SIAG/OPT Views-and-News

A Forum for the SIAM Activity Group on Optimization

No. 3 Fall 1993

NONSMOOTH AND STOCHASTIC OPTIMIZATION AT IIASA

Since the focus of the present issue and the next (Spring, '94) is *computational* nonsmooth and stochastic optimization, this is an opportune time to describe the pivotal role played in these two areas by the International Institute for Applied Systems Analysis (IIASA), in particular, through its System and Decision Sciences Program.

IIASA is an interdisciplinary, non-governmental research institution headquartered at Laxenburg, Austria on the outskirts of Vienna, in a beautifully refurbished Habsburg summer palace provided by the Austrian Government. It was founded in 1972 on the initiative of the USA and the former USSR, and is sponsored by a consortium of National Member Organizations in 15 nations, producing scientific research on economic, environmental, technological and social issues of interest to all members. During its first two decades, IIASA represented a bridge between East and West. Now in its third decade, the Insitute seeks to conduct international and interdisciplinary studies that provide timely and relevant information and options, addressing critical issues of global change in its focus areas.

The theory and methods of systems analysis underpin the foregoing research. This was the province of a group of methodological projects that were collectively organized, during IIASA's first two decades, into a System and Decision Sciences (SDS) Program Area. (Recently, a less structured form of coordination has been adopted at IIASA that permits greater project independence and flexibility.) Under able leadership that included Professors Michel Balinski, Andrzej Wierzbicki and Academician Alexander Kurzhanski (having a surname that ends in 'ki' was *not* a job prerequisite), SDS/IIASA grew into a leading center for methodological research, particularly in the areas of nonsmooth optimization and optimization under uncertainty.

In the standard and broadly applicable mathematical programming model: minimize$_{x \in R^n} f(x)$ subject to real constraints $a_i \leq g_i(x) \leq b_i$, $1 \leq i \leq m$, the functions f and g_i are *smooth* and *deterministic*, usually with no further restrictions, for example, on convexity. This is a positive characterization. In contrast, *non*smooth optimization and optimization under *un*certainty are characterizations in the *negative*. The old Chinese proverb for describing a horse ("it is better to say: that that *is* a horse *is* a horse, rather than to say: that that *is not* a horse *is not* a horse") highlights the fact that in order to develop models and solution methods of more than theoretical significance (the theory is already well developed in these areas), it becomes necessary in practical applications to make a *positive* characterization of the underlying nonsmoothness or uncertainty. *This has led to a multiplicity of relevant models*, in marked contrast to the smooth case.

Different approaches have been emphasized in different countries, often an outgrowth of differing *scientific* cultures. For example, the term 'Russian school' is often attached to very general formulations defined in terms of expectation functions, and *non*monotonic subgradient or stochastic quasi-gradient techniques, which are powerful but difficult to automate. The term 'French school' is attached to deep analysis of the subdifferentiability properties of nonsmooth models, and monotonic 'bundle'-type methods. 'American school' connotes an emphasis on specific models (for example, recourse) solved automatically, on powerful computers, by cleverly adapted methods of large-scale mathematical programming coupled with techniques of approximation/sampling of distribution functions. IIASA's strength, through its core interdisciplinary research projects and outreach networks of scientists, has been to help bridge these gaps. Two pioneering efforts in this regard are: *Nonsmooth Optimization*, C. Lemarechal and R. Mifflin (Eds.), IIASA Proceedings 3, Pergammon (1978), and *Numerical Techniques for Stochastic Optimization*, Y. Ermoliev and R. Wets (Eds.), Springer-Verlag (1988), each representing an important plateau of an IIASA project.

As a result of these efforts and others conducted elsewhere, very significant *practical* breakthroughs have been achieved. One, at Stanford, by Dantzig and Infanger is overviewed in this issue. Others will be mentioned in essays in the next issue, which incidentally are also representative of the different (so-called) 'schools' mentioned above. Interestingly enough, their authors have all been long-term associates of IIASA or active participants in IIASA projects, a further illustration of the important synergistic and catalytic functions of this unique institute.

Contents:

CHAIRMAN'S COLUMN

by A.R. Conn

Like most groups, researchers have their own particular foibles and eccentricities. In particular, writing papers and subjecting them to peer review encourages a certain degree of egotism, paranoia and bloody-mindedness. Most of it is understandable and, except for the occasional ruffled feathers, causes no real harm. However, it sometimes becomes more serious, and in one recent extreme case, an academic in Montreal, denied tenure, claimed that he did all the work in a series of articles, although his more senior colleagues' names (presumably the same ones who contributed to his denied tenure) appear among the authors. The distressed faculty member then attempted to shoot his colleagues. Although unsuccessful in this endeavor, he did, unfortunately, kill a secretary and two other faculty members. Although this is clearly not a normal event, I think we are all aware of abuses of authority/seniority that dictate authorship of articles. Often this is partly a question of precedence in the field — in some areas it is not at all unusual for the head of a laboratory or group to appear as an author on all articles published by the group. Thus for instance, in medicine there are authors who publish at a phenomenal rate. Baboon-liver transplanting Pittsburgh surgeon Dr. Thomas Startzl published every 2.4 days in 1991...and he also found time to write and promote a book about his hectic life (cited from the Toronto Globe & Mail, Science and Academe section, March 6, 1993).

Nevertheless, I believe that intellectual honesty is something that researchers take seriously and, in most cases, when it is lacking, the omission is unintentional.

We are all familiar with tales of stolen ideas and of trespassing on territorial claims. In most instances, I do not believe these are cases of genuine intellectual dishonesty. Typically, I think they are much more likely to be caused by what I call 'cultural confusion', equivalent to say a pure mathematician claiming that the solution to a square non-singular system of linear equations $Ax = b$ is given by $x = A^{-1}b$ — end of story. Thus all the volumes in numerical analysis devoted either directly or indirectly to solving linear systems are either trivial or have already been done. Although usually more subtle than the example just given, I feel that most, but regrettably not all, complaints of plagiarism are in fact more a matter of points of view rather than intentional dishonesty. Similarly, in slide presentations, it is difficult to give references to all the relevant preceding work, if only because of the limited time available for the presentation. I myself, usually give few such references when giving a presentation of less than thirty minutes. Sometimes this appears as intentional dishonesty but in reality it rarely is (in my own case, never, I hope). By contrast, although inevitably it is not trivial to include all the appropriate (especially historical) attributes, I do believe it is essential to endeavor to be as thorough as possible in referencing previous relevant work in written articles. It is certainly intellectually dishonest to expressly exclude references to previous work either because you do not like the authors or disagree with their point of view. Naturally, it is usual to have some cultural biases based upon one's upbringing, and one has to make a conscious effort to avoid, for example, nationalistic biases. At some level such parochialism is inevitable; for instance, the text one used as a student may well introduce certain cultural attitudes. Many of us read few Eastern European articles, not really because we are prejudiced but because we find their style of mathematics different (and less accessible) than our own. Nevertheless, it is often rather astonishing how narrow-minded we can sometimes be. It is not hard to find bibliographies unreasonably biased not only to countries and continents, but also to single institutions and their graduates.

