Fuzzy Markov Chains

Kostya E. Avrachenkov [*] Elie Sanchez [†]

Abstract

General finite state fuzzy Markov chains that are introduced have a finite convergence to a stationary (may be periodic) solution. The Cesaro average and the α-potential for fuzzy Markov chains are defined, then it is shown that the relationship between them corresponds to the Blackwell formula in the classical theory of Markov decision processes. Furthermore, it is pointed out that recurrency does not necessarily imply ergodicity. However, if a fuzzy Markov chain is ergodic, then the rows of its ergodic projection equal the greatest eigen fuzzy set of the transition matrix. Finally, the fuzzy Markov chain is shown to be a robust system with respect to small perturbations of the transition matrix, which is not the case for the classical probabilistic Markov chains.

Keywords: Fuzzy Markov chains, eigen fuzzy sets, ergodicity, robustness.

1 Introduction and basic definitions

In a seminal paper, Bellman and Zadeh [3] first studied a fuzzy decision process. Esogbue and Bellman [7] explored various kinds of fuzzy dynamic programming with finite state spaces and finite action spaces. Fuzzy dynamic programming is different from classical dynamic programming, it uses fuzzy sets instead of reward functions. In [11] Kruse et al introduced fuzzy Markov chains as "a perception of usual Markov chains", based on fuzzy probabilities and, using an ergodic theorem they calculated an effective processor power. Kurano et al [9] studied dynamic fuzzy systems and in [21] Yoshida constructed a Markov fuzzy process, with a transition possibility measure and a general state space. Further, the fuzzy Markov chains have a potential application in fuzzy Markov algorithms proposed by Zadeh in [23].

This paper is a full version of the conference presentation [2].

It will be introduced now the definition of general finite-state fuzzy Markov chains, which will be compared with classical Markov chains based on probability theory. Let $S = \{1, ..., n\}$ be a finite state space.

[*]INRIA Sophia Antipolis, 2004 route des Lucioles, B.P.93, 06902, France, e-mail: k.avrachenkov@sophia.inria.fr

[†]Laboratoire d'Informatique Médicale, Faculté de Médecine, 27 Bd Jean Moulin, 13385 Marseille Cedex5, France, e-mail: sanchez@newsup.univ-mrs.fr

Definition 1 *A (finite)* fuzzy set *or a* fuzzy distribution, on S, *is defined by a mapping* x *from* S *to [0,1], represented by a vector* $x = (x_1, ..., x_n)$, *with* x_i *denoting* $x(i)$, $0 \leq x_i \leq 1, i \in S$. *The set of all fuzzy sets on* S *is denoted by* $\mathcal{F}(S)$.

The fuzzy Markov chains are based on the concept of fuzzy relations and their compositions [14].

Definition 2 *A* fuzzy relation P *is defined as a fuzzy set on the Cartesian product* $S \times S$. P *is represented by a* matrix $\{p_{ij}\}_{i,j=1}^n$, *with* p_{ij} *denoting* $P(i,j)$, $0 \leq p_{ij} \leq 1,\ i,j \in S$.

Equipped with the notion of fuzzy relations, fuzzy Markov chains can now be defined.

Definition 3 *At each time instant t, t=0,1,..., the state of the system is described by the fuzzy set (or distribution)* $x^{(t)} \in \mathcal{F}(S)$. *The transition law of the* fuzzy Markov chain *is given by the fuzzy relation* P *as follows, at time instant t, t=1,2,...*

$$x_j^{(t+1)} = \max_{i \in S}\{x_i^{(t)} \wedge p_{ij}\}, \quad j \in S. \tag{1}$$

We refer to $x^{(0)}$ *as the intial fuzzy set (or the initial distribution).*

Note the kind of similarity between the fuzzy Markov chains and the classical Markov chains based on probability theory. Indeed, the behaviour of the Markov chain with the given transition probability matrix $\{p_{ij}\}_{i,j=1}^n$ is described by the recurrent equation

$$x_j^{(t+1)} = \sum_{i=1}^{n} x_i^{(t)} p_{ij}, \quad j = \overline{1,n}. \tag{2}$$

It is obvious that equations (1) and (2) have a similar structure, the only difference between them is in the employed operations and, of course, the meaning of the terms as fuzzy grades, instead of probabilities. The fuzzy Markov chains are based on the max-min algebra instead of the usual algebra. Namely, equation (1) is obtained from (2) by changing the algebraic summation to the max-operation and the algebraic multiplication to the min-operation, respectively.

Due to the above remark, it can be expected that fuzzy Markov chains will have some properties analogous to the properties of classical Markov chains.

Remark 1 *Note also that the fuzzy Markov chains have been introduced in the most general context. It is quite possible to impose further restrictions. For example as in [9, 10], the fuzzy relation might be required to be normal and/or convex (convex, when fuzzy sets are defined on metric spaces, note that the convexity of fuzzy numbers was already defined in Zadeh's original paper [22]).*

2 Finite convergence

It is natural to define the powers of the fuzzy transition matrix. Namely,

$$p_{ij}^t \triangleq \max_{k \in S}\{p_{ik} \wedge p_{kj}^{t-1}\}, \quad p_{ij}^1 = p_{ij}, \quad p_{ij}^0 = \delta_{ij},$$

where δ_{ij} is a Kronecker delta. Moreover in general $A \circ B$ will denote *max-min matrix multiplication* (in the sense of *max-min composition*) of A and B.

