

Mono-Objective Function Analysis Using an Optimization Approach

J. B. Clempner

Abstract— In this paper we propose an evolutionary technique based in a Lyapunov method for mono-objective optimization, that associate to every ergodic controllable finite Markov Chains a Lyapunov-like mono-objective function. For representing the trajectory-dynamics properties local-optimal policies are defined to minimize the one-step decrement of the cost-function. We propose a state-value function that increase and decrease between states of the Markov decision processes. Then, we show that a Lyapunov mono-objective function, which can only decrease over time, can be built for the system. For illustration purposes, we present a simulated experiment that shows the trueness of the suggested method.

Keywords— Lyapunov, problem solving control methods, search heuristic methods, artificial intelligence, genetic algorithms, optimization, vector optimization.

I. INTRODUCCION

LA OPTIMIZACIÓN de funciones objetivo es un tema importante, tanto para investigadores como para practicantes, por el hecho de que todavía existen problemas abiertos en esta área. En muchos problemas científicos y de ingeniería es dada una función $f(x)$. El objetivo es buscar un valor numérico x^* que minimice (maximice) la solución de la función objetivo, esto es, un valor x^* para el cual la desigualdad $f(x) < f(x^*)$ no sea viable [21]. El concepto de Pareto óptimo fue formulado por Vilfredo Pareto [16] en el siglo XIX y constituye por sí mismo el origen de la investigación en la optimización de funciones objetivo. Decimos que un punto x^* en una región dada es Pareto óptimo si para para toda x sucede una de ambas cosas:

$$\bigwedge_{i \in I} (f(x_i) = f(x_i^*))$$

o, existe al menos una $i \in I$ tal que

$$f(x_i) > f(x_i^*)$$

En palabras, esta definición dice que x_i^* es Pareto óptimo si no existe un vector dado x que pueda hacer decrecer uno de los criterios sin causar un incremento simultáneo en al menos uno de los otros criterios. Desafortunadamente, el criterio de Pareto óptimo no da una solución única, sino un conjunto de soluciones llamada *soluciones no dominadas*.

En la literatura han sido propuestas diferentes

aproximaciones para tratar con funciones que tienen una solución única o soluciones múltiples. En ese sentido, muchos métodos han sido sugeridos que van desde combinaciones suaves de objetivos múltiples en un objetivo único, hasta el próximo y el extremo receptor de la cadena de comunicación.

Una clase muy amplia de procesos estocásticos controlables de Markov (usualmente llamados Modelos de Decisión de Markov) han sido estudiados por practicantes en las áreas de matemáticas, investigación de operaciones, ingeniería eléctrica y economía ([11], [7], [13], [2], [3], [9], [10], [17]), y también en *algoritmos genéticos* ([8], [14], [19] y [20]). El diseño de controles óptimos ha sido usualmente basado en la obtención de ecuaciones de Bellman estocásticas (Programación Dinámica [1]), las cuales, en casos simples de procesos finitos, se ha demostrado que pueden ser transformadas en un problema de programación lineal de horizonte de tiempo infinito. Las ecuaciones de Bellman relacionan la función objetivo, por sí misma, con un problema dinámico. Para los problemas de horizonte con tiempo finito esta aproximación requiere de la aplicación un procedimiento “backward” específico que, en general, es una tarea difícil de ejecutar. Por el contrario, la aplicación de políticas localmente óptimas no requiere de ningún procedimiento “backward” (inclusive aplicando cualquier tipo de optimización local) y utilizan solamente la información actual en el pasado y la información actual de los estados del proceso que deberán ser optimizados. Los resultados más avanzados relacionados con los problemas de optimización han sido obtenidos para una clase específica de procesos finitos de Markov, donde el comportamiento futuro depende en solamente el estado actual del proceso y es independiente de su prehistoria [17].

En este trabajo, se introduce un paradigma de modelado que asocia a todo proceso ergódico una función Lyapunov mono-objetivo para desarrollar la representación de procesos de decisión ([4], [5]). Si el proceso ergódico converge, entonces se alcanza un punto de equilibrio, inclusive uno justificable ([17], [18]). La utilización de una función Lyapunov para el proceso de seguimiento de una trayectoria garantiza la convergencia e inclusive tiene propiedades particulares y fortalecidas. En ese sentido, se presenta la construcción de una función Lyapunov (con un comportamiento monotónico decreciente) que tiene una relación uno a uno con la función de costo dada (ecuación de Bellman). Acotada por abajo una función Lyapunov decreciente asegura la existencia de un punto de equilibrio, o en otras palabras, la convergencia en la aplicación de políticas localmente óptimas y, además, asegura la convergencia de la función de costo a un valor mínimo.

