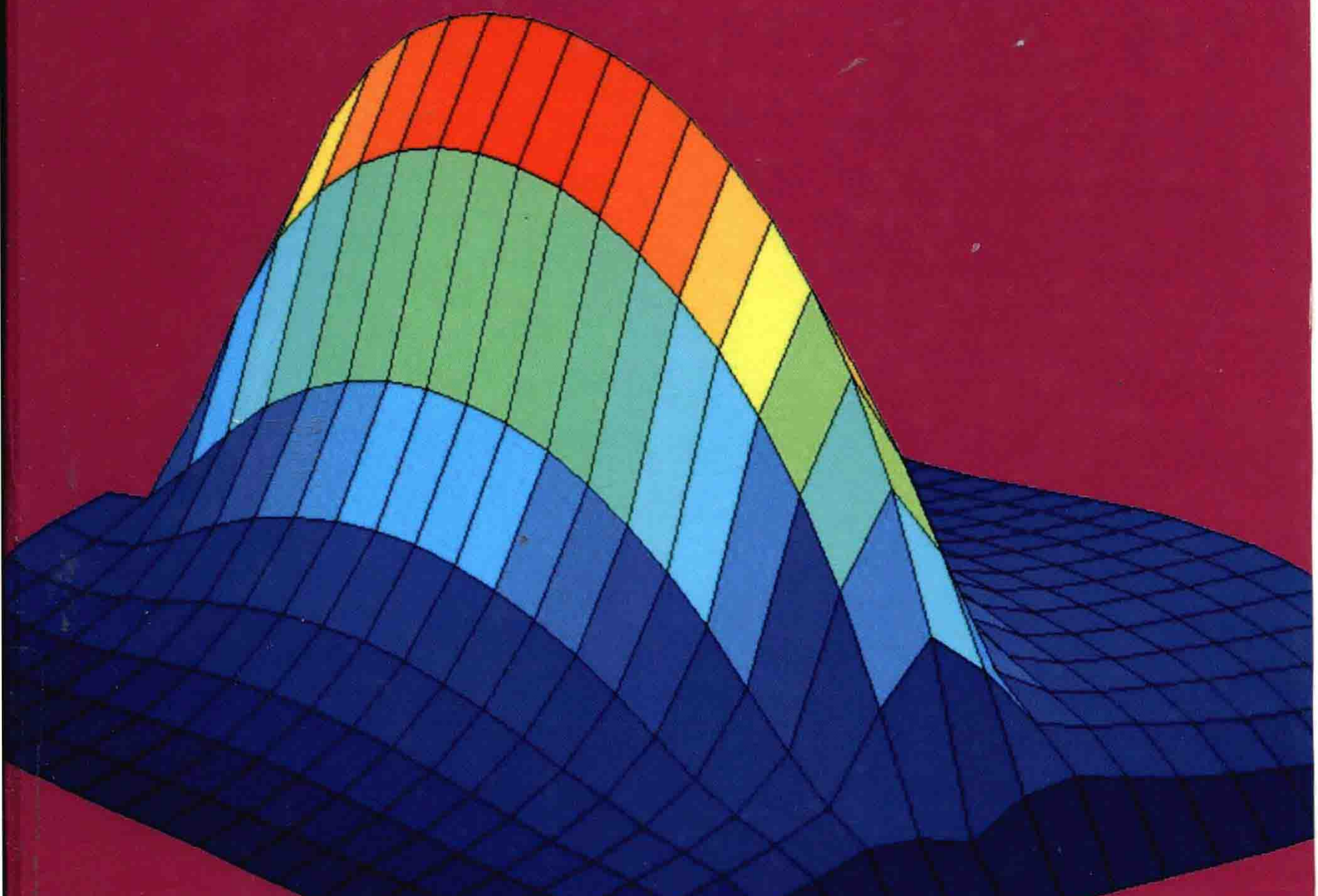
Handbook of Optimisation Techniques in Real Mathematical Analysis



Ashley Brooks
Editor

Handbook of Optimisation Techniques in Real Mathematical Analysis

Ashley Brooks
Editor

AURIS Reference
AURIS REFERENCE LTD.
London, UK

Handbook of Optimisation Techniques in Real Mathematical Analysis

Published by
Auris Reference Ltd., UK
www.aurisreference.com

ISBN: 978-1-78154-470-9

Editor: Ashley Brooks

Printed in UK

10 9 8 7 6 5 4 3 2 1

Cover Design: Cover Lab

For information about Auris Reference Ltd and its publications, visit our website at www.aurisreference.com

