DUALITY THEOREMS FOR CERTAIN NONCONVEX EXTREMAL PROBLEMS

A. L. Fradkov

IDC 519.95

This article investigates the problem posed in [1].

Suppose real functions $F(x)$, $G_1(x)$, ..., $G_m(x)$ are given on an arbitrary set X. Let τ_1, ..., τ_m be real numbers, with $\tau = \|\tau_j\|_{j=1}^m$. Consider the following relations:

$$F(x) \geqslant 0 \text{ for } G_1(x) \geqslant 0, \ldots, G_m(x) \geqslant 0. \tag{0.1}$$

$$\exists \tau_j \geqslant 0, \quad j = 1, \ldots, m : F(x) - \sum_{j=1}^{m} \tau_j G_j(x) \geqslant 0, \quad \forall x \in X. \tag{0.2}$$

It is obvious that (0.1) is implied by (0.2), but that the converse is in general false. The problem posed in [1], and given detailed consideration in [2], consists of giving conditions under which (0.2) is implied by (0.1), i.e., when (0.1) and (0.2) are equivalent. Several problems in the absolute stability of control systems lead to this problem [2, 3]. But it also has applications in the theory of operators on spaces with indefinite metrics [4] and in the variational calculus [5]. Under the assumptions we have made, the functions $F(x)$, $G_1(x)$, ..., $G_m(x)$ obviously depend on several "constructive" parameters,* and we have to find the domain A, say, in the space of these parameters in which (0.1) holds. Finding this domain involves complications, because of the constraints $G_1(x) \geq 0$, ..., $G_m(x) \geq 0$. Therefore the following procedure is often used, called the S-procedure† in [3]. Form the function $S(x, \tau) = F(x) - \sum_{j=1}^{m} \tau_j G_j(x)$, which depends on the extra parameters τ_1, ..., τ_m, and find the domain B in the parameter space for which (0.2) holds. It is always true that $A \subset B$. If the conditions (0.1) and (0.2) are equivalent, then $A = B$.

We shall say, following [2], that the S-procedure for the inequality $F(x) \geq 0$ with the constraints $G_1(x) \geq 0, \ldots, G_m(x) \geq 0$ is advantageous or favorable if (0.1) implies (0.2). Similarly, we may define the advantage or favor of the S-procedure for the inequality $F(x) > 0$ and $F(x) = 0$ under the constraints $G_j(x) > 0$ and $G_j(x) = 0$, for all possible combinations. The advantage of the S-procedure is intimately linked with the validity of a duality theorem in a certain mathematical programming problem. This tie-in is investigated in Paragraphs 1 and 2 of this article. We note that the nonconvexity of the functions met with in applications complicates the application of the usual duality theory [6]. For example, the functions $F(x)$, $G_1(x)$, ..., $G_m(x)$ can be indefinite quadratic forms.

Nonetheless, it is known [2, 4] that for $m = 1$, if $F(x)$ and $G(x)$ are quadratic (Hermitian) forms‡ on the real (complex) space X, with $G(x_0) > 0$ for some $x_0 \in X$, then the S-procedure is favorable, i.e., (0.2) is

*For example, if $F(x)$, $G_1(x)$, ..., $G_m(x)$ are quadratic forms on a Euclidean space X, then these parameters may be the coefficients of the forms.

†We give the S-procedure as described in [1].

‡By a quadratic form $F(x)$ on a real linear space X, we mean a functional of the form $F(x) = B(x, x)$, where $B(x, y)$ is a symmetric bilinear functional on $X \times X$. By an Hermitian form on a complex linear space, we mean a functional $B(x, x)$ where $B(x, y)$ is an Hermitian-symmetric and Hermitian-bilinear functional on $X \times X$ (i.e., linear in the first argument, antilinear in the second, and satisfying the relation $B(x, y) = \overline{B(y, x)}$ for all $x, y \in X$).

Translated from Sibirskii Matematicheskii Zhurnal, Vol. 14, No. 2, pp. 357-383, March-April, 1973. Original article submitted June 21, 1972.

implied by (0.1). If $F(x)$, $G_1(x)$, ..., $G_m(x)$ are quadratic (Hermitian) forms, then when $m > 1$ in the real case and when $m > 2$ in the complex case the S-procedure is not favorable in general. It was proved for the complex case in [7] that when $m = 2$, i.e., when $F(x)$, $G_1(x)$, $G_2(x)$ are Hermitian forms on the linear space X, the S-procedure is not favorable if $G_1(x_0) > 0$, $G_2(x_0) > 0$ for some $x_0 \in X$. In Paragraph 3 of this article we prove these statements by a method which differs from that of [2, 4, 7]. The proofs are based on certain facts from the Euclidean geometry of the space $\mathcal{F}(X)$ of quadratic (Hermitian) forms on the finite dimensional space X.

In Paragraph 4 we reformulate conditions (0.1) and (0.2) in terms of the space $\mathcal{F}(X)$ of Hermitian (quadratic) forms, and derive the necessary and sufficient condition for the S-procedure to be favorable for any $G_j(x) \in \mathcal{F}(X)$, $j = 1, \ldots, m$ when the forms $F(x) \in \mathcal{F}(X)$ are given. The question of finding the classes of forms $G_1(x), \ldots, G_m(x)$ having this property arises in problems of the absolute stability of control systems, where the S-procedure is applied to an a priori unknown form $F(x)$. The results of Paragraph 4 make it possible to resolve the question of the advantage of the S-procedure for quadratic and Hermitian forms in two variables.

Paragraph 5 extends the results of Paragraphs 3 and 4 to a class of functions broader than that of quadratic or Hermitian forms. In the last Paragraph we collect together several examples of unfavorable S-procedures. They show, in particular, why this or that condition of the various theorems of the previous Paragraphs is necessary.

§1. Geometrical Interpretation of the Problem. General Conditions for the S-Procedure to be Favorable

In this Paragraph we give a geometrical interpretation of the relations used to formulate the different variants of the S-procedure. Using this, we obtain general criteria for the S-procedure to be favorable. The line of argument is akin to that of Paragraph 3 of [2], where the S-procedure for the inequality $F(x) \geq 0$ with the constraints $G_1(x) \geq 0, \ldots, G_m(x) \geq 0$ was considered. We note that the results obtained here may easily be generalized to the case of infinitely many constraints. Here we consider in more detail the case of the inequality $F(x) \geq 0$ or $F(x) > 0$ with the constraints $G_1(x) \geq 0, \ldots, G_m(x) \geq 0$ or $G_1(x) = 0, \ldots, G_m(x) = 0$ (the case where the constraints are in the form of an arbitrary combination of equalities and inequalities is treated similarly). We now write out the conditions on the S-procedure for these variants*:

$$F(x) \geqslant 0 \quad \text{for} \quad G_1(x) = 0, \ldots, G_m(x) = 0, \quad x \in X. \tag{1.1}$$

$$\exists \tau : S(x, \tau) \geqslant 0 \quad \forall x \in X. \tag{1.2}$$

$$F(x) > 0 \quad \text{for} \quad G_1(x) \geqslant 0, \ldots, G_m(x) \geqslant 0, \quad x \in X. \tag{1.3}$$

$$\exists \tau \geqslant 0 : S(x, \tau) > 0, \quad \forall x \in X. \tag{1.4}$$

$$F(x) > 0 \quad \text{for} \quad G_1(x) = 0, \ldots, G_m(x) = 0, \quad x \in X. \tag{1.5}$$

$$\exists \tau : S(x, \tau) > 0 \quad \forall x \in X. \tag{1.6}$$

Here X is an arbitrary set, and $F(x)$, $G_1(x)$, ..., $G_m(x)$ are real functions on X.

Consider the map $\varphi : X \to \mathbf{R}^{m+1}$ whose coordinate functions are $G_1(x)$, ..., $G_m(x)$, i.e., $\varphi(x) = (G_1(x), \ldots, G_m(x), F(x))$, $x \in X$. Our conditions locate the set $\mathbf{R}^{m+1}$ in the space $\varphi(X)$. Let $z = (\xi_1, \ldots, \xi_m, \xi_{m+1}) \in \mathbf{R}^{m+1}$. Introduce the sets $Q = \{z : \xi_j \geq 0, j = 1, \ldots, m, \xi_{m+1} < 0\}$, $Q' = \{z : \xi_j = 0, j = 1, \ldots, m, \xi_{m+1} < 0\}$ (see Fig. 1 for $m = 1$). Q and Q' are convex cones.† Now we can write condition (0.1) in the form $\varphi(X) \cap Q = \phi$. Similarly we can rewrite conditions (1.1), (1.3), and (1.5) as: $\varphi(X) \cap Q' = \phi$, $\varphi(X) \cap \overline{Q} = \phi$, $\varphi(X) \cap \overline{Q}' = \phi$.

Now consider (0.2). This condition means that $(z_0, z) \geq 0$ when $z \in \varphi(X)$, for some vector $z_0 = (\xi_{01}, \ldots, \xi_{0m}, \xi_{0m+1}) \in \mathbf{R}^{m+1}$, where $\xi_{0m+1} > 0$, $\xi_{0j} \leq 0$ for $j = 1, \ldots, m$. It can be easily verified that $(z_0, z) < 0$ for $z \in Q$. This, in turn, is equivalent, for $\varepsilon > 0$, to the satisfaction of the inequality $(z_0, z) < 0$

*Here and in what follows, the vector inequality $\tau \geq 0$, where $\tau = \|\tau_j\|_{j=1}^m$, is to be taken component by component, i.e., as the set of inequalities $\tau_j \geq 0$, $j = 1, \ldots, m$. By $S(x, \tau)$ we mean the function $S(x, \tau) = F(x) - \sum_{j=1}^m \tau_j G_j(x)$.

†The set K in the linear space is called a convex cone [8] if $K + K \subset K$ and $\lambda K \subset K$ for $\lambda > 0$.

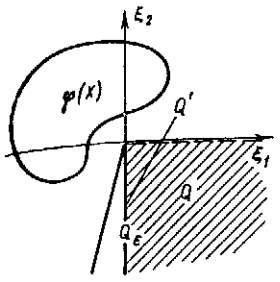

Fig. 1

when $z \in Q_\varepsilon$, where $Q_\varepsilon = \{z \in R^{m+1};\ \xi_j > \varepsilon,\ \xi_{m+1},\ j = 1, \ldots, m,\ \xi_{m+1} < 0\}$ (the set Q_ε is the "conical neighborhood" of the set Q; see Fig. 1).

Let us say that the sets A and B in the linear topological space X are separable (more precisely, linearly separable), if there exists a continuous linear functional $f \in X^*$ and a real number c such that $f(z) \geq c$ for $z \in A$ and $f(z) \leq c$ for $z \in B$. Similarly, A and B are strictly separable if these inequalities can be replaced by strict inequalities. We note that if the set A is open, then the condition $f(z) \geq c$ for $z \in A$ is equivalent to the condition $f(z) > 0$ for $z \in A$. This is because f is a continuous linear functional, and so the set $\{f(z) : z \in A\}$ is open [8].

