

**UNCLASSIFIED**

**AD**

**4 2 0 0 0 5**

**DEFENSE DOCUMENTATION CENTER**

**FOR**

**SCIENTIFIC AND TECHNICAL INFORMATION**

**CAMERON STATION, ALEXANDRIA, VIRGINIA**



**UNCLASSIFIED**

NOTICE: When government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related government procurement operation, the U. S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.

CATALOGED BY JJA

AS AD No. \_\_\_\_\_

42005

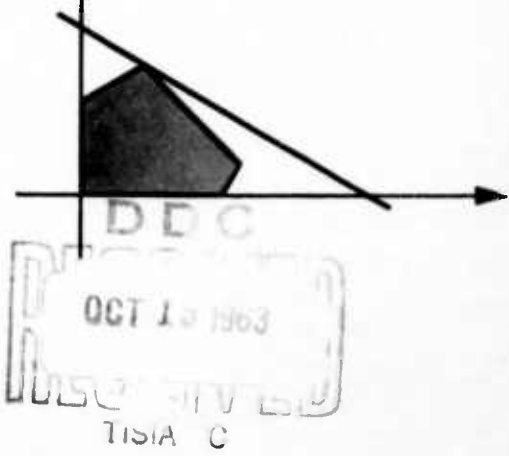
# CONVEX PROGRAMMING— DUAL ALGORITHM

by  
Pierre Huard

ORC 63-21 (RR)  
16 July 1963

OPERATIONS RESEARCH CENTER

INSTITUTE OF ENGINEERING RESEARCH



UNIVERSITY OF CALIFORNIA - BERKELEY

CONVEX PROGRAMMING - DUAL ALGORITHM

by

Pierre Huard\*  
Operations Research Center  
University of California, Berkeley

16 July 1963

ORC 63-21 (RR)

This research has been partially supported by the Office of Naval Research under Contract Nonr-222(83) with the University of California. Reproduction in whole or in part is permitted for any purpose of the United States Government.

\* The author was a visiting research engineer from the Electricite de France in Paris.

## ABSTRACT

This work is the result of a study of G. B. Dantzig's [1] algorithm for nonlinear programs where the objective as well as inequality constraints are convex functions. This algorithm consists of finding an equivalent "generalized linear program" where the standard procedure for determining the candidate to enter the basis is replaced by an auxiliary nonlinear program where the nonlinear constraints of the original problem have disappeared. The algorithm described in this paper is very close to Dantzig's algorithm, but the relations between the two algorithms are not yet well established.

I. Definitions-Properties

The initial program

PROBLEM 1: 
$$\left[ \begin{array}{l} \text{Find } x \text{ such that } \phi_0(x) \text{ is maximum} \\ \text{subject to } \phi_i(x) \geq 0 \quad i = 1, 2, \dots, m. \end{array} \right.$$

where  $\phi_0$  and  $\phi_i$  are continuously differentiable concave functions in  $x$ .

We will call  $\phi(x)$  or  $\phi$  the vector whose elements are  $\phi_i(x)$ .

We assume that the optimal value of  $\phi_0(x)$  is finite.

Let  $\pi = (\pi_1, \pi_2, \dots, \pi_m)$  be a row vector.

Consider the following program:

PROBLEM 3: 
$$\left[ \begin{array}{l} \text{Find } x \text{ such that } \psi(x, \pi) \text{ is maximum} \\ \text{where } \psi(x, \pi) = \pi\phi(x) + \phi_0(x) \text{ and } \pi \text{ is given } (\pi \geq 0). \end{array} \right.$$

Let  $z(\pi)$  be an optimal solution of Problem 3.

Assumption 1:  $z(\pi)$  is continuous in  $\pi$  and finite for all  $\pi \geq 0$ .

Let

$$(1) \quad \theta(\pi) = \pi\phi(z) + \phi_0(z).$$

THEOREM 1: The gradient of  $\theta(\pi)$  in the  $\pi$ -space is equal to  $\phi(z)$ .

