Combined application of drawing and steepest descent in generating initial estimates for subsequent optimization

By K. VARGA and P. FEJES

Introduction

The determination of the unrestricted local minimum of a function of several variables by the "direct search" methods consists in the sequential examination of function values belonging to randomly selected vectors as the independent variables. A comparison of each trial solution with the "best" one up to that time helps locate the approximate value of the minimum. Even the more sophisticated gradient methods [1, 2, 3] cannot dispense with a similar method for generating initial estimates for subsequent iterative optimization. In certain cases (e.g. in the Gauss—Newton method) a favourably chosen initial value is a prerequisite of convergence, whereas in others it allows the gradient method to be used for finding the *absolute* minimum of a function in a bounded region. Finally, the number of iterations can be reduced considerably if a good initial estimate of the minimum is available.

We have found that the effectiveness of this method can be substantially improved if the initial estimate is chosen not as the vector corresponding to the lowest function value but, starting with the drawn vector, a step is performed according to the principles of the steepest descent and the function values are compared *in these modified points*.

Both methods are described from the viewpoint of probability, and some problems of application encountered in practice are discussed.

Statistical basis of the direct search method

For the sake of simplicity let us consider the case of minimization. Let the function to be optimized be the scalar-vector function

$$f(p_1, p_2, \ldots, p_n) \equiv f(p)$$

assumed to be continuous and single-valued in the bounded n-dimensional rectangle $T \subset E_n$. Let us suppose further that function $f(p)$ exhibits a minimum in an *inner* point $p_{\min}$ of the region.

In order to find the point $p_{\min} \in T$ that makes the function to attain its minimum, let us select some optimizing method which unambiguously determines a T_{conv} (eventually multiply interrelated) parameter interval characterized in such a way that any point of this interval could serve as initial value for the method to yield $p_{\min}$, in short: *the method is convergent.* By using the usual set theory notation

$$T_{\text{conv}} \equiv \{p;\ \text{from which the gradient method converges to } p_{\min}\}.$$

If the initial value is selected from the $(T - T_{\text{conv}})$ region, the method "gets stuck" in local minima or some other special points, i.e. it is *divergent.* In other words, this region contains those points of T where such unfavourable properties of an "ill-conditioned" function may be experienced as e.g. saddle points or, if the Hessian matrix is used, points where the Hessian is singular or not positive definite.

If the problem is solved by successive iteration, a preliminary knowledge of some initial value $p_1^* \in T$ is an important condition. Should no other information be available, drawing is to be used for selecting an appropriate p_1^*. Let us suppose that a process is available for drawing n numbers at random with equal probability to represent the components of p_1^*. We shall now examine the probability of drawing a "good" point that lies within T_{conv}.

Let $\mu(X)$ be the Lebesgue measure of some region $X \subseteq T \subset E_n$; then the probability of finding a point within T_{conv} is given by

$$P(p_1^* \in T_{\text{conv}}) = \mu(T_{\text{conv}})/\mu(T) \equiv \varrho_{\text{conv}}.$$

The probability of obtaining a good initial value can be increased by making several independent drawings one after another. The probability that at least one vector from among the set: $p_1^*, \dots, p_m^*$ chosen independently lies within T_{conv} can be written according to the binomial distribution as

$$P_1(m) \equiv P(\text{for at least one } i\text{: } p_i^* \in T_{\text{conv}};\quad i = 1, 2, \dots, m) =$$

$$= 1 - \binom{m}{0} \varrho_{\text{conv}}^0 (1 - \varrho_{\text{conv}})^m = 1 - (1 - \varrho_{\text{conv}})^m. \tag{1}$$

The number of drawings to be made if we want to find at least one point in T_{conv} with a predetermined probability P is

$$m_1'(P) = \log(1 - P)/\log(1 - \varrho_{\text{conv}}) \tag{2}$$

whence the number of drawings is given by

$$m_1(P) = \text{entier}\,[m_1'(P)] + 1$$

$m_1'(P)$ being not an integer number.

Now, provided that ϱ_{conv} is known, in case of some certainty we can perform in principle as many drawings as are necessary for determining at least one suitable initial value. The only problem is: which one of the $m_1(P)$ vectors belongs to T_{conv}?

The generation of initial parameters by drawing for minimization is an evident yet not widely applied possibility just because it is difficult to give a reliable answer to the above question. In practice there is no other possibility of selection from the vectors drawn than to compare the pertaining function values. This method, however, is not reliable because the vector resulting in the lowest function value often does not belong to T_{conv}. Therefore a comparison of this kind, as a method of distinction, might be misleading.

In order to evalute this error characteristic of the direct search methods, let us consider now the probability that the vector with the lowest function value in a set of m vectors is an element of T_{conv}.

