Quantitative Properties of Ideals Arising from Hierarchical Models

Summer School on Geometric and Algebraic Combinatorics

June 24, 2019

Algebraic Statistics Dictionary

Probability/Statistics	**Algebra/Geometry**
Probability distribution	Point
Statistical model	(Semi) Algebraic set
Discrete exponential family	Toric variety
Conditional interference	Lattice points in polytopes
Maximum likelihood estimation	Polynomial optimization
Model selection	Geometry of singularities
Multivariate Gaussian model	Spectral geometry
Phylogenetic model	Tensor networks
MAP estimates	Tropical geometry

From Alg. Stat. book of S. Sullivant

Statistical Models

Hierarchical Models

– record the dependency relationships of random variables

Applications

		Type of Infection		
		Varicella	Influenza	Gastroenteritis
Use Aspirin regularly	Yes	29	21	2
	No	704	188	125

Data on the use of Aspirin from 1070 patients with Reye's Syndrome in US from 1980 to 1997.

Question: *Among patients with Reye's syndrome, is there any relation between the type of infection and the use of Aspirin to treat that infection?*

Running Example

- $\Omega_1 = \{\text{Varicella, Influenza, Gastroenteritis}\}$ $\boxed{r_1 = 3}$
- $\Omega_2 = \{\text{Use Aspirin reg., Don't use Aspirin reg.}\}$ $\boxed{r_2 = 2}$
- $(Z_1, Z_2) \in \Omega_1 \times \Omega_2$
- $P(Z_1 = i, Z_2 = j) = p_{i,j}$
- $\mathcal{P} = \{(p_{i,j}) \mid (i,j) \in \Omega_1 \times \Omega_2, \sum_{i,j} p_{i,j} = 1\}$

$$\boxed{M_1 = \{(p_{i,j}) \in \mathcal{P} \mid p_{i,j} = p_{i,} \cdot p_{,j},\ \text{for all } (i,j) \in \Omega_1 \times \Omega_2\}}$$

$$\boxed{M_2 = \mathcal{P}/M_1}$$

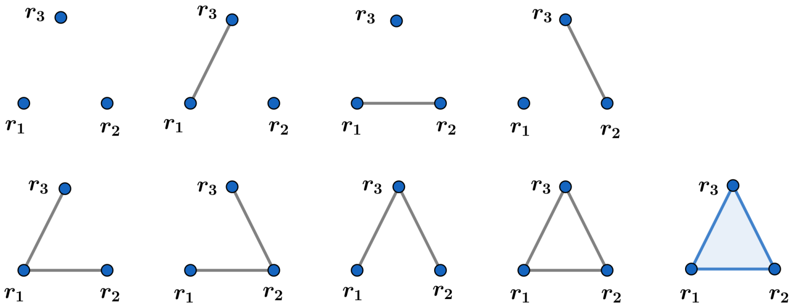
Visualizations of the Models

$\mathrm{M}_1 = \{(p_{i,j}) \mid p_{i,j} = p_{i,} \cdot p_{,j}\}$

$\mathrm{M}_2 = \mathcal{P}/\mathcal{M}_1$

Analyzing Data

- $\Omega_1 = \{\text{Varicella, Influenza, Gastroenteritis}\}$ $\boxed{r_1 = 3}$
- $\Omega_2 = \{\text{Use Aspirin reg., Don't use Aspirin reg.}\}$ $\boxed{r_2 = 2}$
- $\Omega_3 = \{\text{children, teenagers}\}$ $\boxed{r_3 = 2}$



Choose the model that best fits the data.

Hierarchical Models

Definition

A hierarchical model $\mathcal{M}$ on m randon variables consists of

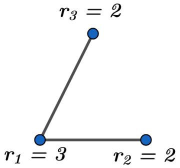
1. a vector $\mathbf{r} = (r_1, r_2, \cdots, r_m)$, where each r_i denotes the number of states for the variable Z_i.
2. a collection $\Delta = \{F_1, F_2, \cdots, F_n\}$, where each $F_j \subset [m]$ in the collection encodes a maximal non independent relation among the parameters indicated in it.

$$\mathcal{M}(\mathbf{r}, \Delta) = \left\{ (p_{i_1 \ldots i_m}) \mid p_{i_1, \ldots, i_m} = \prod_{F 2 \Delta} p_{\mathbf{i}_F} \text{ for all } (i_1 \ldots i_m) \in \prod_{k 2 [m]} [r_k] \right\}$$

How to choose the right model?

