Mathematical Methods in the Applied Sciences

WILEY

RESEARCH ARTICLE OPEN ACCESS

High Relative Accuracy Computations With Covariance Matrices of Order Statistics

Juan Baz[1] | Pedro Alonso[2] | Juan Manuel Peña[3] | Raúl Pérez-Fernández[1]

[1]Departamento de Estadística, I.O. y D.M., Universidad de Oviedo, Oviedo, Spain | [2]Departamento de Matemáticas, Universidad de Oviedo, Oviedo, Spain | [3]Departamento de Matemática Aplicada, Universidad de Zaragoza, Zaragoza, Spain

Correspondence: Pedro Alonso (palonso@uniovi.es)

Received: 18 July 2025 | **Revised:** 23 December 2025 | **Accepted:** 13 February 2026

Keywords: covariance matrix | Gaussian Markov Random Field | high relative accuracy | order statistics | total positivity

ABSTRACT

In many statistical applications, numerical computations with covariance matrices need to be performed. The error made when performing such numerical computations increases with the condition number of the covariance matrix, which is related to the number of variables and the strength of the correlation between the variables. In a recent work, a method for estimating the covariance matrix of a Gaussian Markov Random Field under a total positivity constraint was proposed. This estimation allows for performing many numerical computations with covariance matrices to high relative accuracy (the relative error is of the order of machine precision). However, the necessary conditions for this estimation method to produce a covariance matrix that is close to the population covariance matrix may be too demanding for real-life data. In this paper, we study a particular setting related to order statistics in which these necessary conditions are inherently satisfied. In addition to the theoretical study, an extensive discussion concerning many potential applications is addressed, and a real-life example of an application related to sports data is presented.

MSC2020 Classification: 62G30, 65F30

1 | Introduction

Ordering the values of a random sample is a common procedure in estimation [1, 2] and prediction [3, 4] problems. In case the values of the sample are considered to be independent and identically distributed, which is a common assumption in Statistics, the behavior of the ordered sample has been extensively studied in the literature (see for instance [5]). In such a scenario, we are sometimes interested in the associated covariance matrix, which contains all the information concerning the joint linear variation of the variables with respect to their mean. Moreover, some algebraic computations, such as the inversion and the computation of the eigenvalues of this covariance matrix, have to be computed in some applications related to estimation, reduction of dimension, or the computation of conditional covariances (see Section 5). Unfortunately, sometimes the covariance matrix is ill-conditioned, and the usual algorithms have poor accuracy when performing algebraic computations.

A way to ensure that these computations are performed with high relative accuracy (HRA) when working with covariance matrices arises when these matrices are totally positive [6]. In [7], it is proved that it is possible to achieve HRA when computing the eigenvalues, inverting the matrix, and solving some linear systems of equations with covariance matrices under some sufficient conditions. In addition, statistical methods for estimating a totally positive population covariance matrix from a sample are given in [8]. Unfortunately, these sufficient conditions are quite restrictive, and although they may hold for certain datasets, this is not typically the case. Fortunately, as will be proven throughout this paper, theoretical results demonstrate that, under very weak assumptions, these sufficient conditions are fulfilled when the data of interest is ordinal in nature, and, thus, we deal with order statistics.

More specifically, in this paper, we study conditions under which the covariance matrix of the order statistics is totally positive. Interestingly, these conditions for guaranteeing the total positivity are linked to particular multivariate Gaussian distributions (Gaussian Markov Random Fields) to which the central order statistics converge. As a consequence, this study affirms that, in general, the estimation method for a covariance matrix under a total positivity constraint presented in [8], can be applied to order statistics and, therefore, the inverse and eigenvalues of the estimated covariance matrix can be calculated with HRA.

The remainder of the paper is organized as follows. In Section 2, the main concepts regarding Gaussian Markov Random Fields, totally positive matrices, and their relation are provided. The link between order statistics and total positivity is explained in detail in Section 3. Section 4 is devoted to proving the asymptotic total positivity of the covariance matrix of the order statistics. In Section 5, we present different applications in which HRA computations with order statistics may be useful, whereas an example with real-life data illustrates the benefits of such procedures in Section 6. Finally, some conclusions are presented in Section 7.

2 | Preliminaries

Throughout this paper, we deal with square matrices of dimension n, considering the notation $(M_{i,j})_{i,j\in\{1,\dots,n\}}$. Given $A, B \subseteq \{1,\dots,n\}$, the associated submatrix will be denoted by $M_{A,B}$. If $A = B$, then $M_{A,B}$ will be written simply as M_A. If $A = \{a\}$ and $B = \{b\}$, then the brackets will be dropped, and we will use the notation $M_{a,b}$. Finally, for the complementary sets of A and B, we will consider the notation $M_{-A,-B} = M_{\overline{A},\overline{B}}$. When working with vectors, the same notation will be used for subsets of components related to subsets of indices. In the following, the main notions for the development of the paper concerning Gaussian Markov Random Fields, totally positive matrices, and their relation are presented.

2.1 | Gaussian Markov Random Fields

A random vector $\vec{X}$ (of dimension $n \in \mathbb{N}$) has multivariate Gaussian distribution if its density function $f : \mathbb{R}^n \to \mathbb{R}$ is of the form

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

where $\vec{\mu}$ is its mean vector and Σ is its covariance matrix [9]. The fact that a random vector $\vec{X}$ has a multivariate Gaussian distribution with mean vector $\vec{\mu}$ and covariance matrix Σ is denoted by $\vec{X} \sim N(\vec{\mu}, \Sigma)$. The mean vector $\vec{\mu} \in \mathbb{R}^n$ is a vector in which each component equals the expectation of the same component of the random vector, that is, $\mu_i = E[X_i]$ for any $i \in \{1,\dots,n\}$. The covariance matrix $\Sigma = \left(\Sigma_{i,j}\right)_{1\le i,j\le n}$ presents the covariance between each pair of variables, that is, $\Sigma_{i,j} = Cov(X_i, X_j) = E[X_iX_j] - E[X_i]E[X_j]$ for any $i, j \in \{1,\dots,n\}$. Any covariance matrix is positive semi-definite [9], although throughout this work we will restrict our attention to covariance matrices that are positive definite (if a covariance matrix is not positive definite, there exist some variables that can be expressed as a linear combination of all other variables, thus can be omitted).

For multivariate Gaussian random vectors, the conditional covariance matrix can be computed with a simple formula [9]. Let I denote the set of indices associated with the known variables, the conditional covariance matrix given the variables in I is:

$$\Sigma_{-I\ |\ I} = \Sigma_{-I} - \Sigma_{-I,I}\left(\Sigma_I\right)^{-1}\Sigma_{I,-I}. \tag{1}$$

A Gaussian Markov Random Field (GMRF) is a random vector with a multivariate Gaussian distribution that fulfills some restrictions related to the conditional independence of its components, which can be formulated by means of a

graph. Therefore, we will present some necessary notions related to simple undirected graphs (from now on, simply graphs). A graph $G = (V, E)$ consists of a (finite) set of nodes V and a set of edges E consisting of subsets of V with cardinality 2. Two nodes $i, j \in V$ are called adjacent if $\{i, j\} \in E$. If there exists a sequence $i, v_1, \dots, v_k, j$ such that $\{i, v_1\}, \{v_1, v_2\}, \dots, \{v_{k-1}, v_k\}, \{v_k, j\} \in E$, then i and j are called connected. Such a sequence is called a walk and, if no nodes are repeated, it is called a path. If every two nodes are connected, then the graph is called connected; otherwise, noticing that being connected is an equivalence relation on the set of nodes, one can consider the associated equivalence classes, referred to as the connected components of the graph. Given three pairwise disjoint subsets of nodes $A, B, C \subset V$, it is said that A and B are being separated by C if every path between a node belonging to A and a node belonging to B contains a node belonging to C. A connected graph is called a tree if the number of edges equals the number of nodes minus one, that is, $|E| = |V| - 1$. A tree in which each node is adjacent to at most two other nodes is called a path graph. The following result will be of interest later.

Theorem 2.1 ([10]). *A graph is a tree if and only if there exists a unique path between each pair of nodes.*

Given a graph $G = (V, E)$ with $V = \{1, \dots, n\}$, a GMRF over G is a multivariate Gaussian random vector $\vec{X} = (X_1, \dots, X_n)$ that satisfies the Markov property, that is, for any $i_1, i_2 \in V$ and any $I \subset V$ the fact that I separates i_1 and i_2 implies that X_{i_1} and X_{i_2} are conditionally independent given the value of the components of $\vec{X}$ associated with I [11]. An easier characterization of a GMRF can be provided in terms of the inverse of the covariance matrix.

Theorem 2.2 ([11]). *Let $\vec{X} \sim N(\vec{\mu}, \Sigma)$ be a multivariate Gaussian random vector and let $G = (V, E)$ with $V = \{1, \dots, n\}$ be a graph. $\vec{X}$ is a GMRF over G if and only if $(\Sigma^{-1})_{i,j} = 0$ for every $i, j \in V$ such that $\{i, j\} \notin E$.*

The characterization above implies that any multivariate Gaussian random vector is a GMRF over the graph in which every two nodes are adjacent. If the graph has no edges, then all components of the random vector should be independent. We end this section with a result related to GMRF over tree graphs.

