

Tools for computing primary decompositions and applications to ideals associated to Bayesian networks

Michael Stillman
Cornell University *

1 Introduction

These notes, written for the CIMPA School on “Systems of polynomial equations” (Argentina, 2003), have two goals: to present the underlying ideas and tools for computing primary decompositions of ideals, and to apply these techniques to a recent interesting class of ideals related to statistics.

Primary decompositions are an important notion both in algebraic geometry and for applications. There are several algorithms available (the two closest to what we present are [GTZ88] and [SY96]). A good overview of the state of the art is the paper [DGP99]. Primary decompositions, and related computations, such as finding minimal and associated primes, the radical of an ideal, and the equidimensional decomposition of an ideal, are all implemented in most specialized computer algebra systems, such as CoCoa ([CNR00]), *Macaulay 2* ([GS]), and Singular ([GPS01]). Several years ago, these algorithms and their implementations could handle only very small examples. Now, with improved implementations, and more efficient computers, larger ideals can be handled.

However, if the number of indeterminates is large, the implemented algorithms often are unable to find a primary decomposition, or even to find the minimal primes. This is the case for many of the ideals associated to Bayesian networks that we consider here.

Our first goal in these lectures is to describe some basic methods for manipulating components of an ideal. We put these together into an algorithm for primary decomposition, but, we hope that some of the students will have ideas about novel ways to combine these techniques to a more efficient algorithm!

Our second goal is to define some interesting ideals, called Markov ideals, associated to a Bayesian network. In applications, Bayesian networks have been used in many ways, e.g. in machine learning, in vision and speech recognition, in attempting to reconstruct gene regulatory networks, and in the analysis of DNA micro-array data. These Markov ideals provide a striking link between multivariate statistics and algebra and geometry. In these lectures, we do little more than provide a glimpse into this very interesting and potentially powerful relationship. Here is one short glimpse: hidden variables in some Markov models correspond to secant loci of Segre embeddings of products of projective spaces (see [GSS] for details). Although this relationship is not yet in these notes, if there is time, we will discuss it at CIMPA, as it is related to some of the computational tools we will cover.

*Supported by the NSF.

These Markov ideals often have many components, and can have relatively complicated primary decompositions. We apply the techniques that we have learned to compute some of these primary decompositions. Instead of giving canned algorithms for computing primary decompositions, we will describe several tools that can be used on a given ideal, to help find the primary decomposition “by hand” (although with the help of a computer algebra system!). It is likely that superior algorithms exist. Our hope is that one of our students will find one!

In the first lecture, we will set up the situation, and describe the first two tools of computing primary decompositions: ideal quotients, and splitting principles. As an example, we will find fixed points of some finite dynamical systems.

In the second lecture, we describe this striking new link between algebra/geometry and statistics. Given a Bayesian network, we can associate an ideal, whose primary decomposition is hard to compute, yet very likely carries interesting information. In this lecture, we define Bayesian networks, and give examples of some of their primary decompositions.

In the third lecture, we describe several more tools for computing primary decompositions. We ask several questions: (1) How do we find zero divisors to use in our splitting principles? (2) How do we detect if an ideal is prime, or primary? The tools we develop include birational maps, and the flattener: a polynomial obtained by analyzing the fibers of a projection map. Both of these techniques rely heavily on a Gröbner basis using product orders. A final technique that we address is removing redundancy in the computation as soon as possible.

In the final lecture, we use the tools that we have described to compute “by hand” (but using Macaulay 2!) some of the more interesting primary decompositions of ideals associated to Bayesian networks. We will also present some open problems related to the primary decompositions of the Markov ideals.

Some good reading to get ready for these lectures include: the first chapter of the forthcoming book by Hal Schenck (see [Sch03]), the book [CLO97], and the paper [DGP99]. For delving more deeply into the Bayesian network material, try reading [GSS], and the references contained in there.

In these preliminary notes, proofs are not included for most of the results. It is a good exercise to try to prove them yourself! During the lectures, we will spend more time using these results than proving them, although we will include some proofs.

Example computer sessions are included for Macaulay 2. This is a system that Dan Grayson and I have been working on for almost ten years now. The system is freely available, and easy to install on most computers. The web page is located at <http://www.math.uiuc.edu/Macaulay2>

2 Lecture #1: Algebraic varieties and components

An excellent place to gain the small amount of background necessary for these lectures is the first chapter of the forthcoming book by Hal Schenck [Sch03] which is available on-line. This is a very nice place to see the basics about primary decomposition, with simple examples. In fact, the first part of the CIMPA lecture will cover some of this material.

Throughout these lectures, let k be a field, and let $R = k[x_1, \dots, x_n]$. If $J = (f_1, \dots, f_r) \subset R$ is an ideal, we let

$$V(J) = \{p \in \bar{k}^n \mid f_1(p) = \dots = f_r(p) = 0\},$$

where \bar{k} is the algebraic closure of k . There is a beautiful dictionary which relates the geometry of $X = V(J)$ to algebraic properties of the ideal J . We refer the reader to Schenck([Sch03]) or Cox-Little-O’Shea [CLO97] for the details.

For example, if $J \subset R$ is a *prime* ideal (that is, $fg \in J$ implies $f \in J$ or $g \in J$), then $V(J)$ is irreducible (that is, cannot be written as the intersection $V(I_1) \cap V(I_2)$ of zero sets which properly contain $V(J)$).

Every ideal J in R has a *primary decomposition*, that is, a decomposition

$$J = Q_1 \cap \dots \cap Q_s,$$

where each Q_i is primary (i.e. if $fg \in Q_i$, then $f \in Q_i$ or $g^N \in Q_i$, for some integer N .) The radical

$$P = \sqrt{Q} = \{g \in R \mid g^N \in Q, \text{ for some } N\}$$

is a prime ideal, and Q is called P -primary.

The primary decomposition is called *irredundant* if each $P_i := \sqrt{Q_i}$ is distinct, and if removing any one term Q_i breaks the equality. Every primary decomposition can be pruned to obtain an irredundant primary decomposition.

If the primary decomposition is irredundant, then the P_1, \dots, P_s are called the *associated primes* of J . This set is independent of the particular (irredundant) primary decomposition. The minimal elements of this set of primes (with respect to inclusion) are called the minimal primes of J . The radical of J is the intersection of these minimal primes. If P is a minimal prime, the corresponding primary ideal is unique (i.e. doesn't depend on the specific irredundant primary decomposition). If P is an associated prime, but not minimal, then P is called an *embedded* prime. The primary ideal of an embedded prime is *not* unique.

Example 2.1. Let $J = (x^2, xy) \subset k[x, y]$. For each $N \geq 1$, we obtain a different primary decomposition of J :

$$J = (x) \cap (x^2, y) = (x) \cap (x^2, xy, y^N).$$

The associated primes are $P_1 = (x)$ and $P_2 = (x, y)$, where P_1 is the only minimal prime, and P_2 is embedded. The primary ideal $Q_1 = (x)$ is the same no matter which primary decomposition we use, but the primary ideal Q_2 of P_2 depends on the decomposition. Geometrically, $V(J)$ is simply the line $x = 0$. Thinking algebraically (or, using schemes), the zero set should really be considered as the union of this line, and a “fat” embedded point at the origin.

