Tensors

Notes by Mateusz Michalek and Bernd Sturmfels
for the lecture on June 5, 2018, in the
IMPRS Ringvorlesung *Introduction to Nonlinear Algebra*

This lecture is divided into two parts. The first part, presented by Bernd Sturmfels, gives an introduction to the spectral theory of tensors. This will (definitely) be delivered in German. The second part, presented by Mateusz Michalek, introduces the tensor notions of rank, border rank, real rank and real border rank. This will (definitely) not be delivered in Polish.

1 Eigenvectors of Tensors

Let us begin by reviewing some basics of linear algebra, beginning with the study of symmetric matrices. Symmetric $n \times n$ matrices are important in statistics where they encode the covariance structure of a joint distribution of n random variables. In an algebraic setting, symmetric matrices are important because they uniquely represent quadratic forms.

For instance, consider the following quadratic form in three variables x, y and z:

$$Q \;=\; 2x^2 + 7y^2 + 23z^2 + 6xy + 10xz + 22yz. \tag{1}$$

This quadratic form is represented uniquely by a symmetric 3×3-matrix, as follows:

$$Q \;=\; \begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} 2 & 3 & 5 \\ 3 & 7 & 11 \\ 5 & 11 & 23 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \tag{2}$$

The gradient of the quadratic form Q is a vector of linear forms. It defines a linear map from $\mathbb{R}^3$ to itself. Up to multiplication by 2, this is the map one associates with a square matrix:

$$\nabla Q \;=\; \begin{pmatrix} \partial Q/\partial x \\ \partial Q/\partial y \\ \partial Q/\partial z \end{pmatrix} \;=\; 2 \cdot \begin{pmatrix} 2 & 3 & 5 \\ 3 & 7 & 11 \\ 5 & 11 & 23 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

We call $\mathbf{v} \in \mathbb{R}^n$ an *eigenvector* of Q if $\mathbf{v}$ is mapped to a scalar multiple of $\mathbf{v}$ by the gradient:

$$(\nabla Q)(\mathbf{v}) \;=\; \lambda \cdot \mathbf{v} \quad \text{for some } \lambda \in \mathbb{R}.$$

Geometers often replace $\mathbb{R}^n$ with the projective space $\mathbb{P}^{n-1}$. Two nonzero vectors are identified if they are parallel. From Q we obtain an induced self-map on projective space:

$$\nabla Q \,:\, \mathbb{P}^{n-1} \dashrightarrow \mathbb{P}^{n-1}. \tag{3}$$

We conclude our discussion with the following remark concerning the rational map in (3).

Remark 1. *The eigenvectors of Q are the fixed points $\mathbf{v}$ in $\mathbb{P}^{n-1}$ of its gradient map ∇Q.*

A real $n \times n$-matrix usually has n independent *eigenvectors*, over the complex numbers. When the matrix is symmetric, its eigenvectors have real coordinates and are *orthogonal*. For a rectangular matrix, one considers pairs of *singular vectors*, one on the left and one on the right. The number of these pairs is equal to the smaller of the two matrix dimensions.

Eigenvectors and singular vectors are familiar from linear algebra, where they are taught in concert with *eigenvalues* and *singular values*. Linear algebra is the foundation of applied mathematics and scientific computing. Specifically, the concept of eigenvectors and numerical algorithms for computing them, became a key technology during the 20th century.

Singular vectors are associated to rectangular matrices. We review their definition through the lens of Remark 1. Each rectangular matrix represents a bilinear form, e.g.

$$B \;=\; 2ux + 3uy + 5uz + 3vx + 7vy + 11vz \;=\; \begin{pmatrix} u & v \end{pmatrix} \begin{pmatrix} 2 & 3 & 5 \\ 3 & 7 & 11 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{4}$$

The gradient of the bilinear form defines an endomorphism of the direct sum of the row space and the column space. This fuses left multiplication and right multiplication by our matrix into a single map. In the example, the gradient is the following vector of linear forms

$$\nabla B = \Big(\big(\frac{\partial B}{\partial u}, \frac{\partial B}{\partial v}\big)\,,\, \big(\frac{\partial B}{\partial x}, \frac{\partial B}{\partial y}, \frac{\partial B}{\partial z}\big) \Big). \tag{5}$$

The associated linear map $\nabla B \,:\, \mathbb{R}^3 \oplus \mathbb{R}^2 \rightarrow \mathbb{R}^2 \oplus \mathbb{R}^3$ takes $\big((x,y,z),(u,v)\big)$ to this vector.

