AppliedMath

Article

Computer-Algebra-Software-Assisted Calculus Instruction, Not Calculus for Dummies: Bespoke Applications Necessitate Theory

Daniel A. Griffith

School of Economic, Political, and Policy Science, University of Texas at Dallas, Richardson, TX 75080, USA; dagriffith@utdallas.edu; Tel.: +1-972-429-4817

Abstract: Today, calculus frequently is taught with artificial intelligence in the form of computer algebra systems. Although these software packages may reduce tedium associated with the mechanics of calculus, they may be less effective if not supplemented by the accompanying teaching of calculus theory. This paper presents two examples from spatial statistics in which computer software in an unsupervised auto-execution mode fails, or can fail, to yield correct calculus results. Accordingly, it emphasizes the need to teach calculus theory when using software packages such as Mathematica and Maple.

Keywords: artificial intelligence; calculus; Maple; Mathematica; spatial statistics

Citation: Griffith, D.A. Computer-Algebra-Software-Assisted Calculus Instruction, Not Calculus for Dummies: Bespoke Applications Necessitate Theory. *AppliedMath* **2022**, 2, 261–268. https://doi.org/10.3390/appliedmath2020016

Academic Editor: Tommi Sottinen

Received: 21 May 2022
Accepted: 2 June 2022
Published: 7 June 2022

1. Introduction

As late as the 1960s, calculus books basically contained relevant applications, examples, and exercises from other areas of mathematics, such as geometry, or from physics. One debate that emerged at that time concerned whether or not subjects, such as statistics and other mathematics-ladened topics, could be taught without calculus [1]. This debate coincided with a diversification of calculus textbook examples embracing a much wider range of applied treatments, which often included relevant applications, examples, and exercises from economics, and then biology and medicine [2–5]. A next debate focused on the use of numerically intensive approaches to teach calculus [6]. A current debate concerns the use of artificial intelligence (AI) and computer technology to teach calculus [7–12].

Today, AI abounds, with its evolution over a number of decades yielding technology for teaching mathematics [13] that includes highly developed and sophisticated symbolic software, such as Mathematica [14] and Maple [15], supporting parts of mathematics such as calculus, especially for computer-aided instruction purposes. Such software is employed to teach the subject because it executes the mechanics of calculus with great success [6]. However, the theory of calculus, for example, needs to accompany the mechanics of calculus when teaching this subject. The mid-1990s Intel chip fiasco [16] serves as a prime warning against blindly using AI software for mathematics.

This paper presents an example from spatial statistics [17] in which software defaults render incorrect calculus results. The two employed illustrations represent two of the most commonly encountered situations in spatial statistics, highlighting the potential seriousness of such mistakes. Introducing a spatial statistics problem in this context contributes to the ongoing diversification of relevant applications, examples, and exercises that should be included in calculus textbooks. Nevertheless, regardless of teacher or student awareness of spatial statistics, the serious warning that these cited examples signal merits recognition by teachers and students of calculus alike.

2. Materials and Methods: Spatial Statistics in a Nutshell

Spatial statistics is a specialized branch of statistics similar to time series analysis. However, rather than a sequence across time, spatial statistics involves a sequence across a two-dimensional surface. Accordingly, a focus on spatial autocorrelation, or the correlation among nearby values on a two-dimensional map that arises from their relative locations inducing (dis)similar attribute values to cluster on maps, replaces a focus on serial correlation, or the correlation among an entity's values that are separated by uniformly spaced intervals of time portrayed in a time series scatterplot. Its critical feature is that some response variable, say Y, is on both sides of an equation's equal sign (=), just like in time series analysis. Whereas the left-hand side of an equation contains y_i, the right-hand side of the same equation contains a linear combination of the y_j ($i \neq j$) values; in other words, using matrix notation, **WY**, where **Y** is an n-by-1 vector of y_i values, and **W** is an n-by-n matrix (called a spatial weights matrix) representing the configuration of n locations tied together in a two-dimensional map (e.g., like a bunch of grapes rather than a statistician's separate balls in an urn). This matrix relates to graphic theory adjacency matrices. The analogous matrix, say **C**, for time series, whose values are ordered from the present to the past, is an upper off-diagonal containing ones, and zeroes in all other matrix cells. In both cases, all diagonal entries of the matrices are zero. Matrix **W** also has mostly zeroes. Its positive cell entries are for row and column location labels that are considered geographic neighbors. A matrix **C** with binary 0–1 cell entries may be defined first, similar to that for a time series, but now symmetric rather than simply an upper off-diagonal with entries of one. Next, this matrix **C** can be row-standardized, yielding matrix **W**. This is the most common spatial weights matrix specification in practice and is accompanied by a parameter indexing the nature and degree of spatial autocorrelation latent in geographically distributed attribute data whose positive range is [0, 1) because the maximum eigenvalue of matrix **W** is 1.

