May 16, 2024

Primary decompositions, associated primes, and applications in algebraic statistics

Jutta Rath

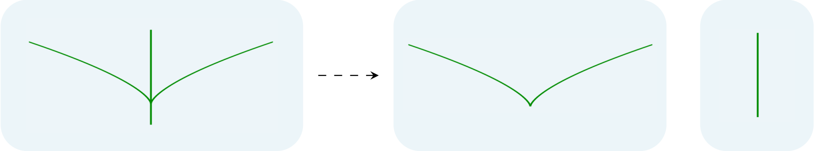
FWF
Der Wissenschaftsfonds.

UNIVERSITÄT KLAGENFURT

Modeling - Analysis - Optimization
DOC.FUNDS DOCTORAL SCHOOL

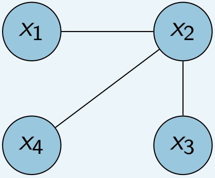
Primary decompositions and associated primes

$180 = 2^2 \cdot 3^2 \cdot 5 \dashrightarrow$ 2 3 5

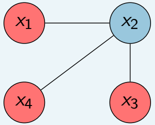


$X^3 - XY^3$ $X^2 - Y^3$ X

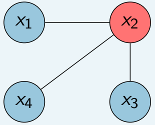
$$\text{associated primes of } 180\mathbb{Z}: \quad \{2\mathbb{Z}, 3\mathbb{Z}, 5\mathbb{Z}\}$$
$$\text{associated primes of } (X^3 - XY^3): \quad \{(X^2 - Y^3), (X)\}$$



$$I = (x_1x_2, x_2x_3, x_2x_4)$$



(x_1, x_3, x_4)



(x_2)

edge ideals $\dashrightarrow$ minimal vertex covers

associated primes of I: $\{(x_2), (x_1, x_3, x_4)\}$

Definition (primary ideal)

An ideal $Q \subsetneq R$ is called primary if whenever $f \cdot g \in Q$, then

- either $f \in Q$, or
- there exists an $n \in \mathbb{N}$ such that $g^n \in Q$.

- every irreducible ideal is primary
- every prime ideal is primary
- if Q is primary, then $\sqrt{Q}$ is prime

Definition (primary decomposition)

Let $I \subseteq R$ be an ideal. A primary decomposition of I is a representation

$$I = Q_1 \cap \cdots \cap Q_r$$

as intersection of finitely many primary ideals Q_i.The decomposition is irredundant if no $Q_i \supseteq \bigcap_{i \neq j} Q_j$ and $\sqrt{Q_i}$ are all distinct.

Example

$$\begin{aligned} I = (x^2, xy) &= (x^2, x) \cap (x^2, y) \\ &= (x) \cap (x^2, y) \\ &= (x) \cap (x^2, cx + y) \end{aligned}$$

for any $c \in \mathbb{R}$.

Theorem (Lasker-Noether)

Every ideal I in a Noetherian ring has an irredundant primary decomposition $I = Q_1 \cap \cdots \cap Q_r$. The ideals in the set

$$\mathrm{Ass}(I) := \left\{ \sqrt{Q_1}, \ldots, \sqrt{Q_r} \right\}$$

are called associated primes of I. $\mathrm{Ass}(I)$ *does not depend on the particular primary decomposition.*

- for every $P \in \mathrm{Ass}(I)$ there exists a $w \in R$ such that $P = I : w$
- w is called a witness of P in I

Example

$I = (x^2, xy) = (x) \cap (x^2, y) = (x) \cap (x^2, x+y)$

$$\begin{aligned}
\mathrm{Ass}(I) &= \left\{ \sqrt{(x)}, \sqrt{(x^2, y)} \right\} = \{(x), (x, y)\} \\
&= \left\{ \sqrt{(x)}, \sqrt{(x^2, x+y)} \right\} = \{(x), (x, y)\}
\end{aligned}$$

$$\begin{aligned}
(x) &= I : (y), \\
(x, y) &= I : (x).
\end{aligned}$$

- y is a witness of (x) in I,
- x is a witness of (x, y) in I.

associated primes of binomial ideals

$$I = (x^{u_1} - \alpha_1 x^{v_1}, \ldots, x^{u_s} - \alpha_s x^{v_s}) \subseteq K[x_1, \ldots, x_r]$$

Theorem (Eisenbud, Sturmfels, 1994)

If I is a binomial ideal, then

- *I has a primary decomposition such that all primary components are binomial,*
- *the radical of I is binomial,*
- *all associated primes of I are binomial.*

associated primes of monomial ideals

Let I be a monomial ideal in $R = K[x_1, \ldots, x_r]$.

