Journal of Computational and Applied Mathematics 430 (2023) 115098

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Gaussian Markov Random Fields and totally positive matrices

Juan Baz [a], Pedro Alonso [b,*], Juan Manuel Peña [c], Raúl Pérez-Fernández [a]

[a] Departamento de Estadística, I.O. y D.M., Universidad de Oviedo, Spain
[b] Departamento de Matemáticas, Universidad de Oviedo, Spain
[c] Departamento de Matemática Aplicada, Universidad de Zaragoza, Spain

ARTICLE INFO

Article history:
Received 21 July 2022
Received in revised form 23 January 2023

MSC:
15A09
65F05

Keywords:
Gaussian Markov Random Field
Graph
Totally positive matrices

ABSTRACT

The present paper focuses on the study of the conditions under which the covariance matrix of a multivariate Gaussian distribution is totally positive, paying particular attention to multivariate Gaussian distributions that are Gaussian Markov Random Fields. More specifically, it is proven that, if the graph over which the Gaussian Markov Random Field is defined consists of path graphs and the covariances between adjacent variables on the graph are non-negative, then there always exists a reordering of the variables that renders the resulting covariance matrix totally positive. Moreover, this reordering is identified and some cases for which the conditions for the covariance matrix of a multivariate Gaussian distribution to be totally positive are necessary and sufficient are provided.

© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

A matrix is called totally positive (resp., strictly totally positive) if all its minors are non-negative (resp., positive). Totally positive matrices arise in many fields of application such as Approximation Theory, Economy, Biology and Computer-Aided Geometric Design (see, e.g., [1–3]). A particularity of non-singular totally positive matrices is that they admit a bidiagonal factorization so that many algebraic computations can be performed to attain High Relative Accuracy (HRA) [4], assuming that the bidiagonal factorization can be obtained with HRA. HRA means that the relative errors of the computations are of the order of machine precision, independently of the size of the matrix condition number. Some examples of such algebraic operations are the computation of the inverse matrix, triangular factorization, computation of eigenvalues and singular values and even the resolution of some linear systems. All of these algebraic operations appear routinely in the field of Statistics when dealing with covariance matrices, for instance when performing Principal Component Analysis, computing conditional distributions and performing the Cholesky factorization of a covariance matrix for simulation purposes.

From an apparently different perspective, Gaussian Markov Random Fields (GRMF) over graphs are a popular statistical tool linking the dependence structure of a multivariate Gaussian distribution to a graph. This type of distribution is widely used in several fields of application such as signal analysis [5], disease control [6], image recognition [7], robotics [8] and general data prediction [9].

In this paper, we provide a link between totally positive matrices and GRMF by proving that, given a GRMF over a graph of paths, there exists a reordering of the variables of the random vector for which the covariance matrix is totally

* Corresponding author.
E-mail addresses: bazjuan@uniovi.es (J. Baz), palonso@uniovi.es (P. Alonso), jmpena@unizar.es (J.M. Peña), perezfernandez@uniovi.es (R. Pérez-Fernández).

https://doi.org/10.1016/j.cam.2023.115098
0377-0427/© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

positive. Moreover, characterizations of the total positivity of the covariance matrix when the graph is acyclic or the GRMF has uniform correlation are provided. These results will be of key interest to practitioners from a computational point of view since they will provide the tools for performing prominent algebraic operations such as inversion and computation of eigenvalues of a covariance matrix with HRA, even in high-dimensional problems. Prototypical stochastic processes for which the presented results are applicable include Gauss–Markov chains and Gaussian processes over the real line, always assuming that the covariance between variables is non-negative.

The remainder of the paper is organized as follows. Section 2 recalls some basic concepts and results on graphs, totally positive matrices and M-matrices, random vectors and GRMF. Section 3 includes the aforementioned results relating GRMF and totally positive matrices and M-matrices. We end with some conclusions in Section 4.

2. Preliminaries

This section is devoted to recall basic notions and to fix the notation used throughout the paper.

2.1. Graph theory

Firstly, we present some basic concepts and results concerning graphs, taking [10] as main reference.

Definition 2.1. A (simple finite) graph G, denoted as $G = (V, E)$, is formed by a finite set of nodes V and a set of edges E, which is a set of subsets of V of cardinality 2. The number of elements of V is called the order of the graph (typically denoted by n) and the number of elements of E is called the size of the graph (typically denoted by m).

Henceforward, we will refer to simple finite graphs simply as graphs.

Definition 2.2. Let $G = (V, E)$ be a graph. Given $u, v \in V$, if $\{u, v\} \in E$, then it is said that u and v are adjacent. The set of adjacent nodes to $u \in V$ is called the neighborhood of u and is denoted by $N(u)$. The cardinality of $N(u)$ is referred to as the degree of incidence of u. The matrix A^G such that $A^G_{i,j} = 1$ if $\{i, j\} \in E$ and $A^G_{i,j} = 0$ otherwise is called the adjacency matrix of G.

The notion of subgraph is of relevance to the present paper.

Definition 2.3. Let $G = (V, E)$ and $G' = (V', E')$ be two graphs. The graph G' is called a subgraph of G if it holds that $V' \subseteq V$ and $E' \subseteq E$.

Also, the notions of walk, path and cycle will be used throughout this paper. In particular, given a graph $G = (V, E)$ and $u_0, u_k \in V$, a sequence of nodes $(u_0, u_1, \ldots, u_k)$ such that $\{u_{i-1}, u_i\} \in E$ for any $i \in \{1, \ldots, k\}$ is called a walk between u_0 and u_k. The number of edges of a walk is referred to as the length of the walk. A walk $(u_0, u_1, \ldots, u_k)$ such that $u_i \neq u_j$ for any $i \neq j$ (except possibly $u_0 = u_k$) is called a path. A walk $(u_0, u_1, \ldots, u_k)$ such that $u_0 = u_k$ is called a cycle.

A graph without any cycle is called acyclic or a forest. If there exists a walk between every pair of nodes, the graph is called connected. An acyclic and connected graph is called a tree. A connected subset of nodes that is maximal on this regard is called a connected component. A well-known property of a tree is that there exists a unique path between any two nodes.