More generally, let B be an $m \times n$-matrix over $\mathbb{R}$. We are interested in the equations

$$B\mathbf{x} = \lambda \mathbf{y} \quad \text{und} \quad B^t \mathbf{y} = \lambda \mathbf{x}, \tag{6}$$

where λ is a scalar, $\mathbf{x}$ is a vector in $\mathbb{R}^n$, and $\mathbf{y}$ is a vector in $\mathbb{R}^m$. These are our unknowns. Given a solution to (6), $\mathbf{x}$ is an eigenvector of B^tB, $\mathbf{y}$ is an eigenvector of BB^t, and λ^2 is a common eigenvalue of these two symmetric matrices. Its square root $\lambda \geq 0$ is a *singular value* of B. Associated to λ are the *right singular vector* $\mathbf{x}$ and the *left singular vector* $\mathbf{y}$. In analogy to Remark 1, the process of solving (6) has the following dynamical interpretation:

Remark 2. *The singular vector pairs $(\mathbf{x}, \mathbf{y})$ of a rectangular matrix are the fixed points of the gradient map, taken on a product of projective spaces, of the associated bilinear form:*

$$\nabla B \,:\, \mathbb{P}^{n-1} \times \mathbb{P}^{m-1} \longrightarrow \mathbb{P}^{m-1} \times \mathbb{P}^{n-1}$$

$$(\mathbf{x}, \mathbf{y}) \;\mapsto\; \Big(\big(\frac{\partial B}{\partial x_1}, \ldots, \frac{\partial B}{\partial x_n}\big)\,,\, \big(\frac{\partial B}{\partial y_1}, \ldots \frac{\partial B}{\partial y_m}\big) \Big).$$

We summarize our review of some linear algebra concepts in the following bullet points:

- Symmetric matrices Q are important because they represent quadratic forms.
- Rectangular matrices B are important because they represent bilinear forms.

- Their gradients define the linear maps one usually identifies with Q and B.
- Fixed points of these maps are called *eigenvectors* and *singular vectors*.
- These fixed points are computed via orthogonal decompositions of our matrices:

$$Q \;=\; O \cdot \mathrm{diag} \cdot O^t \quad \text{and} \quad B \;=\; O_1 \cdot \mathrm{diag} \cdot O_2.$$

  These are known as the *spectral decomposition* and the *singular value decomposition*.

In the age of Big Data, the role of matrices is increasingly played by *tensors*, that is, multidimensional arrays of numbers. Principal component analysis tells us that eigenvectors of covariance matrices $Q = BB^t$ point to directions in which the data B is most spread. One hopes to identify similar features in higher-dimensional data. This has encouraged engineers and scientists to spice up their linear algebra tool box with a pinch of algebraic geometry.

The spectral theory of tensors is the theme of the following discussion. This theory was pioneered around 2005 by Lek-Heng Lim and Liqun Qi. Our aim is to generalize familiar notions, such as rank, eigenvectors and singular vectors, from matrices to tensors. Specifically, we address the following questions. The answers are provided in Examples 7 and 12.

Question 3. *How many eigenvectors does a* $3 \times 3 \times 3$*-tensor have?*

Question 4. *How many triples of singular vectors does a* $3{\times}3{\times}3$*-tensor have?*

A *tensor* is a d-dimensional array $T = (t_{i_1 i_2 \cdots i_d})$. Tensors of format $n_1{\times}n_2{\times}\cdots{\times}n_d$ form a space of dimension $n_1 n_2 \cdots n_d$. For $d = 1, 2$ we get vectors and matrices. A tensor has *rank 1* if it is the outer product of d vectors, written $T = \mathbf{u} \otimes \mathbf{v} \otimes \cdots \otimes \mathbf{w}$, or, in coordinates,

$$t_{i_1 i_2 \cdots i_d} \;=\; u_{i_1} v_{i_2} \cdots w_{i_d}.$$

The problem of *tensor decomposition* concerns expressing T as a sum of rank 1 tensors, using as few summands as possible. That minimal number of summands needed is the *rank* of T.