PROOF: At the optimum,  $z(\pi)$ , of Problem 3 we have

$$(2) \quad \frac{\partial \psi(x, \pi)}{\partial x} = 0 \quad (\text{Maximum with respect to } x)$$

so, at  $x = z(\pi)$

$$(3) \quad \frac{d\theta(\pi)}{d\pi} = \frac{\partial \psi(x, \pi)}{\partial x} \frac{dz}{d\pi} + \frac{\partial \psi(x, \pi)}{\partial \pi} = \phi(z) .$$

THEOREM 2:  $\theta(\pi) = \psi(x, \pi)$  is a convex function in  $\pi$  .

PROOF: Let  $\lambda'$  and  $\lambda''$  be two scalars such that  $\lambda' + \lambda'' = 1$  , and  $\lambda' , \lambda'' \geq 0$  .  
Let  $\pi' , \pi'' \geq 0$  be two values of  $\pi$  ; and  $z' , z''$  the corresponding optimal values for Problem 3. Then,

$$(4) \quad \begin{aligned} \theta'' - \theta' &= \pi'' \phi'' + \phi''_0 - (\pi' \phi' + \phi'_0) \\ &\geq \pi'' \phi' + \phi'_0 - (\pi' \phi' + \phi'_0) \\ &\quad \text{because } \theta(\pi'') = \text{Max}_x \{ \psi(x, \pi'') \} \\ &= (\pi'' - \pi') \phi' \end{aligned}$$

or 
$$\theta'' - \theta' \geq (\pi'' - \pi') \left[ \frac{d\theta}{d\pi} \right]_{\pi=\pi'} \quad \text{by Theorem 1}$$

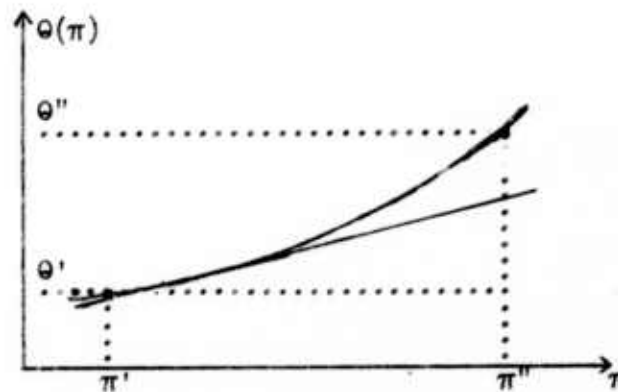


Figure 1

i.e., the variation of  $\theta(\pi)$  is always greater than or equal to that tangent plane, which characterizes a convex function.

Let the following program be:

PROBLEM 2: [Find  $\pi \geq 0$  such that  $\theta(\pi)$  is minimum

Assumption 2: There exists a finite optimal solution  $\hat{\pi}$  to Problem 2.

THEOREM 3: The optimal value of the objective function of Problem 2 is equal to  $\phi_0(\hat{x})$  the optimal value of the objective function of Problem 1.

PROOF: A necessary condition of optimality for Problem 2 is

$$(5) \quad \pi \frac{d\theta}{d\pi} = 0 \quad (\text{Kuhn-Tucker condition})$$

It follows that

$$(6) \quad \hat{\pi} \phi(\hat{\pi}) = 0 ,$$

and

$$(7) \quad \theta(\hat{\pi}) = \phi_0(\hat{x}) .$$

## II. The Algorithm

We replace the initial Problem 1 by Problem 2 of which the optimal solution yields the optimal solution for Problem 1.

Problem 2 is the minimization of a convex function,  $\theta(\pi)$ , where the variable  $\pi$  is subjected to being nonnegative. The expression of the function  $\theta(\pi)$  is not given, but one can compute its value for all  $\pi$ , using the auxiliary program Problem 3. This auxiliary program gives us also the value,  $\phi(z)$  of the gradient of  $\theta(\pi)$  at any given  $\pi$ .

We can solve Problem 2 as follows:

(I) Start with  $\pi = \pi^k \geq 0$ .

Compute  $\theta(\pi)$  and  $\phi[z(\pi^k)]$  using Problem 3.

