PROBABILISTIC OPTIMIZATION PROBLEMS

ANDREI I. KIBZUN
Moscow Aviation Institute
USSR

Technical Report 91-34

November 1991

THE UNIVERSITY OF MICHIGAN

Andrei I. Kibzun, Professor, *Moscow Aviation Institute, USSR*

Lecture # 1

Probabilistic Optimization Problems

Introduction

The lecture is devoted to stochastic optimization problems with criteria in the form of probability (in my lectures later I will refer to these problems as probabilistic problems). The author has developed a confidence approach for solving such problems. First of all, let me describe schematically the place taken by this method among the well-known methods. For this purpose I would like to offer the following classification of optimization problems in presence of uncontrollable factors, though somebody can find it arguable.

In statistics usually the Bayesian and minimax approaches are usually applied separately for estimating the unknown parameter. According to the minimax approach this parameter is assumed to be uncertain, and the a priori information about it is assumed to be given in the form of a range of uncertainty. According to the Bayesian approach, the parameter is assumed to be random and its distribution is assumed to be known. Similar approaches have also been considered in control and filtering theories. Please note that the stochastic and minimax theories corresponding to these approaches have been developing independently as a rule. That is why, I think, the mathematical apparatuses of these theories don't match each other. I would like to stress that, in my opinion, using one approach or the other is a matter of faith rather then a matter of truth. Supporters of the minimax theory just like atomists in philosophy try to split a problem "building bricks" without providing for any chance. They say God does not play dice. Another group of scientists think to the contrary opinion. They say that everything in our world is so interrelated and diverse that the

only way out is to stochastically describe the various phenomena around us. To cut a long story short the situation reminds me the long dispute between idealists and materialists. I hope you will excuse me for such a free interpretation of these approaches. But some scientists, for example Germeyer (USSR) in his monograph, criticize this contrasting the minimax approach with the stochastic one. By the way, Wald's theory of statistical solutions can exemplify a good unity of these approaches. According to this theory, optimization problems were considered with regard to random and uncertain factors, with the maximum of averaged loss function as the optimality criterion.

Similar minimax–stochastic problems have also been examined by Kurzhanskii, Ermoliev, Eliasberg (USSR), Loose, Poor, and Vastola (USA). I would like to call your attention to the fact that these scientists have examined only a formal combination of the minimax approach and the stochastic one. A good deal of effort has been devoted to overcoming the difficulties resulting from the lack of agreement between different mathematical apparatuses of the minimax theory and the stochastic one. While noting the doubtless usefulness of these scientists' works, I'd like to stress the fact that a symbiosis of the stochastic and minimax theories has been used when stating and solving minimax–stochastic problems. These works have failed to take into account the existing relationship between stochastic and minimax problems, although the following fact could have suggested this idea. Equations of the Kalman filter are true for a linear system both at Gaussian random interference and at uncertain interference from an uncertainty set in the form of an ellipsoid.

Now I would like to express the main idea of my research. I think that in the near future new results in the field of optimization with uncontrollable factors should be looked for in the vertical axis, as the figure shows, by revealing links between the minimax and stochastic theories as well as the theory of fuzzy sets, rather than in the horizontal plane, that is, within separate theories. In this case a search for linking bridges between those theories will become very important. Those bridges will help carry the results from one theory into the

other. Today I would like to tell you about my modest attempt to construct such a bridge. This bridge is constructed for stochastic problems with criterion in the form of an objective function quantile or, in short, with quantile criterion. Why do we consider these particular criteria?

For the last twenty years I have been studying the problems of analysis and synthesis of the algorithms of high-precision control and estimation, as well as the problem of using the obtained results in aviation and astronautics. In 1987 a monograph devoted to these problems was published in the USSR which was written by myself in collaboration with Professor Malyshev.

High precision of object position estimation as well as of its motion control is one of the main requirements when designing modern motion control systems. To provide the necessary precision one has to take into account different uncontrollable factors (random, uncertain, and fuzzy) when designing algorithms of estimation and control. Generally, we first consider random factors acting upon an object when it is functioning. For example, when one controls an aircraft, he has to take into consideration atmospheric disturbances (deviations of the atmospheric density or wind gusts), errors of working through control impulses, and deviations of geometric, aerodynamic, and other vehicle characteristics. The necessity of considering these factors influences the problem statement.

Consider, for example, a terminal control problem of a certain hypothetical aircraft. The aim for this aircraft is to reach the given destination with minimal fuel consumption. It can be seen that the formulated problem can only be discussed in a probabilistic sense. Thus, it is natural to choose as the optimality criterion the minimal fuel capacity necessary to keep the probability of reaching the goal above a given value. Therefore we must choose an optimal control strategy guaranteeing that the terminal requirements will be met with the given probability. This situation appears in the majority of control problems. Many probabilistic control quality characteristics such as the probability of a successful landing, the probability of two vehicles rendezvous have become traditional.

In aircraft control problems one wants to know the probability of landing successfully. The information that the control system can land the aircraft "in an average" manner is not satisfactory, because one understands that his particular landing can be unsuccessful as 99 following landings.

Some other examples of probabilistic control quality criteria can be found in the economic applications. Consider the following example, which is referenced to as the "stock-exchange paradox". Lets suppose one is going to buy some shares of stock and there are only two different types of stocks available. Let z_1 be the initial amount of money one is going to invest in the stocks, w_{i1} and w_{i2} be the profits gained in i-th year for the first and the second type of stock respectively. The vectors $w_i = col(w_{i1}, w_{i2})$ are independent and they have the same distribution for any i. Let u_1 and u_2 be the fractions of the amount of money one is going to spend on the shares of the first and the second type respectively, $u_1 + u_2 = 1$. Then the total amount of money after year N will be

$$z_{N+1} = z_1 \prod_{i=1}^{N} (1+u_1 w_{i1} + u_2 w_{i2})$$

The expectation of the total amount will be a criterion of our investment efficiency

$$\Phi(u_1, u_2) = M\left[z_{N+1}(u_1, u_2, w_{11}, w_{12}, \ldots, w_{N1}, w_{N2})\right] =$$

$$= z_1 \prod_{i=1}^{N} (1+u_1 m_1 + u_2 m_2),$$

where $m_1 = M[w_{i1}]$, $m_2 = M[w_{i2}]$. Then, it is obvious that the optimal strategy will be to invest only in the stock that gives the maximal profit, i.e.

if $m_1 < m_2$, then $u_1^0 = 0$, $u_2^0 = 1$

if $m_1 > m_2$, then $u_1^0 = 1$, $u_2^0 = 0$

When using this strategy our criterion is

$$\Phi(u_1^0, u_2^0) = \prod_{i=1}^{N} (z_1 + \max\{m_1, m_2\}) \underset{N \to \infty}{\longrightarrow} 1$$

The paradox is that this strategy leads to a probability of ruin

$$\mathbb{P}\{ z_{N+1} \leq 0 \} \underset{N \to \infty}{\to} 1$$

Thus, it is not sensible to use the expectation as a criterion in this particular case. One can use quantile criterion to avoid this paradox.

Nevertheless, I think that until recently inadequate attention has been paid to probabilistically stated control problems. Apparently, it can be explained by the complexities of taking into consideration probabilistic criteria in comparison with classic characteristics such as mathematical expectation and root-mean-square deviation. Also, there is a lack of methods for solving this type of problems. However, some research has been done in this field, though in a rather uncoordinated manner. In this lecture I will try to analyze the previous experience in this area. I will begin with classification.

Classification of Problems

According to the type of probabilistic criterion, problems can be divided into three large classes: direct problems, inverse problems, and problems with chance constraints. According to the type of object to be controlled, probabilistic optimization problems as well as other optimization problems can be divided into static ones and dynamic ones. The problems, in which the object to be controlled is described by systems of static relations such as equations or inequalities, are called static. If in this case a set of feasible controls is finite-dimensional space, the static probabilistic problems are actually the probabilistically stated problems of mathematical programming.

There are problems where an object under control is described by a system of differential and finite-differential equations. I will refer to them as dynamic probabilistic problems. Though, formally, dynamic problems can be reduced to static ones with an objective function written by the dynamic relation, this division is still useful in practice. The dynamic object state control and estimation problems, when probabilistically stated, can be the examples of dynamic probabilistic problems. In this lecture I will devote most of my

attention to static probabilistic problems. Now I will turn to a formal description of probabilistic problems.

Let an arbitrary probabilistic space $(\Omega, \mathfrak{J}, \mathbb{P})$ be given, where Ω is a space of elementary events ω, $\mathfrak{J}$ is a σ-algebra of subsets from Ω, $\mathbb{P}$ is a probabilistic measure of subsets from $\mathfrak{J}$. Either random variables, random processes, or random fields can play the role of ω. In these cases the space Ω and probabilistic measure $\mathbb{P}$ are to be determined in accordance with the type of random factor. The most frequent is the case when realizations w are vector of finite $-$ dimensional space $\Omega = \mathbb{R}^n$, $\mathfrak{J}$ is a Borel σ-algebra, and a probabilistic measure is given by the probability density $p(x)$. Assume that certain functionals $\Phi(u, w)$ and $Q(u, w)$ are given that characterize the quality of strategies $u \in U$ selected from the set of feasible strategies U. In the future, to simplify the designations, we will use the symbol w for random factors as well as for its particular realization if it will not lead to misunderstanding. When the static problems considered, the functionals $\Phi(u, w)$ and $Q(u, w)$ are given in an explicit form, and when the dynamic problems are considered, the functionals Φ and Q are given in form of integrals, differential or difference equations. Parameters and functions can play the role of a strategy in static problems. Programmed control, positional control(the control law), and an algorithm of estimating can play this role in dynamic problems. The set U of feasible strategies gives not only the type of strategies mentioned above, but it also shows technical constraints imposed on the values of u. For example, the values of programmed control can be restricted by the constraint $|u(t)| \leq 1$ for all $t \in [0, T]$ and the function $u(t)$ can be assumed piecewise-continuous. Let us assume, in future, that the functionals $\Phi(u, w)$ and $Q(u, w)$ are measurable with respect to w for each strategy $u \in U$. This assumption is satisfied in the majority of practical problems and gives us an opportunity to consider the following **probability functional**

$$P_\varphi(u) \triangleq \mathbb{P}(S_{\varphi, u}), \quad S_{\varphi, u} = \{ w : \Phi(u, w) \leq \varphi, \ Q(u, w) \leq 0 \} \qquad (1)$$

where φ is a certain parameter. The probability $P_\varphi(u)$ characterizes such an event that the functional $\Phi(u, w)$ will not

exceed the threshold φ, and $Q(u,w)$ will not be positive. Usually, in technical problems there are some demands which one has to satisfy in the first place. The constraint $Q(u,w) \leq 0$ plays the role of these demands in this case. The fulfillment of the other constraints can be less categorical. At last, one can increase the threshold φ to satisfy the constraint $Q(u,w) \leq 0$. For example, the functional $\Phi(u,w)$ can characterize the consumption of the control resource, and $Q(u,w)$ can characterize the control precision, or vice-versa, depending on the control goal. The considered functional can be interpreted as a convolution of two criteria.

