Logarithmic concave measures with application to stochastic programming

By ANDRÁS PRÉKOPA in Budapest*)

1. Introduction. The problem we are dealing with in the present paper arose in stochastic programming. A wide class of stochastic programming decision rules (see [8], [9]) lead to non-linear optimization problems where concavity or quasi-concavity of some functions is desirable. Let us consider the following special decision problem of this kind for illustration:

Minimize $f(\mathbf{x})$ subject to the constraints:

$$P\{g_1(\mathbf{x}) \geqq \beta_1, \dots, g_m(\mathbf{x}) \geqq \beta_m\} \geqq p, \quad h_1(\mathbf{x}) \geqq 0, \dots, h_M(\mathbf{x}) \geqq 0. \tag{1. 1}$$

Here $\beta_1, \dots, \beta_m$ are random variables, p is a prescribed probability ($0 < p < 1$) and $g_1(\mathbf{x}), \dots, g_m(\mathbf{x}), h_1(\mathbf{x}), \dots, h_M(\mathbf{x}), -f(\mathbf{x})$ are concave functions [1]) in the entire space R^n, where the vectors $\mathbf{x}$ are taken from. If we want to solve Problem (1. 1) numerically then the first thing is to discover the type of the function of the variable $\mathbf{x} \in R^n$:

$$h(\mathbf{x}) = P\{g_1(\mathbf{x}) \geqq \beta_1, \dots, g_m(\mathbf{x}) \geqq \beta_m\}. \tag{1. 2}$$

If this is concave or at least quasi-concave then the numerical solution of Problem (1. 1) is hopeful. We are interested in random variables $\beta_1, \dots, \beta_m$ having a continuous joint probability distribution. Examples show that in the most frequent and practically interesting cases we cannot expect that the function (1. 2) is concave. Surprisingly, however, a special kind of quasi-concavity holds for a wide class of joint probability distributions of the random variables $\beta_1, \dots, \beta_m$. Notably, we show that under some conditions $\log h(\mathbf{x})$ is a concave function in the entire space R^n. This unexpectedly good behaviour of function (1. 2) and problem (1. 1) will result very likely in a frequent application of this and related models.

---

*) This research was supported in part by the Institute of Economic Planning (Budapest).

[1]) From the point of view of numerical solution it is enough to suppose that $h_1(\mathbf{x}), \dots, h_M(\mathbf{x})$ are quasi-concave. A function $h(\mathbf{x})$ defined in a convex set L is quasi-concave if for any $\mathbf{x}_1, \mathbf{x}_2 \in L$ and $0 < \lambda < 1$ we have $h(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) \geqq \min\{h(\mathbf{x}_1), h(\mathbf{x}_2)\}$.

The main theorem in our paper is Theorem 2 which is proved in Section 3. The basic tools for the proof of this theorem are an integral inequality and the Brunn—Minkowski theorem for convex combinations of two convex sets. The integral inequality states that for any measurable non-negative functions f, g we have

$$\int_{-\infty}^{\infty} \sup_{x+y=2t} f(x)g(y)\,dt \geqq \left(\int_{-\infty}^{\infty} f^2(x)\,dx\right)^{\frac{1}{2}} \left(\int_{-\infty}^{\infty} g^2(y)\,dy\right)^{\frac{1}{2}}. \tag{1.3}$$

This will be proved in Section 2.

Let A and B be two convex sets of the space R^n. The Minkowski combination $A+B$ of A and B, and the multiple λA of A (for a real number λ) are defined by $A+B = \{\mathbf{a}+\mathbf{b} \mid \mathbf{a}\in A, \mathbf{b}\in B\}$ and $\lambda A = \{\lambda\mathbf{a} \mid \mathbf{a}\in A\}$.

Theorem of Brunn. *If A and B are bounded convex sets in R^n and* $0<\lambda<1$, *then we have*

$$\mu^{\frac{1}{n}}\{\lambda A+(1-\lambda)B\} \geqq \lambda\mu^{\frac{1}{n}}\{A\}+(1-\lambda)\,\mu^{\frac{1}{n}}\{B\}, \tag{1.4}$$

where μ denotes Lebesgue measure.

This theorem is sharpened by the

Theorem of Brunn—Minkowski. *If the conditions of the theorem of Brunn are fulfilled, moreover both A and B are closed and have positive Lebesgue measures, then equality holds in* (1.4) *if and only if A and B are homothetic.*

Our main theorem contains an inequality similar to that of Brunn. Instead of Lebesgue measure more general measures are involved. Let P be a probability measure [2]) defined on the Borel sets of R^n. We say that the measure P is logarithmic concave if for every convex sets A, B of R^n we have

$$P\{\lambda A+(1-\lambda)B\} \geqq (P\{A\})^{\lambda}(P\{B\})^{1-\lambda} \qquad (0<\lambda<1). \tag{1.5}$$

In section 4 we show that many well-known multivariate probability distributions have this property because they satisfy the conditions of the main theorem.

Inequality (1.5) has an important consequence, namely that the P measure of the parallel shifts of a convex set is a logarithmic concave function of the shift vector. This will be shown in Section 3.

As for the practical applications of the theory presented in this paper the reader is referred to the detailed study [9].

---

[2]) We restrict ourselves to finite measures and, having in mind the applications of our theory, we consider probability measures. The finiteness condition, however, can be dropped as it will be clear from the proofs.

2. An integral inequality. In this section we prove the inequality (1. 3). We formulate it now in the form of a theorem.

Theorem 1. *Let f, g be two non-negative Lebesgue measurable functions defined on the real line R^1. Then the function*

$$r(t) = \sup_{x+y=2t} f(x)g(y) \tag{2. 1}$$

is also measurable and we have

$$\int_{-\infty}^{\infty} r(t)\,dt \geqq \Big(\int_{-\infty}^{\infty} f^2(x)\,dx\Big)^{\frac{1}{2}} \Big(\int_{-\infty}^{\infty} g^2(y)\,dy\Big)^{\frac{1}{2}} \tag{2. 2}$$

(where the value $+\infty$ is also allowed for the integrals).

