Tensors

Paul Breiding
March 7, 2023

Tensors are a generalization of matrices and can be viewed as tables of higher dimensions: A 2×2-matrix is a table that contains 4 numbers aligned in two direction and each direction has dimension 2; a $2 \times 2 \times 2$ tensor is a table with 8 numbers aligned in three directions where each direction has dimension 2.

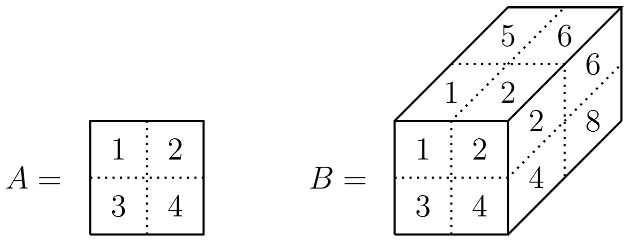


Figure 1: A 2×2 matrix A and a $2 \times 2 \times 2$ tensor B.

For positive integers $n_1, \ldots, n_d$ the set of complex $n_1 \times \cdots \times n_d$–tensors is defined as the set of d-dimensional tables

$$\mathbb{C}^{n_1 \times \cdots \times n_d} := \left\{ (a_{i_1,\ldots,i_d})_{1 \le i_1 \le n_1, \ldots, 1 \le i_d \le n_d} \mid a_{i_1,\ldots,i_d} \in \mathbb{C} \right\}.$$

In particular, tensors of order 2 are matrices. In the literature, elements in $\mathbb{C}^{n_1 \times \cdots \times n_d}$ are also sometimes called *hypermatrices* or *Cartesian tensors* because our definition relies on a choice of basis. The subspace of real tensors is denoted $\mathbb{R}^{n_1 \times \cdots \times n_d}$.

The number d is called the *order* or *power* of the tensor. The tuple $(n_1, \ldots, n_d)$ is called the *format* of the tensor. In the following, $(n_1, \ldots, n_d)$ will be fixed and most of the subsequent notation, even if it involves the n_i, will not refer to them. We may may add tensors and multiply them by scalars so that $\mathbb{C}^{n_1 \times \cdots \times n_d}$ becomes a vector space.

The main focus of this lecture will be on *rank-one tensors*. Let $A = (a_{i_1,\ldots,i_d}) \in \mathbb{C}^{n_1 \times \cdots \times n_d}$ be a tensor. We say that A has rank-one, if there exist vectors $v_i \in \mathbb{C}^{n_i}, 1 \le i \le d$, such that $a_{i_1,\ldots,i_p} = (v_1)_{i_1} \cdot \ldots \cdot (v_d)_{i_d}$ for all tuples of indices $(i_1, \ldots, i_d)$. We write this as $A = v_1 \otimes \cdots \otimes v_d$. The *Segre variety* is then defined as

$$\mathcal{S} := \{ v_1 \otimes \cdots \otimes v_d \mid v_i \in \mathbb{C}^{n_i}, 1 \le i \le d \}.$$

The most immediate way to see that $\mathcal{S}$ is an algebraic variety goes as follows. We can always *flatten* a tensor $A \in \mathbb{C}^{n_1 \times \cdots \times n_d}$ into d matrices $F_1, \ldots, F_d$, where $F_i \in \mathbb{C}^{n_i \times (\prod_{j \neq i} n_j)}$. Then, we have $A \in \mathcal{S}$ if and only if the F_i are all of rank at most 1; i.e., when their 2×2–minors vanish. The real part $\mathcal{S}_{\mathbb{R}}$ of $\mathcal{S}$ consists of tensors $v_1 \otimes \cdots \otimes v_d$ with $v_i \in \mathbb{R}^{n_i}$.

Example 1. Take $v = (v_1, v_2) \in \mathbb{C}^2$ and $w = (w_1, w_2, w_3) \in \mathbb{C}^2$. Then,

$$v \otimes w = \begin{bmatrix} v_1 \cdot w_1 & v_1 \cdot w_2 & v_1 \cdot w_3 \\ v_2 \cdot w_1 & v_2 \cdot w_2 & v_2 \cdot w_3 \end{bmatrix}$$

is the rank one matrix with column space $\mathbb{C} \cdot v$ and row space $\mathbb{C} \cdot w$. Another common notation is vw^T. $\diamond$

Given a d-tuple of matrices $(M_1, \ldots, M_d) \in \mathbb{C}^{k_1 \times n_1} \times \cdots \times \mathbb{C}^{k_d \times n_d}$ we define

$$(M_1, \ldots, M_d).(v_1 \otimes \cdots \otimes v_d) := (M_1 v_1) \otimes \cdots \otimes (M_d v_d), \tag{1}$$

and we extend this action linearly to all of $\mathbb{C}^{n_1 \times \cdots \times n_d}$. The action (1) is called *multilinear multiplication*. It induces a representation of $\mathrm{GL}(n_1) \times \cdots \times \mathrm{GL}(n_d)$ into the general linear group of $\mathbb{C}^{n_1 \times \cdots \times n_d}$.

Example 2. Let $A \in \mathbb{C}^{n_1 \times n_2}$ and $(M_1, M_2) \in \mathbb{C}^{k_1 \times n_1} \times \mathbb{C}^{k_2 \times n_2}$. Then, $(M_1, M_2).A = M_1 A M_2^T$. Thus, multilinear multiplication is a generalization of simultaneous left–right multiplication for matrices.