Besag [18] coined the term auto-model with regard to probability models specified to account for spatial autocorrelation. The auto-normal is the most widely used of these specifications and furnishes the illustrations for this paper. One version of this specification is the simultaneous autoregressive (SAR) model, whose auto-normal likelihood function may be written as, using matrix notation,

$$L = (2\pi)^{-n/2}\left|(\mathbf{I}-\rho\,\mathbf{W})^{T}(\mathbf{I}-\rho\,\mathbf{W})\right|^{1/2}\left(\sigma^{2}\right)^{-n/2}e^{-(\mathbf{Y}-\mu\mathbf{1})^{T}(\mathbf{I}-\rho\,\mathbf{W})^{T}(\mathbf{I}-\rho\,\mathbf{W})(\mathbf{Y}-\mu\mathbf{1})/(2\sigma^{2})} \quad (1)$$

where n is the number of geographic locations, ρ is the spatial autoregressive parameter in a single-parameter model specification, μ is the constant mean, σ^2 is the constant variance of random variable **Y**, **I** is an n-by-n identity matrix, **1** is an n-by-1 vector of ones, and superscript T denotes the matrix transpose operator.

Maximum likelihood techniques [19] furnish parameter estimates for Equation (1) and are one instance in which spatial statistics and calculus interface. The first partial derivative set equal to 0 yields a maximum likelihood estimate (MLE); the second partial derivative determines if this estimate is a maximum (i.e., the second derivative test). MLEs for μ and σ^2 in Equation (1) are

$$\hat{\mu} = \mathbf{1}^{T}(\mathbf{I}-\rho\,\mathbf{W})\mathbf{Y}/[n(1-\rho)], \text{and} \quad (2)$$

$$\hat{\sigma}^{2} = (\mathbf{Y}-\mu\mathbf{1})^{T}(\mathbf{I}-\rho\,\mathbf{W})^{T}(\mathbf{I}-\rho\,\mathbf{W})(\mathbf{Y}-\mu\mathbf{1})/n. \quad (3)$$

Substituting these estimates into the likelihood function renders the following minimization (i.e., nonlinear regression) problem:

$$\text{MIN}: \; e^{-\sum_{i=1}^{n} LN(1-\rho\,\lambda_j)^2/n}(\mathbf{Y}-\mu\mathbf{1})^{T}(\mathbf{I}-\rho\,\mathbf{W})^{T}(\mathbf{I}-\rho\,\mathbf{W})(\mathbf{Y}-\mu\mathbf{1}), \quad (4)$$

where λ_j (j = 1, 2, . . . , n) is the *j*-th rank ordered of the n eigenvalues of matrix **W**. Consequently, the maximum likelihood estimation is implemented as the following nonlinear least squares regression problem:

$$\frac{\mathbf{Y}}{e^{\sum_{i=1}^{n} \mathrm{LN}(1-\rho\ \lambda_j)/n}} = \frac{\rho\mathbf{WY}}{e^{\sum_{i=1}^{n} \mathrm{LN}(1-\rho\ \lambda_j)/n}} + \mu\frac{(1-\rho)\mathbf{1}}{e^{\sum_{i=1}^{n} \mathrm{LN}(1-\rho\ \lambda_j)/n}} + \frac{\boldsymbol{\varepsilon}}{e^{\sum_{i=1}^{n} \mathrm{LN}(1-\rho\ \lambda_j)/n}}, \quad (5)$$

where $\boldsymbol{\varepsilon} \sim \mathrm{N}(\mathbf{0},\ \mathbf{I}\sigma^2)$, with N denoting the normal probability distribution. A considerable number of spatial statistics applications involve estimating the parameters of specifications such as Equation (5).

3. Results

This section presents two spatial statistical examples that demonstrate the argument put forth in this paper.

