ZESZYTY NAUKOWE POLITECHNIKI ŚLĄSKIEJ 1982

Seria: AUTOMATYKA z. 65 Nr kol. 738

Ernest CZOGAŁA, Witold PEDRYCZ

Politechnika Śląska

A CONTRIBUTION TO FUZZY MARKOV CHAINS

Summary. Since the establishment of fuzzy set theory it formed a useful tool for solving a wide class of problems. Nowadays there are some attempts to combine fuzzy set theory and probability theory in order to describe subjective and objective form of uncertainty.

Here we deal with fuzzy Markov chains treating them as a reformulation of well known notion of probability. The paper introduces some basic properties and theorems connected with fuzzy Markov chains (e.g. stationarity of transition matrix of the chain). The applicational example illustrates the utility of discussed method.

1. Preliminaries

Theory of fuzzy sets investigated by many researches played similarly, as the probability theory an important role in description of many problems of science. The theory of probability and statistics, established theoretically by Kolmogorof forms a useful tool for investigation of the nondeterministic phenomena, e.g. optimal control under stochastic disturbances.

There are some papers connected with fuzzy-probability theory, mentioned firstly by Zadeh [5] and continueed by e.g. Hirota [3]. Fuzzy set theory and probability theory were combined in different manner, emphasizing the various aspects of description of complex processes in "hard" and "soft" sciences [6].

Let us now introduce the notion of fuzzy Markov chain forming a reformulation of stochastic Markov chain, discussing some basic properties as stationarity of transition matrix of the chain corresponding to the similar feature in probability theory. The applicational example points out the utility of the discussed method.

2. Fuzzy Markov chains-definitions and theorems

Let us consider a finite collection of states

$$\left\{x_1, x_2, \ldots, x_N\right\} \qquad (1)$$

The transition between the actual state and the past ones is expressed by means of the number:

$$\mu(x_i^n \mid x_j^{n-1}, x_r^{n-2}, \dots, x_s^0) \in [0,1] \tag{2}$$

where upper indices denote discrete time moments and the lower ones stand for a concrete state. In extremal case we have:

$$\mu(x_i^n \mid x_j^{n-1}, x_r^{n-2}, \dots, x_s^0) = \begin{cases} 1, & \text{if the possibility of transition between i-th and s-th state is the highest one,} \\ 0, & \text{if this possibility is equal to 0} \end{cases} \tag{3}$$

Now we could define fuzzy Markov chain.

Def. 1 $\{x_1, x_2, \dots, x_N\}$ forms a fuzzy Markov chain if the following is satisfied:

$$\mu(x_i^n \mid x_j^{n-1}, x_r^{n-2}, \dots, x_s^0) = \mu(x_i^n \mid x_j^{n-1}) \qquad i,j = 1,2,\dots,N \tag{4}$$

From the above definition one can describe the fuzzy Markov chain by means of a transition matrix M_n which completely characterizes this chain:

$$M_n = \begin{bmatrix} \mu(x_1^n \mid x_1^{n-1}) & \mu(x_1^n \mid x_2^{n-1}) & \dots & \mu(x_1^n \mid x_N^{n-1}) \\ \mu(x_2^n \mid x_1^{n-1}) & & & \\ \vdots & & & \\ \mu(x_N^n \mid x_1^{n-1}) & \mu(x_N^n \mid x_2^{n-1}) & \dots & \mu(x_N^n \mid x_N^{n-1}) \end{bmatrix} \tag{5}$$

If we assume the transition between i-th and j-th state does not depend on time, so

$$\mu(x_i^n \mid x_j^{n-1}) = \mu(x_i \mid x_j) \tag{6}$$

is satisfied, we get a stationary Markov chain.

