Weak Matrix-Majorization

Francisco D. Martínez Pería, Pedro G. Massey and Luis E. Silvestre

30th April 2004

Francisco D. Martínez Pería,
Departamento de Matemática, FCE-UNLP, La Plata, Argentina
e-mail: francisco@mate.unlp.edu.ar

Pedro G. Massey,
Departamento de Matemática, FCE-UNLP, La Plata, Argentina
e-mail: massey@mate.unlp.edu.ar

Luis E. Silvestre,
Department of Mathematics, University of Texas at Austin, Austin, U.S.A.
e-mail: lsilvest@mail.ma.utexas.edu

Abstract

Given $X, Y \in \mathbb{R}^{n\times m}$ we introduce the following notion of matrix-majorization, called *weak matrix majorization*,

$$X \succ_w Y \text{ if there exists a row-stochastic matrix } A \in \mathbb{R}^{n\times n} \text{ such that } AX = Y,$$

and consider the relations between this definition, multivariate majorization ($\succ_s$) and directional majorization ($\succ$). It is verified that $\succ_s \Rightarrow \succ \Rightarrow \succ_w$, but none of the reciprocal implications is true. Nevertheless, we study the implication $\succ_w \Rightarrow \succ_s$ under additional hypotheses. We give characterizations of multivariate, directional and weak matrix-majorization in terms of convexity.

We also introduce definitions for majorization of Abelian families of selfadjoint $n \times n$-matrices, $\{X_i\}_{i=1,\dots,m}$ and $\{Y_i\}_{i=1,\dots,m}$, called *joint matrix majorizations*. They are induced by the previously mentioned matrix majorizations. We obtain descriptions of these relations in terms of convex functions.

Keywords: Vector majorization, multivariate and directional matrix-majorizations, majorization between selfadjoint matrices, convex functions.
AMS subject classifications: Primary 47A64, 15A60, 47B63.

1 Introduction

Vector majorization in $\mathbb{R}^n$ has been widely applied both in different branches of mathematics (matrix analysis, operator theory, statistics) and in other sciences like physics and economics. Also, different notions of matrix-majorization for real $n \times m$ matrices have been considered in some recent papers. Among them, we are interested in multivariate ($\succ_s$) and directional ($\succ$) majorization (see [4], [5], [7]). Given $X, Y \in M_{n,m}(\mathbb{R}) = M_{n,m}$ (the vector space of $n \times m$ real matrices) $X \succ_s Y$ if there exists a doubly-stochastic matrix $D \in \mathbb{R}^{n\times n}$ such that $DX = Y$; and $X \succ Y$ if the vector Xv majorizes $Yv \in \mathbb{R}^n$ for every $v \in \mathbb{R}^m$. In [5], G. Dahl gave a different concept of matrix-majorization. For two matrices X and Y having m rows, X majorizes Y (in Dahl's sense) if there is a row-stochastic matrix A such that $XA = Y$. In section 3 we introduce another related concept, *weak matrix majorization*: given $X, Y \in M_{n,m}$

$$X \succ_w Y \text{ if there exists a row-stochastic matrix } A \in \mathbb{R}^{n\times n} \text{ such that } AX = Y.$$

Although our definition of weak matrix-majorization resembles to Dahl's majorization, they are quite different concepts.

The main purpose of this work is to study the following items:

- A criterium to test directional majorization using convexity arguments;
- Conditions to ensure that directional (or weak) majorization implies multivariate majorization.
- Extend to the multivariate context the definition of majorization (to families of vectors and selfadjoint matrices).

We describe in some detail the contents of the paper. In Section 2 we introduce the notations used throughout this notes and state some results of majorization theory in $\mathbb{R}^n$ and nonnegative matrices.

Section 3 is devoted to the study of matrix-majorizations. In section 3.1 we prove that directional matrix-majorization implies weak matrix-majorization and we study some properties of the latter. It turns out that weak matrix-majorization has a simple geometrical interpretation: if we denote $R(X)$ the set of rows of X for $X \in M_{n,m}$, given $X, Y \in M_{n,m}$, $X \succ_w Y$ if and only if $\mathrm{co}(R(Y)) \subseteq \mathrm{co}(R(X))$, where $\mathrm{co}(S)$ is the convex hull of the set $S \subseteq \mathbb{R}^m$. Therefore weak matrix majorization is easy to test and this is one of its advantages. In section 3.2 we use elementary facts of convexity theory in order to obtain new characterizations of matrix majorizations. On one hand, based on the geometrical interpretation of weak matrix-majorization, we get the following alternative description of directional majorization:

Theorem. Let $X, Y \in M_{n,m}$. $X \succ Y$ if and only if, for all $k = 1, 2, \dots, [\frac{n}{2}]$ and $k = n$, the set of averages of k different rows of Y is included in the convex hull of the set of averages of k different rows of X.

This gives us a simple and effective criterium to determine whether $X \succ Y$. We use it along the paper to give some (counter) examples and to prove several remarks (see 3.11 and 3.18). On the other hand, we give a description of the different matrix-majorizations involving the comparison of traces of different families of matrices. For example, given $X, Y \in M_{n,m}$ then, $X \succ_s Y$ if and only if for every $Z \in M_{m,n}$ there exists a permutation matrix $P \in M_n$ such that $\mathrm{tr}(ZPX) \geq \mathrm{tr}(ZY)$. The same can be said for directional matrix-majorization if we restrict the set of matrices $Z \in M_{m,n}$ to those with $\mathrm{rank}(Z) = 1$. We end this section describing the minimal matrices with respect to the different matrix majorizations.

In general, weak matrix majorization is a weaker relation than directional matrix majorization, while the latter is weaker than multivariate majorization. In section 3.3 we study conditions in order to reverse these implications. This problem has been considered in several articles, for example [10], [11], [12], [7].

We prove that if $X \succ Y$ and the convex hull of $R(X)$ has only two extremal points, or if it is a non degenerate simplex, then $X \succ_s Y$. As a consequence $X \succ Y$ implies $X \succ_s Y$ if $X, Y \in M_{n,m}$ for $n \leq 3$. More generally, we get the following result, which is based on a result of Hwang and Pyo [7]:

Theorem. Let $X, Y \in M_{n,m}$ and suppose that $[Y, e][X, e]^\dagger$ has nonnegative entries. Then $X \succ_w Y$ and $e^t X = e^t Y$ if and only if $X \succ_s Y$.

In the previous theorem $[X, e] \in M_{n,(m+1)}$ denote the matrix whose first (ordered) m columns are equal to those of X and its last column is the vector $e = (1, 1, \dots, 1)^t$ and $X^\dagger$ denote the Moore-Penrose pseudo-inverse of the matrix $X \in M_{n,m}$.

In section 3.4 we give a description of the equivalence relations associated to matrix majorizations, i.e. what happens if $X \succ_s Y$ and $Y \succ_s X$ (respectively, $X \succ Y$ and $Y \succ X$; $X \succ_w Y$ and $Y \succ_w X$). It turns out that $X \succ_s Y$ and $Y \succ_s X$ is equivalent to $X \succ Y$ and $Y \succ X$, and they are also equivalent to the existence of a permutation matrix $Q \in M_n$ such that $Y = QX$. In general, the equivalence $X \succ_w Y$ and $Y \succ_w X$ simply says that $\text{co}(R(X)) = \text{co}(R(Y))$ but in the particular case in which X and Y are square invertible matrices $X \succ_w Y$ and $Y \succ_w X$ is equivalent to $X \succ_s Y$ and $Y \succ_s X$. We end this section considering the case of matrix-majorizations between square matrices.

In section 4 many of the results previously obtained are restated in a different context, which we call *joint (matrix or vector) majorizations*. In section 4.1 possible extensions of vector majorization in $\mathbb{R}^n$ are given for pairs of n-tuples of vectors in $\mathbb{R}^n$. In section 4.2 we introduce three different majorizations among Abelian families in $M_n(\mathbb{C})$, where an Abelian family in $M_n(\mathbb{C})$ is an ordered family of mutually commuting selfadjoint matrices in $M_n(\mathbb{C})$. These preorders are induced by the matrix majorizations already mentioned. Indeed, let $(a_i)_{i=1,\dots,m}$ and $(b_i)_{i=1,\dots,m}$ be two Abelian families, then there exist unitary matrices $U, V \in M_n(\mathbb{C})$ such that

$$U^* a_i U = D_{\lambda(a_i)}, \quad V^* b_i V = D_{\lambda(b_i)}, \quad i = 1, \dots, m,$$

where D_x denotes the diagonal matrix with main diagonal $x \in \mathbb{R}^n$. Let A, $B \in M_{n,m}(\mathbb{R})$ be the matrices with column-vectors

$$C_i(A) = \lambda(a_i), \qquad C_i(B) = \lambda(b_i), \qquad i = 1, \dots, m.$$

We say that the family $(a_i)_{i=1,\dots,m}$ *jointly weakly majorizes* (respectively *jointly strongly majorizes*, *jointly majorizes*) the family $(b_i)_{i=1,\dots,m}$ and write

$$(a_i)_{i=1,\dots,m} \succ_w (b_i)_{i=1,\dots,m}$$

(respectively $(a_i)_{i=1,\dots,m} \succ_s (b_i)_{i=1,\dots,m}$, $(a_i)_{i=1,\dots,m} \succ (b_i)_{i=1,\dots,m}$) if $A \succ_w B$ (respectively $A \succ_s B$, $A \succ B$). In section 4.3 and 4.4, based on results previously obtained, some characterizations of these relations are given in terms of convex functions and (multivariate) functional calculus. More explicitely, if $\sigma(a_1, \dots, a_m)$ denotes the set of rows of the the matrix A, recently constructed for a given Abelian family of selfadjoint matrices in $M_n(\mathbb{C})$ $(a_i)_{i=1,\dots,m}$, then we have:

Theorem. Let $(a_i)_{i=1,\dots,m}$ and $(b_i)_{i=1,\dots,m}$ be two Abelian families. Then,

1. $(a_i) \succ_w (b_i)$ if and only if, for every convex function $f : V \subset \mathbb{R}^m \to \mathbb{R}$ it holds that $\|f(a_1, \dots, a_m)\| \geq \|f(b_1, \dots, b_m)\|$;

2. $(a_i) \succ (b_i)$ if and only if, for every convex function $f : \mathbb{R} \to \mathbb{R}$ and $\gamma_1, \ldots, \gamma_m \in \mathbb{R}$ it holds $\operatorname{tr} f(\gamma_1 a_1 + \ldots + \gamma_m a_m) \geq \operatorname{tr} f(\gamma_1 b_1 + \ldots + \gamma_m b_m)$;

3. $(a_i) \succ_s (b_i)$ if and only if, for every convex function $f : V \subset \mathbb{R}^m \to \mathbb{R}$ it holds that $\operatorname{tr} f(a_1, \ldots, a_m) \geq \operatorname{tr} f(b_1, \ldots, b_m)$;

In items 1. and 3. above, $V \subseteq \mathbb{R}^m$ is a convex set containing $\sigma(a_1, \ldots, a_m) \cup \sigma(b_1, \ldots, b_m)$ and

$$f(a_1, \ldots, a_m) := U^* D U$$

where D is the diagonal matrix with main diagonal $D_{ii} = f(R_i(A))$ and U is a unitary matrix that mutually diagonalize the matrices a_i's (see subsection 4.4).

