International Journal of Management, Knowledge and Learning, 9(2), 223–235

On Some Applications of Matrix Partial Orders in Statistics

Iva Golubić
University of Ljubljana, Faculty of Electrical Engineering, Slovenia

Janko Marovt
University of Maribor, Faculty of Economics and Business, Slovenia

In statistics different partial orders appear as useful in several cases. Three of the best known partial orders defined on (sub)sets of real or complex matrices are the Löwner, the minus and the star partial orders. Another two matrix partial orders that are related to the star partial order are the left-star and the right-star partial orders. In the paper we review some of the applications of mentioned partial orders in statistics.

Keywords: matrix partial order, generalized matrix inverse, preserver, statistics, linear model

Introduction

Mathematics is essential for all (serious) branches of science, including natural science, engineering, medicine, finance, and in the last few decades also for social sciences. One can argue that a particular practice becomes a scientific discipline when it starts to obey the postulates of mathematics and adopts the mathematical language and mathematical (especially analytical) way of thinking. Mathematics and statistics are becoming increasingly important in daily operations of various organizations, e.g. for modern knowledge management (i.e. a process of creating, sharing, using and managing the knowledge and information of an organization) the use of mathematics and statistics is crucial (Munje et al., 2020; Phusavat et al., 2009; Priestley & McGrath, 2019). Linear algebra is a branch of mathematics that especially in its subbranch of matrix theory encompasses results which are used in various fields of science and practice. We can not imagine modern (micro and macro) economics and econometrics without the use of matrices. Matrices are for example useful in observing the relationships between individual industries and in calculating the quantities needed to meet the demand for goods produced in the industries of an economy. We can also use matrices in linear programming in management to for example adjust production processes by solving optmization problems such as calculation of the minimum of production costs. In the paper we present particular relations between matrices that have many applications in statistics and thus in other scientific fields, and give an overview of these applications.

Let $\mathbb{F}$ denote the field of all real or complex numbers, i.e. $\mathbb{F}=\mathbb{R}$ or $\mathbb{F}=\mathbb{C}$, and $M_{m,n}(\mathbb{F})$, the set of all $m\times n$ matrices over $\mathbb{F}$. If $m=n$, then we write $M_n(\mathbb{F})$ instead of $M_{n,n}(\mathbb{F})$. Let $A^*\in M_{n,m}(\mathbb{F})$ denote the conjugate transpose of $A\in M_{m,n}(\mathbb{F})$ (if $A\in M_{m,n}(\mathbb{R})$, then $A^*=A^t$, the transpose of A). A generalized inverse or a pseudoinverse of $A\in M_{m,n}(\mathbb{F})$ is a matrix that has some properties of the usual inverse (of $A\in M_n(\mathbb{F})$ with the nonzero determinant) but not necessarily all of them. One of the best known examples of a generalized inverses is the Moore-Penrose inverse. We say that $X\in M_{n,m}(\mathbb{F})$ is *the Moore-Penrose inverse* of $A\in M_{m,n}(\mathbb{F})$ when the following four matrix equations are satisfied:

$$AXA=A,\qquad XAX=X,\qquad (AX)^*=AX,\qquad \text{and}\qquad (XA)^*=XA. \tag{1}$$

It turns out (Mitra et al., 2010) that every $A\in M_{m,n}(\mathbb{F})$ has a Moore-Penrose inverse $X=A^\dagger$ and that $A^\dagger$ is unique. Another example of a pseudoinvese satisfies only the first equation in (1). Namely, we say that $X=A^-$ is *an inner generalized inverse* of $A\in M_{m,n}(\mathbb{F})$ if $A=AA^-A$. Again, every $A\in M_{m,n}(\mathbb{F})$ has an inner generalized inverse A^- however A^- is not necessarily unique. There are many applications of these pseudoinverses. For example, if $A\in M_{m,n}(\mathbb{F})$, $c\in M_{m,1}(\mathbb{F})=\mathbb{F}^m$, and x is the $n\times 1$ vector of variables, then the system

$$Ax=c \tag{2}$$

of m linear equations with n variables has a solution if and only if $AA^-c=c$ for some inner generalized inverse A^- of A. Moreover, if the system (2) has a solution and if A^- is an inner generalized inverse of A, then for every vector $y\in\mathbb{F}^n$

$$x_y=A^-c+(I-A^-A)y, \tag{3}$$

where $I\in M_n(\mathbb{F})$ is the identity matrix, is a solution of (2), and for every solution x_* of (2) there exists a vector y such that $x_*=x_y$ (Schott, 2005).

Both of the above classes of generalized inverses induce partial orders on $M_{m,n}(\mathbb{F})$ (i.e. relations that are reflexive, antisymmetric, and transitive). We say that $A\in M_{m,n}(\mathbb{F})$ is dominated by (or is below) $B\in M_{m,n}(\mathbb{F})$ with respect to *the minus partial order* and write

$$A\leq^- B\qquad \text{when}\qquad A^-A=A^-B\qquad \text{and}\qquad AA^-=BA^- \tag{4}$$

for some inner generalized inverse A^- of A.

