IB Mathematics Higher Level Internal Assessment on Statistics

Approximation of Experimental Physical Values with Non-typical Conditions of Error

David Simon Tetruashvili

Contents

1 Introduction

2 Mathematical Background

Because of the nature of IB Mathematics IAs, the mathematics in this written work extend from the IB syllabus. Some notations, concepts, and calculations have to be explained.

2.1 Notations

During the IB Math HL program, students learn about vectors. 2D-Vectors usually have only magnitude and direction, however, a vector can only be 3D, or even 4D. Essentially, so long that a single value has multiple parameters it is a vector in the same number of dimensions as there are parameters [*]. In this IA, all vectors will be denoted with the variable c, be it with some indexes or accents, as long as a variable has c in its core, it is a vector. Note that this is just a notation that I will use only in this paper. in addition, to save real estate not vertical vector notation will be used, instead I will use the so called ordered set notation of a vector [*] in the form of $c = (c_1, c_2, \dots)$.

Another new to IB notation in this IA is the product notation:

$$\prod_{i=m}^{n} x_i = x_m \cdot x_{m+1} \cdot \ \cdots \ \cdot x_{n-1} \cdot x_n.$$

This essentially works in much of the same way the familiar summation notation we use in class(using the Greek capital letter sigma), except that instead of summing values, it multiplies them (and uses Greek the capital letter pi) [1].

2.2 Concepts

Derivatives are an essential part of the IB curriculum. But when it comes to functions with multiple variables (like in this IA), their derivative is taken differently. First off, the derivative of such functions is called a partial derivative. Denoted using a stylized symbol ∂, it is the derivative of a function with multiple variables, with respect to one of them, when others are held constant [3]. The following example will explain this further:

Find the local minimum of function z with multiple variables:

$$z = f(x, y) = 5(2x - 3)^2 + 4(4y + 1)^2$$

Without calculating anything, we see that the minimum of this function occurs when $x = \frac{3}{2}$,$y = -\frac{1}{4}$.

$$\text{(Now expand this function)}$$
$$z = 20x^2 + 64y^2 - 60x + 32y + 49$$

Now according to normal method of finding a minimum of a function, we equate, in this case, its partial derivatives (one with respect to x, the other to y) to zero and solve for x and y.

$$\begin{aligned} \frac{\partial f(x,y)}{\partial x} &= 40x - 60 = 0 \rightarrow x = \frac{60}{40} = \frac{3}{2} \\ \frac{\partial f(x,y)}{\partial y} &= 128y - 32 = 0 \rightarrow y = -\frac{32}{128} = -\frac{1}{4} \end{aligned}$$

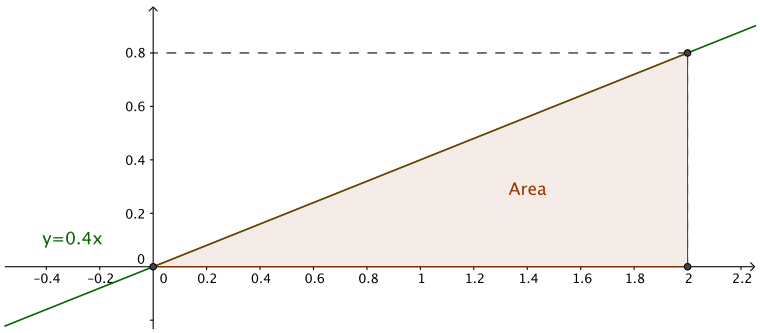
Voila, the partial derivatives gave us the answer as well.

But as much as derivatives a essential, so are integrals. This IA invokes the use of something called

multi integrals. These are basically integral an of integral, where again a function with multiple variables is in question [4]. I.e. a 2-milti integral of some two-variable function $g(x, y)$ is denoted as such:

$$\int\int g(x,y)\mathrm{d}x\mathrm{d}y$$

Yet again let's employ an example:
Find the area of the shaded region



Obviously,

$$A = \frac{1}{2} \cdot 2 \cdot 0.8 = 0.8.$$

Granted, the following method is very unnecessary (and this problem doesn't even require a multi integral), but for the sake of explanation, let's employ multi integrals as well.

$$A = \underbrace{\int_0^2}_{(\text{against } x)} \underbrace{\int_0^{0.4x}}_{(\text{against } y)} \mathrm{d}x\mathrm{d}y = \int_0^2 \mathrm{d}x \cdot [y]_0^{0.4x} = \int_0^2 \mathrm{d}x \cdot [0.4x - 0] = \int_0^2 \mathrm{d}x \cdot 0.4x = 0.4 \int_0^2 x\mathrm{d}x =$$

$$0.4 \cdot \left[\frac{x^2}{2}\right]_0^2 = 0.4 \cdot \left[\frac{2^2}{2} - \frac{0^2}{2}\right] = 0.4 \cdot 2 = 0.8.$$

And, again, the answer is the same. Now with these mathematical concepts explained we formulate the problem of this IA.

3 Formulation of the Problem

A function is one of the most known mathematical objects. An important task which has practical applications, is the approximation of a function or relationship based on some information known about the function or relationship in question. This information may either be determinate or statistical. An example of a determinate piece of information about function $f(x)$ is its range (or the possible values this function may have) on a given interval $[\alpha, \beta]$. Example of a statistical information may be the law of distribution of random errors ξ_i in approximate values $\tilde{y}_i = f(x_i) + \xi_i$ of the function, which in turn can describe a certain physical process (change of temperature over time, for example). In practice, a number n of points x_i can be obtained as results of some kind of physical experiment. Where in this case, the approximation of function $f(x)$ only makes sense if the this function is described by a finite number $m < n$ of parameters (coefficients) c_j, where the true values of said parameters will be denoted as $\dot{c}_j,\ j = 1, 2, \ldots, m$.