Finally, section 4.5 is devoted to the study of the equivalence relations induced by the joint majorizations.

2 Preliminaries

Notations. We denote by $M_{n,m} = M_{n,m}(\mathbb{R})$ (resp. $M_n = M_n(\mathbb{R})$) the vector space of $n \times m$ (resp. $n \times n$) matrices with real entries and $M_{n,m}(\mathbb{C})$ (resp. $M_n(\mathbb{C})$) the vector space of $n \times m$ (resp. $n \times n$) matrices with complex entries. The vectors in $\mathbb{R}^n$ are considered as column vectors and $e_1, \ldots, e_n \in \mathbb{R}^n$ denote the elements of the canonical basis. $GL(n)$ will denote the group of invertible $n \times n$ matrices (with real entries).

Given $S \subseteq \mathbb{R}^n$ we denote by $\operatorname{co}(S)$ the convex hull of S, i.e. the set of convex combinations of elements of S.

For $X \in M_{n,m}$, $R_i(X)$ (or shortly, X_i) denotes the i-th row of X and $C_i(X)$ is the i-th column of X. Also we will consider the sets of rows and columns of X

$$R(X) = \{R_i(X) \,:\, i = 1, \ldots, n\} \qquad \text{and} \qquad C(X) = \{C_i(X) \,:\, i = 1, \ldots, m\}.$$

Given $X \in M_{n,m}$, we denote by $X^t \in M_{m,n}$ the transpose of X and by $X^\dagger \in M_{m,n}$ the *Moore-Penrose* pseudoinverse of X. The dimension of the range of X is noted $\operatorname{rank}(X)$.

The group of permutations of order n is denoted by $\mathbb{S}_n$.

Nonnegative matrices. Let $A = (a_{ij}) \in M_{n,m}$. We say that A is nonnegative (resp. positive) if all $a_{ij} \geq 0$ (resp. $a_{ij} > 0$) and note it $A \geq 0$ (resp. $A > 0$). We shall use the following results regarding nonnegative matrices. In what follows $\rho(A)$ denotes the spectral radius of A.

Theorem (P1). Let $A \in M_n$ be a nonnegative matrix. Suppose that there exists $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ with $x_i > 0$, $i = 1, \ldots, n$, and such that $Ax = \lambda x$. Then $\lambda = \rho(A)$.

There is a well developed theory for *positive* matrices, initiated by O. Perron. Some of the results obtained by Perron also hold for the class of *primitive* matrices i.e, nonnegative matrices A for which there exists $k \in \mathbb{N}$ such that $A^k > 0$. In particular we have the following

Theorem (P2). Let $A \in M_n$ be a nonnegative primitive matrix. Then

$$\lim_{m \to \infty} (\rho(A)^{-1} A)^m = L$$

where $L = xy^t$, with $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$ such that $x_i, y_i > 0$, $i = 1, \ldots, n$ $Ax = \rho(A)x$, $A^t y = \rho(A)y$ and $x^t y = 1$.

A nonnegative matrix $A \in M_n$ with the property that all its row sums are $+1$ is said to be *row-stochastic*. If we denote by $e \in \mathbb{R}^n$ the vector with all components $+1$, the set of row-stochastic matrices in M_n is a compact convex set that can be characterized as follows

$$RS(n) = \{A \in M_n : A \geq 0, Ae = e\}.$$

A row-stochastic matrix $A \in M_n$ with the property that A^t is also row-stochastic is said to be *doubly-stochastic*. The set of doubly-stochastic matrices is also a compact convex set in M_n and is characterized by

$$DS(n) = \{A \in M_n : A \geq 0, Ae = e, A^t e = e\}.$$

The group of permutation matrices in M_n is contained in $DS(n)$. Birkhoff's theorem shows that these are the extremal points of the set of doubly-stochastic matrices.

Theorem (Birkhoff). $A \in M_n$ is a doubly-stochastic matrix if and only if, for some $k \in \mathbb{N}$, there are permutation matrices $P_1, \ldots, P_k \in M_n$ and positive scalars $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ such that $\alpha_1 + \ldots + \alpha_k = 1$ and

$$A = \sum_{j=1}^{k} \alpha_j P_j.$$

Vector majorization. If $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, denote by $x^{\uparrow}$ and $x^{\downarrow}$ the vectors obtained by rearranging the entries of x in increasing and decreasing order, respectively. Given two vectors $x, y \in \mathbb{R}^n$, we say that x *mayorizes* y, and denote it $x \succ y$, if

$$\sum_{i=1}^{k} x_i^{\downarrow} \geq \sum_{i=1}^{k} y_i^{\downarrow} \qquad k = 1, \ldots, n-1 \qquad \text{and} \qquad \sum_{k=1}^{n} x_i = \sum_{k=1}^{n} y_i. \tag{2.1}$$

The next theorem shows some known characterizations of majorization (see, for example, Bhatia's book [3]).

Theorem (P3). Let $x, y \in \mathbb{R}^n$. The following are equivalent:

1. $x \succ y$;
2. $\sum_{i=1}^{n} |x_i - t| \geq \sum_{i=1}^{n} |y_i - t|$, for all $t \in \mathbb{R}$;
3. y belongs to the convex hull of the vectors obtained by permuting the entries of x;
4. there exists a doubly-stochastic $n \times n$ matrix A such that $y = Ax$.

3 Matrix-Majorizations

Given two matrices $X, Y \in M_{n,m}$ we recall some known matrix-majorizations definitions:

- Y is *strongly majorized* by X, noted $X \succ_s Y$, if there exists $A \in DS(n)$ such that $AX = Y$.
- Y is *directionally majorized* by X, noted $X \succ Y$, if for all $v \in \mathbb{R}^m$, $Xv \succ Yv$.

In some previous works (see [2], [7], [9]) strong majorization is called multivariate majorization. Directional majorization has been considered in [7], [9] and [12], for example.

3.1 Weak matrix-majorization

We introduce the following notion of matrix majorization.

Definition. Given two matrices $X, Y \in M_{n,m}$ we say that Y is *weakly majorized* by X, and write $X \succ_w Y$, if there exists $A \in RS(n)$ such that $AX = Y$.

Remark 3.1. When $X, Y \in M_{n,1}$, i.e. X and Y are $\mathbb{R}^n$ vectors, strong and directional matrix-majorization definitions are the same that vector-majorization. But weak majorization $X \succ_w Y$ for $n \times 1$ matrices is equivalent to:

$$\min_{1\leq j\leq n} x_j \leq \min_{1\leq j\leq n} y_j \leq \max_{1\leq j\leq n} y_j \leq \max_{1\leq j\leq n} x_j.$$

$\triangle$

It is clear that strong majorization implies directional majorization. Next we give a characterization of weak majorization and use it to prove that directional majorization implies weak majorization.

Proposition 3.2. *Let $X, Y \in M_{n,m}$. Then,*

(i) $X \succ_w Y$ if and only if $R(Y) \subseteq \mathrm{co}(R(X))$;

(ii) if $X \succ Y$ then $X \succ_w Y$.

Proof.

(i) Let $X, Y \in M_{n,m}$ and $A \in M_n$. Then $AX = Y$ if and only if

$$R_i(Y) = \sum_{k=1}^{n} a_{ik} R_k(X) \quad i = 1, \ldots, n.$$

Therefore, if there exists $A \in RS(n)$ such that $AX = Y$ then $R(Y) \subseteq \mathrm{co}(R(X))$. On the other hand, if $R(Y) \subseteq \mathrm{co}(R(X))$ then, by the equation above, we can construct the rows of a matrix $A \in RS(n)$ such that $AX = Y$.

(ii) Let $X, Y \in M_{n,m}$ such that $X \succ Y$ and suppose that exists $1 \leq i \leq n$ such that $R_i(Y) \notin \mathrm{co}(R(X))$. Then, there exists an hyperplane which separates $R_i(Y)$ of $\mathrm{co}(R(X))$, that is, there exist $v \in \mathbb{R}^m$ and $t > 0$ such that

$$\langle R_i(Y); v\rangle \geq t \quad \text{and} \quad \langle R_j(X); v\rangle < t \;\text{ for all } j = 1, \ldots, n.$$

But this contradicts the vector-majorization $Xv \succ Yv$ because

$$((Yv)^{\downarrow})_1 \geq (Yv)_i = \langle R_i(Y); v\rangle \geq t > ((Xv)^{\downarrow})_1.$$

Therefore, $X \succ_w Y$.

$\square$

Corollary 3.3. *Let $X, Y \in M_{n,m}$. $X \succ_w Y$ if and only if*

$$\max_{1\leq i\leq n} f(X_i) \geq \max_{1\leq i\leq n} f(Y_i)$$

for all convex function $f : V \to \mathbb{R}$ where $V \subseteq \mathbb{R}^m$ is a convex set containing $R(X) \cup R(Y)$.

Proof. It is a consequence of the following remark: given two convex sets V_1 and V_2, $V_1 \subset V_2$ if and only if

$$\max_{x \in V_1} f(x) \leq \max_{x \in V_2} f(x)$$

for every convex function f defined over $V_1 \cup V_2$.

$\square$

The following examples show that, in general, there are no equivalence among the different matrix-majorizations.

Example 1. $X \succ_w Y$ does not imply $X \succ Y$.

