

GAUSSIAN CONDITIONAL INDEPENDENCE RELATIONS HAVE NO FINITE COMPLETE CHARACTERIZATION

SETH SULLIVANT

ABSTRACT. We show that there can be no finite list of conditional independence relations which can be used to deduce all conditional independence implications among Gaussian random variables. To do this, we construct, for each $n > 3$ a family of n conditional independence statements on n random variables which together imply that $X_1 \perp\!\!\!\perp X_2$, and such that no subset have this same implication. The proof relies on binomial primary decomposition.

1. INTRODUCTION

A fundamental question about the nature of conditional independence is whether or not there exists a finite set of conditional independence relations, from which all general conditional independence implications can be deduced. This problem was resolved in the negative by Studený [6], who showed that there can be no such finite characterization. To prove this, he exhibited infinite families of conditional independence implications, using the multi-information function and properties of the submodular cone.

Studený's work leaves open the question of whether or not it is possible to find such a finite axiom characterization over restricted classes of random variables. For instance, if we assume that all random variables are binary, there are more independence relations that hold. Perhaps it is possible that among the many new independence relations there can be a finite axiom system.

In this note, we show that there is no finite characterization in the regular Gaussian case, by exhibiting an infinite family of conditional independence implications which cannot be deduced from any other conditional independence implications. The main result of this note is:

Theorem 1.1. *Let $X \sim \mathcal{N}(\mu, \Sigma)$ with Σ positive definite. Suppose that X satisfies the conditional independence constraints:*

$$X_1 \perp\!\!\!\perp X_2 | X_3, X_2 \perp\!\!\!\perp X_3 | X_4, \dots, X_{n-1} \perp\!\!\!\perp X_n | X_1, X_n \perp\!\!\!\perp X_1 | X_2$$

with $n \geq 4$. Then X also satisfies the marginal independence constraints:

$$X_1 \perp\!\!\!\perp X_2, X_2 \perp\!\!\!\perp X_3, \dots, X_{n-1} \perp\!\!\!\perp X_n, X_1 \perp\!\!\!\perp X_n.$$

However, if X satisfies only a subset of the first CI constraints, it may not satisfy any of the second set.

The proof relies on three simple ideas. First, we use the fact that conditional independence in the Gaussian case is an algebraic restriction on the covariance matrix Σ . So our collection of conditional independence constraints gives an ideal I_n in $\mathbb{R}[\Sigma]$. Next, we compute the primary decomposition of I_n . Since, in our case, I_n will be a binomial ideal, we can exploit the results of [4] to determine the minimal primes of I_n . Then we use Hadamard products (a trick we learned

from [5]) to show that one of these components does not intersect the cone of positive definite matrices and that the other component yields the desired conditional independence implications.

2. THE PROOF

Recall that a regular multivariate Gaussian $X \sim \mathcal{N}(\mu, \Sigma)$ is completely specified by its mean vector $\mu \in \mathbb{R}^n$ and its symmetric positive definite covariance matrix $\Sigma \in PD_n$. First of all, we show that conditional independence corresponds to an algebraic constraint on the covariance matrix Σ .

Proposition 2.1. *Let A, B, C be disjoint subsets of $[n]$. Then the joint distribution for X satisfies the conditional independence constraint $A \perp\!\!\!\perp B | C$ if and only if the submatrix $\Sigma_{A \cup C, B \cup C}$ has rank less than or equal to $\#C$.*

Proof. If $X \sim \mathcal{N}(\mu, \Sigma)$ follows a multivariate Gaussian distribution, then the conditional distribution of $X_{A \cup B} | X_C = x_c$ is also Gaussian with distribution

$$\mathcal{N}\left(\mu_{A \cup B} + \Sigma_{A \cup B, C} \Sigma_{C, C}^{-1} (x_c - \mu_C), \Sigma_{A \cup B, A \cup B} - \Sigma_{A \cup B, C} \Sigma_{C, C}^{-1} \Sigma_{C, A \cup B}\right),$$

