Centrum voor Wiskunde en Informatica

Centre for Mathematics and Computer Science

M. Hazewinkel

On positive vectors, positive matrices and the specialization ordering

Department of Pure Mathematics Report PM-R8407 September

1984

Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

M. Hazewinkel

On positive vectors, positive matrices and the specialization ordering

Department of Pure Mathematics Report PM-R8407 September

The Centre for Mathematics and Computer Science is a research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a nonprofit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

ON POSITIVE VECTORS, POSITIVE MATRICES AND THE SPECIALIZATION ORDERING

M. HAZEWINKEL
Centre for Mathematics and Computer Science, Amsterdam

A brief introductory discussion is given of the specialization partial ordering for positive vectors in connection with the positive rank of nonnegative matrices.

1980 MATHEMATICS SUBJECT CLASSIFICATION: 06A10, 15A48, 15A51
KEY WORDS & PHRASES: specialization order, nonnegative matrices, majorization.
NOTE: This report will be submitted for publication elsewhere.

Report PM-R8407
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

1. Introduction

Let $p = (p_1,...,p_n)$ and $q = (q_1,...,q_n)$ be n-tuples of nonnegative real numbers of the same l_1-norm. I.e. $|p| = p_1+...+p_n = |q| = q_1+...q_n$. Then p is said to be more general than q, or to specialize to q iff $\bar{p}_1 \leqslant \bar{q}_1, \bar{p}_1+\bar{p}_2 \leqslant \bar{q}_1+\bar{q}_2, \ldots, \bar{p}_1+...+\bar{p}_n \leqslant \bar{q}_1+...+\bar{q}_n$. Here $\bar{p} = (\bar{p}_1, \ldots, \bar{p}_n)$ is a reordering of the n-tuple p such that $\bar{p}_1 \geqslant \bar{p}_2 \geqslant \cdots \geqslant \bar{p}_n$.

This is a partial order which I like to call the specialization order. It is denoted $p > q$. Under various names - such as e.g. majorization, Snapper order, Ehresmann order, dominance order, natural order - it occurs and plays important roles in the most diverse parts of mathematics, ranging from algebraic vector-bundle theory to numerical analysis, and from combinatorics, master equations and entropy considerations to representation theory and the geometry of Grassmann and flag manifolds.

Under the name 'majorization' this order is the subject of a book [13], which, in subject matter, is practically disjoint from the survey paper [9] which deals with the same order under the name 'specialization order'. Nor do these two together come anywhere near giving a complete picture of the role of this order in the scheme of things mathematical. The first part of this lecture describes some of the phenomenology of this order. It is largely based on [9].

In the second part I discuss the following problem. Consider a nonnegative $n \times m$ matrix Q and consider all factorizations $Q = AB$ into positive matrices A, B of sizes $n \times r$ and $r \times m$ respectively. What is the smallest r such that this can be done? This r is called the positive rank of Q, and the problem of determining it is of importance in stochastic system theory [14] and elsewhere [4], [16]. There is a flavour about this problem which reminds one - me in any case - of the partial order just described. The second part of this lecture discusses some examples concerning the positive rank of nonnegative matrices and concludes with some remarks on the relations between positive rank and the specialization order.

2. Some of the phenomenology of the specialization order

2.1. Representations of the symmetric groups. Let S_n be the group of permutations on n letters. Let $\kappa = (\kappa_1, \ldots, \kappa_m)$ be a partion of n, and let S_κ be the corresponding Young subgroup $S_\kappa = S_{\kappa_1} \times S_{\kappa_2} \times \cdots \times S_{\kappa_m}$, where S_{κ_i} is seen as the permutation subgroup of S_n acting on the letters $\kappa_1 + \ + \kappa_{i-1} + 1, \ldots, k_1 + \cdots + \kappa_i$. If $\kappa_m = 0$ the factor S_{κ_m} is deleted. Take the trivial representation of S_κ and induce this up to S_n. Let $\rho(\kappa)$ denote the resulting representation which is of dimension $\binom{n}{\kappa}$. It can be easily described as follows. Take m symbols $a_1,...,a_m$, and consider all associative (but noncommutative) words $w_1 w_2 \cdots w_n$ of length n in the symbols $a_1,...,a_m$ such that each a_i occurs precisely κ_i times. Let $W(\kappa)$ denote this set. The group S_n acts on the set $W(\kappa)$ by $\sigma^{-1}(w_1...w_n) = w_{\sigma(1)}...w_{\sigma(n)}$. Let $V(\kappa)$ be the vector space of all $\mathbf{R}$-linear combinations of elements of $W(\kappa)$ and extend the action of S_n linearly. This is the representation $\rho(\kappa)$.