Theorem 2.3 ([12]). *Let $\vec{X}$ be a GMRF over a tree graph. Then, Pearson's correlation coefficient between any two variables is the product of Pearson's coefficients of the adjacent variables in the unique path that connects them.*

2.2 | Totally Positive Matrices

A minor of a matrix is the determinant associated with a square submatrix. More formally, given any two subsets of indices I_1 and I_2 with the same cardinality, the associated minor is $|M_{I_1,I_2}|$. The notion of total positivity is directly related to the sign of these minors.

Definition 2.1 ([6]). A matrix M is totally positive if all its minors are nonnegative.

Totally positive matrices have good computational properties, mainly related to the possibility of performing a bidiagonal decomposition of the matrix. In the next result, we just consider the case of symmetric matrices.

Theorem 2.4 ([13]). *Any symmetric nonsingular totally positive matrix T admits a bidiagonal decomposition of the form*:

$$T = F_{n-1} \dots F_1 D F_1^T \dots F_{n-1}^T ,$$

where D is a diagonal matrix with positive diagonal entries and each F_i has the following structure:

$$F_i = \begin{pmatrix} 1 & & & & & & \\ 0 & 1 & & & & & \\ & \ddots & \ddots & & & & \\ & & 0 & 1 & & & \\ & & & (F_i)_{i+1,i} & 1 & & \\ & & & & \ddots & \ddots & \\ & & & & & (F_i)_{n,n-1} & 1 \end{pmatrix}.$$

The elements of D and $F_1, \dots, F_n$ are all nonnegative and are related to the Neville elimination method (see [14]). Starting with the bidiagonal decomposition of a totally positive matrix, there exist algorithms that allow for computing the eigenvalues, the inverse matrix, and other algebraic computations with HRA. This means that such computations can be done with a relative error of the order of the machine error, regardless of the dimension or the condition number of the matrix. Note that similar approaches have been studied for other types of matrices, such as almost strictly sign regular matrices and Kac-Murdock-Szegö matrices and their generalizations (see, e.g., [15] and [16]).

HRA is achieved by avoiding the subtraction of noninitial data with the same sign. This is because subtracting two nearby numbers can yield a result with significantly smaller magnitude, thereby increasing the relative error. For example, consider the numbers $x_1 = 47.132$ and $x_2 = 47.111$, which have been rounded to the third decimal place. The corresponding absolute error is at most 0.0005, and the relative error is approximately $\frac{0.0005}{47.132} \approx \frac{0.0005}{47.111} \approx 0.00001$. However, their difference is $47.132 - 47.111 = 0.021$, for which the absolute error is at most 0.001, and the relative error becomes $\frac{0.001}{0.021} \approx 0.05$, which is significantly larger. Notably, if the quantities involved are initial data, their values are assumed to be exact, and therefore do not pose a problem in terms of HRA.

2.3 | Gaussian Markov Random Fields and Totally Positive Matrices

In [17], some sufficient conditions for a covariance matrix and its inverse to be totally positive are provided. In particular, these conditions were provided for a GMRF over a graph of paths with natural ordering. This particular type of graph is defined as a graph $G = (V, E)$ with $V = \{1, \dots, n\}$ and such that $\{i, j\} \in E$ implies $|j - i| = 1$.

Theorem 2.5 ([17]). *Let $\vec{X}$ be a GMRF over a graph of paths with natural ordering. It holds that*:

- *The covariance matrix of the GMRF is totally positive if and only if all Pearson's correlation coefficients between adjacent variables are nonnegative.*
- *The inverse of the covariance matrix of the GMRF is totally positive if and only if all Pearson's correlation coefficients between adjacent variables are nonpositive.*

Focusing on the nonnegative case, the bidiagonal decomposition of the covariance matrix can be computed with HRA (see Section 5 in [7]), thus some algebraic operations concerning a totally positive covariance matrix can be computed with HRA.

Theorem 2.6 ([7]). *Let $\vec{X}$ be a GMRF over a graph of paths with natural ordering. If the covariance matrix Σ is nonsingular, Pearson's correlation coefficients between adjacent variables have the same sign and $\Sigma_{i,i}$ for all i are given, the following computations can be done with HRA*:

- *The eigenvalues of Σ,*
- *The inverse of Σ,*
- *The solution of the linear system of equations ($\Sigma\vec{x} = \vec{b}$), with the vector $\vec{b}$ with alternate sign (if the sign of the Pearson's correlation coefficients is nonnegative) or with the vector $\vec{b}$ with constant sign (if the sign of the Pearson's correlation coefficients is nonpositive).*

At this point, it must be noted that, when working with data, one does not know the theoretical covariance matrix associated with the data; thus, the covariance matrix should be estimated. In this regard, it is possible to estimate the covariance matrix of a GMRF over a graph of paths with natural ordering, with nonnegative correlations between adjacent variables assuring that the estimated matrix is still totally positive. The following algorithm is a modification of Algorithm 1 in [8] that only considers the nonnegative case.

The method of maximum likelihood estimation is one of the most prominent methods for estimation of parameters of probability distributions, see [18]. Unfortunately, this estimation method is typically very sensitive to the fulfillment of some necessary assumptions. In this case, these assumptions relate to multivariate normality, conditional independence, and nonnegative correlation between all variables. It is encouraged to previously test for those assumed conditions before accepting the estimate of the algorithm above, since otherwise the estimated covariance matrix might differ greatly from the theoretical covariance matrix. For more information concerning how to test for the fulfillment of the necessary assumptions, the reader is referred to Section 2.5 in [8].

ALGORITHM 1 | Algorithm for computing the maximum likelihood estimator of the covariance matrix of a GMRF over a path graph with nonnegative correlation between adjacent variables.

Input: A positive definite sample covariance matrix S of dimension p.
Output: The MLE of the covariance matrix of a GMRF over a path graph.

1: Construct a complete weighted graph with $V = \{u_1, \dots, u_p\}$ and weights $w(\{u_i, u_j\}) = \log\left(1 - \max\left\{0, \frac{S_{i,j}}{\sqrt{S_{i,i}S_{j,j}}}\right\}^2\right)$.
2: $v_1, \dots, v_p \leftarrow$ The shortest Hamiltonian path of the latter graph.
3: Construct $\hat{\Sigma}$ as $\hat{\Sigma}_{v_i,v_j} = \sqrt{S_{v_i,v_i}S_{v_j,v_j}} \prod_{\ell=\min(v_i,v_j)}^{\max(v_i,v_j)-1} \frac{\max\left(0, S_{v_\ell, v_{\ell+1}}\right)}{\sqrt{S_{v_\ell,v_\ell}S_{v_{\ell+1},v_{\ell+1}}}}$ with $i, j \in \{1, \dots, p\}$

return $\hat{\Sigma}$

The input matrix S is considered to be the sample covariance matrix of the data, defined as:

$$S_{i,j} = \frac{1}{n}\sum_{k=1}^{n}\left(x_{i,k} - \frac{1}{n}\sum_{l=1}^{n}x_{i,l}\right)\left(x_{j,k} - \frac{1}{n}\sum_{l=1}^{n}x_{j,l}\right), \quad i, j \in \{1, \dots, p\},$$

where $x_{i,k}$ denotes the kth observation of the ith variable.

In addition, we recall that the shortest Hamiltonian path is the path containing all the nodes of the graph such that the sum of the weights of the associated edges is minimum. The complexity of Algorithm 1 is $O(p^2 2^p)$, as discussed in Section 3.4 in [8]. In the following, a minimal example of the application of Algorithm 1 is provided; the reader is referred to [8] for more elaborated ones.

Example 2.1. Let S be the sample covariance matrix given in the left-hand side of Equation (2). The weighted graph constructed in the first step of Algorithm 1 has associated weights $w_{1,2} \approx -0.259$, $w_{1,3} \approx -0.0740$ and $w_{2,3} \approx -0.314$. For computing the shortest Hamiltonian path, in this simple case, it is necessary just to identify the edges with a smaller weight, which are $\{1, 2\}$ and $\{2, 3\}$. Therefore, the estimated covariance matrix is the one provided in the right-hand side of Equation (2).

$$S = \begin{pmatrix} 0.917 & 0.446 & 0.532 \\ 0.446 & 0.950 & 1.051 \\ 0.532 & 1.051 & 4.312 \end{pmatrix}, \quad \hat{\Sigma} = \begin{pmatrix} 0.917 & 0.446 & 0.496 \\ 0.446 & 0.950 & 1.051 \\ 0.496 & 1.051 & 4.312 \end{pmatrix}. \tag{2}$$

In contrast to S, $\hat{\Sigma}$ is a totally positive matrix.

