Secant Varieties, Symbolic Powers, Statistical Models

Seth Sullivant

North Carolina State University

November 19, 2012
Given \( V, W \subseteq \mathbb{P}^{n-1} \), the joint is

\[
V \ast W := \bigcup_{v \in V, w \in W} \langle v, w \rangle \subseteq \mathbb{P}^{n-1}
\]

Given \( V \subseteq \mathbb{P}^{n-1} \), the \( r \)-th secant variety of \( V \) is given by

\[
\text{Sec}^r(V) := V \ast \text{Sec}^{r-1}(V) = \bigcup_{v_1, \ldots, v_r \in V} \langle v_1, v_2, \ldots, v_r \rangle \subseteq \mathbb{P}^{n-1}
\]
**Definition**

Let $I, J \subseteq \mathbb{K}[x_1, \ldots, x_n]$ homogeneous ideals. The join of $I$ and $J$ is

$$I \ast J := (I(y) + J(z) + \langle x_i - y_i - z_i : i = 1, \ldots, n \rangle) \cap \mathbb{K}[x_1, \ldots, x_n].$$

i.e. $f \in I \ast J \iff f(y + z) \in I(y) + J(z)$.

**Definition**

The $r$-th secant ideal $I \subseteq \mathbb{K}[x_1, \ldots, x_n]$ is defined by

1. $I^{\{1\}} = I$
2. $I^{\{r\}} = I \ast I^{\{r-1\}}, r > 1.$
Consider $\mathbb{K}[X] := \mathbb{K}[x_{ij} : 1 \leq i \leq m, \ 1 \leq j \leq n]$

Let $I$ be the ideal of $2 \times 2$ minors of generic matrix

$$X = \begin{pmatrix}
    x_{11} & x_{12} & \cdots & x_{1n} \\
    \vdots & \vdots & \ddots & \vdots \\
    x_{m1} & x_{m2} & \cdots & x_{mn}
\end{pmatrix}$$

$V(I) \subseteq \mathbb{P}^{mn-1}$ is the set of rank 1 matrices. (aka $\text{Seg}(\mathbb{P}^{m-1} \times \mathbb{P}^{n-1})$)

$\text{Sec}^r(V(I))$ is the set of all matrices of rank $\leq r$.

$I^\{r\} = I(\text{Sec}^r(V(I)))$ is the ideal of $(r+1) \times (r+1)$ minors of $X$. 

Seth Sullivant (NCSU)  
Secant Varieties, etc.  
November 19, 2012  
4 / 27
When is the Secant Ideal “Nice”? 

Question

For which ideals $I$ does $I^{\{2\}} = \langle 0 \rangle$? i.e. When is $V(I)^{\{2\}} = \mathbb{P}^{n-1}$?

Theorem (Zak 1994)

If $V(I)$ is a linearly nondegenerate, smooth irreducible variety of codimension $\leq \frac{1}{3} n + 1$, then $I^{\{2\}} = \langle 0 \rangle$ unless $V(I)$ is one of

- $\nu_2(\mathbb{P}^2)$,
- $\text{Seg}(\mathbb{P}^2 \times \mathbb{P}^2)$,
- $\text{Gr}(2, 6)$,
- or the Cartan variety.

Question

For which ideals $I$ is $I^{\{r\}}$ generated by polynomials of degree $r + 1$?
Symbolic Powers

Definition

The $r$th symbolic power of $I \subset R$ is the ideal

$$I^{(r)} = (R_I^{-1} \cdot I^r) \cap R$$

where $R_I$ is the complement of the union of the minimal primes of $I$.

- If $P$ is a prime ideal, $P^{(r)}$ is the $P$-primary component of $P^r$.

Theorem ((Special case of) Zariski-Nagata)

If $I \subseteq \mathbb{C}[x_1, \ldots, x_n]$ is radical then

$$I^{(r)} = I^{<r>} := \left\{ f \mid \frac{\partial^{|a|} f}{\partial x^a} \in I \text{ for all } a \in \mathbb{N}^n \text{ with } \sum_{i=1}^{n} a_i \leq r - 1 \right\}.$$ 

Symbolic power gives equations vanishing to high order on $V(I)$.
Consider $\mathbb{K}[X] := \mathbb{K}[x_{ij} : 1 \leq i \leq m, \ 1 \leq j \leq n]$

Let $I$ be the ideal of $2 \times 2$ minors of generic matrix $X$.