For $A,B\in M_{m,n}(\mathbb{F})$ we write

$$A\leq^* B\quad \text{when}\quad A^*A=A^*B \text{ and } AA^*=BA^* \tag{5}$$

and name the relation $\leq^*$ *the star partial order*. It turns out that both relations (4) and (5) are indeed partial orders (Drazin, 1978; Hartwig, 1980). Moreover, the star partial order may also be defined by a generalized inverse. Namely, it is easy to see that for $A, B \in M_{m,n}(\mathbb{F})$ we have

$$A \leq^* B \quad \text{if and only if} \quad A^\dagger A = A^\dagger B \text{ and } AA^\dagger = BA^\dagger$$

where $A^\dagger$ is the Moore-Penrose inverse of A.

Two partial orders that are 'related' to the minus and the star partial orders are *the left-star* and *the right-star partial orders* (Baksalary & Mitra, 1991). Let $\mathrm{Im}A$ denote the image (i.e. the column space) of $A \in M_{m,n}(\mathbb{F})$. For $A, B \in M_{m,n}(\mathbb{F})$ we say that A is dominated by B with respect to the left-star partial order and write

$$A \mathrel{*\!\leq} B \quad \text{when} \quad A^* A = A^* B \text{ and } \mathrm{Im}A \subseteq \mathrm{Im}B. \tag{6}$$

Similarly, we define the right-star partial order: For $A, B \in M_n(\mathbb{F})$ we write

$$A \mathrel{\leq\!*} B \quad \text{when} \quad AA^* = AB^* \text{ and } \mathrm{Im}A^* \subseteq \mathrm{Im}B^*. \tag{7}$$

It is known (Mitra et al., 2010) that for $A, B \in M_{m,n}(\mathbb{F})$, $A \leq^* B$ implies both $A \mathrel{*\!\leq} B$ and $A \mathrel{\leq\!*} B$ and each $A \mathrel{*\!\leq} B$ and $A \mathrel{\leq\!*} B$ implies $A \leq^- B$. The converse implications do not hold in general.

Another well known partial order may be defined on a certain subset of $M_n(\mathbb{F})$. We say that $A \in M_n(\mathbb{F})$ is Hermitian (or symmetric when $A \in M_n(\mathbb{R})$) if $A = A^*$. A Hermitian matrix $A \in M_n(\mathbb{F})$ is said to be positive semidefinite if $x^* Ax \geq 0$ for every $x \in \mathbb{F}^n$. Positive semidefinite matrices have become fundamental computational objects in many areas of statistics, engineering, quantum information, and applied mathematics. They appear as variance-covariance matrices in statistics, as elements of the search space in convex and semidefinite programming, as kernels in machine learning, as density matrices in quantum information, and as diffusion tensors in medical imaging. It is known (Christensen, 1996) that every variance-covariance matrix is positive semidefinite, and that every real positive semidefinite matrix is a variance-covariance matrix of some multivariate distribution. Let now $A, B \in M_n(\mathbb{F})$ be Hermitian matrices. We say that A is dominated by B with respect to the Löwner partial order and write

$$A \leq^L B \quad \text{if} \quad B - A \quad \text{is positive semidefinite.} \tag{8}$$

There are many applications of the above partial orders, especially in statistics. Let us present an example of such an application (Mitra et al., 2010). Let $A, B \in M_n(\mathbb{R})$ be two positive semidefinite matrices and let $A \leq^L B$. Write

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

where A_{ij} and B_{ij} are of the same order for all $i,j \in \{1,2\}$ and A_{11} is a $r \times r$, $r < n$, matrix. Then (Sengupta & Jammalamadaka, 2003)

$$A_{11} - A_{12}A_{22}^{-}A_{21} \leq^{L} B_{11} - B_{12}B_{22}^{-}B_{21}. \tag{9}$$

Consider now a tribal population on which several anthropometric measurements are made. Let y_1 be the vector of measurements on the face and y_2 the vector of measurements on the remaining part of the body. Let the random vector $y = \begin{bmatrix} y_1 & y_2 \end{bmatrix}^t$ have the multivariate normal distributions $N(\mu, V_1)$ in population 1 and $N(\tau, V_2)$ in population 2. Here μ and τ are the mean vectors and V_1 and V_2 are the variance-covariance matrices (also known as dispersion or covariance matrices). Suppose y has a smaller dispersion in population 1 than in population 2. The 'smaller dispersion' condition may be expressed in terms of the Löwner partial order $\leq^L$, i.e. $V_1 \leq^L V_2$. By (9) and by properties of variance-covariance (dispersion) matrices (Sengupta & Jammalamadaka, 2003, p. 59) we have the following: The conditional dispersion of facial measurements given the measurements of the rest of the body, namely $V(y_1 \mid y_2)$, is also smaller in population 1 than in population 2.

