Graphical models in Macaulay2

Luis David García-Puente[*], Sonja Petrović[†], Seth Sullivant[‡]

September 3, 2012

Abstract

The `Macaulay2` package `GraphicalModels` contains algorithms for the algebraic study of graphical models associated to undirected, directed and mixed graphs, and associated collections of conditional independence statements. Among the algorithms implemented are procedures for computing the vanishing ideal of graphical models, for generating conditional independence ideals of families of independence statements associated to graphs, and for checking for identifiable parameters in Gaussian mixed graph models. These procedures can be used to study fundamental problems about graphical models.

1 Graphical models

A graphical model is a statistical model associated to a graph, where the nodes of the graph represent random variables and the edges of the graph encode relationships between the random variables. Graphical models are an important class of statistical models used in many applications (see standard textbooks [5, 8]) because of their ability to model complex interactions between many random variables, by specifying interactions using only local information about connectivity between the vertices in a graph.

There are two natural ways to specify a graphical model, through either conditional independence statements specified by the graph or via a parametric representation (often called a "factorization"). Every distribution which factors according to the graph satisfies the conditional independence statements implied by the graph. This leads to the question: Which distributions satisfy the conditional independence statements implied by the graph, but do not factor?

Once we specify the types of random variables under consideration (e.g. discrete random variables or Gaussian random variables) it is possible to address the questions in the preceding paragraph using (computational) algebraic geometry. Indeed, in these cases, the set of all probability distributions satisfying a family of conditional independence constraints is a semialgebraic set. For discrete random variables, that semialgebraic set is a subset of the probability simplex, and can be represented by a certain homogeneous ideal generated by quadrics. For Gaussian random variables, this set of distributions corresponds to a semialgebraic subset of the cone of positive definite matrices. Similarly, the parametrized family of probability distributions also is a semialgebraic set (of the probability simplex for discrete random variables, and of the cone of positive definite matrices for Gaussian random variables). This algebraic perspective has been studied by many authors [3, 4, 6], and the book [1] provides many details.

The `Macaulay2` package `GraphicalModels` allows the user to compute the ideals of conditional independence statements for any collection of statements for discrete or Gaussian random variables. It can also compute the vanishing ideal of a graphical model in these cases. A number of auxiliary functions are useful for doing further analyses of graphical models.

To be explicit, consider the directed acyclic graph G with five vertices $\{a, b, c, d, e\}$ and edge set $\{a \rightarrow d, b \rightarrow d, c \rightarrow d, c \rightarrow e, d \rightarrow e\}$. The following commands compute the associated conditional independence ideal for the set of global Markov statements, $CI_{\mathrm{global}(G)}$, and the vanishing ideal I_G of the Gaussian graphical model on G.

---

[*]Department of Mathematics and Statistics, Sam Houston State University, Huntsville, TX 77341, lgarcia@shsu.edu

[†]Department of Statistics, Pennsylvania State University, University Park, PA 16802, sonja@psu.edu

[‡]Department of Mathematics, North Carolina State University, Raleigh, NC 27695, smsulli2@ncsu.edu

```
Macaulay2, version 1.4

i1 : loadPackage "GraphicalModels";
i2 : G = digraph{{a,d},{b,d},{c,{d,e}},{d,e}};
i3 : R = gaussianRing G
o3 : PolynomialRing
i4 : vars R
o4 = | s_(a,a) s_(a,b) s_(a,c) s_(a,d) s_(a,e) s_(b,b) s_(b,c) s_(b,d) s_(b,e) s_(c,c)
     s_(c,d) s_(c,e) s_(d,d) s_(d,e) s_(e,e) |
i4 : I = conditionalIndependenceIdeal(R,globalMarkov(G));
i5 : J = gaussianVanishingIdeal(R);
i6 : flatten degrees J
o6 = {1, 1, 1, 2, 3, 3}
```

GraphicalModels uses the package **Graphs** and a number of fundamental constructs and relationships associated with graphs. First we create a polynomial ring for carrying out the computations, which contains the entries of the covariance matrix Σ of a jointly normal random vector as its indeterminates. Information about the underlying graph is stored with the polynomial ring. Hence many functions in this package take only ring as input, but require that it be created with **gaussianRing**, or **markovRing** in the discrete case.

