

An Iterative Method for Curve Adjustment Based on Optimization of a Variable and its Application

Un método iterativo para el ajuste de curvas basado en la optimización en una variable y su aplicación al caso lineal en una variable independiente

ROGELIO ACOSTA^{1,a}, SUITBERTO CABRERA^{2,b}, LUIS MANUEL VEGA^{1,c},
ASELA CABRERA^{2,d}, NERSA ACOSTA^{3,e}

¹DEPARTAMENTO DE MATEMÁTICA, UNIVERSIDAD DE LAS TUNAS, LAS TUNAS, CUBA

²DEPARTAMENTO DE ESTADÍSTICA, INVESTIGACIÓN OPERATIVA APLICADAS Y CALIDAD, UNIVERSITAT POLITÈCNICA DE VALENCIA, VALENCIA, SPAIN

³UNIVERSIDAD DE LAS CIENCIAS INFORMÁTICAS, LA HABANA, CUBA

Abstract

An iterative method for the adjustment of curves is obtained by applying the least squares method reiteratively in functional subclasses, each defined by one parameter, after assigning values to the rest of the parameters which determine a previously determined general functional class. To find the minimum of the sum of the squared deviations, in each subclass, only techniques of optimization are used for real functions of a real variable. The value of the parameter which gives the best approximation in an iteration is substituted in the general functional class, to retake the variable character of the following parameter and repeat the process, getting a succession of functions. In the case of simple linear regression, the convergence of that succession to the least squares line is demonstrated, because the values of the parameters that define each approximation coincide with the values of the parameters obtained when applying the method of Gauss - Seidel to the normal system of equations. This approach contributes to the teaching objective of improving the treatment of the essential ideas of curve adjustment, which is a very important topic in applications, what gives major importance to the optimization of variable functions.

Key words: Curve estimation, Iterative method, Least Squares Method, Linear regression, Teaching materials.

^aProfessor. E-mail: racosta@ult.edu.cu

^bProfessor. E-mail: suicabga@eio.upv.es

^cProfessor. E-mail: racosta@ult.edu.cu

^dProfessor. E-mail: ascabul@eio.upv.es

^eProfessor. E-mail: ndacosta@uci.cu

Resumen

Se obtiene un método iterativo para el ajuste de curvas al aplicar reiteradamente el método de los mínimos cuadrados en subclases funcionales, cada una definida por un parámetro, luego de asignar valores a los restantes parámetros que determinan una clase funcional general, seleccionada previamente. Para hallar el mínimo de la suma de las desviaciones cuadráticas, en cada subclase, solo se utilizan técnicas de optimización para funciones reales de una variable real. El valor del parámetro, que proporciona la mejor aproximación en una iteración, se sustituye en la clase funcional general, para retomar el carácter variable del siguiente parámetro y repetir el proceso, obteniéndose una sucesión de funciones. En el caso de la regresión lineal simple se demuestra la convergencia de esa sucesión a la recta mínimo cuadrática, pues coinciden los valores de los parámetros que definen cada aproximación con los que se obtienen al aplicar el método de Gauss - Seidel al sistema normal de ecuaciones. Este enfoque contribuye al objetivo docente de adelantar el tratamiento de las ideas esenciales del ajuste de curvas, temática muy importante en las aplicaciones, lo que le confiere mayor significación a la optimización de funciones de una variable.

Palabras clave: estimación de curvas, materiales de enseñanza, método de mínimos cuadrados, método iterativo, regresión lineal.

1. Introduction

The method of regression is one of the most important statistical methods for higher education graduates. Its comprehension facilitates obtaining and correctly interpreting the results of different types of models to be applied in their professional careers. Bibliographic research in the scientific literature shows the wide interest and use of regression methods. From such bibliographic analysis, three approaches can be differentiated: The largest one, related to the application of regression methods to different fields and topics of science; see Braga, Silveira, Rodríguez, Henrique de Cerqueira, Aparecido & Barros (2009), Guzmán, Bolívar, Alepuz, González & Martin (2011), Ibarra & Arana (2011) and Santos da Silva, Estraviz, Caixeta & Carolina (2006) a second approach related to the theoretical aspects of the topic; see Núñez, Steyerberg & Núñez (2011), Vega-Vilca & Guzmán (2011), Donal (2001), Ranganatham (2004), Kelley (1999), Schmidt (2005). Lastly, a third group related to the teaching of the method, i.e., how to help students, and professionals in general, in correctly applying regression and interpreting its results see Batanero, Burrill & Reading (2011), Gutiérrez de Ravé, Jiménez-Hornero & Giráldez (2011) and Wei, De Quan & Jian (2001).