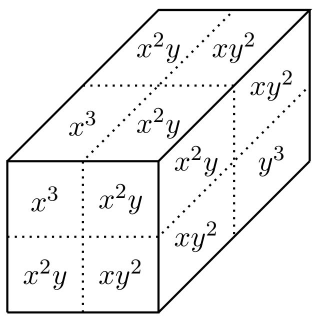
If $n := n_1 = \cdots = n_d$ are all equal, we write $(\mathbb{C}^n)^{\otimes d}$ for the space of order–d tensors of format $n \times \cdots \times n$. The action of the symmetric group $\mathfrak{S}_d$ on $(\mathbb{C}^n)^{\otimes d}$ is defined to be the linear extension of the action $\pi(v_1 \otimes \cdots \otimes v_d) = v_{\pi(1)} \otimes \cdots \otimes v_{\pi(d)}$. This is the generalization of matrix transposition. Let $A \in (\mathbb{C}^n)^{\otimes d}$. We say A is *symmetric* if $\pi(A) = A$ for all $\pi \in \mathfrak{S}_d$. The space of symmetric tensors in $(\mathbb{C}^n)^{\otimes d}$ is denoted $S^d(\mathbb{C}^n)$. Since $S^d(\mathbb{C}^n)$ is the image of the linear map $A \mapsto \sum_{\pi \in \mathfrak{S}_d} \pi(A)$, it is a linear subspace. The dimension of $S^d(\mathbb{C}^n)$ is $\binom{n+d-1}{d}$. Symmetric rank one tensors are written as $v^{\otimes d} := v \otimes \cdots \otimes v$. The *Veronese variety* is the variety of symmetric rank-one tensors

$$\mathcal{V} := \{v^{\otimes d} \mid v \in \mathbb{C}^n\}.$$

Since $\mathcal{V}$ is the intersection of $\mathcal{S}$ with a linear subspace, it is an algebraic variety. Real symmetric tensors are denoted $S^d(\mathbb{R}^n)$. The real part $\mathcal{V}_{\mathbb{R}}$ of $\mathcal{V}$ consists of tensors $v^{\otimes d}$ where v is a real vector.

Example 3. Let $v = (x, y) \in \mathbb{C}^2$. Then,

$v^{\otimes 3} =$



The four independent coordinates of this symmetric tensor give all monomials of degree 3 in x and y. In general, $v^{\otimes d}$ is given by all monomials of degree d in the entries of v. $\diamond$

We define a *Hermitian structure* on $\mathbb{C}^{n_1 \times \cdots \times n_d}$. For $A = (a_{i_1,\ldots,i_d})$ and $B = (b_{i_1,\ldots,i_d})$:

$$\langle A, B \rangle := \sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} \overline{a_{i_1,\ldots,i_d}} \cdot b_{i_1,\ldots,i_d}$$

That is, if we interpret A and B as vectors in $\mathbb{C}^{n_1 \cdots n_d}$, then $\langle A, B \rangle$ is the standard Hermitian inner product. The induced norm is $\|A\| = \sqrt{\langle A, A \rangle}$. The Hermitian structure induces the *Euclidean topology* on $\mathbb{C}^{n_1 \times \cdots \times n_d}$. We also define the bilinear form

$$\beta(A, B) := \sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} a_{i_1,\ldots,i_d} \cdot b_{i_1,\ldots,i_d}.$$

For real tensors $\beta(A, B) = \langle A, B \rangle$, but for complex tensors they are different.

Example 4. Consider the 2×2 matrix

$$A = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{-1} \end{bmatrix}.$$

Then $\langle A, A \rangle = 1 + 1 = 2$, but $\beta(A, A) = 1 - 1 = 0$. $\diamond$

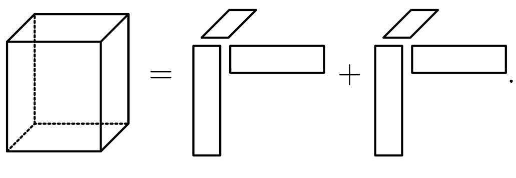
The inner product between rank-one tensors is

$$\begin{aligned} \langle v_1 \otimes \cdots \otimes v_d, w_1 \otimes \cdots \otimes w_d \rangle &= (v_1^T, \ldots, v_d^T).(v_1 \otimes \cdots \otimes v_d) \\ &= \langle v_1, w_1 \rangle \cdots \langle v_d, w_d \rangle. \end{aligned} \tag{2}$$

The Segre variety induces a notion of rank for tensors.

$$\operatorname{rank}(A) := \{ r \mid \text{there exists } A_1, \ldots, A_r \in \mathcal{S} \text{ with } A = A_1 + \cdots + A_r \}.$$

For matrices ($d = 2$) this coincides with the usual matrix rank. We denote tensors of rank at most r by $\Sigma_r := \min \{ A \in \mathbb{R}^{n_1 \times \cdots \times n_d} \mid \operatorname{rank}(A) \le r \}$. A tensor of order 3 and rank 2 can be visualized as follows:



In the case of matrices Σ_r is a variety for every r, defined by $(r + 1) \times (r + 1)$–minors. For $d \ge 3$ and $r \ge 2$, however, Σ_r is not necessarily a variety anymore. This is implied by the following result going back to the work by da Silva and Lim [5].

