Math. Operationsforsch. u. Statist. **5** (1974), Heft 2. pp. 109–116.

PROGRAMMING UNDER PROBABILISTIC CONSTRAINTS WITH A RANDOM TECHNOLOGY MATRIX

András Prékopa
Technological University of Budapest and Computer and Automation Institute of the Hungarian Academy of Sciences, H–1502 Budapest XI. Kende utca 13–17, Hungary

Eingereicht bei der Redaktion: 26. 2. 1973

Abstract

Probabilistic constraint of the type $\mathsf{P}\ (A\mathbf{x} \le \beta) \ge p$ is considered and it is proved that under some conditions the constraining function is quasi-concave. The probabilistic constraint is embedded into a mathematical programming problem of which the algorithmic solution is also discussed.

1 Introduction

We consider the following function of the variable $\mathbf{x}$:

$$G(\mathbf{x}) = \mathsf{P}(A\mathbf{x} \le \boldsymbol{\beta}), \tag{1.1}$$

where A is an $m \times n$ matrix, $\mathbf{x}$ is an n-component, $\boldsymbol{\beta}$ is an m-component vector and it is supposed that the entries of A as well as the components of $\boldsymbol{\beta}$ are random variables. The function (1.1) will be used as a constraining function in a nonlinear programming problem which we formulate below:

$$\begin{aligned} &G(\mathbf{x}) \ \ge p, \\ &G_i(\mathbf{x}) \ge 0, \qquad i = 1, \dots, N, \\ &\mathbf{a}_i\mathbf{x} \quad \ge b_i, \qquad i = 1, \dots, M, \\ &\min f(\mathbf{x}). \end{aligned} \tag{1.2}$$

Here we suppose that $\mathbf{a}_1, \dots, \mathbf{a}_M$ are constant n-component vectors, $b_1, \dots, b_M$ are constants, the functions $G_1, \dots, G_M$ are quasi-concave in $\mathbb{R}^n$ and $f(\mathbf{x})$ is convex in $\mathbb{R}^n$. Many practical problems lead to the stochastic programming decision problem (1.2). Among them we mention the nutrition problem in which case the nutrient contents of the foods are random and the ore buying problem in which case the ferrum contents of the ores are random and in both cases (1.2) is a reasonable decision principle. Regarding the joint distribution of the entries of A and the components of $\boldsymbol{\beta}$, two special cases will be considered. In the first one we suppose that this distribution is normal satisfying further restrictive conditions while in the second one lognormal or some more general distribution is involved. The reason why we deal with these conditions is of mathematical nature: we can prove in these cases the quasi-concavity[1] of the constraining function in the probabilistic constraint.

In Section 2 we prove three theorems and in Section 3 the algorithmic solution of Problem (1.2) is discussed.

The theorems of Section 2 are based on the following theorems of the author proved in [2] resp. [3].

THEOREM 1 *Let $f(\mathbf{x})$ be a probability density function in the n-dimensional* EUCLID*ean space $\mathbb{R}^n$ and suppose that it is logarithmic concave i.e. it has the form*

$$f(\mathbf{x}) = e^{-Q(\mathbf{x})}, \qquad \mathbf{x} \in \mathbb{R}^n, \tag{1.3}$$

where $Q(\mathbf{x})$ is a convex function in the entire space. Let further A, B be two convex subsets of $\mathbb{R}^n$ and $0 < \lambda < 1$. Denoting by $\mathsf{P}(C)$ the integral of $f(\mathbf{x})$ over the measurable subset C of $\mathbb{R}^n$, we have

$$\mathsf{P}(\lambda A + (1-\lambda)B) \geq [\mathsf{P}(A)]^{\lambda}[\mathsf{P}(B)]^{1-\lambda}. \tag{1.4}$$

(The constant multiple of a set and the MINKOWSKI sum of two sets are defined by $kC = \{k\mathbf{c} \mid \mathbf{c} \in C\}$ resp. $C + D = \{\mathbf{c} + \mathbf{d} \mid \mathbf{c} \in C,\ \mathbf{d} \in D\}$, where k is a real number.)