Ideals of Hierarchical Models in Algebra

Constructing the Ideal



$\mathcal{M}[(3,2,2),\{\{1,2\},\{1,3\}\}]$

$$\mathsf{K}[X_{111}, X_{112} \dots X_{322}] \xrightarrow{\phi} \mathsf{K}[Y_{11} \dots Y_{32}, Z_{11} \dots Z_{32}],$$

$$\mathrm{X}_{ijk} \longmapsto Y_{ij} \cdot Z_{ik}$$

$$\boxed{I = \ker(\phi) \text{ is the ideal for } \mathcal{M}.}$$

$$I =< X_{111}X_{122} - X_{112}X_{121},$$

$$X_{211}X_{222} - X_{212}X_{221},$$

$$X_{311}X_{322} - X_{312}X_{321} >$$

General Construction

Let $\mathcal{M}(\mathbf{r}, \Delta)$ be a hierarchical model.

$$\mathsf{K}[X_{i_1 \ldots i_m} \,|\, (i_1, \ldots i_m) \in \prod_{i 2 [m]} [r_i]] \xrightarrow{\phi} \mathsf{K}[Y_{F, \mathbf{j}_F} \,|\, F \in \Delta, \quad \mathbf{j}_F \in \prod_{i 2 F} [r_i]],$$

$$\mathrm{X}_{\mathbf{i}} \longmapsto \prod_{F 2 \Delta} Y_{F, \mathbf{i}_F}$$

Note: $[r_i] = \{1, 2, \ldots, r_i\}$

$$\boxed{I = \ker(\phi) \text{ is the ideal for } \mathcal{M}.}$$

$$I =< \{\overline{\mathrm{x}}^{\mathbf{u}} - \overline{\mathrm{x}}^{\mathbf{v}}\} >$$

Fundamental Theorem of Markov Bases [Diaconis-Sturmfels, 1998]

A subset $\beta \subset ker_{\mathsf{Z}} A$ is a Markov Basis for $\mathcal{M}$ if and only if the corresponding set of binomials $\{\overline{\mathrm{x}}^{b^+} - \overline{\mathrm{x}}^{b^-} \mid b = b^+ - b \quad \in \beta\}$ generates the ideal I.

Note: b^+ and b are respectively positive and negative part of the vector b.

Quantitative Properties of the Ideals

Let R be a graded polynomial ring in finitely many variables over a field K and I a homogeneous ideal in R.

$$R/I = [R/I]_0 \oplus [R/I]_1 \oplus \cdots \oplus [R/I]_d \oplus \dots$$

where $[R/I]_d = \{$all homogeneous polynomials of degree d in $R/I\}$

$$\boxed{\mathrm{H}_{R/I}(t) = \sum_{d \quad 0} \dim_{\mathsf{K}} [R/I]_d \cdot t^d}$$

Hilbert series are rational of the form

$$H_{R/I}(t) = \frac{g(t)}{(1-t)^{\dim R/I}}, \quad g(t) \in \mathsf{Z}[t],\ g(1) \neq 0.$$

Continuing our example

$\mathcal{M}((3,2),\{\{1\},\{2\}\})$

$$I =< X_{1,1}X_{2,2} - X_{1,2}X_{2,1}, \\ X_{1,1}X_{3,2} - X_{1,2}X_{3,1}, \\ X_{2,1}X_{3,2} - X_{2,2}X_{3,1} >$$

$R = \mathsf{K}[X_{1,1}, X_{1,2}, \ldots, X_{3,2}]$.