Let

$$X = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad Y = \begin{pmatrix} 1 & 0 \\ 1/2 & 1/2 \end{pmatrix}.$$

Then, if we take $A = Y \in RS(n)$, it is clear that $AX = Y$. Therefore $X \succ_w Y$. On the other hand, if we consider $v = (2,1)^t$ then $Xv \not\succ Yv$. So that, $X \not\succ Y$.

$\triangle$

Example 2. $X \succ Y$ does not imply $X \succ_s Y$.

It is a known fact. Indeed, there is an example in [11] due to A. Horn. Our example uses smaller matrices. Actually, we shall see in Corollary 3.21 that this is the minimum number of rows and columns required to lack the implication. Let

$$X = \begin{pmatrix} 0 & 0 \\ 3 & -2 \\ -3 & -2 \\ 0 & 4 \end{pmatrix} \qquad \text{and} \qquad Y = \begin{pmatrix} 2 & 0 \\ -2 & 0 \\ 0 & -2 \\ 0 & 2 \end{pmatrix}.$$

Then $X \succ Y$ but $X \not\succ_s Y$. The proof of this fact will be given in Remark 3.11.

$\triangle$

In the next Proposition we state several elementary properties of weak matrix-majorization. Note that all of them also hold for directional and strong majorization. The proof is omitted, it only requires elementary arguments.

Proposition 3.4. *Let $X, Y, Z \in M_{n,m}$. Then,*

1. $X \succ_w X$.
2. *If $X \succ_w Y$ and $Y \succ_w Z$ then $X \succ_w Z$.*
3. *If $X \succ_w Y$ then $X[I] \succ_w Y[I]$ for each $I \subset \{1, \dots, m\}$, where $X[I]$ is the submatrix of X whose columns are the columns of X indexed by the elements in I.*
4. *If $X \succ_w Y$ and $R \in M_{m,p}$ then $XR \succ_w YR$.*
5. *If $X \succ_w Y$ and $P, Q \in M_n$ are permutation matrices, then $PX \succ_w QY$.*
6. *If $X \succ_w Y$ then* $\operatorname{rank}(X) \geq \operatorname{rank}(Y)$.

Proposition 3.5. *Let $X, Y \in M_{n,m}$ and suppose that* $\operatorname{rank}(X) = n$. *The following are equivalent:*

(i) $X \succ_w Y$.

(ii) $YX^\dagger \in RS(n)$ *and* $\ker(X) \subseteq \ker(Y)$.

Proof. Suppose that $X \succ_w Y$. Since $\text{rank}(X) = n$, then $XX^\dagger = I_n$. Hence, $X \succ_w Y$ implies that $I_n = XX^\dagger \succ_w YX^\dagger$. In other words, there exists a matrix $A \in RS(n)$ such that $A = AI_n = YX^\dagger \in RS(n)$. The equation $Y = AX$ clearly implies that $\ker(X) \subseteq \ker(Y)$.

Conversely, if $YX^\dagger \in RS(n)$ and $\ker(X) \subseteq \ker(Y)$, then $X^\dagger X$ is the orthogonal projection onto $\ker X^\perp \supseteq \ker Y^\perp$ and $YX^\dagger X = Y$. Hence $X \succ_w Y$. $\square$

Corollary 3.6. *Suppose that* $X, Y \in M_n$ *and* $X \in GL(n)$. *Then,* $X \succ_w Y$ *if and only if* $YX^{-1} \in RS(n)$.

3.2 Convexity and matrix majorizations

We begin this section recalling a well known characterization of strong majorization in terms of convex functions. A proof of this result can be found in [5].

Proposition 3.7. *Let* $X, Y \in M_{n,m}$. *Then* $X \succ_s Y$ *if and only if, for every convex function* $f : V \to \mathbb{R}$, *it holds*

$$\sum_{j=1}^{n} f(X_j) \geq \sum_{j=1}^{n} f(Y_j)$$

where $V \subseteq \mathbb{R}^m$ *is a convex set such that* $R(X) \cup R(Y) \subseteq V$.

In Proposition 3.2 we proved that $X \succ_w Y$ if and only if the convex hull of $R(X)$ includes the convex hull of $R(Y)$. The following Theorem characterizes directional majorization for matrices in $M_{n,m}$, in terms of the convex hulls of $[\frac{n}{2}] + 1$ sets, where $[r]$ is the greater integer less than r, $r \in \mathbb{R}$.

Remark 3.8. Recall that, given z, $w_i \in \mathbb{R}^m$ with $i = 1, \dots, n$, then

$$z \in \text{co}(\{w_i : i = 1, \dots, n\}) \text{ if and only if } \max_{1 \leq i \leq n} \langle w_i; v \rangle \geq \langle z; v \rangle \quad \text{for all } v \in \mathbb{R}^m.$$

$\triangle$

Theorem 3.9. *Let* $X, Y \in M_{n,m}$. $X \succ Y$ *if and only if, for all* $k = 1, 2, \dots, [\frac{n}{2}]$ *and* $k = n$, *the set of averages of* k *different rows of* Y *is included in the convex hull of the set of averages of* k *different rows of* X.

Proof. Let $X, Y \in M_{n,m}$, and suppose that the set of averages of k different rows of Y is included in the convex hull of the set of averages of k different rows of X. Let $v \in \mathbb{R}^m$, and for $k = 1, \dots, [\frac{n}{2}]$ we have,

$$\begin{aligned}
\sum_{j=1}^{k} (Yv)_j^{\downarrow} &= k \left\langle \frac{1}{k} \sum_{j=1}^{k} Y_{\sigma(j)}; v \right\rangle = k \left\langle \sum_{\mu \in \mathbb{S}_n} c_\mu \left(\frac{1}{k} \sum_{j=1}^{k} X_{\mu(j)} \right); v \right\rangle = \\
&= k \sum_{\mu \in \mathbb{S}_n} c_\mu \left\langle \frac{1}{k} \sum_{j=1}^{k} X_{\mu(j)}; v \right\rangle \leq k \max_{\mu \in \mathbb{S}_n} \left\langle \frac{1}{k} \sum_{j=1}^{k} X_{\mu(j)}; v \right\rangle = \\
&= \sum_{j=1}^{k} (Xv)_j^{\downarrow},
\end{aligned}$$

where $\sigma \in \mathbb{S}_n$ is a permutation such that the coordinates of Yv are arranged in non-increasing order. Moreover the hypothesis for $k = n$ implies

$$\sum_{j=1}^{n} Y_j = \sum_{j=1}^{n} X_j,$$

so that, for $[\frac{n}{2}] < k \leq n$,

$$\begin{aligned}
\sum_{j=1}^{k} (Yv)^{\downarrow}_j &= \left\langle \sum_{j=1}^{n} Y_j; v \right\rangle - \sum_{j=1}^{n-k} \langle Y_{\tau(j)}; v \rangle = \\
&= \left\langle \sum_{j=1}^{n} X_j; v \right\rangle - (n-k) \left\langle \sum_{\mu \in \mathbb{S}_n} c_\mu \left(\frac{1}{n-k} \sum_{j=1}^{n-k} X_{\mu(j)} \right); v \right\rangle \leq \\
&\leq \left\langle \sum_{j=1}^{n} X_j; v \right\rangle - (n-k) \min_{\mu \in \mathbb{S}_n} \left\langle \frac{1}{n-k} \sum_{j=1}^{n-k} X_{\mu(j)}; v \right\rangle = \\
&= \sum_{j=1}^{k} (Xv)^{\downarrow}_j,
\end{aligned}$$

where $\tau \in \mathbb{S}_n$ is a permutation such that the coordinates of Yv are arranged in non-decreasing order. Therefore, $Xv \succ Yv$ and since $v \in \mathbb{R}^m$ was arbitrary then $X \succ Y$.

On the other hand, let us suppose that $X \succ Y$. Given $\mu \in \mathbb{S}_n$,

$$\max_{\sigma \in \mathbb{S}_n} \left\langle \sum_{i=1}^{k} X_{\sigma(i)}; v \right\rangle = \sum_{i=1}^{k} (Xv)^{\downarrow}_i \geq \sum_{i=1}^{k} (Yv)^{\downarrow}_i \geq \left\langle \sum_{i=1}^{k} Y_{\mu(i)}; v \right\rangle \qquad \text{for all } v \in \mathbb{R}^n.$$

Therefore, by Remark 3.8, we have that $\frac{1}{k} \sum_{i=1}^{k} Y_{\mu(i)}$ belongs to the convex hull of $\{\frac{1}{k} \sum_{i=1}^{k} X_{\sigma(i)} : \sigma \in \mathbb{S}_n\}$.

□

The previous Proposition implies the following description of directional majorization in terms of weak majorization.

Corollary 3.10. *Let $X, Y \in M_{n,m}$. $X \succ Y$ if and only if $\overline{X}(k) \succ_w \overline{Y}(k)$ for $k = 1, \ldots, [\frac{n}{2}]$ and $k = n$, where $\overline{X}(k)$ (respectively $\overline{Y}(k)$) is the matrix of $\frac{n!}{k!(n-k)!}$ rows, which are all possible averages of k different rows of X (respectively of Y).*

Note that, in particular, the Corollary proves that if $X \succ Y$ then $X \succ_w Y$.

Remark 3.11. Corollary 3.10 gives a simple way to determine whether $X \succ Y$. Indeed, if X, Y denote the matrices in Example 2, then the following graphics show that $X \succ Y$.

In fact, we only have to verify that $\overline{X}(k) \succ_w \overline{Y}(k)$ for $k = 1, 2, 4$. In first place, $\overline{X}(4) = (0, 0) = \overline{Y}(4) \in M_{1,2}$, so that, $\overline{X}(4) \succ_w \overline{Y}(4)$. Moreover,

$$\overline{X}(2) = \begin{pmatrix} 3/2 & -1 \\ -3/2 & -1 \\ 0 & 2 \\ 0 & -2 \\ 3/2 & 1 \\ -3/2 & 1 \end{pmatrix} \qquad \text{and} \qquad \overline{Y}(2) = \begin{pmatrix} 0 & 0 \\ 1 & -1 \\ 1 & 1 \\ -1 & -1 \\ -1 & 1 \\ 0 & 0 \end{pmatrix}.$$

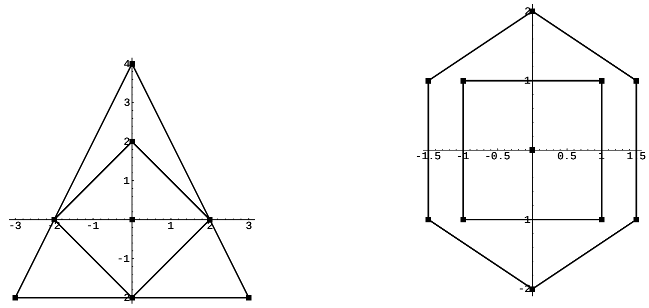

Figure 1: Polygons corresponding to $k = 1$ and $k = 2$

Then, the graphics in Figure 1 show the inclusion of the polygons that prove $\overline{X}(k) \succ_w \overline{Y}(k)$ for $k = 1, 2$. Therefore $X \succ Y$.