This Internal Assessment will focus on the estimation of parameters of the function

$$y = f(\dot{c}, x_i), \quad \dot{c} \in \boldsymbol{R}^m, \quad x \in [\alpha, \beta], \quad \dot{c} = (\dot{c}_1, \dot{c}_2, \ldots, \dot{c}_m) \tag{3.0.1}$$

based on its approximate values

$$\tilde{y}_i = f(x_i) + \xi_i, \quad i = 1, 2, \ldots, n, \tag{3.0.2}$$

when additionally it is also known, that: 1. vector $\dot{c} = (\dot{c}_1, \dot{c}_2, \ldots, \dot{c}_m)$ belongs to a given limited set D, like for example a parallelepiped in $\boldsymbol{R}^m$ dimensions; 2. ξ is a limited continuous random value; the median of which $Med(\xi)$ is equal to zero.

Judging by references in scientific works that I read while researching for this IA [*], the most popular linear model of a studied relationship is

$$f(\dot{c}, x) = \sum_{j=1}^{m} \dot{c}_j \phi_j(x), \tag{3.0.3}$$

specifically in polynomial form, when

$$\phi_1(x) \equiv 1; \quad \phi_j(x) = x^{j-1}, \quad j = 1, 2, \ldots, m. \tag{3.0.4}$$

In practice, it is often the case when it is not only necessary to estimate the parameters of a function, but identifying the type (structure) of this function is needed as well. In other words, a finite number L of alternative structures is given

$$f_l(c; x), c \in \boldsymbol{R}^{m(l)}, \quad l = 1, 2, \ldots, L, \tag{3.0.5}$$

and it is necessary to identify to which of L structures of function $f_l(c; x)$ belongs the function $f(\dot{c}, x)$, and after that estimate the vector $\dot{c}$ of its parameters. In our school program, the class has encountered one such task, when it was said to find out if we were dealing with a linear or exponential relationship, be it in either physics or math. However, then, this problem was solved using the exact (or near to exact) values of both of the relationships, so it was easy to distinguish them.

There are countless papers dedicated to the approximation of functions based on their approximate values (in practice - experimental data). Usually, in such papers the consensus is to use a certain condition. This condition is to assume that the mathematical expectancy of error is equal to zero [*].

$$E(\xi) = 0 \tag{3.0.6}$$

However, in this IA this condition will not be used. Here instead of the condition of mathematical expectancy of error ξ being equal to zero, I will assume that the median of the same error ξ being equal to zero,

$$Med(\xi) = 0 \tag{3.0.7}$$

specifically when the algorithm of evaluation of the parameters of the function is based on ideas from the *method of least squares*.

I justify my interest to the condition $Med(\xi) = 0$ by the case when the traditional condition $E(\xi) = 0$ is unachievable. This happens when measurements are taken close to one of the natural limits of the physical relationship being measured. An example of such natural limit is the inability of some magnitude, such as weight, to be negative. In this case, the absolute value of the error made, can only be large (with respect to other errors made) in the *same sign*, either positive or negative. Figure 1 shows an example of this graphically.
I want to bring attention to the fact, that the argument that this kind of measurement could be withdrawn by human intervention, is invalid for 2 reasons: 1. Any such withdrawal usually leads to loss of information. 2. In cases where the experiment requires high-capacity data collection human intervention may not be possible.

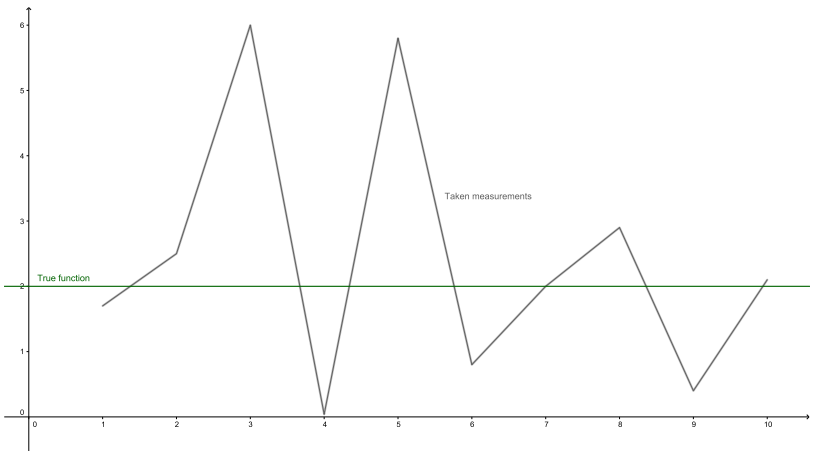


Figure 1: Graphical representation of a case where $E(\xi) = 0$ does not work effectively.

Speaking of the errors, what is meant is not only error that was produced by a faulty measurement, but also any error caused by some factor that was either omitted or unaccounted for in function $f(x)$. Even though both conditions $E(\xi) = 0$ and $Med(\xi) = 0$ are not special cases of each other,it could be argued that from a point of view of solving practical problems, the condition $Med(\xi) = 0$ is the more broad of the two (as in, it is easier to meet). The only requirement for meeting this condition is $P(\xi) > 0$ is 0.5. Hence the condition $Med(\xi) = 0$ allows for some comparatively large random values of error ξ to be on one side of the true function and not on the other, without the approximation to be significantly affected by those large values, unlike the condition $E(\xi) = 0$. With condition $Med(\xi) = 0$ the approximation can account for large peaks in values of ξ.

As mentioned before, I have two aims in this IA. Aim-maximum - to create an algorithm, capable of identifying from a finite number of alternative function structures, to which of those does the relationship in question belong; estimate the parameters of this function accurately; quite securely give an estimate to the accuracy of the calculated approximate values of the function, which belong to the relationship in question. Aim-minimum - to create an algorithm, capable of estimating the parameters of a functional dependence whose structure is known accurately; again, securely give an estimate to the accuracy of the calculated approximate values of the functional dependence.

