SIAG/OPT Views-and-News

A Forum for the SIAM Activity Group on Optimization

No. 5 Fall 1994

THE FIFTEENTH INTERNATIONAL SYMPOSIUM ON MATHEMATICAL PROGRAMMING 14-19 August 1994

by K.G. Murty

This symposium, sponsored by the Mathematical Programming Society (MPS), was held on the Central Campus of the University of Michigan (UM), Ann Arbor. It was well attended with over 1070 registrations from all over the world. There were 1050 talks in all, organized into 22 parallel streams throughout the week.

The truly international character of the symposium was helped by generous grants from the International Science Foundation that made it possible for many participants from countries of the former Soviet Union to come, and by a grant from the US National Science Foundation, which the organizers were able to use to subsidize partly the costs of participants from dollar deficient countries. Grants from General Electric Lighting and Philips Display Components also enabled six of the USA Mathematics Olympiad finalists to attend, helped support other attendees, and host receptions.

The symposium was inaugurated on Sunday evening, 14 August, with an invocation specially composed by the Pulitzer Prize winning composer William Bolcom. His composition, Haunted Laybrinth, based on B.C. Eaves' ghost story interpretation of complementary pivot algorithms for the linear complementarity problem (LCP), was set to the path of iterates for an LCP of order 5 from a paper of K.G. Murty. Welcoming addresses were given by I.B. Sheldon, Mayor of Ann Arbor; and P.M. Banks, Dean of the College of Engineering, UM. The inaugural address was given by J.K. Lenstra, Chairman of MPS.

At the opening session on Monday, 15 August, prizes were awarded. There was a special honorary prize awarded to George B. Dantzig, the father of mathematical programming, to felicitate his 80th birthday later this year. The George B. Dantzig prize sponsored by MPS and SIAM for original research having a major impact on mathematical programming, was awarded jointly to C. Lemaréchal and R.J-B. Wets. The D. Ray Fulkerson prize, sponsored by MPS and AMS for an outstanding paper in discrete mathematics, was awarded to L.J. Billera, G. Kalai, and the group consisting of N. Robertson, P. Seymour and R. Thomas. The Beale-Orchard-Hays prize, sponsored by MPS for excellence in computational mathematical programming, was awarded jointly to A.R. Conn, N.I.M. Gould, and Ph.L. Toint. The finalists of the A.W. Tucker prize, sponsored by MPS for an outstanding paper by a student, were announced to be D. den Hertog, J. Liu, and D.P. Williamson. And after the A.W. Tucker prize session held on Tuesday, 16 August, D.P. Williamson was selected as the prize winner.

We should also mention special sessions that included one organized by J.S. Pang dedicated to R.W. Cottle on the occasion of his 60th brithday, one on integer programming and economics organized by R. Gomory and H. Scarf, and several student sessions.

The symposium banquet held on Wednesday evening, 17 August, at Greenfield Village was attended by 360 people.

There were 23 tutorial talks on recent advances in various branches of mathematical programming. All these talks were outstanding and had very large audiences. Summary articles of most of these tutorial talks were edited by J.R. Birge and K.G. Murty into a 308 page book, *Mathematical Programming: State of the Art 1994*, which was distributed to all the registered participants. Others who want a copy of this book can order it by sending a $20 check, endorsed to the University of Michigan, to Ms. Ruby Sowards, IOE Department, University of Michigan, Ann Arbor, MI 48109-2117, USA. For an additional $5 they can also order the 240 page book of abstracts of all the talks, an updated list of program changes, and a full roster with addresses of all participants.

Of course, each talk at the symposium presented something new, but it has become a tradition with Mathematical Programming Symposia to have talks on some major new breakthroughs. We will briefly summarize some of the major new breakthroughs presented at this symposium.

The new subjects on which there were several talks and tutorials at this symposium are positive-definite and semidefinite programming. Beginning with a technique used by L. Lovasz to establish that the Shanon capacity

Contents:

of C_5 is $\sqrt{5}$, these subjects have mushroomed very recently into major research areas in mathematical programming through the vehicle of interior point methods. The applications of semidefinite programming in solving pure 0-1 integer problems and combinatorial optimization problems is attracting much interest. There were also many talks on interior point methods for linear, quadratic and convex programming problems, and the LCP.

The classical Newton's method figured very prominently in many talks, including a tutorial in which S. Smale summarized his recent proof that Newton's method solves polynomial systems approximately, on the average, in polynomial time.

Large scale computation was the focus of many talks, including the plenary address of W.J. Cook, and those of his co-authors, V. Chvátal and D. Applegate, in which the solution of a 7397 city traveling salesman problem using recent developments in polyhedral combinatorics, was described.

The breakthrough work on balanced 0-1 matrices discussed in the previous symposium, has been extended recently to cover balanced $0, \pm 1$ matrices. This had been the subject of a tutorial and some talks.

Of course, in this brief article it is not possible to describe all the new developments discussed at the symposium, but the interested reader is invited to review the book of abstracts.

****

CHAIRMAN'S COLUMN

by A.R. Conn

I am feeling remarkably mellow these days. Maybe the aftermath of the Ann Arbor meeting or the fact that autumn can be such a wonderful season, especially in the NorthEast.

I hope that some of you felt that the Ann Arbor business meeting of our activity group was worth attending. I certainly thought it was a very useful meeting. I was surprised to discover how unpopular some of my own biases really are — no wonder I have a tendancy to be undemocratic. For those of you that are interested, but were unable to attend, a (very brief) modified summary follows.

We have to have a slate of new officers and I asked for suggestions from the membership.

John Dennis, the founding editor, has now passed over the editorship of the SIAM Journal on Optimization to Michael Overton and both spoke at the meeting.

We had an interesting discussion about the forthcoming SIAM Optimization meeting (Victoria, Canada, May 20-22, 1996) and again I urge the membership to let us know what they want. We want to make this our best ever. Certainly the organisers of the Mathematical Programming Symposium put a great deal of effort into the Ann Arbor meeting. Some obstacles were very difficult to accommodate (for example, the incredible number of talks delivered) but I thought the venue was both excellent and relatively inexpensive.

Another issue raised at the business meeting was our activity group's prize (to be awarded for the first time in Victoria). The award is to be given to the author(s) of the most outstanding paper, as determined by the prize committee, on a topic in optimization published in English in a peer-reviewed journal. It is imperative that we obtain a reasonable number of nominations, so once again, members, we need your active support.

In my first newsletter I tried to argue forcefully why our colleagues should join SIAM, our activity group, subscribe to the journal and encourage their colleagues to do likewise. Once again, the time for renewing memberships is around the corner. To quote myself "Without these societies we would not have the journals and major meetings that are so essential for the well-being of our subject. Moreover, the societies act as significant voices for our interests with respect to governments, industry and universities. I believe they really do an excellent job".

I recently came across an article that I found both provocative and interesting. It was entitled, "Mathematics in Canada:50 years later" by Jonathan M. Borwein and Kenneth R. Davidson and is to be published in a special volume celebrating the 50^{th} anniversary of the Canadian Mathematics Society. Amongst its main theses is the opinion that NSERC (Canada's closest equivalent to NSF), although currently under significant financial and political pressures, is unequaled anywhere as a means of appropriately distributing research funds. One of the most significant differences from the American model is that NSERC does not pay salaries to investigators. They argue that one consequence of this is that NSERC is able to be much more flexible in the level of funding awarded to a given researcher and simultaneously tolerates generous funding of a few and modest funding of many, if not most, genuinely active researchers. They claim that in contrast to Americans or western Europeans, Canadians have a much higher level of autonomy and discretion in how they spend their monies. They are judged by the calibre of the research and not by unreasonable fidelity to a research program (or project) and budget produced several years earlier.

Another significant difference is that there is no 'overhead' component. Thus, each individual mathematician actually sees almost every NSERC grant dollar.