On the other hand, the convex function $f(x, y) = \max\{-y, \frac{y}{2} + x, \frac{y}{2} - x\}$ and proposition 3.7 show that $X \not\succ_s Y$ in Example 2.

$\triangle$

The next theorem gives characterizations of strong, directional and weak matrix-majorization comparing the traces of certain matrices.

Theorem 3.12. *Let $X, Y \in M_{n,m}$. Then,*

1. *$X \succ_s Y$ if and only if for every $Z \in M_{m,n}$ there exists a permutation matrix $P \in M_n$ such that*

$$\operatorname{tr}(ZPX) \geq \operatorname{tr}(ZY).$$

2. *$X \succ Y$ if and only if, for every $Z \in M_{m,n}$ with $\operatorname{rank}(Z) = 1$, there exists a permutation matrix $P \in M_n$ such that*

$$\operatorname{tr}(ZPX) \geq \operatorname{tr}(ZY).$$

3. *$X \succ_w Y$ if and only if for every $w \in \mathbb{R}^m$ and every $1 \leq i \leq n$, there exists a permutation matrix $P \in M_n$ such that*

$$\operatorname{tr}(we_i^t PX) \geq \operatorname{tr}(we_i^t Y)$$

Proof. To prove 1. recall first that $M_{n,m}$ with the inner product given by $\langle X; Y\rangle = \operatorname{tr}(Y^t X)$ can be identified with $\mathbb{R}^{n.m}$ with the usual inner product.

By Birkhoff's theorem $X \succ_s Y$ is equivalent to the fact that Y belongs to the convex hull of the set $\{PX : P \text{ is a permutation matrix in } M_n\}$. By Remark 3.8 this is equivalent to the following: for every $Z \in M_{m,n}$ there exists a permutation matrix $P \in M_n$ such that

$$\operatorname{tr}(ZY) = \langle Y; Z^t\rangle \leq \langle PX; Z^t\rangle = \operatorname{tr}(ZPX).$$

To prove 2. note that, given $v \in \mathbb{R}^m$ then, $Xv \succ Yv$ is equivalent to the fact that Yv belongs to the convex hull of the set $\{PXv : P \text{ is a permutation matrix in } M_n\}$. By Remark 3.8 this is equivalent to the following: for every $w \in \mathbb{R}^n$ there exists a permutation matrix $P \in M_n$ such that $\langle w; PXv \rangle \geq \langle w; Yv \rangle$. Then we have

$$\operatorname{tr}(vw^t PX) = \operatorname{tr}(w^t PXv) = \langle w; PXv \rangle \geq \langle w; Yv \rangle = \operatorname{tr}(vw^t Y).$$

Since every rank one matrix $Z \in M_{m,n}$ can be expressed as $Z = vw^t$ for $v \in \mathbb{R}^m$, $w \in \mathbb{R}^n$ we are done.

Item 3 follows in the same way. Recall that $X \succ_w Y$ is equivalent to $Y_j \in \operatorname{co}(R(X))$ for every $1 \leq j \leq n$ and note that $Y_j = Y^t e_j$. Then, by Remark 3.8, this is equivalent to the following: for every $w \in \mathbb{R}^n$ and every $1 \leq j \leq n$ there exists a permutation matrix $Q \in M_n$ such that $\left\langle w; QX^t e_j \right\rangle \geq \left\langle w; Y^t e_j \right\rangle$. So we have

$$\operatorname{tr}(we_j^t Q^t X) = \left\langle w; QX^t e_j \right\rangle \geq \left\langle w; Y^t e_j \right\rangle = \operatorname{tr}(we_j^t Y),$$

for every $w \in \mathbb{R}^n$. Taking $P = Q^t$ we have the desired result.

□

As a consequence of the geometric interpretation of weak matrix-majorization, we can determine the minimal matrices with respect to the preorders that we have considered so far. In this context, a *minimal* element with respect to a preorder relation $\ll$ in a set P is an element $m \in P$ such that if $n \ll m$ then $m \ll n$.

Proposition 3.13. *$X \in M_{n,m}$ is minimal with respect to any of the preorder $\succ_w$, $\succ$ or $\succ_s$ if and only if $X_1 = \ldots = X_n$, that is, all the rows of X coincide.*

Proof. If $R(X) = \{v\}$, for $v \in \mathbb{R}^m$, then $\operatorname{co}(R(X)) = \{v\}$. Then, if $X \succ_w Y$ it is clear that $X = Y$.

Let $X \in M_{n,m}$ be a matrix with at least to different rows. Then $R(X)$ contains two different points (of $\mathbb{R}^m$). If $D \in DS(n)$ is the matrix with every entry equal to $1/n$ we have that $Y = DX \prec_s X$. Moreover, since $R_1(Y) = R_2(Y) = \ldots = R_n(Y)$, then $\operatorname{co}(R(Y))$ contains only one point, so $R(X) \not\subset \operatorname{co}(R(Y))$. Therefore $Y \not\succ_w X$ and X is not minimal with respect to any of the matrix-majorizations.

□

Given $X \in M_{n,m}$ consider the set $S_X = \operatorname{co}\{R(X)\}$ and notice that:

1. If $X \succ_w Y$ and Y is minimal then $R(Y) = \{v\}$ for some $v \in S_X$. Moreover, for every $v \in S_X$, the matrix $R = ew^t \in RS(n)$ is such that $Y = RX$, where $w \in \mathbb{R}^n$ is the vector that satisfies $w \geq 0$, $\sum_{i=1}^n w_i = 1$ and $\sum_{i=1}^n w_i R_i(X) = v$.

2. If $X \succ Y$ (resp. $X \succ_s Y$) and Y is minimal then $R(Y) = \{v_0\}$, where $v_0 = \frac{X_1 + \ldots + X_n}{n}$. In particular, if every row of X is an extremal of S_X then v_0 is the centroid of S_X.

Note that 1. is a consequence of Proposition 3.2 and 2. follows from Theorem 3.9.

Remark 3.14. Let $R \in RS(n)$ be a primitive matrix, i.e there exists $h \geq 1$ such that $R^h > 0$. Then, for every $X \in M_{n,m}$, $\{R^k X\}_{k \geq 1}$ is a decreasing sequence in $M_{n,m}$ (in weak majorization sense) that converges to a minimal matrix Y such that $R(Y) = \{X^t y\}$, where $y \in \mathbb{R}^n$ satisfies $R^t y = y$, $y > 0$ and $e^t y = 1$. This is a consequence of Theorem (P2) in the preliminaries.

△

3.3 When weak majorization implies strong majorization

In some particular cases directional matrix-majorization implies strong matrix-majorization.

Proposition 3.15. *Let $X, Y \in M_{n,m}$ such that $X \succ Y$. Suppose that* $\mathrm{co}(R(X))$ *has only two extremal points. Then $X \succ_s Y$.*

Proof. Note that, as $\mathrm{co}(R(X))$ has only two extremals, the points in $R(X)$ are contained in a line of $\mathbb{R}^m$. Then, the points of $R(Y)$ also belongs to this line. Let $Z \in M_{n,m}$ be the matrix whose rows are all equal to $R_1(X)$. It is easy to see that $X \succ Y$ (resp. $X \succ_s Y$) if and only if $X - Z \succ Y - Z$ (resp. $X - Z \succ_s Y - Z$).

Therefore, we can suppose that $\operatorname{rank} X \leq 1$ and $\operatorname{rank} Y \leq 1$. If $X = 0$ the result is apparent. If $Y = 0$ and $\operatorname{rank} X = 1$ consider the matrix $D \in DS(n)$ such that $D(Xe_1) = Ye_1 = 0$, then we have that $DX = 0 = Y$ since every column of X is a real multiple of $X_1 = Xe_1$. If $\operatorname{rank} Y = \operatorname{rank} X = 1$, let $x_1, y_1 \in \mathbb{R}^n$ and $x_2, y_2 \in \mathbb{R}^m$ such that $X = x_1 x_2^t$ and $Y = y_1 y_2^t$. Moreover, since $\mathbb{C}y_2 = \mathrm{ran}(Y^t) = \mathrm{ran}(X^t) = \mathbb{C}x_2$, we may assume that $y_2 = x_2$. Note that $Xx_2 = \|x_2\|^2 x_1$ and $Yx_2 = \|x_2\|^2 y_1$.

Since $X \succ Y$, then $x_1 \succ y_1$ and there exists $D \in DS(n)$ such that $Dx_1 = y_1$. Hence

$$DX = Dx_1 x_2^t = y_1 x_2^t = Y$$

and $X \succ_s Y$. □

Given $X \in M_{n,m}$ we will denote $[X, e] \in M_{n,(m+1)}$ to the matrix whose first (ordered) m columns are equal to those of X and its last column is the vector e. In [7], S.-G. Hwang and S.-S. Pyo proved the following theorem.

Theorem. Let X, $Y \in M_{n,m}$ be such that $[Y, e][X, e]^\dagger$ has non-negative entries. Then $X \succ Y$ if and only if $X \succ_s Y$.

We extend this result by replacing $X \succ Y$ by $X \succ_w Y$ plus $e^t X = e^t Y$. Note that if $X \succ Y$ then $X \succ_w Y$ and $e^t X = e^t Y$, but the other implication is not true (see Remark 3.18). Moreover, using the notion of weak matrix-majorization we give a simpler proof.

Theorem 3.16. *Let $X, Y \in M_{n,m}$ and suppose that $[Y, e][X, e]^\dagger$ has nonnegative entries. If $X \succ_w Y$ and $e^t X = e^t Y$ then $X \succ_s Y$.*

In order to prove this theorem we are going to use the following lemma. The proof is straightforward from the definitions.

