INFORMATICA, 2015, Vol. 26, No. 4, 569–591

DOI: http://dx.doi.org/10.15388/Informatica.2015.65

Nonlinear Stochastic Programming Involving *CVaR* in the Objective and Constraints

Valerijonas DUMSKIS[1], Leonidas SAKALAUSKAS[2]*
[1]*Šiauliai University, Šiauliai, Lithuania*
[2]*Institute of Mathematics and Informatics, Vilnius University, Lithuania*
e-mail: valius.du@svajone.su.lt, leonidas.sakalauskas@mii.vu.lt

Received: February 2014; accepted: May 2015

Abstract. The nonlinear stochastic programming problem involving *CVaR* in the objective and constraints is considered. Solving the latter problem in a framework of bi-level stochastic programming, the extended Lagrangian is introduced and the related KKT conditions are derived. Next, the sequential simulation-based approach has been developed to solve stochastic problems with *CVaR* by finite sequences of Monte Carlo samples. The approach considered is grounded by the rule for iterative regulation of the Monte Carlo sample size and the stochastic termination procedure, taking into account the stochastic model risk. The rule is introduced to regulate the size of the Monte Carlo sample inversely proportionally to the square of the stochastic gradient norm allows us to solve stochastic nonlinear problems in a rational way and ensures the convergence. The proposed termination procedure enables us to test the KKT conditions in a statistical way and to evaluate the confidence intervals of the objective and constraint functions in a statistical way as well. The results of the Monte Carlo simulation with test functions and solution of the practice sample of trade-offs of gas purchases, storage and service reliability, illustrate the convergence of the approach considered as well as the ability to solve in a rational way the nonlinear stochastic programming problems handling *CVaR* in the objective and constraints, with an admissible accuracy, treated in a statistical manner.

Key words: stochastic programming, Monte Carlo method, stochastic gradient, CVAR.

1. Introduction

Real life decisions under uncertainty and risk are often modeled by linear or nonlinear stochastic programs. However, the optimization on average, usually realized in the framework of stochastic programming, is limited because it does not take into account the involved risk of possible deviations from the expected value. Indeed, in various practical applications, the random scenarios occur, that are rather unlikely, but which, in the case they do appear, have catastrophic consequences. Because of their low probability such scenarios would not have a significant impact on the expectation value, and finally can provide a decision, what doesn't take into account undesirable scenarios. This motivates the development of risk-averse optimization models involving coherent risk measures, because that can provide decisions more sensitive to harmful, but unlikely scenarios.

*Corresponding author.

Many measures are proposed for risk evaluation. Recently the Conditional Value-at-Risk (*CVaR*), introduced by Rockafellar and Uryasev (2000), has found many applications in the risk management as well as in the financial and engineering optimization. *CVaR* approximately (or exactly, under certain conditions) equals to the average of some percentage of the worst-case loss scenarios. Namely, assume a confidence level $\alpha \in (0, 1)$ and a loss function $f(x, y) : \Re^n \times \Omega \to \Re$, to be given, where x is a decision variable and y represents uncertain factors defined on a probability space (Ω, F, P). While *CVaR* is conceptually defined as the expectation of the random variable $f(x, y)$ in the conditional distribution of its upper α-tail, an operationally convenient definition is given by Rockafellar and Uryasev (2002):

$$CVaR_\alpha(x) = \min_u \left\{ u + \frac{1}{1-\alpha} E\big(f(x, y) - u\big)^+ \big| u \in \Re \right\}, \tag{1}$$

where $t^+ = \max\{0, t\}$ and E stands for the mathematical expectation.

Also, Rockafellar and Uryasev (2000, 2002) proposed an approach for optimizing *CVaR* and showed that the linear programming techniques might be used for optimization in some programs, involving CVaR. The paper of Krokhmal *et al.* (2002) was the first one to deal with portfolio optimization involving *CVaR* in the objective and constraints, and later on, Csaba (2008) considered the two-stage stochastic problems with *CVaR* in the objective and constraints. The issues of *CVaR* application to portfolio planning have been considered in Fortin and Hlouskova (2011), Lim *et al.* (2011), Grechuk and Zabarankin (2014). Application of *CVaR* optimization in engineering and design are described by Geihe *et al.* (2013), Legg *et al.* (2013), and others. It has been shown in Sakalauskas (2004a), Mulvey and Erkan (2006), Sutiene *et al.* (2010), Guigues and Romisch (2012), and Yao and Zhang (2012) that many problems of risk and profit management might be efficiently solved, using the *CVaR* risk measure and the stochastic optimization.

Thus, recently the known methods enable us to treat optimization problems involving *CVaR* as large linear counterparts, however, to solve them it may require huge computing resources. Besides, it is not clear how to solve the optimization problems with nonlinear random objective and/or constraint functions involving *CVaR*, when the distribution of random scenarios is known and continuous. In such a case, the sequential stochastic search by series of dynamically simulated samples looks as a way to develop the practical algorithms for nonlinear stochastic problems with *CVaR*, involved in the objective and/or constraints. The goal here is to simulate the scenarios sample-by-sample, aimed at improving the values of probabilistic objective functions under the given constraints at each sample (see, Cairoli and Dalang, 1996; Shapiro and Homem-de-Mello, 1998; Sakalauskas, 2002, 2004b; Bayraksan and Morton, 2011, etc.). The approach of sequential search by random samples has been developed for two-stage stochastic programming in Shapiro and Homem-de-Mello (1998). Later on the method of stochastic nonlinear programming by series of Monte Carlo samples in the framework of the stochastic gradient descent has been developed in Sakalauskas (2002). Since the amount of simulated scenarios in sampling-based algorithms can achieve a significant volume, the key feature is to

estimate and decrease the number of scenarios needed to obtain a reliable statistical estimate of the probabilistic functions to be optimized. The methods of sequential stochastic search developed in the framework of stochastic gradient descent, converge slowly if the Hessian of Lagrangian of the stochastic nonlinear problem is poorly defined. The idea behind to reinforce the gradient descent is the stochastic variable metric (SVM) approach, using the descent direction computed in the appropriately induced variable metric (see, Uryasev, 1992). Thus, this paper focuses on the development of a method of sequential stochastic search for nonlinear *CVaR* optimization, using the SVM method to improve the convergence. The paper is organized as follows: statement of the risk-aversion optimization problem involving *CVaR* and related Karush–Kuhn–Tucker (KKT) conditions are discussed in the next section, random sampling by the Monte Carlo method is described in Section 3, procedure of the sequential stochastic search itself is presented in Section 4, the results of computer simulation are presented in Section 5, and, finally, the concluding remarks are given in Section 6.

2. KKT Conditions in Stochastic Programming Involving *CVaR*

Risk management and investment optimization require the investor's risk aversion to be specified, because without efficient procedures for identifying the risk aversion, investors can overexpose themselves tom risk and lose profits. Usually the perception of risk is the amount of a security capital which would be set aside from the investments so that under any change of the market, the investor feels himself safe in total of the capital obtained from the market with his security capital. The risk aversion models are efficiently investigated from the stochastic programming point of view. Thus, since any investor's decision is related with a certain value of the expected objective function and risk, the trade-off between objective and risk may be presented through the level of risk-aversion expressed as a weighted sum with the weight θ on the function to be optimized, and with the weight $(1-\theta)$ on the risk measure, where $0 \leqslant \theta \leqslant 1$. Besides, in many decision-making problems under uncertainty it is also crucial to specify the decision makers' risk preferences, based on multiple stochastic risk measures, which is efficiently performed by shaping the distribution of scenarios according to the multiple *CVaR* constraints (see, Krokhmal *et al.*, 2002; Noyan and Rudolf, 2013). Therefore the incorporation of multiple stochastic risk measures into optimization models is a fairly recent research area (Guldman, 1983). The sketch of this problem and relative KKT conditions have been also presented in Dumskis *et al.* (2012).

Thus, assume the random functions $F_i : \Re^n \otimes \Omega \to \Re$, $i = 0, 1, 2, \ldots, m$, to be given, which obey certain conditions on differentiability with respect to $x \in \Re^n$ and integrity with respect to $\zeta \in \Omega$, here (Ω, Σ, P) is the probability space. Now the risk-aversion optimization problem with multiple *CVaR* constraints is stated for sake of simplicity in a following way omitting the constraints without *CVaR*:

$$\begin{aligned} f_0(x) &= \theta E\big[F_0(x,\zeta)\big] + (1-\theta) CVaR_{\alpha_0}\big[F_0(x,\zeta)\big] \to \min_x, \\ f_i(x) &= CVaR_{\alpha_1}\big[F_i(x,\zeta)\big] \leqslant \eta_i, \quad i = 1, 2, \ldots, m, \end{aligned} \tag{2}$$

where *CVaR* follows definition (1), $0 < \alpha_i \leqslant 1$, $i = 0, 1, 2, \ldots, m$. Hereinafter, assume the probabilistic measure to be absolutely continuous and defined by the probability density $p : \Omega \to \Re_+$. Thus, rewrite problem (2) in the following equivalent manner, taking into account definition (1):

$$\begin{aligned} f_0(x) &= \theta \int_{R^l} F_0(x, z)\, p(z)\, dz \\ &\quad + (1-\theta)\left(u_0 + \frac{1}{\alpha_0} \int_{R^l} \left(F_0(x, z) - u_0\right)^+ p(z)\, dz\right) \to \min_{x, u_0}, \\ f_i(x) &= \min_{u_i \in \Re} \left(u_i + \frac{1}{\alpha_i} \int_{R^l} \left(F_i(x, z) - u_i\right)^+ p(z)\, dz\right) \leqslant \eta_i, \quad i = 1, 2, \ldots, m. \end{aligned} \tag{3}$$

Note, that the latter problem turns out to be a stochastic bi-level programming problem or a stochastic Stackelberg game, equivalent to this problem, whereat the *leader* maximizes the risk-averse objective at the upper level choosing the control variables x as well as his *VaR* u_0, and, in turn, *followers* at a lower level optimize individually, under given x, their *CVaR* functions with respect to *VaR*'s u_i, $i = 1, 2, \ldots, m$.