Proposition 5. *For $d \ge 3$ the set of rank at most 2-tensors Σ_2 is not closed in the Euclidean topology.*

Proof. First consider the case $d = 3$. Let $x_1, x_2 \in \mathbb{R}^{n_1}$ and $y_1, y_2 \in \mathbb{R}^{n_2}$ and $z_1, z_2 \in \mathbb{R}^{n_1}$ be three tuples of independent vectors. Define for $\varepsilon > 0$ the tensor $A_\varepsilon \in \Sigma_2$ by

$$A_\varepsilon := \varepsilon \left(x_1 + \varepsilon^{-1} x_2\right) \otimes \left(y_1 + \varepsilon^{-1} y_2\right) \otimes \left(z_1 + \varepsilon^{-1} z_2\right) - \varepsilon x_1 \otimes y_1 \otimes z_1.$$

Then, $A := \lim_{\varepsilon \to 0} A_\varepsilon = x_1 \otimes y_1 \otimes z_2 + x_1 \otimes y_2 \otimes z_1 + x_2 \otimes y_1 \otimes z_1$. To see that this tensor is of rank 3, we first observe that $A = (P_x, P_y, P_z).A$, where P_x is the projection onto the plane spanned by x_1 and x_2, and similar for P_y and P_z. Therefore, we can assume $n_1 = n_2 = n_3 = 2$. Let us consider multilinear multiplication by $X := (\mathbf{1}_2, \mathbf{1}_2, h^T)$, where $h \in \mathbb{C}^2$:

$$X.A = (h^T z_2)\, x_1 \otimes y_1 + (h^T z_1)\, x_1 \otimes y_2 + (h^T z_1)\, x_2 \otimes y_1 \tag{3}$$

Choosing h with $h^T z_1 \neq 0$ yields a matrix of rank 2, which shows that A has rank at least two. Suppose now $A = u_1 \otimes v_1 \otimes w_1 + u_2 \otimes v_2 \otimes w_2$ has rank two. Among the three pairs of vectors there must be at least one that is linearly independent. We assume without restriction that w_1, w_2 are independent. We have

$$X.A = (h^T w_1)\, u_1 \otimes v_1 + (h^T w_2)\, u_2 \otimes v_2. \tag{4}$$

Let us pick h such that $h^T z_1 = 0$. Then, by (3) $X.A$ has rank 1 and we must have $h^T w_1 = 0$ or $h^T w_2 = 0$ by (4). Without restriction we assume $h^T w_1 = 0$. This implies that z_1 is a multiple of w_1, since they both satisfy the linear equation imposed by h. Next, we consider h with $h^T w_1 \neq 0$ but $h^T w_2 = 0$. Then, $h^T z_1 \neq 0$ and so $X.A$ has rank two by (3) and it has rank one by (4). This is a contradiction, so A can't have rank two.

For the case $d \geq 3$ we tensor A with as many factors as needed. □

A decomposition of the form $A = A_1 + \cdots + A_r$, where the $A_i \in \mathcal{S}$, is also called *canonical polyadic decompositions*, or simply rank decomposition. One appealing property of higher order tensors is *identifiability*. That is, many tensor decompositions are actually unique (rank decompositions of matrices are never unique). The following is [3, Theorem 1.1].

Theorem 6. *Let $n_1 \geq \cdots \geq n_d$ with $\prod_{i=1}^d n_i \leq 15000$ and*

$$r_0 = \left\lceil \frac{\dim \mathbb{C}^{n_1 \times \cdots \times n_d}}{\dim \mathcal{S}} \right\rceil = \left\lceil \frac{n_1 \cdots n_d}{1 + \sum_{i=1}^d (n_i - 1)} \right\rceil.$$

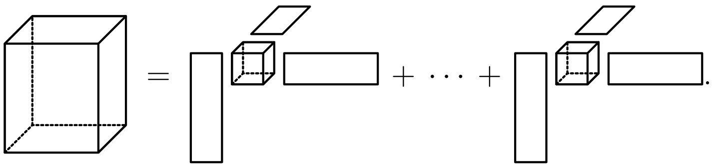
Suppose that $r < r_0$ and $(n_1, \ldots, n_d, r)$ is not one of the following cases

$(n_1, \ldots, n_d)$	r
$(4, 4, 3)$	5
$(4, 4, 4)$	6
$(6, 6, 3)$	8
$(n, n, 2, 2)$	$2n - 1$
$(2, 2, 2, 2, 2)$	5
$n_1 > \prod_{i=2}^d n_i - \sum_{i=2}^d (n_i - 1)$	$r \geq \prod_{i=2}^d n_i - \sum_{i=2}^d (n_i - 1)$

Then, a general tensor in Σ_r has a unique rank $-r$ decomposition.

Remark 7. Another important decomposition is the so-called *block term decomposition*. This is a decomposition of the form $A = A_1 + \cdots + A_r$, where the A_i are low-multilinear rank tensors. A block term decomposition models it a mixture of distributions which allow correlations between the variables, other than the rank decomposition, which models a mixture of independence models. An example, where this is relevant, is detecting epileptic seizures [10]. The interaction between the variables in this case is extremely complex, so that a mixture of independence models is not the appropriate model.

