Proc. Estonian Acad. Sci. Phys. Math., 1992, 41, 1, 1—6
https://doi.org/10.3176/phys.math.1992.1.01

UDC 519.816

Ebu TAMM *

APPROXIMATION OF AN UNCONSTRAINED MINIMUM USING BIASED ESTIMATE OF THE GOAL FUNCTION

(*Presented by G. Vainikko*)

Nonlinear stochastic programming problems usually contain functions of mathematical expectation $u(x)=Ef(x,\xi)$ and probability $v(x)=P\{f(x,\xi)\leqslant 0\}$, where $x\in R^n$ is the vector of decision variables and ξ is an s-dimensional vector of random parameters. Formally these functions are ordinary nonlinear functions of several variables. However, as they both are defined through muliiple integrals, so the calculation of their values and values of their derivatives requires voluminous computational effort, and for this reason direct application of deterministic optimization methods is possible only in very rare cases. If the distribution of ξ is not known then such methods cannot be applied at all and realizations of ξ have to be used in some way. One possibility is to approximate the unknown distribution function of ξ basing on k realizations $\xi_1, \xi_2, \ldots, \xi_k$ of ξ. In other words, the functions $u(x)$ and $v(x)$ are replaced by some statistical estimates (see e. g. [1–4]).

In [5] the solvability of the problem $\min\limits_{x\in R^n} h(x,\xi)$ was considered and the distance between its local solution $x^*(\xi)$ and a local solution x^* of the problem $\min\limits_{x\in R^n} H(x)$ was estimated under the assumption that $Eh(x,\xi)=H(x)$ for every $x\in R^n$. If $\xi_1, \ldots, \xi_k$ are k independent realizations of ξ then $E\frac{1}{k}\sum\limits_{i=1}^{k} f(x,\xi_i)=Ef(x,\xi)$, and the results of [5] were applied in [6] to the pair of problems $\min\limits_{x\in R^n} H(x)=\min\limits_{x\in R^n} Ef(x,\xi)$, $\min\limits_{x\in R^n} h(x,\eta)=\min\limits_{x\in R^n}\frac{1}{k}\sum\limits_{i=1}^{k} f(x,\xi_i)$, where $\eta=(\xi_1, \ldots, \xi_k)$. It was shown that for sufficiently large k the latter problem has a local solution $x_k(\xi_1, \ldots, \xi_k)$ with positive probability, and $\|\bar{x}-x_k(\xi_1, \ldots, \xi_k)\|$ tends to 0 in probability where $\bar{x}$ is a local minimizer of $Ef(x,\xi)$.

For the probability function $v(x)$ a quite natural estimate is $v_k(x, \xi_1, \ldots, \xi_k)=\frac{1}{kh_k}\sum\limits_{i=1}^{k}\int\limits_{-\infty}^{0} K\left(\frac{t-f(x,\xi_i)}{h_k}\right)dt$ which corresponds to the Parzen kernel-type density estimate [7] and where the function $K(\cdot)$ and the smoothing parameter h_k satisfy certain conditions. Unfortunately, $v_k(x, \xi_1, \ldots, \xi_k)$ is only asymptotically unbiased [8], and the results of [5] cannot be directly applied to establish relations between optimization problems containing probability functions and their estimate respectively. In this paper again the pair of general problems $\min\limits_{x\in R^n} H(x)$ and $\min\limits_{x\in R^n} h(x,\xi)$ is considered, but the assumption of unbiasedness is dropped.

* Eesti Teaduste Akadeemia Küberneetika Instituut (Institute of Cybernetics, Estonian Academy of Sciences). 200108 Tallinn, Akadeemia tee 21. Estonia.

In the course of this paper we assume that 1) the conditions for interchanging the order of integration and differentiation are fulfilled and 2) all functions of the parameter ξ are measurable. For the measurability of $x^*(\xi)$ see [9].

To prove the main result we need the following lemma which is a refinement of Lemma 2 in [10].

