Revista Colombiana de Matemáticas
Vol. XXI (1987), págs. 13-24

NONLINEAR DUALITY AND MULTIPLIER THEOREMS

by

Alfonso G. AZPEITIA

Abstract. The main purpose of this paper is to extend the John theorem on nonlinear programming with inequality constraints and the Mangasarian-Fromovitz theorem on nonlinear programming with mixed constraints to any real normed linear space. In addition, for the John theorem assuming Frechet differentiability, the standard conclusion that the multiplier vector is not zero is sharpened to the nonvanishing of the subvector of those components corresponding to the constraints which are not linear affine. The only tools used are generalizations of the duality theorem of linear programming, and hence of the Farkas lemma, to the case of a primal real linear space of any dimension with no topological restrictions. It is shown that these generalizations are direct consequence of the ordiry duality theorem of linear programming in finite dimension.

§1. We begin by stating, in a suitable formulation, the classical duality theorem of linear programming. Let the $m\times n$ matrix A and the m-vector b and the n-vector c be given. Define the sets.

$$C = \{t : At \geqslant b,\ t \in \mathbf{R}^n\}, \qquad D = \{u : uA = c,\ u \in \mathbf{R}^m_+\},$$

and the numbers

$$\alpha = \inf\{ct : t \in C\}, \qquad \beta = \sup\{ub : u \in D\}.$$

Then the duality theorem is as follows.

THEOREM 1. (i) *If* $t \in C$ *and* $u \in D$ *then* $ct \geqslant ub$.
(ii) *If* $\alpha = -\infty$ *then* $D = \emptyset$ *and if* $\beta = +\infty$ *then* $C = \emptyset$.
(iii) *A finite* α *exists if and only if a finite* β *exists and, if that is the case, then* $\alpha = \beta$ *and there exist* $\bar{t} \in C$ *and* $\bar{u} \in D$ *such that* $c\bar{t} = \bar{u}b = \alpha = \beta$.

For direct elementary proofs, independent of the simplex algorithm, the reader can see reference 3, pp.78-82 or reference 4, pp.71-73.

Now we generalize Theorem 1 replacing $\mathbb{R}^n$ with any real linear space $\mathbf{X}$. Let $f_i(x)$, $0 \leqslant 1 \leqslant m$, be real linear functions on $\mathbf{X}$. Define

$$F(x) = (f_1(x), f_2(x), \ldots, f_m(x)),$$

the sets

$$\Gamma = \{x : F(x) \geqslant b\},$$

$$\Delta = \{u : uF(x) = f_o(x),\ x \in \mathbf{X},\ u \in \mathbb{R}^m_+\},$$

and the numbers

$$\alpha = \inf\{f_o(x) : x \in \Gamma\} \quad \text{and} \quad \beta = \sup\{ub : u \in \Delta\}.$$

Then the generalized result is:

THEOREM 2. (a) *If* $x \in \Gamma$ *and* $u \in \Delta$ *then* $cx \geqslant ub$.
(b) *If* $\alpha = -\infty$ *then* $\Delta = \emptyset$ *and if* $\beta = +\infty$ *then* $\Gamma = \emptyset$.
(c) *A finite* α *exists if and only if a finite* β *exists and, if that is the case, then* $\alpha = \beta$, *and there exist* $\bar{x} \in \Gamma$ *and* $\bar{u} \in \Delta$ *such that* $f_o(\bar{x}) = \bar{u}b = \alpha = \beta$.

Proof. Consider the mapping of $\mathbf{X}$ to $\mathbb{R}^{m+1}$

$$\xi_i = f_i(x), \qquad x \in \mathbf{X}, \qquad 0 \leqslant i \leqslant m.$$

The range of the mapping is the linear subspace of $\mathbb{R}^{m+1}$ given by $S = \{\xi : \xi = (\xi_o, \xi_1, \ldots, \xi_m)\}$. If n is the dimension of S, let $\{a^j\}$, $1 \leqslant j \leqslant n$, be a basis for S. Then, if B is the $(m+1) \times n$ matrix with columns a^j and $t \in \mathbb{R}^n$, we have

$$S = \{\xi : \xi = Bt, \quad t \in \mathbb{R}^n\},$$

and if A is the matrix of the last m rows of B, and c is the first row of B, then

$$\{f_o(x) : x \in \Gamma\} = \{ct : At \geqslant b\}.$$

Now the statements (a), (b) and (c) are immediately obtained applying Theorem 1 to the primal program $\min\{ct : At > b\}$ and reformulating the conclusions (i), (ii) and (iii), (notice that $D = \Delta$), in terms of x and the $f_i(x)$, $0 \leqslant i \leqslant m$. ▲

