

ONE METHOD TO SOLVE A PROBLEM OF NONLINEAR TWO-STAGE PERSPECTIVE STOCHASTIC PLANNING

Marina Volkova

Saint-Petersburg State University
199034, Universitetskaya nab., 7-9
Saint-Petersburg, RUSSIA

Abstract: The two-stage problem of perspective stochastic programming is considered. The sufficient conditions for the existence of the finite solution to the problem and simple sufficient conditions for the uniqueness of the solution are found. The effective direct method of stochastic programming based on step by step calculation of solutions of deterministic nonlinear programs appearing when random variables are fixed is proposed.

AMS Subject Classification: 90C15, 90C30, 49M37

Key Words: stochastic programming, iterative method, decision rule, planning, convex

1. Introduction

The main goal of this work is to clarify the existence of a solution often arises in economic models of two-stage stochastic programming problems and getting a simple algorithm for finding the solution [1, 2, 4, 5, 6].

Classic interpretation scheme of simplest one-stage perspective stochastic programming problem is as follows: nonlinear deterministic problem with several parameters (predecessor problem) is described as a non-linear program with the variables of interest, these are the control action. Thus, the minimiza-

Received: January 31, 2017

Revised: February 11, 2017

Published: February 19, 2017

© 2017 Academic Publications, Ltd.

url: www.acadpubl.eu

tion (maximization) of the objective function (functional) of the program with its constraints satisfaction is declared desirable. Further these parameters of the deterministic problem are assumed unknown till the time of control action choice, but in the future they will be random variables with determined distributions. This creates a place for a new deterministic problem named as deterministic equivalent. Usually the objective function of the predecessor problem replaced by its mathematical expectation, as well as some constraint-functions. Also other statistical functions could be used instead of the expectation. For a two-stage stochastic problem a formulation of a illustrative predecessor problem, which contains two stages, is more complicated. Instead, the one-stage predecessor problem with partial randomness will be described.

First let us consider the nonlinear program for: it is required to minimize the objective function

$$f(x, \omega, \widehat{\omega}) := f_1(x, \omega) + \inf_{y \geq 0 \wedge (3)} f_2(y) \longrightarrow \inf_{x \geq 0 \wedge (2)}, \tag{1}$$

with two type constraints

$$g(x, \omega) \leq 0, \tag{2}$$

$$h(x, \widehat{\omega}) - y \leq 0, \tag{3}$$

where $\omega = (\omega_1, \dots, \omega_r)^T \in \Omega \subseteq R^r$ and $\widehat{\omega} = (\widehat{\omega}_1, \dots, \widehat{\omega}_{\widehat{r}})^T \in \widehat{\Omega} \subseteq R^{\widehat{r}}$ are tuples composed of problem parameters, $\Omega, \widehat{\Omega}$ are sets of vector parameters, $f_1 : R_+^n \times \Omega \rightarrow R^1, f_2 : R_+^k \rightarrow R^1, g : R_+^n \times \Omega \rightarrow R^m, R_+^i = \{x \mid x \geq \vec{0}^i\}, h : R_+^n \times \widehat{\Omega} \rightarrow R^k,$

$$g = \begin{pmatrix} g_1(x, \omega) \\ \vdots \\ g_m(x, \omega) \end{pmatrix}, \quad h = \begin{pmatrix} h_1(x, \omega) \\ \vdots \\ h_k(x, \omega) \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_k \end{pmatrix}. \tag{4}$$

Inequalities in (2) and (3) are understood component-wise. It is assumed that f_1, g and h are convex functions with respect to the first argument in R_+^n and continuously differentiable functions with respect to $x \in R_+^n$ for all $\omega, \widehat{\omega}; f_2$ strictly monotonically grows along each argument.

Ordinary nonlinear problem (1)–(4) becomes notably complicated after the time introducing, even if such time consists of two moments.

Formulas of stages. The first stage corresponds to the initial moment of the time, the second stage corresponds to the last moment. It is assumed that ω and $\widehat{\omega}$ are a unknown tuples and x should be chosen at the first stage. At the second stage ω and $\widehat{\omega}$ become known and value of $y = y(x(\omega), \widehat{\omega})$ should be chosen while value of x cannot be changed.

