International Conference on Advances in Convex Analysis and Global Optimization

Honoring the memory of C. Caratheodory (1873-1950)

June 5 - 9, 2000
Pythagorion, Samos, Greece

ABSTRACTS OF TALKS

Abstracts are listed alphabetically by the speaker's name (underlined)

INVITED TALKS

Fast and Reliable Methods for Solving Optimal Control Problems

Roland BULIRSCH, **Rainer CALLIES and Oskar v. STRYK**

In the last decade, considerable progress has been achieved in the numerical optimal control field.

Refined collocation methods can now handle very difficult control problems, and they accomplish this without the need of any a priori complicated analytical preparation.

Improved multiple shooting procedures extend the range of the applicability of numerical methods deep into the regime of those optimal control problems which hitherto were thought as not being numerically solvable at all.

The talk is based on these new procedures and shows illustrative non-trivial examples which demonstrate the power of these methods.

The Role of Optimization in Control Theory

Francis CLARKE

In this nontechnical survey, we begin by defining the principal issue in control theory: to design a feedback law having a certain desired effect on the underlying dynamical system. The possible contexts in which this problem can be given a precise formulation are multiple, depending notably on the type of dynamical system considered, the degree of knowledge of the model that is assumed, and the goal that is desired. The actual and potential applications are also widespread, and very important, encompassing as they do such special cases as linear systems theory and optimal control. Here we shall focus primarily upon the stabilization issue for nonlinear systems of ordinary differential equations, a very active area in the last decade. Besides discussing some early milestones, we shall present more recent work showing how an optimization-based feedback design settles an old problem in the field by proving that every (open-loop) controllable system admits a stabilizing feedback.

Exhausters and Convexificators – New Tools in Nonsmooth Analysis

Vladimir F. DEMYANOV

Positively homogeneous (p.h.) functions play an outstanding role in Nonsmooth Analysis (NSA) and Nondifferentiable Optimization (NDO) since optimality conditions are usually expressed in terms of directional derivatives or their generalizations (the Dini and Hadamard upper and lower directional derivatives, the Clarke derivative, the Michel-Penot derivative etc.). All these derivatives are positively homogeneous functions of direction. In the convex case the directional derivative is convex (and p.h.) and, by the Minkowski duality, optimality conditions can be stated in geometric terms.

Attempts to reduce the problem of minimizing an arbitrary function to a sequence of convex problems were undertaken, among others, by B.N.Pschenichnyi, who introduced the notions of upper convex and lower concave approximations (u.c.a's and l.c.a.'s) and by F.Clarke who introduced generalized derivatives. A.M.Rubinov proposed to consider exhaustive families of upper convex and lower concave approximations. Recently some new tools - upper and lower exhausters and convexificators - closely related to exhaustive families of approximations were used. They represent dual objects and allow to reduce the original optimization problem to a sequence of convex optimization problems. A survey of some results related to these new tools is presented in the lecture. A characterization of the convexity and concavity of a p.h. function is given in terms of minimal convexificators . Representations of a p.h. function in terms of its minimal convexificators are also derived. Examples illustrating the theory are described.

Recent Advances in Deterministic Global Optimization

Christodoulos A. FLOUDAS

Global optimization addresses the computation and characterization of global optima (i.e., minima and maxima) of nonconvex functions constrained in a specified domain. Given an objective function f that is to be minimized and a set of equality and inequality constraints S, the area of *Deterministic Global Optimization* focuses on the following important issues :

(1) Determine a global minimum of the objective function f (i.e., f has the lowest possible value in S) subject to the set of constraints S;

(2) Determine *lower* and *upper bounds* on the global minimum of the objective function f on S that are valid for the whole feasible region S;

(3) Enclose all solutions of the set of equality and inequality constraints S.

In the last decade, we have experienced an explosive growth in the area of Global Optimization. This is complemented by the ubiquitous applications that include important problems such as the structure prediction in protein folding and peptide docking in computational chemistry and molecular biology, phase and chemical equilibrium in thermodynamics, pooling and blending in chemical refineries, robust stability analysis in control, parameter estimation and data reconcilliation, energy recovery systems in process synthesis, and multiple steady states in reactors.

In this plenary talk, we will discuss recent advances in the area of deterministic global optimization. We will focus on (i) constrained optimization problems that are twice-continuously differentiable, (ii) nonlinear and mixed integer optimization models, and (iii) systems with differential and algebraic constraints. The difference of convex functions method, denoted as $\alpha\mathbf{BB}$, which can address general continuous twice-differentiable problems that arise in a variety of engineering and applied science problems, will be presented. The $\alpha\mathbf{BB}$ (i) offers theoretical guarantee of attaining an ϵ-global optimum solution in a finite number of iterations, (ii) provides valid lower and upper bounds on the global solution, and (iii) identifies local optima close to the global minimum. Applications from a variety of disciplines will demonstrate the power of these recent advances.

Image Space Analysis and Maximum Principles

Franco GIANNESSI

A constrained extremum problem is considered in a Hilbert Space (HS), and it is shown how it can be associated with an Image Space (IS). Advantages and drawbacks of the analysis in the IS with respect to that in the HS are discussed with particular attention to the dimensionality of the problems, their regularization and their convexification. By introducing theorems of the alternative and separation theorems, the Lagrangian Theory of Multipliers is interpreted in the HS. The gradient of a separation hyperplane in the IS turns out to be the vector of Lagrange multipliers; this suggests extensions of the theory. In this light, the classic Lagrange Maximum Principle is rivisited, interpreted in geometrical terms, and generalized. Such an analysis has led to show that separation in the IS can be considered as a root for deriving most of the topics of the Theory of Constrained Extrema. Indeed, beside necessary optimality conditions, it is easy to derive existence necessary and sufficient conditions of Weierstrasstype, saddle-point sufficient optimality conditions. If a nonlinear separation is chosen - for instance, exponential one - , then a generalized saddle-point sufficiant condition can be defined and interpreted as a composition af a convexification operator and a linear separation. This ask for a classification in terms of analytical complexity (beside the existing numerical one). Subsequently, it is shown that also Lagrangian Duality Theory and Panalization Theory can be seen as a by-product of separation in the IS; in particular, as concerns the former, the duality gap receives a new geometrical interpretation which leads to generalizations, and, with regard to the latter, we meet an explanation of the fact that augmented penalty functions may be better than the Courant ones and hints for trying to characterize the problems for which the augmented ones are better. Starting again from separation arguments in the IS, it is shown how to derive Maximum Principles; in particular, Caratheodory and Pontryagin Principles are discussed. Within a couple of decades, Farkas, Caratheodory, Hahn, Banach - independently each other - got important ideas, which appear much different in the HS, while they turn out to have much in common if analysed in the IS. Such ideas have laid the foundations for the modern Theory of Constrained Extrema, in particular for the Variational Theory. The final part of this presentation is devoted precisely to some developments of the above mentioned ideas which have led to the Theory of Variational Inequalities. It ill be shown that the IS Analysis lets us find the connections between such a Theory and those mentioned above, and find new approaches and detect a new variational item.

Permanently Going Back and Forth Between the “Quadratic World” and the “Convexity World” in Optimization

Jean-Baptiste HIRIART-URRUTY and **Mounir TORKI**

Quadratic functions naturally appear and play an essential part in various Optimization and Control problems. For instance, the famous S-Procedure, which deals with the nonnegativity of a quadratic form on a set described by quadratic inequalities, provides a powerful tool for proving stability on nonlinear control systems. The *trust region subproblem* which consists in minimizing a quadratic function subject to a norm constraint arises in solving general nonlinear programs.

Some nonconvex quadratic optimization problems enjoy the remarkable property of having a zero duality gap. This phenomenon can be explained by the fact that, in some particular situations, quadratic forms have a *hidden convexity property.* This is closely related to results on the convexity of quadratic maps. The first results in this field are due to DINES (1941) and BRICKMAN (1961). Dines showed that, for any real matrices A, B, the following set

$$\{(\langle Ax, x\rangle, \langle Bx, x\rangle) \mid x \in \mathbb{R}^n\} \tag{1}$$

is a convex cone in $\mathbb{R}^2$, whatever the value of n. Brickman established a finer (and trickier) result: for any real matrices A, B,

$$\{(\langle Ax, x\rangle, \langle Bx, x\rangle) \mid x \in \mathbb{R}^n, \|x\|_2 = 1\} \tag{2}$$

is a convex compact set in $\mathbb{R}^2$, if $n \geq 3$.

These convexity results are closely related to alternative theorems on quadratic maps established by FINSLER (1936) and YUAN (1978). All these results have been extended to the case of m *generalized quadratic forms* where the (vector) variable $x \in \mathbb{R}^n$ is replaced by a matrix of specific size.

The objective of our work is twofold. First, we provide a chronological introduction to the different results concerning convexity of quadratic maps. Next, we aim to enlighten these results from the geometric point of view, to provide new and comprehensible proofs and to immerse them in a more general and abstract context.

We show that the convexity of sets (1) and (2) is essentially a consequence of the special facial structure of the cone of symmetric positive semidefinite matrices. More precisely, we exploit a characterization of the faces of the intersection of two convex sets in order to prove the convexity of the following set

$$\{(\langle a_1, x\rangle, ..., \langle a_m, x\rangle) \mid x \in F_m(K)\}$$

where $a_1, ..., a_m \in E, E$ is a finite dimensional space, K denotes a closed pointed cone in E and $F_m(K)$ is the union of all faces of K whose dimension at most m. Then, using a separation argument, we deduce a *"facial alternative theorem"* for linear mappings. In particular, when K is the cone of positive semidefinite matrices, we retrieve Dine's and Yuan's Theorem as well as their extensions.

Transversal Hypergraphs and Families of Polyhedral Cones

Leonid KHACHIYAN

We discuss the complexity of generating certain geometric configurations related to the classical theorems of Carathéodory and Helly, e.g., all minimal and/or maximal subfamilies of a given family of cones whose sums contain or avoid a prescribed vector. We also survey some complexity results for the closely related problem of generating all minimal transversals for a given hypergraph and some extensions of this problem.

1 Minimal/maximal conic families containing/avoiding a pre scribed vector

Let $\mathcal{K}$ be a finite set of polyhedral cones in $\Re^n$ and let $b \in \Re^n$ be a given vector such that $b \in \sum_{K\in\mathcal{K}} K$. Denote by $MIN(\mathcal{K}, b)$ the family of all minimal sets $\mathcal{S} \subseteq \mathcal{K}$ such that $b \in \sum_{K\in\mathcal{S}} K$. Note that by Carathéodory's theorem, the cardinality of any set $\mathcal{S} \in MIN(\mathcal{K}, b)$ does not exceed n. We assume that b and the cones $K \in \mathcal{K}$ are rational and discuss the complexity of generating $MIN(\mathcal{K}, b)$. We start with the special case where each cone $K \in \mathcal{K}$ is a ray.

Let $\mathcal{K} = \{K_1, \ldots, K_m\}$, where $K_i = cone\{a_i\}$ for given rational vectors $a_i \in \Re^n$. Suppose that the polyhedron $P = \{x \in \Re^m \mid \sum_{i=1}^m a_i x_i = b, \quad x = (x_1, \ldots, x_m) \geq 0\}$ is bounded. Then there is a natural one-to-one correspondence between $MIN(\mathcal{K}, b)$ and the vertex set of P. In particular, if P is the flow circulations polytope for a directed graph $G = (V, E)$ then $MIN(\mathcal{K}, b)$ can be identified with the set of all simple cycles in G. It is well known [6] that uniform sampling from the set of simple cycles of a digraph is a hard computational problem. In fact, using the amplification technique of [6], one can show the following result: *Let $K = 2^{m^{1-\varepsilon}}$, where ε is a positive constant. Suppose there exists a randomized polynomial-time algorithm which, given a rational polytope $P = \{x \in \Re^m | Ax = b,\ x \geq 0\}$, acts as a K-uniform generator on $VERT(P)$. Then any NP-complete problem can be solved in polynomial time by a randomized algorithm with arbitrarily small one-sided failure probability.*

Since random elements of $MIN(\mathcal{K}, b)$ are hard to generate, it is natural to ask whether all elements of $MIN(\mathcal{K}, b)$ can be enumerated by an *output-polynomial* algorithm, i.e., in time polynomial in $|MIN(\mathcal{K}, b)|$ and the input size. For the special case where $\mathcal{K}$ is a system of rays, we arrive at the well-known open question: *Can the vertices of a given rational polytope $P = \{x \in \Re^m \mid Ax = b,\ x \geq 0\}$ be enumerated in output-polynomial time?* For more general systems of cones, however, we have the following negative result:

Theorem 1. Suppose that each cone in $\mathcal{K} = \{K_1, \ldots, K_m\}$ has the form $K_i = cone(a_i, a_i', a_i'')$, $i = 1, \ldots, m$, where a_i, a_i', a_i'' are given rational vectors in $\Re^n$. Then, given a rational vector b and a collection of sets $\mathcal{M} \subseteq MIN(\mathcal{K}, b)$, it is coNP-hard to decide whether the given collection is exhaustive: $\mathcal{M} = MIN(\mathcal{K}, b)$.

Theorem 1 shows that unless $P = NP$, the construction of $MIN(\mathcal{K}, b)$ for trihedral cones cannot be realized in output-polynomial time. This is also true for the family $MAX(\mathcal{K}, b)$ of all maximal sets $\mathcal{S} \subseteq \mathcal{K}$ such that $b \notin \sum_{K \in \mathcal{S}} K$.

Theorem 2. Given a system of rational trihedral cones, a rational vector b and a collection $\mathcal{M} \subseteq MAX(\mathcal{K}, b)$, it is coNP-hard to determine whether $\mathcal{M} = MIN(\mathcal{K}, b)$.

2 Transversal hypergraphs and joint generation of $MIN(\mathcal{K}, b)$ and $MAX(\mathcal{K}, b)$

Even though constructing $MIN(\mathcal{K}, b)$ and $MAX(\mathcal{K}, b)$ is NP-hard, enumerating all elements of $MIN(\mathcal{K}, b) \cup MAX(\mathcal{K}, b)$ is equivalent to a simpler combinatorial problem known as *TRANSVERSAL HYPERGRAPH* (See e.g. [1,2,3,4,5].) Let V be a finite set and let $\mathcal{E}$ be a set family ($\equiv$ hypergraph) on V. A set $X \subseteq V$ is called a *transversal* to $\mathcal{E}$ if X meets all sets $E \in \mathcal{E}$. The hypergraph $\mathcal{E}^d$ consisting of all minimal transversals of $\mathcal{E}$ is called the *transversal* or *dual* hypergraph of $\mathcal{E}$. The transversal hypergraph problem can be stated as follows: *Given a hypergraph $\mathcal{E} \subseteq 2^V$ and a collection of minimal transversals $\mathcal{M} \subseteq \mathcal{E}^d$, either prove that $\mathcal{M} = \mathcal{E}^d$, or find a new transversal $X \in \mathcal{E}^d \setminus \mathcal{M}$.* Note that all the minimal transversals to $\mathcal{E}$ can be enumerated by initializing $\mathcal{M} = \emptyset$ and iteratively solving the transversal hypergraph problem $|\mathcal{E}^d| + 1$ times.

It is not known whether the transversal hypergraph problem can be solved in polynomial time for arbitrary input hypergraphs. This problem, however, is very unlikely to be NP-hard because it can be solved in *quasi-polynomial time* $O(|V|k) + k^{o(\log k)}$, where $k = |\mathcal{E}| + |\mathcal{M}|$ (See [4].) In addition, many classes of hypergraphs can be dualized in incremental polynomial time. (See e.g. references in [3,2].)

The transversal hypergraph problem can be easily embedded in the problem of constructing $MIN(\mathcal{K}, b) \cup MAX(\mathcal{K}, b)$ for an appropriate system of polyhedral cones $\mathcal{K}$. Theorem 3 below asserts that under some natural condition (always satisfied for rational polyhedral cones), constructing $MIN(\mathcal{K}, b) \cup MAX(\mathcal{K}, b)$ can in turn be reduced to the transversal hypergraph problem.

Theorem 3. Consider the following problem J: Given a conic system $\mathcal{K}$, a target vector b, and two set families $\mathcal{M} \subseteq MIN(\mathcal{K}, b)$ and $\mathcal{N} \subseteq MAX(\mathcal{K}, b)$, either find a new set $\mathcal{S} \in \left[MIN(\mathcal{K}, b) \setminus \mathcal{M}\right] \cup \left[MAX(\mathcal{K}, b) \setminus \mathcal{N}\right]$, or show that no such set exists. Suppose that for any input set $\mathcal{S} \subseteq \mathcal{K}$, we can determine in polynomial time whether or not $b \in \sum_{K \in \mathcal{S}} K$. Then problem $\mathcal{J}$ can be reduced in polynomial time to TRANSVERSAL HYPERGRAPH.

Theorem 3 is a special case of the following more general reduction: *Let $f : \{0,1\}^m \to \{0,1\}$ be a monotone Boolean function whose value at any point $x \in \{0,1\}^m$ can be evaluated in polynomial time. Then constructing the set of all minimal true and maximal false vectors of f reduces in polynomial time to the transversal hypergraph problem.* (See [1,5].)

We discuss other examples of transversal NP-hard conic families which are jointly reducible to the transversal hypergraph problem.

References

[1] J. C. Bioch and T. Ibaraki, Complexity of identification and dualization of positive Boolean functions, *Information and Computation* 123 (1995) 50-63.

[2] E. Boros, V. Gurvich, L. Khachiyan, and K. Makino, Generating partial and multiple transversals in a hypergraph, DIMACS Technical Report 1999-62, December 1999, to appear in ICALP 2000.

[3] T. Eiter and G. Gottlob, Identifying the minimal transversals of a hypergraph and related problems, *SIAM Journal on Computing*, 24 (1995) 1278-1304.

[4] M. L. Fredman and L. Khachiyan, On the complexity of dualization of monotone disjunctive normal forms. *Journal of Algorithms*, 21 (1996) 618-628.

[5] V. Gurvich and L. Khachiyan, On generating the irredundant conjunctive and disjunctive normal forms of monotone Boolean functions, *Discrete Applied Mathematics*, 1996-97, issue 1-3, (1999) 363-373.

[6] M. R. Jerrum, L. G. Valiant and V. V. Vazirani, Random generation of combinatorial structures from a uniform distribution *Theoretical Computer Science* 43 (1986) 169-188.

Semidefinite relaxations and Lagrangian duality with application to combinatorial optimization

François OUSTRY and Claude LEMARÉCHAL

We show that it is fruitful to dualize the integrality constraints in a combinatorial optimization problem. First, this reproduces the known SDP relaxations of the max-cut and max-stable problems. The approach can also be applied to general equality-constrained combinatorial problems. The resulting duality gap is smaller than with the classical Lagrangian relaxation, providing that linear constraints are treated carefully.

Generalized Duality in Variational Analysis

Stephen M. ROBINSON

This lecture deals with extensions of the duality structure familiar from convex optimization, to more general problems of variational analysis. We begin with a review of several duality structures that have already proven to be useful in various areas of optimization and equilibrium analysis. Then we show how many of these can be obtained as special cases of a general approach involving inverses of (possibly set-valued) mappings associated with the problems involved. This approach leads to a general procedure that we have elsewhere called *composition duality.*

Duality schemes provide most of their real usefulness through the interplay between the general duality principles and the special structure of a particular problem to which they are applied. We review some examples of this kind of usefulness, and introduce some additional ones suggested by the general duality scheme just mentioned.

Finally, we look at possibilities for extending these schemes to wider classes of problems within the general field of variational analysis.

Generalized envelope representations that propagate Fenchel conjugacy in Hamilton-Jacobi theory

R. T. ROCKAFELLAR

The Legendre-Fenchel transform sets up a one-to-one correspondence between lower semicontinuous proper convex functions g and their conjugate functions g*, as is well known. In Hamilton-Jacobi theory, such a function g at an initial (or terminal) time is propagated through dynamics of cost and control to a function at a general time t. Under basic convexity assumptions about the dynamics, that function at time t is likewise lower semicontinuous, proper and convex, and moreover it can be characterized as conjugate to g* with respect to an extension of the Legendre-Fenchel transform in which the inner product is replaced by a different expression, a certain kernel dependent on t. This kernel function itself propagates the inner product function and is the unique solution to a double Hamilton-Jacobi PDE posed in terms of subgradients.

Global Minimization and Parameter Estimation in Computational Biology

J. B. ROSEN, A. T. PHILLIPS, S. Y. OH and K. A. DILL

Two important problems in computational biology are the prediction of the structure of a protein, given its sequence of amino acids (the protein folding problem), and the docking of a ligand on a protein (drug design). Both of these problems require the computation of the global minimum(or near global minimum) of a differentiable function with many local minima. The Convex Global Underestimator (CGU) algorithm for solving this type of problem, has been developed and implemented, and successfully tested on a variety of structure prediction and docking problems. The function, representing the potential energy of a protein molecule in terms of its dihedral (backbone) angles(a vector Φ), has approximately $2n$ independent variables for a protein consisting of n amino acids. Also, the function contains parameters(α), which must be adjusted so that the structure (determined by the dihedral angles) at the global minimum corresponds to known protein molecular structures. A new algorithm has been developed and implemented which has successfully solved this problem for a number of small test molecules. A structural error function $\rho(\Phi, \alpha)$ is defined which measures the error between the global minimum Φ and the known native state value of Φ. This global minimum is a function of α, and the algorithm finds a minimum in the parameter space of the structural error function ρ. These algorithms, and computational results obtained will be summarized.

Abstract Convexity and Global Optimization

Alex M. RUBINOV

Special tools are required for examination and solving optimization problems. The main tool in the study of local optimization is a local approximation, which can be carried out by classical calculus and its modern generalizations. For global optimization we need to have some pieces of global information about an objective function and/or constraints and we need special tools to present this information in a convenient form. Abstract convexity provides some of these tools.

To explain the idea behind abstract convexity, recall that the main tool in the study of convex (local and simultaneously global) optimization is the subdifferential. First, the subdifferential $\partial f(x)$ accomplishes a local approximation of a convex function f in a neighbourhood of a point x, second, $\partial f(x)$ is a tool for construction of a global affine support of f at x, that is, an affine function, which does not exceed f over the entire space and coincides with f at x. Generalizations of these two traits of the convex subdifferential lead to the development of *nonsmooth analysis* and *abstract convexity*, respectively.

The existence of global affine supports follows from the separation theorem for convex sets. This theorem also leads to the following well-known result from convex analysis: each lower semicontinuous convex function f is the upper envelope of the set of all affine (linear + constant) functions which are minored by f. Thus, the convexity can be presented as a combination of linearity and the envelope presentation.

In some instances the use of the envelope representation is very convenient, even when we consider the upper envelope of non-affine functions. Thus the idea arises to examine main notions of convexity in non-convex environments. First, to examine *abstract convex* functions, that is, functions which can be presented as upper envelopes of subsets of a given set of not necessarily linear *elementary* functions. Second, to study *abstract convex sets*, that is, sets A which enjoy the following not necessarily linear separation property: each point, which does not belong to A, can be separated from A by an elementary function. This idea underlies the rich theory of *convexity without linearity*, known also as *abstract convexity* (see [1, 2, 3]).

For applications of abstract convexity, we need to describe the main objects of this theory in some concrete situations. Just as the set of linear function leads to the theory of convex functions and convex sets, each class of elementary functions should lead to the theory of abstract convex functions and abstract convex sets with respect to this class.

The present talk consists of two parts. In the first part we describe abstract convex functions and sets, and subdifferentials for some concrete sets of elementary functions. We mainly discuss abstract convexity based on classes of elementary functions, consisting of the so-called min-type functions. Instead of linear functions $x \mapsto \sum_{i=1}^{n} l_i x_i$ defined on $\mathbb{R}^n$, we consider functions of the form $x \mapsto \min_{i:l_i>0} l_i x_i$ defined on the cone $\mathbb{R}^n_+$ of vectors with nonnegative coordinates and some kindred classes of elementary functions. Abstract convexity based on these classes leads to the theory of various classes of *monotone functions* and to the theory of *normal* sets. The totality of developed approaches and results obtained in this direction can be considered as certain elements of *monotonic analysis*. It is interesting to note that monotonic analysis is even simpler than convex analysis in some aspects.

Min-type functions of a general form: $l(x) = \min_{i=1,\dots,k}[l_i, x]$, where k is equal to either n or $n+1$, lead to the theory of *star-shaped* and *co-star-shaped* subsets of R^n and to the theory of some classes functions, defined on $\mathbb{R}^n$, in particular *positively homogeneous functions of degree one.* (Here $[l, x]$ stands for the inner product of vectors l and x.) The totality of developed approaches and results obtained in this direction can be considered as certain elements of *star-shaped analysis.*

In the second part of the talk we focus on numerical methods of global optimization based on abstract convexity. We describe conceptual schemes of some numerical methods and possible implementation of these schemes. In particular we consider the so-called *cutting angle method.* We report the results of numerical experiments, which show that these methods are fairly efficient for solving some classes of problems.

References

[1] Pallaschke, D. and Rolewicz, S., *Foundations of Mathematical Optimization (Convex analysis without linearity)*, Kluwer Academic Publishers, Dordrecht (1997).

[2] Rubinov, A.M. *Abstract Convexity and Global Optimization*, Kluwer Academic Publishers, to appear.

[3] Singer, I., *Abstract Convex Analysis.* Wiley-Interscience Publication, New York (1997).