The partial derivatives of minors of a generic matrix are themselves minors of a generic matrix. e.g.

$$\frac{\partial}{\partial x_{11}} \begin{vmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{vmatrix} = \begin{vmatrix} x_{22} & x_{23} \\ x_{32} & x_{33} \end{vmatrix}$$

$I^{(2)}$ generated by $3 \times 3$ minors and products of $2 \ 2 \times 2$ minors of $X$.

Note, if $m = 2$, then $I^{(2)} = I^2$. 
When are the Symbolic Powers “Nice”? I

Definition
A radical ideal \( I \) is normally torsion free if

\[ I^{(r)} = I^r \]

for all \( r \geq 1 \).

Complete intersections, Maximal minors

Theorem (Ein-Lazarsfeld-Smith, Hochster-Huneke)

\( I^{(\text{codim}(I)r)} \subseteq I^r \) for all \( r \) in an equal characteristic regular ring.
Secant Ideals and Symbolic Powers

Proposition

For $I \subseteq \mathbb{C}[x_1, \ldots, x_n]$, \[ I^{<r>} = I \ast \langle x_1, \ldots, x_n \rangle^r. \]

Proposition (Catalano-Johnson)

If $I \subseteq \mathbb{C}[x_1, \ldots, x_n]$, homogeneous and $I \subseteq \langle x_1, \ldots, x_n \rangle^2$, then \[ I^{\{r\}} \subseteq I^{<r>}. \]

Proof.

\[
I^{\{r\}} = I \ast I^{\{r-1\}} \subseteq I \ast \langle x_1, \ldots, x_n \rangle^r = I^{<r>}. \square
\]

Proposition (Landsberg-Manivel, Sidman-Sullivant)

If $I \subseteq \mathbb{C}[x_1, \ldots, x_n]$, homogeneous and $I \subseteq \langle x_1, \ldots, x_n \rangle^d$, then

\[
I^{<(d-1)(r-1)+1}_{(d-1)r+1} = I^{\{r\}}_{(d-1)r+1}
\]
Catalano-Johnson says $I^\{r\} \subseteq I^{<r>}$. 

Product rule gives $I^{<r-i>}I^{<i>} \subseteq I^{<r>}$, $i = 1, \ldots, r - 1$. 

Putting together gives 

$$\sum_{\lambda \vdash r} I^{\{\lambda_1\}} \ldots I^{\{\lambda_l\}} \subseteq I^{<r>}$$

where the sum is over all partitions of $r$.

**Definition**

An ideal $I \subseteq \mathbb{C}[x_1, \ldots, x_n]$ is **differentially perfect** if and only if 

$$I^{<r>} = \sum_{\lambda \vdash r} I^{\{\lambda_1\}} \ldots I^{\{\lambda_l\}} \quad \text{for all } r \geq 1.$$ 

Normally torsion free ideal, determinantal and Pfaffian ideals
Definition

Let $G$ be an undirected graph, with vertex set $[n] = \{1, 2, \ldots, n\}$. The edge ideal of $G$ is

$$I(G) = \langle x_i x_j : ij \in E(G) \rangle.$$
Given a graph $G$ and subset of vertices $V$, $G_V$ is the induced subgraph of $G$ with vertex set $V$.

The **chromatic number** $\chi(G)$ is the smallest number of colors to properly color the vertices of a graph.

**Theorem (Sturmfels-Sullivant)**

The secant ideals of an edge ideal $I(G)$ are generated by

$$I(G)^{\{r\}} = \langle \prod_{i \in V} x_i : \chi(G_V) > r \rangle.$$  

Minimal generators of $I(G)^{\{r\}}$ are minimal obstructions to $r$-coloring $G$.

$$I(G)^{\{2\}} = \langle x_2 x_3 x_4, x_1 x_2 x_4 x_5 x_6 \rangle.$$
"Nice" Secant Ideals of Edge Ideals

Proposition

\[ I(G)\{2\} = \langle 0 \rangle \text{ if and only if } G \text{ is a bipartite graph.} \]

Proposition

\[ I(G)\{r\} \text{ is generated in degree } \leq r + 1 \text{ for all } r \text{ if and only if } G \text{ is a perfect graph.} \]

\[ I(G)\{2\} = \langle x_2x_3x_4, x_1x_2x_4x_5x_6 \rangle \]

\[ I(G)\{2\} = \langle x_1x_2x_5, x_1x_5x_6, x_2x_3x_4, x_2x_4x_5 \rangle \]
The covering number $\tau(G)$ is the size of the smallest vertex cover of $G$.

