Proceedings of WCII 2002 (eds. G. Kern-Isberner, W. Rödder, and F. Kulmann)
LNAI 3301, Springer-Verlag, Berlin, Heidelberg, 152–161.

Conditional independences in Gaussian vectors and rings of polynomials

František Matúš

Institute of Information Theory and Automation, Pod vodárenskou věží 4,
182 08 Prague, Czech Republic
matus@utia.cas.cz

Abstract. Inference among the conditional independences in nondegenerate Gaussian vectors is studied by algebraic techniques. A general method to prove implications involving the conditional independences is presented. The method relies on computations of a Groebner basis. Examples of the implications are discussed.

1 Introduction

Conditional independence (CI) constraints have become a crucial ingredient in building highly structured probabilistic models of artificial intelligence. They provide a language communicating qualitative features and induce decompositions of probability distributions amenable to practical computations. Though the major part of research has been concerned with CI constraints related to graphs, cf. [14, 2], an effort is devoted also to the general problem of discrete structures describing simultaneous occurence of independences within a random vector, for references see [19, 9]. Challenging problems involving CI constraints have emerged also in statistics [7], information theory [21, 20], secret-sharing schemes of cryptography [1], representations of matroid theory [11] and instances of the probabilistic method in combinatorics [17].

For a random vector $\xi = (\xi_i)_{i \in N}$ indexed by a finite set N, the stochastic conditional independences $\xi_I \perp\!\!\!\perp \xi_J | \xi_K$ among its subvectors $\xi_I = (\xi_i)_{i \in I}$, ξ_J, ξ_K, where $I, J, K \subseteq N$ are disjoint, can be encoded by means of the discrete structure

$$[\xi] = \{(ij|K) \in \mathcal{R};\ \xi_i \perp\!\!\!\perp \xi_j | \xi_K\}\,.$$

Here, ij denotes a two-element set $\{i, j\}$ and $\mathcal{R}$ is the family of all ordered pairs $(ij|K)$ with $ij, K \subseteq N$ disjoint. The conditional independence $\xi_I \perp\!\!\!\perp \xi_J | \xi_K$ holds if and only if $(ij|L) \in [\xi]$ for all $i \in I$, $j \in J$ and L satisfying $K \subseteq L \subseteq (I \cup J \cup K) \setminus ij$, see [8, Lemma 3].

Manipulations with the CI constraints can be facilitated when understanding which subset $\mathcal{K}$ of $\mathcal{R}$ can be represented as $\mathcal{K} = [\xi]$ by means of some random vector ξ. This difficult problem remains open even in the discrete setting when

N has more than four elements and ξ takes an arbitrary finite number of values, see [10, 9]. The analogous problem of representability by Gaussian vectors has not been yet investigated in depth. This contribution aims at first steps in this direction, restricting attention to nondegenerate vectors. The degenerate case seems to differ only by additional technicalities. Systematic study of the structures $[\xi]$ is motivated also by a hope to reveal new classes of statistical models for discrete and especially Gaussian vectors, possibly admitting parsimonious parametrizations and effective computations.

Let $A = (a_{i,j})_{i,j=1}^{n}$ be the covariance matrix of a Gaussian vector $\xi = (\xi_1, \ldots, \xi_n)$, $n \geqslant 1$. When the vector is nondegenerate the matrix A is positive definite, in symbols $A > 0$. It is well-known that the variable ξ_1 is conditionally independent of ξ_2 given $(\xi_3, \ldots, \xi_n)$ if and only if the $(1, 2)$-element of the inverse A^{-1} of A vanishes. Denoting submatrices of A by $A_{I,J} = (a_{i,j})_{i \in I, j \in J}$ and determinants by $|A_{I,J}|$, $I, J \subseteq N = \{1, \ldots, n\}$, this element equals $(-1)^{i+j} |A_{N \setminus \{1\}, N \setminus \{2\}}| / |A|$, see for example [15, p. 11]. Therefore, more generally, a conditional independence $\xi_i \perp\!\!\!\perp \xi_j | \xi_K$ for $(ij|K) \in \mathcal{R}$ holds if and only if $|A_{jK,iK}|$, equal to $|A_{iK,jK}|$, vanishes; here iK abbreviates $\{i\} \cup K$ and similarly with jK. Since the independences are determined by the covariance matrix A of ξ one can work directly with positive definite matrices and use the notation

$$\langle\!\langle A \rangle\!\rangle = \{(ij|K) \in \mathcal{R}\,;\; |A_{iK,jK}| = 0\}$$

in place of $[\xi]$. The main problem of this contribution is now to recognize whether a given subset $\mathcal{K}$ of $\mathcal{R}$ can be written as $\mathcal{K} = \langle\!\langle A \rangle\!\rangle$ for some $A > 0$. An equivalent problem is to find for which families $\mathcal{L}, \mathcal{M} \subseteq \mathcal{R}$ the implication

$$\mathcal{L} \subseteq \langle\!\langle A \rangle\!\rangle \quad \Rightarrow \quad \mathcal{M} \cap \langle\!\langle A \rangle\!\rangle \neq \emptyset \tag{1}$$

holds for all $A > 0$. In fact, $\mathcal{K} = \langle\!\langle A \rangle\!\rangle$ for some $A > 0$ if and only if (1) with $\mathcal{L} = \mathcal{K}$ and $\mathcal{M} = \mathcal{R} \setminus \mathcal{K}$ is violated by some $A > 0$.

