Tensors: a geometric view

Open lecture

November 24, 2014

Simons Institute, Berkeley

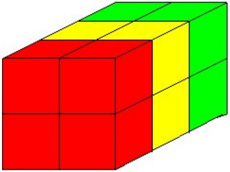
Giorgio Ottaviani, Università di Firenze

Plan of the talk

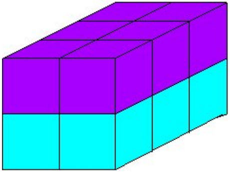
- Introduction to Tensors
- A few relevant applications

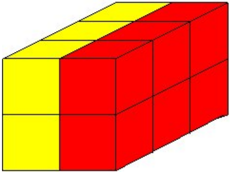
Tensors as multidimensional matrices

A (complex) $a \times b \times c$ tensor is an element of the space $\mathbb{C}^a \otimes \mathbb{C}^b \otimes \mathbb{C}^c$, it can be represented as a 3-dimensional matrix.

Here is a tensor of format $2 \times 2 \times 3$. 

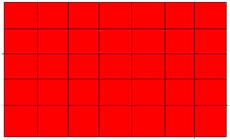
Slices of tensors

Several slices of a $2 \times 2 \times 3$ tensor. 



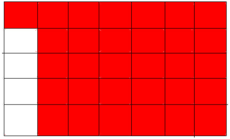
Gaussian elimination

Gaussian elimination consists in simplify a matrix, by adding to a row a multiple of another one. It corresponds to left multiplication by invertible matrices.



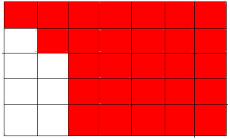
Gaussian elimination

Gaussian elimination consists in simplify a matrix, by adding to a row a multiple of another one. It corresponds to left multiplication by invertible matrices.

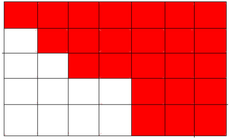


Gaussian elimination

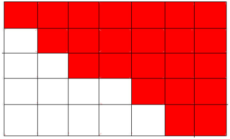
Gaussian elimination consists in simplify a matrix, by adding to a row a multiple of another one. It corresponds to left multiplication by invertible matrices.



Gaussian elimination consists in simplify a matrix, by adding to a row a multiple of another one. It corresponds to left multiplication by invertible matrices.

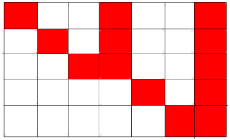


Gaussian elimination consists in simplify a matrix, by adding to a row a multiple of another one. It corresponds to left multiplication by invertible matrices.



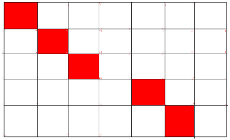
Gaussian elimination and canonical form

adding rows backwards....



Gaussian elimination and canonical form

adding rows backwards....

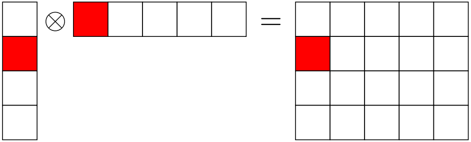


...adding columns we get a canonical form !

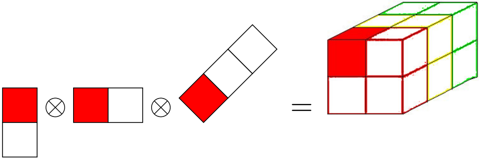
This matrix of rank 5 is the sum of five rank one (or "decomposable") matrices.

The decomposable (rank one) tensors

Here is a decomposable matrix

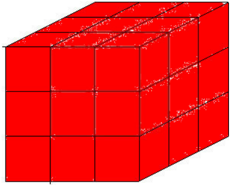


Here is a decomposable tensor



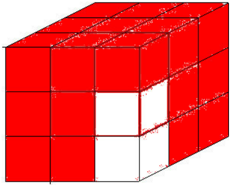
Trying Gaussian elimination on a 3-dimensional tensor

We can add a scalar multiple of a slice to another slice.