Another area where such prejudices do show up from time to time is in refereeing — whether it be of grant proposals or submitted articles. In this arena, associate editors and members of panels have a special responsibility to avoid abuses of their positions of power, whether their own or those of fellow members. Since people in these positions are already often powerful themselves, it can be difficult to oppose them because of the possible adverse consequences. However, if one feels that what one might call 'intellectual harassment' is taking place, it is intellectually dishonest to remain silent — even if that might not be politically prudent.

Finally, prejudice in refereeing brings me to one of my own particular hobby-horses, that of signing referee reports. I have always signed my reports, with the exception of cursory dismissals of clearly unpublishable works (in which case I always volunteer to write a more detailed report that I would be willing to sign, if more details are requested). The usual argument against such a policy is that the author will unjustly blame/hate the referee. However, I contend that just as frequently an author blames a colleague for a report that was actually written by someone else! The referee knows the author so why shouldn't the author know the referee? In particular, I am convinced from my own experience as an author that many reports would be constructively improved if the referee signed them and especially, the all too frequent, superficial and sometimes surprisingly venomous remarks, would disappear.

Well, I hope the above remarks do not appear too pompous and condescending. They are meant to be taken for what they are, my honest, personal opinion. As always, your comments and feedback are welcome. My e-mail address is arconn@watson.ibm.com

FORUM ESSAY

PLANNING UNDER UNCERTAINTY

by George B. Dantzig and Gerd Infanger[1]

Solutions obtained from deterministic planning models are usually unsatisfactory because they fail to hedge against unfavorable events which may occur in the future. Stochastic models address this shortcoming, but in the past have seemed to be intractable because, even for a relatively small number of parameters, subject to uncertainty the size of such problems can get very large. Techniques for solving stochastic mathematical programs were first described in Dantzig (1955) [6]: more recently in numerous articles, some of which are cited in our bibliography.

This essay addresses a common class of stochastic models, namely stochastic linear programs with recourse. We review our approach for solving them efficiently using a blending of classical *Benders Decomposition* with a relatively new technique called *Importance Sampling* (Dantzig and Glynn (1990) [7], Infanger (1992) [21]). Some numerical results from solving large-scale problems in the area of expansion planning of power systems and financial planning are also presented.

1. TWO-STAGE STOCHASTIC LINEAR PROGRAMS

An important class of stochastic models are two-stage stochastic linear programs with recourse. These models are the analog extensions of deterministic dynamic systems which have a staircase structure: x denotes the first $t = 1$, y the second $t = 2$ stage decision variables, A, b represent the coefficients and right hand sides of the first stage constraints and D, d represent the second stage constraints, which together with the transition matrix F, couples the two periods. In the literature D is often referred to as the technology/recourse matrix. The first stage parameters are known with certainty. The second stage parameters are random variables ω that assume certain outcomes with certain probabilities $p(\omega)$. They are known at time $t = 1$ only by their probability distribution of possible outcomes. The actual outcomes will become known later at time $t = 2$. Uncertainty occurs in the transition matrix F and in the right hand side vector d. The second stage costs f and the elements of the technology/recourse matrix D are *assumed* to be known with certainty. We denote an outcome of the stochastic parameters by $\omega \in \Omega$, with Ω being the set of all possible outcomes. $E^\omega(\cdot)$ denotes the expectation with respect to the probability distribution $p(\omega)$. The two-stage stochastic linear program can be written as follows:

$$\begin{array}{lrcll} \min z = & cx & + & E^\omega(\ fy^\omega) & \\ s/t & Ax & & = & b \\ & -F^\omega x & + & Dy^\omega = & d^\omega \\ & x, y^\omega & & \geq & 0, \quad \omega \in \Omega. \end{array}$$

The problem is to find a first stage decision x, if possible, which is feasible for all scenarios $\omega \in \Omega$ and has the minimum expected costs. Note the adaptive nature of the problem: While the decision x is made only with the knowledge of the distribution $p(\omega)$ of the random parameters, the second stage decision y^ω is made later after an outcome ω is observed. The second stage decision compensates for and adapts to different scenarios ω.

Using discrete distributions, one can express a stochastic problem as a deterministically equivalent linear program by writing down the second stage constraints for each scenario $\omega \in \Omega$ one below the other. The objective function carries out the expected value computation by direct summation. Clearly, this formulation could lead to linear programs of enormous sizes if the number K of discrete events ω is large:

$$\begin{array}{lrcccccccl} \min z & = cx & + & p^1 f y^1 & + & \cdots & + & p^K f y^K & & \\ s/t & Ax & & & & & & & = & b \\ & -F^1 x & + & Dy^1 & & & & & = & d^1 \\ & \vdots & & & & \ddots & & & & \vdots \\ & -F^K x & & & & & + & Dy^K & = & d^K \\ & x, & & y^1, & & \ldots, & & y^K & \geq & 0. \end{array}$$

As already noted, the method which we apply to solve large-scale stochastic linear programs uses Benders decomposition and importance sampling. The underlying theory of our approach is developed in Dantzig and Glynn (1990) [7] and in Infanger (1992)[21]. Dantzig and Infanger (1991) [11] report on the solution of large-scale problems. In the following we give a brief review of the concept. First we discuss the concept of Benders decomposition for 2-stage stochastic linear programs. We then introduce the formulation of a special class of multi-stage stochastic linear programs as an extension to the formulation of the two-stage stochastic linear program above. We show how using sampling techniques estimates of the

[1]Research partially supported by the Office of Naval Research Contract N00014-89-J-1659; the National Science Foundation Grants ECS-8906260, DMS-8913089, the Electric Power Research Institute Contract RP 8010-09, CSA-4O05335, and the Austrian Science Foundation, "Fonds zur Förderung der wissenschaftlichen Forschung," Grant J0323-Phy. Authors' address: Department of Operations Research, Stanford University, Stanford, CA 94305-4022, U.S.A.

expected costs and variances can be computed. Here importance sampling is the key to obtaining accurate estimates, i.e. unbiased estimates with low variances, with low sample size. Finally we demonstrate the power of the approach through numerical results from large-scale test problems that we solved.

2. BENDERS DECOMPOSITION

We decompose the 2-stage multi-period stochastic linear program by applying Benders (1962) [2] decomposition. See van Slyke and Wets (1969) [31] for a reference to using Benders decomposition for stochastic linear programs.

Benders decomposition splits the original problem into a master problem and a subproblem, which in turn decomposes into a series of independent subproblems, one for each $\omega \in \Omega$. The latter are used to generate cuts. The master problem, the subproblems and the cuts are summarized as follows.

