ANNALES DE L'I. H. P., SECTION B

FRANTIŠEK MATUS

Combining *m*-dependence with markovness

Annales de l'I. H. P., section B, tome 34, nº 4 (1998), p. 407-423
<http://www.numdam.org/item?id=AIHPB_1998__34_4_407_0>

Ann. Inst. Henri Poincaré,

Vol. 34, n° 4, 1998, p. 407-423. *Probabilités et Statistiques*

Combining m-dependence with Markovness

by

František MATÚŠ

Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
and
Laboratory of Intelligent Systems
University of Economics, Prague

ABSTRACT. – Generally, no stationary sequence of random variables which is Markov of order n but not of order $n-1$ and m-dependent but not $(m-1)$-dependent exists if the state space of the sequence has small cardinality. We show that to ensure the existence for the Markov sequences of order $n = 1$ the number of attainable states must be at least $m+2$ and that this bound is tight. Given a small state space such a sequence exists only for special n and m. On a two-element state space the smallest possible n and m are shown to be 3 and 2, respectively. This results from our parametric description of all binary m-dependent sequences, $m \geq 0$, that are Markov of order 3. © Elsevier, Paris

RÉSUMÉ. – Si l'espace d'états n'est pas suffisamment riche on ne peut pas construire, pour n et m quelconques, une suite aléatoire stationnaire de Markov d'ordre n et pas d'ordre $n-1$ qui est dans le même temps m-dépendante et pas $(m-1)$-dépendante. Nous montrons que pour les chaînes de Markov, $n = 1$, l'espace d'états doit avoir au moins $m+1$ éléments et que ce nombre ne peut pas être amélioré. Pour les suites binaires les plus petits n et m admissibles sont 3 et 2, respectivement. C'est une conséquence

AMS 1991 subject classification. Primary: 60J10; secondary: 60K99, 60G99.

Key words and phrases. Markov chains, m-dependence, sequences with memory, stationary sequences, binary sequences, conditional independence structures.

This research was partially supported by grants No. 275105 and 27564 of Academy of Sciences of the Czech Republic and by the grant No. VS 96008 of Ministry of Education of Czech Republic.

de notre description paramétrique de toutes les suites binaires stationnaires m-dépendantes, $m \geq 0$, de Markov d'ordre 3. © Elsevier, Paris

1. INTRODUCTION

Let $\xi = (\xi_i;\ i \geq 1)$ be a strictly stationary sequence of random variables taking values in a finite state space S; speaking about a sequence we will always assume these properties. The sequence ξ is *Markov of order* $n \geq 0$ if $(\xi_i;\ 1 \leq i \leq k)$ is conditionally independent of $(\xi_i;\ i \geq k+n+1)$ given $(\xi_i;\ k+1 \leq i \leq k+n)$ for all $k \geq 1$. The sequence ξ is *dependent of order* $m \geq 0$ if $(\xi_i;\ 1 \leq i \leq k)$ is unconditionally independent of $(\xi_i;\ i \geq k+m+1)$, $k \geq 1$. For simplicity, we shorten the expressions "Markov of order n" and "dependent of order m" to n-Markov and m-dependent, correspondingly.

The aim of this note is to examine how these two properties interfere under restrictions on the cardinality of the state space. A more precise formulation will use the following notion of an index of a sequence. Let n_ξ be the smallest nonnegative integer n such that a sequence ξ is n-Markov, let m_ξ be the smallest $m \geq 0$ such that ξ is m-dependent and let d_ξ be the cardinality of the set of states which are attained with positive probabilities. Thus, we have $n_\xi \geq 0$, $m_\xi \geq 0$ and $d_\xi \geq 1$ with $n_\xi = 0$ if and only if $m_\xi = 0$. This expresses the sequence ξ is i.i.d. If ξ is Markov of no order $n \geq 0$ it is reasonable to write $n_\xi = \infty$ and similarly with the dependence and m_ξ; we shall, however, not deal with these cases at all. The triple $\langle n_\xi, m_\xi, d_\xi \rangle$ will be called *index* of ξ.

A natural question asks which triple can be equal to the index of a sequence ξ. In other words, given a triple of integers $\langle n, m, d \rangle$ does there exist a (stationary) sequence ξ such that $n_\xi = n$, $m_\xi = m$ and $d_\xi = d$, i.e. in the nontrivial case $n > 0$ and $m > 0$, such that it is n-Markov and not $(n-1)$-Markov, m-dependent and not $(m-1)$-dependent and takes exactly d states with positive probabilities?

We present answers only if $n = 1$ and partially if $d = 2$ here. In the second section devoted to the usual Markov chains (1-Markovness) we prove that a triple $\langle 1, m, d \rangle$ is the index of a sequence if and only if $1 \leq m \leq d-2$. Then we turn our attention entirely to the binary sequences, $S = \{0, 1\}$, and during a technical preparation in the third section we reveal that every $(n, m, 2)$-sequence, this is an abbreviation for n-Markov, m-dependent and two-element state space, is i.i.d. provided $n \leq 2$ or $m \leq 1$. For some years I conjectured this be valid for any n and m nonnegative. It is, however, not the case. We will see in Section 4 that all $(3, m, 2)$-sequences are 3-dependent and therefore the only two candidates for indices of 3-Markov binary sequences are $\langle 3, 2, 2\rangle$ and $\langle 3, 3, 2\rangle$. Both these triples are really indices and, moreover, we provide a complete characterization of the distributions of all $(3, 2, 2)$-sequences in the fifth section and all $(3, 3, 2)$-sequences in the sixth section, respectively.

Though Markov chains is an old topic, Markov chains with 1-dependence appeared for the first time in [1] and then in [2],[6] where the focus was on the structure of block-factors. Notes on binary sequences of this type are in [11] and [12]. Our question is akin to the problems around probabilistic conditional independence structures [8]; [4] and [5] settle an unconditional case for sequences of random variables. The lattest review of the field is in [7]. It is also worthwhile to mention the paper [9] where Markovness was combined with m-$\underline{in}$dependence. That means any m variables of ξ are mutually independent.