The definition of the block term decomposition is as follows: Let $k = (k_1, \ldots, k_d)$ be a vector of integers with $1 \leq k_i \leq n_i$. Let $A \in \mathbb{C}^{n_1 \times \cdots \times n_d}$ and $F_1, \ldots, F_d$ be the flattenings of A. Then, we say that A has multilinear-rank (at most) k, if $\operatorname{rank}(F_j) \leq k_j$ for $1 \leq j \leq d$. Note that the rank decomposition above is the special case $k = (1, \ldots, 1)$. For order-3 tensors a block term decomposition can be visualized as follows:



Identifiability of block term decompositions is less well-studied than for rank-r decompositions. Results exist for the decomposition of tensors into tensors of multilinear rank $(1, k_1, k_2)$ [13], $(1, k, k)$ [4,6], $(k_1, k_2, 1)$ [12], and (k_1, k_2, k_3) [11]. $\diamond$

Proposition 5 implies that for tensors of order $d \geq 3$ the problem of computing $\min_{B \in \Sigma_r} \|A - B\|$ for a tensor A can be ill-posed (the minimizer does not need to exist). In fact, da Silva and Lim [5] prove that there is full-dimensional open subset of tensors such that this problem is ill-posed. This is different for matrices, where the Eckart-Young theorem provides an explicit algorithm for computing the minimizer. By contrast, the Segre– and the Veronese–variety are closed (both in the Euclidean and Zariski topology). Let us study the critical points of the corresponding distance minimzation problem; that is, we study the *Euclidean Distance Degree* of these varieties.

A key property of both the Segre and Veronese variety is that they are *unirational*. One can see that each of them is the image of a polynomial map. Let us denote them by

$$\psi : \mathbb{C}^{n_1} \times \cdots \times \mathbb{C}^{n_d} \to S, \ (v_1, \ldots, v_d) \mapsto v_1 \otimes \cdots \otimes v_d, \tag{5}$$

and

$$\nu : \mathbb{C}^n \to S, \ v \mapsto v^{\otimes d}. \tag{6}$$

These maps are called the Segre– and Veronese–map, respectively.

We first study distance minimzation for the Segre variety. The goal is to understand the Euclidean distance degree of $\mathcal{S}_\mathbb{R}$. For this, we consider a real tensor $A \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ and the optimization problem $\min_{v_i \in \mathbb{R}^{n_i}} \|A - \psi(v_1, \ldots, v_d)\|$. Since ψ is a polynomial map, it is smooth and the critical values are given by $\frac{\mathrm{d}}{\mathrm{d}v_i} \|A - \psi(v_1, \ldots, v_d)\|^2 = 0, i = 1, \ldots, d$. Let us write

$$\|A - \psi\|^2 = \|A\|^2 - 2\langle A, \psi \rangle + \|\psi\|^2.$$

By (2), $\|\psi\|^2 = \|v_1\|^2 \cdots \|v_d\|^2$. This shows, since we are dealing with real tensors,

$$\begin{aligned}\tfrac{\mathrm{d}}{\mathrm{d}v_i}\|A-\psi\|^2 &= -2\tfrac{\mathrm{d}}{\mathrm{d}v_i}\langle A,\psi\rangle + \tfrac{\mathrm{d}}{\mathrm{d}v_i}(\|v_1\|^2\cdots\|v_d\|^2)\\ &= \begin{bmatrix}\beta(A, v_1\otimes\cdots\otimes v_{i-1}\otimes e_1\otimes v_{i+1}\otimes\cdots\otimes v_d)\\ \vdots\\ \beta(A, v_1\otimes\cdots\otimes v_{i-1}\otimes e_{n_i}\otimes v_{i+1}\otimes\cdots\otimes v_d)\end{bmatrix} - 2\Big(\prod_{j\neq i}\|v_j\|^2\Big)\, v_i,\end{aligned}$$

where $e_1,\ldots,e_{n_i}$ denotes the standard basis of $\mathbb{R}^{n_i}$. We have used the bilinear form β instead of $\langle\ ,\ \rangle$, because it turns solving $\frac{\mathrm{d}}{\mathrm{d}v_i}\|A-\psi\|^2$ into solving a system of polynomial equations. Complex solutions to this are called *singular vector tuples* of A. This is summarized in the next definition.

Definition 8. Let $A\in\mathbb{C}^{n_1\times\cdots\times n_d}$. We say that $(u_1,\ldots,u_d)\in\mathbb{C}^{n_1}\times\cdots\times\mathbb{C}^{n_d}$, $u_i\neq 0$, is a *singular vector tuple* for A, if there exists $\sigma_1,\ldots,\sigma_d\in\mathbb{C}$, called *singular values*, with

$$\begin{bmatrix}\beta(A, u_1\otimes\cdots\otimes u_{i-1}\otimes e_1\otimes u_{i+1}\otimes\cdots\otimes v_d)\\ \vdots\\ \beta(A, u_1\otimes\cdots\otimes u_{i-1}\otimes e_{n_i}\otimes u_{i+1}\otimes\cdots\otimes u_d)\end{bmatrix} = \sigma_i u_i$$

for $i=1,\ldots,d$.

For $d=2$ this coincides with the classic definition of singular vector pairs for matrices. The singular value decomposition implies that for a general matrix $A\in\mathbb{R}^{n_1\times n_2}$ there are precisely $\min\{n_1,n_2\}$ singular vector pairs. For higher order tensors their number is given by the following theorem due to Friedland and Ottaviani [8].

