11 : 1 2020

ASTAT

Algebraic Statistics

EDITORIAL: A NEW BEGINNING

THOMAS KAHLE AND SONJA PETROVIĆ

msp

Algebraic Statistics
Vol. 11, No. 1, 2020
https://doi.org/10.2140/astat.2020.11.1

EDITORIAL: A NEW BEGINNING

THOMAS KAHLE AND SONJA PETROVIĆ

Algebraic statistics and *Algebraic Statistics*

The creation of a field that bridges two disparate areas takes both ingenuity and the ability to generate excitement about new interdisciplinary ideas. For that field to continuously evolve over two decades, expanding to include virtually every aspect of the ground fields, as well as a growing number of neighboring research areas, takes a continued and dedicated community effort. *Algebraic Statistics* (AStat) is being established as a journal to be run by and devoted to such a community, representing interdisciplinary researchers in the field coming from all backgrounds.

While algebra has always played a prominent role in statistics, the publication of a couple of seminal works in the late 1990s defined the new direction by connecting modern computational algebraic geometry and commutative algebra to two critical problems in statistics: sampling from discrete conditional distributions [3] and experimental design [7]. With the onset of the 2000s, the use of these techniques in statistics really took off, generating a large body of research papers and several textbooks [1; 2; 4; 5; 6; 8; 9]. In the last decade, the field has seen a massive influx of new people bringing new ideas and perspectives to problems at the intersection of nonlinear algebra, interpreted in broadest possible sense, and statistics.

The term "algebraic statistics" has thus evolved in meaning to include an ever-expanding list of topics. We understand it as an umbrella term for using algebra (multilinear algebra, commutative algebra, and computational algebra), geometry and combinatorics to obtain insights in mathematical statistics as well as for diverse applications of these tools to data science.

The community of algebraic statisticians is quite an active one, organizing many conferences, symposia, seminars, and special sessions at regional and international meetings, and striving for involvement and representation within both nonlinear algebra and statistics. The predecessor community-run journal, which existed for a decade and published ten volumes, has now been discontinued due to a dispute in ownership with a third party interested in a profit oriented future for the journal. The core of the algebraic statistics community strongly supports the establishment of this new journal. It is a leap forward, a fresh start that takes into account historical lessons learned and seeks to grow and expand the research scope. For this endeavor, we are happy to team up with MSP as a publishing partner that is committed to support academic scholarship and to ensuring the long-term success of our research community.

MSC2020: 62R01.
Keywords: algebraic statistics.

The first volume

The first volume, in two issues, contains eleven papers with a mix that represents algebraic statistics well. Mathematical themes include Gröbner bases, both the standard and non-commutative versions, toric and tropical varieties, numerical nonlinear algebra, holonomic gradient descent, and algebraic combinatorics. On the side of statistics, there are models for diverse types of data, parameter estimation under the likelihood principle, covariance estimation, and time series. Applications covered include computational neuroscience, clustering analysis, engineering, material science, and geology.

(1) The paper "Maximum likelihood estimation of toric Fano varieties" showcases likelihood geometry. Its main result explains how properties of likelihood estimation depend on algebraic and geometric features of the underlying toric models.

(2) Linear covariance models are models for Gaussian random variables with linear constraints on the covariance matrix. The paper "Estimating linear covariance models with numerical nonlinear algebra" addresses the problem of maximum likelihood estimation in these models, the related complexity challenges, and introduces an accompanying `Julia` package.

(3) "Expected value of the one-dimensional earth mover's distance" gives explicit formulas for the expected value of a distance between a pairs of one-dimensional discrete probability distributions using algebraic combinatorics, and discusses applications of it in clustering analysis.

(4) In "Inferring properties of probability kernels from the pairs of variables they involve" the authors discuss how inference about inherently continuous and uncountable probability kernels can be encoded in discrete structures such as lattices.

(5) In computational neuroscience, neural codes model patterns of neuronal response to stimuli. The field provides many open problems for mathematics and statistics. "Minimal embedding dimensions of connected neural codes" address a problem from receptive field coding: the embedding of neural codes in low dimension.

(6) The holonomic gradient method in *Holonomic gradient method for two way contingency tables* is a numerical procedure to approximate otherwise inaccessible likelihood integrals. It is here applied in a discrete situation of contingency tables.

(7) *Algebraic analysis of rotation data* studies a well-known model for rotation data using the tools from non-commutative algebra and the holonomic gradient descent method. It also discusses applications to several areas of science and engineering.

(8) *Maximum likelihood degree of the two-dimensional linear Gaussian covariance model* provides explicit formulas for the number of solutions of likelihood equations in special cases of the same problem as in paper (2).

(9) *Tropical gaussians: a brief survey* takes a tour through the analogues of Gaussian distributions over the tropical semiring. This has applications in, for example, economics and phylogenetics.

(10) *The norm and saturation of a binomial ideal, and applications to Markov bases* connects back to the beginnings of algebraic statistics: Markov bases. Here the focus is on the complexity of Markov bases.