Theorem 1 implies that if A is a convex subset of $\mathbb{R}^n$ then $\mathsf{P}(A + \mathbf{x})$ is a logarithmic concave function in the entire space. If we take in particular $A = \{\mathbf{t} \mid \mathbf{t} \leq \mathbf{0}\}$, then we see that the probability distribution function $F(\mathbf{x})$ belonging to the density function $f(\mathbf{x})$ is logarithmic concave in the entire space, since $\mathsf{P}(A + \mathbf{x}) = F(\mathbf{x})$.

A probability measure P in $\mathbb{R}^n$ is said to be *logarithmic concave* if for every pair A, B of convex subsets of $\mathbb{R}^n$ and for every $0 < \lambda < 1$, the inequality (1.4) is satisfied (see [2]).

THEOREM 2 *Let $g_i(\mathbf{x}, \mathbf{y})$, $i = 1, \ldots, r$ be convave functions in $\mathbb{R}^{m+n}$ where $\mathbf{x}$ is an n-component and $\mathbf{y}$ is an m-component vector. Let $\boldsymbol{\xi}$ be an m-component random vector having a logarithmic concave probability distribution. Then the function of the variable $\mathbf{x}$:*

$$\mathsf{P}(g_i(\mathbf{x}, \boldsymbol{\xi}) \geq 0, \quad i = 1, \ldots, r) \tag{1.5}$$

is logarithmic concave in the space $\mathbb{R}^n$.

The formulation of Theorem 2 is slightly different from that given in [3]. Its proof is the same as the proof of the original version only a trivial modification is necessary.

[1]A function $F(\mathbf{x})$ defined on a convex set K is said to be *quasi-concave* if for every pair $\mathbf{x}_1, \mathbf{x}_2 \in K$ we have $F(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) \geq \min(F(\mathbf{x}_1), F(\mathbf{x}_2))$. It is easy to see that a necessary and sufficient condition for this is that the set $\{\mathbf{x} \mid \mathbf{x} \in K,\ F(\mathbf{x}) \geq b\}$ is convex for every real b.

2 Quasi-concavity theorems concerning the function (1.1)

Let $\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_n$ denote the columns of the matrix A and introduce the notation $\boldsymbol{\xi}_{n+1} = -\boldsymbol{\beta}$.

THEOREM 3 *Suppose that the altogether $m(n+1)$ components of the random vectors $\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_{n+1}$ have a joint normal distribution where the cross-covariance matrices of $\boldsymbol{\xi}_i$ and $\boldsymbol{\xi}_j$ are constant multiples of a fixed covariance matrix C i.e.*

$$\mathsf{E}[(\boldsymbol{\xi}_i - \boldsymbol{\mu}_i)(\boldsymbol{\xi}_j - \boldsymbol{\mu}_j)'] = s_{ij}C, \qquad i,j = 1, \dots, n+1, \tag{2.1}$$

where

$$\boldsymbol{\mu}_i = \mathsf{E}(\boldsymbol{\xi}_i), \qquad i = 1, \dots, n+1. \tag{2.2}$$

Then the set of $\mathbf{x}$ *vectors satisfying*

$$\mathsf{P}(A\mathbf{x} \le \boldsymbol{\beta}) \ge p \tag{2.3}$$

is convex provided $p \ge 1/2$.

PROOF. Consider the covariance matrix of $A\mathbf{x} - \boldsymbol{\beta}x_{n+1}$ which is equal to the following

$$\mathsf{E}\left[\left(\sum_{i=1}^{n+1}\boldsymbol{\xi}_i x_i - \sum_{i=1}^{n+1}\boldsymbol{\mu}_i x_i\right)\left(\sum_{i=1}^{n+1}\boldsymbol{\xi}_i x_i - \sum_{i=1}^{n+1}\boldsymbol{\mu}_i x_i\right)'\right] = \sum_{i,j=1}^{n+1} s_{ij}x_i x_j C = \mathbf{z}'S\mathbf{z}C, \tag{2.4}$$