Se considera una clase específica de cadenas controlables de Markov soportadas por una función de costo dada. Las políticas localmente óptimas son definidas para minimizar el decremento “un-paso” de la función de costos. Para formalizar

la premisa, se propone una función no convergente que se incrementa y decrementa entre los estados del proceso de decisión. Entonces demostramos que es posible representar esa función en un formato recursivo usando políticas (fijas) localmente óptimas. Se muestra que una función Lyapunov puede ser construida utilizando la expresión recursiva para un proceso de decisión de Markov, esto es, la función Lyapunov resultante es una función monotónica que solo puede decrecer (o permanecer igual) en el tiempo.

Las principales contribuciones de este artículo consisten en lo siguiente: a) se muestra a través de ejemplos numéricos que el comportamiento de la secuencia de costos, correspondiente a una política localmente óptima, usualmente tiene un carácter no monotónico que no permite probar exactamente la existencia de un punto de equilibrio; b) se sugiere un mapeo uno a uno entre la función de costos actual y la función de energía nueva (Lyapunov) que no crece de manera monotónica sobre las trayectorias del sistema bajo la aplicación de políticas localmente óptimas. Esta función, dada la política no monotónica, puede ser tratada como una función específicamente construida para el proceso considerado. La función (Lyapunov) acotada por abajo converge por el teorema de Weierstrass, lo que implica que la secuencia de costos correspondiente también converge, c) para la clase de cadenas finitas controlables de Markov [17] la convergencia de la política localmente óptima es alcanzada. Además, se provee una fórmula analítica para la realización numérica de la política localmente óptima

El artículo está estructurado de la siguiente manera. La siguiente sección presenta los antecedentes matemáticos y la terminología necesaria para entender el resto del artículo. La sección 3 presenta el modelo de decisión donde todas las suposiciones estructurales son introducidas dando un análisis detallado del proceso de decisión. En la sección 4 se muestra un método para la construcción de la función objetivo de Lyapunov. Un experimento de simulación es presentado en la sección 5. Finalmente, en la sección 6 se presentan conclusiones y trabajos futuros.

II. ANTECEDENTES

El objetivo de esta sección es introducir un modelo de decisión y todas las suposiciones estructurales relacionadas con el modelo de Markov ([17]).

Sea X un conjunto contable, llamado el *espacio de estado*, que consiste de todos lo posible enteros $N \in \mathbb{N}$ de estados $\{x(1), \dots, x(N)\}$. Una *cadena de Markov* es una secuencia de X -valores de variables aleatorias $x_n, n \in \mathbb{N}$, que satisfacen la siguiente condición de *Markov*:

$$P(x_{n+1} = x(j) | x_n = x(i)) = \pi(ij) \quad (1)$$

La cadena de Markov puede ser representada por una gráfica completa cuyos nodos son los estados y cada $(x(i), x(j)) \in X^2$ es etiquetado por la probabilidad de transición (1). La matriz $\Pi = (\pi(ij))_{(x(i), x(j)) \in X} \in [0, 1]^{X \times X}$ y determina la evolución de la cadena: para cada $k \in \mathbb{N}$, la potencia Π^k tiene en cada entrada $(x(i), x(j))$ la probabilidad

de ir desde el estado $x(i)$ al estado $x(j)$ en exactamente k pasos.

Un proceso de *Decisión de Markov* es una 5-tupla

$$PDM = \{X, A, \Theta, \Pi, L\} \quad (2)$$

donde: 1) X es un conjunto contable de estados posibles, $X \subset \mathbb{N}$, dotado de una topología discreta; 2) A es el conjunto de acciones, que es un espacio métrico. Para toda $x \in X$, $A(x) \subset A$ es un conjunto no vacío de acciones admisibles en el estado $x \in X$. Sin perder generalidad se considera $A = \bigcup_{x \in X} A(x)$; 3) $\Theta = \{(x, a) | x \in X, a \in A(x)\}$ es el conjunto de pares estado-acción admisible, que es un subconjunto medible de $X \times A$; 4) $\Pi(k) = [\pi(ij | k)]$ es una matriz de transición estacionaria controlada, donde

$$\pi(ij | k) \equiv P(x_{n+1} = x(j) | x_n = x(i), a_n = a(k))$$

que representa la probabilidad asociada con la transición desde el estado $x(i)$ al estado $x(j)$ bajo una acción $a(k) \in A(x(i))$; 5) $L : X \rightarrow \mathbb{R}$ es una función estado-valor (función de trayectoria), que asocia a cada estado un valor real.