Handbook of
Optimisation Techniques in Real Mathematical Analysis

Preface

Real mathematics is the art of creating, understanding, and exploring the relationships between various mathematical structures. Mathematics is the art f making mathematics, by exploring and understanding mathematical structures. Another insightful view put forth which is that real or pure mathematics is not necessarily applied mathematics: it is possible to study abstract entities with respect to their intrinsic nature, and not be concerned with how they manifest in the real world. Even though the pure and applied viewpoints are distinct philosophical positions, in practice there is much overlap in the activity of pure and applied mathematicians. To develop accurate models for describing the real world, many applied mathematicians draw on tools and techniques that are often considered to be "pure" mathematics. On the other hand, many pure mathematicians draw on natural and social phenomena as inspiration for their abstract research.

Analysis is concerned with the properties of functions. It deals with concepts such as continuity, limits, differentiation and integration, thus providing a rigorous foundation for the calculus of infinitesimals introduced by Newton and Leibniz in the 17th century. Real analysis studies functions of real numbers, while complex analysis extends the aforementioned concepts to functions of complex numbers. Functional analysis is a branch of analysis that studies infinite-dimensional vector spaces and views functions as points in these spaces. In mathematics, computer science, or management science, mathematical optimization is the selection of a best element from some set of available alternatives. In the simplest case, an optimization problem consists of maximizing or minimizing a real function by systematically choosing input values from within an allowed set and computing the value of the function. The generalization of optimization theory and techniques to other formulations comprises a large area of applied mathematics. More generally, optimization includes finding "best available" values of some objective function given a defined domain, including a variety of different types of objective functions and different types of domains. One major criterion for optimizers is just the number of required function evaluations as this often is already a large computational effort, usually much more effort than within the optimizer itself, which mainly has to operate over the N variables. The derivatives provide detailed information for such optimizers, but are even harder to calculate, e.g. approximating the gradient takes at least N+1 function evaluations. For approximations of the 2nd derivatives the number of function evaluations is in the order of N^2. Newton's method requires the 2nd order derivates, so for each iteration the number of function calls is in the order of N^2, but for a simpler pure gradient optimizer it is only N. However, gradient optimizers need usually more iterations than Newton's algorithm. Which one is best with respect to the number of function calls depends on the problem itself. To solve problems, researchers may use algorithms that terminate in a finite number of steps, or iterative methods that converge to a solution, or heuristics that may provide approximate solutions to some problems.

This book is intended first and foremost for students wishing to deepen their knowledge of mathematical analysis, and for lecturers conducting seminars in university mathematics departments.

—*Editor*

Contents

Chapter 1

Applications of Measure Theory

Measurable Cardinal

In mathematics, a measurable cardinal is a certain kind of large cardinal number.

Measurable

Formally, a measurable cardinal is an uncountable cardinal number κ such that there exists a κ-additive, non-trivial, 0-1-valued measure on the power set of κ. (Here the term κ-additive means that, for any sequence A_α, $\alpha<\lambda$ of cardinality $\lambda<\kappa$, A_α being pairwise disjoint sets of ordinals less than κ, the measure of the union of the A_α equals the sum of the measures of the individual A_α.) Equivalently, κ is measurable means that it is the critical point of a non-trivial elementary embedding of the universe V into a transitive class M. This equivalence is due to Jerome Keisler and Dana Scott, and uses the ultrapower construction from model theory. Since V is a proper class, a small technical problem that is not usually present when considering ultrapowers needs to be addressed, by what is now called Scott's trick.

Equivalently, κ is a measurable cardinal if and only if it is an uncountable cardinal with a κ-complete, non-principal ultrafilter. Again, this means that the intersection of any strictly less than κ-many sets in the ultrafilter, is also in the ultrafilter. Although it follows from ZFC that every measurable cardinal is inaccessible (and is ineffable, Ramsey, etc.), it is consistent with ZF that a measurable cardinal can be a successor cardinal. It follows from ZF + axiom of determinacy that ω_1 is measurable, and that every subset of ω_1 contains or is disjoint from a closed and unbounded subset.

The concept of a measurable cardinal was introduced by Stanis law Ulam (1930), who showed that the smallest cardinal κ that admits a non-trivial countably-additive two-valued measure must in fact admit a κ-additive measure. (If there were some collection of fewer than κ measure-0 subsets whose union was κ, then the induced measure on this collection would be a counterexample to the minimality of κ.) From there, one can prove (with the Axiom of Choice) that the least such cardinal must be inaccessible.