By Sylvester's criterion [15, p. 96], the set of positive definite $n \times n$ matrices is a semialgebraic subset of a Euclidean space, that is, a set defined by a finite number of polynomial inequalities. To decide whether for given $\mathcal{L}, \mathcal{M}$ the implication (1) quantified by 'for all $A > 0$' is true or false amounts to decide whether the semialgebraic set

$$\big\{A > 0\,;\; |A_{iK,jK}| = 0 \text{ for } (ij|K) \in \mathcal{L}, \text{ and } |A_{iK,jK}| \neq 0 \text{ for } (ij|K) \in \mathcal{M}\big\}$$

is empty or not, respectively. Algorithms for manipulations with semialgebraic sets are available, see [12], and can recognize whether a given semialgebraic set is empty. Small instances of some statistical problems were attacked by such algorithms in [6], reporting successful treatment of a trivial example of the implication (1), see (5) below. Nevertheless, the general algorithms seem to provide no insight into the structure of this special problem and are of considerable complexity.

Since the construction $A \mapsto \langle\!\langle A \rangle\!\rangle$ makes sense even for a fairly general matrix, for example for any symmetric matrix, the quantification $A > 0$ in the main problem can be replaced by the assumption that A belongs to another set of symmetric matrices. Modified problems obtained in this way approximate the main problem and are of theoretical interest *per se* because they enrich the insight into abstract CI structures. A convenient replacement of $A > 0$ may require A to belong to an algebraic set (a set defined by polynomial equalities) or a constructible set of matrices (finite union of differences of algebraic sets). When matrices are admitted to have the complex numbers as elements numerous methods of the classical algebraic geometry, supported by computational procedures, notably for Groebner bases, are at disposal.

For example, when $A > 0$ the principal minors $|A_{I,I}|$ are positive for $\emptyset \neq I \subseteq N$, and a straightforward relaxation of the quantification $A > 0$ is to require A to be in the set $\mathbb{A}$ of complex symmetric matrices having all proper principal minors nonzero. The discrete structures $\langle\!\langle A \rangle\!\rangle$ for $A \in \mathbb{A}$ keep reasonable properties and remain within the class of semigraphoids, which has been a framework for CI relations, see [14, 7, 9]. When $n = 3$ this relaxation works very well and provides the same class of structures $\langle\!\langle A \rangle\!\rangle$ as if $A > 0$, see Example 1. However, when $n \geqslant 4$ this is not the case, see Example 3. In general, an 'exact' relaxation, of purely theoretical interest, to a constructible set of matrices exists. In fact, the disjoint union of the sets of matrices

$$\big\{A \in \mathbb{A};\ |A_{iK,jK}| = 0 \text{ for } (ij|K) \in \mathcal{K}, \text{ and } |A_{iK,jK}| \neq 0 \text{ for } (ij|K) \in \mathcal{R} \setminus \mathcal{K}\big\}$$

over all $\mathcal{K} \subseteq \mathcal{R}$ admitting the representation $\mathcal{K} = \langle\!\langle A \rangle\!\rangle$ for some $A > 0$ is a constructible set; obviously, (1) is valid for all $A > 0$ if and only if it is valid for all A in the union.

This contribution is organized as follows. In Section 2, examples of the implication (1) are presented for complex matrices. When specialized to positive definite matrices, they summarize basic properties of CI in nondegenerate Gaussian vectors. In Section 3, the implication (1) is analysed when A belongs to the set of matrices that violate each of a given set of polynomial equalities (such sets, for example $\mathbb{A}$ itself, are constructible when working with a finite number of equalities). A necessary and sufficient condition for the validity of such an implication is presented in Proposition 1. The condition can be practically verified by computing a single Groebner basis. Section 4 discusses possible uses of this result: all basic implications of Section 2 can be obtained as consequences of Proposition 1, and two additional implications for CI in Gaussian vectors are derived by a more sophisticated relaxation.

2 Basic properties of conditional independence

In this section, basic properties of CI in nondegenerate Gaussian vectors are derived as consequences of a single determinantal identity. It is convenient to abbreviate $A_{I,I}$ by A_I and take $|A_\emptyset| = 1$.