Note that the fuzzy state $x^{(t)}$ at time instant $t = 1, 2, ...$ can be calculated by the formula

$$x_k^{(t)} = \max_{l \in S}\{x_l^{(0)} \wedge p_{lk}^t\}, \quad k = \overline{1, n},$$

or equivalently in matrix notations

$$x^{(t)} = x^{(0)} \circ P^t.$$

It is now possible to formulate the following result which corresponds to the limit theorem for the classical Markov chains.

Theorem 1 (Giv'on [8], Thomason [18]) *The powers of the fuzzy transition matrix $\{p_{ij}\}_{i,j=1}^n$ either converge to idempotent $\{p_{ij}^\tau\}_{i,j=1}^n$, where τ is a finite number, or oscillate with a finite period ν starting from some finite power.*

PROOF: The proof is quite straightforward. If $\{p_{ij}^t\}_{i,j=1}^n$ does not converge in a finite number of steps, then it must oscillate with a finite period since the number of elements in the transition matrix $\{p_{ij}\}_{i,j=1}^n$ is finite and the max-min operations cannot introduce a number that is not in $\{p_{ij}\}_{i,j=1}^n$ originally.

If P^t does converge to P^τ, then

$$P^\tau = P^\tau \circ P = (P^\tau \circ P) \circ P = P^\tau \circ (P \circ P) = P^\tau \circ P^2 = ... = P^\tau \circ P^\tau,$$

since max-min composition is associative. Thus we conclude that P^τ is idempotent. □

By analogy with the classical Markov chains we give the following definition.

Definition 4 *Let the powers of the fuzzy transition matrix converge in τ steps to a non periodic solution, then the associated fuzzy Markov chain is called nonperiodic (or aperiodic) and $P^* := P^\tau$ is called a limiting fuzzy transition matrix.*

Recall that in the case of classical Markov chains the powers of the probability transition matrix in general converge to a stationary solution (may be periodic) in an *infinite* number of steps. Whereas Theorem 1 states that the sequence of powers of the fuzzy transition matrix has always the *finite* convergence (also possibly to a periodic solution). This is a crucial distinction between classical Markov chains and fuzzy Markov chains. In turn this fact results in the distinction of ergodic properties of the classical and fuzzy schemes. The following example clarifies the latter statement.

Example 1 *Let a Markov chain have the following probability transition matrix*

$$P = \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix}.$$

Note that a fuzzy Markov chain can also be described by this matrix (since its entries are positive and do not exceed one). It is easy to calculate the limit matrix for the classical Markov chain, by solving $\begin{bmatrix} p_1 & p_2 \end{bmatrix} P = \begin{bmatrix} p_1 & p_2 \end{bmatrix}$, *where* $p_1 + p_2 = 1$.

$$P^* = \lim_{t \to \infty} P^t = \begin{bmatrix} 4/7 & 3/7 \\ 4/7 & 3/7 \end{bmatrix}.$$

It can be seen that matrix P^* *has identical rows* $[4/7\ 3/7]$. *This fact has clear probabilistic interpretation : after a long time the probability of the system being in state 1 is equal to 4/7 and the probability of the system being in state 2 is 3/7, respectively. Moreover, the identity of the rows means that the long-run probability distribution does not depend on the initial state. The latter is a quite nice and important property.*

Now we consider the fuzzy Markov chain with the same transition matrix and calculate the powers of matrix P *with respect to the max- and min- operations.*

$$P^2 = P \circ P = \begin{bmatrix} \max\{0.7 \wedge 0.7, 0.3 \wedge 0.4\} & \max\{0.7 \wedge 0.3, 0.3 \wedge 0.6\} \\ \max\{0.4 \wedge 0.7, 0.6 \wedge 0.4\} & \max\{0.4 \wedge 0.3, 0.6 \wedge 0.6\} \end{bmatrix}$$

$$= \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix} = P$$

Hence,

$$P^* = \lim_{t \to \infty} P^t = P = \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix}.$$

In this case the transition matrix P *is the limit matrix itself. And the stationary solution of the fuzzy system* does *depend on the initial state, even though the system can freely move from one state to another. This is a crucial distinction between fuzzy Markov chains and probabilistic Markov chains.*

The above example motivates the following definition of the *ergodic* fuzzy Markov chain.

Definition 5 *The fuzzy Markov chain is called ergodic if it is aperiodic and the limiting transition matrix has identical rows.*

A more detailed discussion on the ergodic properties of fuzzy Markov chains will be given in Section 4.

3 Cesaro average and fuzzy potential

First let us formally define the *Cesaro average* and *α-potential* for the fuzzy Markov chain. It could be defined either for the sequence of fuzzy states or for the sequence of powers of the fuzzy transition matrix. The latter has been chosen, since the fuzzy transition matrix is a more basic characteristic than any particular realization.

Definition 6 *Let P be the fuzzy transition matrix for some fuzzy Markov chain. Then the* Cesaro average *or* Cesaro limit, *is defined as*

$$P^* \triangleq \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} P^t.$$

Note that the Cesaro average always exists for the classical (finite state) Markov chain [4, 6]. Below in Theorem 2 we show that the Cesaro average also always exists for the fuzzy Markov chain. Of course, the Cesaro average is equal to the limiting fuzzy transition matrix (see Definition 4) whenever the fuzzy Markov chain is aperiodic. In particular, this implies that there is no confusion by using P^* to denote both the Cesaro average and the limiting transition matrix.

Definition 7 *For any fuzzy Markov chain with transition matrix P one can define* α-potential *by*

$$V_\alpha \triangleq \sum_{t=0}^{\infty} \alpha^t P^t,$$

where $\alpha \in [0, 1]$.