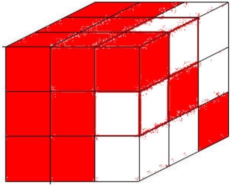


Trying Gaussian elimination on a 3-dimensional tensor

We can add a scalar multiple of a slice to another slice.

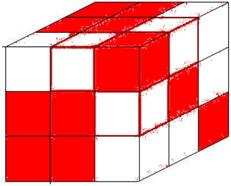


We can add a scalar multiple of a slice to another slice.



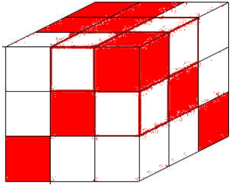
Trying Gaussian elimination on a 3-dimensional tensor

We can add a scalar multiple of a slice to another slice.



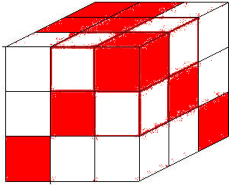
Trying Gaussian elimination on a 3-dimensional tensor

We can add a scalar multiple of a slice to another slice.



Trying Gaussian elimination on a 3-dimensional tensor

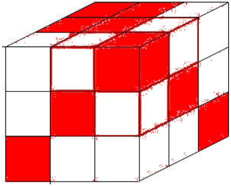
We can add a scalar multiple of a slice to another slice.



How many zeroes we may assume, at most ?

Trying Gaussian elimination on a 3-dimensional tensor

We can add a scalar multiple of a slice to another slice.

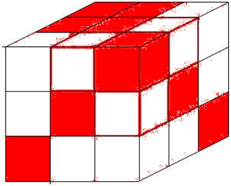


How many zeroes we may assume, at most ?

Strassen showed in 1983 one remains with at least 5 nonzero entries. Even, at least $5 (> 3)$ decomposable summands.

Trying Gaussian elimination on a 3-dimensional tensor

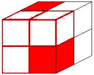
We can add a scalar multiple of a slice to another slice.

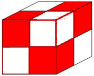


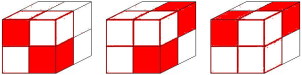
How many zeroes we may assume, at most ?
Strassen showed in 1983 one remains with at least 5 nonzero entries. Even, at least $5(>3)$ decomposable summands.

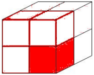
Read the masters !

The six canonical forms of a $2 \times 2 \times 2$ tensor

 general rank 2 .

 hyperdeterminant vanishes.

 support on one slice (only not symmetric)!

 rank 1.

$2 \times 2 \times 2$ is one of the few lucky cases, where such a classification is possible.

A dimensional count shows we cannot expect finitely many canonical forms.

- The dimension of $\mathbb{C}^n \otimes \mathbb{C}^n \otimes \mathbb{C}^n$ is n^3.
- The dimension of $GL(n) \times GL(n) \times GL(n)$ is $3n^2$, *too small for* $n \geq 3$.

The same argument works for general $\mathbb{C}^{n_1} \otimes \ldots \otimes \mathbb{C}^{n_d}$, $d \geq 3$, with a few small dimensional exceptions.

Matrices were created by God, Tensors by the Devil

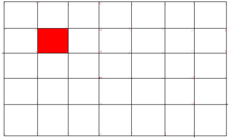
Paraphrasing Max Noether, we can joke that

- Matrices were created by God 

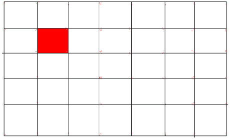
Paraphrasing Max Noether, we can joke that

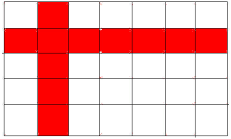
- Matrices were created by God 
- Tensors were created by the Devil 

Rank one matrix and its tangent space

This is a rank one matrix 

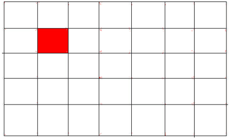
Rank one matrix and its tangent space

This is a rank one matrix 

this is its tangent space  by Leibniz formula.