Lemma 1. *For any complex matrix A, and $L \subseteq N$, $k \in N \setminus L$, and $i, j \in N \setminus kL$*

$$(-1)^{[i,j;k]}\, |A_{ikL,jkL}||A_L| = |A_{iL,jL}||A_{kL}| - |A_{iL,kL}||A_{kL,jL}| \qquad (2)$$

where $[i,j;k]=1$ *if* k *is between* i *and* j*, and* $[i,j;k]=0$ *otherwise.*

Proof. When $L=\emptyset$ this is an obvious formula for determinant expansion. When $L\neq\emptyset$ one can suppose that A_L is invertible, the opposite case following then by continuity. First, let all elements of L be smaller than the minimum of the set $ijk=\{i,j,k\}$. The Schur complement of A_L in $A_{ikL,jkL}$ is the matrix $B_{ik,jk}=A_{ik,jk}-A_{ik,L}(A_L)^{-1}A_{L,jk}$. Then $|A_{ikL,jkL}|$ factors into $|A_L||B_{ik,jk}|$ by [15, Theorem 3.1.1a]. Similarly, the four determinants of the right-hand side of (2) can be factored as $|A_{iL,jL}|=|A_L||B_{i,j}|$, etc. Then (2) follows from the expansion formula for $|B_{ik,jk}|$.

If L and i,j,k are in a general position then a unique permutation π of $ijk\cup L$ exists such that $\pi(l_1)<\pi(l_2)$ for $l_1<l_2$ with l_1,l_2 in L or in ijk, and such that the elements of $\pi(L)=\{\pi(l);\ l\in L\}$ are smaller than the minimum of $\pi(i)\pi(j)\pi(k)$. Let C be the matrix obtained by applying π to the rows and columns of A simultaneously. By the previous argument

$$\begin{aligned}&(-1)^{[\pi(i),\pi(j);\pi(k)]}\,|C_{\pi(ikL),\pi(jkL)}||C_{\pi(L)}|\\&=|C_{\pi(iL),\pi(jL)}||C_{\pi(kL)}|-|C_{\pi(iL),\pi(kL)}||C_{\pi(kL),\pi(jL)}|\,.\end{aligned}$$

Then $[\pi(i),\pi(j);\pi(k)]=[i,j;k]$, $|C_{\pi(L)}|=|A_L|$ and $|C_{\pi(kL)}|=|A_{kL}|$. Denoting by n_i the number of all $l\in L$ greater than i, one can see that $|C_{\pi(ikL),\pi(jkL)}|$ is the $(-1)^{n_i+n_j}$ multiple of $|A_{ikL,jkL}|$. Similar observations concern the expressions $|C_{\pi(iL),\pi(jL)}|$ and $|C_{\pi(iL),\pi(kL)}||C_{\pi(kL),\pi(jL)}|$. Hence, (2) obtains in full generality. □

Remark 1. The identity (2) can be viewed as a technical refinement of its special case

$$|A_N||A_{N\setminus ij}|=|A_{N\setminus i}||A_{N\setminus j}|-|A_{N\setminus i,N\setminus j}||A_{N\setminus j,N\setminus i}| \tag{3}$$

where $ij=\{1,2\}$. Thanks to a personal communication of David Mond, University of Warwick, who kindly drove our attention to [13, 5, 4], the identity (3) can be put into the framework of the identity (1) of [4, p. 156]. In fact, the latter rewrites in our notation as

$$\sum_{j\in M\setminus J}(-1)^j\,|D_{N,jJ}||D_{N,M\setminus jJ}|=0 \tag{4}$$

where $M=\{1,2,\ldots,2n\}$, J is a subset of M of the cardinality $n-1$ and D is any $n\times 2n$ matrix; (4) is a special case of the straightening laws of [5]. Then, for $J=\{2,\ldots,n\}$ and

$$D=\left[\begin{array}{c|c|c} A & \begin{matrix}1&0\\0&1\\\vdots&\vdots\\0&0\end{matrix} & A_{N,\{3,\ldots,n\}}\end{array}\right]$$

a simple computation derives (3) from (4).

Corollary 1. *For any symmetric complex matrix A*

$$\{(ij|kL),(ij|L)\} \subseteq \langle\!\langle A \rangle\!\rangle \Rightarrow (ik|L) \in \langle\!\langle A \rangle\!\rangle \text{ or } (jk|L) \in \langle\!\langle A \rangle\!\rangle, \tag{5}$$
$$\{(ij|kL),(ik|L)\} \subseteq \langle\!\langle A \rangle\!\rangle \Rightarrow (ij|L) \in \langle\!\langle A \rangle\!\rangle, \qquad \text{if } |A_{kL}| \neq 0, \tag{6}$$
$$\{(ij|kL),(ik|L)\} \subseteq \langle\!\langle A \rangle\!\rangle \Rightarrow (ik|jL) \in \langle\!\langle A \rangle\!\rangle, \qquad \text{if } |A_L||A_{kL}| \neq 0, \tag{7}$$
$$\{(ij|kL),(ik|jL)\} \subseteq \langle\!\langle A \rangle\!\rangle \Rightarrow (ij|L) \in \langle\!\langle A \rangle\!\rangle, \qquad \text{if } |A_{jkL}||A_L| \neq 0, \tag{8}$$
$$\{(ij|L),(ik|L)\} \subseteq \langle\!\langle A \rangle\!\rangle \Rightarrow (ij|kL) \in \langle\!\langle A \rangle\!\rangle, \qquad \text{if } |A_L| \neq 0. \tag{9}$$

