Mathematical Programming 4 (1973) pp. 202–221

CONTRIBUTIONS TO THE THEORY OF STOCHASTIC PROGRAMMING* **

András PRÉKOPA
Technological University and Hungarian Academy of Sciences,
Budapest, Hungary

Received: 13 November 1970
Revised manuscript received 4 January 1973

Abstract

Two stochastic programming decision models are presented. In the first one, we use probabilistic constraints, and constraints involving conditional expectations further incorporate penalties into the objective. The probabilistic constraint prescribes a lower bound for the probability of simultaneous occurrence of events, the number of which can be infinite in which case stochastic processes are involved. The second one is a variant of the model: two-stage programming under uncertainty, where we require the solvability of the second stage problem only with a prescribed (high) probability. The theory presented in this paper is based to a large extent on recent results of the author concerning logarithmic concave measures.

1 Introduction

The stochastic programming decision problems are formulated so that we start from a mathematical programming problem, then observe the random nature of some parameters, and finally decide on a certain principle according to which we operate our system under these circumstances. That mathematical programming problem which is the starting point of the stochastic programming model construction will

*This work was supported in part by the Institute of Economic Planning, Budapest

**This paper was presented at the 7th Mathematical Programming Symposium 1970, The Hague, The Netherlands, under the title "Programming under probabilistic constraints and programming under constraints involving conditional expectations".

be called the underlying problem. The underlying problem we consider here is of the following type

$$\begin{aligned} &\text{minimize } g_0(\boldsymbol{x}), \\ &\text{subject to } u_t(\boldsymbol{x}) \geq \beta_t, \quad t \in T; \quad g_i(\boldsymbol{x}) \geq 0, \quad i = 1, \ldots, N. \end{aligned} \tag{1.1}$$

The set of subscripts T can be finite or infinite. In the latter case, T will be supposed to be a finite interval.

Suppose now that certain parameters in the constraining functions $u_t(\boldsymbol{x})$ and bounds β_t are random, and we want to formulate decision principle for the stochastic system. The existing models can be considered static or dynamic, resp. reliability type or penalty type from another point of view.

The first penalty type model was introduced by Dantzig [4] and Beale [1]. The generalization of this given by Dantzig and Madansky [5] is at the same time the first dynamic type stochastic programming model. Programming under probabilistic constraints, i.e. the use or reliability type models, was introduced by Charnes, Cooper and Symonds [3].

At first sight it might seem that the use of reliability and penalty type models are alternative possibilities and exclude each other. Contrary to this opinion, we are convinced that the best way of operating a stochastic system is to operate it with a prescribed (high) reliability and at the same time use penalties to punish discrepancies. This is the general framework of our model constructions.

The theory presented in this paper is based on recent results of the author concerning "logarithmic concave measures". A probability measure $\mathbf{P}$ defined on a σ-field of subsets of $\mathbb{R}^n$, containing all n-dimensional intervals, is said to be *logarithmic concave* if for any pair $A, B \subset \mathbb{R}^n$ of convex sets and for any $0 < \lambda < 1$, we have

$$\mathbf{P}\{\lambda A + (1-\lambda)B\} \geq (\mathbf{P}\{A\})^{\lambda}(\mathbf{P}\{B\})^{1-\lambda}, \tag{1.2}$$

where $+$ denotes Minkowski addition. In [10], we proved the following:

THEOREM 1.1. *Let f be a probability density function in $\mathbb{R}^n$ having the following form*

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

where Q is a convex function in the entire space. Then the probability measure P, defined by

$$\mathbf{P}\{G\} = \int_G f(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}, \tag{1.4}$$

for all Lebesgue measurable subsets G of $\mathbb{R}^n$, is logarithmic concave.

It can be shown that if P is a logarithmic concave measure and A is a convex set, then the function $\mathbf{P}\{A + \boldsymbol{x}\}$ of the variable $\boldsymbol{x}$ is logarithmic concave (see [10, Theorem 3]). In particular, if a probability measure is logarithmic concave, then the distribution function $F(\boldsymbol{x})$ is a logarithmic concave point function since

$$F(\boldsymbol{x}) = \mathbf{P}\{A + \boldsymbol{x}\} \quad \text{for} \quad A = \{t \,:\, t \le 0\}. \tag{1.5}$$

In this paper, we present two stochastic programming models. The first one, formulated in Section 2, contains probabilistic constraints, and constraints involving conditional expectations; further a penalty is incorporated into the objective function. In Sections 3–5, we show that we are lead to a convex programming problem under some conditions. The second one is a variant of the model: two-stage programming under uncertainty. This is formulated in Section 6, while in Section 7 we are dealing with the properties of this model. In Section 8, an algorithmic solution of the models is given, and in Section 9 we give a numerical example for the first model.

2 Programming under probabilistic constraint and constraints involving conditional expectations with penalty incorporated into the objective

We consider the underlying problem (1.1) and formulate the stochastic programming decision model in the following way:

$$\begin{aligned}
&\text{minimize} && \left\{ g_0(\boldsymbol{x}) + \sum_{t=1}^{m} q_t \int_{u_t(\boldsymbol{x})}^{\infty} [z - u_t(\boldsymbol{x})]\,\mathrm{d}F_t(z) \right\}, \\
&\text{subject to} && \mathbf{P}\{u_t(\boldsymbol{x}) \ge \beta_t,\ t = 1, \ldots, m\} \ge p, \\
& && \mathbf{E}\{\beta_t - u_t(\boldsymbol{x}) \mid \beta_t - u_t(\boldsymbol{x}) > 0\} \le l_t,\ t = 1, \ldots, m, \\
& && g_i(\boldsymbol{x}) \ge 0,\ i = 1, \ldots, N,
\end{aligned} \tag{2.1}$$

if T is a finite set, $T = \{1, \ldots, m\}$. If T is an interval, the problem is formulated as

$$\begin{aligned}
&\text{minimize} && \left\{ g_0(\boldsymbol{x}) + \int_T \int_{u_t(\boldsymbol{x})}^{\infty} [z - u_t(\boldsymbol{x})]\,\mathrm{d}F_t(z)\,\mathrm{d}G(t) \right\}, \\
&\text{subject to} && \mathbf{P}\{u_t(\boldsymbol{x}) \ge \beta_t,\ t \in T\} \ge p, \\
& && \mathbf{E}\{\beta_t - u_t(\boldsymbol{x}) \mid \beta_t - u_t(\boldsymbol{x}) > 0\} \le I_t,\ t \in T. \\
& && g_i(\boldsymbol{x}) \ge 0,\ i = 1, \ldots, N.
\end{aligned} \tag{2.2}$$

Problem (2.1) is obviously a special case of Problem (2.2). In (2.1), $q_1, \dots, q_m$ are nonnegative numbers, while in (2.2), $G(t)$ is a nonnegative, nondecreasing and bounded function defined on T. $F_t(z)$ is the probability distribution function of the random variable β_t, $F_t(z) = \mathbf{P}\{\beta_t \le z\}$, $-\infty < z < \infty$.

