AN EXACT SOLUTION TO LINEAR PROGRAMMING USING AN INTERIOR POINT METHOD*

WEI ZI-LUAN (魏紫銮)
(*Computing Center, Academia Sinica, Beijing, China*)

Abstract

This paper presents sufficient conditions for optimality of the Linear Programming (LP) problem in the neighborhood of an optimal solution, and applies them to an interior point method for solving the LP problem. We show that after a finite number of iterations, an exact solution to the LP problem is obtained by solving a linear system of equations under the assumptions that the primal and dual problems are both nondegenerate, and that the minimum value is bounded. If necessary, the dual solution can also be found.

§ 1. Introduction

As is well-known, an optimal solution to the linear programming (LP) problem is obtained at an extreme point of the feasible region under the assumptions that the primal and dual problems are both nondegenerate, and that the minimum value of the LP problem is bounded. Hence, as long as nonbasic variables can be distinguished in the neighborhood of an optimal solution, an exact solution can be found by solving a linear system of equations. This idea will be exploited by use of an interior point method.

This paper presents sufficient conditions for optimality of the LP problem and applies them to an interior point method to obtain an exact solution. First of all, a second order estimate of the dual solutions and Lagrange multipliers in the neighborhood of an optimal solution are given for a standard form of the LP problem under the following assumptions: the primal and dual problems are both nondegenerate, and the minimum value of the LP problem is bounded. Then a vector of quasi-Lagrange multipliers (QLM) is introduced in order to set up the formula to estimate the value of Lagrange multipliers. The nonbasic variables will be discarded based on the estimated value of the Lagrange multipliers, and an exact solution of the LP problem will be achieved by solving a linear systems of equations. The dual solutions can also be obtained at the same time if necessary.

Section 2 shows how to set up the formula of the second-order estimate of the dual solution and Lagrange multiplier vector in the neighborhood of an optimal solution. Section 3 describes the proposed interior point method for solving the LP problem in terms of an affine transformation, and discusses the method's convergence. Section 4 applies the results described in Sections 2 and 3 to the interior point method, and shows that an exact optimal solution is obtained in a finite number of iterations.

* Received February 6, 1986.

§ 2. Optimality Conditions

This section presents the sufficient conditions for optimality of the LP problem in the neighborhood of an optimal solution. A second order estimation formula of the dual solution and Lagrange multipliers is given under certain assumptions that a strictly interior feasible point in the neighborhood of an optimal solution.

First of all, we consider the following standard form of a linear program (LP):

$$\text{Minimize } z = c^T x, \tag{2.1}$$

$$\text{Subject to } Ax = b, \tag{2.2}$$

$$x \geqslant 0, \tag{2.3}$$

where A is an $m\times n$ real matrix with rank m, $m<n$, b and c are real vectors in R^m and R^n, respectively, and x is a real variable in R^n.

Suppose that the LP problem is feasible and nondegenerate. Assume that $\bar{x}$ is a strictly interior feasible point satisfying (2.2)—(2.3). Then, it can be expressed as

$$\bar{x} = x_v + \bar{x} - x_v = x_v + u, \tag{2.4}$$

where x_v is a basic feasible solution of the LP problem, and

$$x_v = \begin{bmatrix} x_B \\ 0 \end{bmatrix}, \tag{2.5}$$

$$u = \begin{bmatrix} u_B \\ u_N \end{bmatrix} = \bar{x} - x_v. \tag{2.6}$$

Let $N(x, \delta)$ denote the Euclidean ball about x of radius δ, and D, D_1, D_2, D_3 denote the $n\times n$, $m\times m$, $m\times m$, and $(n-m)\times(n-m)$ diagonal matrices, containing the components of $\bar{x}$, x_B, u_B, u_N, respectively. Then

$$D = \begin{bmatrix} D_1 + D_2 & 0 \\ 0 & D_3 \end{bmatrix}. \tag{2.7}$$

In order to describe the main results in this section, we define vectors $\bar{y}$, c_p, and $\bar{v}$, by

$$\bar{y} = (AD^2A^T)^{-1}AD^2c, \tag{2.8}$$

$$c_p = Dc - DA^T\bar{y}, \tag{2.9}$$

$$\bar{v} = D^{-1}c_p. \tag{2.10}$$

The vector $\bar{v}$ contains the quasi-Lagrange multipliers.