Lemma. *Let B be a Banach space and $A: B \to B$ an invertible linear operator, $\|A^{-1}\| \leqslant c_1$, $c_1 = \text{const}$. Let $r: B \to B$ be a nonlinear Frechet differentiable operator such that $r(0) = 0$, $\|r'(z)\| \leqslant c_2 \|z\|$, $c_2 = \text{const}$. If $\|a\| \leqslant 1/(4c_1^2 c_2)$ then the equation*

$$Az = a + r(z)$$

has a unique solution z^* in the ball $\|z\| < 1/(c_1 c_2)$ and

$$\|z^*\| \leqslant \frac{2c_1\|a\|}{1+\sqrt{1-4c_1^2 c_2\|a\|}}.$$

The proof is analogous to that of Lemma 2 in [10].

Corollary. *If $\|a\| \leqslant \varepsilon/c_1 - \varepsilon^2 c_2$, $0 < \varepsilon < 1/(2c_1c_2)$, then $\|z^*\| \leqslant \varepsilon$.*

Let us have two problems

$$\min_{x \in R^n} H(x) \tag{1}$$

and

$$\min_{x \in R^n} h(x, \xi), \tag{2}$$

where ξ is a random parameter.

Let the following conditions be fulfilled.

1. The problem (1) has a local solution x^*, i.e. $H(x^*) \leqslant H(x)$ in some neighbourhood $U(x^*, \delta) = \{x \mid \|x - x^*\| < \delta\}$ of x^*.
2. The function $H(x)$ is twice differentiable at x^*.
3. The Hessian $H''(x^*)$ is positively definite, i.e. for some $m > 0$ and for every $u \in R^n$ $u^T H''(x^*) u \geqslant m\|u\|^2$.
4. The function $h(x, \xi)$ is twice differentiable in x for almost every ξ in $U(x^*, \delta)$ and the Lipschitz condition $|h''_{xx}(x^1, \xi) - h''_{xx}(x^2, \xi)\| \leqslant c(\xi)\|x^1 - x^2\|$ is fulfilled, $x^1, x^2 \in U(x^*, \delta)$, where $c(\xi)$ is a random variable with finite mean $\bar{c} = Ec(\xi)$ and variance $\sigma_c = \sigma^2 c(\xi)$.
5. For $x \in U(x^*, \delta)$ the central moments $E\|h'_x(x, \xi) - Eh'_x(x, \xi)\|^2$ and $E\|h''_{xx}(x, \xi) - Eh''_{xx}(x, \xi)\|^2$ are finite.
6. The expectations $Eh'_x(x^*, \xi)$ and $Eh''_{xx}(x^*, \xi)$ are sufficiently close to $H'(x^*)$ and $H''(x^*)$, respectively, i.e. $\|Eh''_{xx}(x^*, \xi) - H''(x^*)\| \leqslant \delta_1 < m$ and $\|Eh'_x(x^*, \xi) - H'(x^*)\| \leqslant (m - \delta_1)/4\bar{c}$.

Under the conditions 1—6 the following theorem holds.

Theorem. *Let the conditions 1—6 be fulfilled. Then*

1) *the problem*

$$\min_{x \in R^n} Eh(x, \xi) \tag{3}$$

has the unique solution $\bar{x}$ in the ball $\|x - x^*\| < \dfrac{m - \delta_1}{\bar{c}}$ and

$$\|\bar{x} - x^*\| \leqslant \frac{2\|H'(x^*) - Eh_x(x^*, \xi)\|}{(m - \delta_1)\left(1 + \sqrt{1 - \bar{c}\|H'(x^*) - Eh'_x(x^*, \xi)\|/(m - \delta_1)^2}\right)};$$

2) *if for some positive constants* δ_2, δ_3, $\delta_1+\delta_2<m$, *the expression*

$$p=1-\frac{E\|h''_{xx}(\bar{x},\xi)-Eh''_{xx}(\bar{x},\xi)\|^2}{\delta_2^2}-\frac{\sigma_c}{\delta_3^2}-\frac{16E\|h'_x(\bar{x},\xi)-Eh'_x(\bar{x},\xi)\|^2(\bar{c}+\delta_3)^2}{(m-\delta_1-\delta_2)^4}$$

is positive then the problem (2) *has a solution* $x^*(\xi)$ *with probability not less than* p;