They also raise the issue of the education of mathematicians and their prospects in our society. One point they make is that North-American school (i.e. pre-university) students do poorly relative to other western countries but 'strikingly, these same students ranked very high in their *perception* of their mathematical performance'. They go on to remark that 'It is a great irony that our society uniformly celebrates elite performance in sports and music while rejecting streaming in the classroom. Who seriously argues that athletes of all ability levels should train together? Or that potential concert pianists need only the same training as a school choir?'

Finally, I would like to mention their comments on postgraduate education and jobs. They claim that "until recently, the training of graduate students has been primarily a way of perpetuating ourselves. Now, however, even very good students are no longer getting academic jobs. Current estimates are that North American universities can only absorb about one sixth of the present record number of doctorates (1,200 per annum). There is a similarly sobering estimate of industry's long–term desire for Ph.D.s".

I recommend that you take the time to read the original article. It raises many issues that are important to our profession and although you may be familar with many of them, they are worth reiterating.

As always, I welcome your opinion. My email address is arconn@watson.ibm.com

****

FORUM ESSAYS

THE TWO FACES OF THE LINEAR COMPLEMENTARITY PROBLEM

by Michael J. Todd [1]

1. TWO CLASSES OF LCP

The linear complementarity problem (LCP) is said to be a unifying problem for convex quadratic programming (including linear programming), equilibrium computation in bimatrix games, and certain problems in engineering. Here I wish to argue that it may be more helpful to think of it as a "duifying" problem: that there are in fact *two* classes of LCP that should be distinguished by the application in which they arise, by the matrix classes to which the corresponding matrices belong, and by the algorithmic approach that should be considered. Two recent books that discuss different algorithmic approaches to the LCP are Cottle, Pang, and Stone [1], which treats pivoting methods in great detail, and Kojima, Megiddo, Noma, and Yoshise [4], which covers interior-point methods.

Given a real $n \times n$ matrix M and a real n-vector q, the LCP (M, q) is to find, if possible, n-vectors x and y such that

$$y = Mx + q, \tag{1}$$

$$x \geq 0, \quad y \geq 0, \tag{2}$$

and *either*

$$x^T y = 0 \tag{3}$$

or

$$x_j y_j = 0, \quad j = 1, \cdots, n. \tag{4}$$

Of course, given that (2) holds, (3) and (4) are "really" equivalent — that is, they are equivalent for vectors x and y of real numbers. We argue that they may not be "really equivalent," i.e., equivalent as formulations of real-world problems.

Let us elucidate this point in two ways. First, $x^T y = \sum_j x_j y_j$ in (3) only makes sense if each product $x_j y_j$ is measured in the same units. (This argument also goes by the name of dimensional analysis.) For example, in the LCP arising from a linear programming problem, each product $x_j y_j$ has the units of the objective function; in

[1] School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY 14853. e-mail: miketodd@cs.cornell.edu
Research was supported in part by NSF, AFOSR, and ONR through NSF grant DMS-8920550.

a standard product mix instance, either x_j is an amount of a certain product and y_j has the units of the profit contribution of that product, or x_j is the shadow price of a scarce resource and y_j has the units of an amount of that resource. Similarly, in the LCP from a bimatrix game, each x_j has the units of a reciprocal payoff, and each y_j is dimensionless. By contrast, (4) only requires that at least one of x_j and y_j be zero for each j, with no necessity that all $x_j y_j$'s be measured in the same units.

A more abstract way to view the same distinction is as follows. We can imagine x and y as lying in abstract n-dimensional vector spaces X and Y respectively. Then (1) says that there is an affine mapping between X and Y, and y is the image of x under this mapping. Requirement (2) implies that there are distinguished bases in X and Y, determining positive cones in which x and y must lie. However, (4) makes sense even if these bases are defined only up to positive scalar multiples. (The components x_j are then the components of x in the X basis, and similarly for y.) On the other hand, (3) requires there to be a scalar product defined on $X \times Y$, or equivalently that Y is (isomorphic to) the dual space X^* of X. (In this case, $x^T y$ is better written $< x, y >$.) The distinguished bases should then be selected to be dual bases, so that a change in scale for the basis in X necessitates a corresponding change in scale for that of Y. (This abstract viewpoint was used to illustrate quasi-Newton updates in [6]; see also the appendix of Gilbert and Lemarechal [3].)

It appears that almost all applications of which the author is aware fall into the latter class, where the scalar product $x^T y$ makes sense — we call these *dual-space* applications. Certainly this is true for applications arising from optimization problems or bimatrix games. In many engineering applications, the x variables are displacements and the y variables stresses, so again all $x_j y_j$'s have the same units.

2. DIAGONAL SCALING

Let D and E be $n \times n$ scaling matrices; that is, diagonal matrices with positive diagonal entries. Then the affine scaling (or change of units) given by $x \to \bar{x} := D^{-1}x$, $y \to \bar{y} := E^{-1}y$ transforms (1)–(2) to

$$\bar{y} = \bar{M}\bar{x} + \bar{q}, \tag{5}$$

$$\bar{x} \geq 0, \quad \bar{y} \geq 0, \tag{6}$$

where

$$\bar{M} := E^{-1}MD, \qquad \bar{q} := E^{-1}q. \tag{7}$$

In the new variables, (4) is equivalent to

$$\bar{x}_j \bar{y}_j = 0, \quad j = 1, \cdots, n, \tag{8}$$

but $\bar{x}^T \bar{y}$ remains equal to $x^T y$ only if $E^{-1} = D$, so that

$$\bar{M} := DMD, \qquad \bar{q} := Dq. \tag{9}$$

Hence, for dual-space applications, matrix classes and algorithms should ideally be invariant with respect to changes in scale leading to new data given by (9). In more general applications, we might ask that the invariance be with respect to more general changes as in (7).

3. MATRIX CLASSES

Much research in linear complementarity theory studies various matrix classes; certain properties of the LCP (M, q) or of algorithms applied to it hold for M belonging to certain of these classes. Thus (M, q) has a unique solution for each q if and only if M lies in P, the class of square matrices with all principal minors positive. Similarly, Lemke's algorithm will process the LCP (M, q) (find a solution or show none exists) for M positive semi-definite (in linear complementarity theory, this is not meant to imply that M is symmetric, only that its symmetric part is positive semi-definite), or more generally copositive-plus.

It is natural to ask for each such matrix class whether it is invariant under the general scaling of (7), or the more restrictive scaling of (9). It seems that almost all classes considered in the literature are in fact invariant under the more general scaling. Thus P-matrices, as well as adequate, sufficient, semi-monotone matrices, and a host of other classes known by various letters of the alphabet, allow arbitrary left and right diagonal scaling. The LCP itself, in the form of (1),(2), and (4), is invariant under such scaling and so questions of existence and uniqueness naturally exhibit this invariance.

The "only" classes of which I am aware for which invariance is only with respect to the more restrictive scaling (9) are those of symmetric, positive semi-definite, and copositive matrices and their variants (skew-symmetric, bisymmetric, positive definite, copositive-plus, etc.). The reason for the quotes is that these are perhaps the most important classes in applications; in particular, positive semi-definite matrices recur frequently.

4. ALGORITHMS

Here I certainly do not want to get into detailed descriptions of particular algorithms. However, I want to distinguish two classes of algorithms: those that rely on pivoting and combinatorial methods (here I include iterative methods that use individual components of vectors like SOR), discussed in detail in [1]; and interior-point methods, many of which approximately follow the so-called *central path* — see [4].

Algorithms in the first class, like Cottle and Dantzig's principal pivoting method and Lemke's method, are invariant under the more general scaling (7). (The artificial vector chosen in Lemke's algorithm must also be scaled by E^{-1}, and if the algorithm searches for, say, a most negative component of y at some point, this must be relative to some vector of positive reference values, also similarly

scaled.)

On the other hand, interior-point methods often try to generate iterates that approximately satisfy

$$y = Mx + q, \quad (10)$$

$$x \circ y = \mu e, \quad (11)$$

$$x \geq 0 \quad , \quad y \geq 0, \quad (12)$$

where e denotes a vector of ones in $\Re^n$, and $x \circ y$ the vector of products $x_j y_j$. Here all the componentwise products are being set equal, so clearly such algorithms assume that they have a dual-space application. (The same holds for potential reduction methods, which use a merit function depending on $x^T y$.) By the same token, these methods are only invariant under the more restrictive scaling (9).

