Subdifferential Calculus: Theory and Applications

A. G. KUSRAEV and S. S. KUTATELADZE

Preface

The subject of the present book is subdifferential calculus. The main source of this branch of functional analysis is the theory of extremal problems. For a start, we explicate the origin and statement of the principal problems of subdifferential calculus. To this end, consider an abstract minimization problem formulated as follows:

$$x \in X, \quad f(x) \to \inf .$$

Here X is a vector space and $f : X \to \overline{\mathbb{R}}$ is a numeric function taking possibly infinite values. In these circumstances, we are usually interested in the quantity $\inf f(x)$, the *value of the problem*, and in a *solution* or an *optimum plan* of the problem (i.e., such an $\bar{x}$ that $f(\bar{x}) = \inf f(X)$), if the latter exists. It is a rare occurrence to solve an arbitrary problem explicitly, i.e. to exhibit the value of the problem and one of its solutions. In this respect it becomes necessary to simplify the initial problem by reducing the task to somewhat more manageable modifications formulated with the details of the structure of the objective function taken in due account. The conventional hypothesis presumed in attempts at theoretically approaching the reduction sought is as follows. Introducing an auxiliary function l, we consider the next problem:

$$x \in X, \quad f(x) - l(x) \to \inf .$$

Furthermore, the new problem is assumed to be as complicated as the initial problem provided that l is a linear functional over X, i.e., an element of the *algebraic dual* $X^{\#}$. In other words, in analysis of the minimization problem for f, we consider as known the mapping $f^* : X^{\#} \to \overline{\mathbb{R}}$ that is given by the relation

$$f^*(l) := \sup_{x \in X} (l(x) - f(x)).$$

The function f^* thus introduced is called the *Young–Fenchel transform* of f. Observe that the quantity $-f^*(0)$ presents the value of the initial extremal problem.

The above-described procedure reduces the problem that we are interested in to that of change-of-variable in the Young–Fenchel transform, i.e., to calculation of the aggregate $(f \circ G)$, where $G : Y \to X$ is some operator acting from Y to X. We emphasize that f^* is a convex function of the variable l. The very circumstance by itself prompts us to await the most complete results in the key case of convexity of the initial function. Indeed, defining in this event the *subdifferential of f at a point $\bar{x}$*, we can conclude as follows. A point $\bar{x}$ is a solution to the initial minimization problem if and only if the next Fermat optimality criterion holds:

$$0 \in \partial f(\bar{x}).$$

It is worth noting that the stated Fermat criterion is of little avail if we lack effective tools for calculating the subdifferential $\partial f(\bar{x})$. Putting it otherwise, we arrive at the question of deriving rules for calculation of the subdifferential of a composite mapping $\partial(f \circ G)(\bar{y})$. Furthermore, the adequate understanding of G as a convex mapping requires that some structure of an ordered vector space be present in X. For instance, the presentation of the sum of convex functions as composition of a linear operator and a convex operator

$$f_1 + f_2 = + \circ (f_1, f_2);$$
$$(f_1, f_2) : X \to \mathbb{R}^2, \quad (f_1, f_2)(x) := (f_1(x), f_2(x)),$$

presumes the introduction into $\mathbb{R}^2$ the coordinatewise comparison of vectors.

Thus, we are driven with necessity to studying operators that act in ordered vector spaces. Among the problems encountered on the way indicated, the central places are occupied by those of finding out explicit rules for calculation of the Young–Fenchel transform or the subdifferential of a composite mapping. Solving the problems constitutes the main topic of subdifferential calculus.

Now the case of convex operators, which is of profound import, appears so thoroughly elaborated that one might speak of the completion of a definite stage of subdifferential analysis.

Research of the present days is conducted mainly in the directions related to finding appropriate local approximations to arbitrary not necessarily convex operators. Most principal here is the technique based on the F. Clarke tangent cone which was extended by R. T. Rockafellar to general mappings. However, the stage of perfection is far from being obtained yet. It is worth nonetheless to mention that key technical tricks in this direction lean heavily on subdifferentials of convex mappings.