2. MARKOV SEQUENCES OF FIRST ORDER

It is not unexpected that a solution of our problem for 1-Markov sequences will be based on an analysis of transition matrices. Let us remind that a sequence ξ with the state space $S = \{1, 2, \ldots, d\}$, $d = d_\xi$, is 1-Markov if and only if the probability of every event $\xi_1 = s_1 \cdots \xi_{k+1} = s_{k+1}$, denoted by $[s_1 \ldots s_{k+1}]$, is equal to $[s_1] p_{s_1 s_2} \cdots p_{s_k s_{k+1}}$, $k \geq 1$, where $[s] > 0$ is the probability of $\xi_1 = s$ and $p_{s,t}$ is the conditional probability of $\xi_2 = t$ given $\xi_1 = s$, $s, t \in S$. The (s, t)-entry of the k-th power of the transition matrix $\mathbf{P} = (p_{s,t};\ 1 \leq s, t \leq d)$ contains the conditional probability of $\xi_{k+1} = t$ given $\xi_1 = s$, $k \geq 1$.

If the sequence ξ is, moreover, m-dependent, $m \geq 0$, then ξ_1 is independent of ξ_{m+2} and $\mathbf{P}^{m+1} = \mathbf{Q}$ where the matrix $\mathbf{Q}$ has constant columns, t-th one containing the probability $[t]$, $t \in S$. It is not difficult to see that, on contrary, this matrix equality implies that ξ is m-dependent. In fact, it implies ξ_k is independent of ξ_{k+m+1} what together with the conditional independence of ξ_k and $(\xi_i;\ i \geq k+m+2)$ given ξ_{k+m+1} yield ξ_k is independent of $(\xi_i;\ i \geq k + m + 1)$. Repeating the same reasoning once again we obtain the desired m-dependence.

LEMMA 1. – *If a Markov sequence of first order with d-element state space, $d \geq 2$, is m-dependent, $m \geq 0$, then it is $(d-2)$-dependent.*

This assertion is nontrivial for $m > d-2$ as it provides reduction of the order of dependence due to "small" state space.

Proof. – Knowing that $\mathbf{P}^{m+1} = \mathbf{Q}$ we deduce that the spectra of both matrices are equal to $\{0, 1\}$. Since $\mathbf{P}$ is primitive (all entries of some of its powers are positive) the number 1 is an eigenvalue of $\mathbf{P}$ with algebraic multiplicity one, see [10]. We can write $\mathbf{P} = \mathbf{T}\mathbf{W}\mathbf{T}^{-1}$ where $\mathbf{T}$ is a regular matrix and $\mathbf{W}$ is the Jordan canonical form of $\mathbf{P}$, see [3]. The matrix $\mathbf{W}$ is block-diagonal. One of the blocks consists of the single eigenvalue 1 and the remaining blocks have zeros on their diagonals and ones on their superdiagonals. The k-th power of such a block of size $b \times b$, $b \geq 2$, is a zero matrix once $k \geq b$. Thus, we can conclude that each matrix $\mathbf{W}^k$, $k \geq d-1$, has only one nonzero entry; obviously it is the eigenvalue 1. Now, if $m+1 \geq d-1$ then $\mathbf{W}^{m+1} = \mathbf{W}^{d-1}$ and consequently $\mathbf{P}^{d-1} = \mathbf{P}^{m+1} = \mathbf{Q}$ what means that the examined sequence is $(d-2)$-dependent. ∎

COROLLARY 1. – *Every (n, m, d)-sequence is, $d \geq 2$, dependent of order $(d^n - n - 1)$.*

Proof. – If ξ fulfils the assumptions, $n > 0$, we consider the sequence $\eta = (\eta_i;\, i \geq 1)$ of the random variables $\eta_i = (\xi_i, \ldots, \xi_{i+n-1})$. This sequence is obviously 1-Markov, $(n+m-1)$-dependent and $d_\eta \leq d_\xi^n \leq d^n$. By Lemma 1 it is $(d_\eta - 2)$-dependent what implies that ξ is dependent of order $(d^n - 2 - (n-1))$. ∎

PROPOSITION 1. – *A triple of integers $\langle 1, m, d\rangle$ is the index of a sequence if and only if $1 \leq m \leq d-2$.*

Proof. – The necessity of the presented condition is a consequence of the previous lemma and the sufficiency will be approved below by a construction of the desired sequences.

Let $\mathbf{x}_1, \ldots, \mathbf{x}_d$ be an arbitrary orthonormal base of the Euclidean space $\mathcal{R}^d$ such that $\mathbf{x}_1$ has all coordinates equal to $d^{-1/2}$. These vectors are taken as rows and their transpositions, columns, are obtained by using the superindex T. For example, $\mathbf{Q} = \mathbf{x}_1^T\mathbf{x}_1$ is a doubly stochastic matrix. If we set $\mathbf{U}_k = \sum_{j=2}^{k}\mathbf{x}_j^T\mathbf{x}_{j+1}$ for $1 \leq k \leq d-1$ then the powers of these matrices are $\mathbf{U}_k^\ell = \sum_{j=2}^{k+1-\ell}\mathbf{x}_j^T\mathbf{x}_{j+\ell}$, $1 \leq \ell \leq k$. This fact can be obtained by a simple induction argument. Note that $\mathbf{U}_k^k$, $1 \leq k \leq d-1$, are zero matrices and that for $\ell < k$ the matrix $\mathbf{U}_k^\ell$ is nonzero owing to $\mathbf{x}_2\mathbf{U}_k^\ell = \mathbf{x}_{2+\ell}$. In addition, $\mathbf{Q}\,\mathbf{U}_k^\ell = \mathbf{U}_k^\ell\,\mathbf{Q}$ are zero matrices for $1 \leq \ell \leq k \leq d-1$, too.

Now, let us have a triple $\langle 1, m, d\rangle$ and $1 \le m \le d-2$. The 1-Markov sequence ξ on a d-element state space with the transition matrix $\mathbf{P} = \mathbf{Q} + \varepsilon\,\mathbf{U}_{m+1}$, $\varepsilon \neq 0$ sufficiently small, and the uniform initial distribution is stationary because $\mathbf{x}_1 \mathbf{P} = \mathbf{x}_1$. In addition, $\mathbf{P}^k = \mathbf{Q} + \varepsilon^k\,\mathbf{U}_{m+1}^k$, $1 \le k \le m+1$, what enables to conclude that ξ is m-dependent but not $(m-1)$-dependent, i.e. $n_\xi = 1$, $m_\xi = m$ and $d_\xi = d$. ■

3. BINARY SEQUENCES: PRELIMINARIES

From now on we fix the state space as $S = \{0, 1\}$. States $s_k, \ldots, s_\ell$ from S, $k \le \ell$, will be concatenated into words and the word $s_k\, s_{k+1} \cdots s_\ell$ will be shortened to s_k^ℓ. The symbol S^k denotes the set of all words made of letters from S which have the length k, $k \ge 0$, e.g. $s_1^k \in S^k$ and $s_1^0 \in S^0$ is the empty word.