5. APPROPRIATE SELECTION OF ALGORITHMS

On the basis of the previous discussion, I would like to argue that interior-point methods are only suitable for dual-space applications. If they are applied to abstract problems (M, q), then they are more likely to be successful if the matrix M belongs to one of the more restrictive classes that are only invariant under the scaling (9). Indeed, the vast majority of such algorithms have been applied to LCP's with positive semi-definite matrices. In the few cases where such methods have been applied to more general matrix classes (mostly described in [4]), the results have been far less satisfactory: global convergence can be proved for column sufficient matrices, but no estimate of computational complexity is known. Similarly, the estimate of the work required to solve an LCP with M a P-matrix is rather complicated and certainly does not lead to a polynomial time bound [5].

One of the most important outstanding questions in LCP theory is indeed whether a polynomial algorithm exists in the P-matrix case. For the reasons above, I am skeptical as to the possibility of an interior-point method providing an answer to this conundrum.

An intriguing matrix class introduced by Kojima, Megiddo, Noma, and Yoshise [4] tends to blur the distinction I have been making. We say M lies in $P_*(\kappa)$ for some nonnegative κ if for every x,

$$(1 + 4\kappa) \sum_{I_+(x)} x_i (Mx)_i + \sum_{I_-(x)} x_i (Mx)_i \geq 0,$$

where $I_+(x) := \{i : x_i(Mx)_i > 0\}$ and $I_-(x) := \{i : x_i(Mx)_i < 0\}$. Then P_* is the union of all such $P_*(\kappa)$'s. (It has recently been shown that all P_*-matrices are sufficient; see Guu and Cottle [2].) Clearly, each $P_*(\kappa)$ is only invariant under the restrictive scaling (9) (note that $P_*(0)$ is just the set of positive semi-definite matrices), but Kojima et al. show that P_* is invariant under the more general scaling (7). They give a polynomial algorithm (with a bound depending polynomially on κ) for LCP's with $M \in P_*(\kappa)$, but only a globally convergent algorithm if $M \in P_*$ (the latter class includes P).

6. CAVEAT

I have been rather dogmatic above in distinguishing two classes of LCP and arguing that algorithms for the more restrictive class should not be applied in the more general setting. Here I want to soften this argument somewhat. Any application can be transformed into a dual-space application by making it dimensionless. Thus, given an arbitrary LCP (M, q) *and* positive vectors $\hat{x}$ and $\hat{y}$, the equivalent LCP $(\bar{M}, \bar{q})$ given by (7) will be dimensionless if we choose $D := \text{diag}(\hat{x})$ and $E := \text{diag}(\hat{y})$. This assumes that any change in scale is reflected in the scaling vectors $\hat{x}$ and $\hat{y}$. In the case of an interior-point method, a similar effect results if e in (11) is replaced by $\hat{x} \circ \hat{y}$, the vector of componentwise products of the "reference vectors." (Unfortunately, algorithms using such a perturbed (11) have worse complexity, roughly by the amount that $\hat{x} \circ \hat{y}$ differs from a multiple of e.) Note that many interior-point algorithms assume that such a pair of vectors (even a pair also satisfying (1)) is available.

References

[1] R.W. Cottle, J.-S. Pang, and R.E. Stone, *The Linear Complementarity Problem*, Academic Press, 1992.

[2] S.-M. Guu and R.W. Cottle, "On a subclass of P_0," to appear in *Linear Algebra and its Applications.*

[3] J.C. Gilbert and C. Lemarechal, "Some numerical experiments with variable-storage quasi-Newton algorithms," *Mathematical Programming* 45 (1989) 407–435.

[4] M. Kojima, N. Megiddo, T. Noma, and A. Yoshise, *A Unified Approach to Interior Point Algorithms for Linear Complementarity Problems*, volume 538 of *Lecture Notes in Computer Science*, Springer Verlag, Berlin, Germany, 1991.

[5] M. Kojima, N. Megiddo, and Y. Ye, " An interior point potential reduction algorithm for the linear complementarity problem," *Mathematical Programming*, 54 (1992) 267–279.

[6] M.J. Todd, "Quasi-Newton updates in abstract vector spaces," *SIAM Review* 26 (1984), 367–377.

****

CONTINUOUS-TIME LINEAR PROGRAMMING

by Malcolm C. Pullan[1]

1. INTRODUCTION

Time is an important element in many areas of optimization. This is particularly true when networks are involved, for example, in transportation, scheduling or routing. One approach to dealing with optimization problems over time is the technique of "Dynamic Programming" introduced by Bellman [7] in his well-known book of 1957. However, many people are not aware that two chapters of this book are also devoted to a production planning problem which Bellman called the "Bottleneck Problem". This problem is an extension of the standard linear program over time in that the costs, right-hand sides and variables are all functions of time instead of just vectors in R^n. Consequently Bellman hoped that an extension of the simplex algorithm could be readily obtained for its solution. However the problem is much more difficult than he first thought. Though a small example problem is solved in [7], Bellman wrote: "The analysis is decidedly difficult and it cannot be said that these problems have in any sense been tamed".

The bottleneck problem, or continuous linear program (CLP) as it is now known, emerged at a time when many other optimization problems were introduced. It is perhaps fair to say that the ones we are most familiar with today are those where significant progress was made on an algorithm for their solution, rather than those that serve as the best models for real problems. Unfortunately no algorithm was developed for CLP. As a result, it is not very well-known, even though it can be an effective model in many situations.

The absence of an algorithm for CLP is in spite of a considerable effort made by a number of authors, such as Lehman [11], Tyndall [18], Levinson [12], Grinold [10], Perold [13] and Anstreicher [6]. It is worth noting that a large part of this research on CLP has been confined to challenging (and often unpublished) Ph.D. theses. Upon completion, most researchers have decided to work on something more productive (the author recalls Anstreicher saying just that, despite CLP being a "great thesis topic"). Both Perold and Anstreicher did make considerable progress towards an algorithm, but their work is still a long way from being able to solve real problems. This put the final nail in the coffin and the prevalent feeling now is that CLP in its full generality is too hard to solve.

In 1978, Anderson [1] introduced a subclass of CLP called "Separated Continuous Linear Programs" (the term "separated" referring to the separation of integral and instantaneous constraints). The given problem data consist of a bounded measurable cost $c(t)$ (vector valued, along with all the problem data), two types of right-hand sides, one, $b(t)$, bounded measurable and the other, $a(t)$, absolutely continuous. The problem also requires two (general) matrices G and H. The variables consist of a vector of bounded measurable functions, $x(t)$, along with two slacks $y(t)$ and $z(t)$. The separated continuous linear program (SCLP) is now written as

$$\begin{aligned} \text{SCLP:}\quad \text{minimize} \quad & \int_0^T c(t)^T x(t)\,dt \\ \text{subject to} \quad & \int_0^t Gx(s)\,ds + y(t) = a(t), \qquad (1) \\ & Hx(t) + z(t) = b(t), \\ & x(t), y(t), z(t) \geq 0, \quad t \in [0,T]. \end{aligned}$$

This is an example of an infinite-dimensional linear program, as there are an infinite number of variables ($x(t)$, $y(t)$ and $z(t)$ for each t) and an infinite number of constraints (several for each time t).

The motivation for the introduction of SCLP by Anderson was to model large job-shop scheduling problems where there are many similar items to process. It then becomes convenient to treat the problem as one of rates of production of different commodities, rather than one of production of individual items. In fact, SCLP serves as a useful model for a variety of time-dependent (that is, dynamic) network problems (single- or multi-commodity, with or without side constraints) where the variables are rates of flow, the costs and demands are time varying and storage is permitted at the nodes. For instance, to model a single-commodity dynamic network program we could set G to be a node-arc incidence matrix and $a(t)$ to be the vector of total supply up to time t in each of the nodes. Then $x(t)$ could represent the rate of flow in the arcs and $y(t)$ the vector of node storages. The matrix H can then be used to represent any side constraints such as upper bounds on the rates of flow (see Anderson and Philpott [4] for a discussion of this model). It is also worth mentioning that it is possible to include storage costs in the model, however these can be removed by substituting (1) into the objective function and integrating by parts.

