Oberwolfach Seminar

METRIC ALGEBRAIC GEOMETRY

(Meeting-ID: 2322b)

Organizers:

Paul Breiding	Osnabrück	pbreiding@uni-osnabrueck.de
Kathlén Kohn	Stockholm	kathlen@kth.se
Bernd Sturmfels	Leipzig	bernd@mis.mpg.de

Where and when:

Mathematisches Forschungsinstitut Oberwolfach, 28 May – 03 Jun 2023.

In the early 19th century, there was no difference between algebraic and differential geometry. The two were part of the same subject. Geometers studied natural properties of curves and surfaces in 3-space, such as curvature, singularities, and defining equations. In the 20th century, the threads diverged. Students of mathematics now first encounter algebraic geometry and differential geometry in rather disconnected courses.

In the present days, the *geometry of data* requires us to rethink that schism. Many applied problems center around metric questions, such as optimization with respect to distances. These require tools from different areas in geometry, algebra and analysis.

Metric Algebraic Geometry offers a path towards integration. This term is a neologism which joins the names *metric geometry* and *algebraic geometry*. It first appeared in the title of Madeleine Weinstein's dissertation [9]. Building on classical foundations, the field embarks towards a new paradigm that combines concepts from algebraic geometry and differential geometry, with the goal of developing practical tools for the 21st century.

Many problems in the sciences lead to solving polynomial equations [8] over the real numbers. The solution sets are real algebraic varieties. Understanding distances, volumes and angles – in short, understanding metric properties – of varieties is important for modeling and analyzing data. Take, as an example, the variety $V \subset \mathbb{R}^4$ defined by the two equations

$$f(p) := \det \begin{bmatrix} p_1 & p_2 \\ p_3 & p_4 \end{bmatrix} = 0 \quad \text{and} \quad g(p) := 1 - (p_1 + p_2 + p_3 + p_4) = 0.$$

This variety forms a statistical model: Points on V that satisfy $0 \leq p_1, p_2, p_3, p_4 \leq 1$ correspond to joint probability distributions of two independent random variables. Suppose that we are given statistical data in the form of an empirical distribution $u \in \mathbb{R}^4$. One goal in statistics is to find a point on V that best explains u, where best means a point $p \in V$ that minimizes a distance function $\ell(p, u)$. For instance, minimizing the log-likelihood function $\ell(p, u) = u_1 \cdot \log(p_1) + u_2 \cdot \log(p_2) + u_3 \cdot \log(p_3) + u_4 \cdot \log(p_4)$ gives rise to the following problem of solving the system of six nonlinear equations in six variables $(p_1, p_2, p_3, p_4, \lambda, \mu)$:

$$f(p) = g(p) = 0 \quad \text{and} \quad \frac{\partial \ell}{\partial p_i} - \lambda \cdot \frac{\partial f}{\partial p_i} - \mu \cdot \frac{\partial g}{\partial p_i} = 0,\ i = 1, \ldots, 4,$$

via Lagrange multipliers. This problem is equivalent to solving a system of polynomial equations. Other metric problems in optimization and statistics involve, for instance, minimizing the Euclidean distance from a variety to a given data point. Furthermore, in topological data analysis, computing the homology of a submanifold depends on curvature and on bottlenecks. Related metric questions arise in the theoretical study of machine learning with neural networks, in the geometry of computer vision, in learning varieties from data, and in the study of Voronoi cells.

Consequently, metric algebraic geometry is motivated by the desire to bring the perspective and tools of algebraic geometry to the objects of metric geometry arising in the context of data analysis. It concerns properties and invariants of real algebraic varieties coming from an extrinsic metric in the ambient space, or possibly from an intrinsic metric inside the variety. Studying metric objects in algebraic geometry is not new, and goes back at least to the work of Salmon in the 19th century [7]. What is new and different to what was done before is the applied context. We see **Metric Algebraic Geometry** as a driver for innovation in the analysis of data and the identification of underlying geometric structures.

This Oberwolfach seminar features a diverse range of topics. Every day focuses on an overall theme that will be explored in three 90-minutes lectures:

Topic	Kohn	Sturmfels	Breiding
Getting Started	Historical context	Critical Equations	Computations
Counting Solutions	Polar Classes	Wasserstein Distance	Euclidean Distance
Metric Geometry	Bottlenecks and Reach	Voronoi Cells	Curvature
Data and Statistics	Machine Learning	Maximum Likelihood	Tensors
Research Trends	Computer Vision	Tropical Limits	Sampling

The first day (Monday) and the last day (Friday) are special. They form the frame of the planned seminar. The theme for Monday is called **Getting Started**. The idea is to introduce the participants to the basic ideas, historical context and computational techniques. On Friday we shall discuss current and upcoming **Research Trends** in Metric Algebraic Geometry and its modern applications. For the other three days we have chosen the overall themes **Counting Solutions** (Tuesday), **Metric Geometry** (Wednesday) and **Data and Statistics** (Thursday).

In addition to the lectures we plan exercise sessions. At least one free hour per day will encourage collaboration, informal discussions, and essential scientific networking. Next to these traditional activities, the workshop will have a “conversation-starter” session on every day. In these 5-minute presentations, participants will offer a scientific teaser about one of their currents projects, or other bit of mathematics that they want to discuss with others during the workshop. Informal and open-ended discussions will take place in afternoons and evenings. The specific material for these casual gatherings will be chosen to respond to the needs and interests arising during the workshop.

We will make a first draft of our lecture notes available to the participants 2-4 weeks before the event. In addition, the participants can have a look at the following references.

References

[1] Paul Breiding: An Algebraic Geometry Perspective on Topological Data Analysis, *SIAM News* **53** (2020).

[2] Paul Breiding, Türkü Özlüm Çelik, Timothy Duff, Alexander Heaton, Aida Maraj, Anna-Laura Sattelberger, Lorenzo Venturello, and Oğuzhan Yürük: Nonlinear Algebra and Applications, *Numerical Algebra, Optimization and Control* (2021).

[3] Jan Draisma, Emil Horobeţ, Giorgio Ottaviani, Bernd Sturmfels and Rekha Thomas: The Euclidean distance degree of an algebraic variety, *Found. Comput. Math.* **16** (2016) 99–149.

[4] Joe Kileel and Kathlén Kohn: Snapshot of Algebraic Vision, `arXiv:2210.11443`.

[5] Kathlén Kohn, Thomas Merkh, Guido Montúfar, Matthew Trager: *Geometry of Linear Convolutional Networks*, SIAM SIAGA **6** (2022) 368–406.

[6] Mateusz Michałek, Bernd Sturmfels: *Invitation to Nonlinear Algebra*. AMS Graduate Studies in Mathematics, 2021.

[7] George Salmon: *Treatise on the Higher Plane Curves*, Hodges-Foster-Figgis, Dublin, 1879.

[8] Bernd Sturmfels: *Solving Systems of Polynomial Equations* CBMS Regional Conference Series in Mathematics, 2002.

[9] Madeleine Weinstein: *Metric Algebraic Geometry*, PhD Dissertation, UC Berkeley, 2021.