For directed acyclic graphs it is known that $V(CI_{\text{global}(G)}) \cap PD_m = V(I_G) \cap PD_m$, in particular, the set of positive definite matrices satisfying the conditional independence constraints equals the set of covariance matrices in the image of the parametrization. Unfortunately, this does not imply that $CI_{\text{global}(G)} = I_G$. In the case of Gaussian random variables, a larger ideal, the trek ideal T_G, generated by all subdeterminants of the covariance matrix that vanish on the model, and satisfying $CI_{\text{global}(G)} \subseteq T_G \subseteq I_G$ is sometimes equal to I_G (see [7]), which happens for our graph of interest G in this case.

```
i7 : isSubset(I,J)
o7 = true
i8 : I == J
o8 = false
i9 : J == trekIdeal(R,G)
o9 = true
```

Similar computations can also be performed for graphical models with discrete random variables, and with other graph families. The mathematical explanation of these graphical models and their associated ideals appear in the remaining sections.

2   Computing conditional independence ideals

Conditional independence constraints on discrete or Gaussian random variables translate to rank conditions on certain matrices associated to the probability densities. We briefly explain these constructions here and how to generate these constraints in **Macaulay2** using **GraphicalModels**. See [1, Ch. 3] for more detail.

Let $X = (X_1, \ldots, X_n)$ be a discrete random vector where each random variable X_i has state space $[d_i] = \{1, 2, \ldots, d_i\}$. Let $d = (d_1, \ldots, d_n)$. A probability distribution for X is a tensor in $\mathbb{R}^{d_1} \otimes \cdots \otimes \mathbb{R}^{d_n}$, all of whose coordinates are nonnegative and sum to one. The set of all such distributions is the probability simplex Δ_d. Let $p_{i_1 \cdots i_n} = \mathrm{P}(X_1 = i_1, \ldots, X_n = i_n)$ denote the probability of a primitive event. The polynomial ring in these quantities is created using the command **markovRing**.

```
i10 : d = (2,3,2); R = markovRing d
o10 : PolynomialRing
i11 : gens R
o11 = {p        , p        , p        , p        , p        , p        , p        , p        , p        ,
         1,1,1    1,1,2    1,2,1    1,2,2    1,3,1    1,3,2    2,1,1    2,1,2    2,2,1
       p        , p        , p        }
         2,2,2    2,3,1    2,3,2
```

For $A \subseteq [n]$, let $X_A = (x_a)_A$ be the subvector indexed by A. Let A, B, C be disjoint subsets of $[n]$. The conditional independence statement $X_A \perp\!\!\!\perp X_B | X_C$ holds if and only if the conditional distribution satisfies

$$\mathrm{P}(X_A = i_A, X_B = i_B | X_C = i_C) = \mathrm{P}(X_A = i_A | X_C = i_C) \cdot \mathrm{P}(X_B = i_B | X_C = i_C)$$

for all i_A, i_B, i_C. This translates into vanishing 2×2 minors of certain matrices in the probabilities $p_{i_1 \cdots i_n}$. Those matrices are computed with the function `markovMatrices`, and the ideal generated by the 2×2 minors is computed with `conditionalIndependenceIdeal`. In the following example, the two conditional independence statements are $X_1 \perp\!\!\!\perp X_2 | X_3$ and $X_1 \perp\!\!\!\perp X_3$ ($:= X_1 \perp\!\!\!\perp X_3 | X_\emptyset$). The ideal of vanishing minors has 7 quadratic generators.

```
i12 : s = {{{1},{2},{3}}, {{1},{3},{}}}
i13 : markovMatrices(R,s)
o13 = {| p_(1,1,1) p_(1,2,1) p_(1,3,1) |, | p_(1,1,2) p_(1,2,2) p_(1,3,2) |,
        | p_(2,1,1) p_(2,2,1) p_(2,3,1) |  | p_(2,1,2) p_(2,2,2) p_(2,3,2) |
     | p_(1,1,1)+p_(1,2,1)+p_(1,3,1) p_(1,1,2)+p_(1,2,2)+p_(1,3,2) |}
     | p_(2,1,1)+p_(2,2,1)+p_(2,3,1) p_(2,1,2)+p_(2,2,2)+p_(2,3,2) |
i14 : I = conditionalIndependenceIdeal(R,s); flatten degrees  I
o14 : Ideal of R
o15 = {2, 2, 2, 2, 2, 2, 2}
```

