The Theorem of the Alternative, the Key-Theorem, and the Vector-Maximum Problem, for Non-Polyhedral Ordering Cones

R. Lehmann *

W. Oettli **

Summary: Consequences of a general formulation of the theorem of the alternative are exploited.

* Mathematisches Institut der Universität Zürich, CH-8032 Zürich, Freiestrasse 36

** Fakultät für Mathematik und Informatik der Universität Mannheim, D-68 Mannheim, Schloss.

1. Introduction

The theorem of the alternative and Tucker's key-theorem are two companion theorems frequently used in Operations Research and Activity Analysis [1], [2]. The theorem of the alternative states that of the two linear systems

$$Bx > 0$$

and

$$B^T\eta = 0,\ \eta \geq 0,\ \eta \neq 0$$

one and only one has a solution. The key-theorem states that the linear system

$$Bx \geq 0,\ B^T\eta = 0,\ \eta \geq 0,\ Bx+\eta > 0$$

always has a solution. Below we discuss a general analogue of these theorems for arbitrary orderings induced by convex cones. The extension of the key-theorem does not seem to have been considered up to now in the literature. From our extension of the theorem of the alternative it is possible to obtain by simple substitutions, mainly based on Farkas' lemma, all the different versions of the theorem of the alternative which occur in the literature (see e.g. [3]). We also obtain in this way a very convenient approach to generalized linear programming as initiated by Duffin [4]. Whereas the validity of the key-theorem seems to be restricted to finite-dimensional spaces the theorem of the alternative holds for more general linear spaces. We state the prerequisites for a reflexive space, though the results proper for the sake of uniformity are all formulated in $\mathbb{R}^n$. As an application we treat a generalized version of the linear vector-maximum problem.

2. Prerequisites

Let $\mathcal{X}$ be a reflexive space, $\mathcal{X}^*$ its dual. As in $\mathbb{R}^n$ we denote the pairing between $\xi \in \mathcal{X}^*$ and $x \in \mathcal{X}$ by $\xi^T x$. Let $P \subset \mathcal{X}$ be a nonvoid, convex cone (i.e. $x, y \in P \Rightarrow \lambda x + \mu y \in P \ \forall \lambda \geq 0, \mu \geq 0$). The polar of P is defined as

$$P^+ = \{\xi \in \mathcal{X}^* \mid \xi^T x \geq 0 \ \forall x \in P\}, \quad P^- = -P^+.$$

P^+ is a closed, nonvoid, convex cone. We list some well-known properties of polars which we need in the sequel. P, Q denote convex cones.

i) $$P \subset Q \Rightarrow P^+ \supset Q^+.$$

ii) $$\begin{pmatrix} P \\ Q \end{pmatrix}^+ = \begin{pmatrix} P^+ \\ Q^+ \end{pmatrix},$$

where $\begin{pmatrix} P \\ Q \end{pmatrix}$ denotes the Cartesian product of P and Q.

iii) $$P^{++} = \overline{P},$$

where $\overline{P}$ denotes the topological closure of P.

iv) Let $A: \mathcal{X} \to \mathcal{Y}$ be a continuous linear transformation, A^T its adjoint, and A^{-1} its (multivalued) inverse, $AQ = \{Ax \mid x \in Q\}$ and $A^{-1}P = \{x \mid Ax \in P\}$. Then

$$(AQ)^+ = (A^T)^{-1} Q^+,$$

$$(A^{-1}\overline{P})^+ = \overline{A^T P^+} \quad \text{(Farkas' lemma)}.$$

The first statement merely expresses the equality of $\eta^T(Ax)$ and $(A^T\eta)^T x$; the second statement follows then from iii) and $A^{TT} = A$. As special cases of the above we have

$$(P+Q)^+ = P^+ \cap Q^+,$$

$$(\overline{P} \cap \overline{Q})^+ = \overline{P^+ + Q^+}.$$

v) Let $Q^o \neq \emptyset$ (Q^o the topological interior of Q). Then $P \cap Q^o = \emptyset$ implies the existence of $\xi \neq 0$ with $\xi \in P^- \cap Q^+$ (weak separation theorem).