A sequence ξ is n-Markov, $n \ge 0$, if and only if for all $k \ge 0$ and $s_1^{n+k} \in S^{n+k}$ the probability $[s_1^{n+k}]$ of the event $\xi_1 = s_1 \cdots \xi_{n+k} = s_{n+k}$ can be factorized as follows

$$[s_1^{n+k}] = [s_1^n] \textstyle\prod_{j=1}^k (s_j^{j+n}).$$

In this formula the numbers (s_j^{j+n}), conditional probabilities, are defined by the equalities $[s_1^{n+1}] = [s_1^n]\,(s_1^{1+n})$. If $[s_1^n] = 0$ the choice of (s_1^{1+n}) is arbitrary and will not affect our next computations.

An n-Markov sequence is m-dependent, $m \ge 0$, if and only if for all $s_1^n,\ s_{n+m+1}^{n+m+n} \in S^n$ the following equality

$$\begin{aligned} 0 &= \square_{n,m}(s_1^n,\ s_{n+m+1}^{n+m+n}) = \\ &= [s_1^n] \left([s_{n+m+1}^{n+m+n}] - \textstyle\sum_{s_{n+1}^{n+m} \in S^m} \prod_{j=1}^{n+m} (s_j^{j+n}) \right) \end{aligned}$$

takes place. That means $(\xi_{k+i};\, 1 \le i \le n)$ is independent of $(\xi_{k+n+m+i};\, 1 \le i \le n)$, $k \ge 0$, and this fact implies m-dependence in a similar way as was done above with $n = 1$. We shall also need the symbol $\square^*_{n,m}(s_1^n, s_{n+m+1}^{n+m+n})$ denoting the difference in parentheses. Sometimes an argument in $\square_{n,m}(s_1^n, \cdot)$ is omitted to work with a function on S^n.

An n-Markov sequence, $n \ge 1$, is $(n-1)$-Markov if and only if

$$0 = \triangle_{n-1}(s_2^n) = [0\, s_2^n\, 0]\,[1\, s_2^n\, 1] - [0\, s_2^n\, 1]\,[1\, s_2^n\, 0]$$

for all $s_2^n \in S^{n-1}$. The equalities express that the variable ξ_k is independent of ξ_{k+n} given $(\xi_{k+i};\, 1 \le i \le n-1)$. We shall also need the symbol

$\triangle^*_{n-1}(s_2^n)$ denoting the above difference with the brackets replaced by parentheses. Thus, $\triangle_{n-1}(s_2^n) = [0\, s_2^n]\,[1\, s_2^n]\,\triangle^*_{n-1}(s_2^n)$.

Beside the foregoing basic observations we want to summarize and label some other useful facts concerning $(n, m, 2)$-sequences, n, $m \geq 1$, $(k \geq 0)$

1. $\square_{n,m-1}(s_2^n\, 0, \cdot) + \square_{n,m-1}(s_2^n\, 1, \cdot) = 0$
2. $[0\, s_2^n]\, \square_{n,k}(1\, s_2^n, \cdot) = [1\, s_2^n]\, \square_{n,k}(0\, s_2^n, \cdot)$ if $\triangle_{n-1}(s_2^n) = 0$
3. $\square^*_{n,m-1}(s_2^n\, 0, \cdot) = \square^*_{n,m-1}(s_2^n\, 1, \cdot) = 0$ if $\triangle_{n-1}(s_2^n) \neq 0$
4. $\square_{n,m-1}(\cdot, 0\, s_2^n) = \square_{n,m-1}(\cdot, 1\, s_2^n) = 0$ if $\triangle^*_{n-1}(s_2^n) \neq 0$

Some comments are in order. The expression in 1. equals the sum of $\square_{n,m}(0\, s_2^n, \cdot)$ and $\square_{n,m}(1\, s_2^n, \cdot)$. The validity of 2. is clear if $[0\, s_2^n]$ or $[1\, s_2^n]$ is zero; if they are both positive we use $(0\, s_2^{n+1}) = (1\, s_2^{n+1})$. To see 3. we write

$$0 = \square_{n,m}(s_1^n, \cdot) = \textstyle\sum_{t \in S}[s_1^n\, t]\, \square^*_{n,m-1}(s_2^n\, t, \cdot)$$

and match these equalities into pairs corresponding to $0\, s_2^n$ and $1\, s_2^n$; the assumption $\triangle_{n-1}(s_2^n) \neq 0$ means that the determinant of two equations in a pair is nonzero. Finally, the validity of 4. is obtained similarly from

$$0 = \square_{n,m}(\cdot, s_{n+m+1}^{n+m+n}) = \textstyle\sum_{t \in S}(s_{n+m}^{2n+m})\, \square_{n,m-1}(\cdot, t\, s_{n+m+1}^{2n+m-1}).$$

LEMMA 2. – *If ξ is a $(n, m, 2)$-sequence where $n \leq 2$ or $m \leq 1$ then ξ is i.i.d.*

Proof. – Using Corollary 1 we can restrict ourselves to $m = 1$ and $n \geq 2$. We shall demonstrate by contradiction that $\triangle_{n-1}$ is identically zero and then apply the induction argument.

