STOCHASTIC REALIZATION OF FINITE-VALUED PROCESSES AND PRIMES IN THE POSITIVE MATRICES

G. Picci and J.H. van Schuppen

ABSTRACT

The stochastic realization problem of finite-valued processes asks for the classification of all minimal finite stochastic systems such that the output process of such a system equals a given process in distribution. This problem is motivated by the use of stochastic models in signal processing, communication, and control. The problem of characterizing the minimal number of states in this problem leads to a factorization problem for positive matrices and hence to the study of positive linear algebra. The results of the paper are a characterization of primes in the positive matrices and examples of such primes.

Key Words: Stochastic realization, finite stochastic system, prime in the positive matrices.

1 INTRODUCTION

The aim of this paper is to show that the stochastic realization problem of finite-valued processes reduces to the problem of factorization of positive matrices. The solution of the latter problem leads to a study of positive linear algebra. The results of the paper concern the classification of primes in the positive matrices.

The motivation of the stochastic realization problem of finite-valued processes is the use of stochastic models in signal processing, communication, and control. For these research areas, stochastic models with finite-valued variables arise naturally or are useful approximate models. For example, in speech processing a ***hidden Markov model*** is used, see the tutorial paper by L.R. Rabiner [7]. Use of stochastic models with finite-valued processes leads to the system identification problem: how, given data for the output process, to determine the parameters of the model such that this model is a good approximation to the data? This problem requires the solution of the stochastic realization problem for finite-valued processes. The results of the corresponding problem for Gaussian processes have proven to be highly useful for system identification in engineering, biomathematics, and economics.

A *finite stochastic system* is defined below as a pair of finite-valued stochastic processes, called the *state process* and the *output process*, satisfying a conditional independence condition. A *weak stochastic realization* of a given process is a stochastic system such that its output process equals the given process in distribution. A stochastic realization of a finite-valued process is called *minimal* if the number of states in the state space is minimal. The *weak finite stochastic realization problem* of a finite-valued process is to show existence of a weak stochastic realization in the form of a finite stochastic system and to classify all minimal stochastic realizations.

An early reference on this problem is [3]. The current status of the finite stochastic realization problem is that it is unsolved. During the 1960's several publications appeared that provide a necessary and sufficient condition for the existence of a finite stochastic realization, see [4, 6]. Unsolved questions are the characterization of minimality of the state space and the classification of all minimal stochastic realizations. The main bottleneck is currently the characterization of minimality of the state space. This point leads to a factorization problem for positive matrices. The solution of the latter problem requires a study of positive linear algebra. To this end the concept of a prime in the positive matrices is used. The results of the paper are contributions to the characterization of primes in the positive matrices. The proofs of these results are deferred to a future publication.

The outline of the paper is as follows. The stochastic realization problem for finite-valued processes is formulated in section 2. In section 3 the factorization problem for positive matrices is posed and primes in the positive matrices are characterized.

2 WEAK STOCHASTIC REALIZATION OF A FINITE VALUED PROCESS

In this paper the set $R_+ = [0, \infty)$ is called the set of the *positive real numbers* and $(0, \infty)$ the set of the *strictly positive real numbers*. Let $Z_+ = \{1, 2, \ldots\}$ denote the set of the *positive integers* and $N = \{0, 1, \ldots\}$ the set of the *natural numbers*. For $n \in Z_+$ let $Z_n = \{1, 2, \ldots, n\}$. Denote by R_+^n the set of n-tuples of the positive real numbers. Denote the *simplex* in R_+^n by

$$S_+^n = \{x \in R_+^n | \sum_{i=1}^{n} x_i = 1\}.$$

The set $R_+^{n \times n}$ will be called the set of the *positive matrices*. Both R_+ and $R_+^{n \times n}$ are examples of a *semi-ring*. They are not rings because they do not admit an inverse with respect to addition. For the same reason R_+^n is not a vector space. The structure $(R_+, R_+^{n \times n})$ may be defined as a *left semi-module*.