vi) $x \in Q^o$, $\xi \in Q^+ \setminus \{0\} \Rightarrow \xi^T x > 0$.

vii) Let P be closed. Sometimes we shall make the simplifying assumption that AP is again closed. This assumption will be true, for example, if A is nonsingular. In $\mathbb{R}^n$ it will be true if P is polyhedral, i.e., is of the form

$$P = \{x | Bx \leq 0\} \text{ or } P = \{x | x = Cu,\ u \geq 0\}$$

(both definitions are equivalent according to theorems of Minkowski and Weyl). If P is polyhedral, then AP and AP^- are also polyhedral and a fortiori closed. Otherwise the literature seems to be poor on practical conditions ensuring that a linear transform of a closed cone is again closed.

3. Theorems

In all what follows the setting is a finite-dimensional linear space.

Theorem 1 (theorem of the alternative): Let P, Q be nonvoid convex cones with $Q^o \neq \emptyset$. Then exactly one of the following two systems (1), (2) has a solution:

(1) $x \in P$, $x \in Q^o$;

(2) $\xi \in P^-$, $\xi \in Q^+$, $\xi \neq 0$.

Proof: If (1) has a solution, then (2) has no solution, since otherwise $x \in P$ and $\xi \in P^-$ implies $\xi^T x \leq 0$, whereas $x \in Q^o$ and $\xi \in Q^+ \setminus \{0\}$ implies $\xi^T x > 0$, by vi). If (1) has no solution, then (2) has a solution by the weak separation theorem v).

Theorem 2 (key-theorem): Let P, Q be nonvoid, closed, convex cones with Q-P being closed. Then the system (3) has a solution:

(3) $x \in P,\ x \in Q,\ \xi \in P^-,\ \xi \in Q^+,\ x+\xi \in (Q+Q^+)^o.$

Proof: Let $K_1 = Q+Q^+$. If K_1 would have empty interior, then, since situated in $\mathbb{R}^n$, it would be contained in a hyperplane H. The polar $K_1^+ = Q^+ \cap Q$ would, by i), contain H^+, which is a line. But then $Q^- \cap Q$ would also contain this line; this is impossible since $Q^- \cap Q = \{0\}$. Therefore K_1 is not empty. Let $K_2 = (P \cap Q)+(Q^+ \cap P^-)$. If the theorem would be wrong we would have $K_1^o \neq \emptyset$, $K_1^o \cap K_2 = \emptyset$, and therefore, by v), the existence of $t \neq 0$ with

$$t \in K_1^- = Q^- \cap -Q,\quad t \in K_2^+ = (P \cap Q)^+ \cap (Q-P).$$

$t \in Q-P$ means the existence of $p \in P$ with $t \in Q-\{p\}$; but then $p \in Q-\{t\}$, and this implies, since $-t \in Q$, that $p \in Q+Q=Q$. Thus $p \in P \cap Q$, and consequently $t \in Q-(P \cap Q)$. Since, however, $t \in (P \cap Q)^+$ and $t \in Q^-$, it follows $t^T t \leq 0$, contradicting $t \neq 0$.

4. Corollaries

We call two systems of conditions dual to each other, if one and only one is consistent. (1) and (2) are examples of dual systems. We derive some more of these. In the remainder of our discussion let A, B, C denote linear transformations and P, Q, S nonvoid convex cones having the properties

$$P \text{ closed},$$
$$Q^o \neq \emptyset,$$
$$S \text{ closed},\ (S^+)^o \neq \emptyset.$$

Corollary 1: The following two systems are dual:

(4) $Ax \in P,\ Bx \in Q^o;$

(5) $B^T \eta \in \overline{A^T P^-},\ \eta \in Q^+ \setminus \{0\}.$

Proof: (4) is equivalent to $y \in V \equiv B(A^{-1}P)$, $y \in Q^o$. This system, by theorem 1 and iv), is dual to $\eta \in V^- \equiv (B^T)^{-1}\overline{A^T P^-}$, $\eta \in Q^+ \setminus \{0\}$. The latter is equivalent to (5).