Its clear that the quality of the solution of this problem is dependant of an array of factors, which include: 1. the ratio between the number n of measurements $\tilde{y}_i$ and the number m of estimated parameters $\dot{c}_j$. 2. the intensity of error ξ_i 3. the number of L alternatives, and more importantly, the degree of similarity of functions $f_l(c; x)$. This means that to confirm my theoretical reasoning, quite an ambitions computational experiment is required. I will proceed with the necessary calculations using custom software.

4 Algorithm

Approximation of a functional dependence taking this new condition $Med(\xi) = 0$ in mind, has been looked at in mathematics [*]. It is believed that in this case it is necessary to minimize the sum of absolute values of deviations of the modelled dependence $f(c^*, x)$ from the unknown true function $f(\dot{c}, x)$, where c^* is the found optimal value of vector c. This method is referred to as the Least Absolute Deviations (LAD). However, through my research I have found no methods of estimating LAD's accuracy. What value does an optimization method have if there is no way to determining the error it made? In addition, LAD does not presume the existence of priori limitations on the vector $\dot{c}$. And i must ask the question: What happens if the vector of parameters c^*, providing the minimum of the sum of modulus of errors, does not belong to the set D?

It is clear, that in every separately taken case (run of an algorithm), the factual accuracy of the model solution (when the true function is known) cannot serve as either a comparative evaluation of two competing algorithms, nor criteria of effectiveness of any given algorithm. It is also clear, that if all, or close to all errors ξ_i have the same sign (the condition $Med(\xi) = 0$, although, the condition $E(\xi) = 0$ as well, allow this, be it with a small probability), then neither method will give any good solutions. And also, with a certain 'layout' of errors ξ_i, a theoretically more sound method might by change give a worse solution that a less sound one. So, when estimating the effectiveness of a method, it is necessary to rely on average results of some number of random solutions. In conjunction with this, the idea lies in the fact that for the quality of constructed approximation $f(c^*, x)$ to $f(\dot{c}, x)$, I take the mathematical expectation

$$E(\rho(c^*, \dot{c})) = \int\limits_D P(c)\rho(c^*, \dot{c})\mathrm{d}c_1\mathrm{d}c_2 \dots \mathrm{d}c_m, \quad c = (c_1, c_2, \dots, c_m) \tag{4.0.1}$$

of proximity (distance) $\rho(c^*, \dot{c})$ of function $f(c^*, x)$ from $f(\dot{c}, x)$ where in the role of distance $\rho(c^*, \dot{c})$, one could take on of the functions

$$\rho_1(c^*, c) = \sum_{j=1}^{m} \left|c_j^* - c_j\right| \tag{4.0.2}$$

$$\rho_2(c^*, c) = \sum_{j=1}^{m} \left(c_j^* - c_j\right)^2 \tag{4.0.3}$$

$$\rho_3(c^*, c) = \sqrt{\frac{1}{n}\sum_{i=1}^{m} \left(y_i - f(c^*, x_i)\right)^2} \tag{4.0.4}$$

In solving the problem, that I have above labelled as 'aim-minimum', criteria (4.0.1) was considered in article [2]. Looking ahead, I say that I will suggest a more constructive algorithm than the one occurring in [2]. I want to note that the problem that I have above labelled as 'aim-maximum' was not looked at in the mentioned paper.

The probability density function $P(c),\ c \in D$ where c is a vector that could be the unknown true vector $\dot{c}$, that (the function) appears in the m-multi integral (4.0.1), can be constructed on the basis of the formula of the binomial distribution of a random value [*]. In fact, let's say: $c \in D$ is one of the vectors which claims that it is the unknown true vector $\dot{c}$ from function (3.0.3); q_i are the elements of the sequence

$$q_1(c) = \tilde{y}_1 - f(c, x_1),\ q_2(c) = \tilde{y}_2 - f(c, x_2), \dots,\ q_n(c) = \tilde{y}_n - f(c, x_n); \tag{4.0.5}$$

where q is a discreet random value, that can assume values

$$r = r(c) = \sum_{i=1}^{n-1} \delta_i(c), \tag{4.0.6}$$

where

$$\delta_i = \delta_i(c) = \begin{cases} 1, \text{if} & q_i(c)q_{i+1}(c) < 0 \\ 0, \text{if} & q_i(c)q_{i+1}(c) \geq 0 \end{cases} \tag{4.0.7}$$

In meaningful terms, the value of r is the number of transitions of sign of the elements of (4.0.5). Where $r \in [0, n-1]$. If it truly happens that $c = \dot{c}$, then the values of q_i would be nothing but the errors ξ_i, and by the condition $Med(\xi) = 0$ the probabilities p_r of events $q = r$ could be written as

$$p_r = \frac{\binom{n-1}{r}}{2^{n-1}}, \quad r = 0, 1, \dots, n-1. \tag{4.0.8}$$

Lets move on from discussing the question of the transition of sign with some one vector c, to the analysis of this situation with regards to the whole set D. Say that there exists a partitioning of set

D into a family of sub-sets $D_1, D_2, \ldots, D_n$ such that the elements (in this case vectors c) of sub-set D_r provide the same number $r-1$ of transitions of sign of elements of (4.0.5). In this way, the priori probability that the unknown true vector $\dot{c} \in D_r$ in given by (4.0.8). However, in any separate case, some of the sub-sets D_r could be empty, meaning that with some vectors of parameters $c \in D$, the number of transitions of sign of (4.0.5) will not equal the given value of r (any one can make sure of this by trying to draw a line which would crossed a given poly-line by guess, in such way that all of the endpoints of this poly-line turned out to be on either side of the guessed line). The priori probabilities p_r that the unknown true vector $\dot{c}$ is in one of the non-empty sub-sets D_r can be recalculated to the posteriori probabilities p_r^*. Let's define I as the set of the numbers of all the non-empty sub-sets D_r. Then

$$p_r^* = \begin{cases} \frac{p_r}{\sum\limits_{s \in I} p_s}, & r \in I. \\ 0, & r \in \{0, 1, \ldots, n-1\} \setminus I \end{cases} \tag{4.0.9}$$