El modelo de control (2) representa un sistema estocástico de tiempo discreto en el cual un estado está disponible en tiempo $n \in \mathbb{N}$. Si se denota por x_n y a_n al estado del sistema y la acción aplicada en tiempo n , respectivamente, la interpretación de la dinámica del *PDM* es la siguiente: para cada tiempo discreto $n \in \mathbb{N}$ el estado del sistema $x_n = x(i) \in X$ está disponible. Para toda acción $a_n = a(k) \in A(x)$, la probabilidad del sistema de encontrarse a sí mismo en el próximo estado x_{n+1} en el tiempo $n+1$ es $P(x_{n+1} | x_n = x(i), a_n = a(k))$. Considerando los estado previos de la trayectoria (camino, orbita, etc.) (x_0, x_1, \dots, x_n) se obtiene el valor de la función de trayectoria L y, entonces el siguiente estado x_{n+1} es seleccionado de acuerdo a L aplicando algún 'criterio'. La propiedad del criterio del proceso de decisión de *Markov* en (2) se dice que está completo si

$$P(x_{n+1} | x_t (t = \overline{1, n-1}), x_n = x(i), a_n = a(k)) = P(x_{n+1} | x_n = x(i), a_n = a(k))$$

Para toda $n \in \mathbb{N}$ el producto $H_n = \Theta^n \times X$ es el conjunto de historias admisibles hasta el tiempo n . El vector $h_n = (x_0, a_0, \dots, x_{n-1}, a_{n-1}, x_n) \in H_n$ denota la historia del proceso hasta el tiempo n . Considerando los estados previos de la trayectoria (x_0, x_1, \dots, x_n) , y para toda acción $a_n \in A(x(i))$, la probabilidad del sistema de encontrarse en el estado $x(j) \in X$ es $\pi(ij | k)$. Una política u es una (probabilidad aleatoria) regla medible para escoger acciones, que dependen del estado actual. La política

$$u_n(k | i) \equiv P(a(k) | x_n = x(i)) \quad (3)$$

representa la probabilidad medible asociada con la ocurrencia de una acción a_n desde el estado $x_n = x(i)$. Si se denota por S el simplex unitario en \mathbf{R}^k y dado $u_n(k|i) \in S$, esto es,

$$S = \{u(k|i) \in \mathbf{R}^k \mid \sum_{k=1}^M u(k|i) = 1, u(k|i) \geq 0\} \quad (4)$$

En este artículo se investiga la clase de políticas localmente óptimas definidas a continuación.

Una política $\{S_n^{loc}\}_{n \geq 0}$ se dice que es localmente óptima si para toda $n \geq 0$ se minimiza la expectativa condicional matemática de la función de costo $L(x_{n+1})$ bajo la condición de que la prehistoria del proceso

$$F_n := \{S_0, P\{x_0 = x(j)\} \mid j = \overline{1, N}, \dots, S_{n-1}, \\ P\{x_n = x(j)\} \mid j = \overline{1, N}\} \quad (5)$$

es fija y no puede ser cambiada en lo sucesivo, esto es, realiza la regla de optimización condicional "un-paso adelante"

$$S_n^{loc} := \arg \min_{S_n} \mathbb{E}\{L'(x_{n+1}) \mid F_n\} \quad (6)$$

III. MODELO DE DECISIÓN

El objetivo de esta sección es introducir el modelo de decisión y todas las suposiciones estructurales relacionadas con el modelo de Markov ([17]). Para abordar este problema proponemos representar la función estado-valor L usando un modelo lineal (con respecto al control $u \in S$). Después obtenemos la política u que resulta en el valor mínimo de la trayectoria. Finalmente, presentamos L en un formato de matriz recursiva.

