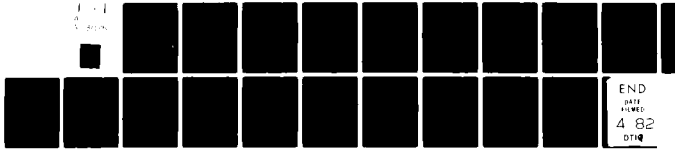
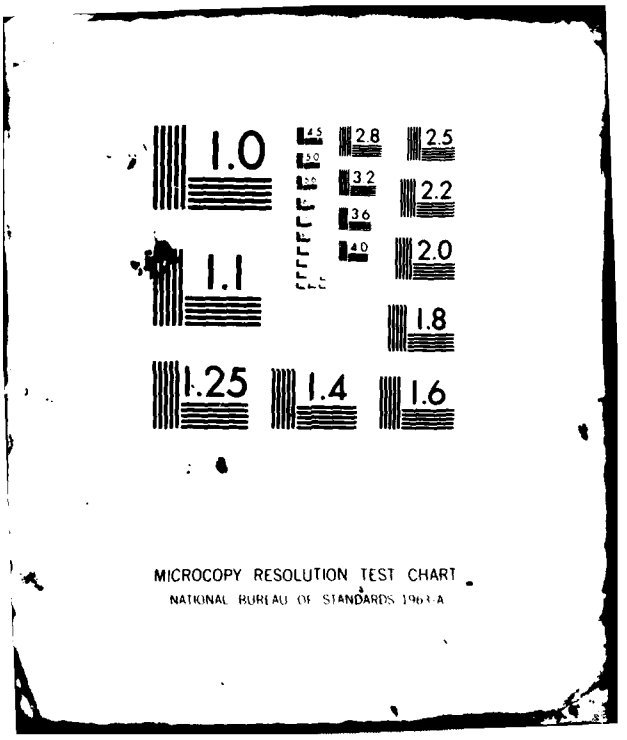
AD-A113 075 WASHINGTON UNIV SEATTLE DEPT OF MATHEMATICS F/G 12/1
OPTIMIZATION PROBLEMS: DUALITY AND MULTIPLIER METHODS.(U)
FEB 82 R T ROCKAFELLAR AFOSR-77-3204

UNCLASSIFIED AFOSR-TR-82-0259 NL



END
DATE FILMED
4 82
DTIC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: AFOSR-TR- 82-0259	2. GOVT ACCESSION NO.: AD-A113075
	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): OPTIMIZATION PROBLEMS: DUALITY AND MULTIPLIER METHODS.	5. TYPE OF REPORT & PERIOD COVERED: Final (Technical Report)
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): R.T. Rockafellar	8. CONTRACT OR GRANT NUMBER(s): AFOSR-77-3204
9. PERFORMING ORGANIZATION NAME AND ADDRESS: University of Washington, Department of Mathematics, Seattle, Washington 98195	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61102F, 2304/A6
CONTROLLING OFFICE NAME AND ADDRESS: Air Force Office of Scientific Research (NM), Bolling AFB, Washington, DC 20332	12. REPORT DATE: February 1982
	13. NUMBER OF PAGES: 21
MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

DTIC ELECTE MAR 0 6 1982

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Nonlinear programming, stochastic programming, subgradient analysis, optimal control

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Research under this grant during the period 1977-1981 resulted in 20 technical articles, two books and a Ph.D. thesis. The work is summarized under the following headings: (1) nonlinear programming algorithms, (2) multistage stochastic optimization, (3) networks and monotropic programming, (4) subgradient analysis and nonsmooth optimization, (5) marginal values and sensitivity analysis, (6) genericity of optimality conditions, (7) optimal control of dynamical systems.

DD FORM 1 JAN 73 1473

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AFOSR-TR- 82-0259

OPTIMIZATION PROBLEMS: DUALITY AND MULTIPLIER METHODS

Final Scientific Report on grant AFOSR-77-3204

R.T. Rockafellar, Principal Investigator

February 19, 1982

Introduction. In the five years covered by this grant, January 1977 through December 1981, a total of twenty research articles were written for publication in technical journals, as well as two books and a Ph.D. thesis. Although these are addressed to a wide variety of optimization problems, they have a common theme: the characterization and computation of solutions by methods based on subgradient analysis and duality. Fundamental advances in theory are embodied in this work. The following topics will be discussed individually below:

A. Multiplier Algorithms in Nonlinear Programming

B. Multistage Stochastic Optimization

C. Networks and Monotropic Programming Methods

D. Generalized Subgradients and Nonsmooth Optimization

E. Marginal Values and Sensitivity Analysis

F. Genericity of Optimality Conditions

G. Optimal Control of Dynamical Systems

References [1], [2], etc., are to work performed under this grant, while references [a], [b], etc., are to other publications; all are listed at the end.