Moreover, it is easy to see that if one of the sets A and B is a cone, then c can be set equal to zero in the definition of linear separability. If both A and B are cones, then c is necessarily zero. Condition (0.2) may now be reformulated as follows: the sets $\varphi(X)$ and Q_ε are separable for some $\varepsilon > 0$, since Q_ε is a cone. Similarly, condition (1.2) is equivalent to the separability of sets $\varphi(X)$ and Q'_ε, where $Q'_\varepsilon = \{z \in R^{m+1}: |\xi_j| < -\varepsilon\,\xi_{m+1},\ j = 1, \ldots, m,\ \xi_{m+1} < 0\}$, and condition (1.4) (respectively (1.6)) is equivalent to the strict linear separability of the sets $\varphi X)$ and Q ($\varphi(X)$ and Q').

LEMMA 1. Let A and B be sets in the linear topological space X. The separability (strict separability) of the sets A and B is equivalent to the separability (strict separability) of the sets coA and coB.* If one of the sets A and B is a cone, then the separability (strict separability) of A and B is equivalent to the separability (strict separability) of the cones $\mathcal{K}(A)$ and $\mathcal{K}(B)$.

The proof of this lemma is obvious.

Let us now give the necessary and sufficient conditions for all variants of an S-procedure to be favorable.

Proposition 1. For (0.1) to imply (0.2), it is necessary and sufficient that $\varphi(X) \cap Q = \phi$ imply that $co\varphi(X) \cap Q_\varepsilon = \phi$ for some $\varepsilon > 0$.

Proposition 2. For (1.1) to imply (1.2), it is necessary and sufficient that $\varphi(X) \cap Q' = \phi$ imply that $co\varphi(X) \cap Q'_\varepsilon = \phi$ for some $\varepsilon > 0$.

Propositions 1 and 2 follow from Lemma 1 and the separability theorem of Edel'geit, which can be formulated as follows (see, e.g., [8]).

LEMMA 2. In order that a convex set A and an open convex set B be separable in a linear topological space, it is necessary and sufficient that $A \cap B = \phi$.

Remark. The set $co\varphi(X)$ in Propositions 1 and 2 can be replaced by the set $\mathcal{K}\{\varphi(X)\}$, by Lemma 1.

The strict separability condition for convex sets is formulated rather cumbrously. Hence the theorems on the equivalence of (1.3) and (1.4), and (1.5) and (1.6) are proved under additional assumptions of the type that the set $co\varphi(X)$ be compact, and this permits us to use the following classical conditions of strict separability [8].

LEMMA 3. Let A and B be closed convex sets in a locally convex space, and let A be compact. For A and B to be strictly separable, it is necessary and sufficient that $A \cap B = \phi$.

Proposition 3. Let the set $\varphi(X)$ be compact. Then, for (1.3) to imply (1.4), it is necessary and sufficient that $\varphi(X) \cap \bar{Q} = \phi$ imply that $co\varphi(X) \cap \bar{Q} = \phi$.

Proof. Since the convex hull of a compact set in a finite dimensional space is compact, (1.4) reduces to the separability of the compact convex set $co\varphi(X)$ and the closed convex set Q. Lemma 3 may now be applied to complete the proof.

Proposition 4. Let the set $\varphi(X)$ be compact. Then, for (1.5) to imply (1.6), it is necessary and sufficient that $\varphi(X) \cap \bar{Q}' = \phi$ imply that $co\varphi(X) \cap \bar{Q}' = \phi$.

Remark 1. The set $co\varphi(X)$ can be replaced by the set $\mathcal{K}\{\varphi(X)\}$ in Propositions 3 and 4.

*We use the notation that coA denotes the convex hull of the set A, i.e., the smallest convex set containing A. By $\mathcal{K}(A)$ we mean the smallest convex cone containing A.

Remark 2. The compactness requirement on $\varphi(X)$ in Propositions 3 and 4 may be weakened by requiring instead that there is a set $X_1 \subset X$ such that $\varphi(X_1)$ is compact, and that the smallest compact cones P and P_1, containing respectively $\varphi(X)$ and $\varphi(X_1)$, coincide. In this case, we will call the cone P compactly generated. In fact, since $\overline{Q}$ and $\overline{Q}'$ are cones, the strict separability, for example, of the sets $\varphi(X)$ and $\overline{Q}$ is equivalent to the strict separability of the sets P and $\overline{Q}$, and this in turn is equivalent to the strict separability of the sets P_1 and $\overline{Q}$.

This remark is useful, for example, in the case where $F(x), G_1(x), \ldots, G_m(x)$ are positively homogeneous continuous functions of the same degree of homogeneity* on the finite-dimensional linear space X (in particular, when they are quadratic or Hermitian forms). Then X_1 may be taken as the unit sphere in the space X.

When Proposition 1 is applied, it is often difficult to check the separability of the sets $\varphi(X)$ and Q_ε. Moreover, it is usually impossible to replace the set Q_ε by the set Q, since the separating hyperplane $(z_0, z) = 0$, $z_0 = (\xi_{01}, \ldots, \xi_{0m}, \xi_{0m+1})$ must have the coefficient ξ_{0m+1} nonzero (see Paragraph 6, Example 1). But the substitution may be justified, if some additional regularity conditions are imposed on the functions $G_1(x), \ldots, G_m(x)$. We will use the so-called Slater condition, well known in mathematical programming (e.g., [6]):

$$\exists x_0 \in X : G_j(x_0) > 0, \quad j = 1, \ldots, m. \tag{1.7}$$

Moreover, when the constraints are equalities, as in (1.1) and (1.2), we must use the modified Slater condition [7]:

$$\forall \tau = \|\tau_j\|_{j=1}^m, \ \tau_j = \pm 1, \ \exists x_\tau \in X: \tau_j G_j(x_\tau) > 0, \ j = 1, \ldots, m. \tag{1.8}$$

When m = 1, for example, (1.8) means that the function $G_1(x)$ changes sign on the set X. We will call the constraints $G_1(x) \geq 0, \ldots, G_m(x) \geq 0$ ($G_1(x) = 0, \ldots, G_m(x) = 0$) regular if (1.7) (respectively (1.8)) holds. The regularity of the conditions involving any arbitrary combination of equality and inequality constraints is similarly defined.

Proposition 1'. If the constraints $G_1(x) \geq 0, \ldots, G_m(x) \geq 0$ are regular, then (0.1) and (0.2) are equivalent if and only if $\varphi(X) \cap Q = \phi$ and $co\varphi(X) \cap Q = \phi$.

Proof. It is sufficient to prove that among the hyperplanes $(z_0, z) = 0$ separating the sets $co\varphi(X)$ and Q there is a hyperplane whose coefficient ξ_{0m+1} is nonzero. So let us assume that there is no such hyperplane: suppose that any separating hyperplane has the form $\sum_{j=1}^{m} \xi_{0j}\xi_j = 0$. Since $\xi_{0j} \leq 0$, $j = 1, \ldots, m$, we have $\sum_{j=1}^{m} \xi_{0j} G_j(x) \leq 0 \ \forall x \in X$, and therefore $\sum_{j=1}^{m} \xi_{0j} G_j(x_0) \leq 0$, which contradicts (1.7). The parallel assertion about (1.1) and (1.2) is proved in exactly the same way.

Proposition 2'. If the constraints $G_1(x) = 0, \ldots, G_m(x) = 0$ are regular, then (1.1) and (1.2) are equivalent if and only if $\varphi(X) \cap Q' = \phi$ and $co\varphi(X) \cap Q' = \phi$ are equivalent.

A simple sufficiency condition for the S-procedure to be favorable in all variants is that the set $\varphi(X)$ be convex. The following assertion follows directly from Propositions 1', 2', 3, and 4.

THEOREM 1. The S-procedure for the inequality $F(x) \geq 0$ under regular constraints is favorable if $\varphi(X)$ is a convex set.

Remark 1. The theorem remains true without any changes even when the constraints are given as any arbitrary combination of equalities and inequalities.

Remark 2. It is easy to show that the regularity condition in Propositions 1 and 2, as well as in Theorem 1, may be replaced by the condition that the convex cone $\mathcal{K}\{\varphi(X)\}$ generated by the set $\varphi(X)$ be closed.

THEOREM 2. The S-procedure for the inequality $F(x) > 0$ is favorable if the set $\varphi(X)$ is convex and the set $\mathcal{K}\{\varphi(X)\}$ is compactly generated.

*The function F(x) on the linear space X is said to be positively homogeneous of degree k, if for any vector $x \in X$ and scalar λ we have $f(\lambda x) = |\lambda|^k f(x)$. In the formulation of (1.3) to (1.6) for homogeneous or positively homogeneous functions, the set X can be replaced by the set $X \setminus \{0\}$.

As an example of the application of Theorems 1 and 2, consider the question as to whether the S-procedure is favorable when $F(x)$, $G_1(x)$, ..., $G_m(x)$ are quadratic or Hermitian forms which are simultaneously diagonalizable.

THEOREM 3. The S-procedure is favorable in all variants if $F(x)$, $G_1(x)$, ..., $G_m(x)$ are quadratic (Hermitian) forms on the space $X = R^n$ (respectively $X = C^n$) which are simultaneously diagonalizable.

Proof. We shall prove that the set $\varphi(X)$ in this case is a closed polyhedral cone. In fact, after the forms $F(x)$, $G_1(x)$, ..., $G_m(x)$ have been reduced to diagonal form, they can be considered as linear functions of the new variables $y_i = |x_i|^2$, $i = 1, \ldots, n$, defined on the closed convex polyhedral cone $K = \{(y_1, \ldots, y_n) \in R^n : y_i \geq 0, i = 1, \ldots, n\}$. The shape of the cone K under the linear mapping also is a closed convex polyhedral cone. The assertion of the theorem now follows from Theorem 2 and Remark 2 of Theorem 1.

The convexity requirement on $\varphi(X)$ in Theorems 1 and 2 can be weakened. Suppose, for example, that the constraints are specified as inequalities $G_1(x) \geq 0, \ldots, G_m(x) \geq 0$. Introduce the sets $\varphi(X) = \varphi(X) - \bar{Q} = \{z \in R^{m+1} : z = z_1 - z_2, z_1 \in \varphi(X), z_2 \in \bar{Q}\}$. It is obvious that every hyperplane separating the sets Q and $\varphi(\tilde{X}) = \varphi(X) - \bar{Q}$ will separate the sets $\varphi(X)$ and Q, and conversely. It is well known [6] that if the set X is convex, if the function $F(x)$ is convex, and if the functions $G_1(x), \ldots, G_m(x)$ are concave, then the set $\varphi(x) - \bar{Q}$ is convex. Thus the following statement is true.

THEOREM 4. Let $F(x)$, $-G_j(x)$, $j = 1, \ldots, m$ be convex functions defined on the convex set X. Then (0.1) implies (0.2) if the constraints are regular, and (1.3) implies (1.4) if the set $\mathcal{K}\{\varphi(X)\}$ is compactly generated.

§2. The S-Procedure and Duality in Mathematical Programming.