$\mathbf{E}$ is the symbol of expectation. Whenever penalty is used to punish deviation of the form $\beta_t - u_t(\boldsymbol{x}) > 0$, or we work with the conditional expectation $\mathbf{E}\{\beta_t - u_t(\boldsymbol{x}) \mid \beta_t - u_t(\boldsymbol{x}) > 0\}$, we always suppose in the whole paper that the expectation of β_t exists.

The functions $u_t(\boldsymbol{x})$, $t \in T$, $-g_0(\boldsymbol{x})$ are supposed to be concave with continuous gradient in an open convex set H. Consider the case where T is an interval. We suppose that $u_t(\boldsymbol{x})$ and $\boldsymbol{\nabla} u_t(\boldsymbol{x})$ are piecewise continuous in $t \in T$ for any $\boldsymbol{x} \in H$. We suppose furthermore that the stochastic process β_t is separable and is piecewise weakly continuous in T. Weak continuity of β_t at the point τ means that β_τ is the limit in probability of β_t whenever $t \to \tau$. It follows that (see [6, Theorem 2.2, p. 54]) the joint occurrence of the events $u_t(\boldsymbol{x}) \ge \beta_t$, $t \in T$, is again an event.

We suppose that all random variables β_t, $t \in T$, are continuously distributed, and denote by $f_t(z)$ the probability density belonging to $F_t(z)$. As $F_t(z)$ is continuous in z for every t, the piecewise weak continuity of the stochastic process β_t implies the piecewise continuity of the function $F_t(z)$ with respect to t for any real z. We suppose further that the function

$$\int_{u_t(\boldsymbol{x})}^{\infty} [z - u_t(\boldsymbol{x})] \, \mathrm{d}F_t(z) \tag{2.3}$$

is piecewise continuous in $t \in T$ for any fixed $\boldsymbol{x} \in H$. This implies the existence of the integral (also in the Riemann–Stieltjes sense) in the objective of Problem (2.2). It is easy to see that the additional term in the objective of Problem (2.1) is convex and has continuous gradient in H.

3 The gradient of the objective of Problem (2.2)

In this section, we prove the following:

Theorem 3.1. *The additional term in the objective of Problem* (2.2) *is a convex function on H. If furthermore there exist functions $w_t^{(1)}(\boldsymbol{x})$, $w_t^{(2)}(\boldsymbol{x})$ of the variables $t \in T$, $\boldsymbol{x} \in H$ such that for every fixed $\boldsymbol{x} \in H$ and for every sufficiently small*

$\Delta x \neq 0$, *we have*

$$\left|\frac{u_t(\boldsymbol{x}+\boldsymbol{\Delta x}_i)-u_t(\boldsymbol{x})}{\Delta x}\right| \leq w_t^{(1)}(\boldsymbol{x}), \qquad t \in T,$$

$$\|\boldsymbol{\nabla} u_t(\boldsymbol{x}+\boldsymbol{\Delta x}_i)-\boldsymbol{\nabla} u_t(\boldsymbol{x})\| \leq w_t^{(2)}(\boldsymbol{x}), \qquad t \in T,$$

where $w_t^{(1)}(\boldsymbol{x})$, $w_t^{(2)}(\boldsymbol{x})$ *have finite integrals over* T *for any* $\boldsymbol{x} \in H$, $\boldsymbol{\Delta x}_i$ *is the vector the* i^{th} *component of which is equal to* Δx *while the others are zero, and the norm is the Euclidean norm, then the additional term has continuous gradient in* H.

Proof. The convexity of the integrand can be proved easily. Integration with respect to t does not disturb this property, hence the additional term is convex.

The difference quotient of the additional term is the following:

$$\begin{aligned}(\Delta x)^{-1} &\int_T \int_{u_t(\boldsymbol{x})}^{u_t(\boldsymbol{x}+\boldsymbol{\Delta x}_i)} [u_t(\boldsymbol{x})-z] f_t(z)\,\mathrm{d}z\,\mathrm{d}G(t) \\ &+ \int_T \frac{u_t(\boldsymbol{x}+\boldsymbol{\Delta x}_i)-u_t(\boldsymbol{x})}{\Delta x}[F_t(u_t(\boldsymbol{x}+\boldsymbol{\Delta x}_i))-1]\,\mathrm{d}G(t).\end{aligned} \tag{3.1}$$

The first term can be majorated as

$$\begin{aligned}&\left|(\Delta x)^{-1} \int_T \int_{u_t(\boldsymbol{x})}^{u_t(\boldsymbol{x}+\boldsymbol{\nabla x}_i)} [u_t(\boldsymbol{x})-z] f_t(z)\,\mathrm{d}z\,\mathrm{d}G(t)\right| \\ &\quad \leq \int_T \left|\frac{u_t(\boldsymbol{x}+\boldsymbol{\Delta x}_i)-u_t(\boldsymbol{x})}{\Delta x}[F_t u_t(\boldsymbol{x}+\boldsymbol{\Delta x}_i))-F_t(u_t(\boldsymbol{x}))]\right|\,\mathrm{d}G(t).\end{aligned}$$

The integrand on the right-hand side tends to 0 as $\Delta x \to 0$, and for small Δx values it is smaller than or equal to $w_t^{(1)}(\boldsymbol{x})$ uniformly in t. Hence by the Dominated Convergence Theorem, the integral on the right-hand side tends to zero as $\Delta x \to 0$. The integrand in the second term in (3.1) is also dominated by $w_t^{(1)}(\boldsymbol{x})$, hence by the same theorem the limit exists whenever $\Delta x \to 0$.