Proof. Rewriting (2) for a symmetric matrix A gives

$$(-1)^{[i,j;k]}\,|A_{ikL,jkL}||A_L| - |A_{iL,jL}||A_{kL}| = -|A_{iL,kL}||A_{jL,kL}|\,.$$

Hence, $|A_{ikL,jkL}| = 0$ and $|A_{iL,jL}| = 0$ imply $|A_{iL,kL}||A_{jL,kL}| = 0$ which proves (5).

Along the same guidelines, the identity (2) allows to conclude $(ij|L) \in \langle\!\langle A \rangle\!\rangle$ in (6).

To prove (7), one switches j and k in Lemma (1)

$$(-1)^{[i,k;j]}\,|A_{ijL,jkL}||A_L| = |A_{iL,kL}||A_{jL}| - |A_{iL,jL}||A_{jL,kL}|\,. \tag{10}$$

and combines this equality with (6).

When $|A_{ikL,jkL}| = 0$ and $|A_{ijL,jkL}| = 0$ the left-hand sides in (2) and (10) vanish. Solving the two homogeneous equations for $|A_{iL,jL}|$ and $|A_{iL,kL}|$, their determinant is equal to $|A_{jL}||A_{kL}| - |A_{jL,kL}||A_{kL,jL}|$. Using Lemma 1 with $i = j$, this determinant equals $|A_{jkL}||A_L|$ and is nonzero by assumption. Hence, $|A_{iL,jL}| = 0$ and (8) follows.

The implication (9) follows directly from (2).

□

A subset $\mathcal{K}$ of $\mathcal{R}$ is called a *semigraphoid* if it fulfills

$$\{(ij|kL),(ik|L)\} \subseteq \mathcal{K} \iff \{(ik|jL),(ij|L)\} \subseteq \mathcal{K}\,, \tag{11}$$

cf. (6) and (7), and a *pseudographoid* if it fulfills the implication $\Rightarrow$ in

$$\{(ij|kL),(ik|jL)\} \subseteq \mathcal{K} \iff \{(ij|L),(ik|L)\} \subseteq \mathcal{K}\,, \tag{12}$$

cf. (8). A semigraphoid which is a pseudographoid is called a *graphoid*. (For the pseudographoid axiom compare [14, (3.6e) and (3.6b)] and [9, p. 107]. For (5)see [14, (3.34f)].)

By Corollary 1, to ensure that $\langle\!\langle A \rangle\!\rangle$ satisfies (11) and (12) it suffices to assume that $|A_I|$ is nonzero for every nonempty proper subset of N or, equivalently to the above notation, $A \in \mathbb{A}$. Since all positive definite matrices belong to $\mathbb{A}$ the conditional independences in nondegenerate Gaussian vectors enjoy (11), (12) and (5) with $\mathcal{K}$ or $\langle\!\langle A \rangle\!\rangle$ replaced by $[\xi]$.

Example 1. For $N = \{1, 2, 3\}$, $n = 3$, Corollary 1 implies that if $A \in \mathbb{A}$ then $\langle\!\langle A \rangle\!\rangle$ is equal to one of the families

$$\emptyset, \{(ij|\emptyset)\}, \{(ij|k)\}, \{(ij|\emptyset), (ik|\emptyset), (ij|k), (ik|j)\} \text{ and } \mathcal{R}\,.$$

The five families are representable by the following positive definite matrices

$$\begin{bmatrix} 1 & \epsilon & \epsilon^2 \\ \epsilon & 1 & \epsilon^4 \\ \epsilon^2 & \epsilon^4 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & \epsilon \\ 0 & 1 & \epsilon \\ \epsilon & \epsilon & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & \epsilon^3 & \epsilon^2 \\ \epsilon^3 & 1 & \epsilon \\ \epsilon^2 & \epsilon & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & \epsilon \\ 0 & \epsilon & 1 \end{bmatrix}$$

($\epsilon \neq 0$ sufficiently small) and the identity matrix, respectively. Therefore, if $n = 3$ the class of structures $\{\langle\!\langle A \rangle\!\rangle;\ A > 0\}$ coincides with the class $\{\langle\!\langle A \rangle\!\rangle;\ A \in \mathbb{A}\}$. This is no longer true when $n > 3$, see Example 3.