3 | Order Statistics and Total Positivity

The computational benefits that will be exposed in the following sections are based on the good properties of totally positive matrices. In Theorem 2.5, sufficient conditions for the total positivity of the covariance matrix of a GMRF are given. Unfortunately, these sufficient conditions are quite restrictive, since one needs a very particular type of distribution: a GMRF over a graph of paths with natural ordering. Interestingly, these sufficient conditions appear naturally for the order statistics when considering identically distributed and independent random variables. In particular, we need to study when the following statements are fulfilled:

1. The order statistics are MRF over a graph of paths (with natural ordering),
2. The correlation between consecutive order statistics is nonnegative,
3. The order statistics follow a multivariate Gaussian distribution.

3.1 | The Markov Property of the Order Statistics

The Markov property is defined in terms of conditional distributions. Fortunately, conditional distributions of order statistics of continuous distributions are easy to compute.

Theorem 3.1 ([5]). *Let $X_1, \dots, X_n$ be independent random variables with the same continuous distribution. The conditional distribution of $(X_{(r+1)}, \dots, X_{(s-1)})$ given $(X_{(1)}, \dots, X_{(r)}, X_{(s)}, \dots, X_{(n)}) = (x_{(1)}, \dots, x_{(r)}, x_{(s)}, \dots, x_{(n)})$ is*

$$f(x_{(r+1)}, \dots, x_{(s-1)} | x_{(1)}, \dots, x_{(r)}, x_{(s)}, \dots, x_{(n)}) = (s-r-1)! \prod_{j=r+1}^{s-1} \frac{f(x_{(j)})}{F(x_{(s)}) - F(x_{(r)})},$$

for any $(x_{(r+1)}, \dots, x_{(s-1)}) \in \mathbb{R}^{s-r-1}$ such that $x_{(r+1)} \leq \dots \leq x_{(s-1)}$.

Notice that the expression of the conditional density depends on the value that $X_{(s)}$ and $X_{(r)}$ take, but not on the value of $X_{(i)}$ when $i < r$ or $i > s$. Therefore, the order statistics of continuous distributions are Markov chains, or equivalently, MRF over the graph $G = (V, E)$ with $V = \{1, \dots, n\}$ such that $\{i, j\} \in E$ if and only $|j - i| = 1$. Trivially, G is a graph of paths with natural ordering. The reader is referred to Section 2.5 in [5] for more information in this regard.

3.2 | Nonnegative Correlation of Consecutive Order Statistics

Consider two consecutive order statistics $X_{(i)}$ and $X_{(i+1)}$ with $i \in \{1, \dots, n-1\}$. The inequality $X_{(i)} \leq X_{(i+1)}$ implies that, if the value of $X_{(i)}$ increases, then the minimum value that $X_{(i+1)}$ can take also increases. This, intuitively, leads to a positive correlation between the order statistics.

Recall that a random vector is said to be exchangeable if its distribution remains invariant when changing the order of the variables [19]. In addition, a continuous distribution is said to be MTP2 (Multivariate Totally Positive of order 2) if its density function satisfies that $f(\vec{x})f(\vec{y}) \leq f(\vec{x} \wedge \vec{y})f(\vec{x} \vee \vec{y})$ [20], where $\wedge$ and $\vee$ denote, respectively, the component-wise minimum and maximum. The following is a slight modification of Theorem 5.2 in [21].

Theorem 3.2. *Let $(X_1, \dots, X_n)$ be an MTP2 exchangeable random vector. The correlation coefficient between any two order statistics is always positive.*

Notice that, as previously stated, we are considering order statistics associated with independent and identically distributed random variables. In such case, exchangeability is trivially fulfilled. For being MTP2, notice that the joint density function $f(\vec{x})$ can be decomposed as $f(\vec{x}) = \prod_{i=1}^{n} g(x_i)$, being g the common marginal density, if the components are independent and identically distributed. Therefore,

$$f(\vec{x})f(\vec{y}) = \prod_{i=1}^{n} g(x_i)g(y_i) = \prod_{i=1}^{n} g(x_i \wedge y_i)g(x_i \vee y_i) = f(\vec{x} \wedge \vec{y})f(\vec{x} \vee \vec{y})$$

and the distribution is MTP2. Keeping in mind the latter comments, the following result is immediate.

Corollary 3.1. *Let $(X_1, \dots, X_n)$ be a random vector with independent and identically distributed components. The correlation coefficient between any two order statistics is always positive.*

In particular, the result above holds for consecutive order statistics.

3.3 | Asymptotic Multivariate Normality of Central Order Statistics

Unfortunately, the order statistics do not usually have a multivariate Gaussian distribution. However, several studies have been carried out to prove the asymptotic normality of order statistics. In particular, several authors have presented results studying a finite or infinite number of order statistics and changing the assumptions needed to reach the asymptotic normality. One of the simplest results in this regard is Theorem B, Section 2.3.3 in [22]. Recall that, given a distribution function F, the associated quantile function F^{-1} is defined as $F^{-1}(p) = \inf\{x \in \mathbb{R} | F(x) \leq p\}$.

Theorem 3.3 ([22]). *Let $k \in \mathbb{N}$ and $p_1, \dots, p_k \in (0, 1)$ with $p_1 < p_2 < \dots < p_k$. Furthermore, let F be a cumulative distribution function with existing density f in the neighborhoods of $F^{-1}(p_1), \dots, F^{-1}(p_k)$, satisfying $f(F^{-1}(p_1)), \dots, f(F^{-1}(p_k)) > 0$. Let $i_1(n), \dots, i_k(n)$ be increasing sequences of natural numbers such that*

$$\lim_{n \to \infty} \frac{i_j(n)}{n} = p_j,$$

for any $j \in \{1, \dots, k\}$. *For independent and identically distributed random variables* $X_1, \dots, X_n$ *with distribution function* F, *it holds that*

$$\sqrt{n}\left(\begin{pmatrix} X_{(i_1(n))} \\ \dots \\ X_{(i_k(n))} \end{pmatrix} - \begin{pmatrix} F^{-1}(p_1) \\ \dots \\ F^{-1}(p_k) \end{pmatrix}\right) \to_d N(\vec{0}, \Sigma),$$

where

$$\Sigma_{i,j} = \frac{\min(p_i, p_j)(1 - \max(p_i, p_j))}{f(F^{-1}(p_i))f(F^{-1}(p_j))},$$

for any $i, j \in \{1, \dots, k\}$.

The latter result assures the convergence of a collection of k order statistics to a multivariate Gaussian distribution when the quotients between the associated indices and the sample size converge. Notice that $p_1 > 0$ and $p_k < 1$, thus the convergence is not true for the minimum and the maximum. For instance, it is well-known that the maximum of n standard Gaussian random variables diverges to infinity.

Apart from the existence of the density function, the assumptions for the distribution are not strong. In particular, any density function f such that the set $\{t \in \mathbb{R} \mid f(t) > 0\}$ is connected fulfills the conditions of Theorem 3.3. Typically, continuous random variables have density functions that are strictly positive over an interval, the real line, or the positive real line (see [23, 24]), thus verifying the conditions of Theorem 3.3.

4 | Asymptotic Total Positivity of the Covariance Matrix of Order Statistics

In the previous section, we have shown that the order statistics, under some assumptions, behave as a Markov chain, their correlations are all positive, and converge to a multivariate Gaussian distribution. Therefore, it is intuitive to expect a convergence of the covariance matrix of the order statistics to a totally positive matrix. In [25], a heuristic formula for the inverse of the covariance matrix of the order statistics is given. In particular, for large n, Σ^{-1} can be approximated by $(n+1)(n+2)DQD$, where D is a diagonal matrix satisfying that $D_{i,i} = f\left(F^{-1}\left(\frac{i}{n+1}\right)\right)$ for any $i \in \{1, \dots, n\}$ and Q is the following tridiagonal matrix:

$$Q = \begin{pmatrix} 2 & -1 & 0 & 0 & 0 & \dots \\ -1 & 2 & -1 & 0 & 0 & \dots \\ 0 & -1 & 2 & -1 & 0 & \dots \\ 0 & 0 & -1 & 2 & -1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Since Q is a tridiagonal M-matrix, its inverse is totally positive [26]. By using the fact that D is a nonnegative diagonal matrix, it follows that the inverse of $(n+1)(n+2)DQD$ is totally positive. However, the formula $\Sigma^{-1} \sim (n+1)(n+2)DQD$ is just a heuristic approximation, not a formal convergence result. In the following, we will prove the total positivity of the asymptotic covariance matrix Σ of Theorem 3.3 by using probabilistic arguments. For that purpose, we first need to present an auxiliary result.

Lemma 4.1. *A multivariate Gaussian random vector* $\vec{X}_V$ *is a GMRF over a tree graph* $G = (V, E)$ *if and only if Pearson's correlation coefficient between any two variables is the product of Pearson's correlation coefficients of the adjacent variables in the unique path that connects them.*

Proof. The fact that in a GMRF over a tree graph the correlation between two variables is the product of Pearson's correlation coefficients of the adjacent variables in the unique path that connects them is stated in Theorem 2.3. For the converse implication, notice that, since the random vector is assumed to be Gaussian, we just need to check the Markov property. Suppose that the nodes $i, j \in V$ are not adjacent. Since G is a tree graph, there exists $k \in V$ such that $\{k\}$ separates $\{i\}$ and $\{j\}$. Moreover, the subgraph induced by $V \setminus \{k\}$ is a forest consisting of two connected components C_i and C_j such that $i \in C_i$ and $j \in C_j$ and k is contained in the unique path that connects any node in C_i with any node

in C_j. Without loss of generality, we rearrange the elements of $\vec{X}_V$ in order to index first the variables associated with C_i, next X_k, and last the variables associated with C_j.

