Taylor expansion.

Defining Functions

- In the first lecture, we used functions without properly defining them.
- Here we dive deeper into some of the mathematical definitions needed for statistics

Definition (Function)

Given two sets A and B , a function from A to B is a rule or mapping that takes element $x \in A$ and associates it with a single element of B . We write

$$f : A \longrightarrow B$$

Given an element $x \in A$, the expression $f(x)$ is used to represent the element of B associated with x by the function f .

Properties of Functions

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

- We now reintroduce some familiar definitions that are useful throughout statistics.
- Consider a function $f : A \rightarrow B$.

Definition (Domain of a Function)

The set A is called the domain of f .

Definition (Range of a Function)

The range is the subset of B given by

$$\{y \in B : y = f(x) \text{ for some } x \in A\}$$

Examples of Parameterized Functions I

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

- Consider the following parameterized function
 $f : \{0, 1\} \rightarrow (0, 1)$

$$f(x|p) = \begin{cases} p & \text{if } x = 1 \\ 1 - p & \text{if } x = 0 \end{cases}$$

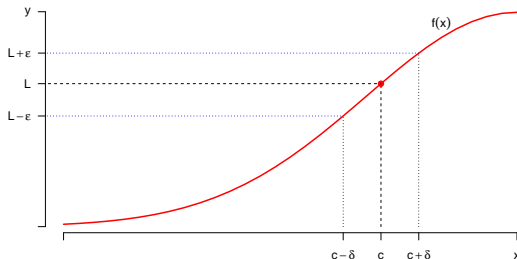
- The domain is the set $A = \{0, 1\}$ and the range is $\{p, 1 - p\}$.
- Note:* When we parameterize a function, we imply that the function is dependent on values of the parameter, which we often “know” beforehand. When we know the value of the parameter θ , then we say that the value is “given” and use the notation $f(\cdot|\theta)$.

Defining the Limit of a Function

Definition

Let $f : A \rightarrow \mathcal{R}$, and let c be a limit point of the domain A . We say that $\lim_{x \rightarrow c} f(x) = L$ provided that for all $\epsilon > 0$, there exists a $\delta > 0$ such that whenever $0 < |x - c| < \delta$ (and $x \in A$) it follows that $|f(x) - L| < \epsilon$.

Limit of a Function



Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

Defining The Exponential Function

- We can use limits to define certain functions, or show convergence of functions throughout statistics.
- One useful relationship is the definition of the exponential function

Definition (exponential function)

$$\exp(a) = e^a = \lim_{n \rightarrow \infty} \left(1 + \frac{a}{n}\right)^n$$

- *(Note: this limit may serve you well through graduate statistics...)*

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

Overview of Differentiation

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

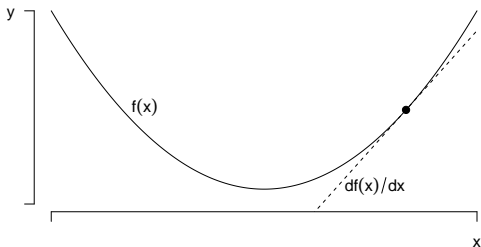
References

- Derivatives play a major role throughout statistics.
- A major area that they become useful is for maximum likelihood estimation.
- Most of frequentist statistics relies on taking derivatives (Score Functions, Information Matrix, etc.)

Intuition of a Derivative

- The geometric interpretation of the derivative of a function f at a point x is that it is the slope of the function at the point x .
- If $f(x)$ is the value of the function evaluated at point x , then the slope at the point x is denoted $f'(x)$.

Derivative of a Function



Defining the Derivative

- We now provide the formal definition of a derivative

Definition (Derivative)

Let $g : A \rightarrow \mathcal{R}$ be a function defined on an interval A . Given a point $c \in A$, the derivative of g at c is defined by

$$g'(c) = \lim_{x \rightarrow c} \frac{g(x) - g(c)}{x - c}$$

provided that the limit exists. If g' exists for all points $c \in A$, we say that g is differentiable on A .