Theorem 9. *Let $A\in\mathbb{C}^{n_1\times\cdots\times n_d}$ be a general tensor. The number of singular vector tuples of the tensor A, up to scaling $(u_1,\ldots,u_d)\mapsto(t_1u_1,\ldots,t_1u_d)$, is the coefficient of the monomial $x_1^{n_1-1}\cdots x_d^{n_d-1}$ in the polynomial*

$$\prod_{i=1}^{d}\frac{f_i^{n_i}-x_i^{n_i}}{f_i-x_i},\quad f_i=\sum_{j\neq i}x_j.$$

Example 10. We consider the formula in Theorem 9 for binary tensors $A\in\mathbb{C}^{2\times\cdots\times 2}$. In this case, $\prod_{i=1}^d\frac{f_i^2-x_i^2}{f_i-x_i}=\prod_{i=1}^d(f_i+x_i)=(x_1+\cdots+x_d)^d$. The coefficient of $x_1\cdots x_d$ in this polynomial is $d!$.

If we follow the approach above for symmetric tensors and the Veronese variety, we arrive at the notion of *eigenpairs of tensors*. The definition comes next.

Definition 11. Let $A\in\mathcal{S}^d(\mathbb{C}^n)$. We say that $u\in\mathbb{C}^n\setminus\{0\}$ is an *eigenvector* of A, if there exists $\lambda\in\mathbb{C}$ with

$$\begin{bmatrix}\beta(A, e_1\otimes u^{\otimes d})\\ \vdots\\ \beta(A, e_n\otimes u^{\otimes d})\end{bmatrix}=\lambda u.$$

The pair (u,λ) is called an eigenpair of A.

Eigenpairs have another interesting interpretation, next to being critical points for the Euclidean distance function on the real Veronese variety. Namely, if $A \in \mathcal{S}^d(\mathbb{C}^n)$ is symmetric and $v \in \mathbb{C}^n$, then

$$F_A(v) := \beta(A, u^{\otimes d}) = \sum_{j=1}^{d} \sum_{i_j=1}^{n_j} a_{i_1,\ldots,i_d}\, v_{i_1} \cdots v_{i_d} \tag{7}$$

is a homogeneous polynomial of degree d in n variables. Moreover, the map $A \to F_A$ is a linear isomorphism between $S^d(\mathbb{C}^n)$ and the vector space of homogeneous polynomials of degree d in n variables with complex coefficients. Eigenpairs thus correspond to fixed points of the rational map $\mathbb{P}^{n-1} \dashrightarrow \mathbb{P}^{n-1}, v \mapsto (\partial F_A/\partial v_1, \ldots, \partial F_A/\partial v_n)$ given by the gradient of F_A.

Example 12. Consider the symmetric matrix

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 0 \end{bmatrix}.$$

Then,

$$F_A(v) = v^T A v = v_1^2 + 4v_1 \cdot v_2$$

On the other hand, we can recover A from the coefficients of F_A. The gradient of F_A is then

$$\begin{bmatrix} \partial F_A/\partial v_1 \\ \partial F_A/\partial v_2 \end{bmatrix} = \begin{bmatrix} 2v_1 + 4v_2 \\ 4v_1 \end{bmatrix} = 2Av,$$

which simply means that eigenvectors of A are given by the equation $Av = \lambda v$. ◇

Example 13. Consider the polynomial $F_A(v) = v_1^3 + v_2^3$, which corresponds to a $2 \times 2 \times 2$–tensor A. The eigenpairs of A are solutions to the system of equations

$$\begin{bmatrix} u_1^2 - \lambda u_1 \\ u_2^2 - \lambda u_2 \end{bmatrix} = 0.$$

If $\lambda = 0$, then $u = 0$, which by definition is not an eigenvector. So, we can set $\lambda = 1$. The other solutions are $u \in \{(1,1),\ (0,1),\ (1,0)\}$. So, we have 3 eigenvectors up to scaling.

In fact, we can remove the assumption that A is symmetric from Definition 11 and define eigenpairs for general tensors, as well. This way, we lose the interpretations above, but we have a more general definition. Let us count the number of eigenpairs of a general tensor. The following theorem was first proved by Cartwright and Sturmfels in [2] using intersection theory. Here, we give a simpler proof.

Theorem 14. *Let $A \in (\mathbb{C}^n)^{\otimes d}$ be a general tensor. The number of eigenvectors of A, up to scaling $u \mapsto tu$, is $\sum_{i=0}^{n-1}(d-1)^i$.*

Observe that for $d = 2$ the formula gives n, which is the number of eigenpairs of a general matrix $A \in \mathbb{R}^{n \times n}$.

Proof of Theorem 14. Consider the tensor $A \in \mathcal{S}^d(\mathbb{C}^n)$ with $F_A(v) = v_1^d + \cdots + v_n^d$ (for $d = 3, n = 3$ this gives the tensor from Example 13). Let us substitute $\lambda = \xi^{d-2}$. The equations for eigenpairs are then

$$\begin{bmatrix} v_1^{d-1} - \xi^{d-2} v_1 \\ \vdots \\ v_n^{d-1} - \xi^{d-2} v_n \end{bmatrix} = 0.$$

We count that this system of homogeneous polynomial equations has precisely $(d-1)^n$ isolated solutions in $\mathbb{P}^n$. By Bézout's theorem, a general system of n polynomials with degrees $(d-1, \ldots, d-1)$ also has $(d-1)^n$ isolated zeros. Therefore, the family of systems $F_A(v) - \xi^{d-2} v$ has $(d-1)^n$ isolated zeros for a general A. Moreover, for general A and a zero $F_A(u) - \xi^{d-2} u = 0$ we must have $\xi \neq 0$, because $F_A(u) \neq 0$ (since it is a general system of n homogeneous equations in n variables). The number of eigenpairs for a general A is

$$\frac{(d-1)^n - 1}{d-2} = \sum_{i=0}^{n-1} (d-1)^i,$$

because we have to remove $u = 0$ from the count and divide the count by $(d-2)$ to account for scaling of ξ with $(d-2)$–th roots of unity. □

For matrices every n-tuple of linearly independent vectors can be eigenvectors of a matrix. For tensors this is not so. Abo, Sturmfels and Seigal proved in [1] that a set of d points $u_1, \ldots, u_d$ in $\mathbb{C}^2$ is the eigenconfiguration of a symmetric tensor in $S^d(\mathbb{C}^2)$ if and only if d is odd, or $d = 2k$ is even and for every of the points u the differential operator $(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2})^k$ annihilates the binary form $f(x, y) = \prod_{i=1}^d (b_i x - a_i y)$, where $u_i = (a_i, b_i)$.

