# Introduction: The Interplay of Symmetric Tensors and Neural Networks

Speaker: Maksym Zubkov, University of British Columbia
Venue: Informal Reading Group on Classical Algebraic Geometry
Date: September 9th, 2025

**Field**: Applied algebraic geometry
**Key Results**: Defining and establishing the connection between homogeneous polynomial and symmetric tensors. Discussing their application to algebraic neural networks.

## 1. Homogeneous polynomials

One of the main objects of our study will be homogeneous polynomials and symmetric tensors. We begin by defining homogeneous polynomials.

Let $F$ be a field of characteristic zero. For the remainder of this reading group, we will work mostly over the complex numbers $\mathbb{C}$ (or over the reals $\mathbb{R}$, when specified).

**Definition 1.1** *A* homogeneous polynomial *of degree $r$ over a field $F$ is a polynomial in which every monomial has degree $r$.*

**Example 1.1** *Let $F = \mathbb{C}$ and $f \in \mathbb{C}[x, y, z]$. Then*

$$f(x, y, z) = x^3 + y^3 + z^3 + 6xyz$$

*is a homogeneous polynomial of degree $r = 3$.*

**Remark 1.1** *The space of all homogeneous polynomials of degree $r$ in $n$ variables with coefficients in $F$ will be denoted by $S^r(F^n)$.*

**Exercise 1.1** *What is the dimension of $S^r(F^n)$ as a vector space over $F$?*

## 2. Symmetric tensors

Our next object of study is a symmetric tensor of order $r$ over an $n$-dimensional vector space. Before introducing symmetric tensor, let us first define a general tensor.

We think of a general tensor as a multiarray, i.e., as an element of the space

$$T \in F^{d_1 \times d_2 \times \cdots \times d_r},$$

where $r$ is called the *order* of the tensor and each $d_i$ is a positive integer. Equivalently, we may write

$$T = (a_{i_1 \dots i_r}),$$

where $T(i_1, \dots, i_r) = a_{i_1 \dots i_r}$ are its entries.

**Example 2.1** *Let $T \in F^{2\times2\times2}$. Then $T$ is a tensor of order 3 whose entries may be displayed as two $2 \times 2$ matrices:*

$$T = \left\{ \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \right\}.$$

This is a somewhat rudimentary definition, but it will suffice for our purposes.

### 2.1 Definition of a symmetric tensor

Let $\mathfrak{S}_r$ denote the symmetric group on $r$ elements. We now define symmetric tensors.

**Definition 2.1 (Symmetric tensor)** *A tensor $T$ of order $r$ is called* symmetric *if it is invariant under any permutation of its indices. That is, for every multi-index $(i_1, \ldots, i_r)$ and every $\sigma \in \mathfrak{S}_r$,*

$$T(i_1, \ldots, i_r) \;=\; T\big(i_{\sigma(1)}, \ldots, i_{\sigma(r)}\big).$$

**Example 2.2** *Let $T \in F^{2\times 2\times 2}$. One example of a symmetric tensor of order 3 is the $2 \times 2 \times 2$ array with entries*

$$T = \left\{ \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}, \begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix} \right\}.$$

**Remark 2.1** *The space of all symmetric tensors of order $r$ on $F^n$ is denoted by $Sym_r(F^n)$.*

## 3. The connection between homogeneous polynomials and symmetric tensors

Next, we want to construct the maps

$$\begin{aligned} L : S^r(F^n) \;&\to\; Sym^r(F^n), \\ R : Sym^r(F^n) \;&\to\; S^r(F^n). \end{aligned} \tag{1}$$

### 3.1 The $L$ map

Let $\partial_{x_i} : F[x_1, \ldots, x_n] \to F[x_1, \ldots, x_n]$ denote the formal partial derivative

$$\partial_{x_i}(f) = \frac{\partial f}{\partial x_i}.$$

If $f \in S^r(F^n)$, we may write

$$f(\mathbf{x}) = \sum_{|I|=r} a_I \mathbf{x}^I,$$

where $I = (i_1, i_2, \ldots, i_n)$ is a multi-index, $a_I = a_{i_1 i_2 \ldots i_n}$, and $\mathbf{x}^I = x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}$.

