Talanta 148 (2016) 641–648

Contents lists available at ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Optimization using the gradient and simplex methods

Víctor Cerdà [a,*], Juan Luis Cerdà [b], Abubakr M. Idris [c]

[a] Chemistry Department, University of the Balearic Islands, 07122 Palma de Mallorca, Spain
[b] Department of Applied Mathematics and Analysis, University of Barcelona, Spain
[c] Department of Chemistry, College of Science, King Khalid University, Abha 61321, Saudi Arabia

ARTICLE INFO

Article history:
Received 11 February 2015
Received in revised form
21 May 2015
Accepted 25 May 2015
Available online 12 June 2015

Keywords:
Multivariate optimization
Simplex
Maximum slope
Gradient

ABSTRACT

Traditionally optimization of analytical methods has been conducted using a univariate method, varying each parameter one-by-one holding fixed the remaining. This means in many cases to reach only local minima and not get the real optimum. Among the various options for multivariate optimization, this paper highlights the gradient method, which involves the ability to perform the partial derivatives of a mathematical model, as well as the simplex method that does not require that condition.

The advantages and disadvantages of those two multivariate optimization methods are discussed, indicating when they can be applied and the different forms that have been introduced. Different cases are described on the applications of these methods in analytical chemistry.

1. Introduction

Optimization is usually a mandatory step when working in experimental sciences and engineering, covering from mathematical functions and industrial processes to new analytical methods.

Despite its familiarity and ease of use, univariate optimization process is time- and reagent-consuming, besides it is unable to consider interaction effect between conditions. Hence, the maximum efficiency of analytical methods might not be obtained. This because univariate is based on optimizing conditions one-by-one by varying levels of one condition while levels of other conditions are held constant. The more optimization effective approach is multivariate, which optimizes all conditions simultaneously by varying levels of all of them. Accordingly, multivariate optimization could obtain the highest efficiency of analytical methods in the shortest time period.

When optimizing a magnitude or a mathematical function, e.g. $U= f(x,y,\ldots,z)=f(X)$, that depends on one or more variables or factors, $X=(x,y,\ldots,z)$, the function is called the target function and the values of the variables giving an optimum value (maximum or minimum) for the target function have to be determined. Thus, a maximum will be the aim when for example optimizing the efficiency of a reaction, the maximum signal of an instrument and the maximum difference between a signal and the instrumental noise. On the contrary, a minimum will be desirable if the aim is to minimize for example interferences or an experimental error. In other cases, targeted values are desirable to be obtained like recovery (100%) in sample treatment and resolution in separation techniques with maintaining short analysis time.

We are going to focus on how to optimize a minimum. However, if a maximum of a function $f(X)$ is the aim, we just have to minimize the opposite function, i.e. $U=-f(X)$.

Although the main aim is to find absolute maxima and minima, mathematical methods frequently provide relative or local maxima and minima depending on the starting point (Fig. 1). Therefore, it is advisable to assure that the maximum or minimum found are the absolute ones repeating the calculation starting from different initial points.

There are different strategies to obtain the optimum values for different optimization cases, which may be simultaneous (e.g. Gradient, Simplex and Evolutionary Operation) [1–3] or sequential (e.g. Box–Behnken, Central Composite, Doehlert and Factorial Design) [3]. The calculation method has to be chosen according to each system [1,4]. Below are described some sequential methods that can efficiently meet most common optimization cases in experimental sciences. The objective function has to be classified in those having derivatives and those not having. Thus, there are two different options,

1. The gradient method is recommended for functions with several variables and obtainable partial derivatives, using a variation of the Newton, Davidon, Fletcher and Powell methods
2. For functions with several variables and unobtainable partial derivatives, the simplex method is then the best option.

* Corresponding author. Fax: +34 971173426.
E-mail address: victor.cerda@uib.es (V. Cerdà).

http://dx.doi.org/10.1016/j.talanta.2015.05.061
0039-9140/

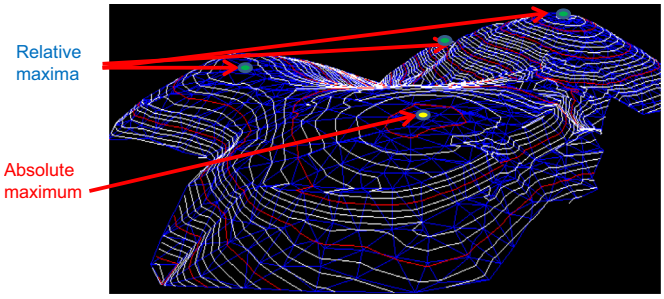

Fig. 1. Level curves with some local maxima and an absolute maximum.