In these lectures, what computations concern us? Given J , we would like to be able to compute (in roughly increasing order of difficulty):

- The radical of J .
- The set of minimal primes of J .
- The set of associated primes of J .
- An irredundant primary decomposition of J .

Finding the P -primary component Q of J , where P is a minimal prime, is another useful computation.

Our plan in these lectures is to present techniques to enable us to compute these.

2.1 Tool #1: Ideal quotients

Ideal quotients form one of the most important constructions in ideal theory.

Definition 2.2 (Ideal quotient and saturation). *If $I \subset R$ is an ideal, and $f \in R$, then define the ideal quotient*

$$(I : f) := \{g \in R \mid gf \in I\},$$

and the saturation of I by f :

$$(I : f^\infty) := \{g \in R \mid gf^N \in I, \text{ for some } N\},$$

This somewhat opaque definition gives little clue of their importance.

Lemma 2.3. *Let Q be a P -primary ideal, and let $f \in R$. Then*

- (a) *If $f \notin P$, then $(Q : f) = Q$.*
- (b) *If $f \in P$, but $f \notin Q$, then $(Q : f)$ is P -primary.*
- (c) *If $f \in Q$, then $(Q : f) = (1)$.*

An elementary fact, which follows directly from the definition, is that

$$(J_1 \cap J_2 \cap \cdots \cap J_r) : f = (J_1 : f) \cap \cdots \cap (J_r : f).$$

Lemma 2.4. *If $J = Q_1 \cap \cdots \cap Q_r$ is an irredundant primary decomposition of J , where Q_i is P_i -primary, and if $f \in Q_j$ only if $j \geq s + 1$, then*

$$J : f = (Q_1 : f) \cap \cdots \cap (Q_s : f)$$

is an irredundant primary decomposition of $J : f$.

Saturations have even simpler behavior.

Lemma 2.5. *Let Q be a P -primary ideal, and let $f \in R$. Then*

- (a) *If $f \notin P$, then $(Q : f^\infty) = Q$.*
- (b) *If $f \in P$, then $(Q : f^\infty) = (1)$.*

Lemma 2.6. *If $J = Q_1 \cap \cdots \cap Q_r$ is an irredundant primary decomposition of J , where Q_i is P_i -primary, and if $f \in P_j$ only if $j \geq s + 1$, then*

$$J : f^\infty = Q_1 \cap \cdots \cap Q_s$$

is an irredundant primary decomposition of $(J : f^\infty)$.

This says that, geometrically, the components of $V(J : f^\infty)$ are precisely the components of $V(J)$ which do not lie on the hypersurface $f = 0$.

What makes ideal quotients so useful is that they may be computed using Gröbner bases.

Proposition 2.7. *Let $J \subset R = k[x_1, \dots, x_n]$ be an ideal, where k is a ring (e.g. a field, or a PID), and let $f \in R$. If $L = J + (tf - 1) \subset k[t, x_1, \dots, x_n]$, then*

$$(J : f^\infty) = L \cap k[x_1, \dots, x_n].$$

This is not always the most efficient method to compute saturations. It also doesn't allow one to compute ideal quotients easily. There are (at least) two further ways to compute ideal quotients which are often used: the reverse lexicographic order, and syzygies. We'll describe the method using the reverse lexicographic order, but we'll leave out the syzygy method.

If $f = x_n$ is a variable, and if J is homogeneous, then $(J : x_n)$ and $(J : x_n^\infty)$ may be computed using a single reverse lexicographic Gröbner basis. The key insight is that if $>$ is the term order in the following proposition, then $x_n|g$ if and only if $x_n|in(g)$.

Proposition 2.8 (Bayer). *Let $J \subset k[x_1, \dots, x_n]$ be a homogeneous ideal, and let $>$ be the graded reverse lexicographic order (GrevLex) with $x_1 > \dots > x_n$. If the Gröbner basis of J is*

$$\{g_1, \dots, g_r, h_{r+1}, \dots, h_s\},$$

where $g_i = x_n^{a_i} h_i$, each $a_i > 1$, and x_n does not divide the h_i , then

- (a) $\{x_n^{a_1-1} h_1, \dots, x_n^{a_r-1} h_r, h_{r+1}, \dots, h_s\}$ is a Gröbner basis of $(J : x_n)$, and
- (b) $\{h_1, \dots, h_s\}$ is a Gröbner basis of $J : f^\infty$.

Exercise 2.9. *This idea can be used to compute $J : f$ and $J : f^\infty$ when f is not a variable.*

(a) *Show that if J is homogeneous, and f is homogeneous of degree d , then Bayer's method applied to the homogeneous ideal $J + (f - z)$, where z is a new variable having degree d can be used to compute $J : f$ and $J : f^\infty$.*

(b) *Show how to compute the homogenization of an ideal by using saturation.*

(c) *Show how to use homogenization and the trick in (a), to compute $J : f$ and $J : f^\infty$ when J and f are not necessarily homogeneous.*

Example 2.10. Consider the ideal $J = (c^2 - bd, bc - ad) \subset \mathbb{Q}[a, b, c, d]$. Notice that the plane $c = d = 0$ is contained in the zero set of J . Let's look at this ideal in *Macaulay 2*.

```
i1 : R = QQ[a..d];
i2 : J = ideal(c^2-b*d, b*c-a*d)
o2 : Ideal of R
```

First, here is the primary decomposition of J :

```
i3 : primaryDecomposition J
-- used 0.02 seconds
-- used 0. seconds
-- used 0.09 seconds
o3 = {ideal (d, c), ideal (c^2 - b*d, b*c - a*d, b^2 - a*c)}
```

The reverse lexicographic order is the default in *Macaulay 2*:

```
i4 : gens gb J
o4 = | c2-bd bc-ad b2d-acd |
o4 : Matrix R <--- R
```

```

i5 : J : d
o5 = ideal (c2 - b*d, b*c - a*d, b2 - a*c)
o5 : Ideal of R
i6 : saturate(J,d)
o6 = ideal (c2 - b*d, b*c - a*d, b2 - a*c)
o6 : Ideal of R
i7 : J == intersect(ideal(c,d),J:d)
o7 = true

```

2.2 Tool #2: Splitting principles

The key technique on which almost all algorithms for primary decomposition are based is the following very simple lemma.

Proposition 2.11. *If $(J : f^\infty) = (J : f^\ell)$, then*

$$J = (J : f^\infty) \cap (J, f^\ell).$$

Proof. Suppose that $g \in (J : f^\infty)$ and also that $g \in (J, f^\ell)$. We want to show that $g \in J$. So $g = a + bf^\ell$, for some $a \in J$ and $b \in R$. However, $gf^\ell \in J$, so $bf^{2\ell} \in J$. Therefore $b \in (J : f^\infty) = (J : f^\ell)$, and so $g \in J$. \square

If we are only interested in finding all of the minimal primes, we may take the radicals of both sides to obtain: for any $f \in R$,

$$\sqrt{J} = \sqrt{J : f^\infty} \cap \sqrt{J, f}.$$

Another useful splitting formula is: if $f_1 f_2 \dots f_r \in J$, then

$$\sqrt{J} = \sqrt{J, f_1} \cap \dots \cap \sqrt{J, f_r}.$$

If we have a way of finding, given an ideal J , a zero divisor $f \bmod J$ (so that $J : f \neq J$), then we may build a recursive algorithm to compute a decomposition of J .