(11) Finally, *Compatibility of distributions in probabilistic models: An algebraic frame and some characterizations* studies the problem when and how two distributions for two sets of variables can be put together to a distribution for the union of the variables and exhibits discrete and algebraic structures in this problem.

Call for submissions

We see AStat as a primary forum serving the broad community in a focused way. As an interdisciplinary endeavor, by definition, a concerted effort will be made for AStat to serve various constituents interested in and interacting with algebraic statistics. Specifically, in our definition, AStat is devoted to algebraic aspects of statistical theory, methodology and applications, seeking to publish a wide range of research and review papers that address one of the following:

- algebraic, geometric and combinatorial insights into statistical models or the behavior of statistical procedures;
- development of new statistical models and methods with interesting algebraic or geometric properties;
- novel applications of algebraic and geometric methods in statistics.

We invite the community to send their best work in algebraic statistics to be considered for publication here. This includes contributions which connect statistical theory, methodology, or application to the world of algebra, geometry, and combinatorics in ways that may not be labeled as traditional.

References

[1] S. Aoki, H. Hara, and A. Takemura, *Markov bases in algebraic statistics*, Springer, 2012.

[2] C. Bocci and L. Chiantini, *An introduction to algebraic statistics with tensors*, Unitext **118**, Springer, 2019.

[3] P. Diaconis and B. Sturmfels, "Algebraic algorithms for sampling from conditional distributions", *Ann. Statist.* **26**:1 (1998), 363–397.

[4] M. Drton, B. Sturmfels, and S. Sullivant, *Lectures on algebraic statistics*, Oberwolfach Seminars **39**, Birkhäuser, 2009.

[5] P. Gibilisco, E. Riccomagno, M. P. Rogantin, and H. P. Wynn, *Algebraic and geometric methods in statistics*, Cambridge Univ. Press, 2010.

[6] L. Pachter and B. Sturmfels, *Algebraic statistics for computational biology*, Cambridge Univ. Press, 2005.

[7] G. Pistone, E. Riccomagno, and H. P. Wynn, *Algebraic statistics: computational commutative algebra in statistics*, Monographs on Statistics and Applied Probability **89**, CRC, Boca Raton, FL, 2001.

[8] S. Sullivant, *Algebraic statistics*, Graduate Studies in Mathematics **194**, American Mathematical Society, 2018.

[9] S. Watanabe, *Algebraic geometry and statistical learning theory*, Cambridge Monographs on Applied and Computational Mathematics **25**, Cambridge University Press, 2009.

THOMAS KAHLE (thomas.kahle@ovgu.de), Otto-von-Guericke Universität Magdeburg
SONJA PETROVIĆ (sonja.petrovic@iit.edu), Illinois Institute of Technology
Managing Editors

msp.org/astat

MANAGING EDITORS

Thomas Kahle — Otto-von-Guericke-Universität Magdeburg, Germany
Sonja Petrovic — Illinois Institute of Technology, United States

ADVISORY BOARD

Mathias Drton — Technical University of Munich, Germany
Peter McCullagh — University of Chicago, United States
Giorgio Ottaviani — University of Florence, Italy
Bernd Sturmfels — University of California, Berkeley, and Max Planck Institute, Leipzig
Akimichi Takemura — University of Tokyo, Japan

EDITORIAL BOARD

Marta Casanellas — Universitat Politècnica de Catalunya, Spain
Alexander Engström — Aalto University, Finland
Hisayuki Hara — Doshisha University, Japan
Jason Morton — Pennsylvania State University, United States
Uwe Nagel — University of Kentucky, United States
Fabio Rapallo — Università del Piemonte Orientale, Italy
Eva Riccomagno — Università degli Studi di Genova, Italy
Yuguo Chen — University of Illinois, Urbana-Champaign, United States
Caroline Uhler — Massachusetts Institute of Technology, United States
Ruriko Yoshida — Naval Postgraduate School, United States
Josephine Yu — Georgia Institute of Technology, United States
Piotr Zwiernik — Universitat Pompeu Fabra, Barcelona, Spain

PRODUCTION

Silvio Levy — (Scientific Editor)
production@msp.org

See inside back cover or msp.org/astat for submission instructions.

Algebraic Statistics (ISSN 2693-3004 electronic, 2693-2997 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

AStat peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY
mathematical sciences publishers
nonprofit scientific publishing
http://msp.org/
© 2020 Mathematical Sciences Publishers

Algebraic Statistics

2020 11 : 1

Editorial: A new beginning 1
THOMAS KAHLE and SONJA PETROVIĆ

Maximum likelihood estimation of toric Fano varieties 5
CARLOS AMÉNDOLA, DIMITRA KOSTA and KAIE KUBJAS

Estimating linear covariance models with numerical nonlinear algebra 31
BERND STURMFELS, SASCHA TIMME and PIOTR ZWIERNIK

Expected value of the one-dimensional earth mover's distance 53
REBECCA BOURN and JEB F. WILLENBRING

Inferring properties of probability kernels from the pairs of variables they involve 79
LUIGI BURIGANA and MICHELE VICOVARO

Minimal embedding dimensions of connected neural codes 99
RAFFAELLA MULAS and NGOC M. TRAN