Let (Ω, F, P) be a probability space consisting of a set Ω, a σ-algebra F, and a probability measure P.

Definition 2.1 *A (discrete-time)* finite stochastic system *is a collection*

$$\{\Omega, F, P, T, Y, X, y, x\}$$

where (Ω, F, P) *is a probability space,* $T = Z$ *an index set,* $Y = Z_m$, $X = Z_n$ *for* $m, n \in Z_+$ *are called respectively the* output space *and the* state space, $y : \Omega \times T \to Y$, $x : \Omega \times T \to X$ *are stochastic processes called respectively the* output process *and the* state process *such that for all* $t \in T$

$$(F_t^{y+} \vee F_t^{x+}, F_t^{x-} \vee F_{t-1}^{y-} | F^{x(t)}) \in CI, \tag{1}$$

meaning that the future of the output and state process at time t *and the past of these processes at time* t *are conditional independent given the state at time* t. *Here*

$$F_t^{x+} = \sigma(\{x(s), \forall s \geq t\}), F_t^{x-} = \sigma(\{x(s), \forall s \leq t\}).$$

A finite stochastic system is called stationary *if the joint process* (x, y) *is stationary.*

The above definition implies that the state process is a Markov process. The output of a finite stochastic system is also called a *hidden Markov process* or a *probabilistic function of a Markov chain.*

Consider a finite stochastic system with $Y = Z_m$. For $t \in Z_+$ denote $v \in Y^t$ by $v = (v_1, v_2, \ldots, v_t)$. Let Y^* denote the set of finite sequences with values in Y,

$$Y^* = \{v | v \in Y^t \text{ for some } t \in Z_+\}.$$

For $v, u \in Y^*$ denote by $uv \in Y^*$ the sequence u followed by the sequence v. Let

$$C(Y^*) = \{ \text{ all probability measures on } Y^*\}.$$

Note that $C(Y^*)$ is a convex set. For $p \in C(Y^*)$ let

$$C_p = \text{ conv } \{p(.|u) | \forall u \in Y^*\}$$

where *conv* denotes the convex hull generated by the specified conditional probability distributions.

Problem 2.2 *The* weak finite stochastic realization problem *for a finite-valued process. Let Y be a finite set, say $Y = Z_m$ for $m \in Z_+$. Let $p \in C(Y^*)$ be a probability distribution on Y^* corresponding to a stationary process.*

a. *Which conditions are necessary and sufficient for the existence of a stationary finite stochastic system such that the distribution of the output process of this system equals the given distribution? If such a system exists it is called a* weak stochastic realization *of the given distribution.*

b. *Determine the minimal number of states of a weak stochastic realization.*

c. *Classify all weak stochastic realizations of the given distribution for which the state space is minimal.*

Theorem 2.3 *[4, 6] Let $p \in C(Y^*)$ be a probability distribution corresponding to a stationary process. There exists a stationary finite stochastic system with state space $X = Z_n$ which is a weak stochastic realization of p iff there exists a convex polyhedral set $C_1 \subset C(Y^*)$ such that:*

1. $C_p \subset C_1$;

2. *C_1 is generated by n distributions, or*

$$C_1 = \textit{ conv } \{q_i(.), i \in Z_n\};$$

3. *C_1 is closed with respect to conditioning, or for any $i \in Z_n$*

$$q_i(.|y) = q_i(.y)/q_i(y) \in C_1.$$

The theorem presented above solves the existence part of problem 2.2. The parts b and c of this problem are unsolved. Part b asks for the characterization of the minimal number of states of a stochastic realization. As argued in [6] this question may be related to the factorization problem of a Hankel matrix into the product of two positive matrices. Such a factorization problem is explored in the next section.