The program (1)–(4) with this formulas of stages is two-stage problem of perspective stochastic programming. During its second stage the search of the posterior decision rule (choice of $y = y^{opt}$) must be done. That is the choice of algorithm solving the following nonlinear problem

$$f_2(y) \longrightarrow \min_{y \geq 0 \wedge h(x, \hat{\omega}) - y \leq 0} . \tag{5}$$

The first stage is most complex. This is search for a priori decision rule. Insufficiency of data (that are $\omega, \hat{\omega}$) does not permit to interpret it as a nonlinear program. Hence the first step of the first stage is some action which is not a part of the total solving algorithm. During the action one formulates the new problem and defines its solution. There are different ways to formulate this problem depending on the type of parameters in the tuple $\hat{\omega}$. If these parameters are interpreted as random variables which have known distributions then usually ordinarily mathematical expectation of initial objective function is appointed as new objective function with fixed ω

$$F(x, \omega) = f_1(x, \omega) + E \left[\min_{\substack{y \geq 0 \\ h(x, \hat{\omega}) - y \leq 0}} f_2(y) \right] \longrightarrow \min_{\substack{x \geq 0 \\ g(x, \omega) \leq 0}} . \tag{6}$$

Stochastic program (2), (3), (4), (6) is a complete problem. For certain ω we understand its solution as a value of objective function F^{opt} together with minimizer x^{opt} which is either an objective function minimum point satisfying (2) and (3) or a minimizing sequence $\{x^i\}_{i=0}^\infty$ such as

$$f_1(x^i, \omega) + E \left[\min_{\substack{y \geq 0 \\ h(x^i, \hat{\omega}) - y \leq 0}} f_2(y) \right] \longrightarrow \inf_{\substack{x \geq 0 \\ g(x, \omega) \leq 0}} F(x, \omega) =: F^{opt}$$

where mathematical expectation is taken only by $\hat{\omega}$ when ω is fixed and besides for all i the following inequalities must be hold

$$\left\{ \begin{array}{l} g(x^i, \omega) \leq 0 \\ (\forall \hat{\omega} \in \hat{\Omega}) \quad h(x^i, \hat{\omega}) \leq y(x^i, \hat{\omega}) \end{array} \right. .$$

Here $y(x^i, \hat{\omega})$ is a minimizer of nonlinear program (5) when $x = x^i$.

Union of minimizers and values of objective functions $f_1(x^{opt}, \omega)$ and $F(x^{opt}, \omega)$ in programs (5) and (6) are understood as general solution of the two-stage problem. Minimizer x^{opt} and minimizer $y(x^{opt}, \hat{\omega})$ termed as decision rule have

different temporal activity: x^{opt} operates from appearance of program (5) solution until the appearance of information about $\omega, \hat{\omega}$, decision rule $y(x^{opt}, \hat{\omega})$ operates after obtaining $\omega, \hat{\omega}$.

Let us consider the tuple ω and it contains parameters of two types. One type of parameters is characterized by random factors such as demand, supply, weather, breakdown equipment, etc. These factors may appear in the future. The other type of parameters corresponds to determined factors with value known only approximately, for example manpower or capacity of the oil field. These parameters are relative to the past and present time. It is convenient to define them as random values with known distribution as in the first type of parameters. Briefly known components of the tuple ω do not contribute anything new to the statement of the problem. Hence they can be neglected by shifting them from parameters to the collection of details forming the functions.

The relation between tuples $\omega, \hat{\omega}$ and constraints (2) and (3) is different. Let us consider that constraints (3) must be fulfilled for every conceivable realization of the parameters (e.g. number of passengers in a taxi). In other words, it is necessary to observe the general natural constriction of the system for every random realization of the problem condition and these constraints cannot be violated at all. Let us consider that constraints (2) should be fulfilled upon the statistical average (e.g. average occupancy of the certain flight should have some concrete level for economic viability of this direction). In other words, long-term contract obligations (raw material supply, output of products, profit, etc.) should be satisfied, though some short-time breach of contract obligations is allowed. Hence, it is reasonable enough that constrain (2) can be changed to

$$E[g(x, \omega)] \leq 0. \quad (7)$$

Problem (5) solution becomes random after introducing random ω in functions f_1 and g and that makes this problem inadequate for modelling some real processes. The similar situation was considered yet for the first stage of the two-stage stochastic program. It is naturally to solve this problem in an analogous way, i.e. to mind desirable optimum as the solution of the nonlinear stochastic program

$$\Phi(x) = E \left[f_1(x, \omega) + \min_{\substack{y \geq 0 \\ h(x, \hat{\omega}) - y \leq 0}} f_2(y) \right] \longrightarrow \min_{\substack{x \geq 0 \\ E[g(x, \omega)] \leq 0}}. \quad (8)$$

Program (8) is often called as determined equivalent of the problem (1)–(4) wherein random nature is given to all its parameters.