Because in our case the vectors c, are those that could be the true vector $\dot{c}$, are continuous values, and not discreet ones, then to use formula (4.0.1) we must proceed from estimates p_r^* of the probabilities of event $\dot{c} \in D_r$ to estimates $p_r^*(c), c \in D_r$, of a probability density function. As there is no information which would allow me to somehow rank/sort the 'preference' of vector c in the bounds of each sub-set D_r, it is logical to assume uniform distribution

$$P_r^*(c) = \frac{p_r^*(c)}{\mu(D_r)}, \quad r \in I, \tag{4.0.10}$$

where $\mu(D_r)$ is the measure (analogous to the volume) of sub-set D_r in $\boldsymbol{R}^m$ dimensions.

Therefore (4.0.1) becomes

$$E(\rho(c^*, \dot{c})) = \sum_{r \in I} \frac{p_r^*(c)}{\mu(D_r)} \int\limits_{D_r} \rho(c^*, \dot{c}) \mathrm{d}c_1 \mathrm{d}c_2 \ldots \mathrm{d}c_m \tag{4.0.11}$$

where $\int_{D_r}(\bullet)$ is a laconic (short) notation of a multi integral (in this case a m-multi integral). In the general case, such as this one, the limits of integration of each univariate integral depend on variables, based on which the integration of the external integral relative to the given univariate integral is carried out [*]. (Refer to the example in section 2). This significantly complicates the calculation of these multi integrals. Yet again, looking ahead, I want to note that in the algorithm created, integration is carried out on multidimensional parallelepipeds, when the limits of integration of each integral remain constant.

4.1 Illustrative example

For it to be clear, that the following bulky equations lead to success, let's consider an illustrative example, where $m = 1$ (i.e. the vector $\dot{c}$ is a scalar value), $y = f(c; x)$ is a function with one parameter, where $\dot{c}$ is a scalar that has to be found. In this case $n = 6$ and $D = \{c : 2 \leq c \leq 6\}$. In this way, the following graphical representations form.

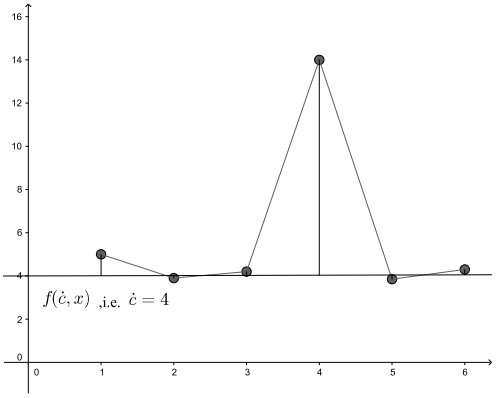
For this example I will use proximity function (4.0.3).

As we already know, we can calculate the priori probabilities using (4.0.8), where $r = 0, 1, 2, 3, 4, 5$ like so:

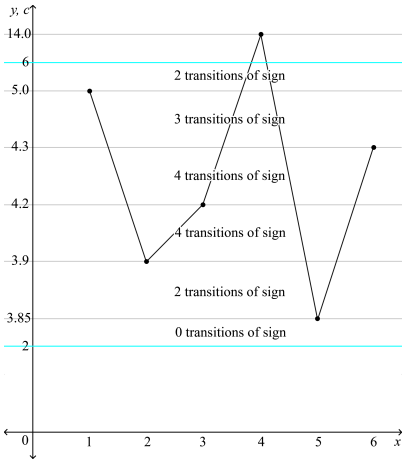
$$p_0 = \frac{1}{32};\ p_1 = \frac{5}{32};\ p_2 = \frac{10}{32};\ p_3 = \frac{10}{32};\ p_4 = \frac{5}{32};\ p_5 = \frac{1}{32}$$

But can only have certain values: $r = 0, 2, 3, 4$. And ,so

$$s = p_0 + p_2 + p_3 + p_4 = \frac{1}{32} + \frac{10}{32} + \frac{10}{32} + \frac{5}{32} = \frac{26}{32}.$$



(a) Graph of proposed example



(b) Illustration of transitions of sign q_i between each $\tilde{y}_i$

Figure 2

Table 1: Table of values for this example

i	x_i	$f(\dot{c}, x_i)$	ξ_i	$\tilde{y}_i = f(\dot{c}, x_i) + \xi_i$
1	1	4	1.0	5.0
2	2	4	-0.1	3.9
3	3	4	0.2	4.2
4	4	4	10	14.0
5	5	4	-0.15	3.85
0	6	4	0.3	4.3

So, to calculate the posteriori probabilities

$$\begin{aligned}
P_0^* &= P_0 : \frac{26}{32} = \frac{1}{32} : \frac{26}{32} = \frac{1}{26} \\
P_1^* &= 0 \\
P_2^* &= P_2 : \frac{26}{32} = \frac{10}{32} : \frac{26}{32} = \frac{10}{26} \\
P_3^* &= P_3 : \frac{26}{32} = \frac{10}{26} \\
P_4^* &= P_4 : \frac{26}{32} = \frac{5}{32} : \frac{26}{32} = \frac{5}{26} \\
P_5^* &= 0
\end{aligned}$$