In the following two sections more applications of the above partial orders in statistics will be presented – we will focus our attention on linear models. In the next section (Linear Models) we will recall the notion of a linear model and then use matrix partial orders to compare different linear models. In the last decades many authors studied *preserves problems* which concern the question of determining or describing the general form of all transformations of a given structure $\mathscr{X}$ which preserve a quantity attached to the elements of $\mathscr{X}$, or a distinguished set of elements of $\mathscr{X}$, or a given relation among the elements of $\mathscr{X}$, etc. It has been recently stated (Dolinar et al., 2020; Golubić & Marovt, in press, 2020; Guillot et al., 2015) that a motivation for the study of preserver problems that concern the above partial orders on certain (sub)sets of real matrices (i.e. $\mathscr{X}$ is a subset of $M_n(\mathbb{R})$) comes from statistics. Let $\mathscr{S}$ be a subset of $M_n(\mathbb{F})$ and let $\leq_G$ be one of the above orders (i.e. $\leq^-$, $\leq^*$, ${}^{*}\!\!\leq$, $\leq\!{}^{*}$, $\leq^L$) on $\mathscr{S}$. We say that the map $\Phi{:}\mathscr{S} \to \mathscr{S}$ preserves the partial order $\leq_G$ in both directions (or is a bi-preserver of $\leq_G$) when for every $A, B \in \mathscr{S}$,

$$A \leq_G B \quad \text{if and only if} \quad \Phi(A) \leq_G \Phi(B). \tag{10}$$

In the last section (Preservers of Partial Orders) we will recall some recent results that were motivated by statistics and that (under some additional assumptions) describe the form of maps Φ with the property (10).

Linear Models

One of the simplest models that is used to illustrate how an observed quantity y can be explained by a number of other quantities, $x_1, x_2, \ldots, x_{p-1}$, is the linear model

$$y=\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_{p-1}x_{p-1}+\epsilon,$$

where $\beta_0,\beta_1,\ldots,\beta_{p-1}$ are constants (real numbers) and ϵ is an error term that accounts for uncertainties. We refer to y as *a response variable* and to $x_1,x_2,\ldots,x_{p-1}$ as *explanatory variables*. For a set of n observations of the response and explanatory variables, the explicit form of the equations would be

$$y_i=\beta_0+\beta_1 x_{i,1}+\beta_2 x_{i,2}+\cdots+\beta_{p-1}x_{i,p-1}+\epsilon_i, \qquad i=1,2,\ldots,n,$$

where for each i, y_i is the i-th observation of the response, $x_{i,j}$ is the i-th observation of the j-th explanatory variable ($j=1,2,\ldots,p-1$), and ϵ_i is the unobservable error corresponding to this observation. These equations can be written in the following matrix form:

$$y=X\beta+\epsilon.$$

Here,

$$y=\begin{bmatrix} y_1\\ y_2\\ \vdots\\ y_n\end{bmatrix}, \quad X=\begin{bmatrix} 1 & x_{1,1} & \cdots & x_{1,p-1}\\ 1 & x_{2,1} & \cdots & x_{2,p-1}\\ \vdots & \vdots & \ddots & \vdots\\ 1 & x_{n,1} & \cdots & x_{n,p-1}\end{bmatrix}, \quad \beta=\begin{bmatrix} \beta_0\\ \beta_1\\ \vdots\\ \beta_{p-1}\end{bmatrix}, \quad \epsilon=\begin{bmatrix} \epsilon_1\\ \epsilon_2\\ \vdots\\ \epsilon_n\end{bmatrix}.$$

We call y *the response vector* (also known as *the observation vector*) and X *the model matrix* (also known as *the design* or *regressor matrix*). In order to complete the description of the model, some assumptions about the nature of the errors have to be made. It is assumed that $E(\epsilon)=0$ and $V(\epsilon)=\sigma^2 D$, i.e. the errors have the zero mean and covariances are known up to a scalar (real number). Here V denotes the variance-covariance matrix. The nonnegative parameter σ^2 and the vector of parameters (real numbers) β are unspecified, and D is a known $n\times n$ (real, positive semidefinite) matrix. We denote this linear model with the triplet $(y,X\beta,\sigma^2 D)$. (It follows that $E(y)=X\beta$ and $V(y)=\sigma^2 D$.) It is known (Mitra et al., 2010, Lemma 15.2.1) that the response vector $y\in \mathrm{Im}(X{:}D)$ with probability 1 where $\mathrm{Im}(X{:}D)$ denotes the image (i.e. the column space) of the partitioned matrix $(X{:}D)$.

Remark An assumption that the errors follow the multivariate normal distribution is often added to the model. Moreover, the matrix X above where all the elements in a first column equal 1 is in fact a special case of a linear model matrix; such model matrices are used in the multiple regression analysis. Models $(y,X\beta,\sigma^2 D)$ where all the elements in the first column of the model matrix do not necessarily equal 1 and the probability distribution of the errors is not necessarily normal are (usually) called general linear models. In the continuation we will deal with general linear models, however, for the sake of simplicity we will use the term "linear model" instead of "general linear model".