By the same substitution which led from theorem 1 to corollary 1 we may conclude from theorem 2 that the system

$$Ax\in P,\ Bx\in Q,\ B^T\eta\in\overline{A^TP^-},\ \eta\in Q^+,\ Bx+\eta\in(Q+Q^+)^o$$

always has a solution, provided Q, $V\equiv B(A^{-1}P)$, $Q-V$, P are closed (Q^o may be empty here). Setting $A=0$ results in Tucker's theorem.

Corollary 2: Let $V\equiv C(A^{-1}P)$ be closed. Then the following two systems are dual:

(6) $$Ax\in P,\ Cx\in S\setminus\{0\};$$

(7) $$C^T\zeta\in\overline{A^TP^-},\ \zeta\in(S^+)^o.$$

Proof: (6) is equivalent to $z\in V$, $z\in S\setminus\{0\}$. This system may be identified with (2). Its dual, according to (1), is $\zeta\in V^-\equiv(C^T)^{-1}\overline{A^TP^-}$, $\zeta\in(S^+)^o$. This is equivalent to (7).

If we drop the closedness assumption on V, then we can merely ascertain that (6) is dual to

(7') $$C^T\zeta\in\overline{A^TP^-+C^TS^-},\ \zeta\in(S^+)^o.$$

This follows from the next corollary upon setting $Q=\mathbb{R}^k$.

Corollary 3: The two systems (8), (9) are dual:

(8) $$Ax\in P,\ Bx\in Q^o,\ Cx\in S\setminus\{0\};$$

(9) $$\begin{cases}\text{(a)}\quad B^T\eta\in\overline{A^TP^-+C^TS^-},\ \eta\in Q^+\setminus\{0\}\\ \qquad\qquad\text{and/or}\\ \text{(b)}\quad B^T\eta+C^T\zeta\in\overline{A^TP^-+C^TS^-},\ \eta\in Q^+,\ \zeta\in(S^+)^o.\end{cases}$$

Proof: Assume that (8) has no solution. Choose $\zeta\in(S^+)^o$ fixed. According to vi) $Cx\in S\setminus\{0\}$ is equivalent to $Cx\in S$, $\zeta^TCx>0$. Therefore with (8) the system

$$x\in V,\ \begin{bmatrix}B\\ \zeta^TC\end{bmatrix}x\in\begin{pmatrix}Q\\ \mathbb{R}_+\end{pmatrix}^o$$

has no solution either, where $V=(A^{-1}P)\cap(C^{-1}S)$, $\mathbb{R}_+=\{\lambda\in\mathbb{R}\,|\,\lambda\geq 0\}$.

According to corollary 1 the dual system

$$B^T\eta+C^T\zeta\mu\in V^-,\ \eta\in Q^+,\ \mu\geq 0,\ (\eta,\mu)\neq 0$$

has a solution, where $V^- = \overline{A^TP^-+C^TS^-}$ by iv). If $\mu = 0$ this means that (9a) has a solution; if $\mu > 0$ we may normalize $\mu = 1$, obtaining thereby a solution of (9b). Thus, if (8) has no solution, (9) has a solution. Conversely (8) and (9) cannot have solutions at the same time. This follows readily from vi). For example (8) and (9b) together result in the contradicting inequalities $\eta^TBx+\zeta^TCx \leq 0$ [since $x\in V$, $B^T\eta+C^T\zeta\in V^-$], $\eta^TBx \geq 0$ [since $Bx\in Q$, $\eta\in Q^+$], $\zeta^TCx > 0$ [since $Cx\in S\setminus\{0\}$, $\zeta\in(S^+)^o$]. Thus (8) and (9) are dual.

We see from the proof that ζ in (9) may or may not be considered as being variable. We also note that (9b) is dual to the system $Ax\in P$, $Bx\in Q^o$, $Cx\in S$. If the latter is supposed to have a solution, then (8) and (9a) are dual.