Note that fuzzy potentials were also studied in [19, 20]. Next theorem exhibits the structure of the Cesaro average.

Theorem 2 *The Cesaro average for the fuzzy Markov chain always exists. Furthermore, let the fuzzy Markov chain oscillate with period ν starting from time instant τ. Then the Cesaro average can be expressed by the formula*

$$P^* = \frac{1}{\nu}(P^\tau + ... + P^{\tau+\nu-1}). \tag{3}$$

PROOF: Theorem 1 implies that in the most general case the sequence of powers of the fuzzy transition matrix P has the following structure

$$P^1, P^2, ..., P^{\tau-1}; P^\tau, ..., P^{\tau+\nu-1}; P^\tau, ..., P^{\tau+\nu-1}; ...$$

In particular, $P^{\tau+\nu} = P^\tau$. Now let us calculate the Cesaro average

$$P^* = \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} P^{(t)} =$$

$$= \lim_{T\to\infty} \left[\frac{1}{T}(P^1 + ... + P^{\tau-1}) + \frac{1}{T}(P^\tau + ... + P^{\tau+\nu-1} + ... + P^T)\right] =$$

$$= \lim_{T\to\infty}\left[\frac{1}{T}(P^1+...+P^{\tau-1})+\frac{1}{T}\left\lfloor\frac{T-\tau}{\nu}\right\rfloor(P^\tau+...+P^{\tau+\nu-1})\right.$$
$$\left.+\frac{1}{T}(P^{\nu\lfloor\frac{T-\tau}{\nu}\rfloor+1}+...+P^T)\right],$$

where $\lfloor a\rfloor$ denotes the greatest integer less than, or equal to, a. Since the sums $P^1 + ... + P^{\tau-1}$, $P^\tau + ... + P^{\tau+\nu-1}$ and $P^{\nu\lfloor\frac{T-\tau}{\nu}\rfloor+1} + ... + P^T$ are bounded; and

$$\lim_{T\to\infty}\frac{1}{T}\left\lfloor\frac{T-\tau}{\nu}\right\rfloor=\frac{1}{\nu}.$$

the Cesaro average exists and we can write

$$P^* = \lim_{T\to\infty}\left[\frac{1}{T}(P^1+...+P^{\tau-1})\right]+\lim_{T\to\infty}\left[\frac{1}{T}\left\lfloor\frac{T-\tau}{\nu}\right\rfloor(P^\tau+...+P^{\tau+\nu-1})\right]$$
$$+\lim_{T\to\infty}\left[\frac{1}{T}(P^{\nu\lfloor\frac{T-\tau}{\nu}\rfloor+1}+...+P^T)\right]=\frac{1}{\nu}(P^\tau+...+P^{\tau+\nu-1})$$

□

We would like to emphasize that in the aperiodic case ($\nu = 1$) the Cesaro average coincides with the limiting fuzzy transition matrix defined in the previous section.

Now it will be shown the relationship between the Cesaro average and the α-potential. The next fact is known in the classical theory of Markov desicion processes [13] as the *Blackwell formula.*

Theorem 3 *Suppose that V_α and P^* are the α-potential and Cesaro average of a fuzzy Markov chain. Then,*

$$\lim_{\alpha\uparrow 1}(1-\alpha)V_\alpha = P^*.$$

PROOF: Let us first rearrange the terms in the series defining the α-potential. It is possible since the series is absolutely convergent.

$$\begin{aligned}
\sum_{t=0}^{\infty}\alpha^t P^t &= I+\alpha P+...+\alpha^{\tau-1}P^{\tau-1}+\\
&\quad \alpha^\tau P^\tau+...+\alpha^{\tau+\nu-1}P^{\tau+\nu-1}+\\
&\quad \alpha^{\tau+\nu}P^\tau+...+\alpha^{\tau+2\nu}P^{\tau+\nu}+...\\
&= I+\alpha P+...+\alpha^{\tau-1}P^{\tau-1}+\\
&\quad (\alpha^\tau P^\tau+\alpha^{\tau+\nu}P^\tau+\alpha^{\tau+2\nu}P^\tau+...)+...\\
&\quad +(\alpha^{\tau+\nu-1}P^{\tau+\nu-1}+\alpha^{\tau+2\nu-1}P^{\tau+\nu-1}+\alpha^{\tau+3\nu-1}P^{\tau+\nu-1}+...)\\
&= I+\alpha P+...+\alpha^{\tau-1}P^{\tau-1}+\\
&\quad \alpha^\tau P^\tau(1+\alpha^\nu+\alpha^{2\nu}+...)+...\\
&\quad +\alpha^{\tau+\nu-1}P^{\tau+\nu-1}(1+\alpha^\nu+\alpha^{2\nu}+...)\\
&= I+\alpha P+...+\alpha^{\tau-1}P^{\tau-1}+\\
&\quad +(1+\alpha^\nu+\alpha^{2\nu}+...)(\alpha^\tau P^\tau+...+\alpha^{\tau+\nu-1}P^{\tau+\nu-1}).
\end{aligned}$$