In this respect we confine the bulk of exposition to the convex case, leaving the vast territory of nonsmooth analysis practically uncharted. The resulting gaps transpire. A slight reassuring apology for us is a pile of excellent recent books and surveys treating raw spots of nonsmooth analysis. The tool-kit of subdifferential analysis is quite full. It contains the principles of classical functional analysis, methods of convex analysis, methods of the theory of ordered vector spaces, measure theory, etc.

Many problems of subdifferential and nonsmooth analysis were recently solved on using nonstandard methods of mathematical analysis (in infinitesimal and Boolean-valued versions). In writing the book, we bear in mind the intention of (and the demand for) making new ideas and tools of the theory more available for a wider readership. The limits of every book (this one inclusively) are too narrow for leaving an ample room for self-contained and independent exposition of all needed facts from the above-listed disciplines.

We therefore choose a compromising way of partial explanations. In their selection we make use of our decade experience from lecture courses delivered in Novosibirsk and Vladikavkaz (North Ossetian) State Universities.

One more point deserves straightforward clarification, namely, the word "applications" in the title of the book. Formally speaking, it encompasses many applications of subdifferential analysis. To list a few, we mention the calculation of the Young–Fenchel transform, justification of the Lagrange principle and derivation of optimality criteria for vector optimization problems. However, much more is left intact and the title to a greater extent reflects our initial intentions and fantasies as well as a challenge to further research.

Chapter 1. Convex Correspondences and Operators

The concept of convexity is among those most important for contemporary functional analysis. It is hardly puzzling because the fundamental notion of the indicated discipline, that of continuous linear functional, is inseparable from convexity. Indeed, the presence of such a nonzero functional is ensured if and only if the space under consideration contains nonempty open convex sets other than the entire space.

Convex sets appear in many ways and sustain numerous transformations without loosing their defining property. Among the most typical should be ranked the operation of intersection and various instances of set transformations by means of affine mappings. Specific properties are characteristic of convex sets lying in the product of vector spaces. Such sets are referred to as convex correspondences. All linear operators are particular instances of convex correspondences. The importance of convex correspondences increased notably in the last decades due to their interpretation as models of production.

Among convex correspondences located in the product of a vector space and an ordered vector space, a rather especial role is played by the epigraphs of mappings. Such a mapping, a function with convex epigraph, is called a convex operator. Among convex operators, positive homogeneous ones are distinguished, entitled sublinear operators and presenting the least class of correspondences that includes all linear operators and is closed under the taking of pointwise suprema. Some formal justification and even exact statement of the preceding claim require the specification of assumptions on the ordered vector spaces under consideration. It is worth stressing that all the concepts of convex analysis are tightly interwoven with various constructions of the theory of ordered vector spaces. Furthermore, the central place is occupied by the most qualified spaces, Kantorovich spaces or K-spaces for short, which are vector lattices whose every above-bounded subset has a least upper bound.

The immanent interrelation between K-spaces and convexity is one of the most important themes of the present chapter. An ample space is also allotted to describing in detail the technique of constructing convex operators, correspondences and sets from the already-given ingredients. An attractive feature of convexity theory is an opportunity to provide various convenient descriptions for one and the same class of objects. The general study of convex classes of convex objects constitutes a specific direction of research, global convex analysis, which falls beyond the limits of the present book. Here we restrict ourselves to discussing the simplest methods and necessary constructions that are connected with the introduction of the Minkowski duality and related algebraic systems of convex objects.

The study if convex sets of operators naturally leads to Some vector spaces that are normed by elements of an auxiliary vector lattice. Under a nonrestrictive assumption of decomposability, this gives rise to the structure of a module over a lattice-ordered ring of orthomorphisms, which in turn allows us to characterize module-discrete elements. The latter are tired with the extreme structure of many convex sets of operators. We address this topic further in Chapter 2.