Proof. First we prove the assertion for such functions f, g which are constant on the subintervals

$$\left[0, \frac{1}{n}\right), \quad \left[\frac{1}{n}, \frac{2}{n}\right), \ldots, \left[\frac{n-2}{n}, \frac{n-1}{n}\right), \quad \left[\frac{n-1}{n}, 1\right]$$

of the interval $[0, 1]$ and vanish elsewhere. Let $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ denote the values of f and g on these subintervals, respectively. Then we have

$$\int_0^1 r(t)\,dt = [A_2 + \max(A_2, A_3) + \cdots + \max(A_{2n-1}, A_{2n}) + A_{2n}]\frac{1}{2n},$$

where

$$A_m = \max_{\substack{i+j=m \\ 1\leqq i,\, j\leqq n}} a_i b_j \qquad (m = 2, 3, \ldots, 2n), \tag{2. 3}$$

and

$$\int_0^1 f^2(x)\,dx = \frac{1}{n}\sum_{i=1}^n a_i^2, \qquad \int_0^1 g^2(y)\,dy = \frac{1}{n}\sum_{i=1}^n b_i^2.$$

Thus the inequality to be proved reduces to the inequality

$$\frac{1}{2}[A_2 + \max(A_2, A_3) + \cdots + \max(A_{2n-1}, A_{2n}) + A_{2n}] \geqq \Big(\sum_{i=1}^n a_i^2\Big)^{\frac{1}{2}} \Big(\sum_{i=1}^n b_i^2\Big)^{\frac{1}{2}} \tag{2. 4}$$

for any sequences of non-negative numbers $a_1, \ldots, a_n$; $b_1, \ldots, b_n$.

First we consider the case

$$a_1 \geqq a_2 \geqq \cdots \geqq a_n, \qquad b_1 \geqq b_2 \geqq \cdots \geqq b_n. \tag{2. 5}$$

This implies $A_2 \geqq A_3 \geqq \cdots \geqq A_{2n}$. It is enough to prove (2. 4) for the special case $a_1 = b_1 = 1$. We prove then that

$$2A_2 + A_3 + \cdots + A_{2n-1} + A_{2n} \geqq \sum_{i=1}^n a_i^2 + \sum_{i=1}^n b_i^2 \tag{2. 6}$$

which is stronger than the required inequality because

$$\frac{1}{2}\left(\sum_{i=1}^{n} a_i^2+\sum_{i=1}^{n} b_i^2\right) \geqq \left(\sum_{i=1}^{n} a_i^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^{n} b_i^2\right)^{\frac{1}{2}}.$$

Let us arrange the numbers $a_2, \ldots, a_n, b_2, \ldots, b_n$ according to their order of magnitude. We may suppose that the first number is a_2. If some a's are equal we keep among these the original ordering and the same is done to the b's. If $a_i = b_j$ for some $i>2$ and $j>1$ then the ordering between these two numbers is b_j, a_i. Let a_r be the first among $a_3, \ldots, a_n$ which is smaller than or equal to b_2. It is possible, of course, that such an a_r does not exist, i.e. $a_n > b_2$. In this case $a_n b_{m-n} \geqq \geqq b_2 b_{m-n} \geqq b_{m-n}^2$ $(m = n+2, \ldots, 2n)$, thus (2.6) follows then from the relations $A_2 = a_1 b_1 = 1$, $A_m \geqq a_{m-1} b_1 = a_{m-1} \geqq a_{m-1}^2$ $(m = 3, \ldots, n+1)$, $A_m \geqq a_n b_{m-n}$ $(m = n+2, \ldots, 2n)$. In the case that a_r exists the following reasoning applies. We associate with each b_j the nearest a to the left: let $a_{i(j)}$ be this number. Similarly, we associate with each a_p the nearest b to the left: let $b_{q(p)}$ be this number. We have

$$a_{i(j)} b_j \geqq b_j^2 \quad (j=2, \ldots, n), \qquad a_p b_{q(p)} \geqq a_p^2 \quad (p=r, \ldots, n).$$

It is easy to see that for any j and p satisfying $2 \leqq j \leqq n$, $r \leqq p \leqq n$, the relation $i(j)+j \neq \neq p+q(p)$ holds. In fact there is no a_p between $a_{i(j)}$ and b_j. Consequently a_p is either to the right from b_j in which case we have $q(p) \geqq j$, $p > i(j)$ or $p \leqq i(j)$ in which case $q(p) < j$. A second remark is that the numbers $i(j)+j$ $(j=2, \ldots, n)$ are different from each other and the same holds for the numbers $p+q(p)$ $(p=r, \ldots, n)$. From these we conclude that

$$A_3 + A_4 + \cdots + A_{2n} \geqq A_3^2 + \cdots + A_r^2 + A_{r+1} + \cdots + A_{2n} \geqq$$

$$\geqq a_2^2 + \cdots + a_{r-1}^2 + \sum_{p=r}^{n} a_p b_{q(p)} + \sum_{j=2}^{n} a_{i(j)} b_j \geqq a_2^2 + \cdots + a_{r-1}^2 + \sum_{p=r}^{n} a_p^2 + \sum_{j=2}^{n} b_j^2.$$

This proves (2. 6) because $A_2 = a_1 b_1 = 1$.