Applications of the S-Procedure

Consider the following mathematical programming problem:

$$F(x) \to \inf, \; G_j(x) \geq 0, \; j = 1, \ldots, m, \; x \in X. \tag{2.1}$$

The functions $F(x)$, $G_1(x)$, ..., $G_m(x)$ are real-valued functions defined on some set X. For each $\tau = \|\tau_j\|_{j=1}^m$ we define the function

$$\psi(\tau) = \inf_{x \in X} \left[F(x) - \sum_{j=1}^{m} \tau_j G_j(x) \right] \tag{2.2}$$

and consider, as in [6] for example, the problem dual to (2.1)*

$$\psi(\tau) \to \sup, \; \tau_j \geq 0, \; j = 1, \ldots, m. \tag{2.3}$$

Denote by v the value of the lower bound in problem (2.1), and by $\tilde{v}$ the value of the upper bound in problem (2.3). It is easy to see that we always have $\tilde{v} \leq v$. This means that we have the duality relation in problems (2.1)-(2.3), if $\tilde{v} = v$, i.e.,

$$\inf_{\substack{G_j(x) \geq 0 \\ j=1,\ldots,m}} F(x) = \sup_{\tau \geq 0} \inf_{x \in X} \left[F(x) - \sum_{j=1}^{m} \tau_j G_j(x) \right]. \tag{2.4}$$

Similar definitions are given for the problem with any arbitrary combination of equalities and inequalities as constraints. For example, when the constraints are equalities, $G_1(x) = 0, \ldots, G_m(x) = 0$, the duality relation has the form:

$$\inf_{\substack{G_j(x) = 0 \\ j=1,\ldots,m}} F(x) = \sup_{\tau} \inf_{x \in X} \left[F(x) - \sum_{j=1}^{m} \tau_j G_j(x) \right]. \tag{2.5}$$

The following simple assertions state the link between the duality relation and the advantage of the S-procedure.

*It is possible that $\psi(\tau) = -\infty$ for some or even for all τ. If $\psi(\tau) \equiv -\infty$, then we naturally take $\sup_\tau \psi(\tau) = -\infty$.

THEOREM 5. The S-procedure for the inequality $F(x) - \alpha \geq 0$ with constraints $G_1(x) \geq 0, \ldots, G_m(x) \geq 0$ is favorable for any $\alpha \in R^1$, if the duality relation holds for the problems (2.1)-(2.3) holds and the upper bound $\tilde{v}$ is attained. Conversely the duality relation in the problems (2.1)-(2.3) holds and the upper bound is attained if (0.1) implies (0.2) for the functions $F(x) - \alpha$, $G_1(x), \ldots, G_m(x)$, where α is any real number.

Proof. Suppose (0.1) holds. Then the lower bound $v \geq \alpha$ in problem (2.1). By the duality relation, $\tilde{v} \geq 0$. So there are numbers $\tau_j \geq 0$, $j = 1, \ldots, m$ such that $\psi(\tau) \geq \alpha$, i.e., (0.2) holds. Thus the first assertion of the theorem is proved. By the definition of v, $F(x) - v \geq 0$ for $G_j(x) \geq 0$, $j = 1, \ldots, m$, i.e., the functions $F(x) - v$, $G_1(x), \ldots, G_m(x)$ satisfy (0.1). Therefore they satisfy (0.2), i.e., there are numbers $\tau_i \geq 0$ such that $\psi(\tau) \geq v$. This means that $\tilde{v} \geq v$, which was to be proved.

THEOREM 6. The S-procedure for the inequality $F(x) > \alpha$ with the constraints $G_1(x) \geq 0, \ldots, G_m(x) \geq 0$ is favorable for any $\alpha \in R^1$ if the duality relation holds in problems (2.1)-(2.3) and the lower bound in the direct problem is attained. Conversely, the duality relation holds in problems (2.1)-(2.3), if (1.3) implies (1.4) for the functions $F(x) - \alpha$, $G_1(x), \ldots, G_m(x)$, where α is any real number.

Proof. Suppose (1.3) holds. Then in problem (2.1), the lower bound $v > 0$ (since it is attained), and therefore $v > 0$. Therefore there are numbers $\tau_j \geq 0$, $j = 1, \ldots, m$ such that $\psi(\tau) > 0$, i.e., (1.4) is fulfilled. Conversely, for a given v, for any $\varepsilon > 0$ the functions $F(x) - v + \varepsilon$, $G_1(x), \ldots, G_m(x)$ satisfy (1.3). Therefore there is a vector $\tau_\varepsilon \geq 0$ such that $\psi(\tau_\varepsilon) \geq v - \varepsilon$, i.e., $\tilde{v} \geq v$, and so $\tilde{v} = v$.

Thus, any pair of extremal problems for which the duality relation holds and the lower bound of the direct problem is attained furnishes an example of the equivalence of (1.3) and (1.4). The attainability of the upper bound in the dual problem means that (0.1) and (0.2) are equivalent. Therefore theorems guaranteeing the attainability of the extremal values in the direct and dual problems would be of some interest. Several assertions of this kind are cited in [6]. If $F(x)$, $G_1(x), \ldots, G_m(x)$ are positively homogeneous functions on a linear space (e.g., quadratic of Hermitian forms), then Theorem 5 can be made more precise.

THEOREM 7. Let the $F(x)$, $G_1(x), \ldots, G_m(x)$ be positively homogeneous functions of the same degree of homogeneity, defined on the linear space X. Then (0.1) implies (0.2) if and only if the duality relation holds for the problems (2.1)-(2.3).

Proof. The set $\varphi(X)$ is a cone for the positively homogeneous functions. Therefore in problem (2.1), either $v = 0$ and (0.1) holds, or $v = -\infty$ and (0.1) does not hold. Similarly, if (0.2) holds, it means that $\tilde{v} = 0$ in the corresponding dual problem. Therefore (0.1) and (0.2) are equivalent if and only if $\tilde{v} = v$.

Let us now go on to give examples of the application of the S-procedure. It is known that every quadratic or Hermitian form is uniquely (up to a scalar multiple) defined by its set of zeros (for Hermitian forms, see [4] for example). For certain forms, a similar assertion is related to the advantage of the S-procedure.

THEOREM 8. Let the quadratic (Hermitian) forms $G_1(x), \ldots, G_m(x)$ on the linear space X satisfy (1.8) so that (1.1) implies (1.2) for any quadratic (Hermitian) form $F(x)$. Then every quadratic (Hermitian) form $F(x)$ whose set of zeros contains the set of common zeros of the forms $G_1(x), \ldots, G_m(x)$ is representable as a linear combination of these forms.

Proof. By hypothesis, both the forms $F(x)$, $G_1(x), \ldots, G_m(x)$ and the forms $F(x)$ $G_1(x), \ldots, G_m(x)$ satisfy (1.1). Applying the S-procedure, we get that for all $x \in X$ the inequalities $F(x) - \sum_{j=1}^{m} \tau_j G_j(x) \geq 0$, $-F(x) - \sum_{j=1}^{m} \tau_j' G_j(x) \geq 0$ are satisfied. If we add these inequalities we obtain $\tau_j' = -\tau_j$, $j = 1, \ldots, m$, i.e., $F(x) \equiv \sum_{j=1}^{m} \tau_j G_j(x)$, as was required to prove.

COROLLARY. Let $G_1(x)$ and $G_2(x)$ be Hermitian forms on the complex linear space X, and let them satisfy (1.8), and let the set of zeros of the Hermitian form $F(x)$ contain the set of common zeros of the forms $G_1(x)$ and $G_2(x)$. Then the form $F(x)$ is a linear combination of the forms $G_1(x)$ and $G_2(x)$.

This comes from Theorem 8 and the results of paragraph 3 of this article (Theorem 17). We remark that the analogous statement for a real space X is false (see Example 4 of Paragraph 6).

The S-procedure can be used to investigate problems on the simultaneous reduction of quadratic or Hermitian forms to diagonal form.

THEOREM 9. Suppose $X = \mathbf{R}^n$ ($X = \mathbf{C}^n$) and that the quadratic (Hermitian) forms $F(x)$ and $G(x)$ do not have common zeros except for $x = 0$. Then there is a nonsingular linear transformation on X which simultaneously reduces $F(x)$ and $G(x)$ to diagonal form.

Proof. By hypothesis, the form $F(x)$ does not change sign on the set of zeros of $G(x)$, i.e., either the forms $F(x)$, $G(x)$ or the forms $-F(x)$, $G(x)$ satisfy (1.3). It is known [2, 4] that if $m = 1$ and $F(x)$, $G(x)$ are quadratic (Hermitian) forms, then (1.3) implies (1.4); another proof of this is given below. Suppose for definiteness that (1.3) holds for $F(x)$, $G(x)$. Then there is a number τ such that the form $S(x) = F(x) - \tau G(x)$ is positive definite. By a well known theorem of linear algebra, the forms $S(x)$ and $G(x)$ can be simultaneously diagonalized. But after this reduction, the form $F(x) = S(x) + \tau G(x)$ is also in diagonal form, as was to be proved.

§3. The Space of Forms and the Convexity of $\varphi(X)$

In the next two paragraphs we consider the S-procedure for the case where $F(x)$, $G_1(x)$, ..., $G_m(x)$ are quadratic (Hermitian) forms on a real (complex) linear space X.

Denote by $\mathcal{F}(\mathbf{R}^n)$ the set of all quadratic forms in n real variables, and by $\mathcal{F}(\mathbf{C}^n)$ the set of all Hermitian forms in n complex variables. It is obvious that $\mathcal{F}(\mathbf{R}^n)$ and $\mathcal{F}(\mathbf{C}^n)$ are real linear spaces; besides, $\dim \mathcal{F}(\mathbf{R}^n) = n(n+1)/2$, $\dim \mathcal{F}(\mathbf{C}^n) = n^2$. In what follows, we will identify the form $F(x)$ with its matrix F with respect to some fixed basis. We transform the spaces $\mathcal{F}(\mathbf{R}^n)$ and $\mathcal{F}(\mathbf{C}^n)$ into Euclidean spaces by defining in them the scalar product $\langle F, G\rangle$ by the formula† $\langle F, G\rangle = \operatorname{Sp} FG$. We note that an orthogonal (unitary) change of variables will induce an orthogonal transformation in the space $\mathcal{F}(\mathbf{R}^n)$ (in the space $\mathcal{F}(\mathbf{C}^n)$) through the relation $\operatorname{Sp} T^*FTT^*GT = \operatorname{Sp} FG$.

All the arguments which follow hold for both the real and the complex cases, unless the contrary is stipulated. Both the spaces $\mathcal{F}(\mathbf{R}^n)$ and $\mathcal{F}(\mathbf{C}^n)$ will be denoted by the same symbol $\mathcal{F}_n$. Consider the convex cone K of the positive definite forms and its closure in the space $\mathcal{F}_n$, the cone of nonnegative forms $\bar{K}$. It is known that the extremal generators [8] of the cone $\bar{K}$ are the forms of rank 1 with matrices of the form $P_x = xx$. The following properties are easily verified:

1. $\langle F, P_x\rangle = (Fx, x)$.

2. $\langle P_x, P_y\rangle = (x, y)^2$.

3. $F \in \bar{K} \Leftrightarrow \forall G \in \bar{K}\ \langle F, G\rangle \geqslant 0$.

3a. $F \in \bar{K} \setminus \{0\} \Leftrightarrow \forall G \in K\ \langle F, G\rangle > 0$.

3b. $F \in K \Leftrightarrow \forall G \in \bar{K} \setminus \{0\}\ \langle F, G\rangle > 0$.

Property 3 states that the cone $\bar{K}$ is selfconjugate.‡ It is easy to state in terms of the space $\mathcal{F}_n$ the convexity criterion of the set $\varphi(X)$ introduced in Paragraph 1.

