KYBERNETIKA — VOLUME *61* (2025), NUMBER 1, PAGES 58–78

# DISCOUNTED MARKOV DECISION PROCESSES WITH FUZZY COSTS

Salvador De-Jesús-Hernández, Hugo Cruz-Suárez,
and Raúl Montes-de-Oca

This article concerns a class of discounted Markov decision processes on Borel spaces where, in contrast with the classical framework, the cost function $\widetilde{C}$ is a fuzzy function of a trapezoidal type, which is determined from a classical cost function $C$ by applying an affine transformation with fuzzy coefficients. Under certain conditions ensuring that the classical (or standard) model with a cost function $C$ has an optimal stationary policy $f_o$ with the optimal cost $V_o$, it is shown that such a policy is also optimal for the fuzzy model with a cost function $\widetilde{C}$, and that the optimal fuzzy value $\tilde{V}_o$ is obtained from $V_o$ via the same transformation used to go from $C$ to $\widetilde{C}$. And these results are obtained with respect to two cases: the max-order of the fuzzy numbers and the average ranking order of the trapezoidal fuzzy numbers. Besides, a fuzzy version of the classical linear-quadratic model without restrictions is presented.

*Keywords:* discounted Markov decision processes, trapezoidal fuzzy costs, max-order, average ranking

*Classification:* 90C40, 93C42

## 1. INTRODUCTION

In the realm of decision-making and uncertainty modeling, fuzzy set theory, introduced by L. Zadeh (see [24]), has emerged as a powerful tool for handling imprecise and ambiguous information. Fuzzy Theory provides a framework to represent and manipulate uncertainty through the concept of fuzzy number (see [9, 23], and [24]). Fuzzy numbers, in a certain sense, can be thought of as a generalization of real numbers along with their order (Remarks 2.3, 2.10, and 2.11). In this way, fuzzy numbers have a structure that can be applied to various problems. Over the years, this theory has found wide applications in various fields (see [8, 10, 13], and [25]). The concept of uncertainty modeled by fuzzy numbers, combined with the uncertainty from probability theory, forms the basis of the fuzzy random variables theory proposed by Puri and Ralescu (see [17] and [18]). Unlike classical random variables, which are real-valued, fuzzy random variables take values in the space of fuzzy numbers. This approach opens new ways for modeling systems where uncertainties are not easily quantifiable in a traditional sense.

DOI: 10.14736/kyb-2025-1-0058

On the other hand, Markov Decision Processes (MDPs) (see [12]) are stochastic models used to represent systems where outcomes are influenced by random factors and can be altered through the selection of specific decision variables or controls. They are valuable for modeling decision-making in uncertain environments. MDPs are applied in various fields to optimize decision-making strategies over time under uncertainty and require the development of decision-making policies that optimize a suitable optimization criterion. The discounted MDP (see [12]), in particular, has garnered significant attention due to its mathematical tractability and practical relevance in modeling decision problems in economics. This article deals with an infinite horizon discounted MDP, $(X, A, \{A(x) : x \in X\}, Q, C)$, where $X$ and $A$ represent the state and action spaces, respectively, both being Borel spaces, $\{A(x) : x \in X\}$ represents the collection of admissible controls for the state $x \in X$, $Q$ represents the transition law and $C$ represents the cost-per-stage function. In this article, the authors will refer to this model as the standard discounted MDP.

If in the standard discounted MDP, the cost-per-stage function $C$ is replaced by the trapezoidal number $\tilde{C}$ (that depends on $C$), a discounted fuzzy MDP is obtained, whose optimization criterion is the expected discounted total cost. The expectation taken in the discounted total cost refers to fuzzy expectation. The solution to the discounted cost optimization problem is considered concerning the order of the $\alpha$-cuts of fuzzy numbers, and its solution is given in terms of the corresponding standard discounted MDP. It is essential to note that in the case where $C$ is considered as a fuzzy number (see Remark 2.3 below) the fuzzy problem reduces to the standard problem. In this sense, the standard theory of discounted MDPs is extended. Besides, a similar analysis is presented for the approach of the average ranking order on the trapezoidal fuzzy numbers. The results obtained contribute to the growing body of literature on fuzzy decision-making and provide valuable insights for practical application. By incorporating trapezoidal fuzzy numbers into the cost-per-stage function, ambiguity is captured in decision-making environments, offering insights into real-world scenarios where precise information may be lacking or difficult to obtain.

Now, there are some comments on the previous works on fuzzy MDPs. Firstly, in [4] and [6] fuzzy MDPs on discrete spaces and with objective functions other than the total discounted cost are presented. Secondly, in [3, 7, 14, 15, 16], and [22], discounted MDPs with different fuzzy characteristics have been provided, for which: (i) Both the state and decision spaces are finite, or (ii) Both the state and the decision spaces are compact sets, or (iii) The standard MDPs considered are deterministic. This excludes, for instance, the fuzzy version of the linear-quadratic (LQ) model provided and solved in Section 6 below, in which neither (i), (ii) nor (iii) holds. In contrast with the models developed in [3, 7, 14, 15, 16], and [22], here, as it was already indicated, fuzzy extensions of the standard discounted MDPs on Borel spaces are obtained which allows to deal with MDPs general enough covering the cases (i), (ii) and (iii).

The organization of this article is as follows: In Section 2, the authors provide a comprehensive overview of fuzzy theory. Section 3 discusses fuzzy random variables and their expectation. Section 4 provides the fact that certain functions are fuzzy random variables in the sense of Puri and Ralescu. In Section 5 an overview of Markov decision processes is given, then fuzzy MDPs are presented, focusing on the discounted framework. Section 6 presents a fuzzy version of the LQ model. Finally, in Section 7 the conclusions are presented.

**Notation 1.** In the article, the following standard mathematical symbols will be distinguished in the fuzzy context with an asterix symbol: " * " (or sometimes the symbol " ** "). That is, in the fuzzy context, "$\leq$", "$+$" and "$\sum$", will be denoted by "$\leq^*$", "$+^*$" and "$\sum^*$", respectively. Similarly, in the fuzzy context, the expectation operator "$\mathbb{E}$", the limit "$lim$" and the infimum "$inf$", will be denoted by "$\mathbb{E}^*$", "$lim^*$" and "$inf^*$", respectively. Also, the notations: "$\leq^{**}$" and "$inf^{**}$" are used for a second comparison between trapezoidal fuzzy numbers referent to the ranking order. It is important to mention that the product of a real number $\lambda$ and a fuzzy number $\Upsilon$ will be simply denoted as $\lambda\Upsilon$. Moreover, some special functions which appear as fuzzy quantities, say, the cost function, the optimal value function, and so on, will be distinguished with a "tilde"; for instance, the fuzzy cost function will be written as $\widetilde{C}$.

## 2. PRELIMINARES ON FUZZY THEORY

The first part of this section presents some definitions and basic results about the fuzzy set theory (see [9, 23], and [24]).

**Basic definitions.** Let $\Lambda$ be a non-empty set. Then a *fuzzy set* $\Gamma$ on $\Lambda$ is defined in terms of the *membership function* $\Gamma'$, which assigns to each element of $\Lambda$ a real value from the interval $[0, 1]$. The *$\alpha$-cut* of $\Gamma$, denoted by $\Gamma_\alpha$, is defined to be the set $\Gamma_\alpha := \{x \in \Lambda \mid \Gamma'(x) \geq \alpha\}$ $(0 < \alpha \leq 1)$ and $\Gamma_0$ is the *closure* of $\{x \in \Lambda \mid \Gamma'(x) > 0\}$ denoted by $cl\{x \in \Lambda \mid \Gamma'(x) > 0\}$.

**Definition 2.1.** A *fuzzy number* $\Gamma$ is a fuzzy set defined on the set of real numbers $\mathbb{R}$ (i. e., taking $\Lambda = \mathbb{R}$ in the previous definition), which satisfies:

a) $\Gamma'$ is normal, i. e., there exists $x_0 \in \mathbb{R}$ with $\Gamma'(x_0) = 1$;

b) $\Gamma'$ is convex, i. e., $\Gamma_\alpha$ is convex for all $\alpha \in [0, 1]$;

c) $\Gamma'$ is upper semicontinuous;

d) $\Gamma_0$ is compact.

The set of the fuzzy numbers will be denoted by $\mathfrak{F}(\mathbb{R})$.

**Definition 2.2.** A fuzzy number $\Gamma$ is called a *trapezoidal fuzzy number* if its membership function has the following form:

$$\Gamma'(x) = \begin{cases} 0 & \text{if } x \leq l \\ \frac{x-l}{m-l} & \text{if } l < x \leq m \\ 1 & \text{if } m < x \leq n \\ \frac{p-x}{p-n} & \text{if } n < x \leq p \\ 0 & \text{if } p < x \ , \end{cases} \tag{1}$$

where $l$, $m$, $n$ and $p$ are real numbers, with $l < m \leq n < p$. A trapezoidal fuzzy number is simply denoted by $(l, m, n, p)$.

**Remark 2.3.** a) The case in which $m = n$ in (1) will be named a *triangular fuzzy number* and it will be simply denoted as $(l, m, p)$. And, considering the degenerated case in which $l = m = p$ in a triangular number, the *fuzzy representation* of the real number $m$ is obtained with the membership function given by:

$$m'(x) = \begin{cases} 1 & \text{if } x = m \\ 0 & \text{if } x \neq m. \end{cases} \tag{2}$$

b) For a trapezoidal fuzzy number $\Gamma = (l, m, n, p)$ the corresponding $\alpha$-cuts are given by $\Gamma_\alpha = [(m - l)\alpha + l, p - (p - n)\alpha]$, $\alpha \in [0, 1]$ (see [21]).