Calculating $E(\rho(c^*, c))$

$$E(\rho(c^*,c)) = \int_2^{3.85} \underbrace{(c^*-c)^2}_{\rho(c^*,c)} \cdot \frac{\overbrace{\frac{1}{26}}^{p_0^*}}{\underbrace{3.85-2}_{\text{similar to } \mu(D_0)}} \mathrm{d}c + \int_{3.9}^{4.3} (c^*-c)^2 \cdot \frac{\frac{5}{26}}{4.3-3.9} \mathrm{d}c + \int_{4.3}^{5.0} (c^*-c)^2 \cdot \frac{\frac{10}{26}}{5.0-4.3} \mathrm{d}c$$

$$+ 1.05 \cdot \left(\int_{3.85}^{3.9} (c^*-c)^2 \cdot \frac{10}{26} \mathrm{d}c + \int_{5.0}^{6.0} (c^*-c)^2 \cdot \frac{10}{26} \mathrm{d}c \right)$$

$$= (c^*)^2 \cdot c - c^* \cdot c^2 + \frac{1}{3} c^3$$

Taking into account that later we will want to find the derivative $\frac{\mathrm{d}E(\rho(c^*,c))}{\mathrm{d}c^*}$, we can omit the summand $\frac{1}{3}c^3$ as it is a constant, and the derivative of a constant is zero. Let's continue.

$$\begin{aligned} E(\rho(c^*,c)) = 0.045 \quad & \left[(c^*)^2 \cdot c - c^* \cdot c^2\right]_{2}^{3.85} \\ +7.323 \quad & \left[(c^*)^2 \cdot c - c^* \cdot c^2\right]_{3.85}^{3.9} \\ +0.481 \quad & \left[(c^*)^2 \cdot c - c^* \cdot c^2\right]_{3.9}^{4.3} \\ +0.549 \quad & \left[(c^*)^2 \cdot c - c^* \cdot c^2\right]_{4.3}^{5.0} \\ +0.366 \quad & \left[(c^*)^2 \cdot c - c^* \cdot c^2\right]_{5.0}^{6.0} \\ = \quad & 1.361(c^*)^2 - 12.061c^* \end{aligned}$$

$$\begin{aligned} \frac{\mathrm{d}E(\rho(c^*,c))}{\mathrm{d}c^*} &= 2.722c^* - 12.061 = 0 \\ c^* &= \frac{12.061}{2.722} = 4.431 \end{aligned}$$

Seeing that $\dot{c} = 4.0$, we finally calculate ρ using this method

$$\rho = |4.431 - 4.0| = 0.431$$

I.e. the error made by this method is 0.431. Now let's see what kind of error will the Least Square Method give:

$$UsingLeastSquareMethod$$

$$\phi(c^*) = \sum_{i=1}^{6}(\tilde{y}_i - c^*)^2 \longrightarrow min$$

$$\phi(c^*) = \sum_{i=1}^{6}(\tilde{y}_i^2 - 2\tilde{y}_i c^* + (c^*)^2)$$

$$\frac{\mathrm{d}\phi(c^*)}{\mathrm{d}c^*} = \sum_{i=1}^{6}(0 - 2\tilde{y}_i + 2c^*) = 0$$

$$\sum_{i=1}^{6} c^* = \sum_{i=1}^{6} \tilde{y}_i$$

$$6 \cdot c^* = \sum_{i=1}^{6} \tilde{y}_i = 35.25$$

$$c^* = \frac{35.25}{6} = 5.875$$

And so the error $\rho = 1.875$

Because this example shows, that in the case of significant non-compliance to condition $E(\xi) = 0$, my method is notably superior to LSM, and because of the limited format of this IA, I will not include further solutions that used LSM. Back to the theory.

Let the proximity $\rho(c^*, c)$ in (4.0.11), be taken in the form (4.0.3). Then the following simplifica-

tion takes place.

$$E(\rho(c^*,\dot{c})) = \sum_{r\in I}\frac{p_r^*}{\mu(D_r)}\int\limits_{D_r}\sum_{j=1}^{m}\left(c_j^*-c_j\right)^2 \mathrm{d}c_1\mathrm{d}c_2\ldots\mathrm{d}c_m =$$

$$expanding\left(c_j^*-c_j\right)^2$$

$$\sum_{r\in I}\frac{p_r^*}{\mu(D_r)}\sum_{j=1}^{m}\left[(c_j^*)^2\underbrace{\int\limits_{D_r}\mathrm{d}c_1\mathrm{d}c_2\ldots\mathrm{d}c_m}_{=\mu(D_r)} - 2c_j^*\int\limits_{D_r} c_j\ \mathrm{d}c_1\mathrm{d}c_2\ldots\mathrm{d}c_m + \int\limits_{D_r} c_j^2\ \mathrm{d}c_1\mathrm{d}c_2\ldots\mathrm{d}c_m\right] =$$

$$\sum_{j=1}^{m}\left(\sum_{r\in I}\frac{p_r^*}{\mu(D_r)}\left[(c_j^*)^2\mu(D_r) - 2c_j^*\int\limits_{D_r} c_j\ \mathrm{d}c_1\mathrm{d}c_2\ldots\mathrm{d}c_m + \int\limits_{D_r} c_j^2\ \mathrm{d}c_1\mathrm{d}c_2\ldots\mathrm{d}c_m\right]\right) =$$

$$\sum_{j=1}^{m}\left((c_j^*)^2\cancel{\mu(D_r)}\underbrace{\sum_{r\in I}\frac{p_r^*}{\cancel{\mu(D_r)}}}_{\sum_{r\in I}p_r^*=1} - 2c_j^*\sum_{r\in I}\frac{p_r^*}{\mu(D_r)}\int\limits_{D_r} c_j\ \mathrm{d}c_1\mathrm{d}c_2\ldots\mathrm{d}c_m + \sum_{r\in I}\frac{p_r^*}{\mu(D_r)}\int\limits_{D_r} c_j^2\ \mathrm{d}c_1\mathrm{d}c_2\ldots\mathrm{d}c_m\right) \tag{4.1.1}$$

