B. BEREANU

Center of Mathematical Statistics
Bucharest

Minimum Risk Criterion in Stochastic Optimization

The paper analyses the advantages of the miniumum risk criterion in stochastic optimization; this criterion is applied in nonlinear programming, and a quasiconvexity condition of the obtained problem is also given. The paper establishes the relation between the level of „ruin" in the minimum risk programming, the ruin minimal probability and the safety coefficient in the optimization with two objective functions (maximizing the expectation and minimizing the objective-function variance, respectively).

REMARKS ON THE PRINCIPLE OF OPTIMIZING EXPECTATION

The theory of classical (deterministic) optimization was initially extended to the stochastic case — which corresponds better to the real conditions wherein decisions are made in practical problems — by replacing the usual objective function through its mean value.

In this procedure, the concern of reducing a new problem to a known one is preponderant, and not the concern of giving the decision problem a formulation adequate to the introduced stochastic framework. As a matter of fact, the selection of an optimal solution, for instance in a maximizing problem, according to the mentioned criterion, arises serious objections.

a) The expectation optimization criterion supposes the possibility of the repetability of the decision-process a great number of times. But this possibility can be annulled in practice by the result of the first selection(s) (for instance, the case of ruin in an investment problem).

b) Even if the possibility of repeating the decision process a great (infinite) number of times is postulated (for instance, assuming the existence of an infinite amount of resources in investment problems), the decision maker is generally interested, in practical problems, in the result of a selec-

tion taking place *only once* in certain existing conditions, hence the expectation of this result does not present interest.

c) The principle of maximizing the expectation of the income does not explain the practice of *diversifying* the investments. According to this principle, the investor would have to select only that investment which maximizes the expectation of the income.

d) It is likely that, from the practical viewpoint, the most important shortcomming of the obtained solution by optimizing the expectation, is represented by the relatively small probability by which the optimal solution assures the achievement of a value of the objective function, at least equal to the corresponding expectation (this probability is 0.5 in the case of a linear objective function and of a symmetrical distribution).

The utility theory and the introduction of the principle of maximizing the utility function expectation, in the works of von Neumann and Morgenstern [1], Friedman and Savage [2], eliminates the objection c) because, by selecting conveniently the utility function, the diversification of the investments can be explained (see Samuelson [3]). Nevertheless, even the principle of maximizing the utility function expectation does not eliminate objection a) and b).

In investment problems and by a more careful analysis, not only in such problems, the main factor to be considered in practice in selecting investments is their quality, from the point of view of their safety, and only afterward the income they produce is taken into account.

All these prove that the maximization of the expectation of the income or of the corresponding utility does not represent an acceptable mathematical model, in many decision problems, in conditions of uncertainty. As a matter of fact, even the necessity of utilizing a function of utility, difficult to build in practice (as it needs a great volume of information, usually nonexistent), makes this approach difficult to apply.

1. MINIMUM RISK CRITERION

The minimum risk criterion was introduced independently, in different contexts, by Bereanu [4], [5], [6], [7]. Charnes and Cooper [8] and Roy [9]. The first author introduced this criterion in connection with stochastic linear optimization problems, which appear when a central planning authority sets up plan indicators for various subordinated enterprises. Charnes and Cooper introduced an analogous model (model P) within stochastic linear programming with probabilistic constraints, and the last author, used this criterion "safety first" in problems of investment selection.

Let us suppose that, instead of the deterministic optimization problem:

$$max\ f(x) \tag{1.1}$$

subject to

$$g_i(x) \leqq 0,\ i = 1, ..., m;\ h_j(x) = 0,\ j = 1, ..., n \tag{1.2}$$

where f, g_i, h_j are real valued functions of the vector $x \in R^n$ (if maximum is reached), the objective functions also depend on a random variable or

vector $\xi(\omega)$, on a probability space (Ω, K, P). We shall denote the new objective function $f(x, \xi(\omega))$ or shorter $f(x, \omega)$. We shall also denote analogously the functions which define the constraints, if these are also stochastic.