**Definition 2.4.** Let $\Gamma$ and $\Upsilon$ be fuzzy numbers. If " $\star$ " denotes the addition or the scalar mutiplication, then it is defined as a fuzzy set on $\mathbb{R}$, $\Gamma \star \Upsilon$, by the membership function:

$$(\Gamma \star \Upsilon)'(u) = \sup_{u = x \star y} \min\{\Gamma'(x), \Upsilon'(y)\},$$

for all $u \in \mathbb{R}$.

As a consequence of Definition 2.4, it is possible to obtain the following result for trapezoidal fuzzy numbers (see [21]).

**Lemma 2.5.** If $H = (b_l, b_m, b_n, b_p)$ and $I = (d_l, d_m, d_n, d_p)$ are two trapezoidal fuzzy numbers and letting $\lambda$ be a positive number, then it follows that

a) $\lambda H = (\lambda b_l, \lambda b_m, \lambda b_n, \lambda b_p)$, and

b) $H +^* I = (b_l + d_l, b_m + d_m, b_n + d_n, b_p + d_p)$. And, it also holds: if $\{(b_l^t, b_m^t, b_n^t, b_p^t) : 0 \leq t \leq N\}$ is a finite set of $N$ trapezoidal fuzzy numbers, then

$$\sum_{t=0}^{N}{}^{*} (b_l^t, b_m^t, b_n^t, b_p^t) = \Big( \sum_{t=0}^{N} b_l^t, \sum_{t=0}^{N} b_m^t, \sum_{t=0}^{N} b_n^t, \sum_{t=0}^{N} b_p^t \Big).$$

**Convergence.** Let $\mathbb{D}$ denote the set of all closed bounded intervals on the real line $\mathbb{R}$. For $\Psi = [b_l, b_u]$, $\Phi = [d_l, d_u] \in \mathbb{D}$ define

$$d_H(\Psi, \Phi) = \max(|b_l - d_l|, |b_u - d_u|). \tag{3}$$

It is possible to verify that $d_H$ defines a metric on $\mathbb{D}$ and that $(\mathbb{D}, d_H)$ is a complete metric space (see [18]). Now, if $\widetilde{\eta} \in \mathfrak{F}(\mathbb{R})$, then $\widetilde{\eta}_\alpha$ is a compact set because its membership function is upper semicontinuous and has a compact support. Therefore, it is defined as $\hat{d} : \mathfrak{F}(\mathbb{R}) \times \mathfrak{F}(\mathbb{R}) \longrightarrow \mathbb{R}$ by

$$\hat{d}(\widetilde{\eta}, \widetilde{\mu}) = \sup_{\alpha \in [0,1]} d_H(\widetilde{\eta}_\alpha, \widetilde{\mu}_\alpha), \tag{4}$$

$\widetilde{\eta}, \widetilde{\mu} \in \mathfrak{F}(\mathbb{R})$. It is straightforward to see that $\hat{d}$ is a metric in $\mathfrak{F}(\mathbb{R})$ (see [18]).

**Definition 2.6.** A sequence $\{\widetilde{\eta}_t\}_{t=0}^{\infty}$ of fuzzy numbers is said to be *convergent* to the fuzzy number $\widetilde{\mu}$, written as $\lim^*_{t\longrightarrow\infty}\widetilde{\eta}_t = \widetilde{\mu}$ if and only if $\hat{d}(\widetilde{\eta}_t, \widetilde{\mu}) \longrightarrow 0$ as $t \longrightarrow \infty$.

**Lemma 2.7.** (Puri and Ralescu [18]) The metric space $(\mathfrak{F}(\mathbb{R}), \hat{d})$ is complete.

**Definition 2.8.** Given a sequence of subsets $\{W_n\}_{n=0}^{\infty}$ in $\mathbb{R}$, the Kuratowski limits are defined as follows:

1. *Kuratowski lower limit*, $\liminf^K W_n := \Big\{x \in \mathbb{R} : \limsup_{n\to\infty} \inf_{w\in W_n} \{|x-w|\} = 0\Big\}$.
2. *Kuratowski upper limit*, $\limsup^K W_n := \Big\{x \in \mathbb{R} : \liminf_{n\to\infty} \inf_{w\in W_n} \{|x-w|\} = 0\Big\}$.

A sequence of sets $\{W_n\}_{n=0}^{\infty}$ is said to *converge in the sense of Kuratowski* to a set $W$ if $\liminf^K W_n = \limsup^K W_n = W$.

**Remark 2.9.** It is important to mention that in $\mathbb{D}$, Kuratowski convergence and convergence with the metric $\hat{d}$ are equivalent ([1]). As a consequence, if a sequence of fuzzy numbers $\{\widetilde{\eta}^{(n)}\}_{n=0}^{\infty}$ converges in the $\hat{d}$ metric to a fuzzy number $\widetilde{\eta}$, then the sequences of $\alpha$-cuts $\{\widetilde{\eta}^{(n)}_{\alpha}\}_{n=0}^{\infty}$ converge in the sense of Kuratowski to $\widetilde{\eta}_{\alpha}$ for all $\alpha \in [0,1]$.

**Orders.** Now, for $\widetilde{\eta}, \widetilde{\mu} \in \mathfrak{F}(\mathbb{R})$, with $\alpha$-cuts $\widetilde{\eta}_{\alpha} = [b_{\alpha}, d_{\alpha}]$ and $\widetilde{\mu}_{\alpha} = [f_{\alpha}, g_{\alpha}]$, $\alpha \in [0,1]$, respectively, define $\widetilde{\eta} \leqslant^* \widetilde{\mu}$ if and only if $b_{\alpha} \le f_{\alpha}$ and $d_{\alpha} \le g_{\alpha}$ for all $\alpha \in [0,1]$ (see [11]). It is not difficult to verify that the order "$\leq^*$ " is, in fact, a partial order on $\mathfrak{F}(\mathbb{R})$.

**Remark 2.10.** Take $w, z \in \mathbb{R}$, and let $\widetilde{w}$ and $\widetilde{z}$ be fuzzy numbers with membership functions given by $(\widetilde{w})'(x)=1$, $x = w$ and $(\widetilde{w})'(x) = 0$, $x \neq w$, and $(\widetilde{z})'(x)=1$, $x = z$ and $(\widetilde{z})'(x) = 0$, $x \neq z$. Then, it is easy to see that $\widetilde{w} \leqslant^* \widetilde{z}$ is equivalent to $w \le z$.

Moreover, the following *comparison* between the trapezoidal fuzzy numbers (see [19] and [20]) is also introduced. Let $H = (b_l, b_m, b_n, b_p)$ be a trapezoidal fuzzy number. Then its *average ranking* $\Re(H)$ is defined as

$$\Re(H) = \frac{b_l + b_m + b_n + b_p}{4}.$$

Now, let $H = (b_l, b_m, b_n, b_p)$ and $I = (d_l, d_m, d_n, d_p)$ be trapezoidal fuzzy numbers. Hence, it is defined that $H \leqslant^{**} I$ if and only if

$$\Re(H) \le \Re(I). \tag{5}$$

**Remark 2.11.** In relation to the last comparison between trapezoidal fuzzy numbers, note that in the degenerate case for which in $H$: $b_l = b_m = b_n = b_p = b$ and in $I$: $d_l = d_m = d_n = d_p = d$, it results that $H \leqslant^{**} I$ if and only if $b \le d$.

## 3. FUZZY RANDOM VARIABLES

**Definitions.** Following [17] and [18], the next definitions on fuzzy random variables and their expectations are established. For this, $\mathfrak{C}(\mathbb{R})$ denotes the class of nonempty compact subsets of $\mathbb{R}$, and if $(\Omega_1, \mathcal{A}_1)$ and $(\Omega_2, \mathcal{A}_2)$ are measurable spaces, then $\mathcal{A}_1 \otimes \mathcal{A}_2$ denotes the corresponding product $\sigma$-algebra associated to the product space $\Omega_1 \times \Omega_2$. Also, $\mathcal{B}(\mathbb{R})$ denotes the Borel $\sigma$-algebra of $\mathbb{R}$.

**Definition 3.1.** Let $(\Omega, \mathcal{A})$ be a measurable space and $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ be the measurable space of the set of real numbers. A function $\widetilde{Y} : \Omega \longrightarrow \mathfrak{F}(\mathbb{R})$ is said to be a *fuzzy random variable* associated with $(\Omega, \mathcal{A})$, if the section $\widetilde{Y}_\alpha : \Omega \longrightarrow \mathfrak{C}(\mathbb{R})$ which is the $\alpha$-level function defined by $\widetilde{Y}_\alpha(\omega) = (\widetilde{Y}(\omega))_\alpha$ for all $\omega \in \Omega$ and $\alpha \in [0, 1]$ satisfies that $Gr(\widetilde{Y}_\alpha) = \{(\omega, x) \in \Omega \times \mathbb{R} \mid x \in (\widetilde{Y}(\omega))_\alpha\} \in \mathcal{A} \otimes \mathcal{B}(\mathbb{R})$, for all $\alpha \in [0, 1]$.

**Definition 3.2.** Given a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, a fuzzy random variable $\widetilde{Y}$ associated to $(\Omega, \mathcal{A})$ is said to be an *integrably bounded fuzzy random variable* with respect to $(\Omega, \mathcal{A}, \mathbb{P})$ if there is a function $h : \Omega \longrightarrow \mathbb{R}$, $h \in L^1(\Omega, \mathcal{A}, \mathbb{P})$ such that $|x| \le h(\omega)$, for all $(\omega, x) \in \Omega \times \mathbb{R}$ with $x \in (\widetilde{Y}(\omega))_0 := \widetilde{Y}_0(\omega)$.