In these terms, a direct probabilistic problem can be formulated as follows. One has to select a strategy u_φ as to maximize the probability functional $P_\varphi(u)$:

$$u_\varphi \triangleq \arg \max_{u \in U} P_\varphi(u), \quad \varphi \in \mathbb{R}^1. \tag{2}$$

By analogy, an inverse probabilistic problem consists of choosing such a strategy u_α as to minimize the quantile functional $\Phi_\alpha(u)$:

$$u_\alpha \triangleq \arg \min_{u \in U} \Phi_\alpha(u), \quad \alpha \in (0,1) \tag{3}$$

where by definition the quantile functional $\Phi_\alpha(u)$ is a minimal threshold φ guaranteed with a probability α:

$$\Phi_\alpha(u) \triangleq \min \{\varphi: P_\varphi(u) \geq \alpha\}. \tag{4}$$

Note that a direct probabilistic problem of optimization can be reduced to a stochastic problem with the optimality criterion in the form of mathematical expectation

$$u_M \triangleq \arg \max_{u \in U} M [\chi(S_{\varphi,u})] \tag{5}$$

if we introduce a characteristic function $\chi(S_{\varphi,u})$ of the set $S_{\varphi,u}$ which is equal to zero when $w \notin S_{\varphi,u}$ and which is equal to unity when $w \in S_{\varphi,u}$. However, methods of optimization of mathematical expectation greatly depend on smoothness of functionals that are under the sign of mathematical expectations. This property is not present in this case due to discontinuity of the function χ. That is why direct problems require a special study. The quantile statement (3) can not be reduced, in principle, to the expectation problem (5). This happens since in order to calculate the quantile functional $\Phi_\alpha(u)$, one has to minimize the parameter

φ, and, moreover, calculate the functional $P_\varphi(u)$ for every fixed φ. A number of works consider a series of functionals

$$P_i(u) = P\{w: Q_i(u,w) \le 0\}, \quad i=\overline{1,n}$$

and demand that each functional be not lower than a certain given threshold: $P_i(u) \ge \alpha_i$ (the so-called per-line-chance-constraints)

A certain deterministic function $\Phi_0(u)$ is a quality criterion of the strategy u in this case. Such statements are called **problems with chance constraints.** One has to stress again that problems with chance constraints are a special case of inverse problems with a quantile criterion. Indeed, if in problem (3) the objective functional $\Phi(u,w)=\Phi_0(u)$, i.e. it does not depend on a random vector w, we arrive at a problem of minimizing the functional $\Phi_0(u)$ under the constraint

$$\mathbb{P}\{w: Q(u,w) \le 0\} \ge \alpha.$$

And vice versa, if we consider the generalized strategy $\overline{u} = (u,\varphi)$, then the problem of minimizing φ with the constraint $P_\varphi(u) \ge \alpha$ is a problem with chance constraint

$$\overline{u}_0 = \arg\min_u \varphi(\overline{u}), \quad P_\alpha(u) \ge \alpha$$

The problem with a chance constraint sometimes can be reduced to the quantile problem without the additional constraint $Q(u,w) \le 0$. In many practical cases, we can obtain this by a change of variables. Such an example will be given in the third lecture.

Now we will discuss a simple example of a problem with a chance constraint.

$$\Phi_0(u_1, u_2) \to \min_{u_1, u_2}$$

$$\mathbb{P}\{-\Phi_0 + u_1 + w_1 \le 0, \ u_1 + u_2 = w_2 \le 0\} \ge \alpha.$$

We can change the variable

$$u_2 = \overline{u}_2 - \Phi.$$

Then the problem will become

$$\Phi_0(u_1, u_2) \to \min_{u_1, \overline{u}_2}$$

$$\mathbb{P}\{u_1 + w_1 \le \Phi_0, \ u_1 + \overline{u}_2 = w_2 \le \Phi_0\} =$$

$$\mathbb{P}\{\max[u_1 + w_1, \; u_1 + \overline{u}_2 = w_2] \le \Phi_0\} \ge \alpha.$$

We can say that we now have the quantile optimization problem
$$\Phi_\alpha(u_1, \; \overline{u}_2) \quad \to \quad \min_{u_1, \overline{u}_2},$$
where $\Phi_\alpha(u_1, \overline{u}_2) = \Phi_0(u_1, \overline{u}_2)$

with an α–level quantile of random variable distribution

$$\Phi(u_1, \; \overline{u}_2, \; w_1, \; w_2) = \max[u_1 + w_1, \; u_1 + \overline{u}_2 + w_2].$$

It seems to me that the problem with chance constraints is more difficult than the quantile optimization problem, which is an ordinary unconditional optimization problem.

Let's consider one more problem statement. The functionals $\Phi(u, w, z)$ and $Q(u, w, z)$ will now depend on the parameter z with the values from some space Ξ. Let's fix $u \in U$ and consider all the values z such that
$$\mathbb{Z}_{\varphi, \alpha}(u) \triangleq \{z: \; P_\alpha(u, z) \ge \alpha\}.$$

It is obvious that if $\mathbb{Z}_{\varphi, \alpha}(u_1) \supset \mathbb{Z}_{\varphi, \alpha}(u_2)$, then the strategy u_1 is better than the strategy u_2. Therefore, we can use the information about the size of $\mathbb{Z}_{\varphi, \alpha}(u)$ to make conclusions about the efficiency of strategy u. The set $\mathbb{Z}_{\varphi, \alpha}(u)$ is called the set of absorption of α–level for strategy $u \in U$. The maximal α–level set of absorption is
$$\mathbb{Z}_{\varphi, \alpha} = \bigcup_{u \in U} \mathbb{Z}_{\varphi, \alpha}(u).$$
The actual purpose of the absorption set $\mathbb{Z}_{\varphi, \alpha}$ is to determine the set of values of z for which it is possible to find a strategy u guaranteeing the fulfillment of a probabilistic constraint
$$\mathbb{P} \{ \; w: \Phi(u, \; w, \; z) \le \varphi, \quad Q(u, \; w, \; z) \le 0\} \ge \alpha.$$

At the end of this section let us consider minimax–probabilistic problems of optimization, in which there are simultaneously random and uncertain factors. For example, an objective functional $\Phi(u, w, \eta)$ can also depend on an uncertain

factor $\eta \in H$. Then, one can consider the following functional instead of probability functional such as (1)

$$P_{\varphi}^{H}(u) = \inf_{\eta \in H} P_{\varphi}^{\eta}(u).$$

A strategy that maximizes a probability functional when there is the worst disturbance η, is considered as an optimal strategy u_{φ}^{H} in this case. Such problems can be called **maximin – probabilistic** ones. **Stochastic differential games** in which a random factor makes a differential game more complicated, are considered to belong to a class of problems like that. In addition to the method mentioned above, there can be another combination of random and uncertain factors. Let a probability measure P be given from some family Π. In this case, the quality of the strategies can be characterized by the functional

$$P_{\varphi}^{\Pi}(u) = \inf_{P \in \Pi} P_{\varphi}(u).$$

A situation when there is a game of two players in mixed strategies, is also possible. In this case, a payoff function will depend on two strategies p and q, which are random factors for which a probability functional $P_{\varphi}(F_{p}, F_{q})$ can be calculated, and distributions of these strategies F_{p} and F_{q} will play the role of strategies u and v. Thus one arrives at the problem of optimizing a functional:

$$P_{\varphi}^{E}(F_{p}) = \inf_{F_{q} \in E} P_{\varphi}(F_{p}, F_{q}).$$

Direct Methods for Solving Probabilistic Problems

Let us consider two main approaches to solving probabilistic problems, which unite the groups of direct and indirect methods.

Methods which are based on the direct calculations of the functionals $P_{\varphi}(u)$ and $\Phi_{\alpha}(u)$ can be considered as direct methods. For example, in the finite–dimensional case, it is sometimes possible to calculate integrals analytically, or the numerical procedure for calculating them is comparatively simpler.

Let's consider the statistical simulation method which is usually used for the numeric calculation of statistical

characteristics which can't be calculated analytically. When using this method, the functionals $P_\varphi(u)$ and $\Phi_\alpha(u)$ are calculated for every feasible strategy $u \in U$, and then the known methods of deterministic optimization, for example, nonlinear programming methods, are used in the case when $U \subset \mathbb{R}^m$.

Let's consider the details of using statistical simulation methods for calculating the functional $\alpha = P_\varphi(u) = \mathbb{P}\{S_\varphi, u\}$, where

$$S_{\varphi,u} = \{w: \Phi(u,w) \leq \varphi, \ Q(u,w) \leq 0\}.$$

The method which is used for the estimation of probability $\mathbb{P}\{S_\varphi, u\}$ is based upon the Muavre–LaPlace theorem, which states that the frequency W_n of a random event $S_{\varphi,u}$ in n trials, as a random variable, is asymptotically Normal as n tends to infinity. Let the trials be independent and the probability of event $S_{\varphi,u}$ in any trial be equal to α. Then we can obtain the lower bound for the probability α. This lower bound is guaranteed with probability β as can be seen in the equation below.

$$\mathbb{P}_W\left\{W_n : \frac{W_n - \alpha}{\sqrt{\alpha(1-\alpha)}}\sqrt{n} \geq p\right\} = \beta$$

In this equation β is the confidence probability (secondary with respect to α). In this equation it is assumed that the number of trials is fixed, but in actual problems one wants to find the number of trials which insures that the frequency W_n is close to the probability α being estimated. If we solve this equation with respect to u then we obtain

$$n_{min} = \left[\frac{2n_s + p^2(1-\alpha) - |p|\sqrt{4n_s(1-\alpha)+p^2(1-\alpha)^2}}{}\right] + 1$$

where $[\cdot]$ is the whole part of the number, n_s is the number of successful trials in a series of n trials, and W_n is n_s divided by n. In this equation the guaranteeing number of trials depends on the confidence probability β (or p) and on the number of successful trials n_s. The function $n_{min}(n_s)$ is a step function and can be plotted as follows (Fig. 1) (for $\alpha = 0.999$, $\beta = 0.9999$). If in m trials we get the point which is under the curve, then the actual probability $P(A)$ is greater than or equal