A. Multiplier Algorithms In Nonlinear Programming.

A general nonlinear programming problem in finitely many variables has the form

$$\text{Minimize } f_0(x) \text{ over all } x \in X \text{ satisfying}$$

$$(1) \qquad f_i(x) \begin{cases} \leq 0 & \text{for } i=1,\ldots,s, \\ = 0 & \text{for } i=s+1,\ldots,m, \end{cases}$$

where $X \subset R^n$ and $f_i : R^n \to R$, $i=0,1,\ldots,m$. A large number of techniques have been proposed over the years for solving such problems, but among the most popular and effective nowadays are the so-called multiplier methods, initiated independently by M.R. Hestenes [a] and M.J.D. Powell [b] in 1969 and developed extensively in the mid-70's, especially by D. Bertsekas [c] and R.T. Rockafellar [d], [e]. The virtue of these methods is that they avoid the difficulties of dealing directly with nonlinear constraints by replacing (1) by a certain sequence of unconstrained minimization problems. In this there is a similarity with penalty methods, and indeed, multiplier methods have largely supplanted the latter because they exhibit the same virtues along with better convergence rates and greater numerical stability.

Multiplier methods involve Lagrange multipliers in addition to a penalty parameter. They are based on the study of the augmented Lagrangian function for problem (1), namely

$$(2) \qquad L(x,y,r) = f_0(x) + \sum_{i=s+1}^{m} \{y_i f_i(x) + \tfrac{1}{2} r f_i(x)^2\} + \sum_{i=1}^{s} \begin{cases} y_i f_i(x) + \frac{1}{2} r f_i(x)^2 & \text{if } f_i(x) \geq -y_i/r \\ -\frac{1}{2r} y_i^2 & \text{if } f_i(x) \leq -y_i/r \end{cases}$$

$$\text{for } x \in X,\ y \in R^m,\ r > 0.$$

This contrasts with the ordinary Lagrangian function

$$(3) \qquad \ell(x,y) = f_0(x) + \sum_{i=1}^{m} y_i f_i(x)$$

$$\text{for } x \in X,\ y \in R_+^s \times R^{m-s}$$

in furnishing saddle point representations of optimal solutions to (1) (augmented duality) even in nonconvex programming. (This was established in [d] by Rockafellar, who also was responsible for showing that formula (2) was the right way to incorporate inequality constraints into the augmented Lagrangian.)

The basic form of the Hestenes-Powell multiplier method begins with a choice of x^0, y^0 and r_0 and in the general step takes

$$(4)\qquad \begin{aligned} x^{k+1} &\approx \operatorname*{argmin}_{x\in X} L(x,y^k,r_k), \\ y^{k+1} &= \nabla_y L(x^{k+1},y^k,r_k), \qquad r_{k+1} \geq r_k. \end{aligned}$$

Here the notation means that x^{k+1} is an approximate minimizer of the function $L(\cdot,y^k,r_k)$ on X. The set X is supposed to have a simple form (e.g. a generalized rectangle), so this minimization, which uses x^k as the starting point, can be effected by means of the highly efficient algorithms now known for (essentially) unconstrained optimization. The main questions concern the stopping rule that should be used in the approximate minimization, the strategy in updating the penalty parameter, and the kinds of convergence that can be obtained. Generally speaking, it is possible to obtain global convergence at an arbitrarily good linear rate, without having $r_k \to \infty$. For this, however, one must use a stopping criterion of the form

$$(5)\qquad F_k(x^{k+1}) - \inf_{x\in X} F_k(x) \leq \varepsilon_k, \quad \text{where} \quad F_k(x) = L(x,y^k,r_k),$$

which requires good estimates of the greatest lower bound for $L(\cdot,y^k,r_k)$ on X, something not always available.

