No. 932
(Received May 2, 1970)

A Heuristic Method for Maxima Searching in Case of Multimodal Surfaces

By
Katsuji UOSAKI, Hideki IMAMURA, Masao TASAKA
and Hiroshi SUGIYAMA
(Department of Applied Physics)

Abstract

Possibility of attaining global maxima has been suggested heuristically by numerical examples, even starting from a single point in our searching region, in case of multimodal response surfaces.

Our method is a modified version of Kiefer-Wolfowitz stochastic approximation process, adding artificial perturbations to the original observations, thus enabling global search.

Our method may be conducive to a variety of practical applications.

1. Introduction

Maximum searching might be one of the most interesting and important problems in the mathematical sciences, because many practical problems in engineering, science, economics, and so on, are reduced to the problem of this category.

Recently, a number of approaches on this subject have been explored[1), 2)] including random search method, steepest ascent method, simplex method, contour elimination method, etc., for deterministic case where no random factors are involved, and Box's hill climbing method and the Kiefer-Wolfowitz method for stochastic case where random variables are introduced in the observation of response values or in the process concerned.

Though, all these maximum searching methods cited above are concerned with the unimodal case primarily.

While, in the practical application of maximum searching methods, we are often concerned with the multimodal surfaces and if we use the above-mentioned searching methods, only the nearest peak can be reached, starting from a specified starting point.

In case of the multimodal case, naturally we wish to search for as many maximum points as possible, thus enabling to attain the global maximum in the searching region. We show, in this paper, possible application of a modified Kiefer-Wolfowitz processes for such global maximum searching, by numerical examples, even starting from a single starting point. Our new device here is to add zero-mean artificial noise with non-zero variance which decreases with time.

2. The Modifications of the Kiefer-Wolfowitz Stochastic Approximation Procedure

Let $f(x)$ be an unknown multimodal function of $x \in R^p$, where R^p denotes p-dimensional Euclidian space. We discuss below the searching method of global maxima of $f(x)$ sequentially, using noisy observations of $f(x)$.

2. 1 Kiefer-Wolfowitz Stochastic Approximation Procedure in Multimodal Case

The Kiefer-Wolfowitz stochastic approximation procedure (K-W procedure) was proposed by Kiefer and Wolfowitz as an extension of the Robbins-Monro stochastic approximation for maximum searching in one dimensional unimodal case and later extended by Blum to the multi-dimensional unimodal case.

We now show the extension of Kiefer-Wolfowitz procedure for multimodal case. In case of multimodal surface, there may be multi-maxima, multi-minima, as well as saddle points, and so a theoretical treatment is very complex. Extensive considerations on this subject have been carried out[3)] and will be appeared in a concrete manner elsewhere in the future. Here we shall confine ourselves to present some results of our consideration in a heuristic way.

We use observations $Y(x_n \pm c_n u_i)$ $(i = 1, \ldots, p)$ of f, where

$$Y(x) = f(x) + Z(x), \tag{2. 1}$$

$Z(x)$ is observational noise, u_i is the p-vector with j-th component δ_{ij}, and $\{c_n\}$ is positive sequence specified later. And we define the iteration procedure as follows. i. e.

$$\begin{aligned} &x_{n+1} = x_n + a_n \Delta_n \\ &x_1 \ : \ \text{arbitrary} \end{aligned} \tag{2. 2}$$

where

$$\Delta_n = \sum_{i=1}^{p} u_i \ (2c_n)^{-1} \left\{ Y(x_n + c_n u_i) - Y(x_n - c_n u_i) \right\} \tag{2. 3}$$

Now we impose in following conditions to the sequences $\{a_n\}$ and $\{c_n\}$,

$$\begin{aligned} &\text{(i)} && \Sigma a_n = +\infty \\ &\text{(ii)} && \Sigma a_n c_n < \infty \\ &\text{(iii)} && \Sigma (a_n / c_n)^2 < \infty \\ &\text{(iv)} && a_n > 0, \quad a_n \to 0 \ (n \to \infty), \qquad a_{n+1} \leq a_n \\ & && c_n > 0, \quad c_n \to 0 \ (n \to \infty) \end{aligned} \tag{2. 4}$$

and for the random noise $Z(x)$,

$$\begin{aligned} &\mathrm{E} \, \| Z(x) \| = 0 \\ &\mathrm{E} \, \| Z(x) \|^2 < \sigma^2 < \infty \end{aligned} \tag{2. 5}$$

