Least squares for generalized Gauss–Laplace distribution of the error in certain nonlinear regressions with perpendicular offsets

Carmen E. Stoenoiu[1] and Lorentz Jäntschi[2]

[1] Department of Electric Machines and Drives, Technical University of Cluj-Napoca, 26-28 Bariţiu Str., 400027 Cluj-Napoca, Cluj, Romania,
carmen.stoenoiu@emd.utcluj.ro, http://carmen.academicdirect.ro
[2] Department of Physics and Chemistry, Technical University of Cluj-Napoca, 103-105 Muncii Blvd., 400641 Cluj-Napoca, Cluj, Romania,
lori@chimie.utcluj.ro, http://lori.academicdirect.org

Abstract. There are mainly two methods of calculating the parameters of the regression equations: the minimization of the squared errors (or least squares, LS) and the maximization of the likelihood. Regarding the distribution of the error of experimental observations, there are several theoretical distributions, but two of them are on the one hand better known and on the other easily generalizable into one (Gauss–Laplace, GL): the normal (or Gaussian) distribution and the double exponential (or Laplace) distribution. In the construction of the squared errors is possible to replace the classical vertical offsets (which are the sides of the squares of the errors) with perpendicular ones, more suited when all variables are equally subjected to experimental errors. In the present work, it is proposed to use an iterative algorithm for the calculation of the regression parameters using LS of perpendicular offsets under the assumption of GL distributed error. The method is exemplified on a non-linear regression model.

Keywords: parameters estimation, least squares, generalized normal distribution, nonlinear regression, perpendicular offsets

1 Introduction

In physics, chemistry and biology and applied sciences symmetric distributions play an essential role [1]. In physics, if it remains unchanged by some space-time transformation, the thing is symmetric, so all conservation principles are examples of symmetry. In chemistry take for instance the water molecule, in biology the radial symmetry of stems, and in medical sciences the bilateral symmetry of body parts. Pillars of logics (true/false logical complements), probability theory (uniform distribution) and statistics (min. vs. max.) are build on symmetry.

Generalized Gauss-Laplace (GL) distribution is a symmetric tri-parametric distribution (parameters: location, scale, shape) for which the parameter estimation via maximum likelihood and the method of moments have been reported in

[2]. Unfortunately the parameters estimates do not have a closed form and must be obtained numerically. Here is used a modified version of it [3] expressed as a natural extension from Gauss's [4] and Laplace's [5] symmetric distributions.

The least squares (LS) method was introduced in the context of identifying the parameters of elliptical planetary motion in [6] and reinvented under a different name in [7]. The classical approach of least squares (named ordinary LS) is to employ vertical offsets [8], which means that the side of the square of the sum of the squares of the errors is the distance between the observed and the estimated values measured on the axis (usually depicted vertically) representing the dependent variable. Unfortunately for the nonlinear case, the parameters estimates do not have a closed form and must be obtained numerically.

When the observations come from an exponential family with identity as its natural sufficient statistics and mild-conditions are satisfied (as in binomial, exponential, normal and Poisson distributions), least-squares and maximum-likelihood estimates are identical [9].

When perpendicular instead of vertical offsets are drawn, the residual is normal to the model [10]. Furthermore, by construction, perpendicular offsets makes the smallest squares and as such maximizes the likelihood.

One interesting fact is that the connection between the perpendicular offsets and the maximization of the likelihood has been discovered [11] even before the formulation of the maximum likelihood method [12, 13]. The solution employing perpendicular offsets for simple linear regression was communicated for the first time in [14] and revised in [15]. The parameters estimates from the perpendicular offsets of the nonlinear regression once again do not have a closed form and must be obtained numerically.

In this work, the least squares method (the correct formulation in this case is least hypercubes) is employed to minimize the residuals constructed from perpendicular offsets in the case of nonlinear regression, assuming that they (residuals) follow a Gauss - Laplace distribution. On a series of paired data coming from current and voltage simultaneous measurements involving a photovoltaic cell a nonlinear regression model was applied and the results are discussed.

2 Material and method

Let's consider a general nonlinear model in two variables (x and y) in an explicit form of it (Eq. (1)).

$$y \sim f(x; c) \tag{1}$$

where c are known coefficients as an array of m values (c_0, ..., c_{m-1}) intended to be used to obtain a best fit with a series of n paired observed data ((x_0, y_0), ..., (x_{n-1}, y_{n-1})). In Eq. 1 semicolon express that the c values are fixed in an deterministic way ($\sim$) usually to provide a best fit with the observed data. In the context used in this paper x and y are one-dimensional (univariate) but the following treatise can be extended to apply for multi-dimensional (multivariate) case. In the classical vertical offsets the residuals are $y_0 - f(x_0; c)$, ..., $y_{n-1} - f(x_{n-1}; c)$.