Now we calculate the limit

$$\begin{aligned}
\lim_{\alpha\uparrow 1}(1-\alpha)\sum_{t=0}^{\infty}\alpha^t P^t &= \lim_{\alpha\uparrow 1}[(1-\alpha)(I+\alpha P+...+\alpha^{\tau-1}P^{\tau-1})]+\\
&\quad \lim_{\alpha\uparrow 1}[(1-\alpha)(1+\alpha^{\nu}+\alpha^{2\nu}+...)(\alpha^{\tau}P^{\tau}+...+\alpha^{\tau+\nu-1}P^{\tau+\nu-1})]\\
&= (P^{\tau}+...+P^{\tau+\nu-1})\frac{1}{\nu}\\
&= P^*,
\end{aligned}$$

since,

$$\lim_{\alpha\uparrow 1}[(1-\alpha)(1+\alpha^{\nu}+\alpha^{2\nu}+...)] = \lim_{\alpha\uparrow 1}\frac{1-\alpha}{1-\alpha^{\nu}} = \lim_{\alpha\uparrow 1}\frac{1-\alpha}{(1-\alpha)(1+\alpha+...+\alpha^{\nu-1})} = \frac{1}{\nu}.$$

□

It is not difficult to show that the α-potential possesses the Laurent series expansion near the point $\alpha = 1$. This Laurent series is called the *Blackwell expansion* in the theory of Markov Decission Processes [13]. Actually the above theorem shows that this series has the pole of order one and the coefficient of the singular term equals to the negative Cesaro average. Moreover, in the aperiodic case the Blackwell expansion for the fuzzy Markov chain has only a finite number of terms. Since in this case $P^* = P^{\tau}$, where τ is some finite number, one gets

$$\begin{aligned}
V_{\alpha} &= (I+\alpha P+...+\alpha^{\tau-1}P^{\tau-1})+\alpha^{\tau}P^*(1+\alpha+\alpha^2+...)\\
&= (I+\alpha P+...+\alpha^{\tau-1}P^{\tau-1})+\frac{\alpha^{\tau}}{1-\alpha}P^*\\
&= (I+[(\alpha-1)+1]P+...+[(\alpha-1)+1]^{\tau-1}P^{\tau-1})-\frac{[(\alpha-1)+1]^{\tau}}{\alpha-1}P^*
\end{aligned}$$

Next we use the binomial formula and collect the terms with the same power of $(\alpha-1)$. This results in

$$V_{\alpha} = -\frac{P^*}{\alpha-1}+(-\tau P^*+P^{\tau-1}+...+I)+...+(-P^*+P^{\tau-1})(\alpha-1)^{\tau-1}.$$

4 Ergodic properties

Before proceeding to the formal analysis, we present the problem with the help of the following example.

Example 2 *Let us consider the fuzzy Markov chain described by the fuzzy transition matrix*

$$P = \begin{bmatrix} 0.4 & 0.9 & 0.4 \\ 0.6 & 0.7 & 0.9 \\ 0.8 & 0.2 & 0.9 \end{bmatrix}$$

It is easy to check that

$$P^{\tau} = P^4 = \begin{bmatrix} 0.8 & 0.8 & 0.9 \\ 0.8 & 0.8 & 0.9 \\ 0.8 & 0.8 & 0.9 \end{bmatrix}$$

Note that the ergodic projection has identical rows. The classical Markov chain with such a property is referred to as the ergodic *chain. We shall apply this definition to the fuzzy Markov chain as well. This example shows that there exists ergodic fuzzy Markov chains. Though Example 1 shows that not every recurrent fuzzy chain is ergodic. And therefore it is of great interest to study the ergodic properties of fuzzy Markov chains.*

The ergodic properties of fuzzy Markov chains can be efficiently investigated by using the notion of eigen fuzzy sets [15, 16].

Definition 8 *Let P be a fuzzy relation given in a matrix form. Then x is called an eigen fuzzy set of P, iff*

$$x \circ P = x.$$

Recall that $\mathcal{F}(S)$ is a lattice with the following partial ordering.

Definition 9 *The fuzzy set $x \in \mathcal{F}(S)$ is* contained *in the fuzzy set $y \in \mathcal{F}(S)$ (written $x \subseteq y$), iff $x_i \leq y_i$ for all $i \in S$.*

Definition 10 *Let $\mathcal{X}$ be the set of eigen fuzzy sets of the fuzzy relation P. Namely,*

$$\mathcal{X} \triangleq \{x \in \mathcal{F}(S) \mid x \circ P = x\}.$$

The elements of $\mathcal{X}$ are the invariants of P according to the $\circ-$compostion (max-min). Then, if there exists $\overset{\vee}{x} \in \mathcal{F}(S)$ such that $x \subseteq \overset{\vee}{x}$ for any $x \in \mathcal{X}$, it is called the greatest eigen fuzzy set *of the fuzzy relation P.*

The existence of the greatest fuzzy eigen set was proven in [16]. With the notations of this paper, here are some basic results on the characterization of eigen fuzzy sets associated with a fuzzy relation, and on the resolution of eigen fuzzy sets equations, followed by an illustrative example.

Definition 11 *Let $\overset{0}{x} \in \mathcal{F}(S)$ be defined as $\overset{0}{x}_i = a_0, i = \overline{1, n}$, where $a_0 = \min_{j\in S} \max_{i\in S} P_{ij}$.*

It is shown that $\overset{0}{x} \in \mathcal{X}$ [15, 16].

