MODEL-BASED OPTIMIZATION OF THE REGRESSION OBJECTS

Mukha V. S., Kako N. F.
Department of Information Technologies of Automated Systems,
Belarusian State University of Informatics and Radioelectronics
Minsk, Republic of Belarus
E-mail: mukha@bsuir.by, kako.nancy@gmail.com

The model-based procedure for determining of the extremum of the regression object is proposed. The numerical comparative analysis of this procedure and Kiefer–Wolfowitz stochastic approximation procedure is performed.

Introduction

Stochastic approximation and random search are most known optimization method at present for extreme objects. They are so called search methods. These approaches do not use the mathematical model of the object. In opposite to the search methods, the model-based approach suppose designing the mathematical model of the controlled object. This allows us to obtain then the mathematical model of the optimality criterion and to find the optimal control action.

I. Stochastic approximation

The stochastic approximation is a method for solving a wide range of estimation problems based on recurrent refinement of the estimate with an increase in the number of observations [1 2, 3]. The first procedure of stochastic approximation was Robbins–Monro stochastic approximation procedure, designed to find the zero of an unknown regression function. The Kiefer–Wolfowitz procedure of stochastic approximation was designed to find the extremum point of an unknown regression function. The stochastic approximation procedures are based on the sequentially estimating, from one observation to next. We consider the Kiefer–Wolfowitz procedure more detail.

Let ϕ be measurable scalar function on R^n. We suppose, that we can measure the realization of the random variable

$$y(X) = \phi(X) + \epsilon(X)$$

in any point $X \in R^n$, where the random errors $\epsilon(X)$ are centered ($E(\epsilon(X)) = 0$) and mutually independent for any $X_0, X_1, ... \in R^n$, is the regression function. The problem consist of the finding the sequence of the points $X_0, X_1, ... \in R^n$ converging to the point X^*:

$$X^* = arg \max_{X \in R^n} \phi(X).$$

The symmetric Kiefer–Wolfowitz algorithm has the following form [2]:

$$X_{k+1} = X_k + \gamma_{k+1}\alpha_{k+1}^{-1} \times$$

$$\times \sum_{i=1}^{n} \left(y(X_k + \alpha_{k+1}e_i) - y(X_k - \alpha_{k+1}e_i)\right) e_i, \qquad (1)$$

where $k = 0, 1, 2, ..., K$ and e_i are orts, i.e. coordinate vectors $e_i = (0_1, ..., 1_i, ..., 0_n)^T \in R^n$. The sequences $\gamma_k = 1/k$, $\alpha_k = 1/\sqrt[3]{k}$ provide the convergence of the Kiefer–Wolfowitz algorithm.

II. Random search

The random search is an optimization method that was proposed for finding of the extremum point of an unknown deterministic inertialess object (function). The term “random search” is attributed to Rastrigin [4]. The distinctive feature of this method is that the points are selected in the search space as random numbers from a certain probability distribution. It is considered that the random search is the most effective method for finding the global extremum. The simplest random search (so named blind scan) for minimization is as follows. The values of the function are calculated in random points and the minimum value is chosen.

The random search can be used in regression experiment for the mathematical model object building. However, the data processing is performed in such an application in package form, i.e. not in the real observation time. That is the irresistible disadvantage of the random search.

III. Model-based method

The model-based approach suppose designing the mathematical model of the controlled object. This allows then to obtain the mathematical model of the optimality criterion and to find the optimal control action. This approach is put into dual control theory [5, 6, 7]. The Box–Wilson method in the framework of the extreme experiment design [8, 9] includes the model-based approach too, but it is rather methodology than a method or algorithm, so it is difficult to implement.

It is most simple to approximate the controlled object by quadratic polynomial.

Let $X = (x_{j_{(q)}})$, $j_{(q)} = (j_1, j_2, ..., j_q)$, be a q-dimensional matrix, that is the argument of a scalar function $\phi(X)$, and this function has the form

$$y = \phi(X) = C_0 + {}^{0,q}(C_1X) + {}^{0,2q}(C_2X^2), \qquad (2)$$

where C_k, $k = 0, 1, 2, ...$, are the kq-dimensional-matrix parameters of the function $\phi(X)$ and C_2 is symmetric relative its last q-multi-indexes [10]. Let it be required to find the extremum of this function.

Optimal value of X can be found from the equation $d\phi(X)/dX = 0$. Differentiating of (2) gives the equation $C_1 + 2\,{}^{0,q}(C_2X) = 0$, from which we get

$$X^* = -{}^{0,q}({}^{0,q}(C_2^{-1}C_1))/2, \qquad (3)$$

where ${}^{0,q}C_2^{-1}$ is the matrix $(0,q)$-inverse to the matrix C_2.