Lemma 3.17. *Let $X, Y \in M_{n,m}$ then*

$$X \succ_w Y \text{ if and only if } [X, e] \succ_w [Y, e]$$
$$X \succ Y \text{ if and only if } [X, e] \succ [Y, e]$$
$$X \succ_s Y \text{ if and only if } [X, e] \succ_s [Y, e]$$
$$e^t X = e^t Y \text{ if and only if } e^t [X, e] = e^t [Y, e]$$

Proof of Theorem 3.16. Let $X_1 = [X, e]$ and $Y_1 = [Y, e]$. Applying lemma 3.17 we only have to prove that if $Y_1 X_1^\dagger$ has nonnegative entries, then $X_1 \succ_w Y_1$ and $e^t X_1 = e^t Y_1$ implies $X_1 \succ_s Y_1$.

Suppose $X_1 \succ_w Y_1$, then there exists a row-stochastic matrix A such that $Y_1 = AX_1$. Multiplying both sides of the equation by $X_1^\dagger$ we obtain:

$$Y_1 X_1^\dagger = AX_1 X_1^\dagger = AP$$

where P is the orthogonal projection onto the range or X_1.

Since $APX_1 = AX_1 = Y_1$, we will conclude that $X_1 \succ_s Y_1$ as soon as we prove that AP is doubly-stochastic. We know by hypothesis that $AP = Y_1 X_1^\dagger$ has nonnegative entries. We are left to show that $APe = e$ and $e^t AP = e^t$.

Since we chose $X_1 = [X, e]$, then e is in the image of X_1 and $Pe = e$. Therefore:

$$APe = Ae = e$$

since A is row-stochastic. By hypothesis, $e^t X_1 = e^t Y_1$, then

$$e^t AP = e^t Y_1 X_1^\dagger = e^t X_1 X_1^\dagger = e^t P = e^t$$

So AP is doubly-stochastic, $(AP)X_1 = Y_1$, and also $(AP)X = Y$.

$\square$

Remark 3.18. The condition $X \succ_w Y$ and $e^t X = e^t Y$ of Theorem 3.16 is weaker than the hypothesis $X \succ Y$ of Hwang-Pyo's theorem. In fact, let $X, Y \in M_{6,2}$ be given by

$$X = \begin{pmatrix} 0 & 1 & 1 & -1 & -1 & 0 \\ 1 & 1 & -1 & -1 & 1 & 1 \end{pmatrix}^t \quad \text{and} \quad Y = \begin{pmatrix} 2/3 & 2/3 & 1 & -1 & -2/3 & -2/3 \\ 1 & 1 & -1 & -1 & 1 & 1 \end{pmatrix}^t.$$

It is easy to show that $X \succ_w Y$ and $e^t X = (0, 2) = e^t Y$. However, Figure 2 shows that $\overline{X}(2) \not\succ_w \overline{Y}(2)$ (where ■ represents the rows of $\overline{X}(2)$ and ▲ represents the rows of $\overline{Y}(2)$). Thus, by Corollary 3.10 $X \not\succ Y$.

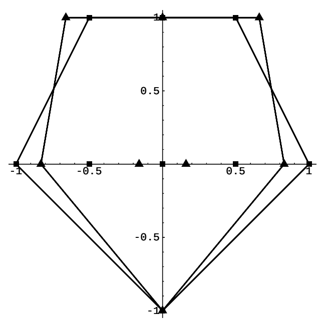

Figure 2: $\mathrm{co}(\overline{Y}(2)) \not\subseteq \mathrm{co}(\overline{X}(2))$

$\triangle$

Corollary 3.19. *Let $X, Y \in M_{n,m}$ and suppose that* $\mathrm{ran}([X, e]) = \mathbb{R}^n$. *If $X \succ_w Y$ and $e^t X = e^t Y$ then $X \succ_s Y$.*

Proof. It is a consequence of Proposition 3.5 and Theorem 3.16. $\square$

Corollary 3.20. *Let $X, Y \in M_{n,m}$ such that the rows of X generate a simplex whose dimension equals $n-1$. If $X \succ_w Y$ and $e^t X = e^t Y$ then $X \succ_s Y$.*

Proof. The fact that the rows of X generate a simplex whose dimension equals $n-1$ is equivalent to say that the set $\{R_2(X) - R_1(X), \ldots, R_n(X) - R_1(X)\}$ is linearly independent. Then, the rank of the matrix

$$Z = \begin{pmatrix} 0 \\ R_2(X) - R_1(X) \\ \vdots \\ R_n(X) - R_1(X) \end{pmatrix}$$

is $n-1$. Therefore the subspace $\mathcal{S}$ spanned by the columns of Z has also dimension $n-1$ and $e \notin \mathcal{S}$. Using that $C_i(Z) = C_i(X) - x_{1i}e$, $1 \le i \le m$, we conclude that the set $\{C_1(X), \dots, C_m(X), e\}$ span $\mathbb{R}^n$. Now, using Corollary 3.19, we get $X \succ_s Y$.

$\square$

Corollary 3.21. *Let $X, Y \in M_{n,m}$ with $1 \le n \le 3$. Then, $X \succ Y$ implies that $X \succ_s Y$.*

Proof. Let $X, Y \in M_{n,m}$, with $1 \le n \le 3$, such that $X \succ Y$. If the convex $\mathrm{co}(R(X))$ is a segment, follows from Proposition 3.15. Otherwise $n = 3$ and we have that $\mathrm{co}(R(X))$ is the triangle contained in $\mathbb{R}^m$ with vertexes $X_i = R_i(X)$, $i = 1, 2, 3$, so we can apply Corollary 3.20.

$\square$

3.4 Equivalence relations associated to matrix-majorizations

As we have already mentioned, matrix-majorizations considered so far are preorder relations. Since $X \succ_w Y$ if and only if $R(Y) \subseteq \mathrm{co}(R(X))$, it is clear that the relation $X \succ_w Y$ and $Y \succ_w X$ is equivalent to $\mathrm{co}(R(X)) = \mathrm{co}(R(Y))$. The next Theorem describes the equivalence relation associated to directional and strong matrix-majorization.

Theorem 3.22. *Let X, $Y \in M_{n,m}$. Then the following are equivalent*

i) There exists a permutation matrix $Q \in M_n$ such that $QX = Y$.

ii) $X \succ_s Y$ and $Y \succ_s X$.

iii) $X \succ Y$ and $Y \succ X$.

Before proving this, we consider the following result.

Lemma 3.23. *Let $X, Y \in M_{n,m}$ be such that $X \succ Y$ and $R_{i_0}(X) = R_{j_0}(Y)$. Let $\tilde{X} \in M_{(n-1),m}$ (respectively $\tilde{Y} \in M_{(n-1),m}$) denote the matrix obtained by deleting the i_0-th row from X (respectively the j_0-th row from Y). Then $\tilde{X} \succ \tilde{Y}$.*

Proof. It follows from the following fact: if $x, y \in \mathbb{R}^r$ then, for every $\lambda \in \mathbb{R}$,

$$x \succ y \iff (x_1, \dots, x_r, \lambda) \succ (y_1, \dots, y_r, \lambda).$$

$\square$

Proof of Theorem 3.22. The implications $i) \Rightarrow ii) \Rightarrow iii)$ are clear. So we only have to prove the implication $iii) \Rightarrow i)$. We use induction on the number of rows of X and Y. If $n = 1$ it is trivial, note that if $X, Y \in M_{1,m}$ then $X \succ Y$ implies $X = Y$.

In case that $n > 1$, note that if $X \succ Y$ and $Y \succ X$ then, $X \succ_w Y$ and $Y \succ_w X$. Therefore the convex hull of $R(X)$ coincides with that of $R(Y)$ and in particular they have the same extremal points. If z is an extremal point of $\mathrm{co}(R(X)) = \mathrm{co}(R(Y))$ then, $z = R_{i_0}(X) = R_{j_0}(Y)$ with $1 \le i_0$, $j_0 \le n$.

If $\tilde{X}$, $\tilde{Y} \in M_{(n-1),m}$ are as in the Lemma, then it holds that $\tilde{X} \succ \tilde{Y}$ and $\tilde{Y} \succ \tilde{X}$. By the inductive hypothesis the rows of $\tilde{X}$ are a reordering of the rows of $\tilde{Y}$. Therefore the rows of X are a reordering of the rows of Y, which implies $i)$.

$\square$

There is more to say about weak matrix-majorization when $X, Y \in M_n$. Most of all in case that $X, Y \in GL(n)$.

Proposition 3.24. *Let $X, Y \in M_n$. If $X \succ_w Y$ then $|\det(X)| \geq |\det(Y)|$. Moreover, if $X \succ_w Y$ and $|\det(X)| = |\det(Y)| \neq 0$ then there exists a permutation matrix $P \in M_n$ such that $Y = PX$.*

Proof. Let $S_X = \mathrm{co}(R(X) \cup \{0\})$ (resp. $S_Y = \mathrm{co}(R(Y) \cup \{0\})$) be the simplex generated by $R(X) \cup \{0\}$ (resp. $R(Y) \cup \{0\}$). $|\det(X)|$ and $|\det(Y)|$ are the volumes of S_X and S_Y, respectively. If $X \succ_w Y$ we have, by Proposition 3.2, that $S_Y \subseteq S_X$, so $|\det(X)| \geq |\det(Y)|$.

Suppose that $X \succ_w Y$ and $|\det(X)| = |\det(Y)| \neq 0$. $|\det(X)| = |\det(Y)| \neq 0$ implies that S_X and S_Y are non-degenerate simplexes and $R(X) \cup \{0\}$ (resp. $R(Y) \cup \{0\}$) are the $n+1$ vertexes of S_X (resp. S_Y), and $X \succ_w Y$ says that $S_Y \subseteq S_X$. But $|\det(X)| = |\det(Y)|$ means that both simplex have the same volume; so they coincide. In particular, they have the same vertexes, meaning that X and Y have the same rows.

□

Remark 3.25. Here we give another possible proof of Proposition 3.24.

If $X \succ_w Y$, there exists $R \in RS(n)$ such that $Y = RX$. The fact that $Re = e$ implies, by Theorem (P1), that the spectral radius $\rho(R)$ equals 1. Therefore $|\det(R)| \leq 1$ and $|\det(Y)| \leq |\det(X)|$.