Applications of regression methods are found in scientific papers related to agriculture, medicine, environment, economics, sociology and different engineering areas. Using a random sample of one hundred papers published during 2012 and obtained by the authors from the Web of Knowledge, in 32% of them there was a direct application of these methods and in almost half of them (46%) there was a reference to regression.

Such significant use of regression supports its inclusion in the largest part of university curricula. It is generally included in the statistics subject.

In terms of teaching, curve adjustment is generally explained once the methods of optimization of real functions of several real variables are known, along with the solution of linear equation systems. This makes possible the support of procedures that permit to determine the values of the parameters characterizing the functional class of the best adjustment curve sought. Usually, in practice, computer packages are used to determine these parameters.

Taking into consideration the importance of curve adjustment for applications, it is advisable to teach the students these ideas much more before it is usually done in university curricula. How can this purpose be achieved if curve adjustment is preceded by mathematical requirements which seem not to be possible to sever? This paper presents an approach that permits to teach in advance the consideration of these basic ideas of regression in at least one semester, what is justified assuming the following hypotheses:

-The methods of optimization for the real functions of many variables are neglected, what has the immediate implication of not requiring the partial derivation.

-A system of linear equations is not stated, so the corresponding theory is not necessary.

-The reiterative application of the least squares method in functional classes determined by only one parameter, so that in each of them, the corresponding sum of the squared deviations is function of a unique variable. Consequently, optimization techniques for real functions of a real variable are only required.

Though sufficient and varied bibliography about the least squares method is available, it was considered necessary to make explicit some of its basic aspects initially, such as the expression that takes the sum of the squared deviations, as well as the normal system of equations that is formed at stating the necessary conditions for extremes, both in the case of the simple linear regression.

Curve adjustment is, possibly, the most frequently used mathematical resource for solving one of the fundamental problems related to numerous scientific areas: “reconstructing” a function starting from experimental data. Essentially, for the case of one variable functions, this problem may be formulated through the following statement:

“Given the set of n points $\{(x_1, y_1); (x_2, y_2); \dots (x_n, y_n)\}$, where n is a natural number and every two x_k abscissas are different, the goal is determining the $y = f(x)$ function which, within a given prefixed class of functions, best adjusts them”.

2. Materials and Methods

2.1. Curve Adjustment and the Least Squares Method

Generally, the prefixed functional class depends on various parameters, and the purpose of the method used for their estimation is to satisfy some criterion of

optimization, which is characteristic of the method; particularly, the objective of the least squares method is to minimize the sum of the squared deviations. Two other alternatives, which are also frequently used are the Maximum Likelihood Method; see Yoshimori & Lahiri (2014), Seo & Lindsay (2013) and Han & Phillips (2013) and for the Bayesian regression method; see Zhao, Valle, Popescu, Zhang & Mallick (2013), Mudgal, Hallmark, Carriquiry & Gkritza (2014) and Choi & Hobert (2013).

In the probably most renowned and significant case of finding the best-adjusting function within the class of linear functions of one independent variable $f(x) = a_1x + a_2$ this problem is solved through the least squares method by determining the values of the a_1 and a_2 parameters (the slope and the intercept with the y-axis, respectively), which provide the minimum value to the sum of the $S(a_1, a_2)$ squared deviations:

$$S(a_1, a_2) = \sum_{k=1}^n (f(x_k) - y_k)^2 = \sum_{k=1}^n (a_1x_k + a_2 - y_k)^2.$$

Determining the minimum of $S(a_1, a_2)$ requires applying optimization techniques for real functions of two real variables, which initially require the use of the necessary condition on extreme points:

$$\frac{\partial S}{\partial a_1} = 2 \sum_{k=1}^n x_k (a_1x_k + a_2 - y_k) = 0; \quad \frac{\partial S}{\partial a_2} = 2 \sum_{k=1}^n (a_1x_k + a_2 - y_k) = 0 \quad (1)$$

Afterwards, it requires the resolution of the system of two linear equations resulting from it with a_1 and a_2 as unknowns. This system is called the normal equation system, which is expressed as follows:

$$a_1 \sum_{k=1}^n x_k^2 + a_2 \sum_{k=1}^n x_k = \sum_{k=1}^n x_k y_k; \quad a_1 \sum_{k=1}^n x_k + n a_2 = \sum_{k=1}^n y_k \quad (2)$$

