On the extreme points of sets of measures defined by moment inequalities

By G. FAZEKAS (Debrecen)

Dedicated to Professor Zoltán Daróczy on his 50th birthday

0. Introduction

In this paper we generalize some results connected with the extreme points of finite dimensional polyhedral sets, for subsets of the space $\mathrm{NBV}[\alpha, \beta]$ defined by moment inequalities. A generalisation of the structure theorem will be proved and, by its use, some theorems on the existence of extreme points. The results obtained are based mainly on [1] and further extend some of the results of [2] and [3].

1. Basic facts and notations

Let $[\alpha, \beta]\subset R$ be a finite closed interval, $\mathrm{NBV}[\alpha, \beta]$ the Banach space of all real valued normed functions of bounded variation defined on $[\alpha, \beta]$. We shall assume that each $\sigma\in\mathrm{NBV}[\alpha, \beta]$ is normed in the following sense:

(i) $\sigma(\alpha)=0$,

(ii) σ is right continuous on (α, β).

Let $\mathrm{NBV}_+[\alpha, \beta]$ be the positive cone of $\mathrm{NBV}[\alpha, \beta]$ consisting of all nondecreasing functions $\sigma\in\mathrm{NBV}[\alpha, \beta]$.

Definition 1.1. For a $\sigma\in\mathrm{NBV}_+[\alpha, \beta]$, $S(\sigma)\subseteqq[\alpha, \beta]$ denotes the set of all increasing points of σ. $S(\sigma)$ is called the spectrum of σ.

Definition 1.2. We define the functions e_λ on $[\alpha, \beta]$ in the following way. Let

$$e_\alpha(t) = \begin{cases} 0, & \text{if } t = \alpha \\ 1, & \text{if } \alpha < t \leqq \beta \end{cases}$$

and let for $\lambda\in(\alpha, \beta]$

$$e_\lambda(t) = \begin{cases} 0, & \text{if } \alpha \leqq t < \lambda \\ 1, & \text{if } \lambda \leqq t \leqq \beta. \end{cases}$$

Proposition 1.1 [1]. (a) *For every* $\lambda\in[\alpha, \beta]$ $e_\lambda\in\mathrm{NBV}_+[\alpha, \beta]$.

(b) *If* $\sigma \in \mathrm{NBV}_+[\alpha, \beta]$ *is a step function with a finite number of jumps, then it can be expressed in the form*

$$\sigma = \sum_{i=1}^{k} \varrho_i e_{\lambda_i}$$

where $\varrho_i > 0$ *for* $i = 1, 2, \ldots, k$ *and* $\alpha < \lambda_1 < \lambda_2 < \ldots < \lambda_k < \beta$. *Moreover, for all continuous functions* $\varphi\colon [\alpha, \beta] \to \mathbf{R}$, *the Stieltjes integral of* φ *with respect to* σ *can be written in the following form:*

$$\int_{\alpha}^{\beta} \varphi(t)\,\mathrm{d}\sigma(t) = \sum_{i=1}^{k} \varrho_i \varphi(\lambda_i).$$

$\mathbf{R}^m$ denotes the vector space of all m-tuples of real numbers and $\mathbf{R}_+^m \subset \mathbf{R}^m$ is the set of vectors with nonnegative components. E stands for the m-order quadratic identity matrix. For $\sigma \in \mathrm{NBV}_+[\alpha, \beta]$ and $w \in \mathbf{R}^m$, we consider the equation:

$$\int_{\alpha}^{\beta} u(t)\,\mathrm{d}\sigma(t) + Ew = b, \tag{1.1}$$

where $u\colon [\alpha, \beta] \to \mathbf{R}^m$ is a given continuous vector-valued function and $b \in \mathbf{R}^m$ is a fixed vector.

Definition 1.3. A solution $(\sigma, w) \in \mathrm{NBV}[\alpha, \beta] \times \mathbf{R}^m$ of (1.1) is called an admissible solution if $\sigma \in \mathrm{NBV}_+[\alpha, \beta]$ and $w \in \mathbf{R}_m^+$.

Let e_i denote the i-th column vector of E $(i = 1, 2, \ldots, m)$ and for each $w \in \mathbf{R}_m^+$, let $P(w)$ denote the set of all indices i, for which the i-th component of w is positive.