3 Implications and polynomial rings

This section focuses on a necessary and sufficient condition for validity of (1) when A is in a certain constructible set of matrices. The condition involves some elementary material on rings of polynomials which is available in an accessible form in [3, 12].

When studying $\langle\!\langle A \rangle\!\rangle$ for a symmetric matrix A with nonzero diagonal elements there is no loss of generality in assuming that all diagonal elements are equal to 1. In fact, multiplying each i-th row and i-th column of A by a number b_i satisfying $b_i^2 = a_{i,i}^{-1}$ one obtains a matrix A' rendering $\langle\!\langle A \rangle\!\rangle = \langle\!\langle A' \rangle\!\rangle$; in addition, A' can be made positive definite or belongs to $\mathbb{A}$ if A was positive definite or belonged to $\mathbb{A}$, respectively. From now on all matrices are supposed to be symmetric and to have the ones on their diagonals.

Let x_{ij} be an indeterminate indexed by a two-element set ij, $1 \leqslant i < j \leqslant n$, and $\mathfrak{R}$ be the ring of polynomials in these indeterminates with coefficients in the field of complex numbers. By little abuse of language we will substitute complex matrices A to polynomials, having in mind that the indeterminates x_{ij} are replaced by the off-diagonal elements $a_{i,j} = a_{j,i}$. For $(ij|K) \in \mathcal{R}$ the determinant $|A_{iK,jK}|$ expands as the sum of products of elements of $A_{iK,jK}$ and this expansion gives rise to a polynomial of $\mathfrak{R}$. Denoting the polynomial by $f_{(ij|K)}$, one has $f_{(ij|K)}(A) = |A_{iK,jK}|$. The notation $f_{(ij|K)}$ is extended to $f_{\mathcal{M}} = \prod_{(ij|K)\in\mathcal{M}} f_{(ij|K)}$ for $\mathcal{M} \subseteq \mathcal{R}$. Similarly, the expansion of $|A_I|$ gives rise to a polynomial denoted by f_I. For example, $f_{(12|3)} = x_{12} - x_{13}x_{23}$, $f_{(13|2)} = x_{12}x_{23} - x_{13}$ and $f_{12} = 1 - x_{12}^2$.

Let p_r, $1 \leqslant r \leqslant t$, be some polynomials in $\mathfrak{R}$ and $\mathbb{B}$ be the set of complex matrices A (symmetric, with ones on diagonals) such that $p_r(A) \neq 0$ for all $1 \leqslant r \leqslant t$. If y_r are additional indeterminates, the polynomials $1 - y_r p_r$ belong to the ring $\mathfrak{T} = \mathfrak{R}[y_1, \ldots, y_t]$ of polynomials in both x_{ij}'s and y_r's. Obviously, a matrix A together with complex numbers $b_1, \ldots, b_t$ solve the polynomial equations $1 - y_r p_r = 0$ if and only if $A \in \mathbb{B}$ and $b_r = p_r(A)^{-1}$. Thus $\mathbb{B}$ is a coordinate projection of an algebraic set and is therefore constructible by [3, pp. 258–9]. In the following assertion another indeterminate z and the ring $\mathfrak{T}[z]$ of polynomials with one more indeterminate come into play.

Proposition 1. *Let $\mathcal{L}, \mathcal{M}$ be subfamilies of $\mathcal{R}$, p_r be a polynomial in $\mathfrak{R}$, $1 \leqslant r \leqslant t$, and $\mathbb{B}$ be the corresponding set of complex matrices. The implication*

$$\mathcal{L} \subseteq \langle\!\langle A \rangle\!\rangle \quad \Rightarrow \quad \mathcal{M} \cap \langle\!\langle A \rangle\!\rangle \neq \emptyset \,, \qquad A \in \mathbb{B} \,, \tag{13}$$

is true if and only if the ring $\mathfrak{T}[z]$ is generated by the polynomials $f_{(ij|K)}$ for $(ij|K) \in \mathcal{L}$, $1-y_r p_r$ for $1 \leqslant r \leqslant t$, and $1-zf_{\mathcal{M}}$. This is equivalent to the incidence of $f_{\mathcal{M}}$ to the radical $\sqrt{I_{\mathcal{L}}}$ of the ideal $I_{\mathcal{L}}$ generated by $f_{(ij|K)}$ for $(ij|K) \in \mathcal{L}$, and $1 - y_r p_r$ for $1 \leqslant r \leqslant t$.