Let us consider (2), (3) as a nonlinear programming problem. Note, that the objective function in the leader problem as well as the functions in the followers problems do not depend on *VaR*'s u_i in an explicit manner. Hereinafter, define the following functions to exploit the latter opportunity:

$$\begin{aligned} s_0(x, u) &= \theta E F_0(x, \zeta) + (1-\theta)\left(u_0 + \frac{E(F_0(x, \zeta) - u_0)^+}{\alpha_0}\right), \\ s_i(x, u) &= u_i + \frac{E(F_i(x, \zeta) - u_i)^+}{\alpha_i}, \end{aligned} \tag{4}$$

which, as it is easy to see, are expectations of the random functions:

$$\begin{aligned} S_0(x, u, \zeta) &= \theta F_0(x, \zeta) + (1-\theta)\left(u_0 + \frac{(F_0(x, \zeta) - u_0)^+}{\alpha_0}\right), \\ S_i(x, u_i, \zeta) &= u_i + \frac{(F_i(x, \zeta) - u_i)^+}{\alpha_i}, \end{aligned} \tag{5}$$

where $i = 1, 2, \ldots, m$.

The concept of the stochastic gradient is a key to explore the differentiability of probabilistic functions. Next propositions help us to study stochastic gradients of functions (4) behind the known fact that locally Lipschitz functions have subgradients, that can be taken as stochastic gradients (see, Clarke, 1983; Michalevitch *et al.*, 1987, etc.).

Proposition 1. *Assume that the objective and constraint functions* $F_i : \Re^n \otimes \Omega \to \Re$, *where* (Ω, Σ, P) *is a probability space, are integrable with respect to* $\zeta \in \Omega$, *and obey locally the Lipshitz property with respect to* $x \in \Re^n$, *besides, the following expectations are*

locally bounded: $|EF_i(x,\zeta)| < \infty$, $|EG_i(x,\zeta)| < \infty$, *where* $G_i(x,\zeta) = \partial_x F_i(x,\zeta)$ *are, respectively, subgradients of the considered random objective and constraint functions,* $i = 0, 1, 2, \ldots, m$. *Then at any* $\zeta \in \Omega$, $u_i \in \Re$, *the random functions in* (5) $S_i(x,u,\zeta)$, $x \in \Re^n$, $u = (u_0, u_1, \ldots, u_m)$, $u \in \Re^{m+1}$, $\zeta \in \Omega$, *obey the Lipshitz property with respect to* x, *and their subgradients are expressed as follows:*

$$\partial_x S_0(x,u,\zeta) = \theta G_0(x,\zeta) + (1-\theta)\frac{G_0(x,\zeta)H(F_0(x,\zeta)-u_0)}{\alpha_0},$$

$$\partial_x S_i(x,u_i,\zeta) = \frac{G_i(x,\zeta)H(F_i(x,\zeta)-u_i)}{\alpha_i}. \tag{6}$$

The values of which can be taken as stochastic gradients of the probabilistic functions (4), *namely:*

$$\partial_x s_i(x,u_i) = E\partial_x S_i(x,u_i,\zeta), \quad i = 0, 1, 2, \ldots, m, \tag{7}$$

where $H : \Re \to \{0,1\}$ *is the Heaviside step function* $H : \Re \to \{0,1\}$, Kanwal *(*1998*).*

Proof. Note, that it is easy to see that

$$\left|\left(F_i(x,\zeta)-u_i\right)^+ - \left(F_i(y,\zeta)-u_i\right)^+\right| \leqslant \left|F_i(x,\zeta) - F_i(y,\zeta)\right|.$$

Next, due to the latter inequality and the assumption on the Lipshitz property of random objective and constraint functions, the random functions $S_i(x,u_i,\zeta)$ in (5) at any $\zeta \in \Omega$, $u_i \in \Re$, are locally Lipshitzian with respect to variables x. Thus, the subgradients of these random functions exist because of the Lipshitz property. Hereupon the property of stochastic gradient (7) is true by virtue of Lebesgue Theorem (see, also, Ermoliev, 1983; Michalevitch *et al.*, 1987; Ermoliev and Wets, 2011, etc.). □

Corollary 1. *If the random objective and constraint functions are given by smoothing*: $F_i(x,\zeta) = F_i(x+\zeta)$, *where the uncertainty is described by an absolutely continuous measure, and functions* $F_i : \Re^n \to \Re$ *obey locally the Lipshitz property, then the functions* $s_i(x,u_i)$ *are differentiable with respect to* $x \in \Re^n$, $i = 0, 1, 2, \ldots, m$.

Indeed, by virtue of Rademacher's Theorem the Lebesgue measure is zero of the set in Ω, where the subgradient (6) is set-valued. Thus, the expression (7) is defined a.s. unambiguously as an point-valued expectation (see, also, Rockafellar and Wets, 1982; Shapiro and Homem-de-Mello, 1998; Bartkute and Sakalauskas, 2007).

Proposition 2. *Let the assumptions of Proposition* 1 *be satisfied, and, besides, the considered probabilistic measure to be absolutely continuous. Then, the subgradients of random functions* (5) *with respect to VaR's* $u_i \in \Re$ *are as follows*:

$$\partial_{u_i} S_i(x,u_i,\zeta) = 1 - \frac{H(F_i(x,\zeta)-u_i)}{\alpha_i}. \tag{8}$$

Besides, functions (4) $s_i(x, u_i)$ *are smoothly differentiable with respect to VaR's* u_i*, and their subgradients turn out almost everywhere to derivatives, expressed by the probabilities of risk* $\Pr(F_i(x, \zeta) \geqslant u_i)$:

$$\frac{ds_i(x, u_i)}{du_i} \equiv E\partial_{u_i} S_i(x, u_i, \zeta) = 1 - \frac{\Pr(F_i(x, \zeta) \geqslant u_i)}{\alpha_i}, \quad i = 0, 1, 2, \ldots, m. \tag{9}$$

Proof. Follows from the Lipschitz property with respect to *VaR*'s u_i of random functions (5), Rademacher's and Lebesgue's theorems, analogously to the proof of the previous proposition. □

Thus, a smooth differentiability of probabilistic objective and constraint functions can be established rather often in practical stochastic problems involving *CVaR*, and the propositions above have presented cases of this kind.

Now let us derive the Karush–Kuhn–Tucker (KKT) conditions of the problem considered, taking into account the latter remark. The standard Lagrangian of this problem is $f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x)$. However, let us introduce extended Lagrange function (LF):

$$\begin{aligned} l(x, u, \lambda) \equiv\ & EL(x, u, \lambda, \zeta) = s_0(x, u) + \sum_{i=0}^{m} \lambda_i s_i(x, u) \\ \equiv\ & \theta \int_{R^l} F_0(x, z) p(z)\, dz \\ & + \sum_{i=0}^{m} \lambda_i \left(u_i - \eta_i + \frac{1}{\alpha_i} \int_{R^n} \left(F_i(x, z) - u_i\right)^+ p(z)\, dz \right), \end{aligned} \tag{10}$$

which, in its turn, can be treated as an expectation of the random Lagrange function:

$$L(x, u, \lambda, \zeta) = \theta F_0(x, \zeta) + \sum_{i=0}^{m} \lambda_i \left(u_i - \eta_i + \frac{(F_i(x, \zeta) - u_i)^+}{\alpha_i} \right), \tag{11}$$

$u = (u_0, u_1, \ldots, u_m)$, $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m)$, $\lambda_0 = 1 - \theta$, $\eta_0 = 0$.

Indeed, under the Lipshitz property of the random objective and constraint functions $F_i(x, \zeta)$, the Lagrangian $L(x, u, \lambda, \zeta)$, $x \in \Re^n$, $\lambda \in \Re^{m+1}$, $u \in \Re^{m+1}$, $\zeta \in \Omega$, obeys the Lipshitz property with respect to x, and the values of its subgradients are as follows:

$$\partial_x L(x, u, \lambda, \zeta) \equiv Q(x, \lambda, u, \zeta) = \theta G_0(x, \zeta) + \sum_{i=0}^{m} \frac{\lambda_i}{\alpha_i} G_i(x, \zeta) H\left(F_i(x, \zeta) - u_i\right), \tag{12}$$

which can be taken as stochastic gradients of the extended Lagrangian (10):

$$\partial_x l(x, u, \lambda) \equiv q(x, u, \lambda) = E\, Q(x, u, \lambda, \zeta). \tag{13}$$

Differentiability of the extended Lagrangian with respect to *VaR*'s vector $u \in \Re^{m+1}$ follows from Proposition 2.