Role of superfluous constraints for improving dual quadratic estimates in combinatorial and multiextremal problems

Naum Z. SHOR

In this lecture we describe the new methods of finding the dual Lagrangian estimates in multiextremal polynomial-type problems which are based on nondifferential optimization algorithms with transformation of argument space. The mentioned classes of problems can be reduced by introducing additional variables and constraints to quadratic type problems which are as a rule nonconvex. Without loss of generality one can represent the generated quadratic-type problem as

$$\text{to find} \quad Q^* = \inf_{x \in X} Q_0(x)$$

$$\text{subject to} \quad Q_i(x) = 0, i = 1, \dots, m,$$

where $X \subseteq E^n$ is a simple in some sense set.

Let $u = \{u_1, \dots, u_m\}$ be the vector of Lagrange multipliers and $L(x, u) = Q_0(x) + \sum_{i=1}^m u_i Q_i(x)$ the Lugrangian function, $\psi(u) = \inf_{x \in X} L(x, u)$. If $\psi(u)$ is a proper concave function then for each $\overline{u} \in \ dom\psi$ one obtains a lower bound $\psi(\overline{u}) \leq Q^*$. To calculate the best dual lower bound $\psi^* = \sup_{u \in \, dom\psi} \psi(u)$ we must use the methods of nonsmooth optimization because *int dom*ψ is a convex set of interior points which is determined by nonsmooth inequality $(\lambda_{\min} K_0(\overline{u}) > 0)$, where $K(\overline{u})$ is a quadratic form in x of $L(x, \overline{u})$, $\lambda_{\min}$ is the minimal eigenvalue of $K(\overline{u})$.

To improve the dual bounds author proposed to modify the Lagrange function by introducing the new superfluous quadratic constraints which are consequences of the given ones. The examples are given which show essential improving of dual estimates by use of superfluous constraints in extremal graph problems and polynomial optimization. For finding dual lower bounds we use certain modifications of subgrafient-type methods with space dilation in the direction of difference of two successive subgradients (see N.Shor, Nondifferentiable optimization and polynomial problems, 1998, Kluwer Academic Publishers).

The results of some numerical experiments are considered.

Numerical Analysis of Sparse Polynomial Systems of Equations

Stephen SMALE

Suppose that a polynomial system of equations is structured to have a large predetermined subset of coefficients zero. This situation is frequently encountered in pure and applied math. How does this structure affect the algorithms, geometry, and complexity of equation solving? Our talk will deal with this question.

Monotonic Optimization: Problems and Solution Aproaches

Hoang TUY

By their very nature, many functions encountered in mathematical modeling of real world systems in a broad range of activities,including economics and engineering, exhibit monotonicity with respect to some variables (partial monotonicity) or to all variables (total monotonicity). The analysis of monotonicity for the purpose of applications to engineering design problems has been previously explored by a number of authors who focussed on finding constraints that must be tight at the optimum, in order to lower the dimension of the problem and reduce it to a form more amenable to effective solution.

A concept closely related to monotonicity is that of *normal set* which was introduced in mathematical economics in connection with the analysis of production activities within an economic system. A normal set is any set $G \subset R^n_+$ such that whenever $a \in G$ then any x such that $0 \leq x \leq a$ also belongs to G. Just as convex sets are essentially lower level sets of quasiconvex functions, normal sets are essentially lower level sets of *increasing functions*, i.e. functions $f(x)$ defined on R^n_+ such that $f(x') \geq f(x)$ whenever $x' \geq x \in R^n_+$.

From the point of view of numerical optimization, a fundamental property of increasing functions is that when seeking a minimizer of such a function over a constraint set, once a feasible solution x^0 is known then all solutions in the cone $x^0 + R^n_+$ can be omitted because no better solution than x^0 can be found in the latter. Such kind of information is very useful and may often help simplify the problem drastically by limiting the search process to a restricted region. Likewise, if the constraint set is a normal set then any infeasible solution z^0 can be strictly separated from the constraint set by a cone $x^0 + R^n_+$, where $x^0 \leq z^0$ is some suitable feasible solution. It is well known that the classical separation property of convex sets is fundamental for many solution strategies in convex and nonconvex optimization. This suggests that the above-mentioned properties of increasing functions and normal sets should play an equally important role in the analysis and solution of monotonic optimization problems.

In this paper we present a theory of monotonic inequalities and monotonic optimization based on a systematic exploitation of these properties. We shall first show that any closed normal set is the intersection of a decreasing sequence of elementary normal sets, called *polyblocks*, so that any normal set can be approximated as closely as desired by an outer polyblock. This approximation of normal sets by polyblocks is similar to the approximation of convex sets by polyhedrons and serves the basis for an analytical characterization of the structure of the solution set to an arbitrary monotonic system. In turn this characterization leads to a general approach for studying monotonic inequalities and solving monotonic optimization problems, and more generally, optimization problems involving *differences of increasing functions.*

Since a host of functions including arbitrary polynomials of several variables, can be represented as differences of increasing functions on the nonnegative orthant, this approach has a wide range of applicability. Furthermore, the characterization of the solutions to a monotonic system as sequences of vectors which can be generated according to simple rules opens the way to hybrid approaches combining nondeterministic methods (simulated algorithm, tabu search, genetic algorithm) with deterministic ones for handling many global optimization problems so far resistant to existing methods.

Confidence Level Solutions for Stochastic Programming

Yu. NESTEROV and Jean-Philippe VIAL

We consider the stochastic programming problem

$$\min_x \{\phi(x) \equiv E_{\xi\in\Xi}(f(x,\xi)) : \ x \in Q\}, \tag{1}$$

where Ξ is a probability space endowed with a probability measure, and Q is subset of R^n. We are concerned with the evaluation of the computational effort to solve the problem. The problem is inherently difficult. Indeed, for general distributions, the computation of the exact value of the expectation is not possible; moreover, the computation of an approximate value with precision ϵ is of order ϵ^{-m}, where m is the dimension of the probability space Ξ. Optimizing a function that, complexity-wise, is in itself difficult to evaluate is thus a challenging issue.

Let us mention two standard approaches. One consists in discretizing the probability distribution, and writing (1) as a deterministic mathematical program. Obviously, this approach is not workable on a multi-dimensional probability space. Another approach consists in performing Monte-Carlo sampling of the random variable ξ. Again, one can build a deterministic mathematical programming problem based on that finite sample. Using statistical theory, one can sometimes give confidence intervals on the optimal value and analyze statistical convergence. However, no complexity estimate seems to be available yet.

We propose an alternative approach to stochastic programming based on Monte-Carlo sampling and stochastic gradient optimization. The candidate solution is the average of the x's generated by the process. Under appropriate convexity and boundedness assumptions we are able to show that the expected value of the function f evaluated at the candidate solution converges to the optimal value of (1). We further give a complexity estimate for this convergence.

Actually, the candidate solution is itself a random variable, since it is generated by a stochastic algorithm. The objective function is thus doubly random, and the expectation in our complexity measure is taken over all possible samples of a given size. Decision-makers are not interested in this average performance. Rather, they wish to know what is the quality of a given candidate solution resulting from a given sample. Thus we introduce the notion of confidence level of that solution. It is measured by the probability that the expected value of the objective associated with that solution departs from the optimal value of (1). We show that it is possible to guarantee a given confidence level, and we give a complexity estimate to achieve this level.

Finally, we compare two ways to improve confidence levels: either by increasing the sample size, or by performing several independent samples. We give complexity estimates for the two approaches and give evidence that repeating independent experiments might be significantly more efficient than prolonging a unique experiment

CONTRIBUTED TALKS

Compound Methods for Approximating Toeplitz Matrix

Suliman AL-HOMIDAN

Hybrid methods for minimizing least distance functions with semi-definite matrix constraints are considered. One approach is to formulate the problem as a constrained least distance problem in which the constraint is the intersection of three convex sets. The Dykstra-Han projection algorithm can then be used to solve the problem. This method is globally convergent but the rate of convergence is slow. However, the method does have the capability of determining the correct rank of the solution matrix, and this can be done in relatively few iterations. If the correct rank of the solution matrix is known, it is shown how to formulate the problem as a smooth nonlinear minimization problem, for which a rapid convergence can be obtained by l_1SQP method. Also this paper studies compound method that attempt to combine the best features of both types of methods. An important feature concerns the interfacing of the component methods. Thus, it has to be decided which method to use first, and when to switch between methods. Difficulties such as these are addressed in the paper. Comparative numerical results are also reported.

References

[1] Dykstra, R. L. [1983]. An algorithm for restricted least squares regression, *J. Amer. Stat. Assoc. 78, pp. 839–842.*

[2] Han, S. P. A successive projection method, *Math. Programming, 40, pp. 1–14.*

Strengthened Semidefinite Relaxations for the Max-Cut Problem

Miguel F. ANJOS and Henry WOLKOWICZ

We present two strengthened semidefinite programming (SDP) relaxations for the max-cut problem (MC) which improve on the well-known SDP relaxation studied by Goemans and Williamson. The first relaxation is obtained by adding nonlinear constraints to the well-known relaxation and then lifting the resulting nonlinear program using Lagrangian duality. The resulting SDP strictly improves the well-known SDP bound.

The second SDP relaxation is a further strengthening of the lifted relaxation previously derived. It inherits all its helpful properties and, furthermore, is a strict improvement over the SDP relaxation obtained by adding all the triangle inequalities to the well-known SDP relaxation.

The objective function and the equality constraints of both strengthened relaxations are sparse and structured. This permits the development of specialized algorithms for these problems. We discuss such algorithms and present numerical results.

Saddle Point Methods for Evolution Equations

Giles AUCHMUTY

This talk will describe the use of saddle point theorems to prove existence - uniqueness results for evolution equations. We first describe certain saddle point functionals associated with evolution equations that involve the subdifferential of a convex functional. Under standard assumptions on these functionals, it is shown how saddle point theorems can be applied to obtain existence uniqueness results of weak solutions.

The results will be illustrated by some specific examples. For ordinary differential equations, this leads to uniqueness theorems without requiring local Lipschitz condituity of the system. For nonlinear parabolic equations this saddle point formulation again leads to results under different hypotheses, and in different spaces, than those usually used.

Cutting angle method for global optimization

A. M. BAGIROV and A. M. RUBINOV

We study the following problem of global optimization:

$$f(x) \longrightarrow \min \text{ subject to } x \in X \tag{1}$$

where a function f in general is assumed to be Lipschitz continuous and $X \subset \mathbb{R}^n$ is compact set describing by equality and inequality constraints. Cutting angle method for solution such problems was introduced and studied in [1, 3]. This method is reduced to the solution of the following auxiliary problem:

$$h(x) \longrightarrow \min \text{ subject to } x \in X \tag{2}$$

where

$$h(x) = \max_{k \le j} (\min_{i \in I(l^k)} l_i^k x_i - c_k),$$

$$l^k = (l_1^k, \dots, l_n^k) \in \mathbb{R}_+^n, \ I(l^k) = \{i : l_i^k > 0\}, \ c_k \in \mathbb{R}, \ k = 1, \dots, j, \ j \ge n.$$

The latter problem is essentially of combinatorial nature. In this talk a survey of various algorithms for the solution of the subproblem (2) is presented. Firstly we assume that the function f is increasing positively homogeneous and the set X is the unit simplex. In this case $c_k = 0$ for all $k \le j$ and therefore we get the following problem

$$h(x) = \max_{k \le j} \min_{i \in I(l^k)} l_i^k x_i \to \min \quad \text{subject to} \quad x \in X. \tag{3}$$

Some algorithms for solving (3) were proposed in [1, 2]. We develop also new algorithms for the solution of this problem. The comparison of algorithms based on the results of numerical experiments are given.

It follows from the results of [3], that the minimization of a Lipschitz function over the unit simplex can be transformed to the minimization of an increasing positively homogeneous function. Using some transformation of variables and penalization we can present the minimization of a Lipschitz function subject to linear (and even some nonlinear) constraints as the minimization of a Lipschitz function over the simplex. We suggest and study some versions of the cutting angle method for the solution of such problems.

A number numerical experiments have been carried out using problems with both increasing positively homogeneous and Lipschitz continuous objective functions. We present results of numerical experiments and discuss them.

References

[1] M. Yu. Andramonov, A. M. Rubinov and B. M. Glover, Cutting angle methods in global optimization, *Applied Mathematics Letters,* 1999, No. 12, pp. 95-100.

[2] A.M. Bagirov and A.M. Rubinov, Global minimization of increasing positively homogeneous functions over the unit simplex. *Annals of Operations Research,* to appear.

[3] A. M. Rubinov and M. Yu. Andramonov, Lipschitz programming via increasing convex-along-rays functions. *Optimization Methods and Software*, 1999, Vol. 10, pp. 763-781.

Differences of Convex Compact Sets: II. Visualization of Directed Sets

Robert BAIER and Elza FARKHI

Differences of convex compact nonempty sets can be embedded in the space of directed sets. For directed sets, generalizations of the Minkowski addition and multiplication with non-negative scalars are available. This new space forms a normed vector space, is complete with the respect of a norm and a vector lattice with respect to an order. The visualization of elements from the closure of the linear hull of all embedded convex compacts, resulting in a Banach and a Riesz space, are studied.

A n-dimensional directed set $\overrightarrow{A}$ is a family of pairs consisting of a $(n-1)$-dimensional directed set $\overrightarrow{A_{n-1}(l)}$ and a scalar value $a_n(l)$, parametrized by unit directions $l \in S_{n-1}$. The description is a generalization of describing convex compacts with the $(n-1)$-dimensional supporting face $Y(l, A)$ and the value of the support function in directions $l \in S_{n-1}$. The basis for this recursive approach is formed by generalized intervals which are well-known in interval analysis.

Special properties of the embedding $J_n(\cdot)$ are reviewed (positive linearity, preservation of the order and addition resp. non-negative scalar multiplication, injectivity, continuity with respect to a new metric, ...). These results will be explained in details in another talk at the conference given by the second author. This talk will focus on the visualization of directed sets, especially on embedded convex compact sets and their differences, and on theoretical properties of this visualization.

The visualization of directed intervals, the one-dimensional directed sets, are presented by attaching outer or inner normals to the end-points to indicate proper intervals (convex sets) or improper ones. This approach is closely connected to generalized intervals introduced by Kaucher and Markov. Simple properties of the boundary and the visualization map are given showing first connections and distinctions to known differences of convex compact intervals.

The visualization set and the boundary are generalized to the n-dimensional case $(n \geq 2)$. Here, the visualization consists not only of a convex and a concave part, additionally, a mixed-type part could appear. The convex and the concave part are both convex sets and are determined only by the scalar function $a_n(\cdot)$, the second component of the directed set. The boundary is the reprojection of the visualizations of all $(n-1)$-dimensional directed sets $\overrightarrow{A_{n-1}(l)}$ being the first component of the directed set. The mixed-type part is formed by all boundary points which are not contained in the boundary of the convex or the concave part. The difference of two embedded convex compacts could yield in visualizations consisting only of the mixed-type part. For each boundary point, the corresponding orientation bundle is defined as the set of all directions l, for which the point lies in the reprojected visualization of the $(n-1)$-dimensional directed set $\overrightarrow{A_{n-1}(l)}$. The orientation bundle serves for a distinction, whether the visualization part contains the convex (outer normals) or the concave part (inner normals).

Results about the visualization, the boundary and the orientation bundle of embedded convex compacts and their inverses are shown. Connections to other differences, i.e. the geometric difference and the Demyanov difference, will be proven. The main result shows that the geometric difference $A \stackrel{*}{-} B$ is always a subset of the nonempty visualization $V_n(J_n(A) - J_n(B))$ of the difference of two embedded convex compacts which itself lies in the Demyanov difference $A \dot{-} B$. The convex part of the visualization equals to the geometric difference $A \stackrel{*}{-} B$, the concave part is the pointwise negative of the geometric difference $B \stackrel{*}{-} A$.

Applications to summands and equivalent pairs of convex compact sets are stated. Necessary and sufficient conditions for directed sets which characterize embedded convex compacts resp. its inverses are studied with the help of the visualization. Some examples and computed visualizations will be shown to exploit the theoretical results.

On the correct formulation of the optimal control problem in relation to the work of Caratheodory

John F. BARRET

In a paper published in 1997, the well known control theorists Sussman and Willems put torward the view that the control theory approach to dynamics is more general than the classical one and that the so-called control Hamiltonian, as defined in connection with Pontryagin's Maximum Principle, should displace the known Hamiltonian of classical dynamics. These authors further considered that doing so leads to a new view of the classical calculus of variations.

An interesting point is that these authors traced back what they call the control approach to the calculus of variations historically to the Brachistochrone problem of John Bernoulli in 1696 which they described at length. They however omitted to mention that it was exactly this problem which was the starting point of Caratheodory's approach in the calculus of variations and that it was he who was the first to generalize the ideas inherent in this problem. The connection of the principle of optimality with the Bernoulli problem had been noted in a previous paper by Pesch and Bulirsch 1994 who described the Caratheodory theory as applied to the classical theory and discussed its relation to the Pontryagin Maximum Principle and Dynamic Programming. As those authors observed, the priority of Caratheodory is not widely acknowledged.

In the present talk we wish to trace in further detail certain aspects of the relation between the Carathedory theory and optimal control and indicate the possibility of a close rapprochment between classical dynamics and optimal control dynamics where the relationship between the classical Hamiltonian and the control Hamiltonian is seen in its proper perspective.

References

[1] H. J. Sussman and J. C. Willems: IEEE Control Systems 17 1997 32-44.

[2] H. J. Pesch and R. Bulirsch: J.O.T.A. 80 1994 199-225.

[3] C. Caratheodory: Ges. Math. Schr., Munchen 1954-7 (Beck)

[4] J. F. Barret: Caratheodory Symposium, Athens 1973.

Concerning Types of Hausdorff-Discontinuity for Convex Polynomial Mappings

Evgeny BELOUSOV

We study the continuity properties of convex polynomial mappings (CP-mappings), i.e. mappings of the form

$$A(\lambda) = \{\mathbf{x} \in E_n \mid f_i(\mathbf{x}) \le \lambda_i,\ i = 1, \ldots, m\}, \qquad \lambda \in \Lambda \tag{1}$$

$$\Lambda = \{\lambda \in E_m \mid A(\lambda) \neq \emptyset\},$$

where $f_i,\ i = 1, \ldots, m$, is fixed system of convex polynomials and Λ is effective domain of $A(\lambda)$.

For arbitrary sets $A, B \subseteq E_n$ we introduce the following notation:

$$\rho(A, B) = \sup_{x \in A} \inf_{y \in B} |\mathbf{x} - \mathbf{y}| \qquad \text{is the excess of A with respect to B.}$$

We say [1, p.166] that mapping $A(\lambda)$ is quasi-bounded if

$$\rho(A(\lambda), A(\lambda')) < +\infty \qquad \forall\ \lambda, \lambda' \in \Lambda.$$

Fix an arbitrary point $\omega \in \Lambda$ and consider the following functions defined on Λ :

$$d_\omega(\lambda) = \rho(A(\omega), A(\lambda)), \qquad d^\omega(\lambda) = \rho(A(\lambda), A(\omega)) \quad \lambda \in \Lambda.$$

The functions d_ω *and* d^ω may take infinite values at some points of Λ. For this reason, we define two sets on which these functions are bounded:

$$\Lambda_\omega = \{\lambda \in \Lambda \mid d_\omega(\lambda) < +\infty\}, \qquad \Lambda^\omega = \{\lambda \in \Lambda \mid d^\omega(\lambda) < +\infty\}$$

and two numerical sets of partial limits of d_ω *and* d^ω :

$$B_\omega = \{b \in E_1 \mid \exists \lambda^k \in \Lambda_\omega,\ \lambda^k \neq \omega,\ \lambda^k \to \omega : \ \ d_\omega(\lambda^k) \to b\},$$

$$B^\omega = \{b \in E_1 \mid \exists \lambda^k \in \Lambda^\omega,\ \lambda^k \neq \omega,\ \lambda^k \to \omega : \ \ d^\omega(\lambda^k) \to b\}.$$

These sets, structure of this sets is main subject of our investigations. If mapping $A(\lambda)$ is Hausdorff-continuous, then the sets B_ω *and* B^ω coincide and contain only number zero.

As shown in [2], set B_ω is convex and closed for a convex and closed mapping. Hence, this set is either a segment $0 \le b \le b_0$ (finite discontinuity from below), or a ray $b \ge 0$ (infinite discontinuity from below), or it just coincides with the number 0 (in this case mappings $A(\lambda)$ is Hausdorff semicontinuous from below at the point ω). By costructing examples, we have shown in [2] that each of those alternatives could actually take place for quasi-bounded CP-mappings as well.

These examples cannot be too simple, because it is well known that CP-mappings are Hausdorff-continuous at all points of Λ if $f_i,\ i = 1, \ldots, m$ in (1) are convex polynomials whose degrees are no greater than four [3], or convex polynomials in no more than three variables [1, p.160].

In contrast to set B_ω set B^ω can have quite complicated structure. We have the following hypotheses: for quasi-bounded CP-mapping set B^ω is an union of finite set of positive numbers und zero.

In [1, p.160], examples were initially constracted that show that quasi-bounded CP-mappings may not be Hausdorff-semicontinuous from above at points of $\Lambda \backslash ri\Lambda$ and for which set B^ω consist of two points.

Note that the minimum degree of convex polynomials $f_i,\ i=1,\dots,m$, for which CP-mapping $A(\lambda)$ of form (1) is not Hausdorff-continuous (i.e., for which the functions d_ω *and* d^ω are not continuous) is not known thus far.

References

[1] Belousov E.G. and Andronov V.G. Solvability and Stability of Polynomial Programming Problems (in Russian), Moscow University Publisher, 1993.

[2] Andronov V.G. and Belousov E.G. On the Continuity of Convex Polynomial Quasi-Bounded Mappings Computational Mathematics and Mathematical Physics. Vol.39, No.6, 1999, p.p.906-918.

[3] Andronov V.G. and Belousov E.G. On the Continuity of the Optimal Mappings of the Convex Programming Problem (in Russian) Vestnik Mosk. Univ., Ser.15: Vychisl. Mat. and Kibern., 1995, No.2, p.p.18-22.

[4] Belousov E.G. Introduction to Convex Analysis and Integer Programming (in Russian), Moscow University Publisher, 1977.

Typical Convex Sets

Efim BRONSHTEIN

Let B be a compact convex subset of Hilbert space. The set $\Psi(B)$ of closed convex subsets of B with Hausdorff metric is considered. Typical subset of metric compact is the set of the second Baire category. V.Klee [1] proved that the set of extreme points of typical element of $\Psi(B)$ is dense if compact B is infinite- dimensional. T.Schwarz and T.Zamfirescu considered the convex cone consisting of convex compact subsets of space R^n and established that the set of extreme points of a typical element of this cone is typical. It has appeared, that this property has general character. Some other properties of typical convex sets are established also. The following result is proved.

THEOREM 1. Let compact K be infinite-dimensional. Subset of $\Psi(B)$ consisting of sets W with properties

1. codimension of set W in B is equal to 0;
2. the set of extreme points $extW$ is typical in W;
3. set W nowhere dense in B;
4. every nonzero linear functional reaches a maximum in unique point of W;
5. no more then one linear functional with norm 1 reaches a maximum in arbitrary point of W

is typical in $\Psi(B)$.

It is interesting to compare this statement to its analogue for finite–dimensional B.

THEOREM 2. Let compact K be finite-dimensional. Subset of $\Psi(B)$ consisting of sets W with properties

1. a dimension of W is equal to the dimension of B;
2. a set W is strictly convex;
3. a boundary of W is smooth

is typical in $\Psi(B)$.

Analogue of the statement 2 of theorem 1 for locally compact convex cones is given in [3].

References

[1] Klee V. Some new results on smoothness and rotundity in normed linear spaces. Math.Ann., **159**:51-63, (1959).

[2] Schwarz T., Zamfirescu T. Typical sets of convex sets. J. Austr. Math. Soc. (Series A), **43**:287-290,(1987).

[3] Bronshtein E.M. On convex cones with Schwarz–Zamfirescu property. Funk. anal. appl, **32**(3):57-59,(1998).

A Proximal Point Method for the Variational Inequality Problem in Banach Spaces

Regina Sandra BURACHIK and Susana SCHEIMBERG

In this paper we prove welldefinedness and weak convergence of the generalized proximal point method when applied to the variational inequality problem in reflexive Banach spaces. The proximal version we consider makes use of Bregman functions, whose original definition for finite dimensional spaces has been here properly extended to our more general framework.

Let B be a reflexive Banach space and $\Omega \subset B$ a nonempty closed and convex set. Given $T: B \to \mathcal{P}(B^*)$ a maximal monotone operator, we consider the *classical variational inequality problem* for T and Ω, $VIP(T, \Omega)$, defined by:

Find $x^* \in \Omega$ such that there exists $y^* \in T(x^*)$ with

$$\langle y^*, x - x^* \rangle \geq 0, \tag{1}$$

for all $x \in \Omega$, where $\langle \cdot, \cdot \rangle$ stands for the dual product in B.

In the particular case in which T is a subdifferential, i.e., $T = \partial\varphi$, where $\varphi : B \to \mathbb{R} \cup \{+\infty\}$ is a proper, convex and lower semicontinuous functional, (1) reduces to the nonsmooth constrained optimization problem:

$$\min_{x \in \Omega} \varphi(x). \tag{2}$$

The following natural hypothesis will be considered in our study:

H_1: $D(T)^0 \cap \Omega \neq \emptyset$ or $D(T) \cap \Omega^0 \neq \emptyset$.