Theorem 2.1. *Suppose that the* LP *problem is feasible and nondegenerate. Assume that x^* is its unique optimal solution, and y^* is the corresponding dual optimal solution. If $\bar{x}$ is a strictly interior feasible point, such that $\bar{x}\in N(x^*, \delta)$, then $\bar{y}$ is a second-order estimate of the dual solution y^*, where δ is a sufficiently small positive scalar.*

Theorem 2.2. *Suppose that the* LP *problem is feasible and nondegenerate. Assume that x^* is its unique optimal solution, and v^* is the vector of Lagrange multipliers with respect to x^*. If $\bar{x}$ is a strictly interior feasible point, such that $\bar{x}\in N(x^*, \delta)$, then the vector $\bar{v}$ is a second order estimate of v^*, where δ is a sufficiently small positive scalar.*

To prove these theorems, we introduce several lemmas.

Lemma 2.3. *Suppose that the* LP *problem is feasible and nondegenerate. Assume that $\bar{x}$ is a strictly interior feasible point satisfying* (2.2)—(2.3). *Then*

$$AD^2A^T = BD_1[I+2D_1^{-1}D_2+D_1^{-2}D_2^2+(BD_1)^{-1}ND_3^2N^T(D_1B^T)^{-1}](BD_1)^T, \quad (2.11)$$

where D, D_1, D_2 D_3 are diagonal matrices as described above, and B is a basic matrix corresponding to a basic solution.

Proof. From the definitions, it follows that

$$AD = (B,\ N)\begin{bmatrix} D_1+D_2 & 0 \\ 0 & D_3 \end{bmatrix}, \quad (2.12)$$

where N is a $(n-m)\times m$ matrix with respect to the nonbasic variable u_N. Hence

$$AD^2A^T = B(D_1+D_2)^2B^T+ND_3^2N^T = B(D_1^2+2D_1D_2+D_2^2)B^T+ND_3^2N^T.$$

From the assumption that the LP problem is nondegenerate, B, D and D_1 are invertible matrices, and consequently

$$AD^2A^T = BD_1[I+2D_1^{-1}D_2+D_1^{-2}D_2^2+(BD_1)^{-1}ND_3^2N^T(D_1B^T)^{-1}](BD_1)^T.$$

This completes the proof of the lemma.

Lemma 2.4. *Using the same assumptions as in Lemma 2.3, it follows that*

$$(BD_1)^{-1}AD^2 = (D_1+2D_2+D_1^{-1}D_2^2,\ D_1^{-1}B^{-1}ND_3^2). \quad (2.13)$$

Proof. From (2.12) and the definitions, the above is straightforward:

$$(BD_1)^{-1}AD = (BD_1)^{-1}[B(D_1+D_2),\ ND_3] = (I+D_1^{-1}D_2,\ D_1^{-1}B^{-1}ND_3), \quad (2.14)$$

and hence

$$(BD_1)^{-1}AD^2 = (D_1+2D_2+D_1^{-1}D_2^2,\ D_1^{-1}B^{-1}ND_3^2).$$

This proves the lemma.

Lemma 2.5. *Let $\|\cdot\|$ denote any matrix norm for which $\|I\|=1$. If $\|E\|<1$, then $(I+E)^{-1}$ exists, and*

$$(I+E)^{-1} = I-E+E^2-\cdots.$$

The proof of the lemma is given in [4].

Lemma 2.6. *Suppose that the LP problem is feasible and nondegenerate. If x_B is a basic feasible solution, δ is a sufficiently small positive scalar, and $\bar{x}$ is a strictly interior point, such that $\bar{x}\in N(x_B,\ \delta)$, then*

$$[I+2D_1^{-1}D_2+D_1^{-2}D_2^2+(BD_1)^{-1}ND_3^2N^T(D_1B^T)^{-1}]^{-1}$$
$$\approx I-2D_1^{-1}D_2+3D_1^{-2}D_2^2-(BD_1)^{-1}ND_3^2N^T(D_1B^T)^{-1}, \quad (2.15)$$

where matrices B, D_1, D_2, D_3 are as described above.

