Mathematical Programming 34 (1986) 235-238
North-Holland

A NONLINEAR EQUATION FOR LINEAR PROGRAMMING

P.W. SMITH
*Department of Mathematical Sciences, Old Dominion University, Norfolk, VA, USA**

H. WOLKOWICZ**
Department of Mathematics and Computer Science, Emory University, Atlanta, GA 30322, USA

Received 8 April 1985
Revised manuscript received 19 July 1985

We present a characterization of the 'normal' optimal solution of the linear program given in canonical form

$$\max\{c^t x\colon Ax = b, x \geqslant 0\}. \tag{P}$$

We show that x^* is the optimal solution of (P), of minimal norm, if and only if there exists an $R > 0$ such that, for each $r \geqslant R$, we have

$$x^* = (rc - A^t\lambda_r)_+.$$

Thus, we can find x^* by solving the following equation for λ_r

$$A(rc - A^t\lambda_r)_+ = b.$$

Moreover, $(1/r)\lambda_r$ then 'converges' to a solution of the dual program.

Key words: Linear Programming, Characterization of Optimality, Dual Program.

1. Introduction

We consider the linear programming problem in canonical form

$$p = \max\{c^t x\colon Ax = b, x \geqslant 0\} \tag{P}$$

where $\cdot^t$ denotes transpose and A is an $m \times n$ matrix. We assume that p is finite. The dual program to (P) is

$$d = \min\{b^t\lambda\colon A^t\lambda \geqslant c\}. \tag{D}$$

We present a characterization of a solution, x^*, of (P) in terms of a nonlinear equation, see (1.1) below.

In Section 2 we present the characterization of the normal optimal solution, x^*, of (P) as the solution of a nonlinear equation, i.e. for $r \geqslant R$, for some fixed $R > 0$,

$$A(rc - A^t\lambda_r)_+ = b \tag{1.1}$$

* On leave from The University of Alberta, Edmonton, Canada. Research partially supported by the National Science and Engineering Research Council of Canada.

** Presently at IMSL, 2500 Park West Tower One, 2500 City West Boulevard, Houston, TX 77042-3020, USA.

and

$$x^* = (rc - A^t\lambda_r)_+, \tag{1.2}$$

where the vector y_+ denotes the projection of y onto the nonnegative orthant, i.e. $(y_+)_i = \max(0, y_i)$. Moreover, if $(1/r)\lambda_r \to \lambda^*$, as $r \to \infty$, then λ^* solves the dual program, (D); while, if λ^* solves (D), then there exists a sequence λ_r, of solutions of (1.2), such that $(1/r)\lambda_r \to \lambda^*$. These results are presented in Theorem 2.1.

The characterization (1.1) follows from the characterization of (P) as a nearest point problem given in [4]. We then apply the approach in [5] to obtain the explicit equation for this characterization.

2. The nonlinear equation

We now present the main result, the characterization of the normal optimal solution of (P) as the solution of a nonlinear equation.

Theorem 2.1. (i) *The point x^* is the optimum of* (P), *of minimum norm, if and only if there exists $R > 0$ such that, for each $r \geq R$, the system*

$$A(rc - A^t\lambda_r)_+ = b \tag{2.1}$$

is solveable for λ_r and

$$x^* = (rc - A^t\lambda_r)_+. \tag{2.2}$$

(ii) *If the solutions of* (2.1) *satisfy*

$$\frac{1}{r}\lambda_r \to \lambda^* \quad \text{as } r \to \infty, \tag{2.3}$$

then λ^* is optimal for (D).

(iii) *For each optimal solution λ^* of* (D), *there exists a sequence of solutions, λ_r, of* (2.1), *such that* (2.3) *holds.*

To prove Theorem 2.1 we combine a result from [5] and one from [4], which we now present. The result in [4] characterizes the normal solution of (P) as the solution of a quadratic program. (Here $\|\cdot\|$ denotes Euclidean norm.)

Theorem 2.2. *The point x^* is the solution of* (P), *of minimum norm, if and only if there exists $R > 0$ such that, for $r \geq R$, we have x^* feasible and*

$$\|rc - x^*\| = \min\{\|rc - x\|: Ax = b, x \geq 0\}. \tag{2.4}$$

Proof. See [4, Theorem 2.1]. □

The result in [5] characterizes the solution of the problem

$$\min\{\|x\|: Ax = b, x \geq 0\}, \tag{2.5}$$

as the solution of (2.1) and (2.2) with $c = 0$. We now prove Theorem 2.1 by obtaining the characterization for the problem (2.4).