Picture of tangent space

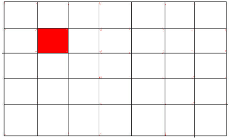
Denote by $\{a_1, \ldots, a_5\}$ the row basis, by $\{b_1, \ldots, b_7\}$ the column basis.



The picture represents $a_2 \otimes b_2 \in A \otimes B$
Consider a curve $a(t) \otimes b(t)$ such that $a(0) = a_2$, $b(0) = b_2$.

Picture of tangent space

Denote by $\{a_1, \ldots, a_5\}$ the row basis, by $\{b_1, \ldots, b_7\}$ the column basis.

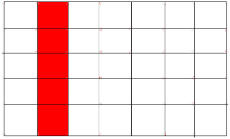


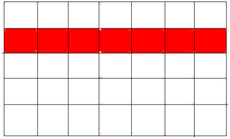
The picture represents $a_2 \otimes b_2 \in A \otimes B$

Consider a curve $a(t) \otimes b(t)$ such that $a(0) = a_2$, $b(0) = b_2$.
Taking the derivative at $t = 0$, get $a'(0) \otimes b(0) + a(0) \otimes b'(0)$, so all derivatives span

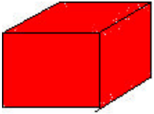
$$A \otimes b(0) + a(0) \otimes B$$

and the two summands corresponds to the following

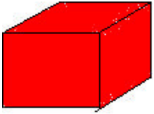


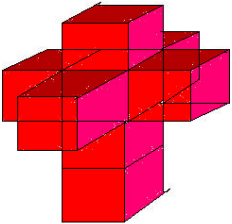


Picture of tangent space in three dimensions

This is a rank one 3-dimensional tensor 

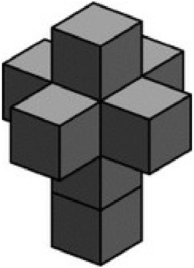
Picture of tangent space in three dimensions

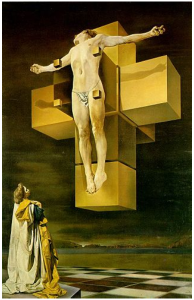
This is a rank one 3-dimensional tensor 

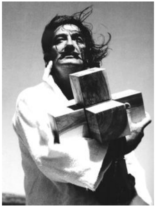
and this is its tangent space 

Salvador Dalì hypercube

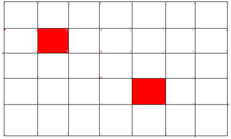
Here are better pictures



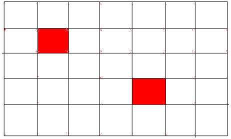


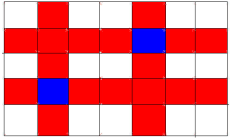


Rank two tangent space at a matrix

This is a rank 2 matrix. 

Rank two tangent space at a matrix

This is a rank 2 matrix. 

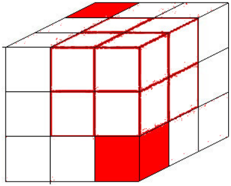
This is the tangent space to this matrix,  thanks

to Terracini Lemma. 

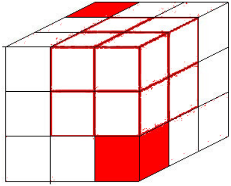
The two summands meet in the two blue boxes.

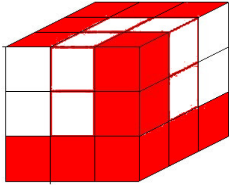
Rank two tangent space at a tensor

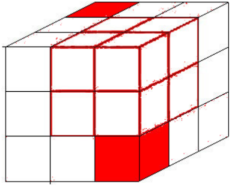
This is a rank 2 three-dimensional tensor.