Proof. The assumptions of nondegeneracy and feasibility imply that B and D_1 are invertible matrices. Thus, the left side of (2.15) is well defined. Let

$$E = 2D_1^{-1}D_2+D_1^{-2}D_2^2+(BD_1)^{-1}ND_3^2N^T(D_1B^T)^{-1}. \quad (2.16)$$

Now we can choose a sufficiently small $\delta>0$, such that for any $\bar{x}\in N(x_B,\ \delta)$, the spectral radius of the matrix E is less than 1. It follows from Lemma 2.5 that

$$(I+E)^{-1} = I-E+E^2-\cdots. \quad (2.17)$$

Substituting (2.16) into (2.17) yields

$$[I+2D_1^{-1}D_2+D_1^{-2}D_2^2+(BD_1)^{-1}ND_3^2N^T(D_1B^T)^{-1}]^{-1}$$
$$= I-2D_1^{-1}D_2-D_1^{-2}D_2^2-(BD_1)^{-1}ND_3^2N^T(D_1B^T)^{-1}+4D_1^{-2}D_2^2+O(\delta^3). \quad (2.18)$$

If $O(\delta^3)$ is ignored, then (2.18) can be written in the form

$$[I+2D_1^{-1}D_2+D_1^{-2}D_2^2+(BD_1)^{-1}ND_3^2N^T(D_1B^T)^{-1}]^{-1}$$
$$\approx I-2D_1^{-1}D_2+3D_1^{-2}D_2^2-(BD_1)^{-1}ND_3^2N^T(D_1B^T)^{-1},$$

which proves the lemma.

Now we prove Theorems 2.1 and 2.2.

The proof of Theorem 2.1 is as follows: Recall that $\bar{y}$ is defined by the following:

$$\bar{y}=(AD^2A^T)^{-1}AD^2c. \tag{2.19}$$

By the result of Lemma 2.3, we have that

$$AD^2A^T=BD_1[I+2D_1^{-1}D_2+D_1^{-2}D_2^2+(BD_1)^{-1}ND_3^2N^T(D_1B^T)^{-1}](BD_1)^T,$$

where now B is an $m\times m$ matrix corresponding to the optimal solution, and D_1, D_2, D_3 are as described above. Hence

$$(AD^2A^T)^{-1}=(D_1B^T)^{-1}[I+2D_1^{-1}D_2+D_1^{-1}D_2+(BD_1)^{-1}ND_3^2N^T(D_1B^T)^{-1}]^{-1}(BD_1)^{-1}. \tag{2.20}$$

Substituting (2.20) into (2.19) gives

$$\bar{y}=(D_1B^T)^{-1}[I+2D_1^{-1}D_2+D_1^{-2}D_2^2+(BD_1)^{-1}ND_3^2N^T(D_1B^T)^{-1}]^{-1}(BD_1)^{-1}AD^2c. \tag{2.21}$$

It follows from (2.13) that (2.21) can be rewritten in the form

$$\bar{y}=(D_1B^T)^{-1}[I+2D_1^{-1}D_2+D_1^{-2}D_2^2$$
$$+(BD_1)^{-1}ND_3^2N^T(D_1B^T)^{-1}]^{-1}(D_1+2D_2+D_1^{-1}D_2^2,\ D_1^{-1}B^{-1}ND_3^2)c. \tag{2.22}$$

From the assumption that δ is a sufficiently small positive scalar, such that $\bar{x}\in N(x^*,\delta)$, the result of Lemma 2.6 holds. So substituting (2.15) into (2.22) gives

$$\bar{y}=(D_1B^T)^{-1}[D_1-(BD_1)^{-1}ND_3^2N^T(D^T)^{-1}-O(\delta^3),\ D_1^{-1}B^{-1}ND_3^2+O(\delta^3)]c. \tag{2.23}$$

By definition, $c^T=(c_B^T, c_N^T)$. Thus, it is straightforward to establish that

$$\bar{y}=(D_1B^T)^{-1}[D_1c_B-(BD_1)^{-1}ND_3^2N^T(B^T)^{-1}c_B+D_1^{-1}B^{-1}ND_3^2c_N+O(\delta^3)]$$
$$=(B^T)^{-1}c_B-D_1^{-2}B^{-1}ND_3^2[(c_B^TB^{-1}N)^T-c_N]+O(\delta^3). \tag{2.24}$$