Under the above conditions for our procedure and under certain regularity conditions on f, it is able to show

$$P\{0=\liminf_{n\to\infty}\|x_n-\eta\|<\limsup_{n\to\infty}\|x_n-\eta\|\}$$

$$+P\{0<\liminf_{n\to\infty}\|x_n-\eta\|<\limsup_{n\to\infty}\|x_n-\eta\|<\infty\}$$

$$+\sum_{l=1}^{n\lambda} P\{\lim_{n\to\infty} x_n=\lambda_l\}=0 \tag{2. 6}$$

where

$$\eta\varepsilon\,\Sigma\equiv\Theta\cup\Lambda\cup T$$

Θ: set of the all local maximum points in R^p, i. e. $(\theta_1,\dots,\theta_{n_\theta})$

Λ: set of the all local minimum points in R^p, i. e. $(\lambda_1,\dots,\lambda_{n_\lambda})$

T: set of the all saddle points in R^p, i. e. $(\tau_1,\dots,\tau_{n_\tau})$

So, taking the compliment case of this statement, it remains the possibility expressed by the following relationship,

$$\sum_{i=1}^{n\theta} P\{\lim_{n\to\infty} x_n=\theta_i\}+\sum_{j=1}^{n\tau} P\{\lim_{n\to\infty} x_n=\tau_j\}+P\{\limsup_{n\to\infty}\|x_n\|=\infty\}=1 \tag{2. 7}$$

2. 2 Kiefer-Wolfowitz Procedure with Perturbation

By the above relationship, if we delete the divergence to infinity and also the remaining possibility of convergence to a saddle point, we are able to force our process to converge to certain maximum points with probability one. Here, note that we do not assert we can force our process to converge to *all* the maxima, starting from a single starting point.

Then, we add, on the top of our observations, perturbations by an artificial noise $\xi_n(x)$ with zero-mean and non-zero variance which decreases with time and whose initial size depends upon the area of the searching region, i. e.

$$E\left(\xi_n(x)\right)=0$$

$$E\left(\xi_n(x)\right)^2\to 0,\ (n\to\infty)$$

and (2. 8)

$$E\left(\xi_1(x)\right)^2: \text{ function of the area of the searching region.}$$

Instead of $Y(x)$ in (2. 3), we use

$$Y^*(x)=Y(x)+\xi_n(x), \tag{2. 9}$$

and we proceed the same procedure described in 2. 1. Note that this modification does not contradict the conditions of the statement in 2.1, and we may search for global maxima including the original one in the sense of the relationship (2. 7).

3. Numerical Examples and Discussion

Here we show our experimental results of our modified K-W process to multi-modal maximum searching. Throughout this section, observational noise is a normal variate with mean zero and variance 0.01 and we used the following coefficients.

$$a_n = \frac{1}{n} \quad (n = 1, 2, \ldots),$$
$$c_n = \frac{1}{n^{\frac{1}{3}}} \quad (n = 1, 2, \ldots) \tag{3. 1}$$

3. 1 Result of Numerical Experiments

(*Experiment* 1) *Bimodal surface in 2–dimension.*

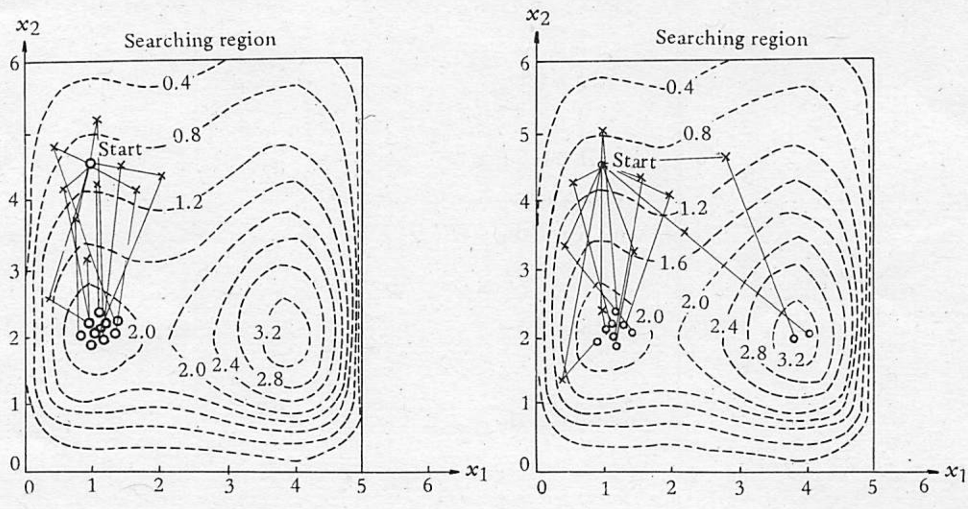
The searching area : $0 \quad x_1 \quad 5$

$$0 \quad x_2 \quad 6 \tag{3. 2}$$

The surface : $f(x_1, x_2) = (-1 + 8x_1 - 7x_1^2 + \frac{7}{3}x_1^3 - \frac{1}{4}x_1^4)(x_2^2 \exp(-x_2))$ (3. 3)

This surface has two maximum point, (1, 2) and (4, 2) and one saddle point, (2, 2) in our searching region.