Classical inference problems related to the linear model $(y, X\beta, \sigma^2 D)$ usually concern a linear parametric function (LPF), $s\beta$ (here s is a $1 \times p$ real vector). We try to estimate it by a linear function of the response zy (here z is a $1 \times n$ real vector). For accurate estimation of $s\beta$, it is desirable that the estimator is not systematically away from the 'true' value of the parameter. We say that the statistic zy is *a linear unbiased estimator* (LUE) of $s\beta$ if $E(zy) = s\beta$ for all possible values of β. A LPF is said to be estimable if it has an LUE.

Let $(y, X\beta, \sigma^2 D)$ be a linear model and let A be a real matrix with p columns. We say that a vector LPF, $A\beta$, is *estimable* if there exists a real matrix C such that $E(Cy) = A\beta$ for all $\beta \in \mathbb{R}^p$. It turns out (Mitra et al., 2010, Theorem 15.2.4) that if A is a real matrix with p columns, then

$$A\beta \quad \text{is estimable if and only if} \quad \operatorname{Im} A^t \subseteq \operatorname{Im} X^t. \tag{11}$$

The best linear unbiased estimator (BLUE) of an estimable vector LPF is defined as the LUE having the smallest variance-covariance matrix. Here, the "variance-covariance" condition is expressed in terms of the Löwner order $\leq^L$: Let $A\beta$ be estimable. Then Ly is said to be BLUE of $A\beta$ if (i) $E(Ly) = A\beta$ for all $\beta \in \mathbb{R}^p$ and (ii) $V(Ly) \leq^L V(My)$ for all $\beta \in \mathbb{R}^p$ and all My satisfying $E(My) = A\beta$.

Let us consider two linear models $L_1 = (y_1, X_1\beta, \sigma^2 D_1)$ and $L_2 = (y_2, X_2\beta, \sigma^2 D_2)$ where the number p of columns of X_1 and X_2 is fixed but arbitrary while the number n of rows may vary from model to model. Then we say that L_1 *is at least as good as* L_2 if for any LUE, $a_2 y_2$, of a parameter $k\beta$ there exists LUE $a_1 y_1$ of this parameter such that $V(a_1 y_1) \leq V(a_2 y_2)$ (here a_1, a_2, k are appropriate vectors, and V denotes the variance). If this condition is satisfied, then we write

$$L_1 \succeq L_2.$$

With the following result, which was proved in Stępniak (1985), Stępniak showed that two linear models L_1 and L_2 may be compared by considering certain matrices that are induced by matrices X_i and D_i, $i \in \{1, 2\}$, and comparing them via the Löwner partial order.

Theorem 1 *Let $L_1 = (y_1, X_1\beta, \sigma^2 D_1)$ and $L_2 = (y_2, X_2\beta, \sigma^2 D_2)$ be two linear models. Then $L_1 \succeq L_2$ is equivalent to $M_2 \leq^L M_1$ where M_i, $i \in \{1, 2\}$, are positive semidefinite matrices defined as*

$$M_i = X_i^t (D_i + X_i X_i^t)^- X_i.$$

Here $(D_i + X_i X_i^t)^-$ *is an inner generalized inverse of* $D_i + X_i X_i^t$.

With the next two results (Mitra et al., 2010, Theorems 15.3.6, 15.3.7) we consider the linear models with model matrices that are related to each other under the minus partial order or the left-star partial order. Let $L_1 = (y, X_1\theta, \sigma^2 D)$ and $L_2 = (y, X_2\beta, \sigma^2 D)$ be two linear models and suppose $X_1 \leq^- X_2$. Note that for any two matrices $A, B \in M_{m,n}(\mathbb{F})$, we have $A \leq^- B$ if and only if $B - A \leq^- B$ (Mitra et al., 2010, Theorem 3.3.16). Let $A = X_2 - X_1$. It follows that then $A \leq^- X_2$ and therefore by (4) there exists an inner generalized inverse A^- of A such that $A^-A = A^-X_2$ and $AA^- = X_2A^-$. Since then $A = AA^-A = AA^-X_2$ and thus $A^t = X_2^t(AA^-)^t$, we may conclude that $\mathrm{Im}\,A^t \subseteq \mathrm{Im}\,X_2^t$ (i.e. $A\beta$ is by (11) estimable in the model L_2). Let the model L_2 be constrained by linear constraints $A\beta = 0$ on the parametric vector $\beta \in \mathbb{R}^p$. Observe that on the one hand $A = X_2 - X_1$ and $AA^- = X_2A^-$ imply

$$X_1 = X_2 - A = X_2 - AA^-A = X_2 - X_2A^-A = X_2(I - A^-A), \tag{12}$$

and on the other hand, $(I - A^-A)\theta$ where $\theta \in \mathbb{R}^p$ is arbitrary are by (3) exactly the solutions of the system $A\beta = 0$ of linear equations (where β is the vector of variables). So, by (12) for each $\beta \in \mathbb{R}^p$ where $A\beta = 0$ there exists $\theta \in \mathbb{R}^p$ such $X_2\beta = X_1\theta$ and for each $\theta \in \mathbb{R}^p$ there exists a solution $\beta \in \mathbb{R}^p$ of $A\beta = 0$ such that $X_1\theta = X_2\beta$. It follows that the model L_1 is the model L_2 constrained by $A\beta = 0$. We may conclude that if $L_1 = (y, X_1\theta, \sigma^2 D)$ and $L_2 = (y, X_2\beta, \sigma^2 D)$ are two linear models with $X_1 \leq^- X_2$, then there exists a matrix A such that $A\beta$ is estimable in the model L_2 and L_1 is the model L_2 constrained by $A\beta = 0$.