Suppose that $F_1(x), \ldots, F_k(x)$ are quadratic (Hermitian) forms on the space $\mathbf{R}^n$ ($\mathbf{C}^n$) and that X is a subset of $\mathbf{R}^n$ (respectively $\mathbf{C}^n$). Suppose that $\psi : X \to \mathbf{R}^k$ is the map defined by the formula $\varphi(x) = (F_1(x), \ldots, F_k(x)) \in \mathbf{R}^k$. Denote by $\tilde{X}$ the set $\{P_x : x \in X\}$ and by L the linear hull of the forms $F_1, \ldots, F_k$.

THEOREM 10. The set $\varphi(X)$ is convex if and only if the set $\operatorname{Pr}_L \tilde{X}$ is convex, where Pr_L is the orthogonal projection onto the subspace $L \subset \mathcal{F}_n$.

Proof. Suppose initially that the forms $F_1, \ldots, F_k$ are orthonormal, i.e., $\langle F_i, F_j\rangle = \delta_{ij}$. Then if $x \in X$, the form $\operatorname{Pr}_L(P_x) \in L$ has in the basis $F_1, \ldots, F_k$ the coordinates $\langle P_x, F_j\rangle = F_j(x)$, $j = 1, \ldots, k$, i.e., the sets $\operatorname{Pr}_L \tilde{X}$ and $\psi(X)$ simply coincide (up to an isomorphism of the Euclidean spaces L and $\mathbf{R}^k$).

Suppose now that the forms $F_1, \ldots, F_k$ are linearly independent. This case is reduced to the first case by a nonsingular linear transformation on the space L which orthonormalizes the forms $F_1, \ldots, F_k$. Moreover, the convexity of the set $\varphi(X)$ is preserved.

In the general case, let $F_{i1}, \ldots, F_{il}$ be a basis for the subspace L. Since the remaining forms are linearly generated by $F_{i1}, \ldots, F_{il}$, the convexity of $\varphi(X)$ is equivalent to the convexity of the set

†We denote the trace of the matrix A by Sp A. By A^* we denote the transposed complex conjugate of the matrix A; if A is real, A^* will be its transpose.

‡The cone $M^* = \{F : \langle F, G\rangle \geq 0\ \forall\ G \in M\}$ will be called the cone conjugate to the cone M.

$\{F_{i_1}(x), \ldots, F_{i_l}(x), x \in X\}$, which in turn is equivalent to the convexity of the projection of the set $\tilde{X}$ on $\{F : F = \sum_{s=1} \alpha_s F_{i_s}, \alpha_s \in R^1\} = L$, as was to be proved.

Remark. The whole space R^n (C^n) or the unit sphere in it is often taken as the set X.

COROLLARY. The set $\varphi(X)$ is convex if the inverse image of all possible straight lines from the projection of the forms of the type P_x, $x \in X$ on the subspace L are linearly related (these preimages are cross sections of the set of forms of the type P_x, $x \in X$ with subspaces of dimension $n(n+1)/(2-m)$ in the real case and $n^2 - m$ in the complex case).

THEOREM 11. (Hausdorff [9], and see also [10].) For all Hermitian forms $F_1(x)$, $F_2(x)$, the image of the unit sphere in the complex pre-Hilbert space H under the map $x \to (F_1(x), F_2(x)) \in R^2$ is a convex set.

Proof. We note that it suffices to prove the theorem under the assumption that H is two dimensional. In fact, let $z_1, z_2 \in \varphi(X)$ in the general case, and then there are points $x_j \in H$, $j = 1, 2$, such that $\|x_j\| = 1$ and $z_j = \varphi(x_j)$, $j = 1, 2$. Without loss in generality, we may assume that x_1 and x_2 are linearly independent. Consider the restriction φ_1 of the map φ to the unit sphere X_1 in the two-dimensional space spanned by x_1 and x_2. If the set $\varphi_1(X_1)$ is convex, i.e., if it contains the whole interval between the points z_1 and z_2, then the set $\varphi(X)$ also is convex, since $\varphi(X) \supset \varphi_1(X_1)$.

Consider now the space $\mathcal{F}_2$ of Hermitian forms on the two-dimensional complex space C_2 ($\dim \mathcal{F}_2 = 4$). The matrix of the form $F \in \mathcal{F}_2$ will be written in the form

$$F = \begin{pmatrix} a_1 + a_2 & b_1 + ib_2 \\ b_1 - ib_2 & a_1 - a_2 \end{pmatrix},$$

where a_1, a_2, b_1, b_2 are real numbers. The coefficients of the forms of the type P_x are numbers satisfying the equation $a_1^2 - a_2^2 = b_1^2 + b_2^2$, while the condition $\|x\| = 1$ is equivalent to $2a_1 = 1$. Thus the set of forms of the type P_x, $\|x\| = 1$ is isometric to a two-dimensional sphere lying in a three-dimensional subspace. The preimage of a straight line from the projection onto the two-dimensional subspace of R^4 is a three-dimensional plane. But two three-dimensional planes in R^4 must intersect along a two-dimensional plane, and the intersection of a two-dimensional plane with a sphere is always linearly connected. An application of the Corollary to Theorem 10 completes the proof.

THEOREM 12 [7]. For any three Hermitian forms $F_1(x)$, $F_2(x)$, $F_3(x)$ on the complex linear space X, the set $\varphi(X)$ is convex.

Proof. Just as in the proof of Theorem 11, it suffices to consider the case of forms on the two-dimensional space C^2. Since the cone of nonnegative forms in $\mathcal{F}_2$ is strictly convex, because every one of its generators is extremal, the set $\{P_x, x \in C^2\}$ is the boundary of the convex set. Therefore the intersection of the set $\{P_x, x \in C^2\}$ with a two-dimensional plane is the boundary of some convex set in the plane and therefore is always linearly connected. An application of the Corollary to Theorem 10 completes the proof.

THEOREM 13. (Dines [11].) For any two quadratic forms $F_1(x)$, $F_2(x)$ on the real linear space X, the set $\varphi(X)$ is convex.

Proof. As in Theorem 11, it is sufficient to consider forms on the two-dimensional space R^2. In the three dimensional space $\mathcal{F}_2$, the set of forms of the type P_x, $x \in R^2$ is the surface of a right circular cone. The intersection of the surface of the cone with any two-dimensional plane is the boundary of a convex set in the plane, and is therefore linearly connected. An application of the Corollary to Theorem 10 now completes the proof.

Theorems 11-13 enable us to establish the advantage of the S-procedure for m = 1 in both the real and the complex case, and for m = 2 in the complex case.

THEOREM 14. Let m = 1 and let $G_1(x)$ be a quadratic (Hermitian) form on the linear space X, where the form is not nonpositive. Then (0.1) implies (0.2) for any quadratic (Hermitian) form $F(x)$.

THEOREM 15. Let m = 1 and let $G_1(x)$ be an indefinite quadratic (Hermitian) form on the linear space X. Then (1.1) implies (1.2) for any quadratic (Hermitian) form $F(x)$.

THEOREM 16. Let m = 1 and let the real (complex) linear space X be finite dimensional. Then the S-procedure for the inequality $F(x) > 0$ under the constraint $G(x) \geq 0$ or $G(x) = 0$ is favorable for any quadratic (Hermitian) forms $F(x)$, $G(x)$.

THEOREM 17. Let $G_1(x)$, $G_2(x)$ be Hermitian forms on the complex linear space X. Then the S-procedure for the inequality $F(x) \geq 0$ with the regular constraints $G_1(x) \geq 0$, $G_2(x) \geq 0$, or $G_1(x) \geq 0$, $G_2(x) \neq 0$, or $G_1(x) = 0$, $G_2(x) = 0$ is favorable for any Hermitian form $F(x)$ on X.

THEOREM 18. Let $m = 2$ and let $G_1(x)$, $G_2(x)$ be Hermitian forms on the finite dimensional complex space X. Then the S-procedure for the inequality $F(x) > 0$ is favorable for any Hermitian form $F(x)$.

Theorems 14-18 are direct consequences of Theorems 1, 2, and 11-13. Propositions resembling Theorems 14-16 have been obtained by several authors independently [2, 4, 11, 12]. Theorem 17 is proved by a different method in [7].

It follows from Theorem 10 that the convexity of $\varphi(X)$ under the map $x \to (G_1(x), \ldots, G_m(x), F(x))$ is not determined by the forms $F(x)$, $G_1(x)$, $\ldots$, $G_m(x)$ themselves, but by the subspace they span in $\mathcal{F}_n$. Thus Theorems 17 and 18 can be strengthened.

THEOREM 17'. Let the Hermitian forms $G_1(x)$, $\ldots$, $G_m(x)$ on the complex linear space X belong to some three-dimensional subspace $L \subset \mathcal{F}(X)$ and satisfy (1.7) (respectively (1.8)). Then (0.1) implies (0.2) (respectively (1.1) implies (1.2)) for any Hermitian form $F(x)$ from the same subspace L.

THEOREM 18'. Let the Hermitian forms $G_1(x)$, $\ldots$, $G_m(x)$ on the space C^n belong to some three-dimensional subspace $L \subset \mathcal{F}(C^n)$. Then (1.3) implies (1.4) (or (1.5) implies (1.6)) for any Hermitian form $F(x)$ from the same subspace L.

COROLLARY. The S-procedure is favorable if $F(x)$, $G_1(x)$, $\ldots$, $G_m(x)$ are real Hermitian forms on C^2 [4].

To prove the corollary, one need only apply Theorems 17' and 18' to the three-dimensional subspace of real Hermitian forms on C^2.

From the theorems proved in this paragraph we get the duality theorems for a series of nonconvex problems in quadratic programming, as well as other corollaries on the advantage of the S-procedure, as in Paragraph 2. We note that the proofs of Theorems 11-13 make it clear that, when $k > 2$ in the real case, and when $k > 3$ in the complex case, the set of points $z \in R^k$ of the form $z = (F_1(x), \ldots, F_k(x))$, $x \in X$ is usually not convex. In fact, let X be a two-dimensional linear space. Then both in the real and the complex case, the set $\varphi(X)$ is the image of the surface of some strictly convex cone under a nonsingular linear transformation, and therefore it is not convex, if there are at least three (four) linearly independent quadratic (Hermitian) forms among the set $F_1(x)$, $\ldots$, $F_k(x)$.

§4. The Space of Forms and the Advantage Condition for the S-Procedure

We now try to put into the language of the space of forms $\mathcal{F}_n$ the conditions which express the formulation of the S-procedure. Moreover, we limit ourselves to the case of the inequality $F(x) \geq 0$ with regular constraints $G_1(x) \geq 0$, $\ldots$, $G_m(x) \geq 0$ or $G_1(x) = 0$, $\ldots$, $G_m(x) = 0$. If nothing is stated to the contrary, the arguments apply to both the real and the complex case.

Conditions (0.1) and (1.1) can be rewritten in the following form:

$$\langle F, P_x \rangle \geqslant 0 \quad \text{for} \quad \langle G_j, P_x \rangle \geqslant 0, \quad j = 1, \ldots, m, \tag{4.1}$$

$$\langle F, P_x \rangle \geqslant 0 \quad \text{for} \quad \langle G_j, P_x \rangle = 0, \quad j = 1, \ldots, m. \tag{4.2}$$

Now consider (0.2) and (1.2). Denote by R the set $\{A \in \mathcal{F}_n: A = \tau_0 F - \sum_{j=1}^{m} \tau_j G_j, \tau_0 > 0, \tau_j \geq 0, j = 1, \ldots, m\}$. The set R is a convex cone. Condition (0.2) can be rewritten in the form $R \cap \bar{K} \neq \phi$. This in turn can be rewritten, by (1.7), as

$$\bar{R} \cap \bar{K} \neq \{0\}, \tag{4.3}$$

where $\bar{R}$ is the closure of the cone R: $\bar{R} = \{A \in \mathcal{F}_n: A = \tau_0 F - \sum_{j=1}^{m} \tau_j G_j, \tau_j \geq 0, j = 0, 1, \ldots, m\}$.