La probabilidad de que el sistema se encuentre en el siguiente estado es la siguiente:

$$P(x_{n+1} = x(j) \mid x_n = x(i)) =$$

$$\sum_{k=1}^M P(x_{n+1} = x(j) \mid x_n = x(i), a(k)) u_n(k|i) = \\ \sum_{k=1}^M \pi(ij|k) u_n(k|i)$$

La función estado-valor L para cualquier política dada u está definida sobre todas las posibles combinaciones de estados y acciones, e indica el valor esperado cuando tomamos la acción a en el estado x siguiendo la política u . Los L -valores para todos los estados de (2) pueden ser expresados por

$$\mathbb{E}^u(L(x_{n+1})) :=$$

$$\sum_{j=1}^N \sum_{i=1}^N \sum_{k=1}^M L(x_{n+1} = x(j) \mid \quad (7)$$

$$x_n = x(i), a_n = a(k)) \cdot \pi(ij|k) u_n(k|i) P(x_n = x(i))$$

donde $L(x_{n+1} = x(i) \mid x_n = x(i), a_n = a(k))$ es una constante en el estado $x(i)$ cuando el control $a(k)$ es aplicado (sin perder generalidad se puede asumir positivo) y $P(x_n)$ para cualquier $P(x_0)$ dado está definido de la siguiente manera:

$$P(x_{n+1} = x(j)) = \sum_{i=1}^N P(x_{n+1} = x(j) \mid x_n = x(i))$$

$$P(x_n = x(i)) = \sum_{i=1}^N \left(\sum_{k=1}^M \pi(ij|k) u_n(k|i) \right) P(x_n = x(i))$$

o, en forma de vector,

$$\mathbf{p}_{n+1} = (\mathbf{\Pi}_n)^T \mathbf{p}_n$$

$$(\mathbf{\Pi}_n)_{ij} := \sum_{k=1}^M \sum_{k=1}^M \pi(ij|k) u_n(k|i)$$

En formato de vector, la fórmula (7) puede ser expresada como

$$\mathbf{L}_{n+1} := \mathbb{E}^u(L(x_{n+1})) =$$

$$\sum_{i=1}^N \left[\sum_{j=1}^N \sum_{k=1}^M L(x_{n+1} = x(j) \mid x_n = x(i),$$

$$a_n = a(k)) \pi(ij|k) u(k|i) \cdot P(x_n = x(i)) \right] = \langle \mathbf{v}_n, \mathbf{p}_n \rangle$$

donde

$$(\mathbf{v}_n)_i := \sum_{k=1}^M \left[\sum_{j=1}^N L(ij|k) \pi(ij|k) \right] u(k|i)$$

$$L(ij|k) := L(x_{n+1} = x(i) \mid a_n = a(k))$$

Entonces, sea S el simplex unitario en \mathbf{R}^k , esto es,

$$S = \{r \in \mathbf{R}^k \mid \sum_{k=1}^M r(k) = 1, r(k) \geq 0\}$$

tenemos que,

$$\min_{r(k) \in S} \sum_{k=1}^M v(k) r(k) = \min_{k=1, \dots, M} v(k) = v(\alpha)$$

y la igualdad es alcanzada por $r(k) = \left(\underbrace{0, 0, \dots, 0}_\alpha, 1, 0, \dots, 0 \right)$.

En efecto, es evidente que

$$\sum_{k=1}^M v(k) r(k) \geq \sum_{k=1}^M (\min v(k)) r(k)$$

$$= \min v(k) \sum_{k=1}^M r(k) = \min v(k) = v(\alpha)$$

y la igualdad es alcanzada por $r(k) = \left(\underbrace{0, 0, \dots, 0}_\alpha, 1, 0, \dots, 0 \right)$.

Entonces se tiene que

$$\begin{aligned} \mathbf{L}_{n+1} &= \langle \mathbf{v}_n, \mathbf{p}_n \rangle = \sum_{i=1}^N (\mathbf{v}_n)_i (\mathbf{p}_n)_i \geq \\ & \sum_{i=1}^N \min_{u_n, (k|i) \in S} (\mathbf{v}_n)_i (\mathbf{p}_n)_i = \\ & \sum_{i=1}^N (\mathbf{p}_n)_i \min_{k=1, \dots, M} \left[\sum_{j=1}^N L(ij | k) \pi(ij | k) \right] \end{aligned}$$

Dada la historia (fija) del proceso $(\mathbf{p}_0, u_0, u_1, \dots, u_{n-1})$

$$\min_{u_n} \mathbf{L}_{n+1} = \sum_{i=1}^N (\mathbf{p}_n)_i \left[\sum_{j=1}^N L(ij | k) \pi(ij | k) \right] \quad (8)$$

la identidad en (8) es alcanzada por la política estacionaria localmente óptima

$$u_n^*(k | i) = \delta_{k^*(i), i} \text{ para toda } n = 0, 1, \dots$$

donde $k^*(i)$ es un índice para el cual

$$\begin{aligned} \sum_{j=1}^N L(ij | k) \pi(ij | k^*(i)) &\leq \sum_{j=1}^N L(ij | k) \pi(ij | k) := \\ V(i | k) \quad \forall k &= 1, \dots, M \end{aligned} \quad (9)$$

Como resultado se puede establecer el siguiente Lema.