Fig. 1
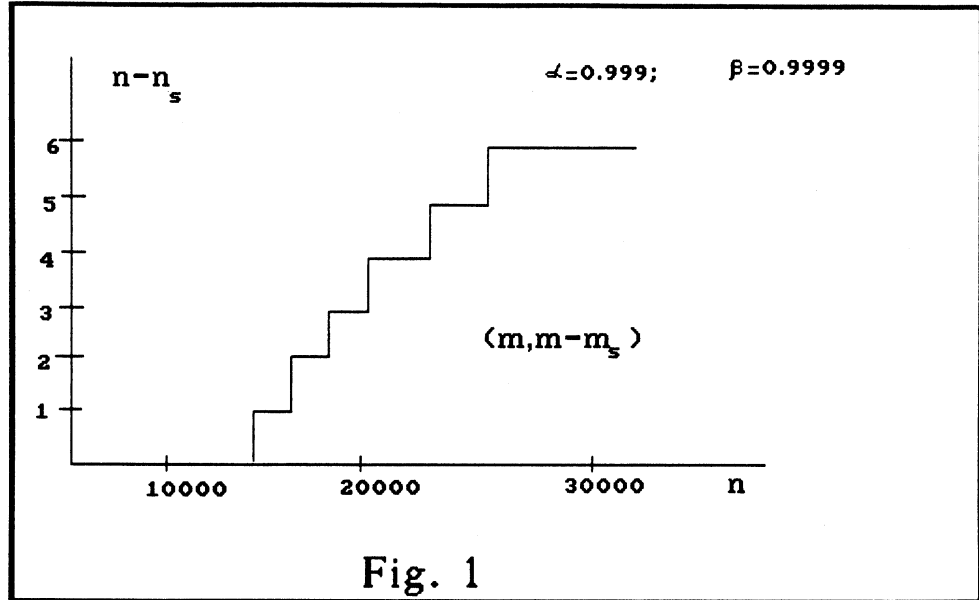
to α, and this statement is true with confidence probability β. It can be seen on the plot that if n_s is equal to n (all of the trials are successful), then fifteen thousand successful trials must be performed in a row. If there is one unsuccessful trial, then n increases considerable to hundreds of thousands.

Why does the number of required trials n increase considerably when the number of successful trials n_s grows? Let us consider a simple example. Suppose random variable $\Phi(u,w)$ has a normal distribution, then the density plot $p(\varphi)$ looks as presented on Fig. 2. From the plot you can see that the quantile Φ_α for distribution $\Phi(u,w)$ is the right bound of figure which has an area equal to α. Suppose that the area is calculated numerically, and the error of this calculation is $\varepsilon=0.01$, the necessary probability level $\alpha=0.99$. In this case it can happen that the quantile estimate $\hat{\Phi}_\alpha$ equals to infinity. This situation happens due to so called asymptotic unstability of quantile Φ_α in parameter α.

If it is necessary to integrate differentiable equations to obtain the result in each trial, as is the case in the estimation

Fig. 2
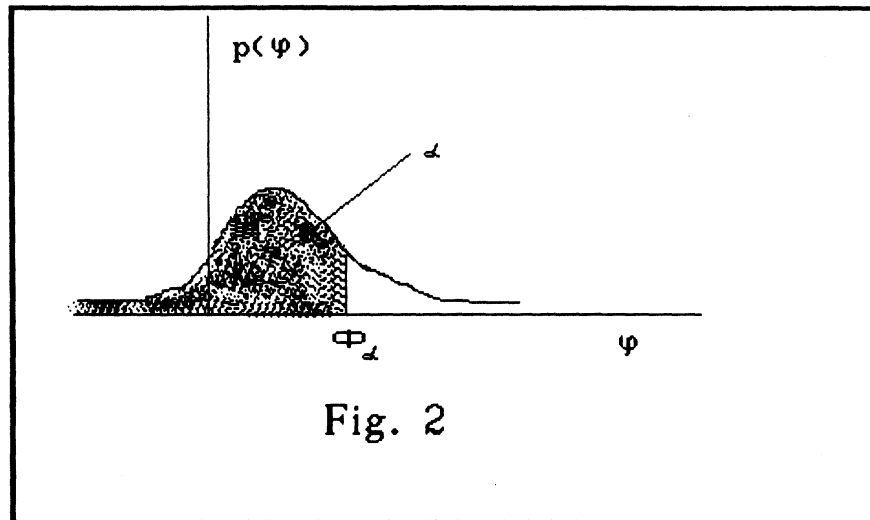
of the probability of successful landings, then the number of calculations will be extremely high.

The other disadvantage of this procedure is the difficulty of setting the confidence probability level β. For example, in an actual problem one may evaluate a probability $\alpha = 0.999$ with a confidence probability β equal to 0.9. However, the number of necessary trials n_{min} considerably depends on the confidence probability level β. In addition, the Muavre–LaPlace theorem states that frequency only has a Normal distribution when n tends to infinity, but for finite n this is not true, and we have to define a new confidence probability (which is tertiary with respect to α).

It is possible to use another method to combine the statistical simulation method with the search optimization method. These methods must be used in parallel, and not consecutively. The group of methods that are constructed in this way is called the stochastic approximation methods. It is not necessary to perform many trials series of the computer simulation for each step of the optimal strategy search. Only one

series is necessary because it is combined with the strategy search. This is the main advantage of this group of methods. The rate of convergence for these algorithms is usually higher than the rate of convergence for optimization methods with the statistical simulation. The main difficulty in designing such a method is in determining the convergency condition. One algorithm from this group, the stochastic quasigradient algorithm, will be discussed in the third lecture.

REFERENCES

1. Kibzun, A.I. and V.V.Malyshev Optimization probabilistic problems. — Izv. AN SSSR. Teknicheskaya Kibernetika, N 6, 1989. (in Russian).
2. Yudin, D.B. Mathematical methods of control under conditions of incomplete information. — Moscow, Sov.Radio, 1974. (in Russian).

THE UNIVERSITY OF MICHIGAN

Andrei I. Kibzun, Professor, *Moscow Aviation Institute, USSR*

Lecture # 2

Indirect Methods for Probabilistic Problem Solving

In the previous lecture, I discussed several direct methods which are being used for probabilistic problem solving. In addition to these methods, I will now consider four indirect methods for probabilistic problem solving. I will apply them to quantile optimization problem solving, though they can also be used for solving other probabilistic problems.

1. The Mean–Square Method

A lot of results have been obtained in the theory of stochastic optimization which are based on the mean–square optimality criteria. Therefore, it is only too natural to use these results in order to solve probabilistic problems. The Mean–square method is based on the Chebyshev's inequality. To make the method clear, let's assume for the simplicity that there is no the constraint $Q(u,\omega) \leq 0$, and a functional $\Phi(u,\omega)$ has the form $\Phi(u,\omega) \overset{\Delta}{=} |\Phi_0(u,\omega) - \lambda_0(u)|$. Consider a mean–square criterion $\Phi_M(u) \overset{\Delta}{=} M\left[\Phi_0(u,\omega) - \lambda_0(u)\right]^2$ instead of a quantile functional. According to the Chebyshev's inequality one has

$$\mathbb{P}\left\{\omega: \ |\Phi_0(u,\omega)-\lambda_0(u)| \geq \varphi \right\} \leq \Phi_M(u)/\varphi^2,$$

if $\lambda_0(u) \triangleq M\left[\Phi(u,\omega)\right]$. As the following inequality is true,

$$\mathbb{P}_\varphi(u)= 1 - \mathbb{P}\left\{\omega: \ |\Phi_0(u,\omega)-\lambda_0(u)| \geq \varphi \right\} \geq 1 - \Phi_M(u)/\varphi^2,$$

the probability functional $\mathbb{P}_\varphi(u)$ is indirectly maximized when the mean–square criterion $\Phi_\alpha(u)$ is minimized in $u \in U$. The quantile functional is estimated similarly.

Note that a mean–square strategy differs greatly from strategies u_α and u_φ, because the Chebyshev's inequality gives a rather rough estimate, especially when the values of α and $\mathbb{P}_\alpha(u)$ approach the unity. Note also that a mean–square strategy u_M does not depend on parameters φ and α, which are present in the probabilistic problems. This fact stresses its roughness. However, in some cases the mean–square strategy is a good approximation and sometimes coincides with u_α and u_φ. This is true when u and φ are vectors, the density $p(x)$ of a random vector ω has a symmetric form $p(x) \triangleq f\left[\|x-m_\omega\|_R\right]$, the objective function may be expressed as follows: $\Phi(u,\omega) \triangleq g\left[\|A\omega + Bu\|_W\right]$. functions $f(.)$, $g(.)$ are continuous and strictly monotonous.

2. The Deterministic Equivalent Method

The deterministic equivalent method was historically the first proposed indirect method. According to this method, a probabilistic constraint $P_\alpha(u) \geq \alpha$, which is an inequality, is replaced by the equality constraint

$$P\left\{\omega: \ \Phi(u,\omega)-\varphi \leq g_1(u), \ Q(u,\omega) \leq g_2(u)\right\} = \alpha,$$

and then the vectors $g_1(u) \leq 0$ and $g_2(u) \leq 0$ are searched for. This method is discussed thoroughly in many references, and thus I will not discuss it in my lecture.

3. The Minimax Method

The main idea of the minimax method is the following: let's fix an arbitrary confidence set E from the elementary event space Ω. The confidence set E is the set with probability measure greater than or equal to α, i.e. $P(E) \geq \alpha$. Such sets form the whole family of sets

$$\mathbb{E}_\alpha \triangleq \left\{ E:\ P(E) \geq \alpha \right\}$$

Now I will assume that an uncontrollable factor ω is not random but an uncertain factor with the uncertainty set E., i.e. $\omega \in E$ (ω belongs to E). The control strategy $u \in U$ (from set U) will be considered fixed. Let's examine the maximum functional:

$$\Psi(E,u) = \sup_{\omega \in E} \Phi(u,\omega)$$

and try to solve the minimax problem (assuming that $Q(u,\omega) \equiv 0$):

$$u^E = \arg \min_{u \in U} \Psi(E,u), \qquad \varphi^E \triangleq \inf_{u \in U} \Psi(E,u),$$

where u^E and φ^E are the minimax control strategy and the minimax value of functional $\Phi(u,\omega)$. Is this minimax problem connected with original quantile problem? For simplicity I shall assume that $Q(u,\omega) \equiv 0$ and that solutions to these problems exist. The answer is "yes". This connection is established in the following theorem.