Let $\triangle_{n-1}(t_2^n) \neq 0$ for some $t_2^n \in S^{n-1}$. By fact 3. we have

$$[s_{n+2}^{2n+1}] = \textstyle\prod_{j=2}^{n+1}(s_j^{j+n})$$

as soon as the word s_2^{2n+1} begins with t_2^n. We multiply both sides by $[s_{n+1}^{2n}]$ and sum over s_{2n+1} what gives

$$[s_{n+1}^{2n}]\,[s_{n+2}^{2n}] = [s_{n+1}^{2n}]\textstyle\prod_{j=2}^{n}(s_j^{j+n}).$$

The conclusion is $[s_{n+1}^{2n+1}]\,[s_{n+2}^{2n}] = [s_{n+1}^{2n}]\,[s_{n+2}^{2n+1}]$ for all $s_{n+1}^{2n+1} \in S^{n+1}$ contradicting the assumption $\triangle_{n-1}(t_2^n) \neq 0$. ∎

4. BINARY 3-MARKOV SEQUENCES

From Corollary 1 we know that every $(3, m, 2)$-sequence is 4-dependent. We will see in a moment that it is even 3-dependent. By Lemma 2 the only nontrivial m's to be examined are then 2 and 3. The aim of this section is to prove some auxiliary results about these two cases.

LEMMA 3. – *For every $(3, m, 2)$-sequence $\triangle_2(01) = 0$ or $\triangle_2(10) = 0$.*

Proof. – Let us suppose that both $\triangle_2(01)$ and $\triangle_2(10)$ are nonzero. By fact 3. we deduce

$$0 = \square^*_{3,m-1}(01t, \cdot) = \square^*_{3,m-1}(10t, \cdot), \quad t \in S.$$

If $\triangle_2(00) = 0$ then by fact 2.

$$0 = [000]\,\square_{3,m-1}(100, \cdot) = [100]\,\square_{3,m-1}(000, \cdot)$$

and since $[100] \neq 0$ (otherwise $\triangle_2(10) = 0$) we employ 1. to obtain

$$0 = \square_{3,m-1}(00t, \cdot), \quad t \in S.$$

If $\triangle_2(00) \neq 0$ we have this equality immediately by 3. The same reasoning applies symmetrically to $\triangle_2(11)$. Thus, we see that the sequence is $(m-1)$-dependent. By induction, it is i.i.d., a contradiction. ■

LEMMA 4. – *A $(3, m, 2)$-sequence is i.i.d. if and only if both numbers $\triangle_2(01)$ and $\triangle_2(10)$ are equal to zero.*

Proof. – One implication is trivial. If $\triangle_2(01) = \triangle_2(10) = 0$ and both $\triangle_2(00)$ and $\triangle_2(11)$ are nonzero we recall 3., 2. and 1. and, similarly as in the proof above, keep lowering of the order of dependence. Hence, by symmetry, let $\triangle_2(00) = 0$. Then $\triangle_2(11) = 0$ would imply the sequence is 2-Markov and by Lemma 2 also i.i.d.

From $\triangle_2(11) \neq 0$ we deduce $\square_{3,m-1}(s1t, \cdot) = 0$ for $s, t \in S$ using 3., 2. and 1. as usually. Further, 1. and 2. enable to write the following four linear equations $(s \in S)$

$$\square_{3,m-1}(s00, \cdot) + \square_{3,m-1}(s01, \cdot) = 0,$$
$$[00s]\,\square_{3,m-1}(10s, \cdot) - [10s]\,\square_{3,m-1}(00s, \cdot) = 0.$$

If the determinant $\triangle_1(0)$ of the system of equations is nonzero then the order of dependence of the sequence decreases. Analogically as soon as

$[00] = 0$. Let $\triangle_1(0) = 0$ and $[00] \neq 0$. Since $[110] \neq 0$ we know that $[s0t] = [s0][0t]/[0] > 0$ for $s,t \in S$ and thus

$$(110s)\,(10st) = (10s)\,(0st) = (0s)\,(00st) = (000s)\,(00st).$$

This equality implies $0 = \square^*_{3,m-1}(110,\cdot) = \square^*_{3,m-1}(000,\cdot)$ and then $0 = \square_{3,m-1}(s0t,\cdot)$ for all $s,t \in S$, having decrease of the order of dependence, too. ∎

LEMMA 5. – *In every* $(3,m,2)$*-sequence* $\triangle_2(00) = 0$ *or* $\triangle_2(11) = 0$.

Proof. – By Lemma 3, Lemma 4 and symmetry we can assume $\triangle_2(10) \neq 0$ and $\triangle_2(01) = 0$. We start from the opposite $\triangle_2(00)\triangle_2(11) \neq 0$ aiming at a contradiction. Note that $[s_1^3]$ is positive for $s_1^3 \in S$. Since $\square_{3,4}(\cdot,\cdot) = 0$ by Corollary 1, one has $\square_{3,3}(s0t,\cdot) = 0$ and $\square_{3,2}(00s,\cdot) = 0$ by means of 3., s, $t \in S$. Owing to 4. $\square_{3,1}(00s,t1u) = 0$ and $\square_{3,0}(00s,t11) = 0$ for any s, t, $u \in S$. The choice $st = 00$ in the latter equality gives $[01] = (0000)(0001)$. Then we substitute $st = 01$ and $st = 11$ and find $(0001) = (1111)$. On the other hand, from $\square_{3,2}(00s,\cdot) = 0$ and $\triangle_2(00) \neq 0$ we have also $\square_{3,1}(00s,t00) = 0$ again by 4., s, $t \in S$. This provides for $st = 01$

$$[100] = (0000)(0001)(010)(0100) + (0001)(011)(0110)(1100)$$

where the left product equals $[0100]$. Thus, $[1100]$ is equal to the right product and then

$$[110] = (0001)(011)(0110).$$

Let us multiply both sides by (0000) whence $(0000)[011] = [0110]$. The contradiction sounds $(0110) = (0000) = (1110)$. ∎

COROLLARY 2. – $\langle 3,4,2\rangle$ *is not an index.*

Proof. – Let a sequence ξ has the index $\langle 3,4,2\rangle$. Then index of the sequence of triples $\eta = ((\xi_i,\xi_{i+1},\xi_{i+2});\ i \geq 1)$ is $\langle 1,6,d\rangle$ whence $d = 8$ by Proposition 1. From the proof of Lemma 1 we know that the transition matrix of η has the rank 7. But, due to Lemma 3 and Lemma 5 this matrix has at least two pairs of equal rows, a contradiction. ∎

LEMMA 6. – *Let* $\triangle_2(10) \neq 0$ *and* $\triangle_2(01) = 0$ *in a* $(3,m,2)$*-sequence* ξ*. Then this sequence is* 2*-dependent if and only if both numbers* $\triangle_2(00)$ *and* $\triangle_2(11)$ *are equal to zero.*

Proof. – Let us observe that $[s_1^3] > 0$ for $[s_1^3] \in S - \{000, 111\}$ and then the 1-Markov sequence η constructed by grouping triples of consequent variables as above has the index $\langle 1, m, d\rangle$ where $6 \leq d \leq 8$. The transition matrix of η has the two rows indexed 001 and 101 identical. If $\triangle_2(ss) = 0$ and $[sss] > 0$ then also the two rows indexed by $0ss$ and $1ss$ coincide, $s \in S$. Hence, the matrix has rank at most 5. Then η is 4-dependent and ξ must be 2-dependent.