LEMMA 4. If K_1 and K_2 are closed convex cones in the finite dimensional linear space L, then

$$K_1 \cap K_2 \neq \{0\} \Leftrightarrow \operatorname{Int} K_1^* \cap \operatorname{Int}(-K_2^*) = \phi. \tag{4.4}$$

Proof. The condition $K_1 \cap K_2 \neq \{0\}$ is equivalent to the condition $(K_1 \cap K_2)^* \neq L$. But for the closed cones K_1 and K_2, the following equality† holds: $(K_1 \cap K_2)^* = K_1^* + K_2^*$, and the condition $K_1^* + K_2^* \neq L$ in turn is equivalent to the fact that the cone $K_1^* + K_2^*$ can be separated from zero by a hyperplane, i.e., there exists a nonzero linear functional $f \in L^*$, such that $f(x) \geq 0$ for $x \in K_1^*$, and $f(x) \leq 0$ for $x \in -K_2^*$. But the necessary and sufficient condition for the separability of the cones K_1^* and $-K_2^*$ is the condition $\operatorname{Int} K_1^* \cap \operatorname{Int} \cdot (-K_2^*) = \phi$ ([13], p. 308), as was required to be proved.

Now apply Lemma 4 to condition (4.3). We get that (4.3) is equivalent to the condition that there is no nonnegative form P such that $P \in \operatorname{Int}(-R^*) = \{A : \langle A, F\rangle < 0, \langle A, G_j\rangle \geq 0, j = 1, \ldots, m\}$, i.e., for any nonnegative form P such that $\langle P, G_j\rangle > 0$, $j = 1, \ldots, m$, we have $\langle P, F\rangle \geq 0$. But on any set M, we have $\inf_{P \in M} \langle P, F\rangle = \inf_{P \in M} \langle P, F\rangle$. Finally, (0.2) can be written in the form

$$\langle F, P\rangle \geqslant 0 \quad \text{for} \quad P \in \overline{K} \text{ and } \langle G_j, P\rangle \geqslant 0, \quad j = 1, \ldots, m. \tag{4.5}$$

Similarly (1.2) can be written as

$$\langle F, P\rangle \geqslant 0 \quad \text{for} \quad P \in \overline{K} \text{ and } \langle G_j, P\rangle = 0, \quad j = 1, \ldots, m. \tag{4.6}$$

Thus the S-procedure problem has been reduced to the following. Let the linear functional $F(P) = \langle F, P\rangle$ be given on the set of forms $\mathcal{F}_n$. It is known that $F(P) \geq 0$ on a set of forms of rank 1, where the forms do not make an obtuse angle with any of the forms $G_1, \ldots, G_m$. We are required to find the condition that $F(P) \geq 0$ for any nonnegative form P not making an obtuse angle with any of the forms $G_1, \ldots, G_m$.

To apply this theory in automatic control theory, as well as in other situations, it is significant that the form F(x) involved in the hypotheses is not known beforehand. So there is a lot of interest in the problem of finding forms $G_1, \ldots, G_m$ such that (0.1) and (0.2) (or that (1.1) and (1.2)) are equivalent for any form $F \in \mathcal{F}_n$.

THEOREM 19. For (0.1) and (0.2) to be equivalent for any form $F \in \mathcal{F}_n$, it is necessary and sufficient that the forms $G_1, \ldots, G_m$ satisfy the following condition (E):

(E) The extremal generators of the set $M = \overline{K} \cap \{A : \langle G_j, A\rangle \geq 0, j = 1, \ldots, m\}$ vanish for the forms of rank 1 in this set.

Proof. Suppose any nonnegative form P which does not make an obtuse angle with any of the forms G_j be represented in the form $P = \sum_{s=1}^{r} \alpha_s P_{x_s}$, $\alpha_s \geq 0$, $P_{x_s} \in M$, and let it satisfy condition (0.1), i.e., $\langle F, P_x\rangle \geq 0$ for any $P_x \in M$. Then $\langle F, P\rangle = \sum_{s=1}^{r} \alpha_s \langle F, P_{x_s}\rangle \geq 0$.

Conversely, suppose that some form $P_0 \in M$ cannot be decomposed into forms of rank 1 from M, i.e., P_0 does not belong to the convex cone M_0 (easily seen to be closed), spanned by the forms of rank 1 from M. Then the separability theorem implies that there is a linear functional $F(A) = \langle F, A\rangle$ such that $\langle F, A\rangle \geq 0$ for $A \in M_0$ and $\langle F, P_0\rangle < 0$, i.e., (0.1) holds for F, while (0.2) does not, which contradicts the hypothesis of the theorem.

The case where the constraints are equalities is done in a completely parallel way.

THEOREM 20. For (1.1) and (1.2) to be equivalent for any form $F \in \mathcal{F}_n$, it is necessary and sufficient that the forms $G_1, \ldots, G_m$ satisfy the following condition (E'):

(E') The extremal generators of the set $M' = \overline{K} \cap \{A : \langle G_j, A\rangle = 0, j = 1, \ldots, m\}$ vanish for the forms of rank 1 in this set.

The condition (E') is weaker than the condition (E), and verification of the condition (E) can be reduced to several verifications of (E').

THEOREM 21. If the forms $G_1, \ldots, G_m$ satisfy (E), then they also satisfy (E').

---

†If K_1 and K_2 are convex cones, then $K_1 + K_2$ denotes the convex cone $\{x \in L : x = x_1 + x_2, x_1 \in K_1, x_2 \in K_2\}$. The cone is considered to be endowed with the topology induced from the smallest subspace of L containing the cone.

Proof. Suppose the form $P \in M'$. Because $M' \subset M$, we get $P \in M$. By (E) the form P can be represented in the form $P = \sum_{s=1}^{r} \alpha_s P_{x_s}$, where $P_{x_s} \in M$. It suffices to prove that if $\alpha_s > 0$, then $P_{x_s} \in M'$. Suppose this is false, i.e., $P_{x_s} \in M \setminus M'$. This means that $\langle P_{x_s}, G_j \rangle \geq 0$ for $j = 1, \ldots, m$ and for some index j_0, $\langle P_{x_s}, G_j^0 \rangle > 0$ holds. Thus $\langle P, G_{j_0} \rangle = \sum_{s=1}^{r} \alpha_s > P_{x_s}, G_{j_0} \rangle > 0$, i.e., $P \notin M'$. The contradiction proves the theorem.

THEOREM 22. If the condition (E') holds for forms $G_j - \varepsilon_j Q_j$ for some $Q_j \in K$ and for all $\varepsilon_j \geq 0$, $j = 1, \ldots, m$, then (E) holds for the forms G_j, $j = 1, \ldots, m$.

Proof. We must show that any form $P \in \bar{K}$, such that $\langle P, G_j \rangle \geq 0$, $j = 1, \ldots, m$, is representable in the form $P = \sum_{s=1}^{r} \alpha_s P_{x_s}$, $\alpha_s \geq 0$, $\langle G_j, P_{x_s} \rangle \geq 0$. It suffices to show this for $P \in K$, because the set of forms of the type P_x (satisfying the condition $\langle G_j, P_{x_s} \rangle \geq 0$), as well as its convex hull, is a closed convex cone. If $P \in K$, then $\langle P, Q_j \rangle > 0$ and $\langle G_j - \varepsilon_j Q_j, P \rangle = 0$ for $\varepsilon_j = \langle G_j, P \rangle / \langle Q_j, P \rangle \geq 0$. By (E'), $P = \sum_{s=1}^{r} \alpha_s P_{x_s}$, where $\langle G_j - \varepsilon_j Q_j, P_{x_s} \rangle = 0$. Therefore $\langle G_j, P_{x_s} \rangle \geq 0$, as was required to be proved.

As an example of an application of the results obtained above, let us consider the question of the advantage of the S-procedure for quadratic (Hermitian) forms acting on a two-dimensional space.

THEOREM 23. Let $G_1(x), \ldots, G_m(x)$ be quadratic forms in two real variables. When $m = 2$, (0.1) and (0.2) are equivalent for any quadratic forms $F(x) \in \mathcal{F}_2$ if and only if the following inequality holds:

$$\begin{vmatrix} a_1 & b_1 \\ b_1 & b_2 \end{vmatrix} \cdot \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} \leqslant \begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}^2, \tag{4.7}$$

where $G_j(x) = x^* \begin{pmatrix} a_j & b_j \\ b_j & c_j \end{pmatrix} x$, $j = 1, 2$, $x = (x_1, x_2)$. When $m \geq 3$, the corresponding condition is that (4.7) hold for each pair of the forms $G_1(x), \ldots, G_m(x)$.

Proof. The set $M = \{A : \langle G_j, A \rangle \geq 0, j = 1, 2\}$ is the intersection of two subspaces, a dihedral angle in the three-dimensional space $\mathcal{F}_2$. All the boundary points of the cone $\bar{K} \subset \mathcal{F}_2$ are extremal. Therefore the intersection of the cone $\bar{K}$ with the dihedral angle is the convex hull of its two forms of rank 1 if and only if the edge of the angle does not intersect the interior of the cone. But the edge of the angle is the straight line defined by the vector product of the forms G_1 and G_2 considered as three-dimensional vectors. Thus (E) is equivalent to the satisfaction of (4.7). When $m \geq 3$, the set M is a polyhedral angle, and (E) is then equivalent to the condition that none of its edges belongs to the cone K, i.e., (4.7) holds for each pair of forms G_{j_1}, G_{j_2}, $1 \leq j_1 < j_2 \leq m$.†

Remark. Condition (4.7) is easily formulated in terms of the set of zeros of the forms G_1 and G_2. This means that the pair of lines $G_1(x) = 0$ and the pair of lines $G_2(x) = 0$ occur sequentially in the plane, and not alternately. It is easy to see that when such a distribution of zeros occurs, $F(x) \geq 0$ for $G_j(x) \geq 0$, $j = 1, 2$ implies that either $F(x) \geq 0$ for $G_1(x) \geq 0$, or $F(x) \geq 0$ for $G_2(x) \geq 0$, and we can take either $\tau_1 = 0$ or $\tau_2 = 0$ in (0.2), i.e., the two constraints are reduced in essence to one.

The following is proved in the same way.