By using original K-W procedure, we only attained one out of the two maxima, starting from a single starting point (1.0, 4.5) and repeating experiments shown in Fig. 1(a). Then, we used our modified K-W procedure with perturbations, adding artifi-



(a) Original K-W procedure (b) K-W procedure with perturbations

Fig. 1 Maximum points searched by original K-W procedure and proposed K-W procedure with perturbations (*Experiment* 1) cial noise $\xi_n(x_1, x_2) \in N(0, 1/n^2)$. In this case, as shown in Fig. 1(b), we attained all the maxima in our searching region.

(*Experiment* 2) *Trimodal surface in 2–dimension*

The searching region : $0 \leq x_1 \leq 4$

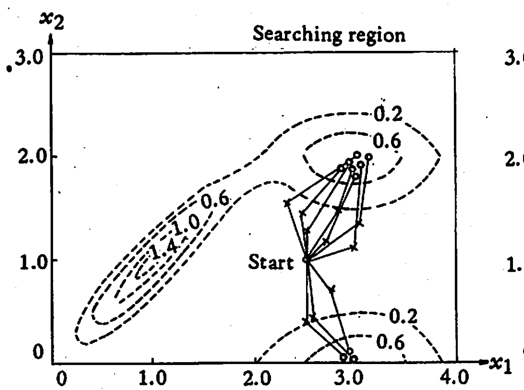
$$0 \leq x_2 \leq 3 \tag{3. 4}$$

The surface : $f(x_1, x_2) = 1.5\, x_1{}^2 \exp(1 - x_1{}^2 - 20.25\,(x_1 - x_2)^2)$

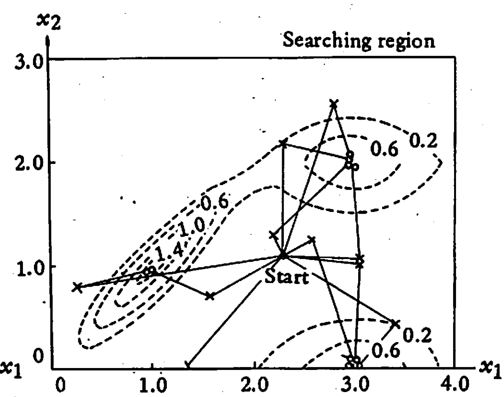
$$+ (0.5\, x_1 - 0.5)^4\, (x_2 - 1)^4 \exp(2 - (0.5\, x_1 - 0.5)^4 - (x_2 - 1)^4) \tag{3. 5}$$

In this case, there are three maxima and one saddle point in the searching region.

Though it was unable to attain all the maxima in the searching region by the original K-W procedure as shown in Fig. 2(a), all the maxima were attained by our modified K-W procedure with perturbations $\xi_n(x_1, x_2) \in N(0, 1/n^2)$, as shown in Fig. 2(b).



(a) Original K-W procedure



(b) K-W procedure with perturbations

Fig. 2 Maximum points searched by original K-W procedure and proposed K-W procedure with perturbations (*Experiment* 2)

In Table 1, number of steps to the maxima by Experiment 2 is shown. We observe that the speed of convergence does not change very much. It is conceivable that this result is due to the rapid decreasing of the perturbations.

Table 1. The number of steps to reach the maxima

	original K-W procedure	proposed K-W procedure with perturbations
Mean (10 experiments)	55.5	56.5
Variance (10 experiments)	1044	760

3. 2 Comparison with Other Methods for Maximum Searching

Using random search method and Box's hill climbing method with random starting points, we made the searching experiments for the maximum points of the same surface listed in (3. 3).

The results of random search method are shown in Fig. 3. We observe that the searched points are close to all of the maximum points. But the convergence rate is not so good, as shown in Fig. 4, comparing with our modified K-W procedure with perturbations. This is due to the difference of utilization of the information concerning our surface.