And finally:

$$E(\rho(c^*,\dot{c})) = \sum_{j=1}^{m}\left((c_j^*)^2 - 2c_j^*\sum_{r\in I}p_r^*\frac{\int\limits_{D_r} c_j\ \mathrm{d}c_1\mathrm{d}c_2\ldots\mathrm{d}c_m}{\mu(D_r)} + \underbrace{\sum_{r\in I}p_r^*\frac{\int\limits_{D_r} c_j^2\ \mathrm{d}c_1\mathrm{d}c_2\ldots\mathrm{d}c_m}{\mu(D_r)}}_{\text{derivative of this with respect to } c_1^*,c_2^*,\ldots,c_m^*=0}\right) \tag{4.1.2}$$

In order to find the vector c^*, which provides the minimum of function (4.1.2) and which I will take as the optimal estimate of the unknown vector $\dot{c}$, it is necessary (like in the case with a simple one-parameter function) to take the first partial derivative $\frac{\partial E}{\partial c_j^*}$ of function (4.1.2) with respect to each variable c_j^* [*]. After that, equate these derivatives to zero and solve the resulting system of linear equations. In this case, this system splits into m separate linear equation with one variable:

$$2c_j^* - 2\sum_{r\in I}p_r^*\frac{\int\limits_{D_r} c_j\mathrm{d}c_1\mathrm{d}c_2\ldots\mathrm{d}c_m}{\mu(D_r)}, \quad j=1,2,\ldots,m. \tag{4.1.3}$$

These equations have the following solutions:

$$c_j^* = \sum_{r\in I}p_r^*\bar{c}_{(j,r)}, \quad j=1,2,\ldots,m, \tag{4.1.4}$$

where

$$\bar{c}_{(j,r)} = \frac{\int\limits_{D_r} c_j\mathrm{d}c_1\mathrm{d}c_2\ldots\mathrm{d}c_m}{\mu(D_r)}. \tag{4.1.5}$$

Note that the equations (4.1.4) and (4.1.5) - even though in some other notations - is the solution of the problem formulated and solved by algorithm in [*].

From the point of view of practical realisation of my algorithm, that is based of formulae (4.1.4),(4.1.5),

I can ask two questions: 1. Is it possible to derive a simple method of constructing the sets D_r; 2. Is it possible to derive a simple method of calculation the integrals of those sets D_r, which appear in the RHS of (4.1.5). What a geometrical construction (Figure [*]) has shown me, is that even in the case where $m = 2$ and $n = 10$, the regions D_r have a relatively complicated geometry (some multi-peaked star-like shapes). However, this problem can be solved with the principal of the 'Gordian knot' - refuse to work with directly with sets D_r, but instead, to use their point (grid) approximation (even without describing these sets concretely) and use an analog of (4.1.4).

The essence of this suggested method of thinking is as follows. Define $k(1), k(2), \dots, k(m)$ as some quite large real values, and a m-dimensional parallelepiped

$$W = \{c : c_j^{(\min)} \le c_j \le c_j^{(\max)}, \quad j = 1, 2, \dots, m\} \subset \boldsymbol{R}^m \tag{4.1.6}$$

which contains the set D (if the set D itself is a parallelepiped, then $W = D$). Cover this parallelepiped with a dense grid Γ with

$$L = \prod_{t=1}^{m} k(t) \tag{4.1.7}$$

number of nodes $c^{(l)} = (c_1^{(l)}, (c_2^{(l)}, \dots, c_m^{(l)}), \quad l = 1, 2, \dots, L$, where each coordinate $c_j^{(l)}, \quad j = 1, 2, \dots, m$, could independently from all the other $m - 1$ coordinates (note that $m - 1$ does not refer to sets of coordinates) assume one of $k(j)$ of values

$$c_j = c_j^{(\min)} + \frac{c_j^{(\max)} - c_j^{(\min)}}{k(j)}(t - \frac{1}{2}), \quad t = 1, 2, \dots, k(j). \tag{4.1.8}$$

Parallel to grid Γ implement the system A of parallelepipeds $W_l \subset \boldsymbol{R}^m, \quad l = 1, 2, \dots, L$, the centres of which are their nodes with sides

$$h_j = \frac{(c_j^{(\max)} - c_j^{(\min)})}{k(j)} : \tag{4.1.9}$$

$$W_l = \{c = (c_1, c_2, \dots, c_m) : c_j^{(l)} - \frac{1}{2}h_j \le c_j \le c_j^{(l)} + \frac{1}{2}h_j, \quad j = 1, 2, \dots, m\}. \tag{4.1.10}$$

These parallelepipeds form the surface of the parallelepiped W. Now we have to chose from the defined L nodes those, the ones who belong to the set D. Let L_0 be the quantity of the chosen nodes. Renumber the chosen nodes, by assigning the first L_0 numbers l to them.