We shall also make the hypothesis that, for every x satisfying the constraints $g_i(x) \leqq 0,\ i = 1, ..., m;\ h_j(x) = 0,\ j = 1, ..., k$, the function $f(x, \omega)$, $x \in R^n, \omega \in \Omega$ is measurable. It is obvious that the problem of determining a vector x^*, which satisfies the constraints and maximizes the function $f(x, \omega)$ makes no sense. Taking into account the objections mentioned in connection with the expectation optimization principle, we may ask ourselves how the management of an enterprise would have to select rationally the production plan x, *if* $g_i(x) \leqq 0,\ i = 1, ..., h_j(x) = 0,\ j = 1, ... k$ would represent the constraints (technological, resources, physical production, etc.), and $f(x, \omega)$ would represent the profit which, according to a planned target would have to be at least d.

It seems adequate to select that plan x, which satisfies the constraints and maximizes the probability that the target be fulfilled. We have hence the following formulation of the "minimum risk solution" in stochastic programming. Let us determine a vector x^* (if such a vector exists), which is the solution of the following non-linear programming problem

$$\begin{aligned} &\max_x P\{\omega \mid f(x, \omega) \geqq d\}, \\ &g_i(x) \leqq 0,\ i = 1 . ..., m, \\ &h_i(x) = 0,\ j = 1, ..., k. \end{aligned} \tag{1.3}$$

In the initial formulation of this problem (Bereanu [5], [6]) one minimizes the probability that the value of the objective function be less than a given number d; hence the term of minimum risk solution. In the terminology of the investment theory (portfolio selection) d is the "ruin level". The minimal risk solution, interpreted within this theory, is the solution that minimizes the probability of ruin.

Remark. From this interpretation, the lack of realism of the supposition of repeating a great number of times, in identical conditions, the optimal selection, implicit in the expectation optimization principle, appears clearly. It is clear that, practically, in all optimization problems, in a stochastic framework, there is such a "ruin" or disaster level (which moreover could be more easily determined than an utility function). This is so either when it comes to practical problems, as the control of a river basin, investments, production plan for export, or when it comes to statistical methods of optimal forecasting, which are in fact stochastic optimization problems.

2. GENERAL CASE

We assumed in (1.3) that the functions $g_i(x), h_j(x),\ i = 1, ..., m; j=1,..,k$ do not depend on the random vector $\xi(\omega)$. But if the constraints are also random we must redefine the constraints, because the condition that they be satisfied for any $\omega \in \Omega$ could be in practice too restrictive.

Let $g_i(x, \omega)$, $h_j(x, \omega)$ be functions with real values on $R^n \times \Omega$, supposed to be measurable for any $x \in R^n$, and $p_i, q_j^+, q_j^-, i = 1, ..., m; j = 1, ..., n$ some nonnegative constants, representing the unit penalties for non satisfying the constraints. A possibility of formulating the minimum risk problem, in this case, is as follows:

$$\min_{x, y, z^+, z^-} P\left\{\omega \left| \left[f(x, \omega) - \left(\sum_{i=1}^{m} p_i y_i + \sum_{j=1}^{k} (q_j^+ z_j^+ + q_j^- z_j^-) \right) \right] \leqslant d \right.\right\} \tag{2.1}$$

$$g_i(x, \omega) - y_i \leqq 0, \; i = 1, ..., m$$

$$h_i(x, \omega) + z_j^- - z_j^+ = 0, j = 1, ..., k \tag{2.2}$$

$$y_i \geqq 0, z_j^- \geqq 0, z_j^+ \geqq 0.$$

This formulation is similar to the one introduced independently by Dantzig [10] and Beale [11] for stochastic linear programming, with the difference that the minimum risk criterion is used in (2.1) instead of minimizing the expectation. It presents surely interest to determine the conditions under which the nonlinear programming problem (2.1), (2.2) is a problem of convex programming (see 3.3). In the following, the "minimum risk problem" will be abreviated through MRP.