Let be g such that ($g(x, y, z; c)$, Eq. 2):

$$g(x, y, z; c) = \sqrt{(x - z)^2 + (y - f(z; c))^2} \tag{2}$$

One approach to get the perpendicular offsets in each of the observation points ((x_k, y_k) for $k = 0, ..., n - 1$) is to find z_k for f, c and y_k given, for which $g(x_k, y_k, z_k)$ is minimum [10].

One should notice that in this instance $g(x_k, y_k, z_k)$ is the residual absolute error (ϵ_k, Eq. 3).

$$\epsilon_k = g(x_k, y_k, z_k; c) \tag{3}$$

Formally z_k values can be added to the (x_k, y_k) pair, but on the other hand, (x_k, y_k) is the pair of observed values while z_k values depends on the values of c coefficients (and by that in any estimation of them from experimental data on all paired values). Let's take a look to the case in which the perpendicular offset is minimum (Fig. 1).

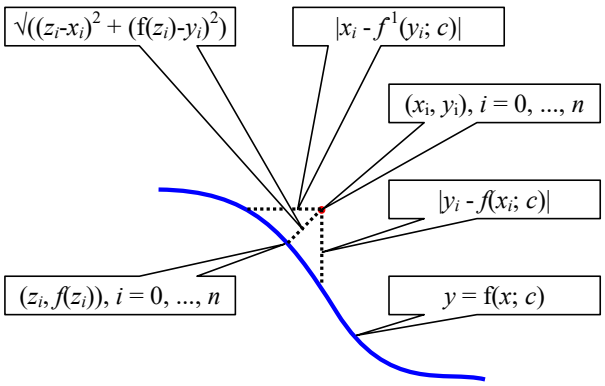

Fig. 1. Perpendicular offset is locally the shortest path from the observation point to the curve

The case illustrated in Fig. 1 doesn't apply for functions with sudden variations. However, there is an alternate route approximating the size of the perpendicular offset when the inverse of the function is available (Eq. 4).

$$\epsilon_k = \frac{s_h s_v}{\sqrt{s_h^2 + s_v^2}}, s_h = |x_k - f^{-1}(y_k; c)|, s_v = |y_k - f(x_k; c)| \tag{4}$$

GL distribution can be expressed [16] in terms of error standard deviation σ and power of the errors q by using the Gamma function (Γ) with Eq. 5.

$$GL(\epsilon, \sigma, q) = \frac{q}{2\sigma} \frac{(\Gamma(3/q))^{1/2}}{(\Gamma(1/q))^{3/2}} \exp\left(-\frac{|\epsilon|^q}{\sigma^q} \left(\frac{\Gamma(3/q)}{\Gamma(1/q)}\right)^{q/2}\right) \tag{5}$$

For convenience, a simple check reveals the two particular casses, Gaussian and Laplace distributions given in Eq. (6), where μ is the population mean.

$$\begin{aligned} GL(x-\mu,\sigma,2) &= \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{|x-\mu|^2}{2\sigma^2}\right), \\ GL(x-\mu,\sigma\sqrt{2},1) &= \frac{1}{2\sigma}\exp\left(-\frac{|x-\mu|}{\sigma}\right) \end{aligned} \tag{6}$$

With the data from [10] function f from Eq. (7) proven to provide a very good fit.

$$f(x;c) = c_0 - \exp(-c_1 + c_2\ln(x)) \tag{7}$$

The same function is considered here. It's inverse is given in Eq. (8):

$$f^{-1}(x;c) = \exp\left(\frac{\ln(c_0-y)+c_1}{c_2}\right) \tag{8}$$

The residual error (RSQ) from GL subjected to minimization can be considered the one from Eq. (9):

$$RSQ = \left(\sum_{k=0}^{n-1}\epsilon_k^q\right)^{q^{-1}} \tag{9}$$

With a powerfull nonlinear optimization library, from hereon, only the statement of the problem in the formalism of the library must be given. In the experience of the author, such libraries are available in Mathcad [17], Mathematica [18], Matlab [19] and open libraries (such as AlgLib [20]) usable under a programming environment (such as FreePascal in [10]). Here only the general procedure is given (for the algorithms to be slightly adapted see [10]):

- Do the initial estimates for the values of the parameters C from classical Gaussian vertical offsets of the selected model (Eq. (7)) of nonlinear regression;
- Add q to the list of parameters ($q \leftarrow 2$ initializes it);
- Run the minimization for RSQ (Eq. (9)) with ϵ_k from Eq. (4) or Eq. (3).

