# Un sistema de inventario modelado como un proceso de decisión de Markov

Eymard Hernández, Maira Madriz , Guadalupe Gaytán y Gabriel Zacarías

E. Hernández, M. Madriz , G. Gaytán y G. Zacarías
Tecnolo´gico de Estudios Superiores del Oriente del Estado de México
Departamento de Matemáticas, UAM- Iztapalapa

M. Ramos., V.Aguilera., (eds.).Ciencias Naturales y Exactas, Handbook -©ECORFAN- Valle de Santiago, Guanajuato, 2014.

**Abstract**

This paper is related to inventories systems with stochastic demand in discrete time and finite horizon. One way to control the system is through the production sequence, with the main goal of reduce the costs. The Markov Decision processes theory allows using the dynamic programming technique to control inventory system.

## 4 Introducción

El presente trabajo está vinculado con sistemas de inventarios estocásticos, a tiempo discreto y con horizonte finito. Un sistema de inventario tiene la siguiente interpretación: en un tiempo determinado se observa el nivel de producción o stock y dependiendo de este se produce (u ordena) cierta cantidad para suplir la demanda en el periodo la cual se supone es estocástica. Esto producirá un nuevo nivel de stock en el siguiente periodo. Todo esto generará costos durante el horizonte de producción ya sean de almacenamiento, producción y penalidades por demandas no suplidas. Una forma de controlar al sistema es mediante la cantidad producida u ordenada para reducir sus costos.

Por otro lado, la teoría de procesos de decisión de Markov (PDMs) permite modelar perfectamente a los sistemas de inventarios. Los PDMs son aquellos procesos que son observados de forma periódica, bajo incertidumbre en sus movimientos y tienen una gran variedad de aplicaciones. En particular, este trabajo se enfoca en el estudio de un sistema de inventario (véase [4] y [5]). Un proceso de decisión de Markov está constituido mediante un modelo conocido como modelo de control de Markov (MCM), cuyas componentes permiten caracterizar su desarrollo en el transcurso del tiempo. El tránsito del sistema puede ser influenciado por medio de la aplicación de acciones en cada periodo de tiempo. A la sucesión de acciones se le conoce como política y, una forma de evaluar su calidad es mediante, un criterio de rendimiento o función objetivo. El problema de control óptimo (PCO) consiste en determinar una política que optimice al criterio de rendimiento.

Una forma de resolver el PCO es mediante la técnica de programación dinámica (PD) iniciada por Richard Bellman a mediados de los años 50's (véase [1]). El método de PD consiste en llevar al PCO a otro equivalente, el cual está dado por medio de una ecuación funcional. Dicha ecuación permite determinar a la política óptima. El objetivo de este trabajo consiste en demostrar que es posible hacer uso de PD para un sistema de inventario estocástico. La metodología es la teoría de PDMs.

### 4.1 Procesos de decisi´on de Markov

Un Modelo de Control de Markov (MCM), estacionario, a tiempo discreto, consiste de una qu´ıntupla:

$(X, A, \{A(x) | x \in X\}, Q, r)$, donde X y A son espacios de Borel, llamados espacio de estados y espacio de acciones (o controles), respectivamente. $\{A(x) | x \in X\}$ es una familia de subconjuntos medibles y no vac´ıos $A(x)$ de A, donde $A(x)$ denota al conjunto de acciones admisibles al correspondiente estado $x \in X$.

El conjunto K de parejas de estado acción admisibles está definido por
$K := \{(x, a) | x \in X, a \in A(x)\}$, $Q(\cdot | \cdot)$ es la ley de transición definida en X dado K, y $c : K \to R$ es una funcio´n medible llamada la funcio´n de costo.

La dinámica que describe a este sistema estocástico actuá de la siguiente manera: si el sistema al tiempo t se encuentra en el estado $x_t = x \in X$, y la acción $a_t = a \in A(x)$ es aplicada, entonces ocurren dos cosas:

- se recibe una recompensa $r(x, a)$; y
- el sistema transita a un nuevo estado $x_{t+1}$ mediante la ley de transicio´n

Una vez hecha esta transición a un nuevo estado, se elige una nueva acción y la dinámica anteriormente descrita se repite.

Una política es una sucesión $\pi = \{\pi_t\}$ de kérneles estocásticos, donde cada $\pi_t$ está definido sobre A dada su correspondiente historia. El conjunto de todas las políticas es denotado por $\Pi$.