The master problem:

$$\begin{array}{rrrrrl} min\ z_M & = & cx & + & \theta & \\ s/t & & Ax & & = & b \\ cuts: & & -G^l x & + \theta & \geq & g^l,\ l = 1, \ldots, L \\ & & x, & \theta & \geq & 0. \end{array}$$

The subproblems:

$$\begin{array}{rrrl} min\ z^\omega & = & p^\omega f y^\omega & \\ s/t\ p^\omega \pi^\omega: & & Dy^\omega & = d^\omega + B^\omega x \\ & & y^\omega & \geq 0,\ \omega \in \Omega, \end{array}$$

where, for example, $\Omega = \{1, 2, \ldots, K\}$, and where $p^\omega \pi^{\omega *}$ is the optimal dual solution of subproblem ω, given x, the first stage decision. To simplify the discussion we assume the subproblems are all feasible.

The parameters of the cuts:

$$\begin{array}{rcrcl} g & = & \sum_\omega p^\omega \pi^{\omega *} d^\omega & = & E(\pi^{\omega *} d^\omega), \\ G & = & \sum_\omega p^\omega \pi^{\omega *} B^\omega & = & E(\pi^{\omega *} B^\omega). \end{array}$$

By solving the master problem, where cuts are initially absent and then sequentially added, we obtain a trial solution $\hat{x}$. Given $\hat{x}$ we can solve K subproblems $\omega \in \Omega$ to compute a cut. The cut is a lower bound on the expected value of the second stage costs represented as a function of x. Cuts are sequentially added to the master problem and new values of $\hat{x}$ are obtained until the optimality criterion is met.

If the expected values z, G, and g are computed exactly, that is, by evaluating all scenarios $\omega \in \Omega$, we refer to it as the universe case. As we will see later the number of scenarios easily gets out of hand if K is large and it is not always possible to solve every subproblem ω. Therefore methods are sought that guarantee a satisfactory solution without having to solve the universe case. We employ Monte Carlo sampling techniques to obtain accurate estimates of the expected values z, G, and g.

3. A CLASS OF MULTI-STAGE STOCHASTIC LINEAR PROGRAMS

Large-scale deterministic mathematical programs, used for operations and strategic planning, often are dynamic linear programs. These problems have a staircase (multi-stage) matrix structure. In general, the size of these stochastic problems can get extremly large because the number of scenarios grows exponentially with the number of periods. We will, however, address a certain restricted class for which the number of scenarios grows linearly with the number of stages: The problem (whose constraints are stated below) breaks down into two parts: a deterministic dynamic part and a stochastic part. We call the deterministic part the master problem. It is a dynamic linear program with T stages. The vectors c_t and b_t, and the matrices B_{t-1} and A_t are assumed to be known with certainty.

$$\begin{array}{rrrrrl} \min z = \sum_{t=1}^T c_t x_t & + & \sum_{t=1}^T E(f_t y_t^{\omega_t}) & & & \\ -B_{t-1} x_{t-1} + A_t x_t & & & & = & b_t \\ -F_t^{\omega_t} x_t & + & & D_t y_t^{\omega_t} & = & d_t^{\omega_t} \\ x_t, & & & y_t^{\omega_t} & \geq & 0, \end{array}$$

where $B_0 = 0$, $t = 1, \ldots, T$, and $\omega_t \in \Omega_t$. Each stage is associated with a stochastic subproblem. Uncertainty appears in the recourse-matrix $F_t^{\omega_t}$ and in the right hand side vector $d_t^{\omega_t}$ where ω_t, denotes an outcome of the stochastic parameters in period t, with Ω_t denoting the set of all possible outcomes in period t. The subproblems in each stage are assumed to be stochastically independent. The subproblem costs f_t and the technology matrix D_t are assumed to be deterministic parameters.

Facility expansion planning is an example of this type of formulation. The master problem models the expansion of the facilities over time. Decision variables are the capacity built and the capacity available at time t. The subproblems model the operation of these capacities in an uncertain environment. Take for example the case of expansion planning of power systems: The expansion or replacement of capacities of generators and transmission lines are determined in the master problem. The capacities at each period t are made available to the system for operation. The subproblems model the power system operation, the optimal sceduling of the available capacities to meet the demand for electricity. The availabilities of generators and transmission lines and the demands are uncertain and not known at the time when the expansion decision is made. The special class of multi-stage problems decomposes as follows:

The master problem:

$$\begin{array}{rrcl} \min\ z_M^L = \sum_{t=1}^T c_t x_t & + \sum_{t=1}^T \theta_t & & \\ -B_{t-1}x_{t-1} + A_t x_t & & = & b_t \\ -G_t^l x_t & + \theta_t & \geq & g_t^l \\ x_t & & \geq & 0, \end{array}$$

where $B_0 = 0$, $t = 1, \ldots, T$, and $l = 1, \ldots, L$.

The subproblems for ω_t in period t:

$$\begin{array}{lrcll} \min\ z_t^{\omega_t}(\hat{x}_t^l) & = & f_t y_t^{\omega_t} & & \\ \pi_t^{\omega_t}(\hat{x}_t^l): & & D_t y_t^{\omega_t} & = & d_t^{\omega_t} + F_t^{\omega_t}\hat{x}_t^l, \quad \omega_t \in \Omega_t \\ & & y_t^{\omega_t} & \geq & 0, \qquad\qquad \hat{x}_t^l \text{ given}, \end{array}$$

where $t = 1, \ldots, T$.

The parameters of the cuts:
for $t = 1, 2, \ldots, T$,

$$G_t^l = E(\pi^{\omega_t} B^{\omega_t}), \quad g_t^l = E(\pi^{\omega_t} d^{\omega_t}),$$

$$z_t(\hat{x}_t^l) = E(z_t^{\omega_t}), \quad \pi_t^{\omega_t} = \pi_t^{\omega_t}(\hat{x}_t^l).$$

4. IMPORTANCE SAMPLING

The difficulty of solving large-scale stochastic problems arises from the need to compute multiple integrals or multiple sums. The expected value of the second stage costs e.g. $z = E(fy^\omega) = E(C)$ is an expectation of functions $C(v^\omega)$, $\omega \in \Omega$, where $C(v^\omega)$ is obtained by solving a linear problem. $v^\omega = (v_1, \ldots, v_h)^\omega$ with $p(v^\omega)$, the corresponding probability, are outcomes of say V (in general), an h-dimensional random vector parameter, e.g. $V = (V_1, \ldots, V_h)$. With P being the probability measure and under the assumption of independence the integral $E\,C(V) = \int C(v^\omega)P(d\omega)$ takes the form of a multiple integral $E\,C(V) = \int \cdots \int C(v)p(v)dv_1 \ldots dv_h$, or, in case of discrete distributions, the form of a multiple sum $E\,C(V) = \sum_{v_1} \cdots \sum_{v_h} C(v)p_1(v_1)\ldots p_h(v_h)$.