2.2. The Snapper, Liebler- Vitale, Lam, Young theorem. The representation $\rho(\kappa)$ is a subrepresentation of the representation $\rho(\lambda)$ iff $\kappa < \lambda$.

2.3. The Gale- Ryser theorem. Let μ and ν be two partitions of n. Then there is a matrix consisting of zeros and ones whose columns sum to μ and whose rows sum to ν iff $\nu > \mu^*$. Here μ^* is the dual partition to μ defined by $\mu_i^* = \# \{j : \kappa_j > i\}$. For example $(2,2,1)^* = (3,2)$.

2.4. Doubly stochastic matrices. An $m \times n$ matrix $Q = (q_{ij})$ is called *stochastic* if $q_{ij} \geqslant 0$ for all i,j and all the columns add up to one. It is called *doubly stochastic* if moreover the rows add up to n/m.

Let x and y be two n-vectors. Then one says that x is an average of y iff there is a doubly stochastic $n \times n$ matrix Q such that $x = Qy$. (This of course implies that $|x| = |y|$.) One now has the theorem that $x > y$ iff x is an average of y.

2.5. Muirheads inequality. One of the best known inequalities is

$$(x_1...x_n)^{1/n} \leqslant n^{-1}(x_1+...+x_n)$$

for positive (or nonnegative) real numbers $x_1,...,x_n$. A far reaching generalization due to Muirhead goes as follows. Given a nonnegative vector $p = (p_1,...,p_n)$, $p_i \geqslant 0$, one defines a symmetrical mean (of the nonnegative variables $x_1,...,x_n$) by the formula

$$[p](x) = (n!)^{-1}\sum_{\sigma} x_1^{p_{\sigma(1)}} \ldots x_n^{p_{\sigma(n)}}$$

where the sum is over all permutations σ in S_n. Now let p and q be two nonnegative vectors of the same l_1-norm. Then Muirheads inequality states that $[p](x) \leqslant [q](x)$ for all nonnegative values of the variables $x_1,...,x_n$ iff p is an average of q; i.e. iff $p > q$. The geometric mean - arithmetic mean inequality thus arises from the maximally crude specialization relation $(1,...,1) > (n,0,...,0)$.

2.6. Completely reachable systems. Let $L_{m,n}^{cr}$ denote the space of all completely reachable pairs of matrices (A,B) of sizes $n \times n$ and $n \times m$ respectively. Let F be the Lie group of all feedback transformations acting on $L_{m,n}^{cr}$. This is the group of all block lower triangular matrices $f = \left(\begin{smallmatrix} S & 0 \\ K & T \end{smallmatrix}\right)$, where S is an invertible real $n \times n$ matrix, T is an invertible real $m \times m$ matrix and K is an $m \times n$ matrix. An element f of F acts on an element (A,B) of $L_{m,n}^{cr}$ according to

$$(A,B)^f = (SAS^{-1} + SBTS^{-1}K, SBT)$$

(F consists of all transformations generated by 'base change in state space', 'base change in input space ' and 'state space feedback'.) It is a theorem of Brunovsky, Kalman and Wonham-Morse that the orbits of F acting on $L_{m,n}^{cr}$ correspond bijectively with all partitions of n into at most m parts.

Let $U(\kappa)$ be the orbit of F acting on $L_{m,n}^{cr}$ labelled by the partition κ. Then a second theorem noted by a fair number of people independantly of each other (Byrnes, Hazewinkel, Kalman, Martin,...) states that the closure $\overline{U}(\kappa)$ of $U(\kappa)$ contains the orbit $U(\lambda)$ iff $\kappa > \lambda$. This theorem says things about what can happen to the controllability indices of a system (A,B) (these are the numbers making up the partion defined by (A,B)), if the system changes, in particular fails.