When applying any of the existing techniques for the resolution of system (2), the result is a single solution $a_1 = a_1^{(0)}$, $a_2 = a_2^{(0)}$, given by the expressions:

$$a_1^{(0)} = \frac{n \sum_{k=1}^n x_k y_k - \sum_{k=1}^n x_k \sum_{k=1}^n y_k}{n \sum_{k=1}^n x_k^2 - \left(\sum_{k=1}^n x_k \right)^2}; \quad a_2^{(0)} = \frac{1}{n} \left(\sum_{k=1}^n y_k - a_1 \sum_{k=1}^n x_k \right) \quad (3)$$

The procedure herein presented is equivalent to applying the Gauss - Seidel Method (McCracken & Dorn 1974) to the normal system of equation (2). This is an iterative method for the resolution of linear equation systems, as it happens with the system in equation (2), or the one resulting from applying the necessary condition in equation (1) to the sum of the squared deviations, when the adjustment takes place in a functional class that is linear with respect to the parameters defining it.

In terms of teaching organization, this approach provides more significance to the optimization methods of real functions of one real variable. At the same time, it permits the introduction of an important application such as curve adjustment, advancing one semester, at least.

2.2. An Iterative Method for the Process of Curve Adjustment

Solving the normal equation system of equation (2) is not possible until a method that permits optimizing a derivable function of two real variables is available. Therefore, a sequence, that only requires applying different variable optimization techniques, one at a time, may be followed. Such method results from realizing the following steps:

1. Prefix the functional class in which the adjustment process will be carried out.

As it is known, the functional class is characterized by a functional expression involving the independent variable and the p parameters that define it, being p a positive integer. This class of functions is denoted by:

$$y = f(x, a_1, a_2, \dots, a_p), \quad (4)$$

where x is the independent variable, and the parameters have been denoted by a_1, a_2, \dots, a_p , for which it is necessary to previously establish an order among them.

2. Keep the variable character of a_1 and assign values to the rest of the parameters.

The values assigned to the parameters are denoted by $a_2^{(0)}, a_3^{(0)}, \dots, a_p^{(0)}$, where the sub-index of each identifies the parameter, and the supra - index 0 indicates that it is the initial assignment. These values may be arbitrary or follow a certain criterion, but this is irrelevant to the method being described. Thus, the set of functions is defined as:

$$y_1 = f_1(x, a_1, a_2^{(0)}, \dots, a_p^{(0)}) \quad (5)$$

which is formed by functions of the independent variable x depending on the parameter a_1 that obviously constitutes a subclass of the pre-fixed functional class.

3. Form the sum of the quadratic differences in $y_1 = f_1(x, a_1, a_2^{(0)}, \dots, a_p^{(0)})$.

Given the set $\{(x_1, y_1); (x_2, y_2); \dots; (x_n, y_n)\}$, of n points, the corresponding sum of the quadratic differences to be minimized is formed, which is a function of the a_1 parameter, and is defined by the expression:

$$S(a_1) = \sum_{k=1}^n \left(f_1(x_k) - y_k \right)^2 = \sum_{k=1}^n \left(f_1(x_k, a_1, a_2^{(0)}, \dots, a_p^{(0)}) - y_k \right)^2$$

4. Apply the necessary extreme condition to $S(a_1)$.

As $S(a_1)$ is a one variable function, it is enough to state $S'(a_1) = 0$ to thus determine the solution of this equation. This gives the value of parameter a_1 , denoted by $a_1^{(1)}$ so that $S(a_1^{(1)})$ is the lowest value of $S(a_1)$. It is important to note that the supra-index 1 in $a_1^{(1)}$ means that this is the first value calculated for a_1 .

3. Results and Discussion

The implementation of the previous process guarantees obtaining the function of better adjustment within the $y_1 = f_1(x, a_1, a_2^{(0)}, \dots, a_p^{(0)})$ subclass, in the initially pre-fixed functional class $y = f(x, a_1, a_2, \dots, a_p)$.

In general, it is not expected that the $y_1^{(1)} = f_1(x, a_1^{(1)}, a_2^{(0)}, a_3^{(0)}, \dots, a_p^{(0)})$ function obtained from $y_1 = f_1(x, a_1, a_2^{(0)}, \dots, a_p^{(0)})$ by substituting $a_1^{(1)}$ by a_1 , to be a good approximation to the better adjustment within the general prefixed class $y = f(x, a_1, a_2, \dots, a_p)$.