In the following discussion we concentrate on discrete distributions. This is not a restriction as the approach can be easily adapted for continuous distributions. Even for h as small as 20 the number of terms in the multiple sum computation gets easily out of hand and the problem is no longer practical to solve by direct summation. This is especially true because function evaluations can be computationally expensive since each term in the multiple sum requires the solution of a linear program.

Monte Carlo Methods are recommended to compute multiple integrals or multiple sums for higher h-dimensional sample spaces (Davis and Rabinowitz (1984) [13], Glynn and Iglehart (1989) [19]). Suppose $C^\omega = C(v^\omega)$ are independent random variates of $v^\omega, \omega = 1, \ldots, n$ with expectation z, where n is the sample size. An unbiased estimator of z with variance $\sigma_{\bar{z}}^2 = \sigma^2/n$, $\sigma^2 = var(C(V))$ is

$$\bar{z} = (1/n)\sum_{\omega=1}^{n} C^\omega.$$

Note that the standard error decreases with $n^{-0.5}$ and the convergence rate of $\bar{z}$ to z is independent of the dimension of the sample space h. We rewrite $z = \sum_{\omega\in\Omega} C(v^\omega)p(v^\omega)$ as

$$\sum_{\omega\in\Omega} \frac{C(v^\omega)p(v^\omega)q(v^\omega)}{q(v^\omega)}$$

by introducing a new probability mass function $q(v^\omega)$ and we obtain a new estimator of z

$$\bar{z} = \frac{1}{n}\sum_{\omega=1}^{n} \frac{C(v^\omega)p(v^\omega)}{q(v^\omega)}$$

by sampling from $q(v^\omega)$. The variance of $\bar{z}$ is given by

$$var(\bar{z}) = \frac{1}{n}\sum_{\omega\in\Omega}\left(\frac{C(v^\omega)p(v^\omega)}{q(v^\omega)} - z\right)^2 q(v^\omega).$$

Chosing $q^*(v^\omega) = C(v^\omega)p(v^\omega)/\sum_{\omega\in\Omega} C(v^\omega)p(v^\omega)$ would lead to $var(\bar{z}) = 0$, which means one could get a perfect estimate of the multiple sum from only one estimation. Practically however, this is useless since to compute $q(v^\omega)$ we have to know $z = \sum_{\omega\in\Omega} C^\omega p(v^\omega)$, which is the value we wanted to compute in the first place. The result, however, helps us to derive a heuristic criterion for choosing a "good" q. It should be proportional to the product $C(v^\omega)p(v^\omega)$ and should have a form that can be integrated easily. Thus a function $\Gamma(v^\omega) \approx C(v^\omega)$ is sought, which can be integrated with less costs than $C(v^\omega)$. Additive and multiplicative (in the components of the stochastic vector v) approximation functions and combinations of these are potential candidates for our approximations. In particular, we have been getting good results using $C(V) \approx \sum_{i=1}^{h} C_i(V_i)$. We compute q as

$$q(v^\omega) \approx \frac{C(v^\omega)p(v^\omega)}{\sum_{i=1}^{h}\sum_{\omega\in\Omega_i} C_i(v^\omega)p_i(v^\omega)}.$$

To understand the motivation for this importance sampling scheme, assume for convenience $C_i(v_i^{\omega_i}) > 0$ and let $\Gamma(v^\omega) = \sum_{i=1}^{h} C_i(v_i^\omega)$. If $\sum C(v^\omega)p(v^\omega)$ were used as an approximation of $\bar{z}$ it can be written

$$\sum_{\omega=1}^{n} \Gamma(v^\omega)p(v^\omega) = \sum_{i=1}^{h} \alpha_i \sum_{\omega=1}^{n} [\frac{C_i(v_i^\omega)}{\alpha_i}]\, p_1(v_1^{\omega_1})\ldots p_h(v_h^{\omega_h}),$$

where $\omega = (\omega_1, \omega_2, \ldots, \omega_h)$ and where we define

$$\alpha_i = \sum_{\omega_i\in\Omega_i} C_i(v_i^{\omega_i})p_i(v_i^{\omega_i}),$$

which is relatively easy to compute since it can be evaluated by summing only h sums each of which is a sum of one of the dimensions of ω. Note that

$$\bar{p}_i(v_i^{\omega_i}) = \frac{C_i(v_i^{\omega_i})p_i(v_i^{\omega_i})}{\alpha_i} \geq 0, \ \ \omega_i \in \Omega_i$$

may be viewed as a modified probability distribution of v_i associated with the i term. It is, of course, a trivial matter to directly sum each term i since each of its factors, being independent probability distributions, sum to one. Suppose, however, one does not notice this fact and decides to estimate the sum by estimating each of the h terms by Monte Carlo sampling. The i-th term would then be evaluated by randomly sampling v_i from the distribution $\bar{p}_i(v_i^{\omega_i})$ and all the rest of the components v_j of v from the distributions $p_i(v_i^{\omega_i})$.

In an analogous manner, we let

$$\rho(\omega) = \frac{C(\omega)}{\Gamma(\omega)}$$

and write

$$\begin{aligned} \bar{z} &= \sum C(\omega)p(\omega) \\ &= \sum \rho(\omega)\Gamma(\omega)p(\omega) \\ &= \sum_{i=1}^{h} \alpha_i \sum_{\omega=1}^{n} \rho(\omega)[\tfrac{C_i(v_i^{\omega})}{\alpha_i}]\, p_1(v_1^{\omega_1}) \ldots p_h(v_h^{\omega_h}). \end{aligned}$$

If our approximation $\Gamma(\omega)$ to $C(\omega)$ is any good, $\rho(\omega)$ will be roughly 1 for almost all scenarios ω. This suggests the heuristic that the sampling be carried out differently for each term i. The importance sampling scheme then is to sample v_i of the i-th term according to the distribution $\bar{p}_i(v_i^{\omega_i})$ and to sample all other components $v_j^{\omega_j}$ of the i-th term according to the distribution $p_j(v_j^{\omega_j})$.

If the additive function turns out to be a bad approximation of the cost function, as indicated by the observed variance being too high, it is easily corrected by increasing the size of the sample. This is done adaptively.