On the other hand, let ξ be 2-dependent. If $\triangle_2(00) \neq 0$ then by the fact 3. we have $\square_{3,1}(s0t, \cdot) = 0$ for $s, t \in S$ and then by 3. again $\square^*_{3,0}(00t, \cdot) = 0$ for $t \in S$. We can argue as in the proof of Lemma 2 to arrive at a contradiction and thus necessarily $\triangle_2(00) = 0$. If $\triangle_2(11) \neq 0$ then by 3. we get $\square^*_{3,1}(11t, \cdot) = 0$, $t \in S$, and by the fact 4.

$$\square^*_{3,0}(11t,\ s_5 1 s_7) = 0, \quad t, s_5, s_7 \in S.$$

The choice $ts_5 = 01$ leads to

$$[11s] = (1101)\,(1011)\,(011s), \quad s \in S,$$

and then (add the above equations) to $[11s] = [11]\,(011s)$. Hence $\triangle_2(11)$ equals zero, a contradiction. ■

Under the assumptions of Lemma 6, the sequence ξ is 3-dependent and not 2-dependent if and only if exactly one of the numbers $\triangle_2(00)$ and $\triangle_2(11)$ is equal to zero. This follows from Corollary 1, Corollary 2, Lemma 5 and Lemma 6.

5. $(3, 2, 2)$-SEQUENCES

In this section we will describe parametrically all binary 2-dependent sequences that are Markov of order 3.

PROPOSITION 2. – *Let ξ be a binary 3-Markov sequence such that*

$$\triangle_2(10) \neq 0 = \triangle_2(00) = \triangle_2(01) = \triangle_2(11) = 0.$$

Then ξ is 2-dependent if and only if $[0st1] = [0s][t1]$, $s,\ t \in S$.

Proof. – If ξ is a $(3, 2, 2)$-sequence with $\triangle_2$'s as above then by 3. and 4. we derive

$$\square^*_{3,1}(10s, \cdot) = \square^*_{3,0}(10s,\ t10) = 0, \quad s, t \in S.$$

Since $[s_1^3]$ are positive if s_1^3 is different from 000 and 111 we obtain $[t10] = (0st)(0st1)(st10)$ and for $st \neq 11$ immediately the desired equalities. But,

$$\textstyle\sum_{s,\,t\in S}[0st1] - [0s][t1] = 0$$

by 2-dependence and we have also $[0111] = [01][11]$.

In the opposite direction, we deduce first from $[0ss1] = [0s][s1]$ that the probabilities [000] and [111] are positive, too. Then

$$[s_5 1 s_7 s_8] = (s_2 0 s_4 s_5)\,(0 s_4 s_5 1)\,(s_4 s_5 1 s_7)\,(s_5 1 s_7 s_8)$$

because $(s_2 0 s_4 s_5) = (0 s_4 s_5)$, $(s_5 1 s_7 s_8) = (1 s_7 s_8)$ and $[0 s_4 s_5 1] = [0 s_4][s_5 1]$. We multiply the above equation by $(s_1 s_2 0 s_4)$, sum over s_4 and s_5 and arrive at $\square^*_{3,2}(s_1^2 0,\, 1 s_7^8) = 0$.

Further, we are going to verify the equality $\square^*_{3,2}(s_1^2 1,\, 1 s_7^8) = 0$ for $s_1^2,\, s_7^8 \in S^2$, which is equivalent to

$$[1 s_7^8] - \sum_{s_5\in S} \frac{[s_5 1 s_7 s_8]}{[s_2 1]\,[s_5 1]} \sum_{s_4\in S} [s_2 1 s_4 s_5]\,(1 s_4 s_5 1) = 0.$$

This will be clear if we show that

$$\nabla(s_2, s_5) = [s_2 1]\,[s_5 1] - \textstyle\sum_{s_4\in S}[s_2 1 s_4 s_5]\,(1 s_4 s_5 1) = 0$$

for every s_2 and s_5 from S. But,

$$[01]^{-1}\nabla(0,1) = [11] - [01]\,(011) - [11]\,(111) = 0$$

is straightforward and

$$\nabla(1,0) = [11]\,[01] - [1100]\,(001) - (111)\,[1101] = 0$$

owing to $(001) = (111)$ which is a consequence of

$$[000][11] - [011][00] = [001]^{-1}\big([0001][11][00] - [0011][00][01]\big) = 0.$$

The fact

$$\textstyle\sum_{s_2\in S}\nabla(s_2, s_5) = [1][s_5 1] - \sum_{s_4\in S}[s_4 s_5 1] - [0 s_4 s_5 1] = 0$$

implies that ∇ is identically zero, indeed.

At this moment we know $\square^*_{3,2}(\cdot, 1s_7^8) = 0$ for $s_7^8 \in S^2$. The multiplication by (s_0^3) and the summation over $s_3 \in S$ will give $\square^*_{3,3}(\cdot, 1s_8^9) = 0$. Now, we sum over s_9 and compare the result with $\square^*_{3,2}(\cdot, 11s_8)$. We have $\square^*_{3,2}(\cdot, 01s_8) = 0$. Repeating the same trick we arrive at $\square^*_{3,2}(\cdot, 001) = 0$. But, the sum of $\square^*_{3,2}(\cdot, s_6^8)$ over $s_6^8 \in S^3$ is zero and thus $\square^*_{3,2}(\cdot, 000) = 0$. Since $\square_{3,2}$ is identically zero the sequence ξ is 2-dependent. ■

THEOREM 1. – *Let α, $\beta \in \mathcal{R}$ satisfy the two inequalities*

$$\pm(4\alpha - 3\beta - \beta^3) \leq 1 - \beta^2.$$

The binary 3-Markov sequence $\zeta^{\alpha,\beta}$, which has its distribution of first four variables proportional to the function given by the following table is 2-dependent.