Proof. If $\mathfrak{T}[z]$ is generated as above then the constant polynomial 1 can be combined from the listed polynomials. Thus, polynomials $h_{(ij|K)}$, h_r, and h in $\mathfrak{T}[z]$ exist such that

$$\sum_{(ij|K) \in \mathcal{L}} h_{(ij|K)} f_{(ij|K)} + \sum_{r=1}^{t} h_r [1 - y_r p_r] + h[1 - zf_{\mathcal{M}}] = 1 \,.$$

Substituting $A \in \mathbb{B}$ for x_{ij}'s and the nonzero numbers $p_r(A)^{-1}$ for y_r's the sum over r equals zero. Then

$$\sum_{(ij|K) \in \mathcal{L}} h^*_{(ij|K)} |A_{iK,jK}| + h^* \Big[1 - z \prod_{(ij|K) \in \mathcal{M}} |A_{iK,jK}|\Big] = 1$$

where $h^*_{(ij|K)}$ and h^* are polynomials involving only the indeterminate z. If $\langle\!\langle A \rangle\!\rangle \subseteq \mathcal{L}$ the sum over $\mathcal{L}$ vanishes. One concludes that the product must be equal to zero, that is, $\mathcal{M} \cap \langle\!\langle A \rangle\!\rangle \neq \emptyset$ and (13) holds.

Suppose (13) is true. When the polynomials $f_{(ij|K)}$ for $(ij|K)$ in $\mathcal{L}$, and $1 - y_r p_r$ for $1 \leqslant r \leqslant t$ vanish after substitution of complex numbers a_{ij} for x_{ij} and b_r for y_r then the symmetric matrix A with the off-diagonal elements $a_{i,j} = a_{ij}$ (and ones on the diagonal) belongs to $\mathbb{B}$. Now, $0 = f_{(ij|K)}(A) = |A_{iK,jK}|$ for $(ij|K)$ in $\mathcal{L}$ and thus $\langle\!\langle A \rangle\!\rangle \subseteq \mathcal{L}$. Due to (13), one has $(ij|K) \in \langle\!\langle A \rangle\!\rangle$ for some $(ij|K) \in \mathcal{M}$. Hence, $\prod_{(ij|K) \in \mathcal{M}} |A_{iK,jK}|$ equals zero and the polynomial $f_{\mathcal{M}}$ vanishes by the same substitution, $f_{\mathcal{M}}(A) = 0$. Hilbert's Nullstellensatz [3, p. 170] implies $f_{\mathcal{M}} \in \sqrt{I_{\mathcal{L}}}$.

If $f_{\mathcal{M}}$ belongs to the radical then $f^m_{\mathcal{M}}$ belongs to $I_{\mathcal{L}}$ for an integer $m \geqslant 1$. Rewriting the constant polynomial 1 as $z^m f^m_{\mathcal{M}} + [1 - z^m f^m_{\mathcal{M}}]$ and further as $z^m f^m_{\mathcal{M}} + h^*[1 - zf_{\mathcal{M}}]$ where $h^* \in \mathfrak{T}[z]$ it combines from a polynomial from $I_{\mathcal{L}}$ and $1 - zf_{\mathcal{M}}$. Thus, the ring $\mathfrak{T}[z]$ is generated by $I_{\mathcal{L}}$ and $1 - zf_{\mathcal{M}}$, cf. [3, p. 176]. □

4 Discussion and examples

When examining an instance of (1) for $A > 0$ Proposition 1 can be useful. Indeed, if each polynomial p_r is chosen so that $p_r(A) \neq 0$ for every $A > 0$ with ones on its diagonal, Proposition 1 provides a sufficient condition for validity of the instance.

For example, since (2) implies

$$-(-1)^{[i,j;k]}\, z f_L f_{(ij|kL)} + z f_{kL} f_{(ij|L)} + \big[1 - z f_{(ik|L)} f_{(jk|L)}\big] = 1$$

Proposition 1 with $\mathcal{L} = \{(ij|kL), (ij|L)\}$, $\mathcal{M} = \{(ik|L), (jk|L)\}$ and $t = 0$ proves validity of (5) for the symmetric matrices with ones on diagonals, and then obviously for the symmetric matrices with nonzero diagonal elements. It is not difficult to see (and we do not give details) that all implications of Corollary 1 quantified by $A \in \mathbb{A}$ can be obtained in this way from Proposition 1 with $t = 1, 2$ and f_I's in the role of p_r's. Hence, all basic properties of CI in nondegenerate Gaussian vectors can be viewed as consequences of Proposition 1.

A more sophisticated relaxation of the quantification $A > 0$ works as follows. Given a matrix A and a permutation π of the set N, the matrix $A^\pi = (a_{\pi(i),\pi(j)})_{i,j=1}^n$ is obtained by permuting rows and columns of A simultaneously. Obviously, $A > 0$ if and only if $A^\pi > 0$. By [15, Theorem 36.2.1] the Hadamard product $A \circ A^\pi = (a_{i,j}\, a_{\pi(i),\pi(j)})_{i,j=1}^n$ of $A > 0$ and A^π is positive definite. Hence, the principal minors of Hadamard products $A^{\pi_1} \circ \ldots \circ A^{\pi_k}$ are positive for $A > 0$, $k \geqslant 1$ and arbitrary permutations $\pi_1, \ldots, \pi_k$ of N. Such minors are of interest when substituted for p_r's in Proposition (1).