In the second part of the lecture we will now address the problem of computing the volumes of the Segre– and of the Veronese variety. This is an important problem in metric algebraic geometry by itself. But it will also have interesting and unexpected applications as we will see. First, we observe that both $\mathcal{S}$ and $\mathcal{V}$ are cones; i.e., closed under scaling. In particular, they are not compact and do not have finite volume. To make a meaningful computation we pass to projective space $\mathbb{P}^N$ where $N = n_1 \cdots n_d - 1$ (or $N = n^d - 1$ if the n_i are all equal) and measure the volume of the projective Segre variety and projective Veronese variety. Let us denote them by

$$\mathcal{S}^{\mathbb{P}} := \{v_1 \otimes \cdots \otimes v_d \in \mathbb{P}^N \mid v_i \in \mathbb{P}^{n_i - 1}, 1 \leq i \leq d\} \quad \text{and} \quad \mathcal{V}^{\mathbb{P}} := \{v^{\otimes d} \in \mathbb{P}^N \mid v_i \in \mathbb{P}^n\}.$$

Furthermore, let us denote the sphere in $\mathbb{C}^n$ by

$$\mathbb{S}^{2n-1} := \{a \in \mathbb{C}^n \mid \langle a, a \rangle = 1\},$$

where, as above in the space of tensors, $\langle a, b \rangle = \overline{a}^T b$ is the standard Hermitian inner product. The sphere is a real manifold of real dimension $\dim_{\mathbb{R}} \mathbb{S}^{2n-1} = 2n - 1$. We have the projection $\pi : \mathbb{S}^{2n-1} \to \mathbb{P}^{n-1}$ that sends a point $a \in \mathbb{S}^{2n-1}$ to its projective class. The tangent space of $\mathbb{S}^{2n-1}$ at a point a is $T_a \mathbb{S}^{2n-1} = \{\tau \in \mathbb{C}^n \mid \mathrm{Re}\langle a, \tau \rangle = 0\}$.

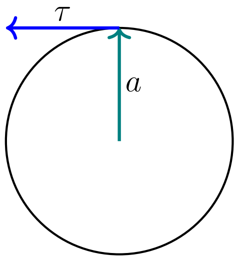


Figure 2: The circle $\mathbb{S}^1$ and a point $a \in \mathbb{S}^1$ together with a tangent vector $\tau \in T_a\mathbb{S}^1$. The tangent vector τ is orthogonal to a; i.e., $\langle a, \tau\rangle = 0$.

We have $\mathrm{Re}\langle a, b\rangle = 0$ if and only if $\langle a, b\rangle = 0$ or $a = \sqrt{-1} \cdot b$. Thus, we choose the following model for the tangent space of projective space:

$$T_x\mathbb{P}^{n-1} = \{\tau \in \mathbb{C}^n \mid \langle a, \tau\rangle = 0\}, \quad x = \pi(a) \text{ for } a \in \mathbb{S}^{2n-1}.$$

The Hermitian structure on $\mathbb{C}^n$ thus turns $\mathbb{P}^{n-1}$ into a Hermitian manifold. Consequently, complex projective space $\mathbb{P}^{n-1}$ is also a Riemannian manifold relative to the Euclidean structure $\mathrm{Re}\langle a, b\rangle$. This induces a notion of volume for subsets of $\mathbb{P}^{n-1}$. More specifically, since the preimage $\pi^{-1}(x)$ for $x \in \mathbb{P}^{n-1}$ is a circle, for $0 \le m \le 2(n-1)$ the m-dimensional real volume of a measurable subset $U \subset \mathbb{P}^{n-1}$ is $\mathrm{vol}_m(U) = \frac{1}{2\pi}\,\mathrm{vol}_{m+1}(\pi^{-1}(U))$. For instance, the volume of projetive space is

$$\mathrm{vol}_{2(n-1)}(\mathbb{P}^{n-1}) = \frac{1}{2\pi}\mathrm{vol}_{2n-1}(\mathbb{S}^{2n-1}) = \frac{\pi^{n-1}}{(n-1)!}. \tag{8}$$

In the following, we will sometimes omit the subscript from vol when the dimension is clear from the context.

We have the following result.