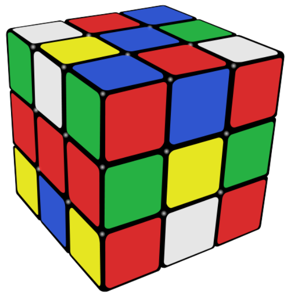
An $n{\times}n{\times}\cdots{\times}n$-tensor $T = (t_{i_1 i_2 \cdots i_d})$ is *symmetric* if it is unchanged under permuting the indices. The space $\mathrm{Sym}_d(\mathbb{R}^n)$ of such symmetric tensors has dimension $\binom{n+d-1}{d}$. It is identified with the space of homogeneous polynomials of degree d in n variables, written as

$$T \;=\; \sum_{i_1,\ldots,i_d=1}^{n} t_{i_1 i_2 \cdots i_d} \cdot x_{i_1} x_{i_2} \cdots x_{i_d}.$$

Example 5. A tensor T of format $3{\times}3{\times}3$ has 27 entries. If T is symmetric then it has ten distinct entries, one for each coefficient of the associated cubic polynomial in three variables. This polynomial defines a cubic curve in the projective plane $\mathbb{P}^2$, as indicated in Figure 1.

Symmetric tensor decomposition writes a polynomial as a sum of powers of linear forms:

$$T \;=\; \sum_{j=1}^{r} \lambda_j \mathbf{v}_j^{\otimes d} \;=\; \sum_{j=1}^{r} \lambda_j (v_{1j} x_1 + v_{2j} x_2 + \cdots + v_{nj} x_n)^d. \tag{7}$$



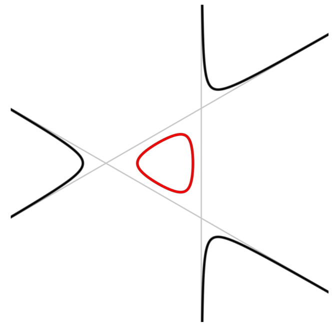


Figure 1: A symmetric 3×3×3 tensor represents a cubic curve in the projective plane.

The *gradient of* T defines a map $\nabla T : \mathbb{R}^n \to \mathbb{R}^n$. A vector $\mathbf{v} \in \mathbb{R}^n$ is an *eigenvector* of T if

$$(\nabla T)(\mathbf{v}) \;=\; \lambda \cdot \mathbf{v} \qquad \text{for some } \lambda \in \mathbb{R}.$$

Eigenvectors of tensors arise naturally in optimization. Consider the problem of maximizing a polynomial function T over the unit sphere in $\mathbb{R}^n$. If λ denotes a Lagrange multiplier, then one sees that the eigenvectors of T are the critical points of this optimization problem.

Algebraic geometers find it convenient to replace the unit sphere in $\mathbb{R}^n$ by the projective space $\mathbb{P}^{n-1}$. The gradient map is then a rational map from this projective space to itself:

$$\nabla T \,:\, \mathbb{P}^{n-1} \dashrightarrow \mathbb{P}^{n-1}.$$

The eigenvectors of T are *fixed points* ($\lambda \neq 0$) and *base points* ($\lambda = 0$) of ∇T. Thus the spectral theory of tensors is closely related to the study of dynamical systems on $\mathbb{P}^{n-1}$.

In the matrix case ($d = 2$), the linear map ∇T is the gradient of the quadratic form

$$T \;=\; \sum_{i=1}^{n} \sum_{j=1}^{n} t_{ij} x_i x_j.$$

By the Spectral Theorem, T has a real decomposition (7) with $d = 2$. Here r is the rank, the λ_j are the eigenvalues of T, and the eigenvectors $\mathbf{v}_j = (v_{1j}, v_{2j}, \ldots, v_{nj})$ are orthonormal. We can compute this by *power iteration*, namely, by applying ∇T until a fixed point is reached.

For $d \geq 3$, one can still use the power iteration to compute eigenvectors of T. However, the eigenvectors are usually not the vectors $\mathbf{v}_i$ in the low rank decomposition (7). One exception arises when the symmetric tensor is *odeco*, or orthogonally decomposable. This means that T has the form (7), where $r = n$ and $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ is an orthogonal basis of $\mathbb{R}^n$. These basis vectors are the attractors of the dynamical system ∇T, provided $\lambda_j > 0$.

Theorem 6. *The number of complex eigenvectors of a general tensor* $T \in \mathrm{Sym}_d(\mathbb{R}^n)$ *is*

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

Example 7. Let $n = d = 3$. The Fermat cubic $T = x^3 + y^3 + z^3$ is an odeco tensor. Its gradient map squares each coordinate: $\nabla T : \mathbb{P}^2 \dashrightarrow \mathbb{P}^2$, $(x : y : z) \mapsto (x^2 : y^2 : z^2)$. This dynamical system has seven fixed points, of which only the first three are attractors:

$$(1:0:0), (0:1:0), (0:0:1),\ (1:1:0), (1:0:1), (0:1:1),\ (1:1:1).$$

We conclude that T has 7 eigenvectors, and the same holds for $3 \times 3 \times 3$-tensors in general.