If the $O(\delta^3)$ term in (2.24) is ignored, then (2.24) can be written in the form

$$y\approx(c_B^TB^{-1})^T+(B^T)^{-1}D_1^{-2}B^{-1}ND_3^2[c_N-c(c_B^TB^{-1}N)^T]\approx y^*+(B^T)^{-1}D_1^{-2}B^{-1}ND_3^2\bar{c}_N, \tag{2.25}$$

where $\bar{c}_N=c_N-(c_B^TB^{-1}N)^T$, $y^*=(c_B^TB^{-1})^T$. Thus, $\bar{y}$ is in fact a second-order estimate of the dual solution y^*.

This completes the proof of Theorem 2.1.

The proof of Theorem 2.2 is as follows: From definition (2.9)

$$c_p=Dc-DA^T\bar{y}. \tag{2.26}$$

By the assumption of nondegeneracy and the fact that $\bar{x}$ is a strictly interior feasible point, it follows from definition (2.10) that

$$\bar{v}=D^{-1}c_p=c-A^T\bar{y}. \tag{2.27}$$

Since δ is a sufficiently small positive scalar, and $\bar{x}\in N(x^*,\delta)$, the result of Theorem 2.1 holds. Thus

$$\bar{y}\approx(c_B^TB^{-1})^T+(B^T)^{-1}D_1^{-2}B^{-1}ND_3^2\bar{c}_N. \tag{2.28}$$

There is no loss of generality in writing $\bar{v}$ in the form

$$\bar{v}=\begin{bmatrix} c_B \\ c_N \end{bmatrix}-\begin{bmatrix} B^T \\ N^T \end{bmatrix}\bar{y}. \tag{2.29}$$

Substituting (2.28) into (2.29), we obtain

$$\bar{v}\approx\begin{bmatrix} c_B \\ c_N \end{bmatrix}-\begin{bmatrix} I \\ (B^{-1}N)^T \end{bmatrix}\cdot(c_B+D_1^{-2}B^{-1}ND_3^2\bar{c}_N)\approx\begin{bmatrix} 0 \\ \bar{c}_N \end{bmatrix}-\begin{bmatrix} D_1^{-2}B^{-1}ND_3^2\bar{c}_N \\ (B^{-1}N)^TD_1^{-2}B^{-1}ND_3^2\bar{c}_N \end{bmatrix}.$$

Hence

$$\bar{v}\approx v^*+O(\delta^2).$$

Thus, $\bar{v}$ is a second order estimate of the vector of Lagrange multipliers v^*.

This completes the proof of Theorem 2.2.

Using analysis similar to that in Theorems 2.1 and 2.2, it is straightforward to show that the conclusions of these two theorems can be generalized to the case where $\bar{x}$ is in the neighborhood of the optimal solution, but is not a feasible point.

§ 3. Algorithm and Convergence

This section describes an interior point method for solving the LP problem, and its convergence.

Suppose that $\bar{x}$ is a strictly interior feasible point. Then an affine transformation and its inverse can be defined as follows:

$$x'=D^{-1}x, \tag{3.1}$$

$$x=Dx'. \tag{3.2}$$

The LP problem is transformed into the following linear programming problem in x'-space

$$\text{Minimize } z=c^TDx', \tag{3.3}$$

$$\text{Subject to } ADx'=b, \tag{3.4}$$

$$x'\geqslant 0, \tag{3.5}$$

where D is a diagonal matrix containing the components of $\bar{x}$.

Obviously, from the definition of D, the point $\bar{x}$ in x-space is mapped into the point $e^T=(1, 1, \cdots, 1)$ in x'-space. So, it is easy to take a large step away from the point e to a new point in x'-space. The new point is transformed back into the x-space, and an iterative point is obtained. To guarantee that the iterative point is well defined, it is necessary to have an additional assumption that the LP problem satisfies the condition F.

Definition. *The* LP *problem satisfies the condition F if it is feasible, and for any feasible point* $\bar{x}$, *the matrix* AD^2A^T *is of full rank, where D is a diagonal matrix containing components of* $\bar{x}$.