Let A_k be the event that k of the selected m points are elements of T_{conv} and introduce the following notation

$$P_a(m, k) \equiv P(\min_{\substack{i \\ p_i^* \in T_{\text{conv}}}} f(p_i^*) < \min_{\substack{j \\ p_j^* \notin T_{\text{conv}}}} f(p_j^*)|A_k).$$

Further on, let p^* be the vector for which

$$f(p^*) \equiv \min_{i=1,2,\ldots,m} f(p_i^*).$$

Now, considering that upon drawing, the probability of finding k points in T_{conv} is

$$P_b(m, k) \equiv P(A_k) \equiv \binom{m}{k} \varrho_{\text{conv}}^k (1-\varrho_{\text{conv}})^{m-k}$$

the probability we are looking for can be expressed as

$$P^*(m) \equiv P(p^* \in T_{\text{conv}}) = \sum_{k=1}^{m} P_a(m, k) \cdot P_b(m, k) =$$

$$= \sum_{k=1}^{m} P(\min_{i=1,2,\ldots,k} f(p_i^*) < \min_{j=k+1,\ldots,m} f(p_j^*)) \binom{m}{k} \varrho_{\text{conv}}^k (1-\varrho_{\text{conv}})^{m-k}. \tag{3}$$

This equation is, however, unsuitable for practical calculations. But even so, it reflects the uncertainty involved in the comparison of function values. A simple consequence of the equation is for example the inequality

$$P^*(m) \leqq P_1(m), \tag{4}$$

which follows from

$$P^*(m) \leqq \sum_{k=1}^{m} \binom{m}{k} \varrho_{\text{conv}}^k (1-\varrho_{\text{conv}})^{m-k} = 1-(1-\varrho_{\text{conv}})^m = P_1(m).$$

This inequality implies that a comparison of the function values provides absolute certainty for the selection of a point in the convergence region (provided that such a point is contained in the set) only under ideal conditions (i.e. if the function to be optimized has appropriate characteristics), whereas in other cases the comparative technique may impair the efficiency of the search by drawing. This means that more than $m_1(P)$ drawings should be carried out in order to be able to single out a point which is an element of T_{conv} with probability P. Although no certain distinction is possible between the points belonging to T_{conv} and others (e.g. those lying in the vicinity of a local minimum), generally there always exists a region $T_{\min} \subset T$, $\mu(T_{\min}) \neq 0$, where the selection on the basis of function values leads to correct results. In other words, there exists a region where no mistake arises, and this is nothing else but the largest neighbourhood around $p_{\min}$ for all the points of which $f(p') < f(p'')$ is true, unless $p' \in T_{\min}$ and $p'' \in (T-T_{\min})$. The exact definition of $T_{\min}$ is

$$T_{\min} \equiv \{p; f(p) < \min_{p \in (T-T_{\text{conv}})} f(p)\}. \tag{5}$$

It follows from the definition that $T_{\min} \subseteq T_{\text{conv}}$.

The probability of finding a point in $T_{\min}$ among m drawn vectors by comparing the function values can be calculated by the binomial distribution; for the selection does not lead to error in this region

$$P_2(m) \equiv P(p^* \in T_{\min}) = 1-(1-\varrho_{\min})^m, \tag{6}$$

using for $\varrho_{\min}$ the simple relation

$$\varrho_{\min} \equiv \mu(T_{\min})/\mu(T).$$

Finally, considering Eq. (5), it can be proved that

$$P_2(m) \leqq P^*(m). \tag{7}$$

The probabilities $P_1(m)$ and $P_2(m)$ are easy to calculate for any value of m and through definitions (4) and (7), they determine the lower and upper bounds of the probability function $P^*(m)$

$$1-(1-\varrho_{\min})^m \leqq P^*(m) \leqq 1-(1-\varrho_{\text{conv}})^m. \tag{8}$$

The inverse $m(P)$ of the function $P^*(m)$ shows how many drawings are to be carried out to get one point lying in the convergence region in case of a predetermined certainty P. Although the inverse function cannot be calculated, an estimate of $m(P)$ can be made using Eq. (8)

$$m_1(P) \leqq m(P) \leqq m_2(P),$$

where

$$m_2(P) = \text{entier}\,[m_2'(P)]+1$$

and

$$m_2'(P) = \log(1-P)/\log(1-\varrho_{\min}), \tag{9}$$

respectively which follows from (6) by analogy to Eq. (2).