Example 2. To prove the implication

$$\{(12|3), (13|4), (14|2)\} \subseteq \langle\!\langle A \rangle\!\rangle \quad \Rightarrow \quad (12|\emptyset) \in \langle\!\langle A \rangle\!\rangle \tag{14}$$

for $A > 0$ observe that $1 - a_{2,3}^2 a_{2,4}^2 a_{3,4}^2$ is a principal minor of the Hadamard product $A \circ A^{\pi_1} \circ A^{\pi_2}$ of three matrices and the polynomials

$$f_{(12|3)}\,,\ f_{(13|4)}\,,\ f_{(14|2)}\,,\ 1 - y_1(1 - x_{23}^2 x_{24}^2 x_{34}^2) \ \text{ and } \ 1 - z x_{12}$$

generate $\mathfrak{I}[z]$ because

$$1 = (1 - z x_{12}) + z x_{12}\big(1 - y_1(1 - x_{23}^2 x_{24}^2 x_{34}^2)\big) + y_1 z x_{12}\big[1 - x_{23}^2 x_{24}^2 x_{34}^2\big]$$

and

$$x_{12}\big[1 - x_{23} x_{24} x_{34}\big] = f_{(12|3)} + x_{23} f_{(13|4)} - x_{23} x_{34} f_{(14|2)}\,.$$

Hence, Proposition 1 with $\mathcal{L} = \{(12|3), (13|4), (14|2)\}$, $\mathcal{M} = \{(12|\emptyset)\}$, $t = 1$ and p_1 equal to $(1 - x_{23}^2 x_{24}^2 x_{34}^2)$ implies (14) for the symmetric matrices A with ones on diagonals and $1 - a_{2,3}^2 a_{2,4}^2 a_{3,4}^2 \neq 0$, in turn, for the positive definite matrices.

Analogously,

$$\{(12|3), (23|4), (34|1), (41|2)\} \subseteq \langle\!\langle A \rangle\!\rangle \quad \Rightarrow \quad (12|\emptyset) \in \langle\!\langle A \rangle\!\rangle \tag{15}$$

holds for $A > 0$, using the principal minor $1 - a_{1,3}^2 a_{2,4}^2$ of the Hadamard product $A \circ A^\pi$ of two matrices and the identity

$$x_{12}\big[1 - x_{13}^2 x_{24}^2\big] = f_{(12|3)} + x_{13} f_{(23|4)} + x_{13} x_{24} f_{(34|1)} - x_{13}^2 x_{24} f_{(14|2)}\,.$$

Remark 2. From the computational point of view, a ring of polynomials is generated by a finite list of polynomials if and only if any Groebner basis of the ideal generated by the list contains a constant (nonzero) polynomial. Therefore, a verification whether $\mathfrak{T}[z]$ is generated as in Proposition 1 has the same complexity as a computation of a single Groebner basis.*

Example 3. For ϵ small and nonzero the matrices A given by

$$\begin{bmatrix} 1 & \epsilon^2 & \epsilon^4 & \epsilon^3 \\ \epsilon^2 & 1 & \epsilon^{-2} & \epsilon \\ \epsilon^4 & \epsilon^{-2} & 1 & \epsilon \\ \epsilon^3 & \epsilon & \epsilon & 1 \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} 1 & \epsilon^2 & \epsilon & \epsilon \\ \epsilon^2 & 1 & \epsilon & \epsilon^{-1} \\ \epsilon & \epsilon & 1 & \epsilon^2 \\ \epsilon & \epsilon^{-1} & \epsilon^2 & 1 \end{bmatrix}$$

belong to $\mathbb{A}$ and render $\langle\!\langle A \rangle\!\rangle$ equal to

$$\{(12|3), (13|4), (14|2)\} \text{ and } \{(12|3), (23|4), (34|1), (41|2)\}\,,$$

respectively. By (14) and (15), these two subsets of $\mathcal{R}$ are not of the form $\langle\!\langle A \rangle\!\rangle$ for some $A > 0$. (Note that the above two matrices are not positive definite.) Therefore, to exclude only vanishing of proper minors of covariance matrices is not sufficient for the algebraic study of CI in Gaussian vectors.

Due to the following assertion, it is not difficult to lift the implications (14) and (15) of Example 2 to a general form.

Lemma 2. *Let $\mathcal{L}, \mathcal{M} \subseteq \mathcal{R}$ and $L \subseteq N$ such that $L \cap (ij \cup K) = \emptyset$ for every $(ij|K)$ in $\mathcal{L} \cup \mathcal{M}$. If (1) is true for all $A > 0$ or $A \in \mathbb{A}$ then it remains true also when $\mathcal{L}$ and $\mathcal{M}$ are replaced by $\mathcal{L}_L = \{(ij|K \cup L);\ (ij|K) \in \mathcal{L}\}$ and $\mathcal{M}_L$, respectively.*

A proof is simple and therefore omitted.