Actually we use a variant of the additive approximation function. By introducing $C(\tau)$, the costs of a base case, we make the model more sensitive to the impact of the stochastic parameters v. Our approximation function is computed as follows:

$$\Gamma(V) = C(\tau) + \sum_{i=1}^{h} \Gamma_i(V_i),$$

$$\Gamma_i(V_i) = C(\tau_1, \ldots, \tau_{i-1}, V_i, \tau_{i+1}, \ldots, \tau_h) - C(\tau).$$

We refer to this as a *marginal cost* approximation. We explore the cost function at the margins, e.g. we vary the random elements v_i to compute the costs for all outcomes v_i while we fix the other random elements at the level of the base case. τ can be any arbitrary chosen point of the set of k_i discrete values of v_i, $i = 1, \ldots, h$. For example we choose τ_i as that outcome of V_i which leads to the lowest costs, ceteris paribus.

Summarizing, the importance sampling scheme has two phases: the preparation phase and the sample phase. In the preparation phase we explore the cost function $C(V)$ at the margins to compute the additive approximation function $\Gamma(V)$. For this process $n_{\text{prep}} = 1 + \sum_{i=1}^{h}(k_i - 1)$ subproblems have to be solved. Using $\Gamma(V)$ we compute the approximate importance density

$$q(v^{\omega}) = \frac{\Gamma(v^{\omega})p(v^{\omega})}{C(\tau) + \sum_{i=1}^{h}\sum_{\omega\in\Omega_i}\Gamma_i(v^{\omega})p(v^{\omega})}.$$

Next we sample n scenarios from the importance density and, in the sample phase, solve n linear programs to compute the estimation of $\bar{z}$ using the Monte Carlo estimator. We compute the gradient G and the right hand side g of the cut using the same sample points at hand from the expected cost calculation. See Infanger (1992) [21] for the computation of the cuts and details of the estimation process.

The additive adaptive importance sampling scheme worked very well in many different kinds of models. In the section "Implementation and Numerical Results" below we review the quality of solutions obtained from models of electric power planning and financial planning. Krishna (1993) [24] developed as part of his Ph.D. thesis a promising importance sampling scheme based on a piecewise linear approximation function.

5. PROBABILISTIC BOUNDS

Based on Benders decomposition and using estimates of the expected second stage costs, z, the gradients, G, and the right-hand sides, g, of the cuts, the objective function value of the master problem gives a lower bound estimate and the total expected costs of a trial solution $\hat{x}_t^l, t = 1, \ldots, T$ gives an upper bound estimate to the objective function value of the problem. If the lower and the upper bound are sufficiently close, which is tested by a Student t-test, the problem is considered to be solved. Lower and upper bounds can be seen as a sum of *i.i.d.* random terms which for sample sizes of 30 or more can be assumed normally distributed with known (derived from the estimation process) variances. A 95% confidence intervall of the optimal solution is computed. See Dantzig and Glynn (1990) [7] and Infanger (1992) [22] for details of the algorithm. Current research involves the development of ways to adaptively improve solutions until a prespecified quality criterion is met, e.g. David Morton (1993) [25] explored as part of his Ph.D. thesis limit results of sampling based improving algorithms.

6. IMPLEMENTATION AND NUMERICAL RESULTS

This method for solving large-scale two-stage stochastic linear programs with recourse has been implemented. The code of MINOS (Murtagh and Saunders (1983) [26]) has been adapted for this purpose as a subroutine for solving both the master problem and the subproblems. When solving large numbers of subproblems we have found it is important for the performance of the algorithm to take advantage of good starting bases. Computation time can be reduced dramatically by solving *first* an expected value problem by replacing the stochastic parameters by their expectations. The expected value solution of the resulting deterministic problem is then used as a starting point for the stochastic solution. Additionally we keep cuts obtained from the expected value problem to initially guide the algorithm. It can be shown that cuts obtained from the expected value problem are valid for the stochastic problem. They are often "weak" and have to be replaced as the algorithm proceeds. The code uses sparce matrix techniques and efficient data structures for handling large-scale problems.

Computational results of the large scale test problems are represented in Table 1. Besides the solution of the stochastic problems, the results from solving the expected value problems are also reported. We also report on the estimated expected costs if the expected value solution is used as the decision in a stochastic environment. The objective function value of the true stochastic solution has to lie between the minimum value of objective function of the deterministic problem and the expected costs of the expected value solution.

Expansion planning of multi-area power systems

WRPM is a multi-area capacity expansion planning problem for the western USA and Canada. The model is very detailed and covers 6 regions, 3 demand blocks, 2 seasons, and several kinds of generation and transmission technologies. The objective is to determine optimum discounted least cost levels of generation and transmission facilities for each region of the system over time. The model minimizes the total discounted costs of supplying electricity (investment and operating costs) to meet the exogenously given demand subject to expansion and operating constraints. A description of the model can be found in Dantzig et. al. (1989). In the stochastic version of the model the availabilities of generators and transmission lines and demands are subject to uncertainty. There are 13 stochastic parameters per time period (8 stochastic availabilities of generators and transmission lines and 5 uncertain demands) with discrete distributions with 3 or 4 outcomes. The model covers a time horizon of 3 future periods of 10 years each. Thus the total number of stochastic parameters is 39. The operating subproblems of each period are stochastically independent. The number of universe scenarios is larger than $5 \cdot 10^6$ per period. In the deterministic equivalent formulation the problem if it were possible to express it exactly would have more than 4.5 billion constraints.

The stochastic WRPM is solved by using a sample size of 100 to generate each cut obtained by sampling. It takes 129 iterations to obtain the expected value solution and additional 68 iterations to compute the stochastic solution. The objective function value of the stochastic solution was estimated as 199017.4 with an amazingly small 95% confidence intervall of 0.029% on the low side and 0.067% on the high side. Thus the optimal solution lies with 95% confidence between $198959.3 \leq z^* \leq 199164.1$. The expected costs of the expected value solution (202590.3) and the objective function value of the stochastic solution differ significantly from the expected costs of the optimal stochastic solution. The problem was solved in 687 minutes on a Toshiba T5200 laptop personal computer. This time includes time to solve 26295 linear subproblems of the size of 302 rows and 289 columns and 197 master problems.

Portfolio Management

LP42 is a portfolio management test-problem, formulated as a network problem. It is a modified version of test-problems found in Mulvey and Vladimirou (1989). The problem is to select a portfolio which maximizes expected returns in future periods taking into account the possibility of revising the portfolio in each period. There are also transaction costs and bounds on the holdings and turnovers. The test problem covers a planning horizon of four future periods. The returns of the stocks in the four future periods are assumed to be independent stochastic parameters, discretely distributed with 3 outcomes each; this formulation differs from that of Mulvey and Vladimirou who restricted the problem size by looking at a certain number of preselected scenarios. Like in Mulvey and Vladimirou the multi-period problem is viewed as a 2-stage problem, where all future periods are included in the second stage. With 13 stocks with uncertain returns, the problem has 52 stochastic parameters. The universe number of scenarios $6 \cdot 10^{24}$ is very large, so that the deterministic equivalent formulation of the problem if it could be expressed explicitely would have more than $1.9 \cdot 10^{27}$ rows. Here, the stochastic parameters appear in the F-matrix as well as in the D-matrix. In this case cuts from the expected value problem are not valid for the stochastic problem. The expected value problem and the stochastic problem are solved seperately. A sample size of 600 was chosen. The solution (objective function value 2.329) is obtained in 4 iterations. Given the large number of stochastic parameters, the 95% confidence intervall is very small namely 0.536% on the low

side and 0.767% on the high side. Thus with 95% confidence the objective function value of the optimal solution lies within $2.316 \leq z^* \leq 2.347$. The expected costs of the expected value solution is significantly different from the expected costs of the stochastic solution.