It is trivial to note that if κ admits a non-trivial κ-additive measure, then κ must be regular. (By non-triviality and κ-additivity, any subset of cardinality less than κ must have measure 0, and then by κ-additivity again, this means that the entire set must not be a union of fewer than κ sets of cardinality less than κ.) Finally, if $\lambda < \kappa$, then it can't be the case that $\kappa \le 2^{\lambda}$. If this were the case, then we could identify κ with some collection of 0-1 sequences of length λ. For each position in the sequence, either the subset of sequences with 1 in that position or the subset with 0 in that position would have to have measure 1. The intersection of these λ-many measure 1 subsets would thus also have to have measure 1, but it would contain exactly one sequence, which would contradict the non-triviality of the measure. Thus, assuming the Axiom of Choice, we can infer that κ is a strong limit cardinal, which completes the proof of its inaccessibility.

If κ is measurable and $p \in V_{\kappa}$ and M (the ultrapower of V) satisfies $\psi(\kappa,p)$, then the set of $\alpha<\kappa$ such that V satisfies $\psi(\alpha,p)$ is stationary in κ (actually a set of measure 1). In particular if ψ is a $Đ_1$ formula and V satisfies $\psi(\kappa,p)$, then M satisfies it and thus V satisfies $\psi(\alpha,p)$ for a stationary set of $\alpha<\kappa$. This property can be used to show that κ is a limit of most types of large cardinals which are weaker than measurable. Notice that the ultrafilter or measure which witnesses that κ is measurable cannot be in M since the smallest such measurable cardinal would have to have another such below it which is impossible.

Every measurable cardinal κ is a 0-huge cardinal because $^{\kappa}M$,IM, that is, every function from κ to M is in M. Consequently, $V_{\kappa+1}$,IM.

Real-valued Measurable

A cardinal κ is called real-valued measurable if there is an atomless κ-additive measure on the power set of κ. They were introduced by Stefan Banach (1930). Banach & Kuratowski (1929) showed that the continuum hypothesis implies that $\mathfrak{c}$ is not real-valued measurable. A real valued measurable cardinal less than or equal to $\mathfrak{c}$ exists if there is a countably additive extension of the Lebesgue measure to all sets of real numbers. A real valued measurable cardinal is weakly Mahlo.

Solovay (1971) showed that existence of measurable cardinals in ZFC, real valued measurable cardinals in ZFC, and measurable cardinals in ZF, are equiconsistent.

Normal Measure

In set theory, a normal measure is a measure on a measurable cardinal κ such that the equivalence class of the identity function on κ maps to κ itself in the ultrapower construction. Equivalently, if $f:\kappa\rightarrow\kappa$ is such that $f(\alpha)<\alpha$ for most $\alpha<\kappa$, then there is a $\beta<\kappa$ such that $f(\alpha)=\beta$ for most $\alpha<\kappa$. (Here, "most" means that the set of elements of κ where the property holds is a member of the ultrafilter, i.e. has measure 1.) Also equivalent, the ultrafilter (set of sets of measure 1) is closed under diagonal intersection.

For a normal measure, any closed unbounded (club) subset of κ contains most ordinals less than κ. And any subset containing most ordinals less than κ is stationary in κ. If an uncountable cardinal κ has a measure on it, then it has a normal measure on it.

Mitchell Order

In mathematical set theory, the Mitchell order is a well-founded preorder on the set of normal measures on a measurable cardinal κ. It is named for William Mitchell. We say that M ∑% N (this is astrict order) if M is in the ultrapower model defined by N. Intuitively, this means that M is a weaker measure than N (note, for example, that κ will still be measurable in the ultrapower for N, since M is a measure on it).

In fact, the Mitchell order can be defined on the set (or proper class, as the case may be) of extenders for κ; but if it is so defined it may fail to be transitive, or even well-founded, provided κ has sufficiently strong large cardinal properties. Well-foundedness fails specifically for rank-into-rank extenders; but Itay Neeman showed in 2004 that it holds for all weaker types of extender.

The Mitchell rank of a measure is the order type of its predecessors under ∑%; since ∑% is well-founded this is always an ordinal.

A cardinal which has measures of Mitchell rank α for each $\alpha < \beta$ is said to be β-measurable.

Valuation (Measure Theory)

In measure theory or at least in the approach to it through domain theory, a valuation is a map from the class of open sets of a topological space to the set positive real numbers including infinity. It is a concept closely related to that of a measure and as such it finds applications measure theory, probability theory and also in theoretical computer science.