Definition 12 *Let $\overset{1}{x} \in \mathcal{F}(S)$ be defined as $\overset{1}{x}_j = \max_{i\in S} P_{ij}, j = \overline{1, n}$.*

One has $\overset{0}{x}_i = a_0 = \min_{j\in S} \overset{1}{x}_j, i = \overline{1, n}$. Moreover, $\overset{0}{x} \subseteq \overset{1}{x}$, and when $\overset{1}{x} \in \mathcal{X}$, then it is the greatest element in $\mathcal{X}$. Defining now a sequence of fuzzy sets $(\overset{n}{x})_{n\succ 0}$ as

$$\overset{2}{x} = \overset{1}{x} \circ P = \overset{1}{x} \circ P^1, \overset{3}{x} = \overset{2}{x} \circ P = \overset{1}{x} \circ P^2, ..., \overset{n+1}{x} = \overset{n}{x} \circ P = \overset{1}{x} \circ P^n, ...$$

it can be shown that this sequence is decreasing and bounded by $\overset{0}{x}$ and $\overset{1}{x}$, that is

$$\overset{0}{x} \subseteq ... \subseteq \overset{n+1}{x} \subseteq \overset{n}{x} \subseteq ... \subseteq \overset{3}{x} \subseteq \overset{2}{x} \subseteq \overset{1}{x} .$$

Theorem 4 (Sanchez) *There exists* $m \in \{1, 2, ..., n\}$ *such that* $\overset{\vee}{x} = \overset{m}{x}$ *is the greatest element in* $\mathcal{X}$. *Moreover,* $\overset{0}{x} \subseteq \overset{\vee}{x} \subseteq \overset{1}{x}$.

The following theorem also holds.

Theorem 5 $\max_{i \in S} P_{ij}^k = (\overset{1}{x} \circ P^{k-1})_j = \overset{k}{x}_j, j = \overline{1, n},$ *for all* $k \succ 0$.

Remark 2 *In general, for fuzzy set inclusion,* $\mathcal{X}$ *is an upper semi-lattice, but not a lattice : if* $x \in \mathcal{X}$ *and* $y \in \mathcal{X}$, *then* $x \cup y \in \mathcal{X}$, *but in general* $x \cap y \notin \mathcal{X}$.

According to previous developments [16], it can be given three algorithmic ways to practically determine $\overset{\vee}{x} = \overset{m}{x}$, the *greatest eigen fuzzy set* of the fuzzy relation P. They are illustrated with an example.

Example 3 *Let* $S = \{1, 2, 3, 4, 5\}$ *and let* P *be given as*

$$P = \begin{bmatrix} 0.1 & 0.7 & 0.2 & \boxed{0.8} & \boxed{0.7} \\ 0.0 & 0.6 & 0.4 & 0.3 & 0.5 \\ \boxed{0.3} & \boxed{1.0} & 0.0 & 0.1 & 0.4 \\ \boxed{0.3} & 0.3 & \boxed{0.8} & 0.1 & 0.0 \\ 0.0 & 0.0 & 0.7 & 0.5 & 0.0 \end{bmatrix}$$

Method I. First determination of $\overset{\vee}{x} = \overset{m}{x}$ **.**

(i) Determine the greatest elements in each column of P (they are framed, in the example) and set them in $\overset{1}{x}$ in their corresponding positions. These elements also allow to define $\overset{0}{x}$. Remember that the greatest eigen fuzzy set is of type $\overset{m}{x}, 1 \leq m \leq 5$, it satisfies $\overset{m}{x} \circ P = \overset{m}{x}$ and $\overset{0}{x} \subseteq \overset{m}{x} \subseteq \overset{1}{x}$.

$$\overset{1}{x} = \begin{bmatrix} 0.3 & 1.0 & 0.8 & 0.8 & 0.7 \end{bmatrix}$$

$$\overset{0}{x} = \begin{bmatrix} 0.3 & 0.3 & 0.3 & 0.3 & 0.3 \end{bmatrix} \subseteq \overset{1}{x} \quad (\text{recall } \overset{0}{x}_i = \min_{j \in S} \overset{1}{x}_j, i = \overline{1, 5}).$$

$\overset{0}{x} \in \mathcal{X}$, but (in this example) $\overset{1}{x} \notin \mathcal{X}$.

(ii) Underline in $\overset{1}{x}$ the smallest of its elements (it is not necessarily in a unique position). It (or they) will be invariant after the next compositions.

$$\overset{1}{x} = \begin{bmatrix} \underline{0.3} & 1.0 & 0.8 & 0.8 & 0.7 \end{bmatrix}$$

(iii) Compute $\overset{2}{x} = \overset{1}{x} \circ P$ and underline in $\overset{2}{x}$ the smallest of the non already underlined elements.

$$\overset{2}{x} = \begin{bmatrix} \underline{0.3} & 0.8 & 0.8 & \underline{0.5} & \underline{0.5} \end{bmatrix}, \quad \overset{0}{x} \subseteq \overset{2}{x} \subseteq \overset{1}{x} \ .$$

(iv) Compute $\overset{3}{x} = \overset{2}{x} \circ P, \overset{4}{x} = \overset{3}{x} \circ P$, etc., underlining at each step the smallest of the non already underlined elements. When all elements are underlined, one gets $\overset{m}{x}$, the greatest eigen fuzzy set. One has $m \leq Card(S)$, so that convergence is fairly fast.

$$\overset{3}{x} = \begin{bmatrix} \underline{0.3} & 0.8 & \underline{0.5} & \underline{0.5} & \underline{0.5} \end{bmatrix}, \overset{0}{x} \subseteq \overset{3}{x} \subseteq \overset{2}{x} \subseteq \overset{1}{x},$$

$$\overset{4}{x}= \left[\begin{array}{ccccc} \underline{0.3} & \underline{0.6} & \underline{0.5} & \underline{0.5} & \underline{0.5} \end{array} \right], \overset{0}{x}\subseteq\overset{4}{x}\subseteq\overset{3}{x}\subseteq\overset{2}{x}\subseteq\overset{1}{x},$$

and it can be checked that $\overset{4}{x}=\overset{4}{x} \circ P$, that is $\overset{4}{x}\in \mathcal{X}$ $(m = 4)$, and it is the greatest eigen fuzzy set.