A similar but weaker result that requires a topological structure on **X** together with the continuity of the functions $f_i(x)$ appears in reference [7] p.68, Theorem 3. 13. 18.

Direct consequences of this theorem are the two following extensions of the original Farkas theorem (reference 1 or reference 9, p.531).

COROLLARY 2.1 *If the* $f_i(x)$, $0 \leqslant i \leqslant m$ *and* $F(x)$ *are defined as before, then the two following statements are equivalent:*

(i) *If* $F(x) \geqslant 0$ *then* $f_o(x) > 0$.

(ii) *There exists* $u \in \mathbb{R}^m_+$ *such that* $f_o(x) = uF(x)$ *for all* $x \in \mathbf{X}$.

Proof. That (ii) ⇒ (i) is obvious. The converse implication follows from Theorem 2.

COROLLARY 2.2 *Assume that*

$$G(x) = (g_1(x), g_2(x), \ldots, g_p(x)),$$

with all the $g_j(x)$, $1 \leqslant j \leqslant p$ *linear. Then with the* f_i *and* F *defined as before the two following statements are equivalent:*

(i) *If* $F(x) \geqslant 0$ *and* $G(x) = 0$ *then* $f_o(x) \geqslant 0$.

(ii) *There exist vectors* $u \in \mathbb{R}^m_+$ *and* $v \in \mathbb{R}^p$ *such that*

$f_o(x) = uF(x)+vH(x)$ *for every* $x \in \mathbf{X}$.

Proof. That (ii) $\Rightarrow$ (i) is again obvious. To prove the converse we replace each constraint $g_j(x) = 0$ by a pair of inequalities $g_j(x) \geqslant 0$, $-g_j(x) \geqslant 0$. By Corollary 2.1, if (i) holds there exist $u \in \mathbb{R}^m_+$, $v^1 \in \mathbb{R}^p_+$ and $v^2 \in \mathbb{R}^p_+$ such that for every $x \in \mathbf{X}$,

$$f_o(x) = uF(x) = v^1G(x) - v^2G(x).$$

Defining $v = v^1 - v^2$ statement (ii) follows and the proof is complete.

§2. In this section we extend the John theorem on non linear programming with inequality constraints, (reference 8, p. 446). It will be assumed in all what follows that $\mathbf{X}$ is a normed real linear space of any dimension.

For terminology, notation and properties of Fréchet differentials, derivatives and continuous differentiability the reader may see Luenberger's book, (reference 5, pp.171-178), with which this paper is consistent. In particular, we use the notation $f'(\bar{x})$ for the derivative of $f(x)$ at $\bar{x}$.

We consider now the non linear case. Let the real functions $f_o(x)$, $f_i(x)$, $i \in I = \{1,2,\dots m\}$, be continuous in an open set U of $\mathbf{X}$. Then we have the following extension of the John theorem.

THEOREM 3. *Assume that the program*

$$\mathbb{P} : \max\{f_o(x) : x \in \Gamma\}$$

$$\Gamma = \{x : f_i(x) \geqslant 0, \quad i \in I\} \cap U$$

has a local solution at $\bar{x}$. *Let*

$$A(\bar{x}) = \{i : f_i(\bar{x}) = 0, \quad i \in I\} = L \cup N$$

with $i \in L$ *if and only if* f_i *is linear affine. Assume further that the functions* f_o *and* f_i, $i \in A(\bar{x})$ *are (Fréchet) differentiable at* $\bar{x}$. *Then there exists a multiplier vector* $u = (u_o, u_1, \ldots, u_m) \in \mathbb{R}_+^{m+1}$ *such that*

(i) *For all* $i \in I$, $u_i f_i(\bar{x}) = 0$.