Domain/Measure Theory Definition

Let (X,τ) be a topological space: a valuation is any map

$$v:\mathcal{T}\to\mathbb{R}^{+}\cup\{+\infty\}$$

satisfying the following three properties

$v(\varnothing)=0$		Strictness property
$v(U)\le v(V)$	if $U\subseteq V \quad U,V\in\mathcal{T}$	Monotonicity property
$v(U\cup V)+v(U\cap V)=v(U)+v(V)$	$\forall U,V\in\mathcal{T}$	Modularity property

The definition immediately shows the relationship between a valuation and a measure: the properties of the two mathematical object are often very similar if not identical, the only difference being that the domain of a measure is the Borel algebra of the given topological space, while the domain of a valuation is the class of open sets. Further details and references can be found in Alvarez-Manilla et al. 2000 and Goulbault-Larrecq 2002.

Continuous Valuation

A valuation (as defined in domain theory/measure theory) is said to be continuous if for *every directed family* $\{U_i\}_{i\in I}$ *of open sets* (i.e. an indexed family of open sets which is also directed in the sense that for each pair of indexes *i* and *j* belonging to the index set *I*, there exists an index *k* such that $U_i\subseteq U_k$ and $U_j\subseteq U_k$) the following equality holds:

$$v\left(\bigcup_{i\in I}U_i\right)=\sup_{i\in I}v(U_i).$$

Simple Valuation

A valuation (as defined in domain theory/measure theory) is said to be simple if it is a finite linear combination with non-negative coefficients of Dirac valuations, i.e.

$$v(U)=\sum_{i=1}^{n}a_i\delta_{x_i}(U)\quad \forall U\in\mathcal{T}$$

where a_i is always greater than or at least equal to zero for all index *i*. Simple valuations are obviously continuous in the above sense. The supremum of a *directed family of simple valuations* (i.e. an indexed family of simple valuations which is also directed in the sense that for each pair of indexes i and *j* belonging to the index set *I*, there exists an index *k* such that $v_i(U) \leq v_k(U)$ and $v_j(U) \subseteq v_k(U)$) is called quasi-simple valuation

$$\overline{v}(U) = \sup_{i \in I} v_i(U) \quad \forall U \in \mathcal{T}.$$

Examples

Dirac Valuation

Let $(X, \mathcal{T})$ be a topological space, and let x be a point of X: the map

$$\delta_x(U) = \begin{cases} 0 & \text{if } x \notin U \\ 1 & \text{if } x \in U \end{cases} \quad \forall U \in \mathcal{T}$$

is a valuation in the domain theory/measure theory, sense called Dirac valuation. This concept bears its origin from distribution theory as it is an obvious transposition to valuation theory of Dirac distribution: as seen above, Dirac valuations are the "bricks" simple valuations are made of.

Filtration (Mathematics)

In mathematics, a filtration is an indexed set S_i of subobjects of a given algebraic structure S, with the index i running over some index set I that is a totally ordered set, subject to the condition that if $i \leq j$ in I then S_i I S_j. The concept dual to a filtration is called a *cofiltration*.

Sometimes, as in a filtered algebra, there is instead the requirement that the S_i besubobjects with respect to certain operations (say, vector addition), but with respect to other operations (say, multiplication), they instead satisfy $S_i \cdot S_j \subset S_{i+j}$, where here the index set is the natural numbers; this is by analogy with a graded algebra.

Sometimes, filtrations are supposed to satisfy the additional requirement that the union of the S_i be the whole S, or (in more general cases, when the notion of union does not make sense) that the canonical homomorphism from the direct limit of the S_i to S is an isomorphism. Whether this requirement is assumed or not usually depends on the author of the text and is often explicitly stated. We are *not* going to impose this requirement in this article.

There is also the notion of a descending filtration, which is required to satisfy $S_i \supseteq S_j$ in lieu of $S_i \subseteq S_j$ (and, occasionally, $\bigcap_{i \in I} S_i = 0$ instead of $\bigcup_{i \in I} S_i = S$). Again, it depends on the context how exactly the word "filtration" is to be understood. Descending filtrations are not to be confused with cofiltrations (which consist of quotient objects rather than subobjects).

Filtrations are widely used in abstract algebra, homological algebra (where they are related in an important way to spectral sequences), and in measure theory and probability theory for nested sequences of σ-algebras.

In functional analysis and numerical analysis, other terminology is usually used, such as scale of spaces or nested spaces.

Examples

Algebra:

Groups: In algebra, filtrations are ordinarily indexed by N, the set of natural numbers. A *filtration* of a group G, is then a nested sequence G_n of normal subgroups of G (that is, for any n we have G_{n+1} I G_n). Note that this use of the word "filtration" corresponds to our "descending filtration".