The realization problem for deterministic positive linear systems is closely related to the stochastic realization problem of finite-valued processes. A *finite-dimensional positive linear system* is a linear dynamical system in which the input, state, and output space are spaces over the positive real numbers, say $X = R^n_+$. Systems in this class are useful models in biomathematics, where they are called *linear compartmental systems*, economics, chemometrics, and other research areas. A recent book on this class of systems is [1]. The realization problem of this class is also unsolved, it leads to a factorization problem for positive matrices that is identical to that for the weak finite stochastic realization problem. For a special case of the realization problem for deterministic continuous-time positive linear systems see [5]. Identifiability of compartmental systems is discussed in [9]. There exists an example that shows that the conditions for minimality of a finite-dimensional linear system are not necessary for the minimality of a finite-dimensional positive linear system.

3 POSITIVE LINEAR ALGEBRA

3.1 FACTORIZATION OF POSITIVE MATRICES

As argued in the previous section, the stochastic realization problem of finite-valued processes leads to a factorization problem of a positive matrix. This problem is defined below.

Definition 3.1 *A* positive matrix factorization *of $A \in R_+^{k\times m}$ is a pair $(B, C) \in R_+^{k\times n} \times R_+^{n\times m}$ such that*

$$A = BC. \tag{2}$$

The minimal $n \in N$ for which such a factorization exists will be called the positive rank *of A and denoted by pos-rank(A). A* minimal positive matrix factorization *is a positive matrix factorization in which $n =$ pos-rank (A).*

Problem 3.2 *The* positive matrix factorization problem. *Let $A \in R_+^{k\times m}$. Determine the positive rank of A and classify all its minimal positive matrix factorizations.*

This problem is unsolved. Its solution requires a study of positive linear algebra. Such a study is started in the next subsection.

3.2 PRIMES IN THE POSITIVE MATRICES

The set of the positive matrices may be considered from several view points, such as: a set of matrices, an algebraic structure, and a geometric structure. The set of positive matrices is a semi-ring. This structure differs from a ring in that it does not have an inverse with respect to addition. A positive matrix may be associated with a convex polyhedral cone. This geometric interpretation will not be explored below because of space limitations.

Classes of positive matrices are introduced next. The set of *permutation matrices* in $R_+^{n\times n}$ is denoted by $P_+^{n\times n}$. The set of *diagonal matrices* in $R_+^{n\times n}$ is denoted by $D_+^{n\times n}$. A *strictly positive diagonal matrix* is a diagonal matrix whose diagonal elements are striclty positive. A *monomial matrix* is a positive matrix such that every row and every column contains exactly one strictly positive element. The set of monomial matrices in $R_+^{n\times n}$ is denoted by $M_+^{n\times n}$. A *doubly stochastic matrix* is a positive matrix such that the sum of the row elements for every row and the sum of the column elements for every column, equals one. The set of doubly stochastic matrices in $R_+^{n\times n}$ is denoted by $DS_+^{n\times n}$. For terminology on positive matrices see [2].

Definition 3.3 *A* prime in the positive matrices *is a positive matrix $A \in R_+^{n\times n}$ such that:*

1. *A is not a monomial matrix;*
2. *if $A = BC$, with $B, C \in R_+^{n\times n}$, then either B or C is a monomial matrix.*

The above definition of a prime has been introduced by D.J. Richman and H. Schneider in 1974 [8]. For an exposition on primes in the positive matrices see [2, section 3.4]. It may be shown that a positive matrix $A \in R_+^{n\times n}$ has an inverse in $R_+^{n\times n}$, or $A^{-1} \in R_+^{n\times n}$, iff A is a monomial matrix. The monomial matrices are therefore the group of units in the positive matrices. The concept of a prime in the positive matrices thus agrees with the algebraic definition.

Problem 3.4 *Classify all primes in the positive matrices.*

Results on the classification of primes in the positive matrices are summarized below.