$$H_{R/I}(t) = \sum_{d2\mathsf{Z}^+} \dim_{\mathsf{K}}[R/I]_d \cdot t^d = 1 + 6t + 18t^2 + 40t^3 + \ldots$$

$$\boxed{\mathrm{H}_{R/I}(t) = \frac{1+2t}{(1-t)^4}}$$

$\mathcal{M}((r_1,r_2),\{\{1\},\{2\}\})$

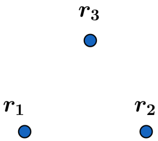
$$H_{R_{\mathbf{r}}/I_{\mathbf{r}}}(t) = ?$$

Running example

$\mathcal{M}((r_1, r_2), \{\{1\}, \{2\}\})$

$$\mathrm{H}_{R_{\mathbf{r}}/I_{\mathbf{r}}}(t) = \frac{\sum_{i=0}^{r_1 \;\; 1} \binom{r_1 \;\; 1}{i} \binom{r_2 \;\; 1}{i} t^i}{(1-t)^{r_1+r_2 \;\; 1}}$$

Conca–Herzog, 1994



$\mathcal{M}((r_1, r_2, r_3), \{1\}, \{2\}, \{3\}\})$

There exists recursive formulas

but not closed formulas

What about more complicated models?

equivariant Hilbert series

Fix Δ and consider $\boldsymbol{I}_\Delta = \{I_\mathbf{r}\}^m_{\mathbf{r}2\mathsf{N}}$, where $I_\mathbf{r} \subset R_\mathbf{r}$ is the ideal for the model $\mathcal{M}(\mathbf{r}, \Delta)$. The equivariant Hilbert series for $\boldsymbol{I}_\Delta$ is the formal power series

$$\mathrm{equivH}_{\boldsymbol{I}_\Delta}(t, \mathbf{s}) = \sum_{r_i \quad 1} H_{R_{\mathbf{r}/I_\mathbf{r}}}(t)\mathbf{s}^\mathbf{r}$$

Note: $\mathbf{s}^\mathbf{r} = s_1^{r_1} s_2^{r_2} \dots s_m^{r_m}$

Theorem (M-Nagel, 2018)

$\Delta = \{\{1\}, \{2\}, \dots, \{m\}\}$ *induces* $\boldsymbol{I}_\Delta = \{I_\mathbf{r}\}_{\mathbf{r}2\mathsf{N}^m}$ *with*

$$\mathrm{equivH}_{\boldsymbol{I}_\Delta}(t, s_1 \dots s_m) = 1 + \frac{s_1 s_2 \dots s_m}{(1 - s_1)(1 - s_2) \dots (1 - s_m) - t}$$

$\mathbf{r} = (r_1, r_2, \dots, r_m), \quad \Delta = \{\{1\}, \{2\}, \dots, \{m\}\}$

$$\mathrm{equivH}(\mathrm{t},\mathrm{s}_1) = \sum_{\mathrm{r}_1 \quad 1} [\frac{1}{(1-\mathrm{t})^{\mathrm{cr}_1}}]\mathrm{s}_1^{\mathrm{r}_1} = \frac{(1-\mathrm{t})^{\mathrm{c}}}{(1-\mathrm{t})^{\mathrm{c}} - \mathrm{s}}$$

Theorem (Nagel-Römer, 2015)

Let $\mathcal{I} = \{I_n \subset R_n\}_{n2\mathsf{N}}$, *where* $R_n = \mathsf{K}[X_{i,j} \mid 1 \leq i \leq c,\ 1 \leq j \in n]$, *is an* Inc*-invariant ltration of ideals. Then*

$$\mathrm{equivH}_{\mathcal{I}}(\mathrm{t},\mathrm{s}) = \sum_{\mathrm{n} \quad 1} \mathrm{H}_{\mathrm{R_n/I_n}}(\mathrm{t}) \cdot \mathrm{s}^{\mathrm{n}}$$

is a rational function.