Our aim is to study the algorithm given by:

1) Take $x^0 \in \Omega \cap D(T)$.

2) Given x^k, define x^{k+1} by the inclusion

$$0 \in (T + N_\Omega + \lambda_k \nabla f)x^{k+1} - \lambda_k \nabla f(x^k), \tag{3}$$

where f is a Bregman function (see [Bre67]), $\lambda_k > 0$, and $N_\Omega(\cdot)$ is the normality operator associated to Ω

3) If $x^{k+1} = x^k$, STOP.

It is easy to see that when $B = \Omega = H$, for H a Hilbert space, and $f = \frac{1}{2}\|\cdot\|^2$, algorithm (3) reduces to the well-known proximal point method in a Hilbert space. The fundamental properties of this method (like weak convergence to a solution, linear convergence rate, etc.) have been established in [Roc76]. In our work we consider the concept of a Bregman function in a Banach space. This notion has its origin in [Bre67] and has been extensively used for convex optimization algorithms and extensions of the proximal point method for finite dimensional spaces (e.g. [DePI86], [CDePEHI90], [Kiw97a], [CZ92], [CT93], [E93], [Ius95], [Kiw97b]). For the infinite dimensional case, Bregman distances have been used in Hilbert spaces in [BurI97], and in Banach spaces in [BI97a], [BI97b], and [BBI97]. A useful tool for comparing the Bregman distance with the distance induced by the norm of the Banach space, leads to the notions of *local modulus of convexity* and *total convexity* of f, both introduced in [BI97a]. We take profit of these new concepts for defining Bregman distances for which the corresponding generalized proximal method converges under minimal hypotheses. It is well known that for the generalized proximal method, strong assumptions must be made on the operator for assuring convergence. For these family of Bregman distances, no extra assumtions on T are required. On the other side, when the distance does not fulfill these technical conditions, we prove convergence for para- and pseudomonotone T, (see [Ius] and [Bro76] for these definitions). The problem of establishing Bregman functions satisfying the assumptions that imply convergence for a general T have been solved recently in [IO].

References

[Bre67] L.M, Bregman. The relaxation method of finding the common points of convex sets and its application to the solution of problems in convex programming. *USSR Computational Mathematics and Mathematical Physics*, 7(3):200–217, 1967.

[BurI97] R.S. Burachik, and A. N. Iusem. Generalized proximal point methods for the variational inequality problem in a Hilbert space. *SIAM Journal on Control* 8:197–216, 1998.

[BI97a] D. Butnariu, A.N. Iusem. Local moduli of convexity and their applications to finding almost common points of measurable families of operators. *Recent developments in Optimization Theory and Nonlinear Analysis, edited by Y. Censor and S. Reich*, *AMS Contemporary Mathematics* 204:61–91, 1997.

[BI97b] D. Butnariu, A.N. Iusem. On a proximal point method for convex optimization in Banach spaces. Submitted, 1997.

[BBI97] R.S. Burachik, D. Butnariu, A.N. Iusem Iterative methods for solving stochastic convexfeasibility problems and applications. (to appear), 1998.

[Bro76] F. E. Browder. Nonlinear operators and nonlinear equations of evolution in Banach spaces. *Proceedings of Symposia in Pure Mathematics, American Mathematical Society*, 18, 2, 1976.

[CDePEHI90] Y. Censor, A. De Pierro, T. Elfving, G.T. Herman, A. N. Iusem. On iterative methods for linearly constrained entropy maximization Accepted for publication *Numerical Analysis and Mathematical Modelling* (A. Waculitz, editor). *Banach Center Publication Series,* Banach Center, Warsaw, 24:145–163, 1990.

[CZ92] Y. Censor, S.A. Zenios. The proximal minimization algorithm with with D-functions. *Journal of Optimization theory and Applications*, 73:451–464, 1992.

[CT93] G. Chen, and M. Teboulle, Convergence analysis of a proximal-like optimization algorithm using Bregman functions. *SIAM Journal on Optimization* 3:538–543, 1993.

[DePI86] A. De Pierro, A. N. Iusem. A relaxed version of Bregman's method for convex programming. *Journal of Optimization theory and Applications*, 51:421–440, 1986.

[E93] J. Eckstein. Nonlinear proximal point algorithms using Bregman functions, with applications to convex programming. *Mathematics of Operations Research*, 18:202–226, 1993.

[Ius] Iusem,A.N. , On some properties of paramonotone operators (to appear).

[Ius95] Iusem, A.N.. On some properties of generalized proximal point methods for quadratic and linear programming *Journal of Optimization Theory and Applications*, 85:593–612, 1995.

[IO] Iusem, A.N., R. Gárciga Otero. , An augmented Lagrangian method in Banach spaces with inexact solution of the subproblems Submitted.

[Kiw97a] K. Kiwiel. Free-steering relaxation methods for problems with strictly convex costs and linear constraints. *Mathematics of Operations Research*, 22:326–349, 1997.

[Kiw97b] K. Kiwiel. Proximal minimization methods with generalized Bregman functions. *SIAM Journal on Control and Optimization*, 35:1142–1168, 1997.

[Roc76] Rockafellar, R.T. Monotone operators and the proximal point algorithm. *SIAM Journal on Control and Optimization*, 14:877–898, 1976.

Variational inequalities and regularity properties of closed sets in infinite dimensional spaces

Giovanni COLOMBO and Vladimir V. GONCHAROV

Let H be a Hilbert space, and let $K \subset H$ be a closed set. It is well known that an equivalent condition for the convexity of K is the (monotonicity) relation

$$\langle z - x, y - x \rangle \leq 0 \,, \tag{1}$$

valid for all $x, y \in K$ and for all $z \in H$ such that x is the projection of z into K. We consider a variational inequality of the same nature of (1): there exists a suitable function $\omega : K \times K \to \mathbb{R}^+$ such that for all $x, y \in K$, for all normal vectors $v \in H$ to K at x, it holds

$$\langle v, y - x \rangle \leq \omega(x, y) \, \|v\| \, \|y - x\| \,. \tag{2}$$

The function ω in (2) controls the lack of monotonicity of the normal cone. Various regularity properties of K, can be characterized through (2), by means of the asymptotic behavior of ω around the diagonal. If we assume ω to be continuous and such that $\omega(x,x) = 0$, then K is normally Fréchet regular (i.e. its Clarke and Fréchet normal cone coincide, see [3]) and the boundary of the normal cone has sequentially s$\times$w-closed graph. This property generalizes at the same time convexity and smoothness of the boundary of K. A more important concept is when $\omega(x,y) = \varphi(x,y)\|y-x\|$, as for example holds if K has a $\mathcal{C}^{1,1}$-boundary. This condition (called φ-convexity) is equivalent to the following: there exists an open set $\mathcal{U} \supset K$ such that the distance d_K from K is of class $\mathcal{C}^{1,1}(\mathcal{U} \setminus K)$ (see [4, 5, 7, 6, 8]). This result can be seen as a local version of Asplund's characterization of convexity [1], and can be understood in the following way: K satisfies an external sphere condition, with locally uniform radius. The size of $\mathcal{U}$ is related to this radius, which is infinite in the convex case.

A consequence of φ-convexity is the following (partial) generalization of the famous Michael's continuous selection theorem: let X be a paracompact space, and let $F : X \to 2^H$ be a continuous multifunction with φ-convex values. Fix $x_0 \in X$ and $y_0 \in F(x_0)$. Then there exists a neighborhood U of x_0 and a continuous selection $f : U \to H$ from F such that $f(x_0) = y_0$.

Some of the above results can be generalized to uniformly convex Banach spaces, through the proximal normal formula in Banach spaces [2].

References

[1] E. Asplund, Chebychev sets in Hilbert spaces, *Trans. Math. Soc.* **144** (1969), 235-240.

[2] J. M. Borwein, H. M. Strojwas, Proximal analysis and boundaries of closed sets in Banach space, Part I: Theory, *Can. J. Math.* **38** (1986), 431-452.

[3] M. Bounkhel, L. Thibault, On various notions of regularity in nonsmooth analysis, to appear.

[4] A. Canino, On p-convex sets and geodesics, *J. Differential Equations* **75** (1988), 118-157.

[5] F. H. Clarke, R. J. Stern, P. R. Wolenski, Proximal smoothness and the lower–$\mathcal{C}^2$ property, *J. Convex Anal.* **2** (1995), 117-144.

[6] G. Colombo, V. V. Goncharov, Variational inequalities and regularity properties of closed sets in Hilbert spaces, *J. Convex Anal.*, submitted.

[7] H. Federer, Curvature measures, *Trans. Am. Math. Soc.* **93** (1959), 418-491.

[8] R. A. Poliquin, R. T. Rockafellar, L. Thibault, Local differentiability of distance functions, *Trans. Amer. Math. Soc.*, to appear.

Relative continuity of lower semicontinuous convex functions

Joel BENOIST and Aris DANIILIDIS

We introduce two notions stronger than cyclic monotonicity, by replacing the finite sums (appearing in the definition of the latter) by appropriate sums of series. This allows us to obtain dual characterizations of relative continuity properties of lower semicontinuous convex functions. In particular it is shown that the restriction of a lower semicontinuous convex function to its domain (resp. domain of subdifferentiability) is continuous if, and only if, its subdifferential is strongly cyclically monotone (resp. σ-cyclically monotone).

Baire's category method in non-convex differential inclusions

F. S. DE BLASI and G. PIANIGIANI

Let $\mathbb{E}$ be a Banach space and let $f : I \times \mathbb{E} \to \mathbb{E}$ be a map belonging to a set, say $\mathcal{M}$. For $f \in \mathcal{M}$ consider the differential equation $x'(t) = f(t, x(t))$. The question whether this equation has properties like existence, uniqueness, continuous dependence, periodic solutions etc., is in general difficult to be answered, especially in infinite dimensions. In this context a number of mathematicians including Orlicz, Peixoto, Markus, Smale started to investigate whether the above properties were, perhaps, generically satisfied in $\mathcal{M}$. Here, if $\mathcal{M}$ is a complete metric space of maps, Baire category enters naturally if we agree to say that a property is *generic* in $\mathcal{M}$ provided that it is satisfied by all f belonging to a residual subset $\mathcal{M}_0$ of $\mathcal{M}$, i.e. with $\mathcal{M}\backslash\mathcal{M}_0$ a set of Baire first category. It is worthwhile to point out that knowing that a property is generic in $\mathcal{M}$ does not furnish, in general, any information whether a specific $f \in \mathcal{M}$ belongs or not to the generic set.

In this note we are concerned with existence results for Cauchy problems for ordinary differential inclusions and Dirichlet problems for partial differential inclusions of the form

$$x'(t) \in F(t, x(t)) \qquad x(0) = x_0$$

$$\begin{cases} \nabla u(x)) \in F(x, u(x)) & x \in \Omega \text{ a.e.} \\ u(x) = \varphi(x) & x \in \partial\Omega \ . \end{cases}$$

The method we discuss is based on the Baire category; yet, unlike the situation considered before, it enables us to establish existence results for specific differential inclusions.

Caratheodory's Theorem on Constrained Optimization and Comparative Statics

Emmanuel M. DRANDAKIS

The paper considers Caratheodory's Theorem on the properties of the Inverse of the Bordered Hessian of an Optimization Problem. After a new proof of the complete theorem, using matrix theory methods, the paper considers the sensitivity of the optimal solution in parameters appearing either in the objective or constraint functions. We also prove a second theorem that compares the symmetric submatrices on the main diagonal of the Inverse Matrix, before and after new constraints are introduced to the problem : such comparisons are essential for Le Chatelier Principle. Our "primal method" of comparative statics is as simple and elegant as any "dual method" that uses the Envelope Theorem.

The Lagrangian Search Method (LSM): Applying lagrangian relaxation to approximately solved restricted linear programs

Pavlos S. EFRAIMIDIS and Paul G. SPIRAKIS

In this work, a new algorithmic framework, the Lagrangean Search Method (LSM) for fast sequential and parallel approximations to Positive Linear Program (PLP) extensions, is presented. The motivation for considering PLP is that it admits very efficient approximations in sequential [1], parallel [1] and distributed settings [2]. An interesting aspect of LSM is that it is an example of introducing the standard lagrangian relaxation (LR) concept to a weaker submodel of LP, the PLP, which additionally is only approximately solved (within any given constant factor).

The original PLP model ([1]), is a restricted submodel of LP, that can directly model only a limited number of combinatorial optimization (CO) problems, including bipartite graph matching and set covering. It is shown that the LSM methodology when applied to the PLP model, achieves relaxed approximation schemes to extended PLP, which can model a substantially larger class of CO problems. The running time of the complete scheme is within a polylogarithmic factor of the original PLP approximation algorithms [1][2].

Finally, LSM is used within a methodology for fast parallel (and sequential) approximations to a class of hard CO problems. The members of the class are NP optimization problems that can be formulated as extended PLP programs and for which either polynomially related upper and lower bounds on their optimum can be found in NC, or they have the polybottleneck property, that is their optimal objective value is always within a polynomial factor of one of the input weights. The methodology is presented on Global Routing in Gate Arrays [3] and on the Discrete s-Median problem [4]. For problems of size N, the approximation algorithms run in polylog(N) time on

O(N) processors and in $\tilde{O}(N)$ time on one processor and achieve logarithmic or polylogarithmic approximation guarantees.

References

[1] M. Luby and N. Nisan, "A parallel approximation algorithm for positive linear programming", Proceedings of the 25th ACM Symposium on Theory of Computing, 1993, p. 448-457.

[2] Y. Bartal, J. Byers and D. Raz, "Global Optimization using Local Information with Applications to Flow Control", 38th IEEE Symposium on Foundations of Computer Science, 1997, p. 303-312.

[3] P. Raghavan and C.D. Thompson, "Randomized Rounding: A technique for provably good algorithms and algorithmic proofs", Combinatorica 7: 365-374, 1987.

[4] Juh-Han Lin and J.S. Vitter, "ϵ-Approximations with minimum Packing Constraint Violation", 24th ACM Symposium on Theory of Computing, 771-782, 1992.

Differences of Convex Compact Sets: I. The Space of Directed Sets

Robert BAIER and Elza FARKHI

The present paper proposes a technique for embedding the class of the convex compact subsets of R^n in a partially ordered and normed vector space providing a visualization of differences of embedded convex compacts.

Denote by $\mathcal{C}(R^n)$ the cone of the convex compact nonempty subsets of R^n, endowed with Minkowski addition and the inclusion order. The idea to extend this cone to a normed vector space has been attracted mathematicians in the last 50 years (e.g. Rådström ('52), Hörmander('54) etc.) The construction of such an abstract linear normed space provides possibilities to define derivatives of set-valued mappings (Banks, Jacobs ('70)), has applications in interpolation and approximation methods for convex set-valued mappings and set-valued integrals and yields general error estimation methods in abstract spaces.

The definition and study of notions of monotonicity and convexity for set-valued mappings at an abstract level can be enabled by introducing an order in the linear space. Note that monotonicity could be defined e.g. by positivity of the first finite differences and convexity by positivity of the second differences.

The constructed linear space of differences of embedded convex compact sets, however, contains new elements, which are no longer convex sets and there were no clear geometrical interpretations of these new objects so far.

To the best authors' knowledge, there are two basic approaches of embedding the cone of all convex compacts in R^n into a normed vector space: The approach of Rådström with pairs of sets exploits the classical algebraic method of extending a commutative semigroup into a group via representing differences as pairs. The obtained linear space contains equivalence classes of pairs with a norm extending the Hausdorff metric.

This approach has some disadvantages and no visualization is available. The problem of choosing a main representative of every equivalence class (such as minimal pairs) stays open for dimensions $n \geq 3$. Moreover,for $n \geq 3$ the minimal pairs may be not unique up to translations.

The "support function" approach of Hörmander is based on the unique representation of every convex compact set by its support function. Then, differences of support functions are embedded into the linear normed space of continuous, positively homogeneous functions, defined on R^n.

The idea to consider the boundary faces having a fixed normal vector and being $(n-1)$-dimensional sets affords to develop a recursive representation of a convex set via its supporting faces. This idea stands behind our construction of the embedding of $\mathcal{C}(R^n)$ into the normed vector space of directed sets and the visualizations of elements in this space. To define our space and the embedding, we use a specific parametrization of the convex sets by their support function and supporting faces.

In the paper the normed and partially ordered vector space $\mathcal{D}(R^n)$ of "directed sets" is constructed, in which the cone $\mathcal{C}(R^n)$ is embedded via a positively linear, order preserving and isometric embedding (with respect to a new metric equivalent to the Demyanov metric).

This space is a Banach and a Riesz space for dimension $n \geq 2$ and a Banach lattice for $n = 1$. The directed sets in R^n are defined recursively with respect to the dimension n by a mapping consisting of a "support" function from R^n to R and a lower dimensional directed "supporting face", defined for each normal direction in R^n.

In the case $n = 1$ the space $\mathcal{D}(R)$ is isomorphic to the space of so-called generalized (Kaucher,'80) resp. directed intervals (Markov,'92,95). Thus, one may say that the generalized/directed intervals form the basis of the recursive definition of directed sets. Let us note that the $n-$dimensional case is more complicated. The directed sets of higher dimensions have a mixed type part (a directed set being neither proper, nor improper), which does not exist in the case $n = 1$.

The difference of embedded directed sets is closely related to known differences of convex sets like geometric or Demyanov difference. These relations and the visualization of directed sets will be presented in the second part of this work.

Multiple Equilibria in Optimal Control Models

Gustav FEICHTINGER

During the last two decades the complexity of optimal paths in control models has been studied. The present paper deals with multiple long-run equilibria, in particular with the occurrence of unstable steady states and the existence of 'thresholds' or 'manifolds' separating different basins of attraction.

The structure of optimal policies in dynamic economic optimisation models is interesting for both mathematicians and economists. Applications cover wide and topic fields such as economic growth (diverging growth rates around the world), population growth, environment (many environmental problems are characterised by threshold properties), infrastructure, and management.

Skiba (1978) studied a one-sector optimal growth model with a convex-concave production function. A first sound proof of the existence of a threshold separating two basins of attraction in a one-state dynamic optimization model was given by Dechert and Nishimura (1983). Since that time many papers have been written dealing with multiple steady states and history dependence in one-dimensional optimal control models (see, e.g., Wirl and Feichtinger, 1999). Recently, Santos analysed a two-sector model of endogenous growth (Ladron-de-Guevara et al., 1999). The inclusion of leisure introduces non-convexities leading to multiple optimal stationary solutions. Although the dynamics is driven by a single state variable, this paper provides an important step to an extension to two-dimensional optimal control models.

In the economics of crime context, multiple steady states and hence threshold policies arise fairly frequently. We describe a dynamic extension of Beckers (1968) general approach to crime and punishment, where the law enforcement agency determines the optimal dynamic trade-off between damages caused by offenders, law enforcement expenditures and cost of imprisonment.

Moreover, we apply dynamic optimisation techniques more specifically to the control of illicit drug consumption. In all these models, there exists a Skiba threshold implying that above the Skiba point the optimal trade-off between social costs implies a steady state with a high level of offenders / drug users, while below the threshold the optimal law enforcement should eradicate crime / drug use. Due to the formal analogies between crime / drug control and, e.g., pollution models (similar control instruments: prevention, treatment, etc.), one can expect that these threshold policies also arise in a variety of other socio-economic contexts, e.g. in environmental planning.

Up to now we know of only a few economic mechanisms being responsible for two (or more) basins of attraction for the equilibria. Convex-concave production technology is one such mechanism, 'enforcement swamping' (cf. Kleiman, 1993) another, and asymmetric effects of the (human) capital to the time spent in different activities is a third. It would be interesting to look for other properties implying instabilities and path dependent behaviour. A promising research strategy is to identify a rather simple economic model exhibiting threshold behaviour. For that we constructed an investment model in which the installation cost depends on the investment to capital ratio. Such a model is simple enough to be studied analytically and shows all desired properties.

Several examples exist with one Skiba threshold defining the border between two basins of attraction (see the description above). Pre-studies suggest that under certain conditions (e.g., introduction of an upper bound on the controls), there may even exist two thresholds with three possible steady states to which the optimal solution may converge. In other words, the optimal steady state solution of such systems is by no means unique. Rather, one out of three equilibria will be approached when pursuing the optimal solution, depending on the state of the system under consideration when control starts.

While for one-dimensional dynamical optimisation models the existence of Skiba points is essentially understood, almost nothing is known for models with more than one state variable. We have numerical and theoretical hints that under certain conditions there exists something like a 'Skiba curve' in the state plane separating two basins of attractions. Our aim is to study socio-economic two-dimensional models to understand the complex behaviour of the optimal trajectories in the four-dimensional canonical system resulting from the first order optimality conditions. Some preliminary work shows that this might be a promising approach. However, to prove the existence of Skiba manifolds (and to study their properties), new methods must be developed and applied.

Another question is the relation of limit cycles as solutions of optimal control models to Skiba points (or curves). Even if it can be shown that the intersection between unstable equilibria and limit cycles is empty, it would be interesting to investigate whether persistent oscillations and path-dependence might occur in one (and the same) model. This is an interesting future research issue.

References

[1] Becker, G. S., 1968, Crime and punishment: an economic approach, Journal of Political Economy 76, 169-217.

[2] Dechert, W.D. and K. Nishimura, 1983, A complete characterization of optimal growth paths in an aggregated model with a non-concave production function, Journal of Economic Theory 31, 332 354.

[3] Kleiman, M. A. R., 1993, Enforcement swamping: a positive-feedback mechanism in rates of illicit activity, Mathematical and Computer Modelling 17(2), 65-75.

[4] Ladron-de-Guevara, A., S. Ortigueira and M. A. Santos, 1999, A two-sector model of endogenous

growth with leisure, Review of Economic Studies 66, 609 631.

[5] Skiba, A. K., 1978, Optimal growth with a convex-concave production function, Econometrica 46, 527 539.

[6] Wirl, F. and G. Feichtinger, 1999, History dependence in concave economies, Working Paper.

Global Optimization for the Graph Coloring Problem

Paola FESTA and Panos PARDALOS

Continuous approaches have been already used for solving several discrete optimization problems, including integer programming, the maximun clique problem, the satisfability problem, the hamiltonian cycle problem, and the Steiner problem.
There are several techniques to formulate a discrete problem as an equivalent continuous problem. The discrete feasible domain of the original problem is embedded in a larger continuous space, whose topological and geometric structure play a fundamental role. In fact, by exploring the properties of the obtained larger continuous space, new methods can be developed that lead to efficient algorithms for computing (sub)optimal solutions to the original discrete optimization problem.
In this talk we consider several continuous approaches for the Graph Coloring Problem, which can be defined as follows.
Let $G = (V, E)$ be an undirected graph, with vertex set $V(G) = \{1, 2, \dots, n\}$ and edge set $E(G)$. G has no loops or multiple edges. Vertices v_i and v_j are said to be **adjacent** if there exists an edge $(i, j) \in E(G)$, that is incident to both.
A **coloring** of G is defined to be an assignment of colors to its vertices so that no pair of adjacent vertices shares identical colors. In other words, a coloring of G is a mapping f of $V(G)$ into $\{1, 2, \dots, n\}$ such that for any $(i, j) \in E(G)$, $f(i) \neq f(j)$. A coloring naturally induces a partition of $V(G)$ such that the elements of each set in the partition are pairwise non-adjacent; these sets are precisely the subsets of vertices assigned the same color.
If there exists a coloring of G that uses no more than k colors, G admits a **k-coloring**, (G is **k-colorable**). The minimal k for which G admits a k-coloring is called its **chromatic number** and it is denoted by $\chi(G)$. Given a graph G, the graph coloring problem is the problem of finding the chromatic number $\chi = \chi(G)$ of G, as well as a partition of the vertex set, induced by a χ-coloring. The graph coloring problem is an **NP-Complete** problem; indeed, Garey and Johnson have shown that obtaining colorings using $s\chi(G)$ colors, where $s < 2$, is **NP-hard**. Thus, while exact solution procedures exist, the complexity of the problem requires the development of heuristics for large problem instances. To evaluate the efficiency of a mathematical programming solution, in terms of both solution quality and speed, extensive testing of the proposed procedure on a wide variety of problem instances is a crucial phase. Moreover, for purposes of benchmarking and comparison of different methods, a collection of problems must be available or readily generated. Since for the graph coloring problem, there are few types of test problems available, in this paper we propose in addition a method for constructing test graphs with a known chromatic number by using the KKT optimality conditions of a continuous nonlinear program.

Canonical Dual Transformation, Generalized Caratheodory Condition and Triality Theory in Nonconvex and Nonsmooth Optimization

David Yang GAO

Duality is a fundamental concept that underlies almost all natural phenomena. In convex optimization, duality theory has been well studied. However, in nonconvex/nonsmooth systems, many problems remain open. In this talk, the speaker will present a newly developed, potentially powerful sequential canonic dual transformation method, by which, many nonconvex/nonsmooth constrained problems in R^n can be converted into certain smooth, unconstrained dual problems in R^m with $n > m$ and without duality gap, and nonlinear partial differential equations in mathematical physics can be transformed into algebraic equations. Therefore, some analytic solutions can be obtained.