For each $l = 1, 2, \dots, L_0$ find elements of $q_i(c^{(l)}), \quad i = 1, 2, \dots, n$ of sequence (4.0.5), and based on their values, using (4.0.6), calculate the number r of transition of sign of elements of (4.0.5). Let $J_r, \quad r = 0, 1, \dots, n-1$ be the set of numbers l of chosen nodes $c^{(l)}$ providing r transitions of sign, and $|J_r|$ - the number of nodes giving that many transitions. In this way, without directly constructing the sub-sets D_r, I approximate each of them with the union

$$\tilde{D}_r = \bigcup_{l \in J_r} W_l. \tag{4.1.11}$$

It is clear, that the assumption, that all vectors c, belonging to the parallelepiped W_l, will provide the same number of transitions of sign as the center c^l of the parallelepiped, will only be incorrect for those parallelepipeds W_l, and do not completely belong to some one sub-set D_r. But with a dense grid Γ, and respectively, quite small-sized parallelepipeds $W_{(l)}$, the percent of such 'debatable' parallelepipeds compared to their general number L_0 will be quite small (so, in figure [*], when $m = 2$, these 'debatable' parallelepipeds are positioned along the lines, and the 'non-debatable' - in the areas). In addition, the error resulting from classification of such 'debatable' parallelepipeds W_l completely to some set D_r, will get (mostly) redeemed/repaid. Summing up all these arguments, it can be said, that the approximations of sets D_r, with the union of relatively 'small' parallelepipeds will give a

good approximation of the mathematical expectancy (4.0.11) of error in estimation of parameters of a studies functional dependence:

$$E(\rho_2(c^*,c)) \approx \sum_{r\in I}\frac{p_r^*}{\mu(\tilde{D}_r)}\int\limits_{\tilde{D}_r}\sum_{j=1}^{m}(c_j^*-c_j)^2\mathrm{d}c_1\mathrm{d}c_2\ldots\mathrm{d}c_m =$$

(integral with limit D_r can be split as the sum of integrals with limit W_l)

$$\sum_{r\in I}\frac{p_r^*}{\mu(\tilde{D}_r)}\sum_{l\in J_r}\int\limits_{W_l}\sum_{j=1}^{m}(c_j^*-c_j)^2\mathrm{d}c_1\mathrm{d}c_2\ldots\mathrm{d}c_m = \tag{4.1.12}$$

(the sum $\sum_{j=1}^{m}$ is facrotised out)

$$\sum_{j=1}^{m}\left[\sum_{r\in I}\frac{p_r^*}{\mu(\tilde{D}_r)}\left(\sum_{l\in J_r}\int\limits_{W_l}((c_j^*)^2-2c_j^*c_j+c_j^2)\right)\mathrm{d}c_1\mathrm{d}c_2\ldots\mathrm{d}c_m\right] =$$

(expanding $((c_j^*)^2-2c_j^*c_j+c_j^2)$)

$$\sum_{j=1}^{m}\left[\sum_{r\in I}\frac{p_r^*}{\mu(\tilde{D}_r)}\left(\sum_{l\in J_r}\left[(c_j^*)^2\mu(W_{(l)})-2c_j^*\int\limits_{W_l}c_j\mathrm{d}c_1\mathrm{d}c_2\ldots\mathrm{d}c_m+\int\limits_{W_l}c_j^2\mathrm{d}c_1\mathrm{d}c_2\ldots\mathrm{d}c_m\right]\right)\right]. \tag{4.1.13}$$

If in a multi integral, the limits of integration for each variable are given by other variables, then this multi integral is equal to the product of included in it, one-dimensional integrals [*]. Now, if we define coordinates of the edges of W_l as $c_j^{(l)}-0.5h_j=a(j,l)$ and $c_j^{(l)}+0.5h_j=b(j,l)$ then it is not hard to be sure that

$$\frac{1}{\mu(W_l)}\int\limits_{W_l}c_j\mathrm{d}c_1\mathrm{d}c_2\ldots\mathrm{d}c_m=\frac{1}{\prod\limits_{s=1}^{m}h_s}\left(\int\limits_{b(j,l)}^{a(j,l)}c_j\mathrm{d}c_j\right)\prod_{\substack{s=1\\(s\neq j)}}^{m}\int\limits_{a(s,l)}^{b(s,l)}\mathrm{d}c_s=c_j^{(l)}. \tag{4.1.14}$$

If in addition we accept that

$$\sum_{r\in I}p_r^*=1, \tag{4.1.15}$$

$$\sum_{l\in J_r}\mu(W_l)=\mu(\tilde{D}_r), \tag{4.1.16}$$

$$\mu(W_l)\equiv\prod_{j=1}^{m}h_j, \tag{4.1.17}$$

then the simplification chain (4.1.13) could be continued

$$E(\rho_2(c^*,c)) \approx \sum_{j=1}^{m}\left[\sum_{r\in I}\frac{p_r^*}{\mu(\tilde{D}_r)}\left(\sum_{l\in J_r}\left[(c_j^*)^2\mu(W_l)-2c_j^*\mu(W_l)c_j^{(l)}+\int\limits_{W_l}c_j^2\mathrm{d}c_1\mathrm{d}c_2\ldots\mathrm{d}c_m\right]\right)\right] =$$

$$\left(\text{ middle step: } \sum_{l\in J_r}^{m}(c_j^*)^2\mu(W_l) = (c_j^*)^2\underbrace{\sum_{l\in J_r}\mu(W_l)}_{=\mu(\tilde{D}_r)} = (c_j^*)^2\mu(\tilde{D}_r) \ \right)$$

$$\sum_{j=1}^{m}\left[(c_j^*)^2\sum_{r\in I}p_r^* - 2c_j^*\sum_{r\in I}\frac{p_r^*}{\mu(\tilde{D}_r)}\sum_{l\in J_r}\mu(W_l)c_j^{(l)} + \sum_{r\in I}\frac{p_r^*}{\mu(\tilde{D}_r)}\sum_{l\in J_r}\int\limits_{W_l}c_j^2\mathrm{d}c_1\mathrm{d}c_2\ldots\mathrm{d}c_m\right] =$$

$$\left(\text{ middle step: } \sum_{r\in I}(c_j^*)^2\frac{p_r^*}{\cancel{\mu(\tilde{D}_r)}}\cancel{\mu(\tilde{D}_r)} = (c_j^*)^2\underbrace{\sum_{r\in I}p_j^*}_{=1} = (c_j^*)^2 \ \right)$$