3. MRP WITH DETERMINISTIC CONSTRAINTS

a) *Case of linear objective function*

$$min\,[P\{c(\omega)\, x\} \leqq d],$$

$$g_i(x) \leqq 0, \; i = 1, ..., m, \tag{3.1}$$

$$h_j(x) = 0, \; j = 1, ..., k.$$

In (3.1) $c(\omega)$ is a random vector $1 \times n$, $d \in R$, and x is a n-dimensional column vector.

3.1. Theorem. *If $c(\omega)$ has a normal non-degenerated distribution, with mean value $\bar{c}$ and covariance matrix V, MRP (3.1) is equivalent to the following non-linear programming problem:*

$$min \frac{d - \bar{c}x}{\sqrt{x'Vx}},$$

$$g_i(x) \leqq 0, i = 1, ..., m, \tag{3.2}$$

$$h_j(x) = 0, j = 1, ..., k,$$

supposing (so as to eliminate trivial cases) that

$$(g_i(x) \leqq 0, \; h_j(x) = 0) \Rightarrow (x \neq 0).$$

Proof. See [6], where the case of linear programming with stochastic objective function is considered.

b) *Case of nonlinear objective function*

If $X = \{x \in R^n;\ g_i(x) \leqq 0,\ i = 1, ..., m\}$ is the set of constraints and the functions $g_i(x)$ are convex on R^n, the set X is convex.

For solving MRP, the convexity properties of the following function present interest:

$$g(x, d) \equiv P\{\omega \mid f(x, \xi(\omega)) \leqq d\}.$$

3.2. Definition. (Prékopa [12]). Let P be a probability measure on the measurable space $(R^n, \mathfrak{B})$ where $\mathfrak{B}$ is the family of the Borel subsets in R^n. The probability measure P is said to be logconcave if, for any pair M, N of convex sets from $\mathfrak{B}$ and any $\lambda \in (0,1)$ the following takes place:

$$P\{\lambda M + (1 - \lambda)N\} \geqq (P\{M\})^{\lambda} (P\{N\})^{1-\lambda}, \tag{3.3}$$

where $\{M + N\}$ represents the vectorial sum of sets M, N i.e. $M + N = \{m + n \mid m \in M, n \in N\}$.

The distribution function corresponding to a logconcave probability measure is said to be logconcave.

3.3. *If the function $f(x, \xi(\omega))$ is concave on X in $(x, \xi(\omega))$, and the distribution of vector $\xi(\omega)$ is logconcave, then, for any $d \in R$, $g(x, d)$ is a quasiconvex function on X.*

Proof. It will be sufficient to show that $-g(x, d)$ or $1\text{-}g(x, d) \equiv h(x, d)$ is quasiconcave on X. But

$$h(x, d) = P\{\omega \mid f(x, \xi(\omega)) > d\}.$$

Let denote $K(x) \equiv \{\xi \mid f(x, \xi(\omega)) > d\}$, $x_1, x_2, \in X$, $K(x_1) \neq \varnothing$, $K(x_2) \neq \varnothing$ and $\xi_1 \in K(x_1)$, $\xi_2 \in K(x_2)$.

As $f(x_1, \xi_1) > d$, $f(x_2, \xi_2) > d$, due to the concavity of $f(x, \xi(\omega))$, $f(\lambda x_1 + (1 - \lambda) x_2,\ \lambda\xi_1 + (1 - \lambda)\xi_2) > d$, i.e.

$$K(\lambda x_1 + (1 - \lambda)x_2) \supset \lambda K(x_1) + (1 - \lambda) K(x_2), \tag{3.7}$$

and the sets $K(x)$ are convex. By applying probabilities in (3.7) we obtain:

$$P\{\omega \mid \xi(\omega) \in K(\lambda x_1 + (1 - \lambda) x_2)\} \geqq P\{\xi \mid \xi \in \lambda K(x_1) + (1 - \lambda) K(x_2)\}. \tag{3.8}$$