2.7. Vectorbundles over the Riemann sphere. Let E by a holomorphic vectorbundle over the Riemann sphere $S^2 = P^1(C)$. Then, according to Grothendieck, E splits as a direct sum of (complex) line bundles

$$E = L(\kappa_1) \oplus L(\kappa_2) \oplus \cdots \oplus L(\kappa_m)$$

where $L(i)$ is the unique (up to isomorphism) line bundle over S^2 of degree i (or first Chern number i). Thus each holomorphic vectorbundle of (complex) dimension m over S^2 is classified by an m-tuple $\kappa(E)$ of integers in decreasing order One now has the following theorem of Shatz concerning the degeneration properties of such bundles. Let E_t be a holomorphic family of holomorhpic bundles over S^2. Then for all small enough t, $\kappa(E_t) > \kappa(E_0)$. And inversely if $\kappa > \lambda$ then there is an holomorphic family E_t such that $\kappa(E_t) = \kappa$ for t small and $\neq 0$ and such that $\kappa(E_0) = \lambda$.

2.8. Orbits of nilpotent matrices. Let N_n be the space of all $n \times n$ complex nilpotent matrices. Consider $GL_n(C)$, the group of all invertible complex matrices of size $n \times n$ acting on N_n by similarity. I.e.

$$A^S = SAS^{-1}, \ A \in N_n, \ S \in GL_n(C).$$

By the Jordan normal form theorem the orbits of this action are labelled by the partitions of n. Let $0(\kappa)$ be the orbit consisting of all nilpotent matrices which are similar to the one consisting of the Jordan blocks $J(\kappa_i)$, $i = 1,...,n$, where $J(r)$ is the $r \times r$ matrix with 1's just above the diagonal and zeros everywhere else. Then the Gerstenhaber-Hesselink theorem says that the closure $\overline{0}(\lambda)$ of the orbit $0(\lambda)$ contains the orbit $0(\kappa)$ iff $\kappa > \lambda$. Note the reversal of the order with respect to the orbit closure-inclusion relation described in 2.6 above.

2.9. Diagonal and eigenvalues of a Hermitian matrix. Let A be a Hermitian matrix; let $a = (a_1,...,a_n)$ be its vector of diagonal elements and let $b = (b_1,...,b_n)$ be its vector of eigenvalues (in any order). Then a well known result of Schur [15] says that $a > b$. A converse was proved by Horn [10]: let a and b be two real n-vectors such that $a_1 + ... + b_n$ and $a > b$, then there exists a real symmetric $n \times n$ matrix A with vector of diagonal elements a and vector of eigenvalues b.

3. Interrelations and generalizations

3.1. Interrelations It turns out that it is no accident that the same partial order turns up in describing inclusion, degeneration or specialization relations in all these seemingly quite unrelated bits of mathematics. There are so to speak compelling reasons why the same order should appear again and again.

For example there is a construction, continuous in the parameter parameters, which associates to every completely reachable pair (A,B) a vectorbundle $E(A,B)$ over the Riemann sphere. This is the socalled Hermann-Martin bundle. This holomorphic bundle is such that its classifying sert of first Chern numbers is precisely the set of controllability indices (or Kronecker indices) of the original pair (A,B). Cf. [9] for more details. This survey paper also describes various interrelations between the matters discussed in 2.2, 2.6, 2.7, 2.8 above.

For interrelations between 2.2, 2.3, 2.4, 2.5 the reader is referred to e.g. [8].

Finally there is a discussion of the Schur-Horn results (in the framework of the specialization order) in [13], section 9.B.

3.2. Generalizations. The interrelations between 2.2, 2.6, 2.7, 2.8 described in [9] bring in and make fundamental use of Grassman manifolds and their Schubert cells. And thus it turns out that the specialization order has much to do with the socalled Bruhat partial order on the Weyl group S_n. These are of course the Weyl groups of the simple Lie algebras of type A_n. There are analogous orderings on the Weyl groups of the other (classical) simple Lie algebras and one wonders if there are suitable analogous of all the object (and their interrelations) occuring in section 2 above 'belonging' to these simple Lie algebras. This is for example the case with 2.2 and 2.8 and the relation between them [7,11]. For a survey of (the role of) the Bruhat on Coxeter groups, in particular Weyl groups, cf. [5].

The Schur-Horn results of 2.9 above can be reinterpreted as a statement about the compact Lie group $U(n)$. It then says something about the orbits of $U(n) = G$ acting on its Lie algebra $L(G)$ (by the adjoint action) compared with the orbits of the Weyl group of G acting on the Lie algebra $L(T)$ of a maximal torus T in G.