where $\mathbf{z}' = (\mathbf{x}', x_{n+1})$ and S is the matrix with entries s_{ij}. If all elements of C are zero, then the vectors $\boldsymbol{\xi}_1, \dots \boldsymbol{\xi}_{n+1}$ are constants with probability 1 and the statement of the theorem holds trivially. If there is at least one nonzero element in C, then there is at least one positive element in its main diagonal. This implies that $\mathbf{z}'S\mathbf{z} \ge 0$ for every $\mathbf{z}$ since (2.4) is a covariance matrix hence all elements in its main diagonal are nonnegative. Thus S is a positive semidefinite matrix. We may suppose that both C and S are positive definite matrices because otherwise we can use a perturbation of the type $C + \varepsilon I_m$, $S + \varepsilon I_{n+1}$, where I_m and I_{n+1} are unit matrices of the size $m \times m$ and $(n+1) \times (n+1)$, respectively, prove the theorem for these positive definite matrices and then take the limit $\varepsilon \to 0$.

Let $x_{n+1} = 1$ everywhere in the sequel. We have

$$\mathsf{P}(A\mathbf{x} \le \boldsymbol{\beta}) = \mathsf{P}\left(\frac{\sum_{i=1}^{n+1}(\boldsymbol{\xi}_i - \boldsymbol{\mu}_i)x_i}{(\mathbf{z}'S\mathbf{z})^{\frac{1}{2}}} \le -\frac{\sum_{i=1}^{n+1}\boldsymbol{\mu}_i x_i}{(\mathbf{z}'S\mathbf{z})^{\frac{1}{2}}}\right). \tag{2.5}$$

Introduce the notation

$$\mathbf{L}(\mathbf{z}) = -\sum_{i=1}^{n+1}\boldsymbol{\mu}_i x_i$$

and denote $L_i(\mathbf{z})$ the ith component of $\mathbf{L}(\mathbf{z})$ for $i = 1, \ldots, m$. If c_{ik} denote the elements of C then (2.5) can be rewritten as

$$\mathsf{P}(A\mathbf{x} \le \boldsymbol{\beta}) = \Phi\left(\frac{L_1(\mathbf{z})}{(c_{11}\mathbf{z}'S\mathbf{z})^{\frac{1}{2}}}, \ldots, \frac{L_m(\mathbf{z})}{(c_{mm}\mathbf{z}'S\mathbf{z})^{\frac{1}{2}}}; R\right), \tag{2.6}$$

where R is the correlation matrix belonging to the covariance matrix C. For every i $(i = 1, \ldots, m)$ we have

$$\Phi\left(\frac{L_i(\mathbf{z})}{(c_{ii}\mathbf{z}'S\mathbf{z})^{\frac{1}{2}}}\right) \ge \Phi\left(\frac{L_1(\mathbf{z})}{(c_{11}\mathbf{z}'S\mathbf{z})^{\frac{1}{2}}}, \ldots, \frac{L_m(\mathbf{z})}{(c_{mm}\mathbf{z}'S\mathbf{z})^{\frac{1}{2}}}; R\right). \tag{2.7}$$

Thus if the right hand side of (2.7) is greater than or equal to p $(\ge 1/2)$ then we conclude $L_i(\mathbf{z}) \ge 0$, $i = 1, ..., m$. It is well-known that $(\mathbf{z}'S\mathbf{z})^{\frac{1}{2}}$ is a convex function in the entire space $\mathbb{R}^{n+1}$. Hence it follows that for every $\mathbf{z}_1, \mathbf{z}_2$,

$$\frac{\mathbf{L}\left(\frac{1}{2}\mathbf{z}_1 + \frac{1}{2}\mathbf{z}_2\right)}{\left[\left(\frac{1}{2}\mathbf{z}_1 + \frac{1}{2}\mathbf{z}_2\right)' S \left(\frac{1}{2}\mathbf{z}_1 + \frac{1}{2}\mathbf{z}_2\right)\right]^{\frac{1}{2}}} \ge \lambda \frac{\mathbf{L}(\mathbf{z}_1)}{(\mathbf{z}_1'S\mathbf{z}_1)^{\frac{1}{2}}} + (1-\lambda)\frac{\mathbf{L}(\mathbf{z}_2)}{(\mathbf{z}_2'S\mathbf{z}_2)^{\frac{1}{2}}}, \tag{2.8}$$

where

$$\lambda = \frac{(\mathbf{z}_1'S\mathbf{z}_1)^{\frac{1}{2}}}{(\mathbf{z}_1'S\mathbf{z}_1)^{\frac{1}{2}} + (\mathbf{z}_2'S\mathbf{z}_2)^{\frac{1}{2}}}.$$