Examples of dynamic network problems are numerous and include water flow management (where we could be concerned with rates of flow along pipes and storage at reservoirs) or the routing of traffic through a network with the minimum of delay. It is also possible that large time-dependent discrete network problems, such as ones of transportation involving a large number of similar small items, may be effectively modelled using SCLP. This would be done by changing the problem from the flow of individual items into one of rates of flow, similarly to the modelling of job-shop scheduling problems in Anderson [1].

[1] St. John's College, Cambridge University, Cambridge CB2 1TP, United Kingdom. e-mail: mcp11@phx.cam.ac.uk

Unfortunately, as with the more general CLP, little progress was made on the solution of SCLP in the years following its introduction. Thus SCLP has never attracted much attention despite its many possible applications. However, recent work on SCLP has resulted in an algorithm for the solution of SCLP under certain assumptions on the problem data. This raises the possibility of SCLP being a useful model for various practical problems. Evolving from the algorithm has also been a large body of results that at last reveal some of the structure of the problem. These results are useful for further algorithmic development but are also interesting in their own right. For instance, the results serve to highlight the similarities between SCLP and its finite-dimensional counter-part. The purpose of this article is to outline these advances and convince the reader of the usefulness of SCLP as a model for various problems.

Before doing this we should answer the following inevitable question. Why not just discretize the problem and solve that? For the single-commodity dynamic network program this idea goes back as far as Ford and Fulkerson [9]. There are several reasons why this is not desirable and we list some of them below.

- It seems sensible, since time is continuous, to allow decisions to be made at any arbitrary time rather than fixed predetermined times.
- An accurate answer may require a large discretization and lots of computation.
- More fundamentally, it is difficult to determine the accuracy of a solution from a discretization.
- In general, solving large time-expansions is numerically difficult.
- There may be a simple answer but the discretization may not show it.
- Important theoretical properties, e.g. sensitivity, may be obscured.

2. RECENT RESULTS ON SCLP

We now give a statement of the main results that have been obtained for SCLP. From a theoretical point of view they are quite complete and are certainly more general than similar results obtained for any other class of continuous linear programs. The first result is a characterisation of the extreme points of the feasible region (or *basic feasible solutions* in normal linear programming terminology). This result is not recent (relatively speaking) but may be found in Anderson *et al.* [3].

Theorem 1 *A feasible solution,* $\omega(t)^T = (x(t)^T, y(t)^T, z(t)^T)$, *for* SCLP *is basic if and only if the columns of*

$$K = \begin{bmatrix} G & I & 0 \\ H & 0 & I \end{bmatrix},$$

corresponding to the support of $\omega(t)$ *(that is,* i *such that* $\omega_i(t) > 0$*) are linearly independent for almost all* $t \in [0, T]$.

As with ordinary linear programming it is not difficult to establish that a non-empty and bounded feasible region ensures the existence of an optimal basic solution for SCLP. However, if we are to solve SCLP using numerical techniques we need more than this. We need to have an optimal solution with some finite structure. The next result does just that by stating that if the problem data are "nice", except at a finite number of points, then so is the optimal solution. This result may be found in Pullan [15]. We use the term "piecewise" to mean only a finite number of pieces.

Theorem 2 *Suppose that the feasible region for* SCLP *is non-empty and bounded and that all the problem data* $a(t)$, $b(t)$ *and* $c(t)$ *are piecewise analytic (with* $a(t)$ *continuous). Then* SCLP *has an optimal basic solution that is piecewise analytic. If* $a(t)$ *and* $b(t)$ *are also piecewise polynomials of degree* $n+1$ *and* n *respectively then* SCLP *has an optimal basic solution with* $x(t)$ *piecewise polynomial of degree* n.

Counter-examples are easy to construct if the assumptions of analyticity are weakened.

Substantial progress has also been made in the area of duality. It is here that the similarities between SCLP and ordinary linear programming can be seen most clearly. In finite dimensions most duality results for non-linear programming require some constraint qualifications whereas such restrictions are not required in linear programming. In infinite-dimensional optimization it seems that all duality results, both linear and non-linear, have required some constraint qualifications. Not only that, but many desirable duality results do not hold, even for infinite-dimensional *linear* programming (see Anderson and Nash [2] for a discussion of this). Recent work on SCLP has removed these problems by the careful consideration of a new dual. This dual, called SCLP*, was first introduced in Pullan [14] and is defined as follows:

$$\begin{array}{lll} \text{SCLP*:} & \text{maximize} & \displaystyle\int_0^T \eta(t)^T b(t)\,dt - \int_0^T d\pi(t)^T a(t) \\ & \text{subject to} & c(t) - G^T\pi(t) - H^T\eta(t) \geq 0, \\ & & \eta(t) \leq 0, \text{ a.e. on } [0,T], \\ & & \pi(t) \text{ monotonic increasing and} \\ & & \text{right continuous on } [0,T] \text{ with} \\ & & \pi(T) = 0. \end{array}$$

Here $\eta(t)$ is a Lebesgue integrable function and $\pi(t)$ monotonic increasing means that each component is

monotonic increasing. Weak duality (V[SCLP*] $\leq$ V[SCLP], where $V[\cdot]$ denotes the optimal value of the linear program) is readily proved. Considerably more difficult is the following strong duality result in Pullan [16].

Theorem 3 *Suppose that the feasible region for* SCLP *is non-empty and bounded and that all the problem data $a(t)$, $b(t)$ and $c(t)$ are piecewise analytic (with $a(t)$ continuous). Then V[SCLP*] = V[SCLP] and both* SCLP *and* SCLP* *have piecewise analytic optimal solutions.*

Again counter-examples may be constructed given non-analyticity.

There are similarities in proof technique between this result and its equivalent for finite-dimensional linear programming. One common way of proving the corresponding result in ordinary linear programming is via the simplex algorithm. The proof of Theorem 3 uses the algorithm for solving SCLP outlined below. Non-linear finite-dimensional duality results are invariably proved using analytical means such as separating hyperplanes. Previous infinite-dimensional duality results, both linear or non-linear, have also tended to be of this latter form.

3. ALGORITHMIC WORK

The main recent breakthrough in the study of SCLP has been the development of a numerical algorithm in Pullan [14] for solving various instances of the problem. The algorithm is fairly simple, in that it does not require a sophisticated mathematical theory to understand or use, and quick implementations require little more than a good LP code. Numerical results suggest that it is also very fast. Craddock and Philpott [8] report very encouraging numerical results using a simplified procedure for single-commodity dynamic network programs.

The algorithm in [14] is developed for piecewise linear costs $c(t)$, piecewise linear and continuous $a(t)$ and piecewise constant $b(t)$. By Theorem 2, such SCLP's have an optimal solution with $x(t)$ piecewise constant. The discontinuities in the optimal solution can (and do!) occur at any time and it is the task of the algorithm to find these times. The algorithm thus works with solutions in which $x(t)$ is piecewise constant and proceeds by producing new piecewise constant solutions with improved cost.

Let $P = \{t_0, t_1, \ldots, t_m\}$ be any partition of $[0, T]$ that contains all the discontinuities of the problem data. A discretization called AP(P) is then introduced which has many important properties. First of all it is an ordinary linear program of the same structure as the particular underlying SCLP problem. For example, if the SCLP problem is a single-commodity dynamic network program then AP(P) will be a single-commodity ordinary network program. Secondly, suppose we have a piecewise constant feasible solution $x(t)$ for SCLP whose discontinuities are in the partition P. Then we may plug the values of $x(t)$ at each of the partition points into the discretization and obtain a feasible solution $\hat{x}$ for the discretization. Moreover the cost of $\hat{x}$ in the discretization agrees with the cost of $x(t)$ in the original SCLP. Conversely (although note the difference), suppose we have a feasible solution $\hat{x}$ for the discretization. Then we may use the values of $\hat{x}$ to construct a piecewise constant feasible solution $x(t)$ for SCLP whose discontinuities occur at points in P as well as the midpoints of the partition, i.e. points of the form $(t_{i-1} + t_i)/2$ for some i. This time the costs are in general *different* between the two solutions.