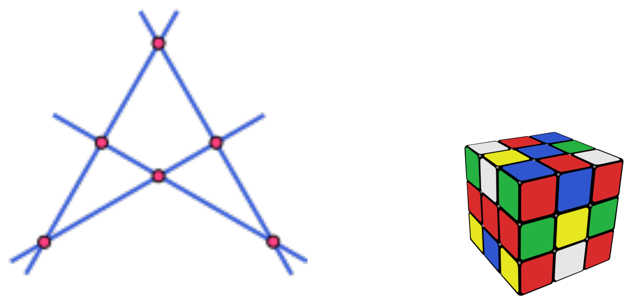


Figure 2: The polynomial $T = xyz(x + y + z)$ represents a symmetric $3{\times}3{\times}3{\times}3$ tensor.

It is known that all eigenvectors can be real for suitable tensors. This was proved in 2017 by Khazhgali Khozhasov, using the theory of *harmonic polynomials*. For $n = 3$, this can be seen by the following simple argument, found earlier by Abo, Seigal and Sturmfels. Let T be a product of linear forms in three unknowns, defining d lines in $\mathbb{P}^2$, then the $\binom{d}{2}$ vertices of the line arrangement are base points of ∇T, and each of the $\binom{d}{2} + 1$ regions contain one fixed point. This accounts for all $1 + (d-1) + (d-1)^2$ eigenvectors, which are therefore real.

Example 8. Let $d = 4$ and fix the product of linear forms $T = xyz(x + y + z)$. Its curve in $\mathbb{P}^2$ is an arrangement of four lines, as shown in Figure 2. This quartic represents a symmetric $3{\times}3{\times}3{\times}3$ tensor. All $13 = 6 + 7$ eigenvectors of this tensor are real. The 6 vertices of the arrangement are the base points of ∇T. Each of the 7 regions contains one fixed point.

For special tensors T, two of the eigenvectors in Theorem 6 may coincide. This corresponds to vanishing of the *eigendiscriminant*, which is a big polynomial in the $t_{i_1 i_2 \cdots i_d}$. In the matrix case ($d = 2$), it is the discriminant of the characteristic polynomial of an $n{\times}n$-matrix. For $3{\times}3{\times}3$ tensors, the eigendiscriminant has degree 24. In general we have the following:

Theorem 9 (Abo-Seigal-Sturmfels)**.** *The eigendiscriminant is an irreducible homogeneous polynomial of degree* $n(n-1)(d-1)^{n-1}$ *in the coefficients* $t_{i_1 i_2 \cdots i_d}$ *of the tensor* T.

Singular value decomposition is a central notion in linear algebra and its applications. Consider a rectangular matrix $T = (t_{ij})$ of format $n_1 \times n_2$. The *singular values* σ of T satisfy

$$T\mathbf{u} = \sigma\mathbf{v} \quad \text{and} \quad T^t\mathbf{v} = \sigma\mathbf{u},$$

where **u** and **v** are the corresponding *singular vectors*. Just like with eigenvectors, we can associate to this a dynamical system. Namely, we interpret the matrix as a bilinear form

$$T \;=\; \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} t_{ij} x_i y_j .$$

The gradient of T defines a rational self-map of a product of two projective spaces:

$$\begin{array}{rccc} \nabla T: & \mathbb{P}^{n_1-1}\times\mathbb{P}^{n_2-1} & \dashrightarrow & \mathbb{P}^{n_1-1}\times\mathbb{P}^{n_2-1} \\ & (\mathbf{u}\,,\,\mathbf{v}) & \mapsto & (T^t\mathbf{v}\,,\,T\mathbf{u}\,) \end{array}$$

The *fixed points* of this map are the pairs of singular vectors of T.