Method II. Second determination of $\overset{\vee}{x}=\overset{m}{x}$.

In the following method it is not needed to evaluate max-min compositions to obtain $\overset{\vee}{x}$. The idea is to get the underlined elements of method I, by replacing P, of order $Card(S)$ by another matrix, a reduction of P, of order strictly less than $Card(S)$, etc. At each step, the invariant elements are exactly the ones of method I in the corresponding step.

(i) As in method I, first determine the greatest element in each column of P.

(ii) Denote r the smallest of these elements (0.3 in our example) and consider the columns containing them (the 1st column in our example).

$$P = \left[\begin{array}{ccccc} 0.1 & 0.7 & 0.2 & \boxed{0.8} & \boxed{0.7} \\ 0.0 & 0.6 & 0.4 & 0.3 & 0.5 \\ \boxed{0.3} & \boxed{1.0} & 0.0 & 0.1 & 0.4 \\ \boxed{0.3} & 0.3 & \boxed{0.8} & 0.1 & 0.0 \\ 0.0 & 0.0 & 0.7 & 0.5 & 0.0 \end{array} \right]$$

(iii) Delete from P these columns *and the same numbered rows* (1st column *and 1st row* in our example), to get the first reduction P' of P. It is important to remark that it is *not* deleted the rows passing through the positions of value 0.3, say row 3 and row 4.

$$P' = \left[\begin{array}{cccc} 0.6 & 0.4 & 0.3 & \boxed{0.5} \\ \boxed{1.0} & 0.0 & 0.1 & 0.4 \\ 0.3 & \boxed{0.8} & 0.1 & 0.0 \\ 0.0 & 0.7 & \boxed{0.5} & 0.0 \end{array} \right]$$

(iv) Set in $\overset{m}{x}$ (m is not known yet) the value that has been found, say $r = 0.3$, in the position of the deleted columns, say the 1st column (only one column has been deleted here).

$$\overset{m}{x}= \left[\begin{array}{ccccc} 0.3 & * & * & * & * \end{array} \right]$$

(v) Return to the first step with P' instead of P and repeat, *but* with the following restriction : if r' denotes the smallest of the greatest elements of each column of P', set $max(r, r')$ in the appropriate position of $\overset{m}{x}$.

From P' it is derived 0.5 in positions 4 and 5, and 0.5 is greater than 0.3, hence

$$\overset{m}{x}= \left[\begin{array}{ccccc} 0.3 & * & * & 0.5 & 0.5 \end{array} \right].$$

Then,

$$P'' = \left[\begin{array}{cc} 0.6 & \boxed{0.4} \\ \boxed{1.0} & 0.0 \end{array} \right]$$

From P'' it is derived 0.4 which is smaller than 0.5, thus 0.5 is set (instead of 0.4) in position 3 of $\overset{m}{x}$.

$$\overset{m}{x}= \left[\begin{array}{ccccc} 0.3 & * & 0.5 & 0.5 & 0.5 \end{array} \right]$$

Finally,

$$P''' = \left[\boxed{0.6}\right]$$

which makes complete $\overset{m}{x}$ as

$$\overset{m}{x} = \overset{4}{x} = \left[\begin{array}{ccccc} 0.3 & 0.6 & 0.5 & 0.5 & 0.5 \end{array}\right],$$

because 0.6 is greater than 0.5.

Method III. Third determination of $\overset{\vee}{x} = \overset{m}{x}$.

The last method presented here is a direct application of theorems 4 and 5.

(i) Determine first $\overset{1}{x}$ with the elements corresponding to the greatest element in each column of P.

(ii) Compute $P^2 = P \circ P$ and determine the greatest elements in each column of P^2 (they are framed below in our example) : they give $\overset{2}{x}$, according to theorem 5, with $k = 2$: $\max_{i \in S} P^2_{ij} = \overset{2}{x}_j, j = \overline{1,5}$.

(iii) Compare $\overset{2}{x}$ with $\overset{1}{x}$: if they are different, compute $P^3 = P \circ P^2$ to get $\overset{3}{x}$ according to $\max_{i \in S} P^3_{ij} = \overset{3}{x}_j, j = \overline{1,5}$.

(iv) Compare $\overset{3}{x}$ with $\overset{2}{x}$: if they are different, compute $P^4 = P \circ P^3$ to get $\overset{4}{x}$, etc. Stop when it is found m such that $\overset{m+1}{x} = \overset{m}{x}$, that is $\overset{m}{x} = \overset{m}{x} \circ P$.

In our example,

$$P^2 = \left[\begin{array}{ccccc} \boxed{0.3} & 0.6 & \boxed{0.8} & \boxed{0.5} & \boxed{0.5} \\ \boxed{0.3} & 0.6 & 0.5 & \boxed{0.5} & \boxed{0.5} \\ 0.1 & 0.6 & 0.4 & 0.4 & \boxed{0.5} \\ \boxed{0.3} & \boxed{0.8} & 0.3 & 0.3 & 0.4 \\ \boxed{0.3} & 0.7 & 0.5 & 0.1 & 0.4 \end{array}\right]$$

and it is derived

$$\overset{2}{x} = \left[\begin{array}{ccccc} 0.3 & 0.8 & 0.8 & 0.5 & 0.5 \end{array}\right]$$

and so on.

Now we are able to state the following important result.