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

Derivative Rules

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

i) $\frac{d}{dx}c = 0$ for any constant c

ii) $\frac{d}{dx}x^n = nx^{n-1}$

iii) $\frac{d}{dx}(f + g)(x) = f'(x) + g'(x)$

iv) $(cf)'(x) = cf'(x)$

v) $(fg)'(x) = f(x)g'(x) + g(x)f'(x)$

vi) $\left(\frac{f}{g}\right)'(x) = \frac{d}{dx}\frac{f(x)}{g(x)} = \frac{g(x)f'(x) - f(x)g'(x)}{(g(x))^2}$, provided that $g(x) \neq 0$ for all $x \in A$.

vii) $f'(x) = \frac{d}{dx}\log(x) = \frac{1}{x}$

viii) $f'(x) = \frac{d}{dx}e^x = e^x$

Derivative Rules

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

- **Chain rule:** If g is differentiable at x and f is differentiable at $g(x)$, then the composite function $(f \circ g)(x) = f(g(x))$ is differentiable at x and is given by the product

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

In alternate notation, if $y = f(\theta)$ and $\theta = g(x)$ are both differentiable functions, then

$$\frac{dy}{dx} = \frac{dy}{d\theta} \frac{d\theta}{dx}$$

Example - Chain Rule in Action I

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

Consider the following parameterized function

$$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$$

and we would like to find the derivative with respect to the variable x .

Example - Chain Rule in Action II

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

We could think of the function in the following way.

$$\begin{aligned}f(g(x)) &= \text{constant} \times \exp \{g(x)\} \\g(x) &= -\frac{(x - \mu)^2}{2\sigma^2}\end{aligned}$$

Example - Chain Rule in Action III

From the above rules, we have that

$$\begin{aligned}f'(g(x)) &= \text{constant} \times \exp\{g(x)\} \\g'(x) &= -\frac{(x - \mu)}{\sigma^2}\end{aligned}$$

Thus by the chain rule,

$$\begin{aligned}f'(x|\mu, \sigma^2) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x - \mu)^2}{2\sigma^2}\right\} \times \frac{-(x - \mu)}{\sigma^2} \\&= -\frac{(x - \mu)}{\sigma^2\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x - \mu)^2}{2\sigma^2}\right\}\end{aligned}$$

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

Multiple Differentiation Example I

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

- As the year progresses we'll be interested in finding derivatives of functions with many parameters
- Specifically, as will be defined later this year the score function and the Fisher Information matrix are defined in terms of derivatives.
- We'll review the notation for derivatives of multiple variables.

Multiple Differentiation Example II

- From Ferguson's 'A Course in Large Sample Theory'
- Consider $f : \mathcal{R}^d \rightarrow \mathcal{R}$, the gradient of f is the length d row vector

$$\nabla f(\mathbf{x}) = \left(\frac{\partial}{\partial x_1} f(\mathbf{x}), \frac{\partial}{\partial x_2} f(\mathbf{x}), \dots, \frac{\partial}{\partial x_d} f(\mathbf{x}) \right)$$

- The Hessian of $f : \mathcal{R}^d \rightarrow \mathcal{R}$ is the $d \times d$ matrix

$$\nabla^2 f(\mathbf{x}) = \begin{pmatrix} \frac{\partial^2}{\partial x_1 \partial x_1} f(\mathbf{x}) & \frac{\partial^2}{\partial x_1 \partial x_2} f(\mathbf{x}) & \dots & \frac{\partial^2}{\partial x_1 \partial x_d} f(\mathbf{x}) \\ \vdots & \ddots & & \vdots \\ \frac{\partial^2}{\partial x_d \partial x_1} f(\mathbf{x}) & \frac{\partial^2}{\partial x_d \partial x_2} f(\mathbf{x}) & \dots & \frac{\partial^2}{\partial x_d \partial x_d} f(\mathbf{x}) \end{pmatrix}$$

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

Multiple Differentiation Example III

- Consider $f(x, y) = x^2 + y^2 - 2xy$. The gradient of the function is the vector

$$\begin{aligned}\nabla f(x, y) &= \left(\frac{d}{dx} f(x, y), \frac{d}{dy} f(x, y) \right) \\ &= (2x - 2y, 2y - 2x)\end{aligned}$$

- The Hessian is the matrix

$$\nabla^2 f(x, y) = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$$

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

Why Optimization?

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

- We now talk about finding extreme values of a function.
- This will become very important for maximum likelihood, where we want to find the value that maximizes our likelihood function.
- Additionally, this can be used to find the 'mode' (or maximum) of probability density functions.

Defining Extreme Values

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

Definition (maximum value)

A function f has an absolute maximum (or global maximum) at a point c if $f(c) \geq f(x)$ for all x in the domain of f . The value of $f(c)$ is called the maximum value.

Definition (minimum value)

A function f has an absolute minimum (or global minimum) at a point c if $f(c) \leq f(x)$ for all x in the domain of f . The value of $f(c)$ is called the minimum value.