3 Results and discussion

The optimization should be feed with some initial data. The initial data can be taken from nonlinear regression with vertical offsets and Gaussian distribution of the error ($q = 2$). Using the experimental data given in [10], the values given in Tab. 1 were used for the initialization of the optimization.

In Tab. 1 also a series of solar cell parameters are derived: short circuit intensity (I_{sc}), open circuit voltage (U_{oc}), current at maximum power point (I_{xp}), voltage at maximum power point (U_{xp}), and power at maximum power point (P_{xp}).

Table 1. Initial estimates for the coefficients.

n	Initial C	Statistics	PV cell parameters
3	$c_0 = 1.83$ $c_1 = 22$ $c_2 = 3.07$ $q = 2.0$	$RSS = 0.0019093$ $r^2_{\text{adj}} = 0.9986$ $F = 9228$	$I_{\text{sc}} = 1.83000$ mA $U_{\text{oc}} = 1576.51$ mV $I_{\text{xp}} = 1.3804$ mA $U_{\text{xp}} = 998.00$ mV $P_{\text{xp}} = 1.3776$ mW

$I_{\text{sc}} = f(0;c)$; $U_{\text{oc}} = f^{-1}(0;c)$; $U_{\text{xp}} = u \,|\, (u{\cdot}f(u;c))'_u = 0$; $I_{\text{xp}} = f(U_{\text{xp}};c)$; $P_{\text{xp}} = U_{\text{xp}} I_{\text{xp}}$

Fig. 2 provides the evolution of parameters values with the optimization advancement starting from Tab. 1 data when Eq. (4) is used.

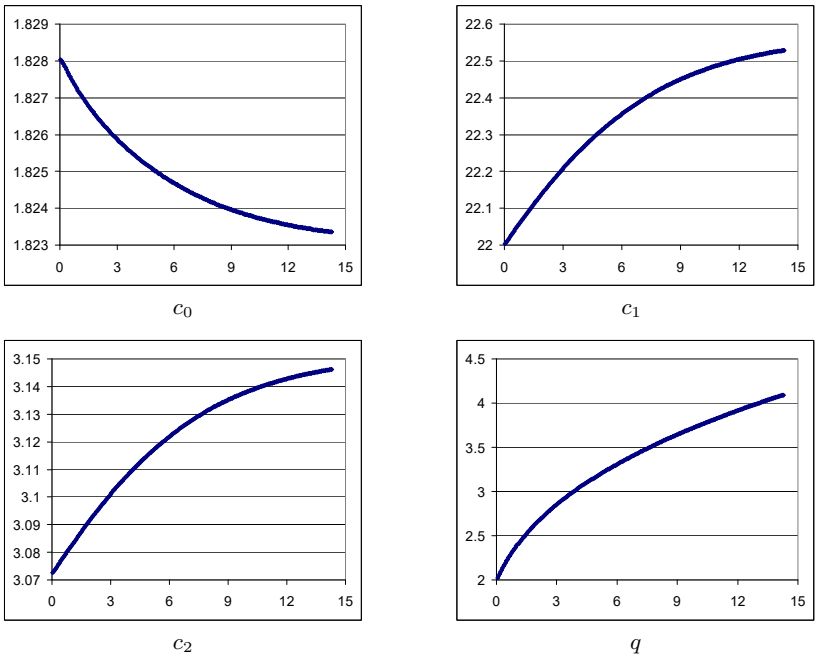

Fig. 2. Parameters of the model (c_0, c_1, c_2) and the power of the error (q) as functions of milions of iterations until convergence with Eq. (4)

Convergence is relatively fast when compared with the alternate algorithm [10] considering that here the double embedded optimization has been reduced to a single optimization (the inner optimization has been replaced by a simple analytic calculation). However, in terms of the number of steps the optimization has been increased in complexity - which is the expected outcome since a new parameter has been added - q - the power of the error (see Eq. (5)).

Optimization is smooth to the convergence (see Fig. 3).