The two articles [3] and [11] on this subject that were produced under this grant propose a new version of the multiplier method that gets around these difficulties with the stopping criterion and, as a byproduct, makes it possible to solve an important class of "extended" convex programming problems, called variational inequalities. The rule in the new version involves another parameter $s_k > 0$:

$$(6)\qquad \begin{aligned} x^{k+1} &\approx \operatorname*{argmin}_{x\in X} \{L(x,y^k,r_k) + (1/2s_k)\|x-x^k\|^2\}, \\ y^{k+1} &= \nabla_y L(x^{k+1},y^k,r_k) \qquad r_{k+1} \geq r_k,\ s_{k+1} \geq s_k. \end{aligned}$$

This is just as easy to execute and leads to the same nice convergence properties. Its big advantage is that it is amenable to a stopping criterion of a much more convenient type:

$$(7)\qquad \|\text{proj } \nabla F_k(x^{k+1})\| \leq \varepsilon_k,$$
$$\text{where } F_k(x) = L(x,y^k,r_k) + (1/2\ s_k)\|x-x^k\|^2,$$

the projection being that of the gradient $\nabla F_k(x^{k+1})$ on the tangent cone to X at x^{k+1} (which for the usual sets X is simple to compute).

A very interesting feature of the modified rule is that everything can be carried out in terms of the mappings ∇F_k alone. The function values $F_k(x)$ don't need to play any role. This being the case, it is possible to replace the gradient ∇f_0 in problem (1) by a much more general kind of mapping $A:R^n \to R^n$. The sequence (x^k,y^k) generated by the algorithm then converges (under mild assumptions) to a solution $(\bar{x},\bar{y})$ to the so-called variational inequality obtained when $A(\bar{x})$ is substituted for $\nabla f_0(\bar{x})$ in writing down the Kuhn-Tucker conditions for optimality in (1).

All this works in particular when A is a monotone mapping in the sense that

$$[A(x') - A(x)]\cdot[x' - x] \geq 0 \quad \text{for all} \quad x', x \in R^n.$$

Indeed, much of the theory of multiplier methods rests on the study of such mappings, so this is a very natural extension. It provides a new computational handle on many problems in partial differential equations that can be represented as variational inequalities.

Much remains to be done in connection with these problems and their special structures. The relationship between the parameters r_k and s_k, and the strategies for updating them, would benefit from further study too.

B. Multistage Stochastic Optimization. A common but difficult situation to deal with in applications of optimization is the kind where decisions must be made here and now, but the outcomes of these decisions will be strongly affected by future events about which there is only statistical information. Usually, recourses are available in the future in order to correct the effects of the here-and-now decisions, after the true situation becomes better known. But the cost and scope of the recourses may depend too on what has to be decided in advance. Multistage stochastic optimization problems, also called stochastic programming problems or recourse problems, are an attempt to model this state of affairs.

To keep things simple, let us imagine a situation where at times $t=1,2,\ldots,N$ a vector x_t must be chosen from a space R^{n_t} in response to an observation w_t (which is a random vector variable with known distribution, at least in the

most elementary versions of the model). Present decisions cannot depend on future observations, so a decision policy must be a function of the special form

$$x(w) = (x_1(w_1), x_2(w_1,w_2), \ldots, x_N(w_1,w_2,\ldots,w_N)).$$

Such a function x is said to be nonanticipative. The problem is to minimize, over all nonanticipative functions x, an expected cost

$$E_w\{f_0(w,x(w))\}$$

subject to constraints of the form

$$(8) \qquad f_i(w,x(w)) \leq 0 \quad \text{almost surely,} \quad i=1,\ldots,m.$$

The theory of such problems was developed by Rockafellar and Wets in the two-stage case ($N=2$) in a series of papers [f], [g], [h], [i]. The foundation for the N-stage case was laid in [j].

Article [1], written under the present grant, derives for the first time the existence of Lagrange multipliers $y_i(w)$ for the constraints (8) as a characterization of the optimal decision function x. It develops special properties in the case of separable constraint functions

$$f_i(w,x(w)) = f_{i1}(w_1,x_1(w_1)) + f_{i2}(w_2,x_2(w_1,w_2)) + \ldots$$

and explores certain connections with stochastic optimal control. Convexity is assumed throughout.

The results are very complete and satisfying as regards optimality conditions and their interpretation, and they can fairly be viewed as a landmark in stochastic optimization on such terms. Nevertheless, they are only theoretical results. They are an important step towards computation, but much work on actual algorithms

will be needed before problems of this highly important kind can be solved practically and efficiently. Large-scale decomposition techniques in terms of the Lagrangian price vectors $y(w)$ will be required. The theory of nonconvex problems will eventually need to be developed too.