Determine the vector  $g(\pi^k)$  defined by

$$(8) \quad \begin{cases} g_i = \phi_i[z(\pi^k)] & \text{if } \phi_i \leq 0 \text{ and/or if } \pi_i > 0. \\ g_i = 0 & \text{otherwise (i.e., if } \phi_i > 0 \text{ and } \pi_i = 0). \end{cases}$$

In order to simplify the notation  $g$  will be considered in some cases to be a row vector and in some other cases to be a column vector.

(2) On the half-line  $\pi = \pi^k + \lambda g(\pi^k)$ ,  $\lambda \geq 0$ , determine a new point  $\pi'^k \geq 0$  at a finite distance corresponding to  $\lambda = \lambda'_k \geq 0$  and satisfying the condition

$$(9) \quad 0 < \lambda'_k \leq \text{Min}_i \left\{ \frac{-\pi_i^k}{g_i} \mid g_i < 0 \right\}$$

(3) Compute  $\theta(\pi'^k)$  using Problem 3

If  $\theta(\pi'^k) < \theta(\pi^k)$ , go to 1, replacing  $\pi^k$  by  $\pi^{k+1} = \pi'^k$ .

If  $\theta(\pi'^k) \geq \theta(\pi^k)$  and if  $\pi^k$  is not the optimal solution, we have necessarily

$$(10) \quad \begin{aligned} (\pi'^k - \pi^k) \cdot \phi(\pi^k) &< 0 \\ (\pi'^k - \pi^k) \cdot \phi(\pi'^k) &> 0. \end{aligned}$$

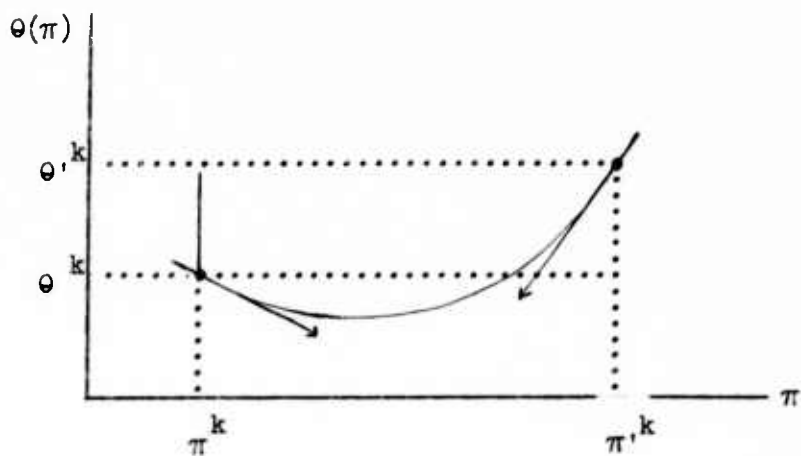


Figure 2

Determine then a new point

$$(11) \quad \pi''^k = \pi^k + (1/2)\lambda'_k g(\pi^k)$$

i.e., the middle point of  $(\pi^k, \pi^{k+1})$  or another intermediate point, taking into account the respective slopes at  $\pi^k$  and  $\pi^{k+1}$ .

Then repeat ③, with  $\pi''^k$  instead of  $\pi^{k+1}$ , and so on until we obtain (in a finite number of steps) a point  $\pi^{k+1}$  such that

$$(12) \quad \theta(\pi^{k+1}) < \theta(\pi^k) .$$

REMARK: To prove the convergence, we have to perform the subdivision of  $\lambda_k$  once more, which yields a point  $\pi^{*k+1}$ , and work with the better of  $\pi^{*k+1}$  and  $\pi^{k+1}$ . In order to simplify the notations, we will call the selected point  $\pi^{k+1}$ , and  $\lambda_k$  the value of  $\lambda$  for that point.

④ Go to ① with  $\pi^{k+1}$  instead of  $\pi^k$ .

### III. Convergence

We can show easily (see ③ of the algorithm), that the variation of  $\Theta(\pi)$ , at each step  $k$ , is strictly negative, except if  $g(\pi^k) = 0$  which implies that  $\phi[z(\pi^k)] \geq 0$  and  $\pi^k \cdot \phi[z(\pi^k)] = 0$ . In these conditions, we achieved the optimum of Problem 2.