For a vector $a \in \mathbb{N}^n$ the parallelization $G_a$ is the graph with each vertex $i$ replicated $a_i$ times.

**Theorem**

$$I(G)^{<r>} = \langle x^a : \tau(G_a) \geq r \rangle$$

*Generators of symbolic powers are minimal obstructions to covering.*
Theorem (Simis-Vasconcelos-Villarreal)

\[ I(G)^{<r>} = I(G)^r \text{ for all } r \text{ if and only if } G \text{ is a bipartite graph.} \]

Theorem (Villarreal, Sullivant)

\[ I(G) \text{ is differentially perfect, i.e.} \]

\[ I(G)^{<r>} = \sum_{\lambda \vdash r} I(G)^{\{\lambda_1\}} \cdots I(G)^{\{\lambda_l\}} \]

for all \( r \) if and only if \( G \) is a perfect graph.
A statistical model is a set of probability distributions or densities.

Definition

Independence model for two discrete random variables:

\[ M_{X \perp Y} = \left\{ P = (p_{ij}) \in \mathbb{R}^{m \times n} : \sum_{i,j} p_{ij} = 1, p_{ij} \geq 0, \text{ rank } P = 1 \right\} \]

(Zariski closure of ) many statistical models have structure of algebraic varieties

Many of those varieties are secant varieties or joins

- Mixture models
- Factor analysis

Can use that algebraic structure to answer statistical questions.
Join vs. Mixture

- Join

- Mixture
Identifiability of Phylogenetic Mixture Models

Theorem (Rhodes-Sullivant)

The unrooted tree and numerical parameters in a r-class, same tree phylogenetic mixture model on n-leaf trivalent trees are generically identifiable, if \( r < 4^{\lceil n/4 \rceil} \).

- Identifiability for Secant Varieties: Does a generic point on the \( r \)th secant variety \( \text{Sec}^r(V) \) lie on a unique secant \( r - 1 \) plane to \( V \)?

- Proof uses:
  - Relation to identifiability problem for \( \text{Sec}^{4r}(\text{Seg}(\mathbb{P}^a \times \mathbb{P}^b \times \mathbb{P}^c)) \)
  - Knowledge of generators of the secant ideals \( I_T^{\{r\}} \)
A centered Gaussian random vector $X \in \mathbb{R}^n$ has density function

$$f(x) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp \left( -\frac{1}{2}x^T \Sigma^{-1} x \right)$$

with $\Sigma \in PD_n \subset \mathbb{R}^{n(n+1)/2}$.

Set of all centered Gaussian r.v.’s parametrized by $PD_n$.

A Gaussian statistical model is a subset $\Theta \subset PD_n$.

**Definition (Factor Analysis Model)**

$$\Theta_{r,n} = \{ \Psi + \Lambda \Lambda^T \mid \Psi > 0 \text{ diagonal } , \ \Lambda \in \mathbb{R}^{r \times n} \} \subset PD_n$$

Diagonal plus rank $\leq r$ PSD matrix.
Model assumptions: $X_1, \ldots, X_n$ are observable covariates. $Y_1, \ldots, Y_r$ are fewer unobservable (hidden) factors, which explain the correlation among $X_1, \ldots, X_n$.

- For all $i, j$, $X_i \perp \perp X_j \mid Y_1, \ldots, Y_r$
- For all $i, j$, $Y_i \perp \perp Y_j$

**Example (g-theory)**

$X_i$ are test scores, and $Y_j$ are underlying “types of intelligence”.

![Diagram of factor analysis model]
Model Selection: How Many Factors? What is $r$?

