Metric Algebraic Geometry Tutorial: Exercises
SIAM AG 25

Paul Breiding & Kathlén Kohn

July 7, 2025

The exercises are divided into two parts. The first part serves as a warm up and applies the material from the lecture to curves in the plane. The second part considers varieties from neural networks.

1 Curves in the plane

The following exercises should be applied to the parabola

$$X_1 = V(y - x^2)$$

and the Trott curve

$$X_2 = V\big(144(x^4 + y^4) - 225(x^2 + y^2) + 350x^2y^2 + 81\big).$$

1.1 Curvature

Compute the following data for both X_1 and X_2.

1. The curvature at $(1, 1), (0, 0) \in X_1$ and $(1, 0) \in X_2$.
2. Inflection points.
3. Critical curvature points.
4. Evolute.

1.2 Offset

Compute the following data for both X_1 and X_2.

1. Offset curve.
2. Bottlenecks.
3. Reach
4. Offset polynomial.
5. Offset discriminant and its decomposition.
6. The second fundamental form from the offset polynomial.

1.3 Voronoi cells

Suppose that $I = \langle f_1, \dots, f_k\rangle \subset \mathbb{Q}[x_1, \dots, x_n]$ is the ideal of a real variety $X = V(I) \subset \mathbb{R}^n$ of codimension c, assumed real radical and prime. Consider variables $u = (u_1, \dots, u_n)$ and

$$N_I(x) := \langle (c + 1) \times (c + 1) \text{ minors of } A\rangle,$$

where $A = \begin{pmatrix} x-u & \nabla f_1(x) \cdots & \nabla f_k(x) \end{pmatrix}$ is the $n \times (k + 1)$ *augmented Jacobian.*

The Voronoi boundary ideal at $y \in X$ is

$$\mathrm{Vor}_I(y) = (C_I(y) : \langle x - y\rangle^\infty) \cap \mathbb{Q}(u),$$

where

$$C_I(y) = N_I(x) + N_I(y) + \langle \|x - u\|^2 - \|y - u\|^2\rangle.$$

1. What is the geometric description of the zero set of $\mathrm{Vor}_I(y)$?
2. Compute the Voronoi boundary ideal at $(1, 1), (0, 0) \in X_1$ and $(1, 0) \in X_2$.

2 Varieties from neural networks

In the following we give four examples of varieties (or, more generally, semi-algebraic sets) describing the neuromanifolds of some neural networks. It would be interesting to study them from the point of view of metric algebraic geometry. Is it possible to compute critical curvature, ED discriminant, bottlenecks, reach, the offset polynomial, or Voronoi cells/boundaries (at smooth or singular points) for these varieties? These are open question, for which we don't know the answer. The goal of this tutorial session is to find out.

2.1 Threefold in 4-space

Consider the following multilayer perceptron (MLP) $\mathbb{R}^2 \to \mathbb{R}$:

$$\begin{bmatrix} e & f \end{bmatrix} \sigma\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}\right), \tag{1}$$

where $\sigma(X) = X^4$ gets applied entrywise. This parametrizes quartic homogeneous polynomials in (x, y):

$$Ax^4 + Bx^3y + Cx^2y^2 + Dxy^3 + Ey^4.$$

The Zariski closure of the set of all polynomials that can be obtained from the MLP is the hypersurface

$$2C^3 - 9BCD + 27AD^2 + 27B^2E - 72ACE = 0,$$

so it is a threefold in a $\mathbb{P}^4$. It is singular along the quartic curve that consists of the polynomials with a quadruple root. For $C = 1$ and $A + B = D + E$, the resulting surface is shown in Figure 1.

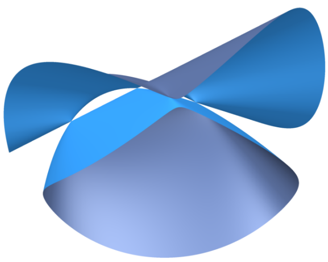

Figure 1: Slice of the Zariski closure of the neuromanifold of the shallow polynomial MLP in (1).
$27xy^2 + 27y^3 - 27y^2z + 27xz^2 - 72x^2 - 72xy + 72xz - 9yz + 2 = 0$

2.2 Smooth quadric in 3-space

If we want smooth surfaces, we can change the MLP to be invariant under permutations:

$$\begin{bmatrix} c & d \end{bmatrix} \sigma\left(\begin{bmatrix} a & b \\ b & a \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}\right), \quad \text{where } \sigma(X) = X^3. \tag{2}$$

This parametrizes cubic polynomials $Ax^3 + Bx^2y + Cxy^2 + Dy^3$. The Zariski closure of the set of all obtainable polynomials is the quadric surface

$$B^2 - 3AC - C^2 + 3BD$$

in $\mathbb{P}^3$. In the affine chart $A + B + C + D = 1$, it is shown in Figure 2.