$$\sum_{j=1}^{m}\left[(c_j^*)^2 - 2c_j^*\sum_{r\in I}p_j^*\frac{1}{\sum\limits_{l\in J_r}\mu(W_l)}\sum_{l\in J_r}\mu(W_j)c_j^{(l)} + \sum_{r\in I}\frac{p_r^*}{\mu(\tilde{D}_r)}\sum_{l\in J_r}\int\limits_{W_l}c_j^2\mathrm{d}c_1\mathrm{d}c_2\ldots\mathrm{d}c_m\right] =$$

$$\sum_{j=1}^{m}\left[(c_j^*)^2 - 2c_j^*\sum_{r\in I}p_j^*\frac{1}{|J_r|\prod\limits_{s=1}^{m}h_s}\sum_{l\in J_r}\left(\prod_{s=1}^{m}h_s\right)c_j^{(l)} + \sum_{r\in I}\frac{p_r^*}{\mu(\tilde{D}_r)}\sum_{l\in J_r}\int\limits_{W_l}c_j^2\mathrm{d}c_1\mathrm{d}c_2\ldots\mathrm{d}c_m\right] =$$

$$\sum_{j=1}^{m}\left[c_j^*)^2 - 2c_j^*\sum_{r\in I}\frac{p_j^*}{|J_r|}\sum_{l\in J_r}c_j^{(l)} + \sum_{r\in I}\frac{p_r^*}{\mu(\tilde{D}_r)}\sum_{l\in J_r}\int\limits_{W_l}c_j^2\mathrm{d}c_1\mathrm{d}c_2\ldots\mathrm{d}c_m\right]. \tag{4.1.18}$$

In order to find vector c^*, which provides the minimum estimate mathematical expectation (4.1.18) of error in the approximation of unknown true parameters of the functional dependence in question, and therefore also of vector $\dot{c}$, we take the same steps as in the case when we calculated the minimum of function (4.1.2). Taking into account, that the last, third, summand in (4.1.18) does not contain any variables c_j^* (which are being optimised), its derivative is again equal to zero, we find its partial derivatives:

$$\frac{\partial E(\rho_2(c^*,c))}{\partial c_j^*} \approx 2c_j^* - 2\sum_{r\in I}\frac{p_r^*}{|J_r|}\sum_{l\in J_r}c_j^{(l)}, \quad j=1,2,\ldots,m. \tag{4.1.19}$$

Equating these partial derivatives (4.1.19) to zero we get, and therefore conclude with:

$$c_j^* = \sum_{r\in I}\frac{p_r^*}{|J_r|}\sum_{l\in J_r}c_j^{(l)}, \quad j=1,2,\ldots,m. \tag{4.1.20}$$

This equation (4.1.20) is the one used in calculations of approximate values of the optimised vector c^* in my custom software, the trails of which are later in this document.

Now, let's compare reviousy suggested in [2] equations (4.1.4) and (4.1.5) with the suggested equation (4.1.20). The superiority of the later is shown in that: 1. it does not require the direct construction of sub-sets D_r , which in fact are quite complex convex polyhedrons [*] in $\boldsymbol{R}^m$ dimensions. Therefore 2. this equation does not require the expansive calculation of multi integrals found in (4.1.4) and (4.1.5). In [2] there is one more problem at hand: even a polygon, not to mention a polyhedron, can not be defined only by its vertexes (those could be connected with line segments in not just one way). The algorithm proposed in this IA to estimate the optimal parameters c^*, only requires the regular (evenly distributed) surface of set D, for the nodes of which the only thing left to do is evaluate the elements of the sequence 4.0.5.

5 Algorithm trials

6 Conclusion and Reflection

First of I want to to conclude with mentioning again the conditions in which this method performs best, and subsequently when I would recommend using it. This method approximates, significantly better than LSM per se (and, it could be argued, that better than LAD), in conditions where among measured values $_i$ there is a significant percentage of such of those values that their absolute value is relatively small (and therefore they are close to the studied functional dependence). We can't know which of them are so, and we don't have to, we just need to know that such values of $_i$ exist. Said that, some values $_i$ can peak quite significantly in one of the signs (+ve or -ve), without the approximation begin significantly affected. It is clear, that all conditions in which this method performs best can not be defined (the limited nature of calculation accounts for this), however, the trails given give a good representation of the success of this method. And thus I have reached my aim.

This is to say, that this investigation is far from perfect. Even though begin more constructive than previous work [2] done in this topic , there are some holes. For example, the algorithm requires that a finite interval $[\alpha, \beta]$ is given to each parameter estimated, a computer can't run a calculation infinitely.

I want to to note that the trials have shown some results that I have not expected. For example, some individual trails (given in the appendix of this IA) have either unreasonably low of high factual error (error actually made by the method). This could be caused by some random 'layout' of the values of $_i$ being either in favor or against a good approximation with any method (say when all or close to all absolute values of $_i$ are large) or a bug in my code (but hopefully not that).

In spite that, now, after writing and researching for this IA, I know that not only I have questioned the validity of a method of approximation. In it's own way deriving methods of approximation is its own mathematical field. Some methods work better than others in different circumstances. Just like the method described in this IA is suited for use even when there are peaks in the given data, other methods deal with other circumstances, like for example LSM, is suited for circumstances without those peaks. And so, like in most fields in mathematics, this is a continuation from what came before, and could (and should) be continued (by others) to find even better (working in more conditions) algorithms of approximation.

7 Bibliography

References

[1] *Summation and Product Notation*. Illinois State University Mathematics Department, 1995.

[2] PI Balk. *Regression-type problems under zero median noise*, volume 81. 2010.

[3] P. C. Matthews. *Vector calculus*. Springer, 1998.

[4] James Stewart. *Calculus: early transcendentals*. Thomson Brooks/Cole, 2008.