Generally speaking, duality theory falls into three catogories [1]: (1) the classical saddle Lagrange (minimax) duality in convex problems, (2) the nice super-Lagrangian bi-duality in geometrically linear systems, and (3) the interesting triality and multi-duality in general global optimization. In convex systems, the well-known Fenchel-Rockafellar duality is essentially equivalent to the saddle Lagrange duality. However, for nonconvex problems, there exists a duality gap between the primal problem and the Fenchel-Rockafellar dual problem. In order to recover this duality gap, a so-called complementary gap function was discovered by Gao and Strang in 1989, which yields a nonlinear Lagrange duality theory in fully nonlinear variational problems.

It was realized recently that if this gap function possesses a negative sign, the nonlinear Lagrangian associated with the nonconvex primal problem is a "super-critical" function. A very interesting triality theory in nonconvex optimization has been proposed, which reveals the intrinsic symmetry in fully nonlinear systems. By introducing the generalized Caratheodory condition, this triality theory can be used to solve many nonconvex, nonsmooth problems in both finite and infinite-dimensional systems (see [1]). In this talk, the speaker will show how to use the sequential canonical dual transformation to solve many problems arising in natural systems. Applications are illustrated by examples in phase transitions with multi-well energy, elastoplastic post-bifurcation with jumping and harderning effects, constrained nonsmooth global optimization and nonconvex dynamical systems.

References

[1] Gao, David Y. (1999). *Duality Principles in Nonconvex Systems: Theory, Methods and Applications*, Kluwer Academic Publishers, 474pp.

The Analytic Center Cutting Quadratic Method

Jean-Louis GOFFIN and Faranak SHARIFI-MOKHTARIAN

We extend the quadratic cut algorithms of Luo and Sun for solving the convex quadratic feasibility problem by placing or translating the quadratic cuts directly through the current approximate center. We also extend the Analytic Center Cutting Quadratic method of Büeller and Lüthi which studies a method where the number of cuts can be infinite, and the cut is central, but the cut is an Euclidean ball given by the constant of strong monotonicity of the quadratic, or in their case, the Jacobian of a variational inequality.

We show that it is possible to introduce an arbitrary quadratic at the analytic center, even in the case of an infinite number of cuts and preserve the convergence results of these papers.

We also prove that the total number of recentering steps required to update from one approximate analytic center to the next is $O(1)$, if the dual logarithmic barrier method is restarted from a point computed by a 1-dimensional curvilinear search. This search can be done efficiently if a bidiagonalization of Dikin's matrix and of the new quadratic is computed.

Solving Second Order Cone Programs by Interior Point Methods in an Efficient and Numerically Stable Manner

Donald GOLDFARB

Second order cone programs can be used to model robust linear programs and convex quadratically constrained quadratic programs as well as many other convex programming problems. If one solves such problems by Nesterov and Todd's interior point method, one has to solve a positive definite symmetric system of equations at each iteration. We prove that under fairly modest assumptions that almost universally hold in practice, that the Cholesky factors of this system have certain nice structural properties that allow one to take advantage of sparsity in the linear equations in the problem in a numerically stable way.

We also present ways to take advantage of sparsity in the Alizadeh-Haeberly-Overton interior point method. In this method a nonsymmetric system of equations has to be solved at each step. Hence the approach needed for it is different than in the Nesterov-Todd case.

The results of computational tests will be reported that show that speed ups of several orders of magnitude are obtained by our techniques over straightforward implementations and that verify the numerical stability of our methods.

A Hybrid Genetic-Tunneling Algorithm for medium size Global Optimization

Susana GOMEZ and Julio SOLANO

In this work we will study the convergence behaviour of these methods, over a set of examples with different characteristics. We will analyze the slow convergence of Genetic Algorithms (GA) near the solution, and will show the robustness of the Tunneling Method (TM).

We will present a Hybrid Method, that combines the capability to explore the space of GA, and the speed of the TM to get the desired precision. Numerical results will show the capabilities of the new method.

Random Unilateral Variational Problems and Random Variational Inequalities

Joachim GWINNER

In this paper we present a symbiosis of results and methods from Calculus of Variations, Measure Theory, and Theory of Partial Differential Equations - fields of analysis, where C. Caratheodory contributed famous achievements.

More specifically, we develop an existence, uniqueness, and regularity theory for problems of the following form.

Let $(\Omega, \mathcal{A}, \mu)$ be a complete σ-finite measure space and $(\mathcal{X}, (\cdot,\cdot), \|\cdot\|)$ be a separable Hilbert space. We are given a closed convex nonvoid subset $\mathcal{K}$ of $\mathcal{X}$ and Caratheodory functions $\lambda : \Omega \times \mathcal{X} \to \mathbb{R}$ and $\beta : \Omega \times (\mathcal{X} \times \mathcal{X}) \to \mathbb{R}$. Moreover, for every $\omega \in \Omega, \lambda(\omega, \cdot)$ and $\beta(\omega, \cdot, \cdot)$ is assumed to be a linear form, repectively a bilinear form; thus $\lambda(\omega, \cdot) \in \mathcal{L}(\mathcal{X}, \mathbb{R}) = \mathcal{X}^*$ and $\beta(\omega, \cdot, \cdot) \in \mathcal{L}(\mathcal{X} \times \mathcal{X}, \mathbb{R})$. Furthermore, for every $\omega \in \Omega$, $\beta(\omega, \cdot, \cdot)$ is assumed to be nonegative, that is, for every $y \in \mathcal{X}$, $\beta(\omega, y, y) \geq 0$ is supposed to hold. In this setting we study the following random variational inequality: Find for every $\omega \in \Omega$, an element $\hat{y} \in \mathcal{K}$ depending on $\omega \in \Omega$ such that

$$\beta(\omega, \hat{y}, z - \hat{y}) \geq \lambda(\omega, z - \hat{y}), \quad \text{for all } z \in \mathcal{K}. \tag{1}$$

In this setting we have the following result.

Theorem. *Let β be a coercive bilinear form. Then there exists a unique solution $\hat{u} : \omega \in \Omega \mapsto \hat{y} = \hat{u}(\omega) \in \mathcal{K}$ of the random variational inequality (1).*
Suppose in addition, that μ is a finite measure and β is uniformly coercive. Moreover suppose, that the real valued functions $\omega \mapsto \|\beta(\omega)\cdot,\cdot)\|_{\mathcal{L}(\mathcal{X}\times\mathcal{X},\mathbb{R})}$ and $\omega \mapsto \|\lambda(\omega,\cdot)\|_{\mathcal{X}^*}$ belong to $L^q(\Omega,\mu)$ for some $1 \leq q \leq \infty$. Then we have $\hat{u} \in L^q(\Omega; \mathcal{X}, \mu)$.

Such a result applies to the following unilateral boundary value problem: Find a function Ψ such that

$$\begin{cases} -\nabla \cdot (A \nabla \Psi) = f & \text{in } G \\ \Psi = 0 & \text{on } \Gamma_D \\ q_\nu := (-A\nabla\Psi) \cdot \nu = h & \text{on } \Gamma_N \\ \Psi \leq 0, q_\nu \geq h, \Psi(q_\nu - h) = 0 & \text{on } \Gamma_S. \end{cases}$$

Here G is a bounded domain in $\mathbb{R}^n (n = 2$ or $n = 3)$ with locally Lipschitz boundary ∂G. Moreover, $\partial G = \overline{\Gamma_D} \cup \overline{\Gamma_N} \cup \overline{\Gamma_S}$ with open parts $\emptyset \neq \Gamma_D, \Gamma_N, \emptyset \neq \Gamma_S$. Randomness enters by A, considered as a random matrix on some probability space $(\Omega, \mathcal{F}, P)$.

Using monotonicity methods, the results above can be further extended to variational problems involving nonlinear elliptic operators with random coefficients.

Theorems for Control Problems with General Boundary Conditions

Andreas H. HAMEL

Several attempts has been made to establish (necessary) conditions for nearly optimal (suboptimal) solutions of optimal control problems. Most of these conditions cover only problems without terminal constraints or admit a violation of these constraints which leads to controls which possibly are not admissible for the original problem. The aim of the talk is to give characterizing conditions for admissible suboptimal controls and, by the way, to weaken the smoothness assumptions of earlier results in this field considerably.
We shall consider the following optimal control problem:

$$\text{minimize} \quad l_0\left(x\left(0\right), x\left(1\right)\right) \tag{1}$$

$$\begin{aligned} \text{subject to} \quad \dot{x}\ (t) \quad &= \quad f\ (t, x\ (t)\ , u\ (t)) \quad \text{a.e.} && (2) \\ u\ (t) \quad &\in \quad U\ (t) \quad \text{a.e.} && (3) \\ (x\ (0)\ , x\ (1)) \quad &\in \quad C && (4) \end{aligned}$$

where the functions $l_0 : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $f : [0,1] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, and the sets $C \subset \mathbb{R}^n \times \mathbb{R}^n$, $U(t) \subset \mathbb{R}^m$ are the given data. We denote by $\inf(P)$ the infimal value of (P). We posit the following assumptions due to Clarke (1976).
(A1) The function $l : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is locally Lipschitz continuous.
(A2) For each x in a neighbourhood of a closed set $X \subset \mathbb{R}^n$, the function $(t,u) \to f(t,x,u)$ is $\mathcal{L} \times \mathcal{B}-$measurable.
(A3) There is a function $\gamma(\cdot) \in L^1(0,1)$ such that for $t \in [0,1]$, $u \in U(t)$, and x, y in a neighbourhood of X,

$$|f(t,x,u) - f(t,y,u)| \leq \gamma(t)\,|x-y|\,.$$

(A4) The graph of the set-valued function $t \to U(t)$ is $\mathcal{L} \times \mathcal{B}-$measurable.
(A5) The set $C \subset \mathbb{R}^n \times \mathbb{R}^n$ is closed.
The main result is the following theorem.

Theorem 1 *(Suboptimality Theorem)*
Assume (A1)-(A4) and $-\infty < \inf(P) < +\infty$. Let $\varepsilon > 0$ be given. Then there are a trajectory $x_\varepsilon \in W^{1,1}$ and a control function u_ε which are admissible for (P) such that there exist multipliers $p_\varepsilon \in W^{1,1}$, $\lambda_\varepsilon \geq 0$ such that $\|p_\varepsilon\| + \lambda_\varepsilon = 1$ as well as

$$l(x_\varepsilon(0), x_\varepsilon(1)) \leq \inf(P) + \varepsilon^2 \tag{5}$$

the adjoint inclusion

$$-\dot{p}_\varepsilon(t) \in p_\varepsilon(t)\,\partial_x f(t, x_\varepsilon(t), u_\varepsilon(t)) \tag{6}$$

the $\varepsilon-$transversality condition

$$(p_\varepsilon(0), -p_\varepsilon(1)) \in \lambda_\varepsilon\left(\partial l_0(x_\varepsilon(0), x_\varepsilon(1)) + \varepsilon B \times \{0\}\right) + \mathcal{N}_C(x_\varepsilon(0), x_\varepsilon(1)) \tag{7}$$

and the $\varepsilon-$maximum principle

$$p_\varepsilon(t)\, f(t, x_\varepsilon(t), u_\varepsilon(t)) \geq p_\varepsilon(t)\, f(t, x_\varepsilon(t), u) - \varepsilon\lambda_\varepsilon \;\forall\; u \in U(t) \;\; a.e. \tag{8}$$

This theorem is a generalization of Ekeland's 1974 first suboptimality theorem in optimal control to nonsmooth data and problems with general boundary conditions. It is much more precise than, for example, the ε-maximum principle of Gabasov/Kirillova/Mordukhovich (1983). The theorem can be utilized to construct approximate solutions of control problems which may have no optimal solution as well as a tool for proving new necessary optimality conditions. We shall give examples for both the applications.

Object Oriented Approach for Direct Methods in Global Optimization

Beidi HAMMA

The work presented in this paper was first motivated by the needs of Object Oriented Approach for optimization algorithms within the project ODESIM, a High Performance Computing and Networking (HPCN) project funded by the European Commission within ESPRIT programme. The project acronym stands for Optimum DESIgn of Multi-Body systems(MBS). The starting point for the implementation of the ODESIM system was the CAD and MBS simulation packages available

within the consortium. They then had to be improved with new modules for HPCN and optimization. For an efficient and reliable implementation, it has been adopted to use standards like UNIX, C++, X-Windows and other internal tools, based on object oriented programming techniques which will guarantee the openness, portability and scalability of the system.
Given that, we designed a general object oriented framework which enables to encapsulate any optimization method.
We considered all optimization problems: Unconstrained, Constrained, availability (or not) of Derivatives , availability (or not) of Second Derivatives. We also considered all type of methods for solving those problems: non-derivative methods, gradients methods, global optimisation methods. This enables to have a very simple and efficient interface which will allow any user to easly define his problem, and to add any new optimization solver to his general framework.
This paper is concened with some examples of OO approach for Direct Methods: Moving Polytopes and Simulated Annealing In the first case we show how the OO approach enables to unify the design and implementation of three Moving Polytope algorithms: the original simplex method of Nelder and Mead [1], the direct search method of Dennis and Torczon [2] and the adaptive search of Hamma [3]. In the second one, we show how a general OO approcah for Simulated Annealing can be acheived. In both cases, the OO approahes enables to plug these packages into any other OO framework and, futhermore, we can easily use a method with its different variants and/or mixt the algorithms inside a same package.

The testing results in this work showed how the use of Object-Oriented approach in the design of numerical optimization libraries, gives the following capabilities:

1. better optimization interface in the case where objective functions (and gradients) are evaluated by the outpout of a simulation that is, as a black-box for the optimizer
2. rapid encapsulation of different optimization routines already written in C, Fortran etc. This encapsulation reduces the user's dependency on the way original routines are implemented and work.
3. rapid evaluation of several optimization routines on a same problem.
4. rapid and efficient combination of several optimization routines in the purpose of global optimization.
5. better parallelization using other object-oriented heteregeneous parallel managers.
6. better code maintenance and re-usability.

Genetic and Stochastic Global Minimization Methods in Shape Optimization

Jaroslav HASLINGER and Tomas KOZUBEK

In this paper we will present how genetic and stochastic type algorithms combined with the so-called fictitious domain methods may be helpful in the numerical realization of optimal shape design problems.

Let us suppose that ω is an admissible domain and $(\mathcal{P}(\omega))$ is the corresponding state problem which is linear. Then its approximation by a standard finite element method leads to the following linear algebraic system:

$$\mathbb{A}(\boldsymbol{\alpha})\mathbf{u}(\boldsymbol{\alpha}) = \mathbf{F}(\boldsymbol{\alpha}), \tag{1}$$

where $\mathbb{A}$ is the stiffness matrix, $\mathbf{F}$ is the load vector and $\boldsymbol{\alpha}$ is the so-called vector of discrete design variables corresponding to ω. After discretization, optimal shape design problems lead to the following *non-linear mathematical programming problems*:

$$(\vec{\mathbb{P}}) \qquad \begin{cases} \textit{Find } \boldsymbol{\alpha}^* \in \mathcal{U} \textit{ such that} \\ \mathcal{L}(\boldsymbol{\alpha}^*, \mathbf{u}(\boldsymbol{\alpha}^*)) \le \mathcal{L}(\boldsymbol{\alpha}, \mathbf{u}(\boldsymbol{\alpha})) \quad \forall \boldsymbol{\alpha} \in \mathcal{U} \end{cases}$$

in which $\mathbf{u}(\boldsymbol{\alpha})$ is the solution of (1), $\mathcal{L}$ is the algebraic representation of a cost functional and $\mathcal{U}$ is the set of discrete design variables corresponding to all admissible domains. A traditional way widely used in practise how shape optimization problems are realized is based on the so-called *boundary variation technique*, when each new term of the minimizing sequence is obtained by a suitable *deformation* of the previous shape. This means, among others, that (1) has to be solved many times for different $\boldsymbol{\alpha}$'s. The main shortcomings of this approach are readily seen: after any change of $\boldsymbol{\alpha}$, we must construct a new partition of any new configuration and then to update the stiffness matrix $\mathbb{A}$ and the load vector $\mathbf{F}$. Only now it is possible to solve (1).

In order to increase the efficiency of the inner optimization level, we propose the so-called *fictitious domain methods* (FDM) for the numerical realization of a state problem. The main advantage of FDM's is the fact that all computations are performed on a fixed, simple shaped domain Ω, using the *same partition* of Ω, which is independent on $\boldsymbol{\alpha}$. As a result, the stiffness matrix $\mathbb{A}$ will not depend on the vector of discrete design variables. This increases the efficiency of the inner optimization level.

Unfortunately the use of the fictitious domain solvers reduces the smoothness of $\mathcal{L}$. The minimized function may become *non-continuously differentiable* and in some cases even *discontinuous*. The lost of the smoothness of $\mathcal{L}$ practically excludes the use of minimization methods employing derivatives of $\mathcal{L}$. To overcome this difficulty, global optimization methods based on function evaluations can be used. In this contribution we show that genetic and stochastic type global optimization methods present a possible way how to overcome all these difficulties. Model examples in 2 and 3-D will be shown.

Convergence and applications of the θ-scheme for finding a zero of the sum of two maximal monotone operators

Sylvianne HAUBRUGE, Van Hien NGUYEN and Jean-Jacques STRODIOT

Many problems of convex programming can be reduced to that of finding a zero of the sum of two maximal monotone operators A and B on $\mathbb{R}^n$. For solving this problem, different splitting methods were studied in the literature as, for example, the forward-backward scheme, the Peaceman–Rachford scheme and the Douglas–Rachford scheme.

In 1989, Glowinski and Le Tallec proposed a new scheme, called θ-scheme, for finding a zero of $A + B$. Its iteration can be expressed as follows :

$$z^{k+1} = J_{\lambda_1 A}(I - \lambda_1 B) J_{\lambda_2 B} \frac{\lambda_2}{\lambda_1}(\alpha J_{\lambda_1 A} - I)(I - \lambda_1 B) z^k,$$

where $\lambda_1, \lambda_2 > 0$, $\alpha = 1 + \frac{\lambda_1}{\lambda_2}$, $J_{\lambda_1 A}$ and $J_{\lambda_2 B}$ denote the resolvents associated with A and B respectively. Glowinski and Le Tallec applied their scheme to some problems such as elastoviscoplasticity problems, liquid crystal theory, eigenvalue computation, ... They mentioned that the scheme seemed to give good numerical results but that they had no general proof of convergence.

In the first part of this talk, we will present some convergence results for the θ-scheme. We will see that the sequence (z^k) converges to a zero of $A+B$ if B^{-1} is strongly monotone. Moreover, we will also see that the rate of convergence is at least linear if, in addition, we assume that either A or B is also strongly monotone.

The second part of the talk will be devoted to some applications of the θ-scheme. We will apply it to some convex programming problem and to the minimization of $m+1$ convex functions in order to derive a new algorithm for solving them. In view of our convergence result for the θ-scheme, we will then give a result for the convergence of these new algorithms.

In the last part of the talk, we consider a perturbed version of the θ-scheme :

$$z^{k+1} = J_{\lambda_1 A^k}(I - \lambda_1 B^k) J_{\lambda_2 B^k} \frac{\lambda_2}{\lambda_1}(\alpha J_{\lambda_1 A^k} - I)(I - \lambda_1 B^k) z^k,$$

where both (A^k) and (B^k) are sequences of maximal monotone operators which approach A and B respectively, in the sense that :

$$\sum_{k \geq 0} \delta_{\lambda_1,\rho}(A^k, A) < +\infty \quad \text{and} \quad \sum_{k \geq 0} \delta_{\lambda_2,\rho}(B^k, B) < +\infty \quad \forall \rho \geq 0$$

where, for maximal monotone operators S and T, $\delta_{\lambda,\rho}(S,T)$ denotes the variational semi-distance between S and T, with parameters $\lambda > 0$ and $\rho \geq 0$. We will see that the sequence (z^k) generated by the perturbed θ-scheme converges also to a zero of $A+B$ if, as for the exact version, B^{-1} is strongly monotone. Finally, we consider the following convex programming problem :

$$\begin{cases} \min & f(x) \\ s.c. & x \in C \end{cases}$$

where $C = \{x \in I\!R^n : g_i(x) \leq 0, 1 \leq i \leq m\}$. The operators A and B are in this case the subdifferentials of f and the indicator Ψ_C of the set C respectively. For this problem, we will give some examples of functions, called "penalty functions", whose subdifferentials $B^k = \partial\varphi_k$ satisfy the proximity assumption

$$\sum_{k \geq 0} \delta_{\lambda_2,\rho}(B^k, B) < +\infty \quad \forall \rho \geq 0$$

required in the convergence result of the perturbed θ-scheme. So, for example, we will see that this is the case for the class of outer penalties (as well for classical outer penalties as for exact outer penalties). Then we will introduce a class of penalties which contains as a particular case the class of exponential penalties and we will see that the proximity assumption is also satisfied in this case.

References

[1] Haubruge, S., Nguyen, V.H., Strodiot, J.J. (1998) : Convergence analysis and applications of the Glowinski-Le Tallec splitting method for finding a zero of the sum of two maximal monotone operators. J. Optim. Theory Appl. **97**, 645–673.

[2] Haubruge, S., Nguyen, V.H., Strodiot, J.J. (2000) : Convergence analysis of a perturbed version of the θ-scheme of Glowinski-Le Tallec. In : Nguyen V.H., Strodiot J.J., Tossings P. (eds.), Optimization. Lecture Notes in Economics and Mathematical Systems, **481**, 149–163. Springer–Verlag,

Berlin.

Pythagoras' Theorem for Areas

J-P. QUADRAT, J. B. LASSERRE and J.-B. HIRIART-URRUTY

Pythagoras' theorem relating the squares of the lengths of the sides of a rectangular triangle is one of the key-results in elementary Geometry. At a more advanced level, one learns that Pythagoras' theorem extends to prehilbertian spaces, but it is still a relation between *lengths* of vectors. We incidentally came across a result relating the areas of a trirectangular tetrahedron (in the 3-dimensional space) and whose analogy with Pythagoras' theorem is striking. In this short communication, we:
- state this result and its historical developments;
- propose a generalization of the result to the n-dimensional space. In short,
"Given a polyrectangular convex compact polyhedron, the square of the area of the hypothenuse facet is equal to the sum of the squares of the areas of the n facets issued from the origin".

Inexact versions of generalized proximal point and augmented Lagrangian algorithms in Banach spaces

Alfredo N. IUSEM

We extend a proximal-like method for finding zeroes of maximal monotone operators in Hilbert spaces with quadratic regularization due to Solodov and Svaiter, making it possible the use of other kind of regularizations and extending it to Banach spaces. In particular, we introduce an appropiate error criterium to obtain an inexact proximal iteration based on Bregman functions and construct a hyperplane wich strictly separates the current iterate from the solution set. A Bregman projection onto this hyperplane is then used to obtain the next iterate. Boundedness of the sequence and optimality of the weak accumulation points are established under suitable assumptions on the regularizing function, which hold for any power greater than one of the norm of any unifomly convex and uniformly smooth Banach space, without any assumption on the operator other than existence of zeroes. We then transpose such methods to generate augmented Lagrangian methods for $\mathcal{L}^p$-constrained convex optimization problems in Banach spaces, obtaining two alternative procedures which allow for inexact solutions of the primal subproblems. Boundedness of both the primal and the dual sequences, and optimality of primal and dual weak accumulation points, are then established, assuming only existence of KKT-pairs.

Exceptional Family of Elements, Leray-Schauder Alternative, Pseudomonotone Operators and Complementarity

George ISAC, Vyacheslav V. KALASHNIKOV

In this paper, the property of a function to be without exceptional family of elements (EFE) is investigated. We show that for a pseudomonotone operator, the solvability of complementarity problem is equivalent to the property of the function to be without EFE. Moreover, we study the strict feasibility of a complementarity problem making use of the Leray-Schauder Alternative and the notion of EFE.

New LMI relaxations for the general nonconvex quadratic problem

Jean B. LASSERE

We consider the general nonconvex quadratic problem and provide a series of LMI relaxations whose sequence of optimal values converges monotonically to the optimal value and contains the well-known Shor's relaxation as its first member. In many cases the global optimal value is obtained for a particular relaxation. Results are presented on simple examples of quadratic problems, 0-1 problems as well as Max-Cut problems.

Semantic Complexity in Linear Programming

Jean-Louis LASSEZ

Work done partly at the IBM T.J. Watson Research Center and while visiting Brown University under a contract from the Office of Naval Research. The complexity measures in Linear Programming are based on the syntactic properties of a set of constraints, such as number of constraints, number of variables or the size of coefficients. We take here a different approach, introducing a measure based on the geometric properties of a polyhedron associated to a linear program. We will argue that such a measure helps us understand better conditions under which a linear program can be executed efficiently by simplex like algorithms. In a first part of the paper we establish a correspondence between a polytope P and a normalized linear program in the space dual to the space of the polytope. This is different from the classical approach where the algebraic and the geometric interpretation are in the same space. We call this linear program quasi dual as it is a normalized form of a dual to an also normalized set of constraints defining P. In primal space we address a solvability problem: is P empty? In dual space we address a minimization problem. An interesting correspondence arises where a basis in the quasi dual corresponds to a simplex in primal space (by simplex we mean here the geometric object, not the algorithm) . The objective function provides the radius of the sphere inscribed in the simplex. A degenerate basis corresponds to a degenerate simplex (a simplex where one or more extreme points are at infinity). As we will show this point of view is quite consistent with Helly's theorem, Fourier's elimination and Farkas lemma. In a second part we derive from the previous discussion a simple solvability algorithm which in the quasi dual moves from basis to basis and in the primal moves from simplex to simplex until the center of the simplex is found to be a solution to P. In the non degenerate case the radii of the spheres inscribed in the simplices decreases by a factor which depends on the angles made by the constraints defining the simplex. We then derive a complexity measure which depends on geometric properties of P. Essentially we find that the algorithm's efficiency depends on how far P is from being degenerate (in our sense, that is the constraints do not form angles too close to zero), and on a measure related to the volume of P. In fact away from degeneracy, the two factors compensate each other leading to quite significant efficiency. In a third part we relate the previous study to analyze the behavior of the simplex algorithm. There are three factors that differentiate the simplex algorithm from the algorithm we discussed: the two normalizations and the use of the objective function. We will argue, that these differences are not necessarily fundamental (one algorithm will have fewer but more costly steps). Using numerous examples and run time pragmatic evidence, we will show that the standard simplex algorithm has a behavior which is not too far removed from the behavior of our proposed algorithm. As a consequence we have a new approach to analyze the

pragmatic complexity of the simplex and derive new heuristics to increase its efficiency.