The gradient of the objective function of Problem (2.2) is the following:

$$\boldsymbol{\nabla} g_0(\boldsymbol{x}) + \int_T \boldsymbol{\nabla} u_t(\boldsymbol{x})[F_t(u_t(\boldsymbol{x}))-1]\,\mathrm{d}G(t). \tag{3.2}$$

The continuity in $\boldsymbol{x}$ of the second term follows easily from the suppositions. Thus Theorem 3.1 is proved. □

4 Properties of the probabilistic constraints of Problems (2.1) and (2.2)

In this section, we show that the functions standing on the left-hand sides of the probabilistic constraints in (2.1) and (2.2) are logarithmic concave under some conditions.

THEOREM 4.1. *Let T be a finite set or a finite interval. Suppose that for $u_t(\boldsymbol{x})$, β_t, $t \in T$, the conditions mentioned in Section 2 are satisfied. Suppose furthermore that the finite-dimensional distributions of the stochastic process β_t are logarithmic concave probability measures. Then the function $\mathbf{P}\{u_t(\boldsymbol{x}) \geq \beta_t,\ t \in T\}$, is logarithmic concave in $\boldsymbol{x} \in H$.*

Proof. Let $t_1, \dots, t_N$ be a finite parameter subset of T. Let further $\boldsymbol{x}, \boldsymbol{y} \in H$ and $0 < \lambda < 1$. Then the concavity of the functions $u_{t_i}(\boldsymbol{x})$, $i = 1, \dots, N$, and the logarithmic concavity of the joint distribution function of β_{t_i}, $i = 1, \dots, N$, implies

$$\begin{aligned} &\mathbf{P}\{u_{t_i}(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \geq \beta_{t_i},\ i = 1, \dots, N\} \\ &\quad \geq \mathbf{P}\{\lambda u_{t_i}(\boldsymbol{x}) + (1-\lambda) u_{t_i}(\boldsymbol{y}) \geq \beta_{t_i},\ i = 1, \dots, N\} \\ &\quad \geq [\mathbf{P}\{u_{t_i}(\boldsymbol{x}) \geq \beta_{t_i},\ i = 1, \dots, N\}]^{\lambda} [\mathbf{P}\{u_{t_i}(\boldsymbol{y}) \geq \beta_{t_i},\ i = 1, \dots, N\}]^{1-\lambda}. \end{aligned} \tag{4.1}$$

If T is a finite set, then the theorem is proved. Let T be a finite interval. Since logarithmic concavity is preserved when taking the limit and

$$\mathbf{P}\{u_t(\boldsymbol{x}) \geq \beta_t,\ t \in T\} = \lim_{N \to \infty} \{u_{t_i}(\boldsymbol{x}) \geq \beta_{t_i},\ i = 1, \dots, N\} \tag{4.2}$$

provided $t_1, t_2, \dots$ is a dense sequence in T (see [6, Theorem 2.2, p. 54]) our theorem is proved. □

One important example for the stochastic process β_t, where the finite dimensional distributions are logarithmic concave measures, is the Gaussian process.

5 The constraints involving conditional expectations

The constraints

$$\mathbf{E}\{\beta_t - u_t(\boldsymbol{x}) \mid \beta_t - u_t(\boldsymbol{x}) > 0\} \leq l_t, \qquad t \in T, \tag{5.1}$$

can be written in the equivalent form

$$h_t(u_t(\boldsymbol{x})) \leq l_t, \qquad t \in T, \tag{5.2}$$

where

$$h_t(z) = \frac{\int_z^\infty (v-z)\,\mathrm{d}F_t(v)}{1-F_t(z)} = \frac{\int_z^\infty v\,\mathrm{d}F_t(v)}{1-F_t(z)} - z. \tag{5.3}$$

The use of a constraint (3.2) is advised only under the following two conditions:

(a) $h_t(z)$ is a nonincreasing function of z,

(b) $\lim\{h_t(z) \,:\, z \to \infty\} = 0$.

The reason of (a) is that $h_t(u_t(\boldsymbol{x}))$ measures the deviation of the type $\beta_t - u_t(\boldsymbol{x}) > 0$, and if for two programs $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ we have $u_t(\boldsymbol{x}_1) < u_t(\boldsymbol{x}_2)$, i.e.

$$\beta_t - u_t(\boldsymbol{x}_1) > \beta_t - u_t(\boldsymbol{x}_2),$$

then it is natural to require $h_t(u_t(\boldsymbol{x}_1)) \geq h_t(u_t(\boldsymbol{x}_2))$. The reason of (b) is that we shall be able to use arbitrary small l_t in (5.1). In view of (a), the constraint (5.2) can be written as

$$u_t(\boldsymbol{x}) \geq h_t^{-1}(l_t), \qquad t \in T, \tag{5.4}$$

provided $h_t^{-1}(z)$ is conveniently defined (if c is the smallest number z for which $h_t(z) = d$, then $h_t^{-1}(d) = c$).

We remark that if β_t has an expectation, then the function (5.3) is continuous for every z for which $F_t(z) < 1$. If $F_t(z)$ reaches the value 1 for finite z and z_0 is the smallest z for which $F_t(z) = 1$, then we define $F_t(z) = 0$ for $z \geq z_0$, and it is easy to see that the function (5.3) is continuous also at the point $z = z_0$. In all cases we have $\lim\{h_t(z) \,:\, z \to -\infty\} = 0$.

If a constraint of the type (5.1) is used and conditions (a), (b) are satisfied, then we convert it into the equivalent form on the basis of tabulated values of the functions (5.3).

There remains to consider the problem for which probability distributions conditions (a) and (b) are satisfied. In connection with this we prove the following:

Theorem 5.1. *Let β be a continuously distributed random variable with logarithmic concave probability density and suppose that $\mathbf{E}\{\beta\}$ exists. Then the function of the variable z*

$$h(z) = \mathbf{E}\{\beta - z \mid \beta - z > 0\}$$

is a nonincreasing function on the entire real line.[1]

[1] V. Zolotariev (Steklov Institute, Moscow) participated in the development of this theorem. The present proof is due to author of this paper.