Definition (extreme values)

The minimum and maximum values are the extreme values of the function.

Interior Extremum Theorem

Theorem (Interior Extremum Theorem)

Let f be differentiable on an open interval (a, b) . If f attains a maximum value at some point $c \in (a, b)$.

(i.e. $f(c) \geq f(x)$ for all $x \in (a, b)$), then $f'(c) = 0$. The same is true if $f(c)$ is a minimum value.

- The Interior Extremum Theorem is one of the fundamental theorems that enable the use of the derivative as a tool for solving applied optimization problems [Abbott, 2001].
- Always remember to consider the domain of the function as well when finding extreme values.

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

Example - Optimizing a Function I

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

Consider the following parameterized function

$$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(x - \mu)^2}{2\sigma^2} \right\}$$

Example - Optimizing a Function II

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

We saw earlier that taking the derivative of this function gave us the following derivative,

$$\begin{aligned} f'(x|\mu, \sigma^2) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{\frac{(x-\mu)^2}{2\sigma^2}\right\} \times \frac{-(x-\mu)}{\sigma^2} \\ &= -\frac{(x-\mu)}{\sigma^2\sqrt{2\pi\sigma^2}} \exp\left\{\frac{(x-\mu)^2}{2\sigma^2}\right\} \end{aligned}$$

This looks like a mess to find $f'(x|\mu, \sigma^2) = 0$.

Example - Optimizing a Function III

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

- Alternatively, we can use a strategy which will help us to find derivatives of functions.
- We appeal to monotonic transformations of functions which allow us to preserve the order of the set.
- That is if $g(\cdot)$ is a monotonic function, then if $f(c)$ is an extreme value of the function f , it is also the extreme value off the function g .
- The main monotonic transformation that we will use is a logarithmic transformation.

Example - Optimizing a Function IV

Applying this

$$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(x - \mu)^2}{2\sigma^2} \right\}$$

$$\log f(x|\mu, \sigma^2) = -\log(\sqrt{2\pi\sigma^2}) - \frac{(x - \mu)^2}{2\sigma^2}$$

$$\frac{d}{dx} \log f(x|\mu, \sigma^2) = -\frac{x - \mu}{\sigma^2}$$

If we set this equal to zero, it will be MUCH easier to solve than the previous function. Thus

$$-\frac{x - \mu}{\sigma^2} = 0$$

$$x - \mu = 0$$

$$x_{\text{extreme}} = \mu$$

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

Example - Optimizing a Function V

Now we must prove that it is a maximum or minimum! We can use the second derivative.

$$\begin{aligned}\log f(x|\mu, \sigma^2) &= -\log\left(\sqrt{2\pi\sigma^2}\right) - \frac{(x - \mu)^2}{2\sigma^2} \\ \frac{d}{dx} \log f(x|\mu, \sigma^2) &= -\frac{x - \mu}{\sigma^2} \\ \frac{d^2}{dx^2} \log f(x|\mu, \sigma^2) &= -\frac{1}{\sigma^2}\end{aligned}$$

We see that this is **negative** and therefore a maximum of the function. Thus,

$$x_{max} = \mu$$

Numerical Optimization: Newton-Raphson I

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

- Consider a function f where for some x , $f(x) = 0$, such an x is a root of the function f .
- We can use numerical methods to find such a point, especially if finding such points are difficult.
- Consider the Taylor series expansion of a function around the point x_0

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n \\ &= f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!} (x - x_0)^2 + \dots \end{aligned}$$

Numerical Optimization: Newton-Raphson II

- This implies that if x and x_0 are close then

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0)$$

- Further if $f'(x_0)$ is invertible, with $f(x) = 0$ for finding the root, we have that

$$0 \approx f(x_0) + f'(x_0)(x - x_0)$$

- Therefore

$$\begin{aligned} -f(x_0) &= f'(x_0)(x - x_0) \\ -(f'(x_0))^{-1}f(x_0) &= x - x_0 \\ x &= x_0 - (f'(x_0))^{-1}f(x_0) \end{aligned}$$

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

Numerical Optimization: Newton-Raphson III

- Now since

$$x = x_0 - (f'(x_0))^{-1}f(x_0)$$

we can use this to create an algorithm to find roots of the function

- Therefore

$$x_{n+1} = x_n - (f'(x_n))^{-1}f(x_n)$$

- Now since a function will have an extremum when $f'(x) = 0$, we can use this algorithm to find extreme values of functions!