(see, for example, [1, §B.6]). The conditional independence statement $A \perp\!\!\!\perp B | C$ holds if and only if $(\Sigma_{A \cup B, A \cup B} - \Sigma_{A \cup B, C} \Sigma_{C, C}^{-1} \Sigma_{C, A \cup B})_{A, B} = 0$. The A, B submatrix of $\Sigma_{A \cup B, A \cup B} - \Sigma_{A \cup B, C} \Sigma_{C, C}^{-1} \Sigma_{C, A \cup B}$ is easily seen to be $\Sigma_{A, B} - \Sigma_{A, C} \Sigma_{C, C}^{-1} \Sigma_{C, B}$ which is the Schur complement of the matrix

$$\Sigma_{A \cup C, B \cup C} = \begin{pmatrix} \Sigma_{A, B} & \Sigma_{A, C} \\ \Sigma_{C, B} & \Sigma_{C, C} \end{pmatrix}.$$

Since $\Sigma_{C, C}$ is always invertible (it is positive definite), the Schur complement is zero if and only if the matrix $\Sigma_{A \cup C, B \cup C}$ has rank less than or equal to $\#C$. \square

Thus, if $\mathcal{M} = \{A_1 \perp\!\!\!\perp B_1 | C_1, \dots, A_m \perp\!\!\!\perp B_m | C_m\}$ is a conditional independence model, we naturally get the *conditional independence ideal*

$$CI_{\mathcal{M}} = \langle \#C + 1 \text{ minors of } \Sigma_{A \cup C, B \cup C} \mid A \perp\!\!\!\perp B | C \in \mathcal{M} \rangle \subseteq \mathbb{R}[\Sigma].$$

To determine which conditional independence statements a collection of independence statements imply, we investigate the primary decomposition of the ideal $CI_{\mathcal{M}}$, and figure out which components of $CI_{\mathcal{M}}$ intersect the positive definite cone PD_n .

Define \mathcal{M}_n to be the cyclic system of conditional independence constraints

$$\mathcal{M}_n = \{1 \perp\!\!\!\perp 2 | 3, 2 \perp\!\!\!\perp 3 | 4, \dots, n-1 \perp\!\!\!\perp n | 1, n \perp\!\!\!\perp 1 | 2\}$$

and let $I_n = CI_{\mathcal{M}_n}$ be the ideal defining this cyclic model. Our goal is to compute the primary decomposition of this ideal.

2.1. Minimal primes of lattice basis ideals. Let B be a finite collection of integral vectors that form a basis for a saturated lattice $\mathcal{L} \subseteq \mathbb{Z}^n$. The *lattice basis ideal* associated to B is the binomial ideal

$$I_B = \langle \mathbf{x}^{\mathbf{u}^+} - \mathbf{x}^{\mathbf{u}^-} \mid \mathbf{u} \in B \rangle \subset \mathbb{K}[\mathbf{x}],$$

where $\mathbf{u} = \mathbf{u}^+ - \mathbf{u}^-$ is the cancellation free representation of \mathbf{u} as the difference of two nonnegative integral vectors, and $\mathbf{x}^{\mathbf{a}} = x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}$ is the monomial vector notation. A thorough study of binomial ideals was done in [2] and a simple combinatorial characterization of the minimal

primes of lattice basis ideals was discovered in [4]. Their combinatorial characterization will be of considerable use, because of the following:

Proposition 2.2. *The conditional independence ideal*

$$I_n = \langle \sigma_{33}\sigma_{12} - \sigma_{13}\sigma_{23}, \sigma_{44}\sigma_{23} - \sigma_{24}\sigma_{34}, \dots, \sigma_{11}\sigma_{n-1,n} - \sigma_{1n-1}\sigma_{n1}, \sigma_{22}\sigma_{1n} - \sigma_{2n}\sigma_{12} \rangle$$

is a lattice basis ideal.