The described process is repeated, leaving the next parameter as arbitrary (in this case a_2) and taking for a_1 the calculated value $a_1^{(1)}$, and for the rest of the parameters the initially assumed values $a_3^{(0)}, \dots, a_p^{(0)}$.

As a result, the value of parameter a_2 , denoted by $a_2^{(1)}$, will be obtained, offering the best adjustment function within the subclass:

$$y_2 = f_2(x, a_1^{(1)}, a_2, a_3^{(0)}, \dots, a_p^{(0)})$$

Once the whole set of parameters has been recovered, by proceeding similarly, the following p functions would be obtained:

$$\begin{aligned} y_1^{(1)} &= f_1(x, a_1^{(1)}, a_2^{(0)}, a_3^{(0)}, \dots, a_p^{(0)}) \\ y_2^{(1)} &= f_2(x, a_1^{(1)}, a_2^{(1)}, a_3^{(0)}, \dots, a_p^{(0)}) \\ &\dots\dots\dots \\ y_p^{(1)} &= f_p(x, a_1^{(1)}, a_2^{(1)}, a_3^{(1)}, \dots, a_p^{(1)}) \end{aligned}$$

where each of them is the best adjustment function within the corresponding functional class.

It can be verified that each of these functions is not a worse, but a better approximation than the previous one. Indeed, $y_2^{(1)}$ is reduced to $y_1^{(1)}$ taking $a_2^{(1)} = a_2^{(0)}$ for it. So with this value for that parameter, function $y_2^{(1)}$ provides a value for the sum of squared differences that is similar to the one given by $y_1^{(1)}$. This proves that $y_2^{(1)}$ is an approximation not worse than $y_1^{(1)}$. This also holds for the rest of the functions and this step completes the first iteration.

As the values of parameters $a_2^{(1)}, a_3^{(1)}, \dots, a_p^{(1)}$ have a similar purpose to the one followed with numbers $a_2^{(0)}, a_3^{(0)}, \dots, a_p^{(0)}$, the above process may be repeated

to find the second iteration, which will be completed once the following new p functions have been determined:

$$\begin{aligned}
 y_1^{(2)} &= f_1(x, a_1^{(2)}, a_2^{(1)}, a_3^{(1)}, \dots, a_p^{(1)}) \\
 y_2^{(2)} &= f_2(x, a_1^{(2)}, a_2^{(2)}, a_3^{(1)}, \dots, a_p^{(1)}) \\
 &\dots\dots\dots \\
 y_p^{(2)} &= f_p(x, a_1^{(2)}, a_2^{(2)}, a_3^{(2)}, \dots, a_p^{(2)})
 \end{aligned}$$

For a third iteration, the calculated values $a_2^{(2)}, a_3^{(2)}, \dots, a_p^{(2)}$ will be used, and the process would be the same successively. A possible stop criterion for the iterative process would be that each of the values determined for the parameters in a given iteration were sufficiently close to the corresponding value in the preceding iteration¹.

Sufficiently close means here, as is common in mathematics, that a certain pre-fixed degree $\epsilon > 0$ of accuracy is fulfilled for all j index ($j = 1, 2, \dots, p$), and inequality $|a_j^{(m)} - a_j^{(m-1)}| < \epsilon$, where m is a natural number indicating the number of order of the iteration.

The objective is proving that this process converges with the function of best adjustment within the pre-fixed functional class $y = f(x, a_1, a_2, \dots, a_p)$, which is the initial step in the iterative method. For the case of pre-fixing the linear functions, the method may be proven to converge.

3.1. Geometrical Interpretation of the Case of Linear Adjustment

If in the initial step of the process the linear function class is pre-fixed in an independent variable $f(x) = ax + b$, where a and b are real numbers, an interesting geometrical interpretation may be given from the described iterative method.

By keeping the notation used in the description of the iterative method, the slope is denoted by a_1 and the intercept with the y axis by a_2 , so that this general functional class is then expressed as:

$$y = a_1x + a_2$$

Prefixing the value of parameter a_2 equal to $a_2^{(0)}$ implies taking the point of coordinates $(0, a_2^{(0)})$ of the y - axis in the system of Cartesian coordinates, and considering the family of all straight lines that pass through such point, except the very axis of the ordinates (y - axis).

This family is formed by the graphics of the $y_1 = a_1x + a_2^{(0)}$ subclass functions, for all the possible values of the a_1 slope.

Once the best adjustment function $y_1^{(1)} = a_1^{(1)}x + a_2^{(0)}$ in $y_1 = a_1x + a_2^{(0)}$ is determined, we come back to the arbitrary character of the intercept with the y

¹This is one of the stop - criteria used in some numeric methods, such as Gauss - Seidel, for the resolution of a linear system.