THEOREM 1. The pair $(u^E,\ \varphi^E)$ is a guaranteed solution for the quantile problem

$$\varphi_\alpha \triangleq \Phi_\alpha(u_\alpha) \leq \Phi_\alpha(u^E) \leq \Psi(E,u^E) \triangleq \varphi^E, \quad \forall\, E \in \mathbb{E}_\alpha,$$

where
$$\varphi_\alpha = \inf_{u \in U} \Phi_\alpha(u), \qquad \alpha \in (0,1),$$

$$u_\alpha = \arg\min_{u \in U} \Phi_\alpha(u).$$

Proof. Let $E \in \mathbb{E}_\alpha$, $\alpha \in (0,1)$ and the pairs $(u_\alpha, \varphi_\alpha)$, (u^E, φ^E) (the solutions in quantile and minimax problems) exist.

Consider a set S_{φ^E, u^E} :

$$S_{\varphi^E, u^E} \triangleq \left\{ \omega\colon \Phi(u^E,\omega) \leq \varphi^E \right\} = \left\{ \omega\colon \Phi(u^E,\omega) \leq \sup_{\omega \in E} \Phi(u^E,\omega) \right\},$$

because $\varphi^E = \Psi(E,u^E) = \sup_{\omega \in E} \Phi(u^E,\omega)$. Thus the set S_{φ^E, u^E} contains the set E. Thus, the probability

$$P_{\varphi^E}(u^E) \triangleq \mathbb{P}(S_{\varphi^E, u^E}) \geq P(E) \geq \alpha,$$

and so $P_{\varphi^E}(u^E) \geq \alpha$. Remember, that the definition of quantile is:

$$\Phi_\alpha(u^E) \triangleq \min\{\,\varphi\colon P_\varphi(u^E) \geq \alpha\,\}$$

The inequality $P_{\varphi^E}(u^E) \geq \alpha$, therefore implies that $\Phi_\alpha(u^E) \leq \varphi^E$. The strategy u^E is not necessarily optimal for the minimization of quantile functional $\Phi_\alpha(u)$, and so $\Phi_\alpha(u_\alpha) \leq \Phi_\alpha(u^E)$.

After combining all the above inequalities we will get the necessary statement of the theorem. ∎

Thus, the minimax methods give the possibility to find solution for the quantile optimization problem. The upper estimation $\varphi^E \geq \varphi_\alpha$ is true for any confidence set E from the family of sets $\mathbb{E}_\alpha$ ($E \in \mathbb{E}_\alpha$). But, if we use not the best E from $\mathbb{E}_\alpha$, then, first of all, the upper estimate obtained will be too rough, and secondly, it can be more difficult to obtain the

solution of the minimax problem than the solution of the original quantile problem.

4. The Generalized Minimax Method.

The possibility to obtain the quantile estimation through solving a minimax problem is demonstrated by the generalized minimax method

$$\tilde{\varphi}_\alpha = \inf_{E \in \mathbb{E}_\alpha} \inf_{u \in U} \sup_{\omega \in E} \Phi(u,\omega).$$

The main idea is that we can choose such a confidence set $E_\alpha \in \mathbb{E}_\alpha$, that the quantile estimation will be equal to the real value of quantile $\varphi_\alpha = \varphi^{E_\alpha}$. This confidence set E_α is called an optimal one. This problem is interpreted in terms of two–person–zero–sum–game theory in the following way. The second player ($\omega \in E$, which is nature) seeks to maximize the damage caused to the first player knowing the first player's strategy and his own opportunities, which are defined by the set E. The first player, without knowing the second player strategy, seeks to minimize his losses by choosing his own strategy and worsening the second player's opportunities. This is achieved by choosing the best confidence set E, i.e. the second player's opportunities. This problem is the problem with dependent variables, and it is known that this type of problem is unstable in criterion. That is why the initial stochastic problem can be solved only with a high consumption of time necessary for many numerical calculations. The most difficult step in the obtained problem is to find the optimal confidence set E . We will discuss the generalized minimax approach before discussing the ways of

choosing the set E_α.

At first, I shall give some definitions. I will say that the minimax problem is equivalent to the quantile problem if the strategy u_α exists and the following three conditions are satisfied:

1. $\Psi(\tilde{E}_\alpha, \tilde{u}_\alpha) = \Phi_\alpha(u_\alpha)$;

2. For any $u_\alpha \; \exists \; E_\alpha$ such that $\Psi(E_\alpha, u_\alpha) = \Phi_\alpha(u_\alpha)$;

3. For any pair (E_α, u_α) it is true that $\Psi(\tilde{E}_\alpha, \tilde{u}_\alpha) = \varphi_\alpha(\tilde{u}_\alpha)$,

where u_α is an optimal strategy in the quantile problem, $(\tilde{E}_\alpha, \tilde{u}_\alpha)$ is an optimal pair in the generalized minimax problem.

THEOREM 2. The quantile problem is equivalent to the generalized minimax problem.

Proof. Let the strategy u_α exist, and so $\varphi_\alpha = \Phi_\alpha(u_\alpha)$. The definition of quantile is

$$\Phi_\alpha(u_\alpha) \triangleq \min \{ \varphi : P_\varphi(u_\alpha) \geq \alpha \}$$

Consider the probability functional of optimal strategy u_α:

$$P_{\varphi_\alpha}(u_\alpha) \triangleq \mathbb{P} \{\omega : \Phi(u_\alpha, \omega) \leq \varphi_\alpha\}$$

Now let us designate $E_\alpha = \{\omega : \Phi(u_\alpha, \omega) \leq \varphi_\alpha\}$ and consider the maximum functional

$$\Psi(\tilde{E}_\alpha, u_\alpha) = \sup_{\omega \in \tilde{E}_\alpha} \Phi(u_\alpha, \omega) = \varphi_\alpha \triangleq \Phi_\alpha(u_\alpha)$$

Thus, the second condition of problem equivalency is satisfied.

Now consider the pair $(\tilde{E}_\alpha, \tilde{u}_\alpha)$ which is optimal for the generalized minimax problem. I will use the minimax estimates obtained in the previous part of this lecture.

$$\varphi_\alpha \triangleq \Phi_\alpha (u_\alpha) \le \Phi_\alpha (u^{\tilde{E}_\alpha}) \le \Psi(\tilde{E}_\alpha, u^{\tilde{E}_\alpha}) \triangleq \tilde{\varphi}_\alpha$$

As the second property is already proved, I can use it to find out that $\varphi_\alpha = \Psi(E_\alpha, u_\alpha)$. The pair $(\tilde{E}_\alpha, \tilde{u}_\alpha)$, as you remember, is optimal in the generalized minimax problem. That is why the last inequality is consistent only if it is a strict equality. And this means that the conditions 1 and 3 of problem equivalency are satisfied and, thus the quantile problem is equivalent to the generalized minimax problem. ∎

The equivalency proved above is a base for the confidence approach.

At the end of this part I will give the formula for building the confidence absorbed set in the case when the function $\Phi(u, \omega, z)$ depends not only on u and on ω, but on z.

$$\mathbb{Z}_{\varphi, \alpha} = \bigcup_{E \in \mathbb{E}_\alpha} \bigcup_{u \in U} \bigcap_{\omega \in E} \mathbb{Z}_\varphi(u, \omega),$$

where $\mathbb{Z}_\varphi(u, \omega) \triangleq \left\{ z: \Phi(u, \omega, z) \le \varphi \right\}$.

This formula is equivalent to the generalized minimax problem in which the search for an upper bound is substituted by the intersection of sets and minimization is substituted by the union of sets. This formula is deduced in the same way as the generalized minimax problem.

5. Bilateral estimates for quantile functional .

Now I shall consider the additional opportunities which the confidence approach gives us when it is used forrp quantile problem solving. First of all, using this approach it is possible to obtain bilateral estimates for a quantile functional

Subject to the different conditions, it is possible to obtain the estimations, which will be given below in Lemmas 1, 2, 3, and 4.

Lemma 1. If the function $\Phi(u,\omega)$ is continuous (both in u and in ω) and quasi-convex in $\omega \in \mathbb{R}^n$ for all $u \in U$, probability measure $\mathbb{P}$ and Lebesque measure are mutually absolutely continuous, and the confidence set $E \in \mathbb{E}_\alpha$ is compact, then

$$\Psi_*(E,u) \triangleq \inf_{\omega \in \partial E} \Phi(u,\omega) \leq \Phi_\alpha(u) \leq \Psi^*(E,u) \triangleq \sup_{\omega \in \partial E} \Phi(u,\omega),$$

where ∂E is the boundary of set E, and the quasi-convex function $\Phi(u,\omega)$ is the function for which the Lebesque set $S_{\varphi,u} \triangleq \left\{ \omega\colon \Phi(u,\omega) \leq \varphi \right\}$ is convex for all φ.

The estimates presented are the base for the numeric procedure of the confidence set transformation. These transformations of the confidence set E_n are aimed to make it as close to optimal set as possible. When using this procedure, in correspondence with the generalized minimax method it is possible to obtain that the estimates

$$\Psi_*(E_n,u) \xrightarrow[n \to \infty]{} \Phi_\alpha(u), \qquad \Psi^*(E_n,u) \xrightarrow[n \to \infty]{} \Phi_\alpha(u),$$

i.e. the upper and lower estimates are coming close while the probability measure of set E remains the same at the every n-step of procedure.

Now let's define the α-order kernel of measure $\mathbb{P}$. The α-order kernel of measure $\mathbb{P}$ is the set $S_\alpha \triangleq \bigcap_{\|a\|=1} E_{a,\alpha}$, which is the intersection of all possible confidence half-spaces

$$E_{a,\alpha} \triangleq \left\{ \omega\colon a^T\omega \leq b(a) \right\}, \quad \|a\| = 1,$$

where parameter $b(a)$ is chosen to provide $P(E_{a,\alpha})=\alpha$. Note that the set S_α belongs to all convex confidence sets E with

probability measure α.

Lemma 2. Let the objective function $\Phi(u,\omega)$ be continuous and quasi–convex in $\omega \in \mathbb{R}^n$ for all $u \in U$, the Probability measure $\mathbb{P}$ and Lebesgue measure are mutually absolutely continuous, $\alpha \in (0.5;1)$, and optimal strategy u_α exist in the quantile problem. There is the following quantile estimation for the kernel S_α:

$$\Psi^*(S_\alpha, u^{S_\alpha}) \leq \Psi^*(S_\alpha, u_\alpha) \triangleq \sup_{\omega \in \partial S_\alpha} \Phi(u_\alpha, \omega) \leq \varphi_\alpha.$$

Thus, the value $\Psi^*(S_\alpha, u^{S_\alpha})$ is the lower estimation for the minimal quantile φ_α.