For a multi-index $I = (i_1, \ldots, i_n)$, define the differential operator

$$\partial_{\mathbf{x}^I} = \big(\partial_{x_1}\big)^{i_1} \circ \cdots \circ \big(\partial_{x_n}\big)^{i_n}.$$

Then the map $L$ associates to $f$ the symmetric tensor $T_f$ of order $r$ defined by

$$T_f(i_1, \ldots, i_n) \;=\; \partial_{\mathbf{x}^I}(f).$$

**Example 3.1** *Let $f \in S^2(\mathbb{C}^3)$ be given by*

$$f(x, y, z) = ax^2 + by^2 + cz^2 + Axy + Bxz + Cyz.$$

*Then*

$$L(f) = \begin{bmatrix} 2a & A & B \\ A & 2b & C \\ B & C & 2c \end{bmatrix}.$$

### 3.2 The $R$ map

Now, given a symmetric tensor $T$, we want to produce a homogeneous polynomial.

Define

$$R(T) = f_T,$$

where

$$f_T(\mathbf{x}) = \sum_{|I|=r} \frac{1}{\binom{n}{i_1, i_2, \ldots, i_n}} T(i_1, \ldots, i_n)\, \mathbf{x}^I.$$

**Exercise 3.1** *If $F$ is a field of characteristic zero, then the maps $L$ and $R$ establish a one-to-one correspondence between the space of homogeneous polynomials $S^r(F^n)$ and the space of symmetric tensors $Sym^r(F^n)$.*

## 4. Rank one symmetric tensors

Let $V$ be a vector space over $F$, and $V^*$ its dual vector space, i.e.

$$V^* = \{\alpha : V \to F \mid \alpha \text{ is a linear functional}\}.$$

A simple tensor $\alpha \otimes v \in V^* \otimes V$ defines a rank-one linear map $V \to V$ by $x \mapsto \alpha(x)\, v$.

**Definition 4.1** *The tensor $\alpha \otimes v \in V^* \otimes V$ is called* rank one *if its corresponding linear map*

$$T_{\alpha,v} : V \to V, \quad T_{\alpha,v}(x) = \alpha(x)\, v$$

*is rank one.*

**Example 4.1** *Let $V = \mathbb{C}^2$ with basis $e_1, e_2$ and dual basis $e^1, e^2$. If $A \in F^{2\times 2}$ with*

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix},$$

*then the associated tensor in $V^* \otimes V$ is*

$$\widetilde{A} \;=\; a_{11}\, e^1 \otimes e_1 \;+\; a_{12}\, e^1 \otimes e_2 \;+\; a_{21}\, e^2 \otimes e_1 \;+\; a_{22}\, e^2 \otimes e_2.$$

In coordinates, if $v, w \in V$, then the rank-one tensor can be written as the outer product

$$v \otimes w = v^T w,$$

where $v, w$ are row vectors.

**Example 4.2** *If*

$$a = \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix},$$

*then*

$$a \otimes a = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix} = \begin{bmatrix} a_1^2 & a_1a_2 & a_1a_3 \\ a_1a_2 & a_2^2 & a_2a_3 \\ a_1a_3 & a_2a_3 & a_3^2 \end{bmatrix}.$$

**Definition 4.2** *If $v_1, \ldots, v_n \in V$ have coordinates $v_{ij}$ (where $i$ indexes the vector $v_i$ and $j$ indexes the $j$th coordinate), then their tensor (outer) product is defined by*

$$v_1 \otimes v_2 \otimes \cdots \otimes v_n = \big(v_{1i_1} v_{2i_2} \cdots v_{ni_n}\big),$$