The probability density function of any nondegenerated normal distribution is logarithmic concave in the entire space. Hence it follows that

$$\log \Phi\left(v_1 c_{11}^{-\frac{1}{2}}, \ldots, v_m c_{mm}^{-\frac{1}{2}}; R\right)$$

is concave in the variables $v_1, \ldots, v_m$. Thus if the value of the function (2.6) is greater than or equal to p $(\ge 1/2)$ for $\mathbf{z} = \mathbf{z}_1$ and also for $\mathbf{z} = \mathbf{z}_2$, then the value of (2.6) will be greater than or equal to p $(\ge 1/2)$ at the point determined by the right hand side of (2.8). Since Φ is monotonically increasing in all variables it follows from (2.8) that (2.6) is greater than or equal to p $(\ge 1/2)$ at the point determined by the left hand side of (2.8). This proves the theorem. $\square$

THEOREM 4 *Let A_i denote the i-th row of A and β_i denote the i-th component of $\boldsymbol{\beta}$, $i = 1, \ldots, m$. Suppose that the random row vectors of $n+1$ components*

$$(A_i - \beta_i), \qquad i = 1, \ldots m \tag{2.9}$$

are independent, normally distributed and their covariance matrices are constant multiples of a fixed covariance matrix C. Then the set of $\mathbf{x}$ vectors on which the function (1.1) *is greater than or equal to p, is convex provided $p \ge 1/2$.*

Proof. We may suppose that C is positive definite. We also suppose that the covariance matrix of (2.9) equals $s_i^2 C$ with $s_i > 0$, $i = 1, \ldots, m$. Let $A_i^{(0)}$, b_i be the expectation of A_i and β_i, respectively, $i = 1, \ldots, m$ and introduce the notation:

$$L_i(\mathbf{z}) = \frac{1}{s_i}\left(-A_i^{(0)}, b_i\right)\mathbf{z}, \quad i = 1, \ldots, m \text{ where } \mathbf{z}' = (x_1, \ldots, x_n, 1).$$

Then we have

$$\begin{aligned} \mathsf{P}(A\mathbf{x} \leq \boldsymbol{\beta}) &= \prod_{i=1}^{m} \mathsf{P}((A_i, -\beta_i)\mathbf{z} \leq 0) = \prod_{i=1}^{m} \mathsf{P}\left(\frac{(A_i - A_i^{(0)}, -\beta_i + b_i)\mathbf{z}}{s_i(\mathbf{z}'C\mathbf{z})^{\frac{1}{2}}} \leq \frac{(-A_i^{(0)}, b_i)\mathbf{z}}{s_i(\mathbf{z}'C\mathbf{z})^{\frac{1}{2}}}\right) \\ &= \prod_{i=1}^{m} \Phi\left(\frac{L_i(\mathbf{z})}{(\mathbf{z}'C\mathbf{z})^{\frac{1}{2}}}\right) = \Phi\left(\frac{L_1(\mathbf{z})}{(\mathbf{z}'C\mathbf{z})^{\frac{1}{2}}}, \ldots, \frac{L_m(\mathbf{z})}{(\mathbf{z}'C\mathbf{z})^{\frac{1}{2}}}; I_m\right) \geq p, \end{aligned}$$

where I_m is the $m \times m$ unit matrix. From here the proof goes in the same way as the proof of Theorem 3.

Let a_{ij}, $i = 1, \ldots, m$; $j = 1, \ldots, n$ denote the elements of the matrix A and for the sake of simplicity suppose that the columns of A are numbered so that those come first which contain random variables. Let r be the number of these columns and introduce the following notations:

- J denotes the set of those ordered pairs (i, j) for which a_{ij} is random variable, $1 \leq i \leq m$, $1 \leq j \leq r$;
- J_i denotes the set of those subscripts j, for which $(i, j) \in J$, where $1 \leq i \leq m$;
- K_i denotes the set $\{1, \ldots, r\} - J_i$;
- L denotes the set of those subscripts i, for which β_i is a random variable, where $1 \leq i \leq m$;
- T denotes the set $\{1, \ldots, m\} - L$.