Consider now an arbitrary d-dimensional tensor T in $\mathbb{R}^{n_1\times n_2\times\cdots\times n_d}$. It corresponds to a multilinear form. The *singular vector tuples* of T are the fixed points of the gradient map

$$\nabla T \;:\; \mathbb{P}^{n_1-1}\times\cdots\times\mathbb{P}^{n_d-1} \dashrightarrow \mathbb{P}^{n_1-1}\times\cdots\times\mathbb{P}^{n_d-1}.$$

Example 10. The trilinear form $T = x_1y_1z_1 + x_2y_2z_2$ gives a $2{\times}2{\times}2$ tensor. Its map ∇T is

$$\begin{array}{ccc} \mathbb{P}^1\times\mathbb{P}^1\times\mathbb{P}^1 & \rightarrow & \mathbb{P}^1\times\mathbb{P}^1\times\mathbb{P}^1, \\ \big((x_1:x_2),(y_1:y_2),(z_1:z_2)\big) & \mapsto & \big((y_1z_1:y_2z_2),(x_1z_1:x_2z_2),(x_1y_1:x_2y_2)\big). \end{array}$$

This map has no base points, but it has six fixed points, namely $\big((1{:}0),(1{:}0),(1{:}0)\big)$, $\big((0{:}1),(0{:}1),(0{:}1)\big)$, $\big((1{:}1),(1{:}1),(1{:}1)\big)$, $\big((1{:}1),(1{:}-1),(1{:}-1)\big)$, $\big((1{:}-1),(1{:}1),(1{:}-1)\big)$, and $\big((1{:}-1),(1{:}-1),(1{:}1)\big)$. These are the triples of singular vectors of the given $2\times2\times2$ tensor.

Here is an explicit formula for the expected number of singular vector tuples.

Theorem 11 (Friedland and Ottaviani)**.** *For a general $n_1{\times}n_2{\times}\cdots{\times}n_d$-tensor T, the number of singular vector tuples (over $\mathbb{C}$) is the coefficient of $z_1^{n_1-1}\cdots z_d^{n_d-1}$ in the polynomial*

$$\prod_{i=1}^{d} \frac{(\widehat{z_i})^{n_i} - z_i^{n_i}}{\widehat{z_i} - z_i} \quad \textit{where} \quad \widehat{z_i} = z_1+\cdots+z_{i-1}+z_{i+1}+\cdots+z_d.$$

We conclude our excursion into the spectral theory of tensors by answering Question 2.

Example 12. Let $d=3$ and $n_1{=}n_2{=}n_3{=}3$. The generating function in Theorem 11 equals

$$(\widehat{z_1}^2{+}\widehat{z_1}z_1{+}z_1^2)(\widehat{z_2}^2{+}\widehat{z_2}z_2{+}z_2^2)(\widehat{z_3}^2{+}\widehat{z_3}z_3{+}z_3^2) \;=\; \cdots + \mathbf{37}z_1^2z_2^2z_3^2 + \cdots$$

This means that a general $3{\times}3{\times}3$-tensor has exactly 37 triples of singular vectors. Likewise, a general $3{\times}3{\times}3{\times}3$-tensor, as illustrated in Figure 2, has 997 quadruples of singular vectors.

2 The Many Ranks of a Tensor

There are several ways to define the rank of a matrix $M \in K^a \times K^b$. It is:

1. the smallest integer r such that all $(r+1) \times (r+1)$ minors vanish,
2. the dimension of the image of the induced linear map $K^a \to K^b$,
3. the dimension of the image of the induced linear map $K^b \to K^a$,
4. the smallest integer r, such that there exist vectors $v_1, \dots, v_r \in K^a$, $w_1, \dots, w_r \in K^b$ for which:
$$M_{ij} = \sum_{k=1}^{r} (v_k)_i (w_k)_j.$$

The first point implies that matrices of rank at most r form a variety. The last point implies that a matrix of rank r is a sum of r matrices of rank one. This is also true for symmetric matrices: a symmetric matrix of rank r is a sum of r symmetric rank one matrices. Another fact is that a real matrix of rank r has also rank r, when regarded as a complex matrix. This seems obvious, but a priori, it is not clear why there is no shorter complex decomposition into rank one matrices. Our aim is to find analogous statements for arbitrary tensors.

From the first part of the lecture we recall the definition of rank one tensor: it is the outer product of d vectors, written $T = \mathbf{u} \otimes \mathbf{v} \otimes \cdots \otimes \mathbf{w}$, i.e.

$$t_{i_1 i_2 \cdots i_d} \;=\; u_{i_1} v_{i_2} \cdots w_{i_d}.$$

Tensors of rank (at most) one form an algebraic variety. It is the affine cone over the Segre product $\mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_d}$. In fact, from Lecture 2 and Lecture 7 we know the equations of this variety! These are binomial quadrics that can be identified with 2×2 minors. In other words, a tensor $T \in V_1 \otimes \cdots \otimes V_d \simeq K^{n_1 \times \cdots \times n_d}$ has rank one if and only if all the induced linear maps/matrices, known as flattenings:

$$K^{\prod_{i \in I} d_i} = \bigotimes_{i \in I} V_i^* \to \bigotimes_{i \in [n] \setminus I} V_i = K^{\prod_{i \in [n] \setminus I} d_i}$$

have rank one, for any subset $I \subset [n]$.

