Symmetric Tensors and Their Decomposition: From Basics to Sylvester's Algorithm

A. Bernardi

January 29, 2025

University of Trento

Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
UNIVERSITÀ DI TRENTO

UNIVERSITÀ DI TRENTO
Dipartimento di Matematica

Introduction

Tensors "by hands"

An order-1 tensor can be represented as a vector (after having chosen basis):

$$v = (a_1, \cdots, a_n) \in \mathbb{C}^n.$$

Tensors "by hands"

An order-2 tensor can be represented as a matrix:

$$A = (a_{i,j})_{\{i=1,\dots,n_1; j=1,\dots,n_2\}} = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n_2} \\ \vdots & & \vdots \\ a_{n_1,1} & \cdots & a_{n_1,n_2} \end{pmatrix} \in \mathbb{C}^{n_1 \times n_2}$$

Remark
After having fixed bases B_1 and B_2 for $\mathbb{C}^{n_1}$ and $\mathbb{C}^{n_2}$ resp., one can associate A to a linear map $f_A : (\mathbb{C}^{n_1})^* \to \mathbb{C}^{n_2}$

$$A = M_{B_1,B_2}(f)$$

So A can be thought as an element of $Hom((\mathbb{C}^{n_1})^*, \mathbb{C}^{n_2})$.

We define for now

$$\mathbb{C}^{n_1} \otimes \mathbb{C}^{n_2} := Hom((\mathbb{C}^{n_1})^*, \mathbb{C}^{n_2})$$

Tensors "by hands"

An order-3 tensor can be represented as a 3-dimensional box (after having fixed basis):

$$A = (a_{i_1,i_2,i_3})_{\{i_j=1,\ldots,n_j;j=1,\ldots,3\}} \in \mathbb{C}^{n_1\times n_2\times n_3}$$

There exist 3 bilinear maps associated to an order-3 tensor (after having fixed basis):

$$f_{A,3} : (\mathbb{C}^{n_1} \times \mathbb{C}^{n_2})^* \to \mathbb{C}^{n_3}$$
$$f_{A,2} : (\mathbb{C}^{n_1} \times \mathbb{C}^{n_3})^* \to \mathbb{C}^{n_2}$$
$$f_{A,1} : (\mathbb{C}^{n_2} \times \mathbb{C}^{n_3})^* \to \mathbb{C}^{n_1}$$

$$A = M_{B_{1,2},B_3}(f_{A,3}) = M_{B_{1,3},B_2}(f_{A,2}) = M_{B_{2,3},B_1}(f_{A,1})$$

So A can be thought as an element of $Hom((\mathbb{C}^{n_1} \otimes \mathbb{C}^{n_2})^*, \mathbb{C}^{n_3})$, $Hom((\mathbb{C}^{n_1} \otimes \mathbb{C}^{n_3})^*, \mathbb{C}^{n_2})$, $Hom((\mathbb{C}^{n_2} \otimes \mathbb{C}^{n_3})^*, \mathbb{C}^{n_1})$.

Tensors "by hands"

We define for now

$$\mathbb{C}^{n_1} \otimes \mathbb{C}^{n_2} \otimes \mathbb{C}^{n_3} := Hom((\mathbb{C}^{n_1} \otimes \mathbb{C}^{n_2})^*, \mathbb{C}^{n_3}) =$$

$$= Hom((\mathbb{C}^{n_1} \otimes \mathbb{C}^{n_3})^*, \mathbb{C}^{n_2}) =$$

$$= Hom((\mathbb{C}^{n_2} \otimes \mathbb{C}^{n_3})^*, \mathbb{C}^{n_1})$$

Tensors "by hands"

An order-k tensor can be seen as k-dimensional box (after having fixed basis):

$$A = (a_{i_1,\dots,i_k})_{\{i_j=1,\dots,n_j; j=1,\dots,k\}} \in \mathbb{C}^{n_1\times\cdots\times n_k}$$

There exist many bilinear maps associated to an order-k tensor (after having fixed basis), for example:

$$f_A : (\mathbb{C}^{n_1} \times \cdots \times \mathbb{C}^{n_{k-1}})^* \to \mathbb{C}^{n_k}.$$

So A can be thought as an element, for example, of

$$Hom(\mathbb{C}^{n_1} \times \cdots \times \mathbb{C}^{n_{k-1}}, \mathbb{C}^{n_k}).$$

We define for now

$$\mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_k} := Hom(\mathbb{C}^{n_1} \times \cdots \times \mathbb{C}^{n_{k-1}}, \mathbb{C}^{n_k})$$

Applications

Topic Models

How could we organize a multitude of documents, e.g., web pages, into a number of clusters of documents about the same topic, without specifying in advance the topics?

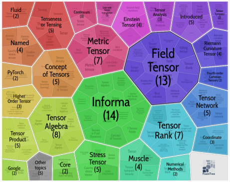


Figure 1: `http://search.carrot2.org`

This is known in machine learning as document clustering in the unsupervised learning setting.

Let's view this as a two-phase statistics problem:

- Estimate the parameters of a statistical model,
- Classify documents based on model.

Topic model: Clustering related fields

Assumption 1

The set of documents is described by a bag of words.

For example, consider the following vocabulary:

$W = \{\text{tensor, decomposition, energy, material, stress, algebra, equation, curvature}\}$

Every document is represented by a bag of words from W, e.g.,

$$
\begin{aligned}
D_1 &= \{\text{tensor, decomposition, algebra, equation, tensor, decomposition}\} \\
D_2 &= \{\text{energy, curvature, equation, energy, curvature, equation}\} \\
D_3 &= \{\text{material, stress, material, equation, stress, material}\} \\
&\ \vdots \\
D_N &= \{\text{tensor, energy, stress, equation}\}
\end{aligned}
$$

The order of appearance is assumed to be irrelevant.

Topic model: Clustering related fields

Assumption 2

The set of documents contains r distinct topics, and the probability of key words is conditional on the topic of the document.

For instance, consider three abstract topics:

$$T = \{ \text{ Mathematics, Physics, Engineering } \}$$

Then, common sense (and statistical analysis) tells us that:

- $P(W = \text{"tensor decomposition"} \mid T = \text{Mathematics}) \gg P(W = \text{"tensor decomposition"} \mid T = \text{Physics})$.
- $P(W = \text{"energy"} \mid T = \text{Physics}) \gg P(W = \text{"energy"} \mid T = \text{Mathematics})$.
- $P(W = \text{"material stress"} \mid T = \text{Engineering}) \gg P(W = \text{"material stress"} \mid T = \text{Physics})$.

Topic model: Independence of words within topics

Assumption 3

The words W in a document are independently and identically distributed (i.i.d.), conditional on the topic.

Let $W = \{w_i\}$ be the words and $T = \{t_i\}$ the topics. Then,

$$P(W = w_i, W = w_j \mid T = t_k) = P(W = w_i \mid T = t_k) \cdot P(W = w_j \mid T = t_k)$$

for all i, j, k.

For example:

- $P(\text{"decomposition"} \mid T = \text{Mathematics}) \cdot P(\text{"algebra"} \mid T = \text{Mathematics})$ gives the probability of both terms appearing together in a mathematical document.

Topic model

With these assumptions, the total probability distribution is

$$P(W) = \sum_{i=1}^{r} P\left(T = t_i\right) \cdot P\left(W \mid T = t_i\right)$$

Let's use the more familiar-looking (at least to me) notation

$$\mathbf{p} = \sum_{i=1}^{r} \alpha_i \mathbf{p}_i$$

Topic model

Statistical parameter inference problem

All that we are handed is:

1. A set of k documents D,
2. (A set of n key words W), and
3. (The number of topics r).

We assume 2 and 3 for convenience, but they can be estimated from D.

We want to find the probability distribution vectors $\mathbf{p}_i$ for each topic t_i and identify the parameters of the model.

Topic model: Probabilities of words and topics

The vocabulary of words is:

$W = \{\text{tensor, decomposition, energy, material, stress, algebra, equation, curvature}\}$

Assume the true probability distributions over the 8 words conditional on the topics T are:

$$\mathbf{p}_{\mathsf{M}} := \begin{bmatrix} 40.00\% \\ 25.00\% \\ 5.00\% \\ 5.00\% \\ 5.00\% \\ 15.00\% \\ 5.00\% \\ 0.00\% \end{bmatrix}, \mathbf{p}_{\mathsf{P}} := \begin{bmatrix} 10.00\% \\ 5.00\% \\ 25.00\% \\ 5.00\% \\ 5.00\% \\ 5.00\% \\ 15.00\% \\ 30.00\% \end{bmatrix}, \mathbf{p}_{\mathsf{E}} := \begin{bmatrix} 5.00\% \\ 5.00\% \\ 10.00\% \\ 25.00\% \\ 35.00\% \\ 5.00\% \\ 10.00\% \\ 5.00\% \end{bmatrix}$$

and that the probability of each topic is:

$$\alpha_{\mathsf{M}} = 35\%, \quad \alpha_{\mathsf{P}} = 40\%, \quad \alpha_{\mathsf{E}} = 25\%.$$

Topic model: Total probability distribution

The total probability distribution is thus:

$$\mathbf{p} = \alpha_{\mathsf{M}}\mathbf{p}_{\mathsf{M}} + \alpha_{\mathsf{P}}\mathbf{p}_{\mathsf{P}} + \alpha_{\mathsf{E}}\mathbf{p}_{\mathsf{E}}$$

Substituting the values:

$$\mathbf{p} = 0.35 \cdot \mathbf{p}_{\mathsf{M}} + 0.40 \cdot \mathbf{p}_{\mathsf{P}} + 0.25 \cdot \mathbf{p}_{\mathsf{E}} = \begin{bmatrix} 18.25\% \\ 16.75\% \\ 15.30\% \\ 11.85\% \\ 14.25\% \\ 8.75\% \\ 10.30\% \\ 4.55\% \end{bmatrix}.$$

Note: While we have the overall distribution $\mathbf{p}$, it is not possible to recover the conditional probabilities (i.e., $\mathbf{p}_{\mathsf{M}}, \mathbf{p}_{\mathsf{P}}, \mathbf{p}_{\mathsf{E}}$) directly from $\mathbf{p}$, because: **Vectors are not *r*-identifiable:** The total distribution $\mathbf{p}$ is a mixture of the r topic-specific distributions, and multiple decompositions of $\mathbf{p}$ could satisfy the equation.

Topic model: Non-uniqueness of factorization

Indeed, there are ∞ possibilities of writing a vector as a linear combination of r vectors.

For example, we can write:

$$\mathbf{p} = \begin{bmatrix} 18.25\% \\ 16.75\% \\ 15.30\% \\ 0\% \\ 0\% \\ 0\% \\ 0\% \\ 0\% \end{bmatrix} + \begin{bmatrix} 0\% \\ 0\% \\ 0\% \\ 11.85\% \\ 14.25\% \\ 8.75\% \\ 0\% \\ 0\% \end{bmatrix} + \begin{bmatrix} 0\% \\ 0\% \\ 10\% \\ 0\% \\ 0\% \\ 0\% \\ 10.30\% \\ 4.55\% \end{bmatrix}$$

which is a completely different factorization.

This highlights that the decomposition of **p** into topic-specific vectors and weights is not unique, and additional constraints are needed for uniqueness.

Topic model

Let $\alpha_l := P\left(T = t_l\right)$.

Recall Assumption 3: the distribution of the words is i.i.d., conditional on the topic. Therefore,

$$\begin{aligned} P\left(W = w_i, W = w_j\right) &:= \sum_{\ell=1}^{r} \alpha_\ell \cdot P\left(W = w_i, W = w_j \mid T = t_\ell\right) \\ &= \sum_{\ell=1}^{r} \alpha_\ell \cdot P\left(W = w_i \mid T = t_l\right) P\left(W = w_j \mid T = t_\ell\right) \end{aligned}$$

so that we get

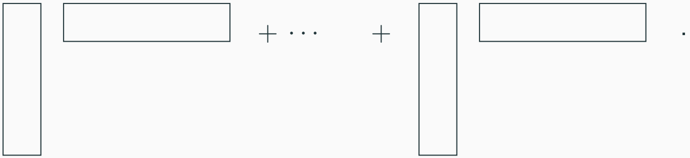
$$P := \sum_{\ell=1}^{r} \alpha_\ell \cdot \mathbf{p}_\ell \mathbf{p}_\ell^T$$

This is a rank- r symmetric matrix decomposition!

Topic model

The total joint probability distribution is

$$P := \sum_{\ell=1}^{r} \alpha_\ell \cdot \mathbf{p}_\ell \mathbf{p}_\ell^T$$

$=$ 

Given k documents D, we determine its population joint probability distribution $\widehat{P}$ by counting word pairs, leading to a matrix

$$\widehat{P} = P + \varepsilon(k), \quad \text{where } \lim_{k\to\infty} \varepsilon \to 0_{n\times n}$$

Topic model

However, stochastic matrices are not r-identifiable!