**Definition 3.3.** Given an integrably bounded fuzzy random variable $\widetilde{Y}$ associated with respect to the probability space $(\Omega, \mathcal{A}, \mathbb{P})$, then the *fuzzy expected value* of $\widetilde{Y}$ in Aumann's sense is the unique fuzzy set of $\mathbb{R}$, $\mathbb{E}^*[\widetilde{Y}]$ such that for each $\alpha \in [0, 1]$:

$$\left(\mathbb{E}^*[\widetilde{Y}]\right)_\alpha = \left\{ \int_\Omega f(\omega)\,\mathrm{d}\mathbb{P}(\omega) \mid f : \Omega \longrightarrow \mathbb{R}, f \in L^1(\Omega, \mathcal{A}, \mathbb{P}), f(\omega) \in (\widetilde{Y}(\omega))_\alpha \text{ a.s. } [\mathbb{P}] \right\}.$$

**Lemma 3.4.** (Puri and Ralescu [18]) Let $\{\widetilde{X}_k\}_{k=0}^\infty$ be a sequence of fuzzy random variables and $\widetilde{X}$ an integrably bounded fuzzy random variable, such that $\widetilde{X}_k \to \widetilde{X}$ a.s. on $\Omega$. If there exists $h \ge 0$, $h \in L^1(\Omega, \mathcal{A}, \mathbb{P})$ such that for all $k \ge 0$ and $\alpha \in [0, 1]$, it holds that

$$\sup_{x \in (\widetilde{X}_k(\omega))_\alpha} |x| \le h(\omega) \text{ for all } k \in \mathbb{N}, \omega \in \Omega, \text{ and } 0 < \alpha \le 1, \text{ then } \mathbb{E}^*[\widetilde{X}_k] \to \mathbb{E}^*[\widetilde{X}].$$

The last convergence is in the Kuratowski sense.

**Lemma 3.5.** Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, $Y : \Omega \longrightarrow \mathbb{R}^+$ a non-negative random variable with $\mathbb{E}[Y] := \int_\Omega Y(\omega)\,\mathrm{d}\mathbb{P}(\omega) < \infty$, and $\Delta = (b_1, b_2, b_3, b_4)$ a trapezoidal fuzzy number, where $0 < b_1 \le b_2 \le b_3 \le b_4$. Then $\widetilde{Y} = Y\Delta$ is a fuzzy random variable and $\mathbb{E}^*[\widetilde{Y}] = \mathbb{E}[Y]\Delta$.

Proof. For each $n \in \mathbb{N}$ let's take the next partition of the image set of $Y$

$$J_n^0 = \left[0, \frac{1}{2^n}\right], \quad J_n^k = \left(\frac{k}{2^n}, \frac{k+1}{2^n}\right], \quad k = 1, \ldots, 2^{2n} - 1.$$

Observe that

$$\bigcup_{k=0}^{2^{2n}-1} J_n^k = \big[0, 2^n\big].$$

If $J_n^{2^{2n}} = \big(2^n, \infty\big)$ is taken, then the collection $\big\{J_n^k\big\}_{k=0}^{2^{2n}}$ is a partition of the interval $[0, \infty]$. It is defined that

$$H_n^k = Y^{-1}\big(J_n^k\big) \in \mathcal{A}, \quad k = 1, \dots, 2^{2n}.$$

The measurability of the random variable $Y$ implies that the collection $\big\{H_n^k\big\}_{k=0}^{2^{2n}}$ is a partition of $\Omega$ and the following is true

$$\begin{aligned} 0 \leq & Y(\omega) < \frac{1}{2^n}, && \omega \in H_n^0 \\ \frac{k}{2^n} < & Y\big(\omega\big) \leq \frac{k+1}{2^n}, && \omega \in H_n^k, \quad k = 1, \dots, 2^{2n} - 1 \\ & Y\big(\omega\big) > 2^n, && \omega \in H_n^{2^{2n}}. \end{aligned}$$

Consider the following sequence of random variables

$$s_n(\omega) = \sum_{k=0}^{2^{2n}} \frac{k}{2^n} \mathbb{I}_{H_n^k}(\omega), \quad \omega \in \Omega.$$

Here, $\mathbb{I}_{H_n^k}$ denotes the indicator function of the set $H_n^k$.

From the construction of the sequence $\{s_n\}_{n=0}^{\infty}$ it is obtained that

$$\begin{aligned} & 0 \leq s_n \leq s_{n+1} \leq Y, \quad n \in \mathbb{N}, \\ & 0 \leq Y(\omega) - s_n(\omega) \leq \frac{1}{2^n}, \quad \omega \in \bigcup_{k=0}^{2^{2n}-1} H_n^k. \end{aligned}$$

This means that the sequence of random variables $\{s_n\}_{n=0}^{\infty}$ converges pointwise to $Y$ and if the random variable $Y$ is bounded, then it also converges uniformly.

Consider the functions $\widetilde{s}_n = \Delta s_n$. Take $\Delta_\alpha = \big[q(\alpha), r(\alpha)\big]$, where $q(\alpha) = \alpha(b_2 - b_1) + b_1$ and $r(\alpha) = b_4 - \alpha(b_4 - b_3)$. Since $s_n(\omega) \geq 0$ for all $\omega \in \Omega$ and for all $n \in \mathbb{N}$, then $\widetilde{s}_n(\omega) = \Delta s_n(\omega) = \big(s_n(\omega)b_1, s_n(\omega)b_2, s_n(\omega)b_3, s_n(\omega)b_4\big)$ is a trapezoidal fuzzy number with membership function $\mu_{\widetilde{s}_n(\omega)}(x)$ ([21]).

$$\mu_{\widetilde{s}_n(\omega)}(x) = \begin{cases} \frac{x - b_1 Y(\omega)}{s_n(\omega)(b_2 - b_1)} & \text{if} \quad x \in \big[b_1 s_n(\omega), b_2 s_n(\omega)\big) \\ 1 & \text{if} \quad x \in \big[b_2 s_n(\omega), b_3 s_n(\omega)\big) \\ \frac{b_4 s_n(\omega) - x}{s_n(\omega)(b_4 - b_3)} & \text{if} \quad x \in \big[b_3 s_n(\omega), b_4 s_n(\omega)\big) \\ 0 & \text{otherwise.} \end{cases}$$

Let's prove that $\widetilde{s}_n$ is a fuzzy random variable:

$$\begin{aligned}
&\Big\{(\omega,x)\in\Omega\times\mathbb{R}\ :\ x\in\big(\widetilde{s}_n\big)_\alpha(\omega)\Big\}\\
&=\Big\{(\omega,x)\in\Omega\times\mathbb{R}\ :\ x\in s_n(\omega)\big[q(\alpha),r(\alpha)\big]\Big\}\\
&=\bigcup_{k=0}^{2^{2n}}\left(\Big\{s_n=\frac{k}{2^n}\Big\}\times\frac{k}{2^n}\big[q(\alpha),r(\alpha)\big]\right)\in\mathcal{A}\otimes\mathcal{B}(\mathbb{R}).
\end{aligned}$$

Now let's verify that the random variables $\widetilde{s}_n$ are integrably bounded. Note that $\big(\widetilde{s}_n\big)_0(\omega)=s_n(\omega)\big[b_1,b_4\big]$. The authors define the function $h:\Omega\longrightarrow\mathbb{R}$, as $h(\omega)=Y(\omega)b_4$. From the construction of the random variables $s_n$ it is obtained that $s_n(\omega)\le Y(\omega)$ for all $\omega\in\Omega$ and with this it is easy to see that $|x|\le h(\omega)$ for all $\omega\in\Omega$ and $x\in s_n(\omega)[b_1,b_4]$. It is also known that $\mathbb{E}\big[h\big]=b_4\mathbb{E}\big[Y\big]<\infty$, then it can be concluded that the random variable $\widetilde{s}_n$ is integrably bounded. It is almost immediate to see that $\mathbb{E}^*\big[\widetilde{s}_n\big]=\mathbb{E}\big[s_n\big]\big(b_1,b_2,b_3,b_4\big)$.

Now consider the fuzzy number $\widetilde{Y}(\omega)=Y(\omega)\Delta,\omega\in\Omega$. Since $Y(\omega)\ge 0$ for all $\omega\in\Omega$, it holds that $\widetilde{Y}(\omega)=\big(Y(\omega)b_1,Y(\omega)b_2,Y(\omega)b_3,Y(\omega)b_4\big)$. It is also known that $\widetilde{Y}(\omega)$ is a trapezoidal number, whose membership function $\mu_{\widetilde{Y}(\omega)}$ is given by

$$\mu_{\widetilde{Y}(\omega)}(x)=\begin{cases}\frac{x-b_1Y(\omega)}{Y(\omega)(b_2-b_1)} & \text{if}\quad x\in\big[b_1Y(\omega),b_2Y(\omega)\big)\\ 1 & \text{if}\quad x\in\big[b_2Y(\omega),b_3Y(\omega)\big)\\ \frac{b_4Y(\omega)-x}{Y(\omega)(b_4-b_3)} & \text{if}\quad x\in\big[b_3Y(\omega),b_4Y(\omega)\big)\\ 0 & \text{otherwise.}\end{cases}$$

From this it is easy to see that its $\alpha$-cuts have the following form

$$\begin{aligned}
\big(\widetilde{Y}\big)_\alpha(\omega)&=\big[\alpha Y(\omega)(b_2-b_1)+b_1Y(\omega),b_4Y(\omega)-\alpha Y(\omega)(b_4-b_3)\big]\\
&=Y(\omega)\big[\alpha(b_2-b_1)+b_1,b_4-\alpha(b_4-b_3)\big]\\
&=Y(\omega)\big(\Delta\big)_\alpha.
\end{aligned}$$

If the function $g$ is defined as $g:\Omega\longrightarrow\mathbb{R}$, $g(\omega)=b_4Y(\omega)$, then $|x|\le g(\omega)$ for all $(\omega,x)\in\Omega\times\mathbb{R}$, $x\in Y(\omega)[b_1,b_4]$. Since $\mathbb{E}\big[g\big]<\infty$ it yields that $\widetilde{Y}$ is integrably bounded.