THEOREM 24. Let $G_1(x), \ldots, G_m(x)$ be Hermitian forms on the space $\mathbf{C}^2$. When $m = 3$, (0.1) and (0.2) are equivalent for any Hermitian form $F(x)$ if and only if the following inequality holds:

$$\begin{vmatrix} a_1 & b_1' & b_1'' \\ a_2 & b_2' & b_2'' \\ a_3 & b_3' & b_3'' \end{vmatrix} \cdot \begin{vmatrix} b_1' & b_1'' & c_1 \\ b_2' & b_2'' & c_2 \\ b_3' & b_3'' & c_3 \end{vmatrix} + \begin{vmatrix} a_1 & b_1' & c_1 \\ a_2 & b_2' & c_2 \\ a_3 & b_3' & c_3 \end{vmatrix}^2 + \begin{vmatrix} a_1 & b_1'' & c_1 \\ a_2 & b_2'' & c_2 \\ a_3 & b_3'' & c_3 \end{vmatrix}^2 \geqslant 0, \tag{4.8}$$

where

$$G_j(x) = x^* \begin{pmatrix} a_j & b_j' + ib_j'' \\ b_j' - ib_j'' & c_j \end{pmatrix} x, \quad j = 1, 2, 3; \ x \in \mathbf{C}^2.$$

When $m > 3$, the corresponding condition is that the inequality holds for any three of the forms $G_1(x), \ldots, G_m(x)$.

†It suffices to check the inequality for pairs of forms orthogonal to the edges of the cone M dual to the convex cone spanned by the forms $G_1, \ldots, G_m$.

COROLLARY. If $G_1(x), \ldots, G_m(x)$ are real Hermitian forms on $\mathbf{C}^2$ satisfying (1.7), then (0.1) implies (0.2) for any (not necessarily real) form $F(x)$.

It is easy to obtain the following theorem from the formulations of (0.1) and (0.2) in the form (4.1), (4.5); this theorem completes our investigation of the problem for quadratic forms on two-dimensional spaces.

THEOREM 25. Suppose that (4.7) does not hold for the quadratic forms $G_1(x)$, $G_2(x)$ on $\mathbf{R}^2$. Let G_0 be a positive form orthogonal to G_1 and G_2: it exists by Theorem 23. Then for (0.1) to imply (0.2) for a given form $F(x)$, it is necessary and sufficient that the inequality $\langle F, G_0\rangle \geq 0$ hold, which is equivalent to the inequality

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \cdot \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a & b & c \end{vmatrix} \geqslant 0, \tag{4.9}$$

where

$$G_j(x) = x^* \begin{pmatrix} a_j & b_j \\ b_j & c_j \end{pmatrix} x, \quad j = 1, 2; \quad F = x^* \begin{pmatrix} a & b \\ b & c \end{pmatrix} x, \quad x \in R^2.$$

To conclude this paragraph, we note that the question of an effective verification procedure for (E) and (E') remains open.

§5. Several Generalizations

Let us consider one of the possible generalizations of the construction studied in Paragraphs 3 and 4. Denote by $\mathcal{F}_p$ the linear hull of the set of p linearly independent real functions $\{\alpha_s(\cdot)\}_{s=1}^p$ selected from some set X_p. Each function $F \in \mathcal{F}_p$ can be uniquely represented in the form $F(x) = \sum_{s=1}^{p} f_s \alpha_s(x)$, and therefore the real linear space $\mathcal{F}_p$ is canonically isomorphic to the Euclidean space of sets of coefficients $\{f_s\}_{s=1}^p$. This isomorphism induces on $\mathcal{F}_p$, in a natural way, the structure of a Euclidean space with the inner product $\langle F, G\rangle = \sum_{s=1}^{p} f_s g_s$. We will no longer differentiate between the function F and the set $\{f_s\}_{s=1}^p$ of its coefficients.

For each point $y \in X$, consider the function δ_y defined by the set of coefficients $\{\alpha_s(y)\}_{s=1}^p$. In other words, the function δ_y is specified by the relation $\delta_y(x) = \sum_{s=1}^{p} \alpha_s(y)\alpha_s(x)$. Then for any function $F \in \mathcal{F}_p$ we have

$$F(y) = \langle F, \delta_y\rangle. \tag{5.1}$$

Consider the convex cone K in the space $\mathcal{F}_p$, where K is spanned by the set of functions of the form δ_y, $y \in X$, i.e., $K = \mathcal{K}\{\delta_y, y \in X\}$, and as well the cone K* of nonnegative functions $K^* = \{F \in \mathcal{F}_p : F(x) \geq 0 \ \forall x \in X\}$. Relation (5.1) shows that K* is the cone dual to the cone K. If the cone K is closed, then the converse relation is true: $K = (K^*)^*$.

Example 1. Let the set X be the Euclidean space $\mathbf{R}^n$ of vectors $x = (x_1, \ldots, x_n)$, and let the basis functions be the quadratic functions $\alpha_{ij}(x) = x_i x_j$, $1 \leq i \leq j \leq n$. It is obvious that $\mathcal{F}_p$, where $p = n(n+1)/2$, is the space of all quadratic forms in n variables, the functions of the form δ_y are the forms of rank 1, and the cone $K = K^*$ is the cone of nonnegative forms, i.e., we have the construction of Paragraphs 3 and 4.

Let us now try to generalize the results of Paragraphs 3 and 4 for the case of arbitrary basis functions $\{\alpha_s(\cdot)\}_{s=1}^p$. Consider the question of the convexity of the image of the set $X_1 \subset X$ de novo, where the map $\varphi: X \to \mathbf{R}^k$ is defined by

$$\varphi(x) = (F_1(x, \ldots, F_k(x)) \in \mathbf{R}^k, \quad F_j \in \mathcal{F}_p, \ j = 1, \ldots, k. \tag{5.2}$$

Denote by $\tilde{X}_1$ the set $\{\delta_y, y \in X_1\}$, and by L the linear hull of the set of functions $\{F_j\}_{j=1}^k$.

THEOREM 26. The set $\varphi(X_1)$ is convex if and only if the set $\mathrm{Pr}_L X_1$ is convex, where Pr_L is the orthogonal projection on the subspace L in $\mathcal{F}_p$.

Proof. Suppose initially that $\langle F_i, F_j\rangle = \delta_{ij}$, $i, j = 1, \ldots, k$. Then for each $x \in X$, the vector $\mathrm{Pr}_L \delta_x \in L$ and has the coordinates $\{F_j\}_{j=1}^k$ in the basis of the functions $\{\langle \delta_{x_1}, F_j\rangle\}_{j=1}^k$. But $\langle \delta_x, F_j\rangle = F_j(x)$, $j = 1, \ldots, k$, and the sets $\mathrm{Pr}_L \tilde{X}_1$ and $\varphi(X_1)$ simply coincide (up to an isomorphism of one-dimensional Euclidean spaces). The general case is reduced to the one already considered by a nonsingular linear transformation on the space L, the transformation also orthogonalizing the functions $\{F_j\}_{j=1}^k$.

As an example of the application of Theorem 26, let us investigate the convexity of the set $\varphi(X_1)$ for several classes of forms of degree four in two real variables. The set $X_1 = X$ in these examples will be the whole space $\mathbf{R}^2$.

Example 2. Consider the set of symmetric forms of degree four in two variables, i.e., functions of the form $F(x) = f_1x_1^4 + f_2x_1^3x^2 + f_3x_1^2x_2^2 + f_2x_1x_2^3 + f_1x_2^4$. The set of all such forms is a three-dimensional linear space $\mathcal{F}$, with basis functions $\alpha_1(x) = x_1^4 + x_2^4$, $\alpha_2(x) = x_1x_2(x_1^2 + x_2^2)$, $\alpha_3(x) = (x_1x_2)^2$. Introduce the new variables

$$u = x_1^2 + x_2^2, \quad v = x_1x_2. \tag{5.3}$$

Equation (5.3) is soluble for x_1, x_2 if and only if u and v satisfy the inequality $u^2 \geq 4v^2$. After the change of variables, we have quadratic forms of the type $F(u, v) = f_1u^2 + f_2uv + (f_3 - 2f_1)v^2$. But the arguments u and v are related by $u^2 - 4v^2 \geq 0$. Therefore if the coordinates of the points of the space $\mathcal{F}$, in the basis of the functions

$$\bar{\alpha}_1(u, v) = u^2, \ \bar{\alpha}_2(u, v) = uv, \ \bar{\alpha}_3(u, v) = v^2 \tag{5.4}$$

are denoted by y_1, y_2, y_3, then the set X is the intersection of the surface of the right circular cone $y_1y_3 = y_2^2$, $y_1 \geq 0$ with the half-space $y_1 - 4y_3$ (Fig. 2).

Simple geometrical arguments show that to verify the convexity of the set $\varphi(\mathbf{R}^2)$ under the mapping $\varphi: \mathbf{R}^2 \to \mathbf{R}^2$ defined by the given forms F_1 and F_2, it suffices to calculate the vector product H of the forms F_1 and F_2 as vectors in $\mathcal{F}$,. If the coordinates of the vector H satisfy the inequalities $y_1 - 4y_3 \geq 0$, $y_1 \geq 0$, $y_1y_3 \geq y_2^2$ (these define a convex cone M), then the set $\varphi(\mathbf{R}^2)$ is not convex, while in the opposite case the set $\varphi(\mathbf{R}^2)$ is convex. It is equally easy to find the conditions which must be imposed on the symmetric form F_1 such that when these conditions are fulfilled the set $\varphi(\mathbf{R}^2)$ is convex for any symmetric form F_2. These conditions are that $\langle F_1, H\rangle \neq 0$ for any form $H \in M$, i.e., $H \in \mathrm{Int}\, M^*$ or $H \in \mathrm{Int}\, M^*$, where M^* is the cone dual to M.

Example 3. Consider the set of even forms of degree four in two variables. They can be represented in the form $F(u, v) = f_1u^2 + f_2uv + f_3v^2$, where $u = x_1^2$, $v = x_2^2$. The variables u and v can take values in the range given by the inequalities $u \geq 0$, $v \geq 0$. If we use the basis functions (5.4) as in the previous example, then the coordinates y_1, y_2, y_3 of the points of the set $\tilde{X}$ will satisfy $y_1y_3 = y_2^2$, $y_2 \geq 0$. Here too, just as in the previous example, the set $\tilde{X}$ is the intersection of the surface of the cone with a half-space, and therefore the convexity condition for $\varphi(\mathbf{R}^2)$ is similar to that of Example 1.

Example 4. Consider the class of forms of the type

$$F(x) = f_1x_1^4 + f_2x_1^3x_2 + f_3x_1^2x_2^2 - f_2x_1x_2^3 + f_1x_2^4. \tag{5.5}$$

Each form of this class can be written as follows: $\tilde{F}(u, v) = f_1u^2 + f_2uv + (f_3 + 2f_1)v^2$, where

$$\begin{aligned} u &= x_1^2 - x_2^2 \\ v &= x_1x_2 \end{aligned} \tag{5.6}$$

Since (5.6) is soluble with respect to x_1 and x_2 for any u and v, the set $\tilde{X}$ in this case will be the whole surface of a right circular cone in $\mathbf{R}^3$ (the basis functions are again as in (5.4)). The projection of the surface on any plane in $\mathbf{R}^3$ is a convex set, and Theorem 26 implies that the set $\varphi(\mathbf{R}^2)$ is convex for all forms F_1 and F_2 of the type (5.5). We note that the forms themselves of type (5.5) are not, in general convex functions, even after the change of variables (5.6).