3) *if* $\|H'(x^*)-Eh'_x(x^*,\xi)\|\leqslant\frac{\varepsilon}{2}(m-\delta_1)-\frac{\varepsilon^2}{4}\bar{c}$, $0<\varepsilon<(m-\delta_1)/2\bar{c}$, *then* $P\{\xi|$ (2) *has a solution* $x^*(\xi)$ *and* $\|x^*(\xi)-x^*\|\leqslant\varepsilon\}\geqslant$

$$\geqslant 1-\frac{E\|h''_{xx}(\bar{x},\xi)-Eh''_{xx}(\bar{x},\xi)\|^2}{\delta_2^2}-\frac{\sigma_c}{\delta_3^2}-\frac{16E\|h_x(\bar{x},\xi)-Eh_x(\bar{x},\xi)\|^2}{\varepsilon^2[2(m-\delta_1-\delta_2)-\varepsilon\bar{c}]^2}.$$

Proof. Let us consider the problem (3). For a point x to be a local solution of (3) it is necessary that it satisfies the equation

$$Eh'_x(x,\xi)=0 \tag{4}$$

which is equivalent to the equation

$$Eh''_{xx}(x^*,\xi)(x-x^*)=H'(x^*)-Eh'_x(x^*,\xi)-r(x-x^*), \tag{5}$$

where

$$r(x-x^*)=Eh'_x(x,\xi)-Eh'_x(x^*,\xi)-Eh''_{xx}(x^*\xi)(x-x^*). \tag{6}$$

Due to Assumption 6 $|u^{\mathrm{T}}Eh''_{xx}(x^*,\xi)u-u^{\mathrm{T}}H''(x^*)u|\leqslant\|Eh''_{xx}(x^*,\xi)-H''(x^*)\|\|u\|^2\leqslant\delta_1\|u\|^2$, which implies $u^{\mathrm{T}}Eh''_{xx}(x^*,\xi)u\geqslant(m-\delta_1)\|u\|^2$ and therefore, $Eh''^{-1}_{xx}(x^*,\xi)$ exists and $\|Eh''^{-1}_{xx}(x^*,\xi)\|\leqslant\frac{1}{m-\delta_1}$. From the expression (6) it is easy to see that $r(0)=0$ and $\|r'(x-x^*)\|\leqslant\leqslant\bar{c}\|x-x^*\|$. Choosing $c_1=1/(m-\delta_1)$ and $c_2=\bar{c}$ it is clear that the assumptions of the Lemma are fulfilled, and so the equation (4) has a unique solution $\bar{x}$ in the ball $\|x-x^*\|<(m-\delta_1)\bar{c}$ and

$$\|\bar{x}-x^*\|\leqslant\frac{2\|H'(x^*)-Eh_x(x^*,\xi)\|}{(m-\delta_1)(1+\sqrt{1-4\bar{c}\|H'(x^*)-Eh'_x(x^*,\xi)\|/(m-\delta_1)^2})}.$$

We have proved that the necessary condition for $\bar{x}$ to be a local solution of (3) is fulfilled. Next let us show that a sufficient condition is fulfilled at the point $\bar{x}$ as well. We have

$u^{\mathrm{T}}Eh''_{xx}(\bar{x},\xi)u\geqslant(m-\delta_1)\|u\|^2$ and $|u^{\mathrm{T}}Eh''_{xx}(\bar{x},\xi)u-u^{\mathrm{T}}Eh''_{xx}(x^*,\xi)u|=$

$$=|u^{\mathrm{T}}r'(\bar{x}-x^*)u|\leqslant\|r'(\bar{x}-x^*)\|\,\|u\|^2\leqslant\bar{c}\|\bar{x}-x^*\|\,\|u\|^2\leqslant$$

$$\leqslant\frac{2\bar{c}\|H'(x^*)-Eh'_x(x^*\xi)\|}{(m-\delta_1)(1+\sqrt{1-4\bar{c}\|H'(x^*)-Eh'_x(x^*,\xi)\|/(m-\delta_1)^2})}\|u\|^2\leqslant$$