(ii) *Not all the numbers* u_o *and* u_i, $i \in N$ *are zero.*

(iii) $u_o f_o'(\bar{x}) + \sum_{i \in I} u_i f_i'(\bar{x}) = 0$.

Proof. We establish in the first place that the following set $S \subset \mathbf{X}$ is empty

$$S = \{x : f_o'(\bar{x})x > 0, \quad f_i'(\bar{x})x \geqslant 0, \quad i \in N, \quad f_i'(\bar{x})x > 0, \quad i \in L\}.$$

Indeed, if there is some point $x_o \in S$ we have, by the differentiability and continuity assumptions, with $x = \bar{x} + tx_o$, $0 < t \in \mathbb{R}$, and small t, that

$$f_o(x) = f_o(\bar{x}) + tf_o'(\bar{x})x_o + t\varepsilon(t) > f_o(\bar{x}),$$

$$f_i(x) = f_i(\bar{x}) + tf_i'(\bar{x})x_o + t\varepsilon_i(t) > f_i(\bar{x}), \quad i \in N,$$

$$f_i(x) = f_i(\bar{x}) + tf_i'(\bar{x})x_o \geqslant f_i(\bar{x}), \quad i \in L,$$

and finally

$$f_i(x) > 0, \quad i \subset I/A(\bar{x}).$$

It follows that $x \in \Gamma$ and $f(x)_o > f_o(\bar{x})$ which is a contradiction and $S = \emptyset$ as claimed. We can conclude now that if $\alpha \in \mathbb{R}$ then the following set of linear inequalities

$$f_o'(\bar{x})x - \alpha \geqslant 0,$$

$$f_i'(\bar{x})x - \alpha \geqslant 0, \quad i \in N,$$

$$f_i'(\bar{x})x \geqslant 0, \quad i \in L,$$

imply the inequality $\alpha \leqslant 0$. By Corollary 2.1 applied in the space $\mathbf{X} \times \mathbf{R}$ of points (x,α) there exist multipliers $u_o \geqslant 0$, $u_i \geqslant 0$, $i \in A(\bar{x})$, such that

$$u_o f_o'(\bar{x}) + \sum_{i\in A(\bar{x})} u_i f_i'(\bar{x}) = 0 \quad \text{and} \quad u_o + \sum_{i\in N} u_i = 1.$$

Defining $u_i = 0$ for $i \in I/A(\bar{x})$ the proof is complete. ▲

If we assume that $N = \emptyset$ we obtain a local extension of Theorem 2 for the case of nonlinear objective.

COROLLARY 3.1 *If* $N = \emptyset$ *then* $u_o = 1$.

Similarly an adequate constraint qualification gives the following variant of the Kuhn-Tucker theorem (reference 4, p.233).

COROLLARY 3.2 *If there exists a vector* $y \in \mathbf{X}$ *such that* $f_i'(\bar{x})y > 0$ *if* $u_i > 0$, $i \in I$, *then* $u_o > 0$.

Proof. If $u_o = 0$ we obtain from (iii) the contradiction

$$0 = \Big[\sum_{i\in I} u_i\, f_i'(\bar{x})\Big]y > 0.$$

REMARK. The same conclusion is also obtained with the more restrictive assumption that the $f_i'(\bar{x})$ involved are linearly independet.

§3. The Mangasarian-Fromovitz theorem (reference 6) can be extended in a similar way. With U, **X**, and the continuous functions $f_o(x)$, $f_i(x)$, $i \in I$, as in Theorem 3, we consider additional equality constraints

$$g_j(x) = 0, \quad j \in J = \{1,2,\ldots,p\},$$

with each $g_j(x)$ defined in U. The result is now as follows.