Indeed, consider the map

$$\mathbb{R}^{r-1} \times \left(\mathbb{R}^{n-1}\right)^{\times r} \to \mathbb{R}^{n\times n},$$
$$\left([\alpha_i]_{i=1}^{r-1}, \widehat{\mathbf{p}}_1, \ldots, \widehat{\mathbf{p}}_r\right) \mapsto \sum_{\ell=1}^{r} \alpha_\ell \begin{bmatrix} \widehat{\mathbf{p}}_\ell \\ 1 - \mathbf{1}^T \widehat{\mathbf{p}}_\ell \end{bmatrix} \begin{bmatrix} \widehat{\mathbf{p}}_\ell & 1 - \mathbf{1}^T \widehat{\mathbf{p}}_\ell \end{bmatrix},$$

where $\alpha_r := 1 - \sum_{\ell=1}^{r-1} \alpha_\ell$.
The domain has dimension $r - 1 + r(n-1)$, but the image is contained in the set of symmetric matrices of rank bounded by r, which only has dimension $\binom{n+1}{2} - \binom{n+1-r}{2}$. For $r \geq 2$ the latter is smaller than the former ($1/2r^2 - 1/2r - 1 > 0$). It follows from the fiber dimension theorem (and the fact that the map above is a polynomial map with an $\mathbb{R}$-variety as source) that every fiber has a positive-dimensional component.

Everything will be much better with "words triplets"

Topic model

Let $\alpha_l := P(T = t_l)$.

Recall Assumption 3: the distribution of the words is i.i.d., conditional on the topic. Therefore,

$$P(W = w_i, W = w_j, W = w_k) :=$$

$$\sum_{\ell=1}^{r} \alpha_\ell \cdot P(W = w_i, W = w_j, W = w_k \mid T = t_\ell)$$

$$= \sum_{\ell=1}^{r} \alpha_\ell \cdot P(W = w_i \mid T = t_l)\, P(W = w_j \mid T = t_\ell)\, P(W = w_k \mid T = t_l)$$

so that we get

$$P := \sum_{\ell=1}^{r} \alpha_\ell \cdot \mathbf{p}_\ell \otimes \mathbf{p}_\ell \otimes \mathbf{p}_\ell$$

This is a rank- r symmetric tensor decomposition, which, if we are not very very unlucky it will have a unique solution.

Topic model

The point that I want to stress now is that the structure in which we stock the information is important:

We were not able to find a unique solution for

1. $v = \sum_{i=1}^{r} \alpha_i v_i \in V$
2. $A = \sum_{i=1}^{r} \alpha_i v_i vi^T \in V \otimes V_2$

while in principle the following has almost always unique solution:

$$T = \sum_{i=1}^{r} \alpha_i v_i^{(1)} \otimes \cdots \otimes v_i^{(k)} \in V_1 \otimes \cdots \otimes V_k.$$

Multilinear algebra and tensors

Multilinear algebra and tensors: Precise definition of Tensors

The tensor product

$$\otimes : V_1 \times \cdots \times V_d \to V_1 \otimes \cdots \otimes V_d$$

is the unique multilinear map that satisfies the universal property:

1. $V_1 \otimes \cdots \otimes V_d = \operatorname{span}\left(\otimes\left(V_1, \ldots, V_d\right)\right)$.
2. If ϕ is a multilinear map $\phi : V_1 \times \cdots \times V_d \to H$ then there exists a unique linear map f such that $\phi\left(\mathbf{x}_1, \ldots, \mathbf{x}_d\right) = f\left(\mathbf{x}_1 \otimes \cdots \otimes \mathbf{x}_d\right)$.
   In other words, there is a unique linear map such that the diagram

$$\begin{tikzcd}
V_1 \times \cdots \times V_d \arrow[r, "\phi"] \arrow[d, "\otimes"'] & H \\
V_1 \otimes \cdots \otimes V_d \arrow[ur, "f"'] &
\end{tikzcd}$$

   commutes.

Multilinear algebra and tensors

- Every multilinear map is a linear map after the proper embedding via the tensor product.
- The vector space $V_1 \otimes \cdots \otimes V_d$ is called the **tensor product** of the V_i's.
- It is uniquely determined by $V_1, \ldots, V_d$ up to isomorphism and it is commutative in the sense that

$$V_1 \otimes \cdots \otimes V_d \simeq V_{\pi_1} \otimes \cdots \otimes V_{\pi_d}$$

  for any permutation π of $[1, d]$.
- The binary tensor product is associative, i.e., $(U \otimes V) \otimes W = U \otimes (V \otimes W)$
- The unique d-factor tensor product is the repeated binary tensor product, i.e., $\otimes (V_1 \times \ldots \times V_d) = \otimes (\cdots \otimes (\otimes (V_1, V_2), V_3) \cdots), V_d)$ which explains the notation.
- So we neither need to specify the number of arguments in the tensor product nor place parenthesis around the binary tensor product.

Multilinear algebra and tensors: back to "tensors by hands"

We call a basis B of $V_1 \otimes \cdots \otimes V_d$ a tensor basis if the basis elements are in the image of the d-factor tensor product map $\otimes$, i.e., if $B = \left\{\mathbf{v}_i^1 \otimes \cdots \otimes \mathbf{v}_i^d \mid \mathbf{v}_i^k \in V_k\right\}_i$.

Let $\left\{\mathbf{e}_i^k\right\}_{i=1}^{n_k}$ be a basis of $V_k, k = 1, \ldots, d$. Then,

$$B = \left\{\mathbf{e}_{i_1}^1 \otimes \cdots \otimes \mathbf{e}_{i_d}^d\right\}_{i_1,\ldots,i_d=1}^{n_1,\ldots,n_d}$$

is the natural tensor basis of $V_1 \otimes \cdots \otimes V_d$.

$$\dim\left(V_1 \otimes \cdots \otimes V_d\right) = \dim V_1 \cdots \dim V_d = n_1 \cdots n_d.$$

Multilinear algebra and tensors: back to "tensors by hands"

The dual basis of the natural basis of $V_1 \otimes \cdots \otimes V_d$ is

$\left\{ \left(\mathbf{e}^1_{i_1} \otimes \cdots \otimes \mathbf{e}^d_{i_d}\right)^* : V_1^* \otimes \cdots \otimes V_d^* \to \mathbb{F} \mid (1, \ldots, 1) \le (i_1, \ldots, i_d) \le (n_1, \ldots, n_d) \right\}$,

where $\left(\mathbf{e}^1_{i_1} \otimes \cdots \otimes \mathbf{e}^d_{i_d}\right)^*$ is defined as

$$\left(\mathbf{e}^1_{i_1} \otimes \cdots \otimes \mathbf{e}^d_{i_d}\right)^* \left(\sum_{i=1}^r \mathbf{v}^1_i \otimes \cdots \otimes \mathbf{v}^d_i \right) = \sum_{i=1}^r \left(\mathbf{e}^1_{i_1}\right)^* \left(\mathbf{v}^1_i\right) \cdots \left(\mathbf{e}^d_{i_d}\right)^* \left(\mathbf{v}^d_i\right)$$

Let $\left\{\mathbf{e}^k_{i_k}\right\}_{i=1}^{n_k}$ be the standard basis of V_k for $k = 1, \ldots, d$, so that the elements $\mathcal{A} \in V_1 \otimes \cdots \otimes V_d$ can always be represented in coordinates as a d-dimensional array:

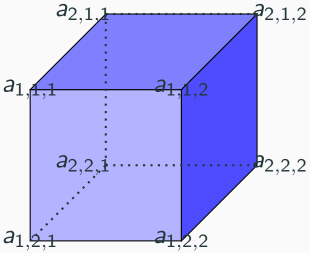
if $\mathcal{A} = \sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} \underbrace{\left(\mathbf{e}^1_{i_1} \otimes \cdots \otimes \mathbf{e}^d_{i_d}\right)^* (\mathcal{A})}_{a_{i_1,\ldots,i_d}} \cdot \left(\mathbf{e}^1_{i_1} \otimes \cdots \otimes \mathbf{e}^d_{i_d}\right)$, then

$\mathcal{A} = [a_{i_1,\ldots,i_d}]_{i_1,\ldots,i_d=1}^{n_1,\ldots,n_d}$ represents $\mathcal{A}$ in the standard tensor basis. No notational distinction is made between abstract multilinear operators $A \in V_1 \otimes \cdots \otimes V_d$ and the d-dimensional arrays representing them (with respect to the standard basis).

Multilinear algebra and tensors: back to "tensors by hands"

The elements of $\mathbb{C}^{n_1 \times \cdots \times n_d}$ can be therefore represented as d-dimensional arrays. The integer $d \geq 2$ is called the **order of the tensor**.

Ex. $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2 \ni \sum_{i_1,i_2,i_3=1,2} a_{i_1,i_2,i_3} e_{i_1} \otimes e_{i_2} \otimes e_{i_3}$



For brevity, we will also call them "tensors," although it would be more accurate to call them a coordinate representation of a tensor with respect to some basis: $[a_{i_1,i_2,i_3}]_{\mathcal{E}}^{i_1,i_2,i_3=1,2}$.

Rank of a tensor

Definition: The rank of a tensor is the minimum number of simple tensors $\mathbf{u}^{(1)} \otimes \mathbf{u}^{(2)} \otimes \cdots \otimes \mathbf{u}^{(d)}$ required to express it as their sum.

For a tensor $T \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$,

$$T = \sum_{i=1}^{r} \mathbf{u}_i^{(1)} \otimes \mathbf{u}_i^{(2)} \otimes \cdots \otimes \mathbf{u}_i^{(d)},$$

where $\mathbf{u}_i^{(j)}$ are vectors in $\mathbb{R}^{n_j}$, and r is the rank of T.

- For matrices ($d = 2$), the rank is the dimension of the column or row space.
- For higher-order tensors ($d \geq 3$), the rank is generally harder to compute and is not directly related to the rank of a matrix.

Rank-1 tensors

$$S_d : \mathbb{P}V_1 \times \ldots \times \mathbb{P}V_d \longrightarrow \mathbb{P}\left(V_1 \otimes \ldots \otimes V_d\right)$$

$$\left([v_1], \ldots, [v_d]\right) \longmapsto [v_1 \otimes \cdots \otimes v_d]$$

Definition
$Im\left(S_d\right) = \{$ rank-1 tensors$\} =$ Segre variety

Rank-1 tensors

How can we determine if a tensor has rank 1 ?

We need equations of Segre variety.

Example 2-factors ($d = 2$):

$$S_2 : \mathbb{P}(V_1) \times \mathbb{P}(V_2) \to \mathbb{P}\left(V_0 \otimes V_2\right)$$

$$\left([v_1], [v_2]\right) \longmapsto [v_1 \otimes v_2]$$

$Im(S_2)$ =rank-1 matrices of order ($\dim V_1 \times \dim V_2$) $\Rightarrow$ the eq's of $Im\left(S_2\right) \subset \mathbb{C}[z_{1,1}, \ldots, z_{n_1,n_2}]$ are the (2×2)-minors of

$$M = \begin{pmatrix} z_{1,1} & \cdots & z_{1,n_1} \\ \vdots & & \vdots \\ z_{n_2,1} & \cdots & z_{n_2,n_1} \end{pmatrix}$$

$$I = \left(z_{i_1,j_1} \cdot z_{i_2,j_2} - z_{i_1,j_2} \cdot z_{i_2,j_1}\right)$$

Claim: The ideal of the Segre variety of 2 factors is I.

Rank-1 tensors

Indeed

$\mathbb{P}(V_1) \times \mathbb{P}(V_2) \to \mathbb{P}\left(V_1 \otimes V_2\right)$

$\left(\left[x_1, \ldots, x_{n_1}\right], \left[y_1, \ldots, y_{n_2}\right]\right) \mapsto \left[x_1, \ldots, x_{n_1}\right] \otimes \left[y_1, \ldots, y_{n_2}\right] =$

$$= \begin{pmatrix} x_1 \\ \vdots \\ x_{n_1} \end{pmatrix} \cdot \begin{pmatrix} y_1 & \cdots & y_{n_2} \end{pmatrix} = \begin{pmatrix} x_1 y_1 & \cdots & x_1 y_{n_2} \\ \vdots & & \vdots \\ x_{n_1} y_1 & \cdots & x_{n_1} y_{n_2} \end{pmatrix}$$

This is a rank-1 matrix so the 2×2 minors vanish on it:

$z_{i_1, j_1} \cdot z_{i_2, j_2} - z_{i_1, j_2} \cdot z_{i_2, j_1} = x_{i_1} y_{j_1} x_{i_2} y_{j_2} - x_{i_2} y_{j_1} x_{i_1} y_{j_2} = 0$

Flattening of a Tensor

Key Idea: Rank-1 tensors can be characterized by the (2×2)-minors of their flattenings. (Set theoretically: Grone 1977, Ideally theorctically: Thai Há 2002)

For a tensor $T \in V_1 \otimes V_2 \otimes \cdots \otimes V_d$, the **flattening** associated with the disjoint subsets of indices $\{1, \ldots, k\}$ and $\{k+1, \ldots, d\}$ is the linear map:

$$T_{\{1,\ldots,k\},\{k+1,\ldots,d\}} : (V_1 \otimes \cdots \otimes V_k)^* \to V_{k+1} \otimes \cdots \otimes V_d,$$

defined by:

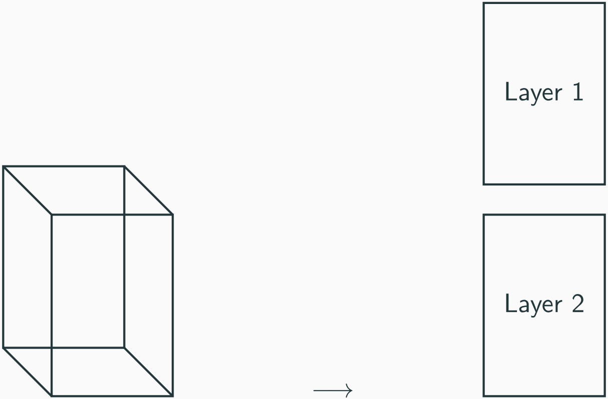
$$T_{\{1,\ldots,k\},\{k+1,\ldots,d\}}(\phi) = \phi \cdot T,$$

where $\phi \in (V_1 \otimes \cdots \otimes V_k)^*$ is a functional that contracts T along the indices $\{1, \ldots, k\}$.