Let's see that the sequence of fuzzy random variables $\{\widetilde{s}_n\}_{n=0}^\infty=\{s_n(b_1,b_2,b_3,b_4)\}_{n=0}^\infty$ converges to the fuzzy number $\widetilde{Y}=Y(b_1,b_2,b_3,b_4)$ with the metric $\hat{d}$.

$$\begin{aligned}
\lim_{n\to\infty}\hat{d}\big(\widetilde{s}_n,\widetilde{Y}\big)&=\lim_{n\to\infty}\sup_{\alpha\in[0,1]}\Big(d_H\big((\widetilde{s}_n)_\alpha,(\widetilde{Y})_\alpha\big)\Big)\\
&=\lim_{n\to\infty}\sup_{\alpha\in[0,1]}\left(d_H\Big(s_n(\omega)\big[q(\alpha),r(\alpha)\big],Y(\omega)\big[q(\alpha),r(\alpha\big]\Big)\right)\\
&=\lim_{n\to\infty}\sup_{\alpha\in[0,1]}\left(\max\Big\{q(\alpha)\big|s_n(\omega)-Y(\omega)\big|,r(\alpha)\big|s_n(\omega)-Y(\omega)\big|\Big\}\right)
\end{aligned}$$

$$
\begin{aligned}
&= \lim_{n\to\infty} \sup_{\alpha\in[0,1]} \left( r(\alpha)\big|s_n(\omega) - Y(\omega)\big| \right) \\
&= \lim_{n\to\infty} b_4\big|s_n(\omega) - Y(\omega)\big| \\
&= 0.
\end{aligned}
$$

Next, it will be proved that $\widetilde{Y}$ is a fuzzy random variable. Let $\alpha \in [0,1]$ and consider the function $\varphi_\alpha : \Omega \times \mathbb{R} \longrightarrow \mathbb{R}$, defined as $\varphi_\alpha(\omega, x) = \inf_{y\in\widetilde{Y}_\alpha(\omega)} |x-y|$. This function represents the distance from the closed set $\widetilde{Y}_\alpha(\omega)$ to the point $x \in \mathbb{R}$. If $\omega$ is fixed, then the map $x \mapsto \varphi_\alpha(\omega, x)$ is continuous: let $y_0 \in \mathbb{R}$ and $\varepsilon > 0$ be an arbitrary number. By the definition of infimum, there exists $z \in \widetilde{Y}_\alpha(\omega)$ such that $|y_0 - z| < \varphi_\alpha(\omega, y_0) + \frac{\varepsilon}{2}$. For any $x \in \mathbb{R}$, it is obtained that $\varphi_\alpha(\omega, x) \le |x - z|$, then $\varphi_\alpha(\omega, x) - \varphi_\alpha(\omega, y_0) \le |x-z| - \varphi_\alpha(\omega, y_0)$. By the triangle inequality, $|x-z| \le |x-y_0| + |y_0-z|$. Using the last two inequalities it is gotten that $\varphi_\alpha(\omega, x) - \varphi_\alpha(\omega, y_0) \le (|x-y_0| + |y_0-z|) - \varphi_\alpha(\omega, y_0)$. Since $|y_0-z| < \varphi_\alpha(\omega, y_0)+\frac{\varepsilon}{2}$, it follows that $\varphi_\alpha(\omega, x)-\varphi_\alpha(\omega, y_0) \le |x-y_0|+\frac{\varepsilon}{2}$. Similarly, it is possible to find that $\varphi_\alpha(\omega, y_0)-\varphi_\alpha(\omega, x) \le |x-y_0|+\frac{\varepsilon}{2}$. Thus, $|\varphi_\alpha(\omega, x)-\varphi_\alpha(\omega, y_0)| \le |x-y_0|+\frac{\varepsilon}{2}$. To prove continuity, let's choose $\delta = \frac{\varepsilon}{2}$. On the other hand, if $x \in \mathbb{R}$ is fixed, the map $\omega \mapsto \varphi_\alpha(\omega, x)$ is $\mathcal{A}$-measurable: three cases are considered based on the value of $x$ relative to the set $Y(\omega)[q(\alpha), r(\alpha)]$. If $x \le Y(\omega)q(\alpha)$, $\varphi_\alpha(\omega, x) = Y(\omega)q(\alpha) - x$ is clearly a measurable function. If $Y(\omega)q(\alpha) \le x \le Y(\omega)r(\alpha)$, it is obtained that $\varphi_\alpha(\omega, x) = 0$. Thus, $\{\omega \in \Omega \mid \varphi_\alpha(\omega, x) = 0\} = \{\omega \in \Omega \mid Y(\omega)q(\alpha) \le x \le Y(\omega)r(\alpha)\} = \{\omega \in \Omega \mid \frac{x}{r(\alpha)} \le Y(\omega) \le \frac{x}{q(\alpha)}\} \in \mathcal{A}$. If $x > Y(\omega)r(\alpha)$, $\varphi_\alpha(\omega, x) = x - Y(\omega)r(\alpha)$ which is a measurable function. Such functions are known as Carathéodory functions (see [1], p. 311), and Lemma 8.2.6 in [1] guarantees that the function $\varphi_\alpha$ is $\mathcal{A}\otimes\mathcal{B}(\mathbb{R})$-measurable. This implies that

$$
\begin{aligned}
&\Big\{(\omega, x) \in \Omega \times \mathbb{R} : \varphi_\alpha(\omega, x) = 0\Big\} \\
=&\Big\{(\omega, x) \in \Omega \times X : x \in \big(\widetilde{Y}\big)_\alpha(\omega)\Big\} \\
=&\Big\{(\omega, x) \in \Omega \times X : x \in Y(\omega)\big[q(\alpha), r(\alpha)\big]\Big\} \in \mathcal{A}\otimes\mathcal{B}(\mathbb{R}).
\end{aligned}
$$

This proves that $\widetilde{Y}$ is a fuzzy random variable. Consider the fuzzy number $\widetilde{U} = \mathbb{E}\big[Y\big]\Delta$ whose $\alpha$-cuts are given by

$$\mathbb{E}\big(Y\big)\big[q(\alpha), r(\alpha)\big], \quad \text{for all } \alpha \in [0,1].$$

To guarantee that $\mathbb{E}^*\big[\widetilde{s}_n\big] \xrightarrow[n\to\infty]{} \widetilde{U}$, the monotone convergence theorem is used:

$$
\begin{aligned}
\lim_{n\to\infty} \hat{d}\Big(\mathbb{E}^*\big[\widetilde{s}_n\big], \widetilde{U}\Big) &= \lim_{n\to\infty} \sup_{\alpha\in[0,1]} \left( d_H\Big(\big(\mathbb{E}^*[\widetilde{s}_n]\big)_\alpha, (\widetilde{U})_\alpha\Big) \right) \\
= \lim_{n\to\infty} \sup_{\alpha\in[0,1]} &\left( d_H\Big(\mathbb{E}\big[s_n\big]\big[q(\alpha), r(\alpha)\big], \mathbb{E}[Y]\big[q(\alpha), r(\alpha\big]\Big) \right)
\end{aligned}
$$

$$
\begin{aligned}
&= \lim_{n\to\infty} \sup_{\alpha\in[0,1]} \left( \max\left\{ q(\alpha)\Big|\mathbb{E}[s_n]-\mathbb{E}[Y]\Big|, r(\alpha)\Big|\mathbb{E}[s_n]-E[Y]\Big| \right\}\right)\\
&= \lim_{n\to\infty} \sup_{\alpha\in[0,1]} \left( r(\alpha)\big|\mathbb{E}[s_n]-\mathbb{E}[Y]\big| \right)\\
&= \lim_{n\to\infty} b_4\big|\mathbb{E}[s_n]-\mathbb{E}[Y]\big|\\
&=0.
\end{aligned}
$$

From Lemma 3.4 and Remark 2.9 the following implication yields:

$$\mathbb{E}^*\big[\widetilde{Y}\big] = \lim_{n\to\infty} \mathbb{E}^*\big[\widetilde{s}_n\big] = \mathbb{E}[Y]\Delta.$$

□

## 4. MEASURABILITY OF A FUZZY AFFINE TRANSFORMATION

The goal of this section is to prove measurability for the case when $\widetilde{Y} = \Delta_1 Y + \Delta_2$, where $\Delta_j = \big(b_1^{(j)}, b_2^{(j)}, b_3^{(j)}, b_4^{(j)}\big), 0 < b_1^{(j)} \le b_2^{(j)} \le b_3^{(j)} \le b_4^{(j)}$, $j = 1, 2$, are two trapezoidal fuzzy numbers with $\alpha$-cuts given by $\big(\Delta_j\big)_\alpha = \big[q_j(\alpha), r_j(\alpha)\big]$, $q_j(\alpha) = \alpha(b_2^{(j)} - b_1^{(j)}) + b_1^{(j)}$ and $r_j(\alpha) = b_4^{(j)} - \alpha(b_4^{(j)} - b_3^{(j)})$. In the same way as in the previous section $Y : \Omega \longrightarrow [0,\infty]$ will represent a random variable with $\mathbb{E}\big[Y\big] < \infty$, $\big\{s_n\big\}_{n=0}^{\infty}$ will represent a sequence of simple, non-negative non-decreasing random variables that converges pointwise to $Y$ and $\big\{\widetilde{s}_n\big\}_{n=0}^{\infty}$, where $\widetilde{s}_n = \Delta_1 s_n$, will represent the sequence of fuzzy random variables that converges with the metric $\hat{d}$ to the fuzzy random variable $\Delta_1 Y$, which satisfy that $\mathbb{E}^*\big[\widetilde{s}_n\big] = \Delta_1 \mathbb{E}\big[s_n\big]$.