Definition 1.4. An admissible solution (σ, w) of (1.1) is called an admissible basic solution if the system of vectors

$$u(\lambda), \quad \lambda \in S(\sigma), \quad e_i, \quad i \in P(w) \tag{1.2}$$

is linearly independent in the space $\mathbf{R}^m$. If $(\sigma, w) = (0, O)$ is a solution of (1.1) then it will be assumed to be an admissible basic solution.

Remark 1.1. Since there exist only m independent vectors in $\mathbf{R}^m$, if (σ, w) is an admissible basic solution of (1.1) then σ has at most m points of increase. In other words, in this case, σ is a step function with at most m jumps.

Definition 1.5. Let $\tilde{V}(b)$ denote the set of all admissible solutions of (1.1) and let

$$V(b) = \left\{\sigma \in \mathrm{NBV}_+[\alpha, \beta] \colon \int_{\alpha}^{\beta} u(t)\,\mathrm{d}\sigma(t) \leqq b\right\} \tag{1.3}$$

2. The characterization of the extreme points of $V(b)$

In this section we shall be concerned with the problem of the characterization of $\operatorname{ext} V(b)$.

Proposition 2.1. *$V(b)/\tilde{V}(b)$ is a convex subset of the space*

$\mathrm{NBV}[\alpha, \beta]/\mathrm{NBV}[\alpha, \beta]\times\mathbf{R}^m$.

Proposition 2.2. *$\sigma\in\operatorname{ext} V(b)$ iff a $w\in\mathbf{R}_m^+$ exists such that $(\sigma, w)\in\operatorname{ext}\tilde{V}(b)$.*
In view of Definition 1.5 these propositions are obvious.

Theorem 2.1. *The set $\operatorname{ext} V(b)$ consists of all admissible basic solutions of* (1.1).

PROOF. Let (σ, w) be an admissible basic solution of (1.1). At first we assume that the sets $S(\sigma)$ and $P(w)$ are nonvoid. Let

$$S(\sigma) = \lambda_1, \dots, \lambda_k$$

and

$$P(w) = j_1, \dots, j_l.$$

Then $k+1\leqq m$ and, by virtue of Definition 1.4, the vectors

$$u(\lambda_1), \dots, u(\lambda_k), \quad e_{j_1}, \dots, e_{j_l} \tag{2.1}$$

are linearly independent in $\mathbf{R}^m$. Moreover, σ has the representation

$$\sigma = \sum_{i=1}^{k} \varrho_i e_{\lambda_i}$$

where $\varrho_i>0$ for $i=1, 2, \dots, k$ and if w has the components $w_1, w_2, \dots, w_m$ then the relation

$$b = \sum_{i=1}^{k} \varrho_i u(\lambda_i) + \sum_{i=1}^{l} w_{j_i} e_{j_i} \tag{2.2}$$

holds. Suppose that there exist (σ, w_1), $(\sigma_2, w_2)\in\tilde{V}(b)$ and $0<v<1$ satisfying the condition

$$(\sigma, w) = v(\sigma_1, w_1)+(1-v)(\sigma_2, w_2).$$

This implies that

$$\sigma = v\sigma_1+(1-v)\sigma_2$$

$$w = vw_1+(1-v)w_2.$$

Therefore, in consequence of $0<v<1$, we get $S(\sigma_i)\subseteqq S(\sigma)$, $P(w_i)\subseteqq P(w)$ for $i=1, 2$. Thus, there exist nonnegative numbers $\varrho_1', \dots, \varrho_k'$ and $\varrho_1'', \dots, \varrho_k''$ such that

$$\sigma_1 = \sum_{i=1}^{k} \varrho_i' e_{\lambda_i}, \quad \sigma_2 = \sum_{i=1}^{k} \varrho_i' e_{\lambda_i}.$$

If $w_1', \dots, w_m'$ and $w'', \dots, w_m''$ denote the components of w_1 and w_2 then we have

$$\sum_{i=1}^{k} \varrho_i' u(\lambda_i) + \sum_{i=1}^{l} w_{j_i}' e_{j_i} = \sum_{i=1}^{k} \varrho_i'' u(\lambda_i) + \sum_{i=1}^{l} w_{j_i} e_{j_i} = b. \tag{2.3}$$

Since the vectors (2.1) are linearly independent, (2.1) and (2.3) imply $\varrho_i'=\varrho_i''=\varrho_i$ for $i=1, 2, \dots, k$ and $w_i'=w_i''=w_i$ for $i\in P(w)$. Consequently we obtain $(\sigma_1, w_1)=(\sigma_2, w_2)=(\sigma, w)$. Therefore (σ, w) is an extreme point of $V(b)$. Observe that in the cases where $S(\sigma)$ and/or $P(w)$ are empty sets, the proof becomes easier.