Una clase particular de pol´ıticas son las que están caracterizadas por sucesiones del siguiente conjunto $F := \{ f : X \to A | \text{ f es medible y para cada } x \in X, f(x) \in A(x)\}$, cuyos elementos de F se les conoce como selectores. Una política $\pi \in \Pi$ descrita por elementos de F se le conoce como pol´ıtica determinista.

Sea ($\Omega$, «) un espacio medible que consiste del espacio muestral canónico $\Omega := (X \times A)^\infty$ y « su correspondiente σ-´algebra producto.

Sean $\pi \in \Pi$ una política arbitraria y $x_0 = x \in X$. Entonces por el teorema de Ionescu-Tulcea (v´ease [3]), existe una u´nica medida de probabilidad $P^\pi_x$ sobre ($\Omega$, «). El proceso estocástico (($\Omega$, «, $P^\pi_x$ ), $\{x_t\}$) es llamado proceso de decisión de Markov (PDM). La esperanza con respecto a $P^\pi$ es denotada por $E^\pi$ .

Una manera de evaluar la calidad de una política es por medio de un criterio de rendimiento. En este trabajo se considera el costo total descontado, definido para cada $\pi \in \Pi$ y cada $x \in X$ como con $\alpha \in (0, 1)$. A $\alpha$ se le conoce como factor de descuento y a N como el horizonte del problema, el cual se supone es finito.

Definición Una política $\pi^* \in \Pi$ es óptima, si para cada $x \in X$, $v(\pi^*, x) = \inf_{\pi\in\Pi} v(\pi, x)$. La función definida para $x \in X$ $V(x) := \inf_{\pi\in\Pi} v(\pi, x)$, es llamada función de valor óptimo.

**El problema de control óptimo consiste en determinar una política óptima.**

**Programación dinámica**

En la literatura existente se encuentra una herramienta esencial para resolver el problema de control óptimo, conocida como programación dinámica (PD) (véase [1], [2] y [3]).
Considere un MCM fijo. La siguiente condición sobre el MCM permite aplicar la técnica de PD, (ver Teorema 2.3, [3]). La Condición 2.2 utiliza los conceptos de infcompacidad y emplea las propiedadades débil y fuerte de un kérnel estocástico, estos conceptos corresponden a las definiciones 6.1 y 6.2 del ap´endice, respectivamente.

**Condición**

- La función de costo c es continua, acotada inferiormente e infcompacta sobre K.

- La ley de transición Q es débilmente continua o fuertemente continua.

## 4.2 Sistema de inventario

Considere un modelo de inventario de alguna produccio´n en el cual la variable xt es el nivel de stock al principio del período t, donde (t = 0, 1, 2 . . . , N ). La variable de acción at es la cantidad pedida (u ordenada) e inmediatamente proporcionada al principio del período, mientras que, la variable de perturbación (o ruido) ξt es la demanda durante ese período. En este caso, el nivel de stock en el siguiente período de tiempo está dada por la siguiente ecuación en diferencias

$x_{t+1} = x_t + a_t - \xi_t,$ (4)
para $t = 0, 1, 2, \ldots, N$, con $x_0 = x$.

Al finalizar el período, se debe pagar un costo determinado por

$c(x_t, a_t, \xi_t) = ba_t + h\,m\acute{a}x\{0, x_{t+1}\} + p\,m\acute{a}x\{0, -x_{t+1}\},$ (3) para $t = 0, 1, 2, \ldots, N$, donde:

* b es el costo de producción por unidad.
* h es el costo por unidad de exceso del inventario (de almacenaje).
* p es el costo por unidad de demanda sin entregar (no suplida). Además, b, h y p son constantes no negativas, con $p \geq b$.

El objetivo principal es minimizar el costo esperado descontado de operación,
es decir, minimizar el valor esperado de N −1. $\alpha^t c(x_t, a_t, \xi_t)$.t=0

## 4.3 Identificación del sistema de inventario como un PDM

El planteamiento del modelo de inventario propone considerar al espacio de estados X = R, al espacio de acciones y al conjunto de acciones admisible como

$A = A(x) = [0, \infty)$, y la ley de transicio´n Q está inducida por la ecuación.

Se supone que $\{\xi_t\}$ son variables aleatorias independientes e idénticamente distribuidas con valores en $S = [0, \infty)$, con función de distribucio´n μ, la cual tiene densidad continua Δ, y además ξ tienen esperanza finita.

