APROXIMACIÓN NUMÉRICA DE UN PROBLEMA DE CONSUMO-INVERSIÓN

RUY ALBERTO LÓPEZ RÍOS [a], HUGO ADÁN CRUZ SUÁREZ [a], FERNANDO VELASCO LUNA [a],
VÍCTOR HUGO VÁZQUEZ GUEVARA [a]

[a]Facultad de Ciencias Físico Matemáticas
Benemérita Universidad Autónoma de Puebla
Av. San Claudio Esq. 18 Sur, San Manuel, C.P. 72570. Puebla, Pue.
e-mail: ruyalberto@gmail.com, hcs@fcfm.buap.mx

El presente trabajo trata sobre la resolución de una aplicación específica de un problema de control óptimo: el Problema de Consumo-Inversión. Se presenta la teoría esencial para el Problema de Control Óptimo, se deducen las Ecuaciones de Programación Dinámica, y se expresan explícitamente las soluciones del problema para una función de utilidad dada. Los resultados son comparados mediante simulación, con el propósito de que, al considerar funciones de utilidad complejas, el problema podría no tener una solución cerrada.

Keywords: Control óptimo estocástico, Bellman, Markov, Procesos de decisión, Programación dinámica

1. Introducción

El Control Óptimo es una herramienta matemática utilizada para resolver problemas de optimización cuya evolución en el tiempo es susceptible a ser influenciado por variables aleatorias exógenas.

En este trabajo se aborda el control óptimo desde el enfoque de la Programación Dinámica.

La definición de un problema de control óptimo, para sistemas estocásticos o deterministas, requieren de tres componentes:

- un modelo de control o decisión,
- un conjunto de políticas de control admisibles, y
- un criterio de rendimiento, o función objetivo.

En la siguiente Sección se presenta la teoría relacionada y se plantea el problema.

2. El Problema de Control Óptimo

Un *Proceso de Decisión de Markov* es un proceso de Markov controlado que transita a través de un espacio de estados X consistente de estaciones discretas. Cada etapa consiste de un estado x para el cual una acción a de un espacio de acciones A es tomada, incurriéndose un costo $R(x,a)$. El proceso se traslada a un nuevo estado de acuerdo a una medida de probabilidad $\mathbb{Q}$, iniciándose una nueva etapa. El proceso se repite mientras un criterio de rendimiento J se va optimizando.

Definición 1. (*Modelo de Control de Markov*) Un Modelo de Control de Markov es una quíntupla

$$(X, A, \{A(x)|x \in X\}, \mathbb{Q}, R) \tag{1}$$

la cual consiste de
a) un espacio de Borel X, llamado *espacio de estados*;
b) un espacio de Borel A, llamado *espacio de control* (o acción);
c) una familia $\{A(x)|x \in X\}$ de subconjuntos medibles $A(x)$ de A, donde $A(x)$ denota el *conjunto de controles* (o acciones) *admisibles* cuando el sistema está en el estado $x \in X$, y con la propiedad de que el conjunto

$$\mathbb{K} = \{(x,a)|x \in X, a \in A(x)\}$$

de parejas estado-acción admisibles es un subconjunto medible de $X \times A$;
d) un kérnel estocástico $\mathbb{Q}$ sobre X dado $\mathbb{K}$ llamada *ley de transición*;
e) una función medible $R : \mathbb{K} \to \mathbb{R}$ llamada *función de recompensa (o costo)*.

Denotaremos a $\mathbb{F}$ el conjunto de todas las funciones medibles $f : X \to A$ satisfaciendo que $f(x) \in A(x)$. A los elementos de $\mathbb{F}$ los llamaremos *selectores*.

2.1. Planteamiento del problema. Consideramos un criterio de rendimiento $J(\pi, x)$ dado de la forma:

$$J(\pi, x) := \mathbb{E}_x^\pi \left\{ \sum_{t=0}^{N-1} R(x_t, a_t) + R_N(x_N) \right\}.$$

Denotamos por J^* la función objetivo, es decir,

$$J^*(x) := \sup_{\Pi} J(\pi, x), \qquad x \in X.$$

El *problema de control óptimo* consiste en encontrar una política $\pi^* \in \Pi$ tal que

$$J(\pi^*, x) = J^*(x), \qquad \forall x \in X. \tag{2}$$

Una opción para resolver problemas de control óptimo es hacer uso del *Teorema de la Programación Dinámica*, el cual provee un algoritmo para encontrar la función objetivo J^* y una política óptima π^*.