To prove the necessity of the condition we assume $(\sigma_0, w_0)\in\operatorname{ext}\tilde{V}(b)$ and let for $i=1, 2, \dots, m+1$

$$r_i = \int_\alpha^\beta u(t)\,\mathrm{d}\big(1+\sigma_0(t)\big)^{1/i}. \tag{2.4}$$

Evidently the vectors defined by (2.4) are dependent, thus there exist real numbers v_i $(i=1, 2, \dots, m+1)$ so that

$$\sum_{i=1}^{m+1} v_i r_i = 0 \tag{2.5}$$

$$\sum_{i=1}^{m+1} v_i = 1. \tag{2.6}$$

With the aid of these numbers, we define the function φ on the half line $0<x<\infty$ by the relation

$$\varphi(x) = \sum_{i=1}^{m+1} v_i[(1+x)^{1/i}-1].$$

In [1] the following properties of φ are shown:

(a) $\varphi(0)=0$.

(b) φ is continuous on $[0, \infty)$.

(c) $\varphi(x_2)-\varphi(x_1)<x_2-x_1$ holds for arbitrary $0<x_1<x_2$.

(d) φ has at most $m+1$ roots in $[0, \infty)$.

Let us now define the functions σ_1 and σ_2 by

$$\sigma_1(t) = \sigma_0(t)+\varphi\big(\sigma_0(t)\big), \quad t\in[\alpha, \beta] \tag{2.7}$$

$$\sigma_2(t) = \sigma_0(t)-\varphi\big(\sigma_0(t)\big), \quad t\in[\alpha, \beta]. \tag{2.8}$$

Using the assertions (a), (b) and (c) it is easy to see that the functions just defined belong to $\mathrm{NBV}_+[\alpha, \beta]$. On the other hand σ satisfies by its definition, by (2.5) and (2.6) the relation

$$\int_\alpha^\beta u(t)\,\mathrm{d}\varphi\big(\sigma_0(t)\big) = O$$

hence in view of (2.7) and (2.8)

$$\int_\alpha^\beta u(t)\,\mathrm{d}\sigma_1(t) = \int_\alpha^\beta u(t)\,\mathrm{d}\sigma_2(t) = \int_\alpha^\beta u(t)\,\mathrm{d}\sigma_0(t).$$

Consequently, we get (σ_1, w_0), $(\sigma_2, w_0)\in\tilde{V}(b)$. (σ_0, w_0) is obviously a nontrivial convex combination of (σ_1, w_0) and (σ_2, w_0), so we can infer by our assumption $(\sigma_0, w_0)\in\operatorname{ext}\tilde{V}(b)$ that $\sigma_0=\sigma_1=\sigma_2$. From this it follows by (2.7) and (2.8) that for all $t\in[\alpha, \beta]$ $\varphi\big(\sigma_0(t)\big)=0$ holds. This means that by (d) the range of σ consists of at most $m+1$ elements, from which we immediately obtain that $S(\sigma_0)$ contains at most m elements.

To complete the proof it remains only to show the independence of the vectors

$$u(\lambda),\quad \lambda\in S(\sigma_0);\quad e_j,\quad j\in P(w_0). \tag{2.9}$$

This is obviously clear in the case of $S(\sigma_0)=\emptyset$. Furthermore the cases of $S(\sigma_0)\neq\emptyset$, $P(w_0)\neq\emptyset$ and $S(\sigma_0)\neq\emptyset$, $P(w_0)=\emptyset$ can be handled in a unified manner. For example, let $S(\sigma_0)=\lambda_1, \lambda_2, \dots, \lambda_k$ and $P(w_0)=j_1, j_2, \dots, j_l$. Then σ_0 may be expressed in the form

$$\sigma_0 = \sum_{i=1}^{k} \varrho_i e_{\lambda_i}$$

where $\varrho_i>0$ for $i=1, 2, \dots, k$. If $w_1, w_2, \dots, w_m$ denotes the components of w_0, then we have

$$b = \sum_{i=1}^{k} \varrho_i u(\lambda_i) + \sum_{i=1}^{l} w_{j_i} e_{j_i}.$$

Suppose now that the real numbers $v_1, v_2, \dots, v_k$ and $\mu_1, \mu_2, \dots, \mu_l$ satisfy

$$\sum_{i=1}^{k} v_i u(\lambda_i) + \sum_{i=1}^{l} \mu_i e_{j_i} = O \tag{2.10}$$

