AD-A081 067 STANFORD UNIV CALIF SYSTEMS OPTIMIZATION LAB F/G 12/1
A GLOBALLY CONVERGENT CONDENSATION METHOD FOR GEOMETRIC PROGRAM--ETC(U)
OCT 79 E ROSENBERG DAAG29-79-C-0110
UNCLASSIFIED SOL-79-11 ARO-16470.2-M NL

END
DATE FILMED
3-80
DDC

16490.2-M

Systems
Optimization
Laboratory

(1)

LEVEL II

AD A081067

DDC FILE COPY

DTIC
ELECTE
FEB 27 1980
S D
C

Department of Operations Research
Stanford University
Stanford, CA 94305

SYSTEMS OPTIMIZATION LABORATORY
DEPARTMENT OF OPERATIONS RESEARCH
Stanford University
Stanford, California
94305

(9) Technical rept.

(6)

A GLOBALLY CONVERGENT CONDENSATION METHOD FOR GEOMETRIC PROGRAMMING.

by

(10) Eric Rosenberg

(14) TECHNICAL REPORT SOL-79-11
(11) October 1979

(12) 28

Research and reproduction of this report were partially supported by the Department of Energy Contract DE-AS03-76SF00326 PA Number DE-AT-03-76ER72018; the National Science Foundation Grant MCS76-20019 Ao1; and the U.S. Army Research Office Contract (15) DAAG29-79-C-0110, DOE-DE-AS03-76SF00326

A GLOBALLY CONVERGENT CONDENSATION METHOD FOR GEOMETRIC PROGRAMMING

by

Eric Rosenberg

Department of Operations Research
Stanford University
Stanford, CA 94305

1. Introduction

Geometric programming is a branch of nonlinear programming that considers certain primal nonlinearly constrained convex programs and their associated dual linearly constrained convex programs. The development of geometric programming is largely the work of Duffin, Peterson, and Zener [5]. The strong duality theory [4,5] of geometric programming states that, under mild conditions, to solve a primal one can solve its dual and then solve a single system of linear equations, known as the invariance conditions. Since many practical problems, especially in optimal engineering design, [16,17], are naturally cast as primal problems, a major computational advantage of this duality is that, when the dual is solved with a feasible direction method, the linear equality constraints restrict the choice of feasible directions. For such methods, the computational effort required to solve the dual increases with the dimension, called the degree of difficulty, of the dual feasible region.

To solve a primal, rather than solving its dual directly, we propose solving the duals of a sequence of approximating primals. Each approximating primal is obtained from the primal by using the current primal solution estimate to condense [4] certain terms in the constraints and the objective

function, as specified by a canonical submatrix of the matrix of exponents. The dual of each approximating primal has fewer degrees of difficulty than the dual of the given primal. Therefore, our scheme is to solve a sequence of problems, each posed in the same low dimensional space, rather than solve one problem in a higher dimensional space.

In this paper, all vectors are column vectors, unless otherwise specified, and x' and A' denote the transpose of the vector x and the matrix A, respectively. If $x, y \in R^m$, by $x = y$, $x \geq y$, and $x > y$, we mean $x_j = y_j$, $x_j \geq y_j$, and $x_j > y_j$, respectively, for $j = 1,2,\ldots,m$. The end of a proof is denoted by $\otimes$.

2. Primal and Dual Geometric Programs

Geometric programming studies pairs (P, D) of optimization problems. The primal P is defined on R^m:

$$\text{minimize} \quad h_0(x)$$

$$\text{subject to} \quad h_k(x) \leq 1, \qquad k = 1,2,\ldots,p,$$

where

$$h_k(x) = \sum_{i \in [k]} c_i e^{a_i x}, \qquad k = 0,1,\ldots,p;$$

$[k]$ are successive blocks of the integers $1, 2, \ldots, n$: $[k] = \{m_k, m_k+1,\ldots,n_k\}$, where $m_0 = 1$, $m_1 = n_0 + 1$, $m_2 = n_1 + 1,\ldots, m_p = n_{p-1}+1, \ldots, n_p = n$; c_i is positive for $i = 1,2,\ldots, n$, and $a_i = (a_{i1}, a_{i2}, \ldots, a_{im})$ is a row vector, with a_{ij} unconstrained in sign for each i and j. Here $a_i x$ means $\sum_{j=1}^{m} a_{ij} x_j$. Each h_k is called a posynomial, and h_k is said to have

$|[k]|$ terms, where $|S|$ denotes the number of elements in set S. The primal is said to have n terms, since $n = \sum_{k=0}^{p} |[k]|$.