Teorema 1. *Sean $J_N, ..., J_1, J_0$ funciones sobre X definidas por*

$$J_N(x) := R_N(x), \tag{3}$$

y para $t = N-1, N-2, ..., 0,$

$$J_t(x) := \sup_{a \in A(x)} \left\{ R(x, a) + \int_X J_{t+1}(y) \mathbb{Q}\{dy|x, a\} \right\}. \tag{4}$$

Suponga que estas funciones son medibles y que, para cada $t = 0, ..., N-1$, existe un selector $f_t \in \mathbb{F}$ tal que $f_t(x) \in A(x)$ alcance el máximo en (4) $\forall x \in X$; entonces la política $\pi^* = \{f_0, f_1, ..., f_{N-1}\}$ es óptima, y la función objetivo J^* es

$$J^*(x) = J_0(x) = J(\pi^*, x), \qquad \forall x \in X.$$

Frecuentemente, se han de considerar criterios de rendimientos de la forma:

$$V(\pi, x) := \mathbb{E}_x^\pi \left\{ \sum_{t=0}^{N-1} \alpha^t u(x_t, a_t) \right\}, \quad 0 < \alpha < 1.$$

siguiéndose los mismos resultados del teorema anterior. Solo que (3) y (4) se convierten en:

$$V_N(x) := 0, \tag{5}$$

$$V_t(x) := \sup_{a \in A(x)} \left\{ R(x, a) + \alpha \int_X V_{t+1}(y) \mathbb{Q}\{dy|x, a\} \right\}. \tag{6}$$

Para otros criterios de rendimiento y variaciones de las Ecuaciones de Programación Dinámica pueden consultarse [2], [3], [10], [11].

3. Problema de Consumo-Inversión

Presentamos un elemental pero muy importante problema, el de consumo-inversión, que consiste en la asignación o distribución que debe hacer un inversor de una riqueza actual x_t en dos partes: 1) una para inversión a_t, y 2) otra para un consumo $x_t - a_t$ en cada período $t = 0, 1, ..., N-1$. Suponemos que el préstamo no está permitido, así que la restricción de inversión, o el conjunto de acciones admisibles es $A(x) = [0, x]$. Consideramos $X = [0, \infty)$.

La relación entre la inversión y el capital acumulado está dado por

$$x_{t+1} = a_t \xi_t, \qquad t = 0, 1, ...,$$

donde las variables aleatorias $\{\xi_t\}$ son iid, independientes de x_0, y además con $\mathbb{E}\{\xi_t\} > 1$, esto último para asegurar que la reinversión pueda ser rentable.

También, una recompensa $R(x, a) := u(x - a)$ se dice que es una utilidad de consumo, para alguna función de utilidad u.

Finalmente, deseamos maximizar la utilidad esperada total descontada de consumo:

$$V(\pi, x) := \mathbb{E}_x^\pi \left\{ \sum_{t=0}^{N-1} \alpha^t u(x_t, a_t) \right\}, \quad 0 < \alpha < 1.$$

Para nuestro problema, las Ecuaciones de Programación Dinámica, para todo $x \in X$ y $t = N-1, N-2, ..., 0$ resultan ser:

$$\begin{aligned} V_N(x) &= 0 \\ V_t(x) &= \max_{a \in A(x)} \{u(x-a) + \alpha \mathbb{E}\{V_{t+1}(a\xi_t)\}\}. \end{aligned}$$

La solución depende de la forma particular de la función u. Cuando se considera

$$u(x, a) := \frac{b}{\gamma}(x-a)^\gamma, \qquad b > 0, \ \ 0 < \gamma < 1.$$

(utilidad marginal isoelástica) afortunadamente se puede encontrar una solución cerrada de las ecuaciones de programación dinámica y de los controles óptimos que son las que se presentan a continuación.