- axis, taking the new subclass within the $y = f(x) = a_1x + a_2$ class, defined by $y_2 = a_1^{(1)}x + a_2$ to carry out the new adjustment process. As a result, the value of parameter a_2 will be determined as $a_2^{(1)}$, which determines the $y_2^{(1)} = a_1^{(1)}x + a_2^{(1)}$ equation, which is a straight line parallel to the one initially determined, $y_1^{(1)} = a_1^{(1)}x + a_2^{(0)}$.

Therefore, the first iteration concludes, when the following functions are ready:

$$y_1^{(1)} = a_1^{(1)}x + a_2^{(0)}, \quad y_2^{(1)} = a_1^{(1)}x + a_2^{(1)}$$

Initially, with $y_1^{(1)} = a_1^{(1)}x + a_2^{(0)}$, we determine the angle of inclination of the straight line that passes through the point of coordinates $(0, a_2^{(0)})$ with the positive direction of the axis of abscissas. This is subsequently transferred parallel to itself until it occupies the graphic position corresponding to $y_2^{(1)} = a_1^{(1)}x + a_2^{(1)}$ that passes through the point of coordinates $(0, a_2^{(1)})$. In turn, this is used to implement the second iteration: the new slope $a_1^{(2)}$ and the new intercept $a_2^{(2)}$, and so on.

3.2. One Example of the Application of the Iterative Method

A table with arbitrary or hypothetical data, which determine five points of integer coordinates: A(1, 1), B(2, 3), C(3, 3), D(4, 5) y E(5, 5), is taken.

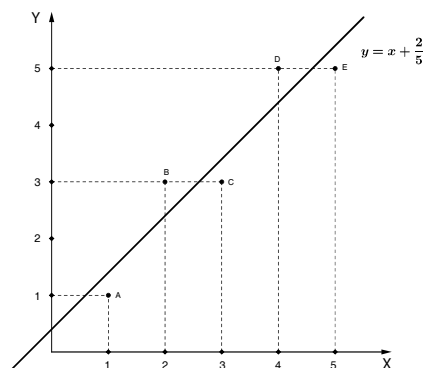


FIGURE 1: Regression line.

In Figure 1, a regression line $y = x + 0.4$ is represented. It was obtained by the the least squares method, in the general functional class $y = a_1x + a_2$, where the parameters are the slope a_1 and the intercept a_2 . The sum of the squared deviations is function of these two parameters, so to obtain $y = x + 0.4$ (it means, $a_1 = 1$ and $a_2 = 0.4$) techniques of optimization for the functions of some variables were required and the exact resolution of the normal system of equations.

In Figure 2, a segment of the first approximation is represented. It is the line that by the origin (of slope $61/55$) better adjusts to the five points. It is optimized in the functional class $y = a_1x$, where the parameter is the a_1 slope, what follows

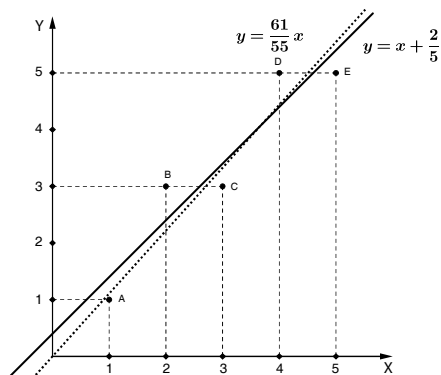


FIGURE 2: Segment of the first approximation.

from assigning, in $y = a_1x + a_2$, the zero value to a_2 parameter. Geometrically, it means that it optimizes in the functional class of all no vertical lines that pass through the origin. The sum of the squared deviations is only function of a_1 , so that this optimum (minimum) is determined by techniques of optimization of functions of one variable (it not even requires the ordinary derivative, observing that the sum of the squared deviations is a quadratic function in a_1 variable, whose graphic is a parabola that opens upwards, so that the optimum (minimum) is reached in the abscissa of the vertex (value $61/55$). It is maintained with purposes of comparison, the segment of the regression line.

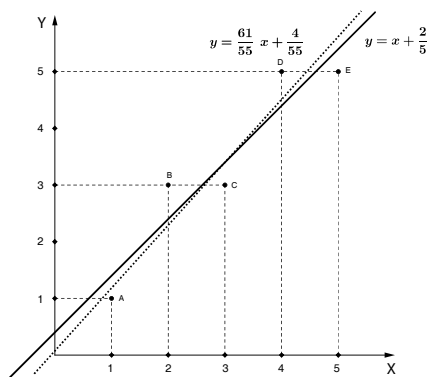


FIGURE 3: Segment of the second approximation.