3.1. A Differential Calculus Example: An Incorrect Derivative

Most geographically distributed socio-economic and demographic variables yield a SAR spatial autocorrelation parameter estimate in the interval [0.4, 0.6]. Accordingly, the case of ρ = 0.5 is of interest.

The SAS statistical software [20] code for estimating Equation (5) is as follows (rather than dividing by J, this computer code could employ the SAS _WEIGHT_ = 1/J option; the results are the same):

```
LINE  COMPUTER CODE
 1    PROC NLIN DATA = STEP1 METHOD = MARQUARDT MAXITER = 2000;
 2    PARMS RHOY = 0.0 B0 = 0;
 3        BOUNDS − 1 < RHOY < 1;
 4        ARRAY LAMBDAJ{400} TLAM1-TLAM400;
 5        JACOB = 0;
 6        DERJ = 0;
 7        DO I = 1 TO 400;
 8              JACOB = JACOB + LOG(1 − RHOY × LAMBDAJ{I});
 9              DERJ = DERJ + -LAMBDAJ{I}/(1 − RHOY × LAMBDAJ{I});
10        END;
11        J = EXP(JACOB/400);
12        DERJ = −DERJ/400;
13        ZY = Y/J;
14        MODEL ZY = (RHOY × WY + B0 × (1 − RHOY))/J;
15        DER.RHOY = ((RHOY × WY + B0 × (1 − RHOY) − Y) × DERJ + WY − B0)/J;
16    RUN;
```

Line #1 specifies the SAS nonlinear regression procedure NLIN, the name of the data file (STEP1), which contains both **Y** and **WY**, the nonlinear optimization technique to use (MARQUARDT), and the maximum number of iterations as one of the stopping criteria. Line #2 sets initial values for the regression equation parameter estimates. Line #3 attaches a constraint to the spatial autocorrelation parameter, ρ, restricting its estimate to the aforementioned feasible parameter space (defined by the extreme eigenvalues of matrix **W**). In this example, n = 400, the number of geographic locations. Line #4 establishes an array for the eigenvalues of matrix **W**; these eigenvalues are arranged in n columns, each with n entries of the same eigenvalue. Lines #5 and #6 initialize DO loop arguments. Lines #7 to #10 constitute the DO loop, with Line #8 calculating the standard calculus Jacobian term (of a transformation from a spatially autocorrelated to an unautocorrelated mathematical space), and Line #9 calculating the derivative of the Jacobian term. Each nonlinear optimization iteration executes this DO loop. Line #11 computes the Jacobian (see Equation (5)], and Line #12 calculates part of its derivative. Line #13 divides the left-hand side, and Line #14 divides the right-hand side, of the model equation by the Jacobian. Line

#15 is the correct derivative that SAS is unable to compute. In contrast, SAS is able to compute the correct derivative with respect to the mean.

Implementation is such that estimation here involves a successive set of linear regressions that conduct a derivative-guided systematic search across the feasible parameter space for the optimal spatial autocorrelation parameter value. Each linear regression employs a different value of the spatial autocorrelation parameter, with this systematic search guided by the SAS derivatives. The DO loop is executed for each one of these iterations. Ideally, the iterative procedure terminates when the minimum error sum of squares reaches its lowest value within the feasible parameter space (defined by Line #13).

To illustrate the problem here, consider a 20-by-20 regular square tessellation of pixels (n = 400) whose spatial weights matrix **W** is defined using the analogy of a rook's moves in chess: $w_{ij} > 0$ if pixels i and j share a common non-zero length boundary (i.e., a border but not solely a single point). Its extreme eigenvalues are ± 1. One hundred replicates were simulated with the following equation:

$$\mathbf{Y}_{\mathrm{j}} = 0 \times \mathbf{1} + (\mathbf{I} - 0.5\mathbf{W})^{-1}\, \boldsymbol{\varepsilon}\ (\mathrm{j} = 1, 2, \ldots, 100), \tag{6}$$

where $\boldsymbol{\varepsilon} \sim \mathrm{N}(\mathbf{0}, \mathbf{I}\sigma^2)$. The estimation results may be summarized as follows, including (the correct derivative) and excluding (the incorrect derivative) Line #15 in the SAS code:

derivative	$\bar{\hat{\rho}}$	range of $\hat{\rho}$	$\widehat{VAR(\hat{\rho})}$
correct	0.492	0.328–0.617	0.002
incorrect	0.823	0.629–0.968	0.005

These results demonstrate that SAS does not calculate the correct default derivative here, although it calculates the correct default derivate for the mean.