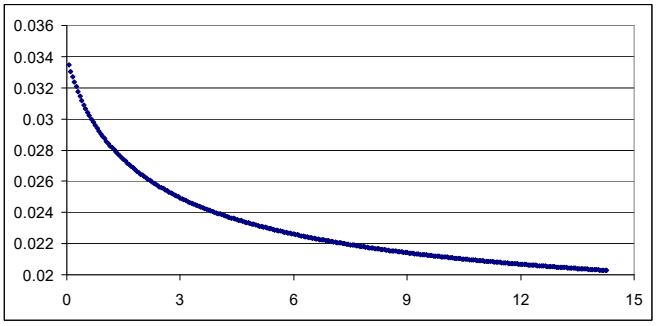

Fig. 3. Residual error RSQ as function of milions of iterations until convergence approximating ϵ_k with Eq. (4)

The final optimized values of the parameters using Eq. (4) are given in Tab. 2.

Table 2. Final (Eq. (4) optimum) estimates for the coefficients (after 14280311 iterations starting from Tab. 1 data).

n	Final C	Statistics	PV cell parameters
3	$c_0 = 1.82335$ $c_1 = 22.5285$ $c_2 = 3.14617$ $q = 4.09287$	$RSS = 0.001163$ $r^2_{\text{adj}} = 0.9986$ $F = 9499$	$I_{\text{sc}} = 1.82335$ mA $U_{\text{oc}} = 1558.56$ mV $I_{\text{xp}} = 1.3836$ mA $U_{\text{xp}} = 994.93$ mV $P_{\text{xp}} = 1.3766$ mW

I_{sc}, U_{oc}, U_{xp}, I_{xp}, P_{xp} derived in same manner as in Tab. 1

If the optimization did not increased the adjusted determination coefficient, it significantly decreased the residual error (RSS, from 0.0019 in Tab. 1 to 0.0012 in Tab. 2). Also the total variance explained by the model has been increased (F value is 9228 in Tab. 1 and 9499 in Tab. 2). Some of the parameters suffered some change after the optimization (take for instance open circuit voltage, with more than 1 % change) a very little change (take for instance the maximum power, with less than 1 ‰ change).

When compared to the previously reported optimization (Tab. 3 in [10]), the statistics did not bring more explanatory power (take for instance F value, which is 9528 in [10] and 9499 reported here).

The alternate (to Eq. (4)) approach is to effectivelly find (by optimization) the normal to the courve (find z_k such that ϵ_k from Eq. (3)) to be minimum. The results regarding the evolution of the values of the parameters staring from the same initial values from Tab. 1 are provided in Fig. 4.

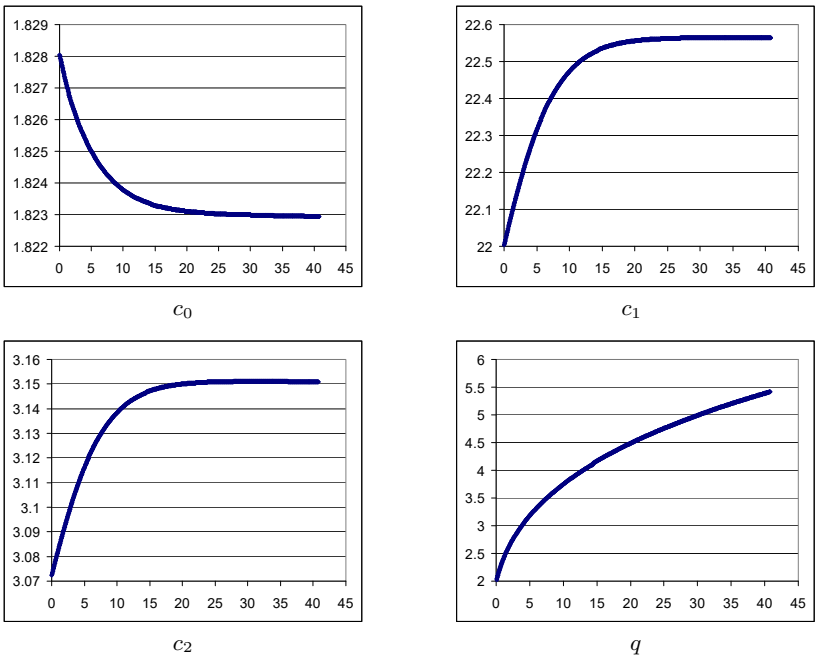

Fig. 4. Parameters of the model (c_0, c_1, c_2) and the power of the error (q) as functions of milions of iterations until convergence with Eq. (3)

Optimization is smooth to the convergence (see Fig. 5).