Now we prove that if we perform independent permutations on the numbers (2. 5) then the left hand side of (2. 4) becomes the smallest at the original non-increasing ordering. Let us consider the following scheme (illustrated in the case $n=3$):

$$\begin{array}{ccccc|c} & & a_1 b_1 & & & A_2 \\ & a_1 b_2 & & a_2 b_1 & & A_3 \\ a_1 b_3 & & a_2 b_2 & & a_3 b_1 & A_4 \\ & a_2 b_3 & & a_3 b_2 & & A_5 \\ & & a_3 b_3 & & & A_6 \end{array} \qquad (2.7)$$

with the row maxima at the right hand side. If in the sequence $a_1, \ldots, a_n$ we interchange a_i and a_j then this means in the scheme (2. 7) that the ith and jth northeast-southwest rows are interchanged. The situation is similar if we interchange b_i and b_j in the sequence $b_1, \ldots, b_n$. Under such transformations the horizontal rows interchange some elements. The following assertion is true, however. The kth largest horizontal row maximum in the original scheme is not larger then the kth largest horizontal row maximum of another scheme obtained from the original by some (independent) permutations of the skew rows. In other terms, if $B_2, \ldots, B_{2n}$ are the horizontal row maxima of the transformed scheme and $B_2^*, \ldots, B_{2n}^*$ denote the same numbers but arranged according to their magnitude, i.e. $B_2^* \geqq B_3^* \geqq \cdots \geqq B_{2n}^*$, then

$$A_i \leqq B_i^* \qquad (i=2, \ldots, 2n). \tag{2. 8}$$

In (2. 8) we already took into account that $A_2 \geqq A_3 \geqq \cdots \geqq A_{2n}$. To prove this statement, suppose that the kth largest horizontal row maximum in the original scheme is realized by the element $a_p b_q$. Then in the rectangle

$$\begin{matrix} a_1 b_1 & a_2 b_1 \ldots a_p b_1 \\ a_1 b_2 & a_2 b_2 \ldots a_p b_2 \\ \cdots & \cdots \\ a_1 b_q & a_2 b_q \ldots a_p b_q \end{matrix} \tag{2. 9}$$

which stands skew in the scheme, all numbers are greater than or equal to $a_p b_q$. We remark that $k = p+q-1$. Now it is easy to see that under any permutations of the skew rows of the original scheme, the numbers (2. 9) cannot be condensed into less than $k = p+q-1$ rows. This means

$$B_{k+1}^* \geqq A_{k+1} (= a_p b_q) \qquad (k = 1, \ldots, 2n-1),$$

which are the required inequalities.

We arrived at the final step of the proof of the inequality (2. 4). From relation (2. 8) we conclude

$$A_2 + \sum_{i=2}^{2n} A_i \leqq B_2^* + \sum_{i=2}^{2n} B_i^* = B_2^* + \sum_{i=2}^{2n} B_i.$$

On the other hand we have for an arbitrary sequence of numbers $B_2, \ldots, B_{2n}$,

$$B_2^* + B_2 + \cdots + B_{2n} = \leqq B_2 + \max(B_2, B_3) + \cdots + \max(B_{2n-1}, B_{2n}) + B_{2n},$$

where B_2^* is the largest among $B_2, \ldots, B_{2n}$. Hence it follows for our non-negative numbers

$$\frac{1}{4}\left[A_2 + A_{2n} + \sum_{i=2}^{2n-1} \max(A_i, A_{i+1})\right]^2 = \frac{1}{4}\left[A_2 + \sum_{i=2}^{2n} A_i\right]^2 \leqq$$
$$\leqq \frac{1}{4}\left[B_2 + B_{2n} + \sum_{i=2}^{2n-1} \max(B_i, B_{i+1})\right]^2.$$

This means that the left hand side of (2. 4) is the smallest at the original permutations of the sequences $a_1, \ldots, a_n$; $b_1, \ldots, b_n$.

If f, g are continuous functions in some closed intervals and are equal to 0 elsewhere then these can be uniformly approximated by such functions for which we already proved the integral inequality (2. 2). Thus (2. 2) holds for these functions f, g too.

If f and g are continuous on the entire real line then first we define

$$f_T(x)=f(x) \quad \text{if} \quad |x|\leqq T, \quad \text{and} \quad f_T(x)=0 \quad \text{otherwise,}$$

$$g_T(y)=g(y) \quad \text{if} \quad |y|\leqq T, \quad \text{and} \quad g_T(y)=0 \quad \text{otherwise.}$$

It follows that

$$r(t) = \sup_{x+y=2t} f(x)g(y) \geqq \max_{x+y=2t} f_T(x)g_T(y) = r_T(t).$$

So we have

$$\int_{-\infty}^{\infty} r(t)\,dt \geqq \int_{-\infty}^{\infty} r_T(t)\,dt \geqq \Big(\int_{-\infty}^{\infty} f_T^2(x)\,dx\Big)^{\frac{1}{2}} \Big(\int_{-\infty}^{\infty} g_T^2(y)\,dy\Big)^{\frac{1}{2}},$$

and hence we infer that (2. 2) also holds.

Let us now prove the theorem for arbitrary non-negative Lebesgue measurable functions. It is enough to consider functions which are bounded and equal to zero outside the interval [0, 1]. We may also suppose that both f and g have a finite number of different values. In fact every measurable bounded function can be uniformly approximated by such functions with arbitrary precision.

The measurability of $r(t) = \sup_{x+y=2t} f(x)g(y)$ will be proved as follows. The space R^2 can be subdivided into a finite number of disjoint rectangular Lebesgue measurable sets $E_1, \ldots, E_N$ each of which has the property that the function of two variables $f(x)g(y)$ is constant on it. The sets

$$H_i = \{t \,|\, 2t = x+y,\ (x, y)\in E_i\} \qquad (i=1, \ldots, N)$$

are clearly measurable. If $E_1, \ldots, E_N$ are arranged so that the values of $f(x)g(y)$ follow each other according to the order of magnitude where the largest value is the first, then $r(t)$ is constant on the sets

$$H_i \setminus \bigcup_{j=i+1}^{N} H_j \quad (i = 1, \ldots, N-1), \quad \text{and} \quad H_N,$$

which proves the measurability of $r(t)$.