Theorem 1. *Assume that the functions $s_i(x, u_i)$, $i = 0, 1, 2, \ldots, m$, defined by* (4)*, are differentiable with respect to variables $x \in \Re^n$ at any $u_i \in \Re$, and the constants on the right side of constraints in* (2) *are chosen so that*

$$\prod_{i=1}^{n} \left\{(x, u) : s_i(x, u) < \eta_i\right\} \neq \emptyset. \tag{14}$$

Now, let $x^* \in \Re^n$ be the solution to problem (2)*, u_i^* be corresponding VaR's in the optimal problems of a leader and followers, and, let the gradients $\nabla_x s_i(x^*, u_i^*)$ be positively linearly independent, $i = 0, 1, 2, \ldots, m$. Then there exist the values $\lambda_i^* \geqslant 0$, $i = 1, \ldots, m$, such that*

$$\begin{aligned}
&\nabla_x s_0\left(x^*, u^*\right) + \sum_{i=1}^{m} \lambda_i^* \nabla_x s_i\left(x^*, u_i^*\right) = 0,\\
&\lambda_i^* s_i\left(x^*, u_i^*\right) = 0, \quad i = 1, 2, \ldots, m,\\
&\Pr\left(F_i\left(x^*, \zeta\right) \geqslant u_i^*\right) = \alpha_i, \quad i = 0, 1, \ldots, m.
\end{aligned} \tag{15}$$

Proof. Indeed, according to assumption (14), the solution point $x \in \Re^n$ can be chosen so that $s_i(x, u_i^*) < \eta_i$, $i = 1, \ldots, m$. Thus, the Slater condition of the problem (2) is satisfied, because in such a case, $f_i(x) = \min_{u_i \in \Re} s_i(x, u_i) < \eta_i$, $i = 1, \ldots, m$. Besides, the objective and constraint functions $f_i(x)$ are differentiable under the given assumptions and the corresponding gradients are positively linearly independent. Then the KKT conditions of problem (2) exist by virtue of the Karush-Kuhn-Tucker theorem (see, e.g., Bertsekas, 1982), which can be written in an equivalent manner as (15) □

3. Random Sampling

Since the expectations and their derivatives, defined by (11), (13), (15), etc., turn out to be complicated multivariate integrals, a random sampling provides the universal technique to solve problems of the considered kind. Now, assume that the Monte Carlo samples of certain size N

$$Z = \left(z^1, z^2, \ldots, z^N\right), \tag{16}$$

where z^i are independent random vectors, realizing random scenarios, identically distributed with the density $p(\cdot) : \Omega \to R_+$, are available at each point $x \in \Re^n$. Now, let us

define Monte Carlo estimators:

$$\tilde{s}_0(x,u) = \frac{\theta}{N}\sum_{j=1}^{N} F_0\big(x,z^j\big) + \frac{1-\theta}{N_0} \sum_{\substack{j=1,\\ F_0(x,z^j)\geqslant u_0}}^{N} F_0\big(x,z^j\big), \tag{17}$$

$$\tilde{s}_i(x,u) = \frac{1}{N_i} \sum_{\substack{j=1,\\ F_i(x,z^j)\geqslant u_i}}^{N} F_i\big(x,z^j\big), \tag{18}$$

$$\Pr_i = \frac{N_i}{N}, \tag{19}$$

where $N_i = \sum_{j=1}^{N} H(F_i(x,z^j) - u_i)$ are the frequencies of events $\{u_i : F_i(x,z^j) \geqslant u_i\}$, that occur in the scenarios of sample (16), $i = 1,2,\ldots,m$. Assume $N_i \neq 0$, without loss of generality, because, in the opposite case, one can always repeat the sampling. By the Strong Law of Large Numbers estimators (17), (18) are consistent estimators of expected values (4), and estimators (19) are, respectively, that of risk probabilities $\Pr(F_i(x,\zeta) \geqslant u_i)$, $i = 0,1,\ldots,m$ (Shapiro, 2003). Let us also define the sampling variances:

$$\tilde{D}_i^2(x,u) = \frac{1}{N_i} \sum_{\substack{j=1,\\ F_i(x,z^j)>u_i}}^{N} \big(F_i\big(x,z^j\big) - \tilde{s}_i(x)\big)^2, \quad i = 1,2,\ldots,m. \tag{20}$$

Naturally, the Monte Carlo estimator of LF is as follows:

$$\tilde{L}(x,\lambda,u) = \tilde{s}_0(x,u) + \sum_{i=1}^{m} \lambda_i \big(\tilde{s}_i(x,u) - \eta_i\big). \tag{21}$$

Let us apply the stochastic gradient technique to approximate the gradients of objective and constraint functions by the Monte Carlo method. Let the subgradients

$$G_i(x,\zeta) = \partial F_i(x,\zeta), \quad i = 0,1,\ldots,m.$$

be available at the scenarios in random sample (16), which can be taken as a stochastic gradient of functions (4), assuming the latter ones to be differentiable. Thus, the sampling estimates of these gradients are as follows:

$$\begin{aligned}
\tilde{g}_0(x,u) &= \frac{1}{N}\sum_{j=1}^{N}\left(\theta G_0\big(x,z^j\big) + \frac{1-\theta}{\Pr_0} G_0\big(x,z^j\big) H\big(F_0\big(x,z^j\big) - u_0\big)\right),\\
\tilde{g}_i(x,u) &= \frac{1}{N_i}\sum_{j=1}^{N} G_i\big(x,z^j\big) H\big(F_i\big(x,z^j\big) - u_i\big),
\end{aligned} \tag{22}$$

where $i = 1,2,\ldots,m$.

Let us write the corresponding sampling estimator of the Lagrangian gradient:

$$q(x,\lambda,u)=\tilde{g}_0(x,u)+\sum_{i=1}^{m}\lambda_i\tilde{g}_i(x,u)\equiv\frac{1}{N}\sum_{j=1}^{N}Q\big(x,\lambda,u,z^j\big), \tag{23}$$

which is the average of the corresponding stochastic Lagrangian gradient:

$$Q\big(x,\lambda,u,z^j\big)=\theta G_0\big(x,z^j\big)+\sum_{i=0}^{m}\lambda_i\frac{G_i(x,z^j)H(F_i(x,z^j)-u_i)}{N_i/N},$$

where $\lambda_0=1-\theta$.

The corresponding sampling covariance matrix is as follows:

$$A(x,\lambda,u)=\frac{1}{N}\sum_{j=1}^{N}\big(Q\big(x,\lambda,u,z^j\big)-q(x,\lambda,u)\big)\big(Q\big(x,\lambda,u,z^j\big)-q(x,\lambda,u)\big)^T, \tag{24}$$

that will be further applied for normalization of stochastic gradient (23).

4. Stochastic Optimization Algorithm

Estimators (17)–(24) might be perfectly applied in solving stochastic problems with *CVaR* involved in the objective and constraints by means of gradient descent. The gradient descent approach for stochastic nonlinear programming by a series of Monte Carlo estimators has been developed in Sakalauskas (2002), however, it converges slowly if the Hessian of the Lagrangian is poorly defined. Let us consider the ideas behind to reinforce the gradient descent, based on the Natural Gradient Descent (Park *et al.*, 2000), and the stochastic variable metric (SVM), using the descent direction computed in the appropriately induced variable metric (see, Uryasev, 1992).

Thus, assume the initial point $x^0\in\Re^n$ and vectors $\lambda^0,u^0\in\Re^m_+$ to be given as well as the random sample (16) of a certain initial size N^0 be simulated, and Monte Carlo estimates (17)–(24) to be computed. Let us consider the following stochastic procedure:

$$\begin{aligned}
&x^{t+1}=x^t-\rho\big(B^t\big)^{-1}q\big(x^t,\lambda^t,u^t\big),\\
&\lambda_i^{t+1}=\max\left[0,\lambda_i^t+\gamma_i\left(\tilde{s}_i(x^t,u^t)-\eta_i+\mu_\beta\frac{\tilde{D}_i(x^t,u^t)}{\sqrt{N^t}}\right)\right],\quad i=1,2,\ldots,m,\\
&u_i^{t+1}=u_i^t-\pi_i\left(1-\frac{\Pr_i}{\alpha_i}\right),\quad i=0,1,\ldots,m
\end{aligned} \tag{25}$$

where $\rho>0,\gamma_i>0,\pi_i>0$ are steplengths of descent, μ_β is the β-quantile of the standard normal distribution, and the matrix B^t induces a variable metric. Two variable metrics

are examined, the first one in the original space, i.e., $B^t = I$, which corresponds to the gradient descent, and the second one is the metric induced using the sampling covariance matrix (24):

$$B^t = \left(A\left(x^t, \lambda^t, u^t\right) + \left(q\left(x^t, \lambda^t, u^t\right)\right)^T q\left(x^t, \lambda^t, u^t\right)\right). \tag{26}$$

Monte Carlo estimates, applied in (25), are random, as usual, and their uncertainty mainly depends on the sample size. However, there is no necessity to simulate large random samples (16) at the beginning of search, because it is more important to have large samples, when the optimum is approaching. According to this idea the sample size can be adjusted by means of the rule:

$$N^{t+1} = \frac{\chi_n^2(\nu)}{(q^t)^T (A^t)^{-1} q^t}, \tag{27}$$

where $\chi_n^2(\nu)$ is the ν-quantile of χ^2 distribution with n degrees of freedom, $q^t = q(x^t, \lambda^t)$ is the estimate of the Lagrange function gradient (23), and $A^t = A(x^t, u^t, \lambda^t)$ is the covariance matrix matrix (24), estimated at the point (x^t, u^t, λ^t), N^{t+1} is the sample size at $t+1$ iteration, $t = 0, 1, 2, \ldots$. Hence, the sample size is taken inversely proportional to the square norm of the Lagrangian stochastic gradient in the metric induced by the sampling covariance matrix (24). Indeed, such a choice of sample size ensures small samples at the beginning of the search, when the current point of solution is far from the optimum, and, therefore the gradient is large, but these samples increase when approaching the optimum, because the gradient becomes small then. Using the stochastic Lyapunov function technique, developed in Sakalauskas (2002), one can make sure that rule (27) under the appropriate choice of algorithm parameters ensures a.s. the increase of sample size and the convergence of the algorithm (25) to the solution of the optimization problem (2)–(3).