Posterior decision rule for the second stage. The very nature of the function f_2 implies its strictly monotone increasing with respect to argument. This provides the existence of minimizer of the nonlinear program (5) and permits to find it easily.

Let us introduce the operation:

$$h^+ := \left(\max(0, h_1), \dots, \max(0, h_k) \right). \tag{9}$$

Then the desired minimizer, i. e. decision rule, has the following form

$$y^{opt}(x, \hat{\omega}) = h^+(x, \hat{\omega}). \tag{10}$$

Applying it to program (8) yields one-stage problem

$$\Phi(x) = \mathbb{E} \left[f_1(x, \omega) + f_2(h^+(x, \hat{\omega})) \right] \longrightarrow \min_{\substack{x \geq 0 \\ \mathbb{E}[g(x, \omega)] \leq 0}} =: \Phi^{opt}, \tag{11}$$

i.e. in (11) it is reputed that ω and $\hat{\omega}$ appear at one moment.

2. Existence and Uniqueness of Two-Stage Stochastic Programming Problem (11)

The following result will be useful from now on.

Lemma 1. *Assume the function $g(x, \omega)$ be uniformly continuous function with respect to x . Then the function*

$$G(x) := \mathbb{E} [g(x, \omega)] = \int_{\Omega} g(x, \omega) \rho(\omega) d\omega,$$

where $\rho(\omega)$ is probability density, is continuous.

Proof. According to the uniform continuity of function g

$$\begin{aligned} (\forall \varepsilon > 0)(\exists \Delta_\varepsilon > 0) \|x - x'\| < \Delta_\varepsilon &\implies \\ (\forall \omega \in \Omega) |g(x', \omega) - g(x, \omega)| < \varepsilon. \end{aligned}$$

Let us choose some ε and Δ satisfying the last implication for this ε . Then

$$\left| \mathbb{E} [g(x', \omega)] - \mathbb{E} [g(x, \omega)] \right| = \left| \int_{\Omega} g(x', \omega) \rho(\omega) d\omega - \int_{\Omega} g(x, \omega) \rho(\omega) d\omega \right| =$$

$$\begin{aligned}
 &= \left| \int_{\Omega} [g(x', \omega) - g(x, \omega)] \rho(\omega) d\omega \right| \leq \\
 &\leq \int_{\Omega} |g(x', \omega) - g(x, \omega)| \rho(\omega) d\omega \leq \int_{\Omega} \varepsilon \rho(\omega) d\omega = \varepsilon.
 \end{aligned}$$

□

Theorem 1. (Existence of a finite solution of the problem). *Let*

- 1) $g(x, \omega)$ be defined over $R^n \times \Omega$,
- 2) Set $M = \{x \mid E[g(x, \omega)] \leq 0\}$ be not empty and bounded,
- 3) $f_1(x, \omega)$, $h(x, \omega)$ be defined and continuous with respect to x over $M \times \Omega$, g be uniformly continuous with respect to x over $M \times \Omega$,
- 4) $f_2(y)$ be defined and continuous over $h^+(M, \Omega)$, and increase strictly monotonically in each y component (h^+ is defined in (9)).

Then the finite solution of the determined equivalent (11) exists. If the set M is unbounded then the finite solution may not exist.

Proof. Lemma 1 gives the continuity of $G(x) = E[g(x, \omega)]$ which implies that set M is closed. Together with condition 2) this gives the possibility to apply Cantor theorem that on a closed and bounded set of finite-dimensional space continuous function is uniformly continuous. Thus we obtain the uniform continuity of $f_1(x, \omega)$ and $f_2(h^+(x, \hat{\omega}))$ with respect to x at any $\omega, \hat{\omega}$. Lemma 1 implies continuity of Φ on M . Hence closedness and boundedness of M leads to existence of finite minimizer of the Φ function. □

Theorem 2. (Uniqueness of the solution of (11)). *Let in addition to the conditions of Theorem 1*

- 1) $f_1(x, \omega)$ be strictly convex at any fixed ω as a function of x ,
- 2) $f_2(y)$ be convex,
- 3) $h(x, \omega)$ be convex at any fixed ω as a function of x .

Then the solution of the problem (11) is unique.