Chapter 2. Geometry of Subdifferentials

There are various connections and interrelations between convex objects which make the latter a convenient tool for investigation of numerous problems. One of the most general form of such relations is granted by the Minkowski duality. It is sufficient to study the duality for some particular class of objects under consideration, for instance, for sublinear operators. Basing on information on their support sets, we can rather easily obtain a description for subdifferentials of arbitrary convex operators, we can find corresponding Young–Fenchel transforms, we can study related extremal problems, etc.

Thus, the principal theme of the present chapter is analysis of the *classical Minkowski duality*, which is the mapping that assigns to a sublinear operator its support set or, in other words, its subdifferential (at zero). The questions arisen in this way relate as regards their form to various branches of mathematics. Indeed, it is necessary to find out the subdifferential of the composition of operators, to look for analogs of the “chain rule” in calculus. The resulting problems are usually attributed to analysis. There also arises a need of describing those algebraic structures in which the laws of subdifferential calculus take place. These problems fall within the competence of algebra. It is of use to explicate the structure of a subdifferential from the standpoint of the classical theory, to study the ways of recovering a subdifferential from its extreme points. The last question lies in the traditional sphere of geometry.

It is very important to emphasize that a specific treat of the problems of subdifferential calculus is incorporated in the synthetic approaches and the variety of the tools for solution, the peculiarities characteristic of functional analysis. However, the leading role is nonetheless played by ideas on the geometric structure of a subdifferential. The problems arising in this connection and above all the problem of describing a subdifferential intrinsically; i.e., in terms not involving the operator that is determined by the subdifferential, form the center of the subsequent exposition. In 2.4.10–2.4.13, an explicit representation of the elements of a subdifferential and its extreme points with the help of a concrete procedure applied to o-extreme points is given.

In the class of sublinear operators we distinguish canonical operators with comparatively simple structure so that only one canonical operator is assigned to each K-space and each cardinality. Any other total sublinear operator is obtained as composition of a canonical operator and a linear operator. Thus there arises a possibility of reducing general questions of the theory of sublinear operators to the analysis of a canonical operator and a linear change of variables in it. This constitutes generally the main idea of the canonical operator method.

According to the scalar theory there is a natural interconnection between discrete functionals and extreme points which leads to the classical Kreĭn–Mil′man theorem. On the other hand, the validity of the Kantorovich extension theorem for a discrete operator suggests that the support sets of general operators are analogous to the usual subdifferentials of convex functions. In fact, such analogy can be drawn on rather deeply and completely. Further we will discuss the analogy in detail. Now we only note that in the case of operators a statement holds which is more delicate than a possibility of reconstruction of support sets from their extreme points.

Studying subdifferentials we come across the algebraic structures with richer structure than that of the initial vector spaces. This can be seen, in particular, in Section 1.5. We should especially emphasize that the support set of a sublinear operator is operator convex rather than simply convex i.e. it satisfies an analog of the usual definition of convexity in which multiplicators serve as scalars. In other words, studying conventional convex objects in vector spaces, we necessarily come to more general analogs of convexity, i.e. to convexity in modules over rings (in particular, over the multiplication ring of a K-space with a strong order unit). There is a more important reason for interest in convex objects in modules. In applications one often encounters problems where the divisibility hypothesis is not acceptable. Such are certainly all problems of integer programming. In this connection of considerable importance is to clarify to what extent it is possible to preserve the subdifferential machinery for arbitrary algebraic systems.

The cones of operators do not have as a rule extreme rays (and thus caps, i.e. nonempty convex weakly compact subsets with convex complement); subdifferentials are compact in no locally convex topology and, at the same time, they can be recovered from their extreme points. The nature of such effects restricting the application of direct geometric methods reveals in Boolean-valued analysis. Indeed, the obstacles turn out to be imaginary to a certain extent and they can be bypass by choice of a suitable Boolean-valued model in which the considered object should be studied. In the previous section this approach was exposed in detail for subdifferentials, i.e. for strongly operator convex pointwise o-closed weakly order bounded sets. The aim of the further presentation is to weaken the boundedness assumption in the spirit of the classical theory of caps which was developed by G. Choquet and his successors. The peculiarity of our approach consists in working with the new notion of operator cap which is not a cap in the classical sense, though coincides with it in the scalar case. The criteria for subdifferentials to be caps and faces of sets of operators are given.