Now, let us more in detail consider the solution with an admissible accuracy of the stochastic programming problem with *CVaR* involved in the objective and constraints, applying the sequence generated by (25), (27). Certainly, a decision on finding the optimum should be made taking into account the sampling error of estimates (17)–(24). Thus, the sequence generated by algorithm (25) might be terminated in a statistical way, testing the statistical hypotheses on the validity of KKT conditions (15), whenever the confidence intervals of objective and constraint functions as well as that of estimated risk probabilities decreased up to an appropriate length.

Note that, since the sample size, chosen with respect to (27), is increasing, the distribution of the considered Monte Carlo estimators can be approximated by the Gaussian law according to Central Limit Theorem (Shapiro, 2003). Such a property helps us to construct statistical means for testing the optimality. For instance, one can apply the well-known Hotelling criterion to verify the hypothesis on the equality to zero of the LF gradient (see, Sakalauskas, 2002, 2004b). Hence, the hypothesis on the equality to zero of the LF gradient is not rejected at some point x with the significance probability μ, if the following condition vanishes:

$$(N-n)\left(q^t\right)^T \left(A^t\right)^{-1} q^t \leqslant \chi_n^2(\zeta), \tag{28}$$

where $\chi_n^2(\zeta)$ is the ζ-quantile of χ^2 distribution with n degrees of freedom. Similarly, the statistical hypotheses on the validity of constraints with probability β are not rejected, if:

$$\tilde{s}_i\left(x^t, u^t\right) - \eta_i + \mu_\beta \frac{\tilde{D}_i(x^t, u^t)}{\sqrt{N^t}} \leqslant 0. \tag{29}$$

Next, the estimated lengths of the confidence interval of the objective function and constraints should not exceed the admissible accuracy ε_i with probabilities β_i, if:

$$2\mu_{\beta_i} \frac{\tilde{D}_i(x^t, u^t)}{\sqrt{N^t}} \leqslant \varepsilon_i, \quad i = 0, 1, \ldots, m, \tag{30}$$

where μ_β denote hereinafter denote the β-quantile of the standard normal distribution, and, analogously, *VaR*'s u_i should be properly chosen in the Monte Carlo *CVaR* estimators, namely:

$$|\mathrm{Pr}_i - \alpha_i| \leqslant \mu_{\sigma_i} \sqrt{\frac{\mathrm{Pr}_i(1 - \mathrm{Pr}_i)}{N}}, \quad i = 1, 2, \ldots, m. \tag{31}$$

The algorithm for optimization by handling *CVaR* in the objective and constraints is described (Algorithm 1).

Thus, if all criteria (28)–(31) are satisfied, then there are no reasons to reject the optimality hypothesis at the current point of the sequence generated according to (25), (27), and, thereby there is a basis for the algorithm termination and decision making about finding the optimum with an permissible accuracy. However, if at least one condition in (28)–(31) is violated, then a new point should be computed according to (25) and next sample (16) of the sample size, adjusted according to (27), should be generated and estimators (17)–(24) should be computed, and so on. Indeed, the probabilities of errors of the first and second kind, when testing the termination criteria (28)–(31), might be regulated by choosing the proper probabilities of quantiles.

5. Counterexamples

Monte Carlo study by solving test problems many times is a widely used way to explore the stochastic optimization algorithms. Let us consider the results of such computer study

Example 1. (Piecewise-linear test functions with *CVaR* in the objective and constraints.) The developed approach has been tested by simulating piecewise-linear test functions:

$$F_i(x, \zeta) = \max_{1 \leqslant k \leqslant k_n} \left(a_{0,k} + \sum_{j=1}^{n} a_{j,k}(x_j + \zeta_j) \right), \quad 0 \leqslant i \leqslant m, \tag{32}$$

hence, for each number of variables $n = 2, 5, 10, 20, 50$, a sample from $M = 100$ sets of normally distributed coefficients $a_{j,k}$ has been simulated with the following values:

Algorithm 1 The algorithm for nonlinear optimization with *CVaR*

Initial conditions: formal description of the objective function and constraints.
Final conditions: vectors of the solution $x^* = (x_1^*, \ldots, x_n^*)$, $u^* = (u_1^*, \ldots, u_m^*)$, the value of the objective function $f_0(x^*, u^*)$, estimated confidence intervals of the objective function and constraints.

1. Fixate parameters of the problem: n is the number of the variables, m is the number of constraints, θ is a risk aversion coefficient, N^0 is the size of the initial Monte Carlo sample, ε_i, $i = 0, 1, \ldots, m$, are the accuracies of evaluation of the objective function and that of constraint functions, α_i, $i = 0, 1, \ldots, m$, are significance levels of *CVaR*'s, β is the probability of the confidence interval, μ is the probability of the optimality hypothesis, ρ, γ_i, π_i are steplengths of descent, *iter* is the maximal number of iterations, $i = 0, 1, \ldots, m$.
2. Fixate that $x^0 = (x_1^0, \ldots, x_n^0)$, $u^0 = (u_1^0, \ldots, u_m^0)$ are the initial vectors for sequential search, $\lambda^0 = (\lambda_1^0, \ldots, \lambda_m^0)$ – initial vector of parameters, set $t = 0$ as the number of iterations.
3. **While the** number of iterations t doesn't exceed *iter* **Do**
   - 3.1 Simulate the Monte Carlo sample $Z = (z^1, z^2, \ldots, z^{N^t})$ (16).
   - 3.2 Compute the frequencies N_i of the events $\{u_i^t : F_i(x, z^j) \geqslant u_i^t\}$ and $\Pr_i = N_i / N^t$ (19). If at least one $N_i = 0$ then go to 3.1.
   - 3.3 Compute estimator $\tilde{s}_0(x^t, u^t)$ (17) of the objective function, estimators $\tilde{s}_i(x^t, u^t)$ (18) of constraints functions, estimator (19) of *CVaR* probabilities, estimators (20) of the variances $\tilde{D}_i^2(x^t, u^t)$ and estimator (21) of LF $\tilde{L}(x^t, \lambda^t, u^t)$.
   - 3.4 Compute the estimator of the stochastic gradient $\tilde{g}_0(x^t, u^t)$ of the objective function, estimators of the gradients $\tilde{g}_i(x^t, u^t)$ of constraint functions (22), and estimator (23) of the gradient of LF $q(x^t, \lambda^t, u^t)$.
   - 3.5 Compute the sampling covariance matrix $A(x^t, \lambda^t, u^t)$ (24).
   - 3.6 Test the termination conditions:
     If (28) & (29) & (30) & (31), **Then** return the vectors $x^* = x^t = (x_1^t, \ldots, x_n^t)$, $u^* = u^t = (u_1^t, \ldots, u_m^t)$ and the value of objective function $f_0(x^*, u^*)$.
     Else set $t = t + 1$ and find the next point x^{t+1}, λ^{t+1}, u^{t+1} (25) and the size N^{t+1} (27) of Monte Carlo sample.
4. **Done.**

$a_{0,0} = \vartheta$, $a_{1,0} = 1 + 3\vartheta$, $a_{i,k} = a'_{i,k} - \frac{1}{k_n} \sum_{k1=1}^{k_n} a'_{i,k1}$, $a'_{0,k} = 2\vartheta$, $a'_{1,k} = \vartheta$, $1 \leqslant k \leqslant k_k$, $m = 0, 1$, where ϑ was the standard norm, variables ζ_k were distributed normally $N(0, 0, 5)$. When computing *CVaR*'s the risk probabilities were taken $\alpha_0, \alpha_1 = 0.1$, initial sample size was $N^0 = 500$. Other details of the test problem as well as some computer simulation results are given in Table 1.

The chosen class of test functions has been chosen due to its universality, because any convex function can be approximated to a desired accuracy by the functions of shape (32), taking rather a large number k_n and choosing appropriate coefficients $a_{j,k}$.