Let F be the class of functions defined on [0, 1] consisting of all non-negative step functions and all functions which can be obtained in the following way: take any non-negative step function $h(x)$, any sequence of intervals $I_1, I_2, \ldots$ with finite sum of lengths and define

$$\text{(2.10)} \qquad k(x) = 0 \quad \text{if} \quad x \in \bigcup_{k=1}^{\infty} I_k, \quad \text{and} \quad k(x) = h(x) \text{ otherwise.}$$

This class of functions has the property that for any pair f, g in F, inequality (2. 2) holds. This statement is trivial for step functions. If f and g are in F and one of them or both are not step functions then

$$f(x) = \lim_{i\to\infty} f_i(x), \quad g(y) = \lim_{i\to\infty} g_i(y),$$

where f_i, g_i are defined so that on the right hand side of (2. 10) we put $h=f$ resp. $h=g$ and write $\bigcup_{k=1}^{i} I_k$ instead of $\bigcup_{k=1}^{\infty} I_k$. It follows that

$$\sup_{x+y=2t} f(x)g(y) = \max_{x+y=2t} f(x)g(y) = \lim_{i\to\infty} \max_{x+y=2t} f_i(x)g_i(y),$$

whence we conclude

$$\int_0^1 \sup_{x+y=2t} f(x)g(y)\,dt = \lim_{i\to\infty} \int_0^1 \max_{x+y=2t} f_i(x)g_i(y)\,dt \geqq$$

$$\geqq \lim_{i\to\infty} \left(\int_0^1 f_i^2(x)\,dx\right)^{\frac{1}{2}} \left(\int_0^1 g_i^2(y)\,dy\right)^{\frac{1}{2}} = \left(\int_0^1 f^2(x)\,dx\right)^{\frac{1}{2}} \left(\int_0^1 g^2(y)\,dy\right)^{\frac{1}{2}}.$$

As the next and final step in the proof we show that every Lebesgue measurable and finitely valued function defined in [0, 1] is the limit in measure of a sequence of functions $f_i \in F$ $(i=1, 2, \ldots)$, where

$$\text{(2.11)} \qquad f_i(x) \leqq f(x) \qquad (0 \leqq x \leqq 1;\ i=1, 2, \ldots).$$

To prove this we denote by $d_1, \ldots, d_n$ $(d_1 < \cdots < d_n)$ the values of the function f and by $D_1, \ldots, D_n$ those measurable sets where f takes on these values. Let us cover $\bar{D}_j = [0, 1] \setminus D_j$ by a sequence of intervals

$$I_{i1}^{(j)}, \quad I_{i2}^{(j)}, \ldots \qquad (i=1, 2, \ldots;\ j=1, \ldots, n),$$

where the sum of the lengths of these intervals tends to the Lebesgue measure of $\bar{D}_j$ as $i\to\infty$. Let us define f_i in the following manner

$$\text{(2.12)} \quad f_i(x) = d_j \quad \text{if} \quad x \notin \bigcup_{k=1}^{\infty} I_{ik}^{(j)} \qquad (j=1, \ldots, n) \quad \text{and} \quad f_i(x)=0 \text{ otherwise.}$$

For every $i=1, 2, \ldots$ we have $f_i \in F$, (2. 11) is fulfilled, and the sequence (2. 12) converges to f in measure.

If the sequence g_i $(i=1, 2, \ldots)$ is defined in a similar way in connection with g then we conclude

$$\int_0^1 \sup_{x+y=2t} f(x)g(y)\,dt \geqq \int_0^1 \sup_{x+y=2t} f_i(x)g_i(y)\,dt \geqq$$

$$\geqq \Big(\int_0^1 f_i^2(x)\,dx\Big)^{\frac{1}{2}} \Big(\int_0^1 g_i^2(y)\,dy\Big)^{\frac{1}{2}} \to \Big(\int_0^1 f^2(x)\,dx\Big)^{\frac{1}{2}} \Big(\int_0^1 g^2(y)\,dy\Big)^{\frac{1}{2}}.$$

This completes the proof of Theorem 1.

3. The main theorems. The main result of this paper is the following

Theorem 2. *Let $Q(\mathbf{x})$ be a convex function defined on the entire n-dimensional space R^n. Suppose that $Q(\mathbf{x}) \geqq a$, where a is some real number. Let $\psi(z)$ be a function defined on the infinite interval $[a, \infty)$. Suppose that $\psi(z)$ is non-negative, non-increasing, differentiable, and $-\psi'(z)$ is logarithmic concave*[3]*. Consider the function $f(\mathbf{x}) = \psi(Q(\mathbf{x}))$ $(\mathbf{x} \in R^n)$ and suppose that it is a probability density*[4]*, i.e.*

$$\int_{R^n} f(\mathbf{x})\,d\mathbf{x} = 1. \tag{3.1}$$

Denote by $P\{C\}$ the integral of $f(\mathbf{x})$ over the measurable subset C of R^n. If A and B are any two convex sets in R^n, then the following inequality holds:

$$P\{\lambda A + (1-\lambda)B\} \geqq (P\{A\})^{\lambda}(P\{B\})^{1-\lambda} \qquad (0<\lambda<1). \tag{3.2}$$

Remark 1. Condition (3. 1) implies that $\psi(z) \to 0$ as $z \to \infty$. Otherwise $f(\mathbf{x})$ would have a positive lower bound contradicting the finiteness of the integral (3. 1).

Remark 2. We supposed that $Q(\mathbf{x})$ is bounded from below. Dropping this assumption and allowing z to vary on the entire real line, where we suppose that $\psi(z)$ satisfies the same conditions as before, we can deduce from the other assumptions of Theorem 2 that $Q(\mathbf{x})$ is bounded from below.

For if $Q(\mathbf{x})$ were unbounded from below then for every real number b the set

$$\{\mathbf{x} \mid Q(\mathbf{x}) \leqq b\} \tag{3.3}$$

would be unbounded and convex. Consequently the Lebesgue measure of (3. 3) would equal infinity. Now the function $\psi(z)$ cannot vanish everywhere because of

[3]) A function $h(\mathbf{x})$ defined on a convex set K is said to be logarithmic concave if for any $\mathbf{x}, \mathbf{y} \in K$ and $0<\lambda<1$ we have $h(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) \geqq [h(\mathbf{x})]^{\lambda}[h(\mathbf{y})]^{1-\lambda}$.