We now attack the conditions for a favorable S-procedure, restricting ourselves, as in Paragraph 4, to considering the conditions (0.1) and (0.2) (or (1.1) and (1.2)), and we assume that the functions $G_1, \ldots, G_m$ satisfy (1.7) (or (1.8)). It is easy to rewrite the two conditions (0.1) and (1.1) in terms of the space $\mathcal{F}_n$, viz,

$$\langle F, \delta_x\rangle \geq 0 \quad \text{for } \langle G_j, \delta_x\rangle \geq 0, \ j = 1, \ldots, m, \tag{5.7}$$

$$\langle F, \delta_x\rangle \geq 0 \quad \text{for } \langle G_j, \delta_x\rangle = 0, \ j = 1, \ldots, m. \tag{5.8}$$

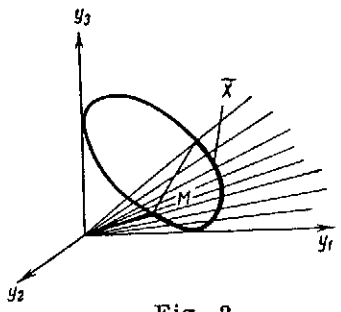

Fig. 2

Introduce the set

$$R = \left\{ A \in \mathcal{F}_p : A = \tau_0 F - \sum_{j=1}^{m} \tau_j G_j \geq 0,\ \tau_i \geq 0,\ j = 0, \ldots, m \right\}.$$

If we use (1.7), then we can rewrite (0.2) in the form

$$R \cap K^* \neq \{0\}. \tag{5.9}$$

By Lemma 4, (5.9) is equivalent to $\operatorname{Int}\{-R^*\} \cap \operatorname{Int} K = \phi$, i.e., there is no function $A \in \operatorname{Int} K$ such that $\langle A, F\rangle < 0$ and $\langle A, G_j\rangle > 0$, $j = 1, \ldots, m$. By repeating the arguments of Paragraph 4, we obtain that (0.2) can be rewritten as follows:

$$\langle F, P\rangle \geq 0 \quad \text{for} \quad P \in \overline{K} \text{ and } \langle G_j, P\rangle \geq 0,\ j = 1, \ldots, m. \tag{5.10}$$

Similarly, (1.2) becomes

$$\langle F, P\rangle \geq 0 \quad \text{for } P \in \overline{K} \text{ and } \langle G_j, P\rangle = 0,\ j = 1, \ldots, m. \tag{5.11}$$

Thus the formulation of (0.1), (0.2), (1.1), and (1.2) is unchanged from Paragraph 4. Nevertheless, (E) and (E') must be formulated somewhat differently. The reason for this is hidden in the fact that the cone K is now not compactly generated in general (see Remark 2 after Proposition 4), and the cone K may not be closed.

THEOREM 27. For (0.1) and (0.2) to be equivalent for any function $F \in \mathcal{F}_p$, it is necessary and sufficient that the functions $G_1, \ldots, G_m$ satisfy the following condition (E):

(E) The set $M = \overline{K} \cap \{A : \langle G_j, A\rangle \geq 0,\ j = 1, \ldots, m\}$ coincides with the closed convex cone spanned by the functions of the type $\delta_X \in M$.

Proof. Suppose (E) holds. We now prove that (0.1) implies (0.2), for any $F \in \mathcal{F}_p$. We have to show that $\langle F, P\rangle \geq 0$ for any $P \in M$. Condition (E) means that each function $P \in M$ can be arbitrarily closely approximated by functions of the form $P_r = \sum_{s=1}^{r} \lambda_s \delta_{X_s}$, where $\lambda_s \geq 0$, $\delta_{X_s} \in M$, $s = 1, \ldots, r$. Therefore

$$\langle F, P\rangle = \lim_{r \to \infty} \langle F, P_r\rangle = \lim_{r \to \infty} \sum_{s=1}^{r} \lambda_s \langle F, \delta_{x_s}\rangle \geq 0.$$

Conversely, suppose some function $P_0 \in M$ does not lie in the closed convex cone M_0 spanned by the functions of the type $\delta_X \in M$. Then the separability theorem implies that there exists a linear functional $F_0(A) = \langle F_0, A\rangle$ in the space $\mathcal{F}_p$ such that $\langle F_0, A\rangle \geq 0$ for $A \in M_0$ and $\langle F_0, P_0\rangle < 0$, i.e., for the function F_0 we have that (0.1) holds, but (0.2) does not, and this contradicts the hypothesis of the theorem.

A similar theorem holds for constraints in the form of equalities.

THEOREM 28. For (1.1) and (1.2) to be equivalent for any function $F \in \mathcal{F}_p$, it is necessary and sufficient that the functions $G_1, \ldots, G_m$ satisfy the following condition (E'):

(E') The set $M' = \overline{K} \cap \{A : \langle G_j, A\rangle = 0,\ j = 1, \ldots, m\}$ coincides with the closed convex cone spanned by the functions of the type $\delta_X \in M'$.

We remark that (E) does not imply (E') in general.

Example 5. Suppose $X = \mathbf{R}^1$. Consider the set of real quadratic polynomials $\alpha_1(x) = 1$, $\alpha_2(x) = x$, $\alpha_3(x) = x^2$. The cone $\overline{K}$ in the three-dimensional space of these polynomials with coordinates y_1, y_2, y_3 is specified by the inequalities $y_1 y_3 \geq y_2^2$, $y_1 \geq 0$. It is easy to convince oneself that the function $G(x) = x$ satisfies (E) but does not satisfy (E').

The above approach is related to the S-procedure in that the change of variables $x \to \delta_X$ transforms all the functions of $\mathcal{F}_p$ into linear functions simultaneously. Moreover, all troubles from the nonlinearity and nonconvexity are pushed into the set K, and the problem is reduced to the study of the geometry of this set and its intersections with subspaces and half-spaces in $\mathcal{F}_p$.

Let us now go on to give a short exposition of another approach to the S-procedure problem. In contrast to the reduction described above for the problem of finite dimensional nonlinear programming, we shall now consider it as an infinite dimensional linear programming problem. Let us assume that the set X is compact in $\mathbf{R}^n$, and that $G_1, G_2, \ldots, G_m$ are continuous real functions on X.

Fix $x \in X$. The inequality $F(x) - \sum_{j=1}^{m} \tau_j G_j(x)$ is linear in the vector $\tau = (\tau_1, \ldots, \tau_m) \in \mathbf{R}^m$. Therefore (0.2) means that a certain system of linear inequalities is consistent. There are as many inequalities in this system as there are points in the compact space X, i.e., in general there are infinitely many of each. Thus (0.2) can be understood as the condition for the existence of admissible solutions to the infinite dimensional linear programming problem (with a fictitious zero objective function). Let us rewrite the problem in standard form:

$$\left.\begin{array}{c} 0(\tau) \to \inf \\ G\tau - f \in K_Y \\ \tau \in K_T \end{array}\right\}. \tag{5.12}$$

Here T is the space $\mathbf{R}^m$, $Y = C(X)$ is the space of continuous functions on X, G is the linear operator from T to Y given by $(G\tau)(x) = -\sum_{j=1}^{m} \tau_j G_j(x)$, K_T is the positive octant in $\mathbf{R}^m$, K_Y is the cone of nonnegative functions in C(X), $0(\cdot)$ is the zero linear functional on $\mathbf{R}^m$, and $f \in C(X)$ is the function defined by $f(x) = -F(x)$. The dual problem for (5.12) is given by (see [6]):

$$\left.\begin{array}{c} \mu(f) \to \sup \\ -G^*\mu \in K_{T^*}^* \\ \mu \in K_{Y^*}^* \end{array}\right\}. \tag{5.13}$$

Here $T = \mathbf{R}^m$, $Y^* = [C(X)]^* \simeq V(X)$ is the space of finite Borel measures on X, $K_{T^*}^*$ is the positive octant in $\mathbf{R}^m$, $K_{Y^*}^*$ the cone of finite measures on X, $G^*: Y^* \to T^*$ is the operator adjoint to G, and $\mu(f) = \int f(x)d\mu$. Since the objective functional in (5.12) is zero, then (5.12) and (5.13) satisfy the duality relation. Therefore (0.2) is equivalent to the finiteness of the optimal value for (5.13), which in turn (since (5.13) is homogeneous) is equivalent to $\mu(f) \leq 0$ for $-G^*\mu \in K_{T^*}^*$ and $\mu \in K_{Y^*}^*$. This condition can be rewritten in the form:

$$\int_X F(x)\,d\mu \geqslant 0 \tag{5.14}$$

for any finite measure μ satisfying

$$\int_X G_j(x)\,d\mu \geqslant 0, \quad j = 1, \ldots, m. \tag{5.15}$$

We now note that (0.1) is equivalent to the satisfaction of (5.14) for measures satisfying (5.15) and concentrated at a point. Denote the measure concentrated at the point $x \in X$ by δ_x, and denote the form bilinear in f and μ, $\int_X f(x)d\mu$, by $\langle f, \mu\rangle$. Then (0.1) and (0.2) finally take the form:

$$\begin{gathered} \langle F, \mu\rangle \geqslant 0 \quad \text{for} \quad \langle C_j, \mu\rangle \geqslant 0, \quad j = 1, \ldots, m, \\ \langle F, \mu\rangle \geqslant 0 \quad \text{for} \quad \langle C_j, \mu\rangle \geqslant 0, \quad j = 1, \ldots, m \quad \text{for} \quad \mu = \delta_x, \ x \in X, \end{gathered} \tag{5.16}$$

$$\langle F, \mu\rangle \geqslant 0 \quad \text{for} \quad \langle C_j, \mu\rangle \geqslant 0, \ j = 1, \ldots, m \quad \text{for any measure} \quad \mu. \tag{5.17}$$

Moreover, (1.1) and (1.2) can be rewritten in a similar way.

Consider the space of finite measures V(X) with the cone of measures mep $\mathcal{M}$ lying inside it. Further, consider the cone $K = \mathcal{M} \cap \{\mu : \langle G_j, \mu\rangle \geq 0, \ j = 1, \ldots, m\}$; it is the intersection of $\mathcal{M}$ with a finite number of half-spaces in V(X). The continuous linear functional $\langle F, \cdot\rangle$ is nonnegative on K if and only if it is nonnegative on the extremal generators of K, which are known to be the measures of the type $\sum_{s=1}^{m+1} \alpha_s \delta_{x_s}$, $\alpha_s \geq 0$. Thus we have the following theorem.

THEOREM 29. For (0.1) to imply (0.2), it is necessary and sufficient that for any $x_s \in X$ and any $\alpha_s \geq 0$, $s = 1, \ldots, m+1$ the condition $\sum_{s=1}^{m+1} \alpha_s F(x_s) \geq 0$ hold for $\sum_{s=1}^{m+1} \alpha_s G_j(x) \geq 0$, $j = 1, \ldots, m$.

A similar theorem holds when the constraints are equalities.

THEOREM 30. For (1.1) to imply (1.2), it is necessary and sufficient that for any $x_s \in X$ and for any $\alpha_s \geq 0$, $s = 1, \ldots, m+1$ the condition $\sum_{s=1}^{m+1} \alpha_s F(x_s) \geq 0$ hold for $\sum_{s=1}^{m+1} \alpha_s G_j(x_s) = 0$, $j = 1, \ldots, m+1$.