$$\leqslant\frac{2\bar{c}(m-\delta_1)^2}{4\bar{c}(m-\delta_1)(1+\sqrt{1-4\bar{c}(m-\delta_1)^2/[4\bar{c}(m-\delta_1)^2]})}\|u\|^2=$$

$$=\frac{1}{2}(m-\delta_1)\|u\|^2$$

and therefore, $u^{\mathrm{T}}Eh''_{xx}(\bar{x},\xi)u = u^{\mathrm{T}}Eh''_{xx}(x^*,\xi)u + u^{\mathrm{T}}Eh''_{xx}(\bar{x},\xi)u -$ $- u^{\mathrm{T}}Eh''_{xx}(x^*,\xi)u \geqslant (m-\delta_1)\|u\|^2 - \frac{1}{2}(m-\delta_1)\|u\|^2 = \frac{1}{2}(m-\delta_1)\|u\|^2.$

In the second part of the proof consider the solvability of the problem (2) and the distance between the local solutions of the problems (2) and (3). As the problem (2) depends on the random parameter ξ, so it may happen that for some values of ξ it has a solution $x^*(\xi)$, yet for some other values of ξ it has not. In other words, we can speek that (2) is solvable with certain probability. Let us estimate this probability. We start again from the equation

$$h'_x(x,\xi) = 0. \tag{7}$$

If for some value of ξ the point $x^*(\xi)$ is a local solution of (2), then the equation (7) is satisfied at $x^*(\xi)$. Analogously to the first part of the proof we transform (7) and obtain

$$h''_{xx}(\bar{x},\xi)(x-\bar{x}) = Eh'_x(\bar{x},\xi) - h'_x(\bar{x},\xi) - \varrho(x-\bar{x},\xi), \tag{8}$$

where

$$\varrho(x-\bar{x},\xi) = h'_x(x,\xi) - h'_x(\bar{x},\xi) - h''_{xx}(\bar{x},\xi)(x-\bar{x}).$$

According to the Lemma, the equation (8) depending on the random parameter ξ has a solution $x^*(\xi)$ if $\xi \in \mathfrak{A} = \mathfrak{A}_1 \cap \mathfrak{A}_2 \cap \mathfrak{A}_3$, where $\mathfrak{A}_1 = \{\xi \mid h''^{-1}_{xx}(\bar{x},\xi)$ exists and $\|h''^{-1}_{xx}(\bar{x},\xi)\| \leqslant d_1\}$, $\mathfrak{A}_2 = \{\xi \mid \|\varrho'_x(x-\bar{x},\xi)\| \leqslant d_2\|x-\bar{x}\|\}$ and $\mathfrak{A}_3 = \{\xi \mid \|h'_x(\bar{x},\xi) - Eh'_x(\bar{x},\xi)\| \leqslant 1/(4d_1^2 d_2)\}$, where d_1, d_2 are some constants. Following the scheme used already in the first part of the proof, we obtain that $h''^{-1}_{xx}(\bar{x},\xi)$ exists if $\|h''_{xx}(\bar{x},\xi) - Eh''_{xx}(\bar{x},\xi)\| \leqslant \delta_2$, $\delta_2 < m - \delta_1$ which implies $\|h''^{-1}_{xx}(\bar{x},\xi)\| \leqslant 1/(m-\delta_1-\delta_2)$. Therefore, using Tchebycheff inequality and choosing $d_1 = 1/(m-\delta_1-\delta_2)$, we have

$$P\{\xi \mid \xi \in \mathfrak{A}_1\} = P\{\xi \mid \|h''^{-1}_{xx}(\bar{x},\xi)\| \leqslant 1/(m-\delta_1-\delta_2)\} \geqslant$$
$$\geqslant P\{\xi \mid \|h''_{xx}(\bar{x},\xi) - Eh''_{xx}(\bar{x},\xi)\| \leqslant \delta_2\} \geqslant 1 - \frac{E\|h''_{xx}(\bar{x},\xi) - Eh''_{xx}(\bar{x},\xi)\|^2}{\delta_2^2}. \tag{9}$$