Let R, R_i and R_j be the correlation matrices of $\vec{X}_V$, $\vec{X}_{C_i}$ and $\vec{X}_{C_j}$, respectively. By using that Pearson's correlation coefficient between two variables is the product of Pearson's correlation coefficients of the adjacent variables in the unique path that connects them, one has that

$$R = \begin{pmatrix} R_i & \vec{\rho}_i & \vec{\rho}_i\vec{\rho}_j^{\,t} \\ \vec{\rho}_i^{\,t} & 1 & \vec{\rho}_j^{\,t} \\ \vec{\rho}_j\vec{\rho}_i^{\,t} & \vec{\rho}_j & R_j \end{pmatrix}, \quad \Sigma = D^\sigma \begin{pmatrix} R_i & \vec{\rho}_i & \vec{\rho}_i\vec{\rho}_j^{\,t} \\ \vec{\rho}_i^{\,t} & 1 & \vec{\rho}_j^{\,t} \\ \vec{\rho}_j\vec{\rho}_i^{\,t} & \vec{\rho}_j & R_j \end{pmatrix} D^\sigma,$$

where D^σ is the diagonal matrix containing the standard deviation of the components of $\vec{X}_V$, $\vec{\rho}_i$ is the vector containing Pearson's correlation coefficients between the variables in C_i and X_k and $\vec{\rho}_j$ is the vector containing Pearson's correlation coefficients between the variables in C_j and X_k. Therefore, the conditional covariance matrix given a value for X_k can be computed using Equation (1):

$$(D^\sigma)_{-k}\begin{pmatrix} R_i & \vec{\rho}_i\vec{\rho}_j^{\,t} \\ \vec{\rho}_j\vec{\rho}_i^{\,t} & R_j \end{pmatrix}(D^\sigma)_{-k} - (D^\sigma)_{-k}\begin{pmatrix} \vec{\rho}_i \\ \vec{\rho}_j \end{pmatrix}\sigma_k\sigma_k^{-2}\sigma_k\begin{pmatrix} \vec{\rho}_i^T & \vec{\rho}_j^T \end{pmatrix}(D^\sigma)_{-k} = (D^\sigma)_{-k}\begin{pmatrix} R_i - \vec{\rho}_i\vec{\rho}_i^{\,t} & M_0 \\ M_0^T & R_j - \vec{\rho}_j\vec{\rho}_j^{\,t} \end{pmatrix}(D^\sigma)_{-k},$$

where M_0 is a null matrix of appropriate dimension. Since for multivariate Gaussian distributions linear independence is equivalent to independence, $\vec{X}_{C_i}$ and $\vec{X}_{C_j}$ are conditionally independent given X_k. Notice that the resulting covariance matrix is block diagonal, with two blocks associated with C_i and C_j. Therefore, its inverse is also a block diagonal, with the two blocks still associated with C_i and C_j. From Theorem 2.2 we conclude that X_i and X_j are conditionally independent given $\vec{X}_{V\setminus\{i,j,k\}}$, so X_i and X_j are conditionally independent given $\vec{X}_{V\setminus\{i,j\}}$. It is concluded that $\vec{X}$ is a GMRF over $G = (V, E)$. □

Therefore, it suffices to check if the asymptotic covariance matrix coincides with the one of a GMRF over a graph of paths with natural ordering and that the correlation between adjacent variables is nonnegative.

Proposition 4.1. *Let f, F^{-1} and $0 < p_1 < \cdots < p_k < 1$ be as in Theorem 3.3. The matrix Σ defined as*

$$\Sigma_{i,j} = \frac{\min(p_i, p_j)(1 - \max(p_i, p_j))}{f(F^{-1}(p_i))f(F^{-1}(p_j))},$$

for any $i, j \in \{1, \ldots, k\}$ is totally positive.

Proof. The correlation matrix R (of dimension k) associated with Σ is:

$$R_{i,j} = \frac{\Sigma_{i,j}}{\sqrt{\Sigma_{i,i}\Sigma_{j,j}}} = \frac{p_i(1-p_j)}{f(F^{-1}(p_i))f(F^{-1}(p_j))}\frac{f(F^{-1}(p_i))f(F^{-1}(p_j))}{\sqrt{p_i(1-p_i)p_j(1-p_j)}} = \frac{\sqrt{p_i(1-p_j)}}{\sqrt{p_j(1-p_i)}}, \quad i < j. \tag{3}$$

Next, consider the graph of paths with natural ordering $G = (V, E)$ with $V = \{1, \ldots, k\}$ and $\{i, j\} \in E$ if and only if $|j - i| = 1$. Given $i, j \in V$, with $i < j$, we denote by $i, v_1, \ldots, v_\ell, j$ the unique path that connects i and j. For R to be the correlation matrix of a GMRF over G, it needs to follow that Pearson's correlation coefficient between two variables is the product of Pearson's correlation coefficients of the adjacent variables in the unique path that connects them. Indeed, for any $i < j$, it follows that:

$$R_{i,j} = \frac{\sqrt{p_i(1-p_{v_1})}}{\sqrt{p_{v_1}(1-p_i)}}\left(\prod_{s=1}^{\ell-1}\frac{\sqrt{p_{v_s}(1-p_{v_{s+1}})}}{\sqrt{p_{v_{s+1}}(1-p_s)}}\right)\frac{\sqrt{p_{v_\ell}(1-p_j)}}{\sqrt{p_j(1-p_{v_\ell})}} = \frac{\sqrt{p_i(1-p_j)}}{\sqrt{p_j(1-p_i)}}.$$

We conclude from Lemma 4.1 that C is the correlation matrix of a GMRF over a graph of paths with natural ordering. In addition, notice that all the correlations/covariances are positive, thus it follows as a result of Theorem 2.5 that Σ is totally positive. □

Remark 4.1. The latter result can also be proven with results from linear algebra by noticing that the matrix Σ is a Green matrix with parameters p_i and $1 - p_i$ with $i \in \{1, \ldots, k\}$ multiplied by two diagonal matrices that multiply rows

and columns by positive quantities. As a result of Theorem 4.2 in [27], the associated Green matrix is totally positive if and only if the sequence $0 < \frac{p_1}{1-p_1} \leq \ldots \leq \frac{p_k}{1-p_k} < 1$ is increasing, which is the case since it holds that $0 < p_1 \leq \ldots \leq p_k < 1$.

As a direct consequence of Theorem 3.3 and Proposition 4.1, the asymptotic covariance matrix of order statistics is a totally positive matrix.

Corollary 4.1. *Let $k \in \mathbb{N}$ and $p_1, \ldots, p_k \in (0,1)$ with $p_1 < p_2 < \cdots < p_k$. Furthermore, let F be a cumulative distribution function with existing density f in the neighborhoods of $F^{-1}(p_1), \ldots, F^{-1}(p_k)$, satisfying $f(F^{-1}(p_1)), \ldots, f(F^{-1}(p_k)) > 0$. Let $i_1(n), \ldots, i_k(n)$ be increasing sequences of natural numbers such that*

$$\lim_{n\to\infty} \frac{i_j(n)}{n} = p_j,$$

for any $j \in \{1, \ldots, k\}$. Then, for independent and identically distributed random variables $X_1, \ldots, X_n$ with distribution function F, the sequence

$$\mathrm{Cov}\left(\sqrt{n}\begin{pmatrix} X_{(i_1(n))} \\ \ldots \\ X_{(i_k(n))} \end{pmatrix}\right)$$

converges to a totally positive matrix.

Remark 4.2. The determinant is a continuous function (see, for instance, Example 6.2.7 in [28]), thus the convergence of the covariance matrix to a totally positive matrix proves that all minors of the associated matrix converge to a nonnegative value.

We end this section by noticing that the formula of the asymptotic covariance matrix of the order statistics is exact for any sample size in the case of the uniform distribution [5]. Therefore, the covariance matrix of the order statistics of the uniform distribution is always totally positive.

Corollary 4.2. *Let $X_1, \ldots, X_n$ be independent random variables with uniform distribution. The covariance matrix of $X_{(1)}, \ldots, X_{(n)}$ is totally positive.*

5 | High Relative Accuracy in Computations With Covariance Matrices of Order Statistics

This section is devoted to discussing the importance of HRA computations when dealing with the covariance matrix of the order statistics. First, we will show that the condition number increases with the sample size, and next, we provide different examples where HRA computations would be useful. Note that within this section, we directly work with the asymptotic covariance matrix, even though we are dealing with finite sample sizes. Therefore, the considered sample size should be large enough (or the underlying distribution should be the uniform distribution, as discussed in Corollary 4.2) for the results to be applicable.