Table 1: Large test problems: computational results power planning

		WRPM	LP42
# iter stoch. (exp. val.)		197 (129)	4(6)
sample size		100	600
exp. val. solution obj		196471.4	1.611
exp. val. solution, exp. cost		202590.3	2.334
stochastic solution		199017.4	2.329
estimated conf. left %		0.0292	0.536
estimated conf. right %		0.067	0.767
solution time (min)		687	209
Problem Size			
Master	rows	128	49
	columns	226	83
	nonzeros	413	133
Sub	rows	302	178
	columns	289	309
	nonzeros	866	570
# stoch parameters		39	52
# univ. scenarios		$15 \cdot 10^6$	$6 \cdot 10^{24}$

7. PARALLEL PROCESSORS

The algorithm lends itself, whether in exact form or using sampling to parallel processing. In collaboration with James K. Ho we have explored how our approach for solving two-stage stochastic linear programs can be effectively implemented on a parallel (Hypercube) multicomputer (Dantzig, Ho and Infanger (1991) [9]). For this work we used an Intel iPSC/2 d6 with 64 nodes at the Oak Ridge National Laboratory. The hypercube has the architecture of loosely coupled multiprocessors, each processor running asynchronously and accessing its private memory (MIMD). Information is exchanged between nodes only by sending messages through the communications network of the hypercube. In our implementation we assign one processor to be the master processor, which, besides its main task of solving the master problem, also controls the computation and synchronizes the algorithm. The other processors were assigned to be sub processors, with the main task of solving subproblems. Numerical results show that significant speed-ups can be achieved using parallel processing. The efficiency of the algorithm increases with larger sample sizes, which makes parallel processing techniques especially worthwhile for solving difficult stochastic problems where large sample sizes are required. For example, solving a reduced version of WRPM with only one sub period, we achieved a speed-up of about 60% using a sample size of 600 and 64 parallel processors. For detailed results, see Dantzig, Ho and Infanger (1991) [9]. Current efforts in a collaboration with IBM involve the implementation of the parallel algorithm on a network of parallel RS/6000 workstations. Here each node is a fast computer itself and is connected with other nodes via the unix network. We expect to be able to quickly solve large practical problems involving large numbers of sample sizes.

8. A GENERAL CLASS OF MULTI-STAGE STOCHASTIC LINEAR PROGRAMS

Encouraged by the promising numerical results for two-stage and a restricted class of multi-stage problems current research involves the development of methods for solving a general class of multi-stage stochastic linear programs. These class of problems arises from deterministic dynamic linear programs and have uncertain parameters in different time periods. As already noted above, the size of these stochastic problems can get extremely large because the number of scenarios grows exponentially with the number of periods. Difficulties in solving these problems also can arise from dependencies of the stochastic parameters both within a certain stage a nd between different stages. Dantzig and Infanger (1991) [10] show how multi-stage portfolio optimization problems can be efficiently solved as multi-stage stochastic linear programs. In Infanger (1992) [22] we developed a theory of solving this general class of multi-stage stochastic linear programs. Our approach includes special sampling techniques for computing upper bounds and methods of sharing cuts between different sub-problems. It will enable us to efficiently solve large-scale multi-stage problems with many stages and numerous stochastic parameters in each stage. The implementation is subject to future research. Preliminary numerical results have turned out to be promising.

9. CONCLUSION

We have presented a review of our approach using Benders decomposition and importance sampling which is capable of solving certain classes of large-scale problems of planning under uncertainty. Numerical results of large problems with numerous stochastic parameters indicate that very accurate solutions of such problems can be obtained using only small sample sizes. Using parallel processors significant speed-ups can be obtained. Numerical results from our parallel implementation on a hypercube multicomputer with 64 processors show an efficiency of about 60%. Current research concentrates on the theory and the implementation of methods for solving a general class of multi-stage problems, where the number of scenarios grows exponentially with the number of

stages. Further research includes improved decomposition techniques for large-scale problems, e.g., optimized tree traversing strategies and passing information based on non-optimal subproblems, improvements to the importance sampling approach, e.g., using different types of approximation functions, improved software, e.g., a parallel implementation of the multi-stage algorithm on distributed workstations, and the testing of the methodology on different practical problems in different areas.

References

[1] Beale E.M.L. (1955): On Minimizing a Convex Function Subject to Linear Inequalities, *J. Roy. Stat. Soc. 17b, 173-184.*

[2] Benders, J.F. (1962): Partitioning Procedures for Solving Mixed-Variable Programming Problems, *Numerische Mathematic 4, 238-252.*

[3] Birge, J.R. (1985): Decomposition and Partitioning Methods for Multi-Stage Stochastic Linear Programming, *Operations Research 33, 989-1007.*

[4] Birge, J.R. and Wets, R.J. (1986): Designing Approximation Schemes for Stochastic Optimization Problems, in Particular For Stochastic Programs with Recourse,*Math. Progr. Study 27, 54-102.*

[5] Birge, J.R. and Wallace, S.W. (1988): A Separable Piecewise Linear Upper Bound for Stochastic Linear Programs, *SIAM J. Control and Optimization, 26, 3.*

[6] Dantzig, G.B. (1955): Linear Programming under Uncertainty, *Management Science 1, 197-206.*

[7] Dantzig, G.B. and Glynn, P.W. (1990): Parallel Processors for Planning Under Uncertainty, *Annals of Operations Research 22, 1-21.*

[8] Dantzig, G.B., Glynn, P.W., Avriel, M., Stone, J., Entriken, R., Nakayama, M. (1989): Decomposition Techniques for Multi-Area Generation and Transmission Planning under Uncertainty, EPRI report 2940-1.

[9] Dantzig, G.B., Ho, J.K. and Infanger, G. (1991): Solving Stochastic Linear Programs on a Hypercube Multicomputer, Technical Report SOL 91-10, Department of Operations Research, Stanford University, Stanford, CA.

[10] Dantzig, G.B., and Infanger, G. (1991): Multi-Stage Stochastic Linear Programs for Portfolio Optimization, Technical Report SOL 91-11, Department of Operations Research, Stanford University, Stanford, CA, to appear in *Ann als of Operations Research.*

[11] Dantzig, G.B. and Infanger, G.(1991): Large-Scale Stochastic Linear Programs: Importance Sampling and Benders Decomposition, Technical Report SOL 91-4, Department of Operations Research, Stanford University, Stanford, CA.