0000	$(1-\beta)^2\,(1+\beta^2-2\alpha)$
1000	$(1-\beta^2)\,(1+\beta^2-2\alpha)$
0100	$(1-\beta)^3\,(1+\beta)+4(1-\beta)\,(\beta-\alpha)$
1100	$(1-\beta^2)^2+8\beta(\beta-\alpha)$
0010	$(1+\beta^2-2\alpha)\,(1-\beta^2+2\beta-2\alpha)$
1010	$(1-\beta^2)^2-4(\alpha-\beta)^2$
0110	$(1-\beta)^2\,(1+\beta^2+2\alpha)$
1110	$(1-\beta^2)\,(1+\beta^2+2\alpha)$
0001	$(1-\beta^2)\,(1+\beta^2-2\alpha)$
1001	$(1+\beta)^2\,(1+\beta^2-2\alpha)$
0101	$(1-\beta^2)^2$
1101	$(1+\beta)^3\,(1-\beta)+4(1+\beta)\,(\alpha-\beta)$
0011	$(1+\beta^2)^2-4\alpha^2$
1011	$(1+\beta^2+2\alpha)\,(1-\beta^2+2\alpha-2\beta)$
0111	$(1-\beta^2)\,(1+\beta^2+2\alpha)$
1111	$(1+\beta)^2\,(1+\beta^2+2\alpha)$

Every $(3,2,2)$-sequence ξ is equal in distribution to some $\zeta^{\alpha,\beta}$ or to the sequence obtained from some $\zeta^{\alpha,\beta}$ by interchanging zeros and ones.

Proof. – The proportionality factor is $1/16$. The conditions imposed on α and β restrict the parameters to be between -1 and 1, see Figure 1, and

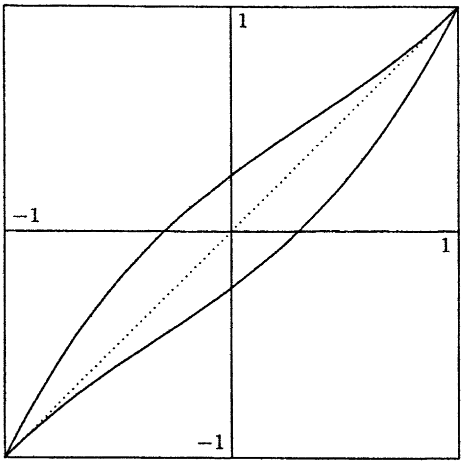

Fig. 1. – Parameters (α, β) are from the region bordered by two cubic curves.

guarantee that all entries of the table are nonnegative; notably the critical inequalities are $[0100] \geq 0$ and $[1101] \geq 0$.

It is easy to verify that $\sum_{t \in S}[t\, s_1^3] - [s_1^3 t] = 0$, $s_1^3 \in S_3$, i.e. that every sequence $\zeta^{\alpha,\beta}$ is strictly stationary. To this end we remark that

$$8\,[000] = (1-\beta)(1+\beta^2-2\alpha) \qquad 8\,[010] = (1-\beta)(1-\beta^2+2\beta-2\alpha)$$
$$8\,[111] = (1+\beta)(1+\beta^2+2\alpha) \qquad 8[101] = (1+\beta)(1-\beta^2+2\alpha-2\beta)$$
$$8\,[100] = (1+\beta)(1+\beta^2-2\alpha) \qquad 8[110] = (1-\beta)(1+\beta^2+2\alpha)$$

and $[100] = [001]$, $[110] = [011]$. It is also not difficult to see that $\zeta^{\alpha,\beta}$, $\alpha \neq \beta$, fulfils the assumptions of Proposition 2, especially

$$(0001) = (1001) = \frac{1+\beta}{2} = (0111) = (1111)$$
$$(0011) = (1011) = \frac{1+\beta^2+2\alpha}{2(1+\beta)}$$

and $\triangle_2(10) \neq 0$. Hence, all sequences $\zeta^{\alpha,\beta}$ are 2-dependent. Note that $\zeta^{\alpha,\alpha}$ are i.i.d., $|\alpha| \leq 1$ (the dotted segment in Figure 1).

In the opposite direction, let ξ be a (3,2,2)-sequence. By Lemmas 3 and 6 we know that up to switching between zeros and ones $\triangle_2(st) = 0$ for $st \neq 10$. If also $\triangle_2(10) = 0$ then ξ is i.i.d. by Lemma 4 and thus equal in distribution to some $\zeta^{\alpha,\alpha}$. If $\triangle_2(10) \neq 0$ then $[0st1] = [0s]\,[t1]$ for $s,\ t \in S$ by Proposition 2. Thus, with $st = 11$ here, the quadratic equation in $[111]$

$$[111]^2 - [11]\,[111] + [11]^2\,[01] = 0$$

has nonnegative discriminant equal to $[11]^2\,(1-4[01])$. We set $\beta^2 = 1-4[01]$ and then we have $2[111] = [11]\,(1+\beta)$. If we take $\alpha = 2[1]-1$ we can compute $4[00] = 1+\beta^2-2\alpha$ and $4[11] = 1+\beta^2+2\alpha$. Using the listed properties of ξ and the stationarity it is easy, but a bit laborious, to compute first the probabilities $[s_1s_2s_3]$ and then to construct the whole table. ■

COROLLARY 3. – *The triple $\langle 3,2,2\rangle$ is the index of $\zeta^{0.1,0}$.*

Remark 1. – Let us mention that the sequence $\zeta^{-\alpha,\alpha}$, $\alpha \neq 0$ small, has the same index as $\zeta^{0.1,0}$. In addition, every its two consequent variables are independent. But, no of the sequences $\zeta^{\alpha,\beta}$, $\alpha \neq \beta$, is 2-independent.

2. From the topological point of view the class of $(3,2,2)$-sequences with the week topology is homeomorphic to two closed circles (disks) pasted together along its diameters; the common diameter corresponds to the i.i.d. sequences and switching between circles to switching between 0 and 1.