If $|\det(X)| = |\det(Y)| \neq 0$ then, by Corollary 3.6, $R = YX^{-1} \in RS(n)$ and $|\det(R)| = 1$. Let $A = RR^t$, then A is positive semi-definite. By Hadamard's theorem

$$1 = \det(A) \leq \prod_{i=1}^{n} A_{ii} = \prod_{i=1}^{n} \|R_i(R)\|_2^2 \leq \prod_{i=1}^{n} \|R_i(R)\|_1^2 = 1.$$

Since R is row stochastic, $\|R_i(R)\|_2 \leq 1$ for all $i = 1, \ldots, n$, then the equation above states that $\|R_i(R)\|_2 = 1 = \|R_i(R)\|_1$ and this implies that $R_i(R) = e_j$ for some $j \in \{1, \ldots, n\}$. Since R is invertible, it must be a permutation matrix. Proposition 3.24 follows from this fact.

△

Corollary 3.26. *Let X, $Y \in M_n$ with $Y \in GL(n)$. Then the following are equivalent*

i) There exists a permutation matrix $Q \in M_n$ such that $QX = Y$.

ii) $X \succ_w Y$ and $Y \succ_w X$.

Proposition 3.27. *Let $X, Y \in M_n$. Suppose that $Y \in GL(n)$, $XY = YX$ and $X \succ Y$. Then, one of the following must hold:*

(i) $Y = PX$ where $P \in M_n$ is a permutation matrix.

(ii) There exists an eigenvector v corresponding to X and Y such that $\langle e; v\rangle = 0$. Moreover, for every common eigenvector v corresponding to different eigenvalues of X and Y, $\langle v; e\rangle = 0$.

Proof. First, if $|\det(X)| = |\det(Y)|$ then $Y = PX$ for some permutation matrix $P \in M_n$ by Proposition 3.24.

Otherwise, note that, as $X, Y \in GL(n)$, $X \succ Y$ is equivalent to $X \succ_s Y$ (see Corollary 3.19). Let $D \in DS(n)$ such that $Y = DX$. Since $XY = YX$ and $|\det(X)| \neq |\det(Y)|$, there is a common eigenvector v of X and Y such that $Xv = \lambda v$ and $Yv = \mu v$ for some $\lambda \in \sigma(X)$, $\mu \in \sigma(Y)$ and $\lambda \neq \mu$. Indeed, consider $\lambda \in \sigma(X)$ and $H_\lambda = \ker(X - \lambda)^k$, where k is the algebraic multiplicity of λ in X. Since $(X - \lambda)^k$ commutes with Y, then H_λ is Y-invariant. Consider, for every $\lambda \in \sigma(X)$, the restriction $Y|_{H_\lambda} : H_\lambda \to H_\lambda$. There is some $\lambda \in \sigma(X)$ such that $\sigma(Y|_{H_\lambda}) \neq \{\lambda\}$ because, otherwise, $\det(X) = \det(Y)$. Let $\mu \in \sigma(Y|_{H_\lambda})$, $\mu \neq \lambda$. If we call $K_\mu = (Y|_{H_\lambda} - \mu)^l$, where l is the algebraic multiplicity of μ in $Y|_{H_\lambda}$, as before K_μ is X-invariant and $\sigma(X|_{K_\mu}) = \{\lambda\}$ since $K_\mu \subseteq H_\lambda$. Thus, there exists a common eigenvector $v \in K_\mu$ of X and Y, and v satisfies our assertion. Moreover

$$\mu \langle v; e\rangle = \langle Yv; e\rangle = \langle DXv; e\rangle = \left\langle Xv; D^t e\right\rangle = \langle Xv; e\rangle = \lambda \langle v; e\rangle$$

and $\mu \neq \lambda$. Therefore $\langle v; e\rangle = 0$. □

4 Joint Majorizations

The goal of this section is to define and characterize several possible extensions of vector majorization in $\mathbb{R}^n$ and operator majorization in $H(n)$ (as introduced by T. Ando [1]), which we call, respectively, joint vector and joint matrix majorizations. In subsection 4.1 we briefly describe joint vector majorizations. The remaining subsections are devoted to state the definitions of joint matrix majorizations and to use some known results to characterize them.

4.1 Joint vector majorization in $\mathbb{R}^n$

In this section we shall consider relations between ordered sets of vectors in $\mathbb{R}^n$, which we call *joint vector majorizations*. These are induced by the different matrix-majorizations and are closely related to the notions of joint matrix majorizations developed in the following subsection.

Let $(x_i)_{i=1,\dots,m}$, $(y_i)_{i=1,\dots,m} \subseteq \mathbb{R}^n$ be two ordered families of vectors and let X, $Y \in M_{n,m}$ be the matrices defined by

$$C_i(X) = x_i, \qquad C_i(Y) = y_i, \qquad i = 1, \dots, m.$$

Definition. Let $(x_i)_{i=1,\dots,m}$, $(y_i)_{i=1,\dots,m} \subseteq \mathbb{R}^n$ and X, $Y \in M_{n,m}$ be as above. We say that $(x_i)_{i=1,\dots,m}$ *jointly weakly majorizes* (respectively *jointly strongly majorizes*, *jointly majorizes*) $(y_i)_{i=1,\dots,m}$ and write

$$(x_i)_{i=1,\dots,m} \succ_w (y_i)_{i=1,\dots,m}$$

(respectively $(x_i)_{i=1,\dots,m} \succ_s (y_i)_{i=1,\dots,m}$, $(x_i)_{i=1,\dots,m} \succ (y_i)_{i=1,\dots,m}$) if it holds

$$X \succ_w Y$$

(resp. $X \succ_s Y$, $X \succ Y$).

Whenever the context makes it clear, we shall not write the subindex corresponding to the set of vectors and simply write $(x_i) \succ_w (y_i)$.

Remark 4.1. The definitions given above can be reinterpreted in the following way

1. $(x_i) \succ (y_i)$ if and only if, for all $\alpha_1, \dots, \alpha_m \in \mathbb{R}$,

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_m x_m \succ \alpha_1 y_1 + \alpha_2 y_2 + \ldots + \alpha_m y_m.$$

2. $(x_i) \succ_s (y_i)$ if and only if there exists $D \in DS(n)$ such that

$$Dx_i = y_i \qquad i = 1, \dots, m.$$

3. $(x_i) \succ_w (y_i)$ if and only if there exists $R \in RS(n)$ such that

$$Rx_i = y_i \qquad i = 1, \dots, m.$$

△

4.2 Joint majorization between Abelian families in $\mathbf{H(n)}$

Let $H(n)$ denote the set of selfadjoint matrices of M_n. In [1] T. Ando considers the majorization relation between pairs of selfadjoint matrices. Indeed, if A, $B \in H(n)$ let $\lambda(A)$, $\lambda(B) \in \mathbb{R}^n$ denote the vectors of eigenvalues of A and B respectively, counted with multiplicity. Then A *majorizes* B (in Ando's sense), denoted $A \succsim B$, if $\lambda(A) \succ \lambda(B)$. Among many others we can cite the following characterizations of majorization between selfadjoint matrices

Theorem 4.2. *Let A, $B \in M_n(\mathbb{C})$ be selfadjoint matrices. Then the following are equivalent*

1. *$A \succsim B$.*
2. *For every convex function $f : (a,b) \to \mathbb{R}$, it holds $tr f(A) \geq tr f(B)$, where $\sigma(A) \cup \sigma(B) \subseteq (a,b)$.*
3. *B belongs to the convex hull of the unitary orbit of A.*

The main purpose of this section is to define and characterize different kinds of *joint majorization* between families of commuting selfadjoint matrices.

Definition. An *Abelian family* is an ordered family $(a_i)_{i=1,\dots,m}$ of selfadjoint matrices in $M_n(\mathbb{C})$ such that

$$a_i\, a_j = a_j\, a_i, \qquad i, j = 1, \dots, m.$$

In order to introduce the notion of joint majorization between Abelian families in $M_n(\mathbb{C})$ we consider the following well known facts.

Let $(a_i)_{i=1,\dots,m}$ and $(b_i)_{i=1,\dots,m}$ be two Abelian families. There exist unitary matrices $U, V \in M_n(\mathbb{C})$ such that

$$U^* a_i U = D_{\lambda(a_i)}, \qquad V^* b_i V = D_{\lambda(b_i)}, \qquad i = 1, \dots, m,$$

where D_x denotes the diagonal matrix with main diagonal $x \in \mathbb{R}^n$. In this case $\lambda(a_i)$ is the vector of eigenvalues corresponding to a_i, counted with multiplicity in some order depending on U. Consider the matrices A, $B \in M_{n,m}$ given by

$$C_i(A) = \lambda(a_i), \qquad C_i(B) = \lambda(b_i), \qquad i = 1, \dots, m.$$

Definition. Let $(a_i)_{i=1,\dots,m}$, $(b_i)_{i=1,\dots,m} \subseteq M_n(\mathbb{C})$, be two Abelian families and consider A, $B \in M_{n,m}$ defined as above. We say that the family $(a_i)_{i=1,\dots,m}$ *jointly weakly majorizes* (respectively *jointly strongly majorizes*, *jointly majorizes*) the family $(b_i)_{i=1,\dots,m}$ and write

$$(a_i)_{i=1,\dots,m} \succ_w (b_i)_{i=1,\dots,m}$$

(respectively $(a_i)_{i=1,\dots,m} \succ_s (b_i)_{i=1,\dots,m}$, $(a_i)_{i=1,\dots,m} \succ (b_i)_{i=1,\dots,m}$) if $A \succ_w B$ (respectively $A \succ_s B$, $A \succ B$).

Whenever the context makes it clear, we shall not write the subindex corresponding to the family of matrices and simply write $(a_i) \succ_w (b_i)$.

Remarks 4.3.

1. A few words concerning the definition are in order. First, note that if U, W are two unitary matrices that diagonalize the family (a_i) simultaneously then there exists a permutation matrix Q such that

$$D_{\lambda(a_i)} = U^* a_i U = Q^* W^* a_i W Q, \qquad i = 1, \dots, n.$$

Thus, if $A' \in M_{n,m}$ denotes the matrix whose columns $C_i(A')$ are the main diagonals of the matrices $W^* a_i W$, then $A = QA'$. That is, the definition above does not depend on the unitary U and the notions are well defined.