- I has a primary decomposition such that all primary components are monomial,
- all associated primes of I are monomial, i.e.,

$$\mathrm{Ass}(I) \subseteq \{(x_1), (x_2), \ldots, (x_r), (x_1, x_2), \ldots, (x_1, \ldots, x_r)\},$$

- all witnesses are monomial, i.e., for every $P \in \mathrm{Ass}(I)$ there exists a monomial x^a such that $P = I : x^a$.

$I = (xy, yz, xz) = (x, y) \cap (x, z) \cap (y, z)$
$I^2 = (x^2y^2, xy^2z, x^2yz, y^2z^2, xyz^2, x^2z^2)$

- $I^2 : x^2y = (y, z)$
- $I^2 : y^2x = (x, z)$
- $I^2 : z^2y = (x, y)$
- $I^2 : xyz = (x, y, z)$

$$\mathrm{Ass}(I^2) \subseteq \left\{ \begin{matrix} (x) & (y) & (z) \\ (x, y) & (x, z) & (y, z) \\ & (x, y, z) & \end{matrix} \right\}$$

The set of associated primes of an ideal changes when looking at its powers.

Associated primes of powers of ideals

Example

$$\begin{aligned} P &= (2 \times 2 \text{ minors of a } 3 \times 3 \text{ matrix}) \\ &= (ae - bd, af - cd, \dots) \\ &\subseteq K[a, b, c, d, e, f, g, h, i, j] \end{aligned} \qquad \begin{pmatrix} a & b & c \\ d & e & f \\ h & i & j \end{pmatrix}$$

is a binomial prime ideal.

P^2 has primary decomposition

$$P^2 = (P^2 + (\det M)) \cap (P^2 + \mathfrak{m}).$$

- Powers of prime ideals are not necessarily primary.
- Associated primes can change when looking at powers of an ideal:

$$\begin{aligned} \mathrm{Ass}(P) &= \{P\}, \\ \mathrm{Ass}(P^2) &= \{P, \mathfrak{m}\}. \end{aligned}$$

Theorem (Kim, Swanson, 2019)

Let $m \geq 3$, $v_1, \ldots, v_m \in \mathbb{N}$. Then there exists a polynomial ring R in $\sum v_i$ variables with a prime ideal P such that for all integers $e \geq 2$, P^e has $\prod v_i$ embedded primes.

Construction of such ideals:

$$
\begin{array}{c}
(x^3 - yz, y^2 - xz, z^2 - x^2 y) \\
\downarrow \textit{spreading of } I \\
(x^3 - yz, y^2 - xz, z^2 - x^2 y, z_1 - z, \ldots, z_{m-3} - z) \\
\downarrow \text{ splitting the variables } x_i \mapsto x_{11} \cdots x_{1v_i} \\
P
\end{array}
$$

Example

There exists a prime ideal P in $11 \cdot 2 = 22$ variables such that P^e has $2^{11} = 2048$ embedded primes for all $e \geq 2$.

Theorem (Brodmann, 1979)

The sequence $(\mathrm{Ass}(I^n))_{n\in\mathbb{N}}$ *stabilizes.*

Definition

stability index of I: smallest $\mathrm{B}^I_= \in \mathbb{N}$ such that for all $n \geq \mathrm{B}^I_=$

$$\mathrm{Ass}(I^n) = \mathrm{Ass}(I^{\mathrm{B}^I_=})$$

- How does the sequence $(\mathrm{Ass}(I^n))_{n\in\mathbb{N}}$ behave? (increasing/decreasing?)
- When does $(\mathrm{Ass}(I^n))_{n\in\mathbb{N}}$ stabilize?
- Can we give an upper bound for $\mathrm{B}^I_=$ for monomial ideals?
- On which parameters does such a bound depend?