*where the entry at $(i_1, \ldots, i_n)$ is the product of the corresponding coordinates.*

### 4.1 Homogeneous polynomials and Veronese map

**Exercise 4.1** *If* $a = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix}$, *then*

$$R(a \otimes \cdots \otimes a) = (a_1 x_1 + \cdots + a_n x_n)^r.$$

In other words, we have a map $\mathcal{V}_r : V \to S^r(F^n)$ given by

$$\mathcal{V}_r(a) = (a_1 x_1 + \cdots + a_n x_n)^r,$$

which is called the *Veronese map.*

**Exercise 4.2** *When $r = 2$, show that the image of $\mathcal{V}_2$ is the set of symmetric matrices of rank one, i.e., the variety defined by the vanishing of all $2 \times 2$ minors.*

## 5. Symmetric Rank

In other words, the Veronese map parametrizes all symmetric rank-one tensors. What is a tensor of rank $m$?

**Definition 5.1** *A symmetric tensor $T$ has* symmetric rank $m$ *if it can be written minimally as a linear combination of $m$ symmetric rank-one tensors, i.e.,*

$$T = v_1^{\otimes r} + v_2^{\otimes r} + \cdots + v_m^{\otimes r},$$

*where each $v_i \in V$, and no such expression exists with fewer than $m$ terms. We write*

$$rank_S(T) = m.$$

**Example 5.1** *If*

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

*then*

$$A = \tfrac{1}{2}\Big((e_1 + e_2) \otimes (e_1 + e_2) \; - \; (e_1 - e_2) \otimes (e_1 - e_2)\Big).$$

**Exercise 5.1** *Show that* $\operatorname{rank}_S(A) = 2$.

### 5.1 Waring Rank

On the level of homogeneous forms, the symmetric rank is called the *Waring rank.* In other words, a form $f \in S^r(F^n)$ has Waring rank $m$ if there exist $m$ nonzero linear forms $l_1, \dots, l_m$ such that

$$f = l_1^r + \cdots + l_m^r.$$

**Example 5.2** *If $f(x, y) = 2xy$, then*

$$f(x, y) = (x + y)^2 - (x - y)^2.$$

## 6. Algebraic Shallow Neural Networks

Let $\mathbf{d} = (n, m, k)$ be a tuple of positive integers. Consider a map $f_{\mathbf{w}} : F^n \to F^k$ defined by

$$f_{\mathbf{w}}(\mathbf{x}) = W_2 \sigma(W_1 \mathbf{x}),$$

where $W_1 \in F^{m \times n}$, $W_2 \in F^{k \times m}$, and $\sigma$ is an algebraic activation function (polynomial, rational, or ReLU) applied componentwise.

### 6.1 Polynomial activation

If $\sigma(x) = x^r$, then $f_{\mathbf{w}} \in (S^r(F^n))^k$.

**Example 6.1** *Let* $\mathbf{d} = (2, 2, 1)$ *and* $r = 2$. *Then the output of the network is*

$$f_{\mathbf{w}}(\mathbf{x}) = \begin{bmatrix} b_1 & b_2 \end{bmatrix} \sigma\left( \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right).$$

*Writing* $\ell_1 = a_{11}x_1 + a_{12}x_2$ *and* $\ell_2 = a_{21}x_1 + a_{22}x_2$, *this becomes*

$$f_{\mathbf{w}}(\mathbf{x}) = \begin{bmatrix} b_1 & b_2 \end{bmatrix} \begin{bmatrix} \ell_1^2 \\ \ell_2^2 \end{bmatrix} = b_1\ell_1^2 + b_2\ell_2^2.$$

*Expanding, we obtain*

$$f_{\mathbf{w}}(\mathbf{x}) = (b_1a_{11}^2 + b_2a_{21}^2)x_1^2 + (2b_1a_{11}a_{12} + 2b_2a_{21}a_{22})x_1x_2 + (b_1a_{12}^2 + b_2a_{22}^2)x_2^2.$$