Proposition 15. *Denote $m := \dim_{\mathbb{R}} \mathcal{S}^{\mathbb{P}}$. The m-dimensional volume of $\mathcal{S}^{\mathbb{P}}$ is*

$$\mathrm{vol}(\mathcal{S}^{\mathbb{P}}) = \mathrm{vol}(\mathbb{P}^{n_1-1}) \cdots \mathrm{vol}(\mathbb{P}^{n_d-1}).$$

Proof. We define the Segre map (5) for projective space: $\psi_{\mathbb{P}} : \mathbb{P}^{n_1-1} \times \cdots \times \mathbb{P}^{n_d-1} \to \mathcal{S}^{\mathbb{P}}$. Then, $\psi_{\mathbb{P}}$ is a smooth embedding. Let $(x_1, \ldots, x_d) \in \mathbb{P}^{n_1-1} \times \cdots \times \mathbb{P}^{n_d-1}$ and $a_i \in \mathbb{S}^{2n_i-1}$ with $\pi(a_i) = x_i$ be a fixed representative for x_i. Let also $\tau_i \in \mathbb{C}^{n_i}$ with $\langle a_i, \tau_i\rangle = 0$. Then, the derivative of $\psi_{\mathbb{P}}$ maps $(\tau_1, \ldots, \tau_d) \in T_{(x_1,\ldots,x_d)}(\mathbb{P}^{n_1-1} \times \cdots \times \mathbb{P}^{n_d-1})$ to

$$\theta := \tau_1 \otimes a_2 \otimes \cdots \otimes a_d + a_1 \otimes \tau_2 \otimes \cdots \otimes a_d + \cdots + a_1 \otimes a_2 \otimes \cdots \otimes \tau_d.$$

It follows from (2) that the terms in this sum are pairwise orthogonal. Therefore

$$\|\theta\|^2 = \|\tau_1\|^2 + \|\tau_2\|^2 + \cdots + \|\tau_d\|^2.$$

This shows that the derivative of $\psi_{\mathbb{P}}$ preserves norms. This implies that $\psi_{\mathbb{P}}$ is volume preserving. $\square$

For the Veronese variety we have the following result.

Proposition 16. *The* $2(n-1)$*-dimensional volume of the projective Veronese variety is*

$$\mathrm{vol}(\mathcal{V}^{\mathbb{P}}) = d^{n-1} \cdot \mathrm{vol}(\mathbb{P}^{n-1}).$$

Proof. The proof is similar to that of Theorem 15. We denote the Veronese map (6) for projective space: $\nu_{\mathbb{P}} : \mathbb{P}^{n-1} \to \mathcal{V}^{\mathbb{P}}$. Then, also $\nu_{\mathbb{P}}$ is a smooth embedding. Let $x \in \mathbb{P}^{n-1}$ and $a \in \mathbb{S}^{2n-1}$ be a representative for x; i.e., $\pi(a) = x$. Let $\tau \in \mathbb{C}^n$ with $\langle a, \tau \rangle = 0$. The derivative of $\nu_{\mathbb{P}}$ maps $\tau \in T_x\mathbb{P}^{n-1}$ to

$$\theta := \tau \otimes a \otimes \cdots \otimes a + a \otimes \tau \otimes \cdots \otimes a + \cdots + a \otimes a \otimes \cdots \otimes \tau,$$

It follows from (2) that the terms in this sum are pairwise orthogonal, so $\|\theta\|^2 = d\|\tau\|^2$. This shows that the derivative of $\nu_{\mathbb{P}}$ scales norms by $\sqrt{d}$. This implies that

$$\mathrm{vol}(\mathcal{V}^{\mathbb{P}}) = \left(\sqrt{d}\,\right)^{\dim_{\mathbb{R}} \mathbb{P}^{n-1}} \cdot \mathrm{vol}(\mathbb{P}^{n-1}) = d^{n-1} \cdot \mathrm{vol}(\mathbb{P}^{n-1}).$$

□

Propositions 15 and 16 can be applied to intersection theory using Howard's *Kinematic Formula* [9]. This is a general formula for the average volume of intersections of submanifolds in homogeneous spaces. For complex projective space we have the following specific formula; see [9, Theorem 3.8 & Corollary 3.9]. Let $M \subset \mathbb{P}^N$ be a smooth manifold of complex dimension m. Then, the Kinematic formula for complex projective space is

$$\mathbb{E}_U \, \#(M \cap U \cdot (\mathbb{P}^{N-m} \times \{0\}^m)) = \frac{\mathrm{vol}_{2m}(M)}{\mathrm{vol}_{2(N-m)}(\mathbb{P}^{N-m})},$$

where the expectation is taken relative to the probability measure on the unitary group $U(n+1)$ induced by the Haar measure. If $X \subset \mathbb{P}^N$ is a smooth algebraic variety of complex dimension m, then the number of intersection points $\#(X \cap U \cdot (\mathbb{P}^{N-m} \times \{0\}^m))$ equals the degree of X for almost all U. Thus, we are taking the expected value of a constant function. This shows

$$\deg(X) = \frac{\mathrm{vol}_{2m}(X)}{\mathrm{vol}_{2(N-m)}(\mathbb{P}^{N-m})}, \quad m = \dim_{\mathbb{C}}(X). \tag{9}$$

Combined with the theorems above we obtain the following result.

Corollary 17.

1. $\deg(\mathcal{S}^{\mathbb{P}}) = \dfrac{(n_1 + \cdots + n_d - d)!}{(n_1 - 1)! \cdots (n_d - 1)!}$.
2. $\deg(\mathcal{V}^{\mathbb{P}}) = d^{n-1}$.