Using (3.3) in the right-hand side of (3.8) we have

$$P\{\omega \mid \xi(\omega) \in K(\lambda x_1 + (1 - \lambda) x_2)\} \geqq (P\{\omega \mid \xi(\omega) \in K(x_1)\})^{\lambda} (P\{\omega \mid \xi(\omega) \in K(x_2)\})^{1-\lambda},$$

or

$$h(\lambda x_1 + (1 - \lambda) x_2) \geqq [h(x_1)]^{\lambda} [h(x_2)]^{1-\lambda}. \tag{3.9}$$

We suppose that $h(x_2) \geqq h(x_1)$ takes place. It results from (3.9) $h(\lambda x_1 + (1 - \lambda)x_2) \geqq h(x_i)$.

i.e. $h(x, d)$ is quasiconcave on X and hence $g(x, d)$ is quasiconvex.

3.4. *Remarks.* Several multidimensional distributions of usual probabilities have the property of being logconcave [12]. So are the normal distributions, Wishart, as well as the multidimensional distributions Dirichlet and Beta. We see from (3.9) that set $X' = \{x \in X \mid h(x) > 0\}$ is convex, and the function $\log h(x)$ is a function concave on X'.

4. CASE OF THE PARTIALLY UNKNOWN DISTRIBUTION FUNCTION

We shall suppose in this paragraph that the distribution function of the random vector $\xi(\omega)$ is unknown. We shall also suppose, following [9], that for any $x \in X$, $m(x)$ and $D(x)$ i.e. the expectation and the variance of the corresponding random variable $f(x, \xi)$, can be determined. For any $d \in R$ and $x \in X$ which satisfy $m(x) - d > 0$, it results from Chebyshev's inequality

$$P(|f(x, \xi) - m(x)| \geq m(x) - d) \leq \frac{D(x)}{(m(x) - d)^2}.$$

Hence

$$P(f(x, \xi) \leq d) = P(m(x) - f(x, \xi) \geq m(x) - d)) \leq \frac{D(x)}{(m(x) - d)^2}. \tag{4.1}$$

Hence, if we do not know the distribution function of the random vector $\xi(\omega)$, but we know $m(x)$ and $D(x)$ for $x \in X$, we can obtain an approximate solution of problem (3.1), by solving the non-linear programming problem

$$\min \frac{d - m(x)}{D(x)}, \tag{4.2}$$

$$g_i(x) \leq 0, \quad i = 1, \ldots, m$$

$$h_j(x) = 0, \quad j = 1, \ldots, k,$$

where $X \equiv \{x \in R^n, g_i(x) \leq 0, \ i = 1, \ldots, m, \ h_j(x) = 0, j = 1, \ldots, k\}$ is compact.

We notice that, in the case considered in theorem 3.1, the solution of problem (4.2) is just the minimum risk solution.

The function $\psi(x) = \dfrac{m(x) - d}{\sqrt{D(x)}}$, is an increasing function of the mean value $m(x)$ and decreasing in the variance $D(x)$. It results that problem (4.2) is a way of approaching an optimization problem with two objective functions, where the maximizing of mean value $m(x)$ and the minimizing of the variance $D(x)$ would be desired. We can establish a more precise connexion between problem (4.2) and an optimization problem with two objective functions $m(x)$ and $D(x)$.

As is well known, the replacement of two objective functions through their linear combination, using certain weights, is a way of treating an optimization problem with two objective functions.

Let hence

$$\max_{x \in X} [m(x) - p\sqrt{D(x)}] \tag{4.3}$$

be the optimization problem, where the set X is compact. We shall call the coefficient p *safety coefficient.* This coefficient measures the aversion towards risk of the decision maker. We have the following result.