Def. 2. We call fuzzy Markov chain a stationary, if

$$M_n = M \quad \forall i = 1,2,\dots \tag{7}$$

holds true.

$$M = \begin{bmatrix} \mu(x_1 \mid x_1) & \mu(x_1 \mid x_2) & \dots & \mu(x_1 \mid x_N) \\ \mu(x_2 \mid x_1) & \mu(x_2 \mid x_2) & \dots & \\ \vdots & & & \\ \mu(x_N \mid x_1) & \mu(x_N \mid x_2) & \dots & \mu(x_N \mid x_N) \end{bmatrix} \tag{8}$$

Now let us consider the problem of calculation of the possibility of transition between j-th and i-th state in finite number of steps (n). We start with n=2, having the situation depicted in Fig. 1. State x_j forming a starting point is fixed. For each $r \in [1,N]$ we could compute the minimal value of possibility between j-th and i-th state via r-th state:

$$\min(\mu(x_i \mid x_r), \ \mu(x_r \mid x_j)) = \mu(x_i \mid x_r) \wedge \mu(x_r \mid x_j) \tag{9}$$

and taking into account the maximal value for all $r \in |1,N|$

$$\mu(x_i \mid x_j) = \max_{1 \le r \le N} \left[\min(\mu(x_i \mid x_r), \ \mu(x_r \mid x_j))\right] =$$

$$= \bigvee_{r=1}^{N} \left[\mu(x_i \mid x_r) \wedge \mu(x_r \mid x_j)\right] \tag{10}$$

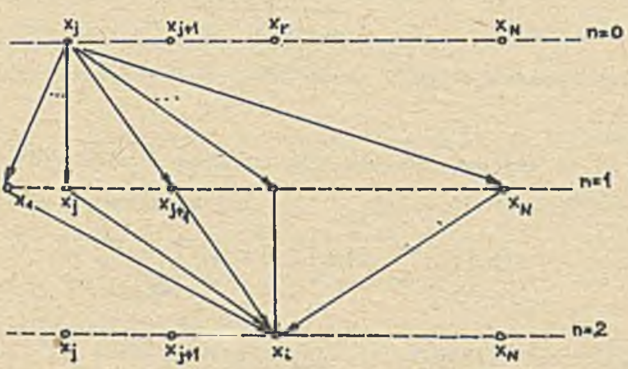

Fig. 1. Illustration of computation of possibility of transition between j-th and i-th state in fuzzy Markov chain

This result is equivalent o the well known max-min composition of transition matrix M:

$$M^2 = M \circ M \tag{11}$$

or generally for every n

$$M^n = \underbrace{M \circ M \circ \ldots \circ M}_{n\text{-times}} \tag{12}$$

M^n is defined recursively:

$$M^0 = I \tag{13}$$

$$I = \begin{bmatrix} 1 & & 0 \\ & 1 \ddots & \\ 0 & & 1 \end{bmatrix} \tag{14}$$

$$M^k = M^{k-1} \circ M \tag{15}$$

In place of max-min composition (eq. 10) we could define another composition:

$$\tilde{\mu}(x_i|x_j) = \bigvee_{r=1}^{N} \left(\mu(x_i|x_r) \cdot \mu(x_i|x_r) \right) \tag{16}$$

"." is an operation of multiplication (max-mult) composition. An equation corresponding with Eq. 12 has a form:

$$\tilde{M}^n = \underbrace{M \cdot M \cdot \ldots \cdot M}_{n\text{-times}} \tag{17}$$

The following theorem gives a comparison of these two kinds of compositions.

<u>Theorem 1.</u> If we discuss max-min and max-mult composition then the following holds true:

$$\left| \mu(x_i|x_j) - \tilde{\mu}(x_i|x_j) \right| \leq \frac{1}{4} \qquad \forall_{1 \leq i,j \leq N} \tag{18}$$

$\mu(.|.)$, $\tilde{\mu}(.|.)$ are elements of transition matrices M^2 and $\tilde{M}^2$ respectively.

<u>Proof.</u> In order to prove the correctness of the theorem it is enough and sufficient to prove the validity of inequality:

$$\left| \min(a,b) - ab \right| \leq \frac{1}{4} \qquad a,b \in [0,1] \tag{19}$$

Let us put down:

$$f(a,b) = \left| \min(a,b) - ab \right|$$

We show that $f(a,b)$ attains its global maximum in point $a=b=\frac{1}{2}$ equal to $\frac{1}{4}$ So we consider:

$$a = \frac{1}{2}, \quad b = \frac{1}{2} + \varepsilon \quad \varepsilon \in \left[0, \frac{1}{2}\right] \tag{20}$$

then we have

$$f(\tfrac{1}{2}, \tfrac{1}{2}+\varepsilon) = \left| \tfrac{1}{2} - \tfrac{1}{2}(\tfrac{1}{2}+\varepsilon)\right| = \left| \tfrac{1}{4} - \varepsilon\right| \leq \tfrac{1}{4} \tag{21}$$