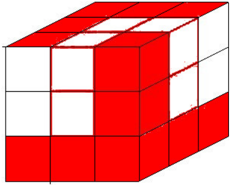


Rank two tangent space at a tensor

This is a rank 2 three-dimensional tensor. 

This is the tangent space to this tensor 

This is a rank 2 three-dimensional tensor. 

This is the tangent space to this tensor 

The two summands do not meet in three dimensions. There are no blue boxes. This has nice consequences for tensor decomposition,

as we will see in a while.  Maybe the Devil was sleeping...

Kronecker-Weierstrass

The general $2 \times b \times c$ tensors were classified in XIX century by Kronecker and Weierstrass. The classification has two steps

1. There are canonical forms for the two cases $2 \times n \times n$ $2 \times n \times (n+1)$
2. The latter are building blocks for all the others.

The case $2 \times n \times n$

These are pencils of square matrices. The general pencil, up to the action of $GL(n) \times GL(n)$, has the following two slices

$$\begin{array}{|c|c|c|c|}\hline 1 & & & \\ \hline & 1 & & \\ \hline & & 1 & \\ \hline & & & 1 \\ \hline \end{array} \qquad \begin{array}{|c|c|c|c|}\hline \lambda_1 & & & \\ \hline & \lambda_2 & & \\ \hline & & \lambda_3 & \\ \hline & & & \lambda_4 \\ \hline \end{array}$$

The case $2 \times n \times (n+1)$

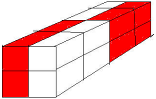
The general $2 \times n \times (n+1)$ pencil, up to the action of $GL(n) \times GL(n+1)$, has the following two slices

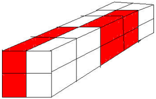
$$\begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & 1 & \end{bmatrix} \qquad \begin{bmatrix} & 1 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & 1 \end{bmatrix}$$

Theorem (Kronecker-Weierstrass)

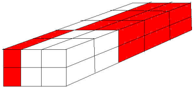
The general $2 \times b \times c$ tensor, with $b < c$, decomposes as a direct sum of several tensors of format $2 \times n \times (n+1)$ and $2 \times (n+1) \times (n+2)$, for some n.

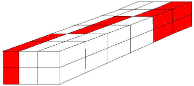
Let see some examples

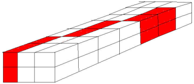
a $2 \times 2 \times 4$ tensor 

a $2 \times 2 \times n$ tensor for $n \geq 4$, padded with zeroes 

$2 \times b \times c$ decompositions

a $2 \times 3 \times 5$ tensor 

a $2 \times 3 \times 6$ tensor 

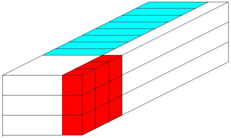
a $2 \times 3 \times n$ tensor for $n \geq 6$, padded with zeroes 

Fibonacci tensors

With three slices, Fibonacci tensors occurr.

Theorem (Kac, 1980)

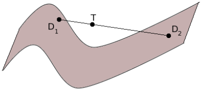
The general $3 \times b \times c$ tensor, with $c \geq \frac{3b+\sqrt{5b^2+1}}{2}$, decomposes as a direct sum of several tensors of format $3 \times F_n \times F_{n+1}$ and $3 \times F_{n+1} \times F_{n+2}$, for some n, where $F_n = 0, 1, 3, 8, 21, 55, \ldots$ are odd Fibonacci numbers.



A tensor decomposition of a tensor T is

$$T = \sum_{i=1}^{r} D_i$$

with decomposable D_i and minimal r (called the rank).



The variety of decomposable tensors is the Segre variety.
In many cases of interest (for $d \geq 3$), tensor decomposition is unique (up to order), this allows to recover D_i from T. From this point of view, tensors may have even more applications than matrices.