Example 13. *A tensor $T = (t_{ijk}) \in V_1 \otimes V_2 \otimes V_3$ induces a linear map:*

$$V_1^* \to V_2 \otimes V_3,$$

given by:

$$e_i^* \to (t_{ijk})_{j,k} = \sum_{j,k} t_{ijk} f_j \otimes g_k,$$

where $(e_i), (f_j), (g_k)$ are respectively bases of V_1, V_2, V_3.

The conclusion is that rank one tensors behave in a very nice way. What is surprising, arbitrary tensors exhibit very strange properties. Recall that the rank of a tensor T is the minimal r such that T is the sum of rank one tensors.

Definition 14. *The following tensor is known in quantum physics as a W-state:*

$$W = e_0 \otimes e_0 \otimes e_1 + e_0 \otimes e_1 \otimes e_0 + e_1 \otimes e_0 \otimes e_0.$$

It plays an important role in quantum information theory. As we will see below it may be defined as a tangent vector to the Segre product $\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1$.

Clearly W has rank at most three. In fact, $\operatorname{rk} W = 3$, as the reader is asked to prove in Exercise 9. However, there exist rank two tensors arbitrary near W! For any $\epsilon \neq 0$ we have:

$$\frac{1}{\epsilon}((e_0 + \epsilon e_1) \otimes (e_0 + \epsilon e_1) \otimes (e_0 + \epsilon e_1) - e_0 \otimes e_0 \otimes e_0) =$$

$$W + \epsilon(e_1 \otimes e_1 \otimes e_0 + e_1 \otimes e_0 \otimes e_1 + e_0 \otimes e_1 \otimes e_1) + \epsilon^2 e_1 \otimes e_1 \otimes e_1.$$

In particular,

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon}((e_0 + \epsilon e_1) \otimes (e_0 + \epsilon e_1) \otimes (e_0 + \epsilon e_1) - e_0 \otimes e_0 \otimes e_0) = W,$$

i.e. $\operatorname{rk} W = 3$, but W can be approximated with arbitrary precision by rank two tensors.

Definition 15. *The* border rank $\operatorname{br} T$ *of the tensor T is the smallest r such that there exist tensors of rank r in any neighbourhood of T.*

We note that the notion of border rank requires a topology on the space of tensors. The geometric locus of tensors of border rank at most r is the closure of the locus of tensors of rank at most r. Over complex numbers, by Chevalley's theorem from Lecture 2, it does not matter if we take Zariski or Euclidean topology: the closures coincide. However, over the real numbers the situation is different.

Example 16. *Consider $\mathbb{R}^2 \otimes \mathbb{R}^2 \otimes \mathbb{R}^2$. In Exercise 10 the reader is asked to prove that the Zariski closure of tensors of rank two is the whole space. However, the Euclidean closure of the locus of rank two tensors is a proper semialgebraic subset. Passing to the projective setting $X = \mathbb{P}^1_{\mathbb{R}} \otimes \mathbb{P}^1_{\mathbb{R}} \otimes \mathbb{P}^1_{\mathbb{R}} \subset \mathbb{P}^7_{\mathbb{R}}$, the union of the tangent spaces to X, known as the tangential variety, is a hypersurface in $\mathbb{P}^7_{\mathbb{R}}$. The sign of the defining equation of the tangential variety determines if the tensor has (real) rank two or three.*

Explicitly, given a tensor T we obtain the associated map:

$$\mathbb{R}^2 \to \mathbb{R}^2 \otimes \mathbb{R}^2.$$

For a general tensor T, the image of this map is a two dimensional linear space S of 2×2 matrices. If T has rank two, i.e. $T = u_1 \otimes v_1 \otimes w_1 + u_2 \otimes v_2 \otimes w_2$ then S must contain two rank one matrices: $v_1 \otimes w_1$ and $v_2 \otimes w_2$. Hence, we ask if S intersects the locus of rank one *matrices in (at least) two points. In projective setting, rank one matrices are defined by the quadric, i.e. the determinant, and coincide with the Segre surface that is the image:*

$$\mathbb{P}^1 \times \mathbb{P}^1 \to \mathbb{P}^3.$$

The line $\mathbb{P}(S)$ must intersect this surface over the field of complex numbers, however, does not have to over the field of real numbers.

