Towards the Salmon Conjecture

Luke Oeding

Texas A&M University

December 16, 2008

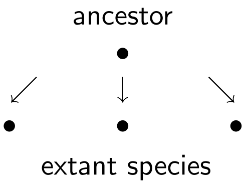
Algebraic Statistics, MSRI, Berkeley, CA

The Salmon Prize

In 2007, E. Allman offered a prize of Alaskan Salmon to anyone who finds the defining ideal of

$$\sigma_4\left(\mathbb{P}^3 \times \mathbb{P}^3 \times \mathbb{P}^3\right).$$

This algebraic variety may be viewed as a statistical model for evolution. (c.f. talk of Marta Casanellas).



Possible observed values $\{A, C, G, T\}$. Assume observations at extant species are independent. Ancestor unknown. Mixture model of 4 independence models.

Invariants of this statistical model $\leftrightarrow$ ideal of the algebraic variety.

The Salmon Prize

The main motivation: Work of Allman-Rhodes implies that solving this problem would provide all invariants for the mixture model of *any* binary evolutionary tree with *any* number of states.

Goals:

- Explain techniques from representation theory and geometry.
- Explain current status of the salmon problem.
- Provide a template for studying similar problems.

Secant Varieties

Let $A = \{a_i\}, B = \{b_j\}, C = \{c_k\}$, be $\mathbb{C}$-vector spaces, then the tensor product is $A \otimes B \otimes C = \{a_i \otimes b_j \otimes c_k\}$, with coordinates p_{ijk}.

- Independence model - aka Segre variety (rank 1 tensors): Defined by

$$\begin{aligned} Seg : \mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C &\longrightarrow \mathbb{P}\big(A \otimes B \otimes C\big) \\ ([a],[b],[c]) &\longmapsto [a \otimes b \otimes c]. \end{aligned}$$

- Mixture model - aka the r^{th} *secant variety* of a variety $X \subset \mathbb{P}^n$:

$$\sigma_r(X) = \overline{\bigcup_{x_1,\dots,x_r \in X} \mathbb{P}(\mathrm{span}\{x_1,\dots,x_r\})} \subset \mathbb{P}^n.$$

(*Can also work over $\mathbb{R}$ or Δ-probability simplex, but not today.)

Symmetry

- The symmetry group of the salmon variety

$$\sigma_4\big(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C\big)$$

  is change of coordinates in each factor, *i.e.* $GL(A) \times GL(B) \times GL(C)$

  (or $GL(A) \times GL(B) \times GL(C) \times \mathfrak{S}_3$ when $A \cong B \cong C$).

- Good news: A large group acts! Can use tools from representation theory!

- This symmetry is a powerful tool if we can exploit it!

Ideals with Symmetry

Recall: projective varieties have homogeneous ideals. This symmetry induces grading by degree.

$$\begin{array}{ccc} \mathbb{C}[p_0,\dots,p_N] & = & \bigoplus_d \mathbb{C}[p_0,\dots,p_N]_d \\ \cup & & \cup \\ \mathcal{I}(X) & = & \bigoplus_d \mathcal{I}_d(X) \end{array}$$

When a larger group G acts on X (and on $\mathcal{I}(X)$), we get a finer decomposition of $\mathcal{I}(X)$ into G-modules using representation theory.

This is good because
Ideal Mantra:

Ideals with Symmetry

Recall: projective varieties have homogeneous ideals. This symmetry induces grading by degree.

$$\begin{array}{ccc} \mathbb{C}[p_0,\dots,p_N] & = & \bigoplus_d \mathbb{C}[p_0,\dots,p_N]_d \\ \cup & & \cup \\ \mathcal{I}(X) & = & \bigoplus_d \mathcal{I}_d(X) \end{array}$$

When a larger group G acts on X (and on $\mathcal{I}(X)$), we get a finer decomposition of $\mathcal{I}(X)$ into G-modules using representation theory.

This is good because
Ideal Mantra: "Polynomials in G-modules are like Musketeers - one for all and all for one!"

What is an irreducible module?