[4]) It would be enough to suppose that the integral of $f(\mathbf{x})$ is finite on the entire space R^n.

(3. 1). Thus if $Q(\mathbf{x})$ is unbounded from below then $f(\mathbf{x})$ is greater than or equal to a positive number on a set of infinite Lebesgue measure. This contradicts (3. 1).

Remark 3. We may allow $Q(\mathbf{x})$ to take on the value ∞. In this case we require that $\psi(\infty)=0$.

Proof of Theorem 2. Consider the one parameter family of sets

$$E(v)=\{\mathbf{x}\mid f(\mathbf{x})\geqq v\}=\{\mathbf{x}\mid Q(\mathbf{x})\leqq\psi^{-1}(v)\}\qquad (v>0), \tag{3. 4}$$

and the corresponding Lebesgue measures $F(v)=\mu\{E(v)\}$ $(v>0)$. As the integral of $f(\mathbf{x})$ is finite over the entire space R^n it follows that the measures $F(v)$ are finite for every v. Furthermore all non-empty sets $E(v)$ $(v>0)$ are convex, thus they must be bounded as well. Finally, the sets (3. 4) are closed because $Q(\mathbf{x})$ is continuous. The integral of $f(\mathbf{x})$ on R^n can be expressed in the form

$$\int_{R^n} f(\mathbf{x})\,d\mathbf{x} = -\int_0^\infty v\,dF(v) = \int_0^\infty F(v)\,dv, \tag{3. 5}$$

where we have used partial integration and the following formulas

$$F(v)=0\big(v>\psi(a)\big),\quad \lim_{v\to 0} vF(v)=0.$$

The first relation is trivial, the proof of the second relation is given below. For any $\varepsilon>0$ we have

$$-\int_0^\infty v\,dF(v)\geqq-\int_\varepsilon^\infty v\,dF(v)=\varepsilon F(\varepsilon)+\int_\varepsilon^\infty F(v)\,dv\geqq\int_\varepsilon^\infty F(v)\,dv.$$

Thus the integral on the right hand side of (3. 5) is finite. Taking this into account we see from the line above that $\lim_{\varepsilon\to 0}\varepsilon F(\varepsilon)$ exists. This limit cannot be positive as $\int_0^\varepsilon F(v)\,dv$ is finite.

Let us introduce the notations

$$K(v)=\{\mathbf{x}\mid Q(\mathbf{x})\leqq v\},\quad L(v)=\mu\{K(v)\}\qquad(-\infty<v<\infty),$$

where μ is again the symbol of Lebesgue measure. Then, for every $v>0$, $E(v)=K(\psi^{-1}(v))$ and $F(v)=L(\psi^{-1}(v))$. Using this notation we can rewrite (3. 5) in the form

$$\int_{R^n} f(\mathbf{x})\,d\mathbf{x}=\int_0^\infty F(v)\,dv=\int_0^{\psi(a)} L(\psi^{-1}(v))\,dv.$$

Applying the transformation $z=\psi^{-1}(v)$ and observing that $\psi^{-1}(0)=\infty$, we obtain that

$$\int_{R^n} f(\mathbf{x})\,dx = \int_a^\infty L(z)[-\psi'(z)]\,dz.$$

The above reasoning can be applied for an arbitrary measurable subset C of R^n with the difference that instead of $E(v)$, $K(v)$ we have to work with the intersections $E(v)\cap C$ and $K(v)\cap C$. Introducing the notation $L(C, v) = \mu\{K(v)\cap C\}$, we can write

$$\int_C f(\mathbf{x})\,d\mathbf{x} = \int_a^\infty L(C, z)[-\psi'(z)]dz. \tag{3.6}$$

By the convexity of the function $Q(\mathbf{x})$ we have for any $v_1 \geqq a$, $v_2 \geqq a$ and $0<\lambda<1$,

$$K(\lambda v_1+(1-\lambda)v_2) \supset \lambda K(v_1)+(1-\lambda)K(v_2). \tag{3.7}$$

Let A and B be any convex sets in R^n. Considering the Minkowski sum $\lambda A+(1-\lambda)B$ with the same λ as in (3. 7), it is easy to see that

$$K(\lambda v_1+(1-\lambda)v_2)\cap[\lambda A+(1-\lambda)B] \supset \lambda[K(v_1)\cap A]+(1-\lambda)[K(v_2)\cap B].$$

By the Theorem of Brunn,

$$[L(\lambda A+(1-\lambda)B,\ \lambda v_1+(1-\lambda)v_2]^{\frac{1}{n}} \geqq \lambda[L(A, v_1)]^{\frac{1}{n}} +(1-\lambda)[L(B, v_2)]^{\frac{1}{n}}. \tag{3.8}$$

We shall use the following consequence of (3. 8):

$$L(\lambda A+(1-\lambda)B,\ \lambda v_1+(1-\lambda)v_2) \geqq [L(A, v_1)]^{\lambda}[L(B, v_2]^{1-\lambda}. \tag{3.9}$$

The function $-\psi'(z)$ is logarithmic concave in the interval $z\geqq a$; hence for any $v_1\geqq a$, $v_2\geqq a$ we have

$$-\psi'(\tfrac{1}{2}(v_1+v_2)) \geqq [-\psi'(v_1)]^{\frac{1}{2}}[-\psi'(v_2)]^{\frac{1}{2}}. \tag{3.10}$$

Putting $\lambda=\frac{1}{2}$ in (3. 9) and multiplying the inequalities (3. 9), (3. 10) we obtain

$$L(\tfrac{1}{2}A+\tfrac{1}{2}B, \tfrac{1}{2}v_1+\tfrac{1}{2}v_2)[-\psi'(\tfrac{1}{2}v_1+\tfrac{1}{2}v_2)] \geqq$$
$$\geqq \{L(A, v_1)[-\psi'(v_1)]\}^{\frac{1}{2}}\{L(B, v_2)[-\psi'(v_2)]\}^{\frac{1}{2}}.$$