Consider the tensor:

$$T := e_1 \otimes f_1 \otimes g_1 - e_1 \otimes f_2 \otimes g_2 - e_2 \otimes f_1 \otimes g_2 - e_2 \otimes f_2 \otimes g_1.$$

The space S is:

$$\begin{pmatrix} a & b \\ b & -a \end{pmatrix},$$

and clearly does not contain real rank one matrices. The argument remains correct in a (Euclidean) neighbourhood of T. On the other hand we obtain two complex rank one matrices, which give rise to the decomposition:

$$T = \frac{1}{2}((e_1 + ie_2)^{\otimes 3} + (e_1 - ie_2)^{\otimes 3}).$$

To conclude, contrary to the case of matrices or rank one tensors:

- tensors of rank at most r may not form a closed set,
- a real tensor may have different (smaller) rank, when regarded as a complex tensor,
- real tensors of bounded real border rank form semialgebraic sets.

We have described rank one tensors as the Segre product of projective spaces. It is natural to ask for a geometric description of tensors of rank at most r.

Definition 17 (Secant Variety)**.** *Let X be a projective (resp. affine) algebraic variety. For a set S let $\langle S \rangle$ be the smallest projective (resp. affine) subspace containing S. The k-th* secant variety *of X is the closure of all k-secant planes:*

$$\sigma_k(X) := \overline{\bigcup_{p_1,\ldots,p_k \in X} \langle p_1, \ldots, p_k \rangle}.$$

In particular,

$$X = \sigma_1(X) \subset \sigma_2(X) \subset \cdots \subset \sigma_{\dim\langle X \rangle}(X) = \langle X \rangle.$$

In fact, the containments must be strict, until $\sigma_r = \langle X \rangle$. If X is the Segre product, then $\bigcup_{p_1,\ldots,p_k \in X} \langle p_1, \ldots, p_k \rangle$ is the locus of tensors of rank at most r. Hence, $\sigma_r(X)$ coincides with the locus of tensors of border rank at most r. It is a major open problem to describe the ideal of $\sigma_r(\mathbb{P}^{a_1} \times \cdots \times \mathbb{P}^{a_n})$, as this would provide an algebraic test for a tensor to be of border rank r. Let us describe the simplest equations. A tensor $T \in V_1 \otimes \dots V_n$ and a subset $I \subset [n]$, induce the flattening map:

$$\bigotimes_{i \in I} V_i^* \to \bigotimes_{i \in [n] \setminus I} V_i.$$

The rank of the flattening map is at most $\operatorname{rk} T$. It follows that $r+1$ minors of the flattening matrix provide (some) equations of $\sigma_r(\mathbb{P}(V_1) \times \cdots \times \mathbb{P}(V_n))$. These are of degree $r+1$ and in fact there are no polynomials in this ideal of strictly smaller degree.

An analogous notion of *symmetric rank* or *Waring rank* can be defined for symmetric tensors. A symmetric tensor T has symmetric/Waring rank one if the following equivalent conditions hold:

1. $\operatorname{rk} T = 1$,
2. $T = v \otimes \cdots \otimes v$ for some vector v,
3. T represented as a polynomial is a power of a linear form.

The symmetric/Waring rank of a symmetric tensor T is the smallest r such that T is a linear combination of r rank one symmetric tensors.

Remark 18. *We do not write that a tensor is a sum of symmetric tensors, as over real numbers, this may be not possible. For example, when T is represented by an even degree polynomial, then sum of even powers of real linear forms always is a polynomial that is nonnegative.*

We also have a concept of symmetric/Waring border rank of a symmetric tensor T; it is the smallest integer r such that T can be approximated by symmetric tensors of symmetric/Waring rank r. Our previous discussion shows that W-state has symmetric rank three and symmetric border rank two. Clearly, for any symmetric tensor its symmetric rank (resp. symmetric border rank) is at least equal to its rank (resp. border rank).

Conjecture 19 (Comon's Conjecture)**.** *For any symmetric tensor its symmetric rank equals its symmetric border rank.*

The conjecture was confirmed in many specific cases, however recently a counterexample was presented by Yaroslav Shitov [2]. Unfortunately, the example is far to complicated to be presented during the lecture. The border rank analogue of Comon's conjecture remains open.