The dual D is defined on R^n:

$$\text{maximize} \quad v(\delta) = \prod_{i=1}^{n} \left(\frac{c_i}{\delta_i}\right)^{\delta_i} \prod_{k=1}^{p} \lambda_k^{\lambda_k}$$

$$\text{subject to} \quad \delta \in D$$

where

$$\lambda_k = \sum_{i \in [k]} \delta_i \quad \text{for} \quad k = 1,2,\ldots, p,$$

$$D = \{\delta | \delta' A = 0, \sum_{i \in [0]} \delta_i = 1, \text{ and } \delta \geq 0\}$$

and

$$A = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

is the $n \times m$ matrix of exponents. Here $\delta' A$ denotes the m-dimensional row vector whose j^{th} coordinate is $\sum_{i=1}^{n} \delta_i a_{ij}$.

Clearly, P is a convex program, as is D when $v(\delta)$ is replaced with $-\log v(\delta)$ [5]. From [4,5] we learn that if x is feasible for P and δ is feasible for its dual D, then $h_0(x) \geq v(\delta)$. Under the same conditions, equality holds if and only if the <u>invariance conditions</u>

$$\delta_i = \begin{cases} \frac{1}{h_0(x)} c_i e^{a_i x}, & i \in [0] \\ \lambda_k c_i e^{a_i x}, & i \in [k], \; k = 1,2,\ldots, p \end{cases}$$

are satisfied. Taking logarithms of each equation generates a system of equations linear in x. Since $\delta_\ell > 0$ for some ℓ in $[k]$ if and only if $\delta_i > 0$ for each i in $[k]$, the invariance conditions are a set of $|[0]| + \sum_{k \in K_+} |[k]|$ nontrivial equations in m unknowns, where $K_+ = \{k | \lambda_k > 0,\ k = 1,2,\ldots,p\}$.

A program is consistent if its feasible region is nonempty. We say P is superconsistent if for some x^0 we have $h_k(x^0) < 1$, $k = 1,2,\ldots, p$. If P is consistent and has a positive infimum ω, then its dual is consistent and has a finite supremum ξ, and $\omega = \xi$. If also P is superconsistent, then ξ is attained at a dual feasible point [4,5].

We are interested in uniqueness of solution in geometric programming. Recall that the posynomial h is strictly convex if its exponent matrix has full rank (see p. 31, Zener [17]). We write

$$A = \begin{pmatrix} A_{[0]} \\ A_{[1]} \\ \vdots \\ A_{[p]} \end{pmatrix}$$

where $A_{[k]}$ contains those rows of A corresponding to h_k, for $k = 0,1,\ldots, p$.

By a Kuhn-Tucker pair (z,u) for P we mean that z solves P and u is a vector of Lagrange multipliers for P.

Proposition 1. For some Kuhn-Tucker pair (z,u) of P, let E be that submatrix of A obtained by deleting $A_{[k]}$ for each k satisfying $u_k = 0$. If rank $E = m$, then z is the unique solution of P.

Proof. Since (z,u) is a Kuhn-Tucker pair for P, then z minimizes

$$\sum_{i \in [0]} c_i e^{a_i x} + \sum_{k=1}^{p} u_k \sum_{i \in [k]} c_i e^{a_i x} .$$

Since $u \geq 0$ and rank $E = m$, it follows that z is the unique minimum of this function. By Corollary 28.1.1, Rockafellar [12], z is the unique solution of P. ⊗

Proposition 2. Suppose P is superconsistent and has the unique Kuhn-Tucker pair (z,u). Let E be as in Proposition 1. If rank $E = m$, then the dual D has a unique solution.

Proof. Each Lagrange multiplier u_k is related to the variable $\lambda_k = \sum_{i \in [k]} \delta_i$ of the dual by $\lambda_k = u_k / h_0(z)$, $k = 1,2,\ldots, p$ (see p. 119, Duffin, Peterson, and Zener [5]). Since each u_k is uniquely determined, then so is each λ_k. By the strong duality theory, the dual has at least one solution. Each dual solution is related to z by the invariance conditions. Since z is the unique solution of P and rank $E = m$, the invariance conditions generate the unique solution of the dual. ⊗

Proposition 1 implies that the ability to determine in advance which primal constraints must be active at any primal solution can prove a priori that the primal has a unique solution. The monotonicity analysis of Wilde [16]

and Papalambros [11] is very useful for determining constraint activity. Proposition 1 is important because proving global convergence of our method requires a priori knowledge that the given primal and each approximating primal have unique solutions. Proposition 2 asserts that, under reasonable conditions, the dual has a unique solution if the primal does. This theorem is useful for our method, which solves only dual problems.

3. Canonical Primals and Degrees of Difficulty

Our method can be applied to a geometric primal P only if P is canonical, that is, if its solution set is nonempty and bounded. It is reasonable to expect a well-formulated problem to be canonical. If P is consistent but not canonical, then the solution set is either empty or unbounded. In the latter case, by deleting terms or variables, we can reduce P to a canonical program [1,5].