2.3 An example: Finite dynamical systems

As an example, let's consider finite dynamical systems: given a prime number p , let \mathbb{F}_p be the finite field with p elements, let $R = \mathbb{F}_p[x_1, \dots, x_n]$, and let $F : \mathbb{F}_p^n \rightarrow \mathbb{F}_p^n$ be defined by

$$a = (a_1, \dots, a_n) \mapsto (f_1(a), \dots, f_n(a)),$$

where $f_i \in R$.

By iterating F , we obtain a directed graph whose vertices are the p^n points of \mathbb{F}_p^n , and there are directed edges from a to $F(a)$.

In this example, we are interested in finding the fixed points of F , or more generally, of F^r (apply F r times) for some integer r . The fixed points of F are the zeros of the ideal $J = (x_1 - f_1, \dots, x_n - f_n)$ which have all coordinates in \mathbb{F}_p . The problem is that there may be solutions

over an extension field of \mathbb{F}_p and we are not particularly interested in these solutions. Notice that all of the elements x of \mathbb{F}_p satisfy $x^p - x = 0$. So, if we include these polynomials, then our zero set will only contain elements of the field we are interested in.

As an added bonus, the resulting ideal is radical, and so we don't have to worry about embedded components or primary components:

Lemma 2.12. *Let $J = (g_1, \dots, g_s, x_1^p - x_1, \dots, x_n^p - x_n) \subset k[x_1, \dots, x_n]$. For any choice of g_i 's, $J = \sqrt{J}$.*

We may use any of these splitting principles to compute the minimal primes (and therefore the primary decomposition) of J , since we have many zero-divisors around: each x_i is (potentially) a zero-divisor!

Example 2.13. Let $R = k[x_1, \dots, x_4]$, where $k = \mathbb{F}_2$. Let $F : k^4 \rightarrow k^4$.

The associated directed graph has $2^4 = 16$ nodes. Let's find the fixed points of one such finite dynamical system, with the aid of *Macaulay 2*. In such a small example, we can compute the fixed points by hand. For larger examples, e.g. $p = 5$, $n = 20$, this is not so easy!

```
i8 : R = ZZ/2[x_1 .. x_4];
i9 : L = ideal(x_1^2 + x_1, x_2^2 + x_2, x_3^2 + x_3, x_4^2 + x_4);
o9 : Ideal of R
Our sample finite dynamical system:
i10 : F = matrix {{x_1*x_2*x_4+x_1+x_4,
                  x_1*x_3*x_4+x_2*x_4+x_2,
                  x_1*x_3+x_3*x_4+x_3,
                  x_1*x_3*x_4+x_1+x_4}}
o10 = | x_1x_2x_4+x_1+x_4 x_1x_3x_4+x_2x_4+x_2 x_1x_3+x_3x_4+x_3 x_1x_4
      |
      1 4
o10 : Matrix R <--- R
```

Fixed points of F are precisely the zeros of the following ideal.

```
i11 : J = L + ideal (vars R - F);
o11 : Ideal of R
i12 : transpose gens gb J
o12 = {-1} | x_1+x_4 |
      {-2} | x_4^2+x_4 |
      {-2} | x_3x_4+x_1 |
      {-2} | x_2x_4+x_3x_4 |
      {-2} | x_3^2+x_3 |
      {-2} | x_2^2+x_2 |
      6 1
o12 : Matrix R <--- R
```

Although we could solve these equations by hand, we instead blindly follow the recursion using indeterminates as (potential) zero divisors. We start with x_1 .

```
i13 : J1 = J : x_1
o13 = ideal (x_4 + 1, x_3 + 1, x_2 + 1, x_1 + 1)
      4 3 2 1
o13 : Ideal of R
```

```
i14 : J2 = ideal gens gb(J + ideal(x_1))
```

```
o14 = ideal (x2, x2, x2 + x2, x2 + x2)
           4 1 3 3 2 2
```

```
o14 : Ideal of R
```

The intersection of these ideals is J .

```
i15 : J == intersect(J1,J2)
```

```
o15 = true
```

The first ideal is already linear, so its zero set is a point. From the description of J_2 we could write down the rest of the solutions, but let's continue. Split using x_3 :

```
i16 : J21 = J2 : x_3
```

```
o16 = ideal (x2, x2 + 1, x2, x2 + x2)
           4 3 1 2 2
```

```
o16 : Ideal of R
```

```
i17 : J22 = ideal gens gb(J2 + ideal(x_3))
```

```
o17 = ideal (x2, x2, x2, x2 + x2)
           4 3 1 2 2
```

```
o17 : Ideal of R
```

Now we can split each of these using x_2 , obtaining 5 solutions total. Already, one can imagine ways to improve the efficiency of even this small example. For larger problems, these improvements can make the difference between obtaining an answer and waiting forever!

We could have computed this directly in *Macaulay 2*. The `decompose` routine provides the list of minimal primes. The `primaryDecomposition` routine provides an irredundant primary decomposition.

```
i18 : C = decompose J;
```

Display these ideals:

```
i19 : C/(I -> (<< toString I << endl));
ideal(x_2,x_3,x_4,x_1)
ideal(x_2+1,x_3,x_4,x_1)
ideal(x_2+1,x_3+1,x_4,x_1)
ideal(x_2,x_3+1,x_4,x_1)
ideal(x_2+1,x_3+1,x_4+1,x_1+1)
```

3 Lecture #2: Bayesian networks and Markov ideals

The emerging field of *algebraic statistics* [GPW01] advocates polynomial algebra as a tool in the statistical analysis of experiments and discrete data. Statistics textbooks define a *statistical model* as a family of probability distributions, and a closer look reveals that these families are often algebraic varieties: they are the zeros of some polynomials in the probability simplex [GHKM01], [SS00].

We begin by reviewing the general algebraic framework for independence models presented in [Stu02, §8]. Let X_1, \dots, X_n be discrete random variables where X_i takes values in the finite set $[d_i] = \{1, 2, \dots, d_i\}$. We write $D = [d_1] \times [d_2] \times \dots \times [d_n]$ so that \mathbb{C}^D denotes the complex

vector space of n -dimensional tables of format $d_1 \times \cdots \times d_n$. We introduce an indeterminate $p_{u_1 u_2 \cdots u_n}$ which represents the probability of the event $X_1 = u_1, X_2 = u_2, \dots, X_n = u_n$. These indeterminates generate the ring $\mathbb{C}[D]$ of polynomial functions on the space of tables \mathbb{C}^D .

A *conditional independence statement* has the form

$$A \text{ is independent of } B \text{ given } C \quad (\text{in symbols: } A \perp\!\!\!\perp B \mid C) \quad (1)$$

where A, B and C are pairwise disjoint subsets of $\{X_1, \dots, X_n\}$. If C is empty then (1) means that A is independent of B . By [Stu02, Proposition 8.1], the statement (1) translates into a set of homogeneous quadratic polynomials in $\mathbb{C}[D]$, and we write $I_{A \perp\!\!\!\perp B \mid C}$ for the ideal generated by these polynomials.

The following example gives the basic idea and method for finding these ideals.