We presented the above argument as an example of how purely linear algebraic techniques can lead to a result that has implications in statistics. It turns out that the converse of the proved implication is true as well (Mitra et al., 2010, proof of Theorem 15.3.6).

Theorem 2 *Let* $L_1 = (y, X_1\theta, \sigma^2 D)$ *and* $L_2 = (y, X_2\beta, \sigma^2 D)$ *be any two linear models. Then* $X_1 \leq^- X_2$ *if and only if there exists a matrix* A *with* $\mathrm{Im}\,A^t \subseteq \mathrm{Im}\,X_2^t$ *and* L_1 *is the model* L_2 *constrained by* $A\beta = 0$.

The following result gives an interpretation of the left-star order in Gauss-Markov linear models, i.e. linear models $(y, X\beta, \sigma^2 D)$ where $D = I$ is the identity matrix.

Theorem 3 *Let* $L_1 = (y, X_1\beta, \sigma^2 I)$ *and* $L_2 = (y, X_2\beta, \sigma^2 I)$. *Then* $X_1 \mathrel{*\!\leq} X_2$ *if and only if: (i) The linear models* L_1 *and* $L = (y, (X_2 - X_1)\beta, \sigma^2 I)$ *have no common estimable linear function of* β*;*
(ii) $X_1\beta$ *is estimable under the model* L_2*;*
(iii) The BLUE of $X_1\beta$ *under the model* L_1 *is also its BLUE under* L_2 *and the*

variance-covariance matrix of the BLUE of $X_1\beta$ under the model L_1 is the same as under the model L_2.

With the next result we give another application of the minus partial order (Baksalary & Puntanen, 1990, p. 366).

Theorem 4 *Consider a linear model $(y, X\beta, \sigma^2 D)$. Then the statistics Fy is BLUE of $X\beta$ if and only if the following conditions hold:*
(i) $FX = X$;
(ii) $\mathrm{Im}(FD) \subseteq \mathrm{Im}X$;
(iii) $V(Fy) \leq^- V(y)$.

Note that $V(Fy)$ and $V(y)$ are positive semidefinite matrices. It is thus natural to ask if there are some characterizations (i.e. equivalent definitions) of the minus partial order on the cone of all positive semidefinite matrices. Observe first that if $A = 0$ is the $n \times n$ zero matrix, then $ACA = A$ for every $C \in M_n(\mathbb{F})$. Take $A^- = 0$ to conclude by (4) that $0 \leq^- B$ for every $B \in M_n(\mathbb{F})$.

Theorem 5 *Let $A, B \in M_n(\mathbb{F})$ be positive semidefinite and $A \neq 0$. Then $A \leq^- B$ if and only if there exists an invertible matrix $S \in M_n$ such that*

$$A = S \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} S^* \quad \text{and} \quad B = S \begin{bmatrix} I_s & 0 \\ 0 & 0 \end{bmatrix} S^*$$

where I_r and I_s are $r \times r$ and $s \times s$, $s \leq n$, identity matrices, respectively, and $r < s$ if $A \neq B$, and $r = s$, otherwise. (In case when $s = n$, the zeros on the right-hand side of the formula for B are absent.)

This purely linear algebraic result (Golubić & Marovt, in press, Theorem 4.1) may now be used with Theorem 4 to obtain the following corollary.

Corollary 1 *Let $(y, X\beta, \sigma^2 D)$ be a linear model. Then the statistics Fy with $V(Fy) \neq V(y)$ is BLUE of $X\beta$ if and only if the following conditions hold:*
(i) $FX = X$;
(ii) $\mathrm{Im}(FD) \subseteq \mathrm{Im}X$;
(iii) There exist an invertible matrix $S \in M_n(\mathbb{R})$ such that

$$V(Fy) = S \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} S^t \quad \text{and} \quad V(y) = S \begin{bmatrix} I_s & 0 \\ 0 & 0 \end{bmatrix} S^t$$

where I_r is a $r \times r$ identity matrix, and I_s is a $s \times s$ identity matrix with $r < s \leq n$.

We conclude this section with another corollary of Theorem 5. Note that for a positive semidefinite matrix $A \in M_n(\mathbb{R})$, the matrix $W^t A W \in M_m(\mathbb{R})$ is still positive semidefinite for any matrix $W \in M_{n,m}(\mathbb{R})$. The following result (Golubić & Marovt, in press, Corollary 4.3) thus follows directly from Theorem 5 and (Baksalary et al., 1992, Theorem 1).