Proof. Let \mathbf{u}_i denote the exponent vector attached the conditional independence statement $X_{i-2} \perp\!\!\!\perp X_{i-1} | X_i$. That is, $\mathbf{u}_i = e_{ii} + e_{i-2,i-1} - e_{i-1,i} - e_{i-2,i}$, where the indices are interpreted modulo n . These vectors are linearly independent because the vector e_{ii} only appears in vector \mathbf{u}_i . Furthermore, this shows that the vectors span a saturated lattice because the submatrix of the matrix $M_n = [\mathbf{u}_1; \dots; \mathbf{u}_n]$ with columns corresponding to the diagonal entries in Σ , is the identity matrix. \square

Since I_n is a lattice basis ideal, we can use the results from [4] to study the primary decomposition of I_n . Let $B = \{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ be a basis for a saturated lattice. We form the matrix $M \in \mathbb{Z}^{k \times n}$ whose rows are the vectors \mathbf{u}_i . From [2], it is known that the minimal primes of a lattice basis ideal are completely determined by the indeterminates that appear in them. In particular, let $S \subseteq [m]$ be a collection indexing indeterminates. Then the prime ideal associated to this subset is the ideal:

$$I_S = \langle x_i \mid i \in S \rangle + \left\langle \mathbf{x}^{\mathbf{u}^+} - \mathbf{x}^{\mathbf{u}^-} \mid \mathbf{u} \in B, \text{supp}(\mathbf{u}) \cap S = \emptyset \right\rangle : \prod_{i \notin S} x_i^\infty.$$

Thus, one must determine the sets S that give minimal primes. When $S = \emptyset$, $I_S = I_B : \prod x_i^\infty$ is the toric ideal associated to the lattice \mathcal{L} . This toric ideal is always a minimal prime of I_B . The main result of [4] is that the other minimal primes of the lattice basis ideal I_B can be read off from the sign patterns in the matrix M .

A matrix M is called *mixed* if every row contains a positive entry and a negative entry. Let S be a subset of $[m]$ indexing a possible minimal prime. After permuting rows and columns of M , and relabeling S , we can assume that $S = \{1, 2, \dots, t\}$ and M has the form

$$M = \left(\begin{array}{c|c} N & B \\ \hline 0 & D \end{array} \right),$$

where N has no all zero rows.

Definition 2.3. A matrix N is called *irreducible* if

- (1) N is a mixed $s \times t$ matrix with $t \leq s$ and
- (2) one cannot bring N into the form

$$N = \left(\begin{array}{c|c} N' & B' \\ \hline 0 & D' \end{array} \right)$$

by permuting rows and columns where N' is a mixed $s' \times t'$ matrix with $t' \leq s'$ and D' is a $(s - s') \times (t - t')$ matrix with $t - t' > s - s'$.

Theorem 2.4. [4] *A set of $S \subseteq [m]$ yields a minimal prime I_S of the lattice basis ideal I_B if and only if the associated matrix N_S is an irreducible matrix.*

Now we will apply Theorem 2.4 to determine the minimal primes of the ideals I_n .

Lemma 2.5. *For $n \geq 4$, the minimal primes of I_n are the toric ideal $I_{A_n} = I_n : \prod \sigma_{ij}^\infty$ and the monomial ideal $\langle \sigma_{12}, \sigma_{23}, \dots, \sigma_{1n} \rangle$.*

Proof. For $n = 4$, this can be checked directly in Macaulay2 [3] or proven directly by a slight variation on the argument below. So assume henceforth that $n \geq 5$.

The toric ideal I_{A_n} is necessarily a minimal prime of I_n . We need to show that the only set of variables S that induces an irreducible decomposition of the lattice basis matrix is the one corresponding to the set of variables $\{\sigma_{12}, \sigma_{23}, \dots, \sigma_{1n}\}$. First, we want to describe the lattice basis matrix M_n , which we need to find irreducible decompositions of. Since the binomials $\sigma_{ii}\sigma_{i-1,i-2} - \sigma_{i-1,i}\sigma_{i-2,i}$ only have index pairs (i, j) with $|i - j| \leq 2 \pmod n$, we only need to use $3n$ columns. We separate these columns into three groups, depending on whether $|i - j| = 0, 1, 2$. For $n = 5$, M_5 is a 5×15 matrix:

$$M_5 = \left(\begin{array}{ccccc|ccccc|ccccc} +1 & 0 & 0 & 0 & 0 & +1 & -1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 & 0 & 0 & +1 & -1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & +1 & 0 & 0 & 0 & 0 & +1 & -1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & +1 & 0 & 0 & 0 & 0 & +1 & -1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & +1 & -1 & 0 & 0 & 0 & +1 & 0 & 0 & 0 & 0 & -1 \end{array} \right)$$

In general, after reordering the columns in each block, we will have an identity matrix, a circulant matrix on the vector $(+1, -1, 0, \dots, 0)$, and minus the identity matrix for the three blocks of size n . Note that the central circulant block, corresponding to the indeterminates $\sigma_{12}, \sigma_{23}, \dots, \sigma_{1n}$, is an irreducible submatrix as it is $n \times n$, mixed, and no nonempty submatrix of it is mixed. This is the desired minimal prime we were seeking. It remains to show that no other subsets of the variables induce an irreducible decomposition of the matrix M .