2.2. Totally positive matrices and M-matrices

Let A be a matrix of dimension $n \times n$. Let $1 \leq i_1 < \cdots < i_k \leq n$ and $1 \leq j_1 < \cdots < j_k \leq n$ (with $k \in \{1 \ldots, n\}$) be two subsets of indices of the matrix. The $k \times k$ submatrix of A containing the rows associated with the indices $i_1, \ldots, i_k$ and the columns associated with the indices $j_1, \ldots, j_k$ is denoted as $A_{\{i_1,\ldots,i_k\},\{j_1,\ldots,j_k\}}$. The determinant of this matrix is denoted as $|A|_{\{i_1,\ldots,i_k\},\{j_1,\ldots,j_k\}}$ and is called the minor associated with $\{i_1, \ldots, i_k\}$ and $\{j_1, \ldots, j_k\}$. A minor is called (leading) principal if $\{i_1, \ldots, i_k\} = \{j_1, \ldots, j_k\} = \{1, \ldots, k\}$.

A popular type of matrix is that of positive-(semi)definite matrices. A matrix A of dimension $n \times n$ is called positive-semidefinite if it is symmetric (or Hermitian) and every (leading) principal minor of A is non-negative. If additionally every (leading) principal minor of A is strictly positive, then A is called positive-definite.

Another popular type of matrix is that of (strictly) totally positive matrices [1,2], which will be of key importance to this paper. The reader should be aware that the terms 'totally non-negative' and 'totally positive' are sporadically used in the literature (see, e.g., [2,4]) for referring to the concepts that will be here referred to as 'totally positive' and 'strictly totally positive'.

Definition 2.4. A matrix A of dimension $n \times n$ is called totally positive if every minor of A is non-negative. If additionally any minor of A is strictly positive, then A is called strictly totally positive.

Totally positive matrices have very interesting properties from a computational point of view. In particular, it is possible to approximate computations with High Relative Accuracy (HRA). HRA means that the relative errors of the computations are of the order of the precision of the machine used for the computations, independently of the size of the condition number of the matrix. In this context, algorithms for tasks such as finding the eigenvalues or eigenvectors, factorization and inverse computation of a totally positive matrix attaining HRA have been developed [4,11].

The next result is consequence of Theorem 3.3(c) in [1].

Proposition 2.1 ([1]). *Let S be a diagonal matrix such that $S_{i,i} = 1$ if i is odd and $S_{i,i} = -1$ if i is even. A matrix M is totally positive if and only if $SM^{-1}S$ is totally positive.*

A Z-matrix is a real matrix whose off-diagonal elements are non-positive. A specific type of Z-matrices, called M-matrices, are closely related to totally positive matrices. Non-singular M-matrices have many equivalent definitions. In fact, Berman and Plemmons (see Theorem 2.3 in Chapter 6 of [12]) list fifty equivalent definitions. We shall use the following equivalent definitions.

Definition 2.5. Let A be a real $n \times n$ Z-matrix. The following concepts are equivalent:

(i) A is a non-singular M-matrix.
(ii) A^{-1} is positive.
(iii) The principal minors of A are strictly positive.

Observe that (iii) of the previous definition may be used to prove that a symmetric positive-definite matrix is an M-matrix. Non-singular M-matrices have important applications, for instance, in economics, numerical analysis, analysis of dynamical systems and mathematical programming (see [12]).

2.3. Multivariate Gaussian distributions

In this subsection, basic concepts of multivariate Gaussian distributions are introduced, taking [13,14] as reference.

A multivariate Gaussian distribution is a probability measure over $\mathbb{R}^n$ which is unequivocally determined by a mean vector and a covariance matrix. A random vector is said to have a multivariate Gaussian distribution if any non-null linear combination of its components has a univariate Gaussian distribution [13]. Obviously, a multivariate Gaussian distribution may also be characterized by its density function. A random vector $\vec{X}$ of dimension n is said to have a multivariate Gaussian distribution if its density function has the following expression:

$$f(\vec{x}) = \frac{1}{\sqrt{|2\pi\Sigma|}} \exp\left(-\frac{(\vec{x}-\vec{\mu})^T \Sigma^{-1}(\vec{x}-\vec{\mu})}{2}\right), \qquad \forall \vec{x} \in \mathbb{R}^n,$$

where $\vec{\mu}$ is the mean vector and Σ is the covariance matrix of $\vec{X}$. The fact that $\vec{X}$ has a multivariate Gaussian distribution with mean vector $\vec{\mu}$ and covariance matrix Σ is denoted as $\vec{X} \sim N(\vec{\mu}, \Sigma)$. Additionally, if $\vec{X} \sim N(\vec{\mu}, \Sigma)$, then the random vector defined as $\vec{Y} = A\vec{x} + \vec{v}$ satisfies that $\vec{Y} \sim N(A\vec{\mu} + \vec{v}, A\Sigma A^T)$.

Given the covariance matrix of the distribution, we may define Pearson's correlation coefficient, which measures the strength of the linear dependence between two components of the random vector.

Definition 2.6. Let $\vec{X}$ be a random vector of dimension n with covariance matrix Σ. Pearson's correlation coefficient $\rho_{i,j}$ between X_i and X_j is defined as:

$$\rho_{i,j} = \frac{\Sigma_{i,j}}{\sqrt{\Sigma_{i,i}\Sigma_{j,j}}} = \frac{\sigma_{i,j}}{\sigma_i \sigma_j}.$$

Given three continuous random vectors $\vec{X}$, $\vec{Y}$ and $\vec{Z}$ of dimensions n_X, n_Y and n_Z, respectively, with (joint) density function $f(\vec{x}, \vec{y}, \vec{z})$, $\vec{X}$ and $\vec{Y}$ are said to be conditionally independent given $\vec{Z}$ if there exists a decomposition $f(\vec{x}, \vec{y}, \vec{z}) = h(\vec{x}, \vec{z})g(\vec{y}, \vec{z})$.

The fact that $\vec{X}_A$ and $\vec{X}_B$ are conditionally independent given $\vec{X}_C$ is denoted by $\vec{X}_A \perp \vec{X}_B \mid \vec{X}_C$. The following property concerning conditional independence will be of key importance for multivariate Gaussian distributions that are GRMF (see the upcoming subsection): Two components X_i and X_j of a random vector with multivariate Gaussian distribution $\vec{X}$ are conditionally independent given the value of all other components of $\vec{X}$ if and only if $\left(\Sigma^{-1}\right)_{i,j} = 0$ (see [13]).