Think: A G-module is a vector space with a G-action

An *irreducible module* is one with no proper nontrivial submodules.
Example: The space of square matrices $V \otimes V$ is not an irreducible $GL(V)$ module since it splits as

$$V \otimes V = S^2 V \oplus \bigwedge^2 V$$

(You already knew this: every square matrix may be written as a sum of a symmetric and a skew symmetric matrix)

Representation Theory Notation

- Module notation: $S^d(A \otimes B \otimes C) = \mathbb{C}[p_{ijk}]_d$.
- Fact: $S^d(A \otimes B \otimes C)$ is a $GL(A) \times GL(B) \times GL(C)$-module.
- The irreducible submodules of $S^d(A \otimes B \otimes C)$ are isomorphic to modules indexed by certain partitions π_1, π_2, π_3 of d:

$$S_{\pi_1} A \otimes S_{\pi_2} B \otimes S_{\pi_3} C,$$

and usually occur with multiplicity - this makes us work harder.

Note: The $S_\pi V$ are irreducible $GL(V)$ modules called Schur modules.

Using Representation Theory

- For the groups we encounter, irreducible G-modules M satisfy

$$\operatorname{span}\{G.f\} = M \quad \text{for} \quad 0 \neq f \in M$$

- In practice, use a distinguished f called a "highest weight vector".
- Can test if an irreducible $M \subset \mathcal{I}(X)$ by testing if $f \in \mathcal{I}(X)$!
- If we have $f \in \mathcal{I}(X)$, can find entire modules in $\mathcal{I}(X)$!

What is a flattening?

Think: $U \otimes V =$ space of matrices.

3 canonical ways to express a tensor $T \in A \otimes B \otimes C$ as a matrix:
$T \in A \otimes (B \otimes C)$ or $T \in B \otimes (A \otimes C)$ or $T \in (A \otimes B) \otimes C$.

For example

$$\begin{pmatrix} p_{111} & p_{121} & p_{131} & | & p_{112} & p_{122} & p_{132} & | & p_{113} & p_{123} & p_{133} \\ p_{211} & p_{221} & p_{231} & | & p_{212} & p_{222} & p_{232} & | & p_{213} & p_{223} & p_{233} \\ p_{311} & p_{321} & p_{331} & | & p_{312} & p_{322} & p_{332} & | & p_{313} & p_{323} & p_{333} \end{pmatrix}$$

is a flattening of $T = [p_{ijk}] \in \mathbb{C}^3 \otimes \mathbb{C}^3 \otimes \mathbb{C}^3$ to $\mathbb{C}^3 \otimes (\mathbb{C}^3 \otimes \mathbb{C}^3) \cong \mathbb{C}^3 \otimes \mathbb{C}^9$.

A familiar G-module

The $GL(A) \times GL(B) \times GL(C)$-module of 3×3 minors of the flattening $A \otimes B \otimes C \to A \otimes (B \otimes C)$ is

$$F := S_{(1,1,1)}A \otimes S_{(1,1,1)}(B \otimes C) = \bigwedge^3 A \otimes \bigwedge^3 (B \otimes C)$$

This module is not irreducible: $F = F_1 \oplus F_2 \oplus F_3 =$

$$(\bigwedge^3 A \otimes \bigwedge^3 B \otimes S^3 C) \oplus (\bigwedge^3 A \otimes S_{(2,1)} B \otimes S_{(2,1)} C) \oplus (\bigwedge^3 A \otimes S^3 B \otimes \bigwedge^3 C)$$

After choosing ordered (or weighted) bases of A, B, C, can define a highest weight. For example, the highest weight vector of $\bigwedge^3 A \otimes \bigwedge^3 B \otimes S^3 C$ is

$$(a_1 \wedge a_2 \wedge a_3) \otimes (b_1 \wedge b_2 \wedge b_3) \otimes (c_1^{\otimes 3}) = \det \begin{pmatrix} p_{111} & p_{121} & p_{131} \\ p_{211} & p_{221} & p_{231} \\ p_{311} & p_{321} & p_{331} \end{pmatrix}$$

Can do ideal membership test for each irreducible module by testing vanishing of its highest weight vector!