In the closing section we give characterization for a more special kind of sublinear operators, namely, those representable as the upper envelope of a family of n-disjoint positive operators. We pay attention to the subdifferentials of these operators and the extreme points of the former.

Chapter 3. Convexity and Openness

By now we have executed our study of subdifferentials on an algebraic level. To put it more precisely, we studied total sublinear operators, or what is the same, the subdifferentials of convex operators at interior points of their domains. Involving topology seems not sufficiently reasonable at this juncture since, in the presence of natural compatibility with the order structure of the domains, all subgradients appear to be automatically continuous in the same sense in which so was the initial sublinear operator. The situation is drastically different for the sublinear operators defined not on the whole space and resulting conventionally as the directional derivatives of convex operators at boundary points of their domains. Here the doors are widely open for all types of pathology. At the same time the study of subdifferentials at boundary points is an absolute necessity in the overwhelming majority of cases. Suffice it to recall that the origin of subdifferential calculus is tied with the modern sections of the theory of extremal problems which provide sophisticated tools for dealing with the intrinsic structure of the set of feasible solutions over which we seek for the optimal value of an objective function is sought.

The central theme of the current chapter is the interaction between convexity and openness in topological vector spaces. Strictly speaking, we study here the conditions under which a convex correspondence is open at a point of its domain. As usual, openness means that open sets containing a point of the domain are transformed by the considered correspondence onto neighborhoods of a fixed element in the image of the point under study. Analysis of the property and its most profound modification leading to the concept of general position for convex sets or convex operators enable us to achieve substantial progress in the problems of subdifferentiation. A matter of fact, we arrive at the opportunity to automatically derive the existence theorems for continuous operators by analyzing only the algebraic version of the problem under consideration.

Using the technique of general position wee can obtain all basic formulas for calculating the polars of conic segments and gauge functions. The problem is solved in the two steps: We, firstly, construct the Minkowski functional of a composite conic segment and, secondly, find the subdifferential of a composite sublinear functional. The first step reduces to simple calculation and the second involves the concept of general position.

It is inconceivable to treat topology and convexity simultaneously without the fundamental concept of duality of vector spaces. Therefore, we develop some apparatus for polar calculus, a polar actually presenting the subdifferential of a Minkowski gauge functional. Also, we give applications of this apparatus to description of open correspondences. A separate important topic is the openness principle for correspondences which summarizes the ideas that stem from the classical Banach open mapping theorem and simplify the applicability of the technique of subdifferentiation.

Application of duality to the analysis of the phenomenon of automatic continuity of convex correspondences reveals the interconnection of the openness principle and the Kreĭn–Shmulyan Theorem about dual characterization of the completeness property of a metrizable locally convex space as well as completeness of various spaces of closed convex sets or, which is the same, hypercompleteness (conic hypercompleteness, perfect completeness, etc.).

Analysis of the proof of the classical open mapping principle which was given by S. Banach himself shows that the method of a “rolling ball” remains applicable in the case when a fundamental sequence of balls of a metrizable topological vector space is replaced with a decreasing sequence of sets refining themselves in an appropriate sense. Systematization of this observation leads to a new concept of webbed correspondence and new results on the automatic continuity of convex correspondences.

Chapter 4. Apparatus of Subdifferential Calculus

The present chapter is the culmination of the book. Here, grounding on the already-developed methods, we deduce the main formulas of subdifferential calculus.

We start with the derivation of the change-of-variable formula for the Young–Fenchel transform. Leaning on them, we then find out formulas for computing ε-subdifferentials which present the generalization of the concept of subdifferential that make it possible to take account of the possibility of solving an extremal problem to within a given ε. It should be emphasized that analysis of ε-subdifferentials converting formally into conventional subdifferentials at $\varepsilon = 0$ has some particularities and subtleties. Complete technical explanations will be given in due course. It suffices now to observe that respective differences are as a matter of fact connected with the truism that the zero element is small in whatever reasonable sense whereas "*a small* ε" can designate a rather large residual.