Proposition 3. Suppose P is consistent. Then the following are equivalent:

1) P is canonical

2) $Ad \leq 0$ implies $d = 0$

3) rank $A = m$ and $\delta^t A = 0$, $\delta > 0$ has a solution.

Proof. 1) <==> 2). See Abrams [1]. Notice that "$Ad \leq 0$ implies $d = 0$" says that the objective function and constraints have no common direction of recession. 2) <==> 3). By Stiemke's Theorem [9], $Ad \leq 0$, $Ad \neq 0$ has no solution if and only if $\delta^t A = 0$, $\delta > 0$ has a solution. Moreover, $Ad = 0$ has no nonzero solution if and only if rank $A = m$. ●

We write

$$A = \begin{pmatrix} A_{[0]} \\ A_{[\sim 0]} \end{pmatrix}, \quad \text{where} \quad A_{[\sim 0]} = \begin{pmatrix} A_{[1]} \\ A_{[2]} \\ \vdots \\ A_{[p]} \end{pmatrix}.$$

If P is consistent, its feasible region is bounded if and only if $A_{[\sim 0]}d \leq 0$ implies $d = 0[1]$.

Notice that, for a consistent primal, the cost coefficients $\{c_1\}$ play no role in determining whether or not P is canonical. This is a wonderful feature, since cost coefficients are usually not known exactly, while the exponent matrix is usually determined from the laws of science and therefore <u>is</u> often known exactly.

From Proposition 3 and the strong duality theory, it is easy to prove that, if P is canonical, its dual is consistent and has $n-m-1$ degrees of difficulty. If, for each dual feasible point, one or more components are always zero, then, as discussed in Abrams [1], these components and the corresponding primal terms can be dropped. In this case, the degree of difficulty is strictly less than $n-m-1$.

Notice that, for a fixed number of primal variables, the degree of difficulty increases with the number of primal terms. Although geometric programming succeeds with duals of small degree of difficulty, turning to the dual can in many cases produce unreasonably large problems. Our method modifies the primal to reduce the number of terms, thus lowering the degree of difficulty of its dual.

4. Condensation and Canonical Submatrices

The key to reducing the number of primal terms is a technique called condensation, discovered by Duffin [4]. Wilde [16] has used condensation to simplify problems in optimal engineering design.

Condensation maps a posynomial into a one-term posynomial. Consider $h(x) = \sum_{i=1}^{n} c_i e^{a_i x}$. For every x, each term of $h(x)$ is positive. Given the point y, form the vector $\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$, where

$$\varepsilon_i = c_i e^{a_i y} \Big/ h(y) , \qquad i = 1,2,\ldots, n.$$

Since $\varepsilon_i > 0$ for each i and $\sum_{i=1}^{n} \varepsilon_i = 1$, the vector ε is called a vector of primal weights for h at y. We define the total condensation of h about y to be the function $\bar{h}(\cdot, y)$ given by $\bar{h}(x,y) = \bar{c} e^{\bar{a}x}$, where

$$\bar{c} = \prod_{i=1}^{n} \left(\frac{c_i}{\varepsilon_i}\right)^{\varepsilon_i} \qquad \text{and} \qquad \bar{a} = \sum_{i=1}^{n} \varepsilon_i a_i .$$

Notice that $\bar{h}(\cdot, y)$ is a one-term posynomial. It is well-known [2,4] that $h(y) = \bar{h}(y,y)$, $\nabla h(y) = \nabla_1 \bar{h}(y,y)$ (where ∇_1 denotes the gradient with respect to the first argument), and $h(x) \geq \bar{h}(x,y)$ for every x.

We now present a scheme for condensing a canonical primal to reduce the degree of difficulty. Section 3 proved that a consistent primal P is canonical if and only if its exponent matrix A satisfies $Ad \leq 0$ implies $d = 0$. Suppose now that P is consistent and canonical. One can often delete rows of A so that the resulting matrix, denoted by A^c, satisifes the following conditions:

1) $A^c d \leq 0$ implies $d = 0$

2) at least two rows of $A_{[k]}$ are deleted for some k in $\{0,1,\ldots, p\}$.

If A^c satisfies 1) and 2), we call A^c a <u>canonical submatrix</u> of A. (This definition differs somewhat from that in Rosenberg [14].)

The concept of a canonical submatrix has a simple geometric interpretation. Consider a canonical unconstrained P whose objective function has two variables and five terms. By Proposition 3, the matrix A must have full rank, and $\delta' A = 0$, $\delta > 0$ must have a solution. Such a situation is shown in Figure 1. Notice that a_1, a_3, and a_5 form a canonical submatrix. In contrast, a_4, a_5, and a_6 do not.