Highlights: Inheritance via an example

Proposition (example of Proposition 4.4 Landsberg–Manivel , '04)

$$\tilde{M}_6 := S_{(2,2,2)}\mathbb{C}^4 \otimes S_{(2,2,2)}\mathbb{C}^4 \otimes S_{(3,1,1,1)}\mathbb{C}^4 \in \mathcal{I}\big(\sigma_4(\mathbb{P}^3 \times \mathbb{P}^3 \times \mathbb{P}^3)\big)$$

if and only if

$$M_6 := S_{(2,2,2)}\mathbb{C}^3 \otimes S_{(2,2,2)}\mathbb{C}^3 \otimes S_{(3,1,1,1)}\mathbb{C}^4 \in \mathcal{I}\big(\sigma_4(\mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^3)\big).$$

Note: $\dim(\tilde{M}_6) = 10^3$ but $\dim(M_6) = 10$, and has basis of polynomials, each with 576 or 936 monomials.

The point: The number of parts of π_1, π_2, π_3 tell us which secant variety to look at. This is a significant dimension reduction.

For $\sigma_4(\mathbb{P}^3 \times \mathbb{P}^3 \times \mathbb{P}^3)$ we only need to consider $S_{\pi_1}A \otimes S_{\pi_2}B \otimes S_{\pi_3}C$ where π_1, π_2, π_3 have 4 parts, and those equations we get from inheritance.

Highlights: Flattenings and subspace varieties

Definition: $$Sub_{p,q,r}(A \otimes B \otimes C) := \{[T] \in \mathbb{P}(A \otimes B \otimes C) \mid \exists \mathbb{C}^p \subseteq A, \mathbb{C}^q \subseteq B, \mathbb{C}^r \subseteq C, \text{ and } [T] \in \mathbb{P}(\mathbb{C}^p \otimes \mathbb{C}^q \otimes \mathbb{C}^r)\}$$

i.e. Tensors that can be written using fewer variables.

Highlights: Flattenings and subspace varieties

Definition: $Sub_{p,q,r}(A \otimes B \otimes C) := \{[T] \in \mathbb{P}(A \otimes B \otimes C) \mid \exists \mathbb{C}^p \subseteq A, \mathbb{C}^q \subseteq B, \mathbb{C}^r \subseteq C, \text{ and } [T] \in \mathbb{P}(\mathbb{C}^p \otimes \mathbb{C}^q \otimes \mathbb{C}^r)\}$

i.e. Tensors that can be written using fewer variables.

Theorem (3.1, Landsberg–Weyman '07)

$Sub_{p,q,r}(A \otimes B \otimes C)$ is normal with rational singularities. Its ideal is generated by the minors of flattenings;

$$\left(\bigwedge^{p+1} A \otimes \bigwedge^{p+1}(B \otimes C)\right) \oplus \left(\bigwedge^{q+1} B \otimes \bigwedge^{q+1}(A \otimes C)\right) \oplus \left(\bigwedge^{r+1}(A \otimes B) \otimes \bigwedge^{r+1} C\right)$$

Fact: $Sub_{r,r,r}(A \otimes B \otimes C) \supseteq \sigma_r(\mathbb{P}A \times \mathbb{P}B \times \mathbb{P}C)$

Key Point: The subspace varieties contain secant varieties, and therefore they give some of the equations of the secant varieties.

Landsberg and Manivel's Reduction

Theorem (Landsberg-Manivel '08 Corollary 5.6)

$\sigma_4\left(\mathbb{P}^3 \times \mathbb{P}^3 \times \mathbb{P}^3\right)$ *is the zero set of:*

1. *Degree* 5 *equations*

$$\begin{aligned} M_5 \quad := \quad & \left(S_{(3,1,1)}\mathbb{C}^4 \otimes S_{(2,1,1,1)}\mathbb{C}^4 \otimes S_{(2,1,1,1)}\mathbb{C}^4\right) \\ & \oplus \left(S_{(2,1,1,1)}\mathbb{C}^4 \otimes S_{(3,1,1)}\mathbb{C}^4 \otimes S_{(2,1,1,1)}\mathbb{C}^4\right) \\ & \oplus \left(S_{(2,1,1,1)}\mathbb{C}^4 \otimes S_{(2,1,1,1)}\mathbb{C}^4 \otimes S_{(3,1,1)}\mathbb{C}^4\right) \end{aligned}$$

2. *Equations inherited from* $\sigma_4\left(\mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^3\right)$

Note: M_5 is a 1728 dimensional irreducible G-module, for $G = GL(4) \times GL(4) \times GL(4) \times \mathfrak{S}_3$ with a natural basis of polynomials with 180 or 360 or 540 monomials.

Key point: It remains to find the equations of $\sigma_4(\mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^3)$!