C. Networks and Monotropic Programming.

A monotropic programming problem is an optimization problem in which a convex function having a representation of the type

$$(9) \qquad F(u_1,\ldots,u_m) = \sum_{j=1}^{n} f_j(a_{j1}u_1+a_{j2}u_2+\ldots+a_{jm}u_m+b_j)$$

is minimized subject to linear equality and inequality constraints on the variables $u_1,\ldots,u_m$. Linear and quadratic programming problems are a special case, as are separable convex programming problems. Indeed, any monotropic programming problem can be reduced to the canonical form

$$(10) \qquad \begin{aligned} &\text{minimize } f_1(x_1)+\ldots+f_N(x_N) \\ &\text{over all } x = (x_1,\ldots,x_N)\in K\subset R^N \text{ satisfying} \\ &\qquad x_j\in C_j \text{ for } j=1,\ldots,N, \end{aligned}$$

where K is a linear subspace of R^N (described by a system of homogeneous linear equations), each C_j is a real interval, and f_j is a closed proper convex function of a single real variable, having C_j as its effective domain. Associated with this is a canonical dual problem of the same sort:

$$(11) \qquad \begin{aligned} &\text{maximize } -g_1(y_1)-\ldots-g_N(y_N) \\ &\text{over all } y\in(y_1,\ldots,y_N)\in L\subset R^N \text{ satisfying} \\ &\qquad y_j\in D_j \text{ for } j=1,\ldots,N, \end{aligned}$$

where L is the linear subspace orthogonal to K (expressible by an adjoint system of equations), and g_j is the convex function on R conjugate to f_j; the interval D_j is the effective domain of g_j.

This kind of duality, which can be utilized almost as fully and explicitly as linear programming duality (which, by the way, it subsumes) is a characteristic feature of monotropic programming. It makes possible a whole range of methods and approaches that otherwise would not be available. This is why such problems need to be recognized and treated as a class apart.

The main theoretical guideline for general monotropic programming comes from network programming, namely the case where x_j represents the flow in the j^{th} arc of a certain directed graph and y_j is the "tension" across the arc. In that setting, K is the space of circulations (flows conserved at every node), and L is the space of tensions representable as potential differences (relative to some potential function defined on the set of nodes of the graph). An enormous number of practical problems in operations research, including logistics, warehousing, project scheduling and the analysis of pipe systems, fall into this category.

Article [13] introduces basic descent methods for monotropic programming problems. It demonstrates that any such method, applied to either problem (10) or (the negative of) problem (11), as is always possible due to total symmetry, will inevitably solve both (10) and (11). This computational circumstance leads to a new theoretical result: a constructive proof of the duality theorem for monotropic programming, i.e. the fact that the optimal values in (10) and (11) must be equal unless both problems fail to be feasible. This theorem is a powerful tool in the design and interpretation of algorithms. It holds a unique position in the duality literature in not requiring either the linearity of objectives or any kind of strict feasibility.

Algorithms in monotropic programming have a distinctly combinatorial nature: descent is in special directions induced by a matroidal substructure associated with the linear subspaces K and L. This subject has until now not been investigated or even recognized as a unified whole (although examples in linear and network programming are well known), and herein lies the novelty and significance of the monograph [16]. There is too much in this work to be described here. For a better idea of the contribution, the preface to [16], the table of contents and the section of comments at the end of each chapter may be consulted. Many new computational methods and conceptual innovations are provided. The book includes the first comprehensive treatment of nonlinear network flow problems and separable convex programming.

D. Subgradient Analysis and Nonsmooth Optimization.

This is another big subject on which far too much has been accomplished in the five years under the present grant for there to be any hope of giving more than a brief indication here. Motivation starts with the fact that optimization problems very frequently involve functions that are not differentiable, at least not everywhere.

In direct terms, one can run into cost functions that are merely piecewise smooth (the derivatives jump at certain breakpoints), as well as "max functions" of the form

$$(12) \qquad h(x) = \max_{i \in I} h_i(x) \qquad (I = \text{some index set}) \ ,$$

whose graphs exhibit "corner points" of a rather complicated sort. Convex functions on R^n are always representable as max functions (12) with each h_i affine (i.e. linear-plus-a-constant), and as this suggests, they are not necessarily

differentiable. Thus economic models in which convexity is postulated, but differentiability is less natural as a fundamental assumption, fall under the heading of "nonsmooth optimization".