Corollary 2 *Let $A=\sum_{i=1}^{k}A_i$ where $A_i \in M_n(\mathbb{R})$ are positive semidefinite matrices, $i=1,2,\ldots,k$. Let the $n \times 1$ random vector x follow a multivariate normal distribution with the mean μ and the variance-covariance matrix V. Let $W=(V{:}\mu)$ be a $n \times (n+1)$ partitioned matrix. Consider the quadratic forms $Q=x^tAx$ and $Q_i=x^tA_ix$, $i=1,2,\ldots,k$. Then the following statements are equivalent.*

(i) Q_i, $i=1,2,\ldots,k$, are mutually independent and distributed as chi-squared variables;

(ii) Q is distributed as a chi-squared variable and there exist invertible matrices $S_i \in M_{n+1}(\mathbb{R})$ such that

$$W^tA_iW=S_i\begin{bmatrix} I_{r_i} & 0 \\ 0 & 0 \end{bmatrix}S_i^t \quad \text{and} \quad W^tAW=S_i\begin{bmatrix} I_s & 0 \\ 0 & 0 \end{bmatrix}S_i^t$$

for every $i=1,2,\ldots,k$, where I_{r_i} are $r_i \times r_i$ identity matrices, and I_s is a $s \times s$ identity matrix with $r_i \le s \le n+1$. (Here $I_{r_i}=0$ if $W^tA_iW=0$ for some $i \in \{1,2,\ldots,k\}$.)

Preservers of Partial Orders

The first example of a solution to a preserver problem dates back to the year 1897 when Frobenius described the form of all bijective, linear maps $\Phi{:}M_n(\mathbb{F}) \to M_n(\mathbb{F})$ that preserve the determinant, i.e.

$$\det\Phi(A)=\det A$$

for every $A \in M_n(\mathbb{F})$. Since then many authors studied various preservers problems (see the monograph by Molnár, 2007, and references therein).

Let $H_n(\mathbb{F})$ denote the set of all Hermitian (i.e. symmetric in the real case) matrices in $M_n(\mathbb{F})$, let $H_n^+(\mathbb{F})$ be the cone of all positive semidefinite matrices in $H_n(\mathbb{F})$. Note that if $A \in M_{m,n}(\mathbb{F})$, then $A^*A \in H_n^+(\mathbb{F})$. We say that two matrices $A,B \in M_{m,n}(\mathbb{F})$ are ordered as

$$A \le_N B \quad \text{if and only if} \quad A^*A \le_L B^*B,$$

i.e. $B^*B-A^*A \in H_n^+(\mathbb{F})$. The relation $\le_N$ has many applications in statistics, e.g. in the study of of probability measures, in linear estimation theory, in the analysis of the power of a binary hypothesis test, etc (Jensen, 1984, Part 2). In some of these applications order-preserving maps are used; e.g. in (Jensen, 1984, Application 3) author uses maps $\Phi{:}M_{m,n}(\mathbb{R}) \to \mathbb{R}$, defined with $\Phi(A)=\varphi(A^tA)$, $A \in M_{m,n}(\mathbb{R})$, where $\varphi{:}H_n^+(\mathbb{R}) \to \mathbb{R}$ is an order-preserving map in one direction with respect to the Löwner partial order, i.e.

$$A \le_L B \quad \text{implies} \quad \Phi(A) \le \Phi(B)$$

for every $A,B \in H_n^+(\mathbb{R})$. It is thus natural to study and try to characterize transformations on $H_n^+(\mathbb{R})$ that have a 'Löwner order-preserving property' (i.e. maps that either preserve the Löwner partial order in one or in both directions, the latter in the sense of (10)), and perhaps have some additional properties. Moreover, in modern high-dimensional probability theory and statistics, transformations are often applied to the entries of variance-covariance matrices in order to obtain regularized estimators with attractive properties (sparsity, good condition number, etc.), see Bickel and Levina (2008). The resulting matrices often serve as ingredients in statistical procedures that require these matrices to be positive semidefinite (Guillot et al., 2015).

Motivated by applications in quantum information theory and quantum statistics Molnár studied preservers that are connected to certain structures of bounded linear operators which appear in mathematical foundations of quantum mechanics, i.e. he studied automorphisms of the underlying quantum structures or, in other words, quantum mechanical symmetries. From one of Molnár's results (Molnár, 2001, Theorem 1) it follows that bijective maps $\Phi : H_n^+(\mathbb{C}) \to H_n^+(\mathbb{C})$, $n \geq 2$, where $A \leq^L B$ if and only if $\Phi(A) \leq^L \Phi(B)$, $A,B \in H_n^+(\mathbb{C})$, are of the form

$$\Phi(A) = TAT^*, \quad A \in H_n^+(\mathbb{C}) \tag{13}$$

where $T \in M_n(\mathbb{C})$ is an invertible matrix. Motivated by possible applications in statistics authors studied in (Golubić & Marovt, in press) bi-preservers on $H_n^+(\mathbb{R})$ of the Löwner partial order. They showed that a similar theorem to Molnár's result (13) holds also in the real matrix case.