It follows from this that

$$L(\tfrac{1}{2}A+\tfrac{1}{2}B, z)[-\psi'(z)] \geqq \sup_{v_1+v_2=2z} \{L(A, v_1)[-\psi(v_1)]\}^{\frac{1}{2}}\{L(B, v_2)[-\psi'(v_2)]\}^{\frac{1}{2}}. \tag{3.11}$$

Now we apply Theorem I for the functions on the right hand side of (3. 11). First taking into account (3. 11) we conclude the following result

$$\int_a^\infty L(\tfrac{1}{2}A+\tfrac{1}{2}B, z)[-\psi'(z)]\,dz \geqq$$

$$\geqq \int_a^\infty \sup_{v_1+v_2=2z} \{L(A, v_1)[-\psi'(v_1)]\}^{\frac{1}{2}} \{L(B, v_2)[-\psi'(v_2)]\}^{\frac{1}{2}}\,dz \geqq$$

$$\geqq \left\{\int_a^\infty L(A, v_1)[-\psi'(v_1)]\,dv_1\right\}^{\frac{1}{2}} \left\{\int_a^\infty L(B, v_2)[-\psi'(v_2)]\,dv_2\right\}^{\frac{1}{2}}.$$

In view of (3. 6) this means

$$P\{\tfrac{1}{2}A+\tfrac{1}{2}B\} = \int_{\frac{1}{2}A+\frac{1}{2}B} f(\mathbf{x})\,d\mathbf{x} \geqq \left[\int_A f(\mathbf{x})\,d\mathbf{x}\right]^{\frac{1}{2}} \left[\int_B f(\mathbf{x})\,d\mathbf{x}\right]^{\frac{1}{2}} = [P\{A\}]^{\frac{1}{2}}[P\{B\}]^{\frac{1}{2}}. \tag{3.12}$$

Thus inequality (3. 2) is proved for $\lambda = \frac{1}{2}$.

The assertion for the case of an arbitrary λ can be deduced from here by a continuity argument. First we remark that if A_1, A_2, A_3, A_4 are arbitrary convex sets in R^n then (3. 12) implies

$$P\{\tfrac{1}{4}A_1+\tfrac{1}{4}A_2+\tfrac{1}{4}A_3+\tfrac{1}{4}A_4\} = P\{\tfrac{1}{2}(\tfrac{1}{2}A_1+\tfrac{1}{2}A_2)+\tfrac{1}{2}(\tfrac{1}{2}A_3+\tfrac{1}{2}A_4)\} \geqq$$

$$\geqq [P\{\tfrac{1}{2}A_1+\tfrac{1}{2}A_2\}]^{\frac{1}{2}}[P\{\tfrac{1}{2}A_3+\tfrac{1}{2}A_4\}]^{\frac{1}{2}} \geqq [P\{A_1\}]^{\frac{1}{4}}[P\{A_2\}]^{\frac{1}{4}}[P\{A_3\}]^{\frac{1}{4}}[P\{A_4\}]^{\frac{1}{4}}.$$

A similar inequality holds for any convex sets C_i $(i=1, \dots, 2^N)$, where N is a positive integer. Define the sets

$$A_i = A \quad (i=1, \dots, j), \qquad B_i = B \quad (i=1, \dots, k),$$

where we suppose that $j+k$ is a power of 2, furthermore

$$\lim_{j,k\to\infty} \frac{j}{j+k} = \lambda. \tag{3.13}$$

Let $j+k = 2^N$. It follows that

$$P\left\{\frac{A_1+\dots+A_j+B_1+\dots+B_k}{2^N}\right\} = P\left\{\frac{j}{2^N}\frac{A_1+\dots+A_j}{j}+\frac{k}{2^N}\frac{B_1+\dots+B_k}{k}\right\} =$$

$$= P\left\{\frac{j}{2^N}A+\frac{k}{2^N}B\right\} \tag{3.14}$$

because A and B are convex sets. On the other hand we have

$$P\{2^{-N}(A_1+\cdots+A_j+B_1+\cdots+B_k)\} \geqq \left(\prod_{i=1}^{j} P\{A_i\}\right)^{2^{-N}} \left(\prod_{i=1}^{k} P\{B_i\}\right)^{2^{-N}} = \tag{3.15}$$
$$= (P\{A\})^{j2^{-N}}(P\{B\})^{k2^{-N}}.$$

Comparing (3. 14) and (3. 15) we conclude

$$P\left\{\frac{j}{2^N}A+\frac{k}{2^N}B\right\} \geqq (P\{A\})^{j2^{-N}}(P\{B\})^{k2^{-N}}. \tag{3.16}$$

Taking into account (3. 16) and the continuity in λ of the function $P\{\lambda A+(1-\lambda)B\}$, we see that (3. 2) holds for arbitrary $0<\lambda<1$. Thus the proof of Theorem 2 is complete.

Theorem 3. *Let $f(\mathbf{x})=\psi(Q(\mathbf{x}))$ be a probability density in R^n satisfying the conditions of Theorem 2 and $A\subset R^n$ a convex set. Then the function*

$$h(\mathbf{t}) = P\{A+\mathbf{t}\} \quad (\mathbf{t}\in R^n) \tag{3.17}$$

is logarithmic concave in R^n.