Theorem 6 *Suppose that the fuzzy Markov chain with transition matrix P is ergodic. Namely, it is aperiodic and the limiting transition matrix $P^* = P^\tau$ has identical rows. Then these rows are equal to $\overset{\vee}{x}$, the greatest eigen fuzzy set of the fuzzy relation defined by P.*

Proof: According to method III above for the computation of $\overset{\vee}{x}$, the greatest eigen fuzzy set of the fuzzy relation P, one has

$$\overset{\vee}{x}_j = \max_{i \in S} P^n_{ij}, \tag{4}$$

where $n = Card(S)$ (in fact $\overset{\vee}{x} = \overset{m}{x}$, where possibly $m \leq n$, but in any case $\overset{m}{x} = \overset{n}{x}$, where $\overset{n}{x} = \overset{n-1}{x} \circ P = \overset{1}{x} \circ P^{n-1}$). However, if the fuzzy Markov chain is ergodic, all elements in the same column of P^n are equal. The latter together with (4) immediately implies that the rows of the ergodic projection are equal to the greatest eigen fuzzy set of P. $\square$

Of course, the above theorem does not provide any conditions for the fuzzy Markov chain to be ergodic. Note that so far fuzzy Markov chains have been explored here in the most general setting. Even no assumptions about normality or convexity were imposed. Therefore the next step in the investigation of fuzzy Markov chains on the finite state space will be to find out the sufficient condition for the ergodicity of fuzzy Markov chains. The less restrictive those conditions would be, the better. A variant of such conditions, namely contraction conditions, were studied in a series of papers [9, 10, 19, 20] on fuzzy Markov chains over a compact metric space. However, in the case of the finite state space those conditions seem to be too restrictive.

5 Robust property of fuzzy Markov chains

In this section it will be shown that the fuzzy Markov chains are very *robust* systems with respect to small perturbations of the transition matrices. Actually, this robust property justifies the name "fuzzy". It means that imprecise data used in the fuzzy model will not cause a totally wrong description of the real object. Of course, it is supposed that the model itself is chosen appropriately.

By contrast, the classical Markov chains are not robust with respect to perturbations of the transition matrix, as demonstrated by the following example [17].

Example 4 *Consider the probabilistic Markov chain with the slightly perturbed transition matrix*

$$P(\varepsilon) = \begin{bmatrix} 1-\varepsilon & \varepsilon \\ \varepsilon & 1-\varepsilon \end{bmatrix}, \tag{5}$$

where ε *(*$0 \le \varepsilon < 1$*) is a small perturbation parameter. It is obvious that if* $\varepsilon = 0$*, then*

$$P^*(0) \triangleq \lim_{t\to\infty} P^t(0) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

However, it turns out that if ε *is arbitrary small but strictly greater than zero, then*

$$P^*(\varepsilon) \triangleq \lim_{t\to\infty} P^t(\varepsilon) = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix} \not\to P^*(0), \ \textit{as } \varepsilon \to 0.$$

So we can see that any arbitrary small errors in the transition matrix can lead to the wrong calculation of the important long-run characteristics of the system. The latter could not happen if we would use the fuzzy model. To demonstrate this, let us consider the fuzzy Markov chain with transition matrix (5). Again calculate the powers of the transition matrix with respect to the max- and min- operations (for ε *"small",* $\varepsilon \le 1-\varepsilon$ *holds).*

$$P^2(\varepsilon) = P(\varepsilon)\circ P(\varepsilon) = \begin{bmatrix} \max\{(1-\varepsilon),\varepsilon\} & \max\{\varepsilon,\varepsilon\} \\ \max\{\varepsilon,\varepsilon\} & \max\{\varepsilon,(1-\varepsilon)\} \end{bmatrix} = \begin{bmatrix} 1-\varepsilon & \varepsilon \\ \varepsilon & 1-\varepsilon \end{bmatrix} = P(\varepsilon)$$

In this example the limiting matrix also coincides with the original transition matrix. Hence,

$$P^*(\varepsilon) = \begin{bmatrix} 1-\varepsilon & \varepsilon \\ \varepsilon & 1-\varepsilon \end{bmatrix} \to \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = P^*(0), \quad as \quad \varepsilon \to 0.$$

The latter means that small errors in the transition matrix have small effect on the limiting matrix.

The results of the above example are generalized as follows.

Proposition 1 *Let $P(\varepsilon)$ be the perturbed transition matrix of a fuzzy Markov chain, that is each element of $P(\varepsilon)$ is a continuous function of ε. With $\varepsilon = 0$ one retrieves the original unperturbed transition matrix $P(0)$. And let $P^*(\varepsilon)$ denote the Cesaro limit of the perturbed chain. Then,*

$$\lim_{\varepsilon \to 0} P^*(\varepsilon) = P^*(0),$$

where $P^*(0)$ is the Cesaro limit of the original fuzzy Markov chain.

PROOF: Using the results of Theorem 2, we get

$$\lim_{\varepsilon \to 0} P^*(\varepsilon) = \lim_{\varepsilon \to 0} \frac{1}{\nu}(P^\tau(\varepsilon) + ... + P^{\tau+\nu-1}(\varepsilon)) = \frac{1}{\nu}(\lim_{\varepsilon \to 0} P^\tau(\varepsilon) + ... + \lim_{\varepsilon \to 0} P^{\tau+\nu-1}(\varepsilon)) =$$

$$= \frac{1}{\nu}(P^\tau(0) + ... + P^{\tau+\nu-1}(0)) = P^*(0).$$

Here we can interchange the limit and power operations, since τ and ν are finite.