In this form Kostant [12] found a generalization valid for all compact Lie groups, and then in turn this can be taken further and placed in the context of symplectic geometry [3].

On the other hand for many of the objects described in section 2 above, especially the more combinatorial ones and the system theoretic one, the right non-A_n -type analogous have not yet been defined.

4. Positive rank of matrices. Examples.

4.1. The positive rank of a nonnegative matrix. Definition. Let Q be a nonnegative (i.e. $q_{ij} \geqslant 0$ for all i,j) $n \times m$ matrix. Consider a factorization

$$Q = AB \tag{4.2}$$

where A and B are nonnegative matrices of sizes $n \times r$ and $r \times m$ respectively. The smallest r for which such a factorization exists is called the positive rank of Q and is denoted $posrk(Q)$. Let $rk(Q)$ denote the ordinary linear rank of Q. Then obviously

$$rk(Q) \leqslant posrk(Q) \leqslant min(m,n) \tag{4.3}$$

and, as we shall see, in general $posrk(Q)$ can assume all intermediate values.

4.4. Geometric interpretation. Obvioulsy if Q has a zero column then its positive rank is unchanged if this column is removed. Similarly if A has a zero column, say, the i-th, then removing the i-th column of A and the i-th row of B produces matrices A' and B' of sizes $n \times (r-1)$ and $(r-1) \times m$ such that $Q = A'B'$. We can therefore assume that Q and A have no zero columns. Multiplying Q on the right by a diagonal matrix $D = diag(\lambda_1, \ldots, \lambda_n)$, $\lambda_i > 0$, also does not change the factorization properties of Q. By inserting also a suitable product DD^{-1} of such matrices between A and B we can therefore restrict ourselves to the consideration of factorizations

$$Q = AB, \ Q \text{ and } A \textit{ are stochastic matrices} \tag{4.5}$$

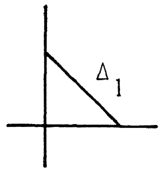
(i.e. all the columns of Q and A sum to one, i.e. are probability vectors). Let a_j be the j-th column vector of A. Then the i-th column vector q_i of Q is equal to $a_1 b_{1i} + \ldots + a_r b_{ri}$, which has length

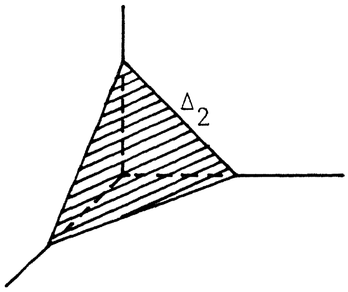
$$1 = |q_i| = b_{li} |a_r| = b_{li} + \cdots + b_{ri}$$

and it follows that the column sums of B are also equal to 1.

4.6. Lemma. Let Q, A be stochastic matrices with the same number of rows. Then there exists a nonnegative matrix B such that $Q = AB$ iff each column vector of Q is a convex linear combination of the column vectors of A.

Let Δ_n denote the simplex in R^{n+1} of all nonnegative vectors of l_1-norm 1. For example Δ_0 is a single point, Δ_1 is a line segment, Δ_2 is a triangle (with its interior) and Δ_3 is a solid tetrahedron.





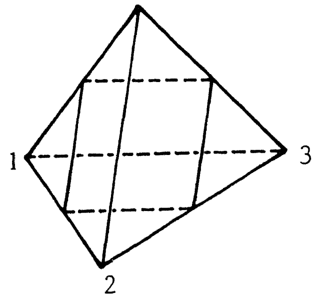
A stochastic $n \times m$ matrix Q is now nicely represented as an m-point set $p(Q)$ in Δ_{n-1}. (Permuting the columns of Q, which is the same as multiplication on the right with a permutation matrix also obviously does not change its factorization properties.) And a relation $Q = AB$ holds iff the m-point set $p(Q)$ is in the convex hull of the r-point set $p(A)$. Thus

4.7. Lemma. The positive rank of Q is equal to the number of elements of the smallest set S in Δ_{n-1} such that Q is contained in the convex hull of S.

4.8. Example. Suppose $rk(Q) = 1$. Then Q has m identical column vectors. Take $r = 1$ and A equal to this column vector. This gives $rk(Q) = 1 \Leftrightarrow posrk(Q) = 1$.