The upper and lower bounds defined in this way are, unfortunately, far from each other because in reality $\mu(T_{\min})$ is several orders of magnitude smaller than $\mu(T_{\text{conv}})$, consequently

$$\varrho_{\min} \ll \varrho_{\text{conv}}.$$

On the other hand, from Eqs. (2) and (9)

$$m_2'(P)\cdot\log(1-\varrho_{\min}) = m_1'(P)\cdot\log(1-\varrho_{\text{conv}}).$$

Now considering that the function $x\log(1-x)$ is negative and monotoneously decreasing in the interval (0, 1), we get

$$\frac{m_2'(P)}{m_1'(P)} = \frac{\log(1-\varrho_{\text{conv}})}{\log(1-\varrho_{\min})} \geqq \frac{\varrho_{\text{conv}}}{\varrho_{\min}},$$

therefore

$$m_2(P) \approx m_2'(P) \gg m_1'(P) \approx m_1(P).$$

The modified method. Search from random sets modified by a step of steepest descent

From the above considerations we conclude that a comparison of the function values belonging to a number of $m_2(P)$ vectors leads to $p_1^* \in T_{\min}$. The smaller set $m_1(P)$ also contains at least one vector belonging to T_{conv}, but we cannot find it owing to *the lack of a perfect method for selection.*

Since the comparison of function values as a principle of selection cannot be replaced with anything else, the error can be reduced only by generating a point within $T_{\min}$. Such a point could be made available only if an $m_2(P)$ number of drawings had been performed. This task becomes especially hard when practically nothing is known about $\varrho_{\min}$; e.g. in the case when, upon increasing the number of drawings from 100 to 1000, still no $p^*_{\min} \in T_{\min}$ can be expected with certainty. We can get out of this apparent deadlock by not generating the parameter vector in $T_{\min}$ *by simple drawing.*

Let us carry out one drawing, then modify this point by moving along the direction-grad $f(p)|_{p_1^*}$ until a minimum of the function at p_1^{**} on this line is found. The number of drawings and searches be altogether $m_1(P)$. It is a practical experience, which can be proved for a number of functions also theoretically, that if $p_i^* \in T_{\text{conv}}$, then $p_1^{**} \in T_{\min}$ is also valid. Therefore, by applying this strategy, an erroneous decision is practically out of the question.

This statement has been verified in many practical applications. In parameter estimations and also in the case of the test functions to be shown later it has been found that the probability $P^{**}(m) \equiv P(p^{**} \in T_{\text{conv}})$ by far exceeded $P^*(m)$. The vector p^{**} is that for which

$$f(p^{**}) = \min_{i=1,2,\dots,m} f(p_i^{**}).$$

As a consequence, if the random vectors are modified by a search for the minimum along the gradient direction, it is sufficient to make only $m_1(P)$ drawings corresponding to the lower bound in Eq. (9).

Examples

We have succesfully used the modified method of generating initial values for optimization in the determination of rate constants in reaction kinetics. The function to be optimized was the sum of squares function

$$f(p) = \sum_{i=1}^{t} (y'_i - y_i(p))^2$$

where

$$y' \equiv (y'_1, y'_2, \dots, y'_t)$$

stands for the experimental data and

$$y(p) \equiv (y_1(p), y_2(p), \dots, y_t(p))$$

for the response function. The values of the response function $y(p)$ in kinetic work can be obtained only after laborious calculations involving expansion, numerical integration, etc. For optimization of the sum of squares functions, we have used the Fletcher—Powell [4] method and a procedure we developed by modifying the Newton—Gauss type of iteration [5]. The modified drawing, when combined with one of the known gradient methods, is well suited according to our experiences for generating initial estimates in practical optimization procedures [5, 6]. Nevertheless, this paper should be confined to a lesser job, i.e. to illustrate the application of the method on two test functions.

In Table I, the results obtained with two *test functions* exhibiting several local minima in the parameter-space are shown. The absolute minimum for both functions lies at $p_{min}=(0, 0)$, where $f(p_{min})=0$. The functions themselves, and the parameter and the convergence regions, expected on the basis of the analytical properties of the functions, are specified in the first part of the table. In the second part of the table are listed the lower and upper bounds evaluated from probability functions (1) and (6), as well as the relative frequencies for both standard and modified drawings obtained as a result of several hundred computer runs. The values of $\varrho_{min,I}=0.01$ and $\varrho_{min,II}=0.13$ used for the computation of the lower bounds have been estimated from the analytical properties of functions I and II, respectively. The upper bounds have been calculated not from the trivial measure of the convergence region defined in the table but from the relative frequencies found for single drawings, making use of the definition of T_{conv}, i.e. putting $\mu(\varrho_{conv})=P^*(1)=P^{**}(1)$.