What went wrong? SAS does not understand spatial statistics. Using notation in the computer code, it correctly calculates $\frac{\partial Y}{\partial B0}\frac{1}{J}$, but it calculates only $\frac{\partial Y}{\partial RHOY}\frac{1}{J}$, which is incomplete because J also is a function of ρ. Furthermore, the other part of this latter derivative, $\frac{\partial J^{-1}}{\partial RHOY}Y$, needs to be subtracted from both sides of the derivative equation in order for SAS to have the correct optimization routine derivative. This subtraction results in the –Y term in the right-hand side of the SAS derivative statement (Line #15). Without an understanding of the product rule of calculus [21] (p. 110), a student or practitioner struggles with, and often fails to obtain, this correct derivative. If one of my students notices the –Y, s/he usually tells me that I have a mistake in my derivative. Meanwhile, I emphasize that SAS does not always calculate correct analytical derivatives: in this case, because it does not understand spatial statistics. One could argue that SAS was not designed to handle this calculus problem. SAS PROC NLIN assumes that a response variable Y is free of model parameters, and, accordingly, the procedure fits a given model specification by taking the derivatives of the right-hand side (RHS), and RHS-only, of the equation with respect to model parameters. However, this is exactly what is meant by the statement that SAS does not understand spatial statistics.

The serious substantive risk here is that because most socio-economic/demographic variables have $0.4 < \hat{\rho} < 0.6$, practitioners could dramatically overestimate the prevailing levels of spatial autocorrelation.

3.2. An Integral Calculus Example: The Jacobian of a Transformation

Most geographically distributed remotely sensed data (e.g., satellite images) yield a SAR spatial autocorrelation parameter estimate in the interval [0.90, 0.99]. Accordingly, the case of ρ = 0.95 also is of interest.

Consider the classical normal likelihood function

$$\mathrm{L} = (2\,\pi)^{-\mathrm{n}/2}(\sigma^2)^{-\mathrm{n}/2}\mathrm{e}^{-\boldsymbol{\varepsilon}^{\mathrm{T}}\boldsymbol{\varepsilon}/(2\sigma^2)}, \tag{7}$$

where $\boldsymbol{\varepsilon}$ is a vector of random variable values with mean zero. Now, $\mathbf{Y} = \mu\mathbf{1} + \boldsymbol{\varepsilon}$ implies

$$\boldsymbol{\varepsilon} = \mathbf{Y} - \mu\mathbf{1}. \tag{8}$$

Substituting Equation (8) into Equation (7) yields

$$L = (2\pi)^{-n/2}(\sigma^2)^{-n/2}e^{-(\mathbf{Y}-\mu\mathbf{1})^T(\mathbf{Y}-\mu\mathbf{1})/(2\sigma^2)}, \tag{9}$$

which is correct. Similarly, rewriting the general form of Equation (6) gives

$$\boldsymbol{\varepsilon} = \mathbf{Y} - \mu\mathbf{1} - \rho\mathbf{WY}. \tag{10}$$

Substituting Equation (10) into Equation (7) yields

$$L = (2\pi)^{-n/2}(\sigma^2)^{-n/2}e^{-(\mathbf{Y}-\mu\mathbf{1})^T(\mathbf{I}-\rho\,\mathbf{W})^T(\mathbf{I}-\rho\,\mathbf{W})(\mathbf{Y}-\mu\mathbf{1})/(2\sigma^2)}, \tag{11}$$

which is incorrect. Symbolic software tends to treat this substitution the same as that for Equation (9). However, preserving integration during a transformation can require the introduction of a Jacobian [21] (p. 204) [22] (p. 454) whose value is other than 1, which is not necessarily performed by symbolic software. This is the purpose of exercises in introductory mathematical statistics courses, for example, that involve calculating normalizing constants for sundry artificial probability density/mass functions.

This Jacobian term, already discussed in the preceding section, actually is like a nuisance parameter: researchers, practitioners, and students alike are not interested in it per se but must calculate it and then essentially ignore it. Nevertheless, the role of this Jacobian term is extremely important. It is a normalizing constant, which guarantees that a probability density/mass function integrates/sums to 1. It also functions as a constraint in spatial statistics, ensuring that the spatial autocorrelation parameter estimate is restricted to its feasible parameter space (Line #15).