□

From the Example 4 we can see that the above assertion does not hold for the classical Markov chains. It follows from the fact that the stationary solution in the classical Markov chains cannot be usually achieved for a finite number of steps, and hence one may not interchange $\lim_{\varepsilon \to 0}$ and $\lim_{t \to \infty}$ in general.

The type of perturbations that leads to the discontinuity of the limiting matrix with respect to perturbation parameter ε is called the *singular* perturbations [1, 5, 12, 17]. Otherwise, the perturbation is said to be *regular*. Thus Proposition 1 states that *all* perturbations of fuzzy Markov chains are regular.

6 Conclusions and Discussion

In this paper a general definition of fuzzy Markov chains over a finite state space has been introduced. The fuzzy Markov chains have a finite convergence to the stationary (may be periodic) solution. The Cesaro average and the α-potential for the fuzzy Markov chains have been defined and the relationship between them, which is known in the classical theory of Markov decision processes as the Blackwell formula, has been established. Furthermore, it has been pointed out that unlike for the classical Markov chains recurrency does not necessarily imply ergodicity in the case of the fuzzy Markov chains. However, if a fuzzy Markov chain is ergodic, then the rows of its limiting transition matrix are equal to the greatest eigen fuzzy set of the fuzzy relation associated with the chain. Finally, it has been shown that the fuzzy Markov chain is a robust system with respect to small perturbations of the transition matrix, which is not the case for the classical probabilistic Markov chains.

Finally, let us point out some open problems, which can be considered as directions towards future research. Probably the most interesting and challenging problem is to find out the most general conditions ensuring the ergodicity of fuzzy Markov chains. Then it is of the great practical importance to find the best way of computing the limiting fuzzy transition matrix and the Cesaro average in the general non-ergodic/periodic case.

Also the general theory of discrete Markov decision processes based on fuzzy sets needs to be developed.

Acknowledgement

The authors are grateful to Prof. A.A. Pervozvanskii for stimulating discussions.

References

[1] M. Abbad and J.A. Filar, Perturbation and stability theory for markov control problems, *IEEE Trans. Auto. Contr.* **37** (1992), 1415-1420.

[2] K.E. Avrachenkov and E. Sanchez, Fuzzy Markov chains: specificities and properties, an extended abstract for IPMU'2000 conference, (1999).

[3] R.E. Bellman and L.A. Zadeh, Decision-making in a fuzzy environment, *Management Science Series B* **17** (1970), 141-164.

[4] K.L. Chung, "Markov chains with stationary transition probabilities", Springer-Verlag, New York, 1960.

[5] F. Delebecque and J.P. Quadrat, Optimal control of Markov chains admitting strong and weak interactions, *Automatica* **17** (1981), 281-296.

[6] J.L. Doob, "Stochastic processes", John Wiley, New York, 1953.

[7] A.O. Esogbue and R.E. Bellman, Fuzzy dynamic programming and its extensions, *TIMS/Studies in the Management Sciences* **20** (1984), 147-167.

[8] Y. Give'on, Lattice matrices, *Information and Control*, **7** (1964), 477-484.

[9] M. Kurano, M. Yasuda, J. Nakagami, and Y. Yoshida, A limit theorem in some fuzzy dynamic systems, *Fuzzy Sets and Systems* **51** (1992), 83-88.

[10] M. Kurano, M. Yasuda, J. Nakagami, and Y. Yoshida, Markov-type fuzzy decision processes with a discounted reward on a closed interval, *European Journal of Operational Research*, **92** (1996), 649-662.

[11] R. Kruse, R. Buck-Emden and R. Cordes, Processor power considerations-An application of fuzzy Markov chains, *Fuzzy Sets and Systems* **21** (1987), 289-299.

[12] A.A. Pervozvanskii and V.G. Gaitsgori, "Theory of suboptimal decisions", Kluwer Academic Publishers, Dordrecht, 1988.

[13] M.L. Puterman, "Markov decision processes", John Wiley, New York, 1994.

[14] E. Sanchez, Resolution of Composite Fuzzy Relation Equations, *Information and Control* **30** (1976), 38-48.

[15] E. Sanchez, Resolution of Eigen Fuzzy Sets Equations, *Fuzzy Sets and Systems* **1** (1978), 69-74.

[16] E. Sanchez, Eigen Fuzzy Sets and Fuzzy Relations, *J. Math. Anal. Appl.*, **81** (1981), 399-421.

[17] P.J. Schweitzer, Perturbation theory and finite Markov chains, *J. App. Prob.* **5** (1968), 401-413.

[18] M.G. Thomason, Convergence of Powers of a Fuzzy Matrix, *J. Math. Anal. Appl.* **57** (1977), 476-480.

[19] Y. Yoshida, M. Yasuda, J. Nakagami, and M. Kurano, A potential of fuzzy relations with a linear structure: The contractive case, *Fuzzy Sets and Systems* **60** (1993), 283-294.

[20] Y. Yoshida, M. Yasuda, J. Nakagami, and M. Kurano, A potential of fuzzy relations with a linear structure: The unbounded case, *Fuzzy Sets and Systems*, **66** (1994), 83-95.

[21] Y. Yoshida, Markov chains with a transition possibility measure and fuzzy dynamic programming, *Fuzzy Sets and Systems* **66** (1994), 39-57.

[22] L.A. Zadeh, Fuzzy sets, *Information and Control* **8** (1965), 338-353.

[23] L.A. Zadeh, Maximizing sets and fuzzy Markoff algorithms, *IEEE Trans. on Systems, Man, and Cybernetics - Part C: Applications and Reviews* **28** (1998), 9-15.