From the definition, it is easy to see that if the feasible region S of the LP problem is bounded, then the condition F is equivalent to nondegeneracy. If, on the other hand, S is unbounded, the condition F implies that for any ray direction v of S the matrix $AD_v^2A^T$ is of full rank, where D_v is a diagonal matrix containing components of v.

As described above, algorithm A creates a sequence of points $x^{(0)}$, $x^{(1)}$ $\cdots$ as follows.

Algorithm A.

$k:=0$. Given $x^{(0)}$, an initial starting point.

(1) Define

$$D_k=\operatorname{diag}(x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)}), \tag{3.6}$$

$$A_k=AD_k. \tag{3.7}$$

(2) Compute the vector $c_p^{(k)}$ and its norm $\|c_p^{(k)}\|_2$ by

$$c_p^{(k)}=[I-A_k^T(A_kA_k^T)^{-1}A_k]D_kc \tag{3.8}$$

and

$$\|c_p^{(k)}\|_2=\sqrt{(c_p^{(k)})^T(c_p^{(k)})}. \tag{3.9}$$

Go to the next step.

(3) Normalize $c_p^{(k)}$

$$p^{(k)}=\frac{c_p^{(k)}}{\|c_p^{(k)}\|_2}. \tag{3.10}$$

Go to the next step.

(4) Determine the largest step which can be taken, and generate a new point:

$$\lambda_k=\frac{1}{q_k}, \tag{3.11}$$

where

$$q_k=\operatorname{Max}[p_i^{(k)}], \tag{3.12}$$

and let

$$x^{(k+1)}=x^{(k)}-\alpha\lambda_kD_kp^{(k)}, \tag{3.13}$$

where $\alpha\in(0, 1)$. k: $k+1$, and go to step (1).

It is clear that the bulk of the computational work at each iteration is from step (2), which ensures the feasibility of the new point.

Concerning the convergence of algorithm A, we have the following results:

Theorem 3.1. *Suppose that the* LP *problem satisfies the condition* F. *If the minimum value of the* LP *problem is bounded, then the sequence* $x^{(k)}$ *generated by algorithm* A *converges to an optimal solution of the* LP *problem.*

Theorem 3.2. *Suppose that the* LP *problem satisfies the condition* F. *If the minimum value of the* LP *problem is unbounded, then there exists a vector* $p\leqslant0$, *such that the sequence* $p^{(k)}$ *defined by* (3.10) *converges to* p.

The proofs of Theorems 3.1 and 3.2 are given in [7].

§ 4. An Exact Solution of the LP Problem

This section applies the results of Sections 2 and 3 to a modified algorithm A in order to obtain an exact solution of the LP problem. To show this, the results related to Sections 2 and 3 are restated as follows.

Theorem 4.1. *Suppose that the* LP *problem satisfies the condition* F. *Assume that* x^* *is its unique optimal solution, and that the sequence* $x^{(k)}$ *generated by algorithm* A *converges to* x^*. *Then*

$$\lim_{k\to\infty}\bar{y}^{(k)}=y^*,$$

where y^* is the dual optimal solution, and $\bar{y}^{(k)}$ is defined by (2.8) *with respect to D_k.*

It is straightforward to prove this result from Theorems 2.1 and 3.1.

Theorem 4.2. *With the same assumptions as in Theorem* 4.1,

$$\lim_{k\to\infty} \bar{v}^{(k)} = v,$$

where v^* is a vector of Lagrange multipliers with respect to x^*, and $\bar{v}^{(k)}$ is defined by (2.10) *with respect to D_k.*

The assertion of the theorem is an immediate consequence of Theorems 2.2 and 3.1.

Let $\bar{v}_i^{(k)}$ and $\bar{v}_j^{(k)}$ denote the vectors within the QLM of basic and nonbasic variables, respectively. Then a corollary follows from Theorem 4.2.

Corollary 4.3. With the same assumptions as in Theorem 4.2, there exists a sufficiently small $s>0$, and an integer $k_0>0$, such that for all $k>k_0$

$$\text{Max}\,|\bar{v}_i^{(k)}| \leqslant s \tag{4.1}$$

and

$$\text{Min}\,\bar{v}_j^{(k)} > s. \tag{4.2}$$

This conclusion shows a very important fact: the result of the corollary can be used for a stopping rule in some kind of iterative algorithm, including interior point methods.