**Lemma 4.1.** Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, $Y : \Omega \longrightarrow \mathbb{R}^+$ a non-negative random variable with $\mathbb{E}[Y] < \infty$, and $\Delta_j = \big(b_1^{(j)}, b_2^{(j)}, b_3^{(j)}, b_4^{(j)}\big)$ where $0 < b_1^{(j)} \le b_2^{(j)} \le b_3^{(j)} \le b_4^{(j)}$, $j = 1, 2$, two trapezoidal fuzzy numbers. Then $\widetilde{Y} = Y\Delta_1 + \Delta_2$ is a fuzzy random variable and $\mathbb{E}^*[\widetilde{Y}] = \mathbb{E}[Y]\Delta_1 + \Delta_2$.

Proof. Consider the sequence of fuzzy numbers $\widetilde{u}_n = \Delta_1 s_n + \Delta_2$. If a number is the sum of two fuzzy numbers, then their $\alpha$-cuts are the sum of the respective $\alpha$-cuts ([21]), so that

$$\big(\widetilde{u}_n(\omega)\big)_\alpha = \Big[q_1(\alpha)s_n(\omega) + q_2(\alpha), r_1(\alpha)s_n(\omega) + r_2(\alpha)\Big].$$

Observe that

$$
\begin{aligned}
&\Big\{(\omega,x)\in\Omega\times X : x\in \big(\widetilde{u}_n(\omega)\big)_\alpha\Big\}\\
=&\left\{(\omega,x)\in\Omega\times X : x\in \Big[q_1(\alpha)s_n(\omega)+q_2(\alpha), r_1(\alpha)s_n(\omega)+r_2(\alpha)\Big]\right\}
\end{aligned}
$$

$$= \bigcup_{k=0}^{2^{2n}} \left( \left\{ s_n = \frac{k}{2^n} \right\} \times \left[ q_1(\alpha)\frac{k}{2^n} + q_2(\alpha), r_1(\alpha)\frac{k}{2^n} + r_2(\alpha) \right] \right) \in \mathcal{A} \otimes \mathcal{B}(\mathbb{R}).$$

The latter proves that $\widetilde{u}_n$ is a fuzzy random variable. Consider the function $h(\omega) = b_4^{(1)}Y(\omega) + b_4^{(2)}$ with $\mathbb{E}\big[h\big] < \infty$. Since $s_n(\omega) \leq Y(\omega)$ for all $\omega \in \Omega$ and $n \in \mathbb{N}$, then $x \leq h(\omega)$ for all $\omega \in \Omega$, $x \in \big(\widetilde{u}_n(\omega)\big)_\alpha$, and $\alpha \in [0,1]$. This implies that the random variables $\widetilde{u}_n$ are integrably bounded. It is almost immediate to obtain that $\mathbb{E}^*\big[\widetilde{u}_n\big] = \Delta_1\mathbb{E}\big[Y\big] + \Delta_2$.

Observe that

$$\big(\widetilde{Y}(\omega)\big)_\alpha = \big[Y(\omega)q_1(\alpha) + q_2(\alpha), Y(\omega)r_1(\alpha) + r_2(\alpha)\big].$$

The authors want to prove that $\Delta_1 Y + \Delta_2$ is a fuzzy random variable. For this purpose the same procedure will be followed as in the proof of the Lemma 3.5. Let $\alpha \in [0,1]$ and consider the function $\varphi_\alpha : \Omega \times \mathbb{R} \longrightarrow \mathbb{R}$, defined by $\varphi_\alpha(\omega, x) = \inf_{y \in \widetilde{Y}_\alpha(\omega)} |x-y|$. This function measures the distance from the closed set $\widetilde{Y}_\alpha(\omega)$ to the point $x \in \mathbb{R}$. It is a Carathéodory function, meaning that if $\omega$ is fixed, the map $x \mapsto \varphi_\alpha(\omega, x)$ is continuous because $\widetilde{Y}_\alpha(\omega)$ is closed; and if $x \in \mathbb{R}$ is fixed, the map $\omega \mapsto \varphi_\alpha(\omega, x)$ is $\mathcal{A}$-measurable. Applying Lemma 8.2.6 from [1], it is ensured that $\varphi_\alpha$ is $\mathcal{A} \otimes \mathcal{B}(\mathbb{R})$-measurable. This implies that

$$\begin{aligned} &\Big\{(\omega, x) \in \Omega \times \mathbb{R} : \varphi_\alpha(\omega, x) = 0\Big\} = \Big\{(\omega, x) \in \Omega \times X : x \in \big(\widetilde{Y}\big)_\alpha(\omega)\Big\} \\ =&\Big\{(\omega, x) \in \Omega \times X : x \in \big[Y(\omega)q_1(\alpha) + q_2(\alpha), Y(\omega)r_1(\alpha) + r_2(\alpha)\big]\Big\} \in \mathcal{A} \otimes \mathcal{B}(\mathbb{R}). \end{aligned}$$

The latter establishes that $\widetilde{Y}$ is a fuzzy random variable. Next, it will be proved that the sequence of fuzzy random variables $\widetilde{u}_n = s_n\Delta_1 + \Delta_2$ converges to $\widetilde{Y} = Y\Delta_1 + \Delta_2$ with the metric $\hat{d}$.

$$\begin{aligned} \lim_{n\to\infty} \hat{d}\big(\widetilde{u}_n, \widetilde{Y}\big) &= \lim_{n\to\infty} \sup_{\alpha\in[0,1]} \Big(d_H\big((\widetilde{u}_n)_\alpha, (\widetilde{Y})_\alpha\big)\Big) \\ &= \lim_{n\to\infty} \sup_{\alpha\in[0,1]} \left( \max\Big\{q_1(\alpha)\big|s_n(\omega) - Y(\omega)\big|, r_1(\alpha)\big|s_n(\omega) - Y(\omega)\big|\Big\} \right) \\ &= \lim_{n\to\infty} \sup_{\alpha\in[0,1]} \left( r_1(\alpha)\big|s_n(\omega) - Y(\omega)\big| \right) \\ &= \lim_{n\to\infty} b_4^{(1)}\big|s_n(\omega) - Y(\omega)\big| \\ &= 0. \end{aligned}$$

Again using the function $h : \Omega \longrightarrow \mathbb{R}, h(\omega) = b_4^{(1)}Y(\omega) + b_4^{(2)}$ it can be verified that $\widetilde{Y}$ is integrably bounded. Consider the fuzzy number $\widetilde{W} = \mathbb{E}\big[Y\big]\Delta_1 + \Delta_2$ with $\alpha$-cuts given by

$$\Big[\mathbb{E}\big[Y\big]q_1(\alpha) + q_2(\alpha) \text{ , } \mathbb{E}\big[Y\big]r_1(\alpha) + r_2(\alpha)\Big].$$

From the monotone convergence theorem it can be guaranteed that $\mathbb{E}^*\left[\widetilde{u}_n\right] \xrightarrow[n\to\infty]{} \widetilde{W}$. Now,

$$\begin{aligned}
&\lim_{n\to\infty} \hat{d}\Big(\mathbb{E}^*\left[\widetilde{u}_n\right], \widetilde{W}\Big)\\
=&\lim_{n\to\infty} \sup_{\alpha\in[0,1]} \left( d_H\Big(\big(\mathbb{E}^*[\widetilde{u}_n]\big)_\alpha, (\widetilde{W})_\alpha\Big)\right)\\
=&\lim_{n\to\infty} \sup_{\alpha\in[0,1]} \left( \max\left\{ q_1(\alpha)\Big|\mathbb{E}[s_n]-\mathbb{E}[Y]\Big|, s_1(\alpha)\Big|\mathbb{E}[s_n]-\mathbb{E}[Y]\Big|\right\}\right)\\
=&\lim_{n\to\infty} \sup_{\alpha\in[0,1]} \left( s_1(\alpha)\big|\mathbb{E}[s_n]-\mathbb{E}[Y]\big|\right)\\
=&\lim_{n\to\infty} b_4^{(1)}\big|\mathbb{E}[s_n]-\mathbb{E}[Y]\big|\\
=&0.
\end{aligned}$$

By Lemma 3.5 and the Remark 2.9, it can be obtained that

$$\mathbb{E}^*\big[\widetilde{Y}\big] = \lim_{n\to\infty} \mathbb{E}^*\big[\widetilde{s}_n\big] = \mathbb{E}\big[Y\big]\Delta_1 + \Delta_2. \tag{6}$$

□

## 5. DISCOUNTED MARKOV DECISION PROCESSES

### 5.1. Markov decision processes (MDPs)

In this subsection, the terminology on MDPs presented in [12] is followed. Consider the crisp Markov decision model denoted by $\mathcal{C} := (X, A, \{A(x) : x \in X\}, Q, C)$. Here, $X$ represents the *state space*, and $A$ denotes the *action space*, both considered as Borel spaces. The family $\{A(x) : x \in X\}$ comprises nonempty measurable subsets $A(x)$ of $A$, representing the *feasible actions* when the system is in state $x \in X$. Furthermore, let $\mathbb{K} := \{(x, a)|x \in X, a \in A(x)\}$ denote the set of *feasible state-action pairs*, assumed to be a measurable subset of $X \times A$. The *transition law* $Q$ is characterized as a stochastic kernel on $X$ given $\mathbb{K}$, i. e. $Q$ satisfies the following conditions: $Q(\cdot|x, a)$ is a probability measure on $X$ for each fixed $(x, a) \in \mathbb{K}$, and $Q(W|\cdot)$ is a measurable function on $\mathbb{K}$ for each fixed $W \in \mathcal{B}(X)$, where $\mathcal{B}(X)$ is the Borel $\sigma$-algebra of $X$. Finally, $C : \mathbb{K} \to \mathbb{R}$ is a measurable function called the *cost-per-stage function.* The dynamics of the system unfold as follows: at time $t$, $t \in 0, 1, 2, \ldots$, let the system occupy state $x_t = x \in X$. At this point, a decision maker or controller selects an action $a_t = a \in A(x)$. Subsequently, a cost $C(x, a)$ is accrued, and the system transitions to a new state $x_{t+1} \in X$ in accordance with the transition law $Q(\cdot|x, a)$. Following this transition, a new decision is made, initiating a recursive process. The following structural assumption will be enforced (see Assumption 4.2.1, p. 46 in [12]).