Interval Analysis Applied to Global Optimization for Solving the Molecular Conformation Problem

Carlile Campos LAVOR

We use a deterministic global optimization algorithm for finding global minima potential energy conformations of molecules based on the Lennard-Jones potential function. This function has a large number of local minima which may grow exponentially with the numbers of atoms of the molecule, thus making the computation of global minima very difficult. The algorithm is based on a branch and bound method that uses techniques of interval analysis to calculate the bounds. The proposed approach was successfully applied to two example problems.

D.C. optimization approaches via Markov models for restoration of signals (1D) and (2D) affected by additive gaussian white noise

Hoai An LE THI and Tao PHAM DINH

This work deals with numerical optimization methods for restoration of signals (1D) and (2D) affected by additive gaussian white noise via the Markov model. The Markov model principe involes Markov and Gibbs fields on finite graphs connected through the Hammersley Clifford theorem. The restaured signal is estimated by the maximum posteriori criterion, which turns out to be a mixed global optimization problem of the form (in case of signal (1D))

$$\inf\{E(u,x)=\sum_{i=1}^{n}(u_i-d_i)^2+\lambda^2\sum_{i=1}^{n-1}(u_i-u_{i-1})^2(1-x_i)+\alpha\sum_{i=1}^{n-1}x_i : u \in R^n,\ x\in\{0,1\}^n\} \quad (P)$$

(d is the signal with noise).

The Gibbs function $E(u,x)$ is convex quadratic in u for each fixed x and affine linear in x while u is fixed. If (x^*,u^*) is an optimal solution of this problem, then u^* is the restaured signal, and the valeurs $x_i^*=1$ correspond to the discontinuous points of this signal.

We propose two methods based on d.c. optimization algorithms (DCA) to solve this problem in the following equivalent forms:

(1) Minimization of a d.c. function on the whole space;

(2) Maximization of a convexe function on $[0,1]^n$.

The DCA introduced by Pham Dinh Tao in 1986 and later developed in our various works is a primal-dual subgradient method for solving a general d.c. program that consists in the minimization of difference of convex function on the whole space. It is actually one of a few algorithms in the local approach which has been successfully applied to many large-scale d.c. optimization problems and proved to be more robust and efficient than related standard methods. Due to its local character it cannot guarantee the globality of computed solutions for general d.c. programs. However, we observe that with a "good" starting point it converges quite often to a global one. This property motivates us to combine our different approaches to solve both (1) and (2). The algorithms are quite simple and easy to implement. Numerical experiments on several test problems are presented which show the efficiency and the robustness of the proposed algorithms.

Energy Landscape Characterization: What Makes a Global Optimization Problem Hard?

Robert H. LEARY

Molecular conformation problems arising in computational chemistry and biology require the global minimization of a non-convex energy function modeling the interactions among the atoms in a molecular system. The global minimum corresponds to the native structure of the molecule, which in the case of biomolecules is of fundamental importance in determining function. Typically such problems exhibit what Stillinger refers to as the 'exponential multiplicity of inherent structures' - that is, the number of structurally distinct local minima grows exponentially with system size and reaches truly enormous numbers even for moderately sized systems.

But exponential multiplicity does not necessarily imply practical difficulty in solving the corresponding global optimization problem. This is the point of the Levinthal paradox for protein folding, which contrasts the relatively fast rates that naturally occuring proteins are observed to fold into their native structures with the impossibly slow rates implied by simple random sampling of the local minima. It is now thought that nature exploits the fact that the local minima are organized into a large scale funnel that is used to guide a trajectory through a descending sequence of local minima to the global minimum. Recently we have modified the Doye-Wales basin-hopping algorithm to mimic this behavior in the context of Lennard-Jones (LJ) clusters, perhaps the best known molecular conformation problem and a benchmark for global optimization algorithms in general. The new algorithm proved remarkably robust, easily finding all known global minima out to N=110 atoms, including several cases that have proved quite difficult or impossible for other algorithms. Remarkably, the algorithm found a new global minimum for N=98 representing a completely novel morphological class, one of only four such classes now known for LJ global mimima. This represents the first instance where a new class has been discovered by global optimization rather than simply constructed based on physical principles and observation of natural forms.

In addition to its effectiveness as a global optimization methodology, the new algorithm also has the interesting property of exposing the funneling (or multi-funneling) nature of the energy landscape, and finding low-energy representatives of the distinct conformational classes represented by each funnel. Thus the algorithm can be used as a tool to characterize the structure of the energy landcape, and thus to provide insight into why, for example, certain problem sizes for the LJ cluster problem are quite easy for many different algorithms, while nearby sizes are consistently difficult. Here we will present some landscape characterizations and observations obtained from large scale numerical experiments on a very large (over 1000-processor) parallel implementation of the algorithm. A variety of models, with a range of landscape structures and consequent behaviors toward global optimization algorithms, will be considered, including LJ clusters, glasses, and simple proteins.

Lower Bound of Network Throughput and Validity of HW Conjectures for Irreducible Closed Markovian Queueing Networks

Kwang-Fu LI

For an irreducible closed Markovian network with N customers, Jin, Ou, and Kumar (1997) obtained pointwise bounds of the network throughput for any fixed integer N by solving linear programs $\overline{T}(N)$ and $\underline{T}(N)$. They also obtained other two linear programs $\overline{T}(\infty)$ and $\underline{T}(\infty)$ that can be used to develop the corresponding linear programs $\overline{FB}(\infty)$ and $\underline{FB}(\infty)$. Then the functional upper bound and lower bound of the network throughput can be derived from the objective values of the linear programs $\overline{FB}(\infty)$ and $\underline{FB}(\infty)$, and the optimal values of the linear programs $\overline{T}(\infty)$ and $\underline{T}(\infty)$ respectively. For convenience, if T is a given linear program, denote its optimum objective value by VT. We prove that $\lim_{N \to \infty} V\underline{T}(N) = V\underline{T}(\infty)$. On the other hand, the limit of the functional lower bound as N approaches to infinity is also equal to $V\underline{T}(\infty)$. Thus the lower bound of the throughput, in heavy traffic, can be obtained by solving $\underline{T}(\infty)$ only. Extend this special linear program $\underline{T}(N)$ to a general linear program, say $G(N)$, where N can be viewed as a variable. We conjecture that if $VG(N)$ is bounded then there exists a limiting program $G(\infty)$ such that $\lim_{N \to \infty} VG(N) = VG(\infty)$. This limiting program $G(\infty)$ can be easily obtained from modifying the original linear program $G(N)$.

Harrison and Wein (HW) (1990) proposed a buffer priority policy for two-station Brownian networks. HW conjectured that buffer priority policy is asymptotically optimal, and the asymptotic loss always has a finite value expression. Jin, Ou, and Kumar (1997) proved that the HW-policy is efficient for all two-station systems. They conjectured that under the "additional condition", all of the conjectures of HW are true for balanced two-station systems. However, a proof of its asymptotic optimality has so far been unavailable. We are able to develop a relation between asymptotic loss and the "additional condition". For a specially constructed system, we can also prove that all non-idling policies are asymptotically optimal. In this specially constructed system, if it is a balanced re-entrant line two-station system, the value of the asymptotic loss is 1.

Classes of Quasiconvex Functions on R, Closed under Addition and Translation

Per Olov LINDBERG

It is well known that many generalizations of convex functions, such as quasiconvex and pseudoconvex functions, have a lot in common with convex functions. However, they do not behave so nicely under functional operations, such as addition.

Therefore, in this talk we study classes of quasiconvex functions, closed under addition and translation.

Trivial examples of such classes are classes of strictly increasing or decreasing functions.

Apart from these examples, we show that for each class there is some real constant m such that $f(1/mlogx)$ is convex for each f in the class, $m = 0$ corresponding to the functions already being convex.

The background of this problem comes partly from a statistical problem. Here one had functions that were scaled translates of a given function, and one wanted to know whether all local optima were global. The other source was an old question of the author, concerning what arc cost functions one could have in network flow problems and still have the property that all local optima are global.

The Analytic Center Quadratic Cut Method (ACQCM) for Strongly Monotone Variational Inequality Problems with Approximate Centers

Hans-Jakob LÜTHI and Benno BÜELER

This paper extends the analysis of Lüthi and Büeler in [1]. Here we replace the exact analytic centers by approximate analytic centers, turning the conceptual algorithm from [1] into a practical algorithm. For strongly monotone variational inequality problems (VIP) convergence of an algorithm is investigated which, at each iteration k, adds a quadratic cut through an approximate analytic center x_k of the subsequently shrinking convex body. First it is shown that the sequence of x_k converges to the unique solution x^* of the VIP at $\mathcal{O}(1/\sqrt{k})$. As an interesting detail note that — for increasingly accurate analytic centers — the complexity constants converge to the quantities obtained for ACQCM with exact centers.

Secondly we show that the arithmetic complexity to update from x_k to x_{k+1} after inserting a quadratic cut through x_k is bounded by a constant number of Newton iterations plus $\mathcal{O}(n \cdot \ln\ln[\zeta \cdot \frac{k^2}{\varepsilon^4}])$, where n is the space-dimension, ε is the final solution accuracy $\|x_k - x^*\|$, and ζ depends on some problem-specific constants only. We want to point to the parallel work of Nesterov, Péton and Vial [2] where a homogeneous analytic center cutting plane method (HACCPM) with approximate centers is analyzed. Due to the construction of their iterates by using a weighted average of analytic centers, the final accuracy determines the accuracy of all previous approximate centers. In our analysis we can avoid this severe restriction. As a disadvantage of our quadratic cuts, however, we have to accept an increase in the computational complexity for updating the analytic center when the iteration number grows. This increase, however, grows at $\mathcal{O}(\ln(\|\tilde{x}_k - x_k\|_{[F_k''(x_k)]})$ where $\tilde{x}_k$ denotes the center of the k-th cut through x_k. In case of HACCPM linearity of the cuts allows to keep the update complexity constant.

References

[1] Hans-Jacob Lüthi and Benno Büeler. The analytic center quadratic cut algorithm for strongly monotone variational inequality problems. *SIAM Journal on Optimizatoion 10: 415 -428, 2000.*

[2] Yu. Nesterov, O. Péton, and Jean-Philippe Vial. Homogeneous analytic center cutting plane methods with approximate centers. *To appear in Optimization Methods and Software.*

A new rule for generating convex functions

Pierre MARÉCHAL

It is well known that a function f of the real variable x is convex if and only if $(x, y) \to y f(y^{-1}x)$ is convex for $y > 0$. This property can be extended, under certain assumptions, to functions of the form $(x, y) \to g(y) f(g(y)^{-1}x)$.

More precisely, let f be a convex function on $\mathbb{R}^n$, let g be any function on $\mathbb{R}^m$, and let $g \rhd f$ be the function defined on $\mathbb{R}^n \times \mathbb{R}^m$ by

$$\big(g \rhd f\big)(\mathbf{x}, \mathbf{y}) := \begin{cases} g(\mathbf{y})\, f\left(\dfrac{\mathbf{x}}{g(\mathbf{y})}\right) & \text{if } |g(\mathbf{y})| \in (0, \infty), \\ f0^+(\mathbf{x}) & \text{if } g(\mathbf{y}) = 0, \\ +\infty & \text{if } |g(\mathbf{y})| = \infty. \end{cases}$$

(Here, $f0^+$ denotes the recession function of f, that is, the support function of $\operatorname{dom} f^\star$.) We shall prove the following theorem:

Theorem 1 *Suppose that f is closed proper and convex and that g is closed proper and concave. Suppose in addition that $f(\mathbf{0}) \leq 0$ and g is non-negative on its domain, then $g \rhd f$ and $f^\star \rhd (-g)^\star$ are conjugate to each other:*

$$(g \rhd f)^\star(\boldsymbol{\xi}, \boldsymbol{\eta}) = (f^\star \rhd (-g)^\star)(\boldsymbol{\eta}, \boldsymbol{\xi})$$
$$\textit{and} \quad (f^\star \rhd (-g)^\star)^\star(\mathbf{y}, \mathbf{x}) = (g \rhd f)(\mathbf{x}, \mathbf{y}).$$

An immediate consequence is that, under the assumptions of Therorem 1 the functions $g \rhd f$ and $f^\star \rhd (-g)^\star$ are closed and convex. This provides a new rule for generating new convex functions from old ones. This will also prove to be useful for the study of the multiplicative potential and penalty functions [1, 2, 5, 6]

$$F(\mathbf{x}) := \frac{(\langle \mathbf{c}, \mathbf{x} \rangle - c_0)^{m+1}}{\prod_{j=1}^{m} (\langle \mathbf{a}_j, \mathbf{x} \rangle - b_j)}, \qquad \Phi(\mathbf{x}) := \frac{(\langle \mathbf{c}, \mathbf{x} \rangle - c_0)^{m}}{\prod_{j=1}^{m} (\langle \mathbf{a}_j, \mathbf{x} \rangle - b_j)}$$

$(\mathbf{c}, \mathbf{a}_j \in \mathbb{R}^n, c_0, b_j \in \mathbb{R})$, which play an important role in linear programming [3, 4]. Additional results on the differentiability and strict convexity of functions of the form $g \rhd f$ will also be given.

References

[1] M. Avriel, W. E. Diewert, S. Schaible and I. Zang, *Generalized Convexity*, (Plenum Press, New York and London, 1988).

[2] J.M. Borwein, "A generalization of Young's l^p inequality," *Mathematical Inequalities and Applications* 1, (1998) 131–136.

[3] I. Imai, "On the convexity of the multiplicative version of Karmarkar's potential function," *Mathematical Programming* 40, (1988) 29–32.

[4] M. Iri and I. Imai, "A multiplicative barrier function method for linear programming," *Algorithmica* 1, (1986) 455–482.

[5] S. Schaible, "Quasi-convex optimization in general real linear spaces," *Zeitschrift fuer Operations Research*, (1972) 205–213.

[6] J.P. Crouzeix, J.A. Ferland and S. Schaible, "Generalized convexity on affine subspaces with an application to potential functions," Mathematical Programming 56, (1992) 223–232.

A Comparison Principle and the Lipschitz Continuity for Minimizers

Carlo MARICONDA and Giulia TREU

We consider the problem of the existence and regularity of the minima of an integral functional of the form

$$I(u) = \int_\Omega L(x, u, \nabla u)\, dx$$

on $\bar{u} + W_0^{1,q}(\Omega)$ $(1 \leq q \leq +\infty)$ where $L(x, z, p) = f(p) + g(x, z)$ is convex in (z, p).

In the classical approach the Lagrangian L is assumed to be smooth, convex in ∇u and to satisfy, together with its derivatives, some growth conditions. This ensures the existence of a minimum of I that satisfies the integral *Euler equation*. The results on the regularity of the minima then follow from those on the regularity of the solutions to quasilinear elliptic equations. There are some recent results that do not involve conditions on the derivatives of L. We mention the papers [CV] and [AAB] for one–dimensional problems and [GG] for the multi–dimentional case. All of them assume some growth conditions on L; namely in [GG] it is required that $L(x, z, p)$ behaves like a power of p for $|p| \to \infty$.

We give some structure assumptions on L (not involving growth conditions) that ensure that the minima of I lying between two Lipschitz functions are actually Lipschitz. Then, in order to study the regularity of the minima, we look for conditions on the boundary datum that guarantee the existence of two Lipschitz functions that bound the minima of I. Since the minima of I are not supposed to satisfy the Euler equation and the Lagrangian is not assumed to be smooth, the classical *Comparison Principle* for partial differential equations cannot be applied. We state a new *Comparison Principle for minima* and we prove that if the boundary datum fulfills a *Generalized Bounded Slope Condition* (GBSC) or a *barrier condition* then the minimum for I in $\bar{u} + W_0^{1,q}(\Omega)$ exists for every $1 \leq q \leq +\infty$, it is unique and Lipschitz (and therefore no Lavrentiev phenomenon occurs).

As a first step we obtain the existence of a minimum in the class of Lipschitz functions, extending a result by M. Miranda for functionals depending only on the gradient. Then, via Clarke's nonsmooth version of the Euler equation, we obtain the existence of minima in Sobolev spaces; uniqueness follows from our Comparison Principle for minima.

In the smooth setting the (GBSC) and the barrier condition are well known and are used to give an estimate of the gradient of regular solutions to partial differential equations. We point out that we exploit them in a different way since we do not know a priori neither whether a minimum exists nor if it is a solution to the Euler equation.

References

[AAB] L. Ambrosio, O. Ascenzi, G. Buttazzo, *Lipschitz regularity for minimizers of integral functionals with highly discontinuous integrands*, J. Math. Anal. Appl.**142** (1989), 301–316.

[CV] F. H. Clarke, R. B. Vinter, *Regularity properties of solutions to the basic problem in the calculus of variations*, Trans. Amer. Math. Soc. **289** (1985), 73–98.

[GG] M. Giaquinta, E. Giusti, *Differentiability of minima of nondifferentiable functionals*, Invent. Math. **72** (1983), 285–298.

[MT1] C. Mariconda, G. Treu, *A Comparison Principle and the Lipschitz continuity for minimizers*, preprint, Dipartimento di Matematica pura e applicata **27** (1999), Università di Padova.

[MT2] C. Mariconda, G. Treu, *A Comparison Principle for minimizers*, to appear, Comptes Rendus de l'Academie des Sciences, Série I (2000).

[MT3] C. Mariconda, G. Treu, *A gradient Maximum Principle for minima*, preprint, Dipartimento di Matematica pura e applicata **14** (2000), Università di Padova.

[MT4] C. Mariconda, G. Treu, *Existence and Lipschitz regularity for minima*, preprint, Dipartimento di Matematica pura e applicata **15** (2000), Università di Padova.

A complementary pivoting approach to the maximum clique problem

Alessio MASSARO and Marcello Pelillo

In this talk we present a new heuristic for approximating the Maximum Clique Problem by transforming the problem into an equivalent indefinite QP using Bomze's refinement of a theorem of Motzkin and Straus.

We consider the LCP arising from the KKT conditions of the aforementioned QP and then solve it with Lemke's algorithm. By exploiting the degeneracy connected with the problem and a least-index-like resolution method, we were able to obtain a very effective heuristic.

Our implementation of Lemke's method enjoys a set of invariants that allow for the reduction of the amount of data needed for processing the LCP and the simplification of certain operations. There is also empirical evidence that the algorithm avoids those KKT points of Bomze's QP that are not local solutions.

The above ideas have been used to obtain a new pivoting-based heuristic that has been tested over a large selection of DIMACS benchmark graphs. The results obtained represent a significant improvement upon other methods based on the Motzkin-Straus QP. In particular the Sanchis graphs seemed extremely vulnerable to our approach, whereas other Motzkin-Straus-based algorithms had to forfeit for much smaller cliques.

The Moments Method

René MEZIAT, Juan José EGOZCUE and Pablo PEDEGRAL

We have devised a method to calculate convex hulls of polynomials which is based on Caratheodory's theorem and some results about algebraic moments used in probability and stochastic processes. The present work shows the success of the method to treat optimization and non convex variational problems involving polynomials. The main feature of this method is its ability to convert non convex optimization problems into linear programing problems under convex constraints. This fact guarantee the convergence to the global optimum eluding stops at local extremes.

As a consequence of the Caratheodory's theorem we can find the values of the convex hull of the polynomial

$$\phi(x) = \sum_{k=0}^{2n} a_k x^k \tag{1}$$

solving the optimization problems

$$co\,\phi(t) = \min_{\mu \in \mathcal{M}_t} \int_R \phi(x)\, d\mu(x) \tag{2}$$

where $\mathcal{M}_t$ is the family of all probability measures with mean at point t. Now we can see that each integral in (2) take the form

$$\sum_{k=0}^{2n} a_k m_k$$

where m_k is the k-th order algebraic moment of measure μ. Therefore we change the original optimization problem (2) into a new programming problem

$$\min_{m_k} \sum_{k=0}^{2n} a_k m_k \tag{3}$$

where the variables m_k must be restricted to be the moments of a probability measure. We will show the equivalence between the solutions of the optimization problems (2) and (3) This equivalence gives an alternative way to treat optimization problems as

$$\min_{t\in R}\phi\left(t\right)$$

where ϕ is a polynomial with the form given in (1)

By other side the method has been proved to treat non convex variational problems as the Bolza's example

$$\min_{u\in W^{1,4}(0,1)}\int_0^1\left\{\phi\left(u'\left(x\right)\right)+u\left(x\right)^2\right\}dx \tag{4}$$

under the boundary conditions

$$u\left(0\right)=0 \qquad u\left(1\right)=0$$

where the non convex potential

$$\phi\left(t\right)=\left(1-t^2\right)^2$$

entails the lack of classical solution. In order to attack this problem we must solve the generalized problem

$$\min_{\nu}\int_0^1\left\{\int_R\phi\left(\lambda\right)d\nu_x\left(\lambda\right)+\left(\int_0^x\int_R\lambda d\nu_s\left(\lambda\right)ds\right)^2\right\}dx \tag{5}$$

under the boundary condition

$$\int_0^1\int_R\lambda d\nu_x\left(\lambda\right)dx=0$$

where ν is a Young Measure or family of probability distributions ν_x. We stress that the optimum ν^* has the information about the oscillatory behavior in the minimizing sequences for the non convex problem at (4)

In this case we have changed the generalized problem (5) in Young Measures into a new optimization problem

$$\min_{m_k(x)}\int_0^1\left\{1-2m_2\left(x\right)+m_4\left(x\right)+\left(\int_0^x m_1\left(s\right)ds\right)^2\right\}dx$$

where the $m_k\left(x\right)$ are restricted to be the algebraic moments of the distribution ν_x. We will show the equivalence between the generalized problem (5) and the new optimization problem (6) next we will compute the solution to the optimization problem (6) as a programming problem with convex constraints.

Necessary Conditions and Location of Solutions

Dumitru MOTREANU

The aim of the paper is two-fold: (i) to find new necessary conditions of optimality in nonlinear mathematical programming problems which involve (possibly unbounded) closed range operators; (ii) to locate the solutions of variational–hemivariational inequalities (in short, VHI) in the sense of P. D. Panagiotopoulos (see, e.g., [2]).

The connection of parts (i) and (ii) is generally given by the fact that the results in (i) are often expressed in terms of VHI. Thus the location of possible solutions to VHI in (ii) applies naturally to problem (i), too.

The optimization problem in (i) is formulated as follows:

$$\begin{aligned}&\text{Minimize } L(y,u)\\ &\text{subject to } Ay = Cu + B(y,u),\end{aligned}$$

where $A : D(A) \subset Y \to E$ and $C : D(C) \subset X \to E$ are (possibly unbounded) closed linear operators with dense domains, $B : V \times U \to E$ is a Gâteaux differentiable map, $L : W \to \mathbb{R}$ is a locally Lipschitz function, where X, Y, E are Banach spaces and V, U, W are open subsets of Y, X, $Y \times X$, respectively, with $W \supset V \times U$. Under a tangency condition for the set of constraints and the assumption for the ranges $R(A - B_y(y,u))$ and $R(C + B_u(y,u))$ to be closed and comparable by inclusion, one obtains that corresponding to each optimal pair (y,u) there exists some $p \in D(A^*) \cap D(C^*)$ such that

$$(-(A - B_y(y,u))^* p, (C + B_u(y,u))^* p) \in \partial L(y,u), \tag{1}$$

where the notation ∂L stands for the generalized gradient of L in the sense of Clarke [1], while B_y and B_u denote the partial derivatives of B.

A model situation with a boundary value problem (see [3], Section 6.3 for other examples) is the following

$$\begin{aligned}&\text{Minimize } \int_\Omega \int_0^{y(x)} g(x,t)\,dt\,dx + \tfrac{1}{2}\int_\Omega u^2(x)\,dx\\ &\text{subject to } -\Delta y = f(y,u) \text{ in } \Omega,\ y = 0 \text{ on } \partial\Omega.\end{aligned}$$

Part (ii) deals with the following VHI on a Banach space V: find $w \in V \setminus \{0\}$ and $\lambda \in \mathbb{R}$ such that

$$\begin{cases} \lambda\langle Aw, v - w\rangle + (-J)^\circ(w; v - w) + \Psi(v) - \Psi(w) \geq 0,\ \forall v \in V,\\ \lambda > \overline{\lambda}, \end{cases} \tag{2}$$

where $\overline{\lambda} > 0$ is a given number, $A : V \to V^*$ is a linear continuous coercive operator, $J : V \to \mathbb{R}$ is a locally Lipschitz functional (denoting by $(-J)^\circ$ Clarke's generalized directional derivative [1] of $-J$) and $\Psi : V \to \mathbb{R} \cup \{+\infty\}$ is a convex, proper and lower semicontinuous function. Relation (1) can take in some situations the form (2) on $V = Y \times X$ (with $w = (y,u)$). In fact, (2) is an eigenvalue problem for VHI with a constraint for the eigenvalue λ.