A canonical submatrix A^c tells which terms of P can be condensed without P becoming noncanonical. For each $k = 0,1,\ldots, p$, collect each term whose exponent vector a_i was deleted from A to form A^c. Add these terms together to build a posynomial, which we call h_k^T. Equivalently, h_k^T is obtained by deleting from h_k each term whose exponent vector appears in A^c.

Given the point y in R^m, totally condense h_k^T about y to form $\bar{h}_k^T(\cdot, y)$. For each $k = 0, 1, \ldots, p$, let

$$\bar{h}_k(x,y) = (h_k(x) - h_k^T(x)) + \bar{h}_k^T(x,y) .$$

That is, $\bar{h}_k(\cdot, y)$ is obtained from h_k by totally condensing about y all terms with exponent vectors not appearing in A^c. Finally, define the <u>condensed primal</u> $\bar{P}(y)$:

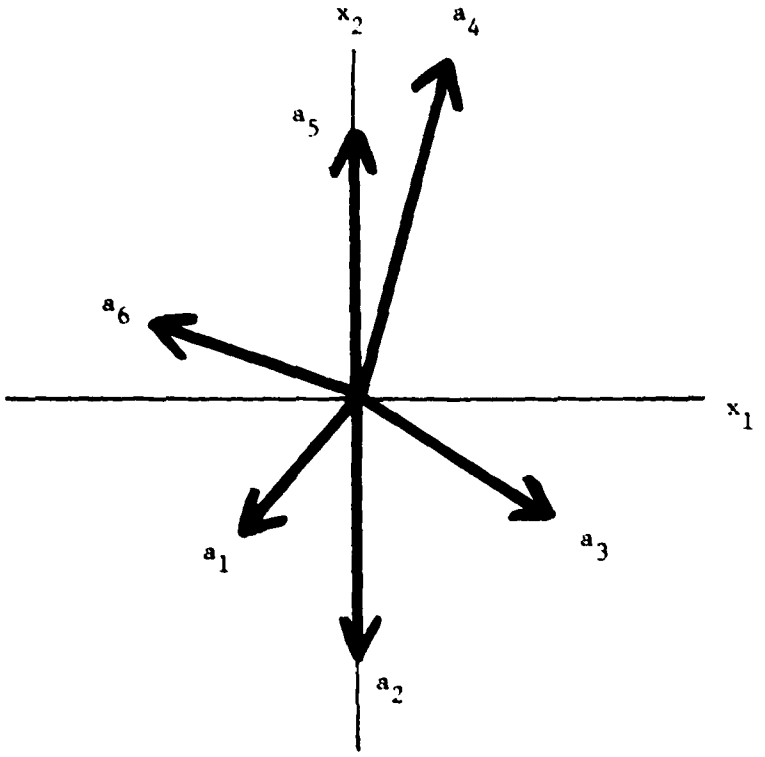

Fig. 1. Exponent matrix of a canonical unconstrained P

$$\underset{x}{\text{minimize}} \quad \bar{h}_0(x,y)$$

$$\text{subject to} \quad \bar{h}_k(x,y) \leq 1, \qquad k = 1,2, \ldots, p\ .$$

It follows immediately from the above remarks on condensation that, for every y, $\bar{P}(y)$ has a larger feasible region and a (pointwise) smaller objective function than P. Moreover, $\bar{P}(y)$ is canonical for every y, since the exponent matrix of $\bar{P}(y)$ has (after possibly reordering the rows) the form

$$\begin{pmatrix} A^c \\ W(y) \end{pmatrix}$$

for some matrix $W(y)$. Since $A^c d \leq 0$ implies $d = 0$, then $\bar{P}(y)$ is canonical, by Proposition 3. Moreover, condition 2) of the definition of a canonical submatrix implies that $\bar{P}(y)$ has fewer terms than P, so that its dual has fewer degrees of difficulty. An example of how to generate a condensed primal is presented in Section 7, along with numerical results.

5. Obtaining a Canonical Submatrix

The following is a procedure that searches for a canonical submatrix. Recall that a matrix A^c obtained by deleting rows of A is a canonical submatrix if

1) $A^c d \leq 0$ implies $d = 0$ and

2) at least two rows of $A_{[k]}$ are deleted for some k in $\{0,1,\ldots,p\}$.

The procedure uses linear programming to search for a submatrix of A satisfying 1). If the submatrix also satisfies 2), then it is a canonical submatrix.

We observed in Section 3 that P is canonical if $\delta'A = 0$, $\delta > 0$ has a solution and rank $A = m$.

Consider the Phase I [3] linear program:

$$\text{LP} \qquad \begin{aligned} \text{minimize} \quad & e_m'w + y \\ \text{subject to} \quad & A'\delta + w = 0 \\ & e_n'\delta + y = 1 \\ & (\delta, w, y) \geq 0 \end{aligned}$$

where $e_r \in R^r$ is a column vector of ones. Notice that LP has m+1 constraints. Clearly LP is consistent. Let γ be the optimal objective function value of LP. Then $\gamma \geq 0$.