To see why there are no other irreducible decompositions, we can just count the number of nonzero entries in each column. Note that there are ≤ 2 nonzero entries in each column. So suppose that N were an irreducible matrix arising from choosing some subsets S of the variables. Each row of N must have at least one $+1$ and one -1 entry, since N is mixed. So if N is an $s \times t$ matrix, the number of nonzero entries is $\geq 2s$. On the other hand, since each column has at most 2 nonzero entries, we know that there are $\leq 2t$ nonzero entries in N . Furthermore, since N is irreducible we have $s \geq t$, so such an N can only exist when $s = t$ and every column used has two nonzero entries. This implies that only the indeterminates corresponding to the circulant submatrix of M can be among the variables associated to the minimal prime. However, we have already seen that this set of variables yields an irreducible submatrix, thus there can be no other minimal primes besides the two we have already found. \square

2.2. Hadamard Products. The next step in the proof depends on analyzing the toric ideal I_{A_n} , and showing that the variety $V(I_{A_n})$ does not intersect the positive definite cone. Ultimately, the basic idea comes from a fact about Hadamard products of positive definite matrices.

Lemma 2.6. *Let Σ and T be $n \times n$ positive definite matrices. Then the Hadamard product $\Sigma * T$ defined by*

$$(\Sigma * T)_{ij} = \sigma_{ij}\tau_{ij}$$

is also positive definite.

Proof. Let $X = (X_1, \dots, X_n)$ and $Y = (Y_1, \dots, Y_n)$ be random vectors both with mean 0 and with covariance matrices Σ and T respectively, and suppose that $X \perp\!\!\!\perp Y$. Also, suppose that X and Y have n -dimensional support. Let $X * Y$ be the new random vector $X * Y = (X_1Y_1, X_2Y_2, \dots, X_nY_n)$. Since X and Y both have mean zero, we compute the covariance Γ of $X * Y$ as

$$\gamma_{ij} = \mathbb{E}[X_iY_iX_jY_j] = \mathbb{E}[X_iX_j]\mathbb{E}[Y_iY_j] = \sigma_{ij}\tau_{ij}.$$

Thus, $\Gamma = \Sigma * T$. Since X and Y have full dimensional support, so does $X * Y$ and thus Γ , the covariance matrix of a random variable with full dimensional support, must be a positive definite matrix. \square

Lemma 2.7. *The variety $V(I_{A_n})$ does not intersect the positive definite cone.*

Proof. Given a permutation $\pi \in S_n$ and an $n \times n$ matrix Σ , let $\pi(\Sigma)$ be the matrix obtained by simultaneously permuting rows and columns by π . Let π denote the cycle $(12 \cdots n)$ and let Σ be a generic matrix. Let the matrix Σ_n be defined by the repeated Hadamard product:

$$\Sigma_n = \Sigma * \pi(\Sigma) * \pi^2(\Sigma) * \cdots * \pi^{n-2}(\Sigma) * \pi^{n-1}(\Sigma).$$

If Σ is positive definite, then so is Σ_n . In particular,

$$\det((\Sigma_n)_{13,13}) > 0.$$

This inequality is equivalent to

$$\prod_{i=1}^n \sigma_{ii}^2 > \prod_{i=1}^n \sigma_{i-2,i}^2.$$

On the other hand, the lattice $\ker_{\mathbb{Z}} A_n$ is spanned by the vectors $\mathbf{u}_i = e_{ii} + e_{i-2,i-1} - e_{i-1,i} - e_{i-2,i}$. In particular, the vector

$$\mathbf{v} = \sum_{i=1}^n \mathbf{u}_i = \sum_{i=1}^n e_{ii} - \sum_{i=1}^n e_{i-2,i}$$