Lema: Sea $PDM = \{X, A, \Theta, \Pi, L\}$ un proceso de decisión de Markov. Dada una política (fija) localmente óptima, los L -valores para todo par estado-acción de (7) en el formato de matriz recursiva se convierte en

$$\mathbf{L}_{n+1} = \langle \mathbf{v}^*, \mathbf{p}_n \rangle \quad (10)$$

donde

$$\mathbf{v}^* := \left((\mathbf{v}^*)_1, \dots, (\mathbf{v}^*)_N \right) \quad (11)$$

$$(\mathbf{v}^*)_i := \sum_{j=1}^N L(ij | k) \pi(ij | k^*(i)) = \min_{k=1, \dots, M} V(i | k)$$

Observación: Bajo la política localmente óptima (9) la probabilidad del vector estado \mathbf{p}_n satisface la siguiente relación

$$\mathbf{p}_{n+1} = \left(\Pi^* \right)^T \mathbf{p}_n = \left(\left(\Pi^* \right)^T \right)^{n+1} \mathbf{p}_0 \quad (12)$$

donde

$$\left(\Pi^* \right)_{ij} := \sum_{k=1}^M \pi(ij | k) \delta_{k^*(i), i} = \pi(ij | k^*(i)) \quad (13)$$

IV. FUNCIÓN MONO-OBJETIVO

El objetivo de esta sección es asociar a cualquier función L_n , gobernada por (10), una función a Lyapunov que decrezca monotónicamente sobre las trayectorias de un sistema dado.

Se representa L_{n+1} como

$$\begin{aligned} \mathbf{L}_{n+1} &= \langle \mathbf{v}^*, \mathbf{p}_n \rangle = \langle \mathbf{v}^*, \mathbf{p}_{n-1} \rangle + \langle \mathbf{v}^*, \mathbf{p}_n - \mathbf{p}_{n-1} \rangle = \\ & \left[\left(1 + \frac{\langle [\Pi^* - I] \mathbf{v}^*, \mathbf{p}_{n-1} \rangle}{L_n} \right) \mathbf{L}_n \right] \end{aligned}$$

y si se denota

$$\begin{aligned} \Phi_n &= \langle [\Pi^* - I] \mathbf{v}^*, \mathbf{p}_{n-1} \rangle \mathbf{L}_n^{-1} = \\ & \frac{\langle [\Pi^* - I] \mathbf{v}^*, \mathbf{p}_{n-1} \rangle}{\langle \mathbf{v}^*, \mathbf{p}_{n-1} \rangle} \end{aligned} \quad (14)$$

así se obtiene

$$\mathbf{L}_{n+1} = (1 + \Phi_n) \mathbf{L}_n \quad (15)$$

Si se define $\tilde{\Phi}_n$ como

$$\tilde{\Phi}_n = \begin{cases} \Phi_n & \text{if } \Phi_n \geq 0 \\ 0 & \text{if } \Phi_n < 0 \end{cases}$$

se obtiene

$$\mathbf{L}_{n+1} = (1 + \Phi_n) \mathbf{L}_n \leq (1 + \tilde{\Phi}_n) \mathbf{L}_n \quad (16)$$

que nos lleva a la siguiente declaración.

Teorema: Sea $PDM = \{X, A, \Theta, \Pi, L\}$ un Proceso de Decisión de Markov y sea la matriz de forma recursiva representada por (15). Entonces, una función Lyapunov posible \mathcal{L}_n (que es monotónicamente no creciente) para el PDM tiene la forma

$$\mathcal{L}_n = \mathbf{L}_n \prod_{l=1}^{n-1} \left(\frac{1}{1 + \tilde{\Phi}_l} \right) = \left(\frac{1 + \Phi_{n-1}}{1 + \Phi_{n-1}} \right) \mathcal{L}_{n-1} \quad (17)$$