Similarly $a = \frac{1}{2}, \quad b = \frac{1}{2} - \varepsilon, \quad \varepsilon \in \left[0, \frac{1}{2}\right]$ (22)

$$f(\tfrac{1}{2}, \tfrac{1}{2}-\varepsilon) = \left| \tfrac{1}{2} - \varepsilon - \tfrac{1}{2}(\tfrac{1}{2}-\varepsilon)\right| = \left|\varepsilon - \tfrac{1}{4}\right| \leq \tfrac{1}{4} \tag{23}$$

so taking into account the symmetry of the discussed function:

$$f(a,b) = f(b,a) \tag{24}$$

it leads to the theorem.

The theorem 1 could be extended to the form which makes it possible to compare $\mu(x_i|x_j)$ and $\tilde{\mu}(x_i|x_j)$ for $n > 2$.
The following inequality is valid:

$$\left|\mu(x_i|x_j) - \tilde{\mu}(x_i|x_j)\right| \leq \frac{1}{2^n} \tag{25}$$

$\mu(x_i|x_j)$ $\tilde{\mu}(x_i|x_j)$ are elements of transition matrices M^n and $\tilde{M}^n$ respectively.
Discussing max-min composition (eq.12) we could use the following results on convergence of powers of fuzzy matrices [4].

<u>Theorem 2</u>. If $M^r \leq M^s$ $(r < s)$ then M converges.

Immediate from this theorem we get:

<u>Proposition 1</u>. If for all $i,j \in [1,N]$ there exists such k that

$$\mu(x_i|x_j) \leq \mu(x_i|x_k) \wedge \mu(x_k|x_j)$$

then M converges to M^t $t \leq N-1$

From this fact we have: - if $I \leq M$ then M converges to M^t $t \leq N-1$.
Of course if $M, M^2, M^3, \ldots, M^t$ forms a convergent series of matrices then we have:

$$\lim \rho(\mu(x_i|x_j), \mu'(x_i|x_j)) = 0 \tag{26}$$

$\mu(\cdot|\cdot)$ stands for n-th power of M, $\mu'(\cdot|\cdot)$ stands for the limit of the power series of M, ρ denotes the distance.
This result corresponds with ergodic property of Markov chain in probability theory [2].

In order to estimate the upper bound transition matrix let us use the closure of fuzzy relation.

$$M^{o} = \begin{cases} M \cup M^2 \cup \ldots \cup M^p, \text{ if } M^p = M^{p+1} = \ldots \\ M \cup M^2 \cup \ldots \cup M^p \cup M^{p+1} \cup \ldots \cup M^{p+o}, \text{ if the series of powers of } M \\ \text{is not convergent and has oscillation with the period equal} \\ \text{to } o. \end{cases} \quad (27)$$

Of course we have:

$$\forall_{t=1,2,\ldots} \quad M^t \subseteq M^o \quad (28)$$

Till now we considered that states are defined precisely and transition between each of the state is given subjectively assigning their transition a concrete value of possibility.

We could extend our considerations allowing a more general situation an initial state is given in the form fuzzy set X^o defined by membership function:

$$\mu_{X^o} : \{x_1, x_2, \ldots, x_N\} \to [0,1] \quad (29)$$

Then using that's X^o we combine it obtaining the state in second time moment as follows:

$$\mu(x_i|X^o) = \bigvee_{m=1}^{N} \left[\mu(x_i|x_m) \wedge \left(\bigvee_{j=1}^{N} \mu(x_m|x_j) \wedge \mu_{X^o}(x_j)\right)\right] \quad (30)$$

or

$$\mu(x_i \mid X^o) = \bigvee_{j=1}^{N} \bigvee_{m=1}^{N} \left[\mu(x_i|x_m) \wedge \left(\mu(x_m|x_j) \wedge \mu_{X^o}(x_j)\right)\right] \quad (31)$$

Eq.30 could be rewritten as follows:

$$\mu(x_i \mid X^o) = \bigvee_{m=1}^{N} \left[\mu(x_i|x_m) \wedge \mu_{X^o \circ M}(x_m)\right] \quad (32)$$

($\circ$ stands for max-min composition), or generally for variable n

$$\mu(x_i|X^o) = \bigvee_{m=1}^{N} \left(\mu(x_i|x_m) \wedge \mu_{X^o \circ \underbrace{M \circ \ldots \circ M}_{n\text{-times}}}(x_m)\right) \quad (33)$$

3. Numerical example

As an illustration of our discussion at the usefulness of the proposed concept let us consider an applicational example such as deterioration of quality of bridge pavement, where this process is discussed with respect to its utilization [1]. There are a lot of factors which play an important role such as:

- increasing of brittleness,
- hydratation,
- freezing and unfreezing.