Example: $T = (t_{i,j,k}) \in \mathbb{C}^{2\times3\times2}$, the flattening $T_{\{1,2\},\{3\}}$ produces a 6×2-matrix:

$$T_{\{1,2\},\{3\}} = \begin{bmatrix} t_{1,1,1} & t_{1,1,2} \\ t_{1,2,1} & t_{1,2,2} \\ t_{1,3,1} & t_{1,3,2} \\ t_{2,1,1} & t_{2,1,2} \\ t_{2,2,1} & t_{2,2,2} \\ t_{2,3,1} & t_{2,3,2} \end{bmatrix}.$$

Flattening of a Tensor



Equations of Rank-1 Tensors

$$T = \sum_{i=1}^{r} a_1^i \otimes \cdots \otimes a_d^i \in V_1 \otimes \cdots \otimes V_d.$$

The **matrix associated with the partial flattening** is:

$$T_{\{1,\dots,k\},\{k+1,\dots,d\}} = \sum_{i=1}^{r} (a_1^i \otimes \cdots \otimes a_k^i) \cdot (a_{k+1}^i \otimes \cdots \otimes a_d^i)^T.$$

Case of rank-1 tensors: If $\mathrm{rank}(T) = 1$, then:

$$T = a_1 \otimes \cdots \otimes a_d.$$

The matrix flattening becomes:

$$T_{\{1,\dots,k\},\{k+1,\dots,d\}} = (a_1 \otimes \cdots \otimes a_k) \cdot (a_{k+1} \otimes \cdots \otimes a_d)^T,$$

which is of the form $w_1 \otimes w_2^T$, i.e., a rank-1 matrix.

Equations of rank-1 TENSORS

We have proved that

If $\text{rank}(T) = 1 \Rightarrow (2 \times 2)$-minors of the flattening vanish.

The reverse implication comes from the fact that if $T \in V_1 \otimes V_2 \otimes \cdots \otimes V_d$ is a tensor such that all its flattenings $T_{\{1,\ldots,k\},\{k+1,\ldots,d\}}$ have vanishing (2×2)-minors, hence each flattening can be written as:

$$T_{\{1,\ldots,k\},\{k+1,\ldots,d\}} = v_1 \otimes v_2,$$

where $v_1 \in V_1 \otimes \cdots \otimes V_k$ and $v_2 \in V_{k+1} \otimes \cdots \otimes V_d$. If $rk(T) \geq 2$ there would exist at least a flattening of rank 2.

Proposition:

$$I(\{\text{rank-1 tensors}\}) = (2 \times 2 \text{ minors or all flattenings})$$

How to characterize tensors of rank > 1?

Analogously: if $rk(T) = 2$ then all the 3-minors of all flattening should vanish.

Question: Does the variety of 3-minors define tensors of rank ≤ 2?

Example: $W = e_1 \otimes e_1 \otimes e_2 + e_1 \otimes e_2 \otimes e_1 + e_2 \otimes e_1 \otimes e_1 \in (\mathbb{C}^2)^{\otimes 3}$.

The flattening $W_{\{1\},\{1,2\}}(\mathbb{C}^2)^* \to \mathbb{C}^2 \otimes \mathbb{C}^2$ in the standard basis can be represented as $(e_1 \otimes e_2 + e_2 \otimes e_1, e_1 \otimes e_2)$ (for any choice of order). So all the minors have rank 2.

With a simple linear system you can check that W doesn't have rank 2 (so $rk(W) = 3$).

Indeed we can express W as the limit of a sequence of tensors of rank 2.

$$W_n = \frac{1}{n}(e_1 + ne_2) \otimes (e_1 + ne_2) \otimes (e_1 + ne_2) - n^2 e_2 \otimes e_2 \otimes e_2.$$

Clearly $\lim_{n \to 0} W_n = W$.

Indeed $W \in T_{e_1^{\otimes 3}} S_3(\mathbb{C}^2)$.

Key: Zero's of polynomials define a closed set (with Zariski topology, Euclidean over $\mathbb{C}$).

Secant varieties

Let $S_k := \{rk-1 \text{ tensors } \in \mathbb{P}(V_1 \otimes \cdots \otimes V_k)\}$ be the Segre variety of k factors.

A tensor T has rank $\leq r$ if $\exists \leq r$ points on S_k, $P_1 \ldots, P_r \in S_k$ s.t. $T \in \langle P_1 \ldots, P_r \rangle$

The set of order k tensors of rank $\leq r$ is:

$$\sigma_r^0(S_k) := \bigcup_{P_1,\ldots,P_r \in S_k} \langle P_1, \ldots, P_r \rangle$$

As we have seen this set may not be closed: $T \in T_Q S_k \subset \overline{\sigma_2^0(S_k)}$ but $\mathrm{rk}(T) > 2$.

If this phenomenon happens we say that the border rank $brk(T) = 2$

Definition
The r-th secant variety of the Segre variety is

$$\sigma_r(S_k) := \overline{\sigma_r^0(S_k)}$$

and $brk(T) = \min\{r \,|\, T \in \sigma_r(S_k)\}$

$\sigma_1(S_k) = S_k \subseteq \sigma_2(S_k) \subseteq \cdots \subseteq \sigma_g(S_k) = \mathbb{P}^N$.

Secant varieties

Finding equations for secant varieties is in general an open problem.

Even though we were able to know the eq's of $\sigma_r(S_k)$ this will give the border rank, the border rank may be different from the rk.

Secant varieties

Finding equations for secant varieties is in general an open problem.

Even though we were able to know the eq's of $\sigma_r(S_k)$ this will give the border rank, the border rank may be different from the rk.

There is one cases (a part from matrices) where everything is known in terms of equations of secant varieties and in terms of an algorithm for the rank:

- symmetric tensors $\in (\mathbb{C}^2)^{\otimes d}$.

Symmetric tensors - Matrix case

Rank 1 matrices

$$v \otimes w = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} y_1 & \cdots & y_n \end{pmatrix} = \begin{pmatrix} x_1 y_1 & \cdots & x_1 y_n \\ \vdots & & \vdots \\ x_n y_1 & \cdots & x_n y_n \end{pmatrix}$$

Symmetric tensors - Matrix case

Rank 1 matrices

$$v \otimes w = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} y_1 & \cdots & y_n \end{pmatrix} = \begin{pmatrix} x_1 y_1 & \cdots & x_1 y_n \\ \vdots & & \vdots \\ x_n y_1 & \cdots & x_n y_n \end{pmatrix}$$

Rank 1 symmetric matrices

$$v \otimes v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} = \begin{pmatrix} x_1^2 & x_1 x_2 & \cdots & x_1 x_n \\ x_2 x_1 & \ddots & & \vdots \\ \vdots & & \ddots & x_{n-1} x_n \\ x_n x_1 & \cdots & x_n x_{n-1} & x_n^2 \end{pmatrix}$$

Symmetric tensors - Matrix case

Rank 1 matrices

$$v \otimes w = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} y_1 & \cdots & y_n \end{pmatrix} = \begin{pmatrix} x_1 y_1 & \cdots & x_1 y_n \\ \vdots & & \vdots \\ x_n y_1 & \cdots & x_n y_n \end{pmatrix}$$

Rank 1 symmetric matrices

$$v \otimes v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} = \begin{pmatrix} x_1^2 & x_1 x_2 & \cdots & x_1 x_n \\ x_2 x_1 & \ddots & & \vdots \\ \vdots & & \ddots & x_{n-1} x_n \\ x_n x_1 & \cdots & x_n x_{n-1} & x_n^2 \end{pmatrix}$$

Symmetric matrix of rank r:

$$A = \sum_{i=1}^{r} v_i \otimes v_i$$

Symmetric tensors - Matrix case

Symmetric matrices $\leftrightarrow$ Quadrics

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{12} & a_{22} & & & \vdots \\ a_{13} & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ a_{1n} & \cdots & & & a_{nn} \end{pmatrix} \leftrightarrow a_{11}x_1^2+2a_{12}x_{21}x_2+2a_{13}x_1x_3+\cdots+a_{nn}x_n^2$$

The rank of the quadric Q is the rank of M_Q the associated symmetric matrix $= \min r$ s.t $M_Q = \sum_{i=1}^{r} v_i \otimes v_i$.

Moreover $Q = \sum_{i=1}^{r} L_i^2$ with $L_i = v_{1,i}x_1 + \cdots + v_{n,i}x_n$ and $v_i = (v_{1,i}, \dots, v_{n,i})$.

In coordinates

$A = (a_{i_1,\ldots,i_d})_{i_j=1,\ldots,n_j, j=1,\ldots,d}$ is a symmetric hypermatrix (tensor) iff

$$a_{i_1,\ldots,i_d} = a_{\sigma(i_1),\ldots,\sigma(i_d)}$$

with $\sigma \in \mathcal{S}_d$ permutation group of d factors.

Symmetric Tensors

Symmetrization:

$$\pi_S : V^{\otimes d} \to S^d(V), \quad v_1 \otimes \cdots \otimes v_d \mapsto \frac{1}{d!} \sum_{\sigma \in \mathcal{S}_d} v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(d)}.$$

Correspondences:

- **Symmetric matrices** $\leftrightarrow$ quadrics.
- **Symmetric tensors of order** d $\leftrightarrow$ **homogeneous polynomials of degree** d:

$$K[x_0, \ldots, x_n]_d \to S^d(K^{n+1}),$$
$$f = \sum_{|\alpha|=d} a_\alpha x^\alpha \mapsto \sum_{|\alpha|=d} \frac{a_\alpha}{\binom{d}{\alpha}} e_{\alpha_1} \otimes \cdots \otimes e_{\alpha_n}.$$

Example:

$$f(x,y) = x^3 + 3x^2y + y^3 \mapsto e_0^{\otimes 3} + 3\, e_0^{\otimes 2} \otimes e_1 + e_1^{\otimes 3}.$$

Working with symmetric tensors is equivalent to work with homogeneous polynomials.

Symmetric Rank

Rank-1 Symmetric Tensor:

$$T = v \otimes \cdots \otimes v \in S^d(V) \subset V^{\otimes d}.$$

Definition: The symmetric rank of a symmetric tensor is the minimum r such that:

$$T = \sum_{i=1}^{r} v_i^{\otimes d},$$

where each $v_i^{\otimes d}$ is a rank-1 symmetric tensor.

Remark: The symmetric rank r and the usual rank r' (in $V^{\otimes d}$) satisfy:

$$r' \leq r.$$

$$T = \sum_{i=1}^{r'} v_1^i \otimes \cdots \otimes v_d^i = \sum_{i=1}^{r} v_i^{\otimes d}.$$

Symmetric Rank

Rank-1 Symmetric Tensor:

$$T = v \otimes \cdots \otimes v \in S^d(V) \subset V^{\otimes d}.$$

Definition: The symmetric rank of a symmetric tensor is the minimum r such that:

$$T = \sum_{i=1}^{r} v_i^{\otimes d},$$

where each $v_i^{\otimes d}$ is a rank-1 symmetric tensor.

Remark: The symmetric rank r and the usual rank r' (in $V^{\otimes d}$) satisfy:

$$r' \leq r.$$

$$T = \sum_{i=1}^{r'} v_1^i \otimes \cdots \otimes v_d^i = \sum_{i=1}^{r} v_i^{\otimes d}.$$

Comon's Question: Is $r' = r$?

We focus on the symmetric rank of symmetric tensors only.