$$\mathcal{L}_0 = \mathbf{L}_0$$

Demostración: Sea considerada la recursión

$$a_{n+1} \leq (1 + \xi_n) a_n + \phi_n$$

con $\xi_n, a_n, \phi_n \geq 0$. Si se define

$$\bar{a}_n := a_n \prod_{l=1}^{n-1} \left(\frac{1}{1 + \xi_l} \right), \quad \bar{\phi}_n := \phi_n \prod_{l=1}^n \left(\frac{1}{1 + \xi_l} \right)$$

y

$$b_n = \bar{\alpha}_n - \sum_{l=1}^{n-1} \phi_l$$

se obtiene

$$b_{n+1} \leq b_n$$

En efecto,

$$\bar{\alpha}_{n+1} = a_n \prod_{l=1}^n \left(\frac{1}{1+\xi_n} \right) \leq$$

$$a_n \left[\prod_{l=1}^n \left(\frac{1}{1+\xi_n} \right) \right] (1 + \xi_n) + \phi_n \prod_{l=1}^n \left(\frac{1}{1+\xi_n} \right) = \bar{\alpha}_n + \tilde{\phi}_n$$

eso implica

$$b_{n+1} = \bar{\alpha}_{n+1} - \sum_{l=1}^n \phi_l \leq$$

$$\bar{\alpha}_n + \tilde{\phi}_n - \sum_{l=1}^n \phi_l = \bar{\alpha}_n + \sum_{l=1}^{n-1} \phi_l = b_n$$

y por lo tanto $b_{n+1} \leq b_n$. Entonces, se obtiene

$$\mathcal{L}_{n+1} = L_{n+1} \prod_{l=1}^n \left(\frac{1}{1+\Phi_l} \right) \leq (1 + \tilde{\Phi}_n) L_n \prod_{l=1}^n \left(\frac{1}{1+\Phi_l} \right) =$$

$$L_n \prod_{l=1}^{n-1} \left(\frac{1}{1+\Phi_l} \right) = \mathcal{L}_n$$

que demuestra el resultado.

V. RESULTADOS NUMÉRICOS

Ejemplo: Sea $N=3$ el número de estados y sea $M=2$ el número de acciones. La función de costos para $k=1,2$ está definida por

$$L(i, j) = \begin{bmatrix} 4.447 & 29.218 & 1.057 \\ 6.154 & 19.382 & 1.355 \\ 7.919 & 1.763 & 22.169 \end{bmatrix}$$

y sea la matriz de transición para $k=1,2$ definida de la siguiente manera

$$\pi(i, j, 1) = \begin{bmatrix} 50.4548 & 40.3912 & 10.1540 \\ 40.1900 & 9.7958 & 400.1062 \\ 90.2172 & 80.0370 & 70.7457 \end{bmatrix}$$

$$\pi(i, j, 2) = \begin{bmatrix} 0.9893 & 0.0098 & 10.6008 \\ 95.1833 & 990.5029 & 0.3138 \\ 90.1871 & 850.4189 & 0.0040 \end{bmatrix}$$

Suponemos que el perfil inicial es uniforme, esto es, $P(x_0 = j) = 0.3$ y sus estado $j=1,2,3$. De lo establecido anteriormente sigue que en el caso ergódico este perfil puede ser seleccionado arbitrariamente sin ninguna influencia en el punto de equilibrio final.

Para las políticas localmente óptimas u^{1*} , u^{2*} , u^{3*} (9) los siguientes resultados fueron obtenidos: a) en la Fig. 1 el comportamiento de la función de estado-valor muestra un comportamiento completamente no monotónico; b) en la Fig. 2 la función de Lyapunov correspondiente (17) es graficada mostrando definitivamente un comportamiento monotónico decreciente. El resultado del método muestra claramente que bajo la misma política localmente óptima la función original de costos converge no monotónicamente al valor 2.8725 y la función de Lyapunov correspondiente converge al valor 2.6837, los cuales, obviamente, son muy cercanos.

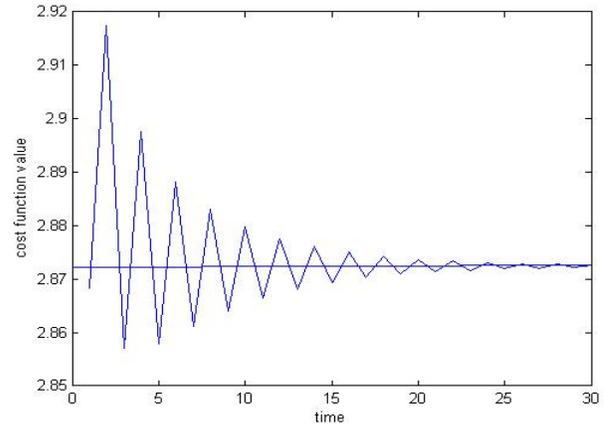


Figura 1. Comportamiento no monotónico de la función de costos dada.