2.4. Gaussian Markov random fields

The concept of Markov Random Field (MRF) links the conditional (in)dependence structure of a random vector to the adjacency matrix of a graph. In particular, given a random vector $\vec{X} = (X_1, \ldots, X_n)$ and a graph $G = (V, E)$ with $V = \{1, \ldots, n\}$, the following three properties are of interest:

- The pairwise Markov property: $X_i \perp X_j \mid \vec{X}_{-\{i,j\}}$ for any $i, j \in V$ such that $(i, j) \notin E$ and $i \neq j$, where $\vec{X}_{-\{i,j\}}$ denotes all components of $\vec{X}$ but X_i and X_j.
- The local Markov property: $X_i \perp \vec{X}_{-\{i\}\cup N(i)} \mid \vec{X}_{N(i)}$ for any $i \in V$.
- The global Markov property: $\vec{X}_A \perp \vec{X}_B \mid \vec{X}_C$, for any pairwisely disjoint $A, B, C \subset V$ with $A, B \neq \emptyset$ and such that C separates A and B.

If the random vector has a multivariate Gaussian distribution, the three properties above are equivalent [15]. Additionally, if any of the three properties above holds, the multivariate Gaussian random vector is called a GRMF over G. From a matrix perspective, the following characterization comes up handy.

Theorem 2.1 ([16]). *Let $G = (V, E)$ be a graph with $V = \{1, 2, \ldots, n\}$ and $\vec{X}_V$ be a random vector with multivariate Gaussian distribution and covariance matrix Σ. It holds that $\vec{X}_V$ is a GMRF over G if and only if:*

$$\{i, j\} \notin E \quad \Longrightarrow \quad \left(\Sigma^{-1}\right)_{i,j} = 0.$$

This result above is key to the theory of GMRFs and firstly appeared in a covariance selection problem introduced by Dempster [17]. It is important to highlight that the theorem above does not apply in general for other distributions, thus justifying why, even today, theoretical properties [18–20] and computational aspects [21,22] of GMRFs are still attracting the interest of the research community.

Let us also introduce a particular type of GMRF that is useful in particular topics, see [23–25], in which Pearson's correlation coefficient between adjacent variables in the graph is always the same fixed value.

Definition 2.7. Let $G = (V, E)$ be a graph with $V = \{1, 2, \ldots, n\}$ and $\rho_0 \in (-1, 1)$. A random vector $\vec{X}_V$ is called a GMRF with uniform correlation ρ_0 over G if its covariance matrix is positive-definite and there exist $\sigma_1, \ldots, \sigma_n > 0$ such that:

- $\Sigma_{i,i} = \sigma_i^2$,
- $\Sigma_{i,j} = \sigma_i \sigma_j \rho_0$, if $\{i, j\} \in E$,
- $(\Sigma^{-1})_{i,j} = 0$ if $\{i, j\} \notin E$, if $i \neq j$.

We end this section by providing a couple of basic results regarding the correlation structure of GMRFs over acyclic graphs.

Corollary 2.1 ([26]). *Let $\vec{X}$ be a GMRF over a tree $G = (V, E)$. Pearson's correlation coefficient between two components X_i and X_j of $\vec{X}$ is the product of Pearson's correlation coefficients between adjacent variables of the unique walk that connects X_i and X_j.*

The latter result may be extended to acyclic graphs. It is only necessary to note that Pearson's correlation coefficient between variables of different connected components is zero. Thus, considering the correct order of the variables, the covariance matrix is block-diagonal, where each block is associated with a connected component of the graph.

In general, a GMRF over a graph G does not need to be a GMRF over a subgraph of G. However, if G is a tree, we still have a GMRF when removing edges $\{i, j\}$ associated with components of the GMRF that are uncorrelated.

Proposition 2.2. *Let $\vec{X}$ be a GMRF over a tree $G = (V, E)$. If there exists an edge $\{i, j\}$ such that $\rho_{i,j} = 0$, then $\vec{X}$ is a GMRF over the graph $G = (V, E\backslash\{i, j\})$.*

Proof. Consider an edge $\{i, j\}$ such that $\rho_{i,j} = 0$. For any $v \in V$, we introduce the notation $c_v = (v, \ldots, i)$ to represent the unique path from v to i. We separate the set of nodes (that are all connected since G is a tree) in those that are connected to i by a walk that does not pass through j, resulting in the set $C_i = \{v \in V \mid j \notin c_v\}$, and those that are connected to i by a walk that passes through j, resulting in the set $C_j = \{v \in V \mid j \in c_v\}$. It trivially holds that $C_i \cup C_j = V$ and $C_i \cap C_j = \emptyset$. Since there exists a unique path between any two nodes of a tree, it follows that any walk between $v_i \in C_i$ and $v_j \in C_j$ contains the edge $\{i, j\}$, therefore, as a consequence of Corollary 2.1, it holds that $\vec{X}_{C_i} \perp \vec{X}_{C_j}$. This implies that $\vec{X}$ is a GMRF over a graph in which C_i and C_j are not connected, and in particular over $G = (V, E\backslash\{i, j\})$. $\square$

3. Gaussian Markov random fields and totally positive matrices

As introduced in Section 2.2, totally positive matrices are matrices where all minors are non-negative. In this section, we will study the conditions under which a covariance matrix is totally positive and its relation with GMRFs. Let us start with a very simple example that shows that it is easy to find positive-definite matrices (even with positive elements) that are not totally positive:

Example 3.1. Consider the following positive-definite matrices M and T:

$$M = \begin{pmatrix} 1 & 0.5 & 0.6 \\ 0.5 & 1 & 0.4 \\ 0.6 & 0.4 & 1 \end{pmatrix}, \qquad T = \begin{pmatrix} 1 & 0.5 & 0.2 \\ 0.5 & 1 & 0.4 \\ 0.2 & 0.4 & 1 \end{pmatrix}.$$

On the one hand, M is not totally positive since $\begin{vmatrix} 0.5 & 0.6 \\ 1 & 0.4 \end{vmatrix} < 0$. On the other hand, it is easy to check that T is totally positive.