**Assumption 5.1.** a) The one-stage cost $C$ is lower semicontinuous, nonnegative and inf-compact on $\mathbb{K}$, i. e. the set $\{a \in A(x) : C(x, a) \leq \lambda\}$ is compact for every $x \in X$ and $\lambda \in \mathbb{R}$.

b) The transition law $Q$ is strongly continuous.

In the model $\mathcal{C}$ outlined previously, the *admissible history* of a Markov Decision Process (MDP) up to the $n$th transition, denoted as $h_n$, is defined for each $n = 0, 1, \ldots$. This history is constructed as $h_n = (x_0, a_0, \ldots, x_{n-1}, a_{n-1}, x_n)$, where $(x_k, a_k) \in \mathbb{K}$ for $k = 0, 1, \ldots, n-1$, and $x_n \in X$. A *control policy* $\pi = \pi_n$ is delineated as a sequence of stochastic kernels on $A$ given $\mathbb{H}_n$, adhering to the constraint $\pi_n(A(x_n)|h_n) = 1$ for every $h_n \in \mathcal{H}_n$, $n = 0, 1, \ldots$. The collection of all policies is denoted by $\Pi$. Let $\mathbb{F}$ denote the collection of all measurable functions $f : X \to A$, ensuring that $f(x) \in A(x)$ holds for every $x \in X$. Consequently, a *Markov policy* is a sequence $\{f_t\}$ such that $f_t \in \mathbb{F}$, for $t = 0, 1, \ldots$. In particular, a Markov policy $\pi = \{f_t\}$ is said to be *stationary* if $f_t$ is independent of $t$, i.e. $f_t = f \in \mathbb{F}$, for all $t = 0, 1, \ldots$. In such instances, $f_t$ is simply denoted as $f$, and $\mathbb{F}$ is naturally identified as the set of stationary policies.

Consider the measurable space $(\Omega, \mathcal{A})$, comprising the canonical sample space $\Omega := (X \times A)^\infty$ and its corresponding product $\sigma$-algebra $\mathcal{A}$. Here, elements of $\Omega$ take the form of sequences $\omega = (x_0, a_0, x_1, a_1, \ldots)$, where $x_t \in X$ and $a_t \in A$ for all $t = 0, 1, 2, \ldots$. Given that the initial state $x_0 = x \in X$ and the policy $\pi \in \Pi$ used to drive the system, via the theorem of Ionescu-Tulcea (see C.10 Proposition in [12]), there is a unique probability measure $\mathbb{P}_x^\pi$ on $(\Omega, \mathcal{A})$ which is supported on $\mathbb{H}^\infty$, i.e., $\mathbb{P}_x^\pi(\mathbb{H}^\infty) = 1$. The stochastic process $(\Omega, \mathcal{A}, \mathbb{P}_x^\pi, \{x_t\})$ is called a *discrete-time Markov decision process.* The expectation operator concerning $\mathbb{P}_x^\pi$ is denoted by $\mathbb{E}_{x,\pi}$.

Now, the optimal control problem associated with the crisp Markov Decision model $\mathcal{C}$ will be introduced. For this purpose, consider $\pi \in \Pi$ and $x \in X$, and define the *total expected discounted cost* as follows

$$v(x, \pi) := \mathbb{E}_{x,\pi} \left[ \sum_{t=0}^{\infty} \beta^t C(x_t, a_t) \right], \tag{7}$$

where $\beta \in (0, 1)$ represents a predetermined discount factor.

Consider the performance criterion (7), then the *optimal control problem* consists of determining a policy $\pi_o$, such that

$$v(x, \pi_o) = \inf_{\pi \in \Pi} v(x, \pi), \tag{8}$$

$x \in X$, and $\pi_o$ will be called an *optimal policy.* The function $V_o$ defined by

$$V_o(x) := \inf_{\pi \in \Pi} v(x, \pi),$$

$x \in X$, will be called the *optimal value function.*

The following assumption is necessary to guarantee the finiteness property of the optimal value function (see Assumption 4.2.2, p. 46 in [12]).

**Assumption 5.2.** There exists a policy $\pi \in \Pi$ such that $v(x, \pi) < \infty$ for each $x \in X$.

Now, the authors are poised to introduce the following result, which ensures the existence of stationary optimal policies. Moreover, it outlines a dynamic programming procedure for determining such a policy and its corresponding optimal value function (see Theorem 4.2.3 in [12]).

**Theorem 5.3.** Under Assumptions 5.1 and 5.2, the following statements hold.

a) The optimal value function $V_o$ satisfies the next dynamic programming equation for each $x \in X$

$$V_o(x) = \min_{a \in A(x)} \left[ C(x,a) + \beta \int_X V_o(y) Q(\mathrm{d}y|x,a) \right]. \tag{9}$$

b) There exists $f_o \in \mathbb{F}$ such that $f_o(x) \in A(x)$ attains the minimum in (9), i.e. for each $x \in X$

$$V_o(x) = C(x, f_o(x)) + \beta \int_X V_o(y) Q(\mathrm{d}y|x, f_o(x)), \tag{10}$$

and $f_o$ is optimal.

In the following section, this result will be applied to characterize the optimal policy for the fuzzy control problem.

### 5.2. Fuzzy MDPs

Let $\mathcal{C} := (X, A, \{A(x) : x \in X\}, Q, C)$ denote the model introduced in the previous subsection. In this subsection, the fuzzy version of this model will be defined. To accomplish this, the following conditions on the fuzzy cost function will be considered.

**Assumption 5.4.** Let $B, D, F, G, B_1, D_1, F_1$, and $G_1$ be nonnegative real numbers such that $B < D \le F < G$ and $B_1 < D_1 \le F_1 < G_1$. It will be assumed that

$$\begin{aligned} \widetilde{C}(x,a) &= (B,D,F,G)C(x,a) +^* (B_1, D_1, F_1, G_1) \\ &= (BC(x,a) + B_1, DC(x,a) + D_1, FC(x,a) + F_1, GC(x,a) + G_1), \end{aligned}$$

$x \in X, a \in A(x)$.

Thus, the fuzzy Markov decision model is given by $\mathcal{M} := (X, A, \{A(x) : x \in X\}, Q, \widetilde{C})$.

As a consequence of Lemma 4.1, the following result holds.

**Lemma 5.5.** Suppose that Assumption 5.4 is valid. Take $x \in X$, $\pi \in \Pi$, $t \ge 0$, and suppose that $\mathbb{E}_{x,\pi}[C(x_t, a_t)] < \infty$. Then, $\widetilde{C}(x_t, a_t)$ is a fuzzy random variable and

$$\mathbb{E}^*_{x,\pi}\left[\widetilde{C}(x_t,a_t)\right] = \mathbb{E}_{x,\pi}[C(x_t,a_t)](B,D,F,G) +^* (B_1, D_1, F_1, G_1).$$

**Definition 5.6.** Let $\mathcal{M} := (X, A, \{A(x) : x \in X\}, Q, \widetilde{C})$ be the fuzzy Markov decision model. For each policy $\pi \in \Pi$ and state $x \in X$, it is defined that

$$\widetilde{v}_T(x,\pi) := \sum_{t=0}^{T-1}{}^{*} \beta^t \mathbb{E}^*_{x,\pi}\left[\widetilde{C}(x_t,a_t)\right], \tag{11}$$

where $T$ is a positive integer and $\beta \in (0,1)$ is a predetermined discount factor, and it is assumed that $\mathbb{E}_{x,\pi}[C(x_t,a_t)] < \infty$, $t = 0, 1, \ldots, T$. $\widetilde{V}_T$ is called the *fuzzy total expected discounted cost with finite horizon $T$*.

**Remark 5.7.** Observe that $\widetilde{v}_T$ in (11) is well-defined due to the arguments exposed in Section 4 and Lemma 2.5. Moreover, the following equality holds for each $x \in X$ and $\pi \in \Pi$:

$$\begin{aligned}\tilde{v}_T(x,\pi) = (Bv_T(x,\pi) + B_1\frac{1-\beta^T}{1-\beta}, Dv_T(x,\pi) + D_1\frac{1-\beta^T}{1-\beta},\\ Fv_T(x,\pi) + F_1\frac{1-\beta^T}{1-\beta}, Gv_T(x,\pi) + G_1\frac{1-\beta^T}{1-\beta}),\end{aligned}$$

as a consequence of (6), where $v_T(x,\pi)$ is the crisp total expected discounted cost with finite horizon $T$, i. e.

$$v_T(x,\pi) := \mathbb{E}_{x,\pi}\left[\sum_{t=0}^{T-1}\beta^t C(x_t,a_t)\right]. \tag{12}$$

Notice that for each $x \in X$ and $\pi \in \Pi$,

$$v(x,\pi) = \lim_{T\to\infty} V_T(x,\pi). \tag{13}$$

**Lemma 5.8.** Suppose that Assumption 5.4 holds. Then, for each $x \in X$ and $\pi \in \Pi$ such that $v(x,\pi) < \infty$, the fuzzy sequence $\{\widetilde{v}_T(x,\pi) : T \in \{0,1,\dots\}\}$ converges (see Definition 2.6) to

$$\begin{aligned}\widetilde{v}(x,\pi) \;&:=\; (Bv(x,\pi), Dv(x,\pi), Fv(x,\pi), Gv(x,\pi)) +^* \frac{1}{1-\beta}(B_1, D_1, F_1, G_1)\\ &= \left(Bv(x,\pi) + \frac{B_1}{1-\beta}, Dv(x,\pi) + \frac{D_1}{1-\beta}, Fv(x,\pi) + \frac{F_1}{1-\beta}, Gv(x,\pi) + \frac{G_1}{1-\beta}\right).\end{aligned}$$