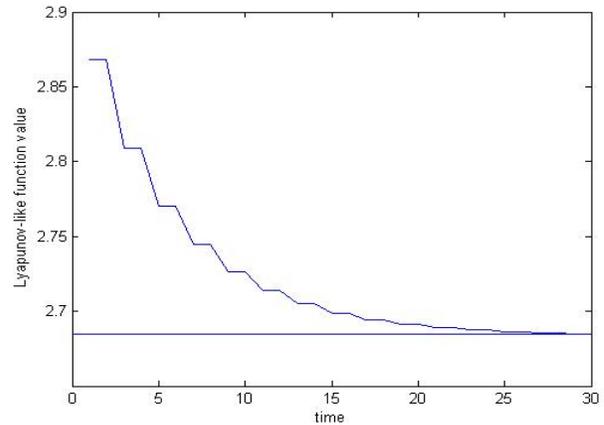


Figura 2. Comportamiento de la función mono objetivo.

VI. CONCLUSIONES.

Este artículo es una contribución teórica a las técnicas evolutivas para la optimización de un mono-objetivo que considera un Proceso de Decisión de Markov. El marco de optimización propuesto y los formalismos proveen una diferencia significativa en la conceptualización del dominio del problema. Con la introducción de una función de Lyapunov mono-objetivo como un concepto de solución para un proceso de decisión, se propone un modelo que es natural y que garantiza la existencia de un punto de equilibrio y, además, es computacionalmente tratable.

Como puede ser visto en el experimento numérico, el comportamiento de la función de pérdida L_n , generada por (10) bajo una política localmente óptima (9), no es siempre monotónicamente decreciente. Este hecho no permite establecer la convergencia de $\{L_n\}_{n \geq 0}$ basados directamente en el comportamiento de la secuencia costo-valor. Todo esto requiere otro concepto para analizar las propiedades asintóticas de esta secuencia de pérdida. Existen dos aproximaciones posibles para analizar la convergencia de las políticas localmente óptimas (9): a) el análisis de la ergodicidad de las cadenas homogéneas de Markov correspondientes con la matriz de transición Π^* dada por (12); b) la construcción de una función Lyapunov (energía) \mathcal{L}_n que es monotónicamente decreciente (no-creciente) en la trayectoria de la dinámica considerada (12).

Sin duda existen varios retos teóricos que requieren ser considerados en trabajos futuros bajo la teoría basada en Lyapunov para Procesos de Decisión de Markov. Como se mencionó anteriormente el uso de una función Lyapunov puede producir mejores resultados para encontrar la trayectoria más corta para un tipo de problemas dado. Por otro lado, la incorporación de más de un objetivo en la función de trayectoria se ha convertido en un área popular de investigación, como consecuencia una aproximación evolutiva nueva y variaciones de las técnicas existentes pueden ser desarrolladas considerando el método de Lyapunov propuesto.

RECONOCIMIENTO

El autor desea agradecer al Dr. Alexander Poznyak por ser un excelente mentor y por su invaluable aportación en el desarrollo de este artículo, su perseverancia, crítica, pensamiento analítico, disponibilidad constante, retroalimentación inmediata y estándares rigurosos.