3. Reversing time in a $(3,2,2)$-sequence indexed by integer numbers one obtains again a $(3,2,2)$-sequence. In our parametrization this corresponds to the transition $0 \leftrightarrow 1$ and $(\alpha,\beta) \leftrightarrow (-\alpha,-\beta)$ simultaneously.

6. $(3,3,2)$-SEQUENCES

In this section we will describe parametrically all binary 3-dependent sequences that are Markov of order 3. Since all of them which are 2-dependent were investigated in the previous section we concentrate on the non-2-dependent ones. This will close the description of all $(3,m,2)$-sequences.

PROPOSITION 3. – *Let ξ be a binary 3-Markov sequence satisfying $\triangle_2(1s) \neq 0 = \triangle_2(0s)$ for $s \in S$. Then ξ is 3-dependent if and only if $[111] = (001)^3$, $[0s01] = [0s][01]$, $s \in S$, and $(11st) = (00t)$, $s,t \in S$.*

Proof. – Obviously, $[s_1^3] > 0$ for $s_1^3 \neq 000$. If ξ is 3-dependent then by 3. $\square_{3,2}(1st,\cdot) = 0$ and by 4. $\square^*_{3,1}(1st,u1v) = 0$ and $\square^*_{3,0}(1st,u11) = 0$, $s,t,u,v \in S$. For $su = 00$ the last equation gives $[000] > 0$ and $[011] = (0t0)(0t01)(011)$, i.e. $[0t][01] = [0t01]$; as a useful consequence we have $[01] = (000)(001)$. The choice $su = 01$ provides $[01] = (11t0)(1t01)$ what reads as $(000) = (1100)$ if $t = 0$ and as $(000) = (1110)$ if $t = 1$. And finaly, $stu = 111$ leads to $[111] = (1111)^3 = (001)^3$.

For the reverse implication we will first demonstrate $\square^*_{3,0}(1st,u11) = 0$. For $u = 0$ this is equivalent to $[01] = (1st0)(st01)$ which certainly holds if $s = 0$; if $s = 1$ we have $[01] = (000)(001)$. For $u = 1$ we want to see that $[111] = (1st1)(st11)(t111)$ or, rewritten, $[t11] = (1st1)(st11)(111t)$. If $t = 0$ this means $[011] = (001)(011)(000)$. If $t = 1$ this amounts $[s11] = (1s11)(111s)(1111)$ which is fulfilled for $s = 1$ as $(1111) = (001)$ and which holds also for $s = 0$ having $[011] = (011)(000)(001)$.

The next step is to show $\square^*_{3,1}(1st, u1v) = 0$. For $u = 1$ this can be obtained from $\square^*_{3,0}(1st, u'11) = 0$ by multiplication with $(u'11v)$ and summation over u'. For $u = 0$ we want to verify

$$[01] = \textstyle\sum_{v \in S}(1stv)(stv0)(tv01)$$

which is clear for $t = 0$: the product equals $(1s0v)[01]$. If $t = 1$ then we rewrite it into

$$(000) = \textstyle\sum_{v \in S}(1s1v)(s1v0)$$

being satisfied for $s = 1$. For $s = 0$ we have

$$\begin{aligned}(000)[01] &= [0100] + [0110] = [01] - [0101] - [0111]\\ &= [01] - [01]^2 - (1110)[111],\end{aligned}$$

which can be casted into $(001)[01] = [01]^2 + (000)(001)^3$ and $(001) = [01] + (001)^2$.

Knowing that $\square^*_{3,1}(1st, u1v) = 0$ we can obtain $\square^*_{3,2}(1st, 1vu) = 0$ and $\square^*_{3,2}(s1t, u1v) = 0$. These two equalities imply $\sum_{t \in S} \square^*_{3,2}(11s, 00t) = 0$ where both summands must equal zero due to $(001)\square^*_{3,2}(\cdot, 000) = (000)\square^*_{3,2}(\cdot, 001)$ (cf. the fact 2.). Hence, $\square^*_{3,2}(11s, \cdot) = 0$ and then $\square^*_{3,3}(11s, \cdot) = 0$ and $\square^*_{3,3}(s11, \cdot) = 0$ arguing as usually. But owing to the equality $(110t)(10tu) = (s00t)(00tu)$ we have $\square^*_{3,2}(110, \cdot) = \square^*_{3,2}(s00, \cdot)$ whence $\square^*_{3,3}(s00, \cdot) = 0$ and $\square^*_{3,4}(s00, \cdot) = 0$. We can write now

$$0 = \square_{3,4}(000, \cdot) + \square_{3,4}(100, \cdot) = \square_{3,3}(000, \cdot) + \square_{3,3}(001, \cdot)$$

and deduce $\square^*_{3,3}(s01, \cdot) = 0$. Thence $\square_{3,3}(\cdot, \cdot) = 0$ and ξ is 3-dependent. ■

THEOREM 2. – *Let* α, $\beta \in \mathcal{R}$ *satisfy the two inequalities*

$$-(1 - \beta^2)^2 \leq 8\alpha - 8\beta \leq (1 - \beta)^3 (1 + \beta).$$

The binary 3*-Markov sequence* $\theta^{\alpha,\beta}$, *which has its distribution of first four variables proportional to the function given by the following table is* 3*-dependent.*

0000	$(1-\beta)^2\,(1+\beta^2-2\alpha)$
1000	$(1-\beta^2)\,(1+\beta^2-2\alpha)$
0100	$(1-\beta)^3\,(1+\beta)+8(\beta-\alpha)$
1100	$(1-\beta^2)^2+4(1-\beta)\,(\alpha-\beta)$
0010	$(1+\beta^2-2\alpha)\,(1+3\beta-\beta^2+\beta^3-4\alpha)/(1-\beta)$
1010	$(1-\beta^2+2\alpha-2\beta)\,(1+3\beta-\beta^2+\beta^3-4\alpha)/(1-\beta)$
0110	$(1-\beta^2)^2+8(\alpha-\beta)$
1110	$(1-\beta^2)\,(1+\beta)^2$
0001	$(1-\beta^2)\,(1+\beta^2-2\alpha)$
1001	$(1+\beta)^2\,(1+\beta^2-2\alpha)$
0101	$(1-\beta^2)^2$
1101	$(1+\beta)^3\,(1-\beta)+4(1+\beta)\,(\alpha-\beta)$
0011	$(1+\beta^2-2\alpha)\,(1-3\beta-\beta^2-\beta^3+4\alpha)/(1-\beta)$
1011	$(1-\beta^2+2\alpha-2\beta)\,(1-3\beta-\beta^2-\beta^3+4\alpha)/(1-\beta)$
0111	$(1-\beta^2)\,(1+\beta)^2$
1111	$(1+\beta)^4$

Every $(3,3,2)$-sequence ξ which is not 2-dependent equals in distribution to some $\theta^{\alpha,\beta}$, $\alpha \neq \beta$, up to the switching of zeros and ones or up to the time reversal.