Example (Weinstein, Swanson, 2020)

For every $d \in \mathbb{N}$:

$$I = (a^{d+2}y, a^{d+1}by, ab^{d+1}y, b^{d+2}y, a^d b^2 xy) \subseteq K[a, b, x, y]$$

$$\mathrm{Ass}(I^n) = \begin{cases} \{(a, b), (y), (a, b, x)\}, & \text{for } n < d \\ \{(a, b), (y)\}, & \text{for } n \geq d \end{cases}.$$

$$\mathrm{B}^I_= = d$$

degree

Example (Martínez-Bernal, Morey, Villarreal, 2012)

Edge ideals of odd cycles of length $2s+1$:

$$I = (x_1x_2, x_2x_3, x_3x_4, \ldots, x_{2s}x_{2s+1}, x_{2s+1}x_1) \subseteq K[x_1, \ldots, x_{2s+1}].$$

- $n \leq s$: $\operatorname{Ass}(I^n) = \{\text{prime ideals generated by } s+1 \text{ variables}\}$
- $n > s$: $\operatorname{Ass}(I^n) = \operatorname{Ass}(I) \cup \{\mathfrak{m}\}$

$$\mathrm{B}^I_= = s$$

number of generators and variables

persistence index of I: smallest integer $\mathrm{B}^I_{\subseteq}$ such that

$$\mathrm{Ass}(I^n) \subseteq \mathrm{Ass}(I^{n+1}) \text{ for all } n \geq \mathrm{B}^I_{\subseteq}.$$

copersistence index of I: smallest integer $\mathrm{B}^I_{\supseteq}$ such that

$$\mathrm{Ass}(I^n) \supseteq \mathrm{Ass}(I^{n+1}) \text{ for all } n \geq \mathrm{B}^I_{\supseteq}.$$

$$\text{stability index} = \max\{\mathrm{B}^I_{\subseteq}, \mathrm{B}^I_{\supseteq}\}$$

I monomial ideal in $K[x_1, \ldots, x_r]$

- r – number of variables
- s – number of generators
- d – maximal total degree of the generators

Theorem (Hoa, 2006)

- $\mathrm{B}^I_{\subseteq} \leq s^{r+3}(s+r)^4 d^2 \left(2d^2\right)^{s^2-s+1}$
- $\mathrm{B}^I_{\supseteq} \leq d(rs+s+d) \left(\sqrt{r}\right)^{r+1} (\sqrt{2}d)^{(r+1)(s-1)}$

Example

$I = (a^6, b^6, a^5b, ab^5, ca^4b^4, a^4xy^2, b^4x^2y) \subseteq K[a, b, c, x, y]$

- $r = 5,\ s = 7,\ d = 9$
- upper bound $\approx 10^{108}$
- stability index: 4

Theorem (Heuberger, R., Rissner, 2024)

$$\mathrm{B}^{I}_{\supseteq} \leq (rs + r + 2)(\sqrt{r})^{r+2}(d+1)^{rs} := \sigma_1$$

Hoa: $\mathrm{B}^{I}_{\supseteq} \leq d(rs + s + d)\left(\sqrt{r}\right)^{r+1}(\sqrt{2}d)^{(r+1)(s-1)} := \sigma_2$

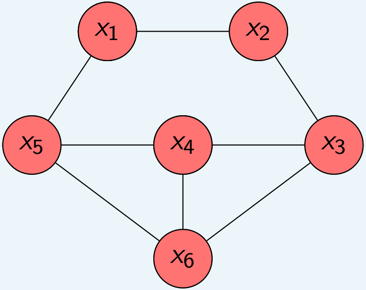
$$\sigma_2 \geq \left(\frac{d\sqrt{2}}{\sqrt{d^2+1}}\right)^{rs} \frac{1}{\sqrt{2r}} \cdot \sigma_1$$

Example

$I = (a^6, b^6, a^5 b, ab^5, ca^4 b^4, a^4 xy^2, b^4 x^2 y) \subseteq K[a, b, c, x, y]$

- $\sigma_1 \approx 3 \cdot 10^{37}$
- $\sigma_2 \approx 3 \cdot 10^{44}$

Squarefree monomial ideals: edge ideals and cover ideals



$J = ($ minimal vertex covers $)$
$= (x_1x_2x_4x_6\ , x_1x_3x_4x_5, \ldots)$

$$\mathrm{Ass}(J\ ) = \Big\{ \begin{array}{ll} (x_1, x_5) & (x_1, x_2) \\ (x_3, x_4) & (x_2, x_3) \\ (x_4, x_5) & (x_3, x_6) \\ (x_5, x_6) & (x_4, x_6) \end{array} \Big\}$$