The rules, which yield a formal apparatus for accounting for the measure of precision in dealing with subdifferentials (for instance, in the analysis of convex extremal problems, do not completely correlate with the practice of "neglecting infinitesimals" used in many applied works. The rules for approximate calculation are in perfect accord with the routine infinitesimal conception that the sum of two infinitesimals is infinitesimal. In other words, the practical methods of using ε-subgradients correspond to treating ε as an actual *infinitesimal*.

In modern mathematics, such conceptions are substantiated in the context of infinitesimal analysis called sometimes by expressive but slightly arrogant term "nonstandard analysis." By applying the indicated approach, a convenient apparatus can be developed for dealing with approximate, infinitesimal subdifferentials, which adequately reflects the rules for calculating "practical" optima.

While studying the Young–Fenchel transform, we are confronted with the question of whether it acts as involution. In the language of extremal problems we are talking about the absence of the duality gap. In view of utmost theoretical and practical importance of the indicated phenomenon, we discuss several ways of approaching and settling the problem.

Of paramount importance is the question of validity for the analog of the "chain rule" of the classical calculus: the subdifferential of a composition equals the composition of the subdifferentials of the composed mappings. Clearly, the rule fails in general. However,the rule is operative when we sum, integrate or take a finite supremum. The technique of treating the effect was titled *disintegration*. The apparatus of disintegration is closely related to the positive operators that preserve order intervals, i.e. that meet the Maharam condition. Study of order continuous operators with the property (they are referred to as Maharam operators) is of profound independent import for the general theory of Kantorovich spaces. General methods of disintegration unify, in a conventional form of the rules of calculus, various facts of the theory of K-spaces which are based on the Radon–Nikodým theorem. Here an analogy can be established with the fact that the calculus of support sets provides a uniform approach to different variants of the extension principles based on the application of the

Hahn–Banach–Kantorovich theorem.

Chapter 5. Convex Extremal Problems

The conventional field of application for convex analysis is the theory of extremal problems. The respective tradition ascends to the classical works of L. V. Kantorovich, Karush, and Kuhn and Tucker. Now we will touch the section of the modern theory of extremal problems which is known as convex programming. The exposition to follow is arranged so that everywhere we deal with multiple criteria optimization, i.e. the extremal problems with vector-valued objective functions are treated, whereas the bulk of the presented material is of use for analyzing scalar problems (those with a single target).

The characteristic particularity of the problems of multiple criteria optimization consists in the fact that, while seeking for an optimum solution, we must take account of different utility functions contradictory to each other. At this juncture it is as a rule impossible to distinguish a separate objective without ignoring the others and thus changing the initial statement of the problem. The indicated circumstance leads to the appearance of specific questions that are not typical of the scalar problems: what should be meant by a solution to a vector program; how can different interests be harmonized; is such a harmonization possible in principle; etc.? At this juncture we discuss various conceptions of optimality for multiple criteria problems; the ideal and generalized optima, the Pareto optimum, as well as the approximate and infinitesimal optima.

The apparatus of subdifferential calculus presents an effective tool for analyzing extremal problems. The change-of-variable formulas for the Young–Fenchel transform are applied to justification of numerous versions of the Lagrange principle: an optimum in a multiple criteria optimization problems is a solution to an unconstrained problem for a suitable Lagrangian. With the aid of ε-subdifferential calculus we deduce optimality criteria for approximate and infinitesimal solutions together with those for Pareto optima.

We also establish a vector-valued version of the Ekland variational principle and give some applications to generalized solutions and ε-subdifferentials.

We pay the main attention to the general conceptual aspects, leaving aside those that are thoroughly dealt with in the vast literature on the theory of extremal problems.

Chapter 6. Quasidifferentials

A mapping is quasidifferentiable at an interior point of the domain of definition if the directional derivative at this point exists and may be presented as the difference of sublinear operators. The quasidifferential is introduced to be the appropriate element of the space of sets by using a natural extension of the Minkowski duality. This opens ways of abstracting the tools of subdifferential calculus to a rather wide class of quasidifferentiable mapping which contains convex and concave operators.