P r o o f. Let $\pi \in \Pi$, $x \in X$ and $\alpha \in [0,1]$ be fixed. To simplify the notation in this proof, write $v = v(x,\pi)$ and $v_T = v_T(x,\pi)$. Then, the $\alpha$-cut of $\tilde{v}_T(x,\pi)$ (see Remark 5.7) is given by

$$\Delta_\alpha^T = [q_T(\alpha), r_T(\alpha)], \tag{14}$$

where

$$q_T(\alpha) = B(1-\alpha)v_T + B_1(1-\alpha)\frac{1-\beta^T}{1-\beta} + \alpha Dv_T + \alpha D_1\frac{1-\beta^T}{1-\beta}$$

and

$$r_T(\alpha) = G(1-\alpha)v_T + G_1(1-\alpha)\frac{1-\beta^T}{1-\beta} + \alpha Fv_T + \alpha F_1\frac{1-\beta^T}{1-\beta}.$$

In a similar way, the $\alpha$-cut of $\tilde{v}(x,\pi)$ is

$$\Delta_\alpha = [B(1-\alpha)v + \frac{B_1(1-\alpha)+\alpha D_1}{1-\beta} + \alpha Dv, G(1-\alpha)v + \frac{G_1(1-\alpha)+\alpha F_1}{1-\beta} + \alpha Fv].$$

Then, applying (3) and due to the identity $\max(c,b) = (c+b+|b-c|)/2$ with $b, c \in \mathbb{R}$, it yields that

$$d_H(\Delta_\alpha^T, \Delta_\alpha) = H(1-\alpha)(v - v_T) + H_1(1-\alpha)\frac{\beta^T}{1-\beta} + \alpha F(v - v_T) + F_1\alpha\frac{\beta^T}{1-\beta}.$$

Consequently, through (4),

$$\begin{aligned} \hat{d}(\tilde{v}_T, \tilde{v}) &= \sup_{\alpha\in[0,1]} d_H(\Delta_\alpha^T, \Delta_\alpha) \\ &= H(V - V_T) + H_1\frac{\beta^T}{1-\beta} + F(V - V_T) + F_1\frac{\beta^T}{1-\beta}. \end{aligned} \tag{15}$$

Then, when $T \longrightarrow \infty$ in (15), using (13) and the fact that $\beta^T \longrightarrow 0$ when $T \longrightarrow \infty$, it concludes that

$$\lim_{T\longrightarrow\infty} \hat{d}(\widetilde{v}_T, \widetilde{v}) = 0.$$

Therefore, since $\pi$ and $x$ are arbitrary, the result follows. □

Let $\mathcal{M} = (X, A, \{A(x) : x \in X\}, Q, \widetilde{C})$ be the fuzzy Markov model. For each policy $\pi \in \Pi$ and state $x \in X$ such that $v(x,\pi) < \infty$, the *fuzzy total expected discounted cost* is defined as follows:

$$\tilde{v}(\pi, x) = \sum_{t=0}^{\infty}{}^{*} \beta^t \mathbb{E}_{x,\pi}^{*}\left[\widetilde{C}(x_t, a_t)\right].$$

Now, the corresponding *fuzzy optimal control problem with respect to the max-order* is as follows: determine $\pi_o \in \Pi$ (if it exists) such that:

$$\tilde{v}(x, \pi) \leq^{*} \tilde{v}(x, \pi_o) \tag{16}$$

for all $x \in X$ and $\pi \in \Pi$. In this case, it is possible to write

$$\tilde{v}(x, \pi_o) = \inf_{\pi\in\Pi}{}^{*} \tilde{v}(x, \pi),$$

$x \in X$, and it is said that $\pi_o$ is *optimal* with respect to the max-order. Moreover, the function $\widetilde{V}_o(x) = \tilde{v}(x, \pi_o)$, $x \in X$, will be called the *optimal fuzzy cost function.*

**Remark 5.9.** Observe that when in the decision model $\widetilde{C}(x,a)$ has a membership function given by:

$$(\widetilde{C}(x,a))'(x) = \begin{cases} 1 & \text{if } z = C(x,a) \\ 0 & \text{if } z \neq C(x,a), \end{cases} \tag{17}$$

for all $x \in X$ and $a \in A(x)$, and considering Remark 2.10, it follows that the fuzzy optimal control problem given in (16) is reduced to the optimal control problem described in (8).

**Theorem 5.10.** Suppose that Assumptions 5.1, 5.2, and 5.4 hold. Then, the following statements hold.

a) The optimal policy of the fuzzy decision problem coincides with the optimal policy of the crisp control problem $f_o$ (see Theorem 5.3).

b) The optimal fuzzy cost function is given by

$$\tilde{V}_o(x) = (B, D, F, G)\, v(x, f_o) +^* \frac{1}{1-\beta}\, (B_1, D_1, F_1, G_1) \tag{18}$$

$x \in X$.

Proof. Set fixed $\pi \in \Pi$, $x \in X$, and $\alpha \in [0, 1]$. In order to simplify the exposition of the demonstration, consider the following notation:

$$\theta_1 = \frac{B_1}{1-\beta}, \theta_2 = \frac{D_1}{1-\beta}, \theta_3 = \frac{F_1}{1-\beta}, \theta_4 = \frac{G_1}{1-\beta},$$

and

$$v = v(\pi, x).$$

Then the $\alpha$-cut $\Delta_\alpha$ of $\widetilde{v}(\pi, x)$ is given by

$$\Delta_\alpha \quad = \quad [(1-\alpha)(Dv + \theta_1) + \alpha(Dv + \theta_2), (1-\alpha)(Gv + \theta_4) + \alpha(Fv + \theta_3)].$$

Now, since $v \geq v(x, f_o)$, the left side of the interval $\Delta_\alpha$ satisfies

$$(1-\alpha)(Bv + \theta_1) + \alpha(Dv + \theta_2) \geq (1-\alpha)(Bv(x, f_o) + \theta_1) + \alpha(Dv(x, f_o) + \theta_2),$$

and, similarly, the right side of the interval $\Delta_\alpha$ fulfill

$$(1-\alpha)(Gv + \theta_4) + \alpha(Fv + \theta_3) \geq (1-\alpha)(Gv(x, f_o) + \theta_4) + \alpha(Fv(x, f_o) + \theta_3).$$

Since

$$[(1-\alpha)(Bv(x, f_o) + \theta_1) + \alpha(Dv(x, f_o) + \theta_2), (1-\alpha)(Gv(x, f_o) + \theta_4) + \alpha(Fv(x, f_o) + \theta_3)],$$

is the $\alpha$-cut of $\widetilde{v}(x, f_o)$ and as $\pi$, $x$ and $\alpha$ are arbitrary, it results that

$$\widetilde{v}(x, f_o) \leq^* \widetilde{v}(x, \pi).$$

Then, $f_o$ is an optimal policy for the fuzzy control problem, see (16). Finally, (21) is a direct consequence of Lemma 5.8. Therefore, Theorem 5.10 follows. □

Similarly, fuzzy optimal policies can be defined with respect to ranking comparisons. Thus, the *fuzzy optimal problem with respect to the ranking order* ranking is as follows: Determine $\pi_o^r \in \Pi$ (if it exists) such that:

$$\widetilde{v}(x, \pi_o^r) \leq^{**} \widetilde{v}(x, \pi) \tag{19}$$

for all $x \in X$, and $\pi \in \Pi$. In this case, it is possible to write

$$\widetilde{v}(x, \pi_o^r) = \inf_{\pi \in \Pi}{}^{**} \tilde{v}(x, \pi), \tag{20}$$

$x \in X$, and it is said that $\pi_o^r$ is *optimal.* Moreover, the function $\widetilde{V}_o^r(x) = \widetilde{v}(x, \pi_o^r)$, $x \in X$, will be called the *optimal fuzzy cost function with respect to the ranking.*

**Remark 5.11.** If in Assumption 5.4 the degenerated case is considered in which $B = D = F = G = 1$ and $B_1 = D_1 = F_1 = G_1 = 0$, it results that

$$\widetilde{C}(x,a) = C(x,a)\widetilde{1}$$

for all $x \in X$ and $a \in A(x)$, then

$$\widetilde{v}(x,\pi) = v(x,\pi)\widetilde{1}$$

and that

$$\Re(\widetilde{v}(x,\pi)) = v(x,\pi)$$

for all $\pi \in \Pi$ and $x \in X$. It follows that the optimal control problem described in (19) is reduced to the standard optimal control problem given in (8).

**Theorem 5.12.** Under assumptions 5.1, 5.2, and 5.4, the following statements hold.

a) The optimal policy of the fuzzy decision problem is $f_o$ (see Theorem 5.3).

b) The optimal fuzzy cost function is given by

$$\tilde{V}_o(x) = (B, D, F, G)\, v(x, f_o) +^* \frac{1}{1-\beta}\,(B_1, D_1, F_1, G_1)\,, x \in X. \tag{21}$$

P r o o f. Fix $\pi \in \Pi$ and $x \in X$. Then, it is obtained that

$$\Re(\widetilde{v}(x,\pi)) = v(x,\pi)\Re((B,D,E,F)) + \Re((B_1,D_1,E_1,F_1)).$$

Now, since $v(x,\pi) \geq v(x,f_o) \geq 0$, it follows that

$$\Re(\widetilde{v}(x,\pi)) \geq v(x,f_o)\Re((B,D,E,F)) + \Re((B_1,D_1,E_1,F_1)),$$

thus,

$$\Re(\widetilde{v}(x,\pi)) \geq \Re((\widetilde{v}(x,f_o)).$$

Since $\pi$ and $x$ are arbitrary, $f_o$ is optimal with respect to the ranking comparison (see (19), (20)). Therefore, Theorem 5.12 follows from Lemma 5.8. □

## 6. A FUZZY VERSION OF THE LINEAR-QUADRATIC MODEL

In this section, the results presented using a linear system with quadratic one-stage cost will be illustrated. This example is a variation of the standard linear-quadratic (LQ) problem (see [2]). Of course, the standard LQ problem could also be solved using the same technique.