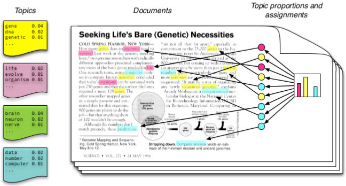
The Comon conjecture

Conjecture

A symmetric tensor T has a minimal symmetric decomposition, that is $T = \sum_{i=1}^{r} D_i$ where also D_i are symmetric.

Topic models

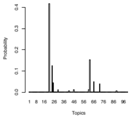
Topic models are statistical model for discovering the abstract "topics" that occur in large archives of documents.



Christos Papadimitriou and others gave in 1998 a probabilistic algorithm to attack this problem.

Tensor decomposition for Topic models

A. Bhaskara , M. Charikar , A. Vijayaraghavan  
propose in 2013 to compute probabilities for the occurrence of three words, then they use tensor decomposition of a 3-tensor.
Here uniqueness of tensor decomposition is crucial.



"Genetics"	"Evolution"	"Disease"	"Computers"
human	evolution	disease	computer
genome	evolutionary	host	models
dna	species	bacteria	information
genetic	organisms	diseases	data
genes	life	resistance	computers
sequence	origin	bacterial	system
gene	biology	new	network
molecular	groups	strains	systems
sequencing	phylogenetic	control	model
map	living	infectious	parallel
information	diversity	malaria	methods
genetics	group	parasite	networks
mapping	new	parasites	software
project	two	united	new
sequences	common	tuberculosis	simulations

The cocktail party problem

Suppose we have

- p persons in the room, speaking,
- p microphones spread over the room,
- we record sound at N discrete time steps.

Because of the superposition of sound waves, we can model this as

$$Y = MX,$$

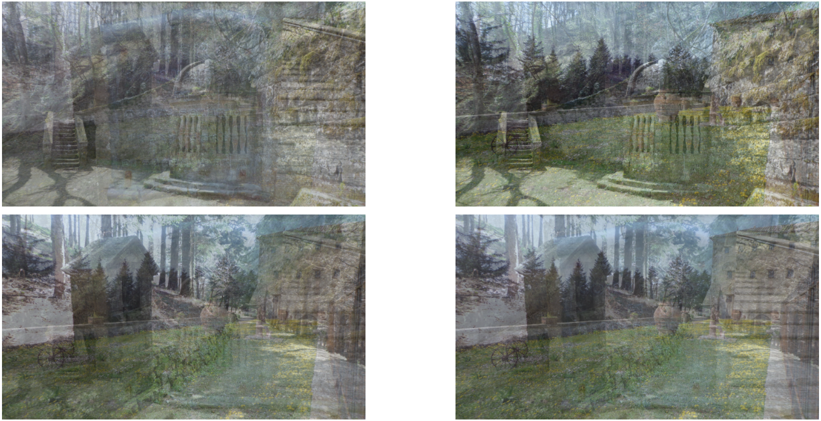
where,

- row i of X contains the sound produced by person i at each of the time steps.
- row i of Y contains the sound measured at microphone i at each of the time steps.

Police wants to recover individual speeches.

The cyclically broken camera problem

courtesy by Nick Vannieuvenhowen 



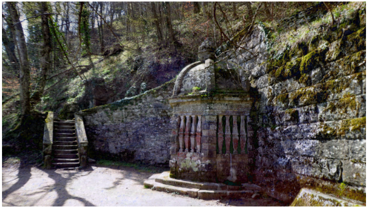
The cyclically broken camera problem

Recovered

Original

The cyclically broken camera problem

Recovered



Original

The stochastic matrix completion problem

Problem

Given some entries of a matrix, is it possible to add the missing entries so that the matrix has rank 1*, its entries sum to one, and it is nonnegative?*

Example

For example, the partial matrix

$$\begin{pmatrix} 0.16 & & & \\ & 0.09 & & \\ & & 0.04 & \\ & & & 0.01 \end{pmatrix}$$