Before starting the study of GMRFs with a totally positive covariance matrix, it is important to clarify the difference with another similar concept with the term *totally positive* appearing in its name. A multivariate distribution is said to be multivariate totally positive of order 2 (MTP2) if its density function $f(\vec{x})$ is such that $f(\vec{x})f(\vec{y}) \leq f(\vec{x} \wedge \vec{y})f(\vec{x} \vee \vec{y})$ for any $\vec{x}, \vec{y} \in \mathbb{R}$, where $\wedge$ and $\vee$ denote, respectively, the component-wise maximum and minimum [27]. Interestingly, a multivariate Gaussian distribution is MTP2 if and only if the inverse of the covariance matrix is a symmetric M-matrix [27]. As a counterexample for the equivalence of both the covariance matrix being totally positive and the inverse of the covariance matrix being an M-matrix, we introduce an example of a matrix that is not totally positive but such that its inverse is an M-matrix:

$$\Sigma = \begin{pmatrix} 1 & 0.5 & 0.5 \\ 0.5 & 1 & 0.5 \\ 0.5 & 0.5 & 1 \end{pmatrix}, \qquad \Sigma^{-1} = \begin{pmatrix} 1.5 & -0.5 & -0.5 \\ -0.5 & 1.5 & -0.5 \\ -0.5 & -0.5 & 1.5 \end{pmatrix}.$$

3.1. Properties of totally positive covariance matrices

In this subsection, we will assume that we are dealing with a positive-definite matrix that is also totally positive and find the transformations that do not alter these properties. Moreover, we will link these transformations to covariance matrices of multivariate Gaussian distributions.

For instance, if a random vector $\vec{X}$ has a multivariate Gaussian distribution with a totally positive covariance matrix, then any random vector obtained from $\vec{X}$ as a result of a translation and/or a (non-negative) rescaling of each of the components also has a multivariate Gaussian distribution with a totally positive covariance matrix.

Proposition 3.1. *Let $\vec{X} \sim N(\vec{\mu}, \Sigma)$ be a random vector with multivariate Gaussian distribution with Σ being totally positive and D be a diagonal matrix with non-negative elements. It follows that the covariance matrix of $\vec{Y} = D\vec{X} + \vec{v}$ is totally positive.*

Proof. From the properties of the multivariate Gaussian distribution, it follows that $\vec{Y} \sim N(D\vec{\mu} + \vec{v}, D\Sigma D^T)$. Since D and D^T are diagonal matrices with non-negative elements, they are totally positive. In addition, as Σ is totally positive, it finally follows that $D\Sigma D^T$ is totally positive since the product of totally positive matrices is also totally positive (see Theorem 3.1 in [1]). □

However, a reordering of the variables of the random vectors does not necessarily maintain the total positivity of the covariance matrix, as can be seen in the following example. Interestingly, this same type of transformation does not alter the positive-definiteness of a matrix [28].

Example 3.2. Let T' be the matrix obtained from permuting the first and second rows and columns of the matrix T from Example 3.1:

$$T' = \begin{pmatrix} 1 & 0.5 & 0.4 \\ 0.5 & 1 & 0.2 \\ 0.4 & 0.2 & 1 \end{pmatrix}.$$

This matrix is no longer totally positive since $\begin{vmatrix} 0.5 & 1 \\ 0.4 & 0.2 \end{vmatrix} = -0.3$ is negative. Note that both T and T' are positive-definite.

From the example above, it is concluded that the ordering of the indices of the rows and columns is important when asking a covariance matrix to be totally positive. In particular, we may have a covariance matrix that is not totally positive, yet it admits a reordering of the indices with a covariance matrix that is totally positive. When dealing with a multivariate Gaussian distribution, this type of transformations is just a reordering of the components of the random vector [13]. In this direction, it is interesting to change the main question of the present study from *identifying the conditions under which a covariance matrix is totally positive* to *identifying the conditions under which a covariance matrix admits a reordering of the indices that renders the covariance matrix totally positive*.

3.2. Characterization of totally positive covariance matrices

From now on, we will focus on the study of necessary and sufficient conditions that a covariance matrix must fulfill in order to be totally positive (or to admit a reordering of the indices that renders the covariance matrix totally positive). In particular, we will prove that, given a multivariate Gaussian distribution $\vec{X} \sim N(\vec{\mu}, \Sigma)$, its covariance matrix is totally positive if the random vector is a GMRF over a particular type of graph in which any connected component is a path and Pearson's correlation coefficient between adjacent variables in the graph is non-negative.

As illustrated previously, the ordering of the variables is important when studying the total positivity of a covariance matrix. In this direction, let us define a type of graph with a particular ordering of its nodes.

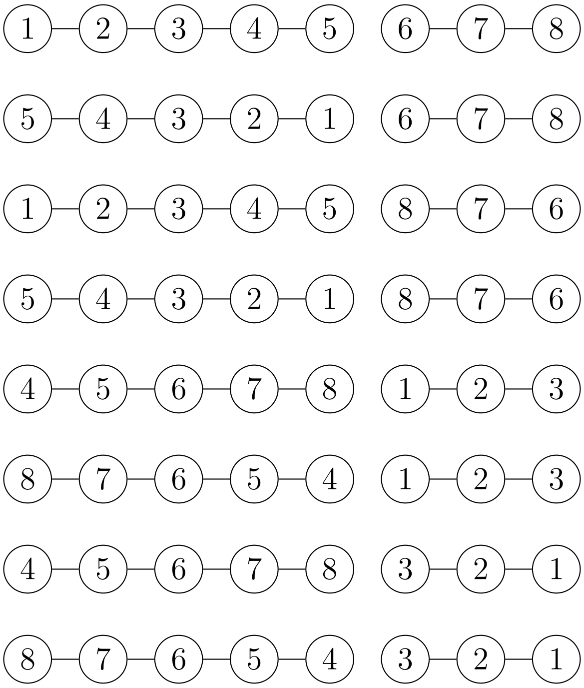

Fig. 1. All natural orderings for a graph consisting of a path of order 5 and a graph of order 3.

Definition 3.1. Let $G = (V, E)$ with $V = \{1, \dots, n\}$ be a graph such that any connected component is a path graph. If for any $i, j \in V$ with $i < j$ and $\{i, j\} \in E$ it holds that $j - i = 1$, then G is called a graph of paths with natural ordering.