4.9. Example. Suppose $rk(Q) = 2$. The columns of Q then span a plane with intersects Δ_{n-1} in a line segment. Let S be the two-point set consisting of the endpoints of this line segment. Then $p(Q)$ is contained in the convex hull of S. Thus $rk(Q) = 2 \Leftrightarrow posrk(Q) = 2$.

4.10. Example. Things change for $rk(Q) \geqslant 3$. Consider for example 4×4 matrices Q of rank 3. The intersection of the three dimensional hyperplane spanned by the columns of Q with the tetrahedron Δ_3 is either a triangle or a quadrangle (or a lower dimensional degeneration).



For instance the intersection may well look like the one depicted above. If $p(Q)$ consists of the four corner points of this intersection quadrangle then it is visibly impossible to find a three point set S such that $p(Q) \subset convex\ hull(S)$. Thus $posrk(Q) = 4$, $rk(Q) = 3$ in such cases. The case illustrated is

$$Q = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}.$$

This also shows that $posrk(Q) = 4$, $rk(Q) = 3$ is not a rare phenomenon. Indeed let V_3 be the space of all nonnegative 4×4 matrices of rank 3. Then $\{Q \in V_3 : posrk(Q) = 4\}$ has a nonempty interior.

Let us now analyze the case $rk(Q) = 3$ systematically. As defined Δ_n is a subset of R^{n+1}. Let $W_n = \{x \in R^{n+1} : x_1 + \cdots + x_{n+1} = 1\}$. Then W_n is affinely isomorphic with R^n and $\Delta_n \subset R^n$.

4.11. Lemma. Let V be a plane in $W_n = R^n$. Then the intersection with Δ_n is a convex polygon with at most $(n+1)$ sides. (It may degenerate into a line segment (i.e. one side), or a point (0 sides), or the intersection may also be empty (i.e. -1 sides by convention).)

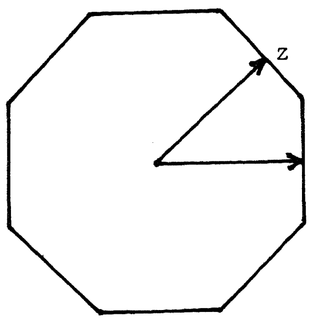
Proof. Δ_n is the intersection of $n+1$ halfspaces in $W_n = R^n$. The intersection of such a halfspace with V is a halfplane or empty.

4.12. Lemma. There are planes V in $W_n = R^n$ such that the intersection of V with Δ_n is a convex polygon with precisely $n+1$ sides.

Proof. Consider the regular $(n+1)$-polygon D in $R^2 = C$ centered at zero given by $\{x \in R^2 : <x, z^k> \leqslant 1, k = 0, \ldots, n\}$ where z is a primitive $(n+1)$-st root of unity. Consider $C = R^2$ as embedded in $R^n = W_n$ by $(x_1, x_2) \rightarrow (x_1, x_2, 0, \ldots, 0)$. Let T_k be the hyperplane defined by

$$T_k = \{x \in R^n : <x, h_k> = 1\}$$

where h_k is a vector of the form $h_k = (z^k, a_k)$, $a_k \in R^{n-2}$.



Then $T_k \cap R^2$ is the line through the k-th segment of the regular $(n+1)$-polygon. Now the points $1, z, \ldots, z^n$ are in general position in R^2 (i.e. not all are on one line). It follows that one can choose $a_0, \ldots, a_n$ such that the $h_0, \ldots, h_n$ are in general position, i.e. not all in some $(n-1)$-plane. Let H_k be the halfspace defined by T_k such that $0 \in H_k$. Then $\cap H_k \neq \emptyset$ and in fact $(\cap H_k) \cap R^2 = D$. Moving the H_k slightly if necessary we will have $dim(\cap H_k) = n$ so that $\cap H_k$ is an n-simplex and it will still be true that $(\cap H_k) \cap R^2$ is a convex $(n+1)$-polygon (because it has to be near the original regular $(n+1)$-polygon).

Now all simplices in R^n are affinely isomorphic. The image of R^2 under the affine isomorphism which takes Δ_n onto $(\cap H_k)$ is the desired plane V.

4.13. Corollary and example. Let $n > 3$. Then there exist nonnegative $n \times n$ matrices Q with $rk(Q) = 3$ and $posrk(Q) = n$.