In Figure 3 a segment of the second approximation, which best adjusts to the five points among all the lines with slope $61/55$, is represented. It is optimized in the functional class $y = (61/55)x + a_2$, which is obtained from the functional class $y = a_1x + a_2$ replacing $a_1 = 61/55$ and retaking the variable character of a_2 (notice that for a_2 the value 0 was initially assumed). Geometrically it means that the line of equation $y = (61/55)x$ is paralleled displaced itself up to a position that betters

the adjustment (provides the minimal for the sum of the squared deviations which now depends on a_2). The resulting value for the parameter is $a_2 = 4/55$.

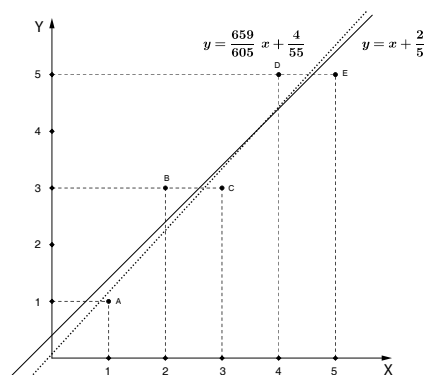


FIGURE 4: Segment of the third approximation.

For a new approximation $a_2 = 4/55$ in the general class $y = a_1x + a_2$ is replaced to optimize in the subclass $y = a_1x + 4/55$, in which the slope a_1 is variable again, so that what is looked for is the line that better adjusts to the data (in the sense of minimizing the corresponding sum of squared deviations) among all those that pass through the axis point $(0, 4/55)$. The result is the new slope value $a_1 = 659/605$, what permits the line with best equation adjustment $y = (659/605)x + 4/55$, a segment of which is represented in Figure 4 together with the one of the least regression line.

The process continues similarly, so that the new adjustment would take place in the functional subclass $y = (659/605)x + a_2$, where the variable character of the second of the parameters is retaken.

3.3. The Gauss - Seidel Method and Convergence in the Case of Linear Adjustment

The issue related to the convergence of the described iterative method has an affirmative answer in the case of linear adjustment, if the set of points fulfills the initially described characteristics; i.e., if within the full set, every pair of points has different abscissas.

As the best adjusting straight line, with the equation $y = a_1^{(0)}x + a_2^{(0)}$, does exist, and the parameters are analytically determined as the only solution by (3) in the standard equation system, an iterative method convergent to the solution of such system would obviously provide, after an adequate number of iterations, an $\alpha \approx a_1^{(0)}$, $\beta \approx a_2^{(0)}$ approximation. This offers the possibility of defining a $y = \alpha x + \beta$ approximation for the best adjustment equation $y = a_1^{(0)}x + a_2^{(0)}$.

One of the simplest iterative methods for the resolution of a linear equation system, easily programmed for its computerized application, is the Gauss - Seidel Method.

For the case of a two - equation system with two unknowns:

$$\begin{aligned} b_{11}a_1 + b_{12}a_2 &= c_1 \\ b_{21}a_1 + b_{22}a_2 &= c_2 \end{aligned}$$

The method is described as follows:

Supposing that in the coefficients matrix, those of the main diagonal are not null, it is possible to find the unknowns a_1 and a_2 :

$$a_1 = \frac{1}{b_{11}}(c_1 - b_{12}a_2); \quad a_2 = \frac{1}{b_{22}}(c_2 - b_{21}a_1) \quad (6)$$

An arbitrary approximation is now defined for the $a_1 = a_1^{(0)}$, $a_2 = a_2^{(0)}$ solution, and it is used to find a new approximation for the a_1 unknown value stemming from the first of the expressions in (6):

$$a_1^{(1)} = \frac{1}{b_{11}}(c_1 - b_{12}a_2^{(0)})$$

The $a_1^{(1)}$ calculated value is substituted in the second of the expressions (6) to determine an approximation to the a_2 unknown value:

$$a_2^{(1)} = \frac{1}{b_{22}}(c_2 - b_{21}a_1^{(1)})$$

At this point the first iteration is fulfilled.

The second iteration is implemented by taking the calculated approximation $a_1^{(1)}$, $a_2^{(1)}$ with the same role played by $a_1 = a_1^{(0)}$, $a_2 = a_2^{(0)}$ in the first iteration.