Lemma 3. Let the objective function $\Phi(u,\omega)$ be continuous and quasi–convex in $\omega \in \mathbb{R}^n$ for all $u \in U$, and confidence set E be a convex polyhedron. Then

$$\Phi_\alpha(u) \leq \max_{\omega \in \mathbb{J}(E)} \Phi(u,\omega) = \Psi^*(E,u),$$

where $\mathbb{J}(E)$ is the set of all vertexes of the polyhedron E.

In this case one has to look over all the vertexes of polyhedron E in order to find out the value of the maximum function $\Psi(E,u)$. And if the function $\Phi(u,\omega)$, in addition to the conditions mentioned above, is convex and piece wise linear in **u**, then maximum function $\Psi(E,u)$ will also be piece–wise linear in **u**. If, besides that, the set of feasible strategies **U** is a convex polyhedron, then the minimax problem can be reduced to a linear programming problem. In this case, one can easily obtain the upper estimation of a quantile $\varphi_\alpha \leq \varphi^E$, while the solution of a quantile problem is very difficult.

Now let's consider the deviation of the set E_r from the set

E, i.e. $\quad d(E_r, E) \overset{\Delta}{=} \underset{\omega_1 \in E_r \backslash E}{\sup} \ \underset{\omega_2 \in E}{\inf} \ \|\omega_1 - \omega_2\|.$

Lemma 4. Let the objective function $\Phi(u, \omega)$ be continuous and quasi-convex in $\omega \in \mathbb{R}^n$ for all $u \in U$ and also satisfy the Lipschitz type condition, i.e.

$$|\Phi(u, \omega_1) - \Phi(u, \omega_2)| \leq k \ \| \omega_1 - \omega_2 \|^\gamma$$

for any ω_1, ω_2 such that $\|\omega_1\| > R$, $\|\omega_2\| > R$, $\|\omega_1 - \omega_2\| < \delta$, where $k, \gamma, R, \delta > 0$. Let the probability measure is a Gaussian measure corresponding to the standard normal distribution with zero expectation and the unit matrix of covariances, confidence set E be a sphere E_r with radius r, $\mathbb{P}(E_r) = \alpha$.

Then under the assumptions mentioned above the following estimates are valid:

$$\varphi^{E_r} = \Psi(E_r, u^{E_r}) \geq \varphi_\alpha \geq \Psi(E_r, u^{E_r}) - kd^\gamma(E_r, E),$$

and moreover,

$$\Psi(E_r, u^{E_r}) - \varphi_\alpha \rightarrow 0 \quad \text{as } \alpha \rightarrow 1.$$

This lemma states that when α is increasing, the lower and the upper quantile estimates tend to each other. It means that when α is close enough to 1, the minimax estimate φ^{E_r} becomes very close to quantile φ_α and is still an upper estimation. It is interesting that the minimax estimate is more precise when probability level α is higher, and the complexity of the solution does not depend on α. It is not correct when one uses the statistical simulation method or the mean-square method, because their errors increase as α increases.

The usefulness of the estimates obtained will be demonstrated in the third lecture when designing the 2nd and the 3rd algorithms for quantile function minimization.

Remark. The proofs of lemmas presented above can be found in the book: *V.V.Malyshev, A.I.Kibzun.*

REFERENCES

1. Yudin, D.B. Mathematical methods of control under conditions of incomplete information. – Moscow, Sov.Radio, 1974. (in Russian).

2. Malyshev, V.V. and A.I.Kibzun Analysis and Synthesis of High Precision Aircraft Control. – Moscow, Mashinostroenie, 1987. (in Russian)

THE UNIVERSITY OF MICHIGAN

Andrei I. Kibzun, Professor, *Moscow Aviation Institute, USSR*

Lecture #3

Three Algorithms for Quantile Function Minimization

The methods for solving probabilistic problems which were considered in the previous two lectures give us the opportunity to begin the design of algorithms for solving them. Three algorithms will be discussed. The first one is based on the direct method; the second and the third ones are based on the confidence method. Two examples which show the particularities of these algorithms will be given.

1. The Direct Algorithm.

At first, I will consider an algorithm based on the direct method. let the following conditions hold:

A) The objective function $\Phi(u,\omega)$ depends monotonically on a quasi–convex function $t(u)$ for any ω and is continuous in $u \in U$ and ω;

B) The random vector ω has the Standard Normal distribution $N(0,I)$;

C) The set U of feasible strategies is compact in $\mathbb{R}^m$. Now I impose no constraints on the dependence of the objective function $\Phi(u,\omega)$ on ω except continuity. These requirements will be set below after introducing some new constraints. Under the given assumption, the quantile of the objective function will be continuous and quasi–convex function in u.

Suppose for any u it is possible to obtain the sample of values $\{\Phi(u,\omega_i)\}^t_{i=1}$ of function $\Phi(u,\omega)$, where $\{\omega_i\}^t_{i=1}$ are independent realizations of the random vector ω. Let's designate them as $\{\Phi_i(u)\}^t_{i=1}$ the **ordered series** of the sample, i.e. the

values of the objective function $\Phi(u, \omega_i)$ arranged in increasing order. Under the assumptions made above, the random variable $\Phi(u, \omega)$ has a continuous cumulative distribution for any $u \in U$. According to the Central Limit Theorem, the middle terms of the **ordered series** can be represented as t tends to infinity as:

$$\Phi_{[\alpha t]+1}(u) \approx \Phi_\alpha(u) + \frac{\nu}{p_u(\Phi_\alpha(u))} \sqrt{\frac{\alpha(1-\alpha)}{t}}$$

where $[\alpha \cdot t]$ is a whole part of the value $\alpha \cdot t$, $\alpha \in (0,1)$; ν is a random variable with Standard Normal probability density; $p_u(\varphi)$ is a probability density function for the random function $\Phi(u, \omega)$; $\Phi_\alpha(u)$ is an α-level quantile of it's distribution. From the representation above it follows, that the statistic

$$\widetilde{\Phi}_t(u) \triangleq \Phi_{[\alpha t]+1}(u)$$

are asymptotically unbiased and $\sqrt{t}$ – consistent estimate of the quantile $\Phi_\alpha(u)$ for any $u \in U$. Moreover, the bias of the estimate

$$z_t(u) \triangleq |\Phi_\alpha(u) - M[\widetilde{\Phi}_t(u)]|$$

has an order $o(1/t)$ uniformly on any compact set U, i.e. there exists a constant $c < \infty$ which doesn't depend on t such that

$$z_t(u) \leq \frac{c}{t}$$

for any $u \in U$. The estimate $\widetilde{\Phi}_t(u)$ turns out to be suitable when working with high order quantiles ($\alpha > 0.99$). That is why the necessary sample size cannot be less than $T_\alpha \triangleq [\frac{1}{1-\alpha}] + 1$. Otherwise the estimate is insensitive to variations of α. Moreover, since the asymptotic representation is valid only for middle terms, it is necessary for the extreme terms to be in mean greater than the value of the minimal quantile. And it means that the sample size has to be at least several times more than T_α. In addition, every time one is going to get the quantile estimate, he has to build the empirical cumulative distribution function for random variable $\Phi(u, \omega)$.

Let's consider the estimate of the other type which is based on the use of asymptotic properties of extremal order statistics, i.e. the largest terms of the ordered series of sample $\{\Phi_j(u)\}^t_{j=t-k}$, $k \ll t$. Just as the distribution of the middle terms of the ordered series can be found (and is Normal), the limiting external order statistics distribution can also be found in some cases (but it is not Normal). In particular, for the most well known distributions of the random variable Φ (for example, the

Normal distribution, Exponential distribution, Gamma–distribution etc.) the asymptotical behavior of extremal order statistics obeys the law:

$$F_{\Phi}(\varphi) = \exp(-\exp(-\varphi)), \varphi \in \mathbb{R}^1$$

In this case the random variable Φ is a random variable of the **exponential type**. Totally there exists three types of limiting distributions for the extremal statistics, but the exponential type is the most common one.

I will assume later that in addition to condition A the condition A' also holds:

A') The nature of dependence $\Phi(u,\omega)$ on ω is such that the random variable $\Phi(u,\omega)$ is a random variable of the exponential type for any $u \in U$.

The knowledge of these laws gives us the opportunity to calculate the estimate of the quantile function using the sample with the size less than or equal to T_{α} in a relatively simple manner. Let's consider the estimate

$$\hat{\Phi}_t(u) \triangleq \Phi_t(u) - \mu[\Phi_t(u) - \Phi_{t-1}(u)]$$

where $\Phi_t(u)$ and $\Phi_{t-1}(u)$ are the last two terms of the ordered series of the sample of the size $t > 1$, $\mu = 0.5772$ is the Euler constant.

Lemma 1. If a random variable $\Phi(u,\omega)$ is of the exponential type for any $u \in U$, then the error of the quantile estimate

$$z_t(u) \triangleq |\, M\, [\hat{\Phi}_t(u)] - \Phi_{\alpha}(u)\, | = o(1/t)$$

is of order $o(1/t)$ uniformly in $u \in U$. Here $t \triangleq [\frac{1}{1-\alpha}] + 1$. Thus, the statistic $\Phi_t(u)$ calculated using a sample of the size t can be used as a quantile estimate, and moreover its bias $z_t(u)$ will be less as the quantile order α tends to one. It should be noted that the obtained estimate $\hat{\Phi}_t(u)$ is considerably simpler than the previous $\tilde{\Phi}_t(u)$ since when taking the sample it is necessary to remember only the two worst realizations to calculate the estimate $\hat{\Phi}_t(u)$. On the other hand, when calculating the estimate $\tilde{\Phi}_t(u)$, it is necessary to remember them and put the whole sample in order. This is a difficult problem when the sample size is considerable. The simplicity of the mentioned above estimate gives us the opportunity to use it in stochastic recurrent procedures in which it is necessary to estimate the quantile function repeatedly for several values of u.