Proof. Let f, F be the density and the distribution functions, respectively, of the random variable β. As $h(z) = 0$ if $F(z) = 1$, it is enough to consider such z values for which $F(z) < 1$. Since $\mathbf{E}\{\beta\}$ exists, it follows that the integrals below exist and

$$\int_z^\infty x f(x)\,\mathrm{d}x = \int_z^\infty [1 - F(x)]\,\mathrm{d}x + z[1 - F(z)].$$

Using this, we derive the equality

$$h(z) = \frac{\displaystyle\int_z^\infty (x - z) f(x)\,\mathrm{d}x}{1 - F(z)} = \frac{\displaystyle\int_z^\infty [1 - F(x)]\,\mathrm{d}x}{1 - F(z)}. \tag{5.5}$$

Let $G(z)$ and $-g(z)$ denote the numerator resp. the denominator on the right-hand side in (5.5). Then $h(z) - G(z)/(-g(z))$ and $G'(z) = g(z)$. By Theorem 1.1, $1 - F(x)$ is logarithmic concave on the entire real line. In fact, if $A = (0, \infty)$, then $1 - F(x) = \mathbf{P}\{A + x\}$, and the function on the right-hand side has the mentioned property. Again by Theorem 1.1, $G(z)$ is also logarithmic concave in $\mathbb{R}^1$. Thus $\mathrm{d}\,\log G(z)/\mathrm{d}z = g(z)/G(z)$ must be nonincreasing, which proves that $h(z)$ is also nonincreasing.

We required also that $\lim\{h(z) : z \to \infty\} = 0$. We prove now that if β has a normal distribution, then this condition is also satisfied. We may suppose that $\mathbf{E}\{\beta\} = 0$ and $\mathbf{E}\{\beta^2\} = 1$. The function $h(z)$ has the form

$$h(z) = \frac{\varphi(z)}{1 - \Phi(z)} - z,$$

where

$$\varphi(z) = (2\pi)^{-1} \exp\left[-\frac{1}{2}z^2\right], \qquad \Phi(z) = \int_{-\infty}^z \varphi(x)\,\mathrm{d}x,$$
$$-\infty < z < \infty.$$

Since $h(z) > 0$ for every z, it follows that

$$\frac{\varphi(z)}{1 - \Phi(z)} > z, \qquad -\infty < z < \infty,$$

hence applying the l'Hospital rule for the right-hand side of (5.5), we obtain

$$\lim_{z\to\infty} h(z) = \lim_{z\to\infty} (1 - \Phi(z))/\varphi(z) = 0.$$

We may introduce a constraint involving conditional expectation also in the case where the technological coefficients are also random. If the good situation is the fulfillment of the inequality $\boldsymbol{\alpha}'\boldsymbol{x} = \alpha_1 x_1 + \ldots + \alpha_n x_n \geq \beta$, where $\alpha_1, \ldots, \alpha_n$ are random variables, then we can measure deviations into the wrong direction by the quantity $\mathbf{E}\{\beta - \boldsymbol{\alpha}'\boldsymbol{x} : \beta - \boldsymbol{\alpha}'\boldsymbol{x} > 0\}$ but in some other ways too. An example is the following: First we define

$$\delta = \frac{\beta - \boldsymbol{\alpha}'\boldsymbol{x}}{\sqrt{(\boldsymbol{z}'V\boldsymbol{z})}},$$

where $\boldsymbol{z}' = (1, x_1, \ldots, x_n)$, V is the covariance matrix of the random variables β, $-\alpha_1, \ldots, -\alpha_n$, and then introduce the measure of deviation together with its upper bound as follows:

$$\mathbf{E}\{\delta \mid \delta > 0\} \leq l, \qquad l > 0. \tag{5.6}$$

If the random variables have a joint normal distribution with expectations $\mathbf{E}\{\beta\} = d$, $\mathbf{E}\{\alpha_i\} = a_i$, $i = 1, \ldots, n$, and $\boldsymbol{a}$ is the vector with components a_i, then (5.6) has the following form:

$$\begin{aligned} h(t) &= \mathbf{E}\{\delta \mid \delta > 0\} \\ &= \mathbf{E}\left\{ \frac{\beta - d - (\boldsymbol{\alpha} - \boldsymbol{a})'\boldsymbol{x}}{\sqrt{(\boldsymbol{z}V\boldsymbol{z})}} + \frac{d - \boldsymbol{a}'\boldsymbol{x}}{\sqrt{(\boldsymbol{z}'V\boldsymbol{z})}} \,\middle|\, \frac{\beta - d - (\boldsymbol{\alpha} - \boldsymbol{a})'\boldsymbol{x}}{\sqrt{(\boldsymbol{z}'V\boldsymbol{z})}} + \frac{d - \boldsymbol{a}'\boldsymbol{x}}{\sqrt{(\boldsymbol{z}'V\boldsymbol{z})}} > 0 \right\} \\ &= \mathbf{E}\{\gamma - t \mid \gamma - t > 0\} \leq l, \qquad t = \frac{\boldsymbol{a}'\boldsymbol{x} - d}{\sqrt{(\boldsymbol{z}'V\boldsymbol{z})}}. \end{aligned} \tag{5.7}$$

We shall show that for normally distributed random variables conditions (a), (b) are satisfied. Thus if $\beta, -\alpha_1, \ldots, -\alpha_n$ have a joint normal distribution, then γ has an $N(0,1)$ distribution for any $\boldsymbol{x}$ vector, hence (5.7) can be converted into the equivalent form

$$h^{-1}(l)\sqrt{(\boldsymbol{z}'V\boldsymbol{z})} + d - \boldsymbol{a}'\boldsymbol{x} \leq 0, \tag{5.8}$$

and on the left-hand side there stands a convex function of the variable $\boldsymbol{x} \in \mathbb{R}^n$.