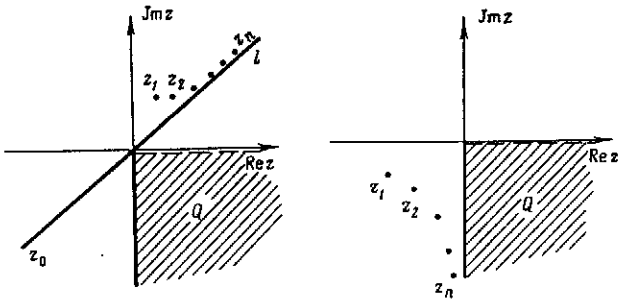

Fig. 3

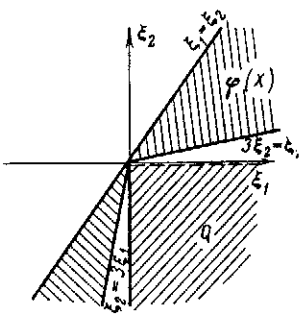

Fig. 4

Application of Theorems 29 and 30 to the case where $F, G_1, \ldots, G_m$ are quadratic or Hermitian forms makes it possible to establish the following algebraic facts, by using the results of paragraph 3.

COROLLARY 1. Let F and G be quadratic forms on $\mathbf{R}^n$ satisfying (0.1) and (1.7). Then for any $x_1, \ldots, x_N \in \mathbf{R}^N$, the inequality $\sum_{s=1}^{N} F(x_s) \geq 0$ holds for $\sum_{s=1}^{N} G(x_s) \geq 0$.

COROLLARY 2. Let F_1, G_1, G_2 be Hermitian forms on $\mathbf{C}^n$ satisfying (0.1) and (1.7). Then for any $x_1, \ldots, x_N \in \mathbf{C}^n$, the inequality $\sum_{s=1}^{N} F(x_s) \geq 0$ holds for $\sum_{s=1}^{N} G_1(x_s) \geq 0$ and $\sum_{s=1}^{N} G_2(x_s) \geq 0$.

§6. Counter Examples

In this paragraph we collect together several examples of functions for which the S-procedure is unfavorable. They show that this or that condition is essential in the theorems of the preceding paragraphs.

Example 1. Let G(x) be a quadratic form on the space $\mathbf{R}^n$ which is nonpositive but is not negative definite ($G \leq 0$ but not $G < 0$). We now construct a quadratic form F(x) for which (0.1) does not imply (0.2) and (1.1) does not imply (1.2). By a nonsingular transformation, G(x) can be reduced to the form x^*Gx, where $G = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}$, $A < 0$. Consider the form $F_0(x) = x^*F_0x$, where† $F_0 = \begin{pmatrix} 0 & C \\ C^* & 0 \end{pmatrix}$. The condition (0.1) for the form G(x) coincides with (1.1) and holds for the form $F_0(x)$. On the other hand, the form $F_0(x) - \tau G(x)$ has the matrix $F_0 - \tau G = \begin{pmatrix} -\tau A & C \\ C^* & 0 \end{pmatrix}$, and if $C \neq 0$ it is an alternating matrix for any real τ.

This example shows the necessity for the conditions (1.7) and (1.8) in Theorems 14 and 15. A similar example is easily constructed for Hermitian forms.

Example 2. Let G(x) be an Hermitian form on the infinite dimensional Hilbert space H, but not a form which acts on a finite dimensional subspace (i.e., its rank is infinite). Then there is an Hermitian F(x) such that for the forms G(x), F(x) condition (1.3) is valid, but (1.4) is not. Moreover, if the symmetric operator G corresponding to the form G(x) is bounded, and if the form G(x) takes positive values, then the operator F corresponding to the form F(x) may be assumed to be bounded. If G is completley continuous then F too can be assumed to be completely continuous.

The construction of the required form F(x) is based on the following theorem (see [10], p. 117).

THEOREM 31. If A is a normal operator on H, then the closure of the set of complex numbers of the type (Ax, x), $\|x\| = 1$ coincides with the convex hull of the spectrum of the operator A.

Let us now describe the construction when the spectrum of the operator G is discrete. When there is a continuous component in the spectrum, the argument is similar. The operator F may also be assumed to have a (real) discrete spectrum and eigenvectors identical with those of the operator G. Consider the

†The blocks of the matrix F_0 have the same orders as the corresponding blocks in the matrix G.

operator $A = G + iF$. It is normal, and so Theorem 31, together with the results of paragraph 2, clearly imply that for the construction of the required example there must be imaginary parts in its eigenvalues for specified real parts, so that the set $P = \{\alpha z : \alpha > 0,\ z = (Ax, x),\ \|x\| = 1\}$ does not intersect the set $\overline{Q} = \{z : \mathrm{Re}\, z \geq 0,\ \mathrm{Im}\, z \leq 0\}$ but may be strictly separated from it by a vertical line. It can be assumed that G has at least one negative eigenvalue λ_0,* and a countable set $\{\lambda_k\}_{k=1}^{\infty}$ of eigenvalues of the one sign. Suppose $\lambda_k > 0$, $k = 1, 2, \ldots$. Then we draw an arbitrary line $l = \{z : \mathrm{Im}\, z - \tau \mathrm{Re}\, z = 0\}$ in the complex plane, where $\tau > 0$, and we take the corresponding eigenvalues μ_k, $k = 0, 1, \ldots$ of the operator F so that the point $z_0 = \lambda_0 + i\mu_0$ lies on the line l, while μ_k/λ_k tends to τ from above as $k \to \infty$ (Fig. 3a). If all the $\lambda_k < 0$, then the numbers μ_k, $k = 1, 2, \ldots$ are chosen so that the sequence $\{\mu_k/\lambda_k\}$ contains a subsequence which tends to $K - \infty$ (Fig. 3b). It is easy to see that in both cases the set P satisfies the necessary requirements. The numbers μ_k are constructed in a similar manner in the case where the operator G is bounded or completely continuous.

This example shows that the condition that the domain of definition of F(x) and G(x) be finite dimensional in Theorem 16 is necessary in the usual sense, and that it may not be replaced, generally speaking, by conditions on the continuity or the complete continuity of the operators G and F.

Example 3. Suppose $m = 2$ and that the quadratic forms $G_1(x)$, $G_2(x)$ are of rank 2 on the space R^4 and have the form $G_1(x) = x_1^2 - x_2^2$ $G_2(x) = x_3^2 - x_4^2$, $x = (x_1, x_2, x_3, x_4)$ in some basis.† Consider the quadratic form $F(x) = x_1^2 + x_3^2 + x_1(x_3 + x_4) + x_2(x_3 - x_4)$. The forms F(x), $G_1(x)$, $G_2(x)$ satisfy (0.1) since $G_1(x) \geq 0$, $G_2(x) \geq 0$ imply that $F(x) \geq x_1^2 + x_3^2 - |x_1|(|x_3 + x_4| + |x_3 - x_4|) \geq x_1^2 + x_3^2 - 2|x_1||x_3| \geq 0$. Nevertheless the S-procedure for $F(x) \geq 0$ with the constraints $G_1(x) \geq 0$, $G_2(x) \geq 0$ is unfavorable for the given case. In fact, from the nonnegativity of the principal minors of first and second orders of the matrix of the form $F(x) - \tau_1 G_1(x) - \tau_2 G_2(x)$ it follows that $\tau_1 = \tau_2 = 1/2$. But $F(x) - [G_1(x) + G_2(x)]/2 = -1$ for $x = (1.1 - 1.0)$, i.e., the forms F(x), $G_1(x)$, $G_2(x)$ do not satisfy even condition (0.2).

Example 4. Let $X = R^3 = \{(x_1, x_2, x_3)\}$ $G_1(x) = x_1^2 - 3x_2^2 - x_3^2$, $G_2(x) = 2x_1x_2 - x_3^2$, $F(x) = x_1^2 - 9x_2^2 + x_3(x_1 - 3x_2)$. Direct verification shows that $F(x) = 0$ for $G_1(x) = G_2(x) = 0$, but it is obvious that F(x) is not a linear combination of the forms $G_1(x)$ and $G_2(x)$.

Example 5. Let $m = 1$, and suppose that the functions F(x), G(x) are forms of degree four in two variables. Consider the forms $F(x) = -x_1^2x_2^2 + x_1x_2(x_1^2 + x_2^2)$, $G(x) = x_1^2x_2^2 + x_1x_2(x_1^2 + x_2^2)$. If $G(x) \geq 0$, then $x_1x_2 \geq 0$, and it follows from the inequality $2x_1x_2 \leq x_1^2 + x_2^2$ that $F(x) = x_1x_2(x_1^2 + x_2^2 - x_1x_2)$, i.e., (0.1) and (1.1) hold. But $F(x) - \tau G(x) = -x_1^2x_2^2(1 + \tau) + x_1x_2(x_1^2 + x_2^2)(1 - \tau)$ and the requirement $F(x) - \tau G(x) \geq 0$, for example, lead to the inconsistent inequalities $\tau \geq 3$, $\tau \leq 1/3$ for $x_1 = \pm x_2$ i.e., (0.2) and (1.2) do not hold.

It is easy to show that in this example the set $\varphi(R^2)$ consists of two sectors of the plane, and these are located as in Fig. 4.

The author is deeply grateful to V. A. Yakubovich for his support and interest, and to A. M. Vershik for his stimulating discussions.

LITERATURE CITED

1. F. P. Gantmacher and V. A. Yakubovich, "Absolute stability of nonlinear control systems," Proceeds of the Second All-Union Session on Theoretical and Applied Mechanics [in Russian], Nauka, Moscow (1966).
2. V. A. Yakubovich, "S-procedures in nonlinear control theory," Vest. Leningr. Univ., No. 1, 62-77 (1971).
3. M. A. Aizerman and F. P. Gantmacher, Absolute Stability of Nonlinear Control Systems, Izd. AN SSSR, Moscow (1963).
4. M. G. Krein and Yu. L. Shmul'yan, "Plus-operators on spaces with indefinite metrics," Matem. Issledovaniya, 1, No. 1, 131-161 (1966).

*If not, then (1.3) always implies (1.4).

†The question of the advantage of the S-procedure for the forms $G_1(x)$ and $G_2(x)$ of the type specified allows one to investigate the stability of control systems with two nonlinearities [2]. A positive answer to this question would mean that the Popov frequency condition encompasses all conditions which can be obtained via the use of all possible Lyapunov functions of the type: "quadratic form plus the sum of integrals of the nonlinearities."

5. M. R. Hestenes and E. J. McShane, "A theorem on quadratic forms and its application in the calculus of variations," Trans. Amer. Math. Soc., 47, 501-520 (1940).
6. E. G. Gol'shtein, The Theory of Duality in Mathematical Programming and Its Applications [in Russian], Nauka, Moscow (1971).
7. A. L. Fradkov and V. A. Yakubovich, "S-procedures and the duality relation in nonconvex probler of quadratic programming," Ves. Leningr. Univ., No. 1, 71-76 (1973).
8. N. Bourbaki, Topological Vector Spaces [Russian translation], IL, Moscow (1959).
9. F. Hausdorf, "Der Wertvorrat einer Bilinearform," Math. Z., 3, 314-316 (1919).
10. P. Halmos, Hilbert Space Problems [Russian translation], Mir (1970).
11. L. L. Dines, "On the mapping of quadratic forms," Bull. Amer. Math. Soc., 47, 494-498 (1941),
12. P. Finsler, "Über das Vorkommen definiter und semidefiniter Formen in Scharen quadratischen Formen," Comment. Math. Helv., 9, 188-192 (1937).
13. I. M. Glazman and Yu. I. Lyubich, Problems in Finite Dimensional Linear Analysis [in Russian], Nauka, Moscow (1969).