Without loss of generality, we can assume that $v_i<\varrho_i$ for $i=1, 2, \dots, k$ and $\mu_i<w_{j_i}$ for $i=1, 2, \dots, l$. Then it is easy to see that the pairs (σ_1, w_1) and (σ_2, w_2) defined by

$$\sigma_1 = \sum_{i=1}^{k} (\varrho_i+v_i) e_{\lambda_i};\quad w_1 = w_0 + \sum_{i=1}^{l} \mu_i e_{j_i}$$

and

$$\sigma_2 = \sum_{i=1}^{k} (\varrho_i-v_i) e_{\lambda_i};\quad w_2 = w_0 - \sum_{i=1}^{l} \mu_i e_{j_i}$$

belong to $\tilde{V}(b)$, furthermore $(\sigma_0, w_0)=1/2(\sigma_1, w_1)+1/2(\sigma_2, w_2)$. Hence in consequence of $(\sigma_0, w_0)\in\operatorname{ext}\tilde{V}(b)$, we get $\sigma_1=\sigma_2=\sigma_0$, $w_1=w_2=w_0$, which means, by the definition of σ_1, σ_2, w_1, w_2, that all the numbers v_i and μ_i are zero. Therefore, we have shown that (2.10) implies $v_1=v_2=\dots=v_k=0$ and $\mu_1=\mu_2=\dots=\mu_l=0$, i.e. the system (2.9) is indeed linearly independent. *This completes the proof of Theorem 2.1.*

As an immediate consequence of the above theorem and the Proposition 2.1 we find the following

Corollary 2.1. *$\sigma\in\operatorname{ext} V(b)$ iff there exists a $w\in\mathbf{R}^m_+$ so that (σ, w) is an admissible solution of* (1.1).

Remark 2.1. The notion of admissible (basic) solution of the equation

$$\int_{\alpha}^{\beta} u(t)\,\mathrm{d}\sigma(t) = b \tag{2.11}$$

can be defined analogously to Definitions 1.3 and 1.4. Let $V(b)$ denote the set of all admissible solutions of (2.11). The extreme points of $V(b)$ were first examined in [3] under certain additional conditions, for which the components of u constitute a Chebyshev system on $[\alpha, \beta]$. In [1] we proved, without additional conditions, that $\operatorname{ext} V(b)$ consists of all admissible basic solutions of (2.11). The same result can be obtained from our above considerations replacing in (2.11) each equation by two inequalities.

Remark 2.2. Let A be a real (m, n) matrix. It is well known that the set of extreme points of the finite dimensional polyhedron L which is defined by $L=\{x\in\mathbf{R}^m\colon\ Ax\leqq b,\ x\geqq 0\}$ consists of all admissible basic solutions of the equation $Ax\leqq b$. Our result may be regarded as an infinite dimensional extension of this assertion.

3. The existence of extreme points

As a first application of Theorem 2.1 we shall give a simple proof for the existence of the extreme points. For this, the following propositions will be necessary:

Proposition 3.1 *Let* Co U *denote the convex hull of the curve* $U=\{u(t)\colon t\in[\alpha, \beta]\}$. *Then*

$$\mathrm{Co\,U} = \Big\{r = \int_\alpha^\beta u(t)\,\mathrm{d}\sigma(t)\colon\ \sigma\in\mathrm{NBV}_+[\alpha,\beta],\ \int_\alpha^\beta \mathrm{d}\sigma(t) = 1\Big\}.$$

Proposition 3.2 [4]. *Each* $r\in\mathrm{Co}\,U$ *can be written as a finite convex combination of points of* U.