6 A variant of the model: two stage programming under uncertainty

The original two stage programming under uncertainty model [5] is formulated concerning the following underlying problem: minimize $c'\boldsymbol{x}$ subject to the constraints $A\boldsymbol{x} = \boldsymbol{b}$, $T\boldsymbol{x} = \boldsymbol{\beta}$, $\boldsymbol{x} \geq 0$, which is a special case of (1.1). We suppose that $\boldsymbol{\beta}$ is a random vector and consider the so called second-stage problem

$$M\boldsymbol{y} = \boldsymbol{\beta} - T\boldsymbol{x}, \quad \boldsymbol{y} \geq \boldsymbol{0}, \quad \min \boldsymbol{q}'\boldsymbol{y}, \tag{6.1}$$

where $\boldsymbol{x}$ is fixed, $\boldsymbol{\beta}$ is random but it is also fixed at some realization. The optimum value depends on $\boldsymbol{\beta}$, hence it is also a random variable. The first-stage problem specifying $\boldsymbol{x}$ is the following:

$$\begin{gathered} A\boldsymbol{x} = \boldsymbol{b}, \qquad \boldsymbol{x} \geq \boldsymbol{0}, \\ \min\{c'x + \mathbf{E}\{\min \boldsymbol{q}'\boldsymbol{y} \,:\, M\boldsymbol{y} = \boldsymbol{\beta} - T\boldsymbol{x},\ \boldsymbol{y} \geq \boldsymbol{0}\}\}. \end{gathered} \tag{6.2}$$

The second-stage problem must be solvable for any realization of the random vector $\boldsymbol{\beta}$. This shows that in some cases this model imposes a very strong condition concerning the structure of the second-stage problem. If e.g. $\boldsymbol{\beta}$ has a nondegenerated normal distribution, then the set of realizations of $\boldsymbol{\beta}$ is the entire space and the right-hand side of the first constraint in (6.1) varies also in the entire space. In our modification of this very important model formulation it will not be required that the second-stage problem be solvable for all realizations of $\boldsymbol{\beta}$. Before presenting the modification we have to make some remarks.

Wets proved in [12] that the set of feasible $\boldsymbol{x}$ vectors is a convex polyhedron. ($\boldsymbol{x}$ is *feasible* if $A\boldsymbol{x} = \boldsymbol{b}$, $\boldsymbol{x} \geq \boldsymbol{0}$; further for every realization of $\boldsymbol{\beta}$ there exists a $\boldsymbol{y} \geq \boldsymbol{0}$ satisfying $T\boldsymbol{x} + M\boldsymbol{y} = \boldsymbol{\beta}$.) We shall use his proof for the development of our theory. Consider the convex polyhedral cone

$$C = \{\boldsymbol{z} \,:\, \boldsymbol{z} = M\boldsymbol{y},\ \boldsymbol{y} \geq \boldsymbol{0}\}. \tag{6.3}$$

By the theorem of Weyl there exist vectors $\boldsymbol{d}_1, \ldots, \boldsymbol{d}_v$ such that

$$C = \{\boldsymbol{z} \,:\, \boldsymbol{d}_i'\boldsymbol{z} \leq 0,\ i = 1, \ldots, v\}. \tag{6.4}$$

Let $\boldsymbol{x}$ be a vector satisfying $A\boldsymbol{x} = \boldsymbol{b}$, $\boldsymbol{x} \geq \boldsymbol{0}$. Then problem (6.1) is solvable if and only if $\boldsymbol{\beta} - T\boldsymbol{x} \in C$. Thus the condition of the solvability of the second-stage problem is

$$\boldsymbol{d}_i'\boldsymbol{\beta} \leq \boldsymbol{d}_i'T\boldsymbol{x}, \qquad i = 1, \ldots, v. \tag{6.5}$$

If this holds for every realization of $\boldsymbol{\beta}$, then $\boldsymbol{x}$ is feasible. The probability of the fulfillment of the events in (6.5) gives the probability of solvability of the second-stage problem in case of the given $\boldsymbol{x} \in \mathbb{R}^n$.

Now we define the new second-stage problem as follows

$$\begin{aligned} & M\boldsymbol{y} + \boldsymbol{y}^+ - \boldsymbol{y}^- = \boldsymbol{\beta} - T\boldsymbol{x}, \\ & \boldsymbol{y} \geq 0, \quad \boldsymbol{y}^+ \geq \boldsymbol{0}, \quad \boldsymbol{y}^- \geq \boldsymbol{0}, \\ & \min\{\boldsymbol{q}'\boldsymbol{y} + \boldsymbol{q}^{+\prime}\boldsymbol{y}^+ + \boldsymbol{q}^{-\prime}\boldsymbol{y}^-\}, \end{aligned} \tag{6.6}$$

where it is reasonable to suppose that $\boldsymbol{q}^+ \geq \mathbf{0}$, $\boldsymbol{q}^- \geq \mathbf{0}$. It is also reasonable to choose $\boldsymbol{q}^+$ and $\boldsymbol{q}^-$ in such a way that we obtain automatically $\boldsymbol{y}^+ = \mathbf{0}$, $\boldsymbol{y}^- = \mathbf{0}$ if the originally formulated second-stage problem is solvable. If the original second-stage problem is not solvable, then the optimum value of (6.6) expresses a certain distance between the given $\boldsymbol{x}$ and the set defining the solvability,

$$\{\boldsymbol{x} \,:\, \boldsymbol{d}_i'\boldsymbol{\beta} \leq \boldsymbol{d}_i'T\boldsymbol{x}, \qquad i = 1, \ldots, v\}, \tag{6.7}$$

where now $\boldsymbol{\beta}$ is fixed at some realization. Our model is formulated in the following way:

$$\begin{aligned} A\boldsymbol{x} = \boldsymbol{b}, \qquad \boldsymbol{x} \geq \mathbf{0}, \\ \mathbf{P}\{\boldsymbol{d}_i'\boldsymbol{\beta} \leq \boldsymbol{d}_i'T\boldsymbol{x}, \quad i = 1, \ldots, v\} \geq p, \\ \mathbf{E}\{\boldsymbol{d}_i' - \boldsymbol{d}_i'T\boldsymbol{x} \mid \boldsymbol{d}_i' - \boldsymbol{d}_i'T\boldsymbol{x} > 0\} \leq l_i, \qquad i = 1, \ldots, v, \\ \min\{c'\boldsymbol{x} + \mathbf{E}\{\mu(\boldsymbol{x})\}\}, \end{aligned} \tag{6.8}$$

where $\mu(\boldsymbol{x})$ is the random optimum value of the modified second-stage problem, p is a given probability, $0 < p < 1$, and $l_1, \ldots, l_v$ are given constants.