5.1 | Condition Number of the Covariance Matrix of the Order Statistics

The condition number of the covariance matrix of a GMRF over a path graph increases depending on three factors: the dimension, the difference between variances and the strength of the correlation between adjacent variables. In [7], it has been shown that the effect that the dimension has on the condition number of the covariance matrix, when the correlation is fixed, is moderate. On the contrary, even for a small dimension, correlation coefficients close to 1 in absolute value increase the condition number at a high rate (see [7]).

In the case of the order statistics, the correlation matrix associated with the asymptotic covariance matrix was computed in Equation (3). When considering all order statistics, that is, all $p_i = \frac{i}{n+1}$, the correlation between consecutive variables is

$$R_{i,i+1} = \sqrt{\frac{\frac{i}{n+1}\left(1-\frac{i+1}{n+1}\right)}{\frac{i+1}{n+1}\left(1-\frac{i}{n+1}\right)}} = \sqrt{\frac{p_i\left(1-p_i-\frac{1}{n+1}\right)}{\left(p_i+\frac{1}{n+1}\right)(1-p_i)}},$$

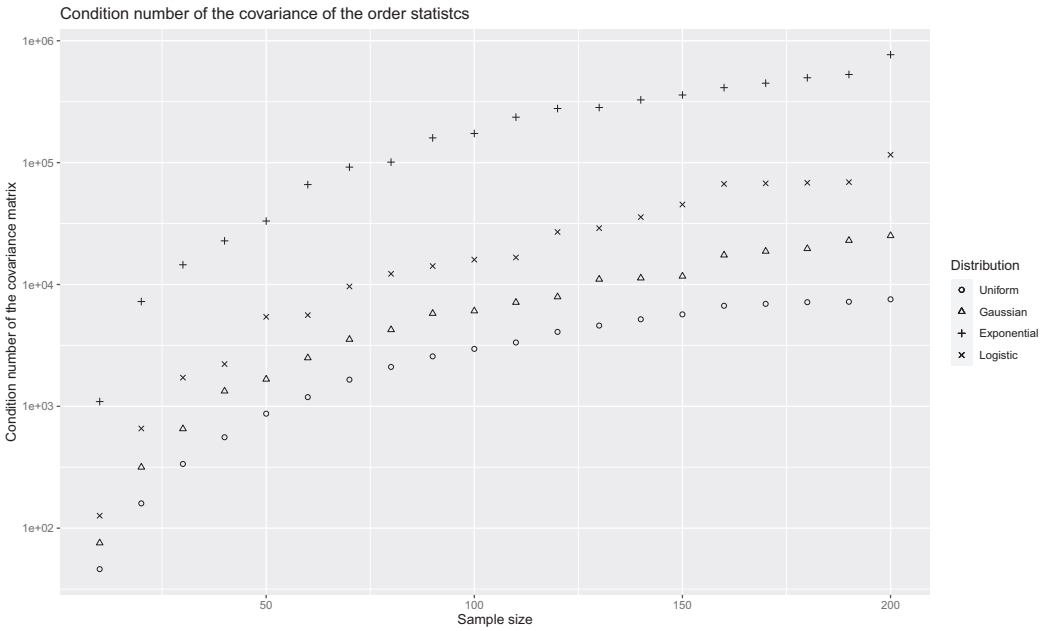

FIGURE 1 | Condition number of the asymptotic covariance matrix of the order statistics for different distributions and sample sizes. The condition number has been averaged over 10^5 independent samples for each distribution and sample size.

for any $i \in \{1, \dots, n-1\}$. Therefore, for a fixed p_i, the correlation between consecutive order statistics increases as the dimension increases. Notice that, if one considers a set of k consecutive order statistics, for instance the k central ones, all correlations tend to 1 as the dimension tends to infinity, so the associated submatrix of the correlation matrix tends to a matrix of dimension $k \times k$ with all values equal to 1, which has rank equal to 1 regardless of the value of k.

The diagonal values of the asymptotic covariance matrix, which coincide with the asymptotic variances of the associated variables, are $\Sigma_{i,i} = \frac{p_i(1-p_i)}{f(F^{-1}(p_i))^2}$. The term $p_i(1-p_i)$ results in a greater variance for the central order statistics, while quantiles associated with regions with smaller density also have more variability. Notice that $f(F^{-1}(p_i)) > 0$ should be fulfilled, as an assumption of Theorem 3.3.

In order to illustrate the behavior of the condition number of the asymptotic covariance matrix of the order statistics as the sample size increases, different covariance matrices of the order statistics for the uniform, exponential, normal and logistic distributions (see [23, 24]) have been simulated for different samples sizes ranging from 10 to 200. The obtained results can be found in Figure 1, showing that the exponential distribution leads to the largest condition numbers and the uniform distribution leads to the smallest ones for all sample sizes. This could be caused by the fact that the density of the exponential is not symmetric and decays for bigger values, while the uniform distribution has a constant density. In any case, even for a small sample size of 30, the condition number is of an order of magnitude between 10^3 and 10^4, increasing rapidly for larger sample sizes.

5.2 | Mean Estimation for Symmetric Distributions Using L-Statistics

Given a random variable X with mean μ, it is quite common in Statistics to use linear combinations of order statistics, known as L-statistics, in order to estimate μ. This approach started at the beginning of the last century with the first foundational papers [29–32], and it is still widely used, see [33–35] for some recent applications of such a procedure. In particular, the considered estimator is defined by

$$\sum_{i=1}^{n} \lambda_i X_{(i)},$$

where $\vec{\lambda}$ is a vector that should be determined. A common choice for $\vec{\lambda}$ is based on the minimization of the following quantity,

$$E\left[\left(\mu - \sum_{i=1}^{n} \lambda_i X_{(i)}\right)^2\right] = E\left[\mu - \sum_{i=1}^{n} \lambda_i X_{(i)}\right]^2 + Var\left(\sum_{i=1}^{n} \lambda_i X_{(i)}\right).$$

If the distribution of X is symmetric, $E[X_{(i)}] + E[X_{(n-i+1)}] = \mu$ for any $i \in \{1, \dots, n\}$ and the covariance matrix Σ of the order statistics is persymmetric ($\Sigma_{i,j} = \Sigma_{n-i+1,n-j+1}$ for any $i, j \in \{1, \dots, n\}$) [36]. Under such conditions, it is possible to prove that the best linear unbiased estimator minimizes the following equation [37]:

$$\operatorname{Var}\left(\sum_{i=1}^{n} \lambda_i X_{(i)}\right) = \vec{\lambda}^T \Sigma \vec{\lambda} \,,$$

under the restriction $\sum_{i=1}^{n} \lambda_i = 1$. The solution of such a problem is [37]:

$$\vec{\lambda} = \frac{\Sigma^{-1}\vec{1}}{\vec{1}^{\,t}\Sigma^{-1}\vec{1}}, \tag{4}$$

where $\vec{1}$ is a vector of ones.

The computation of the optimal vector $\vec{\lambda}$, since it involves the inversion of Σ, is quite sensitive to big condition numbers. Considering the asymptotic matrix, its inverse can be computed with HRA, see Theorem 2.6. Unfortunately, the operation $\Sigma^{-1}\vec{1}$ cannot be performed with HRA, since Σ^{-1} is a tridiagonal M-matrix with positive and negative elements, so sums of numbers of different signs appear.

However, it is possible to design an algorithm that allows us to compute the numerator and denominator in Equation (4) such that, in both cases, only one subtraction in the final step is needed, being able to perform all the previous operations with HRA. More precisely, for computing Σ^{-1} with HRA, it is possible to first sum all the positive elements in each row and, analogously, all the negative elements in each row. These operations can also be performed with HRA; thus, it only remains to compute a final subtraction in order to achieve each component of $\Sigma^{-1}\vec{1}$. Similarly, the quantity $\vec{1}^T\Sigma^{-1}\vec{1}$ can be computed with all steps but the last one with HRA. If necessary, these last subtractions can be performed with quadruple precision in order to improve the result. Finally, the quotient can be computed as well with HRA. Therefore, the components of the optimal vector $\vec{\lambda}$ can be computed with just two problematic operations in terms of HRA, one subtraction associated with the numerator and another one with the denominator.

5.3 | Computation of Conditional Covariances of Order Statistics

Suppose that the values of some order statistics are known, and we want to work with the remaining order statistics. In the limit, as a result of the asymptotic multivariate normality, it is possible to use Equation (1). Unfortunately, the subtraction of the latter term does not allow for computing such a matrix with HRA. However, as a consequence of Theorem 2.5 in [11], such computation can also be expressed as

$$\Sigma_{-I\,|\,I} = \left(\left(\Sigma^{-1}\right)_{-I}\right)^{-1} .$$

As already commented before, Σ^{-1} can be computed with HRA, as well as $\left(\Sigma^{-1}\right)_{-I}$. Unfortunately, the matrix $\left(\Sigma^{-1}\right)_{-I}$ is not totally positive, thus in principle its inverse cannot be computed with HRA. However, if we denote by A the diagonal matrix with elements $1, -1, 1, -1, \dots$, it follows that:

$$\Sigma_{-I\,|\,I} = AA\Sigma_{-I\,|\,I}AA = AA\left(\left(\Sigma^{-1}\right)_{-I}\right)^{-1}AA = A\left(A\left(\Sigma^{-1}\right)_{-I}A\right)^{-1}A \,.$$

Notice that $A\left(\Sigma^{-1}\right)_{-I}A$ is the inverse of the covariance matrix of a GMRF over a graph of paths with nonpositive correlations. Therefore, it follows from Theorem 2.6 that its inverse, $\left(A\left(\Sigma^{-1}\right)_{-I}A\right)^{-1}$, can be computed with HRA. Trivially, the same holds for $A\left(A\left(\Sigma^{-1}\right)_{-I}A\right)^{-1}A$, thus the conditional covariance matrix $\Sigma_{-I\,|\,I}$ can be computed with HRA.