Proof. – The proportionality factor is $1/16$. The conditions imposed on α and β restrict the parameters to be strictly between -1 and 1, see Figure 2, and guarantee that all entries of the table are nonnegative (the critical inequalities are $[0100] \geq 0$ and $[0110] \geq 0$). By continuity, $\theta^{1,1}$ will be a constant sequence. The dashed curves in Figure 2 remind the restrictions from Theorem 1 which can be here interpreted as $[010] \geq 0$ and $[011] \geq 0$.

It is easy to verify that the sequences $\theta^{\alpha,\beta}$ are strictly stationary. For this purpose we write down

$8\,[000] = (1-\beta)(1+\beta^2-2\alpha)$ $\quad 8\,[111] = (1+\beta)^3$

$8\,[100] = (1+\beta)(1+\beta^2-2\alpha)$ $\quad 8\,[010] = (1+3\beta-\beta^2+\beta^3-4\alpha)$

$8\,[101] = (1+\beta)(1-\beta^2+2\alpha-2\beta)$ $\quad 8\,[110] = (1-3\beta-\beta^2-\beta^3+4\alpha)$.

From

$$(0001) = (1001) = \frac{1+\beta}{2} = (1101) = (1111)$$

$$(0011) = (1011) = \frac{1-3\beta-\beta^2-\beta^3+4\alpha}{2(1-\beta^2)}$$

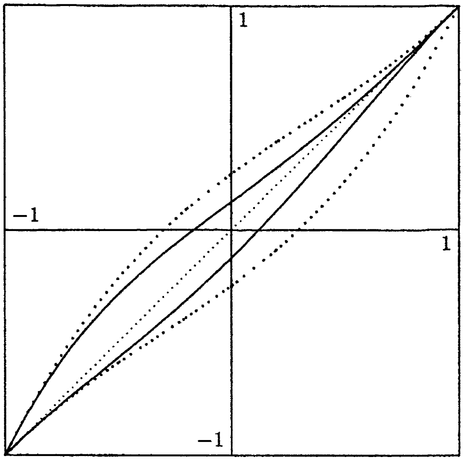

Fig. 2. – Parameters (α, β) are from the region bordered by two biquadratic curves.

we see that the sequences $\theta^{\alpha,\beta}$ satisfy the assumptions of Proposition 3 and are thus 3-dependent.

Let ξ be a (3,3,2)-sequence which is not 2-dependent. Switching zeros and ones and reversing time, if necessary, we know that $\triangle_2(0s) = 0$ and $\triangle_2(1s) \neq 0$, $s \in S$; see the lemmas of Section 4. By Proposition 3, $[000][001] = [0001][00] = [00]^2[01]$, which can be casted into a quadratic equation in [000] similarly as in the proof of Theorem 1. The equation has nonnegative discriminant equal to $[11]^2 (1 - 4[01])$. We set $\beta^2 = 1 - 4[01]$, $\alpha = 2[1] - 1$ and then we can easily compute [000], [001] and [101]. Using $[111] = (011)^3$ we obtain [111], [110] and [010]. From $(11st) = (00t)$ and $(s0tu) = (0tu)$ the whole table can be computed with a little effort. ■

COROLLARY 4. – *The triple* $\langle 3, 3, 2 \rangle$ *is the index of* $\theta^{0,0.1}$.

Remark. – From the topological point of view the union of the classes of $(3, m, 2)$-sequences over $m \geq 0$ finite, with the week topology, is homeomorphic to six closed disks pasted together along its diameters. The common diameter corresponds to the i.i.d. sequences, two disks to the $(3, 2, 2)$-sequences and the remaining four disks without their common diameter to the $(3, 3, 2)$-sequences that are not 2-dependent.

ACKNOWLEDGEMENT

This is to thank to my colleague A. Otáhal for fruitful discussions on the topic. He proposed me, and independently H. Berbee, CWI Amsterdam, the trick with grouping of random variables used in Corollary 1.

REFERENCES

[1] Aaronson J., Gilat D. and Keane M. S., On the structure of 1-dependent Markov shifts, *J. Theor. Probab.* Vol. **5**, 1992, pp. 545-561.
[2] Burton R. M., Goulet M. and Meester R. W. J., On 1-dependent processes and k-block factors, *The Annals of Probability* Vol. **21**, 1993, pp. 2157-2168.
[3] Horn R. A. and Johnson C. R., *Matrix Analysis*, Cambridge University Press, Cambridge.
[4] Matúš F., Independence and Radon projections on compact groups, PhD. theses (in Slovak), Institute of Information Theory and Automation, Prague, 1988.
[5] Matúš F., Stochastic independence, algebraic independence and abstract connectedness, *Theor. Comp. Science* Vol. **134**, 1994, pp. 455-471.
[6] Matúš F., On two-block-factor sequences and one-dependence, *Proc. Amer. Math. Soc.* Vol. **14**, 1996, pp. 1237-1242.
[7] Matúš F., Conditional independence structures examined via minors, *Annals of Mathematics and AI* Vol. **21**, 1997, pp. 99-128.
[8] Pearl J., *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufman, San Meteo, California, 1988.
[9] Rosenblatt M. and Slepian D., Nth order Markov chains with every N variables independent, *J. Soc. Indust. Appl. Math.* Vol. **10**, 1962, pp. 537-549.
[10] Seneta E., *Non-Negative Matrices*, G. Allen & Unwin Ltd., London, 1973.
[11] Valk V. de, One-dependent processes, PhD. theses, Delft Univ. Technology, Delft Univ. Press, 1988.
[12] Valk V. de, One-dependent processes: two-block-factors and non-two-block-factors, CWI Tract 85, Amsterdam, 1994.

(Manuscript received on July 10 1995;
Revised on April 29, 1997.)