Proof. Take a plane V in $R^{n-1} = W_{n-1}$ which intersects $\Delta_{n-1} \subset W_{n-1}$ in a convex polygon of n sides as in lemma 4.12. Let Q be such that $p(Q)$ is the set of corners of this n-polygon. Suppose that $Q = AB$ with A a stochastic $n \times r$ matrix and $r<n$. The columns of A (or more precisely the elements of $p(A)$) span an $(r-1)$-simplex in Δ_{n-1}. The intersection of this simplex with V is a polygon with $\leqslant r$ sides (Lemma 4.11). Now if Convex hull $p(A) \supset p(Q)$, then also Convex hull $p(A) \cap V \supset p(Q) \cap V = p(Q)$, and

$$\textit{Convex hull } p(A) \cap V \supset \textit{Convex hull } p(Q) = \Delta_{n-1} \cap V \supset$$
$$\textit{Convex hull } p(A) \cap V.$$

So we would have

$$\textit{Convex hull } p(A) \cap V = \textit{Convex hull } p(Q).$$

But the left hand side of this is a polygon with $\leqslant r < n$ sides and the right hand side is a convex polygon with n sides. A contradiction.

4.14. Corollary. *From these examples it is straightforward to construct examples for all* $n,m,d \geqslant 3$, $d \leqslant i \leqslant min(n,m)$ *of nonnegative* $n \times m$ *matrices* Q *such that* $rk(Q) = d, posrk(Q) = i$.

5. Concluding remarks. Relations between specialization order and positive rank

Let me conclude with some remarks pertaining to interrelations and similarities between the material discussed in sections 2,3 and that of section 4.

5.1. Characterizations by means of 'measuring' functions. There are various ways of characterizing the specialization order $x > y$ in terms of 'measuring' functions or 'gauge' functions. E.g. (Hardy, Littlewood and Polya; cf. [13], Prop. 4.B.1) one has that $x > y$ iff $\sum_i g(x_i) \leqslant \sum_i g(y_i 0$ for all continuous convex functions g. Another characterization is (cf. [13], Prop. 4.B.8): $x > y$ iff $\max_\pi \sum t_i x_{\pi(i)} \leqslant \max_\pi \sum t_i y_{\pi(i)}$ where π runs through all permutations, for all $t \in R^n$.

For that matter the very definition of $x > y$ is in terms of a finite number of measuring functions.

Now it is a rather simple fact (easily seen geometrically) that a vector x is a convex linear combination of vectors $x^{(1)},...,x^{(r)}$ iff for all vectors $t \in R^n$ $<t,x> \leqslant \max_i <t,x^{(i)}>$. Using this one sees that a factorization $Q = AB$ of nonnegative matrices with Q and A (and hence B) stochastic holds iff for all $t \in R^n$, $\max_i <t,q_i> \leqslant \max_i <t,a_i>$.

Given a probability vector x let X be the matrix consisting of the columns Px where P runs through all the $n \times n$ permutation matrices. Combining the facts above one now obtains

5.2. Observation. For two probability vectors x,y one has $x > y$ iff there is a factorization relation $X = YB$ with a nonnegative B.

This is of course practically the same as the result of Rado, 1952 that $x > y$ iff x is a convex linear combination of the Py, where P runs through all the permutation matrices.

5.3. Primes. Let Q be a nonnegative $n \times n$ matrix which is invertible. Suppose that Q^{-1} is also nonnegative. Let C be the cone $C = \{x \in R^n : x_i \geqslant 0\}$. It follows that $QC = C$, and by linearity it follows that Q must take extreme rays into extreme rays. From this one sees that such Q are necessarily of the form $Q = DP$ (or PD), where P is a permutation matrix and D is diagonal $D = diag(c_1,...,c_n)$, $c_i > 0$. Let us call such Q's units.

Obviously the factorization properties of Q do not change if it is pre- or post-multiplied with a unit. Now almost every nonnegative matrix Q can be written in the form $D_1 Q_{ds} D_2$ with Q_{ds} doubly stochastic and the D_1,D_2 nonnegative, invertible and diagonal. (This depends on the zero pattern of Q; the remaining matrices are decomposable; cf. [4] and [17] for precise statements.)

A nonnegative matrix Q is said to be prime if for every factorization $Q = AB$ into nonnegative matrices, A or B is a unit, and Q itself is not a unit.

Obviously a good understanding of the primes will help for a theory of factorization of nonnegative matrices and, by the $Q = D_1 Q_{ds} D_2$ result, understanding the primes has much to do with (the semigroup of) doubly stochastic matrices, which, inturn, are related to the specialization order by 2.4 above.