Proposition 4. If P is canonical, then $\gamma = 0$.

Proof. If P is canonical, then $\delta'A = 0$, $\delta > 0$ has a solution. Therefore $\delta'A = 0$, $e_n'\delta = 1$, $\delta \geq 0$ has a solution, so that every solution of LP has $w = 0$ and $y = 0$. Hence $\gamma = 0$. ∞

By Proposition 4, if $\gamma > 0$ then P is not canonical; hence no canonical submatrix exists. In this case, reduction may be desirable (see Abrams [1]). Notice that

$$A = \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix}$$

provides a counterexample to the converse of Proposition 4: P is not canonical, since rank $A = 1$, yet $\delta = (1/2, 1/2)'$, $w = (0,0)'$, $y = 0$ solves LP and $\gamma = 0$.

Suppose the simplex method [3] used to solve LP gives $\gamma = 0$. If $m+1$ components of δ are positive in the optimal solution of LP, then the corresponding rows of A constitute a canonical submatrix if condition 2) is satisfied. Unfortunately, degenerate solutions of LP are possible, even when P is canonical. For example, in Figure 1, where $a_2 = -a_5$, the solution $\delta_2 = \delta_5 = 1/2$; $\delta_i = 0$, $i \neq 2, 5$ is a degenerate solution. Termination in the degenerate case does not automatically imply failure. For it may be possible to find alternative nondegenerate solutions of LP by, for example, bringing into the basis a column with zero reduced cost. It is possible to systematically determine all basic solutions of LP (see [6] and references therein) but such a procedure is generally unnecessary.

6. The Algorithm and a Stopping Criterion

This section presents the condensation algorithm and discusses some of the details of implementation. We also describe how the algorithm automatically generates a sequence of lower bounds, thus suggesting a stopping criterion that may substantially reduce computation.

To execute the algorithm for solving P, we require that P be superconsistent and that a canonical submatrix A^c be available. If

P satisfies only these two conditions then the algorithm can be executed. However, to prove the algorithm globally convergent, we must impose additional assumptions. Failure to establish that some given P satisfies these extra assumptions should not discourage use of the algorithm. To see how the algorithm might still be of use, suppose that we take an arbitrary point x as the first solution estimate. Since a canonical submatrix is available, then $\bar{P}(x)$ is canonical and has a dual easier to solve than the dual of P. It might be that a solution of $\bar{P}(x)$, obtained after relatively small effort by solving the dual of $\bar{P}(x)$, also qualifies as an acceptable solution of P, so that no further computation is necessary. The important word here is "acceptable." "He that knoweth when he hath enough is no fool." (John Heywood (1546) as quoted in Wilde [16].) For example, it is foolish to compute an extremely accurate solution to a problem formulated with inaccurate data.

We now discuss the extra assumptions needed to insure global convergence. First is that P have a bounded feasible region. As mentioned in Section 3, this is true if and only if

$$A_{[\sim 0]}d \leq 0 \qquad \text{implies} \qquad d = 0 ,$$

where $A_{[\sim 0]}$ is that submatrix of A obtained by deleting the exponent vector of each objective function term. If our problem fails to satisfy this condition, we could impose additional constraints bounding the feasible region so that no solution of P violates the additional constraints. Also,

P and <u>each</u> $\bar{P}(x)$ (that is, for <u>any</u> x) must be known <u>in advance</u> to have unique solutions. Sufficient conditions for a primal to have a unique solution have been derived in Section 2. As mentioned there, verifying that a primal has a unique solution can often be accomplished by determining constraint activity.

A solution of $\bar{P}(x)$ defines a descent direction of an exact penalty function. These interesting functions have been well studied (see [8], [13], and references therein). Here we simply define θ_ρ, the exact penalty function of the primal P, by

$$\theta_\rho(x) = h_0(x) + \rho \sum_{k=1}^{p} \max(0, h_k(x) - 1) ,$$

where ρ is a fixed positive number.

Notice that θ_ρ is convex and nondifferentiable. Our method for solving P requires a line search on θ_ρ at each iteration. Fibonacci or golden section [15] or superlinearly convergent [10] line search routines can be utilized.

Our algorithm is the following. Let the positive numbers ρ and β be given. Choose any x_0 in R^m.

<u>Algorithm</u> *GPA*. For $i = 0, 1, 2, \ldots$

1) solve the condensed primal $\bar{P}(x_i)$ to obtain a solution z_i; let

   $d_i = z_i - x_i$

2) find an α_i such that

$$\alpha_i \in \arg\min_{0 \leq \alpha \leq \beta} \theta_\rho(x_i + \alpha d_i) \ ;$$

let $x_{i+1} = x_i + \alpha_i d_i$

3) STOP if $x_{i+1} = x_i$; otherwise, go to 1) with i replaced by $i+1$.