There are many different sets of conditional independence statements associated to a graph G, whose nodes correspond to random variables. For example, the *local Markov statements* of an undirected graph G is the set of conditional independence statements of the form $X_i \perp\!\!\!\perp X_{V \setminus \{i \cup N(i)\}} | X_{N(i)}$, where $N(i)$ is the set of neighbors of i in the graph G. The functions `pairMarkov`, `localMarkov`, and `globalMarkov` compute the pairwise, local, and global Markov statements, respectively, for both directed and undirected graphs.

```
i16 : G = graph{{1,2},{2,3},{3,4},{4,5},{1,5}};
i17 : localMarkov G
o17 = {{{1}, {3, 4}, {5, 2}}, {{1, 2}, {4}, {5, 3}}, {{1, 5}, {3}, {4, 2}},
      {{2, 3}, {5}, {4, 1}}, {{2}, {4, 5}, {1, 3}}}
```

For example, the first conditional independence statement produced is $X_1 \perp\!\!\!\perp (X_3, X_4) | (X_2, X_5)$. In the context of conditional independence, the graphical model consists of all distributions satisfying one of these collections of independence statements associated to the graph G.

A Gaussian random vector, $X = (X_1, \ldots, X_n) \sim \mathcal{N}(\mu, \Sigma)$, is an n-dimensional random vector with state space $\mathbb{R}^n$ and density function

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

where $\mu \in \mathbb{R}^n$ and $\Sigma = (\sigma_{s,t}) \in PD_n$, the cone of $n \times n$ symmetric positive definite matrices. The Gaussian random vector X satisfies the conditional independence statement $X_A \perp\!\!\!\perp X_B | X_C$ if and only if the submatrix $\Sigma_{A \cup C, B \cup C} := (\sigma_{s,t})_{s \in A \cup C, t \in B \cup C}$ has rank $\leq \#C$. Hence the set of all Gaussian random vectors satisfying a given collection of conditional independence statements naturally corresponds to a subset of PD_n, and are studied algebraically by investigating the corresponding determinantal ideals in the polynomial ring in the $\sigma_{s,t}$ indeterminates. The corresponding ring is generated using the command `gaussianRing`. Computations involving conditional independence ideals with Gaussian random variables were exemplified in Section 1.

3   Computing the vanishing ideal of a model

The fact that graphical models can be described in two possible ways (either by a recursive factorization of probability distributions or by conditional independence statements) corresponds to the algebraic principle that varieties can be presented either parametrically or implicitly. The *vanishing ideal* of a model is the set of homogeneous polynomial relations in the joint probability distributions (for discrete random variables) or

in the variance-covariance parameters (for Gaussian random variables). `GraphicalModels` has the capability of computing the vanishing ideals of graphical models on directed graphs (for discrete random variables) and also of graphical models on directed, undirected, or mixed graphs (for Gaussian random variables). The vanishing ideal of an undirected graphical model for discrete random variables is a toric ideal and should be computed using the `Macaulay2` package `FourTiTwo`.

The method `discreteVanishingIdeal` implements this capability for graphical models on discrete random variables. For a directed acyclic graph G on discrete random variables, the graphical model consists of all distributions satisfying the recursive factorization property

$$p(X = i) = \prod_v \mathrm{P}(X_v = i_v | X_{\mathrm{pa}(v)} = i_{\mathrm{pa}(v)}),$$

where the product runs over all vertices v of G and $\mathrm{pa}(v)$ is the set of parents of v. Our implementation of this method does not compute the kernel of the corresponding ring map. Instead, the vanishing ideal is computed recursively using the factorization

$$\mathrm{P}(X = i) = \mathrm{P}(X_1 = i_1, \ldots, X_{n-1} = i_{n-1}) \cdot \mathrm{P}(X_n = i_n | X_{\mathrm{pa}(n)} = i_{\mathrm{pa}(n)}),$$

where $1, \ldots, n$ is a topological ordering of the vertices of the directed acyclic graph G.

The following example computes the vanishing ideal of the graphical model $1 \rightarrow 2 \rightarrow 3 \rightarrow 4$ on four binary random variables. The vanishing ideal is minimally generated by 20 quadratic binomials.

```
i18 : G = digraph {{1,{2}}, {2,{3}},{3,{4}},{4,{}}};
i19 : R = markovRing (2,2,2,2);
i20 : I = discreteVanishingIdeal (R,G);
o20 : Ideal of R
i21 : betti mingens I
            0  1
o21 = total: 1 20
        0: 1  .
        1: . 20
```