The random vector ξ_t will be called the stochastic quasigradient of the quantile function

$$\xi_t(u,a) \equiv \frac{1}{2a} \sum_{i=1}^{m} [\hat{\Phi}_t(\tilde{u}_1,...,\tilde{u}_i+a,...,\tilde{u}_n) - \hat{\Phi}_t(\tilde{u}_1,...,\tilde{u}_i-a,...,\tilde{u}_m)]$$

where $\tilde{u}_i$ are independent random variables uniformly distributed upon segments $[u_i-a, u_i+a]$, e_i – are unit co-ordinate vectors, $i=\overline{1,m}$, a is a smoothing constant. Consider the stochastic quasigradient algorithm for quantile function minimization.

$$u^{k+1} = \Pi_U(u^k - \rho_k \xi_{t_k}(u^k, a_k)), u^0 \in U$$

where Π_U is an operator which projects a vector from space $\mathbb{R}^m$ on its subset U.

Theorem 1. Let all the conditions mentioned above A, A', B, C hold. Moreover, let the parameters of algorithm obey with the following conditions:

$$\sum_{k=1}^{\infty} \rho_k = \infty \; , \; \sum_{k=1}^{\infty} (\rho_k)^2 < \infty \; , \; \sum_{k=1}^{\infty} (\rho_k)/(a_k t_k) < \infty$$

$$a_k \to 0 \; , \; t_k \to \infty \; , \; \frac{k}{a_k t_k} \to 0 \; , \; \rho_k > 0 \; , \; a_k > 0$$

Then all of the limit points of sequence of the stochastic quasigradient algorithm vectors u^k, with probability 1 belong to a set of extremal points of the quantile function.

The algorithm described above belongs to a class of stochastic approximation algorithms that are similar to the algorithm of the projective gradient which is used for the minimization of deterministic functions. In the stochastic quasigradient algorithm the quasigradient $\xi_t(u,a)$ plays the role of a gradient. The quasigradient is in mean turned to the direction of the quantile function gradient (if this gradient exists). The parameter ρ_k plays the part of a step length. The setup parameters of the algorithm ρ_k, a_k, t_k can be chosen by using the considerations equivalent to the considerations used during the derivation of the stochastic approximation algorithm The step length ρ_k has to be long enough to provide the opportunity to reach any point in $\mathbb{R}^m$ from the initial point $u^1 \in \mathbb{R}^m$ (the condition $\sum_{k=1}^{\infty} \rho_k = \infty$) but on the other hand the length of step has to tend to 0 (the condition $\sum_{k=1}^{\infty} \rho_k^2 < \infty$). The sample size

has to increase to infinity since the biased quantile estimate $\hat{\Phi}_k(u)$ is used and the bias of this estimate tends to 0 as the sample size increases (the condition $t_k \to \infty$). The length a_k of test step that is used for calculating the quasigradient $\xi_{t_k}(u_k, a_k)$ has to tend to 0 (the condition $a_k \to 0$) but slower than the sample size increases (the condition $k/(a_k t_k) \to \infty$). Moreover, the length of current test step must be more than the algorithm step length ρ_k (the condition $\sum_{=1}^{\infty}(\rho_k)/(a_k t_k) < \infty$).

The given conditions hold, for example, when $\rho_k = \rho_0 k^{-1}, t_k = t_0 k^{2/3}$, and $a_k = a_0 k^{-1/3}$ where ρ_0 is the length of the initial step, t_0 is the initial sample size, and a_0 is the initial test step length. It follows that the sample size increases quickly enough. It's essential for the effectiveness of the algorithm to choose good enough initial parameters of the algorithm u_0, ρ_0, t_0, a_0. The successful choice of these parameters gives us the opportunity to obtain a good estimate for the quantile problem quickly. The method for choosing these parameters depends considerably on the problem statement and so it is impossible to give universal recommendations.

At the end of this section I would like to make an important remark concerning the connection between the quantile optimization problems and minimax ones. Suppose the following minimax problem is under the consideration:

$$u_0 = \arg \min_{u \in U} \max_{\omega \in E} \Phi(u, \omega) \overset{\Delta}{=} \arg \min_{u \in U} \psi(E, u)$$

where E is a set of feasible values of uncertain vector $\omega \in \mathbb{R}^n$. A set E is convex and compact. Let the function $\Phi(u, \omega)$ satisfy conditions A and C given at the beginning of this lecture.

The solution to this problem may be very difficult because of the complex dependence of the objective function in ω. Instead of solving this problem it is possible to look for the solution to a quantile problem with the same objective function $\Phi(u, \omega)$ and with a random vector ω distributed Normally with the parameters:

$$m \overset{\Delta}{=} M[\omega], \quad K \overset{\Delta}{=} M[(\omega - m)(\omega - m)^T]$$

The covariance matrix K and the vector of expectations m are chosen to provide the best approximation of the uncertainty set E by the ellipsoid of concentration:

$$E \overset{\Delta}{=} \{ \omega : (\omega - m)^T K (\omega - m) \leq r \}$$

with the probability measure that is equal to α. The value α will be considered as the level of objective function quantile. The algorithm described above is used to minimize it. The quantile estimate $\Phi_\alpha(u_\alpha)$ has to be chosen in some sense to minimize the value of the maximum function $\Psi(E, u_0)$. Thus, a minimax problem can be transformed to a quantile optimization one. But in some cases it is advantageous to act conversely. Let's consider the following algorithm.

2. The Minimax Algorithm.

Let's now consider the minimization algorithm based on the minimax estimates of the quantile function. We understood from the perfunctory analysis of the algorithm in the first section that the convergency rate is not high. The cause of this situation is the high sensitivity of quantile criterion to computational errors. It was noted in lecture #2. Using the minimax estimates of the quantile, we can hope to find the approximate solution of an original problem, and the convergency rate to this approximation will be higher. Moreover, the conditions for using the direct algorithm are rigid enough. In particular, it is required for an objective function $\Phi(u, \omega)$ to depend monotonically on a quasi-convex function $t(u)$. Now I will try to weaken this requirement.

Let the random vector ω still have standard Normal distribution $N(0, I)$. The set of feasible strategies be compact, i.e. conditions B and C formulated above hold. However, the following condition must hold instead of A and A':

A*) The objective function $\Phi(u, \omega)$ is a continuous function both in $u \in U$ and in ω. Let it be quasi-convex in $u \in U$ for any ω and pseudo-convex in ω for any $u \in U$.

The pseudo-convexity of a function $\Phi(u, \omega)$ in ω means that the objective function is differentiable, i.e. its **gradient** $\nabla_\omega \Phi(u, \omega)$ exists in any point ω and the following condition holds:

$$\text{if } \nabla_\omega^T \Phi(u, \omega_1)(\omega_2 - \omega_1) \geq 0 \text{ then } \Phi(u, \omega_1) \geq \Phi(u, \omega_2)$$

Under the given assumptions, the maximum function

$$\psi(u) \overset{\Delta}{=} \max_{\omega \in \partial E} \Phi(u, \omega)$$

where ∂E is a surface of a confidence sphere E, i.e. $P(E)=\alpha$, will be quasi-convex in u. In lecture #2 I showed that the maximum function $\psi(u)$ is an upper estimate for the quantile function $\Phi_\alpha(u)$, i.e. $\psi(u) \geq \Phi_\alpha(u)$. And furthermore, the estimate becomes closer to the real value $\Phi_\alpha(u_\alpha)$ as the value of probability α tends to 1, i.e. $\psi(u) - \Phi_\alpha(u) \to 0$ as $\alpha \to 1$. That leads us to beleive that a maximum function might be a good estimate for quantile function. Now I will describe the stochastic quasigradient algorithm for quantile function minimization which is similar to the one described above.

The following function will be used as a statistical estimate of a quantile function :

$$\hat{\psi}_t(u) \equiv \psi_t(u) + \frac{n-1}{2} [\psi_t(u) - \psi_{t-1}(u)]$$

where n is a dimension of the space $\Omega = \mathbb{R}^n$, $\psi_t(u)$ and $\psi_{t-1}(u)$ are the two last terms of the ordered series of the sample of objective function values $\Phi(u, \omega_j), j = \overline{1, t}$, t is the sample size; ω_j are random vectors uniformly distributed on the surface ∂E of the confidence ball E. Thus the statistics $\hat{\psi}_t(u)$ are estimates of the quantile function $\Phi_\alpha(u_\alpha)$ since $\psi(u) \geq \Phi_\alpha(u)$ and $\hat{\psi}_t(u)$ is an estimate of $\psi(u)$.

Recall that when the quantile function statistical estimate was found, as in the previous section, it was assumed that the random variable $\Phi(u, \omega)$ is a random variable of the exponential type. And it means that the range of values of the objective function$\Phi(u, \omega)$ is the whole real line $\mathbb{R}^1$. In this case, the statistical estimate of the maximum function $\psi(u)$ is considered. If the function $\Phi(u, \omega)$ is continuous in ω, then the maximum function $\psi(u)$ will be bounded above and below on the sphere E. Thus the random variable $\Phi(u, \omega)$ with ω chosen on the surface ∂E can not be a variable of the exponential type. In this case it is possible to show that the random variable $\Phi(u, \omega)$ is a variable of the truncated exponential type, i.e. the limiting distribution of the extremal ordered statistic $\psi_t(u)$ is the following:

$$F_\varphi = \begin{cases} \exp(-(-\varphi^\gamma)), & \varphi \leq 0 \\ 1, & \varphi > 0 \end{cases}$$

where $\gamma \triangleq \frac{m-1}{2}$, m is a dimension of the space $\Omega = \mathbb{R}^m$. For

simplicity, it will be assumed that in the future the level surface $\{\omega: \Phi(u,\omega)=\varphi\}$ has first order tangency with the sphere ∂E. Otherwise, the constant γ is equal to $\frac{m-1}{r+1}$ where r is an order of tangency of these surfaces. This statement can be formulated more precisely in this way :

Lemma 2. If the conditions A*,B,C are fulfilled and the random vector u is uniformly distributed on the sphere ∂E, then the objective function $\Phi(u,\omega)$ is a random variable of the truncated exponential type, and the error of the maximum function estimate

$$z_t(u) \stackrel{\Delta}{=} |\ M\ [\ \hat{\psi}_t(u)-\psi(u)\ ]\ |$$

is of order $o(1/t)$ uniformly in $u \in U$.

Let's now consider the stochastic quasigradient of the maximum function

$$\xi_t(u,a) \stackrel{\Delta}{=} \frac{1}{2a}\sum_{i=1}^{m}[\hat{\psi}_t(\tilde{u}_1,...,u_i+a,...,\tilde{u}_m)-\hat{\psi}_t(\tilde{u}_1,...,u_i-a,...,\tilde{u}_m)]e_i$$

where $\tilde{u}_i$ are independent random variables uniformly distributed on the segments $[u_i-a,u_i+a]$; e_i are unit coordinate vectors in the space $\mathbb{R}^m$; a is a smoothing constant.