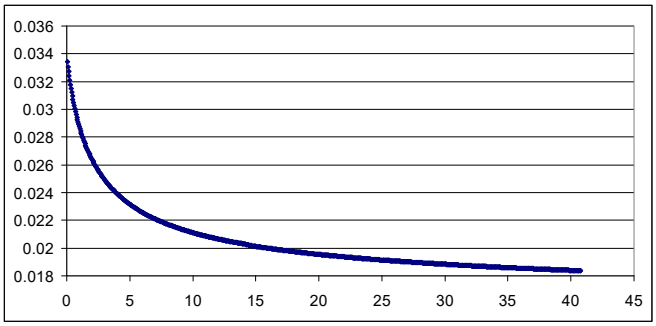

Fig. 5. Residual error as function of milions of iterations until convergence calculating ϵ_k with Eq. (3)

A significantly higher number of iterations were necessary to get the exact solution (over 40 millions interations in Fig. 5, less than 15 millions iterations in Fig. 3). The exact solution is even more time consumming than the ratio of the

two numbers shows, considering that to get the exact values of the perpendicular oofsets is another problem of optimization (for a detailed discussion please see [10]).

The final optimized values of the parameters using Eq. (3) are given in Tab. 3.

Table 3. Final (Eq. (3) optimum) estimates for the coefficients (after 40764550 iterations starting from Tab. 1 data).

n	Final C	Statistics	PV cell parameters
3	$c_0 = 1.82294$ $c_1 = 22.5634$ $c_2 = 3.15092$ $q = 5.14599$	$RSS = 0.001190$ $r^2_{\text{adj}} = 0.9986$ $F = 9477$	$I_{\text{sc}} = 1.82294$ mA $U_{\text{oc}} = 1558.48$ mV $I_{\text{xp}} = 1.3838$ mA $U_{\text{xp}} = 992.01$ mV $P_{\text{xp}} = 1.3727$ mW

Tab. 3 data reveals that the statistics for the regression optimum model in the supposition of generalized Gauss-Laplace distribution of the error have RSS and F values smaller than the values calculated under the supposition of Gaussian distribution of the error [10]. Even more, it can be mathematically proved that also r^2_{adj} from generalized Gauss-Laplace distribution supposition is less or equal with r^2_{adj} from Gaussian supposition, and this is simply because everywhere in these formulas (RSS, r^2_{adj}, and F) is about squared quantities, like in Gaussian distribution. This expected outcome does not imply that the obtained regression model is less performant, since the two models are derived in two different statistical suppositions. On the same argument, the residual sums of the Eq. (9) form with the optimum coefficients from Gaussian supposition are bigger (or equal when $q = 2$) than the residuals obtained with the optimum coefficients from generalized Gauss-Laplace supposition (for the sake of the argument, the residual sums of the Eq. (9) form, when q is 5.14599 - with the coefficients from Tab. 3 in [10] is 0.019428 while the same with our Tab. 3 coefficients - also represented in Fig. 5 - is 0.018369).

4 Conclusions

Generalized Gauss-Laplace distribution of the experimental error is always able to reduce further the residual error (by adding also one more parameter to be estimated, the power of the resuidual error), but as has been shon in the considered example case this addition does not always bring more explanatory power.

A negative result is also communicated here - about the simplification of the route to the perpendicular offets by the use of the approximation of them as the height of the triangle formed by the vertical and the horrizontal offsets (Eq. (4)

in the paper) - even in a case with smooth variations and points close to the model, the approximation do not accuarely provide the optimum point.

The main result reported here is that a iterative optimization has been succesfully applied to the problem of generalized Gauss-Laplace distribution of the error in a case of paired data sets when both paired variables are affected by errors and perpendicular offsets were involved to identify the parameters for a nonlinear regression model.

Authors Statements

Author Contributions C.E.S.: formal analysis, investigation, resources, validation, writing—original draft preparation; L.J.: writing—review and editing, conceptualization, drawings, methodology, supervision.

Funding This research received no external funding.

Data Availability Data used in this paper is fully available online in Table 1 of [10].

Conflicts of Interest The authors declares no conflict of interest.

Acknowledgments The authors wish to express their gratitude to the ICMSQUARE 2023 organizers for providing the oportunity to present this research under the frame of the event. Help from the reviewers in improving the work was highly appreciated.