If $A^TP^-+C^TS^-$ is closed, then, because of $C^TS^++C^T(S^+)^o = C^T(S^+)^o$, (9) can be written as

$$(9')\quad \begin{cases} \text{(a)} \quad 0 = A^T\xi+B^T\eta+C^T\zeta,\ \xi\in P^+,\ \eta\in Q^+\setminus\{0\},\ \zeta\in S^+ \\ \qquad\qquad \text{and/or} \\ \text{(b)} \quad 0 = A^T\xi+B^T\eta+C^T\zeta,\ \xi\in P^+,\ \eta\in Q^+,\ \zeta\in(S^+)^o. \end{cases}$$

We turn to the case of inhomogeneous systems.

Corollary 4: The following two systems are dual:

$$(10)\qquad Ax+a\in P,\ Bx+b\in Q^o;$$

$$(11)\qquad \begin{pmatrix} B^T\eta \\ b^T\eta+\mu \end{pmatrix} \in \overline{\begin{bmatrix} A^T \\ a^T \end{bmatrix} P^-},\ \eta\in Q^+,\ \mu\geq 0,\ (\eta,\mu)\neq 0.$$

Under the assumption that $Ax+a\in P$ has a solution the condition $(\eta,\mu)\neq 0$ in (11) may be replaced by $\eta\neq 0$.

Proof: (10) is equivalent with

$$Ax+a\lambda\in P, \begin{pmatrix} Bx+b\lambda \\ \lambda \end{pmatrix} \in \begin{pmatrix} Q \\ \mathbb{R}_+ \end{pmatrix}^o ,$$

which is dual to (11), by corollary 1. If $Ax+a\lambda\in P$, $\lambda > 0$ has a solution, then the dual system

$$\begin{pmatrix} 0 \\ \mu \end{pmatrix} \in \overline{\begin{bmatrix} A^T \\ a^T \end{bmatrix} P^-}, \ \mu \geq 0, \ \mu \neq 0$$

has no solution. This means that (11) can be satisfied only with $\eta \neq 0$.

5. The vector-maximum problem: an application.

We apply corollary 4 to a generalized version of the linear vector-maximum problem [5]. We define the sets

$$\Gamma = \{b \,|\, Ax+a\in P,\ Bx+b=0\},$$

$$\Delta = \{b \,|\, \begin{pmatrix} B^T\eta \\ b^T\eta \end{pmatrix} \in \overline{\begin{bmatrix} A^T \\ a^T \end{bmatrix} P^-},\ \eta\in Q^+\setminus\{0\}\},$$

where b is considered now as being variable. We call $\bar{b}\in\Gamma$ Q-minimal, ib $b-\bar{b} \notin -Q^o$ $\forall b\in\Gamma$; we call $\bar{b}\in\Delta$ Q-maximal, if $b-\bar{b} \notin Q^o$ $\forall b\in\Delta$. The following can be said:

1) $b_1 \in \Gamma$, $b_2 \in \Delta \Rightarrow b_2-b_1 \notin Q^o$.

Proof: Since $\begin{pmatrix} x \\ 1 \end{pmatrix} \in [A,a]^{-1}P \equiv V$ and $\begin{pmatrix} B^T\eta \\ b_2^T\eta \end{pmatrix} \in V^-$ we have

$$(b_2-b_1)^T\eta = b_2^T\eta + x^TB^T\eta = (x^T,1)\begin{pmatrix} B^T\eta \\ b_2^T\eta \end{pmatrix} \leq 0,$$

whereas $b_2-b_1 \in Q^o$, $\eta \in Q^+\setminus\{0\}$ would result in $(b_2-b_1)^T\eta > 0$.

2) $\bar{b} \in \Gamma$, $b-\bar{b} \notin -Q^o$ $\forall b\in\Gamma \Rightarrow \bar{b} \in \Delta$.

Proof: The system $Ax+a\in P$, $Bx+\bar{b}\in Q^0$ has no solution. According to corollary 4 the system

$$\begin{pmatrix} B^T\eta \\ \bar{b}^T\eta+\mu \end{pmatrix} \in \overline{\begin{bmatrix} A^T \\ a^T \end{bmatrix} P^-}, \quad \eta\in Q^+, \quad \mu \geq 0, \quad (\eta,\mu) \neq 0$$

has a solution. But since the system $Ax+a\in P$, $Bx+\bar{b}=0$ has a solution, $\bar{x}$ say, we have

$$0 \geq (\bar{x}^T,1)\begin{pmatrix} B^T\eta \\ \bar{b}^T\eta+\mu \end{pmatrix} = \mu;$$

therefore $\mu = 0$, and $\bar{b}\in\Delta$.