Symmetric Rank-1 Tensors

Correspondence Between Symmetric Tensors and Polynomials:

- For $d = 1$: $V^{\otimes 1} \to S^1(V), \quad [a_0, \ldots, a_n] \mapsto a_0x_0 + \cdots + a_nx_n.$

Symmetric Rank-1 Tensors

Correspondence Between Symmetric Tensors and Polynomials:

- For $d = 1$: $V^{\otimes 1} \to S^1(V), \quad [a_0, \ldots, a_n] \mapsto a_0 x_0 + \cdots + a_n x_n$.
- For $d = 2$:

$$V^{\otimes 2} \to S^2(V), \quad v^{\otimes 2} \mapsto \begin{bmatrix} a_0^2 & a_0 a_1 & \cdots & a_0 a_n \\ a_0 a_1 & a_1^2 & \cdots & a_1 a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_0 a_n & a_1 a_n & \cdots & a_n^2 \end{bmatrix}.$$

  This corresponds to the polynomial:

$$a_0^2 x_0^2 + 2a_0 a_1 x_0 x_1 + \cdots + a_n^2 x_n^2 = (a_0 x_0 + \cdots + a_n x_n)^2.$$

Symmetric Rank-1 Tensors

Correspondence Between Symmetric Tensors and Polynomials:

- For $d = 1$: $V^{\otimes 1} \to S^1(V), \quad [a_0, \ldots, a_n] \mapsto a_0x_0 + \cdots + a_nx_n.$
- For $d = 2$:

$$V^{\otimes 2} \to S^2(V), \quad v^{\otimes 2} \mapsto \begin{bmatrix} a_0^2 & a_0a_1 & \cdots & a_0a_n \\ a_0a_1 & a_1^2 & \cdots & a_1a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_0a_n & a_1a_n & \cdots & a_n^2 \end{bmatrix}.$$

  This corresponds to the polynomial:

$$a_0^2x_0^2 + 2a_0a_1x_0x_1 + \cdots + a_n^2x_n^2 = (a_0x_0 + \cdots + a_nx_n)^2.$$

- For general d:

$$v^{\otimes d} \mapsto (a_0x_0 + \cdots + a_nx_n)^d.$$

Definition of Symmetric Rank: The symmetric rank of a polynomial F is the minimum r such that:

$$F = \sum_{i=1}^{r} L_i^d, \text{ where } L_i \text{ are linear forms.}$$

Veronese Variety

$$\mathbb{P}V \times \cdots \times \mathbb{P}V \to \mathbb{P}(V^{\otimes d}) = \{\text{Space of tensors}\}$$

$$([v_1], \ldots, [v_d]) \mapsto [v_1 \otimes \cdots \otimes v_d] = \{\text{Segre variety}\}$$

Veronese Variety

$$\mathbb{P}V \times \cdots \times \mathbb{P}V \to \mathbb{P}(V^{\otimes d}) = \{\text{Space of tensors}\}$$

$$([v_1], \ldots, [v_d]) \mapsto [v_1 \otimes \cdots \otimes v_d] = \{\text{Segre variety}\}$$

$$\nu_d : \mathbb{P}(V) \to \mathbb{P}(S^d V) = \{\text{Space of Symm tensors}\}$$

$$[v] \mapsto [v^{\otimes d}] = \{\text{Veronse variety}\}$$

Veronese Variety

$$\mathbb{P}V \times \cdots \times \mathbb{P}V \to \mathbb{P}(V^{\otimes d}) = \{\text{Space of tensors}\}$$

$$([v_1], \ldots, [v_d]) \mapsto [v_1 \otimes \cdots \otimes v_d] = \{\text{Segre variety}\}$$

$$\nu_d : \mathbb{P}(V) \to \mathbb{P}(S^d V) = \{\text{Space of Symm tensors}\}$$

$$[v] \mapsto [v^{\otimes d}] = \{\text{Veronse variety}\}$$

$$\nu_d(\mathbb{P}^n) = \text{Segre}(\mathbb{P}^n, d) \cap \mathbb{P}S^d V$$

Example

$$\nu_3 : \mathbb{P}^1 \to \mathbb{P}(S^3\mathbb{C}^2)$$

$$[v] = [1, t] \mapsto [(x + ty)^3] = [x^3 + 3tx^2y + 3t^2xy^2 + t^3y^3] = (1, t, t^2, t^3)$$

Example

$$\nu_3 : \mathbb{P}^1 \to \mathbb{P}(S^3\mathbb{C}^2)$$

$$[v] = [1, t] \mapsto [(x + ty)^3] = [x^3 + 3tx^2y + 3t^2xy^2 + t^3y^3] = (1, t, t^2, t^3)$$

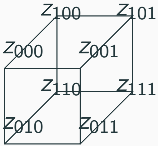
This is a general fact: the parameterization of the Veronese variety in 2 variables (= rational normal curve of degree d) is

$$\nu_d : \mathbb{P}^1 \to \mathbb{P}(S^d\mathbb{C}^2)$$

$$[v] = [1, t] \mapsto [(x + ty)^d] = (1, t, t^2, \ldots, t^d)$$

Equations of Segre variety

- Equations of rank-1 tensors (Segre variety) $=$ (2×2)-minors of flaktenings filled with variables



$$\begin{pmatrix} z_{000} & z_{001} & z_{100} & z_{101} \\ z_{010} & z_{011} & z_{110} & z_{111} \end{pmatrix}, \begin{pmatrix} z_{000} & z_{100} & z_{001} & z_{101} \\ z_{010} & z_{110} & z_{011} & z_{111} \end{pmatrix},$$

$$\begin{pmatrix} z_{000} & z_{100} & z_{010} & z_{110} \\ z_{001} & z_{101} & z_{011} & z_{111} \end{pmatrix}$$

Equation of Veronese variety

We have seen:

1. the set of tensors of $rk = 1$ is closed $\Rightarrow \exists$ equations defining exactly rk-1 tensors
2. Those equations are the (2×2)-minors of the flattenings of a generic tensor (filled with variables)
   $\Rightarrow$ What about rk-1 symmetric tensors?

Equation of Veronese variety

We have seen:

1. the set of tensors of $rk = 1$ is closed $\Rightarrow \exists$ equations defining exactly rk-1 tensors
2. Those equations are the (2×2)-minors of the flattenings of a generic tensor (filled with variables)
   $\Rightarrow$ What about rk-1 symmetric tensors?

$$\nu_d(\mathbb{P}(\mathbb{C}^n)) = \mathrm{Segre}(\mathbb{P}(\mathbb{C}^n) \times \cdots \times \mathbb{P}(\mathbb{C}^n)) \cap \mathbb{P}(S^d(\mathbb{C}^n))$$

$\{(2 \times 2) -$ minors of flattenings of generic tensors$\} \cap$
$\{$symmetric tensors$\} =$
$\{(2 \times 2) -$ minors of flattenings of generic symmetric tensors$\}$

Last time

$\nu_3(\mathbb{P}^1) = \text{Segre}(\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1) \cap \mathbb{P}(S^3\mathbb{C}^3)$

Coordinates $\{z_{000}, z_{001}, \ldots, z_{111}\}$

Imposing symmetries:

- $z_{001} = z_{010} = z_{100}$
- $z_{110} = z_{101} = z_{011}$

Coordinates $\{z_{000}, z_{001}, z_{011}, z_{111}\}$

eg. of one flattening: $\begin{pmatrix} z_{000} & z_{010} & z_{100} & z_{110} \\ z_{001} & z_{011} & z_{101} & z_{111} \end{pmatrix}$: the 2-minors of all flattenings give equation of Segre

eg. of a symmetric flattening: $\begin{pmatrix} z_{000} & z_{001} & z_{001} & z_{011} \\ z_{001} & z_{011} & z_{011} & z_{111} \end{pmatrix}$ the two minors of all symm. flattenings give equations of Veronese.

Let's check that the (2×2)-minors of

$$\begin{pmatrix} z_{000} & z_{001} & z_{001} & z_{011} \\ z_{001} & z_{011} & z_{011} & z_{111} \end{pmatrix} \tag{1}$$

are the equations of $\nu_3\left(\mathbb{P}^1\right) \ni F = (ax + by)^3$

- First notice that the 2^{nd} and the 3^{th} colums are equal as the 2-minors of (1) vanish if and only if the (2×2)-minors of $\begin{pmatrix} z_{000} & z_{001} & z_{011} \\ z_{001} & z_{011} & z_{111} \end{pmatrix}$ vanish

$$\mathbb{P}^1 \xrightarrow{\nu_3} \mathbb{P}\left(S^3\mathbb{C}^2\right) \quad \operatorname{Im}(\nu_3) = v_3\left(\mathbb{R}^2\right)$$
$$[1, t] \mapsto [1, t, t^2, t^3]$$

Coordinates on $\mathbb{P}\left(S^3\mathbb{C}^2\right)$ are $\{z_{000}, z_{001}, z_{011}, z_{111}\}$

$$\begin{pmatrix} z_{000} & z_{001} & z_{011} \\ z_{001} & z_{011} & z_{111} \end{pmatrix} \begin{pmatrix} 1 & t & t^2 \\ t & t^2 & t^3 \end{pmatrix}$$

$$\begin{cases} z_{000}z_{011} - z_{001}^2 = 0 \\ z_{000}z_{111} - z_{011}z_{001} = 0 \\ z_{001}z_{111} - z_{011}^2 = 0 \end{cases} \qquad \begin{cases} t^2 - t^2 = 0 \\ t^3 - t^3 = 0 \\ t^2 - t^2 = 0 \end{cases}$$

This proves that $\nu_3(\mathbb{P}^1) \subseteq V((2 \times 2) - \text{minors of } (1))$.

The prof for $\nu_d(\mathbb{P}^1)$ is straightforward.

In order to see the other containment $\nu_3\left(\mathbb{P}^1\right) \supseteq V((2 \times 2)$-minors of (1)=V) we need some algebraic geometry:

- $\dim \nu_d\left(\mathbb{P}^1\right) = 1 = \dim V$
- they are both irreducible

$$\Rightarrow \nu_3\left(\mathbb{P}^1\right) = V\left(\begin{array}{c} (2 \times 2)- \text{ minors of flattenings} \\ \text{of a symm. order-3} \\ \text{generic tensor} \end{array}\right)$$

Equations of the rational normal curve

This works in general:

$$\nu_d\left(\mathbb{P}^1\right) = V\left(\begin{array}{l}(2\times 2)-\text{ minors of}\\ \text{flattenings of}\\ \text{symmetric order-d}\\ \text{generic tensor}\end{array}\right)$$

$\mathbb{P}^1 \xrightarrow{\nu_d} \mathbb{P}\left(S^d\mathbb{C}^2\right)$, $[1,t] \longmapsto \left[1, t, t^2, t^3, \ldots, t^d\right]$

basis of degree d polynomials in 2 variables: $\{x^d, x^{d-1}y, \ldots, y^d\}$

$\nu_d : x + ty \mapsto (x+ty)^d = x^d + tx^{d-1}y + t^2x^{d-2}y^2 + \cdots + t^dy^d$

coordinates on $\mathbb{P}(\mathbb{C}^2) : \{z_1, \ldots, z_d\}$

$\mathbb{P}^1 \xrightarrow{\nu_d} \mathbb{P}\left(S^d\mathbb{C}^2\right)$, $[1,t] \longmapsto \left[1, t, t^2, t^3, \dots, t^d\right]$

basis of degree d polynomials in 2 variables: $\{x^d, x^{d-1}y, \dots, y^d\}$

$\nu_d : x + ty \mapsto (x+ty)^d = x^d + tx^{d-1}y + t^2x^{d-2}y^2 + \cdots + t^dy^d$

coordinates on $\mathbb{P}(\mathbb{C}^2) : \{z_1, \dots, z_d\}$

The flattening of symmetric tensors turn out to be:

$\begin{pmatrix} z_0 & z_1 & z_2 & \cdots & z_{d-1} \\ z_1 & z_2 & z_3 & \cdots & z_d \end{pmatrix}$. Evaluated on $\nu_d(\mathbb{P}^1)$ one gets
$\begin{pmatrix} 1 & t & t^2 & t^3 & \dots & t^{d-1} \\ t & t^2 & \dots & \dots & \dots & t^d \end{pmatrix}$ whose (2×2)-minors are

$$\begin{vmatrix} t^i & t^j \\ t^{i+1} & t^{j+1} \end{vmatrix} = t^{i+j+1} - t^{j+i+1} = 0$$

So $\nu_d(\mathbb{P}^1) \subseteq V$ The other containment is given by the previous argument of algebraic geometry.

- Equations of rational normal curve: (2×2)-minors of symmetric tensors of order d in 2 variables filled with variables (respecting symmetries).

Equations of Veronese variety

It is a general fact that the equations of any Veronese variety $\nu_d(\mathbb{P}^n)$ are the (2×2)-minors of all the flattenings of orderd-d symmetric tensors.