The term 'natural ordering' has been chosen because it is the most intuitive way to index the vertices of a graph consisting of different path graphs. In particular, we index consecutively all vertices within the same connected component, starting from one end of the path to the other one. This process is repeated for all connected components. As an illustrative example, in Fig. 1 we provide all the possible natural orderings of a graph of paths consisting of a path of order 5 and another one of order 3.

Remark 3.1. A graph of paths with natural ordering satisfies the following properties:

(i) Given two connected components C_i and C_j, it either holds that $\ell_i < \ell_j$ for any $\ell_i \in C_i$ and $\ell_j \in C_j$ or $\ell_i > \ell_j$ for any $\ell_i \in C_i$ and $\ell_j \in C_j$.
(ii) The adjacency matrix is tridiagonal.
(iii) The graph is acyclic.

Considering the previous remark, it follows that the inverse of the covariance matrix of a GMRF over a graph of paths with natural ordering is tridiagonal.

In the next result, we will prove that any GMRF over a graph of paths with natural orderings has a totally positive covariance matrix in case all covariances between adjacent variables in the graph are non-negative.

Proposition 3.2. *Let $\vec{X}_V = (X_1, \dots, X_n)$ be a GMRF over a graph of paths with natural ordering $G = (V, E)$ with $V = \{1, \dots, n\}$. If the covariance between adjacent variables in the graph is non-negative, then its covariance matrix is totally positive.*

Proof. On the one hand, since we are dealing with a GMRF over an acyclic graph and the covariance between adjacent variables in the graph is non-negative, all elements of the covariance matrix are non-negative, see Corollary 2.1 and the subsequent comments. From Proposition 5.3 in [27], it follows that we are dealing with a distribution that is MTP2 and, in particular, that Σ^{-1} is a non-singular and symmetric M-matrix. On the other hand, Σ^{-1} is a tridiagonal matrix since we are dealing with a GMRF over a graph of paths with natural ordering. Since Σ^{-1} simultaneously is a non-singular M-matrix and a tridiagonal matrix, it finally follows from Theorem 2.2 in [29] that Σ is a totally positive matrix. □

Due to Proposition 2.1, the latter result may be reformulated in terms of the inverse of the covariance matrix.

Corollary 3.1. *Let $\vec{X}_V = (X_1, \ldots, X_n)$ be a GMRF over a graph of paths with natural ordering $G = (V, E)$ with $V = \{1, \ldots, n\}$. If the covariance between adjacent variables in the graph is non-positive, then the inverse of its covariance matrix is totally positive.*

Proof. Consider the transformation $S\vec{X}$ with S the matrix of Proposition 2.1. The covariance matrix of $S\vec{X}$ is $S\Sigma S$. Since S is diagonal, the resulting random vector still is a GMRF over G. In addition $S^T \Sigma S$ is totally positive (see Proposition 3.2). Consider $i, j \in V$ with $i < j$ and $\{i, j\} \in E$. This implies that $j - i = 1$ and therefore $(S\Sigma S)_{i,j} = -\Sigma_{i,j} \geq 0$, i.e., the transformation changes the sign of the covariance between adjacent variables. Finally, noticing that $S^{-1} = S^T = S$, it follows from Proposition 2.1 that $S(S^T \Sigma S)^{-1} S = SS\Sigma^{-1}SS = \Sigma^{-1}$ is totally positive. $\square$

The latter results provide sufficient conditions for (the inverse of) a positive-definite matrix to be totally positive. However, these conditions are not necessary, as can be seen in the following example.

Example 3.3. The following matrix M is positive-definite and totally positive. However, since the inverse matrix M^{-1} of M does not contain any zeros, therefore M cannot be the covariance matrix of any graph of paths with natural ordering.

$$M = \begin{pmatrix} 1 & 0.7 & 0 \\ 0.7 & 1 & 0.7 \\ 0 & 0.7 & 1 \end{pmatrix}, \qquad M^{-1} = \begin{pmatrix} 25.5 & -35 & 24.5 \\ -35 & 50 & -35 \\ 24.5 & -35 & 25.5 \end{pmatrix}.$$

In general, we cannot find necessary and sufficient conditions for (the inverse of) a positive-definite matrix to be totally positive, and, in particular, we cannot find necessary and sufficient conditions for (the inverse of) the covariance matrix of a GMRF to be totally positive. However, when restricting our search to GMRFs over acyclic graphs, the conditions that were proven above to be sufficient are now also necessary.

Theorem 3.1. *Let $\vec{X}_V = (X_1, \ldots, X_n)$ be a GMRF over an acyclic graph $G = (V, E)$ with $V = \{1, \ldots, n\}$.*

(i) The covariance matrix Σ of $\vec{X}$ is totally positive if and only if $\vec{X}$ is a GMRF over a graph of paths with natural ordering and the covariance between adjacent variables in the graph is non-negative.
(ii) The inverse Σ^{-1} of the covariance matrix Σ of $\vec{X}$ is totally positive if and only if $\vec{X}$ is a GMRF over a graph of paths with natural ordering and the covariance between adjacent variables in the graph is non-positive.

Proof. (i) The right-to-left implication has already been proven in Proposition 3.2. We now prove the left-to-right implication. Notice that any element of the covariance matrix must be non-negative, since a negative element will result in a negative minor. Thus, the covariance between adjacent variables in the graph is non-negative. Without loss of generality (see Proposition 3.1), we can work with the correlation matrix S instead of working directly with the covariance matrix.

If all nodes have degree of incidence smaller than or equal to two, the graph consists of connected components that are path graphs. Otherwise, let $i_0 \in V$ be a node with degree of incidence greater than or equal to three and let i_1, i_2, i_3 be three nodes that are adjacent to i_0 such that $i_1 < i_2 < i_3$. Since G is acyclic, from Corollary 2.1, it follows that $\rho_{i_1,i_2} = \rho_{i_0,i_1}\,\rho_{i_0,i_2}$, $\rho_{i_1,i_3} = \rho_{i_0,i_1}\,\rho_{i_0,i_3}$ and $\rho_{i_2,i_3} = \rho_{i_0,i_2}\,\rho_{i_0,i_3}$.