7 Properties of the model introduced in Section 6

In connection with problem (6.8), we have to analyze the functions in the probabilistic constraint, in the constraints containing conditional expectations, and in the objective function. Our aim is to show that problem (6.8) is a convex programming problem (or equivalent to a convex programming problem) under some conditions. First we prove:

THEOREM 7.1. *Suppose that the random vector $\boldsymbol{\beta}$, standing in the probabilistic constraint of problem* (6.8)*, has continuous distribution and its probability density is logarithmic concave in the entire space $\mathbb{R}^m$. Then as a function of $\boldsymbol{x}$, $\mathbf{P}\{\boldsymbol{d}_i'\boldsymbol{\beta} \leq \boldsymbol{d}_i'T\boldsymbol{x},\ i = 1, \ldots, v\}$ is logarithmic concave on $\mathbb{R}^n$.*

Proof. If we define

$$\begin{aligned} A &= \{\boldsymbol{z} \,:\, \boldsymbol{d}_i'\boldsymbol{z} \leq \boldsymbol{d}_iT\boldsymbol{x},\ i = 1, \ldots, v\}, \\ B &= \{\boldsymbol{z} \,:\, \boldsymbol{d}_i'\boldsymbol{z} \leq \boldsymbol{d}_iT\boldsymbol{y},\ i = 1, \ldots, v\}, \\ C &= \{\boldsymbol{z} \,:\, \boldsymbol{d}_i'\boldsymbol{z} \leq \boldsymbol{d}_iT(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}),\ i = 1, \ldots, v\}, \end{aligned}$$

where $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $0 < \lambda < 1$, we have that $C \supset \lambda A + (1-\lambda)B$. Hence applying Theorem 3.1 completes the proof. □

Concerning the random variables $\boldsymbol{d}_i'\boldsymbol{\beta}$, $i = 1, \ldots, v$, we prove the following lemma.

LEMMA 7.2. *If $\boldsymbol{\beta}$ is a continuously distributed random vector with logarithmic concave density $f(x)$ and $\mathbf{d} \neq \mathbf{0}$ is a constant vector having the same number of components as $\boldsymbol{\beta}$, then the random variable $\mathbf{d}'\boldsymbol{\beta}$ has a logarithmic concave density.*

Proof. We may suppose that $d_1 \neq 0$. The distribution function of $\boldsymbol{d}'\boldsymbol{\beta}$ is given by

$$G(z) = \int_{\boldsymbol{d}'\boldsymbol{x} \leq z} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = |d_1|^{-1} \int_{y_1 \leq z} f(D^{-1}\boldsymbol{y})\,\mathrm{d}\mathbf{y},$$

where we applied the transformation

$$\begin{aligned} d_1 x_1 + d_2 x_2 + \ldots + d_m x_m &= y_1, \\ x_2 \qquad\qquad &= y_2, \\ &\vdots \\ x_m &= y_m, \end{aligned}$$

and D is the matrix of the coefficients standing on the left-hand side. It follows that the probability density g of $\boldsymbol{d}'\boldsymbol{\beta}$ equals

$$g(y_1) = |d_1|^{-1} \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} f(D^{-1}\boldsymbol{y})\,\mathrm{d}y_2 \ldots \mathrm{d}y_m,$$
$$-\infty < y_1 < \infty.$$

Since all marginal densities belonging to a logarithmic concave density are logarithmic concave (see [11, Theorem 8]), it follows that g is a logarithmic concave function and the lemma is proved. □

Taking into account Theorem 5.1, the constraints involving conditional expectations in problem (6.8) can be converted into equivalent linear constraints provided $\boldsymbol{\beta}$ has a logarithmic concave density.

The objective function of problem (6.8) does not require new investigation because problem (6.6), where the random optimum value comes from, is a special second-stage problem of the type (6.1), hence the results concerning two-stage programming under uncertainty are applicable here.

8 Solutions of the problems formulated in Sections 2 and 6

First we make some remarks concerning the gradients of the functions standing in the probabilistic constraints of problems (2.1), (2.2) and (6.8).

In general, the gradient of the probability distribution function $F(\boldsymbol{x})$ of a continuous distribution can be obtained as follows:

$$\frac{\partial F(\boldsymbol{x})}{\partial x_i} = F(x_j,\ j = 1, \dots, i-1,\ i+1, \dots, n \mid x_i) f_i(x_i), \qquad i = 1, \dots, n, \quad (8.1)$$

where $F(\cdot \mid \cdot)$ stands for conditional distribution function and $f_i(x_i)$ is the probability density of the conditioning variable. For many multivariate probability distributions, the function $F(\cdot \mid \cdot)$ is of the same type as $F(\boldsymbol{x})$. This is very important from the point of view of computation because the same subroutine can be used to compute the values of the gradient of $F(\boldsymbol{x})$, which is used to compute the values of $F(\boldsymbol{x})$. Let $\Phi(x; R)$ denote the probability distribution function belonging to the n-dimensional normal distribution, where the random variables have expectations equal to zero, variances equal to 1 and R is their correlation matrix. In this case, we obtain

$$\frac{\partial \Phi(\boldsymbol{x}\,;\,R)}{\partial x_i} = \Phi\left(\frac{x_j - r_{ji}x_i}{\sqrt{(1-r_{ji}^2)}},\ j = 1, \dots, i-1,\ i+1, \dots, n\,;\, R_i\right)\varphi(x_i),$$
$$i = 1, \dots, n, \quad (8.2)$$

where R_i is the following $(n-1) \times (n-1)$ correlation matrix

$$\begin{aligned} R_i &= (s_{jk}), \\ s_{jk} &= \frac{r_{jk} - r_{ji}r_{ki}}{\sqrt{(1-r_{ji}^2)}(\sqrt{(1-r_{ki}^2)}}, \qquad j, k = 1, \dots, i-1, i+1, \dots, n. \end{aligned} \quad (8.3)$$

The gradient of the function standing on the left-hand side in the probabilistic constraint of problem (2.1) follows from this easily. This formula can be applied even in the case where $|R| = 0$, i.e., the distribution is degenerated but all elements of R outside the diagonal have absolute values less than 1. The function standing on the left-hand side in the probabilistic constraint of problem (6.8) will be frequently of this type. In fact, if $\boldsymbol{d}_i \neq \boldsymbol{0}$, and no pair from $\boldsymbol{d}_1, \dots, \boldsymbol{d}_v$ is linearly dependent, further if $\boldsymbol{\beta}$ has a nondegenerated normal distribution, then $\boldsymbol{d}_1'\boldsymbol{\beta}, \dots, \boldsymbol{d}_v'\boldsymbol{\beta}$ have a joint normal distribution which satisfies the above requirements, but this joint distribution is surely degenerated if r is greater than the dimension of $\boldsymbol{\beta}$. This implies the continuity of the gradient of the mentioned function in (6.8) under the above suppositions.