THEOREM 4. *Assume that the program*

$$\mathbb{P} : \max\{f_o(x) : x \in \Gamma\},$$

$$\Gamma = \{x : f_i(x) \geq 0,\ i \in I,\ g_j(x) = 0,\ j \in J\} \cap U$$

has a local solution at $\bar{x}$. *Let*

$$A(\bar{x}) = \{i : f_i(x) = 0,\ i \in I\}$$

and assume further that $f_o(x)$ *and* $f_i(x)$, $i \in A(\bar{x})$, *are differentiable at* $\bar{x}$ *and that the* $g_j(\bar{x})$, $j \in J$, *are continuously differentiable in some open neighborhood of* $\bar{x}$. *Then there exists a multiplier vector* $u = (u_o, u_1, \ldots, u_m, v_1, v_2, \ldots v_p)$ *such that*

(i) $u \neq 0$;

(ii) $u_o \geq 0$ *and* $u_i \geq 0$, $i \in I$;

(iii) $u_i f_i(\bar{x}) = 0$, $i \in I$;

(iv) $u_o f_o'(\bar{x}) + \sum_{i \in I} u_i f_i'(\bar{x}) + \sum_{j \in J} v_j g_j'(\bar{x}) = 0$;

(v) *If the derivatives* $g_j'(\bar{x})$, $j \in J$, *are linearly independent then* u *can be chosen such that the numbers* u_o, u_i, $i \in A(\bar{x})$, *are not all zero.*

Proof. If the $g_j'(\bar{x})$, $j \in J$, are linearly dependent the theorem is trivial. Otherwise, we show first that the following set S is empty:

$$E = \{x : f_o'(\bar{x})x > 0, \quad f_i'(\bar{x})x > 0, \quad i \in A(\bar{x}),\ g_i'(\bar{x})x = 0,\ j \in J\}.$$

Assume that, on the contrary, there is some $x_o \in E$. Then $x_o \neq 0$ and since the $g_j'(\bar{x})$ are linearly independent there exist linearly independet vectors $x_j \in \mathbf{X}$, $j \in J$ such that the determinant

$$\Delta = \det\left[g_\lambda'(\bar{x})x_\mu\right] , \quad \lambda \in J, \quad \mu \in J,$$

is different from zero. We introduce now the p + 1 real variables $S, S_1, S_2, \ldots, S_p$ and the p functions G_j, $j \in J$, with domain $\mathbb{R}^{p+1}$, given by

$$G_j (S, S_1, \ldots, S_p) = g_j(\bar{x}+Sx_o+S_1x_1+\ldots+S_px_p), \quad j \in J.$$

The functions G_j are continuously differentiable in a neighhood of the origin, satisfy the conditions $G_j(0,0,\ldots,0) = 0$, $j \in J$, and their Jacobian at $(0,0,\ldots,0)$ is $\Delta \neq 0$. If we set the equations

$$G_j(S,S_1,\ldots,S_p) = 0, \quad j \in J, \tag{1}$$

it therefore follows from the implicit function theorem (see for example reference [2], p. 148) that for some number $\sigma > 0$ there exists a continuously differentiable mapping from the interval $(-\sigma,\sigma) \to \mathbb{R}^p$; that is, a curve

$$\gamma = \{(S_1,S_2,\ldots,S_p) : S_j = \gamma_j(S), \ -\sigma < S < \sigma\}$$

with $\gamma_j(0) = 0$, $j \in J$, and with all the continuously differentiable functions $\gamma_j(S)$ satisfying identically the equations (1) for all S in the interval $(-\sigma,\sigma)$. Consequently there is another continuously differentiable curve $\gamma^* \in \mathbf{X}$ given by

$$\gamma^* = \{x : x = \phi(S) = \bar{x}+Sx_o+\gamma_1(S)x_1+\ldots+\gamma_p(S)x_p, \ -\sigma < S < \sigma\}$$

for which $\phi(0) = \bar{x}$ and $\phi'(0) = x_o$. Now, for small $S > 0$ we have

$$f_o(x)-f_o(\bar{x}) = f_o(\phi(S))-f_o(\phi(0))$$

$$= f'(\phi(0))\phi'(0)S + \varepsilon(S)S = \left[f_o'(\bar{x})x_o + \varepsilon(S)\right]S > 0. \tag{2}$$