Proof. Let $\mathbf{t}_1$, $\mathbf{t}_2$ be arbitrary vectors in R^n and let $0<\lambda<1$. Then we have

$$\lambda(A+\mathbf{t}_1)+(1-\lambda)(A+\mathbf{t}_2) = A+[\lambda\mathbf{t}_1+(1-\lambda)\mathbf{t}_2].$$

In fact if $\mathbf{x}\in A$, $\mathbf{y}\in A$ then

$$\lambda(\mathbf{x}+\mathbf{t}_1)+(1-\lambda)(\mathbf{y}+\mathbf{t}_2) = [\lambda\mathbf{x}+(1-\lambda)\mathbf{y}]+[\lambda\mathbf{t}_1+(1-\lambda)\mathbf{t}_2]$$

and we supposed that A is convex. Thus by Theorem 2

$$P\{A+[\lambda\mathbf{t}_1+(1-\lambda)\mathbf{t}_2]\} = P\{\lambda(A+\mathbf{t}_1)+(1-\lambda)(A+\mathbf{t}_2)\} \geqq$$
$$\geqq (P\{A+\mathbf{t}_1\})^{\lambda}(P\{A+\mathbf{t}_2\})^{1-\lambda},$$

which means that

$$h(\lambda\mathbf{t}_1+(1-\lambda)\mathbf{t}_2) \geqq [h(\mathbf{t}_1)]^{\lambda}[h(\mathbf{t}_2)]^{1-\lambda}.$$

Theorem 4. *Let $F(\mathbf{x})$ be a continuous multivariate probability distribution function the probability density of which is of the form $f(\mathbf{x})=\psi(Q(\mathbf{x}))$ and satisfies the conditions of Theorem 2. Then $F(\mathbf{x})$ is a logarithmic concave function in R^n.*

Proof. Apply Theorem 3 to the set $A=\{\mathbf{z}|\mathbf{z}\leqq\mathbf{0}\}$ and take into account that $F(\mathbf{x}) = P\{A+\mathbf{x}\}$ for $\mathbf{x}\in R^n$.

4. Examples of probability measures satisfying the conditions of Theorem 1. The most important multivariate probability distribution is the normal distribution. Its density is given by

$$(4.1)\qquad f(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{n}{2}}|C|^{\frac{1}{2}}} e^{-\frac{1}{2}(\mathbf{x}-\mathbf{m})'C^{-1}(\mathbf{x}-\mathbf{m})} \quad (\mathbf{x}\in R^n),$$

where $\mathbf{m}\in R^n$ is an arbitrary vector and C is a positive definite matrix the determinant of which is denoted by $|C|$. Vectors are considered as column matrices as well and the prime means transpose. This function satisfies the conditions of Theorem 2. In fact $f(\mathbf{x})$ can be written as

$$f(\mathbf{x}) = \psi(Q(\mathbf{x})) \quad (\mathbf{x}\in R^n)$$

with

$$(4.2)\qquad \psi(z) = Ke^{-z^{\alpha}} \quad (z \geqq 0) \quad \text{and} \quad Q(\mathbf{x}) = \left[\frac{1}{2}(\mathbf{x}-\mathbf{m})'C^{-1}(\mathbf{x}-\mathbf{m})\right]^{1/\alpha},$$

where α is any fixed number satisfying $1 \leqq \alpha \leqq 2$ further K is the constant standing on the right hand side in (4. 1). That $\psi(z)$ has the required property, is trivial. Only $Q(\mathbf{x})$ needs a remark. It is well known that a function

$$(\mathbf{x}'D\mathbf{x})^{\frac{1}{2}} \qquad (\mathbf{x}\in R^n)$$

is convex in the entire space provided D is positive semidefinite. This implies the convexity of $Q(\mathbf{x})$ in (4. 2).

Three further probability distributions will be discussed. In all cases we shall show that the probability densities are logarithmic concave in the entire space R^n.

The probability density $f(X)$ of the Wishart distribution is defined by

$$f(X) = \frac{|X|^{\frac{N-p-2}{2}} e^{-\frac{1}{2}\mathrm{Sp}\, C^{-1}X}}{2^{\frac{N-1}{2}p} \pi^{\frac{p(p-1)}{4}} |C|^{\frac{N-1}{2}} \prod_{i=1}^{p} \Gamma\left(\frac{N-i}{2}\right)}$$

if X is positive definite, and $f(X)=0$ otherwise. Here C and X are $p\times p$ matrices, C is fixed and positive definite while X contains the variables. In view of the symmetry of the matrix, the number of independent variables is $n = \frac{1}{2}\, p(p+1)$. We suppose that $N \geqq p+2$. It is well known that the set of positive definite [5] $p\times p$ matrices is convex in the $n = \frac{1}{2}p(p+1)$-dimensional space.

[5]) Any positive definite (or semi-definite) matrix is supposed to be symmetrical in this paper.

We show that $f(X)$ is logarithmic concave on this set [6]). To this it is enough to remark that for any $0<\lambda<1$ and any pair X_1, X_2 of positive definite matrices the inequality

$$|\lambda X_1+(1-\lambda)X_2| \geqq |X_1|^{\lambda}|X_2|^{1-\lambda} \tag{4.3}$$

holds, where we have a strict inequality if $X_1 \neq X_2$ (see [1]).

The multivariate beta distribution has the probability density $f(X)$ defined by

$$f(X) = \frac{c(n_1, p)c(n_2, p)}{c(n_1+n_2, p)} |X|^{\frac{1}{2}(n_1-p-1)}|I-X|^{\frac{1}{2}(n_2-p-1)},$$

if X and $I-X$ are positive definite, and $f(X)=0$ otherwise (see [7]), where

$$\frac{1}{c(k, p)} = 2^{\frac{pk}{2}} \pi^{\frac{p(p-1)}{4}} \prod_{i=1}^{p} \Gamma\left(\frac{k-i+1}{2}\right),$$

I is the unit matrix, I and X are of order $p \times p$, p is a positive integer. We suppose that $n_1 \geqq p+1$, $n_2 \geqq p+1$. The number of independent variables of the function $f(X)$ is equal to $n = \frac{1}{2}p(p+1)$.

It is clear that the set of positive definite matrices X for which $I-X$ is also positive definite, is convex. The function $f(X)$ is zero outside this set and is logarithmic concave on this set which can be seen very easily on the basis of (4. 3).