Proof. The second formula follows directly from Proposition 16 and (9). For the second formula, we recall from (8) the volume of projective space: $\mathrm{vol}_{2(n-1)}(\mathbb{P}^{n-1}) = \frac{\pi^{n-1}}{(n-1)!}$. Denote $k = n_1 + \cdots + n_d - d = \dim_{\mathbb{C}}(\mathcal{S}^{\mathbb{P}})$. Using Proposition 15 and (9) we then have

$$\deg(\mathcal{S}^{\mathbb{P}}) = \frac{\mathrm{vol}(\mathbb{P}^{n_1-1}) \cdots \mathrm{vol}(\mathbb{P}^{n_d-1})}{\mathrm{vol}(\mathbb{P}^k)} = \frac{\pi^{\sum_{i=1}^d (n_i-1)}}{\pi^k} \cdot \frac{k!}{(n_1-1)! \cdots (n_d-1)!}$$
$$= \frac{(n_1 + \cdots + n_d - d)!}{(n_1-1)! \cdots (n_d-1)!}.$$

□

By (7), a linear equation on $\deg(\mathcal{V}^{\mathbb{P}})$ corresponds to the evaluation of a homogeneous polynomial of degree d in n variables. Thus, $\deg(\mathcal{V}^{\mathbb{P}}) = d^{n-1}$ means that a general system of $n-1$ homogeneous polynomials of degree d has d^{n-1} zeros.

Remark 18. Howard's Kinematic Formula from [9] also provides the following result for real projective space. Let $\mathbb{P}^N_{\mathbb{R}}$ denote real projective space and let $M \subset \mathbb{P}^N_{\mathbb{R}}$ be a real submanifold of real dimension m. Then, $\mathbb{E}_U \, \#(M \cap U \cdot (\mathbb{P}^{N-m}_{\mathbb{R}} \times \{0\}^m)) = \frac{\mathrm{vol}_m(M)}{\mathrm{vol}_{N-m}(\mathbb{P}^{N-m}_{\mathbb{R}})}$, where here the expectation is taken relative to the probability measure on the orthogonal group $O(n+1)$. Thus, the volume of real projective varieties can be interpreted as an "average degree". We can use the same proof strategies as above to show that the volume of the real Segre variety $\mathcal{S}^{\mathbb{P}}_{\mathbb{R}}$ is equal to $\mathrm{vol}(\mathbb{P}^{n_1-1}_{\mathbb{R}}) \cdots \mathrm{vol}(\mathbb{P}^{n_d-1}_{\mathbb{R}})$ and the volume of the real Veronese $\mathcal{V}^{\mathbb{P}}_{\mathbb{R}}$ is $\sqrt{d^{n-1}} \cdot \mathrm{vol}(\mathbb{P}^{n-1}_{\mathbb{R}})$. The latter result was first observed by Edelman and Kostlan in their seminal paper [7] to show that a homogeneous polynomial of degree d in n variables has on the average $\sqrt{d}$ many real zeros.

References

[1] H. Abo, A. Seigal, and B. Sturmfels. *Eigenconfigurations of tensors*, pages 1–25. 01 2017.

[2] D. Cartwright and B. Sturmfels. The number of eigenvalues of a tensor. *Linear Algebra and its Applications*, 438(2):942–952, 2013. Tensors and Multilinear Algebra.

[3] L. Chiantini, G. Ottaviani, and N. Vannieuwenhoven. An algorithm for generic and low-rank specific identifiability of complex tensors. *SIAM J. Matrix Anal. Appl.*, 35(4):1265–1287, 2014.

[4] L. De Lathauwer. Blind Separation of Exponential Polynomials and the Decomposition of a Tensor in Rank-$(L_r, L_r, 1)$ Terms. *SIAM J. Matrix Anal. Appl.*, 32(4):1451–1474, 2011.

[5] V. de Silva and L.-H. Lim. Tensor rank and the ill-posedness of the best low-rank approximation problem. *SIAM Journal on Matrix Analysis and Applications*, 30(3):1084–1127, 2008.

[6] I. Domanov and L. Lathauwer. On Uniqueness and Computation of the Decomposition of a Tensor into Multilinear Rank-$(1, L_r, L_r)$ Terms. *SIAM J. Matrix Anal. Appl.*, 41(2):747–803, 2020.

[7] A. Edelman and E. Kostlan. How many zeros of a random polynomial are real? *Math. Soc. Mathematical Reviews*, 32:1–37, 05 1995.

[8] S. Friedland and G. Ottaviani. The number of singular vector tuples and uniqueness of best rank-one approximation of tensors. *Foundations of Computational Mathematics*, 14(6):1209–1242, 2014.

[9] R. Howard. The kinematic formula in Riemannian homogeneous spaces. *Mem. Amer. Math. Soc.*, 106(509):vi+69, 1993.

[10] B. Hunyadi, D. Camps, L. Sorber, W. V. Paesschen, M. De Vos, S. V. Huffel, and L. D. Lathauwer. Block term decomposition for modelling epileptic seizures. *EURASIP J. Adv. Signal Process.*, (1):139, 2014.

[11] L. D. Lathauwer. Decompositions of a higher-order tensor in block terms - Part II: Definitions and uniqueness. *SIAM J. Matrix Anal. Appl.*, 30:1033–1066, 2008.

[12] M. Sorensen and L. De Lathauwer. Coupled Canonical Polyadic Decompositions and (Coupled) Decompositions in Multilinear Rank-$(L_r, n, L_r, n, 1)$ Terms—Part I: Uniqueness. *SIAM J. Matrix Anal. Appl.*, 36(2):496–522, 2015.

[13] M. Yang. On partial and generic uniqueness of block term tensor decompositions. *Annali dell'universita' di Ferrara*, 60(2):465–493, 2014.