Proof. 0. One knows that the sum of convex functions is convex and strictly convex if at least one of them is strictly convex.

1. Therefore $E(f_1(x, \omega))$ is strictly convex function.

2. $h^+(x, \hat{\omega})$ is convex in x for all $\hat{\omega}$ due to the convexity $h(x, \hat{\omega})$. Therefore, superposition $\varphi(x, \hat{\omega}) := f_2(h^+(x, \hat{\omega}))$ is also convex for all $\hat{\omega}$. Indeed,

$$\begin{aligned}
 \varphi(\lambda x' + (1 - \lambda)x'') &= f_2(h^+(\lambda x' + (1 - \lambda)x'', \hat{\omega})) \leq \\
 &\leq f_2(\lambda h^+(x', \hat{\omega}) + (1 - \lambda)h^+(x'', \hat{\omega})) \leq
 \end{aligned}$$

$$\leq \lambda f_2(h^+(x', \hat{\omega})) + (1 - \lambda) f_2(h^+(x'', \hat{\omega})) = \lambda \varphi(x', \hat{\omega}) + (1 - \lambda) \varphi(x'', \hat{\omega}).$$

As we have mentioned in paragraph 0, the expectation $\varphi(x, \hat{\omega})$ is convex function.

3. Again due to the point 0, the function (11) is strictly convex. Therefore if there is a finite minimum of Φ then the minimum point is unique. \square

3. Iterative Solution of the Problem (11)

Problem (11) can be solved with the direct iterative algorithm under the assumption that functions f_1, g and h are continuous with respect to x and function $f_2(y)$ is strictly monotonically increasing in each component.

Algorithm 1

Step 0. Set some precision $\varepsilon \geq 0$ and natural number K . Starting point x^0 is chosen from some additional concepts and it is assumed that $z = x^0, t = 1, S = 0$.

Step 1. Generate stochastic vectors $\omega = \omega^t = (\omega_1^t, \dots, \omega_r^t)$ and $\hat{\omega} = \hat{\omega}^t = (\hat{\omega}_1^t, \dots, \hat{\omega}_r^t)$.

Step 2. Solve nonlinear program

$$f_1(x, \omega^t) + f_2(h^+(x, \hat{\omega}^t)) \longrightarrow \inf_{x \geq 0 \wedge g(x, \omega^t) \leq 0 \wedge \|x - z\| \leq \rho^t}, \tag{12}$$

by some method of nonlinear programming. Choice of the weights $\{\rho_t\}_1^\infty$ in (12) is performed according to axioms (13) – (15) (see below). Method of quadratic optimization [7] for the program (12) shows good efficiency. The convergence of the method can be accelerated and stabilized by the exact relaxations [8].

Step 3. If $\|x - z\| > \varepsilon$ then $\{S := 0; z := x; \text{go to Step 4}\}$

$$S := S + \|x - z\|.$$

If $S > K\varepsilon$ then output x ; exit from the algorithm.

Step 4. Assume $t := t + 1 \quad z := x$ and $z := x$. Go to Step 1.

Notes to the algorithm 1

Weight generation. Weights $\{\rho_t\}_1^\infty$ are stated using the classic scheme [1, 2, 4], according the axioms:

$$\rho_t > 0 \quad \forall t, \tag{13}$$

$$\sum_{t=1}^\infty \rho_t^2 < \infty, \tag{14}$$

$$\sum_{t=1}^n \rho_t \xrightarrow{n \rightarrow \infty} \infty. \quad (15)$$

Such weights exist (for example, one may be $\rho_t = 1/t$). On the one hand they provide a possibility to tend from starting point x^0 toward problem (11) solution independently on how far from the starting point the solution is located (axiom (15)). On the other hand from axiom (14) it follows that

$$(\forall N) i \rightarrow \infty \implies \sum_{t=i}^{N+i} \rho_t \rightarrow 0.$$

Let $\{x^t\}_{t=1}^{\infty}$ be sequence of iterations originated by algorithm without exit in the Step 3. If $\{x^t\}_{t=1}^{\infty}$ contains the subsequence $\{x^{t_i}\}_0^{\infty}$ such that it converges monotonically to a certain α and $t_{i+1} - t_i \leq N$, $i = 1, 2, \dots$, then the sequence also converges to α . Indeed, increase of the distance from x^t to α while k runs from t_i to $t_{i+1} - 1$ does not exceed

$$\sum_{k=t_i}^{t_{i+1}-1} \rho_k \leq \sum_{k=t_i}^{N+t_i-1} \rho_k \rightarrow 0 \quad i \rightarrow \infty.$$

Therefore,

$$\|x^t - \alpha\| \leq \|x^{t+i} - \alpha\| + \left\| \sum_{t=t_i}^{N+t_i} \rho_t \right\|,$$

i.e. oscillations of the sequence of elements $\{x^t\}_1^{\infty}$ with respect to α will asymptotically converge to zero. Hence the sequence does converge but not necessary monotonically. The natural condition of shift from current iterative point toward obtained minimum contains in axiom (13).