I. Table

1. Optimizations	I.	II.
test functions	$f(p) = (25-p_1^2)\sin^2 \Pi p_1+p_2^2$	$f(p) = p_1^2((p_2^2-4)^2+1)+p_2^2$
parameter region	$T = \{p;\lvert p_1\rvert \leqq 4.5, \lvert p_2\rvert \leqq 1.5\}$	$T = \{p;\lvert p_1\rvert \leqq 2.5, \lvert p_2\rvert \leqq 2.5\}$
convergence region	$T_{conv} \cong \{p;\lvert p_1\rvert<0.5, \lvert p_2\rvert<1.5\}$	$T_{conv} \cong \{p;\lvert p_1\rvert<1.2, \lvert p_2\rvert<2.5\}$

2. Probability of convergence depending on number of drawings

number of drawings	1	3	6	9	12	1	2	4	6	8
lower limit	0,010	0,030	0,059	0,087	0,114	0,130	0,243	0,427	0,566	0,672
simple drawing	0,137	0,35	0,50	0,63	0,68	0,550	0,57	0,74	0,84	0,88
modified drawing	0,137	0,35	0,55	0,71	0,83	0,550	0,78	0,92	0,97	1,00
upper limit	0,137	0,357	0,587	0,735	0,843	0,550	0,798	0,959	0,992	0,998

By comparing the lines of the table, it becomes obvious that the modified drawing, as expected, is more efficient than the simple one. Also, the relative frequencies for modified drawings are in good agreement with the theoretical maxima, thus the method seems to be suitable for the elimination of the error involved in simple selection.

A similar result has been obtained in multiparameter fits. The illustration, however, would be more complicated in this case, owing to the features of the sum of squares functions mentioned above and the excessive computer time needed for setting up a similar table.

Though the modified method practically eliminates the error in the comparison of function values, the problem of generating initial estimates for optimization is far from being solved. As yet, no satisfactory answer has been found to the main question: how many drawings have to be made in a given case (with or without modification of the vectors drawn) in order to obtain good initial values. No $m(P)$ inverse can be given for the theoretical function in Eq. (3), and the limits $m_1(P)$ and $m_2(P)$, which might be of theoretical interest and were applied successfully in this work, cannot be calculated in practical problems. (It is easy to show that the computation of ϱ_{conv} and ϱ_{min} would be a more complex problem than finding the optimum itself.) The only definite statement which can be made is that the number of drawings needed to assure convergence in the modified search is always of a lower order of magnitude than that needed in the direct search. In practice it proved to be a good strategy to try to find the initial value by modified search from as many drawings as there are parameters involved and to repeat the whole of optimization in case of divergence.

Summary

When one tries to determine the unrestricted local minimum of a function of several variables by an iterative algorithm, it frequently happens that the algorithm is successful only if a sufficiently good estimate of the starting vector can be provided. Authors consider the following process: generate n random vectors, and apply one iteration of the steepest descent method for each of them; select as starting vector for subsequent optimization one that yields the least function value. The paper deals with the probability theory foundation of the modified drawing method, and with the discussion of the experiences of its application. It is proved that this strategy enhances the probability of convergence in practical optimization procedures.

Совместное применение разыгрыша и метода "steepest descent" в задаче определения начальных значений для дальнейшей оптимизации

Если локальный минимум функции от нескольких переменных нужно определить итеративным алгоритмом, тогда операция в большинстве случаев только в том случае удачная, если можно предписать относительно хорошие начальные значения со стороны переменных. Авторы и предлагают следующий метод для определения таких начальных значений:

после генерации n случайных векторов, исходя из них, осуществляем по одной итерации методом "steepest descent" и из полученных векторов тот нужно выбрать для исходного значения, который является меньшим значением функции. Статья занимается новым методом, как теоретическим сформулированием задачи теории вероятностных исчислений, и дает результат практических опытов, которые доказывают, что такая стратегия в действительности увеличивает вероятность конвергенции метода оценки параметров.

INSTITUTE OF RADIOCHEMISTRY
JÓZSEF ATTILA UNIVERSITY
SZEGED, HUNGARY

Literature

[1] RALSTON, A., H. S. WILF, *Mathematical methods for digital computers*, John Wiley and Sons, Inc., 1960.
[2] LINNIK, YU. V., *Methods of least squares and principles of the theory of observations*, Pergamon Press, 1961.
[3] BARD, Y., L. LAPIDUS, *Catalysis Reviews*, v. 2, 1968, p. 67.
[4] FLETCHER, R., M. J. D. POWELL, *Comput. J.* v. 6, 1963, p. 163.
[5] VARGA, K., *Acta Chim. Acad. Sci. Hungar.* v. 73, 1972, p. 411.
[6] FEJES, P., P. KÖNIG, *Acta Chim. Acad. Sci. Hungar.* v. 71, 1972, p. 217.

(Received May 16, 1972)