Therefore, $S_{\{i_1,i_2,i_3\}}$ has the following structure:

$$S_{\{i_1,i_2,i_3\}} = \begin{pmatrix} 1 & \rho_{i_0,i_1}\,\rho_{i_0,i_2} & \rho_{i_0,i_1}\,\rho_{i_0,i_3} \\ \rho_{i_0,i_1}\,\rho_{i_0,i_2} & 1 & \rho_{i_0,i_2}\,\rho_{i_0,i_3} \\ \rho_{i_0,i_1}\,\rho_{i_0,i_3} & \rho_{i_0,i_2}\,\rho_{i_0,i_3} & 1 \end{pmatrix}.$$

Computing one of the minors, we obtain:

$$|S|_{\{i_1,i_2\},\{i_2,i_3\}} = \begin{vmatrix} \rho_{i_0,i_1}\,\rho_{i_0,i_2} & \rho_{i_0,i_1}\,\rho_{i_0,i_3} \\ 1 & \rho_{i_0,i_2}\,\rho_{i_0,i_3} \end{vmatrix} = \rho_{i_0,i_1}\,\rho_{i_0,i_3}\left(\rho_{i_0,i_2}^2 - 1\right).$$

Since $\rho_{i_0,i_1}, \rho_{i_0,i_3} \geq 0$ and $\left(\rho_{i_0,i_2}^2 - 1\right) < 0$ and the latter minor must be non-negative, it either holds that $\rho_{i_0,i_1} = 0$ or $\rho_{i_0,i_3} = 0$ (or both). Thus, the edge $\{i_0, i_1\}$ or the edge $\{i_0, i_3\}$ can be deleted and $\vec{X}_V$ will still be a GMRF over G, as a direct consequence of Proposition 2.2.

We may repeat this process until the node i_0 has degree of incidence equal to two. Analogously, we may repeat this process for all nodes with degree of incidence greater than two until we obtain an acyclic graph in which all nodes have at most incidence 2. This graph consists of connected components that are path graphs.

Now, let us prove that the ordering of the nodes must be the natural ordering. Consider $i, j \in V$ with $i < j$ such that $\{i, j\} \in E$ but $j - i > 1$. There exists $k \in V$ such that $i < k < j$. The node k must be contained in the same connected component as i and j, as mentioned in Remark 3.1. Without loss of generality, suppose that the distance between i and k is smaller than the distance between j and k. Thus, the unique path between k and i passes through j and, as a result of Corollary 2.1, it holds that $\rho_{i,k} = \rho_{i,j}\rho_{j,k}$. The structure of $S_{\{i,k,j\}}$ is the following:

$$S_{\{i,k,j\}} = \begin{pmatrix} 1 & \rho_{i,j}\,\rho_{j,k} & \rho_{i,j} \\ \rho_{i,j}\,\rho_{j,k} & 1 & \rho_{j,k} \\ \rho_{i,j} & \rho_{j,k} & 1 \end{pmatrix}.$$

The minor associated with $\{i, k\}$ and $\{k, j\}$ has the following expression:

$$|S|_{\{i,k\},\{k,j\}} = \begin{vmatrix} \rho_{i,j}\rho_{j,k} & \rho_{i,j} \\ 1 & \rho_{j,k} \end{vmatrix} = \rho_{i,j}\left(\rho_{j,k}^2 - 1\right).$$

Since $\rho_{i,j} \geq 0$, the only option that renders the latter minor non-negative is $\rho_{i,j} = 0$.

Analogously to the previous cases, the edge $\{i, j\}$ can be eliminated, as a result of Proposition 2.2. Repeating the process, the resulting graph is such that, for any $i, j \in V$ with $i < j$ and with $\{i, j\} \in E$, the equality $j - i = 1$ holds.

(ii) The result follows as a consequence of (i) and Proposition 2.1. □

As a consequence, it is concluded that a GMRF over an acyclic graph admits a reordering of its variables that renders its covariance matrix totally positive if and only if any component of the graph is a path and the covariance between all variables is non-negative.

An immediate consequence is that the covariance matrix of a GMRF over a graph of paths with natural ordering is totally positive if and only the inverse matrix of the covariance matrix is a tridiagonal M-matrix. The result can be further generalized to acyclic graphs as follows.

Proposition 3.3. *Let $\vec{X}$ be a GMRF over an acyclic graph G with covariance matrix Σ. It holds that G is a graph of paths with natural ordering and Σ is totally positive if and only if Σ^{-1} is a tridiagonal M-matrix.*

Proof. For the left-to-right implication, notice that if Σ is totally positive, from Proposition 2.1, it follows that $S\Sigma^{-1}S$ is totally positive. In addition, since G is a graph of paths with natural ordering, it holds that Σ^{-1} is tridiagonal (see Remark 3.1). Thus, since Σ^{-1} is tridiagonal and the off-diagonal elements of $S\Sigma^{-1}S$ are non-negative, it is concluded that Σ^{-1} is a Z-matrix. Finally, since Σ^{-1} is also positive-definite, it follows by Definition 2.5 that Σ^{-1} is an M-matrix.

For the right-to-left implication, assume that Σ^{-1} is a tridiagonal M-matrix. It follows from Theorem 2.2 in [29] concerning M-matrices whose inverses are totally positive that Σ is totally positive. From Theorem 3.1, it follows that G is a graph of paths with natural ordering. □

As a direct consequence of Theorem 3.1 and Proposition 3.3, any GMRF over an acyclic graph with a totally positive matrix is always MTP2. The converse implication is not true in general, since we also need the inverse of the covariance matrix to be tridiagonal.

Corollary 3.2. *Let $\vec{X}$ be a GMRF over an acyclic graph G with a totally positive covariance matrix Σ. It holds that $\vec{X}$ is MTP2.*

3.3. Characterization with uniform correlation

In the previous section, a characterization of the total positivity of covariance matrices of GMRFs over acyclic graphs was provided. This subsection provides general results (not restricted to acyclic graph) for a specific type of GMRF: GMRFs with uniform correlation.