2. Recall the different notions of joint vector majorizations introduced in subsection 4.1. Then we can consider the above joint (matrix) majorizations as those induced by the corresponding joint vector majorizations. Indeed, if

$$U^* a_i U = D_{\lambda(a_i)}, \qquad V^* b_i V = D_{\lambda(b_i)}, \qquad i = 1, \ldots, m,$$

as above, then (a_i) jointly (resp. weakly, strongly) majorizes (b_i) if and only if the vectors $(\lambda(a_i))$ jointly (resp. weakly, strongly) majorizes $(\lambda(b_i))$. On the other hand, it is easy to see that if (x_i), $(y_i) \subseteq \mathbb{R}^n$ then (x_i) jointly (resp. weakly, strongly) majorizes (y_i) if and only if the diagonal matrices (D_{x_i}) jointly (resp. weakly, strongly) majorizes (D_{y_i}) in $H(n)$.

4.3 Characterizations of joint majorizations

In this subsection we consider some characterizations of the different notions of joint majorization introduced so far. We begin with the following elementary lemma.

Lemma 4.4. *Let $\mathcal{D}$ be the diagonal algebra in $M_n(\mathbb{C})$ and let $T : \mathcal{D} \to \mathcal{D}$ be a trace preserving, positive unital map. Then there exists $A \in DS(n)$ such that*

$$T(D_x) = D_{Ax} \tag{4.1}$$

On the other hand, if T is as in (4.1) for some $A \in DS(n)$, then T is a trace preserving positive unital map.

Proof. We fix the canonical basis $\{e_i\}_{i=1,\ldots,n}$ of $\mathbb{C}^n$ and identify $\mathcal{D}$ with $\mathbb{C}^n$ as vector spaces by the map $D \mapsto \sum_{i=1}^n D_{ii} e_i$. Therefore T induces a linear transformation $\tilde{T}$ in $\mathbb{C}^n$ by $\tilde{T}x = \sum_{i=1}^n T(D_x)_{ii} e_i$. Let A be the matrix of $\tilde{T}$ with respect to the canonical basis in $\mathbb{C}^n$. Then A satisfies $\operatorname{tr}(Ax) = \operatorname{tr}(x)$, $Ae = e$ and $Ax \geq 0$ whenever $x \geq 0$, where $y \geq 0$ means that all coordinates of $y \in \mathbb{C}^n$ are non-negative and $\operatorname{tr}(y) = y_1 + \ldots + y_n$. Therefore $A \in DS(n)$ and $T(D_x) = D_{Ax}$. The converse is clear.

□

Recall that a *system of projections* in $M_n(\mathbb{C})$ is a family $\{P_i\}_{i=1,\ldots,k}$ such that P_i is an orthogonal projection and $\sum_{i=1}^k P_i = 1$. Given such a family, we consider the associated *pinching*, $\mathcal{C} : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ given by

$$\mathcal{C}(A) = \sum_{i=1}^k P_i A P_i.$$

In particular, if P_i is the orthogonal projection onto $\mathbb{C}e_i$, $i = 1, \ldots, n$, then the pinching associated to this system of projections is called the *diagonal pinching.* Note that $\mathcal{C}$ is a trace preserving positive unital map.

Given $\mathcal{A}$ a unital C^*-subalgebra of $M_n(\mathbb{C})$, it is well known that there exists a system of projections $\{P_i\}_{i=1,\ldots,k}$ such that, for every $A \in M_n(\mathbb{C})$, $A \in \mathcal{A}$ if and only if $\mathcal{C}(A) = A$, where $\mathcal{C}$ is the pinching associated to the system $\{P_i\}_{i=1,\ldots,k}$. Then, we conclude the following well known result.

Lemma 4.5. *Let $\mathcal{A} \subseteq M_n(\mathbb{C})$ be a unital C^*-subalgebra of $M_n(\mathbb{C})$. Then there exists a trace preserving positive unital map $\mathcal{C} : M_n(\mathbb{C}) \to \mathcal{A}$ such that $\mathcal{C}(A) = A$ for all $A \in \mathcal{A}$.*

Theorem 4.6. *Let $(a_i)_{i=1,\ldots,m}$ and $(b_i)_{i=1,\ldots,m}$ be two Abelian families in $M_n(\mathbb{C})$. Then*

1. *$(a_i) \succ_s (b_i)$ if and only if there exists a trace preserving positive unital map*

$$T : C^*(a_1, \dots, a_m) \to C^*(b_1, \dots, b_m)$$

such that $T(a_i) = b_i$.

2. *$(a_i) \succ (b_i)$ if and only if, for all $\gamma_1, \dots, \gamma_m \in \mathbb{R}$ it holds*

$$\gamma_1 a_1 + \ldots + \gamma_m a_m \succsim \gamma_1 b_1 + \ldots + \gamma_m b_m$$

in Ando's sense (see [1]).

3. *$(a_i) \succ_w (b_i)$ if and only if*

$$\mathrm{co}(\sigma(b_1, \dots, b_m)) \subseteq \mathrm{co}(\sigma(a_1, \dots, a_m))$$

where $\mathrm{co}(S)$ denotes the convex hull of the set $S \subseteq \mathbb{R}^m$.

Proof. We have to prove only 1, since 2 and 3 follow from the definitions. Let $U, V \in M_n(\mathbb{C})$ be unitary matrices such that

$$U^* a_i U = D_{\lambda(a_i)}, \qquad V^* b_i V = D_{\lambda(b_i)}, \qquad i = 1, \dots, m$$

where $\lambda(a_i) \in \mathbb{R}^n$. Then, if $a \in C^*(a_1, \dots, a_m)$, $U^* a U \in \mathcal{D}$.

Suppose there exists a trace preserving, positive unital map $T : C^*(a_1, \dots, a_m) \to C^*(b_1, \dots, b_m)$ such that $T(a_i) = b_i$ for all $i = 1, \dots, m$. Let $\tilde{T} : \mathcal{D} \to \mathcal{D}$ be defined by

$$\tilde{T}(D) = V^* T(\Psi(UDU^*))V,$$

where $\Psi : M_n(\mathbb{C}) \to C^*(a_1, \dots, a_m)$ is the map obtained in Lemma 4.5. Note that $\tilde{T}$ is a trace preserving, positive unital such that $\tilde{T}(D_{\lambda(a_i)}) = D_{\lambda(b_i)}$, $i = 1, \dots, m$. By lemma 4.4 we conclude that there exists $A \in DS(n)$ such that

$$A\lambda(a_i) = \lambda(b_i), \qquad i = 1, \dots, m.$$

On the other hand, if $(a_i) \succ_s (b_i)$ there exists $A \in DS(n)$ such that $A\lambda(a_i) = \lambda(b_i)$. Let $T : C^*(a_1, \dots, a_m) \to C^*(b_1, \dots, b_m)$ be defined by

$$T(UD_xU^*) = \Phi(VD_{Ax}V^*)$$

where $\Phi : M_n(\mathbb{C}) \to C^*(b_1, \dots, b_m)$ is a trace preserving, positive unital map such that $\Phi(b) = b$ for all $b \in C^*(b_1, \dots, b_m)$ obtained as in lemma 4.5 and $UD_xU^* \in C^*(a_1, \dots, a_m)$.

$\square$

Corollary 4.7. *Let $(a_i)_{i=1,\dots,m}$ be an Abelian family in $M_n(\mathbb{C})$ and let $\mathcal{C}$ denote the diagonal pinching. Then $(a_i) \succ_s (\mathcal{C}(a_i))$.*

Remark 4.8. Recall the Schur-Horn's theorem for vector majorization in $\mathbb{R}^n$: for $x, y \in \mathbb{R}^n$, $x \succ y$ if and only if there exists a unitary matrix $U \in M_n(\mathbb{C})$ such that

$$\mathcal{C}(U^* D_x U) = D_y, \tag{4.2}$$

where $\mathcal{C}$ is the diagonal pinching. In other words, there exists $a \in H(n)$ such that $x = \lambda(a)$ is the vector of eigenvalues of a and $D_y = \mathcal{C}(a)$, i.e. y is the main diagonal of a.

An orthostochastic matrix of order n is a matrix $B \in M_n$ of the form $B = U \circ \bar{U}$, where $U \in M_n$ is a unitary matrix and "$\circ$" denotes the Schur matrix product. It is well known that the existence of a unitary matrix as in (4.2) is equivalent to the existence of an orthostochastic matrix B of order n such that $Bx = y$.

We can ask whether there is an analogue of the Schur-Horn's theorem for the strong joint vector majorization introduced in section 4.1. More explicitly, we can ask whether it holds the following: given $(x_i)_{i=1,\dots,m}$, $(y_i)_{i=1,\dots,m} \subseteq \mathbb{R}^n$, $(x_i) \succ_s (y_i)$ if and only if there exists a unitary matrix $U \in M_n(\mathbb{C})$ such that

$$\mathcal{C}(U^* D_{x_i} U) = D_{y_i}, \qquad i = 1, \dots, m$$

or equivalently, if there exists an orthostochastic matrix B of order n such that $Bx_i = y_i$ for $i = 1, \dots, m$.

The corollary above establishes one of these implications. Indeed let $U \in M_n(\mathbb{C})$ be unitary and let $D_{y_i} = \mathcal{C}(U^* D_{x_i} U)$, $i = 1, \dots, m$, where $\mathcal{C}$ denotes the diagonal pinching. Let $\Phi : M_n(\mathbb{C}) \to C^*(D_{y_1}, \dots, D_{y_m})$ be a trace preserving, positive unital map such that $\Phi(c) = c$ for all $c \in C^*(D_{y_1}, \dots, D_{y_m})$ obtained as in lemma 4.5. Then $T : C^*(D_{x_1}, \dots, D_{x_m}) \to C^*(D_{y_1}, \dots, D_{y_m})$ given by $T(D) = \Phi(\mathcal{C}(U^* D U))$ is a trace preserving, positive unital map such that

$$T(D_{x_i}) = D_{y_i} \qquad i = 1, \dots, m.$$

Therefore $(D_{x_i}) \succ_s (D_{y_i})$ and, by Remark 4.3, $(x_i) \succ_s (y_i)$.