REFERENCIAS

- [1] R. E. Bellman, "Dynamic Programming. Princeton", Univ. Press, Princeton N.J., 1957.
- [2] D. P. Bertsekas, "Dynamic Programming: Deterministic and Stochastic Models", Prentice-Hall, Englewood Cliffs, N.Y., 1987.
- [3] D. P. Bertsekas and J. N. Tsitsiklis, "Parallel and Distributed Computation: Numerical Methods", Prentice-Hall, Englewood Cliffs, N.Y., 1989.
- [4] J. B. Clempner and A. S. Poznyak, "Convergence Properties and Computational Complexity Analysis for Lyapunov Games", International Journal of Applied Mathematics and Computer Science, vol. 21, no. 2, 349-361, 2011.
- [5] J. B. Clempner and A. S. Poznyak, "Analysis of Best-Reply Strategies in Repeated Finite Markov Chains Games", 52nd IEEE Conference on Decision and Control, Florence, Italy, IEEE, 2013.
- [6] C. Coello, "A Comprehensive Survey of Evolutionary-Based Multiobjective Optimization Techniques", Knowledge and Information Systems, vol. 1, no. 3, pp. 269-308, 1999.
- [7] C. Derman, "Finite State Markovian Decision Processes", Academic Press, N.Y., 1970.
- [8] A. E. Eiben, E.H.L. Aarts and K. M. van Hee, "Global Convergence of Genetic Algorithms: A Markov Chain Analysis", in H.P. Scheffel and R. Männer, editors, Parallel Problem Solving From Nature, Berlin Heidelberg, pp. 4-12, 1991.
- [9] O. Hernández-Lerma and J.B. Lasserre, "Discrete-Time Markov Control Process: Basic Optimality Criteria"--- Berlin, Germany : Springer, 1996.
- [10] O. Hernández-Lerma, G. Carrasco and R. Pérez-Hernández, "Markov Control Processes with the Expected Total Cost Criterion: Optimality, Stability and Transient Model", Acta Applicandae Mathematicae, vol. 59, no. 3, pp. 229-269, 1999.

- [11] R. A. Howard, "Dynamic Programming and Markov Processes", MIT Press, Cambridge, 1960.
- [12] R. E. Kalman and J. E. Bertram, "Control System Analysis and Design Via the Second Method of Lyapunov", Journal of Basic Engineering, vol. 82, D, pp. 371-393, 1960.
- [13] H. Kushner, "Introduction to Stochastic Control", Holt, Rinehart, and Winston, N.Y., 1971.
- [14] A. S. Poznyak and M. D. Vose, "Modeling Genetics Algorithms with Markov Chains", in Annals of Mathematics and Artificial Intelligence, vol. 5, pp. 79-88, 1992.
- [15] M. Osborne and A. Rubinstein, "A Course in Game Theory", The MIT Press, 1994.
- [16] V. Pareto, "Cours D'Economie Politique Volume I and II", F. Rouge, Lausanne, 1896.
- [17] A. S. Poznyak, K. Najim and E. Gomez-Ramirez, "Self-learning control of finite Markov chains", Marcel Dekker, Inc., New York, 2000.
- [18] A. S. Poznyak, E. N. Sanchez, and W. Yu, "Differential Neural Networks for Robust Nonlinear Control: Identification, State Estimation and Trajectory Tracking", World Scientific, New Jersey, London, 2001.
- [19] G. Rudolph, "Convergence Properties of Canonical Genetic Algorithms", IEEE Transactions on Neural Networks, vol. 5, no. 1, pp. 96-101, 1994.
- [20] G. Rudolph, "Convergence Properties of Evolutionary Algorithms", Verlag Dr. Kovac, 1997.
- [21] B. Toth and V. Kreinovich, "Verified Methods for Computing Pareto Sets: General Algorithmic Analysis", International Journal of Applied Mathematics and Computer Science, vol. 19, no. 3, pp. 369-380, 2009.



Dr. Julio Clempner has more than ten years experience in the field of management consulting. Currently, he specializes in the application of high technology. He also conducted projects related to IT Government, Project Management, Analysis and Design of Software (RUP, UML), Development of Software (add-hoc and products), Information Technology Strategic Planning, Balanced Score Card, Evaluation of Software and Business Process Reengineering (SIX-SIGMA) for infusing advanced computing technologies into diverse lines of businesses. Dr Julio Clempner is now involved in providing e-Business solutions including: e-Business strategy, management of technology, operations strategy, project management, continuous quality improvement, and managing the total quality transformation. Dr. Clempner research interests are focused on justifying and introducing the Lyapunov equilibrium point in shortest-path decision processes and shortest-path game theory. This interest has led to several streams of research. One stream is on the use of Markov decision processes for formalizing the previous ideas, changing the Bellman's equation by a Lyapunov-like function which is a trajectory-tracking function, and also it is an optimal cost-to-target function for tracking the net. A second stream is on the use of Petri nets as a language for modeling decision process and game theory introducing colors, hierarchy, etc. Petri nets are used for process representation taking advantage of the formal semantic and the graphical display. The final stream examines the possibility to meet modal logic, decision processes and game theory. Dr. Julio Clempner is a member of the Mexican National System of Researchers (SNI).