Proposition 3.4. *Let $\vec{X}_V$ be a GMRF with uniform correlation ρ_0 over $G = (V, E)$. If the covariance matrix Σ of $\vec{X}_V$ is totally positive, then G must be acyclic.*

Proof. If the covariance matrix Σ of $\vec{X}_V$ is totally positive, then all elements are non-negative and, therefore, $\rho_0 \geq 0$. If $\rho_0 = 0$, the covariance matrix is diagonal (see Theorem 1 in [16]), thus the distribution is a GMRF over the graph with no edges, which is acyclic. We now prove that $\rho_0 \in (0, 1)$, assuming that G is cyclic and reaching a contradiction.

Without loss of generality (see Proposition 3.1), we can work with the correlation matrix S instead of working directly with the covariance matrix. Let $(i_1, i_2, \ldots, i_{k-1}, i_1)$ be a cycle in G of length $k - 1 \geq 3$. Since $\{i_1, i_2\}, \{i_2, i_3\} \in E$, it holds that $\rho_{i_1,i_2} = \rho_{i_2,i_3} = \rho_0$ and, therefore, the minor associated with $\{i_1, i_2\}$ and $\{i_2, i_3\}$ is the following:

$$|S|_{\{i_1,i_2\},\{i_2,i_3\}} = \begin{vmatrix} \rho_0 & \rho_{i_1,i_3} \\ 1 & \rho_0 \end{vmatrix} = \rho_0^2 - \rho_{i_1,i_3}.$$

Since S is totally positive, $\rho_{i_1,i_3} \geq 0$ and, therefore, it is necessary that $\rho_{i_1,i_3} \leq \rho_0^2$. If $i_1 = i_4$, then we reach a contradiction since $\{i_1, i_3\} \in E$ implies $\rho_{i_1,i_3} = \rho_0$ (and $\rho_0 > \rho_0^2$ since $\rho_0 \in (0, 1)$). If $i_1 \neq i_4$, consider the minor associated with $\{i_1, i_3\}$ and $\{i_3, i_4\}$:

$$|S|_{\{i_1,i_3\},\{i_3,i_4\}} = \begin{vmatrix} \rho_{i_1,i_3} & \rho_{i_1,i_4} \\ 1 & \rho_0 \end{vmatrix} = \rho_0\,\rho_{i_1,i_3} - \rho_{i_1,i_4}.$$

Since $\rho_{i_1,i_3} \leq \rho_0^2$ for the minor $S_{\{i_1,i_2\},\{i_2,i_3\}}$ to be non-negative, it is necessary that $\rho_{i_1,i_4} \leq \rho_0\,\rho_{i_1,i_3} \leq \rho_0^3$. If $i_1 = i_5$, then we reach a contradiction since $\{i_1, i_4\} \in E$ implies $\rho_{i_1,i_4} = \rho_0$ (and $\rho_0 > \rho_0^3$ since $\rho_0 \in (0, 1)$). If $i_1 \neq i_5$, we may proceed iteratively and prove that $\rho_{i_1,i_j} \leq \rho_0^{j-1}$ with $j \in \{1, \ldots, k - 1\}$. However, since $\{i_1, i_{k-1}\} \in E$, $\rho_{i_1,i_{k-1}} = \rho_0$, we reach a contradiction since $\rho_{i_1,i_{k-1}} \leq \rho_0^{k-2} < \rho_0$. □

Finally, a characterization of the total positivity of the covariance matrix of GMRFs with uniform correlation is obtained.

Theorem 3.2. *Let $\vec{X}_V = (X_1, \ldots, X_n)$ be a GMRF with uniform correlation ρ_0 over a graph $G = (V, E)$ with $V = \{1, \ldots, n\}$.*

(i) The covariance matrix Σ of $\vec{X}$ is totally positive if and only if G is a graph of paths with natural ordering and $\rho_0 \geq 0$.
(ii) The inverse Σ^{-1} of the covariance matrix Σ of $\vec{X}$ is totally positive if and only if G is a graph of paths with natural ordering and $\rho_0 \leq 0$.

Proof. (i) The left-to-right implication is a direct result of Proposition 3.2. The right-to-left implication is a result of the fact that G is acyclic due to Proposition 3.4 and, therefore, Σ is totally positive due to Theorem 3.1.
(ii) The result follows as a consequence of (i) and Proposition 2.1. □

It is concluded that a GMRF with uniform correlation ρ_0 admits a reordering of the variables that renders its covariance matrix totally positive if and only if the connected components of the graph are paths and $\rho_0 \geq 0$.
We end this section by concluding that there does not exist a GMRF with uniform correlation of dimension greater than or equal to two with a covariance matrix that is strictly totally positive.

Corollary 3.3. *Let $\vec{X}_V = (X_1, \ldots, X_n)$ be a GMRF with uniform correlation ρ_0 over a graph $G = (V, E)$ with $V = \{1, \ldots, n\}$ and $n \geq 3$. None of Σ and Σ^{-1} can be strictly totally positive.*

Proof. If Σ is strictly totally positive, then it is totally positive. From Theorem 3.2, it follows that $\vec{X}$ is a GMRF over a graph of paths with natural ordering. If this graph of paths with natural ordering is not connected, then there exist $i, j \in V$ belonging to two different connected components, thus $\Sigma_{i,j} = 0$ and Σ cannot be strictly totally positive. If this graph of paths with natural ordering is connected, we assume without loss of generality that the variances of the variables are 1 (see Proposition 3.1). Let $|\Sigma|_{\{1,2\},\{2,3\}}$ be the minor associated with the indices $\{1, 2\}$ and $\{2, 3\}$:

$$|\Sigma|_{\{1,2\},\{2,3\}} = \begin{vmatrix} \rho_0 & \rho_0^2 \\ 1 & \rho_0 \end{vmatrix} = \rho_0^2 - \rho_0^2 = 0.$$

It is concluded that Σ cannot be strictly totally positive. For Σ^{-1}, it suffices to note that the inverse of the covariance matrix of a GMRF over a graph of paths with natural ordering is tridiagonal . □

4. Conclusions and future research

A sufficient condition for a multivariate Gaussian distribution to have a totally positive covariance matrix has been given. In particular, in such case the multivariate Gaussian distribution needs to be a GMRF over a graph of paths with natural ordering and the covariance between adjacent variables needs to be non-negative. Moreover, when restricting the study to GMRFs over acyclic graphs or to GMRFs with uniform correlation, the latter condition is also proven to be necessary. Similar results have been obtained for the inverse of the covariance matrix, but requiring instead that the covariance between adjacent variables needs to be non-positive.
A potential consequence of these results is that recurrent operations such as eigenvalue computation, matrix inversion or Cholesky factorization of covariance matrices that are totally positive may be performed with HRA methods. In this direction, the design of numerical algorithms to take advantage of the definite-positivity, total positivity and sparseness of the inverse of this type of matrices is left for a future study.
Another interesting open problem concerns the study of necessary and sufficient conditions for the total positivity of the covariance matrices of GMRFs that are not acyclic and do not have uniform correlation.