Nondifferentiable functions also arise indirectly. In linear programming, for instance, the optimal value

$$(13) \qquad \varphi(b) = \inf\{cx \mid x \geq 0,\ Ax \leq b\}$$

is only piecewise linear with respect to the vector b. The role of the optimal solutions to the dual problem, as vectors of "shadow prices" associated with the resources in the primal problem, cannot be understood without reference to this potential lack of differentiability of φ. More generally, the quantity

$$(14) \qquad p(v) = \inf\{f_0(v,x) \mid f_1(v,x) \leq 0,\ldots,f_m(v,x) \leq 0\},$$

giving the optimal value in a constrained minimization problem in x which depends on a parameter vector v, is generally not differentiable with respect to v, even if the functions $f_0,f_1,\ldots,f_m$ themselves are infinitely smooth.

Exact penalty methods for solving nonlinear programming problems, as well as decomposition techniques and duality-based computational schemes of the sort that is now very popular in branch-and-bound approaches to combinatorial problems, typically lead to the consideration of auxilliary functions that are not smooth. Sometimes these functions take on quite a complicated form, as in the case of problems of engineering design where specifications can be met within certain tolerances by a "tuning" process after basic manufacture; see E. Polak [k] (also introduction to [12]).

Nondifferentiable convex functions have been treated successfully for many years. Many of the techniques were developed by R.T. Rockafellar and presented in his book Convex Analysis [ℓ]. The big breakthrough for nonconvex functions, however, came with the thesis work of F.H. Clarke under a predecessor of the present grant. Clarke was able to define subgradients of arbitrary lower semi-continuous functions on R^n is a manner totally in harmony with the convex case and the classical analysis of smooth functions; see [m]. Clarke's approach was somewhat roundabout, though, and his definitions seemed to depend unduly on the Euclidean norm in R^n, which tended to hamper applications, not to mention extensions to problems in infinite-dimensional spaces.

One of the main accomplishments under the present grant has been the further development and strengthening of the theory of generalized subgradients of non-convex functions, especially with an eye towards certain applications that will be discussed in the next section. Deep, fundamental results were obtained in [6], [7], [8] and [18]. These are long papers, and as mentioned above, it is impossible to go into the details here. Fortunately that isn't necessary, since the recently published monograph [12], also written under this grant, provides a readable survey, in fact the very first to become available on this burgeoning subject.

E. Marginal Values and Sensitivity in Nonlinear Programming.

The generalized subgradient analysis described above is ideally suited to elucidating the properties of the sort of nonsmooth function appearing in formulas (13) and (14). Let us imagine more generally a problem of the form

$$(P_{u,v}) \quad \text{minimize } f_0(v,x) \text{ over all } x \in D(v)$$
$$\text{satisfying } f_i(v,x) + u_i \begin{cases} \leq 0 & \text{for } i=1,\ldots,s, \\ = 0 & \text{for } i=s+1,\ldots,m, \end{cases}$$

where $v \in R^d$ and $u=(u_1,\ldots,u_m) \in R^m$. Denote the optimal value in $(P_{u,v})$ by $p(u,v)$; then p is an extended-real-valued function on $R^m \times R^d$, and under mild assumptions p is lower semicontinuous.

Although *no* amount of differentiability assumptions on $f_0,f_1,\ldots,f_m$ will imply differentiability of p, there are special situations where it has been known for some time that for a particular $(\bar{u},\bar{v})$, the gradient $\nabla p(\bar{u},\bar{v})$ exists. These tend to be situations where problem $(P_{\bar{u},\bar{v}})$ has a unique globally optimal solution $\bar{x}$, and this $\bar{x}$ happens to satisfy second-order optimality conditions of the strongest kind. The interesting thing is that in such situations

$$(15) \qquad \nabla p(\bar{u},\bar{v}) = (\bar{y},\bar{z}) \quad \text{with} \quad \bar{z} = \nabla_v \ell(\bar{v},\bar{x},\bar{y}),$$

where $\bar{y}$ is the unique Lagrange multiplier vector associated with $\bar{x}$, and

$$(16) \qquad \ell(v,x,y) = f_0(v,x) + \sum_{i=1}^{m} y_i f_i(v,x).$$

The reason this is so important is that it indicates a fundamental connection between the dual variables that occur in optimality conditions for problem $(P_{\bar{u},\bar{v}})$ and the possible rates of change of the function p at $(\bar{u},\bar{v})$.