4.1. Theorem. *The vector x_0 is an optimal solution of problem*

$$\max_{x \in X} \psi(x) \tag{4.4}$$

if and only if it is also an optimal solution of problem (4.3) *with $p = p_0 = \psi(x_0)$ and we have*

$$m(x_0) - p_0\sqrt{D(x_0)} = d. \tag{4.5}$$

Proof. Let x_0 be an optimal solution in (4.4.)

$$\text{We have } p_0 = \frac{m(x_0) - d}{\sqrt{D(x_0)}} \geqq \frac{m(x) - d}{\sqrt{D(x)}}, \quad \forall x \in X.$$

Hence

$$m(x) - p_0\sqrt{D(x)} \leqq d, \quad \forall x \in X$$

and

$$m(x_0) - p_0\sqrt{D(x_0)} = d,$$

hence x_0 is optimal solution in (4.3), if we take $p = p_0$ so that equality (4.5) be satisfied.

From $m(x) - p_0\sqrt{D(x)} \leqq d$, $\forall x \in X$ and (4.5) it results

$$\frac{m(x) - d}{\sqrt{D(x)}} \leqq p_0, \quad \forall x \in X$$

and

$$\frac{m(x_0) - d}{\sqrt{D(x_0)}} = p_0,$$

i.e. x_0 is optimal solution in (4.4), and p_0 is the value of optimum.

4.2. This theorem offers us an intuitive interpretation of the connection between the "ruin" level d and the minimal probability π of being ruined, as well as between this probability and the safety coefficient.

This connection will be illustrated by means of problem (3.2).

Let have

$$\max_{x \in X} \frac{\bar{c}x - d}{\sqrt{x'Vx}} = \frac{\bar{c}x_0 - d}{\sqrt{x_0'Vx_0}} = p_0.$$

Then the minimum in (3.1) is given by $\pi_0 = \Phi(-p_0)$, where $\Phi(\cdot)$ is Laplace's function.

Hence $p_0 = -\Phi^{-1}(\pi_0)$. The safety coefficient furnishing in (4.3) a value at the utmost equal to the ruin level d is $-\Phi^{-1}(\pi_0)$, where π_0 is the corresponding ruin probability. We note that, if we take $p = 0$, problem (4.3) reduces itself to the problem of maximizing the expectation. Hence if $d = \max_{x \in X} \bar{c}x$, the corresponding ruin probability is 1/2, which corresponds to intuition. The minimal ruin probability is a decreasing function of the safety coefficient p and it obviously decreases when d increases. When $p_0 \to +\infty$, $\pi_0 \to 0$.

5. CASE OF DISCRETE PROGRAMMING

All observations concerning the advantages of the minimal risk criterion in comparison to that of optimizing the expectation remain valid also in the case of a discrete set of feasible solutions. The most frequent case is that in which the components of the decision vector x are integer numbers. Often the decision is of the type "yes or not"", i.e. the components of the vector x take values 0 or 1.

We shall exemplify this case by an investment problem. We suppose to have n possible investment projects, project i, needing investment c_i, $i = 1, ..., n$, and p_i, total product (investment + profit) at the end of the time period taken into account, being a random variable. We suppose that we know the expectation $\bar{p}$ of the random vector $p = (p_1, ..., p_n)$ and the matrix of its covariances V.

Let x_i, $i = 1, ..., n$ be a decision vector which takes value 0 or 1, according as the investment is made in project i or not.

If M is the total sum we dispose of for the investments, and d is the minimal level of the total product fixed for the end of the planned time period, the minimal risk criterion leads us to the following bivalent stochastic programming problem:

$$\min_x p\left\{\omega \,\middle|\, \sum_{i=1}^{n} p_i x_i \leqq d\right\}, \tag{5.1}$$

$$\sum_{i=1}^{n} c_i x_i \leqq M$$

$$x_i = 0 \text{ or } 1, \; i = 1, ..., n.$$

In the case of a multinormal distribution of vector p, the problem comes up to the solving of the following non-linear programming problem with bivalent variables:

$$\max_x \frac{\bar{p}x - d}{\sqrt{x'Vx}},$$

$$\sum_{i=1}^{n} c_i x_i \leqq M$$

$$x_i = 0 \text{ or } 1, \; i = 1, ..., n.$$

Remarks. Methods of solving the minimal risk problem (3.2) with linear constraints are given in [5], [7], [13], and [14]. In [15] an extension of the minimum risk solution to the linear case, with several objective functions is considered.