Probably the most important result is that the linear programming dual of AP(P), AP*(P), is a discretization of SCLP*, the dual of SCLP. Thus any feasible solution for AP*(P) may be used to construct a feasible solution for SCLP*. Somewhat surprisingly the costs of the two solutions agree. Thus by strong duality for finite-dimensional linear programming, V[AP(P)] = V[AP*(P)] $\leq$ V[SCLP*]. We may thus deduce the following:

- By weak duality, V[AP(P)] $\leq$ V[SCLP], i.e. the optimal value of AP(P) provides a lower bound on the optimal value of SCLP. More generally, a lower bound on the optimal value of AP(P) will provide a lower bound on the optimal value of SCLP.
- If $x(t)$ is a piecewise constant feasible solution for SCLP and the corresponding AP(P) solution $\hat{x}$ mentioned above is optimal for AP(P), then $x(t)$ is optimal for SCLP.

This lower bound property is very useful. As seen above, it provides an optimality test for a SCLP solution. Another use is that it enables a partition P to be constructed, before any discretization is solved, so that AP(P) will generate a feasible SCLP solution whose cost is within any desired tolerance of the optimal value.

The algorithm presented in Pullan [14] for solving SCLP uses these properties of the discretization AP(P). However this does not mean to say that the algorithm is a discretization method, rather that the properties of the discretization are used to generate better solutions. We now summarise the algorithm.

- Start off with a piecewise constant feasible solution $x(t)$ for SCLP whose discontinuities are in P.
- Generate the corresponding $\hat{x}$ feasible for AP(P).
- If $\hat{x}$ is optimal for AP(P) then stop as $x(t)$ is optimal for SCLP (see above).
- Otherwise, find $\hat{\hat{x}}$, an improved feasible solution for AP(P).
- Use this to generate a corresponding $\tilde{x}(t)$ feasible for SCLP.

- "Patch" $x(t)$ and $\tilde{x}(t)$ together to obtain an improved solution for SCLP.

This patching together process is quite simple and involves defining the new solution $\bar{x}(t)$ to take the value $x(t)$ for some t and $\tilde{x}(t)$ for other t. Explicit and simple formulae are given which ensure that the new solution obtained is an improved feasible solution for SCLP. This new solution has different discontinuities to the old solution and we thus obtain a new partition for the next iteration.

Much numerical work remains to be done to find the best way for refining the partition obtained for use in the next iteration. However, even using a simple implementation the partition generally tends to the discontinuities in the optimal solution. Thus when solving problems, partitions are usually kept small and the discretizations are solved quickly. In the introduction we mentioned that numerical difficulties often occur when solving large discretizations. These tend to occur when the optimal solution is constant over many time intervals in the partition. The algorithm thus avoids this problem by keeping points in the partition to a minimum.

As well as numerical fine-tuning, there is also the possibility of extending the algorithm to include more general problem data. This work has begun in Pullan [17], where the structure of an algorithm for solving SCLP with general piecewise analytic costs is worked out using the duality theory in Pullan [16]. This algorithm has many properties of the simplex algorithm for ordinary linear programming. It also requires the use of a purification step, that is, a step to produce a good basic feasible solution from a given non-basic feasible solution. One possible purification step is developed in Anderson and Pullan [5].

References

[1] Anderson, E. J., 1978. *A Continuous Model For Job-Shop Scheduling.* Ph.D. thesis, University of Cambridge, U.K.

[2] Anderson, E. J. and Nash, P., 1987. *Linear Programming in Infinite-Dimensional Spaces.* Wiley-Interscience, Chichester, U.K.

[3] Anderson, E. J., Nash, P. and Perold, A. F., 1983. "Some properties of a class of continuous linear programs", *SIAM J. Control Optim.* 21: 758–765.

[4] Anderson, E. J. and Philpott, A. B., 1989. "A continuous-time network simplex algorithm", *Networks* 19: 395–425.

[5] Anderson, E. J. and Pullan, M. C., 1994. "Purification for separated continuous linear programs," submitted to *Z. Oper. Res.*

[6] Anstreicher, K. M., 1983. *Generation of feasible descent directions in continuous time linear programming.* Tech. Report SOL 83-18, Department of Operations Research, Stanford University, Stanford, CA.

[7] Bellman, R. E., 1957. *Dynamic Programming.* Princeton University Press, Princeton, NJ.

[8] Craddock, M. and Philpott, A. B., 1993. "An adaptive discretization method for a class of continuous network problems", submitted to *Networks.*

[9] Ford, L. R. and Fulkerson, D. R., 1962. *Flows in Networks.* Princeton University Press, Princeton, NJ.

[10] Grinold, R.C., 1968. *Continuous Programming.* Ph.D. thesis, Operations Research Center, University of California, Berkeley, CA.

[11] Lehman, R. S., 1954. *On the continuous simplex method,* RM-1386, Rand Corporation, Santa Monica, CA.

[12] Levinson, N., 1966. A class of continuous linear programming problems, *J. Math. Anal. Appl.* 16: 73–83.

[13] Perold, A. F., 1978. *Fundamentals of a continuous time simplex method.* Tech. Report SOL 78-26, Department of Operations Research, Stanford University, Stanford, CA.

[14] Pullan, M. C., 1993. "An algorithm for a class of continuous linear programs", *SIAM J. Control Optim.* 31: 1558–1577.

[15] Pullan, M. C., 1993. "Forms of optimal solutions for separated continuous linear programs", to appear in *SIAM J. Control Optim.*

[16] Pullan, M. C., 1993. "A duality theory for separated continuous linear programs", to appear in *SIAM J. Control Optim.*

[17] Pullan, M. C., 1994. "An extended algorithm for separated continuous linear programs", submitted to *Math. Programming.*

[18] Tyndall, W. F., 1965. "A duality theorem for a class of continuous linear programming problems", *SIAM J. Appl. Math.* 13: 644–666.

****

MAXIMUM ENTROPY METHODS: CONVEX ANALYSIS IN ACTION

by Jonathan M. Borwein and Mark A. Limber[1]

1. INTRODUCTION

Many applied problems can be reduced to the question of how to "best" solve a (possibly underdetermined) consistent system of linear equations $Ax = b$, where $b \in \mathbb{R}^n$, and x lies in some appropriate space. Often, the unknown, x, most appropriately lives in a function space, rather than $\mathbb{R}^m$, but by discretization one reduces the problem to a finite dimensional setting where A is then a $m \times n$ matrix. *We believe that in many cases, it is*

[1]CECM, Simon Fraser University, Burnaby, B.C. V5A 1S6, CANADA, e-mail: jborwein@cecm.sfu.ca, malimber@cecm.sfu.ca

better to address the problem in its function space home, discretizing only as necessary for computation.

Consider, for example, the problem of extrapolating an autocorrelation function $R(t)$ given several sample measurements. It is well known (see e.g. [11]) that the Fourier transform $S(z)$ of the autocorrelation function is the power spectrum of the data. Fourier moments of the power spectrum are the same as samples of the autocorrelation function, so by computing several values of $R(t)$ directly from the data, we are in essence computing moments of $S(z)$. If we compute a finite number of moments of S, we can then estimate S from these moments, compute more moments from the estimate $\hat{S}$ by direct numerical integration, thereby affording an extrapolation of R, without directly computing R from the potentially noisy data. Here, as in many situations, the problem is, *how do we estimate S from a finite number of its moments?* This is a underdetermined linear inversion problem where the unknown is most naturally a function, not a vector in $\mathbb{R}^m$.