Rates of change of p are called <u>marginal values</u>. They are significant in the economic analysis of optimization models where the components of u and v represent production coefficients, costs and resource availabilities that may be subject to fluctuation. They also have a role in determining the stability of computational procedures which could be at the mercy of errors in the specification of u and v. Furthermore, the ability to calculate, or at least estimate, such rates of change is valuable in decomposition techniques.

For example, the real problem to be solved may be one in which only u is given:

(Q_u) minimize $f_0(v,x)$ over all $(v,x)\in E$

$$\text{satisfying } f_i(v,x)+u_i \begin{cases} \leq 0 & \text{for } i=1,\ldots,s, \\ = 0 & \text{for } i=s+1,\ldots,m. \end{cases}$$

For each v, the corresponding subproblem of minimizing in x can be identified with problem $(P_{u,v})$ in the case of $D(v) = \{x|(v,x)\in E\}$. The master problem then consists of minimizing $p(u,v)$ with respect to v for fixed u. A decomposition of this sort may be very attractive in cases where $(P_{u,v})$ is particularly easy to solve for each (u,v) (a well known technique due to Bender). However, it does necessitate the minimization of a nonsmooth function p. Obviously, any information about directional rates of change of p is crucial to the success of such an approach.

In certain situations in convex programming, it has been known that formula (15) could be stated in a more general way in terms of subgradients rather than gradients, such subgradients being a way of describing one-sided directional derivatives. The challenge of the work under the present grant was to extend this somehow to nonconvex programming. Since one-sided derivatives of p in the ordinary sense do not necessarily exist, even under smoothness assumptions on the functions f_i and set D, basic theoretical developments were needed. These have been described in the preceding section.

Article [14] provided a key by giving an exact formula for the subgradient set $\partial p(\bar{u},\bar{v})$ in Clarke's sense in terms of extended limits of Lagrange multiplier vectors y^k associated with optimal solutions x^k to neighboring problems (P_{u^k,v^k}), as $(u^k,v^k) \to (\bar{u},\bar{v})$. In fact, the multiplier vectors in question satisfy the saddle point condition for the augmented Lagrangian for (P_{u^k,v^k}). Thus the augmented Lagrangian described in the first section of this report was

shown to have theoretical powers much beyond what might be expected simply from its role in computations.

The next paper, [15], developed this formula into estimates for $\partial p(\bar{u},\bar{v})$ not just in terms of limits of multipliers for neighboring problems, but certain multiplier vectors associated with optimality conditions for $(P_{\bar{u},\bar{v}})$ itself. Actually, this was a two-way process: the mathematical machinery that had been devised was sensitive enough to allow optimality conditions to be stated for solutions $\bar{x}$ to $(P_{\bar{u},\bar{v}})$, even when the functions f_i are not smooth and the multifunction D is merely of closed graph. These conditions were shown to be necessary on the basis of differential properties of p, a new technique in nonconvex optimization that sheds much light on the subject of "constraint qualifications". In particular, multiplier rules of Clarke [m] and Hiriart-Urruty [o] were sharpened in this way.

Many consequences will be obtained from the results in [15], due to their depth and far-reaching generality. This work is the culmination of much effort.

Article [17] deals with certain more abstract versions of the formulas in [15], true in part for infinite-dimensional problems. (The framework in [15] is intrinsically finite-dimensional.)

An application to second-order conditions is carried out in [19]. The formulas in [15] are refined in terms of second derivative information, and in this way new results on necessary conditions for optimality are again obtained.

F. Genericity of Optimality Conditions

Are the standard optimality conditions in nonlinear programming "usually" satisfied? This is the question tackled in article [9]. The question is significant because it is not possible, as a practical matter in most applications, to check whether a given nonlinear programming problem (1) satisfies the constraint qualifications and strengthened forms of the second-order optimality conditions on which the analysis of many algorithms, etc., depends. One often hears the argument that it is all right to base results on the assumption of such conditions, because they hold in "typical" problems. But what does that assertion really mean?