Theorem 3.1. *If* $V(b)$ *is nonvoid, then it has an extreme point.*

Proof. Let $\sigma\in V(b)$. It is clear that if $\sigma=0\in V(b)$, then $\sigma\in\operatorname{ext} V(b)$. Therefore we can assume $\sigma\neq 0$. In this case $\sigma(\beta)>0$ and, by virtue of Proposition 3.1, the vector $\big(1/\sigma(\beta)\big)r$, where r is defined by

$$r = \int_\alpha^\beta u(t)\,\mathrm{d}\sigma(t)$$

belongs to Co U. From this we get by Proposition 3.2 that this vector may be expressed in the form:

$$\big(1/\sigma(\beta)\big)\,r = \sum_{i=1}^{k} \varrho' u(\lambda_i),$$

where $\varrho_i'>0$, $\lambda_i\in[\alpha, \beta]$ for $i=1, 2, \ldots, k$ and $\sum_{i=1}^{k}\varrho_i=1$. Let A denote the (m, k) matrix which has the column vectors $u(\lambda_1), \ldots, u(\lambda_k)$, and let us consider the equation

$$Ax+Ey = b. \tag{3.1}$$

It obviously has a solution, namely the one defined by $x=\big(\sigma(\beta)\,\varrho_1', \ldots, \sigma(\beta)\,\varrho_k'\big)$, $y=b-r$. A classical theorem of linear programming states, that it has then an admissible basic solution as well. Let (ϱ, w) be an admissible solution of (3.1), and let us now define σ by

$$\sigma = \sum_{i=1}^{k} \varrho_i e_{\lambda_i}$$

where $\varrho_1, \ldots, \varrho_k$ mean the components of ϱ. This (σ, w) evidently presents an admissible basic solution of (1.1). Therefore by Theorem 2.1 and by Corollary 2.1 our assertion is proved.

4. The maximum of linear functions

If a real valued linear function, defined on a closed convex linefree subset of a finite dimensional Euclidean space, achieves its maximum, then it achieves this maximum at least at one extreme point of the convex set. We shall prove the same statement for the linear function defined on $V(b)$ by

$$\Phi(\sigma) = \int_{\alpha}^{\beta} \omega(t)\,\mathrm{d}\sigma(t),$$

where $\omega\colon [\alpha, \beta] \to \mathbf{R}$ is an arbitrary fixed continuous function.

Theorem 4.1. *If Φ achieves its maximum on $V(b)$ then this maximum will be attained at least at one extreme point of $V(b)$.*

Proof. Let $H(b)$ denote the set of all $\sigma \in V(b)$, for which the maximum of Φ will be attained. By our assumption $H(b)$ is nonvoid. On the other hand $H(b)$ may be regarded as the set of all $\sigma \in \mathrm{NBV}_+[\alpha, \beta]$ satisfying each of the inequalities

$$\int_{\alpha}^{\beta} u(t)\,\mathrm{d}\sigma(t) \leqq b,\ \int_{\alpha}^{\beta} \omega(t)\,\mathrm{d}\sigma(t) \leqq \xi,\ \int_{\alpha}^{\beta} -\omega(t)\,\mathrm{d}\sigma(t) \leqq -\xi,$$

where ξ means the maximal value of Φ on $V(b)$. Thus, applying Theorem 3.1, we obtain that $\operatorname{ext} H(b)$ is nonvoid. Using the extreme property of the elements of $H(b)$ it is easy to see, that $\sigma \in \operatorname{ext} H(b)$ implies $\sigma \in \operatorname{ext} V(b)$. *This completes the proof.*

Remark 4.1. The above statement and Theorem 2.1 imply that the problem of finding the maximum of Φ on $V(b)$ may be reduced to the problem of finding the maximal value of Φ on a subset of $V(b)$ consisting of all step functions with a finite number of jumps. A specialisation of this fact was essentially used in [1] to construct an algorithm for the approximative solution of the Chebyshev—Markov problem.

References

[1] G. Fazekas, Eine Verallgemeinerung des Simplexalgorithmus zur Lösung des Tschebyscheff—Markoff Problems, *Publicationes Mathematicae, Debrecen,* **28**/1981.

[2] K. Glashoff and S.-A. Gustafson, Linear Optimization and Approximation, Springer-Verlag, *New York, Heidelberg, Berlin,* 1983.

[3] S. J. Karlin and W. J. Studden, Tschebyscheff Systems: with Applications in Analysis and Statistics, *Interscience Publishers, New York—London—Sydney,* 1966

[4] М. Г. Крейн и А. А. Нудельман, Проблема моментов Маркова и экстремальные задачи, *Изд. «Наука», Москва,* 1973.

(Received April 10, 1988)