**Statistical Approach**
- Wald-Type Test
- Likelihood Ratio Test
- Information Criteria (AIC, BIC, WAIC)

**Mathematical Problem**
- Find generators of $I(\Theta_{r,n})$
- Compute tangent cone $TC_p(\Theta_{r,n})$ at $p \in \Theta_{r-1,n}$.
- Determine the Real Log-Canonical Threshold of singular fibers of $\Theta_{r,n}$. 
The Second Hypersimplex

Observation

\[ \Theta_{r,n} = \text{Sec}^r(\Theta_{1,n}) \]

\( \Theta_{1,n} \) is a toric variety called the second hypersimplex.

\[ \phi : \mathbb{P}^{2n-1} \rightarrow \mathbb{P}^{n(n+1)/2-1} \]

\[ \phi_{ij}(\psi, \lambda) = \begin{cases} 
\lambda_i \lambda_j & \text{if } i \neq j \\
\psi_i + \lambda_i^2 & \text{if } i = j 
\end{cases} \]

Definition

\[ I_{1,n} = I(\text{im } \phi) \] toric ideal of second hypersimplex

\[ I_{r,n} = I_{1,n}^{\{r\}} = \text{vanishing ideal of } r\text{-factor model} \]

- Quadratic Gröbner basis of \( I_{1,n} \) computed by De Loera-Sturmfels-Thomas
Theorem (Sullivant)

The secant ideal \( I_{2,n} = I_{1,n}^{\{2\}} \) has a Gröbner basis, with squarefree initial terms, consisting of polynomials of all odd degrees between 3 and \( n \), with respect to a circular term order.

Proof idea:

- Degenerative strategy: Hope that \( \text{in}_{\prec}(I_{1,n}^{\{2\}}) = (\text{in}_{\prec}(I_{1,n}))^{\{2\}} \)

Lemma (Simis-Ulrich)

For any \( I \) and any term order \( \prec \)

\[
\text{in}_{\prec}(I^{\{r\}}) \subseteq (\text{in}_{\prec}(I))^{\{r\}}
\]

- Just need to construct polynomials in \( I^{\{r\}} \) whose initial terms generate \( (\text{in}_{\prec}(I))^{\{r\}} \).
The initial ideal $\in_{\prec}(I_{1,n})$ is an edge ideal.

The graph $G_n$ such that $\in_{\prec}(I_{1,n}) = I(G_n)$ is the non-crossing pair graph of the cyclically embedded complete graph $K_n$. [De Loera-Sturmfels-Thomas]

$\in_{\prec}(I_{1,n})^{\{2\}}$ is generated by the odd cycles in $G_n$.

Construct polynomials in $I_{2,n}$ having those cycles as initial terms.

\[
\sigma_{12}\sigma_{15}\sigma_{23}\sigma_{34}\sigma_{45} - \sigma_{12}\sigma_{13}\sigma_{25}\sigma_{34}\sigma_{45} - \sigma_{12}\sigma_{14}\sigma_{23}\sigma_{35}\sigma_{45} + \sigma_{12}\sigma_{14}\sigma_{25}\sigma_{34}\sigma_{35} + \sigma_{12}\sigma_{13}\sigma_{24}\sigma_{35}\sigma_{45} - \sigma_{12}\sigma_{15}\sigma_{24}\sigma_{34}\sigma_{35} \\
+ \sigma_{13}\sigma_{14}\sigma_{23}\sigma_{25}\sigma_{45} - \sigma_{13}\sigma_{14}\sigma_{24}\sigma_{25}\sigma_{35} - \sigma_{13}\sigma_{15}\sigma_{23}\sigma_{24}\sigma_{45} + \sigma_{13}\sigma_{15}\sigma_{24}\sigma_{25}\sigma_{34} - \sigma_{14}\sigma_{15}\sigma_{23}\sigma_{25}\sigma_{34} + \sigma_{14}\sigma_{15}\sigma_{23}\sigma_{24}\sigma_{35}
\]

Prove that these generalized pentads are in $I_{1,n}^{\{2\}}$ using symbolic powers.
Secant varieties and symbolic powers are intimately related.

- Symbolic powers shed light on equations for secant varieties.
- In special cases, symbolic powers’ structure complete determined by secant ideals.
- Nice combinatorial and computational structure.

Secant varieties make frequent appearance in statistics.

- Mixture models
- Factor analysis model

Theoretical advances on secant varieties lead to advances in statistics.