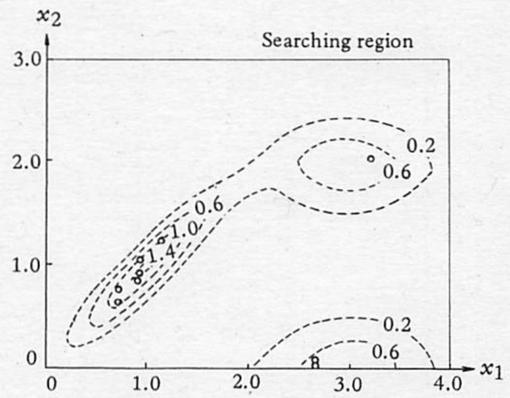


Fig. 3 Maximum points searched by random search method

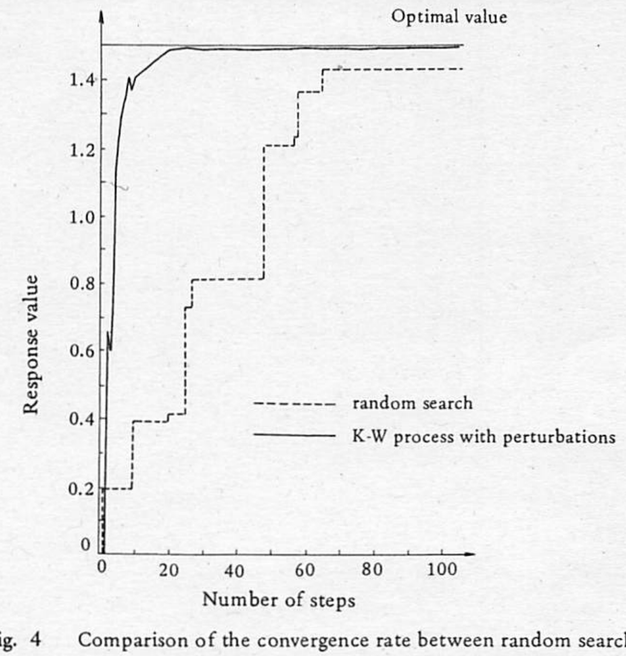


Fig. 4 Comparison of the convergence rate between random search method and proposed K-W procedure with perturbations

In Fig. 5, results of the maximum searching by Box's hill climbing method with random starting points are presented. Though all the maxima were also searched by this method, there were many failure cases in estimating maximum point. However we consider that this method may be useful for acceleration of convergence when used on the top of other methods, such as our proposed procedure.

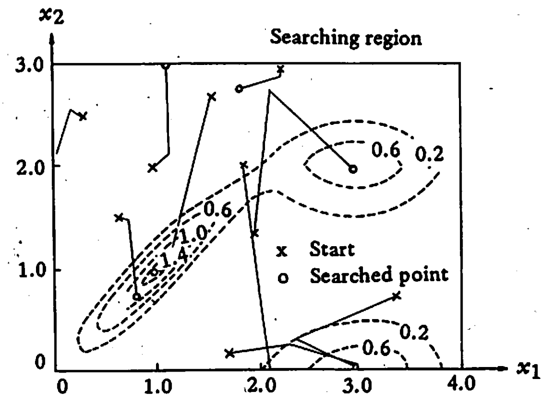


Fig. 5 Maximum points searched by Box's hill climbing method with random starting point

4. Conclusion

We have shown the possibility of attaining all the maxima in the searching region without affecting the speed of convergence by our modified Kiefer-Wolfowitz process with perturbations, starting from a single starting point and repeating the experiments. After attaining all the local maximum with our searching procedure, we can select the global maximum in the region. It is noted there is still great room for improvement of our procedure for the acceleration of convergence in case of response surface of higher dimension.

In closing, we would like to mention an application of this procedure. Some of us applied this procedure to automatic lens design and the results will be published elsewhere[4].

References

1) D. J. Wilde, "Optimum Seeking Methods", Prentice-Hall Inc., New Jersey (1964)

2) J. Kowalik and M. R. Osborne, "Methods for Unconstrained Optimization Problems", Elsevier, New York (1969)

3) S. Kusano et al., (unpublished paper),
M. Tasaka, *Osaka Tokei Danwakai Hokoku* (1969) (*in Japanese*)

4) H. Imamura, K. Uosaki, M. Tasaka and T. Suzuki, Preprints of IFAC Symposium on Systems Engineering Approach to Computer Control, Kyoto (1970)

5) I. Kita, K. Uosaki, M. Tasaka and H. Sugiyama, 13th JAACE Convention Record, 110, Osaka (1969) (*in Japanese*)