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

Numerical Optimization: Multivariate Newton-Raphson

When optimizing a multivariate function $f : \mathbb{R} \rightarrow \mathbb{R}^d$ (for which second order partial derivatives exist), we can use a multivariate form of the Newton-Raphson method with update step:

$$\mathbf{x}_{n+1} = \mathbf{x}_n - [\nabla^2 f(\mathbf{x}_n)]^{-1} \nabla f(\mathbf{x}_n)$$

- Newton-Raphson is nice because it does not require selecting a step-size or other tuning.
- However, it does require inverting the Hessian matrix, which can be expensive in high-dimensions.
- This computational expense has motivated the development of many other numerical optimization methods, such as gradient descent and stochastic gradient descent.

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

Numerical Optimization: Gradient Descent

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

- Suppose $f : \mathbb{R} \rightarrow \mathbb{R}^d$ and that its first order partial derivatives exist.
- Gradient descent uses an iterative update step similar to Newton-Raphson, but which does not require computing or inverting the Hessian.
- The trade off for this reduction in computation is the need to choose a step size γ .

The gradient descent update step is

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \gamma_n \nabla f(\mathbf{x}_n)$$

Stochastic Gradient Descent

- Often we wish to optimize a function of the form

$$f(\mathbf{x}) = \frac{1}{m} \sum_{i=1}^m f_i(\mathbf{x}).$$

- When n is very large, even gradient descent can be computationally impractical. In this situation, stochastic gradient descent (SGD) is a useful alternative method.
- SGD finds the optimum of f iteratively as in gradient descent, but uses a random sample of the functions f_i when computing the gradient, so as to save on computation.

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

Stochastic Gradient Descent

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

- A simple implementation of SGD is
 - 1 Choose an initial value \mathbf{x}_0 and a learning rate γ .
 - 2 Until convergence, repeat:
 - 1 Select a random subset of $k < m$ functions f_i , and let A denote the corresponding set of indices.
 - 2 Update

$$\mathbf{x}_n = \mathbf{x}_n - \gamma \sum_{i \in A} f_i(\mathbf{x}_n)$$

Why Integration?

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

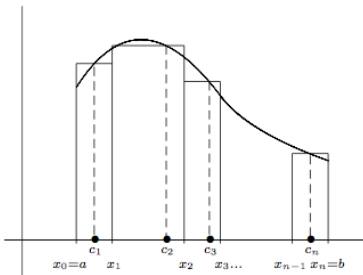
References

References

- Integration additionally plays an important role in statistics.
- While, the derivative is the slope of a function a point and often represents the rate of change of the function. The integral represents the area under the curve
- Properties of the random variables are defined based on integration, such as the cumulative distribution function
- Additionally, most of Bayesian statistics relies on integration to find estimates for parameters.

Intuition of the Integral

- Integration and 'area' go hand in hand. The integral of a function allows you to find the total area under the curve.



- You'll remember from your calculus class the idea of Riemann sums.
- Those of you in the Ph.D. program will further these ideas with measure theoretic integration.

Fundamental Theorem of Calculus

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

- We'll state the fundamental theorem of calculus as a reminder and then provide some common integration strategies that you learned in your undergraduate work.
- As a reminder the Fundamental Theorem of Calculus allows us to relate integration and differentiation.
- Again this isn't everything you'll need to get through the year, just a reminder of the mathematical tricks that you've learned along the way.

Fundamental Theorem of Calculus

Theorem (Fundamental Theorem of Calculus)

i) If $f : [a, b] \rightarrow \mathcal{R}$ is integrable, and $F : [a, b] \rightarrow \mathcal{R}$ satisfies $F'(x) = f(x)$ for all $x \in [a, b]$, then

$$\int_a^b f = F(b) - F(a)$$

ii) Let $g : [a, b] \rightarrow \mathcal{R}$ be integrable and define

$$G(x) = \int_a^x g$$

for all $x \in [a, b]$. Then G is continuous on $[a, b]$. If g is continuous at some point $c \in [a, b]$ then G is differentiable at c and $G'(c) = g(c)$.