Our approach in obtaining the location of eigensolutions $(w,\lambda) \in V \times \mathbb{R}$ of (2) relies on the use of an appropriate function $\beta \in C^1(\mathbb{R})$ closely related to the data A, J, Ψ, $\overline{\lambda}$. The location of eigensolutions (w,λ) is then described under suitable assumptions by a kind of parametric representation

$$\begin{cases} \langle Aw, w\rangle^{\frac{1}{2}} = s^{-1}(-\beta'(s))^{\frac{1}{q}},\\ \lambda = \overline{\lambda} + s^3(-\beta'(s))^{\frac{q-2}{q}}, \end{cases} \tag{3}$$

with a positive parameter s, where $q \geq 2$ is the exponent in the growth condition for $-J + \Psi$. Relations (3) allow to deduce various qualitative properties for the eigensolutions (w,λ) of (2). For example, the imposed conditions allow to obtain a sequence (w_n,λ_n) of eigensolutions of (2) such that $w_n \to 0$ in V, $\lambda_n \to +\infty$ and $\lambda_n\|w_n\|^2 \to 0$ as $n \to \infty$.

References

[1] F. H. Clarke, *Optimization and Nonsmooth Analysis*, John Wiley, New York, 1983.

[2] D. Motreanu and P. D. Panagiotopoulos, *Minimax Theorems and Qualitative Properties of the Solutions of Hemivariational Inequalities and Applications*, Kluwer Academic Publishers, Nonconvex Optimization and Its Applications, Vol. 29, Dordrecht, Boston, London, 1999.

[3] D. Motreanu and N. H. Pavel, *Tangency, Flow Invariance for Differential Equations and Optimization Problems*, Pure and Applied Mathematics, Vol. 219, Marcel Dekker, Inc., New York, Basel, 1999.

A Global Concavity Condition for Decisions with Several Constraints

Christophe MULLER

The negativity of the substitution matrix is a fundamental property in consumer theory and general equilibrium. At the optimum, it is associated with sufficient second order conditions of the optimisation programme. Globally, it is the consequence of the rong quasiconcavity of the utility function and play an important role in applied economics for the confrontation of estimated demands with theoretical restrictions. With several nonlinear constraints, similar results incorporating a global generalised concavity condition is known. However, locally the second order conditions of the optimisation programme can still be considered as theoretical restrictions. They ensure the local unicity and local stabi ty of the equilibrium. They still correspond at the optimum to the negativity of a generalised substitution matrix (NGSM). In general models, the quasiconcavity or the strong quasiconcavity of the objec ve function are sometimes debatable restrictions to impose, and different generalised concavity conditions, and weaker one at least at the optimum, are desirable. We propose such a generalised conca ty condition that not only implies the NGSM, but also the global uniqueness and the smoothness of the decision functions.

On a global optimization problem involving integrands expressed as a minimum of convex functions

Zdzisław NANIEWICZ

We consider a nonconvex, nonsmooth global optimization problem for the class of integral functionals involving integrands expressed by means of two convex functions. Our problem reads

$$(P) \qquad \inf\Big\{\int_\Omega \min\big\{\tfrac{1}{2}a\,|\varepsilon(u)+C|^2\,,\ \tfrac{1}{2}b\,|\varepsilon(u)+D|^2\big\}\,dx\colon u\in H_0^1(\Omega;\mathbb{R}^n)\Big\}=\alpha,$$

where $\Omega \in \mathbb{R}^n$ is a bounded region in $\mathbb{R}^n$, $\varepsilon(u)=\frac{1}{2}(\nabla u+\nabla u^T)$ is the symmetrized gradient of u, $C, D \in L^\infty(\Omega;\mathbb{R}^{n\times n}_{sym})$ and where $a, b \in L^\infty(\Omega)$ are such that $a(x)$, $b(x) \geq \delta$ for a.e. $x \in \Omega$, $\delta > 0$ is a positive constant. The integrand as being not quasiconvex does not ensure the existence of a solution of the problem under consideration. However the minimizing sequence $u^k \in H_0^1(\Omega;\mathbb{R}^n)$ for (P) can be constructed together with functions $\chi_a^k : \Omega \to \{0,1\}$ and $\chi_b^k : \Omega \to \{0,1\}$, $\chi_a^k + \chi_b^k \equiv 1$, such that

(a) $u^k \to u$ weakly in $H_0^1(\Omega;\mathbb{R}^n)$ as $k \to \infty$,

(b) $\chi_a^k \to \chi_a$, $\quad \chi_b^k \to \chi_b$ weak* in $L^\infty(\Omega)$ as $k\to\infty$, where $\chi_a\colon \Omega \to [0,1]$, $\chi_b\colon \Omega\to[0,1]$ with $\chi_a+\chi_b\equiv 1$,

(c) $\int\limits_\Omega \frac{1}{2}\chi_a^k a\left|\varepsilon(u^k)+C\right|^2+\frac{1}{2}\chi_b^k b\left|\varepsilon(u^k)+D\right|^2\,dx \le \int\limits_\Omega \frac{1}{2}\chi_a^k a\left|\varepsilon(w)+C\right|^2+\frac{1}{2}\chi_b^k b\left|\varepsilon(w)+D\right|^2 dx,\ \forall\, w\in H_0^1(\Omega;\mathbb{R}^n)$,

(d) $\int\limits_\Omega \frac{1}{2}\chi_a^k a\left|\varepsilon(u^k)+C\right|^2+\frac{1}{2}\chi_b^k b\left|\varepsilon(u^k)+D\right|^2 dx=\alpha^k\to\alpha$ as $k\to\infty$.

Using the Fenchel duality theory the dual problem of (c) can be formulated in the form

$$\begin{aligned}\beta^k=-\inf\Big\{\int\limits_\Omega\Big\{&\tfrac{1}{2}\tfrac{1}{a\chi_a^k+b\chi_b^k}\left|q-(\tfrac{C+D}{2}+(\chi_b^k-\chi_a^k)\tfrac{D-C}{2})\right|^2\\&-\tfrac{a|C|^2+b|D|^2}{4}-(\chi_b^k-\chi_a^k)\tfrac{b|D|^2-a|C|^2}{4}\Big\}\,dx,\quad q\in\operatorname{Ker}L^\star\Big\},\end{aligned}\tag{1}$$

where $L: H_0^1(\Omega;\mathbb{R}^n)\to L^2(\Omega;\mathbb{R})$ is a linear operator defined by $Lv:=\varepsilon(v)$ and $L^\star$ is its transpose (the divergence operator). Denoting by $p^k\in\operatorname{Ker}L^\star$ the solution of (1) one can suppose that $p^k\to p$ weakly in $L^2(\Omega;\mathbb{R})$ as $k\to\infty$ for some $p\in\operatorname{Ker}L^\star$. Moreover, by the Fenchel theorem we have $\alpha^k=\beta^k$. The question arises what is the interrelation between the limits α, u, p, χ_a and χ_b? To answer this question let $\Omega_0=\{x\in\Omega\colon a(x)=b(x)\}$. The main result states that if $\frac{C-D}{b-a}\in L^2(\Omega\setminus\Omega_0;\mathbb{R}^{n\times n}_{sym})$ then there exists a constant $\theta\in[0,1]$ such that

$$\begin{aligned}\alpha=&\int\limits_{\Omega_0}\Big\{\big(\tfrac{C+D}{4}+(\chi_b-\chi_a)\tfrac{D-C}{4}\big)\cdot\varepsilon(u)+\tfrac{a|C|^2+b|D|^2}{4}+(\chi_b-\chi_a)\tfrac{b|D|^2-a|C|^2}{4}\Big\}\,dx\\&+\tfrac{1}{2}\int\limits_{\Omega\setminus\Omega_0}\Big\{\tfrac{ab(C-D)}{b-a}\cdot\varepsilon(u)+\tfrac{bD-aC}{b-a}\cdot p+\tfrac{ab(|C|^2-|D|^2)}{2(b-a)}-(\chi_b-\chi_a)\,\tfrac{ab|C-D|^2}{2(b-a)}\Big\}\,dx\\&-\tfrac{\theta}{2}\int\limits_{\Omega_0}\chi_a\chi_b\,a\,|C-D|^2\,dx.\end{aligned}$$

The last integral can be regarded as the relaxation term and it vanishes whenever $\chi_a=1$ or $\chi_b=1$. The other formulas are available and will be discussed in this presentation.

Box reduction for systems of nonconvex quadratic constraints

Arnold NEUMAIER

Methods from linear algebra and convex geometry are combined with Lagrangian techniques to reduce a box to a subbox without losing points of the box satisfying a system of convex or nonconvex quadratic constraints.

Together with branch and bound techniques and local optimization, this allows to decide in finite time whether a system of quadratic constraints has solutions in a given box.

The matrix calculations work within a fixed sparsity pattern, so that larger problems can utilize sparse matrix techniques.

New Search Directions Through Riemannian Metrics

J. X. da Cruz Neto, O. P. Ferreira and Paulo Roberto OLIVEIRA

Let M be a Riemannian manifold of dimension m, with metric $\langle,\rangle$. Denote by $X(M)$ the space of vector fields over M and by T_xM the tangent space of M at x, and by $F(M)$ the ring of functions over M. The metric $\langle,\rangle$ induces a map in $f \in F(M) \to \operatorname{grad} f \in X(M)$, which associates to each f its gradient via the rule $\langle \operatorname{grad} f, V\rangle = \mathrm{d}f(V)$, where $V \in X(M)$. Now let $\Omega, \varphi, \xi = (\xi_1, ..., \xi_m)$ a local coordinate system in a neighborhood of a point $x \in M$ and let $\{\partial/\partial\xi_i\}$ be the natural basis of the tangent space to M. The metric in M, corresponding to that coordinate system is given by the real, symmetric and positive definite $m \times m$ matrix $G(y)$, with $G(y)_{ij} = \langle \partial/\partial x_i\,|_y, \partial/\partial x_j\,|_y\rangle$ for all $y \in \Omega$. Then the gradient vector field is given by $\operatorname{grad} f(y) = G^{-1}(y)\, f'(y)$,where $f'(y) = (\partial f/\partial\xi_1, ..., \partial f/\partial\xi_m)$.

Let $N \subset M$ be a Riemannian submanifold of M with induced metric. For $x \in N$, let $\Pi_x : T_xM \to T_xN$ be the orthogonal projection onto T_xN with respect to the metric of the manifold. Now if $f\,|_N$ is the restriction of f to N, then the gradient of $f\,|_N$ with respect to the induced metric to N is given by $\operatorname{grad} f\,|_N(x) = \Pi_x \operatorname{grad} f(x)$ .From now on we will refer to it as the steepest descent direction.

Consider the optimization problem $\min_{x\in N} f(x)$, where $N \subseteq M$. By endowing M with the Riemannian metric G, the steepest descent direction of the objective function, given by the gradient formulas above, with respect to this metric, can generate an extensive family of search directions. Let Ω be a subset of R^n with interior Ω^o nonempty. Consider the optimization problem $\min f(x)$, $x \in \Omega$. Let $h : \Omega^o \to R$ be a real function, such that $h''(x)$ is positive definite for all $x \in \Omega^o$. Define a Riemannian metric G in Ω^o by setting $G(x) = h''(x)$. Then, the steepest descent direction with respect to the metric G is given by $\operatorname{grad} f(x) = (h''(x))^{-1} f'(x)$. In this talk, through different choices of h, we will produce some families of minimization algorithms for unconstrained problems.

Next, we consider constrained problems. Let Ω and Ω^o as above. Let $a \in R^n$ be a regular value of a function $\Gamma : \Omega^o \to R^m$ and $S = \Gamma^{-1}(a)$. Consider the constrained optimization problem $\min f(x)$, $x \in S$. Let $h : \Omega^o \to R$ be a real function, such that $h''(x)$ is positive definite for all $x \in \Omega^o$. Define a Riemannian metric G in Ω^0 by setting $G(x) = h''(x)$. If $S = \Gamma^{-1}(a)$ then $S \subset \Omega^o$ is a Riemannian submanifold of dimension $m - n$ with induced metric. In this case, the orthogonal projection $\Pi_x : R^n \to T_xS$ onto the tangent space $T_xS = \{v \in R^n : \Gamma'(x)\, v = 0\}$ with respect to the metric G, omitting the point x, is given by the known formula $\Pi = I - G^{-1}\Gamma'^T \left(\Gamma' G^{-1} \Gamma'^T\right)^{-1} \Gamma'$, where I is the $n \times n$ identity matrix. Therefore the steepest descent search direction to the function $f\,|_S$ with respect to the induced metric in S by G is $\operatorname{grad} f\,|_S = G^{-1/2}\left(I - G^{-1/2}\Gamma'^T \left(\Gamma' G^{-1}\Gamma'^T\right)^{-1} \Gamma' G^{-1/2}\right) G^{-1/2} f'$. Analogously to the unconstrained case, we propose some search directions for some classes of constrained optimization problems.

The third part of this talk is the interpretation of a lot of known methods as gradient ones, e. g., multiplicative, affine scaling, among others.

In conclusion, we present convergence theorems, for the general structures that were considered, including convex and non-convex functions. Also, a (linear) rate of convergence theorem is given.

The Convex Partition Problem

Shmuel ONN, N. ALON, F. HWANG, U. ROTHBLUM, L. SCHULMAN and B. STURMFELS

The *convex partition problem* is to find an ordered partition $\pi = (\pi_1, \ldots, \pi_p)$ of a given set $A = \{a_1, \ldots, a_n\}$ of n vectors in real d-space into p parts that maximizes an objective function which is convex on the sum of vectors in each part. The problem has broad expressive power and applications in diverse areas - see [2] and references therein. One particularly appealing recent application, for symbolic computation, is an efficient algorithm for constructing *universal Gröbner bases* of vanishing ideals of point configurations in fixed dimension [3].

In its full generality, the problem instantly captures NP-hard problems even if either p or d is fixed, hence is presumably intractable. In this talk I will describe our recent work [4] which provides, for every fixed p and d, a strongly polynomial time algorithm using $O(n^{d(p-1)-1})$ arithmetic operations for the problem.

This result improves upon the previous bound of $O(n^{d\binom{p}{2}})$ provided in [2] for the more general *shaped* partition problem. The solution method in [2] was based on the *p-partition polytope* $\mathcal{P}_A^p$ defined to be the convex hull of all $d \times p$ matrices $A^\pi := \sum_{j=1}^{p} \sum_{a \in \pi_j} a \otimes e_j$ associated with p-partitions π of A, and on the enumeration of all *separable partitions* of A. But, this method can not be accelerated since, as shown in [1], for each $p \geq 2, d \geq 3$, the number of separable partitions is $\Theta(n^{d\binom{p}{2}})$. Our new method in [4] is based on the introduction of the *p-signing zonotope* $\mathcal{Z}_A^p$ of A and on showing that it is a refinement of the p-partition polytope $\mathcal{P}_A^p$ of the same set of points.

References

[1] N. Alon and S. Onn, Separable partitions, *Discrete Applied Mathematics*, 91:39–51, 1999.

[2] F. Hwang, S. Onn and U. Rothblum, A polynomial time algorithm for shaped partition problems, *SIAM Journal on Optimization*, 10:70–81, 1999.

[3] S. Onn and B. Sturmfels, Cutting corners, *Advances in Applied Mathematics*, 23:29–48, 1999.

[4] S. Onn and L.J. Schulman, The convex partition problem, preprint, 2000.

The closedness of the linear image: weakest constraint qualification revisited

Gabor PATAKI

A fundamental question in convex analysis is the following: given a closed convex cone K, and a linear mapping M, under what conditions is MK closed ?

We consider the case, when K is "nice", that is, its faces satisfy a certain condition. Nice cones include the semidefinite cone, and cones arising from p-norms; practically all cones occurring in optimization.

For this class, we give a necessary, and sufficient condition for MK to be closed, and discuss several implications in the duality theory of conic linear programs.

A General Class of Interior-Point Algorithms for Semidefinite Programming with Polynomial Complexity and Superlinear Convergence

Florian A. POTRA

Semidefinite programming (SDP) has been one of the most active areas of optimization research in the 1990s. SDP consists in minimizing a linear functional of a matrix subject to linear equality and inequality constraints, where the inequalities are generated by the cone of positive semidefinite matrices. SDP includes as special cases linear programming, convex quadratic programming and other optimization problems. Semidefinite programming has numerous applications in such diverse areas as optimal control, combinatorial optimization, structural optimization, pattern recognition, trace factor analysis in statistics, matrix completions, etc. However, It was only after the advent of interior-point methods, that efficient solution methods for SDP problems were available. During the past few years an impressive number of interior-point methods for SDP have been proposed. Some of them have been successfully implemented and used to solve important application problems. However the theory and practice of interior-point methods for SDP has not yet reached the level of maturity of interior-point methods for LP, QP, and LCP. One reason that the study of interior-point methods for SDP is extremely important is that while LP, QP, and LCP can also be solved by other methods (e.g. the simplex method or Lemke's method), interior-point methods appear to be the only efficient methods for solving general SDP problems presently known.

A recent paper of Todd is witness to the intensity of research in SDP interior-point methods: It describes twenty techniques for obtaining search directions for SDP. The analysis of interior-point algorithms for SDP is considerably more difficult than for linear programming, as both the global convergence and the local convergence properties seem to depend on the choice of the search direction. Superlinear convergence is especially important for SDP since no finite termination schemes exist for such problems. As predicted by theory and confirmed by numerical experiments, the condition number of the linear systems defining the search directions increases as we approach the solution. Therefore an interior-point method that is not superlinearly convergent is unlikely to obtain high accuracy in practice. On the other hand, superlinearly convergent interior-point methods often achieve good accuracy in substantially fewer iterations than indicated by general computational complexity results.

We propose a general class of infeasible-start predictor-corrector algorithms for semidefinite programming problems, using the Monteiro-Zhang unified direction. The algorithms are direct generalizations of the Mizuno-Todd-Ye predictor-corrector algorithm for linear programming. We show that the algorithms belonging to this class are globally convergent, provided the problem has a solution, and have optimal computational complexity. For any positive number M, the algorithms can also detect in a finite number of steps that the problem has no solution of norm less than M. This feature can be used to detect infeasible problems. We also give simple sufficient conditions for superlinear convergence. Our results generalize and unify several recent results in the field.

Faithful representations of convex Hamilton-Jacobi equations

Franco RAMPAZZO

In this talk by a control theoretic *representation* of a Hamilton-Jacobi equation

$$u_t + H(t,x,\nabla_x u) = 0 \qquad (t,x) \in [0,T] \times R^n \tag{E}$$

we mean a pair (f,l), where $f = f(t,x,c)$ is a vector field on R^n, the *dynamics*, and $l = l(t,s,c)$ is a scalar map, the *lagrangian*, such that

$$H(t,x,p) = \max_{c\in C}\{-p\cdot f(t,x,c) - l(t,x,c),$$

C being a *control set.*

If that is the case, the value function of a control problem having f and l as dynamics and lagrangian, respectively, turns out to be a solution (in some weak sense) of (E). Of course, a necessary condition for such a representation is that the hamiltonian H is convex in adjoint variable p. Moreover, under very general assumptions, a representation always exists via dualization: in that case $l = H^*$, the polar of H. However, H^* generally can take the value $+\infty$, while there are several reasons for looking for representations as regular as possible. For example, this would allow one to establish sharp regularity results for the solution of (E) (when this is unique).

We present a kind of control theoretical representations such that:

i) the dynamics is locally Lipschitz continuous in x

ii) if the Hamiltonian H is $\alpha-$Holder continuous, then the lagrangian is $\alpha-$Holder continuous as well.

From the point of view of Holder regularity this is the best which can be done, for hamiltonians deriving from $\alpha-$Holder continuous lagrangians and locally Lipschitz continuous dynamics turn out to be $\alpha-$Holder continuous. Our result improves Ishi's ones (1985), where the obtained dynamics was not locally Lipschitz and the lagrangian turned out to be just $(\alpha/2)$ Holder continuous.

An infeasible active set method for convex problems with simple bounds

K. KUNISCH and Franz RENDL

We present an active set approach to strictly convex quadratic problems with simple bounds. The algorithm generates iterates that satisfy the first order condition of optimality and also the complementarity conditions, but neither primal nor dual feasibility is maintained, i.e. the dual multiplier need not have the right sign, and primal inequalities may be violated. We show finite convergence under a certain 'diagonal dominance condition', and exhibit the practical behaviour on a variety of problems. It turns out that this approach is often significantly better than interior-point based approaches.

Semidefinite Relaxations of Fractional Programs via Novel Techniques for Constructing Convex Envelopes of Nonlinear Functions

Mohit TAWARMALANI and Nick SAHINIDIS

In a recent work, we introduced the concept of convex extensions for lower semi-continuous functions and studied their properties. In this work, we present new techniques for constructing convex and concave envelopes of nonlinear functions using the theory of convex extensions. In particular, we develop the convex envelope and concave envelope of $z = x/y$ over a hypercube. We show that the convex envelope is strictly tighter than previously known convex underestimators of x/y. We propose a new relaxation technique for fractional programs which includes the derived envelopes. The resulting relaxation is shown to be a semidefinite program. Finally, we derive the convex envelope for a class of functions of the type $f(x, y)$ over a hypercube under the assumption that f is concave in x and convex in y.

Global Optimization Algorithms for Stochastic Integer Programming

Shabbir AHMED, Mohit TAWARMALANI and Nick SAHINIDIS

Stochastic integer programming lies at the intersection of two optimization areas, integer programming and stochastic programming, each of which is computationally challenging on its own. Whereas a large amount of algorithmic work has been published in the optimization literature on integer programming and a comparable amount of work is found on stochastic programming, very little has been done for problems that are integer *and* stochastic at the same time. Nonetheless, a large number of applications in production, location, transportation, finance, and other areas demand stochastic integer programming approaches.

We develop exact algorithms for two important classes of stochastic integer optimization problems:

1. *Multi-stage stochastic integer programs.*

   Owing to the discontinuous and nonconvex nature of the value function of integer programs, upon projection of a two-stage stochastic integer problem onto the space of the first stage variables, one is confronted with a difficult discontinuous multi-extremal optimization problem. Thus, the standard Benders decomposition approaches that work nicely for multi-stage stochastic linear programs, break down when integer variables are present in stages other than the first one. We develop a continuous global optimization algorithm for this problem. Key to the development of this algorithm is the observation that a variable transformation leads to a value function that possesses discontinuities only orthogonal to variable axes. This allows for a natural partitioning scheme that results into a finite branch and bound algorithm for this class of mathematical programs.

2. *Stochastic programs with discrete decisions and decision-dependent uncertainties.*

   A fundamental assumption in standard stochastic programming approaches is that the distributions of the uncertain parameters are independent of the decision variables. There have been limited attempts in the literature to address the case of decision-dependent uncertainties. This is the subject of the second part of this work. We show that a class of single-stage stochastic programs with discrete decisions and decision-dependent uncertainties can be formulated as $0-1$ hyperbolic programs. We develop lower bounding schemes and optimization algorithms for these programs. This work involves developing characterizations of convex extensions of lower semicontinuous functions of $0-1$ variables.

In preliminary computational experiments, we show that the proposed exact algorithms solve some heretofore open problems from the stochastic integer programming literature.

A Generalized Version of the Auxiliary Problem Method for Paramonotone Multivalued Mappings

Geneviève SALMON, Van Hien NGUYEN and Jean-Jacques STRODIOT

Let F be a maximal monotone multivalued operator in a Hilbert space H and let $\varphi : H \to I\!R \cup \{+\infty\}$ be a lower semicontinuous proper convex function. We consider the following general variational inequality problem:

$$(P) \begin{cases} \text{find } x^* \in H \text{ and } r(x^*) \in F(x^*) \text{ such that, for all } x \in H, \\ \langle r(x^*), x - x^* \rangle + \varphi(x) - \varphi(x^*) \geq 0. \end{cases}$$

When F is the subdifferential operator of a finite-valued convex continuous function f defined on H, problem (P) reduces to the nondifferentiable convex optimization problem:

$$(OP) \quad \min_{x \in H} \{f(x) + \varphi(x)\}.$$

For solving problem (P), Cohen developed, several years ago, the auxiliary problem method. Let $K : H \to I\!R$ be an auxiliary function continuously differentiable and strongly convex and $\{\lambda_k\}_{k \in I\!N}$ be a sequence of positive numbers. The problem considered at iteration k is the following:

$$(P^k) \begin{cases} \text{choose } r(x^k) \in F(x^k) \text{ and find } x^{k+1} \in H \text{ such that, for all } x \in H, \\ \langle r(x^k) + \lambda_k^{-1}(\nabla K(x^{k+1}) - \nabla K(x^k)), x - x^{k+1} \rangle + \varphi(x) - \varphi(x^{k+1}) \geq 0. \end{cases}$$

Recently, this method has been coupled with a perturbation of the function φ. This allows to consider, for example, barrier functions or interior approximations of the feasible set. Convergence results emphasizing the global and local assumptions to be satisfied by these perturbed functions can be found in [1] for the singlevalued case in finite dimension. In [2], the perturbed scheme for the multivalued case has been relaxed by taking $r(x^k)$ in the enlargement of F at x^k introduced by Burachik, Iusem and Svaiter. For the optimization problem (OP) with φ the indicator function of a closed convex subset of H, this procedure has been already used in [3] to relax the classical subgradient method. In this case, the enlargement of $F = \partial f$ is the ϵ-subdifferential of f at x^k.