Equations of Veronese variety

Example: $\left(\nu_3\left(\mathbb{P}^2\right)\right) \ni (x + ty + uz)^3$

$$\mathbb{P}^2 \xrightarrow{\nu_3} \mathbb{P}\left(s^3\mathbb{C}^3\right)$$

$$[1, t, u] \longmapsto \left[1, t, u, t^2, tu, u^2, t^3, t^2u, tu^2, u^3\right]$$

$$[x + ty + uz]^3 \mapsto (x + ty + uz)^3 = x^3 + 3tx^2y + \cdots + u^3z^3$$

$$\begin{pmatrix} 000 & 001 & 002 & 011 & 012 & 022 \\ 001 & 011 & 012 & 111 & 121 & 122 \\ 002 & 012 & 021 & 121 & 122 & 222 \end{pmatrix}$$

$$\{x^3, x^2y, x^2z, xy^2, xyz, \quad xz^2, y^3, \quad y^2z, \quad yz^2, \quad z^3\}$$

$$\{z_0, z_1, z_2, z_3, z_4, z_5, z_6, z_7, z_8, z_9\}$$

$$(2 \times 2) - \text{minors of} \begin{pmatrix} z_0 & z_1 & z_2 & z_3 & z_4 & z_5 \\ z_1 & z_3 & z_4 & z_6 & z_7 & z_8 \\ z_2 & z_4 & z_5 & z_7 & z_8 & z_9 \end{pmatrix} = I\left(\nu_3\left(\mathbb{P}^2\right)\right)$$

$$\nu_d\left(\mathbb{P}^n\right) = V\left(\begin{array}{c}(2\times 2)\text{-minors of flattening}\\ \text{of order-d symmetric}\\ \text{tensor}\end{array}\right)$$

There is another characterization of those minors.

Catalecticant maps

Example: $\nu_3(\mathbb{P}^2) \subset \mathbb{P}(S^3\mathbb{C}^3) \ni F$

$$C_F^{2.1}\left(S^2 C^3\right)^* \longrightarrow S^1\mathbb{C}^3$$
$$\mathbb{C}\left[\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right]_2 \quad \mathbb{C}[x, y, z]$$
$$\partial \longmapsto \partial(F)$$

eg: $C_{x^3}^{2,1}\left(\frac{\partial}{\partial x}\frac{\partial}{\partial y}\right) = \left(\frac{\partial}{\partial x}\frac{\partial}{\partial y}\right)(x^3) = 0$

$$F = \underset{z_0}{ax^3} + \underset{z_1}{bx^2y} + \underset{z_2}{cx^2z} + \underset{z_3}{dxy^2} + \underset{z_4}{exyz} + \underset{z_5}{fxz^2} + \underset{z_6}{gy^3} + \underset{z_7}{hy^2z} + \underset{z_8}{iyz^2} + \underset{z_9}{ez}$$

$$C_F^{2,1} : (S^2\mathbb{C}^3)^* \to S^1\mathbb{C}^3$$

Fix the corresponding basis:
$B_1 = \{\partial/\partial x^2, \partial/\partial xy, \dots, \partial/\partial z^2\}$ and $B_2 = \{x, y, z\}$

$$\begin{aligned}
\partial/\partial x^2 &\longrightarrow 6z_0x + 2z_1y + 2z_2z = (6z_0, 2z_1, 2z_2)\\
\partial/\partial xy &\longmapsto 2z_1x + 2z_3y + z_4z = (2z_1, 2z_3, z_4)\\
\partial/\partial xz &\longrightarrow 2z_2x + 2z_5z + z_4y = (2z_2, z_4, 2z_5)\\
\partial/\partial y^2 &\longmapsto 2z_3x + 2z_7z + 6z_6y = (2z_3, 6z_6, 2z_7)\\
\partial/\partial yz &\longrightarrow z_4x + 2z_7y + 2z_8z = (z_4, 2z_7, 2z_8)\\
\partial/\partial z^2 &\longrightarrow 2z_5x + 2z_1y + 6z_9z = (2z_5, 2z_8, 6z_9)
\end{aligned}$$

Matrix representing $C_F^{2,1}$ in the choosen basis:

$$\begin{pmatrix} 6z_0 & 2z_1 & 2z_2 & 2z_3 & z_4 & 2z_3 \\ 2z_1 & 2z_3 & z_4 & 6z_6 & 2z_7 & 2z_3 \\ 2z_2 & 2z_4 & 2z_5 & 2z_7 & 2z_8 & 6z_3 \end{pmatrix}$$

modulo the coefficients, which can be fixed with the choice of the basis it coincides with the previous matrix obtained from the sym. flattenings

$$\begin{pmatrix} z_0 & z_1 & z_2 & z_3 & z_4 & z_5 \\ z_1 & z_3 & z_4 & z_6 & z_7 & z_8 \\ z_2 & z_4 & z_5 & z_7 & z_8 & z_9 \end{pmatrix}$$

Catalecticant Matrix representing $C_F^{2,2}$ in the chosen basis:

$$\begin{pmatrix} z_0 & z_1 & z_2 & z_3 & z_4 & z_5 \\ z_1 & z_3 & z_4 & z_6 & z_7 & z_3 \\ z_2 & z_4 & z_5 & z_7 & z_8 & z_3 \end{pmatrix}$$

Symmetric flattening :

$$\begin{pmatrix} z_0 & z_1 & z_2 & z_3 & z_4 & z_5 \\ z_1 & z_3 & z_4 & z_6 & z_7 & z_8 \\ z_2 & z_4 & z_5 & z_7 & z_8 & z_9 \end{pmatrix}$$

Catalecticant matrices

$$F \in S^d(\mathbb{C}^m), 0 \leqslant i \leqslant d$$

$$C_F^{i,d-i} : \mathbb{C}\left[\frac{\partial}{\partial x_1} \cdots \cdots \frac{\partial}{\partial x_m}\right]_i \longrightarrow \mathbb{C}[x_1, \dots x_n]_{d-i}$$

$$\partial \longrightarrow \partial(F)$$

- Linear map $\leftrightarrow$ Associated matrix:= catalecticant matrix.
- Catalecticant matrices are defined for all $i \in \{1, \dots \deg(F)\}$.
- They have different sizes in dependence on i
- the first catalecticant coincides with the first flattening:

$$(C_F^{1,d-1})_{i,j} = (\text{coeff. of the j-th monomial of } \partial F/\partial x_i =$$

$$= (\text{coeff. of the } (j+1)-\text{th. coeff. of } F)$$

- The first catalecticant allows to test if a polynomial has rank 1: i.e. $F = L^d$.

Algorithm for rank-1 poly's

- Given $F \in S^d V$
- compute the rk of $C_F^{1,d-1}$
- If it is $= 1$ then $\mathrm{rk}(F) = 1$ otherwise $\mathrm{rk}(F) > 1$

Example

$F = (x + y + z)^3$

$$\begin{aligned} &C_F^{1,2} : \mathbb{C}[\partial/\partial x, \partial/\partial y, 0/\partial z]_2 \to \mathbb{C}[x, y, z]_2 \\ &\partial/\partial x \longrightarrow 3(x + y + z)^2 \\ &\partial/\partial y \longmapsto 3(x + y + z)^2 \\ &\partial/\partial z \longmapsto 3(x + y + z)^2 \end{aligned}$$

$$C_F^{1,2} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

$rk(C_F^{1,2}) = 1.$

Example

$F = x^3 + y^3 + z^3$

$$C_F^{1,2} : \mathbb{C}[\partial/\partial x, \partial/\partial y, \partial/\partial z]_1 \to \mathbb{C}[x, y, z]_2$$
$$\partial/\partial x \longmapsto 3x^2$$
$$\partial/\partial y \longrightarrow 3y^2$$
$$\partial/\partial z \longrightarrow 3z^2$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$ $\mathrm{rk}(C_F^{1,2}) = 3 \neq 1$ so $F = x^3 + y^3 + z^3 \neq L^3$.

Example

$F = x^3 + y^3 + z^3$

$$\begin{aligned} C_F^{1,2} &: \mathbb{C}[\partial/\partial x, \partial/\partial y, \partial/\partial z]_1 \to \mathbb{C}[x,y,z]_2 \\ \partial/\partial x &\longmapsto 3x^2 \\ \partial/\partial y &\longrightarrow 3y^2 \\ \partial/\partial z &\longrightarrow 3z^2 \end{aligned}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathrm{rk}(C_F^{1,2}) = 3 \neq 1 \text{ so } F = x^3 + y^3 + z^3 \neq L^3.$$

The (2×2)-minors of a catalecticaunt matrix characterize rank-1 polynomials.

What about the (3×3)-minors?

In there any relation with the rank-2 polynomials?

Two variables case

Example: $F = x^4 + y^4$

$$C_F^{1,3} : \mathbb{C}[\partial/\partial x, \partial/\partial y]_1 \to \mathbb{C}[x, y]_3$$

$$\partial/\partial x \mapsto 4x^3$$

$$\partial/\partial y \mapsto 4y^3$$

$$C_F^{1,3} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}$$

$\mathrm{rk}(C_F^{1,2}) = 2$ but NO (3×3)-minors.

Two variables case

$F = x^4 + y^4$

$$C_F^{2,2} : \mathbb{C}[\partial/\partial x, \partial/\partial y]_2 \to \mathbb{C}[x,y]_2$$

$$\partial/\partial x^2 \mapsto 12x^2$$

$$\partial/\partial x \partial y \mapsto$$

$$\partial y^2 \mapsto 12y^2$$

$$C_F^{2,2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$\mathrm{rk}(C_F^{2,2}) = 2$ and the (3×3)-minors vanish on that matrix.

Two variables cases

Fact 1 : The existence of the (3×3)-minors is just a matter of size, in fact for all fixed $F \in S^d\mathbb{C}^n$ all its catalecticant have the same rank (if they are big enough to show the maximum).

Two variables cases

Fact 1 : The existence of the (3×3)-minors is just a matter of size, in fact for all fixed $F \in S^d\mathbb{C}^n$ all its catalecticant have the same rank (if they are big enough to show the maximum).

Example: $F = x^6 + y^6 + (x+y)^6$

$$\begin{aligned} C_F^{1\cdot 5} : \mathbb{C}\left[\partial_x, \partial_y\right]_1 &\longrightarrow \mathbb{C}[x,y]_5 \\ \partial_x &\longrightarrow 6x^5 + 6(x+y)^5 \\ \partial_y &\longrightarrow 6y^5 + 6(x+y)^5 \end{aligned}$$

$C_F^{1,5}$ has only two (l.i.) columns $\Rightarrow$ it's rank is 2 (it cannot be higher)

$$C_f^{1,5} = \begin{pmatrix} \vdots & \vdots \\ \vdots & \vdots \end{pmatrix}$$

Two variables case

Example: $F = x^6 + y^6 + (x+y)^6$

$$C_F^{2,4} : \mathbb{C}\left[\partial_x, \partial_y\right]_2 \longrightarrow \mathbb{C}[x,y]_4$$
$$\partial_x^2 \longmapsto 30x^2 + 30(x+y)^2$$
$$\partial_{xy} \longmapsto 30x^2 + 60xy + 30y^2$$
$$\partial y^2 \longmapsto 30y^2 + 30(x+y)^2$$

$\mathrm{rk}(C_F^{2,4}) = 3 > 2$ (no (4×4)-minors)

$$C_F^{2,4} = \begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$$

Two variables case

Example: $F = x^6 + y^6 + (x+y)^6$

$$\begin{aligned} C_F^{3,3} : \mathbb{C}\left[\partial_x, \partial_y\right]_3 &\longrightarrow \mathbb{C}[x,y]_3 \\ \partial_{x^3} &\longmapsto 120x^3 + 120(x+y)^3 \\ \partial_{x^2y} &\longmapsto 120(x+y)^3 \\ \partial_{xy^2} &\longmapsto 120(x+y)^3 \\ \partial_{y^3} &\longmapsto 120y^3 + 120\left(x+y\right)^3 \end{aligned}$$

$\mathrm{rk}(C_F^{3,3}) = 3$. Big enough to

- have rank 3
- see the (4×4)-minors

$\mathrm{rk}(C_F^{2,4}) = 3$. Big enough to

- have rank 3
- but not to see the (4×4)-minors

$\mathrm{rk}(C_F^{1,5}) = 2$. Too small to have rank 3.

Two variables case

Fact: The existence of the (3×3)-minors is just a matter of size, in fact for all fixed $F \in S^d\mathbb{C}^n$ all its catalecticant have the same rank (if they are big enough to show the maximum).

The rank of the catalecticants stabilizes as soon as the size of the catalecticant is big enough.

If we don't know in advance the rank of $C_F^{i,d-i}$ and we are interested in it, it is more convenient to write the "biggest" (the most square one) catalecticant and check its rank: $C_F^{\lfloor d/2 \rfloor, \lceil d/2 \rceil}$

Two variables case

Fact: The existence of the (3×3)-minors is just a matter of size, in fact for all fixed $F \in S^d\mathbb{C}^n$ all its catalecticant have the same rank (if they are big enough to show the maximum).

The rank of the catalecticants stabilizes as soon as the size of the catalecticant is big enough.

If we don't know in advance the rank of $C_F^{i,d-i}$ and we are interested in it, it is more convenient to write the "biggest" (the most square one) catalecticant and check its rank: $C_F^{\lfloor d/2 \rfloor, \lceil d/2 \rceil}$

Question: We have already seen that the vanishing of the (3×3)-minors of the catalecticant $\Rightarrow \mathrm{brk}(F) \leq 2$. What can we say about the rank?