Stopping rule (for Step 3). We must have the final error of the method no greater than $\varepsilon > 0$. After t -th iteration another several steps will be performed so that their integrated length exceeds ε in K times ($K \gg 1$). In other words,

$$\sum_{j=t}^{t+N-1} |x^{j+1} - x^j| > K\varepsilon.$$

If in the process all iterations from i -th to $(i + N)$ -th remain inside full-sphere with radius ε and center x^t , i.e. $\{x^j\}_{j=t}^{N+t} \subset \mathbf{S}_{x^t}^{\varepsilon}$, then iteration process should be stopped. The last iteration point of set $\{x^j\}_{j=t}^{N+t}$ can be taken as the solution.

Theorem 3. *Let the conditions of Theorem 2 be satisfied. Then Algorithm 1 with a choice of weights according to the axioms (13) – (15) almost certain converges.*

Proof. Indeed, since transition from x^t and x^{t+1} corresponds to descent in generalized gradient direction, presented algorithm with the specified choice of the weights fall within scope of the theorem of Yu. M. Yermolyev [1], and this ensures the almost certain convergence of the algorithm. \square

4. Conclusion

The problem of perspective stochastic programming with two type random, two type variables, and two type constraints has been considered. The first type variables are selected before input of randomness. Contribution to the objective function of second type variable has not explicit dependence on the first type randomness. They are chosen then according to the information about already assigned first type variables and to the observation of implemented randomness of the second type. The first type constraints should be satisfied upon the statistical average and have the most general form. The second type constraints must be satisfied for every conceivable realization of the random parameters. Their constraint functions have general form depending on the first type parameters and the second type randomness, but only linear in the second type variables. Thus, the specification of the first type variables occurs a priori, and the second type ones occurs posterior. The proposed iterative algorithm is quite adequate to the considered problem. Its convergence has some drawbacks common to the classical direct methods of stochastic programming, such as slowing convergence with increasing index of iteration step. The advantage of the proposed method consists in an easier step by step solution of the nonlinear program (12) while other methods require the calculation of the generalized gradient with subsequent finding of its projection on the admissible set [1, 2, 3]. This method can be recommended for solving such stochastic programs where finding of generalized gradient and its projection is difficult.

References

- [1] Y. M. Ermol'ev, *Methods of Stochastic Programming*, Moscow, Nauka, 1976, In Russian.
- [2] Y. M. Ermol'ev, A. I. Yastremkii, *Stochastic Models and Methods in Economic Planning*, Moscow, PhysMatGiz, 1979, In Russian.

- [3] V. I. Norkin, Y. M. Ermol'ev, G. Fischer, Convergence of a method for computing economic equilibria, *Cybernetics and Systems Analysis*, **33**, No. 6 (1997), 854-866.
- [4] V. Kolbin, K. M. Donets, *Decision Making and Programming*, World Scientific Publishing Co., 2003, **doi:** 10.1142/5283.
- [5] V. V. Kolbin, Generalized mathematical programming as a decision model, *Applied Mathematical Sciences*, No. 69-72 (2014), 3469-3476, **doi:** 10.12988/ams.2014.44231.
- [6] S. E. Miheev, Competitiveness of similar productions, In: *2015 International Conference on "Stability and Control Processes" in Memory of V.I. Zubov, SCP 2015 – Proceedings*, art. no. 7342179 (2015), 486-487, **doi:** 10.1109/SCP.2015.7342179.
- [7] S. E. Mikheev, *Quadratic optimization*, LAP LAMBERT Academic Publishing, 2011, In Russian.
- [8] S. E. Miheev, Exact relaxation of multy point iterative methods, In: *2014 International Conference on Computer Technologies in Physical and Engineering Applications, ICCTPEA 2014 – Proceedings*, art. no. 6893311 (2014), 116-117, **doi:** 10.1109/ICCTPEA.2014.6893311.