The problem of expressing the quasidifferential of a composite mapping through the quasidifferentials of its constituents splits into the three steps: (1) finding the explicit formula for the directional derivative of a composite mapping through the directional derivatives of the constituents; (2) representation of the so-obtained derivative as the difference of some sublinear operators; and (3) calculation of the quasidifferential of the mapping under study through the quasidifferentials of the constituents. The first step consists in calculation of the respective limits and involves slight modification of the arguments of the classical calculus. The second step is often not straightforward and requires inventing some extra tricks. The third step rests on the Minkowski duality extended to the class of quasilinear operators representable as differences of sublinear operators. Using this scheme, we derive all calculus rules for quasidifferentials, The formulas for the quasidifferentials of the sum, product, composition, sum, supremum, and infimum of quasilinear operators.

We have seen in Chapter 4, that only on some special occasions we observe the full analog of the *chain rule* of the classical calculus: the subdifferential of a composite mapping is the composition of the subdifferentials of the constituents. These special occasions rest on the disintegration technique and Maharam operators. Part of this translates to quasidifferentials.

Quasidifferential calculus enables us to derive necessary optimality conditions for multiple criteria extremal problems with constraints using quasidifferentiable mappings. This is achieved mainly with the same tools as described in Chapter 5.

Chapter 7. Local Convex Approximations

Modern research into general nonsmooth optimization problems pays much attention to finding convenient approximations for sufficiently wide classes of functions and sets. One of the available types of local approximation is the use of quasidifferentials which was addressed in Chapter 6. Its advantages notwithstanding, the concept of quasidifferential is far from being the most convenient and effective tool in all cases. Various authors have invented a good deal of local approximations: contingents, hypertangent directions, upper (lower) convex approximations, Hadamard, Bouligand and Clark cones, Dini, Hadamard, and Rockafellar derivatives fail to exhaust the list of available approximations. However, no universal type of approximations exists which suits all classes of problems. Various local convex approximations complement one another. Moreover, each of the available types may be best in a particular class of extremal problems.

An important signpost of the theory of local convex approximations Is the definitions by F. Clark of the subdifferential of a locally Lipschitz function and the respective tangent cone to a set. The idea behind the F. Clarke definition has an infinitesimal origin. His observation reads as follows: if one collects all directions that are feasible for all points arbitrarily closed to the point under study, then a convex cone arises which approximates the initial set so closely that it can be successfully employed in deriving necessary conditions for an extremum. The introduction of the Clark tangent cone entailed a turmoil of research in nonsmooth analyses as well as spreading of new ideas and methods which have essentially influenced progress in the theory of extremal problems. However, this area is still far from the state of completeness and understanding which is achieved in the subdifferential calculus of convex operators.

The approaches to local approximation of sets and functions which are involved in subdifferential calculus and nonsmooth analysis are often connected with complex and bulky formulas. The relevant concepts of hypertangents, Rockafellar's limits, and Clark's derivatives may bewilder the reader at a first glance since their intrinsic meaning remains totally obscure. Nonstandard analysis suggests effective alleviating procedures. The new external concepts, using actual infinites and infinitesimals, "kill quantifiers," which simplifies the complexity of understanding and applying the standard intricate constructions of nonsmooth analysis.

We freely and happily evoke the technique of robinsonian nonstandard analysis for developing the theory of one-sided tangents to arbitrary functions and sets. A special attention we pay to the infinitesimal status of the Clark cone and the relevant reqularizing and approximating cones as well as the technique of operations over these cones.

Audience

This book is conceived as bridging the gap between the theoretical core of modern functional analysis and its applicable sections invoked by optimization, optimal control, mathematical programming, economics and related subjects. The book will be of service to mathematicians with theoretical background seeking for new applications as well as for those keen on applications and looking for contemporary powerful theoretical tools.

Оглавление