Finally we consider the Dirichlet distribution (see e.g. [11]) the probability density of which is given by the formula

$$f(\mathbf{x}) = Kx_1^{p_1-1} \dots x_n^{p_n-1}(1-x_1 \cdots -x_n)^{p_{n+1}-1}$$

if $x_i>0$ $(i=1, \dots, n)$, $x_1+\cdots+x_n < 1$, and $f(\mathbf{x})=0$ otherwise. Here we have set $K = \dfrac{\Gamma(p_1+\cdots+p_{n+1})}{\Gamma(p_1)\cdots\Gamma(p_{n+1})}$. The logarithm of this function in the positivity domain is

$$\log f(\mathbf{x}) = \log K+\sum_{i=1}^{n}(p_i-1)\log x_i+(p_{n+1}-1)\log(1-x_1-\cdots-x_n). \tag{4.4}$$

We suppose that $p_i \geqq 1$ $(i=1, \dots, n+1)$. This implies that the function (4. 4) is concave. In fact the second term is trivially concave while $\log(1-x_1-\cdots-x_n)$ is an increasing and concave function of a linear function. Hence the assertion.

5. Application to stochastic programming. Let us now return to Problem (1.1) and consider the x-function in the first constraint which is given separately in (1. 2). We show if the random variables $\beta_1, \dots, \beta_m$ have a continuous joint distribution

[6]) If a function is logarithmic concave on a convex set and equal to zero elsewhere then the function is logarithmic concave in the entire space.

satisfying the conditions of Theorem 2, then the function $h(\mathbf{x})$ is logarithmic concave in the entire space R^n. We recall that the functions $g_1(\mathbf{x}), \dots, g_m(\mathbf{x})$ are supposed to be concave in R^n.

Let $\mathbf{x}, \mathbf{y} \in R^n$ and $0<\lambda<1$. In view of the concavity of the functions $g_1(\mathbf{x}), \dots, g_m(\mathbf{x})$ we have

$$g_i(\lambda\mathbf{x}+(1-\lambda)\mathbf{y}) \geqq \lambda g_i(\mathbf{x})+(1-\lambda)g_i(\mathbf{y}) \qquad (i=1, \dots, m).$$

The function $P\{\beta_1 \leqq z_1, \dots, \beta_m \leqq z_m\}$ of the variables $z_1, \dots, z_m$ is logarithmic concave by Theorem 4, and also a probability distribution function; hence it is monotonic non-decreasing in all variables. Taking these into account we conclude

$$h(\lambda\mathbf{x}+(1-\lambda)\mathbf{y}) = P\{g_1(\lambda\mathbf{x}+(1-\lambda)\mathbf{y}) \geqq \beta_1, \dots, g_m(\lambda\mathbf{x}+(1-\lambda)\mathbf{y}) \geqq \beta_m\} \geqq$$
$$\geqq P\{\lambda g_1(\mathbf{x})+(1-\lambda)g_1(\mathbf{y}) \geqq \beta_1, \dots, \lambda g_m(\mathbf{x})+(1-\lambda)g_m(\mathbf{y}) \geqq \beta_m\} \geqq$$
$$\geqq [P\{g_1(\mathbf{x}) \geqq \beta_1, \dots, g_m(\mathbf{x}) \geqq \beta_m\}]^\lambda [P\{g_1(\mathbf{y}) \geqq \beta_1, \dots, g_m(\mathbf{y}) \geqq \beta_m\}]^{1-\lambda} =$$
$$= [h(\mathbf{x})]^\lambda [h(\mathbf{y})]^{1-\lambda};$$

what was to be proved.

Considering Problem (1. 1), we may take the logarithm of both sides of the first constraint. Then we obtain a convex programming problem. For some reason we may leave it in the original form (the computational solution may prefer this form), then we have a quasi-convex programming problem because a logarithmic concave function is always quasi-concave. Any of these versions can be solved by non-linear programming methods (see e. g. [4], [8], [12]). We emphasize again that this short remark concerning the application of the theory presented in this paper is just for illustration and to disclose the origin of the problem.

Bibliography

[1] E. F. Beckenbach and R. Bellman, *Inequalities*, Springer-Verlag (Berlin—Göttingen—Heidelberg, 1961).

[2] H. Brunn, Über Ovale und Eiflächen, *Inaugural-Dissertation* (München, 1887).

[3] G. L. Dirichlet, Sur une nouvelle méthode pour la détérminations des intégrales multiples, *C. R. Acad. Sci. Paris*, **8** (1839), 156—160.

[4] A. V. Fiacco and G. P. McCormick, *Nonlinear programming: sequential unconstrained minimization techniques*, Wiley (New York—London, 1968).

[5] H. Hadwiger, *Vorlesungen über Inhalt, Oberfläche und Isoperimetrie*, Springer-Verlag (Berlin—Göttingen—Heidelberg, 1957).

[6] H. Minkowski, *Geometrie der Zahlen*, Teubner (Leipzig—Berlin, 1910).

[7] I. Olkin and H. Rubin, Multivariate beta distributions and independence properties of the Wishart distribution, *Annals Math. Stat.*, **35** (1964), 261—269.

[8] A. Prékopa, On probabilistic constrained programming, *Proc. Princeton Symp. Math. Programming* (Princeton, N. J., 1970), 113—138.

[9] A. PRÉKOPA, Contributions to the theory of stochastic programming, *Math. Programming* (to appear).

[10] B. SZ.-NAGY, *Introduction to Real Functions and Orthogonal Expansions*, Akadémiai Kiadó (Budapest, 1964).

[11] S. S. WILKS, *Mathematical Statistics*, Wiley (New York—London, 1962).

[12] J. WISHART, The generalized product moment distribution in samples from a normal multivariate population, *Biometrika*, **20A** (1928), 32—52.

[13] G. ZOUTENDIJK, *Methods of feasible directions*, Elsevier (Amsterdam—London—New York—Princeton, 1960).

TECHNOLOGICAL UNIVERSITY, BUDAPEST, AND
COMPUTING CENTER, HUNGARIAN ACADEMY OF SCIENCES

(Received June 15, 1970)