We consider now the function $\mathbf{P}\{u_t(\boldsymbol{x}) \geq \beta_t,\ t \in T\}$, where β_t is a Gaussian process with $\mathbf{E}\{\beta_t\} = 0$, $\mathbf{E}\{\beta_t^2\} = 1$, $t \in T = [a, b]$. Let $t_1, t_2, \dots$ be a sequence of elements of T, supposed to be dense in the same interval. Using (8.2), we have

$$\frac{\partial \mathbf{P}\{u_{t_j}(\boldsymbol{x}) \geq \beta_{t_j},\ j=1,\dots,N\}}{\partial x_k}$$
$$= \sum_{i=1}^{N} \mathbf{P}\left\{ \frac{u_{t_j}(\boldsymbol{x}) - r(t_j,t_i)u_{t_i}(\boldsymbol{x})}{\sqrt{(1-r^2(t_j,t_i))}} \geq \beta_{t_j}^{(i)},\ j=1,\dots,N,\ j\neq i \right\}$$
$$\times \varphi(u_{t_i}(\boldsymbol{x}))\frac{\partial u_{t_i}(\boldsymbol{x})}{\partial x_k};\quad k=1,\dots,n, \tag{8.4}$$

where $\beta_t^{(i)}$, $t\in T$, is again a Gaussian process with identically 0 expectation function and identically 1 variance function. Its correlation function is given by

$$r_i(s,t) = \frac{r(s,t) - r(t_i,s)r(t_i,t)}{\sqrt{(1-r^2(t_i,s))}\sqrt{(1-r^2(t_i,t))}},\qquad i=1,\dots,N, \tag{8.5}$$

where $r(s,t)$ is the correlation function of the Gaussian process β_t. The gradient of the function $\mathbf{P}\{u_t(\boldsymbol{x}) \geq \beta_t,\ t\in T\}$ is the limit of (8.4) when $N\to\infty$, provided the differentiation with respect to the components of $\boldsymbol{x}$ and the limit can be interchanged. If T is a finite set, $T=\{t_1,\dots,t_N\}$, then (8.4) gives the gradient of the considered function.

There are many nonlinear programming methods which can be applied for the solution of our problems. We have computer experience concerning the solution of problem (2.1) by the application of the method of feasible directions and the sequential unconstrained minimization technique procedure. The convergence proof of the former method [14, procedure P2, p. 74] was generalized in [9] for the case of quasi-concave constraints under some regularity assumptions. Thus we can solve the following problem:

$$\begin{aligned}
&\text{minimize } g_0(\boldsymbol{x}) \\
&\text{subject to } G_i(\boldsymbol{x}) \geq p_i, && i\in I_C=\{1,\dots,m_1\},\\
&\qquad\qquad \boldsymbol{a}_i'\boldsymbol{x} \geq b_i, && i\in I_L=\{1,\dots,m_2\},
\end{aligned}$$

where $g_0(\boldsymbol{x})$ is convex and $G_1(\boldsymbol{x}),\dots,G_{m_1}(\boldsymbol{x})$ are quasi-concave. The procedure starts with an arbitrary feasible $\boldsymbol{x}_1$. If $\boldsymbol{x}_1,\dots,\boldsymbol{x}_k$ are already determined, then we solve the linear programming problem

$$\begin{aligned}
G_i(\boldsymbol{x}_k) + \nabla G_i(\boldsymbol{x}_k)(\boldsymbol{x}-\boldsymbol{x}_k) + \theta_i y &\geq p_i, && i\in I_C,\\
\boldsymbol{a}_i'\boldsymbol{x} &\geq b_i, && i\in I_L,\\
\nabla g_0(\boldsymbol{x}_k)(\boldsymbol{x}-\boldsymbol{x}_k) &\geq y,\\
\min y,
\end{aligned} \tag{8.6}$$

denote by $\boldsymbol{x}_k^*$ an optimal solution, minimize $g_0(\boldsymbol{x})$ on the feasible part of the ray $\boldsymbol{x}_k + \lambda(\boldsymbol{x}_k^* - \boldsymbol{x}_k)$, $\lambda \geq 0$, and define $\boldsymbol{x}_{k+1}$ as an arbitrary minimizing point. The gradients in case of problems (1.1) and (1.2) are given in this section and in Section 3.

The application of the SUMT methods offers particularly good posibilities if we use logarithmic penalty function. In fact, our functions standing in the constraints are logarithmic concave for a wide class of the random variables β_t, $t \in T$; thus the penalty function will be convex, which simplifies its minimization.

When solving problem (6.8), the explicit knowledge of the vectors $\boldsymbol{d}_1, \ldots, \boldsymbol{d}_v$ is supposed. We remark that the values of the functions and their gradients standing in the probabilistic constraints of problems (2.1), (2.2) and (6.8) can be obtained by simulation. The same holds for $\mathbf{E}\{\mu\}$ standing in the objective function of problem (6.8).

9 A numerical example and sensitivity analysis of the constraining functions with respect to correlations

We consider the following problem:

$$\begin{aligned} &\mathbf{P}\{x_1 + x_2 - 3 \geq \beta_1,\ 2x_1 + x_2 - 4 \geq \beta_2\} \geq 0.8, \\ &x_1 + 4x_2 \geq 4,\ 5x_1 + x_2 \geq 5,\ x_1 \geq 0,\ x_2 \geq 0, \\ &\min(3x_1 + 2x_2), \end{aligned} \tag{9.1}$$

where β_1, β_2 have joint normal distribution with $\mathbf{E}\{\beta_1\} = \mathbf{E}\{\beta_2\} = 0$, $\mathbf{E}\{\beta_1^2\} = \mathbf{E}\{\beta_2^2\} = 1$, further $\mathbf{E}\{\beta_1\beta_2\} = r = 0.2$. We apply the method of feasible directions. Given a feasible $\boldsymbol{x}_k$, the direction finding problem is the following (where $L_1(\boldsymbol{x}) = x_1 + x_2 - 3$, $L_2(\boldsymbol{x}) = 2x_1 + x_2 - 4$):