5.4 | Computation of the Entropy

The entropy of a random vector $\vec{X}$ is a quantity $H(\vec{X})$ related to its expected surprisal, see [38]. In the case of multivariate Gaussian distributions, its formula is simple:

$$H(\vec{X}) = \frac{1}{2}\ln(|\Sigma|) + \frac{n}{2}(1 + \ln(2\pi)) \,,$$

where n is the dimension of the random vector.

In such a formula, the determinant of the covariance matrix is involved, a quantity that can present computation problems when the condition number is too large. The first obstacle to computing the entropy with HRA is that the sign of both terms is unknown. However, by using properties of the logarithm, it follows that

$$H(\vec{X}) = \frac{1}{2}\ln(|\Sigma|) + \ln\left(e^{\frac{n}{2}(1+\ln(2\pi))}\right) = \ln\left(\sqrt{|\Sigma|}e^{\frac{n}{2}(1+\ln(2\pi))}\right).$$

Note that the determinant of the covariance matrices presented in this paper can be computed with HRA by considering the bidiagonal decomposition given in [7] since it coincides with the product of the diagonal of the matrix D (see Theorem 2.4), which can be computed with HRA. The square root can also be computed with HRA, so it only remains to check whether the evaluation of the logarithm can be computed with HRA.

Quite conveniently, the logarithm can be expressed in terms of the arc hyperbolic tangent as $\ln(x) = 2\,arctanh\left(\frac{x-1}{x+1}\right)$ [39]. Therefore, substituting *arctanh* by its Taylor series leads to

$$\ln(x) = 2\sum_{k=0}^{\infty}\frac{1}{2k+1}\left(\frac{x-1}{x+1}\right)^{2k+1}.$$

Although the latter sum involves an infinite number of summands, the logarithm can be approximated by considering a sufficiently large number of terms. For any $x \in (0, \infty)$, the terms of the sum have the same sign, thus this sum can be computed with HRA. It is concluded that the entropy can be computed with HRA.

5.5 | Additional Computational Problems

In Data Analysis, one of the most used procedures for reducing the dimension of random vectors is Principal Component Analysis (PCA), which involves the computation of the eigenvalues of the associated covariance or correlation matrix [40]. In general, this method is especially useful when there are some eigenvalues that are significantly greater than the rest, so the variability of the random vector can be summarized by using a few random variables. When considering independent and identically distributed random variables, this method is not useful, since the covariance matrix is the identity matrix multiplied by a factor, thus all the eigenvalues are the same. However, if the order statistics are considered, the eigenvalues vary from one another. As an example, the quotient of each eigenvalue and the sum of all eigenvalues of the covariance and correlation matrices of the order statistics of dimension 10 for the uniform, Gaussian, exponential, and logistic distributions can be found in Table 1.

Notice that the eigenvalues of the correlation matrices are similar, since the correlation matrices for all distributions converge to the same matrix. However, the eigenvalues of the covariance matrices are more diverse, since the probability density varies from one distribution to another. Since some of the eigenvalues are very small, it makes sense to use PCA to reduce the dimension when working with the order statistics. In this sense, the eigenvalues should be computed, a task that can be performed with HRA for the type of matrices considered in this paper as a consequence of Theorem 2.6. Admittedly, HRA is not especially useful in this case, since the most relevant eigenvalues are the greatest ones, which are the ones that benefit less from HRA procedures (see, e.g., Table 1 in [7]).

TABLE 1 | Quotient of the i-th eigenvalue and the sum of all eigenvalues (c_i) for the covariance and correlation matrices of the order statistics of dimension 10 for different distributions.

Distribution	Matrix	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}
Uniform	Covariance	0.618	0.157	0.073	0.043	0.029	0.022	0.018	0.015	0.014	0.013
	Correlation	0.551	0.183	0.092	0.055	0.037	0.026	0.020	0.015	0.012	0.010
Gaussian	Covariance	0.479	0.218	0.118	0.068	0.041	0.027	0.018	0.013	0.010	0.008
	Correlation	0.535	0.183	0.096	0.059	0.040	0.028	0.021	0.016	0.012	0.010
Exponential	Covariance	0.676	0.191	0.067	0.030	0.015	0.009	0.005	0.003	0.002	0.001
	Correlation	0.532	0.183	0.095	0.060	0.040	0.029	0.021	0.016	0.013	0.010
Logistic	Covariance	0.438	0.260	0.132	0.069	0.039	0.024	0.016	0.011	0.007	0.005
	Correlation	0.525	0.184	0.097	0.062	0.042	0.029	0.022	0.016	0.013	0.010

As commented before, the condition number needs to be computed in some scenarios. For instance, in [41], the condition number is used to estimate covariance matrices with a constraint that renders the estimated covariance matrix ill-conditioned, since, as commented before, it increases the relative error of many algebraic computations (see Section 6 in [7]). Since the condition number is the quotient of the largest and the smallest eigenvalues, it can be computed for the type of matrices considered in this paper as a consequence of Theorem 2.6.

Finally, notice that the bidiagonal decomposition of the type of matrices considered in this paper, which can be computed with HRA, can be used to compute the LDL^T decomposition of the covariance matrix. Since L can be expressed as the product of totally positive matrices, it can also be computed with HRA. The values of L allow us to express the i-th order statistic as a sum of the smaller order statistics and an independent random variable. Moreover, such an expression is directly related to the so-called spacings, defined as $X_{(i+1)} - X_{(i)}$ for $i \in \{1, \dots, n-1\}$. Spacings are relevant in the study of the time between two consecutive failures in reliability theory, see [42].

6 | A Real-Life Example Related to Sports Data

In this section, an example using real-life data is used to illustrate the presented results related to order statistics. Consider the results of the Men's 100-meter freestyle swimming competition held during the 2024 Olympic Games in Paris, France. The data, which is presented in Table 2 and has been retrieved from [43], show the results of each athlete in the different races. Notice that the data for each race is ordered from smallest to largest, so the order statistics appear naturally.

If we ignore the observation related to Heat 1, since it has incomplete information and abnormally large values, the sample covariance matrix is the following:

$$S \approx \begin{pmatrix} 2.958 & 2.946 & 2.983 & 3.034 & 2.998 & 2.976 & 3.055 & 3.816 \\ 2.946 & 3.103 & 3.142 & 3.200 & 3.147 & 3.178 & 3.261 & 3.998 \\ 2.983 & 3.142 & 3.200 & 3.258 & 3.200 & 3.246 & 3.326 & 4.080 \\ 3.034 & 3.200 & 3.258 & 3.330 & 3.267 & 3.315 & 3.398 & 4.145 \\ 2.998 & 3.147 & 3.200 & 3.267 & 3.215 & 3.255 & 3.338 & 4.076 \\ 2.976 & 3.178 & 3.246 & 3.315 & 3.255 & 3.348 & 3.428 & 4.130 \\ 3.055 & 3.261 & 3.326 & 3.398 & 3.338 & 3.428 & 3.514 & 4.234 \\ 3.816 & 3.998 & 4.080 & 4.145 & 4.076 & 4.130 & 4.234 & 5.526 \end{pmatrix}.$$

TABLE 2 | Results of the Men's 100-meter freestyle swimming competition of the 2024 Olympic Games in Paris.