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

Common Integration Rules I

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization**Integration**

Sequences

Series

Useful Series
and Partial
Sums

References

References

Substitution Rule

If $u = g(x)$ is a differentiable function whose range is an interval I and f is continuous on I , then

$$\int f(g(x))g'(x)dx = \int f(u)du$$

For definite integrals, if g' is continuous on $[a, b]$ and f is continuous on the range of $u = g(x)$, then

$$\int_a^b f(g(x))g'(x)dx = \int_{g(a)}^{g(b)} f(u)du$$

Common Integration Rules II

Integration by Parts

- Based upon the product rule of differentiation we can create a similar rule for integration, namely integration by parts.
- This rule allows us to calculate the integral of the product of two functions and is particularly useful when one of the functions is easy to integrate.

By the product rule of differentiation,

$$\frac{d}{dx} [f(x)g(x)] = f(x)g'(x) + g(x)f'(x)$$

Which in term implies that

$$\int [[f(x)g'(x) + g(x)f'(x)] dx = f(x)g(x)$$

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization**Integration**

Sequences

Series

Useful Series
and Partial
Sums

References

References

Common Integration Rules III

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

Rearranging

$$\int f(x)g'(x)dx = f(x)g(x) - \int g(x)f'(x)dx$$

For definitely integrals this becomes,

$$\int_a^b f(x)g'(x)dx = f(x)g(x) \Big|_a^b - \int_a^b g(x)f'(x)dx$$

Example - Integration by Parts I

- Consider the function $h(x) = xe^x$ which appears to be unruly to integrate at first glance.
- Using integration by parts, we can actually solve this:
Consider,

$$h(x) = f(x)g'(x) = xe^x$$

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

Recall that $\int e^x dx = e^x$ and $f'(x) = 1$, thus using integration by parts

$$\begin{aligned} \int xe^x dx &= xe^x - \int e^x dx \\ &= xe^x - e^x + C \\ &= (x - 1)e^x + C \end{aligned}$$

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization**Integration**

Sequences

Series

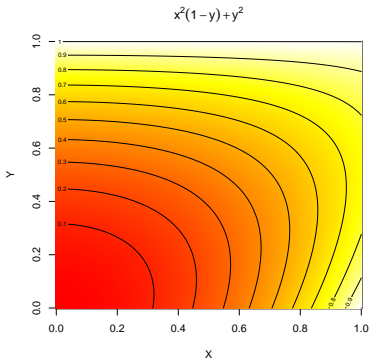
Useful Series
and Partial
Sums

References

References

Multiple Integration Example I

- Consider the function $f(x, y) = x^2(1 - y) + y^2$ for $x \in (0, 1)$ and $y \in (0, 1)$.
- We want to find the area under the surface created by the function.



Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

Multiple Integration Example II

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

$$\begin{aligned} & \int_0^1 \int_0^1 x^2(1-y) + y^2 dy dx \\ &= \int_0^1 x^2 \left(y - \frac{y^2}{2} \right) + \frac{y^3}{3} \Big|_0^1 dx \\ &= \int_0^1 \frac{x^2}{2} + \frac{1}{3} dx \\ &= \frac{x^3}{6} + \frac{1}{3}x \Big|_0^1 = \frac{1}{2} \end{aligned}$$

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

Background Material

Defining a Sequence

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

Definition (sequence)

A sequence is a function whose domain is the natural numbers $\mathbf{N} = \{1, 2, 3, 4, 5, \dots\}$.

- Given a function $f : \mathbf{N} \rightarrow \mathcal{R}$, $f(n)$ is the n^{th} term of the list. We note that \mathcal{R} is the real numbers.

Examples of Sequences

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

- Below are some examples of sequences

i) $(1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \dots)$

ii) $(\frac{n+1}{n})_{n=1}^{\infty} = (\frac{2}{1}, \frac{3}{2}, \frac{4}{3}, \frac{5}{4}, \dots)$

iii) (a_n) , where $a_n = 2^n$ for $n \in \mathbf{N}$

iv) (x_n) , where $x_1 = 3$ and $x_{n+1} = \frac{x_n+1}{2}$

Convergence of a Sequence

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

Definition (Convergence of a Sequence)

A sequence (x_n) converges to a number c if for every positive number ϵ , there exists an $N \in \mathbf{N}$ such that whenever $n \geq N$ it follows that $|x_n - c| < \epsilon$

- The general idea is that a “challenger” chooses some value ϵ , then you must produce a value N such that for all numbers $n \geq N$ it follows that $|x_n - c| < \epsilon$

Definition (Divergent Sequence)

A sequence that does not converge is said to diverge.