From 1) and 2) follows

<u>Theorem 3:</u> Let $\bar{b}\in\Gamma$. Then $\bar{b}$ is Q-minimal if and only if $\bar{b}\in\Delta$ (equivalently: if and only if $\bar{b}$ is Q-maximal).

3) Under a certain closedness assumption we may infer from $\Gamma \neq \emptyset$ and $\Delta \neq \emptyset$ that $\Gamma\cap\Delta \neq \emptyset$. The argument runs as follows:

Let η be feasible for Δ. With $\rho = B^T\eta$ it is sufficient to show that the system

$$(12) \qquad Ax+a\in P, \quad \begin{pmatrix} \rho \\ -\rho^T x \end{pmatrix} \in \overline{\begin{bmatrix} A^T \\ a^T \end{bmatrix} P^-}$$

has a solution x, for $b = -Bx$ is then in $\Gamma\cap\Delta$. The system

$$(13) \qquad \begin{cases} -A^Tw + \rho\delta = 0 \\ Az + a\delta \in P \\ w \in P^- \\ \delta \geq 0 \\ a^Tw + \rho^Tz > 0 \end{cases}$$

cannot have a solution. $\delta > 0$ would result in the contradiction

$$0 \geq w^T(Az+a\delta) > z^T(A^Tw-\rho\delta) = 0.$$

$\delta = 0$ would result in

$$0 \geq w^T(Ax+a) = a^T w > -\rho^T z$$

for any x which is feasible for Γ. But since $\rho \in \overline{A^T P^-}$ and $z \in A^{-1}P$ we must have $\rho^T z \leq 0$, again a contradiction. According to corollary 1 the system

$$\begin{pmatrix} a \\ \rho \\ 0 \end{pmatrix} \in \overline{\begin{bmatrix} -A & 0 & I & 0 \\ 0 & A^T & 0 & 0 \\ \rho^T & a^T & 0 & 1 \end{bmatrix} \begin{pmatrix} \mathfrak{X} \\ P^- \\ P \\ \mathbb{R}_- \end{pmatrix}},$$

dual to (13), has then a solution. If we can omit here the closure operation, this means that

$$Ax+a \in P, \begin{pmatrix} \rho \\ -\rho^T x+\mu \end{pmatrix} \in \begin{bmatrix} A^T \\ a^T \end{bmatrix} P^-, \ \mu \geq 0$$

has a solution. As under 2) we conclude that $\mu = 0$, and (12) has a solution.

Note: If b is a scalar, then we may replace theorem 3 by the more general statement that

$$\bar{b} = \inf \{b \mid Ax+a \in P, \ Bx+b=0\}$$

if and only if

$$\bar{b} = \max \{b \mid \begin{pmatrix} B^T \\ b \end{pmatrix} \in \overline{\begin{bmatrix} A^T \\ a^T \end{bmatrix} P^-}\},$$

provided the inf is finite and $Ax+a \in P$ has a solution. This follows readily from the duality of the two systems

$$Ax+a \in P, \ Bx+b > 0$$

and

$$\begin{pmatrix} B^T \\ b+\mu \end{pmatrix} \in \overline{\begin{bmatrix} A^T \\ a^T \end{bmatrix} P^-}, \ \mu \geq 0.$$

References

[1] H. Nikaido: Convex Structures and Economic Theory, pp.33-39. Academic Press, New York, 1968.

[2] O. Mangasarian: Nonlinear Programming, pp.16-37. McGraw-Hill, New York, 1969.

[3] A. Berman, A. Ben-Israel: Linear inequalities, mathematical programming and matrix theory. Math. Programming 1 (1971), 291-300.

[4] R.J. Duffin: Infinite programs. Ann. of Math. Studies 38 (1956), 157-170.

[5] P. Schönfeld: Effizienz und Dualität in der Aktivitätsanalyse. Dissertation, Freie Universität Berlin, 1964.