Theorem 6 *Let $n \geq 2$ be an integer. Then $\varphi : H_n^+(\mathbb{R}) \to H_n^+(\mathbb{R})$ is a surjective bi-preserver of the Löwner partial order $\leq^L$ if and only if there exists an invertible matrix $S \in M_n(\mathbb{R})$ such that*

$$\varphi(A) = SAS^t$$

for every $A \in H_n^+(\mathbb{R})$.

The following observation is connected to the theory of comparison of linear models and was presented in (Golubić & Marovt, in press).

Remark Let $L_1 = (y_1, X_1\beta, \sigma^2 D_1)$ and $L_2 = (y_2, X_2\beta, \sigma^2 D_2)$ be two linear models. Here $X_1 \in M_{n,p}(\mathbb{R})$, $X_2 \in M_{m,p}(\mathbb{R})$, $D_1 \in H_n^+(\mathbb{R})$, and $D_2 \in H_m^+(\mathbb{R})$. Recall that $L_1 \succeq L_2$ means that the model L_1 is at least as good as the model L_2 and that by Theorem 1

$$L_1 \succeq L_2 \quad \text{if and only if} \quad M_2 \leq^L M_1$$

where $M_i = X_i^t\left(D_i + X_iX_i^t\right)^- X_i$, $i \in \{1,2\}$. Moreover, Stępniak noted in (Stępniak, 1985) that when $\operatorname{Im} X_i \subseteq \operatorname{Im} D_i$, $i \in \{1,2\}$, we may replace $X_i^t\left(D_i + X_iX_i^t\right)^- X_i$ with

$X_i^t D_i^- X_i$. When $D_i = X_i$, $i \in \{1,2\}$, these matrices may be further simplified to $M_i = X_i^t D_i^- X_i = D_i^t D_i^- D_i = D_i D_i^- D_i = D_i$. For models $L_1 = (y_1, D_1\beta, \sigma^2 D_1)$ and $L_2 = (y_2, D_2\beta, \sigma^2 D_2)$ we thus have

$$L_1 \succeq L_2 \quad \text{if and only if} \quad D_2 \leq^L D_1.$$

Let $n > 1$. For a random $n \times 1$ vector of observed quantities y_i, an unspecified $n \times 1$ vector β_i, and an unspecified nonnegative scalar σ_i^2, let $\mathscr{L}_i$ be the set of all linear models $L_i = (y_i, D\beta_i, \sigma_i^2 D)$ where $D \in H_n^+(\mathbb{R})$ may vary from model to model. Define a map $\psi: \mathscr{L}_1 \to \mathscr{L}_2$ with $\psi((y_1, D\beta_1, \sigma_1^2 D)) = (y_2, \varphi(D)\beta_2, \sigma_2^2 \varphi(D))$ where $\varphi: H_n^+(\mathbb{R}) \to H_n^+(\mathbb{R})$ is a surjective map. Suppose

$$L_{1_a} \succeq L_{1_b} \quad \text{if and only if} \quad \psi(L_{1_a}) \succeq \psi(L_{1_b})$$

for every $L_{1_a}, L_{1_b} \in \mathscr{L}_1$. This assumption may be reformulated as $D_{1_b} \leq^L D_{1_a}$ if and only if $\varphi(D_{1_b}) \leq^L \varphi(D_{1_a})$, $D_{1_a}, D_{1_b} \in H_n^+(\mathbb{R})$, and therefore Theorem 6 completely determines the form of any such a map ψ.

Let $A, B \in H_n(\mathbb{F})$. Since then $A^*A = A^*B$ if and only if $(A^*A)^* = (A^*B)^*$ if and only if $A^2 = BA$ which is equivalent to $AA^* = BA^*$, we may conclude (compare (5) with (6) and (7)) that the star, the left-star, and the right-star partial orders are the same partial order on $H_n(\mathbb{F})$. They are however different to the minus partial order even on $H_n^+(\mathbb{F})$ (see Golubić & Marovt, 2020, for a counterexample). Motivated by applications of the minus and the left-star partial orders in the linear model theory (see Theorems 3, 4) authors characterized in (Golubić & Marovt, in press, 2020) the surjective, additive minus and star partial order bi-preservers on $H_n^+(\mathbb{R})$, $n \geq 3$. We present here a result concerning the star partial order bi-preservers. Recall that $A \in M_n(\mathbb{R})$ is called an orthogonal matrix when $A^t A = AA^t = I$, i.e. when $A^t = A^{-1}$ where A^{-1} denotes the usual inverse of an invertible matrix A.

Theorem 7 *Let $n \geq 3$ be an integer. Then $\varphi: H_n^+(\mathbb{R}) \to H_n^+(\mathbb{R})$ is a surjective, additive bi-preserver of the star partial order if and only if there exists an orthogonal matrix $R \in M_n(\mathbb{R})$ and $\lambda > 0$ such that*

$$\varphi(A) = \lambda RAR^t$$

for every $A \in H_n^+(\mathbb{R})$.