Convergence of the perturbed and inexact version of the auxiliary problem method is established in [2] under the key assumption that F is strongly monotone over dom φ. However, this condition is not satisfied in many applications. In this talk, we present convergence results under weaker monotonicity assumptions on F. First, when F is paramonotone on dom φ and satisfies a kind of continuity property, we obtain that at least one weak limit point of the sequence generated by the algorithm is a solution of problem (P). In general, paramonotonicity is strictly weaker than strict monotonicity and even than cocoercivity in the singlevalued case. In a second part, we prove weak convergence under a condition satisfied, for example, if F is singlevalued and paramonotone, or multivalued and strongly monotone, or is the subdifferential of a lower semicontinuous proper convex function. For this result, the enlargement that we use is included in the ϵ^k-enlargement and depends on the properties of F. Finally, we present a strong convergence result.

References

[1] *Salmon, G., Nguyen, V.H., Strodiot, J.J.* (2000) Coupling the Auxiliary Problem Principle and the Epiconvergence Theory to Solve General Variational Inequalities. J. Opt. Theory Appl. **104**, pp. 629–657.

[2] *Salmon, G., Nguyen, V.H., Strodiot, J.J.* (2000) A Perturbed and Inexact Version of the Auxiliary Problem Method for Solving General Variational Inequalities with a Multivalued Operator.

In: Nguyen, V.H., Strodiot, J-J., Tossings, P. (eds.), Optimization, Lecture Notes in Economics and Mathematical Systems **481**, pp. 396–418, Springer–Verlag, Berlin.

[3] *Alber, Ya.I., Iusem, A.N., Solodov, M.V.* (1998) On the Projected Subgradient Method for Nonsmooth Convex Optimization in a Hilbert Space. Math. Program. **81**, pp. 23–35.

Interior–point algorithms for Dantzig and Wolfe decomposition principle

Paulo Roberto OLIVEIRA and Marcos Augusto dos SANTOS

We present algorithms to solve problems originated from the application of the Dantzig and Wolfe decomposition principle to linear programming problems. The main features are the stability and the partial use of the analytic center method.

One algorithm requires the sub–problems (slaves) entirely solved at each iteration, and the other uses sub–optimal solutions. All of then solve partially the master problem at each iteration. The evolution of the bundle of information is evaluated and we use a build-down strategy to keep it in a manageable size.

A crucial aspect in a method that uses analytic center in a cutting plane scheme is the recovering of the (approximate) center when a new cut is generated. We propose a method to recover the proximity to the analytic center when the oracle generates multiple cuts.

The convergence of the algorithms is established.

Theoretical and computational results for a linear bilevel problem

Susana SCHEINBERG and Manoel CAMPÊLO

In this paper we consider a linear bilevel programming problem (LBP). We derive theoretical results concerning the existence or absence of solutions of problem (LBP). The properties we obtain are useful from a numerical point of view. We characterize global and local solutions of this problem as well as we identify unbounded and infeasible situations, which are not usually considered in the literature.

Our approach consists of examining problem (LBP) by the analysis of two related programs. The first one (P) is obtained by replacing the inner linear problem by its Karush-Kuhn-Tucker conditions. It is often used instead of (LBP), see for example [2, 4, 5]. Indeed, they are globally equivalent, in the sence that a global solution of the first problem, if it exists, corresponds to a global solution of the latter one, and vice-versa. The correspondence between these problems also holds in the cases of infeasibility and unboundedness. However, a local solution of (P) may not be associated with a local solution of (LBP) [3].

Let us note that the feasible set of problem (P) is nonconvex due to the complementary slackness constraints. So, the second auxiliary problem $P(M)$ is obtained by penalizing the complementary slackness conditions with a penalty paremeter $M \geq 0$. This nonconvex penalty function has been considered in the literature (see for example [1, 4, 5]) as an approach to identify global optima of (LBP) under restrictive assumptions.

Now, we show that this penalty is still useful in a wider context involving unbounded and infeasible situations. Also, we introduce the notion of equilibrium point of the penalized problem to characterize the local solutions of (LBP). Moreover, we present necessary and sufficient conditions for strict optimality. We point out that the penalty parameter can be implicitly considered, which is computationally attractive.

Finally, we propose algorithms to find global and local optima of (LBP). They are based on the concept of equilibrium point. We also present computational results.

References

[1] M. Amouzegar and K. Moshirvaziri (1998). A penalty method for linear bilevel programming problems. *Multilevel optimization: algorithms and applications* 251-271, Nonconvex Optimization and its Applications 20, Kluwer Acad. Plublishers.

[2] J.F. Bard (1984). An investigation of the linear three level programming problem. *IEEE Transactions on Systems, Man and Cybernetics* 14(5), 711-717.

[3] M. Campêlo and S. Scheimberg (1998) Characterizing local and global solutions of a linear bilevel problem, Tech. Report ES-551/99, *Engenharia de Sistemas e Computação, Universidade Federal do Rio de Janeiro.*

[4] H. Önal (1993). A modified simplex approach for solving bilevel linear programming problems. *European J. of Operational Research* 67, 126-135.

[5] D.J. White and G. Anandalingam (1993). A penalty function approach for solving bi-level linear programs. *J. of Global Optimization* 3, 397-419.

Stochastic perturbations methods for affine restrictions

M. BOUHADI, R. ELLAIA and Eduardo SOUZA DE CURSI

This work considers the following global optimization problem:

$$x^* = \arg\min_{C} \; J \;, \quad C \; = \{x \in E = \mathbb{R}^n \mid Ax \leq b\}$$

where J is a *regular* but *non convex* function; A is a $m \times n$ given matrix; b is $m \times 1$ given matrix. A wide range of deterministic methods for the numerical resolution of this problem has been developed and can be found in the literature. For instance, in the framework of the methods using information about the derivatives, we can consider the *projected gradient method*: an initial feasible guess $x_o \in C$ is given and a sequence $\{x_k\}_{k \geq 0} \subset C$ is generated by using iterations of the general form

$$\forall \; k \geq 0 \qquad x_{k+1} = Q_k\left(x_k\right) = x_k + \eta_k d_k \quad . \tag{1}$$

The choice of the descent direction must simultaneously maintains feasibility and improves the value of the objective function J: d_k must be a *feasible direction*, i.e., there exists $\delta > 0$ such that

$$\forall \; \eta \in [0, \delta] : \; x_k + \eta d_k \in C \quad . \tag{2}$$

The optimal choice for η_k is

$$\eta_k = \text{argmin}\left\{j_k(\eta) : 0 \leq \eta \leq \delta\right\} \quad ; \quad j_k(\eta) = J(x_k + \eta d_k) \qquad . \tag{3}$$

Fundamental difficulty arises from the lack of convexity of J : global convergence is not ensured and the sequence may converge to a point which is not a global minimum.

In order to prevent from convergence to local minima, various modifications of the basic methods have been introduced in the literature. We consider in this paper a different approach, inspired from the method of random perturbations for unconstrained minimization of continuously differentiable functions. *From the mathematical standpoint, we establish that the adjunction of suitable random perturbations leads to a globally convergent method:* the sequence of real numbers $\{x_k\}_{k\geq 0}$ is replaced by a sequence of random variables $\{X_k\}_{k\geq 0}$ involving a random perturbation P_k of the deterministic iterations (1). The following three operations are performed at at each step $k \geqslant 0$:

- *unperturbed descent:* generation of a point Y_k by using the basic descent method Q_k. We assume that Y_k is feasible, i. e., $Y_k \in C$. This assumption is satisfied by the projected gradient method.
- *perturbation:* generation of feasible trial points by using a random variable $T_k \in C$, related to the random perturbation P_k and Y_k. We assume that T_k spans Y_k, i. e., Y_k is a value from T_k. this assumption is satisfied by all perturbations introduced below.
- *dynamics*: the definition of X_{k+1}, which is usually a function of the actual feasible values $\mathcal{P}_k = \{T_k, X_k\}$. We denote this function by S_k. Then, $X_{k+1} = S_k(\mathcal{P}_k)$.

From the numerical standpoint, random perturbation algorithms start at the initial value $X_0 = x_0 \in C$. At step $k \geqslant 0$, X_k is known and X_{k+1} is determined by performing the following three operations:

- *unperturbed descent:* generate a point $Y_k = Q_k(X_k)$.
- *perturbation:* Since Y_k is a feasible point and T_k spans Y_k, we set $T_k^0 = Y_k$. In addition, we generate a NT supplementary values from T_k, denoted by $T_k^i : 1 \leq i \leq NT$. Then, $\mathcal{T}_k = \left(T_k^0, T_k^1, ..., T_k^{NT}\right)$ is a sample of the random variable T_k. This sample is used in order to approach T_k. The case $NT = 0$ corresponds to the unperturbed descent method.
- *dynamics*: the available feasible points form the set $\mathcal{P}_k = \{X_k\} \cup \mathcal{T}_k$. X_{k+1} is obtained by applying a rule S_k, *i. e.*, $X_{k+1} = S_k(\mathcal{P}_k)$.

We shall present numerical experiments involving different perturbations and two dynamics (elitist and Metropolis).

ε-Feedback design in state constrained optimal control

Ronald STERN, F. H. CLARKE and L. RIFFORD

In problems of optimal control, it is generally the case that effective continuous feedback controls do not exist. In this talk, it will be explained how nonsmooth analysis can be employed in order to produce near-optimal feedback laws in a general Mayer problem with state constraints. The discussion will include the fact that the value function for this problem satisfies a certain nonsmooth version of the Hamilton-Jacobi equation, and that this is equivalent to the known notion of constrained viscosity solution.

ε-Enlargements of maximal monotone operators in Banach spaces

Regina Sandra BURACHIK and Benar Fux SVAITER

Let A and B be arbitrary sets and $F : A \to 2^B$ a multifunction. Denote by R_+ the nonegative real numbers. By an *enlargement* or *extension* of F we mean a multifunction $E : R_+ \times A \to 2^B$ such that

$$F(x) \subseteq E(b,x) \quad \forall b \geq 0,\, x \in A.$$

A well known and most important example of extension of a multifunction is the ε-subdifferential. Given a proper convex function f on a Banach space X, $f : X \to R \cup \{+\infty\}$, the *subdifferential* of f at x, i.e., the set of *subgradients* of f at x, denoted by $\partial f(x)$, is given by

$$\partial f(x) = \{u \in X^* : f(y) \geq f(x) + \langle u, y - x\rangle \geq 0 \text{ for all } y \in X\}.$$

The ε-subdifferential enlargement (of ∂f) was introduced in [1]. It is defined as

$$\partial_\varepsilon f(x) := \{u \in X^* : f(y) \geq f(x) + \langle u, y - x\rangle \geq -\varepsilon \text{ for all } y \in X\},$$

for any $\varepsilon \geq 0$, $x \in X$. Note that $\partial f = \partial_0 f$ and in general (i.e., for $\varepsilon \geq 0$) $\partial f(x) \subseteq \partial_\varepsilon f(x)$. If f is proper closed convex, then ∂f is a maximal monotone operator (see [6]). So, we have an example of an enlargement of *certain* maximal monotone operators (those which are subdifferentials).

For an arbitrary maximal monotone operator $T : R^N \to 2^{(R^N)}$, Burachik and Svaiter [2] proposed the following enlargement: Given $\varepsilon \geq 0$ and $x \in R^N$

$$T^\varepsilon(x) = \{u \in R^N \mid \langle v - u, y - x\rangle \geq -\varepsilon,\, \forall y \in R^N, v \in T(y)\}.$$

The enlargement defined above can be formulated in a straightforward manner for operators defined on Hilbert spaces. In the case of a Banach space X, the scalar product is replaced by the usual "dual product" on $X \times X^*$[5]. Hence, for an arbitrary maximal monotone operator $T : X \to 2^{X^*}$, we consider

$$T^\varepsilon(x) = \{u \in X^* \mid \langle v - u, y - x\rangle \geq -\varepsilon,\, \forall y \in X, v \in T(y)\},$$

for any $\varepsilon \geq 0$, $x \in X$. We will discuss the basic properties of this enlargement for maximal monotone operators defined in Banach spaces.

An important property concerning any extension of a maximal monotone operator T is whether an element in the graph of the extension of T can be approximated by an element in the graph of the original operator. This question has been successfully solved for the extension $\partial_\varepsilon f$ by Brøndsted & Rockafellar in [1]: Given $\varepsilon \geq 0$ and $v_\varepsilon \in \partial_\varepsilon f(x_\varepsilon)$, for any $\eta > 0$, there exists x and $v \in \partial f(x)$ such that

$$\|x - x_\varepsilon\| \leq \varepsilon/\eta, \quad \|v - v_\varepsilon\| \leq \eta.$$

We say that an enlargement E of T satisfies the *Brøndsted & Rockafellar property* if any element in the graph of E can be approximated in a similar way, i.e.: Given $b \geq 0$, $v_b \in E(b, x_b)$, for any $\eta > 0$ there exists x and $v \in T(x)$ such that

$$\|x - x_b\| \leq \varepsilon/\eta, \quad \|v - v_b\| \leq \eta.$$

When X is a Hilbert space, it has been proved in [3], that the enlargement defined above satisfies the Brøndsted & Rockafellar property. If X is a reflexive Banach spac, T^ε has still this property.

Another key question concerning an extension of T is how to construct an element of the extension using elements of T. For the extension $\partial_\varepsilon f$, the tool for doing this is the so-called "transportation formula" For the extension T^ε, an "alternative" transportation formula also holds.

Finally, the application $(\varepsilon, x) \hookrightarrow T^\varepsilon(x)$ is locally bounded by an affine function of ε on the interior of the domain of T, as well as Lipschitz continuous, for T defined in a Banach space.

Algorithmic applications of this enlargement will be discussed.

References

[1] Brøndsted, A. and Rockafellar, R.T.: On the subdifferentiability of convex functions, *Proceedings of the American Mathematical Society* **16** (1965), 605–611.

[2] Burachik, R.S., Iusem, A.N. and Svaiter, B.F.: Enlargements of maximal monotone operators with application to variational inequalities, *Set Valued Analysis* **5** (1997), 159–180.

[3] Burachik, R.S., Sagastizábal, C.A. and Svaiter, B.F.: epsilon-Enlargement of Maximal Monotone Operators with Application to Variational Inequalities, in *Reformulation - Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods.* Fukushima, M. and Qi, L. (editors), Kluwer Academic Publishers, 1997.

[4] Burachik, R.S., Sagastizábal, C.A. and Svaiter, B.F.: Bundle methods for maximal monotone operators, in *Ill–posed Variational Problems and Regularization Techniques.* Tichatschke, R. and Thera, M. (editors), LNCIS Springer, 1999 (to appear).

[5] Burachik, R.S., Svaiter, B.F.: "epsilon-Enlargements of maximal monotone operators in Banach Spaces" *Set-Valued Analysis,* 7(2):117-132, June 1999.

[6] Rockafellar, R.T.: On the maximal monotonicity of subdifferential mappings, *Pacific Journal Math.* **33** (1970), 209–216.

Approximate Convex Functions

Michel THERA

The class of Lipschitz convex functions on a Banach space possesses the following important properties: *(a) it is stable under finite sums and finite suprema; (b) if ∂f stands for the classical subdifferential of convex functions then the optimality condition $0 \in \partial f(x)$ implies that, for every $\epsilon > 0$, x is a local minimum of the functions $y \to f(y) + \epsilon\|y - x\|$; (c) equality holds for the sum rule: $\partial(f_1 + f_2)(x) = \partial f_1(x) + \partial f_2(x)$; and (d) the integration property holds: if $\partial f_1(x) = \partial f_2(x)$ for all $x \in X$, then $f_1 - f_2$ is a constant.*

As was pointed out by Ioffe in [1], it is a challenging problem to find out whether the above class can be extended to nonconvex functions so that it still verifies properties (a)-(d) with a suitably choosen subdifferential, and contains all continuously differentiable functions. Of course, if such an extended class exists, it cannot contain all Lipschitz functions because most of the known subdifferentials do not satisfy equality in the sum rule for these functions. A smaller class, consisting of Lipschitz and primal lower nice functions [Poliquin 3], [Thibault-Zagrodny 4] on a Hilbert space also verifies properties (a)-(d), but it does not contain all continuously differentiable functions.

The purpose of the present talk is to introduce a new class consisting of generalized convex functions on a Banach space, called approximate convex functions, which meets the above requirements. The main feature of these functions is twofold: (1) they include the convex functions as well as the continuously differentiable functions;

(2) in this class, the subdifferential in the sense of Clarke and Mordukhovich coincides with the Ioffe geometric subdifferential, as well as with the Ioffe approximate subdifferential when the functions are Lipschitz.

The so-called ϵ-convex functions introduced by Jofré-Luc-Théra [2] serve as the main tool for defining approximate convex functions and deriving their properties.

References

[1] Ioffe A. D., *Approximate Subdifferential and Application, Part 1,* Trans. Amer. Math. Soc., **281**,1,389-416,1984.

[2] Jofré A., Luc D. T., Théra M., *$\epsilon-$Subdifferential and $\epsilon-$Monotonicity,* Nonlinear Analysis, **33**, 71-90, 1998.

[3] Poliquin R. A., *Integration of Subdifferentials of Nonconvex Functions*, Nonlinear Analysis, **17**, 385-398, 1991.

[4] Thibault L., Zagrodny D., *Integration of Subdifferentials of lower semicontinuous Functions on Banach spaces,* J. of Math. Anal. and Appl., **189**, 33-58, 1995.

Polarity in DEA models

Jørgen TIND

The models in data envelopment analysis (DEA) compare the performance of decision making units subject to some assumptions about the *possibility space* of input-output vectors. The original models by Banker [1] and Banker, Charnes and Cooper [2] assume free disposability and convexity of the entire production possibility space. These models have been widely used, even though the convexity assumption has been criticized for being too strict. The model proposed by Tulkens [7] only assumes free disposability which implies that comparison is only possible via domination. Petersen proposed in [6] a model requiring convexity of the projections into the input and output space, respectively. This assumption does not in general imply convexity of the possibility space and represents therefore a relaxation of the original convexity assumptions. Further work along this line has been done by Bogetoft [3].

Bogetoft et al. [4] proposes a method which subject to free disposability constructs the smallest possibility set having convex projections in input and output space. It operates with a selection of pairs of input-output vectors such that any vector of a pair constitutes a feasible input-output combination. The pairs are formed by a sequence of alternating intersection and convex hull operations.

For a given input vector the convex projection in the output space is obtained as the largest output part from all pairs containing the input vector. By construction this consists of all outputs that may obtained by the given input and is therefore called the *production set* by standard DEA terminology. Symmetrically one may also construct the *consumption set* consisting of all inputs from which a given output vector may be reached.

The present analysis gives a dual description of the production and consumption sets. The blocking and antiblocking theory, originally developed by Fulkerson [5] for the analysis of certain problems in extremal combinatorics, appears to offer an appropriate framework. This is due to the assumption about free disposability which implies that a consumption set undertakes the form of a blocking polyhedron and a production set the form of an anti-blocking polyhedron. This establishes an involutory polarity which gives additional flexibility by enabling an interchange of the above intersection and convex hull operations as well as an economic interpretation.

In DEA the efficiency of a specific decision making unit may be measured by a productivity index which subject to the usual convexity assumptions is calculated by linear programming. An extended linear programming model is proposed for the calculation of a productivity index subject to the present relaxed convexity assumptions.

References

[1] Banker, R.D., "Estimating Most Productive Scale Size Using Data Envelopment Analysis", *European Journal of Operational Research* 17 (1984) 35–44.

[2] Banker, R., A. Charnes, and W. Cooper, "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis", *Management Science* (1984) 1078–1092.

[3] Bogetoft, P., "DEA on Relaxed Convexity Assumptions", *Management Science* 42 (1996) 457 – 465.

[4] Bogetoft, P., J. Tama and J. Tind, "Convex Input and Output Projections of Nonconvex Production Possibility Sets", Working Paper, Department of Operations Research, University of Copenhagen, Denmark, revised 1999. To appear in *Management Science*.

[5] Fulkerson, D.R., "Blocking and Anti-blocking Pairs of Polyhedra", *Mathematical Programming* 1 (1971) 168 - 194.

[6] Petersen, N.C., "Data Envelopment Analysis on a Relaxed Set of Assumptions", *Management Science* 36 (1990) 305 – 214.

[7] Tulkens, H., "On FDH Efficiency Analysis: Some Methodological Issues and Applications to Retail Banking, Courts, and Urban Transit", *The Journal of Productivity Analysis* 4 (1993) 183–210.

A GRASP for computing approximate solutions for the Three-Index Assignment Problem

R.M. AIEX, P.M. PARDALOS, M.G.C. RESENDE and Geraldo TORALDO

A greedy randomized adaptive search procedure (GRASP) is presented for computing approximate solutions for the NP-hard three-index assignment problem (AP3). Extensive computational testing of this GRASP is done on several problem instances of the AP3. A hybrid strategy that combines GRASP and path relinking is also proposed and tested for the problem. In the hybrid strategy, a pool of elite solutions is kept to be used as targets in the path relinking phase. Computational results indicate that the GRASP and the hybrid heuristic provide good approximate solutions to a variety of AP3 instances.

On Convexity Recognition of the Union of Polyhedra

Alberto BEMPORAD, Komei FUKUDA and Fabio D. TORRISI

In this paper, we study the following polyhedral computation problem. Let P and Q be convex polyhedra in the d-dimensional Euclidean space $\mathbb{R}^d$. Our problem is to check whether their union $P \cup Q$ is convex. To be more precise, this problem needs to be specified further depending on the form of input. We shall consider three natural cases: (1) when P and Q are H-polyhedra (given by inequalities), (2) when they are V-polyhedra (given by vertices) and (3) when they are VH-polyhedra (given by both), where we set the output representation to be the same as the input representation. We also investigate the extension of the problems to the union of k polyhedra.

While our problem of efficient convexity recognition does not seem to have been investigated before, there are various studies on the union of polyhedra. Franklin [3] studied union and intersection operations of 2 and 3 dimensional nonconvex polyhedra. Aronov, Sharir and Tagansky [1] studied the complexity of the union of k convex 3-polyhedra and proposed a randomized algorithm for computing it. Balas [2] presented a nice formulation of an H-polyhedra containing the convex hull of several H-polyhedra for the purpose of solving the mixed integer 0/1 program.

Our algorithms are based on the following characterizations of the convexity of the union of polyhedra.

Theorem 2 (H-Theorem) *For two H-polyhedra $P = \{x \in \mathbb{R}^d : \ Ax \leq \alpha\}$ and $Q = \{x \in \mathbb{R}^d : Bx \leq \beta\}$,*

$$P \cup Q \text{ is convex } \Leftrightarrow P \cup Q = \mathrm{env}(P, Q),$$

where $\mathrm{env}(P,Q) := \{x \in \mathbb{R}^d : \ \overline{A}x \leq \overline{\alpha}, \overline{B}x \leq \overline{\beta}\}$, *$\overline{A}x \leq \overline{\alpha}$ is the subsystem of $Ax \leq \alpha$ obtained by removing from all inequalities violated by Q, and $\overline{B}x \leq \overline{\beta}$ is defined in the same way w.r.t. P.*

Theorem 3 (V-Theorem) *Let P, Q be V-polytopes with V-representations V_P and V_Q. Then*

$$P \cup Q \text{ is convex } \Leftrightarrow [v, w] \subseteq P \cup Q, \ \forall v \in V_P, \ \forall w \in V_Q.$$

Moreover, the stronger statement holds:

$$P \cup Q \text{ is nonconvex } \Leftrightarrow \exists v \in V_P, w \in V_Q \text{ such that } \ (v, w) \cap (P \cup Q) = \emptyset,$$

where $[v, w]$ and (v, w) represent the closed and the open line segments connecting v and w.

It follows quite naturally from the two theorems that the convexity recognition of the union of two H-(V-)polyhedra can be verified in polynomial time via linear programming [4]. For the H-polyhedra case, one can generalize H-Theorem and show the polynomial solvability for any fixed number of H-polyhedra. On the other hand, we do not know of any natural extension of V-Theorem even for three polytopes and thus the efficient solvability is still open for k V-polyhedra with $k \geq 3$. Finally we present a strongly polynomial algorithm for two VH-polyhedra.

References

[1] B. Aronov, M. Sharir, and B. Tagansky. The union of convex polyhedra in three dimensions. *SIAM J. Comput.*, 26:1670–1688, 1997.

[2] E. Balas. On the convex hull of the union of certain polyhedra. *Operations Research Letters*, 7:279–283, 1988.

[3] Wm. R. Franklin. Efficient polyhedron intersection and union. In *Proc. Graphics Interface*, pages 73–80, 1982.

[4] A. Schrijver. *Theory of Linear and Integer Programming.* Wiley, 1986.