Table 1
Data of the test problem and some results.

n	u_0	u_1	η_1	ε	k_n	min, *it*	max, *it*	aver., *it*
2	3.34	3.46	4.5	0.015	5	8	16	9.4
5	4.93	6.82	7	0.1	11	8	22	12.5
10	8.33	8.63	9	0.1	21	6	14	7.7
20	12.42	9.59	11.5	0.075	31	14	53	26.6
50	20.84	14.01	15	0.1	76	44	88	59.9

The simulation results obtained in this way, enable us to draw conclusions that can be applied when solving similar practical problems. Initial approximations $x^0 = (x_1^0, \ldots, x_n^0)$, $u^0 = (u_1^0, \ldots, u_m^0)$ were taken as the Expected Value solutions (EVS) (Birge and Louveaux, 2011). EVS of *CVaR* threshold variables u_0 and u_1, taken as initial, are given in Table 1. All the test problems have been solved by the approach considered in the SVM framework, i.e., problems have been solved in the metric, induced by matrix (26), with dimensionalities $n = 2, 5, 10, 20, 50$, i.e. 500 test problems were solved in total. The termination conditions in (28)–(31) have been tested with probabilities of significance $\beta = \beta_i = \zeta = 0.05$, the quantile in (27) was chosen with the probability $\nu = 0.99$. In order to avoid very large samples, simulated according to (27), the sample simulation was interrupted, whenever the confidence intervals became smaller than the prescribed values ε (given in Table 1).

The averaged results of solving the test problems are presented in Fig. 1 and some results are also given in Table 1. Figure 1 presents the averaged dependencies of characteristics of the method, depending on the number of iterations t, that illustrate the convergence of the approach developed. Namely, the averaged dependencies of the objective function value as well as probabilities of *CVaR*, depending on the number of iterations are presented in Fig. 1(a)–(c), illustrating the convergence of the approach developed. The frequency of termination according to the number of iterations is presented in Fig. 1(d). The ratio of criterion (28) with the respective quantile of χ^2 distribution depending on the number of iterations is shown in Fig. 1(e), which illustrates how this ratio tends to critical termination value 1. The averaged number of the Monte Carlo sample size at each iteration is presented in Fig. 1(f), which shows the adaptation of this sample during the optimization process. Note, that all the test problems were terminated according to (28)–(31) after some number of iterations. The maximal, minimal, and averaged number of iterations, used to solve test problems with an admissible accuracy ε, is given in Table 1. The averaged Monte Carlo sample size is presented in Fig. 1(f), which illustrates the adaptation of this sample during optimization as well. For instance, one can see from (30) and (31) that the length of the confidence interval depends on the variance. However, since the simulated objective and constrained functions occurred having small variances in cases $n = 5$ and $n = 50$, the local increase of sample size is noticed on figures before achieving the zone of convergence.

Actually, when solving the problem with a higher accuracy, the number of iterations and the total amount of simulated scenarios are increased. Efficiency on the dimensionality

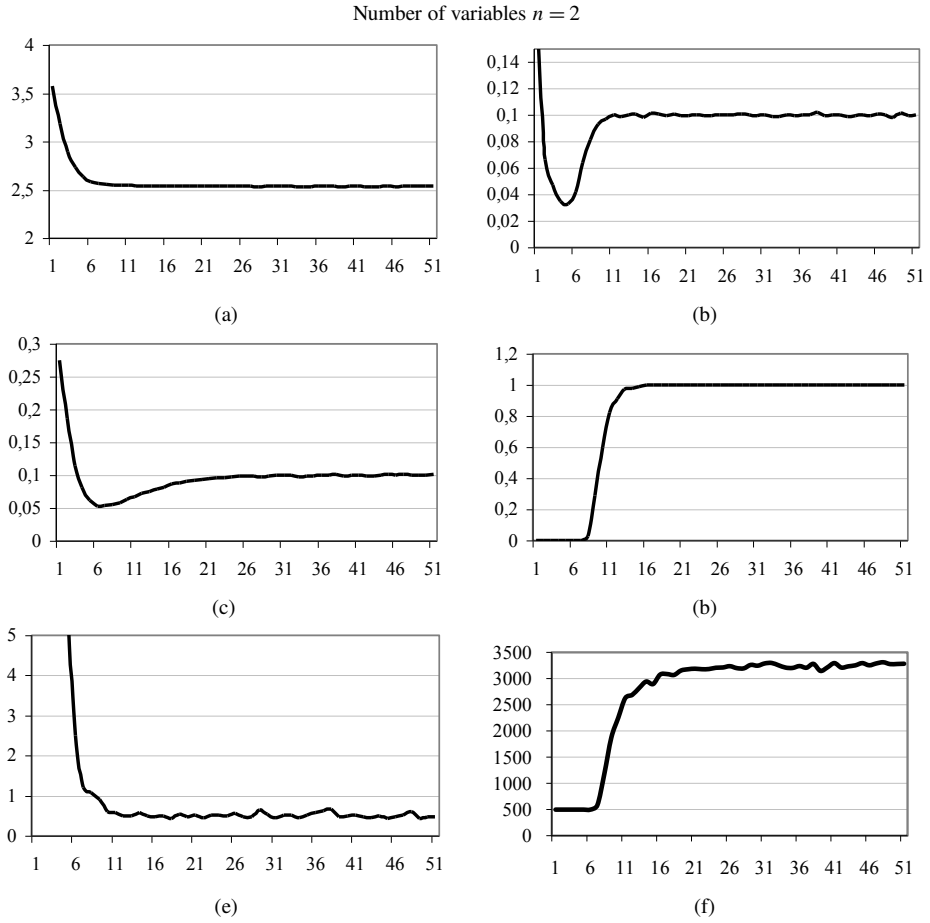


Fig. 1. Averaged results of simulation: (a) objective function value; (b) frequency Pr_0; (c) frequency Pr_1; (d) frequency of termination; (e) ratio of Hotelling criterion (28) with 0.95-quantile of χ^2 distribution; (f) Monte Carlo sample size at each iteration.

of the developed method was not explored, because it was not clear how the variance of estimators depends on dimensionality, however one can see, that the computer time, needed to solve the problem has increased, if the number of variables increased.

The simulation results allow us to conclude that the approach considered enables us to solve with an admissible accuracy, the nonlinear stochastic problems handling *CVaR* in the objective and constraints, using the rational amount of computer resources. Indeed, the computer time mainly depends on the number of simulated random scenarios, where the stochastic approach is applied. Thus, if the termination condition (28) is satisfied with probability 0.95, the objective function is evaluated with accuracy, given in Table 1, and the number of variables is varying as $n = 2, 5, 10, 20, 50$, the number of iterations, needed to solve the problem, varies from ≈ 8 to ≈ 60, and the total amount of simulated scenarios varies from ≈ 950 to ≈ 6000.

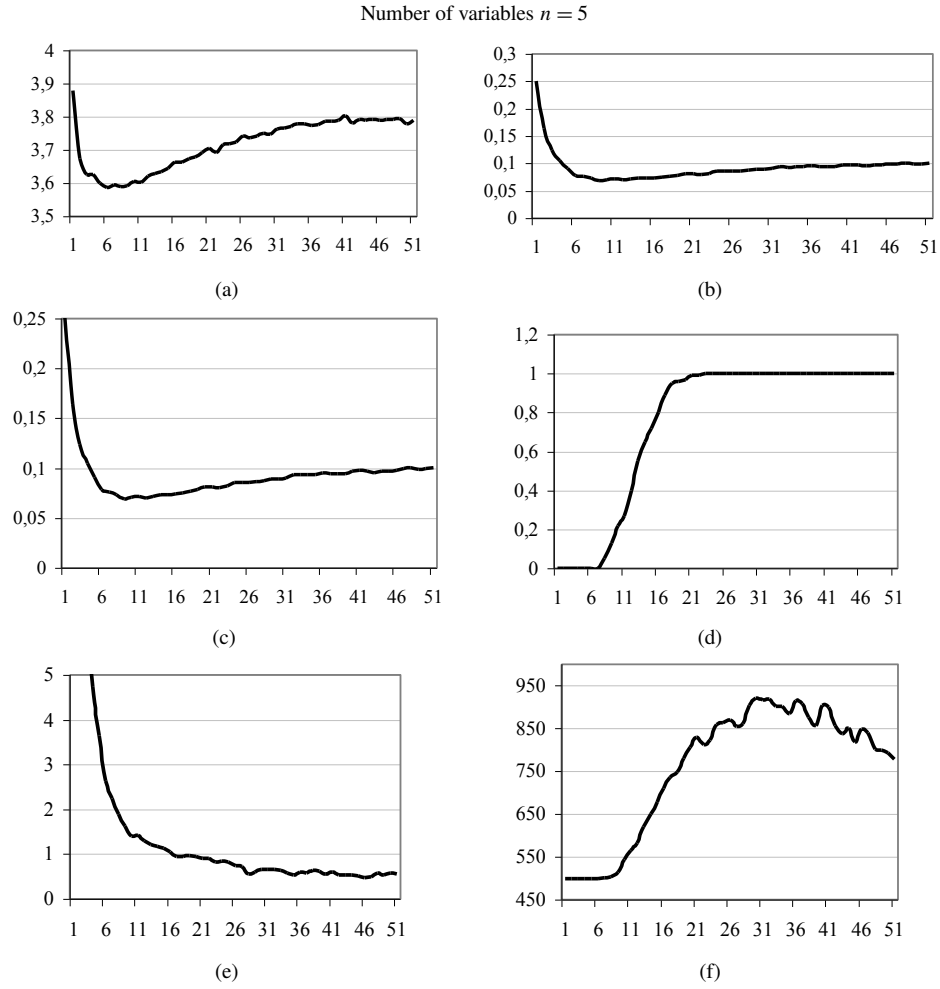


Fig. 1. (*continued*)

EXAMPLE 2. (Trade-offs in gas purchase, storage and service reliability.) Let us consider a practice example of the trade-offs problem on purchase, storage, and service reliability decisions faced by distribution utilities of natural gas, when the short-term demand for gas fluctuates randomly because of the weather (see for details and references in Ermoliev and Wets, 2011; Sakalauskas, 2002). To encourage load leveling, the pipeline transmission companies that supply utilities use a demand contract, which charges utilities based on their peak day needs, and often charges for the minimum daily purchase requirement, whether that the purchase is made or not. The policy variables, available to the utility, include increase in its storage capacity, and provision of interruptible service to some customers lowering thereby the reliability of service.