The states given here having its linguistic representation are as follows:

x_1 - smooth
x_2 - quite good
x_3 - uneven
x_4 - not useful for exploitation.

It is an obvious fact that from engineering point of view we have some problems which are interesting from essential design problems. Transition matrix has a form:

$$M = \begin{bmatrix} .8 & .4 & .1 & 0 \\ .2 & .7 & .3 & .1 \\ 0 & 0 & .2 & .9 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Calculating sequentially $M, M^2, \ldots$ we have:

$$M^2 = \begin{bmatrix} .8 & .4 & .3 & .1 \\ .2 & .7 & .3 & .3 \\ 0 & 0 & .2 & .9 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$M^3 = \begin{bmatrix} .8 & .4 & .3 & .3 \\ .2 & .7 & .3 & .3 \\ 0 & .0 & .2 & .9 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$M^4 = M^3$ and M^5 is equal to M^3.

Let us consider the following situation, when we take X^* as a fuzzy set with membership function

$$\begin{bmatrix} 1 & .1 & .05 & 0 \end{bmatrix}$$

which describes the initial state of pavement as quite good. We calculate:

$$X^{\infty} = X^{0} \circ M^{0} = [.8 \quad .4 \quad .3 \quad .3]$$

It could be noticed that the possibility of x_1 decreased and the possibility of x_2, x_3, x_4 considerably increased pointing the fact of deterioration of quality of the pavement.

REFERENCES

[1] Benjamin J.R., Cornell C.A.: Probability, Statistics and Decision for Civil Engineers. Mc Graw-Hill, N. York, 1970.

[2] Hadley H.: Introduction to Probability and Statistical Theory, Holden-Day, San Francisco, 1967.

[3] Hirota K.: Extended fuzzy expressions of probabilistic sets, in Advances in fuzzy set theory and applications, edited by M. Gupta et al, North Holland, 1979.

[4] Thomason M.G.: Finite fuzzy automata regular fuzzy languages and pattern recognition. Pattern Recognition, 1975, 5, 383-390.

[5] Zadeh L.A.: Probability measure of fuzzy events, J. Math Anal and Appl. 1968, 23, 421-427.

[6] Zadeh L.A.: Outline of a new approach to the analysis of complex systems and decision processes. IEEE Trans. SMC, 1973, s, 28-44.

Recenzent: Prof. dr hab. inż. Andrzej Tylikowski

РАСПЛЫВЧАТЫЕ ЦЕПИ МАРКОВА

Р е з ю м е

Теория расплывчатых множеств дала практические методы решений широкого класса проблем. Нынче существует несколько попыток объединения теории расплывчатых множеств с теорией вероятностей и описание таким образом субъективной и обективной формы неуверенности.

Работа касается расплывчатых цепей Маркова, аналогичных цепям Маркова. Введено некоторые основные свойства (например стационарность) и доказано связанные с ними теоремы. Исследования иллюстрирует практический пример.

ROZMYTE ŁAŃCUCHY MARKOWA

S t r e s z c z e n i e

Teoria zbiorów rozmytych dostarczyła użytecznego narzędzia rozwiązywania szerokiej klasy problemów. Istnieje obecnie kilka prób mających na celu połączenie teorii zbiorów rozmytych i probabilistyki, umożliwiając w ten sposób opis subiektywnej i obiektywnej formy niepewności.

Praca dotyczy rozmytych łańcuchów Markowa stanowiących analogię łańcuchów Markowa. Wprowadzono niektóre podstawowe własności (np. stacjonarność), udowodniono twierdzenia z nimi związane. Przykład aplikacyjny stanowi ilustrację przeprowadzonych rozważań.