The stochastic quasigradient algorithm for the maximum function minimization will be the following:

$$u^{k+1}=\Pi_U(u^k-\rho_k\xi_{t_k}(u^k,a_k)),\ \ u^0 \in U$$

where Π_U is an operator which projects the vector from space $\mathbb{R}^m$ on its subset U

Theorem 2. Let conditions A*,B, and C for the problem statement hold, and the parameters of the algorithms obey the following conditions:

$$\sum_{k=1}^{n}\rho_k = \infty\ \ ,\ \ \sum_{k=1}^{n}\rho_k^2 < \infty\ ,\ \ \sum_{k=1}^{n}\rho_k/(a_kt_k) < \infty\ ,$$

$$a_k \to 0\ ,\ t_k \to \infty\ ,\ a_kt_k \to \infty\ ,\ \ \frac{k}{a_k^2t_k^{4/(m-1)}} \to 0\ ,\ a_k > 0\ ,\ \rho_k > 0$$

Then all of the limit points of the sequence of the stochastic quasigradient algorithm vectors u^k belong with probability 1 to a set of extremal points of a maximum function.

I must remark that the constraints on the parameters are similar to those in the previous section. The differences here are the requirement that both the test step length a_k and the sample size t_k increase simultaneously, i.e. $k/(a_k^2 \cdot t_k^{4/(m-1)}) \to 0$.

The last requirement depends on the dimension n of the random vector space $\mathbb{R}^n$. When the dimension is greater than 3, this requirement results in the sample size increasing at a higher rate than analogous parameter in the quantile function minimization algorithm (when all other parameters are the same).

3. Landing Runway Area Minimization.

Now a comparatively simple example will be considered. This example will be used to illustrate the possibility of reducing the problem with the chance constraint to the quantile optimization problem. Peculiarities of the direct and the minimax algorithms discussed above will also be shown.

Let L_0 be a nominal length of the aircraft landing runway. Let ξ_L and ξ_Z be random longitudinal and cross errors of landing the aircraft caused by wind influence during landing. These errors are nonlinearly connected with the longitudinal (ω_L) and cross (ω_Z) components of the horizontal wind ω ($\omega=\mathrm{col}(\omega_L,\omega_Z)$). For simplicity, it will be assumed that magnitudes of components ω_L and ω_Z are not changed during every landing. If the aircraft is controlled only through the angle of it's roll then the error components at the moment of landing will be

$$\xi_L=a_{11}\omega_L\sigma_L+a_{12}|\omega_Z|\sigma_Z, \quad \xi_Z=a_{22}\omega_Z\sigma_Z, \quad a_{21}=0$$

where a_{11},a_{12},a_{22} are given scalar coefficients, σ_L and σ_Z are the mean square deviations of the longitudinal and the cross wind components, ω_L and ω_Z are independent random variables with standard Normal distributions, i.e. with zero expectations and unit variances. A similar model for describing a reaction under disturbances is used in a "Buran" shuttle spacecraft control system. Now it is possible to state the problem of choosing the positive parameters L_1,L_2,Z of the runway which play the role of "reserve" in case of undershoot, overshoot, and cross miss, respectively. These parameters have to provide that the minimal area of the runway :

$$S(L_1,L_2,Z)=2Z(L_0+L_1+L_2) \to \min_{L_1,L_2,Z}$$

When the constraint guaranteeing a successful landing with a probability not lower than given α holds:

$$\mathbb{P}\{\omega_L,\omega_Z: -L_1 \le \xi_L \le L_2, \ |\xi_Z| \le Z\} \ge \alpha$$

- 9 -

A reduction in the total area of the runway results in a lower cost of construction. This problem is an example of a typical problem with chance constraints. Now it is possible to show how this problem can be reduced to an unconditional quantile optimization problem. First, replace the variables:

$$u_1 \equiv \frac{L_1}{L_2} \ , \quad u_2 \equiv \sqrt{\frac{L_0 + L_1 + L_2}{2Z}} \ , \quad \Phi \equiv \sqrt{2Z(L_0 + L_1 + L_2)}$$

Then,

$$Z(u_2, \Phi) = \frac{\Phi}{2u_2} \ ,$$

$$L_1(u_1, u_2, \Phi) = \frac{(u_2 \Phi - L_0)u_1}{1 + u_1} \ ,$$

$$L_2(u_1, u_2, \Phi) = \frac{u_2 \Phi - L_0}{1 + u_1}$$

Now re-write the original constraints with the new notation:

$$\Phi_1(u_1, u_2, \omega_L, \omega_Z) \equiv -\frac{u_1 + 1}{u_1 u_2}(\sigma_L a_{11}\omega_2 + a_{12}\sigma_Z |\omega_Z|) + \frac{L_0}{u_2} \leq \Phi$$

$$\Phi_2(u_1, u_2, \omega_L, \omega_Z) \equiv \frac{u_1 + 1}{u_2}(\sigma_L a_{11}\omega_2 + a_{12}\sigma_Z |\omega_Z|) + \frac{L_0}{u_2} \leq \Phi$$

$$\Phi_3(u_2, \omega_Z) \equiv 2a_{22}u_2\sigma_Z\omega_Z \leq \Phi$$

$$\Phi_4(u_2, \omega_Z) \equiv -2a_{22}u_2\sigma_Z\omega_Z \leq \Phi$$

and introduce the objective function as the maximal of the four functions mentioned above:

$$\Phi(u_1, u_2, \omega_L, \omega_Z) = \max_{i = \overline{1,4}}\{\Phi_i(u_1, u_2, \omega_L, \omega_Z)\}$$

Now it is possible to reduce the original problem with chance constraint to the unconditional quantile optimization problem

$$\Phi_\alpha(u_1, u_2) \to \min_{u_1, u_2}$$

where $\Phi_\alpha(u_1, u_2)$ is the α-order quantile of the random variable $\Phi(u_1, u_2, \omega_L, \omega_Z)$ which can be found from the inequality:

$$\mathbb{P}\{\omega : \Phi(u_1, u_2, \omega_L, \omega_Z) \leq \Phi_\alpha\} \geq \alpha$$

The square of a quantile function is equal to the original criterion

$$\Phi_\alpha^2(u_1, u_2) = S(L_1, L_2, Z)$$

Since the area of the runway is a positive value and the function $\Phi_\alpha = \sqrt{S}$ is strictly monotonic, and thus, minimization of S is equivalent to minimization $\Phi_\alpha = \sqrt{S}$.

This quantile optimization problem was solved by using both of the algorithms described above. I want to remark here that the minimax problem could also be solved, in this case, without using the stochastic quasigradient method. In fact, thanks to the piece-wise linear structure of an objective function, its maximum on the confidence circle ∂E can be found analytically. It is necessary to consider the points of tangency between the confidence circle and the level curves of an objective function. After the maximum function $\psi(u)$ is found, one can use convex programming methods for minimizing it. It is easy to check that, in this case, the objective function is convex in u_1 and in u_2. This means that the maximum function $\psi(u)$ will be convex also.

This example is used here for illustrative purposes. The aim is to analyze the peculiarities of the stochastic quasigradient algorithms. An example of this is the comparison of time consumption of the algorithms.

The problem has been solved by IBM-PC/AT-286 when:

$a_{11}= 20$ sec.,$a_{12}=-20$ sec.,$a_{22}= 3$ sec.,$L_0= 1500$ m,$\sigma_L=\sigma_Z= 5$ m/sec. The table written below indicates the results:

α	algorithm	$S[km^2]$	time of counting [sec.]	u_1	u_2
0.99	direct	0.17	85	1.88	26.3
	minimax	0.202	20	1.48	24.4
0.999	direct	0.236	210	1.96	23.3
	minimax	0.265	30	1.48	21.3
0.9999	direct	0.304	720	2.03	23.3
	minimax	0.324	55	1.48	21.3
0.99999	direct	0.366	3600	2.12	19.2
	minimax	0.38	70	1.48	16.3
0.999999	direct	0.422	25200	2.2	18.1
	minimax	0.432	90	1.48	17.4

The table shows that as α increases the upper estimate of the quantile of runway area becomes closer to the true value of quantile. However the convergency rate is not high. Use of the direct algorithm provides an opportunity to obtain the optimal solution through the direct minimization of a quantile function The estimate found by the minimax algorithm is greater than the true value of the runway area quantile. But the minimax problem

solution is weakly dependent on the probability level α as soon as this level is essential for time consumption on direct algorithm. For example, to make calculations using the direct method when α is equal to 0.999999, it was necessary to use a more powerful computer than the IBM-PC/AT-286.

In practice, it is more useful to combine the two algorithms. For example, one can quickly obtain a good initial approximate solution using the minimax algorithm and then improve the solution using the direct algorithm.

4. The Linear-Quadratic Algorithm.

Consider the following algorithm which is an advanced minimax algorithm. Suppose that the objective function:

$$\Phi(u,\omega) = \max_{1 \le i \le n} \Phi_i(u,\omega)$$

where $\Phi_i(u,\omega) \triangleq (u^T a_i + b_i^T)\omega + (u+d_i)^T C_i(u+d_i) + q_i$, $i=\overline{1,n}$ is piece-wise linear in the random vector ω and piece-wise quadratic in the strategy u. As usual, I will assume that the random vector ω has a standard Normal distribution $N(0,I)$. To estimate the quantile $\Phi_\alpha(u_\alpha)$ of the random variable $\Phi(u,\omega)$, I will use the minimax estimate

$$\Phi_\alpha(u_\alpha) \le \inf_{u \in U} \sup_{\omega \in \partial E} \max_{i=\overline{1,n}} \Phi_i(u,\omega)$$

where ∂E is the sphere of the confidence ball E, $\mathbb{P}(E)=\alpha$. Suppose that the set U of feasible strategies u is a convex polyhedron. As soon as the functions $\Phi_i(u,\omega)$ are convex and continuous in u and in ω for any $i=\overline{1,n}$, then the maximum function $\Phi(u,\omega)$ will also be convex and continuous in u and in ω. Thus the upper bound of an objective function $\Phi(u,\omega)$ on the sphere ∂E is attained and, therefore, the supremum can be replaced by the maximum. Moreover, the maximum function

$$\psi(u) \triangleq \max_{\omega \in \partial E} \Phi(u,\omega)$$

will also be continuous and convex in u. Thus the minimum will be attained on a convex polyhedron U. The original minimax problem is reduced in this way to a convex programming problem

$$\psi^* = \min_{u \in U} \psi(u)$$

In this case the maximum function can be written explicitly. The maximum of the function $\Phi(u,\omega)$ on the sphere ∂E is attained at

the points of contact with the hyperplanes when the strategy u is fixed.