According to [3], the vanishing ideal of a graphical model on discrete random variables is the *distinguished component* of the conditional independence ideal described by the Markov statements of the model. For the directed path in our previous example, the conditional independence ideal of the local Markov statements is a radical ideal with 3 associated primes. However, since G is a directed tree, the conditional independence ideal of the global Markov statements is a prime ideal and it equals the vanishing ideal of G.

```
i22 : J = conditionalIndependenceIdeal (R, localMarkov G);
o22 : Ideal of R
i23 : I == J
o23 = false
i24 : K = conditionalIndependenceIdeal (R, globalMarkov G);
o24 : Ideal of R
i25 : I == K
o25 = true
```

The method `gaussianVanishingIdeal` computes the vanishing ideal of a Gaussian graphical model on a graph, digraph, or mixed graph. If G is a mixed graph, the vanishing ideal I_G corresponds to kernel of the polynomial parametrization of the variance-covariance parameters σ_{ij} in terms of the *direct causal effect* parameters λ_{ij} associated to the directed edges of G, and the *noise* parameters ψ_{ij} associated to the bidirected edges of G. This parametrization is given by

$$\Sigma = (I - \Lambda)^{-T} \Psi (I - \Lambda)^{-1},$$

where Σ is the variance-covariance matrix, Λ is the strictly upper triangular matrix with $\Lambda_{ij} = \lambda_{ij}$ if $i \rightarrow j$ is a directed edge in G and 0 otherwise, and Ψ is a positive definite matrix of parameters ψ_{ij} with zeros in each entry Ψ_{ij} if there is no bidirected edge in G between i and j, and $i \neq j$.

The following example computes the vanishing ideal of the Gaussian graphical model on the mixed graph with directed edges $\{1 \to 2, 1 \to 3, 2 \to 3, 3 \to 4\}$ and bidirected edges $\{1 \leftrightarrow 2, 2 \leftrightarrow 4\}$. This ideal is a principal ideal generated by one quartic polynomial with 8 terms. This ideal is not *determinantal*, i.e., it is not generated by the determinantal equations defining the trek ideal, which in this case is the zero ideal.

```
i26 : G = mixedGraph(digraph {{1,{2,3}},{2,{3}},{3,{4}}},bigraph {{1,2},{2,4}});
i27 : R = gaussianRing G;
i28 : I = gaussianVanishingIdeal R;
o28 : Ideal of R
i29 : J = trekIdeal (R,G)
o29 = 0
```

An important problem in these models consists in finding which parameters are *identifiable* or *generically identifiable*, see [2]. The method `identifyParameters` can be used to solve the identifiability problem for Gaussian graphical models on mixed graphs (also known as *structural equation models*). The following example shows that the parameter ψ_{24} is generically identifiable by the formula $\psi_{24} = (\sigma_{13}\sigma_{24} - \sigma_{14}\sigma_{23})/\sigma_{13}$.

```
i30 : H = identifyParameters R;
i31 : H#(p_(2,4))_0
o31 = p    s    + s    s    - s    s
       2,4 1,3    1,4 2,3    1,3 2,4
```

In this model there are three non-generically identified parameters. `identifyParameters` produces a hash table whose entries are indexed by the parameters and contain ideals that can be used to find explicit rational functions for every parameter that is generically identifiable.

Acknowledgments

References

[1] M. Drton, B. Sturmfels and S. Sullivant, *Lectures on Algebraic Statistics*, Oberwolfach Seminars, 39. Birkhäuser Verlag, Basel, 2009.

[2] L. D. Garcia-Puente, S. Spielvogel and S. Sullivant: Identifying causal effects with computer algebra, *Proceedings of the 26th Conference of Uncertainty in Artificial Intelligence*.

[3] L. D. Garcia, M. Stillman and B. Sturmfels: Algebraic geometry of Bayesian networks, *J. Symbolic Comput.* **39** (2005) 331–355.

[4] D. Geiger, C. Meek and B. Sturmfels: On the toric algebra of graphical models, *Ann. Statist.* **34** (2006), no. 3, 1463–1492.

[5] S. Lauritzen, *Graphical models*, Oxford Statistical Science Series, 17. Oxford University Press, New York, 1996.

[6] S. Sullivant: Algebraic geometry of Gaussian Bayesian networks, *Adv. in Appl. Math.* **40** (2008) 482–513.

[7] S. Sullivant, K. Talaska, and J. Draisma: Trek separation for Gaussian graphical models, *Ann. Statist.* **38** (2010) 1665–1685.

[8] Whitaker. Graphical Models in Applied Multivariate Statistics. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Ltd., Chichester, 1990.