Two variables cases

Ex: $F = x^4 + y^4$

$C_F^{2,2} : \mathbb{C}[\partial_x, \partial_y]_2 \to \mathbb{C}[x,y]_2$

$\partial_{x^2} \mapsto 12x^2$

$\partial_{xy} \mapsto 0$

$\partial_{y^2} \mapsto 12y^2$

$$C_F^{2,2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$rk(C_F^{2,2}) = 2$$

The rank of F is clearly 2.

Ex: $W = x^3 y$

$C_G W^{2,2} : \mathbb{C}[\partial_x, \partial_y]_2 \to \mathbb{C}[x,y]_2$

$\partial_{x^2} \mapsto 6xy$

$\partial_{xy} \mapsto 3x^2$

$\partial_{y^2} \mapsto 0$

$$C_W^{2,2} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$rk(C_W^{2,2}) = 2$$

What about W?

Two variables case

Claim $W = x^3 y \in T_{[x^4]}\nu_4\left(\mathbb{P}^4\right)$ Which is the structure of $T_{x^4}C_4 =$?

$$\mathbb{P}^1 \xrightarrow{\nu_4} \mathbb{P}\left(S^4\mathbb{C}^2\right)$$

$$(x + tM) \longmapsto (x + tM)^4$$

$$\lim_{t\to 0} \frac{d}{dt}(x + tM)^4 = \lim_{t\to 0} 4M(x + tM)^3 = 4Mx^3$$

This shows that all the points on the tangent line to $\nu_4(\mathbb{P}^1)$ at $[x^4]$ are of type x^3M where M in any binary linear form:

$$T_{[x^4]}(C_4) = \langle x^3M,\ M \in \mathbb{C}[x, y]_1\rangle$$

So our original $W = x^3 y \in T_{[x^4]}\nu_4\left(\mathbb{P}^4\right)$

Two variables case

Claim $W = x^3y \in T_{[x^4]}\nu_4\left(\mathbb{P}^4\right)$ Which is the structure of $T_{x^4}C_4 =$?

$$\mathbb{P}^1 \xrightarrow{\nu_4} \mathbb{P}\left(S^4\mathbb{C}^2\right)$$

$$(x + tM) \longmapsto (x + tM)^4$$

$$\lim_{t\to 0} \frac{d}{dt}(x + tM)^4 = \lim_{t\to 0} 4M(x + tM)^3 = 4Mx^3$$

This shows that all the points on the tangent line to $\nu_4(\mathbb{P}^1)$ at $[x^4]$ are of type x^3M where M in any binary linear form:

$$T_{[x^4]}(C_4) = \langle x^3M,\ M \in \mathbb{C}[x, y]_1\rangle$$

So our original $W = x^3y \in T_{[x^4]}\nu_4\left(\mathbb{P}^4\right)$

Both $F = x^4 + y^4$ and $W = x^3y$ have border rank 2, but $F \in \sigma_2^0(C_4)$ and $W \in T_{[x^4]}(C_4)$

If the rank-2 polynomials in 2 variables kill the (3×3)-minors of the catalecticants (not proved yet) then any border rank 2 polynomial will be killed by the (3×3)-minors.

Apolarity

Definition
The Apolarity is the perfect pairing:

$$\mathbb{C}[\partial_1, \dots, \partial_n]_d \times \mathbb{C}[x_1, \dots, x_n]_d \to \mathbb{C}$$

$$(\partial, F) \mapsto \partial F$$

It allows to consider

$$\mathbb{C}[\partial_1, \dots, \partial_n]_d \simeq (\mathbb{C}[x_1, \dots, x_n]_d)^*$$

Definition
$g \in \mathbb{C}[\partial_1, \dots, \partial_n]_k$ is apolar to $F \in \mathbb{C}[x_1, \dots, x_n]_d$ if $k \leq d$ and $g(F) = 0$.

Example
$L = 2x + y$, $F = L^3$, $g = \partial^3/\partial x^3$, $g(F) \neq 0$ so g is not apolar to F.
$\tilde{g} = 2\partial_y - \partial_x$, $\tilde{g}(F) = 6(2x + y)^2 - 6(2x + y)^2 = 0$ so $\tilde{g}$ is apolar to F.

Apolarity

Definition
Let $F \in S^d\mathbb{C}^n$. The annihilator of F is

$$F^{\perp} := \{\, g \in (S\mathbb{C}^n)^* \,|\, g(F) = 0\}.$$

Example
$L = 2x + y$, $F = L^3$, $\tilde{g} = 2\partial_y - \partial_x \in F^{\perp}$. So $\tilde{g}M \in F^{\perp}$ for all forms M, i.e. $(\tilde{g}) \subseteq F^{\perp}$.

Let $g = a\partial_x + b\partial_y \in F^{\perp}$, so $g(F) = 0$, i.e. $(a\partial_x + b\partial_y)(2x + y)^3 = 0$, $6a(2x + y)^2 + 3b(2x + y)^2 = 0$, $(6a + 3b)(2x + y)^2 = 0$ iff $b = -2a$ therefore the only possibility for g is $g = \tilde{g}$. Therefore $F^{\perp} \subseteq (\tilde{g}) \Rightarrow$

$$F^{\perp} = (\tilde{g}) = I([1, 2])$$

Apolarity

Definition
Let $F \in S^d\mathbb{C}^n$. The annihilator of F is

$$F^\perp := \{\, g \in (S\mathbb{C}^n)^* \mid g(F) = 0\}.$$

Example
$L = 2x + y$, $F = L^3$, $\tilde{g} = 2\partial_y - \partial_x \in F^\perp$. So $\tilde{g}M \in F^\perp$ for all forms M, i.e. $(\tilde{g}) \subseteq F^\perp$.

Let $g = a\partial_x + b\partial_y \in F^\perp$, so $g(F) = 0$, i.e. $(a\partial_x + b\partial_y)(2x + y)^3 = 0$, $6a(2x + y)^2 + 3b(2x + y)^2 = 0$, $(6a + 3b)(2x + y)^2 = 0$ iff $b = -2a$ therefore the only possibility for g is $g = \tilde{g}$. Therefore $F^\perp \subseteq (\tilde{g}) \Rightarrow$

$$F^\perp = (\tilde{g}) = I([1, 2])$$

Remark: The linear form defined by $I([1, 2]) = F^\perp$ is precisely L.

This is not a chance!

Apolarity

Proposition: If $F = L^d = (x_0 + a_1 x_1 + \cdots + a_n x_n)^d$, then
$F^\perp \supseteq I([1, a_1, \ldots, a_n]^*) = I(L)$

We can prove an even more general fact:

Theorem (Apolarity Lemma)

$$F = \lambda_1 L_1^d + \cdots + \lambda_r L_r^d \Leftrightarrow I([L_1], \ldots, [L_r]) \subseteq F^\perp$$

for $\lambda_i \in \mathbb{C}$ and $L_i \in S^1\mathbb{C}$.

Proof.
If $F = \sum_{i=1}^r \lambda_i L_i^d$, take $g \in I([L_1], \ldots, [L_r])$, such a g kills all the L_i's, so g kills $\sum_{i=1}^r \lambda_i L_i^d$, i.e. $g \in F^\perp$.

Viceversa: $\langle L_1^d, \ldots, L_r^d \rangle \subseteq \ker(L_1^\perp \circ \cdots \circ L_r^\perp)$. And the showing that they have the same dimension. $\square$

Apolarity

Theorem (Apolarity Lemma)

$$F = \lambda_1 L_1^d + \cdots + \lambda_r L_r^d \Leftrightarrow I([L_1], \ldots, [L_r]) \subseteq F^\perp$$

for $\lambda_i \in \mathbb{C}$ *and* $L_i \in S^1\mathbb{C}$.

Since $F^\perp = \{g \in (S\mathbb{C}^n)^*\}$, we have an algebraic way of computing $F^\perp$, i.e. via

Apolarity

Theorem (Apolarity Lemma)

$$F = \lambda_1 L_1^d + \cdots + \lambda_r L_r^d \Leftrightarrow I([L_1], \ldots, [L_r]) \subseteq F^\perp$$

for $\lambda_i \in \mathbb{C}$ *and* $L_i \in S^1\mathbb{C}$.

Since $F^\perp = \{g \in (S\mathbb{C}^n)^*\}$, we have an algebraic way of computing $F^\perp$, i.e. via

$$\ker(C_f^{i,d-i})$$

Apolarity

Theorem (Apolarity Lemma)

$$F = \lambda_1 L_1^d + \cdots + \lambda_r L_r^d \Leftrightarrow I([L_1], \ldots, [L_r]) \subseteq F^\perp$$

for $\lambda_i \in \mathbb{C}$ *and* $L_i \in S^1\mathbb{C}$*.*

Since $F^\perp = \{g \in (S\mathbb{C}^n)^*\}$, we have an algebraic way of computing $F^\perp$, i.e. via

$$\ker(C_f^{i,d-i})$$

Corollary:

$$F = \lambda_1 L_1^d + \cdots + \lambda_r L_r^d \Leftrightarrow I([L_1], \ldots, [L_r]) \subseteq \ker(C_F^{i,d-i})$$

for $\lambda_i \in \mathbb{C}$ and $L_i \in S^1\mathbb{C}$.

Apolarity

This shows that:

1. Either the rank of $C_F^{i,d-i}$ is maximum and $\ker(C_F^{i,d-i}) = \{0\}$, or the rank of all $C_F^{i,d-i}$ are all equal after certain index.
2. We have a way of computing the rank and a decomposition of a given F.

Apolarity

Example

$F = (x+y)^3 + (x-y)^3 = 2x^3 + 6xy$

$C_F^{2,1} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$

$\ker(C_F^{2,1}) = \langle(-1,0,1)\rangle \subset \mathbb{C}[\partial_x, \partial_y]_2 = \mathbb{C}[u_0, u_1]_2$. The vector $(-1,0,1)$ corresponds to the form $-u_0^2 + u_1^2 = (u_0 - u_1)(u_0 + u_1)$. The ideal $I((u_0 - u_1)(u_0 + u_1))$ defines the variety $V(I) = \{[1,1],[1,-1]\}$ from which we get $L_1 = x + 1$ and $L_2 = x - y$. So

$$F = \lambda_1(x+y)^2 + \lambda_2(x-y)^2.$$

In order to find λ_1 and λ_2 one simply has to solve a linear system.

Apolarity

Let's study an example that we already know it can be more subtle.

Example

$F = x^4 y \in T_{[x^5]} C_5$ (we already know that $F \in \sigma_2(C_5)$ but its rank is bigger than 2).

$$C_F^{2,3} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},\ rk(C_F^{2,3}) = 2 \text{ and } \ker(C_F^{2,3}) = \langle (0,0,1) \rangle \text{ which}$$

corresponds to $u_1^2 = 0$. The ideal $I = (u_1^2)$ does not define $2 = rk(C_F^{2,3})$ distinct points: $V(I) = \{[1,0]\}$ counted with multiplicity 2.

We already knew that this example could have been problematic. Let's append it for a while and let's focus on equations of $\sigma_2(C_d)$.

Equations of $\sigma_2(C_d)$

If $P \in \sigma_2(C_d)$ then

1. either $P \in C_d \Rightarrow rk(P) = 1$,
2. either $P \in \sigma_2^0(C_d) \Rightarrow rk(P) = 2$,
3. or $P \in \sigma_2(C_d) \setminus \sigma_2^0(C_d) \Rightarrow P = L^{d-1}M$.

1. If $rk(P) = 1 \Rightarrow rk(C_P^{i,d-i} = 1)$,
2. If $rk(P) = 2 \Rightarrow P = L_1^d + L_2^d$, if we choose the basis of $\mathbb{C}[x,y]_d$ to be $\{L_1^d, L_1^{d-1}L_2, \ldots, L_2^d\}$, then in that basis $C_P^{2,d-2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 1 \end{pmatrix}$ and $rk(C_P^{2,d-2}) = 2$,
3. If $P = L_1^{d-1}L_2$, if we choose the same basis as above, then $C_P^{2,d-2} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \end{pmatrix}$ and $rk(C_P^{2,d-2}) = 2$,

Equations of $\sigma_2(C_d)$

We have just proved that for all possible elements of $\sigma_2(C_d)$, the $rk(C_P^{2,d-2}) \leq 1$. This proves that

$$\sigma_2(C_d) \subseteq V(3 \times 3 - \text{minors of } C_P^{2,d-2}).$$

Equations of $\sigma_2(C_d)$

We have just proved that for all possible elements of $\sigma_2(C_d)$, the $rk(C_P^{2,d-2}) \leq 1$. This proves that

$$\sigma_2(C_d) \subseteq V(3 \times 3 - \text{minors of } C_P^{2,d-2}).$$

Also the other inclusion is true (not difficult).