A result of Strassen

Theorem (Strassen 1988 (reinterpreted))

The ideal of the hypersurface $\sigma_4(\mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^2) \subset \mathbb{P}^{26}$ *is generated in degree* 9 *by a nonzero vector in the* 1 *dimensional module*

$$S_{(3,3,3)}\mathbb{C}^3 \otimes S_{(3,3,3)}\mathbb{C}^3 \otimes S_{(3,3,3)}\mathbb{C}^3$$

This polynomial has $9,216$ monomials.

Can get more polynomials via representation theory and geometry: Inheritance implies that
$M_9 := S_{(3,3,3)}\mathbb{C}^3 \otimes S_{(3,3,3)}\mathbb{C}^3 \otimes S_{(3,3,3)}\mathbb{C}^4 \in \mathcal{I}(\sigma_4(\mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^3))$

Note: $\dim(M_9) = 20$, has natural basis of polynomials with $9,216$ or $25,488$ or $43,668$ monomials!

What is known about $\mathcal{I}\left(\sigma_4\left(\mathbb{P}^2\times\mathbb{P}^2\times\mathbb{P}^3\right)\right)$?

- General theory: $\mathcal{I}_s(\sigma_k(\mathbb{P}A\times\mathbb{P}B\times\mathbb{P}C))=0$ for $s\leq k$.
- Computational tests: (Please download my Maple code and double check this work!)

$$\mathcal{I}_5\left(\sigma_4\left(\mathbb{P}^2\times\mathbb{P}^2\times\mathbb{P}^3\right)\right)=0$$

$$M_6 := \mathcal{I}_6\left(\sigma_4\left(\mathbb{P}^2\times\mathbb{P}^2\times\mathbb{P}^3\right)\right)=S_{(2,2,2)}\mathbb{C}^3\otimes S_{(2,2,2)}\mathbb{C}^3\otimes S_{(3,1,1,1)}\mathbb{C}^4 \qquad *$$

- Strassen:
$M_9 := S_{(3,3,3)}\mathbb{C}^3\otimes S_{(3,3,3)}\mathbb{C}^3\otimes S_{(3,3,3)}\mathbb{C}^4\in\mathcal{I}(\sigma_4(\mathbb{P}^2\times\mathbb{P}^2\times\mathbb{P}^3))$

Do M_6 and M_9 suffice to cut out $\sigma_4\left(\mathbb{P}^2\times\mathbb{P}^2\times\mathbb{P}^3\right)$?

*(correction)

Status of the salmon conjecture

Known equations of $\sigma_4\left(\mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^3\right)$:

$$M_6 = S_{(2,2,2)}\mathbb{C}^3 \otimes S_{(2,2,2)}\mathbb{C}^3 \otimes S_{(3,1,1,1)}\mathbb{C}^4$$
$$M_9 = S_{(3,3,3)}\mathbb{C}^3 \otimes S_{(3,3,3)}\mathbb{C}^3 \otimes S_{(3,3,3)}\mathbb{C}^4$$

Shape of partitions implies that $\langle M_9 \rangle \not\subset \langle M_6 \rangle$.

It is known that $\mathcal{V}(M_6) \supseteq \sigma_4\left(\mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^3\right) \cup Sub_{2,3,4} \cup Sub_{3,2,4} \cup Sub_{3,3,3}$. Does equality hold? If not, what are the other components?

If M_9 eliminates the “extra” components, this would resolve the salmon problem (at least set theoretically).

A template for finding invariants:

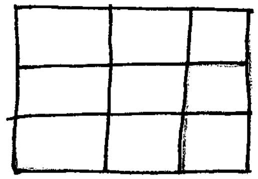
The salmon variety has been studied via the following:

1. Input: statistical model.
2. Find the corresponding algebraic variety X.
3. Find the largest symmetry group G acting X.
4. Study $\mathcal{I}(X)$ as a G-module using representation theory.
5. Use computational tools to study modules potentially in $\mathcal{I}(X)$. (works well for low degree) - See me for Maple implementations. Note: Representation theory tells where to look for invariants as well as how to get new invariants from old.
6. Try to make geometric reductions to show that the known invariants suffice.

This template should be useful for studying other statistical models.

Notes from Media provided

Partition (3,3,3)



Partition (3,1,1,1)