Analogously we obtain

$$P\{\xi \mid \xi \in \mathfrak{A}_2\} = P\{\xi \mid \|\varrho'_x(x-\bar{x},\xi)\| \leqslant (\bar{c}+\delta_3)\|x-\bar{x}\|\} \geqslant$$
$$\geqslant P\{\xi \mid \|h''_{xx}(x,\xi) - h''_{xx}(\bar{x},\xi)\| \leqslant (\bar{c}+\delta_3)\|x-\bar{x}\|\} \geqslant$$
$$\geqslant P\{\xi \mid c(\xi) \leqslant \bar{c}+\delta_3\} \geqslant 1 - \sigma_c/\delta_3^2, \tag{10}$$

and

$$P\{\xi \mid \xi \in \mathfrak{A}_3\} = P\left\{\xi \mid \|h'_x(\bar{x},\xi) - Eh'_x(\bar{x},\xi)\| \leqslant \frac{(m-\delta_1-\delta_2)^2}{4(\bar{c}+\delta_3)}\right\} \geqslant$$
$$\geqslant 1 - \frac{16(\bar{c}+\delta_3)^2 E\|h'_x(\bar{x},\xi) - Eh'_x(\bar{x},\xi)\|^2}{(m-\delta_1-\delta_2)^4}. \tag{11}$$

As $P\{\xi \mid \xi \in \mathfrak{A}\} = P\{\xi \mid \xi \in \mathfrak{A}_1 \cap \mathfrak{A}_2 \cap \mathfrak{A}_3\} \geqslant 1 - \gamma_1 - \gamma_2 - \gamma_3$ if $P\{\xi \mid \xi \in \mathfrak{A}_i\} \geqslant$ $\geqslant 1 - \gamma_i$, $i = 1, 2, 3$, then from the inequalities (9), (10) and (11) it follows that

$$P\{\xi \mid \xi \in \mathfrak{A}\} \geqslant p, \quad p = 1 - \frac{E\|h''_{xx}(\bar{x},\xi) - Eh''_{xx}(\bar{x},\xi)\|^2}{\delta_2^2} - \frac{\sigma_c}{\delta_3^2} - \frac{16(\bar{c}+\delta_3)^2 E\|h'_x(\bar{x},\xi) - Eh'_x(\bar{x},\xi)\|^2}{(m-\delta_1-\delta_2)^4}.$$

We have proved that with probability not less than p there exists a point $x^*(\xi)$ where the necessary condition $h'_x(x^*(\xi),\xi) = 0$ of a local minimum of (2) is fulfilled. Analogously to the first part of the proof it can be shown that for $\xi \in \mathfrak{A}$ the sufficient condition $u^{\mathrm{T}} h''_{xx}(x^*(\xi),\xi) u \geqslant M\|u\|^2$ with some constant M holds for every $u \in R^n$, and so the second statement of the theorem has been proved.

The third statement follows from the Corollary of the Lemma. Indeed, according to the Corollary, the inequality $\|H'(x^*) - Eh'_x(x^*,\xi)\| \leqslant \leqslant \frac{\varepsilon}{2}(m-\delta_1) - \frac{\varepsilon^2}{4}\bar{c}$ implies $\|\bar{x} - x^*\| \leqslant \varepsilon/2$. Similarly, if $\|h'_x(\bar{x},\xi) - Eh'_x(\bar{x},\xi)\| \leqslant \frac{\varepsilon}{2}(m-\delta_1-\delta_2) - \frac{\varepsilon^2}{4}(\bar{c}+\delta_3)$, then $\|x^*(\xi) - \bar{x}\| \leqslant \frac{\varepsilon}{2}$ and, therefore, $P\{\xi \mid$ the problem (2) has a solution $x^*(\xi)$ and $\|x^*(\xi) - x^*\| \leqslant \leqslant \varepsilon\} \geqslant P\{\xi \mid$ the problem (2) has a solution $x^*(\xi)$, and $\|x^*(\xi) - \bar{x}\| \leqslant \leqslant \varepsilon/2\} \geqslant P\{\xi \mid \xi \in \mathfrak{A}_1 \cap \mathfrak{A}_2 \cap \mathfrak{A}_{3,\varepsilon}\} \geqslant 1 - \frac{E\|h''_{xx}(\bar{x},\xi) - Eh''_{xx}(\bar{x},\xi)\|^2}{\delta_2^2} - \frac{\sigma_c}{\delta_3^2} - \frac{16E\|h'_x(\bar{x},\xi) - Eh'_x(\bar{x},\xi)\|^2}{\varepsilon^2[2(m-\delta_1-\delta_2) - \varepsilon(\bar{c}+\delta_3)]^2}$, where $\mathfrak{A}_{3,\varepsilon} = \Big\{\xi \mid \|h'_x(\bar{x},\xi) - Eh'_x(\bar{x},\xi)\| \leqslant \frac{\varepsilon}{2}(m-\delta_1-\delta_2) - \frac{\varepsilon^2}{4}(\bar{c}+\delta_3)\Big\}$.