Example 3.1. Let X_1, X_2, X_3 be three random variables, with $d_1 = d_3 = 2$ and $d_2 = 3$. Let's write down the ideal in $k[p_{u_1 u_2 u_3}]$ (12 variables) which defines the set of probability distributions which satisfy $X_1 \perp\!\!\!\perp X_2 \mid X_3$.

A probability distribution satisfies this independence condition if

$$Pr(X_1 = u_1, X_2 = u_2 \mid X_3 = u_3) = Pr(X_1 = u_1 \mid X_3 = u_3)Pr(X_2 = u_2 \mid X_3 = u_3),$$

for all choices of $u_i \in [d_i]$. By removing the conditional probabilities, and multiplying by $Pr(X_3 = u_3)$, we obtain

$$p_{++u_3} p_{u_1 u_2 u_3} = p_{u_1 + u_3} p_{+u_2 u_3},$$

where we have replaced Pr by p , and a “+” means sum over all possible values in that variable (i.e. marginalize over that variable). For example,

$$p_{1+2} = p_{112} + p_{122} + p_{132}.$$

It is a simple exercise in determinants to show that the ideal generated by

$$\{p_{++u_3} p_{u_1 u_2 u_3} - p_{u_1 + u_3} p_{+u_2 u_3} \mid \text{all } u_1, u_2, u_3\},$$

is the same as the ideal generated by the six 2 by 2 minors of the matrices M_1 and M_2 , where

$$M_i = \begin{pmatrix} p_{11i} & p_{12i} & p_{13i} \\ p_{21i} & p_{22i} & p_{23i} \end{pmatrix}$$

Note that all 12 indeterminates appear, and each matrix has 6 of them.

The general case goes the same way: The ideal $I_{A \perp\!\!\!\perp B \mid C}$ is generated by the 2 by 2 minors of matrices M_i , for $i = 1..c$, where c is the number of possible values of C . Each matrix is obtained by making an $a \times b$ matrix where the j, k th entry is the linear polynomial in the $p_{u_1 \dots u_n}$ which represents $Pr(A = j, B = k, C = i)$.

Proposition 3.2. *For any choice of A, B , and C , the ideal $I_{A \perp\!\!\!\perp B \mid C}$ is prime.*

The interesting part begins when we have more than one independence statement.

Definition 3.3. *If $\mathcal{M} = \{A_1, A_2, \dots, A_r\}$ is a set of independence statements, define*

$$I_{\mathcal{M}} = I_{A_1} + \cdots + I_{A_r}.$$

Example 3.4 (The contraction lemma). In statistics, there is a lemma that says that any probability distribution which satisfies the two independence statements $X_1 \perp\!\!\!\perp X_2 \mid X_3$, and $X_2 \perp\!\!\!\perp X_3$, also satisfies $X_2 \perp\!\!\!\perp \{X_1, X_3\}$.

In this example, we investigate the algebraic analog of this statement. Let $\mathcal{M} = \{X_1 \perp\!\!\!\perp X_2 \mid X_3, \{X_2 \perp\!\!\!\perp X_3\}\}$. Let's suppose for now that $d_1 = d_2 = d_3 = 2$, i.e. we have three binary random variables. The first independence statement translates into two quadratics:

$$\phi_1 = \det \begin{pmatrix} p_{111} & p_{121} \\ p_{211} & p_{221} \end{pmatrix}, \quad \phi_2 = \det \begin{pmatrix} p_{112} & p_{122} \\ p_{212} & p_{222} \end{pmatrix}$$

The second statement translates into a single determinant:

$$\phi = \det \begin{pmatrix} p_{+11} & p_{+12} \\ p_{+21} & p_{+22} \end{pmatrix},$$

where for example $p_{+11} = p_{111} + p_{211}$.

So $I_{\mathcal{M}} = (\phi_1, \phi_2, \phi)$.

If we change coordinates in $\mathbb{C}[D]$ by replacing each unknown p_{1jk} by $p_{+jk} = \sum_{i=1}^{i=d_1} p_{1jk}$, this ideal $I_{\mathcal{M}}$ transforms to a *binomial* ideal in $\mathbb{C}[D]$, i.e. generated by polynomials which are differences of two monomials. Binomial ideals enjoy many nice properties. For instance, all of their reduced Gröbner bases consist of binomials, and they have primary decompositions where each associated prime and primary ideal is binomial. For more details, see [ES96].

The algebraic analog of the contraction lemma is the primary decomposition of this ideal. The ideal $I_{\mathcal{M}}$ has 3 components in its primary decomposition (all prime).

$$I_{\mathcal{M}} = P_1 \cap P_2 \cap I_{X_2 \perp\!\!\!\perp \{X_1, X_3\}},$$

where $P_1 = (p_{+11}, p_{+21}, \phi_2)$, and $P_2 = (p_{+12}, p_{+22}, \phi_1)$. This implies that any probability distribution which satisfies the two independence statements M also satisfies the statement: $X_2 \perp\!\!\!\perp \{X_1, X_3\}$. The algebraic picture is more complicated: outside of the probability simplex, these two zero sets differ.

As a warmup for computing primary decompositions later, try

Exercise 3.5. (a) Prove that this is a primary decomposition of $I_{\mathcal{M}}$.

(b) Consider the same \mathcal{M} , but now suppose that $d_1 = d_2 = 2$ and $d_3 = 3$. Write down the ideal $I_{\mathcal{M}}$ and find a primary decomposition for $I_{\mathcal{M}}$. Is this ideal radical? What if $d_3 \geq 4$?

3.1 Bayesian networks and associated ideals

A *Bayesian network* is an acyclic directed graph G with vertices X_1, \dots, X_n .

For a given node X_i , let $pa(X_i)$ denote the set of parents of vertex X_i in G , and let $nd(X_i)$ be the set of non-descendants of X_i , *excluding* the parents of X_i . (A non-descendent of X_i is a vertex X_j such that there is no directed path from X_i to X_j .)

The *local Markov property* on G is the set of independence statements

$$\text{local}(G) = \{X_i \perp\!\!\!\perp nd(X_i) \mid pa(X_i) : i = 1, 2, \dots, n\},$$

The *global Markov property*, $\text{global}(G)$, is the set of independence statements $A \perp\!\!\!\perp B \mid C$, for any triple A, B, C of subsets of pairwise disjoint vertices of G such that A and B are *d-separated* by C .

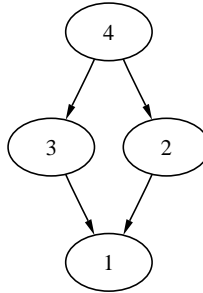


Figure 1: A Bayesian network on 4 vertices

The notion of d -separated (“directed separated”) is a bit technical. Since we don’t really need the definition for these lectures, we refer to [GSS] or to [Lau96] for the definition.

For any Bayesian network G , we have $local(G) \subset global(G)$. Therefore we have inclusions $I_{local(G)} \subset I_{global(G)}$, and $V_{global(G)} \subset V_{local(G)}$.

Example 3.6. Let G be the network on four binary random variables shown in 1.