We describe a maximum entropy approach to solving an underdetermined system of equations where the unknown, x, is a function, typically living in a Hilbert space, or more generally, in a space of functions. This technique essentially picks a "best" representative from the infinite set of feasible functions (functions that possess the same n moments as the sampled function) by minimizing an integral functional, f, of the unknown. This is a well known technique in the applied literature. Applications appear in constrained spline fitting, tomographic reconstruction algorithms, statistical moment fitting, and time series analysis. However, often the derivations and mathematics are fraught with subtle errors. *We will discuss some of the difficulties inherent in infinite dimensional calculus*, and provide a simple theoretical algorithm for correctly deriving maximum entropy-type solutions.

2. ENTROPIES

Let X be our function space, typically the Hilbert space $L^2(\Omega)$, or the function space $L^1(\Omega)$. For $p > 0$, we define

$$L^p(\Omega) = \left\{ x \text{ measurable } : \int_\Omega |x(t)|^p dt < \infty \right\}.$$

It is well known that $L^2(\Omega)$ is a Hilbert space with inner product $\langle x, y \rangle = \int_\Omega x(t)y(t)dt$.

The continuous linear map $A : X \to \mathbb{R}^n$ is defined by $(Ax)_i = \int x(t)a_i(t)dt$ for $i = 1, \ldots, n$ and $a_i \in X^*$ the topological dual of X (L^2 in the Hilbert space case, L^∞ in the L^1 case). To pick a solution from the infinity of possibilities, we define what we mean by "best". A very common approach is to pick the solution with minimum norm, by solving the Gram system $AA^T\lambda = b$, and the solution is then $\hat{x} = A^T\lambda$. This solution solves the following optimization problem:

$$\min \left\{ \int_\Omega x(t)^2 dt : Ax = b \;\; x \in X \right\}.$$

We generalize this by replacing the norm with a strictly convex functional f as in

$$\min \{ f(x) : Ax = b, \;\; x \in X \}, \qquad (P)$$

where f is what we call, an *entropy functional*, $f : X \to (-\infty, +\infty]$. For this note, we restrict attention to the case where f is a strictly convex integral functional of the form

$$f(x) = \int_\Omega \phi(x(t))dt.$$

The functional f can be include other constraints, including nonnegativity, by appropriate use of $+\infty$. For example,

$$f(x) = \begin{cases} \int_0^1 x(t)^2 dt & \text{if } x \geq 0 \\ +\infty & \text{else} \end{cases}$$

is the nonnegatively constrained L^2 norm functional, used in [7].

The two most common choices for f are the Boltzmann-Shannon entropy popular in image processing [6], $f(x) = \int x \ln[x]$, and the Burg [5] entropy from time series analysis, $f(x) = -\int \ln[x]$. Both implicitly assume that there is a nonnegativity (positivity in the Burg case) constraint. There has been much debate about which entropy is best in which situation, but we do not address that issue here. See [2,8] for some comparisons.

3. WHAT CAN GO WRONG

For a concrete example, consider the problem of solving the linear system of equations $Ax = b$, where, $b \in \mathbb{R}^n$ and $x \in L^2[0,1]$. Assume further that A is a continuous linear map, so that our system can be represented by $b_i = \int x(t)a_i(t)dt$ for $i = 1, \ldots, n$ and $a_i \in L^2[0,1]$. As L^2 is infinite dimensional, and $\mathbb{R}^n$ is finite dimensional, the null space of A is infinite dimensional, and thus if there are any solutions to $Ax = b$, there are an infinite number. We pick our solution to minimize the functional $f(x) = \int \phi(x(t))dt$.

We can consider the Lagrangian

$$L(x, \lambda) = \int_0^1 \phi(x(t))dt + \sum_{i=1}^n \lambda_i \left(b_i - \langle x, a_i \rangle \right),$$

and the associated dual problem

$$\max_{\lambda \in \mathbb{R}^n} \min_{x \in X} \{ L(x, \lambda) \}. \qquad (D)$$

At the solution $\hat{x}$ to (P), the derivative of $L(\hat{x}, \cdot)$ should be zero, which implies

$$\begin{aligned} \hat{x}(t) &= (\phi')^{-1}\left(\sum_{i=1}^{n} \lambda_i a_i(t)\right) \\ &= (\phi')^{-1}\left(A^T\lambda\right). \end{aligned}$$

There are two immediate problems with this approach.

The first is the assumption that a solution $\hat{x}$ exists. For example, consider the problem

$$\inf_{x \in L^1[0,1]} \left\{ \int_0^1 x(t)dt : \int_0^1 tx(t)dt = 1, x \geq 0 \right\}.$$

It is shown in [3] that the optimal value for this problem is not attained. Additional side conditions on ϕ are needed to insure solutions exist. The solution is actually the absolutely continuous part of a measure in $C(\Omega)^*$. Again, see [3] for details.

The second problem is the assumption that the Lagrangian is differentiable. In the above example, the functional f is $+\infty$ for every function x negative on a set of positive measure. This implies that the Lagrangian is $+\infty$ on a dense subset of L^1, the set of functions *not* nonnegative a.e.; the Lagrangian is nowhere continuous, much less differentiable.

One approach to circumvent the differentiability problem, is to pose the problem in $L^\infty(\Omega)$, or in $C(\Omega)$, the space of essentially bounded, or continuous, functions. However, in these spaces, even with additional side qualifications, we are not necessarily assured solutions to (P) exist. In [3], there is an example of a problem where $\Omega \subset \mathbb{R}^3$, the moments are fourier coefficients, and the entropy is Burg's, yet no solutions exist. Another example, Minerbo[10] poses the problem of tomographic reconstruction in $C(\Omega)$ with the Boltzmann-Shannon entropy. However, there the functions a_i are characteristic functions of strips across Ω, and the solution is piecewise constant, not continuous.

4. CONVEX ANALYSIS

A correct derivation of the form of solution is given in detail in [3]. For brevity, we only state the theorem that guarantees that the form of solution found in the above faulty derivation, $\hat{x} = (\phi')^{-1}(A^T\hat{\lambda})$, is, in fact, correct.

First, we introduce the Fenchel conjugate of a function $\phi : \mathbb{R} \to (-\infty, +\infty]$:

$$\phi^*(u) = \sup_{v \in \mathbb{R}} \{uv - \phi(v)\}.$$

In many cases, this can be computed explicitly, using ordinary calculus. We say ϕ possess *regular growth* if either $d = \infty$, or $d < \infty$ and $k > 0$, where $d = \lim_{u\to\infty} \phi(u)/u$ and $k = \lim_{v\uparrow d}(d - v)(\phi^*)'(v)$. The *domain* of a convex function is $\mathrm{dom}(\phi) = \{u : \phi(u) < +\infty\}$. A function is *proper* if $\mathrm{dom}(\phi) \neq \emptyset$.

Let $\iota = \inf \mathrm{dom}(\phi)$ and $\sigma = \sup \mathrm{dom}(\phi)$. Our *constraint qualification*, (CQ), reads

$$\begin{gathered} \exists \overline{x} \in L^1(\Omega), \text{ such that } A\overline{x} = b, \\ f(\overline{x}) \in \mathbb{R}, \ \iota < \overline{x} < \sigma \ \ a.e. \end{gathered}$$

We mention that in many cases, (CQ) in fact reduces to feasibility, and trivially holds.

The *dual problem* for (P) is

$$\sup \left\{ \langle b, \lambda \rangle - \int_\Omega \phi^*(A^T\lambda(t))dt \right\}. \qquad (D)$$

Theorem 1 *Suppose Ω is a finite interval, μ is Lebesgue measure, each a_k is locally Lipschitz (or in particular, continuously differentiable), and ϕ is proper, strictly convex with regular growth. Suppose (CQ) holds and*

$$\exists \tau \in \mathbb{R}^n \text{ such that } \sum_{i=1}^{n} \tau_i a_i(t) < d \quad \forall t \in [a, b], \qquad (1)$$

then the unique solution to (P) is given by

$$\hat{x}(t) = (\phi^*)'(\sum_{i=1}^{n} \hat{\lambda}_i a_i(t)) \qquad (2)$$

where $\hat{\lambda}$ is any solution to the dual problem (D).