The equations

$$(u^T a_i + b_i^T)\omega + (u+d_i)^T C_i (u+d_i) + q_i = \varphi, \quad i=\overline{1,n}$$

describe hyperplanes of equal value of the function $\Phi(u,\omega)=\varphi$. If r is the radius of a sphere E, then it is possible to find n points where the objective function $\Phi(u,\omega)$ reaches it's maximal value

$$\omega^i = \frac{r \ (a_i u + b_i)}{\|a_i u + b_i\|} \quad i=\overline{1,n}$$

Thus

$$\psi(u) = \max_i \ [\|a_i u + b_i\| \cdot r + \|u+d_i\|_{C_i}^2 + q_i]$$

where $\|u+d_i\|_{C_i}^2 \triangleq (u+d_i)^T C_i (u+d_i)$. To solve the convex programming problem of $\psi(u)$ minimization, it is possible to use known methods of the undifferential optimization. Here I will not discuss them in detail. But, for example, one method of solving this problem is the subgradient algorithm

$$u^{k+1} = \Pi_U [u^k - \frac{\psi(u^k)-q}{\|\nabla\psi(u^k)\|} \nabla\psi(u^k)]$$

where $\nabla\psi(u^k)$ is a sub gradient of function $\psi(u)$ in point u^k, $q = \min_{1\le i\le n} q_i$ is a minimal value of the free parameter in this function. It is known that a convergency rate of this algorithm decreases if the function $\psi(u)$ is not differentiable at the minimal point.

To solve this problem, it is also possible to use the generalized simplex method, which is especially effective in this case.

I would like to note that in particular cases the convex programming problem will become simpler. For example, if the functions

$$\Phi_i(u,\omega) \triangleq u_{k_i} a_i^T \omega + c_i^T u + q_i, \quad i=\overline{1,n}$$

where u_{k_i} is a certain k_i-th component of vector u, then the maximum function $\psi(u)$ is piece wise linear:

$$\psi(u) = \max_{1\le i\le n} \ [u_{k_i} \|a_i\| r + c_i^T u + q_i]$$

Thus the programming problem of maximization of this function on a convex polyhedron reduces to a linear programming problem

$$\psi_r = \min_{u \in U} \psi(u)$$

It is possible to use the simplex method to solve this problem. The efficiency of this method is comparatively high and the solution is found after finite number of steps.

I would like to call your attention to the fact that the obtained minimax estimate ψ_r can be considerably excessive with respect to the real quantile $\Phi_\alpha(u_\alpha)$. The error of this estimation can be evaluated in an indirect way by calculating the difference between the upper estimate ψ_r and the lower estimate ψ_s. The latter can be found in the following way. Since the objective function in this case is a convex function and, therefore, it is possible to consider the α-level kernel of Gaussian measure. In this case, it will be a sphere S_α with a radius ρ defined by the equation

$$\mathbb{P}\{\, \omega : \omega_1 \leq \rho \,\} = \alpha$$

Then, according to lemma 3 from the previous lecture

$$\psi_\rho \triangleq \min_{u \in U} \max_{\omega \in \partial S_\alpha} \Phi(u,\omega) \leq \Phi_\alpha(u_\alpha)$$

For calculation of the lower estimate it is possible to use the same algorithm that was used for the calculation of the upper estimate ψ_r, but the radius of the confidence ball will be ρ instead of r.

If it turns out that the difference between the two estimates is close to zero, i.e. $\psi_r - \psi_\rho \approx 0$, then the minimax estimate can be considered satisfactory. Otherwise, it is possible to use the following method to improve this estimate: Let u_ρ and u_r be the strategies corresponding to the upper and the lower quantile estimates. These strategies are found when solving the corresponding minimax problems. One can use a statistical simulation method for calculating the quantiles $\Phi_\alpha(u_\rho)$ and $\Phi_\alpha(u_r)$. Next, one chooses a new sphere with radius $x_1 = (r+\rho)/2$. Solve the minimax problem again for this sphere. The result is a new strategy u_x, and a new quantile $\Phi_\alpha(u_x)$. Then the dichotomy method is used until the minimax value of a quantile is found:

$$\psi_\alpha = \min_{\rho \leq x \leq r} \Phi_\alpha(u_k), \quad R = \arg \min_{\rho \leq x \leq r} \Phi_\alpha(u_x)$$

Let's denote by u_R the minimax strategy found for the sphere S_R with a radius R. This procedure is based on uni-modal dependence of quantile function $\Phi_\alpha(u_x)$ on radius x. The obtained quantile

estimate $\psi_\alpha \overset{\Delta}{=} \Phi_\alpha(u_R)$ is still the upper estimate of the desired value $\Phi_\alpha(u_\alpha)$, but it is closer to $\Phi_\alpha(u_\alpha)$ than the minimax estimate ψ_r. In practical problems, only a few steps are needed for the improvement of the minimax estimate ψ_r.

5. Water–Supply Optimization For a Desert Region.

Consider the model of water–supply for a desert region. Suppose that the region can be supplied with water carried from other regions or with water obtained using solar batteries. The model describing the way that the water reserves change in time can be stated in the following way:

$$z_{i+1} = \min \{z_i, u_{N+1}\} + u_i + u_{N+1}\omega_i - h_i, \quad z_1=0, \quad i=\overline{1,N}$$

where z_{i+1} is the remainder of water at the end of i–th month; u_{N+1} is an area of solar battery; u_{N+2} is a tank volume where the remaining water is kept; u_i, $i=\overline{1,N}$ is the amount of water carried to the region during the i–th month; h_i is the known consumption of water during the i–th month; ω_i is an output of the solar battery during the i–th month. It is supposed that ω_i are random and mutually independent variables, and they are Normally distributed N (m_i, σ_i^2). The requirement is to keep the water reserve grater than zero during all the months and the probability of the event that there is some reserve must be grater than or equal to α:

$$\mathbb{P} \{ w_i: \min_{i=\overline{1,N+1}} z_i \geq 0 \} \geq \alpha.$$

The other requirement is to provide the minimal expenditures,

i.e.
$$\Phi_0(u) = c_0 \sum_{i=1}^{N} u_i + c_1 u_{N+1} + c_2 u_{N+2} \to \min_{u \in U},$$

where $U = \{ u : u_i \geq 0, i=\overline{1,N+2} \}$; $u=\text{col}(u_1,...,u_{N+2})$; c_0 is the cost of the delivered water; c_1 is the cost of a solar battery; c_2 is the cost of a tank. This problem is the problem with chance constraint. But it can be reduced to a quantile problem by changing the variables. For simplicity I will consider the case of two months, i.e. N=2. Change the variables:

$$\tilde{u}_4 = \Phi_0 - 2u_4 a_2, \quad u_4 = (\Phi_0 - \tilde{u}_4)/2a_2; \quad 0 \leq \tilde{u}_4 \leq \Phi_0$$

Then the probabilistic constraint transfers to the following:

$$\mathbb{P} \{\omega_1, \omega_2 : \max_{i=\overline{1,3}} \Phi_i(u,\omega) \leq \Phi_0 \} \geq \alpha,$$

where

$$\Phi_1(u,\omega) = 2a_0 u_2 + 2a_1 u_3 - 2a_0 u_3 \omega_1 - \tilde{u}_4 \, 2a_0 h_1$$

$$\Phi_2(u,\omega) = -2a_2 u_2 - 2a_2 u_3 \omega_2 - \tilde{u}_4 \, 2a_0 h_2$$

$$\Phi_3(u,\omega) = 2a_1 u_3 - 2a_0 u_3 (\omega_0 + \omega_3) - \tilde{u}_4 \, 2a_0 (h_1 + h_2).$$

In this case Φ_0 is an α–order quantile of random variable $\max\limits_{i=\overline{1,3}} \Phi_i(u_1 w)$.

The three above presented algorithms were used to solve this problem. In the third algorithm due to the structure of an objective function one gets a linear programming problem. It happens since all the functions $\Phi_i(u,\omega)$ contain the same control parameter u_3 multiplied by the random parameters.

The results are presented in the table below. The input data in the problem were the following:

$$a_0 = 25 \ [rbl/m^3]; \quad a_1 = 5 \ [rbl/m^2]; \quad a_2 = 10 \ [rbl/m^3];$$

$$h_1 = 80 \ [m]^3, \ h_2 = 120 \ [m]^3, \ m_1 = 0.12 \ [m]; \ m_2 = 0.15 \ [m];$$

α	the cost Φ_0 [rbl]			time consumption for solving using IBM PC/AT [seconds]		
	direct	minimax	linear	direct	minimax	linear
0.9	4151	4258	4196	360	210	10
0.99	4389	4478	4407	1200	370	40
0.999	4614	4671	4629	4000	630	160
0.9999	4783	4854	4789	12500	800	600
0.99999	4930	5000	4939	40000	1100	1200

As you can see from this table, the minimax algorithm gives the overestimate for the quantile, but the time necessary for calculations is only slightly dependent on the probability α. If the probability level close to 0.99999 then the time consumptions for the minimax and the linear algorithms are approximately the same. It happens since when using the linear algorithm it is necessary to make statistical simulation in every step of the algorithm and the amount of calculations increases as α increases. But the time consumption when using linear algorithm

is considerably less than when using direct algorithm as in linear algorithm only one parameter is chosen, and in direct algorithm four parameters are varying. The linear algorithm gives the overestimate also, but the errors are smaller than when using minimax algorithm. A more powerful computer than IBM PC/AT was used to find the solution for the probability level $\alpha=0.99999$ using direct method. I must also note that if the statistical simulation is used consecutively with the random search method it results in increasement of the time consumptions that is necessary for the direct algorithm.

REFERENCES

1. Kibzun, A.I. and V.Yu.Kurbakovskiy Guaranteeing approach to solving quantile optimization problems.- Annals of Operations research, 30, 1991.
2. Kibzun, A.I. and V.Yu.Kurbakovskiy Numerical algorithms for quantile optimization and their application for solving probabilistic problems.- Izv. AN SSSR, Teknicheskaya Kibernetika, N 1, 1991. (in Russian).
3. Galambos J. The Asymptotic Theory of Extreme Order Statistics, Wiley, 1978.
4. Ermoliev, Yu.M. Methods of Stochastic Programming, Nauka, 1976 (in Russian)
5. Kibzun, A.I. Stochastic Control of Descrete-Time Dynamic Systems. - Moscow Aviation Institute, 1991 (in Russian)