To illustrate the problem here, consider the same geographic landscape as before. Now one hundred replicates are simulated with the equation

$$\mathbf{Y}_j = 0 \times \mathbf{1} + (\mathbf{I} - 0.95\mathbf{W})^{-1}\boldsymbol{\varepsilon}\ (j = 1, 2, \ldots, 100). \tag{12}$$

The estimation results may be summarized as follows:

derivative	$\bar{\hat{\rho}}$	range of $\hat{\rho}$	$\widehat{VAR(\hat{\rho})}$
correct	0.948	0.916–0.977	0.0002
incorrect	1.058	1.012–1.104	0.0004

These results demonstrate that ignoring the Jacobian term can yield not only incorrect, but also nonsensical, parameter estimates. The spatial autocorrelation parameter value cannot exceed 1, yet without a Jacobian term, it always is greater than 1 in this example simulation experiment. This outcome is like obtaining a bivariate statistics Pearson product moment correlation coefficient greater than 1. Furthermore, this incorrect result relates to a divergent series: rather than $(\mathbf{I} - \rho\mathbf{W})^{-1}$ producing a spatial autocorrelation that dampens with increasing distance separating locations (convergence), spatial autocorrelation actually amplifies with increasing distance.

What went wrong? Calculus theory was overlooked, and the analysis guided by a standard type of substitution went astray. The serious risk here is that practitioners could obtain nonsense spatial autocorrelation results, at least in more extreme cases.

4. Discussion

A principal implication here is that although symbolic software removes tedium from many calculus exercises, students in the applied mathematical sciences still need to know calculus theory. I encourage my students to double check their calculus solutions with

Mathematica and/or Maple. However, I also encourage them to know and understand the calculus theory underlying the solutions they obtain. The two more common than rare examples presented in this paper illustrate situations in which symbolic software can fail when in its unsupervised auto mode. Many real-world problems are sufficiently complicated such that they may not align with the simpler mechanics of AI software. In addition to this cautionary remark, the two examples detailed in this paper offer an opportunity to further diversify relevant applications, examples, and exercises appearing in contemporary calculus textbooks, especially those targeting applied mathematics audiences.

Another important debate question meriting a dialogue here concerns the extent to which evolving symbolic algebra software development, of which the many existing systems [23] almost continuously undergo, can avoid such categories of automated errors. This is a more general issue epitomized by people's frustration with the near-universal incorporation of spelling autocorrection in word-processing types of programs, which habitually changes false positive typographical errors. Nevertheless, the statistics literature, for one, extends guidance from its computerization experience that began more than half a century ago.

Arguably, Longley [24] initialized the statistics effort by submitting a simple regression problem to several widely used computer routines; his goal was to assess the efficiency and accuracy of ordinary least squares (OLS) algorithms. One outcome of this type of endeavor was a differentiation between conceptual and computational coding representations. Shortly thereafter, Allerbeck [25] proposed a set of statistical software package benchmarks, including "solutions must not be tailored for some special problem" (p. 25); this is the weakness—software design for only the relatively simpler calculus calculations—underscored in this paper. Next, because OLS involves matrix inversion, investigative attention shifted to routines computing this particular matrix operation [26], as well as the estimation of a much wider range of commonly reported sample statistics [27]. A proposal stemming from this collection of work promotes the establishment of a library of readily available test datasets with algorithm-challenging nuances and idiosyncrasies. IMSL [28] (pp. 1309–1310), for example, followed this line of reasoning with quality evaluation of its pseudorandom number generators—a problem of interest even today [29]—reporting replicable goodness-of-fit test results for them. A similar publicly available open access library of simple looking but tricky calculus problems should be established accompanied by exercises for applied calculus pupils that familiarize these students with common inconspicuous solution eccentricities that confuse symbolic software. Interestingly, aligning with the application theme of this paper, Simon and LeSage [30], the latter scholar a prominent spatial econometrician, discuss benchmark datasets for indexing the numerical accuracy of statistical algorithms, essentially endorsing the creation of the aforementioned libraries. Therefore, a freely accessible test suite of problems might help certify calculus AI packages, as well as target categories of problems for which students/practitioners potentially would need to intervene and override automated algebraic manipulations in order to steer the software to a correct solution. Unfortunately, such certification most likely would need to be ongoing to maintain pace with new developments in computer algebra software.