REFERENCES

1. NEUMANN VON J., MORGENSTERN O.: *Theory of games and economic behaviour.* Princeton: Princeton Unibersity Press, 1947.
2. FRIEDMAN, M., SAVAGE L.J.: *The utility analysis of choices involving risk.* Journal of Political Economy, vol. 56, 279—304, 1948.
3. SAMUELSON P. : *General proof that diversification pays.* Journal of Financial and Quantitative Analysis, 2, 1—13, 1967.
4. BEREANU B.: *Asupra programării liniare stochastice* (On stochastic linear programming), Paper to the Conference of the theory of probabilities, Braşov, September, 1962.
5. BEREANU B.: *Problema funcţiei de repartiţie în programarea stochastică şi soluţii cu risc minim* (The distribution problem in stochastic linear programming and minimum risk solution). Dissertation, Universitatea Bucureşti, 1963.
6. BEREANU B.: *Programme de risque minimal en programmation linéaire stochastique.* Comptes Rendus Academie des Sciences, Paris, *259*, 5, 981—983, 1964.
7. BEREANU B.: *Soluţii cu risc minim în programarea liniară* (Solutions with minimum risk in linear programming). Analele Universităţii Bucureşti, *13*, 121—140, 1964.
8. CHARNES A., COOPER W.W.: *Deterministic equivalent for optimizing and satisficing under chance constraints.* Operations Research, *11*. 18—39, 1963.
9. ROY A.D.: *Safety first and the holding of assets.* Econometrica, *37*, 181—196, 1963.
10. DANTZIG G.B.: *Linear programming under uncertainty*, Management Science, 1, 3—4, 197—206, 1955.
11. BEALE M.L.: *On minimizing a convex function subject to linear inequalities.* Journal of Royal Statistical Society, Ser. B., 17, 173—184, 1955.
12. PRÉKOPA A.: *Logarithmic concave measures with applications to stochastic programming.* Acta Scientiarum Mathematicarum, 32, 3—4, 301—316, 1971.
13. BERGTHALLER C.: *A quadratic equivalent for the minimum risk problem,* Rev. Roumaine Math. Pures et Appl., *15*, 1, 17—23, 1970.
14. DRAGOMIRESCU M.: *An algorithm for the minimum risk problem of stochastic programming.* Operations Research, *20*, 154—164, 1972.
15. STANCU-MINASIAN I.M.: *Programarea stocastică cu mai multe funcţii obiectiv* (Stochastic programming with several objective functions). Editura Academiei R.S.R., Bucureşti, 1980.

LE CRITERE DU RISQUE MINIME DANS L'OPTIMISATION STOCHASTIQUE

On fait l'analyse des avantages du critère du risque minime dans l'optimisation stochastique et on applique ce critère dans la programmation non-linéaire offrant aussi une condition de quasiconvexité du problème résulté. On établit la relation entre le niveau de „ruine" de la programmation à risque minime, la probabilité minime de ruine et le coefficient de sûreté de l'optimisation à deux fonctions-objectif (la maximisation de la valeur moyenne, respectivement la minimisation de la dispersion de la fonction objectif).

КРИТЕРИИ МИНИМАЛЬНОГО РИСКА ПРИ СТОХАСТИЧЕСКОЙ ОПТИМИЗАЦИИ

Проанализируются преимущества критерия минимального риска при стохастической оптимизации и применяется этот критерий в нелинейном программировании, а также приводится условие квазивыпуклости полученной задачи. Устанавливается отношение между уровнем «разорения» при программировании с минимальным риском, минимальной вероятностью разорения и коэфициентом уверенности при оптимизации по двум целевым функциям (максимизация среднего значения и минимизация дисперсии целевой функции).

3.04.1981