We (Luis Garcia, Bernd Sturmfels and myself) have written code in *Macaulay 2* and in *Singular* for computing these independence conditions, and the resulting ideals. These files will be available for your use.

```
i20 : load "markov.m2"
```

A graph is input as a list of parents of each node. The above graph is represented as the following list.

```
i21 : G = {{},{1},{1},{2,3}};
```

The Markov conditions come as a list of triples of sets of integers. Each triple represents a single independence statement.

```
i22 : LM = localMarkovConditions G;
```

```
i23 : LM/print;
{Set {1}, Set {4}, Set {2, 3}}
{Set {3}, Set {2}, Set {4}}
```

```
i24 : GM = globalMarkovConditions G;
```

```
i25 : GM/print;
{Set {1}, Set {4}, Set {2, 3}}
{Set {3}, Set {2}, Set {4}}
```

Note that for this example, $local(G)$ and $global(G)$ are both the same set:

$$\{1 \perp\!\!\!\perp 4 \mid \{2, 3\}, 2 \perp\!\!\!\perp 3 \mid 4\}.$$

```
i26 : R = probring(2,2,2,2);
```

```
i27 : numgens R
```

```
o27 = 16
```

```

i28 : gens R
o28 = {p1,1,1,1, p1,1,1,2, p1,1,2,1, p1,1,2,2, p1,2,1,1, p1,2,1,2, p1,2,2,1, p1,2,2,2, ...}
o28 : List
i29 : C = indepMatrices((2,2,2,2),set{1},set{4},set{2,3});
i30 : C2 = indepMatrices((2,2,2,2),set{2},set{3},set{4});
i31 : C = join(C,C2);
i32 : C/(m -> (<< m << endl << endl));
| p_(1,1,1,1) p_(1,1,1,2) |
| p_(2,1,1,1) p_(2,1,1,2) |
| p_(1,1,2,1) p_(1,1,2,2) |
| p_(2,1,2,1) p_(2,1,2,2) |
| p_(1,2,1,1) p_(1,2,1,2) |
| p_(2,2,1,1) p_(2,2,1,2) |
| p_(1,2,2,1) p_(1,2,2,2) |
| p_(2,2,2,1) p_(2,2,2,2) |
| p_(1,1,1,1)+p_(2,1,1,1) p_(1,1,2,1)+p_(2,1,2,1) |
| p_(1,2,1,1)+p_(2,2,1,1) p_(1,2,2,1)+p_(2,2,2,1) |
| p_(1,1,1,2)+p_(2,1,1,2) p_(1,1,2,2)+p_(2,1,2,2) |
| p_(1,2,1,2)+p_(2,2,1,2) p_(1,2,2,2)+p_(2,2,2,2) |

```

The ideal, after making the change of coordinates discussed above (so that p_{1111} refers to p_{+111} , and p_{2111} still refers to p_{2111}):

```

i33 : J = trim localIdeal G;
o33 : Ideal of R
i34 : transpose generators J
o34 = {-2} | p_(1,2,2,2)p_(2,2,2,1)-p_(1,2,2,1)p_(2,2,2,2) |
        {-2} | p_(1,2,1,2)p_(2,2,1,1)-p_(1,2,1,1)p_(2,2,1,2) |
        {-2} | p_(1,1,2,2)p_(2,1,2,1)-p_(1,1,2,1)p_(2,1,2,2) |
        {-2} | p_(1,1,1,2)p_(2,1,1,1)-p_(1,1,1,1)p_(2,1,1,2) |
        {-2} | p_(1,1,2,2)p_(1,2,1,2)-p_(1,1,1,2)p_(1,2,2,2) |
        {-2} | p_(1,1,2,1)p_(1,2,1,1)-p_(1,1,1,1)p_(1,2,2,1) |
o34 : Matrix R <--- R
        6      1

```

The ideal J is minimally generated by 6 binomial quadrics.

One of the most useful aspects of Bayesian networks is that they provide a useful factorization of the joint probability distribution of the n random variables. In this example, note that

$$\begin{aligned}
Pr(X_1 = u_1, \dots, X_4 = u_4) &= \\
&Pr(X_4 = u_4) \times Pr(X_3 = u_3 \mid X_4 = u_4) \times Pr(X_2 = u_2 \mid X_3 = u_3, X_4 = u_4) \\
&\times Pr(X_1 = u_1 \mid X_2 = u_2, X_3 = u_3, X_4 = u_4) \\
&= Pr(X_4 = u_4) \times Pr(X_3 = u_3 \mid X_4 = u_4) \times Pr(X_2 = u_2 \mid X_4 = u_4) \\
&\times Pr(X_1 = u_1 \mid X_2 = u_2, X_3 = u_3)
\end{aligned}$$

If we set $Pr(X_4 = 1) := a$ and $Pr(X_4 = 2) := 1 - a$, and similarly let $Pr(X_3 = 1 \mid X_4 = k) := b_k$, let $Pr(X_2 = 1 \mid X_4 = k) := c_k$, and $Pr(X_1 = 1 \mid X_2 = j, X_3 = k) := d_{jk}$, then the joint

probabilities factor. For example, $p_{1111} = ab_1c_1d_{11}$, $p_{1112} = (1-a)b_2c_2d_{11}$, $p_{1121} = a(1-b_1)c_1d_{12}$, and so on. Instead of requiring 15 parameters, such a probability distribution may be specified using 10 numbers. This is a small example; when the number of vertices is large, and the graph is sparse, the savings is dramatic.

If we denote $\mathbb{C}[E] := \mathbb{C}[a, b_1, b_2, c_1, c_2, d_{11}, \dots, d_{22}]$, we may define a ring map

$$\Phi : \mathbb{C}[D] \longrightarrow \mathbb{C}[E].$$

In what follows we shall assume that every edge (i, j) of the Bayesian network G satisfies $i > j$. In particular, the node 1 is always a sink and the node n is always a source.

For any integer $r \in [n]$ and $u_i \in [d_i]$ as before, we abbreviate the *marginalization* over the first r random variables as follows:

$$p_{++\dots+u_{r+1}\dots u_n} := \sum_{i_1=1}^{d_1} \sum_{i_2=1}^{d_2} \cdots \sum_{i_r=1}^{d_r} p_{i_1 i_2 \dots i_r u_{r+1} \dots u_n}.$$

This is a linear form in our polynomial ring $\mathbb{C}[D]$. We denote by \mathbf{p} the product of all of these linear forms.

As in the example, given a Bayesian network G , we obtain a factorization map:

$$\Phi : \mathbb{C}[D] \longrightarrow \mathbb{C}[E].$$

The main theorem, which is the algebraic analog of the factorization for the joint probabilities for a Bayesian network is the following:

Theorem 3.7. *The prime ideal $\ker(\Phi)$ is a minimal primary component of both of the ideals $I_{\text{local}(G)}$ and $I_{\text{global}(G)}$. More precisely,*

$$(I_{\text{local}(G)} : \mathbf{p}^\infty) = (I_{\text{global}(G)} : \mathbf{p}^\infty) = \ker(\Phi). \quad (2)$$

For a proof, see [GSS].

This result suggests many questions, most of them unsolved. For example:

Problem 3.8. *Find conditions on G so that $I_{\text{global}(G)}$ is a prime ideal (and therefore equal to $\ker \Phi$).*

Problem 3.9. *Find the primary decomposition of $I_{\text{local}(G)}$ or of $I_{\text{global}(G)}$.*

In the remainder of these lectures, we will develop the tools needed to answer these questions computationally, for small networks G .