Now we prove the following

Theorem 5 *Suppose that the random variables a_{ij}, $(i, j) \in J$ are positive with probability* 1 *and the constants a_{ij}, $1 \leq i \leq m$, $1 \leq j \leq r$, $(i, j) \notin J$ are nonnegative. Suppose further that the joint distribution of the random variables*

$$\alpha_{ij}, \quad (i, j) \in J, \quad \beta_i, \quad i \in L$$

is a logarithmic concave probability distribution, where $\alpha_{ij} = \log a_{ij}$, $(i, j) \in J$. Under these conditions the function

$$h(x_1, \ldots, x_r, x_{r+1}, \ldots, x_n) = G(e^{x_1}, \ldots, e^{x_r}, x_{r+1}, \ldots, x_n) \tag{2.10}$$

is logarithmic concave in the entire space $\mathbb{R}^n$, where $G(\mathbf{x})$ is the function defined by (1.1).

PROOF. Consider the following functions

$$
\begin{aligned}
g_i &= -\sum_{j\in J_i} e^{u_{ij}+x_j} - \sum_{j\in K_i} a_{ij}e^{x_j} - \sum_{j=r+1}^{n} a_{ij}x_j + v_i, \qquad i \in L,\\
g_i &= -\sum_{j\in J_i} e^{u_{ij}+x_j} - \sum_{j\in K_i} a_{ij}e^{x_j} - \sum_{j=r+1}^{n} a_{ij}x_j + b_i, \qquad i \in T.
\end{aligned}
\tag{2.11}
$$

These are all supposed to be functions of the variables

$$u_{ij},\ (i,j)\in J; \quad x_j, \quad 1\le j\le n; \quad v_i,\ i\in L,$$

though not every variable appears explicitely in every function. The functions $g_1,\dots,g_m$ are clearly concave in the entire space. Substituting u_{ij} by α_{ij} for every $(i,j)\in J$ and v_i by β_i for every $i\in L$, instead of the functions $g_1,\dots,g_m$ we obtain random variables which we denote by $\gamma_1,\dots,\gamma_m$, respectively. By Theorem 2, the function

$$\mathsf{P}(\gamma_1\ge 0,\dots,\gamma_m\ge 0)$$

is a logarithmic concave function of the variables $x_1,\dots,x_n$. On the other hand we have obviously

$$\mathsf{P}(\gamma_1\ge 0,\dots,\gamma_m\ge 0) = G(e^{x_1},\dots,e^{x_r},x_{r+1},\dots,x_n),$$

thus the proof of the theorem is complete. □

3 Algorithmic solution of Problem 1.2

In this section we deal with the algorithmic solution of Problem (1.2) under the conditions of Theorem 3, Theorem 4 and Theorem 5, respectively. Consider the case of Theorem 3. We suppose that C and S are positive definite matrices. The other cases can be treated in a similar way. We suppose further that C is a correlation matrix i.e. its diagonal elements are equal to 1. The random vectors $\boldsymbol{\xi}_1,\dots,\boldsymbol{\xi}_{n+1}$ can always be defined so that the constraint $A\boldsymbol{x}\le\boldsymbol{\beta}$ remains equivalent and this condition holds. Using the notation $\mathbf{L}(\mathbf{z})$ introduced in Section 2, we can write

$$
\begin{aligned}
G(\mathbf{x}) &= \mathsf{P}(A\mathbf{x}\le\boldsymbol{\beta}) = \mathsf{P}\left(\sum_{i=1}^{n+1}(\boldsymbol{\xi}_i-\boldsymbol{\mu}_i)x_i \le \mathbf{L}(\boldsymbol{z})\right)\\
&= \mathsf{P}\left(\frac{1}{(\mathbf{z}'S\mathbf{z})^{\frac{1}{2}}}\sum_{i=1}^{n+1}(\boldsymbol{\xi}_i-\boldsymbol{\mu}_i)x_i \le \frac{1}{(\mathbf{z}'S\mathbf{z})^{\frac{1}{2}}}\mathbf{L}(\mathbf{z})\right) = \Phi\left(\frac{1}{(\mathbf{z}'S\mathbf{z})^{\frac{1}{2}}}\mathbf{L}(\mathbf{z});C\right),
\end{aligned}
\tag{3.1}
$$