Proof of Theorem 2.1. Let $R>0$ be found using Theorem 2.2 and let $r \geqslant R$. Now let us solve (2.4), after replacing the norm, $\|\cdot\|$, by its square and dividing by 2. The Karush–Kuhn–Tucker conditions, e.g. [2], yield the system

$$x^* = rc - A^t\lambda + s, \quad s \geqslant 0, \quad s^t x^* = 0. \tag{2.6}$$

If $(rc - A^t\lambda)_i > 0$, then $x_i^* > 0$ and $s_i = 0$, i.e. $x_i^* = (rc - A^t\lambda)_i$. If $(rc - A^t\lambda)_i < 0$, then $s_i > 0$ and $x_i^* = 0$. If $(rc - A^t\lambda)_i = 0$, then $s_i = 0$ and $x_i^* = 0$. Thus, we have shown (2.2). Since x^* must necessarily be feasible, we see that (2.1) must hold. This proves the characterization (2.1) and (2.2) in (i).

Now, if x^* solves (2.4), the optimality conditions (2.6) yield

$$c^t x^* - \frac{1}{r}\|x^*\|^2 - b^t \frac{1}{r}\lambda = 0, \quad Ax^* = b, \quad x^* \geqslant 0,$$

$$c - \frac{1}{r}x^* - A^t\left(\frac{1}{r}\lambda\right) + \frac{1}{r}s = 0, \quad s \geqslant 0, \quad s^t x^* = 0.$$

Letting $r \to \infty$, $(1/r)\lambda \to \lambda^*$, shows that x^* and λ^* are feasible for (P) and (D), respectively, with the optimal values $c^t x^* = b^t \lambda^*$, i.e. they are both optimal. This proves (ii).

Now suppose that λ^* solves the dual (D) and x^* is the minimum norm solution of (P). Then the optimality conditions for (P) yield

$$A^t(r\lambda^*) - rc = r\bar{s} \geqslant 0, \quad \bar{s}^t x^* = 0. \tag{2.7}$$

Moreover, x^* solves the program

$$\min\{z^t x^*: Az = b, c^t z \geqslant p, z \geqslant 0\}$$

and so

$$x^* + A^t\lambda + \beta c = s \geqslant 0, \quad s^t x^* = 0, \quad \beta \leqslant 0. \tag{2.8}$$

Upon adding (2.7) and (2.8) we get

$$x^* + A^t(r\lambda^* + \lambda) - (r - \beta)c = r\bar{s} + s \geqslant 0, \quad (r\bar{s} + s)^t x^* = 0. \tag{2.9}$$

This shows that $\lambda_r = (r + \beta)\lambda^* + \lambda$ solves (2.6) and so also (2.1). Since (2.3) holds, we have shown (iii). □

Remarks. The optimality conditions for (P) (e.g. (2.7) along with feasibility of x^*) yield $n+m+1$ equations with $2n$ inequalities. These characterize the optimum solutions of (P).

The solution of only m equations

$$A(rc - A^t\lambda_r)_+ = b \tag{2.1}$$

for λ_r provides the solution

$$x^* = (rc - A^t\lambda_r)_+ \tag{2.2}$$

for the linear program (P). The equation (2.1) is nonlinear and nondifferentiable due to the plus. Also one needs to find an estimate for R in order to find $r \geqslant R$ in (2.1).

At the points where (2.1) is differentiable, the Jacobian is

$$J(\lambda) = -AD_\lambda A^t,$$

where D_λ is a diagonal, zero-one matrix with diagonal elements

$$d_{ii} = \begin{cases} 0 & \text{if } rc_i < \sum_{j=1}^{m} a_{ji}\lambda_j, \\ 1 & \text{otherwise.} \end{cases}$$

Note that the points of nondifferentiability occur exactly when $rc_i = \sum_{j=1}^m a_{ji}\lambda_j$. One can now try Newton type methods to solve (2.1).

Consider the abstract linear program (P) where the constraint $x \geqslant 0$ is replaced by $x \geqslant_S 0$, the operator $A: X \to Y$, X and Y are normed spaces, S is a convex cone, and $x \geqslant_S y$ if and only if $x - y \in S$. A duality theory for (P) is given in [1]. The characterization (2.1) of (P) can be extended to include these abstract linear programs.

The above ideas will be presented in a forthcoming study.

Acknowledgement

We are grateful to two referees for their constructive remarks. In particular, we would like to thank one referee for improving our proof of Theorem 2.1 (ii).

References

[1] J. Borwein and H. Wolkowicz, "A simple constraint qualification in infinite dimensional programming", Research Report 14, Emory University (Atlanta, GA, 1984).

[2] O.L. Mangasarian, *Nonlinear programming* (McGraw-Hill, New York, 1969).

[3] O.L. Mangasarian, "Uniqueness of solution in linear programming", *Linear Algebra and its Applications* 25 (1979) 151-162.

[4] O.L. Mangasarian, "Normal solutions of linear programs", *Mathematical Programming Study* 22 (1984) 206-216.

[5] C.A. Micchelli, P.W. Smith, J. Swetits and J.D. Ward, "Constrained L_p approximation", *Journal of Constructive Approximation* 1 (1985) 93-102.