In this way it is possible to reach the order m iteration defined by the expressions:

$$a_1^{(m)} = \frac{1}{b_{11}}(c_1 - b_{12}a_2^{(m-1)}), \quad a_2^{(m)} = \frac{1}{b_{22}}(c_2 - b_{21}a_1^{(m-1)}) \quad (7)$$

A sufficient condition for convergence to the solution of the iterations produced through the Gauss - Seidel Method lies in the matrix of the coefficients being diagonally dominant, which means in this case that the $|b_{11}b_{22}| > |b_{21}b_{12}|$ inequality has to be fulfilled (McCracken & Dorn 1974).

If we now define $b_{11} = \sum_{k=1}^n x_k^2$; $b_{12} = b_{21} = \sum_{k=1}^n x_k$; $a = a_1$, and $b = a_2$ then the standard equation system (2) can be expressed as:

$$\begin{aligned} b_{11}a_1 + b_{12}a_2 &= c_1 \\ b_{21}a_1 + b_{22}a_2 &= c_2 \end{aligned}$$

where $c_1 = \sum_{k=1}^n x_k y_k$, $c_2 = \sum_{k=1}^n y_k$ so that the expressions in (7) allow determining an approximation to its solution.

It is not difficult to verify that the values given by the expressions (7) for the unique solution of the standard system (2) match in each m iteration, those provided for the a_1 slope and the a_2 intercept by each iteration of the iterative adjustment process here described. Neither is it difficult to prove that the system matrix (2) is diagonally dominant if in the $\{(x_1, y_1); (x_2, y_2); \dots; (x_n, y_n)\}$ set of points, where n is a natural number, every two of the x_k abscissas are different, which means that the $n \sum_{k=1}^n x_k^2 > \left(\sum_{k=1}^n x_k\right)^2$ inequality is fulfilled.

Indeed, according to Bronshtein & Semendiaev (1971), in the inequality (which is strict if there are at least two different x_k values):

$$\frac{|x_1 + x_2 + \dots + x_n|}{n} \leq \sqrt{\frac{x_1^2 + x_2^2 + \dots + x_n^2}{n}}$$

it would suffice to square both sides to obtain, first:

$$\frac{(x_1 + x_2 + \dots + x_n)^2}{n^2} \leq \frac{x_1^2 + x_2^2 + \dots + x_n^2}{n}$$

Then, multiplying the two sides by n^2 and expressing the sums in a compact form, it results:

$$n \sum_{k=1}^n x_k^2 \geq \left(\sum_{k=1}^n x_k\right)^2$$

As every two of the x_k numbers are supposed to be different, the fulfillment of the $n \sum_{k=1}^n x_k^2 > \left(\sum_{k=1}^n x_k\right)^2$ inequality is finally guaranteed. This proves that the matrix of the standard equation system (2) is diagonally dominant, and in turn implies that the expressions (7), obtained by applying the Gauss - Seidel Method, converge to the unique solution of such system.

At the same time, each iteration of the method was observed to coincide with the parameter values that result, in each step, from the function of better adjustment within the corresponding subclass. Therefore, a conclusion can be advanced so that these functions converge to the least squares straight line of equation $y = a_1^{(0)}x + a_2^{(0)}$.

4. Conclusions

An iterative method has been proposed to obtain an approximation of the best adjustment function to a given set of points, consisting of determining the best adjustment function within a certain subclass of the pre-fixed functional class each time. Each subclass is dependent on a single parameter.

As optimization is used only on one variable, it is not required to explicitly write the standard equation system.

For the case of linear adjustment with one independent variable, the iterative method is revealed to be equivalent to the application of the standard equation

system of the Gauss - Seidel Method, which permits to show its convergence to the lowest quadratic straight line of the $y = a_1^{(0)}x + a_2^{(0)}$ equation.

Everything suggests that for other linear functional classes, with respect to the parameters that define them, similar results should be obtained, in the sense of the equivalence between the iterative method and the Gauss - Seidel one. Also, in this way, it may be possible to show that the iterative method is convergent to the best adjustment function obtained when applying the least squares method.

The proposed approach offers the possibility of focusing the least squares method along with that of curve adjustment, as an application of optimization techniques of the real functions of one real variable, developed during the first semester of higher education diplomas. This would permit speeding the approach of the significant topic of curve adjustment.