Part of this preceding outlined promising approach would entail experienced users constructing and contributing to a common library the necessary illuminating specialized example problems in various applications areas, such as the one focused on in this paper for spatial statistics/econometrics. The current R Project [https://www.r-project.org/ and https://r-dir.com/reference/datasets.html (accessed on 19 May 2022)] furnishes an excellent model for this initiative. This theme definitely is a topic for future applied mathematics research.

5. Conclusions

One applied statistics discipline undergoing an audience expansion is spatial statistics. Its extra mathematical complexities evaded pedagogic efforts for many decades after its initial conceptualization in the very early 1900s, and its popularizing by Cliff and Ord [31] in the latter third of the 1900s. Today it offers new examples for inclusion in a myriad of applied mathematics textbooks, including those for teaching calculus. Its tricky equations that include a response variable on both sides of an equal sign are one source of more unusual complications. Interestingly, this is not the only mathematics arena benefiting from rather recent spatial statistics advances. Another is Markov chain Monte Carlo (MCMC) techniques, which frequently have been necessary for the practical implementation of spatial statistical inferential tools, especially within the context of auto-models and Bayesian map analysis. By motivating new spatial stochastic model specifications, spatial statistics in turn motivated the devising of improved MCMC algorithms. Besag and Green [32], for example, furnished an historical overview of this connection. In conclusion, similar papers can, and probably should, be written about other spatial statistics and applied mathematics topics, such as MCMC, for future research undertakings. After all, any mathematics curriculum, not just those for calculus, should teach associated theory.

Funding: This research received no external funding.

Informed Consent Statement: Not applicable.

Data Availability Statement: For the pseudorandom number generators used to create test data see https://documentation.sas.com/doc/en/pgmsascdc/9.4_3.2/lefunctionsref/p0fpeei0opypg8n1b06qe4r040lv.htm (last accessed on 19 May 2022).

Conflicts of Interest: The University of Texas at Dallas subscribes to an annual license for the latest version of Mathematica. The University of Texas System subscribes to an annual license for SAS. Otherwise, the author declares he has no personal circumstances or interest that may be perceived as inappropriately influencing the representation or interpretation of reported research results. Because the research summarized here had no funding support, no funders had a role in: the design of the study; collection, analyses, or interpretation of data; writing of the manuscript; or decision to publish the results.

References

1. Mathematical Association of America (MAA) Panel on Statistics. *Introductory Statistics without Calculus*; Committee on the Undergraduate Program in Mathematics: Berkeley, CA, USA, 1972.
2. Shockley, J. *The Brief Calculus: With Applications in the Social Sciences*; Holt, Rinehart and Winston: New York, NY, USA, 1971.
3. Saltz, D. *A Short Calculus: An Applied Approach*, 2nd ed.; Goodyear Publishing: Santa Monica, CA, USA, 1977.
4. Newmark, J. *The Usefulness of Calculus for the Behavioral, Life, and Managerial Sciences*; Canfield Press: San Francisco, CA, USA, 1978.
5. Bittinger, M.; Brand, N.; Quintanilla, J. *Calculus for the Life Sciences*, 1st ed.; Pearson/Addisson WesIey: Boston, MA, USA, 2006.
6. Frey, J. Technology and mathematics education: A survey of recent developments and important problems. *Educ. Stud. Math.* **1989**, *20*, 237–272.
7. Kabaca, T.; Aksoy, Y.; Aktumen, M.; Mirasyediglu, S. The use of computer algebra systems in calculus teaching: Principles and sample applications. In *Technology Education and Development*; Lazinica, A., Calafate, C., Eds.; InTech: Rijeka, Croatia, 2009; Chapter 30; pp. 479–500. Available online: http://www.intechopen.com/books/technology-education-and-development/the-use-of-computer-algebrasystems-in-calculus-teaching-principles-and-sample-applications (accessed on 19 May 2022).
8. Pierce, R.; Ball, L. Perceptions that may affect teachers' intention to use technology in secondary mathematics classes. *Educ. Stud. Math.* **2009**, *71*, 299–317. [CrossRef]
9. Bressoud, D.; Mesa, V.; Rasmussen, C. *Insights and Recommendations from the MAA National Study of College Calculus*; The Mathematical Association of America: Washington, DC, USA, 2015.
10. Karjanto, N.; Husain, H.S. Not another computer algebra system: Highlighting *wxMaxima* in calculus. *Mathematics* **2021**, *9*, 1317. [CrossRef]
11. Slavíčková, M. Implementation of digital technologies into pre-service mathematics teacher preparation. *Mathematics* **2021**, *9*, 1319. [CrossRef]
12. Ziatdinov, R.; Valles, J. Synthesis of modeling, Visualization, and programming in GeoGebra as an effective approach for teaching and learning STEM topics. *Mathematics* **2022**, *10*, 398. [CrossRef]
13. Amado, N.; Domingos, A.; Jones, K. Editorial: Re-visioning teaching and learning with technology in mathematics: Selected papers from ICTMT12. *Teach. Math. Its Appl.* **2016**, *35*, 51–52. [CrossRef]
14. Wolfram. *Mathematica Quick Revision History*, Mathematica latest version; Wolfram Worldwide Headquarters: Champaign, IL, USA, 2022; Available online: http://www.wolfram.com/mathematica/quick-revision-history.html (accessed on 19 May 2022).
15. MapleSoft. *Maple User Manual: Maple 18*; Waterloo Maple Inc.: Waterloo, ON, Canada, 2021; Available online: https://www.maplesoft.com/documentation_center/maple2021/UserManual.pdf (accessed on 19 May 2022).