Example

For example, the partial matrix

$$\begin{pmatrix} 0.16 & & & \\ & 0.09 & & \\ & & 0.04 & \\ & & & 0.01 \end{pmatrix}$$

has a unique completion:

$$\begin{pmatrix} 0.16 & 0.12 & 0.08 & 0.04 \\ 0.12 & 0.09 & 0.06 & 0.03 \\ 0.08 & 0.06 & 0.04 & 0.02 \\ 0.04 & 0.03 & 0.02 & 0.01 \end{pmatrix}.$$

Diagonal stochastic partial matrices


Theorem (Kubjas,Rosen)

Let M be an $n \times n$ partial stochastic matrix, where $n \geq 2$, with nonnegative observed entries along the diagonal:

$$M = \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix}.$$

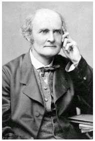
Then M is completable if and only if $\sum_{i=1}^{n} \sqrt{a_i} \leq 1$. In the special case $\sum_{i=1}^{n} \sqrt{a_i} = 1$, the partial matrix M has a unique completion.

Theorem (Kubjas,Rosen)

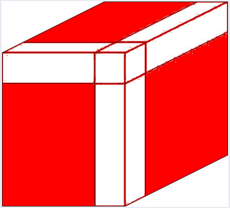
Suppose we are given a partial probability tensor $T \in (\mathbb{R}^n)^{\otimes d}$ with nonnegative observed entries a_i along the diagonal, and all other entries unobserved. Then T is completable if and only if

$$\sum_{i=1}^{n} a_i^{1/d} \leq 1.$$

Hyperdeterminant



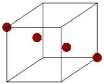
Theorem (Cayley)

A tensor A can be reduced to the form  *if and only if* $\mathrm{Det}(A) = 0$ *where* (Det) *is the equation of a irreducible hypersurface, called hyperdeterminant.*

When the triangle inequality $n_j \leq \sum_{i \neq j} n_i$ is satisfied, the hyperdeterminant gives the dual variety of Segre variety of decomposable tensors.

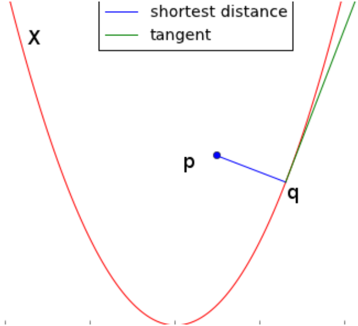
Orthogonal setting again.

E. Robeva gives (quadratical) equations for the tensors that can be reduced to diagonal form, up to orthogonal transformation.



The minimum distance problem

Let $X \subset \mathbb{A}^n_{\mathbb{R}}$ be an algebraic variety, let $p \in \mathbb{A}^n_{\mathbb{R}}$. We look for the points $q \in X$ which minimize the euclidean distance $d(p, q)$.
A necessary condition, assuming q is a smooth point of X, is that the tangent space T_qX be orthogonal to $p - q$, this is the condition to get a critical point for the distance function $d(p.-)$.



The case of matrices of bounded rank

There is a relevant case when the minimum distance problem has been solved.

Consider the affine space of $n \times m$ matrices, and let X_k be the *variety of matrices of rank* $\leq k$.

X_1 is the cone over the Segre variety $\mathbb{P}^{n-1} \times \mathbb{P}^{m-1}$. X_k is the k-secant variety of X_1. The matrices in X_k which minimize the distance from A, are called the *best rank k approximations of A*.

Critical points in any rank.

Lemma (SVD revisited)

Let $A = XDY$ be the SVD decomposition of A. The critical points of the distance function $d_A = d(A, -)$ from A to the variety X_k of rank $\leq k$ matrices are given by

$$\Sigma_{j\in\{i_1,\ldots,i_k\}}\sigma_j x_j \otimes y_j$$

for any subset of indices $\{i_1, \ldots, i_k\}$, where x_i are the columns of X, y_j are the rows of Y and D has σ_i on the diagonal.