Let $X = A = A(x) = \mathbb{R}, x \in X$. The cost function and the dynamic of the system are given by

$$C(x,a) = x^2 + a^2 + ax,$$

$$x_{t+1} = x_t + a_t + \xi_t,$$

$$t = 0, 1, 2, \ldots, (x,a) \in \mathbb{K}.$$

The random variables $\xi_t, t = 0, 1, \ldots$, are assumed to be i.i.d. with common continuous density $\Theta$, with zero mean and finite variance $\sigma^2$. In this case, the transition law is induced as follows

$$Q(W|x,a) = \int_W \Theta(s - x - a)\,\mathrm{d}s,$$

for $W \in \mathcal{B}(X)$ and $(x, a) \in \mathbb{K}$.

Assumptions 5.1 and 5.2 are thoroughly demonstrated in [5]. Additionally, the proof of the following result can also be found in [5].

**Lemma 6.1.** The optimal value function and the optimal policy for the crisp optimal control problem are given by

$$V_o(x) = Kx^2 + L,$$
$$f_o^{LQ}(x) = -\frac{1 + 2\beta K}{2 + 2\beta K}x,$$

respectively, where $x \in X$, $L = [\beta K/(1-\beta)]\sigma^2$ and $K$ is the positive solution of the Ricatti's equation:

$$4\beta K^2 + (4 - 4\beta)K - 3 = 0.$$

Suppose that the authors aim to model the cost to be approximately within the interval $[C(x,a) + 2\epsilon, C(x,a) + 3\epsilon]$, for a given positive number $\epsilon$ and $(x,a) \in \mathbb{K}$. In this scenario, it is possible to consider the following cost function:

$$\begin{aligned}\widetilde{C}(x,a) &= (1/2, 1, 1, 2)\, C(x,a) +^* (\varepsilon, 2\varepsilon, 3\varepsilon, 4\varepsilon) \\ &= \left(\frac{1}{2}C(x,a) + \varepsilon, C(x,a) + 2\varepsilon, C(x,a) + 3\varepsilon, 2C(x,a) + 4\varepsilon\right).\end{aligned} \tag{22}$$

The fuzzy cost function in (22) satisfies Assumption 5.4, then the following statements hold:

a) $f_o^{LQ}$ represents the fuzzy optimal policy in terms of the max-order, see Theorem 5.10.

b) $f_o^{LQ}$ is the fuzzy optimal policy for ranking order, see Theorem 5.12.

c) The optimal fuzzy cost function is given by:

$$\widetilde{V}_o(x) = \left(\frac{1}{2}V_o(x) + \varepsilon, V_o(x) + 2\varepsilon, V_o(x) + 3\varepsilon, V_o(x) + 4\varepsilon\right).$$

$x \in X$.

## 7. CONCLUSIONS

MDPs on Borel spaces with trapezoidal fuzzy costs and discounted objective functions were analyzed. These trapezoidal costs model the fact that the cost of a standard MDP is approximately in a certain interval. The optimal solutions for the discounted MDPs studied have been obtained, and these optimal solutions have been related to optimal solutions of suitable classical MDPs. The results presented extend the classical discounted MDPs theory in two cases which allows to consider, under general state and action spaces: (i) fuzzy discounted MDPs with respect to the max-order, and (ii) fuzzy discounted MDPs with respect to the average ranking order.

## 8. ACKNOWLEDGMENTS

This works was partially supported by CONAHCyT-México under Grant No. CF-2023-I-1362.

(Received October 31, 2024)

## REFERENCES

[1] J. P. Aubin and H. Frankowska: Set-Valued Analysis. Birkhäuser, Boston 2009.

[2] D. Bertsekas: Dynamic programming and optimal control: Volume I. Athena Sci. (2012).

[3] K. Carrero-Vera, H. Cruz-Suárez, and R. Montes-de-Oca: Finite-horizon and infinite-horizon Markov decision processes with trapezoidal fuzzy discounted rewards. Commun. Comput. Inf. Sci., Springer, Cham *1623* (2022), 171–192.

[4] K. Carrero-Vera, H. Cruz-Suárez, and R. Montes-de-Oca: Markov decision processes on finite spaces with fuzzy total reward. Kybernetika *58* (2022), 2, 180–199. DOI:10.14736/kyb-2022-2-0180

[5] H. Cruz-Suárez and R. Montes-de-Oca: An envelope theorem and some applications to discounted Markov decision processes. Math. Oper. Res. *67* (2008), 299–321. DOI:10.1007/s00186-007-0155-z

[6] H. Cruz-Suárez, R. Montes-de-Oca, and R. I. Ortega-Gutiérrez: An extended version of average Markov decision processes on discrete spaces under fuzzy environment. Kybernetika *59* (2023), 1,160–178. DOI:10.14736/kyb-2023-1-0160

[7] H. Cruz-Suárez, R. Montes-de-Oca, and R. Ortega-Gutiérrez: Deterministic discounted Markov decision processes with fuzzy rewards/costs. Fuzzy Inf. Engrg. *15* (2023), 3, 274–290. DOI:10.26599/FIE.2023.9270020

[8] J. de Andrés-Sánchez: A systematic review of the interactions of fuzzy set theory and option pricing. Expert Syst. Appl. *223* (2023), 119868. DOI:10.1016/j.eswa.2023.119868

[9] P. Diamond and P. Kloeden: Metric Spaces of Fuzzy Sets: Theory and Applications. World Scientific, Singapore 1994.

[10] J. C. Figueroa-García, G. Hernández, and C. Franco: A review on history, trends and perspectives of fuzzy linear programming. Oper. Res. Perspect. *9* (2022), 100247. DOI:10.1016/j.orp.2022.100247

[11] N. Furukawa: Parametric orders on fuzzy numbers and their roles in fuzzy optimization problems. Optimization *40* (1997), 171–192. DOI:10.1080/02331939708844307

[12] O. Hernández-Lerma and J. B. Lasserre: Discrete-Time Markov Control Processes: Basic Optimality Criteria. Springer-Verlag, New York, 1996.

[13] S. Kambalimath and P. C. Deka : A basic review of fuzzy logic applications in hydrology and water resources. Appl. Water Sci. *10* (2020), 8, 1–14. DOI:10.1007/s13201-020-01276-2

[14] M. Kurano, M. Yasuda, J. Nakagami, and Y. Yoshida: Markov-type fuzzy decision processes with a discounted reward on a closed interval. Eur. J. Oper. Res. *92* (1996), 3, 649–662. DOI:10.1016/0377-2217(95)00140-9

[15] M. Kurano, M. Yasuda, J. Nakagami, and Y. Yoshida: Markov decision processes with fuzzy rewards. J. Nonlinear Convex Anal. *4* (1996), 1, 105–116.

[16] M. Kurano, M. Hosaka, J. Song, and Y. Huang: Controlled Markov set-chains with discounting. J. Appl. Prob. *35* (1998), 3, 293–302. DOI:10.1006/cogp.1998.0683

[17] M. López-Díaz and D. A. Ralescu: Tools for fuzzy random variables: embeddings and measurabilities. Comput. Statist. Data Anal. *51* (2006), 109–114. DOI:10.1016/j.csda.2006.04.017

[18] M. L. Puri and D. A. Ralescu: Fuzzy random variable. J. Math. Anal. Appl. *114* (1986), 402–422.

[19] D. Rani, and T. R. Gulati : A new approach to solve unbalanced transportation problems in imprecise environment. J. Transp. Secur. *7* (2014), 3, 277–287. DOI:10.1007/s12198-014-0143-5

[20] D. Rani, T. R. Gulati, and A. Kumar : A method for unbalanced transportation problems in fuzzy environment. Sadhana *39* (2014), 3, 573–581. DOI:10.1007/s12046-014-0243-8

[21] S. Rezvani and M. Molani: Representation of trapezoidal fuzzy numbers with shape function. Ann. Fuzzy Math. Inform. *8* (2014), 89–112.

[22] A. Semmouri, M. Jourhmane, and Z. Belhallaj: Discounted Markov decision processes with fuzzy costs. Ann. Oper. Res. *295* (2020), 769–786. DOI:10.1007/s10479-020-03783-6

[23] A. Syropoulos and T. Grammenos: A Modern Introduction to Fuzzy Mathematics. Wiley, New Jersey 2020.

[24] L. Zadeh: Fuzzy sets. Inform. Control *8* (1965), 338–353. DOI:10.1016/S0019-9958(65)90241-X

[25] W. Zhou, D. Lou, and Z. Xu: Review of fuzzy investment research considering modelling environment and element fusion. Int. J. Syst. Sci. *53* (2022), 9, 1958–1982. DOI:10.1080/00207721.2022.2031340

*Salvador De-Jesús-Hernández, Departamento de Matemáticas, Universidad Autónoma Metropolitana-Iztapalapa, Av. Ferrocarril San Rafael Atlixco 186, Col. Leyes de Reforma 1ª Sección, Alcaldía Iztapalapa, C.P. 09310, CDMX. México.*
*e-mail: salvador.hernandez@xanum.uam.mx*

*Hugo Cruz-Suárez, Facultad de Ciencias Físico Matemáticas, Benemérita Universidad Autónoma de Puebla, Av. San Claudio y Río Verde, Col. San Manuel, CU, Puebla, Pue. 72570. México.*
*e-mail: hcs@fcfm.buap.mx*

*Raúl Montes-de-Oca, Departamento de Matemáticas, Universidad Autónoma Metropolitana-Iztapalapa, Av. Ferrocarril San Rafael Atlixco 186, Col. Leyes de Reforma 1ª Sección, Alcaldía Iztapalapa, C.P. 09310, CDMX. México.*
*e-mail: momr@xanum.uam.mx*