where $\Phi(\mathbf{y};\mathbf{C})$ denotes the probability distribution of the m-dimensional normal distribution with $\mathrm{N}(0,1)$ marginal distributions and correlation matrix C. It is well-known that

$$\frac{\partial\Phi(\mathbf{y},\mathbf{C})}{\partial y_i} = \Phi\left(\frac{y_i-c_{ji}y_i}{1-c_{ji}^2},\quad j=1,\dots,i-1,i+1,\dots,m;\ C_i\right)\varphi(y_i), \tag{3.2}$$

where

$$\varphi(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}, \quad -\infty < y < \infty \tag{3.3}$$

and C_i is an $(m-1) \times (m-1)$ correlation matrix with the entries

$$\frac{c_{jk} - c_{ji}c_{ki}}{\sqrt{1-c_{ji}^2}\sqrt{1-c_{ki}^2}}, \quad j,k = 1,\ldots,i-1,i+1,\ldots,m. \tag{3.4}$$

With the aid of (3.1)–(3.3) the gradient of the function $g(\mathbf{x})$ can be obtained by an elementary calculation.

If a nonlinear programming procedure solves the problem with quasi-concave constraints and convex objective function is to be minimized, it can be applied to solve our problem. Such a method is the method of feasible directions (see [6]) the convergence proof of which for the case mentioned above is given in [4] and in a more detailed form in [5].

The case of Theorem 4 does not present new difficulty. In fact we see in the proof of Theorem 4 that the function $\mathsf{P}(A\mathbf{x} \leq \boldsymbol{\beta})$ has the same form as the function standing on the right hand side in (3.1).

In the case of Theorem 5 we substitute x_i by e^{x_i} everywhere in Problem (1.1) for $i = 1,\ldots,r$. Then the first constraining function becomes logarithmic concave in all variables. It may happen that the other constraining functions are concave while the objective function is convex after this substitution. If this is the case and the original problem contained constraints of the form

$$x_i > D_i, \qquad i = 1,\ldots,r, \tag{3.5}$$

where D_i, $i = 1,\ldots,r$ are positive constants, then we can solve the problem by convex programming procedures e.g. by the SUMT method (see [1]) which seems to be very suitable due to the logarithmic concavity of the first constraining function. In fact using the logarithmic penalty function, the unconstrained problems are convex problems. The values of the probabilities can be obtained by simulation. If instead of (3.5) we had nonnegativity constraints for $x_1,\ldots,x_r$ in the original problem then using the constraints (3.5) instead of the constraints $x_1 \geq 0,\ldots,x_r \geq 0$ in the original problem, by a suitable selection for the constants D_i, $i = 1,\ldots,r$ we can come arbitrarily near the original optimum value.

References

[1] Fiacco, A. V. and G. P. McCormick (1968). *Nonlinear Programming: Sequential Unconstrained Minimization Techniques.* Wiley, New York.

[2] Prékopa, A. (1971). Logarithmic concave measures with application to stochastic programming. *Acta Sci. Math.*, Szeged, **32**, 301–316.

[3] PRÉKOPA, A. (1972). A class of stochastic programming decision problems. *Math. Operationsforsch. Statist.*, **3**, 349–354.

[4] PRÉKOPA, A. (1970). On probabilistic constrained programming. *Proc. Princeton Sympos. Math. Programming*, Princeton Univ. Press, Princeton, New Jersey, 113–138.

[5] PRÉKOPA, A. (1974). Eine Erweiterung der sogenannten Methode der "zulässigen Richtungen" der nichtlinearen Optimierung auf den Fall quasikonkaver Restriktionsfunktionen. *Math. Operationsforsch. und Statistik*, **5**, 281–293.

[6] ZOUTENDIJK, G. (1960). *Methods of Feasible Directions*. Elsevier, Amsterdam.