We now present the global convergence theorem.

Theorem 1. Suppose that P is superconsistent and has a bounded feasible region. Suppose also that P and $\bar{P}(x)$ have unique solutions for any x. Then there is a positive number ρ_0 such that, whenever $\rho \geq \rho_0$, algorithm *GPA* either stops at the unique solution of P or generates a sequence $\{x_i\}$ converging to it.

Proof. Since condensation preserves function value and first partial derivatives at the point of condensation, and since the objective function h_0 is everywhere positive, the result follows immediately from this author's convergence theory [13]. ⊗

Notice that the success of algorithm *GPA* depends on choosing a sufficiently large value for ρ. We offer no a priori advice on how to choose ρ, since the threshold value ρ_0 is actually a function of the solution set of P, of course unknown.

An important feature of algorithm *GPA* is that it automatically produces a sequence of lower bounds on θ_ρ, the exact penalty function associated with P. We now characterize this sequence and explain how to use it as a very effective stopping criterion.

At each iteration i of *GPA*, we solve $\bar{P}(x_i)$ (via its dual) to obtain a solution z_i. The optimal objective function value of $\bar{P}(x_i)$ is $\bar{h}_0(z_i,x_i)$. Since each function defining $\bar{P}(x_i)$ underestimates the corresponding function defining P, it must be that $h_0(z) \geq \bar{h}_0(z_i,x_i)$, where $h_0(z)$ is the optimal objective function value of P. Also, since $\{\theta_\rho(x_i)\}$ descreases monotonically to $\theta_\rho(z) = h_0(z)$, then $\theta_\rho(x_i) \geq h_0(z)$. Combining the above, we have $\theta_\rho(x_i) \geq h_0(z) \geq \bar{h}_0(z_i,x_i)$. This shows that the sequences $\{\theta_\rho(x_i)\}$ and $\{\bar{h}_0(z_i,x_i)\}$ bound $h_0(z)$ from above and below, respectively. Since z solves $\bar{P}(z)$, the bounds are tight, that is, $\theta_\rho(z) = h_0(z) = \bar{h}_0(z,z)$.

We cannot say that the sequence $\{\bar{h}_0(z_i,x_i)\}$ is monotone increasing. However, non-monotonic behavior occurred in only one test problem, suggesting that non-monotonic behavior is the exception rather than the rule.

The lower bounds provide an excellent stopping criterion: accept x_i as a solution of P if $\theta_\rho(x_i) - \max_{0 \leq j \leq i} \bar{h}_0(z_j,x_j) \leq \epsilon_1$, for some positive number ϵ_1. This test is especially useful when the objective function is not known to great accuracy, as in optimal design problems. This criterion might be used in conjunction with the requirement that x_i not violate any constraint by more than ϵ_2, for some positive number ϵ_2. Notice that these lower bounds are <u>not</u> generated when the primal is solved by the well-known technique of solving quadratic subproblems in an active set strategy [7]; indeed, the generation of these lower bounds is a very special feature of our method.

7. Computational Experience

This section discusses computational experience with algorithm *GPA*. The algorithm was programmed for the Stanford University IBM 360/70 computer in the language APL, especially suitable because of the ease with which it handles vectors and matrices. Moreover, since APL is interactive, changing the data or parameters and conducting sensitivity analysis is easy.

We solve each condensed dual by obtaining a feasible point and applying a projected Newton-like method (see Chapter 3, Gill and Murray [7]). This feasible direction method requires a line search on the condensed dual objective function at each iteration. We employ the golden section method [15] to solve the line search. The original interval of uncertainty is chosen to be min(A,B), where

A = the minimum step along the condensed dual search direction that causes some component of the solution estimate to hit zero

and

B = 100 times the length of the condensed dual search direction.

Each line search is stopped when the values of the function at either end of the current interval of uncertainty differ by no more than 10^{-8}. Each condensed dual optimization is stopped when successive line searches yield condensed dual objective function value estimates that differ by no more than 10^{-8}.

Our experience indicates that these tolerances cannot be greatly increased, since changing them to 10^{-3} led to jamming at x_1. That is, we obtained $x_1 = x_2 = x_3 = \cdots$.

We also used the golden section method for each exact penalty function line search. We choose $\beta = 10$; this value is more than adequate. Each exact penalty function line search is stopped when the values of the function at either end of the current interval of uncertainty differ by no more than 10^{-8}. Our experience indicates that this tolerance may be increased to 10^{-3} with no significant detrimental effect. However, we chose to use the former value.