The number of critical points for A of rank $r \geq k$ is $\binom{r}{k}$. For a general $m \times n$ matrix, assuming $m \leq n$, it is $\binom{m}{k}$.

Rank one approximation for tensors

L.-H. Lim gave in 2005 a variational construction of critical rank one tensors, including a best rank one approximation.

This includes singular t-ples and eigentensors (defined independently by L. Qi).

E. Robeva work quoted before, fits actually in this area.

Counting eigentensors and singular d-ples of vectors

The number of eigentensors in $\mathrm{Sym}^d\mathbb{C}^n$ is

$$\frac{(d-1)^m - 1}{d-2} \quad \text{(Cartwright-Sturmfels)}$$

The number of singular d-ples in $\bigotimes^d \mathbb{C}^n$ is the coefficient of $(t_1 \ldots t_d)^{m-1}$ in the polynomial

$$\prod_{i=1}^{d} \frac{\left(\sum_{j\neq i} t_j\right)^m - t_i^m}{\left(\sum_{j\neq i} t_j\right) - t_i} \quad \text{(O-Friedland)}$$

These two computations has been performed by using respectively Toric Varieties and Chern Classes.
Both these computations have been generalized with the concept of ED degree.

ED degree has the more general to measure the complexity of computing the distance from an object, which is in some sense a good measure of the complexity of the object itself.

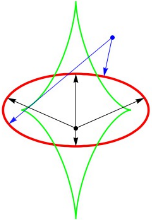
Joint work with J. Draisma, E. Horobet, B. Sturmfels, R. Thomas.

Definition

- The Euclidean Distance(ED) degree of an affine variety X is the number of critical points of the euclidean distance d_p from p to X, for general p.
- The ED degree of a projective variety X is the ED degree of its affine cone.

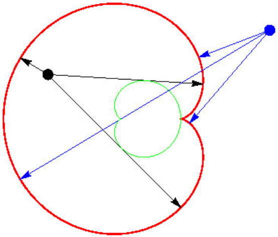
ED degree has a long history

(Apollonius)



- A line has ED degree $= 1$.
- A circle has ED degree $= 2$.
- A parabola has ED degree $= 3$.
- A general conic has ED degree $= 4$.

The cardioid picture



Duality property

The dual variety of $X \subset \mathbb{P}(V)$, can be identified, thanks to the metric, to a variety in the same $\mathbb{P}(V)$.

Theorem (Draisma-Horobet-O-Sturmfels-Thomas)

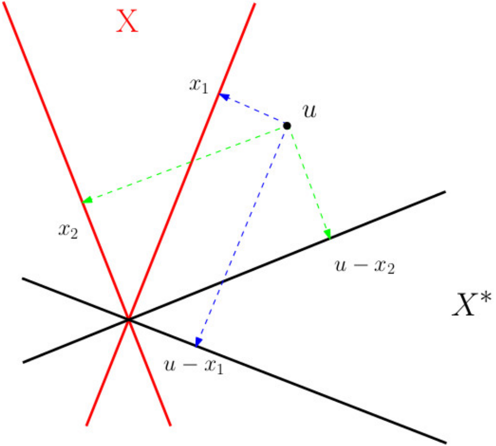
Bijection of critical points for X and $X^\vee$ Let $X^\vee$ be the dual variety of $X \subset \mathbb{P}(V)$. Let $u \in \mathbb{P}(V)$. If $x \in X$ is a critical point for $d(p, -)$ on X, then $u - x$ belongs to $X^\vee$ and it is a critical point for $d(p, -)$ on $X^\vee$. In particular

$$ED\ degree\ X = \ ED\ degree\ X^\vee.$$

Corollary

Let u be a tensor and u^* be its best rank one approximation. Then the hyperdeterminant of $u - u^*$ is zero.

Thanks



Thanks for your attention !