Robust and Interior Point Optimization Methods in Support Vector Machine Training

Theodore B. TRAFALIS

Although the subject of support Vector Machines (SVMs) can be said to have started in the late seventies it is only now receiving increasing attention in the machine learning and computer vision research communities. The reason is that we want to analyse massive data and SVMs turn out to be an effective tool for this kind of analysis. Unfortunately, SVMs are not popular yet in the engineering community. A basic reason is that the training of SVMs is slow, especially for large problems.

Recently a lot of attention has been given to applications of mathematical programming to machine learning and neural networks. There are several challenges for mathematical programming. In this paper we consider the following two issues related to the use of mathematical programming techniques to SVMs:

1. Scaling to large sets of constraints and variables.

Our emphasis will be given to Interior Point Methods (IPMs) and how they can be used for efficient training of SVMs with a huge amount of data. Specifically, training a SVM is equivalent to solving a linearly constrained quadratic programming (QP) problem in a number of variables equal to the number of data points. This problem becomes quite challenging when the size of the data becomes of the order of some thousands. IPMs have been shown quite promising for solving large scale linear and quadratic programming problems. Since the training of SVMs can be cast as a quadratic programming problem it seems natural to investigate the use of IPMs for training SVMs. IPMs have been already used by the author for training artificial neural networks (ANNs) with good results. The training of ANNs is a highly nonconvex optimization problem in contrast to the SVMs training problem which is a convex optimization problem. Therefore IPMs can be used quite effectively in SVM training.

There are several classes of IPMS from which to choose to solve the SVMs QP problem.Some preliminary work using primal-dual methods has been investigated by Smola. He solved moderately sized problems up to approximately 3,000 samples. He didn't exploit the special structure of the problem which is more appropriate for column generation interior point techniques. Here it is proposed to use potential reduction and column generation techniques. Investigation of the special structure of the SVM QP problem will be examined by using decomposition techniques of mathematical programming. Linear Programming formulations of the SVM training problem are also examined.

2.Modeling noise. We investigate the robust convex optimization approach to uncertainty (Ben-Tal and Nemirovski with applications to SVM learning. This approach deals with a "decision environment" which is characterized by crude knowledge of the uncertainty: all that is known about the uncertain data vector v is that it belongs to a given uncertainty set U. We show that in the case that the set U is an ellipsoid the SVM learning is equivalent to a convex quadratic optimization problem with conic quadratic constraints.

Semidefinite programming with sparse matrices

Stephen VAVASIS

Semidefinite programming has emerged as a powerful approach to approximating optima of combinatorial optimization problems. The original discovery in the regard was the Goemans-Williamson [3] semidefinite relaxation for the MAX-CUT problem. The MAX-CUT problem is: Given an undirected graph G, find the partition (U, V) of the nodes that maximizes the number of edges crossing from U to V. Although this problem is NP-hard, Goemans and Williamson showed that in polynomial time one can compute a solution whose objective is 0.878 (or better) times the optimal objective.

In many applications of this and other combinatorial problems, the underlying graph is sparse. In particular, if the graph has n nodes, then the number of edges is expected to be much less than n^2. But unfortunately, all known polynomial-time algorithms for semidefinite programming require formation of a fully-populated $n \times n$ matrix followed by solution of linear systems on this matrix. The inability to take advantage of sparsity has diminished the attraction of semidefinite programming.

In this talk I will explore and compare several recent approaches to preserving sparsity in semidefinite relaxation of sparse combinatorial problems. One approach [4] uses the Erisman-Tinney idea [2] of sparse inverses to compute the gradient of the SDP dual barrier function. This algorithm requires time and space proportional to the time and space proportional of Cholesky factorization of the original sparse matrix. For many classes of sparse graphs, the Cholesky factor is also very sparse. This algorithm so far not known to be polynomial time because the gradient by itself is insufficient for the standard Newton step of semidefinite programming.

A second approach is due to Burer et al. [1]. This approach parametrizes the set of positive semidefinite matrices in terms of entries of the Cholesky factor. The original SDP problem is now a nonconvex optimization problem in the smaller number of variables. Their algorithm is also not known to be polynomial time.

References

[1] S. Burer, R. D. C. Monteiro, and Y. Zhang. Solving semidefinite programs via nonlinear programming part I: Transformations and derivatives. Technical Report TR99-17, Department of Computational and Applied Mathematics, Rice University, 1999.

[2] A. Erisman and W. Tinney. On computing certain elements of the inverse of a sparse matrix. *Comm. ACM*, 18:177–179, 1975.

[3] M. X. Goemans and D. P. Williamson. Improved approximation algorithms for maximum cut and satisfiability using semidefinite programming. *J. ACM*, 42:1115–1145, 1995.

[4] S. Vavasis. A note on efficient computation of the gradient in semidefinite programming. Available on-line from `ftp: // ftp.cs.cornell.edu / pub / vavasis / papers / fastgrad2.ps.gz`, 1999.

Global Search Methods for Efficient Neural Network Training

V. P. PLAGIANAKOS, G. D. MAGOULAS and M. N. VRAHATIS

In neural network batch training the objective is to minimize a cost function defined as the multi–variable error function of the network. In particular, supervised training of a feed–forward neural network (FNN) can be viewed as the minimization of an error function that depends on the weights of the network. This perspective gives some advantage to the development of effective training algorithms, because the problem of minimizing a function is well studied in the field of numerical analysis.

The commonly used training methods are gradient based algorithms such as the widely used back–propagation (BP) algorithm [9]. These are local minimization methods and have no mechanism that allows them to escape the influence of a local minimum. The existence of local minima is due to the fact that the error function is the superposition of nonlinear activation functions that may have minima at different points, which sometimes results in a nonconvex error function. When a learning rule is trapped in a local minimum of the error function, lead to slow training and often yields suboptimal solutions [3].

In this work global search methods for feedforward neural network batch training are investigated. These methods are expected to lead to "optimal" or "near-optimal" weight configurations by allowing the network to escape local minima during training. Of course, escaping from local minima is not always possible, however these methods exhibit a better chance in locating appropriate solutions. Specifically, in this contribution various techniques are analyzed that alleviate the problem of occasional convergence to local minima in neural network training, such as simulated annealing [1, 2, 4] and a recently proposed deflection procedure [5] as well as genetic [6] and evolutionary algorithms [7, 8]. Simulations results and comparisons with the standard backpropagation algorithm and some well known modifications are exhibited. Preliminary results on notorious for their local minima problems are promising.

References

[1] M. Burton Jr., G.J. Mpitsos, "Event dependent control of noise enhances learning in neural networks", Neural Networks, 5, 627–637, (1992).

[2] A. Corana, M. Marchesi, C. Martini, S. Ridella, "Minimizing multimodal functions of continuous variables with the Simulated Annealing algorithm", ACM Trans. Math. Soft., 13, 262–280, (1987).

[3] M. Gori, A. Tesi, "On the problem of local minima in backpropagation", IEEE Trans. Pattern Analysis and Machine Intelligence, 14, 76–85, (1992).

[4] S. Kirkpatrick, C.D. Gelatt Jr., M.P. Vecchi, "Optimization by simulated annealing", Science, 220, 671–680, (1983).

[5] G.D. Magoulas, M.N. Vrahatis, G.S. Androulakis, "On the alleviation of local minima in backpropagation", Nonlinear Analysis, Theory, Methods and Applications, 30, 4545–4550, (1997).

[6] Z. Michalewicz, "Genetic algorithms + data structures = evolution programs", Springer, (1996).

[7] V.P. Plagianakos and M.N. Vrahatis, Training Neural Networks with 3-bit Integer Weights, In: Proceedings of Genetic and Evolutionary Computation Conference (GECCO'99), Orlando, 910–915, (1999).

[8] V.P. Plagianakos and M.N. Vrahatis, Neural Network Training with Constrained Integer Weights, In: Proceedings of the Congress On Evolutionary Computation (CEC'99), Washington D.C., 2007–2013, (1999).

[9] D.E. Rumelhart, G.E. Hinton, R.J. Williams, "Learning internal representations by error propagation", Parallel Distributed Processing: Explorations in the Microstructure of Cognition 1, D.E. Rumelhart, J.L. McClelland Eds., MIT Press, 318–362, (1986).

Recently Proposed Global Optimization Methods for Effective Neural Network Training

K.E. PARSOPOULOS, V.P. PLAGIANAKOS, G.D. MAGOULAS and M. N. VRAHATIS

Neural networks are usually trained using local, gradient–based algorithms [3, 4, 5]. Such methods frequently find suboptimal solutions, as they get trapped in local minima. Optimization of neural structure and global optimization methods applied to neural networks have strong influence on all aspects of network performance.

Our will for knowing the optimal weights to be used in a neural network is fulfilled by using global optimization schemes suitable for neural network training. We need to say "suitable" because, except the training performance, one method will promote other "good" features of the network too, such as generalization and lower requirements for memory and computational power.

Recently, genetic and evolutionary algorithms [6, 7], tabu search [1] and particle swarm [2] are frequently combined with neural network methods to select the best architecture and avoid drawbacks of local optimization training methods. Many other global optimization methods are also suitable for that purpose, although they are used rather rarely in this context. In this paper we provide a survey of global methods suitable for neural network training and compare the most promising among them.

References

[1] R. Batiti and G. Tecchiolli, Training Neural Nets with the Reactive Tabu Search, *IEEE Trans. on Neural Networks*, **6**, 1185–1200, (1995).

[2] R.C. Eberhart and Y. Shi, Particle Swarm Optimization, In: Proceedings of International Conf. on Neural Networks and the Brain, Beijing, China, PL5–PL13, (1998).

[3] R.A. Jacobs, Increased rates of convergence through learning rate adaptation. *Neural Networks*, **1**, 295–307, (1988).

[4] D.E. Rumelhart, G.E. Hinton, & R.J. Williams, Learning internal representations by error propagation, Parallel Distributed Processing: Explorations in the Microstructure of Cognition 1, D.E. Rumelhart, J.L. McClelland Eds., MIT Press, 318–362, (1986).

[5] T.P. Vogl, J.K. Mangis, J.K. Rigler, W.T. Zink, & D.L. Alkon, Accelerating the convergence of the back-propagation method, *Biological Cybernetics*, **59**, 257–263, (1988).

[6] V.P. Plagianakos and M.N. Vrahatis, Training Neural Networks with 3-bit Integer Weights, In: Proceedings of Genetic and Evolutionary Computation Conference (GECCO'99), Orlando, 910–915, (1999).

[7] V.P. Plagianakos and M.N. Vrahatis, Neural Network Training with Constrained Integer Weights, In: Proceedings of the Congress On Evolutionary Computation (CEC'99), Washington D.C., 2007–2013, (1999).

Global Learning Rate Adaptation in Stochastic Neural Network Training

G.D. MAGOULAS, V.P. PLAGIANAKOS and M. N. VRAHATIS

The efficient supervised training of Feedforward Neural Networks (FNNs) is a subject of considerable ongoing research and numerous algorithms have been proposed to this end. Applications of supervised training can be divided into two categories: batch training and on–line training.

Batch training is consistent with the theory of unconstrained optimization, since the information from all the training set is used. It can be viewed as the minimization of the error function E; that is to find a minimizer $w^* = (w_1^*, w_2^*, \dots, w_n^*) \in \mathbb{R}^n$, such that:

$$w^* = \min_{w \in \mathbb{R}^n} E(w),$$

where E is the batch error measure defined as the sum–of–squared–differences error function over the entire training set. This minimization corresponds to updating the weights by epoch and to be successful it requires a sequence of weight iterates $\{w^k\}_{k=0}^{\infty}$, where k indicates epochs, which converges to a minimizer w^*. The rapid computation of such a minimizer is a rather difficult task since, in general, the number of network variables is large and the corresponding nonconvex error function possesses multitudes of local minima and has broad flat regions adjoined with narrow steep ones. Two of the most used algorithms of this class are the batch backpropagation (BP) [2] and the batch adaptive BP with adaptive momentum [7].

On–line training is probably faster for large training sets and FNNs, it helps escaping local minima and provides a more natural approach for learning non–stationary tasks. However, it suffers from several drawbacks such as sensitivity to learning parameters [3]. Given the inherent efficiency of stochastic gradient descent various schemes have been proposed recently [1, 3, 4, 5, 6].

In this paper a method for adapting a global learning rate, i.e. a common learning rate for all the network weights, for on–line training is presented. The proposed technique derives from the global learning rate adaptation method proposed in [1]. The main modification of our approach consists in taking into consideration previously computed pieces of information regarding the learning rate adaptation procedure.

The proposed algorithm has been implemented and tested in various problems with large data sets and networks. Our experience is that the algorithm behaves predictably and reliably and the results were quite satisfactory.

References

[1] L.B. Almeida, T. Langlois, J.D. Amaral, A. Plankhov, Parameter adaptation in Stochastic Optimization, In: On–line Learning in Neural Networks, D. Saad, ed., 111–134, Cambridge University Press, 1998.

[2] D.E. Rumelhart and J.L. McClelland, Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1: Foundations, MIT Press, 1986.

[3] D. Saad, *On–line Learning in Neural Networks*, Cambridge University Press, 1998.

[4] N.N. Schraudolph, Online Local Gain Adaptation for Multi–layer perceptrons, Technical Report, IDSIA-09-98, IDSIA, Lugano, Switzerland, 1998.

[5] N.N. Schraudolph, Local Gain Adaptation in Stochastic Gradient Descend, Technical Report, IDSIA-09-99, IDSIA, Lugano, Switzerland, 1999.

[6] R.S. Sutton, Adapting Bias by Gradient Descent: an Incremental Version of Delta–Bar–Delta, In: *Proc. of the Tenth National Conference on Artificial Intelligence*, MIT Press, 171–176, 1992

[7] T.P. Vogl, J.K. Mangis, J.K. Rigler, W.T. Zink and D.L. Alkon, Accelerating the Convergence of the Back–propagation Method, Biological Cybernetics, **59**, 257–263, 1988.

Characteristic inclusions for a class of Hamilton-Jacobi equations

Peter R. WOLENSKI

We consider a class of optimal control problems $(P(\tau,\xi))$ of the form

$$\min \quad g\big(x(0)\big) \;+\; \int_0^\tau L\big(x(t),\dot{x}(t)\big)\,dt,$$

where the minimization is over absolutely continuous arcs $x(\cdot)$ satisfying $x(\tau)=\xi$. We focus on the *Fully Convex* case in which the Langrangian $L : R\times R \to (-\infty,\infty]$ is convex in both variables (x,v) jointly, which is equivalent to saying the Hamiltonian defined by

$$H(x,y) := \sup\big\{\langle v,y\rangle - L(x,v)\big\}$$

is concave/convex. If $g(\cdot) : R \to (-\infty,\infty]$ is assumed to be convex, then it is immediate that the value function $V(\tau,\cdot)$ is also convex. Our main result is that the set of endpoints of all the trajectories $\big(x(\cdot),y(\cdot)\big)$ for the Hamiltonian inclusion

$$\big(-\dot{y}(t),\dot{x}(t)\big) \in \partial H\big(x(t),y(t)\big) \quad \text{a.e.}\, t\in[0,\tau]$$

with inital values $\big(x(0),y(0)\big) \in \operatorname{gr}\partial g(\cdot)$ $(:= \{\xi,\eta) : \eta \in \partial g(\xi)\})$ is precisely $\operatorname{gr} V(\tau,\cdot)$, the graph of $\partial V(\tau,\cdot)$. This is a direct generalization of the classical method of characteristics, but holds globally due to the convexity assumpions.

A Fast Newton's Algorithm for Entropy Maximization for Solving the Phase Problem in Protein X-ray Crystallography

Zhijun WU, **George PHILLIPS**, **Richard TAPIA** and **Yin ZHANG**

A long-standing problem in X-ray crystallography, known as the phase problem, is to determine the phases for a large set of complex variables, called structure factors, given their magnitudes obtained from X-ray diffraction experiments. We introduce a statistical phase estimation approach to the problem. This approach requires solving a special class of entropy maximization problems repeatedly to obtain the joint probability distribution of the structure factors. The entropy maximization problem is an infinite dimensional convex program, but can be solved in a finite dual space by using a standard Newton's method. The Newton's method converges quadratically, but is costly in general, requiring $O(n^3)$ floating point operations in every iteration, where n is the number of variables. We describe a fast Newton's method which can exploit the problem structure and solve the problem with much less computational cost. More specifically, in this method, a Sherman-Morrison formula is used to reduce the inverse of the Hessian to mainly the inverse of a special matrix called Karle-Hauptman matrix in X-ray crystallography. In terms of matrix theory, a Karle-Hauptman matrix is a convolution matrix. Its inverse can be computed by using fast Fourier transform. It is well known that a fast Fourier transform can be carried out in $O(n \log n)$ floating point operations. Therefore, the fast Newton's method requires only $O(n \log n)$ floating point operations for each of its iterates, yet has the same convergence rate as the standard Newton. We describe the algorithm and related computational issues. Numerical results on a simple test case will also be presented to demonstrate the behavior of the algorithm.

Positive Semidefinite Relaxations for Distance Geometry Problems

Yasutoshi YAJIMA

In this paper, we consider the distance geometry problem for determining r–dimensional position $v_1, v_2, \ldots, v_n \in R^r$ of n points from a given set of the distances between the points. This problem is a fundamental problem in molecular biology for finding the structure of proteins from NMR (Nuclear Magnetic Resonance) data.

A distance geometry problem is specified by a subset $\mathcal{S}$ of pairs of points and by the distances d_{ij} between points i and j for $(i,j) \in \mathcal{S}$. We formulate the problem as the minimization of an error function defined by the sum of the absolute differences, i.e.,

$$f(V) = \sum_{(i,j)\in\mathcal{S}} w_{ij} \left| \|v_i - v_j\|^2 - d_{ij}^2 \right|, \tag{1}$$

where

$$V = \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n^T \end{bmatrix} \in R^{n\times r},$$

and w_{ij} are positive weights.

We note that the function (1) is non-convex and may have many local minimizers. Therefore, finding a global minimizer of $f(V)$ is a global optimization problem.

It is obvious to see that a global minima of $f(V)$ is obtained by solving the following problem:

$$\left|\begin{array}{ll} \text{Minimize} & \sum\limits_{(i,j)\in\mathcal{S}} w_{ij}\left(2t_{ij} - g_{ij}(V) + d_{ij}^2\right) \\ \text{Subject to} & -t_{ij} + g_{ij}(V) \le d_{ij}^2, \\ & t_{ij} \ge 0, \end{array}\right. \tag{2}$$

where $g_{ij}(V) = \|v_i - v_j\|^2$. Since $g_{ij}(V)$ are convex, (2) is a concave minimization problem over a convex region.

Associated with (2), let us consider the following positive semidefinite problem:

$$\left|\begin{array}{ll} \text{Minimize} & F(X) = \sum\limits_{(i,j)\in\mathcal{S}} w_{ij}\left(2t_{ij} - \langle A_{ij}, X\rangle + d_{ij}^2\right) \\ \text{Subject to} & -t_{ij} + \langle A_{ij}, X\rangle \le d_{ij}^2, \\ & t_{ij} \ge 0, \quad X \succeq 0, \end{array}\right. \tag{3}$$

where X is a variable of a symmetric matrix of order n, and for each $(i,j) \in \mathcal{S}$, A_{ij} is defined by $A_{ij} = (e_i - e_j)(e_i - e_j)^T$.

Let X^* be an optimal solution of (3). We have shown that $F(X^*)$ gives a lower bound of the global minimum of $f(V)$, and that X^* generates a global minimizer of $f(V)$ if $rank(X^*) \le r$, which implies that one can obtain an optimal solution of (2) when $r = n - 1$.

We will show computational results on various types of problems.

New properties of uniformly convex and uniformly smooth convex functions

Constantin ZĂLINESCU

Uniformly convex functions were introduced by E.S. Levitin & B.T. Polyak (1966) generalizing strongly convex functions introduced by B.T. Polyak (1966). Some properties of uniformly convex functions were established by A.A. Vladimirov, Y.E. Nesterov & Y.N. Chekanov (1978), while C. Zălinescu gave several characterizations of this class of functions. D. Azé & J.-P. Penot (1995) completed these characterizations and introduced the class of uniformly smooth convex functions. In this note we add some new characterizations for these classes of functions and show that a Banach space on which is defined at least one uniformly convex function whose domain has nonempty interior is reflexive.

References

[1] D. Azé and J.-P. Penot, *Uniformly convex and uniformly smooth convex functions*, Ann. Fac. Sci. Toulouse Math. (6) **4** (1995), no. 4, 705–730.

[2] J.-B. Hiriart-Urruty, *Optimisation et analyse convexe*, Presses Universitaires de France, Paris, 1998.

[3] E. S. Levitin and B. T. Polyak, *Convergence of minimizing of sequences in the conditional-extremum problem*, Dokl. Akad. Nauk SSSR **168** (1966), 997–1000, Russian.

[4] B. T. Polyak, *Existence theorems and convergence of minimizing of sequences in extremum problems with constraints*, Dokl. Akad. Nauk SSSR **166** (1966), 287–290, Russian.

[5] A. A. Vladimirov, Y. E. Nesterov, and Y. N. Chekanov, *On uniformly convex functionals*, Vestnik Moskov. Univ. Ser. XV Vyčisl. Mat. Kibernet. **3** (1978), 12–23, Russian.

[6] C. Zălinescu, *On uniformly convex functions*, J. Math. Anal. Appl. **95** (1983), no. 2, 344–374.

Asymptotic behaviour of the renormalised steepest descent

Anatoly ZHIGLJAVSKY, Luc PRONZATO and Henry P. WYNN

The behaviour of the standard steepest-descent algorithm for a quadratic function in R^d is investigated. We show that by rescaling the iterates to remain always on the unit sphere one can reveal special features of this behaviour. The renormalized algorithm is shown to converge to a two-point cycle on the unit circle. The cycle depends on the starting point in a complicated manner (the set of points converging to the same cycle is fractal), but all cycles belong to a particular plane, given by certain eigenvectors of the Hessian matrix of the objective function. The stability of the attractor is analysed. The rate of convergence of the algorithm is investigated. It is shown that the worst value of this rate is obtained only for some particular starting points. The introduction of a relaxation coefficient in the steepest-descent algorithm completely changes its behaviour, which may become chaotic. Different attractors are presented. We show that relaxation allows a significantly improved rate of convergence.

For a general smooth function $f(\cdot)$ the steepest-descent algorithm is

$$x^{(k+1)} = x^{(k)} - \alpha_k \nabla f(x^{(k)}) \,, \tag{1}$$

where

$$x^{(k)} = \left(x_1^{(k)}, \ldots, x_d^{(k)}\right)^T$$

is the k-th iterate of the algorithm,

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_d}\right)^T$$

is the gradient of $f(\cdot)$, and

$$\alpha_k = \arg\min_{\alpha} f(x^{(k)} - \alpha \nabla f(x^{(k)}) \,.$$

Among the other things, in [1] the authors studied the asymptotic behaviour of this algorithm for a quadratic function in R^d:

$$f(x) = \frac{1}{2}(x - x^*)^T A (x - x^*) \,,$$

where $x, x^* \in R^d$ and A is a positive definite $d \times d$ matrix. This seemingly simple problem has some very complex aspects which are uncovered by the following renormalization idea. Define

$$v^{(k)} = \frac{x^{(k)} - x^*}{||x^{(k)} - x^*||} \,, \tag{2}$$

which has the effect of rescaling the iterates to remain always on the unit sphere: $||v|| = 1$.

The authors studied the behaviour of the process $\{v^{(k)}\}$ as $k \to \infty$. Except for pathological cases, in the limit the process is found to belong to the two dimensional space spanned by the eigenvectors (u_1, u_d) corresponding to the minimal and maximum eigenvalues λ_1 and λ_d of the matrix A under the assumption

$$0 < \lambda_1 < \lambda_2 \leq \ldots \leq \lambda_{d-1} < \lambda_d \,,$$

where λ_i is the i-th ordered eigenvalue of A. Ordering the $v_i^{(k)}$ compatibly with (2), the conjecture states that $v_i^{(k)} \to 0$ for $i = 2, \ldots, d-1$ as $k \to \infty$, and that the algorithm, in its renormalised form, converges to a two-point set on the circle $\{||v_1||^2 + ||v_d||^2 = 1\}$ in the plane spanned by u_1 and u_d. Numerical simulations show that the same convergence property holds for a convex differentiable function locally quadratic around its minimum x^*. It has been known for many years that the asymptotic convergence rate of the algorithm depends on the ratio $\rho = \lambda_d/\lambda_1$. Thus for any $x^{(k)} \in R^d$

$$f(x^{(k+1)}) \leq \left(\frac{\rho - 1}{\rho + 1}\right)^2 f(x^{(k)}) \,. \tag{3}$$

The asymptotic behaviour of the steepest-descent algorithm en route to the limiting circle is fractal in nature, and it still remains open how the limiting asymptotic rate depends on the starting values. It is the worst-case rate, given by (3), not the actual rate, which simply depends on λ_1 and λ_d. However we are able to show that the algorithm attracts to two "conjugate" points on the circle. It is easy to check that *starting* on the circle the renormalized algorithm oscillates between the points, but without *attraction* to the circle this limiting behaviour is only a conjecture. We also consider the limiting behaviour of the steepest-descent algorithm with relaxation.

References

[1] Pronzato L., Wynn H., Zhigljavsky A.A. (1999) *Dynamical Search: Applications of Dynamical Systems in Search and Optimisation* , Chapman & Hall / CRC.