For our test problems we chose $\rho = 400$. This choice was found to be adequate. According to the theory in [13], nothing is gained by increasing ρ beyond the threshold value ρ_0. Indeed, increasing ρ to 1000 in one problem produced a sequence of solution estimates identical to the fifth decimal place.

Representative computer results are given below. By θ_ρ we mean the exact penalty function corresponding to the given primal P. All entries in the table are rounded to five significant figures. At each iteration i, we supply

a. $\theta_\rho(z_i)$, where z_i solves $\bar{P}(x_i)$

b. $\theta_\rho(x_{i+1})$, where x_{i+1} solves the line search on θ_ρ

c. the objective function and constraints of P at x_{i+1}

d. the lower bound $\bar{h}_0(z_i, x_i)$.

Problem

$$\text{minimize} \quad e^{x_1+x_2}$$

$$\text{subject to} \quad \frac{1}{11} e^{-x_1} + \frac{2}{11} e^{x_1+2x_2} + \frac{3}{11} e^{x_2-x_1} + \frac{4}{11} e^{2x_1+x_2} \leq 1$$

$$\frac{5}{27} e^{-x_2} + \frac{6}{27} e^{x_1-x_2} + \frac{7}{27} e^{x_1-2x_2} + \frac{8}{27} e^{2x_1-x_2} \leq 1 .$$

This problem is superconsistent, canonical, and has six degrees of difficulty. Both constraints are active at the unique solution.

Method. Condense the last three terms of each constraint. Given y, $\bar{P}(y)$ is

$$\text{minimize} \quad e^{x_1+x_2}$$

$$\text{subject to} \quad \frac{1}{11} e^{-x_1} + \bar{c}_1 e^{\bar{a}_1 x} \leq 1$$

$$\frac{5}{27} e^{-x_2} + \bar{c}_2 e^{\bar{a}_2 x} \leq 1$$

where

$$\bar{c}_1 = \left(\frac{2/11}{\epsilon_1}\right)^{\epsilon_1} \left(\frac{3/11}{\epsilon_2}\right)^{\epsilon_2} \left(\frac{4/11}{\epsilon_3}\right)^{\epsilon_3} ,$$

$$\bar{a}_1 = (\epsilon_1, \epsilon_2, \epsilon_3) \begin{pmatrix} 1 & 2 \\ -1 & 1 \\ 2 & 1 \end{pmatrix},$$

$$(\varepsilon_1, \varepsilon_2, \varepsilon_3) = S^{-1}\left(\frac{2}{11}\, e^{y_1+2y_2}, \frac{3}{11}\, e^{y_2-y_1}, \frac{4}{11}\, e^{2y_1+y_2}\right),$$

$$S = \frac{2}{11}\, e^{y_1+2y_2} + \frac{3}{11}\, e^{y_2-y_1} + \frac{4}{11}\, e^{2y_1+y_2},$$

and

$$\bar{c}_2 = \left(\frac{6/27}{\beta_1}\right)^{\beta_1} \left(\frac{7/27}{\beta_2}\right)^{\beta_2} \left(\frac{8/27}{\beta_3}\right)^{\beta_3},$$

$$\bar{a}_2 = (\beta_1, \beta_2, \beta_3) \begin{pmatrix} 1 & -1 \\ 1 & -2 \\ 2 & -1 \end{pmatrix},$$

$$(\beta_1, \beta_2, \beta_3) = T^{-1}\left(\frac{6}{27}\, e^{y_1-y_2}, \frac{7}{27}\, e^{y_1-2y_2}, \frac{8}{27}\, e^{2y_1-y_2}\right),$$

$$T = \frac{6}{27}\, e^{y_1-y_2} + \frac{7}{27}\, e^{y_1-2y_2} + \frac{8}{27}\, e^{2y_1-y_2}.$$

Each $\bar{P}(y)$ is superconsistent, canonical, and has two degrees of difficulty.

We chose $\rho = 400$ and $x_0 = (4,6)$ which yields $\theta_\rho(x_0) = 8.2 \times 10^8$, $h_1(x_0) = 2.0 \times 10^6$, and $h_2(x_0) = 2.2$. The following table shows that four iterations produce an x_4 feasible for the original problem such that $h_0(x_4)$ is within 0.1 percent of the optimal objective function value. Notice that the lower bounds $\bar{h}_0(z_i, x_i)$ proved very useful in determining when an acceptable solution is at hand.

i	$\theta_\rho(z_i)$	$\theta_\rho(x_{i+1})$	$h_0(x_{i+1})$	$h_1(x_{i+1})$	$h_2(x_{i+1})$	$\bar{h}_0(z_i,x_i)$
1	535.13	30.636	0.21623	1.0761	0.49853	0.017098
2	29.496	7.7871	0.089265	1.0192	0.85597	0.065315
3	0.67405	0.25104	0.073065	1.0004	1.0000	0.072943
4	0.078688	0.073136	0.073136	1.0000	0.99989	0.073124

In conclusion, we have shown that solving condensed geometric programs in our globally convergent scheme, combined with a lower bound stopping criterion using the knowledge of how accurate a solution it is worthwhile to compute, provide a powerful new means of solving geometric programs.