4 Lecture #3: Tools for computing primary decompositions

In order to use the techniques we have already considered to make an algorithm for computing a primary decomposition, we must answer these questions three.

- Question #1: How do we find zero divisors to use with one of our splitting principles?
- Question #2: How can we detect that an ideal is prime or primary?

- Question #3: Practice shows that the splitting tree is highly redundant. How should we fight this problem?

We will provide answers to these questions. But: keep your mind open. You might find better methods yourself!

Example 4.1. As a running example throughout this lecture, let's consider the simple example which occurred in the contraction lemma in the second lecture. This is an ideal generated by 3 quadrics, in 8 unknowns. Let's rename the unknowns so that we can avoid indices.

```
i35 : R = QQ[a..h];
i36 : J = ideal(a*d-b*c, e*h-f*g, a*f-b*e);
o36 : Ideal of R
```

Just so we know the answer ahead of time, here is the primary decomposition:

```
i37 : (primaryDecomposition J)/print;
-- used 0.02 seconds
-- used 0.01 seconds
-- used 0.22 seconds
ideal (b, a, f*g - e*h)
ideal (f, e, b*c - a*d)
ideal (f*g - e*h, d*g - c*h, b*g - a*h, d*e - c*f, b*e - a*f, b*c - a*d)
```

There are three primary components. In many ways, this is too simple of an example: all of the components have the same dimension (5), and all of the primary components are prime, so this is a radical ideal. The example still provides a good picture of the different tools and also some of the problems which occur.

4.1 Finding zero divisors

Given an ideal $J = (f_1, \dots, f_r) \subset k[x_1, \dots, x_n]$, how can we find a zero divisor $g \bmod J$ (i.e. an element g for which $J : g \neq J$)? One method that often works is to examine the generators f_i and see if they factor. If so, use a factor as the zero divisor g . Often no f_i will factor. In this case, one may start computing a Gröbner basis, and examine each new Gröbner basis element g_i . If g_i factors, use this factorization to split the ideal (This is the basic description of what is known as the factoring Gröbner basis algorithm). The exact details of how best to use this are not clear, and vary with the problem domain. There is definitely room for improvement here in existing algorithms!

Suppose that you cannot find a factor with one of these methods, or, perhaps, are unwilling or unable to look there for zero divisors? What then? Our answer is obtained by analyzing projection maps.

4.2 Projections and elimination of variables

Let $R = k[x] = k[x_1, \dots, x_n]$, where k is a field. Choose a subset of variables

$$t = \{t_1, \dots, t_d\} \subset x = \{x_1, \dots, x_n\},$$

and let $u = x \setminus t$. The inclusion $k[t] \subset k[u, t] = k[x]$ corresponds geometrically to the projection map $k^n \rightarrow k^d$ defined by sending a point $(u, t) = (u_1, \dots, u_{n-d}, t_1, \dots, t_d)$ to $t \in k^d$. The map of rings $\phi : k[t] \rightarrow k[u, t]/J$ corresponds to the projection map $\pi : V(J) \subset k^n \rightarrow k^d$, and the map of rings $k[t]/J_1 \hookrightarrow k[u, t]/J$ corresponds to the projection map $\pi : V(J) \rightarrow V(J_1) = \overline{\pi(V(J))}$, where

$J_1 = \ker(\phi)$. If J is not a radical ideal, or k is not an algebraically closed field such as \mathbb{C} , then this correspondence between the algebra and geometry needs to be defined more carefully: this is where schemes enter the algebraic geometry picture. For us though, we will think geometrically, but work algebraically, and so we won't be concerned with these subtleties.

Recall that we can compute $L = \ker(\phi)$ by using Gröbner bases. A term order on $k[x, t]$ is called an *elimination order* (eliminating x) if $\text{in}(f) \in k[t]$ implies that $f \in k[t]$.

Proposition 4.2. *If $>$ is an elimination order eliminating u , and $J \subset k[u, t]$ is an ideal, with Gröbner basis $\{f_1, \dots, f_r, h_1, \dots, h_s\}$, where $h_i \in k[t]$, but each $f_i \notin k[t]$, then $\{h_1, \dots, h_s\}$ is a Gröbner basis (and therefore a generating set) of $J_1 = J \cap k[t]$.*

For the purpose of analyzing projection maps, the *product order* $u \gg t$ is a good choice (this is sometimes called a block order): $u^a t^b > u^c t^d$ if $u^a >_{\text{grevlex}} u^c$, or $u^a = u^c$ and $t^b >_{\text{grevlex}} t^d$.

Example 4.3. Continuing Example 4.1, suppose that $t = \{a, b, c, d\}$ and $u = \{e, f, g, h\}$.

```
i38 : R1 = QQ[e,f,g,h,a,b,c,d, MonomialOrder=>ProductOrder{4,4}];
i39 : L = substitute(J,R1)
o39 = ideal (- b*c + a*d, - f*g + e*h, - e*b + f*a)
o39 : Ideal of R1
i40 : transpose gens gb L
o40 = {-2} | bc-ad |
      {-2} | eb-fa |
      {-2} | fg-eh |
      {-3} | ead-fac |
      4      1
o40 : Matrix R1 <--- R1
```

So $J \cap k[a, b, c, d] = (ad - bc)$.

4.3 Tool: Birational projections

Suppose that J contains an element f which is linear in a variable, say, x_1 . Write $f = gx_1 + h$, where g, h don't involve x_1 . If g is a non-zero divisor on J , then the projection map $k[t]/J_1 \rightarrow k[t, x_1]/J$ is called *birational* (where $t = \{x_2, \dots, x_n\}$ and $J_1 = \ker(k[t] \rightarrow k[x_1, t]/J)$). Geometrically, this means that for almost all points p of $V(J_1) \subset k^{n-1}$, there is a unique point $(p_1, p) \in V(J)$ which maps to it. If $g(p) \neq 0$, then this value is $p_1 = -\frac{h(p)}{g(p)}$.

Birational maps are well-behaved with respect to primary decompositions:

Proposition 4.4. *Let $J \subset k[x_1, \dots, x_n]$ be an ideal, containing a polynomial $f = gx_1 + h$, with g, h not involving x_1 , and g a non-zero divisor modulo J . Let $J_1 = J \cap k[x_2, \dots, x_n]$ be the elimination ideal. Then*

- (a) $J = (\langle J_1, gx_1 + h \rangle : g^\infty)$,
- (b) J is prime if and only if J_1 is prime.
- (c) J is primary if and only if J_1 is primary.
- (d) Any irredundant primary decomposition of J_1 lifts to an irredundant primary decomposition of J .

This tool may often be used to prove that an ideal is prime (if it is!), and can sometimes simplify the work to look for zero divisors. However, caution is required: the resulting ideal J_1 , although it is an ideal in one fewer variable, can often be much more complicated than J .