Similarly

$$f_i(x) = f_i(x) - f_i(\bar{x}) > 0, \quad i \in A(\bar{x}). \tag{3}$$

Since $\|x-\bar{x}\|$ is small and all the functions f_i are continuous at $\bar{x}$ we also have

$$f_i(x) > 0, \quad i \in I/A(\bar{x}). \tag{4}$$

Finally, for all $j \in J$

$$g_j(x) = g_j(\phi(S)) = 0. \tag{5}$$

From (2), (3), (4) and (5) we conclude that $\bar{x}$ is not a local solution of $\mathbb{P}$ and the claim $E = \emptyset$ is established. Now, as in the case of Theorem 3, it follows that with $\alpha \in \mathbb{R}$ the set of linear conditions

$$f_o'(\bar{x})x - \alpha \geqslant 0 \ ;$$

$$f_i'(\bar{x})x - \alpha \geqslant 0, \qquad i \in A(\bar{x}) \ ;$$

$$g_j'(\bar{x})x = 0, \quad j \in J \ ;$$

implies $\alpha \leqslant 0$.

By Corollary 2.2 applied to the space $\mathbf{X} \times \mathbb{R}$ of points (x,α) it follows that there exist numbers $u_o \geqslant 0$, $u_i \geqslant 0$, $i \in A(\bar{x})$ and v_j, $j \in J$, such that

$$u_o f_o'(\bar{x}) + \sum_{i \in A(\bar{x})} u_i f_i'(\bar{x}) + \sum_{j \in J} v_j \, g_j'(\bar{x}) = 0,$$

and that

$$u_o + \sum_{i \in A(\bar{x})} u_i = 1.$$

Defining now $u_i = 0$ if $i \in I/A(\bar{x})$ the theorem is proved. ▲

Finally we have the following corollary similar to Corollary 3.2.

COROLLARY 4.1 *If the* $g_j'(\bar{x})$, $j \in J$, *are linearly independent and if there exists a vector* $y \in \mathbf{X}$ *such that*

(a) $g'(\bar{x})y = 0$, $j \in J$,

(b) $f_i'(\bar{x})y > 0$ *if* $u_i > 0$, $i \in I$, *then* $u_o > 0$.

Proof. Let $T = \{i : u_i > 0, i \in I\}$; then if $u_o = 0$ it follows from (iv) that

$$\sum_{i \in T} u_i \left[f_i'(\bar{x})y\right] = 0 ,$$

which by (b) implies the contradiction $u_i = 0$ for $i \in T$, unless $T = \emptyset$, in which case (iv) contradicts the linear independence assumption. ▲

A remark similar to the one made after Corollary 3.2 applies here as well.

REFERENCES

[1] Farkas, J., *Theorie der einfachen Ungleichungen*, J.Reine Angew. Math. 124 (1901) 1-27.

[2] Fleming, W., *Function of Several Variables*, 2nd ed., Springer-Verlag, New York, Heidelberg, Berlin, 1977.

[3] Gale, D., *The Theory of Linear Economic Models*, McGraw Hill, New York, (1960).

[4] Luenberger, D.G., *Introduction to linear and nonlinear programming*, Addison-Wesley, Reading Massachusetts, 1972.

[5] Luenberger, D.G., *Optimization by Vector Space Methods*, Wiley & Sons, New York, 1969.

[6] Mangasarian, O.L. and Fromovitz, S., *The Fritz John necessary optimality conditions in the presence of equality and inequality constraints*, J. Math. Anal. Appl., 17 (1967) 37-47.

[7] Ponstein, J., *Approaches to the Theory of Optimization*, Cambridge Tracts in Mathematics Nº 77, Cambridge Univ. Press, Cambridge, 1980.

[8] Pourciau, B.H., *Modern multiplier rules*, Amer. Math. Montly, 87 (1980) 433-452.

[9] Prekopa, A., *On the development of Optimization Theory*, Amer. Math. Monthly, 87 (1980) 527-542.

Department of Mathematics and Computer Science
University of Massachusetts, Boston
Boston, Massachusetts 02125
U. S. A.

(Recibido en julio de 1986).