5.4. The specialization order for n-tuples of vectors. The specialization order compares n-tuples of real numbers with n-tuple tuples of real numbers. There is a generalization (which has to do with dissipative mottion in the framework of the C^*-algebra approach to thermodynamics [1,2]) which compares n-tuples of r-vectors with n-tuples (same n) of s-vectors. The definition is as follows. Let $x^{(1)},...,x^{(n)} \in R^r$; $y^{(1)},...,y^{(n)} \in R^s$, then $(x^{(1)},...,x^{(n)}) > (y^{(1)},...,y^{(n)})$ iff there exists a stochastic matrix T such that $x^{(i)} = Ty^{(i)}$, $i = 1,...,n$.

Recall that $T:R^s \rightarrow R^r$ is stochastic if it takes probability vectors into probability vectors, or, inother words, if the columns of (the matrix of) T all sum to one.

This generalizes the specialization order for n-tuples of real numbers in the following sense: if $x,y \in R^n$, then $x>y$ iff $(x,e_n) > (y,e_n)$ where e_n is the n-vector $(n^{-1},...,n^{-1}) \in R^n$. Indeed, the condition that T be stochastic and $Te_n = e_n$ means that T is doubly stochastic.

Interpreted in terms of this generalized specialization order a factorization relation $Q = AB$ thus means that the n-tuple of column vectors of Q is more general than the n-tuple of column vectors of B. (Here we have again, as we can, assumed that the matrix A is stochastic.) Note that here the matrices Q and B are being compared, rather than the matrices A and Q as in the geometric considerations of section 4.

References

[1] Alberti, P. & A. Uhlmann, *Dissipative motion in state spaces,* Teubner, 1981.

[2] Alberti, P. & A. Uhlmann, *Stochasticity and partial order,* Reidel Publ. Cy., 1982.

[3] Atiyah, M.F., *Convexity and commuting Hamiltonians,* Bull. London Math. Soc. 14 (1982), 1-15.

[4] Berman, A. & R.J. Plemmons, *Nonnegative matrices in the mathematical sciences,* Acad. Pr., 1979.

[5] Björner, *Orderings of Coxeter groups,* preprint, 1983.

[6] Brualdi, R.A., S.V. Parter & H. Schneider, *The diagonal equivalence of a nonnegative matrix to a stochastic matrix,* J. Math. Anal. and Appl. 16 (1966), 31-50.

[7] De Concini, C. & C. Procesi, *Symmetric functions, conjugacy classes and the flag variety,* Inv. Math. 64 (1981), 203-220.

[8] Harper, L. & G.-C. Rota, *Matching theory: an introcuction,* In: P. Ney (ed.), Adv. in Prob. Vol. 1, Marcel Dekker, 1971, 171-215.

[9] Hazewinkel, M. & C.F. Martin, *Representation of the symmetric groups,* the specialization order, systems and Grassmann manifolds, Ens. Math. 29 (1983), 53-87.

[10] Horn, A., *Doubly stochastic matrices and the diagonal of a rotation matrix,* Amer. J. Math. 76 (1954), 620-630.

[11] Kraft, H.-P., *Conjugacy classes and Weyl group representations,* Astérisque 87/88 (1981), 191-206.

[12] Kostant, B., *On convexity, the Weyl group and the Iwasawa decomposition,* Ann. Ec. Norm. Sup. (4) 6 (1973), 413-455.

[13] Marshall, A.W. & I. Olkin, *Inequalities: theory of majorization and its applications,* Acad. Pr., 1979.

[14] Picci, G. & J. van Schuppen, *On the weak finite stochastic realization problem,* Preprint BW 184/83, CWI, Amsterdam.

[15] Schur, I., *Uber eine Klasse von Mittelbildungen mit Anwendungen auf der Determinantentheorie,* Sitzungsberichte Berliner Math. Gesellschaft 22 (1923), 9-20.

[16] Schwarz, S., *On the structure of the semigroup of stochastic matrices,* Publ. Math. Inst. Hung. Acad. Sci., Ser. A, 9 (1964), 297-311.

[17] Sinkhorn, R., *Diagonal equivalence to matrices with perscribed row and column sums. II,* Proc. Amer. Math. Soc. 45 (1977), 195-198.