References

1. Jäntschi, L.: Introducing structural symmetry and asymmetry implications in development of recent pharmacy and medicine. Symmetry **14**(8), 1674 (2022). DOI 10.3390/sym14081674. URL http://doi.org/10.3390/sym14081674
2. Varanasi, M.K., Aazhang, B.: Parametric generalized Gaussian density estimation. J. Acoust. Soc. Am. **86**(4), 1404–1415 (1989). DOI 10.1121/1.398700. URL https://doi.org/10.1121/1.398700
3. Jäntschi, L., Bolboacă, S.D.: Observation vs. observable: maximum likelihood estimations according to the assumption of generalized gauss and laplace distributions. Leonardo Electron. J. Pract. Technol. **8**(15), 81–104 (2009). URL http://lejpt.academicdirect.org/A15/081_104.pdf
4. Gauss, C.O.: Theoria motus corporum coelestium. Perthes et Besser, Hamburg (1809)
5. Laplace, P.: Thèorie analytique des probabilitès. Courcier, Paris (1812)
6. Legendre, A.M.: Sur la méthode des moindres quarrés, pp. 72–75. Firmin Didiot, Paris (1805). URL http://books.google.com/books?id=FRcOAAAAQAAJ

7. Griliches, Z., Ringstad, V.: Error-in-the-variables bias in nonlinear contexts. Econometrica **38**(2), 368–370 (1970). DOI 10.2307/1913020. URL http://doi.org/10.2307/1913020
8. Jäntschi, L., Bolboacă, S.D.: Szeged matrix property indices as descriptors to characterize fullerenes. Ovidius Univ. Ann. Chem. **27**(2), 73–80 (2016). DOI 10.1515/auoc-2016-0010. URL http://doi.org/10.1515/auoc-2016-0010
9. Charnes, A., Frome, E.L., Yu, P.L.: The equivalence of generalized least squares and maximum likelihood estimates in the exponential family. J. Am. Stat. Assoc. **71**(353), 169–171 (1976). DOI 10.1080/01621459.1976.10481508. URL http://doi.org/10.1080/01621459.1976.10481508
10. Jäntschi, L.: Symmetry in regression analysis: Perpendicular offsets–the case of a photovoltaic cell. Symmetry **15**(4), 948 (2023). DOI 10.3390/sym15040948. URL http://doi.org/10.3390/sym15040948
11. Adcock, R.J.: Note on the method of least squares. Analyst **4**(6), 183–184 (1877). DOI 10.2307/2635777. URL http://jstor.org/stable/2635777
12. Fisher, R.A.: On an absolute criterion for fitting frequency curves. Messenger Math. **41**, 155–160 (1912). URL http://hdl.handle.net/2440/15165
13. Fisher, R.A.: On the mathematical foundations of theoretical statistics. Philos. Trans. Royal Soc. A **222**(594–604), 309–368 (1922). DOI 10.1098/rsta.1922.0009. URL http://royalsocietypublishing.org/doi/abs/10.1098/rsta.1922.0009
14. Adcock, R.J.: A problem in least squares. Analyst **5**(2), 53–54 (1878). DOI 10.2307/2635758. URL http://jstor.org/stable/2635758
15. Kummell, C.H.: Reduction of observation equations which contain more than one observed quantity. Analyst **6**(4), 97–105 (1879). DOI 10.2307/2635646. URL http://www.jstor.org/stable/2635646
16. Jäntschi, L., Pruteanu, L.L., Cozma, A.C., Bolboacă, S.D.: Inside of the linear relation between dependent and independent variables. Comput. Math. Methods Med. **2015**, 360,752 (2015). DOI 10.1155/2015/360752. URL http://doi.org/10.1155/2015/360752
17. Jäntschi, L.: A test detecting the outliers for continuous distributions based on the cumulative distribution function of the data being tested. Symmetry **11**(6), 835 (2019). DOI 10.3390/sym11060835. URL http://doi.org/10.3390/sym11060835
18. Jäntschi, L.: Detecting extreme values with order statistics in samples from continuous distributions. Mathematics **8**(2), 216 (2020). DOI 10.3390/math8020216. URL http://doi.org/10.3390/math8020216
19. Joiţa, D.M., Tomescu, M.A., Bàlint, D., Jäntschi, L.: An application of the eigenproblem for biochemical similarity. Symmetry **13**(10), 1849 (2021). DOI 10.3390/sym13101849. URL http://doi.org/10.3390/sym13101849
20. Bochkanov, S.: Alglib project. ©1994-2017. http://alglib.net, last accessed 2023/07/19 (2017)