Acknowledgment. The author is grateful to Douglass Wilde for his helpful suggestions on a preliminary draft of this paper.

REFERENCES

[1] R. A. Abrams, "Projections of Convex Functions with Unattained Infima," SIAM J. Control, 13, 1975, pp. 706-718.

[2] C. S. Beightler and D. T. Phillips, Applied Geometric Programming, Wiley, New York, 1976.

[3] G. B. Dantzig, Linear Programming and Extensions, Princeton University Press, 1963.

[4] R. J. Duffin, "Linearizing Geometric Programs," SIAM Review, 12, 1970, pp. 211-227.

[5] R. J. Duffin, E. L. Peterson, and C. Zener, Geometric Programming, Wiley, New York, 1967.

[6] M. E. Dyer and L. G. Proll, "An Algorithm for Determining all Extreme Points of a Convex Polytope," Mathematical Programming, 12, 1977, pp. 81-96.

[7] P. E. Gill and W. Murray, Numerical Methods for Constrained Optimization, Academic Press, New York, 1974.

[8] S.-P. Han and O. L. Mangasarian, "Exact Penalty Functions in Nonlinear Programming," Mathematics Research Center Technical Report No. 1897, University of Wisconsin at Madison, 1978.

[9] O. L. Mangasarian, Nonlinear Programming, McGraw Hill, New York, 1969.

[10] W. Murray and M. L. Overton, "Steplength Algorithms for Minimizing a Class of Nondifferentiable Functions," Stanford University Computer Science Dept., Report STAN-CS-78-679, 1978.

[11] P. Papalambros, Doctoral Dissertation, Dept. of Mechanical Engineering, Stanford University, 1979.

[12] R. T. Rockafellar, Convex Analysis, Princeton University Press, 1970.

[13] E. Rosenberg, "Globally Convergent Algorithms for Convex Programming," Stanford University, Dept. of Operations Research, Report 79-1, 1979 (submitted for publication).

[14] E. Rosenberg, "Minimizing Canonical Posynomials by Condensation to Zero Degrees of Difficulty," (submitted for publication).

[15] D. J. Wilde, Optimum Seeking Methods, Prentice-Hall, Englewood Cliffs, N.J., 1964.

[16] D. J. Wilde, Globally Optimal Design, Wiley, New York, 1971.

[17] C. Zener, Engineering Design by Geometric Programming, Wiley, New York, 1971.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER SOL 79-11	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A GLOBALLY CONVERGENT CONDENSATION METHOD FOR GEOMETRIC PROGRAMMING		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL REPORT 6. PERFORMING ORG. REPORT NUMBER 79-11
7. AUTHOR(s) Eric Rosenberg		8. CONTRACT OR GRANT NUMBER(s) ARO DAAG29-79-C-0110
9. PERFORMING ORGANIZATION NAME AND ADDRESS Operations Research Department - SOL Stanford University Stanford, CA 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Army Research Office Mathematics Div., Box CM Duke Station Durham, NC 27706		12. REPORT DATE October 1979 13. NUMBER OF PAGES 24
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

This document has been approved for public release and sale; its distribution is unlimited.

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

18. SUPPLEMENTARY NOTES

THE VIEW, OPINIONS, AND/OR FINDINGS CONTAINED IN THIS REPORT ARE THOSE OF THE AUTHOR(S) AND SHOULD NOT BE CONSTRUED AS AN OFFICIAL DEPARTMENT OF THE ARMY POSITION, POLICY, OR DECISION, UNLESS SO DESIGNATED BY OTHER DOCUMENTATION

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

Geometric Programming
Condensation
Global Convergence

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

SEE ATTACHED

over

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

SOL 79-11 Eric Rosenberg

A GLOBALLY CONVERGENT CONDENSATION METHOD FOR GEOMETRIC PROGRAMMING

To solve a posynomial geometric program, rather than solve its dual directly, one solves the duals of a sequence of approximating geometric programs. Each approximating program, obtained by condensing certain terms of the given problem, has fewer degrees of difficulty than the given problem. Therefore, one method is to solve a sequence of problems, each posed in the same low dimensional space, rather than solve one problem in a higher dimensional space. The method requires a special submatrix of the matrix of exponents, and procedures to find such a submatrix are presented. A line search on an exact penalty function is employed. Under mild conditions, from any arbitrary starting solution estimate the method generates a sequence of estimates converging to a solution. The method also generates a sequence of lower bounds, thus providing an effective stopping criterion. A numerical example is presented.