16. Cipra, B. How number theory got the best of the Pentium chip. *Science* **1995**, *267*, 175. Available online: http://www.sciencemag.org/content/267/5195/175 (accessed on 20 May 2022). [CrossRef] [PubMed]
17. Gelfand, A.; Diggle, P.; Fuentes, M.; Guttorp, P. (Eds.) *Handbook of Spatial Statistics*; CRC Press: Boca Raton, FL, USA, 2010.
18. Besag, J. Spatial interaction and the statistical analysis of lattice systems. *J. R. Stat. Soc. Ser. B Stat. Method* **1974**, *36*, 192–225. [CrossRef]
19. Ward, M.; Ahlquist, J. *Maximum Likelihood for the Social Sciences: Strategies for Analysis*; Cambridge University Press: New York, NY, USA, 2018.
20. SAS Institute Inc. The NLIN procedure. In *SAS/STAT(R) 15.1 User's Guide*; SAS Institute Inc.: Cary, NC, USA, 2022; Available online: https://documentation.sas.com/doc/en/pgmsascdc/v_027/statug/titlepage.htm?fromDefault= (accessed on 19 May 2022).
21. Johnson, R.; Kiokemeister, F. *Calculus with Analytic Geometry*, 3rd ed.; Allyn and Bacon: Boston, MA, USA, 1965.
22. Lindgren, B. *Statistical Theory*, 3rd ed.; Macmillan: New York, NY, USA, 1968.
23. Wikipedia. List of Computer Algebra Systems. Available online: http://en.wikipedia.org/wiki/List_of_computer_algebra_systems (accessed on 31 May 2022).
24. Longley, J. An appraisal of least squares programs for the electronic computer from the point of view of the user. *J. Am. Stat. Assoc.* **1967**, *62*, 819–841. [CrossRef]
25. Allerbeck, K. Data analysis systems: A user's point of view. *Comput. Data Processing/Ordinat. Traitement Données* **1971**, *10*, 23–35. [CrossRef]
26. Beaton, A.; Rubin, D.; Barone, J. The acceptability of regression solutions: Another look at computational accuracy. *J. Am. Stat. Assoc.* **1976**, *71*, 158–168. [CrossRef]
27. Wilkinson, L.; Dallal, G. Accuracy of sample moments calculations among widely used statistical programs. *Am. Stat.* **1977**, *31*, 128–131.
28. IMSL. *IMSL Fortran Library User's Guide MATH/LIBRARY*; Visual Numerics, Inc.: San Ramon, CA, USA, 2003; Volume 2.
29. James, F.; Moneta, L. Review of high-quality random number generators. *Comput. Softw. Big Sci.* **2020**, *4*, 2. [CrossRef]
30. Simon, D.; LeSage, J. Benchmarking numerical accuracy of statistical algorithms. *Comput. Stat. Data Anal.* **1988**, *7*, 197–209. [CrossRef]
31. Cliff, A.; Ord, J. *Spatial Autocorrelation*; Pion: London, UK, 1973.
32. Besag, J.; Green, P. Spatial statistics and Bayesian computation. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **1993**, *55*, 25–37; discussion 53–102. [CrossRef]