Example 4.5. Continuing Example 4.1, all variables occur linearly, and so we may choose any one we wish, e.g. a . The corresponding coefficient is d .

```
i41 : use R;
```

In this example, d is not a zero divisor:

```
i42 : J : d == J
```

```
o42 = true
```

We use the *Macaulay2* “eliminate” package for eliminating variables.

```
i43 : load "eliminate.m2"
```

```
i44 : I1 = eliminate(J,a)
```

```
o44 = ideal (f*g - e*h, b*d*e - b*c*f)
```

```
o44 : Ideal of R
```

The variable f occurs linearly, with coefficient g . It so happens that g is also a non-zero-divisor:

```
i45 : I1 : g == I1
```

```
o45 = true
```

So I_1 is birational to

```
i46 : I2 = eliminate(I1,f)
```

```
o46 = ideal(b*d*e*g - b*c*e*h)
```

```
o46 : Ideal of R
```

This single element has three factors:

```
i47 : factor I2_0
```

```
o47 = (b)(- d*g + c*h)(e)(-1)
```

```
o47 : Product
```

The original ideal J is birational to I_2 . Therefore, the original ideal has three components, all prime. We could use this factorization to produce the three primary components. Exercise: how?

4.4 Tool: The flattener of a projection

Let $J \subset k[x_1, \dots, x_n]$ be an ideal. A set of variables $t = \{x_{i_1}, \dots, x_{i_d}\}$ is called a *maximal independent set* if $J \cap k[t] = (0)$ and t has maximal cardinality over all such subsets with this property.

Proposition 4.6. *Let $in(J)$ be the initial monomial ideal of $J \subset k[x]$ with respect to some arbitrary term order. Then every maximal independent set of $in(J)$ is also a maximal independent set of J .*

The cardinality d of a maximal independent set of J is called the *dimension* of J .

Geometrically, if $J \cap k[t] = (0)$, the map $V(J) \rightarrow k^d$ is *dominant*, i.e. the closure of the image is all of k^d . In this case, every component of J which also maps dominantly to k^d must have the same dimension d as J . A component of J which maps into a subvariety of k^d (algebraically: a primary ideal Q of J for which $Q \cap k[t] \neq (0)$) can either have dimension d , or have smaller dimension.

Suppose that $t \subset x$ is a maximal independent set for $in(J)$ and therefore for J , let $u = x \setminus t$, and let $>$ be the product order $u \gg t$ defined above. Let $\{g_1, \dots, g_r\}$ be a reduced Gröbner basis for the ideal J , where

$$g_i = \alpha_i(t)u^{A_i} + \text{lower terms in } u \text{ variables.}$$

Since $J \cap k[t] = (0)$, each of the monomials $u^{A_i} \neq 1$. Define $in_u(J) = (u^{A_1}, \dots, u^{A_r}) \subset k[u]$.

Let $h \in k[t]$ be any non-zero element such that for each minimal generator u^A of the monomial ideal $in_u(J)$, there is a Gröbner basis element g_i of J , such that $u^{A_i} = u^A$, and $\alpha_i(t)|h(t)$. For example, we could take $h = lcm\{\alpha_1, \dots, \alpha_r\} \in k[t]$. Any such element h is called a *flattener* of the inclusion $k[t] \subset k[u, t]/J$.

(The reason that h is called a flattener, comes from commutative algebra. One can prove that the inclusion of localized rings $k[t]_h \subset k[u, t]_h/J$ is a flat extension. Caution though: our element h enjoys more properties than an arbitrary element that satisfies this flatness).

Proposition 4.7. *If $h \in k[t]$ is a flattener for the inclusion $k[t] \hookrightarrow k[u, t]/J$ defined above, and if $J = Q_1 \cap \dots \cap Q_m$ is an irredundant primary decomposition such that $Q_i \cap k[t] = (0)$, iff $1 \leq i \leq \ell$, then*

$$(J : h^\infty) = Q_1 \cap \dots \cap Q_\ell$$

is an irredundant primary decomposition of the saturation $J : h^\infty$.

Furthermore, if $d = \dim k[x, t]/J$, then every component Q_1, \dots, Q_ℓ dominates k^d , and has dimension d .

Algebraically, h allows us to compute the “generic fiber”:

Proposition 4.8. *If $h \in k[t]$ is a flattener as defined above, then*

$$(J : h^\infty) = Jk(t)[u] \cap k[u, t].$$

This is great! It allows us to use the results from David Cox’s lectures on computing the primary decomposition of zero dimensional ideals, since the extended ideal $Jk(t)[u]$ is a zero dimensional ideal in $k(t)[u]$. Given a primary decomposition of this extended ideal, we get a primary decomposition of $(J : h^\infty)$:

Proposition 4.9. *If $Jk(t)[u] \cap k[u, t] = J$, and if $\tilde{Q}_1 \cap \dots \cap \tilde{Q}_s$ is an irredundant primary decomposition of $Jk(t)[u]$, then if $Q_i = \tilde{Q}_i \cap k[u, t]$, then $J = Q_1 \cap \dots \cap Q_s$ is an irredundant primary decomposition of J .*

Applying this proposition to $J : h^\infty$ (instead of to J) gives a primary decomposition of $J : h^\infty$, if h is the flattener defined above.

Example 4.10. Let’s use the flattener method to compute the primary decomposition of the ideal of Example 4.1. Even though this is a simple example, it highlights several possible efficiency problems. First, we find a maximal independent set of J .

```
i48 : independentSets J
o48 = {a*b*d*f*h, a*c*d*f*h, a*c*e*f*h, c*d*e*f*h, a*b*d*g*h, a*c*d*g* ...
o48 : List
```

This finds 8 such sets, each monomial represents one independent set. For example, the first set found is $t = \{a, b, d, f, h\}$.

```
i49 : R1 = QQ[c, e, g, a, b, d, f, h, MonomialOrder=>ProductOrder{3,5}];
i50 : L = substitute(J, R1)
o50 = ideal (- c*b + a*d, e*h - g*f, - e*b + a*f)
o50 : Ideal of R1
```

```

i51 : gens gb L
o51 = | eh-gf eb-af cb-ad gbf-afh caf-ead |
      1      5
o51 : Matrix R1 <--- R1

```

By examining the lead terms and coefficients, we see that $in_u(J) = (c, e, g)$, and that the α_i corresponding to c is a , to e are b and h , and to g is bf . Therefore abf is a flattener. Let $F = abf$.

```

i52 : use R
o52 = R
o52 : PolynomialRing
i53 : J1 = saturate(J,a*b*f)
o53 = ideal (f*g - e*h, d*g - c*h, b*g - a*h, d*e - c*f, b*e - a*f, b* ...
o53 : Ideal of R
i54 : J1 == J : (a*b*f)
o54 = true

```

So $J = J_1 \cap J_2$, where

```

i55 : J2 = trim(J + ideal(a*b*f))
o55 = ideal (f*g - e*h, b*e - a*f, b*c - a*d, a*b*f)
o55 : Ideal of R
i56 : J == intersect(J1,J2)
o56 = true

```

As it turns out, J_1 is a prime ideal. How can we see this? Since the initial ideal $in_u(J_1) = (c, e, g)$, this means that the projection map is birational, and therefore the ideal J_1 is prime and even more, is rational.

```

i57 : Q1 = J1;
o57 : Ideal of R

```

Now let's decompose J_2 .

```

i58 : independentSets J2
o58 = {c*d*e*f*h, a*b*d*g*h, a*c*d*g*h, c*d*e*g*h}
o58 : List