is in the lattice $\ker_{\mathbb{Z}} A_n$. This implies that the binomial

$$\prod_{i=1}^n \sigma_{ii} - \prod_{i=1}^n \sigma_{i-2,i}$$

belongs to the toric ideal I_{A_n} . This implies that any $\Sigma \in V(I_{A_n})$ satisfies

$$\prod_{i=1}^n \sigma_{ii}^2 = \prod_{i=1}^n \sigma_{i-2,i}^2.$$

Thus, $V(I_{A_n}) \cap PD_n = \emptyset$. \square

As the last tool in the proof, we need to show that if X satisfies some, but not all, of the conditional independence statements $X_i \perp\!\!\!\perp X_{i+1} | X_{i+2}$, then none of the independence statements $X_i \perp\!\!\!\perp X_{i+1}$ need to be satisfied. This is explained in the following lemma.

Lemma 2.8. *Suppose that $X \sim \mathcal{N}(\mu, \Sigma)$ and X satisfies all the cyclic conditional independence statements except for $X_{n-1} \perp\!\!\!\perp X_n | X_1$. Then X need not satisfy any conditional independence statements of the form $X_i \perp\!\!\!\perp X_{i+1}$.*

Proof. It suffices to exhibit a positive definite covariance matrix that satisfies all the cyclic conditional independence constraints except $\sigma_{11}\sigma_{n-1,n} - \sigma_{1,n}\sigma_{1,n-1}$ that does not have any zero entries. To this end, let $0 < a, e < \frac{1}{n}$. Define $\sigma_{ii} = 1$ for all i , $\sigma_{i-1,i} = a^{n-i+1}$ for $i \in [n]$, $\sigma_{i-2,i} = a$ for all $i \in [n]$, and $\sigma_{ij} = e$ for all other values of i, j . Note that we take the indices cyclically modulo n , so $\sigma_{1,n} = a^n$.

First of all, note that the Σ defined in this way is positive definite because it is diagonally dominant. Furthermore, for all i except $i = 1$, we have

$$\sigma_{ii}\sigma_{i-1,i-2} - \sigma_{i-1,i}\sigma_{i,i-2} = 1 \cdot a^{n-(i-1)+1} - a^{n-i+1} \cdot a = 0$$

and thus, X satisfies all the cyclic conditional independence statements except $X_{n-1} \perp\!\!\!\perp X_n | X_1$. Finally, by construction, all elements of Σ are nonzero. \square

Proof of Theorem 1.1: Lemmas 2.5 and 2.7 show that $V(I_n) \cap PD_n = V(\langle \sigma_{12}, \sigma_{23}, \dots, \sigma_{1n} \rangle) \cap PD_n$ which implies the desired implication of conditional independence statements (reinterpreting those polynomial constraints back into CI constraints). Lemma 2.8 shows that if even one of the n initial CI constraints is omitted, there exists positive definite covariance matrices that satisfy the indicated constraints and do not imply any of the marginal CI constraints $X_i \perp\!\!\!\perp X_{i+1}$. This completes the proof of the theorem. \square

REFERENCES

- [1] P. Bickel and K. Doksum. *Mathematical Statistics. Vol 1*. Prentice-Hall, London, 2001.
- [2] D. Eisenbud and B. Sturmfels. Binomial ideals. *Duke Math. J.* **84** (1996), no. 1, 1–45.
- [3] D. Grayson and M. Stillman. Macaulay 2, a software system for research in algebraic geometry. Available at <http://www.math.uiuc.edu/Macaulay2/>
- [4] S. Hoşten and J. Shapiro. Primary decomposition of lattice basis ideals, *Journal of Symbolic Computation* **29** (2000) 625-639.
- [5] F. Matúš. Conditional independences in Gaussian vectors and rings of polynomials. *Proceedings of WCII 2002* (eds. G. Kern-Isberner, W. Rödder, and F. Kulmann) LNAI 3301, Springer-Verlag, Berlin, Heidelberg, 152-161.
- [6] M. Studený. Conditional independence relations have no finite complete characterization. In *Information Theory, Statistical Decision Functions and Random Processes*. 1992, 377-396.

DEPARTMENT OF MATHEMATICS AND SOCIETY OF FELLOWS, HARVARD UNIVERSITY, CAMBRIDGE, MA 02138