Race	1st	2nd	3rd	4th	5th	6th	7th	8th
Final	46.40	47.48	48.57	49.67	50.98	52.38	53.94	55.52
Semifinal 1	47.58	47.61	47.66	47.80	48.25	48.76	49.29	50.13
Semifinal 2	47.21	47.66	47.68	47.94	47.95	48.06	48.10	48.78
Heat 1	53.85	55.56	57.79	60.28	65.81	75.28	88.64	—
Heat 2	52.22	52.73	52.95	52.99	53.10	53.19	53.38	56.19
Heat 3	51.21	51.77	51.91	52.29	52.32	52.35	52.52	52.94
Heat 4	50.35	50.39	50.39	50.58	50.81	50.84	51.12	51.42
Heat 5	48.84	50.10	50.42	50.63	50.64	51.46	51.54	51.55
Heat 6	48.84	49.12	49.44	49.51	49.60	49.65	49.70	49.92
Heat 7	48.49	48.66	48.82	48.82	49.06	49.24	49.28	49.35
Heat 8	47.57	47.70	48.01	48.24	48.40	48.41	48.53	49.34
Heat 9	47.92	47.93	48.25	48.34	48.47	48.82	48.88	49.38
Heat 10	48.07	48.25	48.34	48.35	48.40	48.40	48.46	48.79

As can be seen in the sample covariance matrix, the variability seems to be smaller for the first positions. All correlations are positive and close to 1, since covariances and variances are similar. The condition number of the matrix is 9×10^4. In order to obtain the estimation of the covariance matrix under a total positivity constraint, one has to compute the weights of the graphs in Algorithm 1. These weights are provided in the following matrix:

$$W = \begin{pmatrix} - & \mathbf{-2.916} & -2.814 & -2.725 & -2.906 & -2.246 & -2.281 & -2.217 \\ \mathbf{-2.916} & - & \mathbf{-5.222} & -4.699 & -4.954 & -3.577 & -3.698 & -2.693 \\ -2.814 & \mathbf{-5.222} & - & \mathbf{-5.508} & -5.450 & -4.080 & -4.128 & -2.836 \\ -2.725 & -4.699 & \mathbf{-5.508} & - & \mathbf{-5.909} & -4.232 & -4.314 & -2.717 \\ -2.906 & -4.954 & -5.450 & \mathbf{-5.909} & - & \mathbf{-4.148} & -4.319 & -2.740 \\ -2.246 & -3.577 & -4.080 & -4.232 & \mathbf{-4.148} & - & \mathbf{-6.693} & -2.548 \\ -2.281 & -3.698 & -4.128 & -4.314 & -4.319 & \mathbf{-6.693} & - & \mathbf{-2.567} \\ -2.217 & -2.693 & -2.836 & -2.717 & -2.740 & -2.548 & \mathbf{-2.567} & - \end{pmatrix}.$$

The shortest Hamiltonian path, computed using the TSP package [44], is $(1, 2, 3, 4, 5, 6, 7, 8)$. This implies that the variables that are not supposed to be conditionally independent given all other variables are the consecutive order statistics. Notice that this is coherent with what is stated in Section 3.1.

Next, we test for the fulfillment of the necessary assumptions, that is, the multivariate normality, the conditional independence between nonconsecutive variables and the fact that correlations are positive (the reader is referred to Section 2.5 in [8] for more details). The Henze Zirkler test of multivariate normality has been performed, obtaining a p-value of 0.105. Therefore, multivariate normality is not rejected. Moreover, conditional independence and tailed correlation tests have been performed considering the Benjamini and Hochberg correction. The results show that the conditional independence cannot be rejected for any pair of variables, and it is concluded that all correlations are positive. In conclusion, the necessary conditions for the estimation to yield a covariance matrix that is close to the population covariance matrix are reasonable. The associated estimation is the following:

$$\hat{\Sigma} = \begin{pmatrix} 2.958 & 2.946 & 2.984 & 3.038 & 2.981 & 3.018 & 3.090 & 3.723 \\ 2.946 & 3.103 & 3.142 & 3.199 & 3.139 & 3.178 & 3.254 & 3.921 \\ 2.984 & 3.142 & 3.200 & 3.258 & 3.196 & 3.236 & 3.313 & 3.992 \\ 3.038 & 3.199 & 3.258 & 3.330 & 3.267 & 3.308 & 3.387 & 4.081 \\ 2.981 & 3.139 & 3.196 & 3.267 & 3.215 & 3.255 & 3.332 & 4.015 \\ 3.018 & 3.178 & 3.236 & 3.308 & 3.255 & 3.348 & 3.428 & 4.130 \\ 3.090 & 3.254 & 3.313 & 3.387 & 3.332 & 3.428 & 3.514 & 4.234 \\ 3.723 & 3.921 & 3.992 & 4.081 & 4.015 & 4.130 & 4.234 & 5.526 \end{pmatrix}.$$

Notice that the resulting $\hat{\Sigma}$ is similar to the usual sample covariance matrix S, but $\hat{\Sigma}$ is totally positive [8]. The largest differences are associated with the most distant variables, since the correlation is computed by considering the product of correlation coefficients over longer paths (see Lemma 4.1).

The errors in the computation of the eigenvalues of the estimated matrix when using the function *eig* implemented in Matlab [45] and the routine implemented in TNTool [46–48] (HRA) are provided in Table 3. As expected, the relative error in the computation of all eigenvalues is lower for the method with HRA. For the real values of the eigenvalue s symbolic computation in Mathematica [49], with 60 digits of precision, has been considered. As expected, the relative error in the computation of all six eigenvalues is lower for the method with HRA. In addition, the inverse matrix has been computed using the function *inv* implemented in Matlab and the routine implemented in TNTool. As the real value of the inverse, the formula given in Proposition 3.1 in [8] is considered, as well as symbolic computation with 60 digits of precision. The relative error considering the 1-norm (defined as $||A||_1 = \max_{1 \le j \le n} \sum_{i=1}^{n} |A_{ij}|$) is $9.050784 \cdot 10^{-14}$ for the Matlab routine and $5.525399 \cdot 10^{-17}$ for the HRA method. Finally, a similar procedure has been carried out for the condition number, considering the routine *cond* in Matlab and the quotient of the smallest and largest eigenvalues computed with HRA, obtaining relative errors of $4.0563 \cdot 10^{-14}$ and $3.854015 \cdot 10^{-17}$, respectively. Such differences will increase with the condition number of the matrix.

TABLE 3 | Eigenvalues and relative error of the function eig from Matlab and the routine TNEigenValues in TNTool for the estimated covariance matrix.

λ	e_{Matlab}	e_{HRA}
27.521168031	$3.872706646791 \cdot 10^{-16}$	$3.872706646791 \cdot 10^{-16}$
0.4138924389	$9.388382635143 \cdot 10^{-16}$	$8.047185115837 \cdot 10^{-16}$
0.1581954879	$5.614435856193 \cdot 10^{-15}$	$1.754511205060 \cdot 10^{-16}$
0.0697814206	$9.744880463186 \cdot 10^{-15}$	$5.966253344808 \cdot 10^{-16}$
0.0181454295	$3.499001164994 \cdot 10^{-14}$	$1.912022494532 \cdot 10^{-16}$
0.0062455862	$9.138037659067 \cdot 10^{-14}$	$2.777519045309 \cdot 10^{-16}$
0.0033462618	$2.203227097457 \cdot 10^{-15}$	$3.888047819041 \cdot 10^{-16}$
0.0020733431	$3.304883617246 \cdot 10^{-14}$	$4.183396983856 \cdot 10^{-16}$

7 | Conclusion

In this paper, we studied the application of a method for the estimation of the population covariance matrix of a Gaussian Markov Random Field under a total positivity constraint to samples consisting of order statistics of identically distributed random variables. The presented theoretical results state that the order statistics are Markov chains, the correlation coefficients between consecutive order statistics are nonnegative, and that, asymptotically, their joint distribution is multivariate Gaussian. Moreover, it has been proved that the covariance matrix of a finite number of order statistics converges to a totally positive matrix.

Notice that, when working with order statistics, the necessary conditions for the estimation method for the population covariance matrix of a Gaussian Markov Random Field under a total positivity constraint to yield a covariance matrix that is close to the population covariance matrix are not as restrictive as in [8]. More precisely, in the context of order statistics, it is only required that we are dealing with independent observations of a sufficiently regular continuous distribution.

The results presented in this paper enable the application of HRA methods in the context of covariance matrices across a wide range of tasks, including the definition of certain mean estimators, the computation of conditional covariances, entropy computation, and data reduction (PCA). We conclude by demonstrating the proposed method using sports data, where a natural ordering of the observations is inherent. The findings reveal a difference of two orders of magnitude in the relative errors when computing the inverse matrix, the condition number, and eigenvalues associated with the empirical covariance matrix.

As an open problem, it would be interesting to study the case in which the number of order statistics k increases with the sample size n, since one may be tempted to consider a larger number of order statistics when the sample size is greater.

Author Contributions

Juan Baz: conceptualization; investigation; software; validation; data curation; writing – original draft. **Pedro Alonso:** conceptualization; investigation; formal analysis; validation; supervision; writing – review and editing; funding acquisition. **Juan Manuel Peña:** conceptualization; investigation; supervision; writing – review and editing; formal analysis. **Raúl Pérez-Fernández:** conceptualization; investigation; supervision; validation; writing – review and editing; formal analysis.

Acknowledgments

This research was partially supported through the Spanish research grants PID2022-138569NB-I00 (MCIU/AEI), PID2022-139886NB-I00 (MCIU/AEI), PID2022-140585NB-I00 (MICIU/AEI/10.13039/501100011033 and "FEDER/UE") and RED2022-134176-T (MCI/AEI), Gobierno de Aragón (E41_23R), and Principado de Asturias (Programa Severo Ochoa, BP21042).