REFERENCES

1. *Shapiro, A.* Ann. Statist., 1989, **17**, 2, 841—858.
2. *Kaňková, V.* Problems Control Inform. Theory, 1989, **18**, 4, 251—260.
3. *Dupačova, J., Wets, R.* Ann. Statist., 1988, **16**, 4, 1517—1549.
4. *Kall, P., Ruszczyński A., Frauendorfer, K.* Numerical Techniques for Stochastic Optimization. Berlin etc., Springer, 1988, 33—64.
5. *Tamm, E.* Math. Operationsforsch. Statist. Ser. Optim., 1980, **11**, 3, 487—497.
6. *Тамм Э.* Изв. АН ЭССР. Физ. Матем., 1978, **27**, 4, 448—450.
7. *Parzen, E.* Ann. Math. Statist., 1962, **33**, 3, 1065—1076.
8. *Тамм Э.* Изв. АН ЭССР. Физ. Матем., 1979, **28**, 1, 23—30.
9. *Engl, H. W.* Appl. Math. Optim., 1979, **5**, 4, 271—281.
10. *Поляк Б. Т.* Ж. вычисл. мат. и мат. физ., 1971, **11**, 1, 3—11.

Received
June 26, 1991

Ebu TAMM

KITSENDUSTETA MIINIMUMÜLESANDE APROKSIMEERIMINE SIHIFUNKTSIOONI NIHUTATUD HINNANGU ABIL

On vaadeldud diferentseeruva funktsiooni $H(x)$ minimeerimist eukleidilises ruumis R^n. On eeldatud, et juhuslikust parameetrist ξ sõltuva funktsiooni $h(x, \xi)$ gradiendi $h'_x(x, \xi)$ ja teiste tuletiste maatriksi $h''_{xx}(x, \xi)$ matemaatilised ootused on ülesande $\min_{x \in R^n} H(x)$ lokaalse lahendi x^* kohal üsna lähedal vastavalt funktsiooni $H(x)$ gradiendile ja teiste tuletiste maatriksitele selles punktis. On leitud tingimused, millal ülesandel $\min_{x \in R^n} h(x, \xi)$ eksisteerib positiivse tõenäosusega lokaalne lahend $x^*(\xi)$ ja on hinnatud tõenäosust $P\{\|x^*(\xi) - x^*\| \leqslant \varepsilon\}$.

Эбу ТАММ

АППРОКСИМАЦИЯ БЕЗУСЛОВНОГО МИНИМУМА С ПОМОЩЬЮ СМЕЩЕННОЙ ОЦЕНКИ ЦЕЛЕВОЙ ФУНКЦИИ

Рассматривается минимизация дифференцируемой функции $H(x)$ в евклидовом пространстве R^n. Пусть, кроме того, дана функция $h(x, \xi)$, зависящая от случайного параметра ξ. Предполагается, что задача $\min_{x \in R^n} H(x)$ имеет точку локального минимума x^* и математические ожидания градиента $h'_x(x^*, \xi)$ и гессиана $h''_{xx}(x^*, \xi)$ отличаются от $H'(x^*)$ и $H''(x^*)$ соответственно достаточно мало. Найдены условия, при которых задача $\min_{x \in R^n} h(x, \xi)$ с положительной вероятностью имеет точку локального минимума $x^*(\xi)$ и оценена вероятность $P\{\|x^*(\xi) - x^*\| \leqslant \varepsilon\}$.