Connection to graph theory

Let G be a graph and I its edge ideal:

graph theoretical tools		
minimal vertex covers	$\longrightarrow$	minimal associated primes of I
matching number	$\longrightarrow$	$(\mathrm{Ass}(I^n))_{n\in\mathbb{N}}$ is increasing
generalized ear decompositions	$\longrightarrow$	fully describe $(\mathrm{Ass}(I^n))_{n\in\mathbb{N}}$

Let H be a hypergraph and J its cover ideal:

chromatic number	$\longrightarrow$	lower bound for the stability index
graph colorings	$\longrightarrow$	fully describe $(\mathrm{Ass}(J^n))_{n\in\mathbb{N}}$

Some known results about the changes of $\mathrm{Ass}(I^n)$

- edge ideals [Martínez-Bernal, Morey, Villarreal, 2012]
- cover ideals of perfect graphs [Francisco, Hà, Tuyl, 2011]
- ideals with all powers integrally closed [Ratliff, 1984]

$(\mathrm{Ass}(I^n))_{n\in\mathbb{N}}$ is increasing

- ideals can be constructed with
  - $(\mathrm{Ass}(I^n))_{n\in\mathbb{N}}$ not increasing
  - $(\mathrm{Ass}(I^n))_{n\in\mathbb{N}}$ not monotone [McAdam, Eakin, 1979]
  - $\mathrm{B}^I_=$ arbitrarily large [Hà, Nguyen, Trung, Trung, 2021]

- conjecture [J. Herzog]: if I square-free, $\mathrm{B}^I_= \leq r-1$
- upper bound for $\mathrm{B}^I_=$ of general monomial ideals

Algebraic statistics: primary decompositions of conditional independence ideals

What is algebraic statistics?

- many questions in statistics are fundamentally problems of algebra and algebraic geometry
- apply tools from
  - algebraic geometry,
  - commutative algebra,
  - combinatorics, and
  - symbolic computation

  to problems in probability theory, statistics, and their applications

Some history

First connections between algebra and statistics:

- Raj Chandra Bose, 1947: first link between the geometry of finite fields and construction of designs

  > "It is a startling idea that Galois fields might be helpful to provide people with more and better food"– Levi

- Ulf Grenander, 1963: algebraic structures to describe central limit theorems in complex settings
- Persi Diaconis, 1988: representation theoretic methods in the analysis of discrete data

"Algebraic statistics" started with

- Persi Diaconis and Bernd Sturmfels, 1998: Algebraic algorithms for sampling from conditional distributions
- Giovanni Pistone, Eva Riccomagno, and Henry P. Wynn, 2001: Algebraic Statistics

An introductory example (part 1)

- X_1, X_2, X_3 random variables on $\{0, 1\}$
- probability that $X_1 = i$, $X_2 = j$ and $X_3 = k$ is

$$P(X_1 = i, X_2 = j, X_3 = k) =: p_{ijk}$$

- joint distribution of X_1, X_2 and X_3 is a point

$$(p_{000}, p_{100}, p_{010}, p_{001}, p_{110}, p_{101}, p_{011}, p_{111}) \in \mathbb{R}^8$$

probability distribution $\longleftrightarrow$ point

Notation and some definitions

$X = (X_1, \ldots, X_m)$ m-dimensional random vector

- values in $\mathcal{X} = \prod_{i=1}^{m} \mathcal{X}_i$
- assume that the joint probability distribution of X has a density function f

For $A \subseteq [m]$, write

- $X_A := (X_a)_{a \in A}$, and
- $\mathcal{X}_A := \prod_{a \in A} \mathcal{X}_a$

Definition (marginal density)

$$f_A(x) := \int_{\mathcal{X}_{[m]\setminus A}} f(x_A, x_{[m]\setminus A}) d\nu_{[m]\setminus A}(x_{[m]\setminus A}), \quad x_A \in \mathcal{X}_A.$$

Example

$X = (X_1, X_2, X_3)$ discrete random vector,

- X_i takes values in $[r_i]$, $r_i \in \mathbb{N}$
- X takes values in $[r_1] \times [r_2] \times [r_3]$.
- $P(X_1 = i, X_2 = j, X_3 = k) = p_{ijk}$

If $A = \{1, 2\}$, then

$$P(X_1 = i, X_2 = j) = \sum_{k \in [r_3]} p_{ijk} =: p_{ij+}$$

Definition (conditional density)