Funding

This work was supported by Gobierno de Aragón, Agencia Estatal de Investigación, Ministerio de Ciencia Innovación y Universidades, and Principado de Asturias (Programa Severo Ochoa, BP21042) and Spanish Research grants PID2022-138569NB-I00

(MCIU/AEI), PID2022-139886NB-I00 (MCIU/AEI), PID2022-140585NB-I00 (MICIU/AEI/10.13039/501100011033 and "FEDER/UE") and RED2022-134176-T (MCI/AEI) and Principality of Asturias (Severo Ochoa Program: BP21042; Asturian Agency for Science, Business Competitiveness and Innovation: SEK-25-GRU-GIC-24-018).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The authors confirm that the data supporting the findings of this study are available within the article and/or its Supporting Information.

References

1. U. Shahzad, I. Ahmad, I. M. Almanjahie, N. H. Al-Noor, and M. Hanif, "A Novel Family of Variance Estimators Based on l-Moments and Calibration Approach Under Stratified Random Sampling," *Communications in Statistics: Simulation and Computation* 52, no. 8 (2023): 3782–3795.

2. N. Tiwari, G. Chandra, S. Bhari, and J. Banerjie, "Estimation of Location and Scale Parameters of Lognormal Distribution Using Median With Extreme Ranked Set Sampling," *Sankhya. Series B* 87, no. 1 (2025): 27.

3. A. Figuerola-Wischke and A. M. Gil-Lafuente, "Forecasting the Real Average Retirement Benefit in the United States Using Owa Operators," *Technological and Economic Development of Economy* 30, no. 4 (2024): 956–975.

4. M. Flores-Sosa, E. Avilés-Ochoa, J. M. Merigó, and J. Kacprzyk, "The Owa Operator in Multiple Linear Regression," *Applied Soft Computing* 124 (2022): 108985.

5. H. A. David and H. N. Nagaraja, *Order Statistics* (John Wiley & Sons, 2004).

6. T. Ando, "Totally Positive Matrices," *Linear Algebra and Its Applications* 90 (1987): 165–219.

7. J. Baz, P. Alonso, J. M. Peña, and R. Pérez-Fernández, "Gaussian Markov Random Fields Over Graphs of Paths and High Relative Accuracy," *Journal of Computational and Applied Mathematics* 453 (2025): 116142.

8. J. Baz, P. Alonso, J. M. Peña, and R. Pérez-Fernández, "Estimation of the Covariance Matrix of a Gaussian Markov Random Field Under a Total Positivity Constraint," *Journal of Computational and Applied Mathematics* 464 (2025): 116543.

9. K. V. Mardia, J. T. Kent, and C. C. Taylor, *Multivariate Analysis* (John Wiley & Sons, 2024).

10. W. Kocay and D. L. Kreher, *Graphs, Algorithms, and Optimization* (Chapman and Hall/CRC, 2016).

11. H. Rue and L. Held, *Gaussian Markov Random Fields: Theory and Applications* (Chapman and Hall/CRC, 2005).

12. J. Baz, I. Díaz, S. Montes, and R. Pérez-Fernández, "Some Results on the Gaussian Markov Random Field Construction Problem Based on the Use of Invariant Subgraphs," *TEST* 31, no. 3 (2022): 856–874.

13. M. Gasca and J. M. Peña, "On Factorizations of Totally Positive Matrices," in *Total Positivity and Its Applications* (Springer, 1996), 109–130.

14. M. Gasca and J. M. Peña, "Total Positivity and Neville Elimination," *Linear Algebra and Its Applications* 165 (1992): 25–44.

15. P. Alonso, J. Jiménez, J. M. Peña, and M. L. Serrano, "Depth of Almost Strictly Sign Regular Matrices," *Mathematical Methods in the Applied Sciences* 46, no. 1 (2023): 732–744.

16. J. Delgado and J. M. Peña, "High-Relative-Accuracy Computations With Kac-Murdock-Szegö Matrices and Their Generalizations," *Computational and Applied Mathematics* 44, no. 4 (2025): 188.

17. J. Baz, P. Alonso, J. M. Peña, and R. Pérez-Fernández, "Gaussian Markov Random Fields and Totally Positive Matrices," *Journal of Computational and Applied Mathematics* 430 (2023): 115098.

18. V. K. Rohatgi and A. M. E. Saleh, *An Introduction to Probability and Statistics* (John Wiley & Sons, 2015).

19. M. Shaked and J. G. Shanthikumar, *Stochastic Orders* (Springer, 2007).

20. S. Fallat, S. Lauritzen, K. Sadeghi, C. Uhler, N. Wermuth, and P. Zwiernik, "Total Positivity in Markov Structures," *Annals of Statistics* 45 (2017): 1152–1184.

21. J. Navarro and N. Balakrishnan, "Study of Some Measures of Dependence Between Order Statistics and Systems," *Journal of Multivariate Analysis* 101, no. 1 (2010): 52–67.

22. R. J. Serfling, *Approximation Theorems of Mathematical Statistics* (John Wiley & Sons, 2009).

23. N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, vol. 1 (John Wiley & Sons, 1994).

24. N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, vol. 2 (John Wiley & Sons, 1995).

25. M. Stephens, "Asymptotic Calculations of Functions of Expected Values and Covariances of Order Statistics," *Canadian Journal of Statistics* 18, no. 3 (1990): 265–270.

26. J. M. Peña, "M-Matrices Whose Inverses Are Totally Positive," *Linear Algebra and Its Applications* 221 (1995): 189–193.

27. A. Pinkus, *Totally Positive Matrices. No. 181* (Cambridge University Press, 2010).

28. C. D. Meyer, *Matrix Analysis and Applied Linear Algebra* (SIAM, 2023).

29. E. H. Lloyd, "Least-Squares Estimation of Location and Scale Parameters Using Order Statistics," *Biometrika* 39, no. 1/2 (1952): 88–95.

30. A. E. Sarhan, "Estimation of the Mean and Standard Deviation by Order Statistics," *Annals of Mathematical Statistics* 25, no. 2 (1954): 317–328.

31. A. E. Sarhan, "Estimation of the Mean and Standard Deviation by Order Statistics. Part II," *Annals of Mathematical Statistics* 26, no. 3 (1955): 505–511.

32. A. E. Sarhan, "Estimation of the Mean and Standard Deviation by Order Statistics. Part III," *Annals of Mathematical Statistics* 26, no. 4 (1955): 576–592.

33. M. Ahsanullah and A. Alzaatreh, "Parameter Estimation for the Log-Logistic Distribution Based on Order Statistics," *REVSTAT-Statistical Journal* 16, no. 4 (2018): 429–443.

34. A. Dytso, M. Cardone, M. S. Veedu, and H. V. Poor, "On Estimation Under Noisy Order Statistics," in *Proceedings of the 2019 IEEE International Symposium on Information Theory* (IEEE, 2019), 36–40.

35. D. Kumar, M. Kumar, and J. S. Joorel, "Estimation With Modified Power Function Distribution Based on Order Statistics With Application to Evaporation Data," *Annals of Data Science* 9, no. 4 (2022): 723–748.

36. G. H. Golub and C. F. Van Loan, "Cf Matrix Computations, Johns Hopkins Studies in the Mathematical Sciences," in *The Johns Hopkins University Press*, vol. 90 (The Johns Hopkins University Press, 1996), 93–94.

37. S. Zozor, E. Moisan, and P.-O. Amblard, "Revisiting the Estimation of the Mean Using Order Statistics," *Signal Processing* 68, no. 2 (1998): 155–173.

38. J. V. Michalowicz, J. M. Nichols, and F. Bucholtz, *Handbook of Differential Entropy* (CRC Press, 2013).

39. M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*, vol. 55 (Courier Corporation, 1965).

40. R. Bro and A. K. Smilde, "Principal Component Analysis," *Analytical Methods* 6, no. 9 (2014): 2812–2831.

41. J. H. Won and S.-J. Kim, "Maximum Likelihood Covariance Estimation With a Condition Number Constraint," in *Proceedings of the 2006 Fortieth Asilomar Conference on Signals, Systems and Computers* (IEEE, 2006), 1445–1449.

42. S. C. Kochar, "Stochastic Comparisons of Spacings and Order Statistics," in *Frontiers in Reliability* (World Scientific, 1998), 201–216.

43. Le Monde, "Results 100m Freestyle Swimming Paris 2024," (2024), https://www.lemonde.fr/en/sport/jo-2024/results/swimming/100m-freestyle-swimming-m/.

44. M. Hahsler and K. Hornik, "TSP: Traveling Salesperson Problem (TSP), R Package Version 1.2–4," 2023.

45. The MathWorks, "Matlab Reference Guide,"

46. P. Koev, "Accurate Eigenvalues and SVDs of Totally Nonnegative Matrices," *SIAM Journal on Matrix Analysis and Applications* 27, no. 1 (2005): 1–23.

47. P. Koev, "Accurate Computations With Totally Nonnegative Matrices," *SIAM Journal on Matrix Analysis and Applications* 29, no. 3 (2007): 731–751.

48. P. Koev "Tntool," (2024), https://math.mit.edu/~plamen/software/TNTool.html.

49. S. Wolfram, *The Mathematica Book* (Wolfram Research, Inc., 2003).