One approach is to consider families of problems that depend on parameters, like $(P_{u,v})$ in the preceding section. These parameters can be imagined as random variables with known distributions. The question can then be phrased as follows: consider the set of all pairs (u,v) such that $(P_{u,v})$ has a locally optimal solution which fails to satisfy certain conditions, and ask whether this set represents an event of probability zero. Now as long as the distributions are continuous, this can be subsumed by a much simpler question that doesn't involve knowledge of the particular statistical distributions of the parameters, namely, whether the exceptional set of pairs (u,v) is negligible (i.e. of measure zero in the Lebesgue sense).

An affirmative answer to this question was given in [9] and [10] for a fundamental class of parameterizations of nonlinear programming problems. This was complemented by results in [14] on the genericity of uniqueness of optimal solutions.

J. E. Spingarn in his Ph.D. thesis [21] developed a more complete theory. It was necessary to consider other classes of parameterizations in order to have a practical result, but it was not clear until his work, how to identify the ones with the desired property that "almost all" problems in the parameterized family

were well-behaved. He used differential topology and other mathematical tools to show how certain kinds of constraint structures could be kept fixed (unparameterized) without unbalancing the abundance of "good" problems over the "bad" in a given family. These results have been published by Spingarn in [22] and [23].

Besides serving to justify certain approaches to computation, it is expected that these ideas will have a role to play in multistage stochastic programming (see B above). In that subject one has to treat, as a matter of course, nonlinear programming subproblems which depend on random variables, and whose optimal solutions therefore are random variables too. It would be impossible to get very far without theoretical assurance that such optimal solution random variables can be analyzed in terms of nice kinds of multiplier conditions almost surely.

G. Optimal Control of Dynamical Systems.

Three publications under the present grant come under this heading, [4], [5] and [21]. In [4] the subject of duality in problems of optimal control is surveyed, and also a number of recent developments concerning the existence of optimal arcs and the conditions which characterize them. This exposition provides a good introduction to the general approach to optimal control that can be made in terms of extended-real-valued hamiltonians and subdifferential calculus.

Paper [5] describes in terms of models of optimal economic growth a number of results and open questions concerning control problems over an infinite time interval. The main question in such problems is what kind of behavior is naturally optimal in a "self-sustaining" sense, i.e. in a steady-state manner that could be prolonged indefinitely. The concepts that arise in this connection are interesting for several basic reasons especially as a description of limiting behavior in

various situations, even though real problems never involve infinite time.

State constraints are the subject of the most recent article [21]. It has been shows that such constraints in an optimal control problem can cause jumps (discontinuities) in the adjoint variables. Conversely, the possibility of jumps in the primary variables can be linked to inherent state constraints on the adjoint variables. This is what is proved in [21] through detailed analysis of a particular class of models of interest in economics.

Publications of R.T. Rockafellar written under Grant AF-AFOSR-77-3204 during 1977-1981

[1] "The optimal recourse problem in discrete time: L^1-multipliers for for inequality constraints" (with R.J.B. Wets), SIAM J. Control + Opt. 16(1978), 16-36.

[2] "Higher derivatives of conjugate convex functions," International J. Appl. Analysis 1(1977), 41-43.

[3] "Monotone operators and augmented Lagrangian methods in nonlinear programming," Nonlin. Programming 3(1978), O.L. Mangasarian (ed.), 1-25.

[4] "Duality in optimal control," Mathematical Control Theory, W.A. Coppel (ed.), Springer-Verlag Lecture Notes in Math. 680(1978), 219-257.

[5] "Convex processes and Hamiltonian dynamical systems," in Convex Analysis + Math. Econ., Springer-Verlag Lecture Notes in Economics and Math. Systems 168(1979), 122-126.

[6] "Clarke's tangent cones and the boundaries of closed sets in R^n," Nonlinear Analysis Th. Meth. Appl. 3(1979), 145-154.

[7] "Generalized directional derivatives and subgradients of nonconvex functions," Canadian J. Math. 32(1980), 157-180.

[8] "Directionally Lipschitzian functions and subdifferential calculus," Proc. London Math. Soc. 33(1979), 331-335.

[9] "The generic nature of optimality conditions in nonlinear programming," (with J.E. Spingarn), Math. of Op. Research 4(1979), 425-430.

[10] "Almost sure existence of Lagrangian price vectors in nonlinear programming," Math. Programming and its Applications, G. Castellani and P. Mazzoleni (eds.), Angeli, Milano (1981), 255-262.