A, $B \subseteq [m]$ disjoint and $x_B \in \mathcal{X}_B$. The conditional density of X_A given $X_B = x_B$ is

$$f_{A|B}(x_A \mid x_B) = \begin{cases} \frac{f_{A\cup B}(x_A, x_B)}{f_B(x_B)}, & \text{if } f_B(x_B) > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Example

$X = (X_1, X_2, X_3)$ as before, $A = \{1, 2\}$, $B = \{3\}$

$$P(X_1 = i, X_2 = j \mid X_3 = k) = \begin{cases} \frac{p_{ijk}}{p_{++k}}, & \text{if } p_{++k} > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Conditional independence

Definition

A, B, $C \subseteq [m]$ pairwise disjoint; X_A is conditionally independent of X_B *given* X_C, if

$$f_{A\cup B|C}(x_A, x_B \mid x_C) = f_{A|C}(x_A \mid x_C) \cdot f_{B|C}(x_B \mid x_C).$$

Write $X_A \perp\!\!\!\perp X_B \mid X_C$ (sometimes also $A \perp\!\!\!\perp B \mid C$).

If $X_A \perp\!\!\!\perp X_B \mid X_C$ and x_C such that $f_C(x_C) > 0$, then

$$\begin{aligned} f_{A|B\cup C}(x_A \mid x_B, x_C) &= \frac{f_{A\cup B\cup C}(x_A, x_B, x_C)}{f_{B\cup C}(x_B, x_C)} \\ &= \frac{f_{A\cup B|C}(x_A, x_B \mid x_C) f_C(x_c)}{f_{B|C}(x_B \mid x_C) f_C(x_c)} = f_{A|C}(x_A \mid x_C). \end{aligned}$$

"given X_C, knowing X_B does not give any information about X_A"

An introductory example (part 2)

- X_1, X_2, X_3 Markov chain on $\{0,1\}$, i.e., $X_3 \perp\!\!\!\perp X_1 \mid X_2$, or

$$P(X_3 = k \mid X_1 = i, X_2 = j) = P(X_3 = k \mid X_2 = j).$$

- That is,

$$\frac{p_{ijk}}{p_{ij+}} = \frac{p_{+jk}}{p_{+j+}} \qquad \text{for all } i, j, k \in \{0,1\}$$

- ... expanding and simplifying gives

$$p_{000}p_{101} - p_{001}p_{100} = 0, \text{ and}$$
$$p_{010}p_{111} - p_{011}p_{110} = 0.$$

An introductory example (part 2)

A vector $(p_{000}, p_{100}, p_{010}, p_{001}, p_{110}, p_{101}, p_{011}, p_{111}) \in \mathbb{R}^8$ is the probability distribution from the Markov chain model iff:

- $p_{ijk} \geq 0$ for all i, j, $k \in \{0, 1\}$,
- $\sum_{i,j,k} p_{ijk} = 1$,
- $p_{000}p_{101} - p_{001}p_{100} = 0$, and
- $p_{010}p_{111} - p_{011}p_{110} = 0$.

statistical model $\longleftrightarrow$ (semi)algebraic set

Dictionary (Seth Sullivant: Algebraic Statistics, 2018)

Probability/Statistics	**Algebra/Geometry**
probability distribution	point
statistical model	(semi)algebraic set
exponential family	toric variety
conditional inference	lattice points in polytopes
maximum likelihood estimation	polynomial optimization
model selection	geometry of singularities
multivariate gaussian model	spectrahedral geometry
phylogenetic model	tensor networks
MAP estimates	tropical geometry

Conditional independence ideals

Q: Given a list of conditional indepenence statements, what other constraints must the same random vector satisfy?

- assuming we do not know the density (otherwise we could test all constraints)
- Which implications hold regardless of the distribution?
- A few obvious implications:

$$\begin{aligned} X_A \perp\!\!\!\perp X_B \mid X_C &\Longrightarrow X_B \perp\!\!\!\perp X_A \mid X_C \quad \text{Symmetry} \\ X_A \perp\!\!\!\perp X_{B\cup D} \mid X_C &\Longrightarrow X_A \perp\!\!\!\perp X_B \mid X_C \quad \text{Decomposition} \\ &\vdots \end{aligned}$$

- In general, finding such implications is difficult, and
- it is impossible to find a finite set of axioms from which all CI statements can be deduced (Milan Studený, 1992)