Data availability

No data was used for the research described in the article.

Acknowledgments

This research was partially supported through the Spanish research grant PGC2018-096321-B-I00 (MCIU/AEI), Gobierno de Aragón, Spain (E41_20R) and the Spanish Ministry of Science and Technology (TIN-2017-87600-P).

References

[1] T. Ando, Totally positive matrices, Linear Algebra Appl. 90 (1987) 165–219.
[2] S.M. Fallat, C.R. Johnson, Totally Nonnegative Matrices, Princeton University Press, Princeton, 2011.
[3] A. Pinkus, Totally Positive Matrices, Cambridge University Press, Cambridge, UK, 2010.
[4] P. Koev, Accurate computations with totally nonnegative matrices, SIAM J. Matrix Anal. Appl. 29 (3) (2007) 731–751.
[5] R. Torkamani, H. Zayyani, Statistical graph signal recovery using variational Bayes, IEEE Trans. Circuits Syst. II 68 (6) (2020) 2232–2236.

[6] K.A. Alene, C.A. Gordon, A.C.A. Clements, G.M. Williams, D.J. Gray, X. Zhou, Y. Li, J. Utzinger, J. Kurscheid, S. Forsyth, J. Zhou, Z. Li, G. Li, D. Lin, Z. Lou, S. Li, J. Ge, J. Xu, X. Yu, F. Hu, S. Xie, D.P. Mcmanus, Spatial analysis of schistosomiasis in hunan and jiangxi provinces in the People's Republic of China, Diseases 10 (4) (2020) 93.
[7] S. Huadong, Z. Pengfei, Z. Yingjing, Multi-angle face recognition based on GMRF, in: International Conference on Business Intelligence and Information Technology, 2021, pp. 366–378.
[8] L. Nguyen, S. Kodagoda, R. Ranasinghe, G. Dissanayake, Mobile robotic sensors for environmental monitoring using Gaussian Markov random field, Robotica 39 (5) (2021) 862–884.
[9] J. Oskarsson, P. Sidén, F. Lindsten, Scalable deep gaussian markov random fields for general graphs, in: Proceedings of the 39th International Conference on Machine Learning, 2022, pp. 17117–17137.
[10] W. Kocay, D.L. Kreher, Graphs, Algorithms, and Optimization, CRC Press, Boca Ratón, 2016.
[11] P. Koev, Accurate eigenvalues and SVDs of totally nonnegative matrices, SIAM J. Matrix Anal. Appl. 27 (1) (2005) 1–23.
[12] A. Berman, R.J. Plemmons, Nonnegative Matrices in the Mathematical Sciences, Academic Press, New York, 1979.
[13] K.V. Mardia, J.T. Kent, J.M. Bibby, Multivariate Analysis, Academic Press INC, San Diego, 1979.
[14] V.K. Rohatgi, An Introduction to Probability Theory and Mathematical Statistics, John Wiley and Sons, New York, 1976.
[15] H. Rue, L. Held, Gaussian Markov Random Fields: Theory and Applications, CRC press, Boca Ratón, 2005.
[16] T.P. Speed, H.T. Kiiveri, Gaussian Markov distributions over finite graphs, Ann. Statist. 14 (1986) 138–150.
[17] A.P. Dempster, Covariance selection, Biometrics 28 (1972) 157–175.
[18] V. Borovitsky, I. Azagulov, A. Terenin, P. Mostowsky, M. Deisenroth, N. Durrande, Matérn Gaussian processes on graphs, in: International Conference on Artificial Intelligence and Statistics, 2021, pp. 2593–2601.
[19] A.L.M. Levada, On the Kullback–Leibler divergence between pairwise isotropic Gaussian-Markov random fields, Stat. Probab. Lett. (2022) http://dx.doi.org/10.2139/ssrn.4089590, in press.
[20] Y.C. MacNab, On coregionalized multivariate Gaussian Markov random fields: construction, parameterization, and Bayesian estimation and inference, TEST (2022) http://dx.doi.org/10.1007/s11749-022-00832-z, in press.
[21] D. Bolin, J. Wallin, Efficient methods for Gaussian Markov random fields under sparse linear constraints, Adv. Neural Inf. Process. Syst. 34 (2021) 9882–9894.
[22] P. Sidén, L. Fredrik, Deep gaussian markov random fields, in: International Conference on Machine Learning, 2020, pp. 8916–8926.
[23] J. Baz, I. Díaz, S. Montes, R. Pérez-Fernández, A fuzzy order for graphs based on the continuous entropy of Gaussian Markov random fields, in: 19th World Congress of the International Fuzzy Systems Association (IFSA), 12th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT), and 11th International Summer School on Aggregation Operators, AGOP, Atlantis Press, 2021, pp. 509–516.
[24] P. Csikvári, B. Szegedy, On Sidorenko's conjecture for determinants and Gaussian Markov random fields, Random Struct. Algorithms 62 (2023) 335–375, http://dx.doi.org/10.1002/rsa.21112.
[25] U. Grenander, Y.-S. Chow, D.M. Keenan, Hands: A Pattern Theoretic Study of Biological Shapes, Volume 2, Springer Science & Business Media, Luxembourg, 2012.
[26] J. Baz, I. Díaz, S. Montes, R. Pérez-Fernández, Some results on the Gaussian Markov random field construction problem based on the use of invariant subgraphs, TEST 31 (2022) 1–19.
[27] S. Lauritzen, C. Uhler, P. Zwiernik, Maximum likelihood estimation in Gaussian models under total positivity, Ann. Statist. 47 (2019) 1835–1863.
[28] R. Bhatia, Positive Definite Matrices, Princeton University Press, Princeton, 2007.
[29] J.M. Peña, $M-$Matrices whose inverses are totally positive, Linear Algebra Appl. 221 (1995) 189–193.