Theorem
$I(\sigma_2(C_d)) = (3 \times 3 - \textit{minors of } C_P^{2,d-2}).$

So the (3×3)-minors of a catalecticant (in the 2-variables case) reveals if the border rk of a given polynomial is 2.

Theorem
$I(\sigma_r(C_d)) = ((r+1) - \textit{minors of } C_P^{\lfloor d/2 \rfloor, \lceil d/2 \rceil}).$

You can also choose the smallest catalecticant where the $(r+1)$-minors can be computed.

Questions

$$\sigma_r(C_d) = \overline{\cup_{P_1,\ldots,P_r \in C_d} \langle P_1, \ldots, P_r \rangle}$$

1. The rank of the catalecticant tells in which $\sigma_r(C_d)$ a given $F \in S^d\mathbb{C}^2$ belongs.
2. How can we distinguish if F either has rk$= r$ or if $F \in \sigma_r(C_d) \setminus \sigma_r^0(C_d)$?
3. How to compute the rk and a decomposition of such an F?

Questions

Question 2. can be answered thanks to Apolarity Lemma.

Once we know the border rank r of F (via rk of catalecticant) we only need to understand if either the rk of F is actually r as the border rk or not.

By Apolarity Lemma F has rank r if and only if there exists an ideal of r distinct simple points inside $F^\perp$.

If not, then $F \in \sigma_r(C_d) \setminus \sigma_r^0(C_d)$ and therefore $rk(F) > r$.

Sylvester Algorithm (original version)

Input: $F \in S^d\mathbb{C}^2$.

Output: A symm-rk decomposition of F.

1. Initialize with $i = 1$,
2. Compute $C_F^{i,d-i}$,
3. Compute $rk(C_F^{i,d-i}) = r$, if $r = i$ then $F \in \sigma_r(C_d)$, otherwise restart with $i + 1$.
   (We already know that if we started with $C_F^{\lfloor d/2 \rfloor, \lceil d/2 \rceil}$ we immediately get the correct starting rank.)
4. Compute $K := \ker(C_F^{i,d-i})$
5. If $K \supseteq$ an ideal of r distinct points $L_1, \ldots, L_r$, done: $F = \sum_{i=1} \lambda_i L_i^d$ for certain $\lambda \in \mathbb{C}$, otherwise re-start from 4 with $r = i + 1$.

The procedure will finish because of the Apolarity Lemma.

Tangential example

Example
$F = x^4 y$.

We already know that $rk(C_F^{2,3}) = rk(C_F^{3,2}) = 2$ but their kernels do not define 2 distinct points.

$C_F^{4,1} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}$,
$\ker(C_F^{4,1}) = \langle (0,0,1,0,0), (0,0,0,1,0), (0,0,0,0,1) \rangle$. A generic element of the kernel is $u_0^2 u_1^2 + x_0 u_1^3 + u_1^4 = u_1^2(u_0^2 + u_0 u_1 + u_1^2)$ which does not have 4 distinct solutions.

$C_F^{5,1} = (0,1,0,0,0)$, $\ker(C_F^{5,1}) = \langle e_0, e_2, e_3, e_4, e_5 \rangle$. A generic element of the kernel is of type $u = u_0^5 + u_0^3 u_1^2 + u_0^2 u_1^3 + u_0 u_1^4 + u_1^5$ for which we can compute 5 distinct numerical roots: $(-0.31, -i1.05)$, $(-0.31, +i1.05)$, $(0.44, i0.68)$, $(0.44, -i0.68)$, $(-1, 249)$. So

$$rk(x^4 y) = 5.$$

There is a more efficient algorithm, but before doing it we will understand the behaviour of $rk(x^{d-1}y) \in T_{[x^d]}(C_d)$ and more in general of the points $\sigma_r(C_d) \setminus \sigma_r^0(C_d)$.

Rank of points on tangent lines to rational normal curves

Example (already seen)

$$T = e_1 \otimes e_1 \otimes e_2 + e_1 \otimes e_2 \otimes e_1 + e_2 \otimes e_1 \otimes e_1$$

$$T \in T_{[e_1 \otimes e_1 \otimes e]} \,\mathrm{Segre}\left(\left(\mathbf{C}^2 \right)^{\times 3} \right)$$

We saw that $rk(T) = 3$
$e_1 \otimes e_1 \otimes e_1$ is symmetic $\longleftrightarrow x^3$
T is symmetric $\longleftrightarrow x^2 y$:

$$e_1 \otimes e_1 \otimes e_2 + e_1 \otimes e_2 \otimes e_1 + e_2 \otimes e_1 \otimes e_1$$

$$x \cdot x \cdot y + x \cdot y \cdot x + y \cdot x \cdot x$$

We have already seen that $x^2 y \in T_{[x^3]} C_3 \subset \sigma_2(C_3) \setminus C_3$ so $\mathrm{rk} x^2 y \geq 2$.

Which is the symmetric rank of $x^2 y$?

We look for an ideal of at least 2 points in $(x^2 y)^{\perp}$, so we can start from the catalecticant whose kernel is contained in a space of polynomials of degree at least 2.

Rank of points on tangent lines to rational normal curves

$F = x^2 y$

$$\begin{aligned} C_F^{2,1} :& \mathbb{C}\left[\partial_x, \partial_y\right]_2 \to \mathbb{C}[x,y]_1 \\ & \partial_{x^2} \mapsto 2y \\ & \partial_{xy} \mapsto 2x \\ & \partial_{y^2} \mapsto 0 \end{aligned} \qquad C_F^{2,1} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

$\ker(C_F^{2,1}) = \langle (0,0,1) \rangle$, which in the basis $\{u_0^2, u_0 u_1, u_1^2\}$ becomes $u_1^2 = 0$ which root is $[(1,0)]$: NO 2 distinct roots so $\mathrm{rk}(x^2 y) > 2$.

Rank of points on tangent lines to rational normal curves

$F = x^2 y$

$$C_F^{2,1} : \mathbb{C}\left[\partial_x, \partial_y\right]_2 \to \mathbb{C}[x, y]_1$$

$$\partial_{x^2} \mapsto 2y$$

$$\partial_{xy} \mapsto 2x \qquad\qquad C_F^{2,1} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

$$\partial_{y^2} \mapsto 0$$

$\ker(C_F^{2,1}) = \langle(0,0,1)\rangle$, which in the basis $\{u_0^2, u_0 u_1, u_1^2\}$ becomes $u_1^2 = 0$ which root is $[(1,0)]$: NO 2 distinct roots so $\mathrm{rk}(x^2 y) > 2$.

$$C_F^{3,0} : \mathbb{C}\left[\partial_x, \partial_y\right]_3 \to \mathbb{C}$$

$$\begin{pmatrix} 0 & 1 & 0 & 0 \end{pmatrix} \Rightarrow \ker C_F^{3,0} = \langle(1,0,0,0),(0,0,1,0),(0,0,0,1)\rangle$$

A generic element in the kernel is $u = u_0^3 + u_0 u_1 + u_1^3$. We may look for solutions of type $[1, u_1]$, so $u = 1 + u_1 + u_1^3$ is a cubic polynomial: 3 complex solutions and they are all different because it does not have quadratic factors, so $rk(x^2 y) = 3$.

Rank of points on tangent lines to rational normal curves

We can see that rk $\left(x^2y\right) = 3$ also in a more geometric way.

$$\begin{aligned} \nu_3 : \mathbb{P}^1 &\to \mathbb{P}\left(S^3\mathbb{C}^2\right) \\ [1, t] &\longmapsto \left[t^3, t^2, t, 1\right] \end{aligned}$$

Definition
The *degree* of a curve is the number of points in the intersection of the curve with a generic hyperplane.

Proposition: $Im(\nu_3) = C_3$ is a curve of degree 3.

Proof.

$$\begin{aligned} &\mathbb{P}^3 \supset H = V\left(a_0 z_0 + \cdots + a_3 z_3\right) : \text{ hyperplane} \\ &H \cap C_3 : a_0 t^3 + a_2 t^2 + a_2 t + a_3 = 0 \end{aligned}$$

Polynomial of degree 3 over the complex numbers $x \Rightarrow 3$ solutions
H generic: the solutions are distinct

$$\Rightarrow \deg\left(C_3\right) = 3$$

Rank of points on tangent lines to rational normal curves

More generally:

Proposition: $Im(\nu_d) = C_d$ is a curve of degree d.

Proof.
$\mathbb{P}^d \supset H = V\left(a_0 z_0 + \cdots + a_d z_d\right) =$ hyperplane
$H \cap C_d : a_0 t^d + a_1 t^{d-1} + \cdots + a_d = 0$ poly of degree d over the complex
$\Rightarrow d$ solutions
H generic : solutions are distinct $\Rightarrow \deg\left(C_d\right) = d$ $\square$

Rank of points on tangent lines to rational normal curves

$F = x^2 y \in T_{[x^3]} C^3$,
$\deg C_3 = 3 \Rightarrow \exists H \simeq \mathbb{P}^2 = \langle P_1, P_2, P_3 \rangle$, $P_i \in C_3$ s.t. $F \in H$, so
$F = \alpha_1 P_1 + \alpha_2 P_2 + \alpha_3 P_3 = \alpha_1 L_1^3 + \alpha_2 L_2^3 + \alpha_3 L_3^3$.
This only proves that $rk(F) \leq 3$ but we already know that $rk(F) > 2$ so $rk(F) = 3$.

$F = x^{d-1} y \in T_{[x^d]} C^d$,
$\deg C_d = d \Rightarrow \exists H \simeq \mathbb{P}^d = \langle P_1, \dots, P_d \rangle$, $P_i \in C_d$ s.t. $F \in H$, so
$F = \alpha_1 P_1 + \cdots + \alpha_d P_d = \alpha_1 L_1^d + \cdots + \alpha_d L_d^d$.
This only proves that $rk(F) \leq d$.
To prove the $=$ we need another argument.

Rank of points on tangent lines to rational normal curves

Proposition: $rk(x^{d-1}y) = d$

Proof. We have already proved that $rk(F) \leq d$

Suppose that $\mathrm{rk}(F) < d$

$\Rightarrow \quad \exists L_1, \ldots, L_{d-1}$, linear forms s.t $F \in \underbrace{T_{[x^d]} C_d}_{\text{line}} \cap \underbrace{\langle [L_1^d], \ldots, [L_{d-1}^d] \rangle}_{\mathbb{P}^{d-2}}$.

In $\mathbb{P}^d$ a line intersects a $\mathbb{P}^{d-2}$ if and only if they are linearly dependent, i.e. $[x^d], [(x+\epsilon y)^d]$ together with $[L_1^d], \ldots, [L_{d-1}^d]$ are linearly dependent.

But these are $d - 1 + 2 = d + 1$ points on C_d which cannot be linearly independent. Which is a contradiction $(\rightarrow)$

Rank of points on tangent lines to rational normal curves

Claim: Any choice of $d+1$ points on a C_d is linearly independent.

Proof.
$P_1, \ldots, P_d \in C_d \Rightarrow P_i = [t_i^d, t_i^{d-1}, \ldots, t, 1],\ i = 1, \ldots, d+1$

$$\det \begin{pmatrix} t_1^d & t_1^{d-1} & \cdots & t_1^1 & 1 \\ t_2^d & t_2^{d-1} & \cdots & t_2^1 & 1 \\ \vdots & & & & \vdots \\ t_{d+1}^d & t_{d+1}^{d-1} & \cdots & t_{d+1}^1 & 1 \end{pmatrix} \neq 0$$

This is a Vandermonde. Equivalently $P_1, \ldots, P_{d+1}$ are linearly independent. $\square$

Summing up

- Belonging to $\sigma_2(C_d)$ (i.e. $\mathrm{brk}(F) = 2$)) is equivalent to the vanishing of the (3×3)-minors.
- If also the (2×2)-minors vanish, this is equivalent to $F \in C_d$, i.e. $rk(F) = 1$;
- $F \in \sigma_2^0(C_d)$ is equivalent to $rk(F) = 2$ and this can be distinguished from the next case by finding a polynomial in $\ker(C_F^{2,d-2})$ with 2 distinct roots;
- $F \in T_Q(C_d)$ is equivalent to the fact that $rk(F) = d$, so if one is interested in the decomposition, one has to go to elements of $\ker(C_F^{d,0})$.

Summing up

For points of border rank ≤ 2 we have an efficient procedure to compute their rank and decomposition.

Algorithm:
Input: $F \in S^d\mathbb{C}^2$.
Output: $rk(F)$ and L_i's s.t. $F = \sum_{i=1}^r \lambda_i L_i^d$, $\lambda_i \in \mathbb{C}$.