All cited these are numerical methods, which are based on refining initial values. These initial values are estimated according to the knowledge of the particular problem, e.g. when estimating the pK of a functional group, or by values obtained with other calculation methods of approximation, either graphical or analytical ones.

If possible, it is advisable to take advantage of algorithms using derivatives to achieve a better reliability and rapid convergence to the optimum (gradient method), avoiding direct search despite their simplicity (simplex method).

The aim of this review is to provide a general knowledge on the gradient and simplex methods in the matter of optimization in the experimental sciences, avoiding an excessive use of mathematical language, and providing an intuitive justification of the proposed methods. More specialized texts in the matter can be found in the literature [5–7]. A good introduction to numerical methods of analysis can be found in bibliography [8]. Below are described a selection of useful methods to apply in the experimental sciences field to achieve an optimum.

2. The gradient method

In the "gradient", "maximum slope" or "steepest-ascent" method, a level set, or a level curve in the case of having two variables of $U=f(X)$, is a set of points where the studied function has always the same value. Furthermore, not only U does not vary in this set, but also it is easily demonstrated that the variation of U respect to the parameters is the maximum in the direction perpendicular to the level set. This direction is determined by the gradient vector $G(X)$ (Eq. (1)), whose components are first partial derivatives of function "f". It has to be taken into account that the gradient vector size shouldn't be constant, otherwise the maximum won't be reached, it will be surpassed (Fig. 2a). The criterion to be followed is to reduce the gradient vector size as the slopes do so (Fig. 2b).

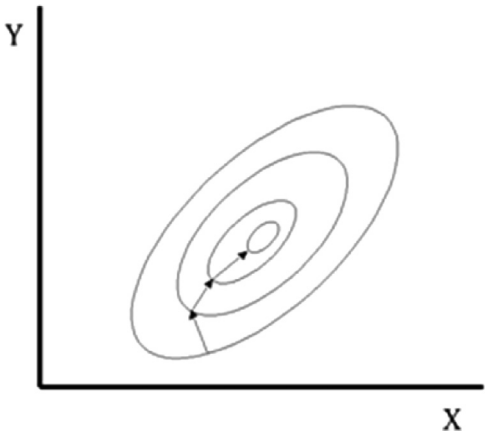

Fig. 3. The gradient method for two variables.

$$G(X) = (f_x(X), \ldots, f_z(X)), \quad (1)$$

In the case of having two variables affecting the target function, if we follow the gradient direction, i.e. a way perpendicular to the level curves (Fig. 3), we will be following the curve with maximum slope. Thus, we will pass from point (X_m) to the next (X_{m+1}) considering Eq. (2) for maximizing, and Eq. (3) for minimizing, being the step "k" always $k>0$. It will be clearer if the points X_m, X_{m+1}, etc. are presented in Fig. 4.

$$X_{m+1} = X_m + k\ G(X_m) \quad (2)$$

$$X_{m+1} = X_m - k\ G(X_m) \quad (3)$$

2.1. Examples of application of the gradient method

The gradient method has been widely used in programs to refine thermodynamic parameters, such as equilibrium constants, reaction enthalpies and entropies, where the sum of the residuals squares is minimized (Table 1).

All programs listed in Table 1 need initial parameters values. As closer these initial values to the optimum, as faster will be the convergence to it and the risk to diverge will also be lower. However, the chemical criterion has to be always in mind, otherwise absurd results will be obtained, e.g. to obtain three constants of protonation for a diprotic acid because the residuals value is lower.

With the last four programs shown in Table 1, two different processes can be applied to reach the optimum, i.e. a process

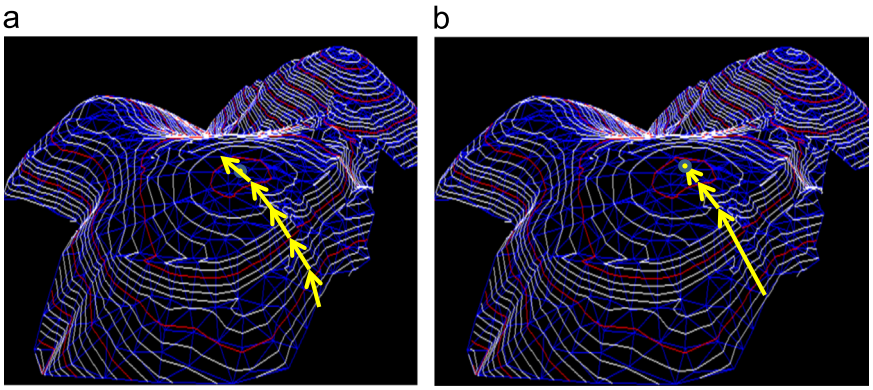

Fig. 2. The gradient method with (a) constant vector size and (b) variable vector size.