[12] Dantzig, G.B. and Madansky M. (1961): On the Solution of Two-Staged Linear Programs under Uncertainty, *Proc 4th Berkeley Symp. on Mathematical Statistics and Probability I, ed. J. Neyman, 165-176.*

[13] Davis, P.J., and Rabinowitz, P. (1984): Methods of Numerical Integration, Academic Press, London.

[14] Ermoliev, Y. (1983): Stochastic Quasi-gradient Methods and Their Applications to Systems Optimization, *Stochastics 9, 1-36.*

[15] Ermoliev,Y. and R.J. Wets Eds. (1988): Numerical Techniques for Stochastic Optimization, Springer Verlag.

[16] Frauendorfer, K. (1988): Solving SLP Recourse Problems with Arbitrary Multivariate Distributions – The Dependent Case, *Mathematics of Operations Research, Vol. 13, No. 3, 377-394.*

[17] Frauendorfer K. and Kall, P. (1988): Solving SLP Recourse Problems with Arbitrary Multivariate Distributions – The Independent Case, *Problems of Control and Information Theory, Vol 17 (4), 177-205.*

[18] Frauendorfer, K. (1992): *Stochastic Two-Stage Programming,* Lecture Notes in Economics and Mathematical Systems 392, Springer Verlag, Berlin, Germany.

[19] Glynn, P.W. and Iglehart, D.L. (1989): Importance Sampling for Stochastic Simulation, *Management Science, Vol 35, 1367-1392.*

[20] Higle, J.L. and Sen, S. (1991): Stochastic Decomposition: An Algorithm for Two-Stage Linear Programs with Recourse, *Mathematics of Operations Research 16/3, 650–669.*

[21] Infanger, G. (1992): Monte Carlo (Importance) Sampling within a Benders Decomposition Algorithm for Stochastic Linear Programs, *Annals of Operations Research 39, 69-95.*

[22] Infanger, G. (1992): Planning Under Uncertainty – *Solving Large-Scale Stochastic Linear Programs,* Habilitationsschrift, to be published by The Scientific Press.

[23] Kall, P. (1979): Computational Methods for Two Stage Stochastic Linear Programming Problems, *Z. angew. Math. Phys. 30, 261-271.*

[24] Krishna, A.S. (1993): Importance Sampling Techniques for Stochastic Linear Programming, Dissertation in progress, Systems Optimization Laboratory, Department of Operations Research, Stanford University, Stanford, CA.

[25] Morton, D.P. (1993): Algorithmic Advances in Multistage Stochastic Programming, Dissertation in progress, Systems Optimization Laboratory, Department of Operations Research, Stanford University, Stanford, CA.

[26] Murtagh, B.A. and Saunders, M.A. (1983): MINOS User's Guide, SOL 83-20, Department of Operations Research, Stanford University, Stanford CA 94305.

[27] Pereira, M.V., Pinto, L.M.V.G., Oliveira, G.C. and Cunha, S.H.F. (1989): A Technique for Solving LP-Problems with Stochastic Right-Hand Sides, CEPEL, Centro del Pesquisas de Energia Electria, Rio de Janeiro, Brazil.

[28] Pereira, M.V., Pinto, L.M.V.G. (1989): Stochastic Dual Dynamic Programming, Technical Note, DEE- PUC/RJ – Catholic University of Rio de Janeiro, Caixa Postal 38063 Gávea, Rio de Janeiro RJ CEP 22452 Brazil.

[29] Rockafellar, R.T. and Wets, R.J-B. (1989): Scenario and Policy Aggregation in Optimization Under Uncertainty, *Mathematics of Operations Research 16, 119–147.*

[30] Ruszczynski, A. (1986): A Regularized Decomposition method for Minimizing a Sum of Polyhedral Functions, *Mathematical Programming 35, 309-333.*

[31] Van Slyke and Wets, R.J. (1969): L-Shaped Linear Programs with Applications to Optimal Control and Stochastic Programming, *SIAM Journal of Applied Mathematics, Vol 17, 638-663.*

[32] Wets R.J. (1984): Programming under Uncertainty: The Equivalent Convex Program, *SIAM J. on Appl. Math. 14, 89-105.*

****

BULLETIN BOARD

GEORGE B. DANTZIG PRIZE 1994

Nominations are solicited for the George B. Dantzig Prize, administered jointly by the Mathematical Programming Society (MPS) and the Society for Industrial and Applied Mathematics (SIAM). This prize is awarded to one or more individuals for original research which, by virtue of its originality, breadth, and depth, is having a major impact on the field of mathematical programming. The contributions eligible for consideration must be publicly available and may address any aspect of mathematical programming in its broadest sense. Strong preference is given to contributions by individuals under 50 years of age.

The prize will be presented at the Mathematical Programming Society's triennial symposium to be held August 15-19, 1994, in Ann Arbor, Michigan, USA. Past Dantzig Prize recipients have been: M.J.D. Powell and R.T. Rockafellar in 1982, E.L. Johnson and M.W. Padberg in 1985, M.J. Todd in 1988, and M. Groetschel and A.S. Nemirovsky in 1991.

The Prize Committee members are: Michael J. Todd, Chair, Martin Groetschel, Ellis L. Johnson, and R. Tyrrell Rockafellar.

Please send nominations to Michael J. Todd, School of Operations Research & Industrial Engineering, 206 Engineering & Theory Center Building, Cornell University, Ithaca, New York, 14853-3801, USA, or electronically to miketodd@cs.cornell.edu. Nominations are due by September 30, 1993 and should provide a brief one or two page description of the nominee's outstanding contributions and, if possible, a current resume including a list of the nominee's publications.

A.W. TUCKER PRIZE 1994

The Mathematical Programming Society invites nominations for the A.W. Tucker Prize for an outstanding paper authored by a student. The award will be presented at the International Symposium on Mathematical Programming in Ann Arbor (15-19 August 1994). All students, graduate and undergraduate, are eligible. Nominations of students who have not yet received the first university degree are especially welcome. In advance of the Symposium an award committee will screen the nominations and select at most three finalists. The finalists will be invited, but not required, to give oral presentations at a special session of the Symposium. The award committee will se-

lect the winner and present the award prior to the conclusion of the Symposium. The members of the committee for the 1994 A.W. Tucker Prize are : Thomas M. Liebling, Swiss Federal Institute of Technology, Lausanne; Andrew R. Conn, Thomas J. Watson Research Center, Yorktown Heights, William H. Cunningham, University of Waterloo, Clovis Gonzaga, COPPE, Federal University of Rio de Janeiro and Jean-Philippe Vial, University of Geneva.