Now in order to obtain an exact solution to the LP problem, we present the modified algorithm A′ as follows:

Algorithm A′.

$k:=0$. Given $x^{(0)}$, an initial starting point.

(1) Define

$$D_k = \text{diag}(x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)}),$$

$$A_k = AD_k. \tag{4.3}$$

(2) Compute the vectors $c_p^{(k)}$ and $\bar{v}^{(k)}$

$$c_p^{(k)} = [I - A_k^T(A_kA_k^T)^{-1}A_k]D_kc, \tag{4.4}$$

and

$$\bar{v}^{(k)} = D_k^{-1}c_p^{(k)}. \tag{4.5}$$

If there is a subset $\beta=[i_1, i_2, \cdots, i_m]$. such that $|\bar{v}_{i_l}^{(k)}|<s$, $i_l\in\beta$, and $\bar{v}_{j_l}^{(k)}>s$, $j_l\in\bar{N}-\beta$, where $\bar{N}=[1, 2, \cdots, n]$, then let $B=[a_{i_1}, a_{i_2}, \cdots, a_{i_m}]$, $i_l\in\beta$, and solve the linear system of equations

$$Bx^* = b. \tag{4.6}$$

x^* is the optimal solution of the LP problem. If no such β exists, go to the next step.

(3) Normalize $c_p^{(k)}$

$$p^{(k)} = \frac{c^{(k)}}{\|c_p^{(k)}\|_2}. \tag{4.7}$$

If

$$v_i^{(k)} \leqslant s, \quad i=1, 2, \cdots, n, \tag{4.8}$$

then the minimum value of the LP problem is unbounded, stop. Otherwise, go to the next step.

(4) Determine the largest step which can be taken and generate a new point

$$\lambda_k = \frac{1}{q_k}, \tag{4.9}$$

where

$$q_k = \text{Max}[p_i^{(k)}],$$

and let

$$x^{(k+1)} = x^{(k)} - \alpha\lambda_k D_k p^{(k)}, \tag{4.10}$$

where $\alpha \in (0, 1)$, k: $k+1$, and go to step (1).

It follows from algorithm A′, and Corollary 4.3, that an exact solution of the LP problem can be achieved in a finite number of iterations.

Theorem 4.4. *Suppose that the* LP *problem satisfies the condition F. Assume its minimum value is bounded, and the optimal solution x^* is unique. Then x^* can be obtained by algorithm* A′ *in a finite number of iterations.*

Numerical tests have shown that the total number of iterations is reduced by using algorithm A′ instead of algorithm A. It follows from Theorem 2.2 that the reduced number of iterations depends on the value of the optimal solution.

Acknowledgement. The author expresses his appreciation to Prof. I. Adler of the University of California, Berkeley for many helpful discussions, to Prof. W. Murray for valuable help during the author's visit at Stanford University, and to J. Doucet for helpful editorial suggestions.

References

[1] G. B. Dantzig, *Linear Programming and Extensions*, PUP, Princeton, NJ, 1963.

[2] P. E. Gill, W. Murray, *Numerical Methods for Constrained Optimization*, Academic Press Inc., New York, London, 1974.

[3] N. Karmarkar, *A New Polynomial-time Algorithm for Linear Programming*, Combinatorica, 1984.

[4] P. Lancaster, M. Tismenetsky, *The Theory of Matrices with Applications*, Academic Press, Inc., Orlando, Florida, 1985.

[5] M. J. Todd, B. P. Burrell, *An Extension of Karmarkar's Algorithm for Linear Programming Using Dual Variables*, Tech. Report No. 648, School of Operations Research & Industrial Engineering, College of Engineering, Cornell University, Ithaca, NY 14850.

[6] J. A. Tomlin, *An Experimental Approach to Karmarkar's Projective Method for Linear Programming*, manuscript, Ketron Inc., Mountain View, CA 94040, 1985.

[7] Z. L. Wei, *An Interior Point Method for Linear Programming, to follow.*