Just as rank one tensors correspond to Segre products $\mathbb{P}^{a_1} \times \cdots \times \mathbb{P}^{a_n}$, symmetric rank one tensors correspond to Veronese reembeddings $v_d(\mathbb{P}^n) \subset \mathbb{P}^{\binom{d+n}{n}-1}$. Here, we interpret $\mathbb{P}^n$ as a space of linear forms in $n+1$ variables and $\mathbb{P}^{\binom{d+n}{n}-1}$ as the space of degree d homogeneous polynomials. The map v_d sends a linear form l to its d-th power l^d. This is the same Veronese map as discussed in Lecture 7, up to scaling the coordinates. Further, the locus of symmetric tensors of Waring rank at most r is precisely $\sigma_r(v_d(\mathbb{P}^n))$.

Remark 20. *For each integer k there exists a minimal number r, such that for any n there exist positive integers $a_1, \ldots, a_r$ such that*

$$n = \sum_{i=1}^{r} a_i^k.$$

Waring's original problem is, to determine r as a function of k.

The problem for polynomials that we are facing is to represent a homogeneous polynomial of degree d as a linear combination of powers of linear forms. Thus, by analogy, the minimal number of linear forms that is needed is called the Waring rank. By seminal work of Alexander and Hirschowitz we know Waring ranks of general polynomials (of any degree in any number of variables). In other words we know the maximal border rank that a homogeneous polynomial may have. For usual (nonsymmetric) tensors, the problem of determining maximal border rank (or rank) in general remains open.

Although, it is easy to prove that general tensors have high rank and border rank it is extremely hard to find explicit examples. Here, we do not want to dive into precise definition of 'explicit', let us just say one seeks a tensor with not too big integer entries. In particular, it is not know how to provide examples of tensors $T \in \mathbb{C}^n \otimes \mathbb{C}^n \otimes \mathbb{C}^n$ of either:

- rank greater than $3n$,
- border rank greater than $2n$.

Still, by Exercise 12, a general tensor in this space has border rank quadratic in n.

Exercises

1. Fix the quadratic form Q in (1). Compute the maxima and minima of Q on the unit 2-sphere. Find all fixed points of the map $\nabla Q : \mathbb{P}^2 \to \mathbb{P}^2$. How are they related?
2. Compute all fixed points of the map $\nabla B : \mathbb{P}^2 \times \mathbb{P}^1 \to \mathbb{P}^2 \times \mathbb{P}^1$ given by B in (4).
3. Consider the $3 \times 3 \times 2 \times 2$ tensor defined by the multilinear form $T = x_1y_1z_1w_1 + x_2y_2z_2w_2$? Determine all quadruples of singular vectors of T.
4. For $d = 2, 3, 4$, pick random symmetric tensors of formats $d \times d \times d$ and $d \times d \times d \times d$. Compute all eigenvectors of your tensors.
5. Prove Theorem 6.
6. Write down an explicit $3 \times 3 \times 3 \times 3$ tensor with precisely 13 <u>real</u> eigenvectors.
7. What is the number of singular vector tuples of your tensors in Problem 4?

8. Compute the eigendiscriminants for tensors of format 2×2 and $2 \times 2 \times 2$ and $2 \times 2 \times 2 \times 2$. Write them explicitly as homogeneous polynomials in these entries of an unknown tensor.

9. By showing that a particular system of polynomial equations has no solutions, prove that $\operatorname{rk} W = 3$.

10. Prove that the Zariski closure of tensors of rank two in $\mathbb{R}^2 \otimes \mathbb{R}^2 \otimes \mathbb{R}^2$ is the whole space (e.g. by computing the dimension of the locus of such tensors).

11. Find the equation of the tangential variety to $\mathbb{P}^1 \otimes \mathbb{P}^1 \otimes \mathbb{P}^1 \subset \mathbb{P}^7$.

12. Prove that in $\mathbb{C}^n \otimes \mathbb{C}^n \otimes \mathbb{C}^n$:

    (a) there exists a tensor of border rank at least $\frac{1}{3}n^2$,

    (b) every tensor has rank at most n^2.

References

[1] J.M. Landsberg: *Tensors: Geometry and Applications*, Graduate Studies in Mathematics, 128, American Mathematical Society, Providence, RI, 2012.

[2] Y. Shitov: *A counterexample to Comon's conjecture*, arXiv preprint arXiv:1705.08740 (2017).

[3] L. Qi, H. Chen and Y. Chen: *Tensor Eigenvalues and their Applications*, Advances in Mechanics and Mathematics, Springer-Verlag, New York, 2018.