Theorem 3.5 *a. The matrix $A \in R_+^{n\times n}$ is a prime in the positive matrices iff*

$$A = M_1 \begin{pmatrix} S & 0 \\ 0 & I \end{pmatrix} M_2, \tag{3}$$

with $M_1, M_2 \in M_+^{n\times n}$, $n_1, n_2 \in N$, $n_1 + n_2 = n$, $S \in DS_+^{n_1\times n_1}$ is fully indecomposable [2] and a prime in the positive matrices, and $I \in R^{n_2\times n_2}$ is the identity matrix.

b. If

$$A = M_1 \begin{pmatrix} S_1 & 0 \\ 0 & I \end{pmatrix} M_2 = M_3 \begin{pmatrix} S_2 & 0 \\ 0 & I \end{pmatrix} M_4,$$

are two factorizations as displayed in (3) with $S_1 \in DS_+^{n_1\times n_1}$ and $S_2 \in DS_+^{n_3\times n_3}$ then $n_1 = n_3$ and

$$S_1 = P_1 S_2 P_2$$

for $P_1, P_2 \in P_+^{n_1\times n_1}$. Thus in (3) A determines S up to permutation equivalence.

Theorem 3.5 reduces the classification of primes in the positive matrices to the classification of fully indecomposable doubly stochastic matrices which are primes in the positive matrices.

The structure $(DS_+^{n\times n}, \cdot, I)$ is also a monoid. The group of units in this structure is the set of permutation matrices. One may then define a prime in the doubly stochastic matrices analogously to that of a prime in the positive matrices, see definition 3.3.

If a doubly stochastic matrix is a prime in the positive matrices then it also a prime in the doubly stochastic matrices. The converse of this statement is conjectured and it seems to be true in several cases. The classification of primes in the doubly stochastic matrices has been developed to quite an extent. Because of space limitations the result is not stated here. Below follow examples of primes in the positive matrices.

Proposition 3.6 *a. The matrix $A \in R_+^{3\times 3}$ is a prime in the positive matrices iff*

$$A = M_1 \begin{pmatrix} 0 & 1-s & s \\ s & 0 & 1-s \\ 1-s & s & 0 \end{pmatrix} M_2, \tag{4}$$

for $M_1, M_2 \in M_+^{3\times 3}$, and $s \in (0,1)$.

b. The following positive matrices are prime in the positive matrices:

$$M_1 \begin{pmatrix} s & 0 & 0 & 1-s \\ 1-s & s & 0 & 0 \\ 0 & 1-s & s & 0 \\ 0 & 0 & 1-s & s \end{pmatrix} M_2 \tag{5}$$

for $M_1, M_2 \in M_+^{4\times 4}$, $s \in (0,1)$,

$$M_3 \begin{pmatrix} 0 & 1-s & s & 0 \\ s & 0 & 1-s & 0 \\ 1-s & s & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} M_4, \tag{6}$$

for $M_3, M_4 \in M_+^{4\times 4}$, $s \in (0,1)$.

References

[1] A. Berman, M. Neumann, and R.J. Stern. *Nonnegative matrices in dynamic systems.* John Wiley & Sons, New York, 1989.

[2] A. Berman and R.J. Plemmons. *Nonnegative matrices in the mathematical sciences.* Academic Press, New York, 1979.

[3] D. Blackwell and L. Koopmans. On the identifiability problem for functions of finite markov chains. *Ann. Math. Statist.*, 28:1011–1015, 1957.

[4] A. Heller. On stochastic processes derived from markov chains. *Ann.Math. Statist.*, 36:1286–1291, 1965.

[5] Y. Ohta, H. Maeda, and S. Kodama. Reachability, observability, and realizability of continuous positive systems. *SIAM J. Control Optim.*, 22:171–180, 1984.

[6] G. Picci. *On the internal structure of finite-state stochastic processes*, volume 162 of *Lecture Notes in Economics and Mathematical Systems*, pages 288–304. Springer-Verlag, Berlin, 1978.

[7] L.R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *Proc. IEEE*, 77:257–286, 1989.

[8] D.J. Richman and H. Schneider. Primes in the semigroup of non-negative matrices. *Linear and Multilinear Algebra*, 2:135–140, 1974.

[9] E. Walter. *Identifiability of state space models*, volume 46 of *Lecture Notes in Biomathematics*. Springer-Verlag, Berlin, 1982.