Example of Showing Convergence I

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

Claim:

$$\lim_{n \rightarrow \infty} \left(\frac{n+1}{n} \right) = 1$$

Our sequence here is $\left(\frac{n+1}{n} \right)_{n=1}^{\infty} = \left(\frac{2}{1}, \frac{3}{2}, \frac{4}{3}, \frac{5}{4}, \dots \right)$, where it appears that for large numbers the numerator and the denominator will be close enough that the value is almost 1. Our goal is find N so that for any ϵ that someone chooses we can satisfy the requirements of the definition for convergence.

Example of Showing Convergence II

Thus we want

$$\left| \frac{n+1}{n} - 1 \right| < \epsilon$$

$$\left| \frac{n+1}{n} - \frac{n}{n} \right| < \epsilon$$

$$\left| \frac{n+1-n}{n} \right| < \epsilon$$

$$\left| \frac{1}{n} \right| < \epsilon$$

Now since n is always positive it follows that $1/n < \epsilon$ and $n > 1/\epsilon$. Therefore we can pick N larger than $1/\epsilon$ and then every $n \geq N$ will satisfy the definition. See Abbott page 42 for a formal proof.

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

Infinite Series I

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

Definition (Infinite Series)

Let (b_n) be a sequence. An infinite series is a formal expression of the form

$$\sum_{n=1}^{\infty} b_n = b_1 + b_2 + b_3 + \dots$$

Definition (Partial Sums)

Define the sequence of partial sums (s_m) as

$$s_m = b_1 + b_2 + \dots + b_{m-1} + b_m$$

Definition (Convergence of Series)

Infinite Series II

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

We say that the series $\sum_{n=1}^{\infty} b_n$ converges to B if the sequence (s_m) converges to B . We write $\sum_{n=1}^{\infty} b_n = B$

Partial Sum of Integers I

- Let $(a_n) = n$ with $n \in \mathbf{N}$. Then the sequence of partial sums b_m is

$$b_m = \sum_{n=1}^m n = 1 + 2 + \cdots + (m-1) + m = \frac{m(m+1)}{2}$$

Note that this does not converge.

- To show this consider that

$$\begin{aligned} b_m &= 1 + 2 + \cdots + (m-1) + m \\ &= m + (m-1) + \cdots + 2 + 1 \end{aligned}$$

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

Partial Sum of Integers II

- Therefore if we consider $2b_m$, adding the top line and the bottom line of the above we get

$$2b_m = (m + 1) + (m + 1) + \cdots + (m + 1) = m(m + 1)$$

- Thus,

$$b_m = \frac{m(m + 1)}{2}$$

- Additionally, the partial sums of squared integers is

$$\sum_{n=1}^m n^2 = \frac{m(m + 1)(2m + 1)}{6}$$

we leave this for you to figure out.

Harmonic Series

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

- Define the harmonic series as $\sum_{n=1}^{\infty} \frac{1}{n}$ with $n \in \mathbf{N}$.
- The sequence of partial sums b_m for the harmonic series is

$$b_m = \sum_{n=1}^m \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{m}$$

- Note that this does not converge nor does it have a closed form solution.

Geometric Series

- The geometric series can be written

$$\sum_{n=1}^{\infty} ar^{n-1}$$

- The partial sum of the geometric series is

$$\sum_{n=1}^m ar^{n-1} = \frac{1 - r^m}{1 - r}$$

therefore the series converges to

$$\sum_{n=1}^{\infty} ar^{n-1} = \frac{a}{1 - r}$$

when $|r| < 1$, otherwise it diverges.

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

Taylor Series I

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

- Another important approximation in mathematics is approximating functions by their Taylor series expansion.
- Taylor series expansions play an important role in asymptotic statistics.

Taylor Series II

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

Definition (Taylor Series from Wikipedia)

The Taylor series of a real or complex valued function $f(x)$ that is infinitely differentiable at a real or complex value a is the power series

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

where $n!$ is the factorial of n and $f^{(n)}(a)$ is the n^{th} derivative of $f(x)$ evaluated at the point a .

Taylor Series III

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

- One of the beauties of the Taylor series is that it can be used to approximate functions at a point.
- To approximate a function we take the first few terms of the Taylor series and consider the rest as being some 'small' remainder.

References

Functions

Limits of
Functions

Differentiation

Optimization

Numerical
Optimization

Integration

Sequences

Series

Useful Series
and Partial
Sums

References

References

Stephen Abbott. *Understanding analysis*. Springer, 2001.

George Casella and Roger L Berger. *Statistical inference*.
Second edition, 2002.