[11] "Lagrange multipliers and variational inequalities," Variational Inequalities and Complementarity Problems, R.W. Cottle (ed.), Wiley(1980), 303-322.

[12] The Theory of Subgradients and its Applications to Problems of Optimization: Convex and Nonconvex Functions, originally published in French translation by the University of Montreal Press (1979), recently published also in English (with some updating) by Helderman Verlag, W. Berlin(1981).

[13] "Monotropic programming: descent algorithms and duality," Nonlin. Programming 4(1981), O.L. Mangasarian (ed.), Academic Press 327,366.

[14] "Proximal subgradients, marginal values, and augmented Lagrangians in nonconvex optimization," Math. of Op. Res. 6(1981), 424-437.

[15] "Lagrange multipliers and subderivatives of optimal value functions in nonlinear programming," Math. Programming Study 15(1982).

[16] Optimization in Networks and Monotropic Systems, Princeton Univ. Press, in process.

[17] "Augmented Lagrangians and marginal values in parametric optimization problems," in Generalized Lagrangian Methods in Optimization, A. Wierzbicki (ed.), Pergamon Press, in process.

[18] "Favorable classes of Lipschitz continuous functions in subgradient optimization," in Nondifferentiable Optimization. E. Nurminski (ed.), Pergamon Press, in process.

[19] "Marginal values and second-order conditions for optimality," Math. Programming, in process.

[20] "Optimality conditions for convex control problems with nonnegative states and the possibility of jumps," Game Theory and Math. Economics, O. Moeschlin (ed.), North-Holland(1981), 339-349.

Publications of J.E. Spingarn written under this grant

[21] Generic Conditions for Optimality in Constrained Minimization Problems, University of Washington, 1977 (Ph.D. Thesis).

[22] "Fixed and variable constraints in sensitivity analysis," SIAM J. Control Opt. 18(1980), 297-310.

[23] "On optimality conditions for structured families of nonlinear programming problems," Math. Programming 22(1982), 82-92.

Other References

[a] M.R. Hestenes, "Multiplier and gradient methods," J. Opt. Theory Appl. 4 (1969), 303-320.

[b] M.J.D. Powell, "A method for nonlinear constraints in minimization problems," Optimization, R. Fletcher (ed.), Academic Press, 1969, 283-298.

[c] D. Bertsekas, "On penalty and multiplier methods for constrained minimization," SIAM J. Control 13(1975).

[d] R.T. Rockafellar, "Augmented Lagrange multiplier functions and duality in nonlinear programming." SIAM J. Control 12(1974), 268-285.

[e] R.T. Rockafellar, "Augmented Lagrangians and applications of the proximal point algorithm in convex programming," Math. of Op. Research 1(1976), 97-116.

[f] R.T. Rockafellar and R.J.B. Wets, "Stochastic convex programming: basic duality," Pacific J. Math. 62(1976), 173-195.

[g] R.T. Rockafellar and R.J.B. Wets, "Stochastic programming: singular multipliers and extended duality," Pacific J. Math. 62(1976), 507-522.

[h] R.T. Rockafellar and R.J.B. Wets, "Stochastic convex programming: Kuhn-Tucker conditions," J. Math. Econ. 2(1975), 349-370.

[i] R.T. Rockafellar and R.J.B. Wets, "Stochastic convex programming: relatively complete recourse and induced feasibility," SIAM J. Control + Opt. 14(1976), 574-589.

[j] R.T. Rockafellar and R.J.B. Wets, "Nonanticipativity and L^1-martingales in stochastic optimization problems," Stochastic Systems: Modeling, Identification and Optimization, Math. Prog. Study 6(1976), 170-187.

[k] E. Polak and A. Sangiovanni-Vincentelli, "Theoretical and computational aspects of the optimal design centering, tolerancing and tuning problem," research report UCB/ERL M79/6, Electronics Research Lab., Berkeley, 1979.

[ℓ] R.T. Rockafellar, Convex Analysis, Princeton Univ. Press, 1970.

[m] F.H. Clarke, "Generalized gradients and applications," Trans. Amer. Math. Soc. 205(1975), 247-262.

[n] F.H. Clarke, "A new approach to Lagrange multipliers," Math. of Operations Research 1(1976), 165-174.

[o] J.B. Hiriart-Urruty, "Refinements of necessary optimality conditions in nondifferentiable programming, I," Appl. Math. + Opt. 5(1979), 63-82.


DAT
FILM