Obsérvese que el costo determinado en la ecuación (3) depende, además del estado y acción, de la demanda aleatoria, es decir,

$c(x, a, \xi) = ba + h\,máx\{0, x + a - \xi\} + p\,máx\{0, \xi - x - a\}$ (4.1)

Sin embargo, si se considera para cada $(x, a) \in K$

cr(x, a) = E( c(xt, at, ξt)| xt = x, at = a)¸= c(x, a, s)μ(ds)¸=c(x, a, s)Δ(s)ds,

donde E denota la esperanza de ξ. Entonces cr resulta ser la función de costo para el MCM.

Finalmente, el objetivo es minimizar al criterio de rendimiento con $\pi \in \Pi$ y $x \in X$.
v(π, x) := Eπ .N −1 .
. αtc(xt, at, ξt) ,
t=o

Lema 4.1. Bajo las suposiciones antes mencionadas sobre el sistema de inventario se satisface la Condicio´n 2.2.

Demostración. Es obvio que la funcio´n de costo cr está acotada inferiormente.

Para mostrar que cr es continua en K, consideramos a

$$W(x, a) := hE[\mathrm{ma\acute{}x}\{0, x + a - \xi\}] + pE[\mathrm{ma\acute{}x}\{0, \xi - x - a\}], \quad (4.2)$$

es una función continua en K.

Sean $\{xn\}$ y $\{an\}$ sucesiones en X y A, respectivamente, tales que, $xn \to xr$
y $an \to ar$. Definiendo a gn y g como

$$gn(s) := |xn + an - s|\,\Delta(s), \quad g(s) := |xr + ar - s|\,\Delta(s). \quad (4.3)$$

Se tiene que $gn(s) \to g(s)$ para cada $s \in S$. Asimismo, sucede que

$$gn(s) \le (|xn| + |an| + s)\Delta(s)$$

$$\le (M + s)\Delta(s) =: f(s), \quad (4.4)$$

donde M es la cota $|xn| + |an|$. De manera que,

$$(M + s)\Delta(s)ds = M + E(\xi) < +\infty. \quad (4.5)$$

En consecuencia, f es integrable y por el Teorema de la Convergencia Dominada, se concluye que w es continua.

Por otra parte, se afirma que cr es inf-compacta sobre K. En efecto, sean
$x \in X$, $\lambda \in R$ y considere al conjunto

$$A\lambda(x) := \{ a \in A(x) |\, cr(x, a) \le \lambda\} \quad (4.6)$$

N´otese que $A\lambda(x)$ es acotado, de lo contrario, existe una sucesio´n $\{ar_n\}$ sucesión

enAλr∞, implicando que $\lim_{n\to\infty} cr(x, ar_n) = \infty$,

y dada la continuidad de cr se obtiene que $\infty \le \lambda$, lo que es una contradiccio´n. Por otro lado, si $\{ar\}$ es una sucesio´n en $A\lambda(x)$, tal que ar
$\to a \in A$, y

### 4.4 Conclusiones

En general, la t´écnica de programación dinámica, sirve como herramienta para resolver el problema de control óptimo. Para el modelo de inventario estocástico presentado, esta técnica permite determinar no sólo a la política óptima, sino también al valor óptimo de una forma recursiva. Como trabajo a futuro se pretende utilizar simulación numérica para su aproximación.

### 4.5 Agradecimientos

Tecnológico de Estudios Superiores del Oriente del Estado de México. Paraje San Isidro s/n Barrio de Tecamachalco, La Paz C.P.56400, Tel. ( 55 ) 59863497 ext. 101.

### 4.6 Referencias

Bellman R., (1957). Dynamic Programming, Princeton University Press, Princeton.

Bertsekas D. P., (1987). Dynamic Programming: Deterministic and Stochas- tic Models, Prentice-Hall, Belmont.

Hern´andez-Lerma O. and Lasserre J. B., (1996). Discrete-Time Markov Con- trol Processes: Basic Optimality Criteria, Springer-Verlag, New York.

Love S.F., (1979). Inventory Control Theory, Mc. Graw-Hill.

Porteus E. L., (1990). Stochastic inventory theory, in Handbooks in OR & MS, Vol. 2, Eds. D. P. Heyman, M. J. Sobel, Elsevier Science Publishing B. V., North Holland.

Sundaram R. K., (1997). A First Course in Optimization Theory, Cambridge University Press.