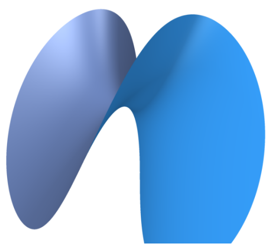

Figure 2: Zariski closure of the neuromanifold of the invariant MLP of degree 3 in (2).

2.3 Smooth surface in 4-space

If you think quadric surfaces are too boring, you can use the invariant MLP in (2), but with $\sigma(X) = X^4$. That way, we get quartic polynomials in (x, y) as in Section 2.1. The Zariski closure of the set of polynomials that the MLP parametrizes is now of codimension two in $\mathbb{P}^4$, defined by the following equations:

$$\begin{aligned} 4C^2 - 9BD - 9D^2 + 24CE &= 0, \\ BC - 6AD - CD + 6BE &= 0, \\ 3B^2 - 8AC - 3D^2 + 8CE &= 0. \end{aligned}$$

This has degree three and is smooth. Hence, by projecting from a point, we get a (singular) surface in $\mathbb{P}^3$ that we can look at. For instance, after eliminating C, we obtain:

$$B^3 - 16A^2D - B^2D - BD^2 + D^3 + 16ABE + 16ADE - 16BE^2 = 0.$$

This is depicted in Figure 3 in the affine chart $A + B + D + E = 1$.

Figure 3: Projection of the Zariski closure of the neuromanifold of the invariant MLP of degree 4.

2.4 Attention is all you need

The *self-attention mechanism* is the key ingredient of popular transformer architectures (used e.g. in ChatGPT). A single-layer network of unnormalized self-attention parametrizes cubic functions of the form

$$\begin{aligned} \mathbb{R}^{d\times t} &\to \mathbb{R}^{d'\times t}, \\ X &\mapsto VXX^\top K^\top QX, \end{aligned} \tag{3}$$

where $V \in \mathbb{R}^{d'\times d}$ and $K, Q \in \mathbb{R}^{a\times d}$ are the weight matrices of the network. So the neuromanifold is a semi-algebraic subset of $\mathrm{Sym}_3(\mathbb{R}^{d\times t})^{d'\times t}$.

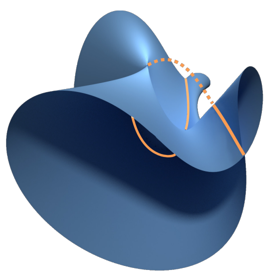

Figure 4: Slice of the neuromanifold of unnormalized self-attention mechanisms.

Let us consider the example $d' = 1$ and $d = t = a = 2$. Then, the neuromanifold is a subset of the 40-dimensional vector space $\mathrm{Sym}_3(\mathbb{R}^{2\times 2})^2$. The neuromanifold itself has been shown to be 5-dimensional in [1]. We can see it as a hypersurface in a 6-dimensional ambient space as follows: Writing x_1 and x_2 for the two columns of $X \in \mathbb{R}^{2\times 2}$ and setting $A := K^\top Q$, equation (3) becomes

$$(x_1, x_2) \mapsto \begin{pmatrix} Vx_1 \cdot x_1^\top Ax_1 + Vx_2 \cdot x_2^\top Ax_1 \\ Vx_1 \cdot x_1^\top Ax_2 + Vx_2 \cdot x_2^\top Ax_2 \end{pmatrix}. \tag{4}$$

The term that is quadratic in x_2 and linear in x_1 is $Vx_2 \cdot x_2^\top Ax_1$. Note that this term uniquely determines the function (4). Writing $x_1 = (x_{1,1}, x_{1,2})^\top$, denoting by α_i the linear form that takes the inner product with the i-th column of A, and setting similarly $\nu(x) := Vx$, we can express that term as

$$Vx_2 \cdot x_2^\top Ax_1 = x_{1,1} \cdot \alpha_1(x_2) \cdot \nu(x_2) + x_{1,2} \cdot \alpha_2(x_2) \cdot \nu(x_2).$$

This function is uniquely determined by the two products of linear forms $(\alpha_1\nu, \alpha_2\nu) \in \mathrm{Sym}_2(\mathbb{R}^2)^2$. Note that the dimension of the latter space is 6, as promised. (It is in fact generally true that neuromanifolds of single-layer attention networks are linearly isomorphic to *Segre varieties on certain linear forms* [1, Prop. A.5].)

To summarize, the 5-dimensional neuromanifold can be interpreted as the image of the map

$$(\mathbb{R}^2)^* \times (\mathbb{R}^2)^* \times (\mathbb{R}^2)^* \to \mathrm{Sym}_2(\mathbb{R}^2)^2 \cong \mathbb{R}^6,$$
$$(\alpha_1, \alpha_2, \nu) \mapsto (\alpha_1\nu, \alpha_2\nu).$$

This semi-algebraic set is not Zariski closed and it has singularities. A generic slice is shown in Figure 4.

References

[1] Nathan W Henry, Giovanni Luca Marchetti, and Kathlén Kohn. Geometry of lightning self-attention: Identifiability and dimension. In *International Conference on Learning Representations*, 2025.