Corollary

Let Δ *be any simplicial complex and* $\mathcal{I}_\Delta = \{I_r\}_{r2\mathsf{N}}$ *a family of ideals arising from* $\mathcal{M}(\mathbf{r},\Delta)$ *where* $\mathbf{r}$ *has all but one component xed. Then* $\mathrm{equivH}_{\mathcal{I}_\Delta}(\mathrm{t},\mathrm{s})$ *is rational.*

Note: $s^{\mathbf{r}} = s_1^{r_1} s_2^{r_2} \ldots s_m^{r_m}$

$$\mathrm{equivH}(\mathrm{t}, \mathrm{s}_1, \mathrm{s}_2) = \sum_{\mathrm{r}_1, \mathrm{r}_2 \quad 1} \left(\frac{1}{(1-\mathrm{t})^{\mathrm{r}_1 \mathrm{r}_2}}\right) \mathrm{s}_1^{\mathrm{r}_1} \mathrm{s}_2^{\mathrm{r}_2} = \text{ no rational presentation}$$

Set $T = \{t \in [m] \mid r_t \in \mathsf{N}\}$. Given Δ and the fixed values $\{r_i, i \notin T\}$, one considers the family of ideals $\mathcal{I}_{\Delta, \mathbf{r}_{[m]\setminus T}} = \{I_{\mathbf{r}_T}\}_{r_t 2\mathsf{N}}$.

Under what conditions on Δ and T is $\mathrm{equivH}_{\mathcal{I}_{\Delta, \mathbf{r}_{[m]\setminus T}}}(\mathrm{t}, \mathbf{s})$ rational?

Theorem (M-Nagel, 2018)

The equivariant Hilbert series for $\mathcal{I}_{\Delta, \mathbf{r}_{[m]\setminus T}}$ is rational if

1. $F_i \cap F_j = \emptyset$ *for any* $F_i, F_j \in \Delta$.
2. $|F \cap T| \leq 1$ *for any* $F \in \Delta$.

Sketch of the proof:

1. Reduce the problem to Δ being a graph.
2. Study $H_{\mathrm{im}(\phi_{\mathbf{r}})}(t)$, where $\mathrm{im}(\phi_{\mathbf{r}})$ is an algebra over K generated by $\phi_{\mathbf{r}}(x_{\mathbf{i}})$, for all $\mathbf{i} \in [r_1] \times [r_2] \times \cdots \times [r_m]$.
3. Determine a regular language $\mathcal{L}$ and a weight function ρ such that

$$P_{L,\rho}(t, \mathbf{s}) = \mathrm{equivH}_{\mathcal{I}_{\Delta,\mathbf{r}_{[\mathrm{m}]\setminus\mathrm{T}}}}(\mathrm{t}, \mathbf{s})$$

4. Theorem [Honkala, 1989]: Let $\mathcal{L}$ be a regular language and let ρ be a weight function on $\mathcal{L}$. Then the power series

$$H_{L,\rho} = \sum_{w 2 L} \rho(w)$$

is a rational function.

References:

1. E. Belay, J. Bresee and +4, *Reye's Syndrome in the United States from 1981 through 1997*, The New Eng. J. of Med. **340** (1999), 1377–1382.
2. A. Conca, J. Herzog, *On the Hilbert function of determinantal rings and their canonical module*, Proc. Amer. Math. Soc. **112** (1994), 677–681.
3. P. Diaconis, B. Sturmfels, *Algebraic algorithms for sampling from conditional distributions*, Ann. Statist. **26(1)** (1998), 363–397.
4. J. Honkala, *A necessary condition for the rationality of the zeta function of a regular language*, Theor. Comput. Sci. **66 (3)** (1989), 341-347.
5. C. Hillar, S. Sullivant, *Finite Grobner bases in in nite dimensional polynomial rings and applications*, Adv. Math. **229** (1998), 1–25.
6. R. Krone, A. Lynkyn, A. Snowden, *Hilbert series of symmetric ideals in in nite polynomial rings via formal languages Regular languages*, J. of Alg. **485** (2015), 353–362.
7. U. Nagel, T. Römer, *Equivariant Hilbert series in non-Noetherian polynomial rings*, J. Algebra **486** (2015), 204â-245.
8. S. Sullivant, *Algebraic Statistics*, **194** American Mathematical Society, Providence, RI (2018).

Thank you!