To explore trade-offs, the following cost minimization problem in the risk-averse framework has been solved. Choose the annual plan of gas purchases $x = (x_1, \ldots, x_{12})$,

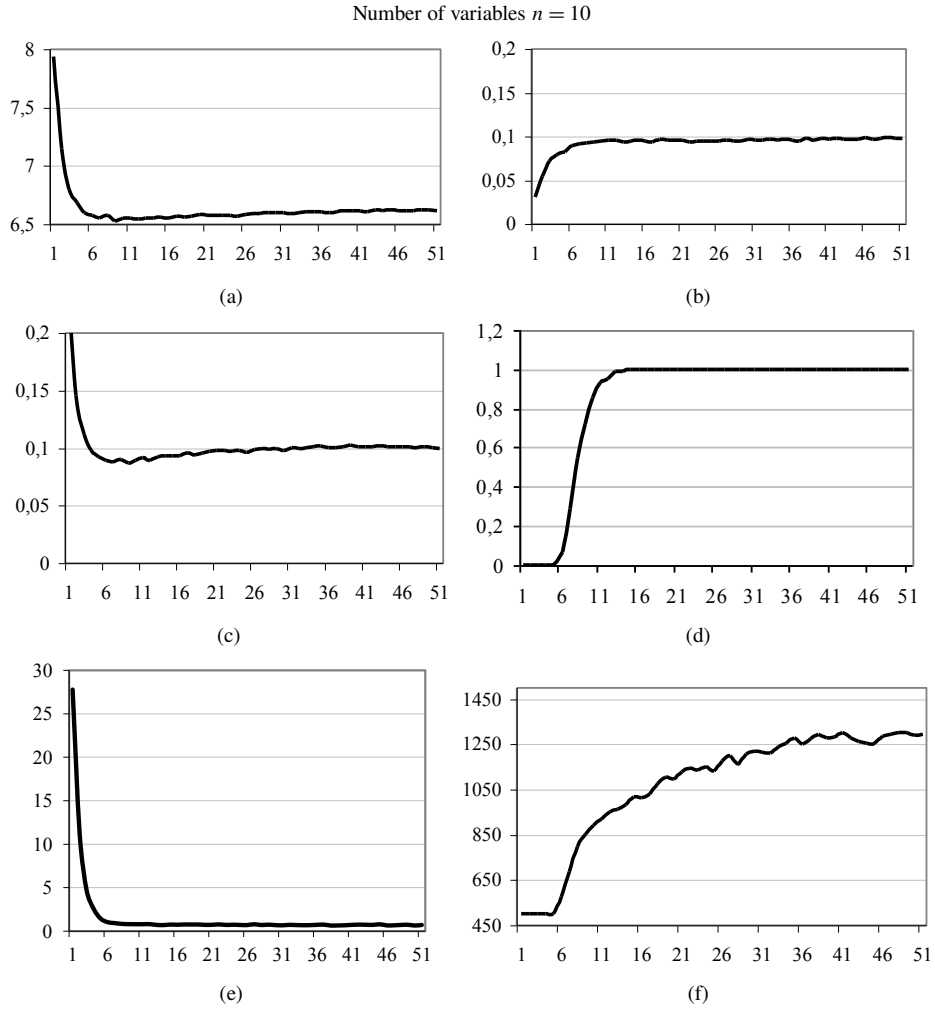


Fig. 1. (*continued*)

where x_i is the amount of gas ordered from the pipeline per month t under the cost c_t in order to satisfy the actual gas consumption ω_t, considered as independent and normal $N(\mu_t, \sigma_t)$, aimed to minimize the expected cost of purchases and storage operations. Note that the additional cost c_3 of the maximal monthly gas order $\max_i\{x_i\}$, applied to countervail the supply level, and the cost c_0 of transportation to storage of monthly gas surplus or shortage $|x_t - \omega_t|$, where $t = 1, 2, \ldots, 12$, should be taken into account as well.

In order to meet technical restrictions on the maximum storage gas flows, the increment of the gas storage z, and the gas amount v supplied from an additional source in order to provide the interruptible service, is planned under respective costs. Taking into account

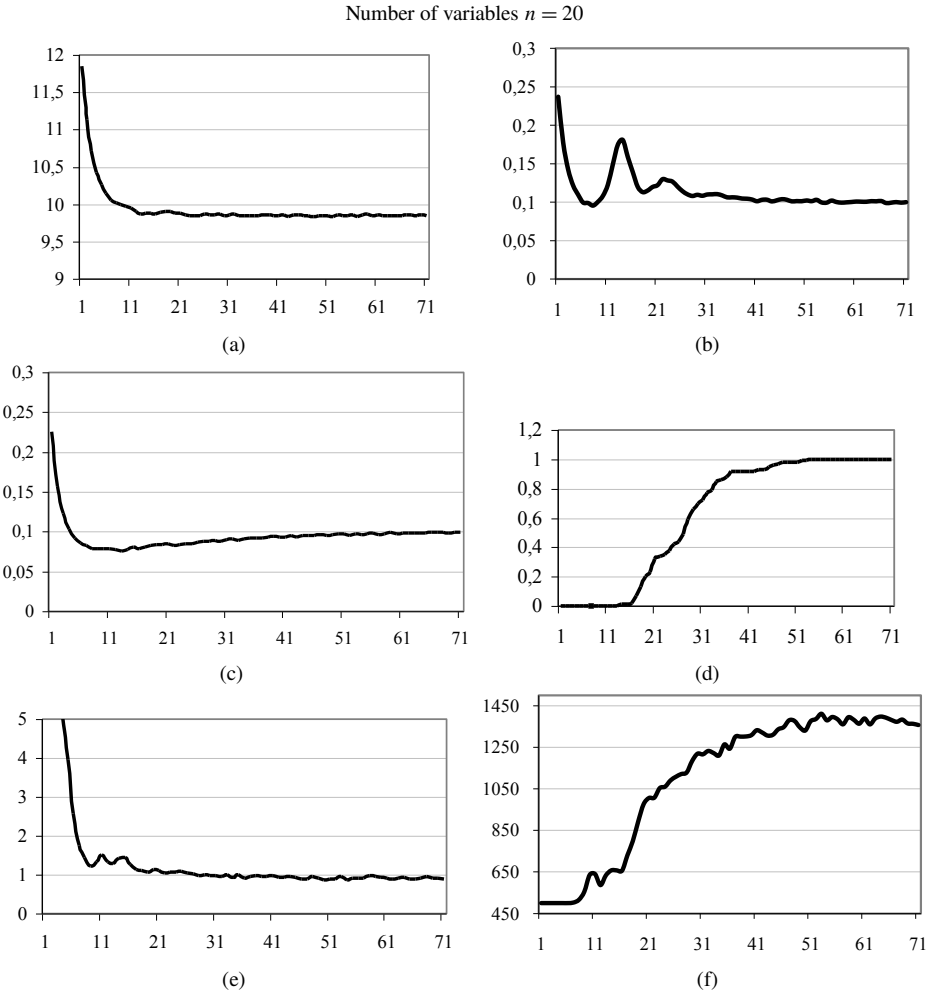


Fig. 1. (*continued*)

these assumptions, the latter technical restrictions are as follows:

$$
\begin{aligned}
&(x_t - \omega_t) - a_1 \sum_{s=1}^{t-1} (x_s - \omega_s) \leqslant b_1 + a_2 z,\\
&-(x_t - \omega_t) - a_3 \sum_{s=1}^{t-1} (x_s - \omega_s) \leqslant b_2 + a_4 z,\\
&\sum_{s=1}^{t} (x_s - \omega_s) \leqslant b_3 + a_5 z,\\
&\sum_{s=1}^{t} (x_s - \omega_s) \geqslant b_4 - a_6 v, \quad t = 1, 2, \ldots, 12.
\end{aligned} \tag{33}
$$