This theorem generalizes to cover the case $\Omega \subset \mathbb{R}^p$. The results can be found in [3,4].

What Theorem 1 means in practice is that the form of the maximum entropy solution can be found without going through the (incorrect) derivation, by simply validating the easily checked conditions of Theorem 1.

Also, any solution to $Ax = b$ of the form in Eq. (2) is automatically a solution to (P). Thus, finding solutions to (P) is equivalent to solving the nonlinear system of equations

$$\langle (\phi^*)'(A^T\lambda), a_i \rangle = b_i, \quad i = 1, \ldots, n \qquad (3)$$

for $\lambda \in \mathbb{R}^n$, a *finite dimensional* system of equations. One can then apply a standard nonlinear equation solver, like Newton's method, to this system, to find the optimal λ.

It can be shown that in many cases, $(\phi')^{-1} = (\phi^*)'$, and so the incorrectly derived solution agrees with the solution given in the theorem.

Finally, notice that discretization is only needed to compute the terms in Eq. (3), and in some cases, these integrals can be computed exactly, see [9] for an example in tomography. *This is what we mean by not discretizing before necessary.* By waiting to see what form the dual problem takes, one can customize one's integration scheme to match the problem at hand.

As part of the $\mathcal{M}om\mathcal{E}nt+$ project here at the Centre for Experimental and Constructive Mathematics (CECM) at Simon Fraser University, we have implemented much of the above theory in a setting that allows general entropies and moment generating functions. There are several publications and more information available electronically via our gopher server:

`gopher gopher.cecm.sfu.ca`

our anonymous ftp site,

`ftp ftp.cecm.sfu.ca`

and our World Wide Web server

`http:\\www.cecm.sfu.ca`

which can be accessed via www, lynx or xmosaic, for example.

References

[1] J. Barzilai and J. M. Borwein. "Two point step size gradient methods." *IMA J. Numer. Anal.*, 8:141–148, 1988.

[2] J. M. Borwein, R. K. Goodrich, and M. A. Limber. "A comparison of entropies in the underdetermined moment problem." Submitted for publication.

[3] J. M. Borwein and A. S. Lewis. "Duality relationships for entropy–like minimization problems." *SIAM Journal of Control Theory and Optimization*, 29(2):325–338, March 1991.

[4] J. M. Borwein and A. S. Lewis. "Partially-finite programming in L_1: entropy maximization." *SIAM Journal of Optimization*, 3: 248-267, 1993.

[5] J. P. Burg. "Maximum entropy spectral analysis." *Paper presented at 37th meeting of the Society of Exploration Geophysicists, Oklahoma City*, 1967.

[6] N. J. Dusaussoy and I. E. Abdou. "The extended MENT algorithm: a maximum entropy type algorithm usin g prior knowledge for computerized tomography." *IEEE Transactions on Signal Processing*, 39(5):1164–1180, May 1991.

[7] R. K. Goodrich and A. Steinhardt. "L_2 spectral estimation." *SIAM J. Appl. Math.*, 46:417–426, June 1986.

[8] R. W. Johnson and J. E. Shore. "Which is the better entropy expression for speech processing: $-s \log s$ or $\log s$?" *IEEE Trans on Acoustics, Speech and Signal Processing*, February 1984.

[9] M. A. Limber, T. A. Manteuffel, S. F. McCormick, and D. S. Sholl. "Optimal resolution in maximum entropy image reconstruction from projections with multigrid acceleration." In *Proceedings of the Sixth Annual Copper Mountain Conference on Multigrid Methods*, 1993.

[10] G. Minerbo. "MENT: A maximum entropy algorithm for reconstructing a source from projection data." *Computer Graphics and Image Processing*, 10:48–68, 1979.

[11] M. B. Priestly. *Non-linear and non-stationary time series analysis.* Academic Press, London, 1988.

BULLETIN BOARD

MEMBERSHIP RENEWAL

SIAM sent out membership renewals in late August. If you cannot locate yours, you will have received a reminder in October. We hope that you will renew promptly. Your membership helps to support the products and services provided by SIAM. (Vickie Kearn)

THE SIAG/OPT PRIZE

The SIAM Activity Group on Optimization (SIAG/OPT) Prize, established in 1991, is awarded to the author(s) of the most outstanding paper, as determined by the prize committee, on a topic in optimization published in English in a peer-reviewed journal. The prize will be awarded at the SIAM Optimization Conference in 1996. The deadline for nominations will be announced in 1995.

Candidate papers must be published in English in a peer-reviewed journal bearing a publication date within the award period. They must contain significant research contributions to the field of optimization, as commonly defined in the mathematical literature, with direct or potential applications.

A nomination must include:

1. a copy of the paper being nominated,

2. a description of the significance of the paper, the important questions that have been resolved and/or raised in the paper, the applications of the work,

3. a general description of the contributions of the author(s) and how the nominated paper fits in.

Other supporting information may, at the option of the nominator, be included.

A formal call for nominations will be issued in mid-1995, but it's not too early to begin preparing a case for a nomination. (Tim Kelley)

SUMMER RESEARCH CONFERENCE ON CG-RELATED METHODS

An AMS-IMS-SIAM Joint Summer Research Conference on Linear and Nonlinear Conjugate Gradient-Related Methods co-chaired by L. Adams (UW) and J.L. Nazareth (WSU) will be held on the campus of the University of Washington in Seattle, Washington, USA from Sunday, July 9, 1995 through Thursday, July 13, 1995. The purpose of this meeting is to bring together re-

searchers and practitioners from the linear and nonlinear conjugate gradient (CG) communities for an exchange of ideas and to foster communication. Invited speakers for one-hour talks include A.R. Conn (IBM, Watson), A. Greenbaum (NYU), T. Manteuffel (Colorado), S. Mehrotra (Northwestern), J. Nocedal (Northwestern), D. O'Leary (Maryland), M.A. Saunders (Stanford), D.P. Young (Boeing).

The meeting will seek to maintain a balance between formal presentations of research (a substantial number of 1/2 hour slots will be available to participants) and the opportunity for informal and more spontaneous interaction. There will be no parallel sessions.

Additional detail can be found in the *Notices of the American Mathematical Society*, Vol. 41, No. 8, pp 1014-1017, October, 1994, or obtained by sending e-mail to Wayne Drady (wsd@math.ams.org) at AMS or one of the co-chairs.

SELECTED UPCOMING ARTICLES FOR SIAM J. OPTIMIZATION

Large-Scale, Nonlinearly Constrained Optimization on a 1024-Processor nCube *J. H. Glick and J. B. Rosen*

Evaluation of Large-Scale Optimization Problems on Vector and Parallel Architectures *Brett M. Averick and Jorge J. More*

Serial and Parallel Multicategory Discrimination *Kristin P. Bennett and O. L. Mangasarian*

An Extension of the DQA Algorithm to Convex Stochastic Programs *Adam J. Berger, John M. Mulvey, and Andrzej Ruszczynski*

Problem Formulation for Multidisciplinary Optimization *Evin J. Cramer, J. E. Dennis, Jr., Paul D. Frank, Robert Michael Lewis, and Gregory R. Shubin*

Coordination in Coarse-Grained Decomposition *R. DeLeone, R. R. Meyer, S. Kontogiorgis, A. Zakarian, and G. Zakeri*

Parallel Branch-and-Bound Algorithms for General Mixed Integer Programming on the CM-5 *Jonathan Eckstein*

Parallel Variable Distribution *M. C. Ferris and O. L. Mangasarian*

Parallel Factorization of Structured Matrices Arising in Stochastic Programming *Elizabeth R. Jessup, Dafeng Yang, and Stavros A. Zenios*

Structured Linear Least-Squares Problems in System Identification and Separable Nonlinear Data Fitting *Linda Kaufman, Garrett S. Sylvester, and Margaret H. Wright*