Corollary 2. *For any positive definite matrix A, $L \subseteq N$ and $i, j, k, l \in N \setminus L$ different*

$$\{(ij|kL), (ik|lL), (il|jL)\} \subseteq \langle\!\langle A \rangle\!\rangle \quad \Rightarrow \quad (ij|L) \in \langle\!\langle A \rangle\!\rangle\,, \tag{16}$$

$$\{(ij|kL), (jk|lL), (kl|iL), (li|jL)\} \subseteq \langle\!\langle A \rangle\!\rangle \quad \Rightarrow \quad (ij|L) \in \langle\!\langle A \rangle\!\rangle\,. \tag{17}$$

A prototype of the implication (16) with $(ij|L)$ replaced by $(ik|jL)$ goes back to [16], dealing with relational databases. Its validity for random vectors taking finite number of values was established in [18].

Acknowledgement

This research was supported by the grant A 1075104 of GA AV and the grant 201/01/1482 of GA ČR.

* For example, in the software package Mathematica, the following input `GroebnerBasis[{x12−x13x23, x13−x14x34, x12x24−x14, 1−y1[1−x23^2x24^2x34^2], 1−z x12}, {x12, x13, x14, x23, x24, x34, y1, z}]` is evaluated as the list $\{1\}$ consisting of the constant polynomial 1, cf. Example 2.

References

1. L. Csirmaz (1996) The dealer's random bits in perfect secret-sharing schemes. *Studia Scient. Math. Hungarica* **32** 429–437.
2. R.G. Cowell, A.P. Dawid, S.L. Lauritzen and D.J. Spiegelhalter (1999) *Probabilistic Networks and Expert Systems.* Springer: New York.
3. D. Cox, J. Little and D. O'Shea (1997) *Ideals, Varieties, and Algorithms.* Undergraduate Texts in Mathematics, Springer: New York.
4. C. DeConcini, D. Eisenbud and D. Procesi (1980) Young diagrams and determinantal varieties, *Inventiones math.* **56** 129-165.
5. P. Doubilet, G.C. Rota and J. Stein (1974) On the foundations of the combinatorial theory IX: combinatorial methods in invariant theory, *Studies in Applied Math.* **53** 185-216.
6. D. Geiger and Ch. Meek (1999) Quantifier Elimination for statistical problems. *Proceedings of Fifteenth Conference on Uncertainty in Artificial Intelligence* (Eds. K. Laskey and H. Prade), Morgan Kaufmann, 226-235.
7. S.L. Lauritzen (1996) *Graphical Models.* Oxford University Press: Oxford.
8. F. Matúš (1992) On equivalence of Markov properties over undirected graphs. *Journal of Applied Probability* **29** 745–749.
9. F. Matúš (1997) Conditional independence structures examined via minors. *Annals of Mathematics and Artificial Intelligence* **21** 99–128.
10. F. Matúš (1999) Conditional independences among four random variables III: final conclusion. *Combinatorics, Probability and Computing* **8** 269–276.
11. F. Matúš (1999) Matroid representations by partitions. *Discrete Mathematics* **203** 169–194.
12. B. Mishra (1993) *Algorithmic Algebra.* Texts and monographs in Computer Science, Vol. XIV, Springer: New York.
13. D. Mond and R. Pellikaan (1989) Fitting ideals and multiple points of analytic mappings. In: Algebraic and Analytic Geometry, Patzcuaro 1987, E. Ramirez de Arellano ed., Springer Lecture Notes in Mathematics 1414, pp. 107–161.
14. J. Pearl (1988) *Probabilistic Reasoning in Intelligent Systems.* Morgan Kaufman: San Meteo, California.
15. V.V. Prasolov (1994) *Problems and Theorems in Linear Algebra.* AMS Translations of Mathematical Monographs, Vol. 134, AMS: Providence, Rhode Island.
16. Y. Sagiv and S.F. Walecka (1988) Subset dependencies and a completeness result for a subclass of embedded multivalued dependencies. *Journal of ACM* **29** 103-117.
17. J. Spencer (1994) *Ten Lectures on the Probabilistic Method.* SIAM: Philadelphia, Pennsylvania.
18. M. Studený (1991) Conditional independence relations have no finite complete characterization. *Transactions of the 11-th Prague Conference on Information Theory, Statistical Decision Functions and Random Processes, Vol.B*, Academia, Prague, 377-396 (also Kluwer, Dordrecht).
19. M. Studený (2002) On stochastic conditional independence: the problems of characterization and description. *Annals of Mathematics and Artificial Intelligence* **35** 323–341.
20. R.W. Yeung (2002) *A First Course in Information Theory.* Kluwer Academic/Plenum Publishers: New York.
21. Z. Zhang and R.W. Yeung (1997) A non-Shannon-type conditional inequality of information quantities. *IEEE Trans. Information Theory* **43** 1982–1986.