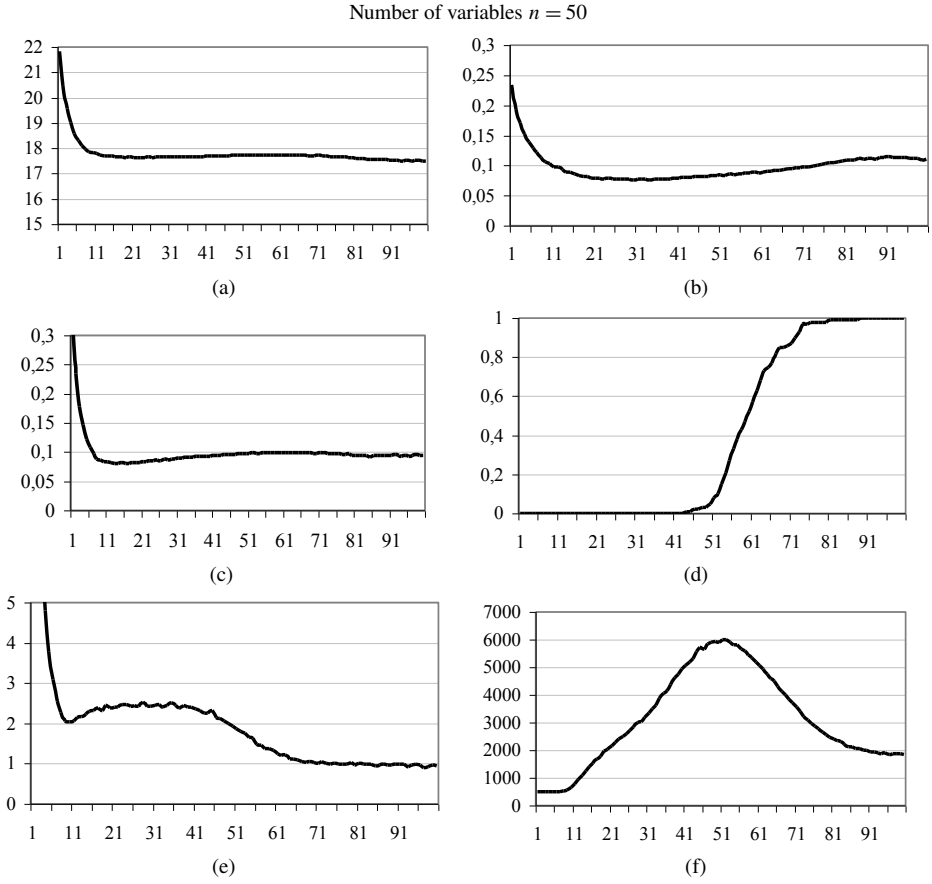


Fig. 1. (*continued*)

Now the optimization problem, involving *CVaR* in the objective and constraints with the confidence level α, is as follows:

$$
\begin{aligned}
F(x) &= \sum_{t=1}^{12} c_t x_t + c_0 \max_i \{x_i\} + c_{13} E \sum_{t=1}^{12} |x_t - \omega_t| \\
&\quad + c_{14} CVaR_\alpha \left(\sum_{t=1}^{12} |x_t - \omega_t| \right) \to \min_{x \geqslant 0} \\
\text{s.t.:} \quad & CVaR_\alpha\big(Z(x,\omega)\big) \leqslant \eta, \\
& CVaR_\alpha(V(x,\omega)) \leqslant \gamma,
\end{aligned} \tag{34}
$$

where:

$$
\begin{aligned}
&Z(x,\omega) \\
&\quad = \max_{1 \leqslant t \leqslant 12} \left[\frac{(x_t - \omega_t)(1 + a_1) - a_1 \sum_{s=1}^{t} (x_s - \omega_s) - b_1}{a_2}, \right.
\end{aligned}
$$

Table 2
Data of the gas trade-off problem.

T	1	2	3	4	5	6	7	8	9	10	11	12
c_t	1.2024	1.2024	1.2024	1.2024	1.2024	1.2024	1.2024	1.2993	1.2993	1.2993	1.2993	1.2993
μ_t	33.433	23.980	16.747	15.302	15.591	19.308	28.494	40.969	54.102	59.081	53.177	47.550
σ_t	3.311	3.230	1.054	0.344	0.516	1.540	3.333	3.132	5.334	4.735	4.214	4.588
x_t^{50}	35.451	25.692	17.210	15.432	15.782	19.863	29.754	41.755	57.758	60.887	54.017	48.716

$$\frac{-(x_t-\omega_t)(1-a_3)-a_3\sum_{s=1}^{t}(x_s-\omega_s)-b_2}{a_4}, \frac{\sum_{s=1}^{t}(x_s-\omega_s)-b_3}{a_5}\Bigg],$$

$$V(x,\omega)=\max_{1\leqslant t\leqslant 12}\frac{b_4-\sum_{s=1}^{t}(x_s-\omega_s)}{a_6}.$$

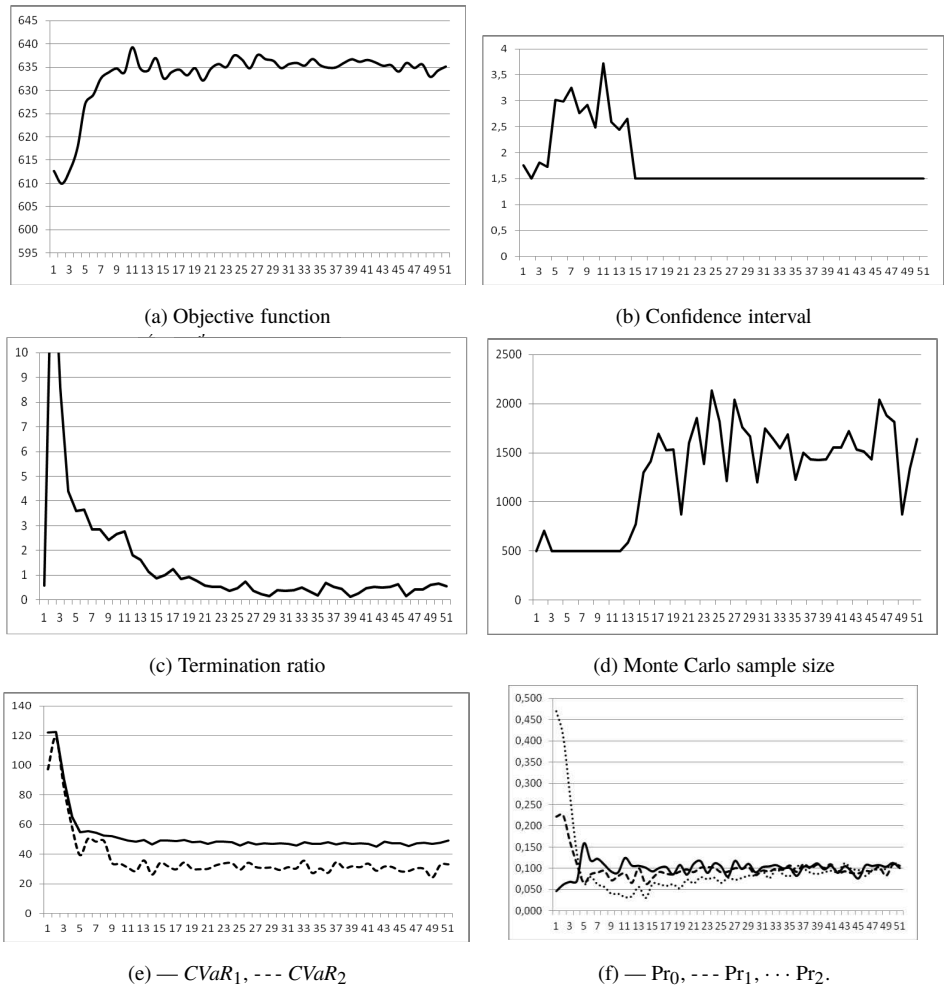
The data of the problem are as follows: $a_1 = 0.078$, $a_2 = 0.8$, $a_3 = 0.15$, $a_4 = 0.049$, $a_5 = 0.41$, $a_6 = 0.37$, $b_1 = 118.0752$, $b_2 = 7.2321$, $b_3 = 30.5153$, $b_4 = 8.34$, $c_{13} = 0.03323$, $c_{14} = 1.6782$, $c_0 = 0.392$, $\eta = 35$, $\gamma = 50$, $\alpha = 0.1$, other data are given in Table 2. The task considered has been solved using the developed approach with the following parameters: $\beta = \beta_i = \zeta = 0.05$, $\nu = 0.999$, $\rho = 0.05$, $\gamma_1 = 2.5$, $\gamma_2 = 1$, $\gamma_3 = 2.5$, $\pi_i = 0.005$, $i = 0, 1, 2$. The initial approximation of decision vector was $x^0 = \mu$, the inital approximations of *VaR*'s were $u^0 = (42, 10, 40)$.

The results of solving the task are presented in Figs. 2(a)–(f). The objective function and its confidence interval change in Figs. 2(a), (b) illustrate the convergence of the method. The change of the ratio of criterion (28) with the respective quantile of χ^2 distribution depending on the number of iterations presented in Fig. 2(c), shows, how this ratio tends to the critical termination value 1. The Monte Carlo sample size at each iteration is given in Fig. 2(d). Figures 2(e) and (f) show the change of *CVaR*'s and *VaR*'s according to the constraints in problem (33). Only a few iterations were indispensable to establish the optimal decision, i.e., to fulfill the termination conditions.

Thus, the results of the Monte Carlo study with test functions and solution of the practice example of trade-offs of gas purchases, storage, and service reliability, illustrate the convergence of the approach considered as well as the ability to solve nonlinear stochastic programming problems in a reasonable way, when handling *CVaR* in the objective and constraints, with an admissible accuracy, treated in a statistical manner.

6. Discussion and Conclusions

Thus, the approach of sequential Monte Carlo search to nonlinear stochastic programming involving *CVaR* in the objective and multiple constraints has been developed and investigated in the paper. The theoretical background for the approach is created by introducing the extended Lagrangian and deriving the related KKT conditions. The approach developed is grounded on a rule of the iterative regulation of Monte Carlo sample size and termination of the optimization in a statistical manner, taking into account the statistical simulation risk. The proposed termination procedure allows us to verify the statistical



(a) Objective function

(b) Confidence interval

(c) Termination ratio

(d) Monte Carlo sample size

(e) — $CVaR_1$, - - - $CVaR_2$

(f) — Pr_0, - - - Pr_1, $\cdots$ Pr_2.