Since the minimum of  $\Theta(\pi)$ ,  $\pi \geq 0$  is equal to  $\phi_0(\mathcal{R})$ , optimum of Problem 1 which is assumed to be finite, the sequence of points  $\pi^k$  gives a sequence of convergent values  $\Theta(\pi^k)$  because we have a lower bound for these values. Moreover, we will make the following assumption:

Assumption 3: The second partial derivatives of  $\Theta(\pi)$  are bounded above in all directions by a scalar  $D > 0$ .

One can show that:

$$(13) \quad \Theta(\pi^k) - \Theta(\pi^{k+1}) \geq (1/2) \lambda_k |g^k|^2 \quad \text{if } \lambda_k \leq 1/D,$$

$$(14) \quad \Theta(\pi^k) - \Theta(\pi^{k+1}) \geq 1/4 \left| \frac{g^k}{D} \right|^2 \quad \text{if } \lambda_k \geq 1/D$$

where  $|g|$  = modulus of  $g$ , letting  $g^k = g(\pi^k)$ .

Since  $\lambda_k$  is either determined by  $(g)$  or equal to a fraction of the first value  $\lambda'_k$  that we tried at the beginning of each step which can be taken equal to a constant  $h$ , for all  $k$ .

We obtain for all  $\lambda_k$  such that  $0 \leq \lambda_k \leq h$ .

$$(15) \quad \Theta(\pi^k) - \Theta(\pi^{k+1}) \geq K \cdot \lambda_k |g^k|^2,$$

with  $K = \text{Min} \{1/2, 1/4hD\}$

from which we deduce that the series  $\sum_k \lambda_k |g^k|^2$  converges and that

$$(16) \quad \lambda_k |g^k|^2 \rightarrow 0 \quad \text{as } k \rightarrow \infty$$

For each step  $k$ , by the convexity of  $\theta(\pi)$  we have

$$(17) \quad 0 \leq \theta(\pi^k) - \theta(\hat{\pi}) \leq (\pi^k - \pi) \cdot \phi^k$$

where  $\hat{\pi}$  is an optimal solution of Problem 2.

By definition of  $g^k$ :

$$(18) \quad (\pi^k - \hat{\pi}) \phi^k = -(\pi^k - \hat{\pi}) g^k + (\pi_N^k - \hat{\pi}_N) \cdot \phi_N^k$$

where  $\phi_N$  is a vector containing only the components of  $\phi$  such that  $\phi_i > 0$ ,  $\pi_i = 0$ . Then

$$(19) \quad (\pi_N^k - \hat{\pi}_N) \cdot \phi_N^k = -\hat{\pi}_N \cdot \phi_N^k \leq 0,$$

and finally

$$(20) \quad 0 \leq \theta(\pi^k) - \theta(\hat{\pi}) \leq -(\pi^k - \hat{\pi}) g^k,$$

and in particular

$$(21) \quad (\pi^k - \hat{\pi}) \lambda_k \cdot g^k \leq 0$$

which means that the vector  $(\pi^{k+1} - \pi^k)$  has an acute angle with  $(\hat{\pi} - \pi^k)$

It follows

$$(22) \quad |\hat{\pi} - \pi^{k+1}|^2 \leq |\hat{\pi} - \pi^k|^2 + |\pi^{k+1} - \pi^k|^2 = \\ |\hat{\pi} - \pi^k|^2 + \lambda_k^2 |g^k|^2.$$

Summing with respect to  $k$ :

$$(23) \quad |\hat{\pi} - \pi^k|^2 \leq |\pi - \pi^0|^2 + \sum_{k'=0}^k \lambda_{k'}^2 |g^{k'}|^2$$

Since  $\sum_{k'=0}^k \lambda_{k'} |g^{k'}|^2$  converges and  $\lambda_{k'} \leq h$ , the series

$$\sum_{k'=0}^k \lambda_{k'} |g^{k'}|^2$$

converges when  $k \rightarrow \infty$ , from which we obtain an upper bound for the distance

$$|\pi^k - \bar{\pi}|.$$

We have to consider two cases:

CASE 1:  $g^k \rightarrow 0$ , at least for a finite or infinite subsequence of  $k$ . Then the corresponding subsequence

$$(24) \quad g^k (\pi^k - \bar{\pi}) \rightarrow 0 \quad \text{and} \quad \theta(\pi^k) \rightarrow \theta(\bar{\pi})$$

For these conditions, the algorithm converges to an optimal solution.