```

We'll use the first one.

```

i59 : R1 = QQ[a,b,g, c,d,e,f,h,MonomialOrder=>ProductOrder{3,5}];
i60 : L = substitute(J2,R1)
o60 = ideal (g*f - e*h, - a*f + b*e, - a*d + b*c, a*b*f)
o60 : Ideal of R1
i61 : gens gb L
o61 = | gf-eh af-be ad-bc bde-bcf bge-afh b2e b2cf abeh a2eh2 |
      1      9
o61 : Matrix R1 <--- R1

```

In this case $in_u(J_2) = (a, b, g)$ (So the saturation will again be rational and prime, as before). One choice for a flattener is $f(ce - df)$.

```

i62 : use R
o62 = R
o62 : PolynomialRing
i63 : Q2 = saturate(J2,f*(c*e-d*f))
o63 = ideal (f*g - e*h, b*g - a*h, b*e - a*f, b*c - a*d, a*b, a2)
o63 : Ideal of R
i64 : J3 = trim(J2 + ideal(f*(c*e-d*f)))
o64 = ideal (f*g - e*h, b*e - a*f, b*c - a*d, c*e*f - d*f2, a*b*f)
o64 : Ideal of R
i65 : J == intersect(Q1,Q2,J3)
o65 = true

```

One more time. Let's decompose J_3 .

```

i66 : independentSets J3
o66 = {a*b*d*g*h, a*c*d*g*h}
o66 : List
i67 : R1 = QQ[c,e,f, a,b,d,g,h, MonomialOrder=>ProductOrder{3,5}]
o67 = R1
o67 : PolynomialRing
i68 : L = substitute(J3,R1)
o68 = ideal (- e*h + f*g, e*b - f*a, c*b - a*d, c*e*f - f*d2, f*a*b)
o68 : Ideal of R1
i69 : transpose gens gb L
o69 = {-2} | eh-fg |
      {-2} | eb-fa |
      {-2} | cb-ad |
      {-3} | fbg-fah |
      {-3} | cfa-ead |
      {-3} | fab |
      {-3} | cef-f2d |
      {-4} | fa2h |
      {-4} | f2a2 |
      {-4} | fa2d |
      {-4} | cf2g-f2dh |
      {-4} | efad-f2bd |
      {-4} | e2ad-f2ad |
      {-5} | f2adg-f2bdh |
      {-5} | f2b2d |
      {-5} | ea2d2 |
      16 1
o69 : Matrix R1 <--- R1

```

This time, $in_u(J_3) = (c, e, f)$, and so once again the saturation will be a prime rational ideal. A flattener that works this time is ab .

```
i70 : use R
o70 = R
o70 : PolynomialRing
i71 : Q3 = saturate(J3,a*b)
o71 = ideal (f, e, b*c - a*d)
o71 : Ideal of R
i72 : Q3 == J3 : (a*b)
o72 = false
i73 : Q3 == J3 : (a*b)^2
o73 = true
```

This time,

```
i74 : J4 = trim(J3 + ideal(a^2*b^2))
o74 = ideal (f*g - e*h, b*e - a*f, b*c - a*d, c*e*f - d*f^2, a*b*f, a^2 b^2)
o74 : Ideal of R
```

But notice that

```
i75 : J == intersect(Q1,Q2,Q3)
o75 = true
```

Therefore, we may avoid the primary decomposition of J_4 , since it will only consist of redundant terms. You should check, but the primary decomposition of J_4 has 7 terms, and J_4 is not a radical ideal.

Exercise 4.11. *Apply this technique to other Bayesian network examples, such as the example from the contraction lemma.*

How can we tell if an ideal J is prime or primary? One way that always works (if the computation finishes, that is) is the zero dimensional decomposition combined with the flattener.

As in the last example, if the projection is birational to an obviously prime ideal (e.g. the zero ideal!), then the ideal is prime. This method works very well for the Markov ideals.

5 Lecture #4: Putting it all together

In the lectures, and final lecture notes, we will include algorithms based on these ideas. However, better than presenting an algorithm right away, write down your own algorithm, perhaps based on the flattener approach and example above. Try to improve the approach taken in the above example. Maybe you will find better optimizations than we will present!

The questions you should consider are: (1) how to consider and compute with as few redundant components as possible? (2) The precise element which one splits the ideal by has a dramatic effect on the complexity of the computation. Is there any way to control this?

We will also apply these (improved) algorithms to see some interesting behavior of the primary decompositions of Markov ideals.

References

- [CLO97] David Cox, John Little, and Donal O’Shea. *Ideals, varieties, and algorithms*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, second edition, 1997. An introduction to computational algebraic geometry and commutative algebra.
- [CNR00] A. Capani, G. Niesi, and L. Robbiano. *CoCoA, a system for doing Computations in Commutative Algebra*. Available via anonymous ftp from cocoa.dima.unige.it, 4.0 edition, 2000.
- [DGP99] Wolfram Decker, Gert-Martin Greuel, and Gerhard Pfister. Primary decomposition: algorithms and comparisons. In *Algorithmic algebra and number theory (Heidelberg, 1997)*, pages 187–220. Springer, Berlin, 1999.
- [ES96] David Eisenbud and Bernd Sturmfels. Binomial ideals. *Duke Math. J.*, 84(1):1–45, 1996.
- [GHKM01] Dan Geiger, David Heckerman, Henry King, and Christopher Meek. Stratified exponential families: graphical models and model selection. *Ann. Statist.*, 29(2):505–529, 2001.
- [GPS01] G.-M. Greuel, G. Pfister, and H. Schönemann. SINGULAR 2.0. A Computer Algebra System for Polynomial Computations, Centre for Computer Algebra, University of Kaiserslautern, 2001. <http://www.singular.uni-kl.de>.
- [GPW01] E. Riccomagno G. Pistone and H. Wynn. *Algebraic Statistics: Computational Commutative Algebra in Statistics*. Chapman and Hall, Boca Raton, 2001.
- [GS] Daniel R. Grayson and Michael E. Stillman. *Macaulay 2, a software system for research in algebraic geometry*. Available at <http://www.math.uiuc.edu/Macaulay2/>.
- [GSS] Luis David Garcia, Michael Stillman, and Bernd Sturmfels. Algebraic Geometry of Bayesian Networks. arXiv:math.AG/0301255.
- [GTZ88] Patrizia Gianni, Barry Trager, and Gail Zacharias. Gröbner bases and primary decomposition of polynomial ideals. *J. Symbolic Comput.*, 6(2-3):149–167, 1988. Computational aspects of commutative algebra.
- [Lau96] S.L. Lauritzen. *Graphical Models*. Oxford University Press, 1996.
- [Sch03] H. Schenck. Computational algebraic geometry. Available at <http://www.math.tamu.edu/~schenck/99r.html>, 2003.
- [SS00] Raffaella Settini and Jim Q. Smith. Geometry, moments and conditional independence trees with hidden variables. *Ann. Statist.*, 28(4):1179–1205, 2000.
- [Stu02] B. Sturmfels. *Solving Systems of Polynomial Equations*. CBMS Lecture Series. American Mathematical Society, 2002.
- [SY96] Takeshi Shimoyama and Kazuhiro Yokoyama. Localization and primary decomposition of polynomial ideals. *J. Symbolic Comput.*, 22(3):247–277, 1996.