Acknowledgments

The authors acknowledge the financial support of AECID (Spain) under Project N^o A2/039476/11 - "Institutional Strengthening in the Teaching of Subjects of Statistics, Operative Investigation, Liability and Quality, and Scientific Application in Topics of Regional Interest".

[Recibido: agosto de 2013 — Aceptado: febrero de 2014]

References

- Batanero, C., Burrill, G. & Reading, C. (2011), *Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education: A Joint ICMI/IASE Study*, Vol. 18, 2 edn, Springer, New York.
- Braga, G., Silveira, C., Rodríguez, V., Henrique de Cerqueira, P., Aparecido, W. & Barros, F. (2009), 'Quantifying herbage mass on rotationally stocked palisadegrass pastures using indirect methods', *Scientia Agricola* **66**(1), 127–131.
- Bronshtein, I. & Semendiaev, K. (1971), *Mathematics Manual*, Mir, Moscow.
- Choi, H. M. & Hobert, J. P. (2013), 'Analysis of MCMC algorithms for Bayesian linear regression with Laplace errors', *Journal of Multivariate Analysis* **117**(0), 32–40.
- Donal, R. J. (2001), 'A taxonomy of global optimization methods bases on response surfaces', *Journal of Global Optimization* **21**, 345–383.
- Gutiérrez de Ravé, E., Jiménez-Hornero, F. J. & Giráldez, J. V. (2011), 'A computer application for interpolation algorithms of curves', *Computer Applications in Engineering Education* **19**(1), 40–47.

- Guzmán, K. P., Bolívar, I., Alepuz, M. T., González, D. & Martín, M. (2011), 'Impacto en el tiempo asistencial y el estadio tumoral de un programa de diagnóstico y tratamiento rápido del cáncer colorrectal', *Revista Española de Enfermedades Digestivas* **103**(1), 13–19.
- Han, C. & Phillips, P. C. (2013), 'First difference maximum likelihood and dynamic panel estimation', *Journal of Econometrics* **175**(1), 35 – 45.
- Ibarra, M. & Arana, P. (2011), 'Crecimiento del camarón excavador *Parastacus pugnax* (Poëppig, 1835) determinado mediante técnica de marcaje', *Latin American Journal of Aquatic Research* **39**(2), 378–384.
- Kelley, C. T. (1999), *Iterative Methods for Optimization*, Society for Industrial and Applied Mathematics, Philadelphia, U.S.
- McCracken, D. D. & Dorn, W. S. (1974), *Métodos Numéricos y Programación Fortran.*, 1 edn, Editorial Pueblo y Educación, La Habana, Cuba.
- Mudgal, A., Hallmark, S., Carriquiry, A. & Gkritza, K. (2014), 'Driving behavior at a roundabout: A hierarchical Bayesian regression analysis', *Transportation Research Part D: Transport and Environment* **26**(0), 20–26.
- Núñez, E., Steyerberg, E. W. & Núñez, J. (2011), 'Estrategias para la elaboración de modelos estadísticos de regresión', *Revista Española de Cardiología* **64**(6), 501–507.
- Ranganatham, A. (2004), 'The Levenberg - Marquardt algorithm', *Tutorial on LM Algorithm* .
- Santos da Silva, L. M., Estraviz, L. C., Caixeta, J. V. & Carolina, S. B. (2006), 'Fitting a Taper function to minimize the sum of absolute deviations', *Scientia Agricola* **63**(5), 460–470.
- Schmidt, M. (2005), *Least Squares Optimization with L1-Norm Regularization*, CS542B Project Report.
- Seo, B. & Lindsay, B. G. (2013), 'Nearly universal consistency of maximum likelihood in discrete models', *Statistics and Probability Letters* **83**(7), 1699–1702.
- Vega-Vilca, J. & Guzmán, J. (2011), 'Regresión PLS y PCA como solución al problema de multicolinealidad en regresión múltiple', *Revista de Matemática: Teoría y Aplicaciones* **18**(1), 09–20.
- Wei, W., De Quan, L. & Jian, L. (2001), 'On the system of multiple linear regression of higher education tuition in China', *Journal Advanced Materials Research* **211-212**, 752–755.
- Yoshimori, M. & Lahiri, P. (2014), 'A new adjusted maximum likelihood method for the Fay-Herriot small area model', *Journal of Multivariate Analysis* **124**(0), 281 – 294.

Zhao, K., Valle, D., Popescu, S., Zhang, X. & Mallick, B. (2013), 'Hyperspectral remote sensing of plant biochemistry using Bayesian model averaging with variable and band selection', *Remote Sensing of Environment* **132**(0), 102–119.