CASE 2:  $g^k$  does not converge to zero for any subsequence of  $k$ . This means  $g^k$  is different from zero for all  $k$ , and moreover  $|g^k|$  has a lower bound  $> 0$ . For these conditions, the convergence of the

series  $\sum \lambda_k |g^k|^2$  implies the convergence of the series

$\sum \lambda_k |g^k|$  and  $\sum \lambda_k$  and therefore implies the convergence of the point  $\pi^k$  which was defined by

$$(25) \quad \pi^k = \sum_{k'=0}^k \lambda_{k'} g^{k'}$$

Let  $\bar{\pi}$  be the limit of  $\pi^k$ , and let  $\bar{g}$  be the limiting value of  $g^k$  at that point. (This limit exists by Assumption 1.)

The convergence of the series  $\sum \lambda^k$  implies that  $\lambda_k \rightarrow 0$  and by Assumption 3,  $\lambda_k$  is determined by the relation (9) for  $K$  sufficiently large. There exists at least one index  $i$  and two subsequences  $K'$  and  $K''$

corresponding to the steps  $k$  such that,

$$(26) \quad k' \in K'' \quad \begin{cases} g_1^{k'} < 0 & ; & \phi_1^{k'} > 0 \\ \pi_1^{k'} > 0 & ; \\ \lambda_{k'} = \frac{\pi_1}{-g_1} \end{cases}$$

and

$$(27) \quad k'' \in K'' \quad \begin{cases} g_1^{k''} > 0 & ; & \phi_1^{k''} < 0 \\ \pi_1^{k''} = 0 \end{cases}$$

So  $\phi_1^{k'}$  and  $\phi_1^{k''}$  go to zero, and we have by the fact that the second order derivatives of  $\phi$  have an upper bound,  $D$ ,

$$(28) \quad -g_1^{k'} \leq D \cdot \pi_1^{k'}$$

If  $j$  is the index of a nonzero component of  $\tilde{g}$ , we have for  $k'$  sufficiently large,

$$(29) \quad \lambda_{k'} |g_j^{k'}| = -\frac{\pi_1^{k'}}{g_1^{k'}} |g_j^{k'}| \geq \frac{1}{D} |g_j^{k'}| \cong \frac{1}{D} |\tilde{g}_j| > 0$$

We have here a contradiction with the fact that  $\pi^k$  converges to  $\tilde{\pi}$  which implies that  $\lambda_{k'} g_j^{k'} \rightarrow 0$ . Thus the components of  $\tilde{g}$  must be zero, which contradicts the assumption of a possible second case.

#### IV. REMARKS

1) Nature of the Algorithm: Our initial program is

Problem 1  $\left[ \begin{array}{l} \text{Find } x \text{ such that } \phi_0(x) \text{ is maximum} \\ \text{subject to } \phi_1(x) \geq 0 \end{array} \right.$

Let us consider the dual [2] of Problem 1:

$$(D) \begin{cases} \text{Find } \pi \geq 0 \text{ such that } \psi(x, \pi) \text{ is minimum} \\ \text{subject to } \frac{\partial}{\partial x_j} \psi(x, \pi) = 0 \quad j = 1, 2, \dots, n. \end{cases}$$

We see that solving the auxiliary program Problem 3 for a given nonnegative  $\pi$ , is equivalent to finding a feasible solution  $(x, \pi)$  for the dual program (D). Problem 2, which minimizes  $\theta(\pi)$  solves (D) by a discrete gradient method.

2) On Assumption 1: Assumption 1 seems at first very restrictive, nevertheless, if the optimum of Problem 1 is finite, one can always remove the constraints  $\phi_1(x)$  and the functional  $\phi_0(x)$  far enough from the origin in order to have  $z(\pi)$  at a finite distance. For instance, this modification can consist in joining the surfaces  $y_1 = \phi_1(x)$  and  $y_0 = \phi_0(x)$  to the vertical cylinder  $|x| = \text{constant}$  if the constant is chosen large enough. This would reduce the search for an optimum of Problem 3 to the interior of a sphere,  $(x) = \text{constant}$ . One could imagine other compact spaces.