Given a group G and a filtration G_n, there is a natural way to define a topology on G, said to be associated to the filtration.

A basis for this topology is the set of all translates of subgroups appearing in the filtration, that is, a subset of G is defined to be open if it is a union of sets of the form aG_n, where a"G and n is a natural number.

The topology associated to a filtration on a group G makes G into a topological group.

The topology associated to a filtration G_n on a group G is Hausdorff if and only if $\cap G_n = \{1\}$.

If two filtrations G_n and $G2_n$ are defined on a group G, then the identity map from G to G, where the first copy of G is given the G_n-topology and the second the $G2_n$-topology, is continuous if and only if for any n there is an m such that G_m I$G2_n$, that is, if and only if the identity map is continuous at 1.

In particular, the two filtrations define the same topology if and only if for any subgroup appearing in one there is a smaller or equal one appearing in the other.

Rings and Modules: Descending Filtrations

Given a ring R and an R-module M, a *descending filtration* of M is a decreasing sequence of submodules M_n. This is therefore a special case of the notion for groups, with the additional condition that the subgroups be submodules. The associated topology is defined as for groups.

An important special case is known as the I-adic topology (or J-adic, etc.). Let R be a commutative ring, and I an ideal of R.

Given an R-module M, the sequence $I^n M$ of submodules of M forms a filtration of M. The *I-adic topology*o n M is then the topology associated to this filtration. If M is just the ring R itself, we have defined the *I-adic topology* on R.

When R is given the I-adic topology, R becomes a topological ring. If an R-module M is then given the I-adic topology, it becomes a topological R-module, relative to the topology given on R.

Rings and modules: ascending filtrations

Given a ring R and an R-module M, an *ascending filtration* of M is an increasing sequence of submodules M_n. In particular, if R is a field, then an ascending filtration of the R-vector space M is an increasing sequence of vector subspaces of M. Flags are one important class of such filtrations.

Sets

A maximal filtration of a set is equivalent to an ordering (a permutation) of the set. For instance, the filtration $\{0\} \subset \{0,1\} \subset \{0,1,2\}$ corresponds to the ordering $(0,1,2)$. From the point of view of the field with one element, an ordering on a set corresponds to a maximal flag (a filtration on a vector space), considering a set to be a vector space over the field with one element.

Measure Theory

In measure theory, in particular in martingale theory and the theory of stochastic processes, a filtration is an increasing sequence of σ-algebras on a measurable space. That is, given a measurable space $(\Omega, \mathcal{F})$, a filtration is a sequence of σ-algebras $\{\mathcal{F}_t\}_{t \geq 0}$ with $\mathcal{F}_t \subseteq \mathcal{F}$ for each t and

$$t_1 \leq t_2 \Rightarrow \mathcal{F}_{t_1} \subseteq \mathcal{F}_{t_2}.$$

The exact range of the "times" t will usually depend on context: the set of values for t might be discrete or continuous, bounded or unbounded. For example,

$$t \in \{0,1,\ldots,N\}, \mathbb{N}_0, [0,T] \text{ or } [0,+\infty).$$

Similarly, a filtered probability space (also known as a stochastic basis) is a probability space with a filtration of its σ-algebra.

It is also useful (in the case of an unbounded index set) to define $\mathcal{F}_\infty$ as the σ-algebra generated by the infinite union of the $\mathcal{F}_t$'s, which is contained in $\mathcal{F}$:

$$\mathcal{F}_\infty = \sigma\left(\bigcup_{t\geq 0}\mathcal{F}_t\right) \subseteq \mathcal{F}.$$

A σ-algebra defines the set of events that can be measured, which in a probability context is equivalent to events that can be discriminated, or "questions that can be answered at time t". Therefore a filtration is often used to represent the change in the set of events that can be measured, through gain or loss of information. A typical example is in mathematical finance, where a filtration represents the information available up to and including each time t, and is more and more precise (the set of measurable events is staying the same or increasing) as more information from the evolution of the stock price becomes available.

Lebesgue Integration

In mathematics, Lebesgue integration, named after French mathematician Henri Lebesgue (1875-1941), refers to both the general theory of integration of a function with respect to a general measure, and to the specific case of integration of a function defined on a subset of the real line or a higher dimensional Euclidean space with respect to the Lebesgue measure. This article focuses on the more general concept.

Lebesgue integration plays an important role in real analysis, the axiomatic theory of probability, and many other fields in the mathematical sciences.

The integral of a non-negative function can be regarded in the simplest case as the area between the graph of that function and the x-axis. The Lebesgue integral is a construction that extends the integral to a larger class of functions defined over spaces more general than the real line.