Fig. 2. Results of optimization of trade-offs of gas purchases, storage and service reliability.

hypothesis on the validity of KKT conditions and to evaluate the confidence intervals of the objective and constraint functions as well. Although in stochastic programming it is common to use sampling approximations, the amount of simulated samples can achieve a significant volume, and therefore the key feature is to estimate and decrease the number of scenarios needed to obtain the reliable statistical estimate of probabilistic functions to be optimized. The proposed regulation of sample size, when this size is taken inversely proportional to the square of the norm of the gradient of the Monte Carlo estimate, enables us to solve stochastic problems with *CVaR* rationally from the computational viewpoint and ensures the convergence. In order to reinforce the gradient descent, the stochastic variable metric approach is proposed, using the descent direction, computed in the metric, induced by the sampling covariance matrix of stochastic gradient. The results of Monte Carlo simulation and solution of a practice example of trade-offs in gas delivery plan-

ning, have illustrated the convergence of the approach considered as well as the ability to solve stochastic programming problems with an admissible accuracy, treated in a statistical manner, by using reasonable computer resources, when *CVaR* is involved in the objective and constraints.

Aknowledgment. We express our sincere gratitude to Prof. Vincent Guigues from IMPA—National Institute of Pure and Applied Mathematics, Rio de Janeiro, Brazil, on statement of problem and valuable discussion.

References

Alexander, G., Baptista, A., Jan, S. (2012). When more is less: using multiple constraints to reduce tail risk. *Journal of Banking & Finance*, 36(10), 2693–2716.

Bayraksan, G., Morton, D.P. (2011). A sequential sampling procedure for stochastic programming. *Operations Research*, 59(4), 898–913.

Bartkute, V., Sakalauskas, L. (2007). Simultaneous perturbation stochastic approximation for nonsmooth functions. *European Journal of Operational Research*, 181(3), 1174–1188.

Bertsekas, D.P. (1982). *Constrained Optimization and Lagrange Multiplier Methods*. Academic Press, San Diego.

Birge, J.R., Louveaux, F. (2011). Introduction to Stochastic Programming. *Springer Series in Operations Research*. Springer, New York.

Cairoli, R., Dalang, R.C. (1996). *Sequential Stochastic Optimization*. Wiley Series in Probability and Statistics. Wiley, New York.

Clarke, F.H. (1983). *Optimization and Nonsmooth Analysis*. John Wiley & Sons.

Csaba, I.F. (2008). Handling *CVaR* objectives and constraints in two-stage stochastic models. *European Journal of Operational Research*, 191, 888–911.

Dumskis, V., Guigues, V., Sakalauskas, L. (2012). On risk aversion optimization by Monte-Carlo method. In: *Proceedings of International Workshop of Stochastic Programming for Implementation and Advanced Applications (STOPROG 2012), July 3–6, 2012, Lithuania*. Technika, Vilnius, 2012, pp. 19–23.

Ermoliev, Yu. (1983). Stochastic quasigradient methods and their application to system optimization. *Stochastic*, 9, 1–36.

Ermoliev, Yu., Wets, R.J.-B. (2011). *Numerical Techniques for Stochastic Optimization*. Springer Series in Computational Mathematics. Springer, Berlin.

Fortin, I., Hlouskova, J. (2011). Optimal asset allocation under linear loss aversion. *Journal of Banking & Finance*, 35(11), 2974–2990.

Geihe, B., Lenz, M., Rumpf, M., Schultz, R. (2013). Risk averse elastic shape optimization with parametrized fine scale geometry. *Mathematical Programming, Ser. A*, 141, 383–403.

Grechuk, B., Zabarankin, M. (2014). Inverse portfolio problem with mean-deviation model. *European Journal of Operational Research*, 234(2), 481–490.

Guigues, V., Romisch, W. (2012). Sampling-based decomposition methods for multistage stochastic programs based on extended polyhedral risk measures. *SIAM Journal on Optimization*, 22 286–312.

Guldman, J. (1983). Supply, storage and service reliability decisions by gas distribution utilities: a chance-constrained approach. *Management Science*, 29(8), 884–906.

Kanwal, R.P. (1998.) *Generalized Functions: Theory and Technique*. Birkhäuser, Basel.

Krokhmal, P., Palmquist, J., Uryasev, S. (2002). Portfolio optimization with conditional value-at-risk objective and constraints. *Journal of Risk*, 4, 11–27.

Michalevitch, V.S., Gupal, A.M., Norkin, V.I., 1987. *Methods of Nonconvex Optimization*. Nauka, Moscow (in Russian).

Mulvey, J.M., Erkan, H.G. (2006). Applying *CVaR* for decentralized risk management of financial companies. *Journal of Banking & Finance*, 30, 627–644.

Legg, S.W., Wang, C., Benavides-Serrano, A.J., Laird, C.D. (2013). Optimal gas detector placement under uncertainty considering conditional-value-at-risk. *Journal of Loss Prevention in the Process Industries*, 26(3), 410–417.

Lim A.E.B., Shanthikumar, J.G., Vahn, G.-Y. (2011). Conditional value-at-risk in portfolio optimization: Coherent but fragile. *Operations Research Letters*, 39(3), 163–171.

Noyan, N., Rudolf, G. (2013). Optimization with multivariate conditional value-at-risk constraints. *Operations Research*, 61(4), 990–1013.

Park, H., Amari, S.-I., Fukumizu, K. (2000). Adaptive natural gradient learning algorithms for various stochastic models. *Neural Networks*, 13(7), 755–764.

Rockafellar, R.T., Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, 2, 21–41.

Rockafellar, R.T., Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking and Finance*, 26(7), 1443–1471.

Rockafellar, R.T., Wets, R.J.-B. (1982). On the interchange of subdifferentiation and conditional expectation for convex functional. *Stochastics*, 7, 173–182.

Sakalauskas, L. (2002). Nonlinear stochastic programming by Monte Carlo estimators. *European Journal of Operational Research*, 137, 547–555.

Sakalauskas, L. (2004a). Financial risk in open economies. *European Journal of Operational Research*, 155(2), 267–532.

Sakalauskas, L. (2004b). Application of the Monte Carlo method to nonlinear stochastic optimization with linear constraints. *Informatica*, 15(2), 271–282.

Shapiro, A. (2003). Monte Carlo sampling methods. In: Ruszczynski, A., Shapiro, A. (Eds.), *Stochastic Programming*, *Handbooks in Operations Research and Management. Science*, vol. 10. Elsevier, Amsterdam.

Shapiro, A., Homem-de-Mello, T. (1998). A simulation-based approach to two-stage stochastic programming with recourse. *Mathematical Programming*, 81, 301–325.

Sutiene, K., Makackas, D., Pranevicius, H. (2010). Multistage K-means clustering for scenario tree construction. *Informatica*, 21(1), 123–138.

Uryasev, S. (1992). A stochastic quasi-gradient algorithm with variable metric. *Annals of Operations Research*, 39, 251–267.

Yao, F., Zhang, P. (2012). The measurement of operational risk based on *CVaR*: a decision engineering technique. *Systems Engineering Procedia*, 4, 438–447.

V. Dumskis, PhD, lecturer at the Institute of Informatics, Mathematics and E-Studies of Šiauliai University.

L. Sakalauskas, Prof. Habil. Dr., he is the Principal Researcher and Head of Department of Operational Research of the Institute of Mathematics and Informatics of Vilnius University, President of Lituanian Operational Research Society, member of several European working groups. His scientific interests: stochastic programming, queueing theory, statistical simulation, operational research.

Netiesinis stochastinis programavimas įterpiant sąlyginę riziką į tikslo funkciją ir ribojimus

Valerijonas DUMSKIS, Leonidas SAKALAUSKAS

Nagrinėjama netiesinio stochastinio programavimo problema su sąlygine rizika, įterpta į tikslo funkciją ir ribojimus. Įvestas išplėstinis Lagranžianas ir išvestos atitinkamos KKT sąlygos, sprendžiant šią problemą, kaip dviejų lygių stochastinio programavimo uždavinį. Toliau plėtojamas būdas stochastinėms problemoms su sąlygine rizika spręsti nuosekliai generuojant baigtines Monte-Carlo sekas. Nagrinėjamas būdas remiasi Monte-Karlo imčių tūrio iteratyvinio reguliavimo taisykle ir stochastine stabdymo taisykle, atsižvelgiant į modelio riziką. Monte-Karlo imčių tūris yra imamas atvirkščiai proporcingu stochastinio gradiento normos kvadratui, nes šitoks reguliavimas užtikrina konvergavimą ir leidžia spręsti uždavinį, racionaliai panaudojant kompiuterio išteklius. Pasiūlyta stabdymo taisyklė leidžia tikrinti optimalumo sąlygas statistiniu būdu, kartu įvertinant tikslo funkcijos ir ribojimų tikėtinumo intervalus. Testinių funkcijų modeliavimas Monte-Karlo metodu ir dujų tiekimo kompanijos investicijų planavimo praktinio uždavinio sprendimas parodė sukurto metodo konvergavimą bei galimybę efektyviai spręsti stochastines problemas su sąlygine rizika, įterptą į tikslo funkciją ir ribojimus.