Definiendo

$$\delta := (\alpha \mathbb{E}\{\xi_0^\gamma\})^{\frac{1}{\gamma - 1}}$$

y una función D de manera recursiva que será útil para los cálculos:

$$\begin{aligned} D_{N-1} &:= 1 \\ D_t &:= \frac{\delta^{\gamma-1} D_{t+1}}{\left(1 + \delta (D_{t+1})^{\frac{1}{\gamma-1}}\right)^{\gamma-1}}, \qquad t < N-1. \end{aligned}$$

Se pueden obtener finalmente las soluciones cerradas:

$$\begin{aligned} V_t(x) &= \frac{b}{\gamma} D_t x^{\gamma}. \\ f_t(x) &= \frac{x}{1+\delta(D_{t+1})^{\frac{1}{\gamma-1}}}. \\ V^*(x) &= V_0(x) = \left(\frac{b}{\gamma}\right)\left(\frac{\delta^{N-1}(1-\delta)}{1-\delta^N}\right)^{\gamma-1} x^{\gamma}. \end{aligned}$$

4. Ejemplo

Para fines comparativos, se desarrolla un ejemplo con parámetros particulares que se resolverá analíticamente. Posteriormente se resuelve mediante simulación, sirviendo para casos complejos en los que no es posible encontrar soluciones cerradas.

En ambos casos se considera una condición inicial $x_0 = 10$, y una función de utilidad

$$u(x,a) := \frac{b}{\gamma}(x-a)^{\gamma},$$

con $b = 1$ y $\gamma = 0.1$.

El factor de descuento tomado es $\alpha = 0.5$, y las variables aleatorias $\{\xi_t\}$ se toman con distribución $LogNormal(0.5, 0.2)$, calculándose $\delta = 2.04293 > 1$.

Consideremos $N = 11$ etapas, y se obtienen los valores D_i mostrados en la Tabla 1.

Table 1. Valores de D_i.

D_{10}	1.00000
D_9	1.43131
D_8	1.63696
D_7	1.73656
D_6	1.78508
D_5	1.80877
D_4	1.82036
D_3	1.82603
D_2	1.82880
D_1	1.83016
D_0	1.83083

Se encuentra el rendimiento óptimo analítico, dato que almacenamos para comparación posterior.

$$V^*(x) = V_0(x_0) = \frac{b}{\gamma} D_0(x_0)^{\gamma} = \mathbf{23.0487}$$

Pasando a la solución mediante simulación, se consideran los mismos datos y criterios anteriores, y un número de 500 repeticiones simuladas mediante computadora. En la Tabla 2 se puede ver, al menos, una realización de la dinámica que sigue el sistema en cada etapa.

Table 2. Evolución del sistema.

i	x_i	a_i	c_i
0	10.00000	4.89295	5.10705
1	7.50625	3.67123	3.83502
2	7.45630	3.64365	3.81265
3	5.85973	2.85840	3.00133
4	5.08803	2.47295	2.61508
5	3.63638	1.75409	1.88229
6	2.59052	1.22980	1.36072
7	1.97405	0.904905	1.06915
8	1.44318	0.608537	0.934646
9	1.02445	0.336666	0.687787
10	0.404347	0.00000	0.404347

En la Tabla 3, se observan las utilidades netas y las acumuladas descontadas para el sistema específico descrito en la Tabla 2.

Table 3. Utilidades netas y descontadas acumuladas.

i	$u(x_i, a_i)$	$\sum \alpha^k u(x_k - a_k)$
0	11.7711	11.7711
1	11.4387	17.4905
2	11.4320	20.3485
3	11.1617	21.7437
4	11.0090	22.4317
5	10.6529	22.7646
6	10.3128	22.9258
7	10.0671	23.0044
8	9.82088	23.0428
9	9.63264	23.0616
10	9.13430	*23.0705*

La réplica de 500 simulaciones de cálculos realizados tanto en la Tabla 2 como en la Tabla 3 arrojan un promedio de las sumas del tipo $\sum_{k=0}^{N-1} \alpha^k u(x_k - a_k)$:

$$\bar{V} = \mathbf{23.0475}$$

que es comparable con el resultado obtenido analíticamente.

5. Conclusiones

Se ha abordado, al menos, la resolución de un problema de control óptimo para una función de utilidad isoelástica y un criterio de rendimiento específicos, mediante dos formas, encontrando la solución cerrada, y mediante simulación para fines comparativos.