The paper may concern any aspect of mathematical programming; it may be original research, an exposition or survey, a report on computer routines and computing experiments, or a presentation of a new and interesting application. The paper must be solely authored, and completed after January 1991. The paper and the work on which it is based should have been undertaken and completed in conjunction with a degree program.

Nominations must be made in writing to the chairman of the award committee, namely, Thomas M. Liebling, Swiss Federal Institute of Technology, Department of Mathematics, MA(Ecublens), CH-1015 Lausanne, Switzerland, by a faculty member at the institution where the nominee was studying for a degree when the paper was completed. Letters of nomination must be accompanied by four copies each of : the student's paper; a separate summary of the paper's contributions, written by the nominee, and no more than two pages in length; and a brief biographical sketch of the nominee.

Nominations must be sent to the chairman no later than December 31, 1993. (Postmark on recommended letter).

OPT–NET: A COMMUNICATION FACILITY FOR SCIENTISTS INTERESTED IN OPTIMIZATION

The Special Interest Group for OPTimization (SIGOPT) of the Deutsche Mathematiker–Vereinigung (DMV) has started a new electronic service: the operation of the OPT–NET. This facility is not limited to SIGOPT, it is also open to all scientists and students of mathematics, computer science, economics, electrical engineering etc. who are interested in optimization or who want to obtain the relevant information.

The basic services provided by OPT–NET are:

- a moderated forum for discussion with a weekly digest
- a unique E–mail address for each subscriber (virtual mailbox)
- a WhitePage service with "whois"–command
- an OPT–NET archive within eLib (a library for mathematical software)

This facility has been implemented at the Konrad–Zuse–Zentrum in Berlin (ZIB), which provides the technical infrastructure for its operation. Utilization of OPT–NET is free of charge.

For a first start send a message with the text "help" in the subject line to

"opt–net–request@zib–berlin.de".

Then you will receive automatically a document which gives you the information on all services offered by OPT–NET, in particular on how to register with OPT–NET, how to obtain the digest and, of course, how to become a member of SIGOPT.

Contributions to the discussion forum should be sent to

"opt–net@zib–berlin.de".

The subject line of a contribution should contain the senders name (followed by a colon) followed by a brief description of the contents of the contribution.

All contributions will be examined by a moderator, presently Prof. Dr. Uwe Zimmermann (TU Braunschweig), who selects those of general interest and combines them into the weekly digest, which will be distributed to all members of OPT–NET.

PRACTITIONERS' OPTIMIZATION WORKSHOP

Date: Spring, 1994
Given by:
Logistics Optimization Center
School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0205
Tel: (404) 894-4550 or 894-2379
FAX: (404) 894-2301
e-mail: POWorkshop@akula.isye.gatech.edu
Description: The workshop deals with modeling and solution techniques for large scale optimization problems, especially linear and mixed integer programming problems. The mathematical modeling software used is AMPL, and the optimization software used is OSL. The presentations and practical sessions are geared toward practitioners applying mathematical programming in the workplace.
Lecturers: Lloyd Clarke, Ellis Johnson,
George Nemhauser, and Martin Savelsbergh
Prerequisites: Practical optimization experience or POW I (held in early September, 1994); Working knowledge of FORTRAN or C; Experience with OSL and UNIX is helpful, but not required.
Objective: Understand the special techniques needed to solve large and difficult LP & IP problems, and how to implement them using OSL.
Price (the bottom line): $750 per course.

SELECTED UPCOMING ARTICLES FOR SIAM J. OPTIMIZATION

Convex Functions with Unbounded Level Sets and Applications to Duality Theory *A. Auslender, R. Cominetti, and J.-P. Crouzeix*

Reducing Matching to Polynomial Size Linear Programming *Francisco Barahona*

A Quadratically Convergent Polynomial Algorithm for Solving Entropy Optimization Problems *Florian Potra and Yinyu Ye*

Higher Order Predictor-Corrector Interior Point Methods with Application to Quadratic Objectives *Tamra J. Carpenter, Irvin J. Lustig, John M. Mulvey, and David F. Shanno*

A Global Optimization Algorithm for Concave Quadratic Programming Problems *Immanuel M. Bomze and Gabriele Danninger*

A Note on K Best Solutions to the Chinese Postman Problem *Yasufumi Saruwatari and Tomomi Matsui*

Generating Fenchel Cutting Planes for Knapsack Polyhedra *E. A. Boyd*

On Mizuno's Rank One Updating Algorithm for Linear Programming *Robert A. Bosch*

Primal-Dual Projected Gradient Algorithms for Extended Linear-Quadratic Programming *Ciyou Zhu and R. T. Rockafellar*

Accelerated Stochastic Approximation *Bernard Delyon and Anatoli Juditsky*

The D_2^*-Triangulation for Continuous Deformation Algorithms to Compute Solutions of Nonlinear Equations *Chuangyin Dang*

α-Lower Subdifferentiable Functions *J. E. Martínez-Legaz and S. Romano-Rodríguez*

Parallel Projected Aggregation Methods for Solving the Convex Feasibility Problem *Ubaldo Garcia-Palomares*

A Central Cutting Plane Algorithm for Convex Semi-Infinite Programming Problems *K. O. Kortanek and Hoon No*

On Sensitivity Analysis of Nonlinear Programs in Banach Spaces: The Approach via Composite Unconstrained Optimization *Alexander Ioffe*

A Nonconvex Duality with Zero Gap and Applications *Phan Thien Thach*

A Path-Following Projective Interior Point Method for Linear Programming *Dong Shaw and Donald Goldfarb*

Fast Approximation Schemes for Convex Programs with Many Blocks and Coupling Constraints *Michael D. Grigoriadis and Leonid G. Khachiyan*

On the Use of Product Structure in Secant Methods for Nonlinear Least Squares Problems *J. Huschens*

Existence and Differentiability of Metric Projections in Hilbert Spaces *Alexander Shapiro*

Approximating Oracle Machines for Combinatorial Optimization *Shmuel Onn*

Strong Rotundity and Optimization *J. M. Borwein and A. S. Lewis*

Local Minimizers of Quadratic Functions on Euclidean Balls and Spheres *Jose Mario Martinez*

On the Maximization of a Concave Quadratic Function with Box Constraints *Ana Friedlander and Jose Mario Martinez*

Short Steps with Karmarkar's Projective Algorithm for Linear Programming *J.-L. Goffin and J.-Ph. Vial*

On the Convergence of a Class of Infeasible Interior-Point Methods for the Horizontal Linear Complementarity Problem *Yin Zhang*

CONTRIBUTIONS TO THE V&N

The set of essays previously planned for this issue could not all be accomodated. Therefore, the remaining essays will appear in the Spring, '94 issue, which will continue the focus on nonsmooth optimization and optimization under uncertainty

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 Spring, '94 issue is February 1, 1994.

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