Let us to find the minimum value $y^* = \phi(X^*)$ of the function $\phi(X)$. Because of the fact that ${}^{0,2q}(C_2X^2) = {}^{0,q}({}^{0,q}(C_2X)X)$ and ${}^{0,q}(C_2X^*) = -C_1/2$, we have ${}^{0,2q}(C_2(X^*)^2) = -{}^{0,q}(C_1X^*)/2$ and

$$y^* = \phi(X^*) = C_0 +^{0,q} (C_1X^*). \qquad (4)$$

Substituting X^* (3) into (4), we obtain

$$y^* = C_0 -^{0,q} (C_1\,{}^{0,q}({}^{0,q}C_2^{-1}C_1))/4.$$

We can now perform K experiments in K given points $X_1, X_2, ..., X_K \in R^n$ and find the estimations $\hat{C}_0, \hat{C}_1, \hat{C}_2$ of the parameters C_0, C_1, C_2. Then we get the following point of the extremum:

$$X^* = -\frac{1}{2}^{0,q}({}^{0,q}(\hat{C}_2^{-1}\hat{C}_1)).$$

The points of experiments $X_1, X_2, ..., X_K$ may be regular or random as in the random search approach.

IV. Computer simulation

We have performed computer check of the symmetric Kiefer–Wolfowitz algorithm. We simulated the regression object with the regression function of the form

$$\phi(X) = 1 + x_1 + 2x_2 + 2x_1^2 + 4x_2^2$$

and with normal distributed inner noise $\epsilon(X)$ with zero mean value and variance equal to 0.2. The parameters of this regression function in multidimensional-matrix form (2) are as follows:

$$C_0 = 1, \quad C_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad C_2 = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}.$$

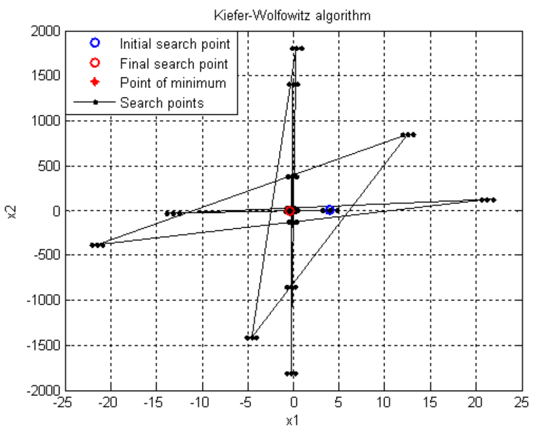

Рис. 1 – The points of experiments in the symmetric Kiefer–Wolfowitz algorithm

Figure 1 shows the search process when iterations number equal to 25 ($K = 25$ in (1)). The real minimum point of the regression function equal to $(-0.25, -0.25)$, the found point equal to $(-0.03, -0.26)$. We can see that Kiefer–Wolfowitz algorithm searches the minimum point quite well. However, the Kiefer–Wolfowitz algorithm requires very big number of experiments (125 provided $K = 25$). This makes serious difficulties for utilization it in a real production control process, since each experiment is accompanied by a product release. It can be very expensive.

Figure 2 shows the points of measurements in model-based approach with the number of the measurements equal to 25. The measurements were processed by classical least square method. The accuracy is very high.

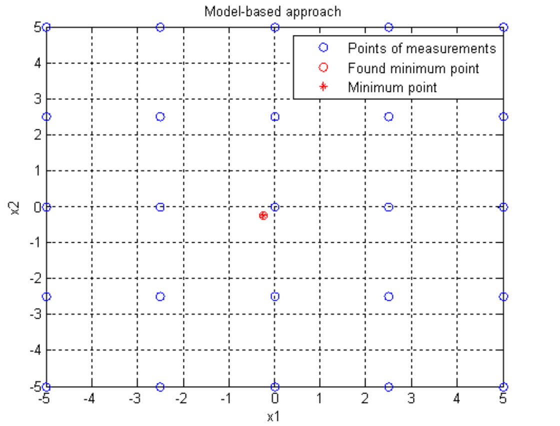

Рис. 2 – The points of experiments in the model-based approach

V. References

1. Probability and mathematical statistics: Encyclopedia. Editor-in-chief Yu.V. Prokhorov. Moscow, Big Russian Encyclopedia, 1999, 910 p. In Russian.
2. Ermakov S.M., Zhiglyavskii A.A. Mathenatical Theory of the Optimal Design. Moscow, State Press for Physico-Mathematical Literature, 1987, 318 p. In Russian.
3. Wasan M.T. Stochastic Approximation. Cambridge university press, 2004, 216 p.
4. Rastrigin L.A. Systems of extreme control. Moscow, Nauka, 1974. 630 p. In Russian.
5. Feldbaum A.A. Optimal Control Systems. Academic Press, New York and London, 1965. 452 p. Mathematics in Science and Engineering. A series of monographs and textbooks. Edited by Richard Bellman. Volume 22.
6. Mukha V.S. On the dual control of the inertialess objects. Proceedings of the LETI. Issue 130, 1973. Pp. 31–37. In Russian.
7. Mukha V.S., Sergeev E.V. Dual control of the regression objects. Proceedings of the LETI. Issue 202, 1976. Pp. 58–64. In Russian.
8. Box G.E.P., Wilson K.B. On the Experimental Attainment of Optimum Conditions. Journal of the Royal Statistical Society. 1951. Vol. 13, No. 1. Pp. 1–45.
9. Box G.E.P., Hunter W.G. and Hunter J.S. Statistics for experimenters. New York: Wiley, 1978. 653 p.
10. Mukha V.S. Analysis of multidimensional data. Minsk, Technoprint, 2004. 368 p. In Russian.