$$\begin{gathered} \Phi(L_1(\boldsymbol{x}_k), L_2(\boldsymbol{x}_k); 0, 2) + \boldsymbol{G}_k'(\boldsymbol{x} - \boldsymbol{x}_k) + \theta y \geq p, \\ x_1 + 4x_2 \geq 4, \quad 5x_1 + x_2 \geq 5, \quad x_1 \geq 0, \quad x_2 \geq 0, \\ \boldsymbol{F}_k(\boldsymbol{x} - \boldsymbol{x}_k) \leq y, \quad \min y, \end{gathered} \tag{9.2}$$

where $\boldsymbol{G}_k$ and $\boldsymbol{F}_k$ are the following gradient vectors

$$\boldsymbol{G}_k = \begin{pmatrix} 2\Phi\left(\dfrac{L_1(\boldsymbol{x}_k) - 0.2L_2(\boldsymbol{x}_k)}{0.4\sqrt{6}}\right)\varphi(L_2(\boldsymbol{x}_k)) + \Phi\left(\dfrac{L_2(\boldsymbol{x}_k) - 0.2L_1(\boldsymbol{x}_k)}{0.4\sqrt{6}}\right)\varphi(L_1(\boldsymbol{x}_k)) \\ \Phi\left(\dfrac{L_1(\boldsymbol{x}_k) - 0.2L_2(\boldsymbol{x}_k)}{0.4\sqrt{6}}\right)\varphi(L_2(\boldsymbol{x}_k)) + \Phi\left(\dfrac{L_2(\boldsymbol{x}_k) - 0.2L_1(\boldsymbol{x}_k)}{0.4\sqrt{6}}\right)\varphi(L_1(\boldsymbol{x}_k)) \end{pmatrix},$$

$$\boldsymbol{F}_k = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

Altogether, 30 iterations have been performed. The program was run on a CDC 3300 computer. The values of the bivariate normal distribution were obtained by simulation. Most of the computer time was consumed by this. As the solution we obtained[2] $x_{1\,\mathrm{opt}} = 1.055$, $x_{2\,\mathrm{opt}} = 3.200$ and $f(\boldsymbol{x}_{\mathrm{opt}}) = 9.565$. The same problem was solved by the SUMT, interior point method. The computer time was less than in the former case.

Another, much larger problem was also solved on the same computer, where the number of linear inequality constraints was 50, the number of variables was also 50 and there was one probabilistic constraint constructed out of 4 constraints of the underlying problem with normally distributed and dependent random variables β_1, β_2, β_3, β_4. The objective function was linear. The problem was solved by the method of feasible directions combined with the simulation of the normal distribution, in 25 minutes. Seven iterations had to be performed, and in this case more time was spent for the nonlinear programming procedure than for the simulation.[3]

Finally, we mention a useful formula concerning the derivative of the normal distribution function. We consider $\varphi(z_1, \ldots, z_m; R)$, $\Phi(z_1, \ldots, z_m; R)$, where $R = (r_{ik})$. It is well-known that (see e.g. [8])

$$\frac{\partial \Phi(z_1, \ldots, z_m; R)}{\partial r_{12}} = \int_{-\infty}^{z_3} \cdots \int_{-\infty}^{z_m} \varphi(z_1, z_2, u_3, \ldots, u_m; R)\, \mathrm{d}u_3 \ldots \mathrm{d}u_m. \tag{9.3}$$

A similar formula holds for the case of r_{ik}. It follows from (9.3) that

$$\begin{aligned} \frac{\partial \Phi(z_1, \ldots, z_m; R)}{\partial r_{12}} &= \frac{\partial \Phi(z_1, \ldots, z_m; R)}{\partial z_1\ \partial z_2} \\ &= \frac{\partial}{\partial z_2} \Phi\left(\frac{z_2 - r_{12} z_1}{\sqrt{(1 - r_{12}^2)}}, \ldots, \frac{z_m - r_{1m} z_1}{\sqrt{1 - r_{1m}^2}}; R_1 \right) \varphi(z_1), \end{aligned}$$

where R_1 is the correlation matrix defined by (8.5). The derivative on the right-hand side can be obtained by (8.4).

References

[1] E.M.L. Beale, "On minimizing a convex function subject to linear inequalities", *Journal of the Royal Statistical Society Ser.* B 17 (1955) 173–184.

---

[2] I express my thanks primarily to Dr. I. Deák, who developed the computer program, and also to Mr. T. Rapcsák, who gave important help in the programming of the SUMT method.

[3] This was an economic problem formulated for the electrical energy sector of the Hungarian economy.

[2] E.F. Beckenbach and R. Bellman, *Inequalities* (Springer, Berlin, 1961).

[3] A. Charnes, W.W. Cooper and G.H. Symonds, "Cost horizons and certainty equivalents: an approach to stochastic programming of heating oil production", *Management Science* 4 (1958) 236–263.

[4] G.B. Dantzig, "Linear programming under uncertainty", *Management Science* 1 (1955) 197–206.

[5] G.B. Dantzig and A. Madansky, "On the solution of two stage linear programs under uncertainty", in: *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability* (University of California Press, Berkeley, Calif., 1960) 165–176.

[6] J.L. Doob, *Stochastic processes* (Wiley, New York, 1953).

[7] A.V. Fiacco and G. P. McCormick, *Nonlinear programming: Sequential unconstrained minimization technique* (Wiley, New York, 1968).

[8] S.S. Gupta, "Probability integrals of multivariate normal and multivariate t", *Annals of Mathematical Statistics* 34 (1963) 829–838.

[9] A. Prékopa, "On probabilistic constrained programming", in: *Proceedings of the Princeton Symposium on Mathematical Programming* (Princeton University Press, Princeton, N.J., (1970) 113–138.

[10] A. Prékopa, "Logarithmic concave measures with application to stochastic programming", *Acta Scientiarum Mathematicarum* (*Szeged*) 32 (1971) 301–316.

[11] A. Prékopa, "On logarithmic concave measure and functions", *Acta Scientiarum Mathematicarum* (*Szeged*), 34 (1973) 339–343.

[12] R. Wets, "Programming under uncertainty: the solution set", *SIAM Journal of Applied Mathematics* 14 (1966) 1143–1151.

[13] S.S. Wilks, *Mathematical statistics* (Wiley, New York, 1962).

[14] G. Zoutendijk, *Methods of feasible directions* (Elsevier, Amsterdam, 1960).