However the other implication is not true, as the following example shows. Let

$$A = \frac{1}{2} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

It is well known that although $A \in DS(n)$, A is not ortho-stochastic. Let $x_i = e_i \in \mathbb{R}^3$, $i = 1, 2$ and $y_i = Ax_i$, $i = 1, 2$. Then $(x_1, x_2) \succ_s (y_1, y_2)$ and A is the only matrix in $DS(3)$ such that $Ax_i = y_i$, $i = 1, 2$. So there is no orthostochastic B of order 3 such that $Bx_i = y_i$, $i = 1, 2$.

$\triangle$

4.4 Joint majorizations and convex functions

Joint spectrum and multivariate functional calculus. Let $(a_i)_{i=1,\dots,m}$ and A be as in the definition above. Then the set of rows $R(A)$ is called the *joint spectrum* of the family and noted $\sigma(a_1, \dots, a_m)$. Note that if $C^*(a_1, \dots, a_m)$ is the unital C^*-algebra generated by $(a_i)_{i=1,\dots,m}$ then $C^*(a_1, \dots, a_m)$ is $*$-isomorphic to $C(\sigma(a_1, \dots, a_m))$, the set of continuous functions over $\sigma(a_1, \dots, a_m)$. By item 1. of Remark 4.3 $\sigma(a_1, \dots, a_m)$ does not depend on the unitary U.

Let $(a_i)_{i=1,\dots,m}$ be an Abelian family in $M_n(\mathbb{C})$ and let U be a unitary matrix that simultaneously diagonalize this family so that

$$U^* a_i U = D_{\lambda(a_i)}, \qquad i = 1, \dots, m.$$

Let $f : V \to \mathbb{R}$ be a function such that $\sigma(a_1, \dots, a_m) \subseteq V$ (and therefore $f \in C(\sigma(a_1, \dots, a_m))$ since $\sigma(a_1, \dots, a_m)$ is a discrete set). Then we consider

$$f(a_1, \dots, a_m) = UDU^*$$

where D is the diagonal matrix with main diagonal

$$\operatorname{diag}(D) = (f(R_1(A)), \dots, f(R_m(A)))$$

that does not depend on U (see Remark 4.3). We say that the selfadjoint matrix $f(a_1, \dots, a_m)$ is obtained from the family $(a_i)_{i=1,\dots,m}$ by functional calculus.

The following proposition is a restatement of Proposition 3.7 in this context.

Proposition 4.9. *Let $(a_i)_{i=1,\dots,m}$ y $(b_i)_{i=1,\dots,m}$ be two Abelian families. Then, $(a_i) \succ_s (b_i)$ if and only if, for every convex function $f: V \to \mathbb{R}$ it holds that*

$$\operatorname{tr} f(a_1, \dots, a_m) \geq \operatorname{tr} f(b_1, \dots, b_m),$$

where $V \subseteq \mathbb{R}^m$ is a convex set containing $\sigma(a_1, \dots, a_m)$ and $\sigma(b_1, \dots, b_m)$.

The following Proposition is an immediate consequence of theorem 4.2.

Proposition 4.10. *Let $(a_i)_{i=1,\dots,m}$ and $(b_i)_{i=1,\dots,m}$ be two Abelian families. Then, $(a_i) \succ (b_i)$ if and only if, for all $\gamma_1, \dots, \gamma_m \in \mathbb{R}$ and every convex function $f: \mathbb{R} \to \mathbb{R}$ it holds*

$$\operatorname{tr}\ f(\gamma_1 a_1 + \ldots + \gamma_m a_m) \geq \operatorname{tr}\ f(\gamma_1 b_1 + \ldots + \gamma_m b_m).$$

Let $(a_i)_{i=1,\dots,m} \subseteq M_n(\mathbb{C})$, the *(first) joint numerical range* is defined by

$$W(a_1, \dots, a_m) = \{(v^* a_1 v, \dots, v^* a_m v),\ v \in \mathbb{C}^n,\ v^* v = 1\},$$

(see [8]). We shall relate the (first) joint numerical range $W(a_1, \dots, a_m)$ to the joint spectrum $\sigma(a_1, \dots, a_m)$ of an Abelian family.

Lemma 4.11. *Let $(a_i)_{i=1,\dots,m}$ be an Abelian family. Then,*

$$W(a_1, \dots, a_m) = \operatorname{co}(\sigma(a_1, \dots, a_m)).$$

Proof. Note that $W(a_1, \dots, a_m)$ is invariant under unitary conjugation of the a_i's by a fixed unitary $U \in M_n$. So we can suppose that the a_i's are diagonal, i.e $a_i = D_{\lambda(a_i)}$, $i = 1, \dots, m$. If $v^* v = 1$ we have

$$\begin{aligned} (v^* a_1 v, \dots, v^* a_m v) &= \left(\sum_{j=1}^{n} |v_j|^2 \lambda_j(a_1)\,,\ \dots\,,\ \sum_{j=1}^{n} |v_j|^2 \lambda_j(a_m) \right) \\ &= \sum_{j=1}^{n} |v_j|^2 \left(\lambda_j(a_1), \dots, \lambda_j(a_m)\right) \end{aligned}$$

where $\sum_{j=1}^{n} |v_j|^2 = 1$. Then, the Lemma follows from this fact.

□

Proposition 4.12. *Let $(a_i)_{i=1,\dots,m}$ and $(b_i)_{i=1,\dots,m}$ be two Abelian families. Then, the following are equivalent:*

1. *$(a_i) \succ_w (b_i)$.*
2. *$W(b_1, \dots, b_m) \subseteq W(a_1, \dots, a_m)$.*
3. *For every convex function $f: V \to \mathbb{R}$ it holds*

$$\|f(a_1, \dots, a_m)\| \geq \|f(b_1, \dots, b_m)\|.$$

where $V \subseteq \mathbb{R}^m$ is a convex set containing $\sigma(a_1, \dots, a_m)$ and $\sigma(b_1, \dots, b_m)$.

Proof. 1. ⇔ 2. follows from Lemma 4.11 and item 3. of Theorem 4.6. On the other hand, 1. ⇔ 3. follows from Corollary 3.3.

□

4.5 Equivalence relations associated to joint majorizations

The joint majorizations considered so far are preorder relation among Abelian families in $M_n(\mathbb{C})$. The next Proposition, which follows from Theorem 4.6 and Theorem 3.22, describes the equivalence relations associated to some of these preorders.

Theorem 4.13. *Let $(a_i)_{i=1,\dots,m}$ and $(b_i)_{i=1,\dots,m}$ be two Abelian families in $M_n(\mathbb{C})$. Then the following are equivalent:*

i) There exists a unitary matrix $W \in M_n$ such that $W^* a_i W = b_i$ for all $i = 1, \dots, m$.

ii) $(a_i) \succ_s (b_i)$ and $(b_i) \succ_s (a_i)$.

iii) $(a_i) \succ (b_i)$ and $(b_i) \succ (a_i)$.

Proof. Note that the inner automorphism $\alpha : C^*(a_1, \dots, a_m) \to C^*(b_1, \dots, b_m)$ induced by W is a trace preserving, positive unital map. Therefore $i)$ implies $ii)$. Clearly $ii) \Rightarrow iii)$.

On the other hand, if $(a_i) \succ (b_i)$ and $(b_i) \succ (a_i)$, by Theorem 3.22 there exits a permutation matrix $Q \in M_n$ such that

$$V(Q^t(U^* a_i U)Q)V^* = b_i, \qquad i = 1, \dots, m$$

where U, $V \in M_n$ are as in the proof of Theorem 4.6. Therefore, by taking $W = UQV$ we have completed the proof.

□

Next Proposition is an immediate consequence of Proposition 4.12, so we omit the proof.

Proposition 4.14. *Let $(a_i)_{i=1,\dots,m}$ and $(b_i)_{i=1,\dots,m}$ be two Abelian families in $M_n(\mathbb{C})$. Then the following are equivalent:*

i) $(a_i) \succ_w (b_i)$ and $(b_i) \succ_w (a_i)$.

ii) $W(a_1, \dots, a_m) = W(b_1, \dots, b_m)$.

On the other hand, we describe the Abelian families that are minimal with respect to the joint majorizations. The following is a restatement of Proposition 3.13.

Proposition 4.15. *Suppose that $(a_i)_{i=1,\dots,m}$ is an Abelian family in $M_n(\mathbb{C})$, which is minimal with respect to any of the joint majorizations. Then, $a_i = \lambda_i I$, $i = 1, \dots, m$, where $\lambda_i \in \mathbb{R}$ for all $i = 1, \dots, m$.*

Acknowledgement

We would like to thank Professor Demetrio Stojanoff for his advise and suggestions regarding the results herein.

References

[1] T. Ando, *Majorization, doubly stochastic matrices and comparison of eigenvalues*, Lecture Note, Hokkaido Univ., 1982.

[2] L. Beasley, S. Lee, *Linear operators preserving multivariate majorization*, Linear Algebra and Appl. 304 (2000), 141-159.

[3] R. Bhatia, *Matrix Analysis*, Springer-Verlag New York, 1997.

[4] G.-S. Cheon, Y.-H. Lee, *The doubly stochastic matrices of a multivariate majorization*, J. Korean Math. Soc. 32 (4) (1995), 857-867.

[5] G. Dahl, *Matrix majorization*, Linear Algebra and Appl. 288 (1999), 53-73.

[6] R. Horn, C. Johnson, *Matrix Analysis*, Cambridge Univ. Press, 1985.

[7] S.-G. Hwang, S.-S. Pyo, *Matrix majorization via vector majorization*, Linear Algebra and Appl. 332-334 (2001), 15-21.

[8] C.-K. Li, Y.-T. Poon, *Convexity of the Joint Numerical Range*, SIAM J. Matrix Analysis Appl. 21 (1999), 668-678.

[9] C.-K. Li, E. Poon, *Linear Operators Preserving Directional Majorization*, Linear Algebra and Appl. 325 (2001), no. 1-3, 141-146.

[10] S. Sherman, *On a conjecture concerning doubly stochastic matrices*, Proc. Amer. Math. Soc. 3 (1952), 511-513.

[11] S. Sherman, *A correction to "On a conjecture concerning doubly stochastic matrices"*, Proc. Amer. Math. Soc. 5 (1954), 998-999.

[12] S. Shreiber, *On a result of Sherman concerning doubly stochastic matrices*, Proc. Amer. Math. Soc. 9 (1958) 350-353.