1. Compute the rank of $C_F^{\lceil d/2 \rceil, \lfloor d/2 \rfloor} := C$,
2. If $rk(C) = 1$, then $rk(F) = 1$. Pic $u \in ker(C_F^{1,d-1}) \subset S^1\mathbb{C}^2$, $V(u) = L$, so $F = L^d$.
3. If $rk(C) = 2$ then the border rank of F is 2.
    3.1 Pic $u \in \ker(C^{2,d-2}) \subset S^2\mathbb{C}^2$. If u has 2 distinct solutions L_1, L_2 then $F = \lambda_1 L_1^d + \lambda_2 L_2^d$ for certain $\lambda_i \in \mathbb{C}$, otherwise
    3.2 $\mathrm{rk}(F) = d$. Pic $u \in \ker(C_F^{d,0})$, such an u will have d distinct roots $L_1, \ldots, L_r$ which will give rise to $F = \sum_{i=1}^r \lambda_i L_i^d$for certain $\lambda_i \in \mathbb{C}$.
4. If $rk(C) > 2$ then $F \notin \sigma_2(C_d)$.

Rank of binary forms with given border rank

Theorem
Let $F \in S^d\mathbb{C}^2$ be of border rank r, then $rk(F) = r, d - r + 2$.

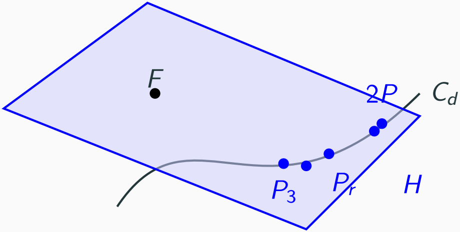
Proof. $brk(F) = r \Rightarrow F \in \sigma_r(C_d) \setminus \sigma_{r-1}(C_d) \Rightarrow rk(F) \geq r$.

If $\sigma_r^0(C_d)$ then the $rk(F) = r$.

Let's assume that $F \in \sigma_r(C_d) \setminus \sigma_r^0(C_d)$. This implies

- $rk(F) > r$,
- there exists a $\mathbb{P}^{r-1} := H$ limit of a sequence of $\mathbb{P}^{r-1}$'s which are r-secant to C_d.

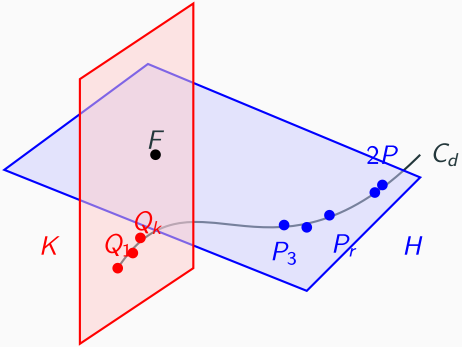
Since H is not r-secant ($\sharp(H \cap C_d) < r$ without multiplicity) and since H is limit of r-secant spaces $\sharp(H \cap C_d) = r$ counted with multiplicities.
Set $Z := H \cap C_d$. Z contains at least one point of multiplicity 2: let's call it $2P$: $Z = \langle 2P, P_3, \ldots, P_r \rangle$ with $P, P_3, \ldots, P_r \in C_d$



Rank of binary forms with given border rank

Let $rk(F) = k$, so there exists $Q_1, \ldots, Q_k \in C_d$ s.t. $F \in \langle Q_1, \ldots, Q_k \rangle = \mathbb{P}^{k-1} := K$. Since $F \in H \cap K$, the linear spaces H, K have to be l.d., i.e. $S := \{2P, P_3, \ldots, P_r, Q_1, \ldots, Q_k\} \subset C_d$ are l.d. Therefore, all together (counted with multiplicity) they have to be $\sharp S \geq d + 2$. Since the $\sharp\{P, P_i\} = r$, one has that $\sharp\{Q_i\} \geq d - r + 2$. Therefore

$$rk(F) \geq d - r + 2.$$

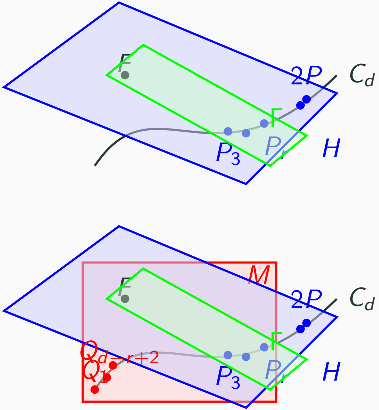


We are left with the proof of the other inequality: $rk(F) \leq d - r + 2$. A way of doing it is to exhibit exactly $d - r + 2$ points of C_d whose span contains F.

Rank of binary forms with given border rank

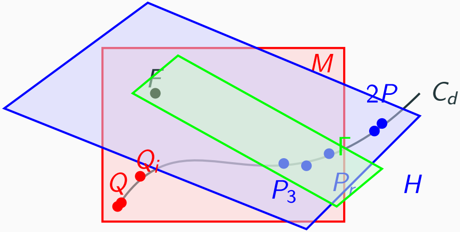
Let's now prove that
$rk(F) \leq d - r + 2$.
Consider $\Gamma := \langle Z \setminus 2P, F\rangle \simeq \mathbb{P}^{r-2}$
because $F \notin \langle Z \setminus 2P\rangle$.

Now consider all the hyperplanes containing Γ.
Each one of them intersects C_d in d points counted with multiplicity,
$r-2$ of those points belong to Γ:
$P_3, \dots, P_r$ and the other $d-r+2$ counted with multiplicity.



Rank of binary forms with given border rank

Claim: There exists at least a hyperplane containing Γ which intersects C_d in d distinct points.
Proof of the Claim: Assume that this is not the case: any hyperplane $M \supset \Gamma$ intersects C_d in at least a double point $2Q$.



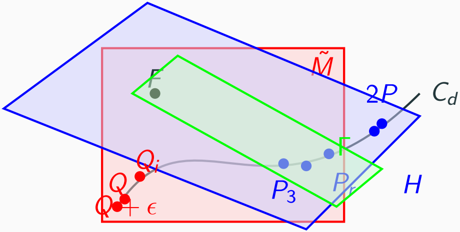
Rank of binary forms with given border rank

Consider the hyperplane

$$\tilde{M} := \langle \Gamma, Q_{r-1}, \ldots, Q_{d-2}, Q, Q{+}\epsilon \rangle =$$

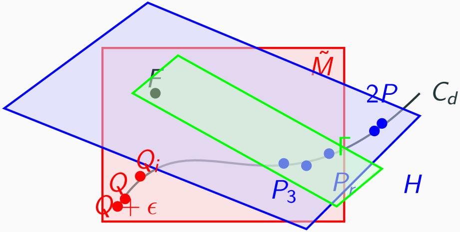
$$= \langle P_3, \ldots, P_r, Q_{r-1}, \ldots, Q_{d-2}, Q, Q{+}\epsilon \rangle.$$

$\tilde{M}$ is a hyperplane which contains Γ and which intersects C_d in $d - r + 2$ distinct points not belonging to Γ. This concludes the proof of the Claim.



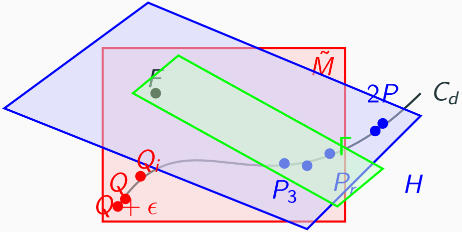
Rank of binary forms with given border rank

Now if we prove that $F \in \langle Q_{r+1}, \dots, Q_{d-2}, Q, Q+\epsilon \rangle$ we will be done because $Q_{r+1}, \dots, Q_{d-2}, Q, Q+\epsilon \in C_d$.
$F \notin \langle Z \setminus 2P \rangle$, so F is not a linear combination of $P_3, \dots, P_r$. But $F \in \Gamma = \langle Z \setminus 2P, F \rangle \subset \tilde{M}$, so F is a linear combination of the points of $\tilde{M} \cap C_d \setminus \{Z \setminus 2P\}$.
This concludes the proof of the Theorem.



Rank of binary forms with given border rank

Now if we prove that $F \in \langle Q_{r+1}, \ldots, Q_{d-2}, Q, Q+\epsilon \rangle$ we will be done because $Q_{r+1}, \ldots, Q_{d-2}, Q, Q+\epsilon \in C_d$. $F \notin \langle Z \setminus 2P \rangle$, so F is not a linear combination of $P_3, \ldots, P_r$. But $F \in \Gamma = \langle Z \setminus 2P, F \rangle \subset \tilde{M}$, so F is a linear combination of the points of $\tilde{M} \cap C_d \setminus \{Z \setminus 2P\}$.
This concludes the proof of the Theorem.



If we want to perform an efficient version of Sylvester algorithm we can compute firstly the border rank (via the minors of the most square catalecticant) and then we check if $rk(F) = brk(F)$, if not then we have just proved we ha$rk(F) = d - r + 2$.

Sylvester Algorithm - Efficient version

Input: $F \in S^d\mathbb{C}^2$

Output: rk, brk and decomposition of F.

1. Compute the rank of $C_F^{\lceil d/2 \rceil, \lfloor d/2 \rfloor} =: r$.
   Then the border rank of F is r.
2. Compute the $\ker(C_F^{r,d-r}) := K$
   2.1 If a generic element of K has r distinct roots, $L_1, \ldots, L_r$ then $rk(F) = r$ and $F = \sum_{i=1}^r \lambda_i L_i^d$
   2.2 otherwise $rk(F) = d - r + 2$ and go to next step
3. Compute a generic element $u \in \ker(C_F^{d-r+2,r-2}) := \tilde{K}$, the $d - r + 2$ distinct roots $L_1, \ldots, L_{d-r+2}$ of u give a decomposition of $F = \sum_{i=1}^{d-r+2} \lambda_i L_i^d$

Example

$F = x^2y^4$

1. $C_F^{3,3} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 3 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$, $rk(C_F^{3,3}) = 3$, so $brk(F) = 3$, therefore $rk(F) = 3, 5 (= 6 - 3 + 2)$.
2. Check if $rk(F) = 3$:
   $\ker(C_F^{3,3}) = \langle (1,0,0,0) \rangle \subset S^3\mathbb{C}^2$. A generic element $u \in \ker(C_F^{3,3})$ has not distinct roots, so $rk(F) = 5$.
3. $C_F^{5,1} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \end{pmatrix}$ $\ker(C_F^{5,1}) \ni u_0^5 + u_0^4u_1 + u_0^3u_1^2 + u_1^5$ (numerically) 5 distinct roots.

Example

$F = (x+y)^2(x-y)^2 = x^4 - 2x^2y^2 + y^4$

1. $C_F^{2,2} = \begin{pmatrix} 3 & 0 & -1 \\ 0 & -1 & 0 \\ -1 & 0 & 3 \end{pmatrix}$, $rk(C_F^{2,2}) = 3$ so $brk(F) = 3$ so $rk(F) = 3, 3$.
2. $C_F^{3,1} = \begin{pmatrix} 3 & 0 & -1 & 0 \\ 0 & -1 & 0 & 3 \end{pmatrix}$,
   $\ker(C_F^{3,1}) = \langle (1,0,3,0), (0,3,0,1) \rangle \ni u_0^3 + 3u_0^2u_1 + 3u_0u_1^2 + u_1^3$ which does NOT have distinct roots.

It seems that there are not 3 distinct solutions. BUT we know that the rank of F is 3. Let's try another element in the same kernel.

$2(1,0,3,0) + 1(0,3,0,1) = (2,3,6,1)$ which corresponds to
$2u_0^3 + 3u_0^2u_1 + 6u_0u_1 + u_0^3$

Numerical solutions are $(1, 0.23 + i0.55), (1, 0.23 - i0.55), (1, -5.52)$.

The problem was that the previous element in the kernel was not generic. When we test if $rk(F) = brk(F)$ we should take a random element in the kernel.

Summary

- Rank-1 (symmetric) tensors: easy to be detected thank to equations of (Veronese) Segre variety: 2-minors of flattenings
- Rank > 1 Sylvester algorithm can do everything for homogeneous polynomials in 2 variables (tensors in $S^d((\mathbb{C}^2)^{\otimes d})$)

Summary

- Rank-1 (symmetric) tensors: easy to be detected thank to equations of (Veronese) Segre variety: 2-minors of flattenings
- Rank >1 Sylvester algorithm can do everything for homogeneous polynomials in 2 variables (tensors in $S^d((\mathbb{C}^2)^{\otimes d})$)

What's next?

- Any polynomial:
  - If $rk(F) < rkC_F^{\lfloor d/2\rfloor,\lceil d/2\rceil}$, then $\ker(C_F^{\lfloor d/2\rfloor,\lceil d/2\rceil})$ is non trivial and reveals the apolar points (already observed by [Iarrobino-Kanev], comuptational way of finding them [Brachat-Comon-Mourrain-Tsidgaridas])
  - If $rk(F) \geq rkC_F^{\lfloor d/2\rfloor,\lceil d/2\rceil}$, then $\ker(C_F^{\lfloor d/2\rfloor,\lceil d/2\rceil}) =$ is trivial and it doesen't reveal the apolar points.
    - One need to build $\tilde{F} \in S^{D>d}C^n$ such that $\partial_\alpha \tilde{F}$ and if $F = \sum_{i=1}^r L_i^d$ then $\tilde{F} = \sum_{i=1}^r L_i^D$ (follow Mourrain's lectures!))