We end the paper with a remark that in a very recent paper (Dolinar et al., 2020) the forms of general (not necessarily additive) surjective bi-preservers of the left-star partial order and the right-star partial order on $M_n(\mathbb{F})$, $n \geq 3$, were described. The results, which were expressed by using the Moore-Penrose inverse, are rather technical and hence we omit them. Nevertheless we mention that it was noted in (Dolinar et al., 2020) that given the model $M = (y, X\beta, \sigma^2 I)$ one might rather work with the transformed model $\hat{M} = (y, \hat{X}\beta, \sigma^2 I)$ because the matrix $\hat{X} \in M_n(\mathbb{R})$ has more attractive properties than $X \in M_n(\mathbb{R})$ (e.g. elements of X that are very close to zero are transformed to zero), and thus it is natural to demand that the transformed model still retains most of the properties of the original model (e.g. has similar relations to other transformed models). Thus, in view of Theorem 3, it is interesting to know what transformations on $M_n(\mathbb{R})$ preserve the left-star partial order in both directions.

We believe that preservers of various relations on sets of matrices hold a great potential for applications in statistics and hope that our review of 'preserver results' might encourage some statisticians and/or mathematicians to find further connections between certain bi-preservers and statistics.

References

Baksalary, J. K., Hauke, J., & Styan, G. P. H. (1992, March). On some distributional properties properties of quadratic forms in normal variables and some associated matrix partial orderings. Paper presenteed at the International Symposium on Multivariate Analysis and its Applications, Hong-Kong.

Baksalary J. K., & Mitra, S. K. (1991). Left-star and right-star partial orderings. *Linear Algebra and its Applications, 149,* 73–89.

Baksalary, J. K., & Puntanen, S. (1990). Characterizations of the best linear unbiased estimator in the general Gauss-Markov model with the use of matrix partial orderings. *Linear Algebra and its Applications, 127,* 363–370.

Bickel, P. J., & Levina, E. (2008). Covariance regularization by tresholding. *Annals of Statistics, 36*(6), 2577–2604.

Christensen, R. (1996). *Plane answers to complex questions: The theory of linear models.* Springer.

Dolinar, G., Halicioglu, S., Harmanci, A., Kuzma, B., Marovt, J., & Ungor, B. (2020). Preservers of the left-star and right-star partial orders. *Linear Algebra and its Applications, 587,* 70–91.

Drazin, M. P. (1978). Natural structures on semigroups with involution. *Bulletin of the American Mathematical Society, 84,* 139–141.

Golubić, I., & Marovt, J. (2020). Monotone transformations on the cone of all positive semidefinite real matrices. *Mathematica Slovaca, 70*(3), 733–744.

Golubić, I., & Marovt, J. (In press). Preservers of partial orders on the set of all variance-covariance matrices. *Filomat.*

Guillot, D., Khare, A., & Rajaratnam, B. (2015). Complete characterization of Hadamard powers preserving Loewner positivity, monotonicity, and convexity. *Journal of Mathematical Analysis and Applications, 425*(1), 489–507.

Hartwig, R. E. (1980). How to partially order regular elements. *Japanese Journal of Mathematics, 25,* 1–13.

Jensen, D. R. (1984). Invariant ordering and order preservation: Inequalities in statistics and probability. *IMS Lecture Notes, 5,* 26–34.

Mitra, S. K., Bhimasankaram, P., & Malik, S. B. (2010). *Matrix partial orders, shorted operators and applications.* World Scientific.

Molnár, L. (2001). Order-automorphisms of the set of bounded observables. *Journal of Mathematical Physics, 42,* 5904–5909.

Molnár, L. (2007). *Selected preserver problems on algebraic structures of linear operators and on function spaces* (Lecture Notes in Mathematics, No. 1895). Springer.

Munje, P. N., Tsakeni, M., & Jita, L. C. (2020). School heads of departments' roles in advancing science and mathematics through the distributed leadership framework. *International Journal of Learning, Teaching and Educational Research, 19*(9), 39–57.

Phusavat, K., Anussornnitisarn, P., Tangkapaisarn, P., & Kess, P. (2009). Knowledge management and performance improvement: The roles of statistics and mathematics. *International Journal of Innovation and Learning, 6*(3), 306–322.

Priestley, J., & McGrath, R. J. (2019). The evolution of data science: A new mode of knowledge production. *International Journal of Knowledge Management, 15*(2), 97–109.

Schott, J. R. (2005). *Matrix analysis for statistics*. Wiley.

Sengupta, D., & Jammalamadaka, S. R. (2003). *Linear models, an integrated approach.* World Scientific.

Stępniak, C. (1985). Ordering of positive semidefinite matrices with application to comparison of linear models. *Linear Algebra and its Applications, 70,* 67–71.

Iva Golubić is a lecturer of mathematics and statistics at Universitiy of Applied Sciences Velika Gorica and Algebra University College in Croatia. Her fields of interests are linear algebra and functional analysis with applications in linear models. *ig4290@student.uni-lj.si*

Janko Marovt Janko Marovt is an associate professor of mathematics at University of Maribor. His additional work and employment is at Institute of Mathematics, Physics, and Mechanics (IMFM), Ljubljana. His research interests are functional analysis and linear algebra with applications. He is an author or coauthor of more than 30 research papers. *janko.marovt@um.si*