Estos algoritmos serán implementados a otras utilidades. Desafortunadamente, existen problemas cuya solución de políticas no pueden presentarse en forma cerrada, por lo que se requiere de un procedimiento numérico que aproxime estas elecciones.

Por ejemplo, cuando se presentan funciones de costo exponenciales, que requerirá de simulación estocástica para la aproximación de soluciones.

Las discretizaciones de espacios abren paso a los *métodos multigrid*, que consisten en la variación de un parámetro que regule la discretización, y que es útil, por ejemplo, para funciones de utilidad más complejas. Métodos como este inician el trabajo futuro a seguir. (Ver [5], [6], [7]).

Para otros métodos de aproximación y tópicos adicionales sobre métodos multigrid (Ver [10], [14], [16], [18], [20]).

Referencias

[1] Bellman, R., *Dynamic programming*, Princeton University Press, USA, 1972.

[2] Bertsekas, D., Shreve, S., *Stochastic optimal control: The discrete-time case*, Athena Scientific, USA, 1996.

[3] Bertsekas, D., *Dynamic programming and optimal control*, Vol. I & II, Athena Scientific, USA, 1995.

[4] Bertsekas, D., *Convergence of discretization procedures in dynamic programming*, IEEE Transactions On Automatic Control, pp. 415-419, 1975.

[5] Chow, C., *Multigrid algorithms and complexity results for discrete-time stochastic control and related fixed-point problems*, Tesis de doctorado, MIT, USA, 1989.

[6] Chow, C., Tsisiklis, J., *An optimal one-way multigrid algorithm for discrete-time stochastic control*, IEEE Transactions on Automatic Control, Vol. 36, No. 8, 1991.

[7] Denardo, E., *Contraction mappings in the theory underlying dynamic programming*, SIAM Review, Vol. 9, No. 2, pp 165-177, 1967.

[8] Fleming, W., Rishel, R., *Deterministic and stochastic optimal control*, Springer-Verlag, USA, 1975.

[9] Fox, B., *Discretizing dynamic programs*, Journal of Optimization Theory and Applications, Vol. 11, No. 3, 1973.

[10] Han, D., Wan, J., *Multi-Grid methods for second order Hamilton-Jacobi-Bellman equations and Hamilton-Jacobi-Bellman-Isaacs equations*, Journal On Scientific Computing, Vol. 35, No. 5, pp. S323-S344, SIAM, 2013.

[11] Hernández-Lerma, O., *Adaptive Markov control processes*, Applied Mathematical Sciences, Vol. 79, Springer-Verlag, USA, 1989.

[12] Hernández-Lerma, O., Laserre, J. B., *Discrete-time Markov control processes. Basic optimality criteria*, Applications of Mathematics, Stochastic Modelling and Applied Probability, Vol. 30, Springer, USA, 1996.

[13] Hernández-Lerma, O., *Lectures on continuous-time Markov control processes*, Aportaciones Matemáticas, Sociedad Matemática Mexicana, México 1994.

[14] Hoppe, R., *Multi-Grid methods for Hamilton-Jacobi-Bellman equations*, Numerische Mathematik, Vol. 49, pp. 239-254, Springer-Verlag, 1986.

[15] Lions, P., Mercier, B., *Approximation numérique des équations Hamilton-Jacobi-Bellman*, RAIRO Analyse Numérique, Vol. 14, No. 4, pp. 369-393, France, 1980.

[16] Munos, R., Moore, A., *Variable resolution discretization in optimal control*, Machine Learning, 49, 291–323, Kluwer Academic Publishers. Netherland. 2002.

[17] Puterman, M., *Markov decision processes. Discrete stochastic dynamic programming*, John Wiley & Sons, USA, 2005.

[18] Schal, M., *On dynamic programming: compactness of the space of policies*, Stochastic Processes and their Applications, Vol. 3, pp. 345-364, North Holland Publishing Company, 1975.

[19] Whitt, W., *Approximations of dynamic programs I*, Mathematics of Operations Research, Vol. 3, No. 3, USA, 1978.

[20] Zou, Z., *A new scheme for discrete HJB equations*, Applied Mathematics, Vol. 5, pp. 2643-2649, Scientific Research, 2014.