Serial and Parallel Computation of Karush-Kuhn-Tucker Points via Nonsmooth Equations *Jong-Shi Pang*

A Stochastic Method for Constrained Global Optimization *Klaus Ritter and Stefan Schaffler*

Infeasible-Interior-Point Primal-Dual Potential Reduction Algorithms for Linear Programming *Shinji Mizuno, Masakazu Kojima, and Michael J. Todd*

A Fast Heuristic Method for Polynomial Moment Problems with Boltzmann-Shannon Entropy *J. M. Borwein and W. Z. Huang*

Interior Point Methods in Semidefinite Programming with Applications to Combinatorial Optimization *Farid Alizadeh*

On the Primal-Dual Steepest Descent Algorithm for Extended Linear-Quadratic Programming *Ciyou Zhu*

Symmetric Quasidefinite Matrices *Robert J. Vanderbei*

An All-Inclusive Efficient Region of Updates for Least Change Secant Methods *Henry Wolkowicz and Qing Zhao*

A Positive Algorithm for the Nonlinear Complementarity Problem *Renato D. C. Monteiro, Jong-Shi Pang, and Tao Wang*

A Practical Interior-Point Method for Convex Programming *Florian Jarre and Michael A. Saunders*

An Optimal Positive Definite Update for Sparse Hessian Matrices *Roger Fletcher*

Trust Region Algorithms for Solving Nonsmooth Equations *Liqun Qi*

Why a Pure Primal Newton Barrier Step may be Infeasible *Margaret H. Wright*

CONTRIBUTIONS TO THE V&N

The next issue (Spring, '95) will have as its theme *Optimization Algorithmica Esoterica*, and will include essays by Olvi Mangasarian (Wisconsin; on optimization and machine learning) and Virginia Torczon (Rice; on pattern search).

Articles contributed by SIAG/OPT members are always welcome and can take one of two forms:
a) *Views*: short, scholarly, N^3 (Not Necessarily Noncontroversial) essay-type articles, say 2 to 4 pages long, on any topic in optimization and its interfaces with the sciences, engineering and education.
b) *News*: brief items for the Bulletin Board Section.

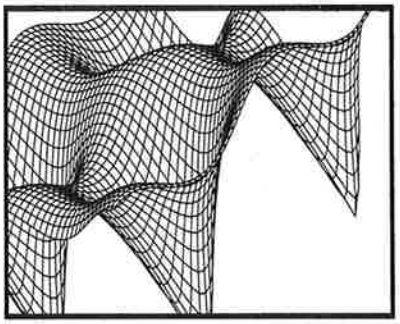
Our first preference is that a contribution take the form of a LaTeX file sent by email to the editor at the address given below. (If possible try it out in two-column format.) However, other forms of input are also acceptable.

The Bulletin-Board deadline for the next issue is March 15, 1995.

Larry Nazareth, Editor
Department of Pure and Applied Mathematics
Washington State University
Pullman, WA 99164-3113
email: nazareth@alpha.math.wsu.edu
or nazareth@amath.washington.edu

Optimization Software Guide

Jorge J. Moré and Stephen J. Wright



Frontiers in Applied Mathematics 14

1993
xii + 154 pages / Softcover
ISBN 0-89871-322-6
List Price $24.50
SIAM Member Price $19.60
Order Code FR14

"Moré and Wright have produced a compact, well-researched guide providing material and information that can be used by graduate students and researchers familiar with optimization problems. . . . "
– D.E. Bentil, University of Massachusetts at Amherst, *Choice,* June 1994.

"This is the book for the busy man or woman who is looking for a piece of software without having the time to learn everything about mathematical programming. . . .The book should be on the desk of everybody who has to solve a practical optimization problem, e.g., in engineering sciences, operations research, natural sciences."
– Klaus Schittkowski, Mathematisches Institut, Universitaet Bayreuth, January 1994.

Developments in optimization theory, including emphasis on large problems and on interior-point methods for linear programming, have begun to appear in production software. Here is a reference tool that includes discussions of these areas and names software packages that incorporate the results of theoretical research. After an introduction to the major problem areas in optimization and an outline of the algorithms used to solve them, a data sheet is presented for each of the 75 software packages and libraries in the authors' survey. These include information on the capabilities of the packages, how to obtain them, and addresses for further information.

Standard optimization paradigms are addressed – linear, quadratic, and nonlinear programming; network optimization; unconstrained and bound-constrained optimization; least-squares problems; nonlinear equations; and integer programming. The most practical algorithms for the major fields of numerical optimization are outlined, and the software packages in which they are implemented are described.

Contents

to ORDER

Use your credit card (AMEX, MasterCard, and VISA):
Call toll free in USA: 800-447-SIAM
Outside USA call: 215-382-9800
Fax: 215-386-7999
E-mail: service@siam.org

Or send check or money order to:
SIAM, Dept. BKAG94, P.O. Box 7260
Philadelphia, PA 19101-7260

Payments may be made by wire transfer to SIAM's bank:
PNC Bank, 3535 Market Street, Philadelphia, PA 19104; ABA Routing #031000053; Account Name: Society for Industrial and Applied Mathematics; Account #509-704-5

Shipping and Handling
USA: Add $2.75 for the first book and $.50 for each additional book.
Canada: Add $4.50 for the first book and $1.50 for each additional book.
Outside USA/Canada: Add $4.50 per book. All overseas delivery is via airmail.

siam®

Science and Industry
Advance with Mathematics

For information on all SIAM services, access *gopher.siam.org.*

Nonlinear Programming

Olvi L. Mangasarian

Audience
Undergraduates with an advanced calculus background and graduate students in computer science, industrial engineering, operations research, electrical engineering, economics, business, mathematics, and civil and mechanical engineering will find this book of great use. It will also be of interest to researchers in oil, investment, chemical, and software companies, as well as banks and airlines.

Classics in Applied Mathematics 10

November 1994
xv + 220 pages / Softcover
ISBN 0-89876-341-2
List Price $28.50
SIAM Member Price $22.80
Order Code CL10

This reprint of the 1969 book of the same name is a concise, rigorous, yet accessible, account of the fundamentals of constrained optimization theory. Many problems arising in diverse fields such as machine learning, medicine, chemical engineering, structural design, and airline scheduling can be reduced to a constrained optimization problem. This book provides readers with the fundamentals needed to study and solve such problems.

Beginning with a chapter on linear inequalities and theorems of the alternative, basics of convex sets and separation theorems are then derived based on these theorems. This is followed by a chapter on convex functions that includes theorems of the alternative for such functions. These results are used in obtaining the saddlepoint optimality conditions of nonlinear programming without differentiability assumptions. Properties of differentiable convex functions are derived and then used in two key chapters of the book, one on optimality conditions for differentiable nonlinear programs and one on duality in nonlinear programming. Generalizations of convex functions to pseudoconvex and quasiconvex functions are given and then used to obtain generalized optimality conditions and duality results in the presence of nonlinear equality constraints.

The book has four useful self-contained appendices on vectors and matrices, topological properties of n-dimensional real space, continuity and minimization, and differentiable functions.

Contents

About the Author

Olvi L. Mangasarian is the John von Neuman Professor of Mathematics and Computer Sciences at the University of Wisconsin, Madison.

to ORDER

Use your credit card (AMEX, MasterCard, and VISA): Call toll free in USA: 800-447-SIAM
Outside USA call: 215-382-9800
Fax: 215-386-7999; E-mail: service@siam.org

Or send check or money order to:
SIAM, Dept. BKAG94, P.O. Box 7260
Philadelphia, PA 19101-7260

Payments may be made by wire transfer to SIAM's bank: PNC Bank, 3535 Market Street, Philadelphia, PA 19104
ABA Routing #031000053; Account Name: Society for Industrial and Applied Mathematics; Account #8550970454

Shipping and Handling
USA: Add $2.75 for the first book and $.50 for each additional book.
Canada: Add $4.50 for the first book and $1.50 for each additional book.
Outside USA/Canada: Add $4.50 per book.
All overseas delivery is via airmail.

siam®

Science and Industry Advance with Mathematics