The continuity of  $z(\pi)$  for  $\pi$  can be achieved, e.g., by assuming the optimum of Problem 3 is always unique. This is the case if  $\psi(x, \pi)$  is strictly concave in  $x$ , i.e., if one at least of the functions  $\phi_1(x)$  or  $\phi_0(x)$  is strictly concave.

## REFERENCES

1. Dantzig, G. B., Linear Programming and Extensions, Chapter 24, Princeton University Press, Princeton, New Jersey, 1963.
2. Huard, P., "Dual Programs," Proceedings, Fifth Symposium on Mathematical Programming, Chicago, Illinois, June 1962.

BASIC DISTRIBUTION LIST  
FOR UNCLASSIFIED TECHNICAL REPORTS

Head, Logistics and Mathematical  
Statistics Branch  
Office of Naval Research  
Washington 25, D. C.

C. O., ONR Office  
Navy No. 100, Box 39, F. P. O.  
New York, New York

ASTIA Document Service Center  
Building No. 5  
Cameron Station  
Alexandra, Virginia

Institute for Defense Analyses  
Communications Research Division  
von Neumann Hall  
Princeton, New Jersey

Technical Information Officer  
Naval Research Laboratory  
Washington 25, D. C.

C. O., ONR Branch Office  
1030 East Green Street  
Pasadena 1, California  
ATTN: Dr. A. R. Laufer

Bureau of Supplies and Accounts  
Code OW, Department of the Navy  
Washington 25, D. C.

Professor Russell Ackoff  
Operations Research Group  
Case Institute of Technology  
Cleveland 6, Ohio

Professor Kenneth J. Arrow  
Serra House, Stanford University  
Stanford, California

Professor G. L. Bach  
Carnegie Institute of Technology  
Planning and Control of Industrial  
Operations, Schenley Park  
Pittsburgh, 13, Pennsylvania

Professor L. W. Cohen  
Mathematics Department  
University of Maryland  
College Park, Maryland

Professor Donald Eckman  
Director, Systems Research Center  
Case Institute of Technology  
Cleveland, Ohio

Professor Lawrence E. Fouraker  
Graduate School of Business  
Harvard University  
Cambridge, Massachusetts

Professor David Gale  
Department of Mathematics  
Brown University  
Providence 12, Rhode Island

Dr. Murray Geisler  
The RAND Corporation  
1700 Main Street  
Santa Monica, California

Professor L. Hurwicz  
School of Business Administration  
University of Minnesota  
Minneapolis 14, Minnesota

Dr. James R. Jackson, Director  
Western Management Sciences Institute  
University of California  
Los Angeles 24, California

Professor Samuel Karlin  
Mathematics Department  
Stanford University  
Stanford, California

Professor C. E. Lemke  
Department of Mathematics  
Rensselaer Polytechnic Institute  
Troy, New York

Professor W. H. Marlow  
Logistics Research Project  
The George Washington University  
707 - 22nd Street, N. W.  
Washington 7, C. D.

BASIC DISTRIBUTION LIST  
FOR UNCLASSIFIED TECHNICAL REPORTS

Professor A. Charnes  
The Technological Institute  
Northwestern University  
Evanston, Illinois

Professor R. Radner  
Department of Economics  
University of California  
Berkeley 4, California

Professor Murray Rosenblatt  
Department of Mathematics  
Brown University  
Providence 12, Rhode Island

Mr. J. R. Simpson  
Bureau of Supplies and Accounts  
Navy Department (Code W31)  
Washington 25, D. C.

Professor A. W. Tucker  
Department of Mathematics  
Princeton University  
Princeton, New Jersey

Professor J. Wolfowitz  
Department of Mathematics  
Lincoln Hall, Cornell University  
Ithaca 1, New York

C. O., ONR Branch Office  
346 Broadway  
New York 13, New York  
ATTN: J. Laderman

Professor Oskar Morgenstern  
Economics Research Project  
Princeton University  
92 A Nassau Street  
Princeton, New Jersey