APLICACION DEL MODELO

de Política Iterativa para la Solución de Procesos Decisionarios en Serie de Naturaleza Markoviana a la toma de Decisiones en Sistemas Nacionales *

Por: Max Alberto Soto

Introducción

El propósito de este ensayo es el de investigar procesos decisionarios que han sido desarrollados en otros campos —especialmente el de Investigaciones Operativas— y aplicarlos a la toma de decisiones de sistemas nacionales con el propósito de estimular el estudio de nuevos enfoques para la solución de problemas viejos. Específicamente, se brega aquí con

* Este trabajo está basado en los desarrollos presentados por R. Howard en su libro *Dynamic Programming and Stochastic Processes*, MIT Press, 1960, y no reclama originalidad excepto posiblemente en el área de aplicación.

procesos de naturaleza markoviana que requieren soluciones en "secuencia", es decir, una serie de decisiones, cada una sucediendo a otra anterior. En términos generales, lo que se trata de obtener es una serie de decisiones a tomar bajo diferentes circunstancias o "estados", que mejoraría las probabilidades del sistema en cuestión de lograr sus objetivos. Primero se revisarán conceptos fundamentales, como el de procesos Markov y el de decisiones secuenciales, y luego se explicará el método de "política iterativa" para los casos de una y de varias "cadenas recurrentes". También se discutirán ejemplos en el contexto de sistemas nacionales.

Decisiones en secuencia para procesos estocásticos

Un proceso es estocástico si se desarrolla probabilísticamente a través del tiempo. Es decir, se define como una colección de variables aleatorias (X_t) en las que el índice t puede tomar valores contenidos en un conjunto T, que frecuentemente se considera como el conjunto de los números enteros positivos. (X_t) representa una característica de interés de un sistema dado en el tiempo t.[1]

En problemas de programación, se trata de determinar los valores de las variables controlables, o de decisión, que obtuviesen una función dada (el objetivo). Dichos valores dependerán de los parámetros del problema, tales como los recursos disponibles, los coeficientes tecnológicos, etc. Generalmente al variar los parámetros cambian los valores de las variables "de control" que optimizan la solución. El problema con el que tropiezan los métodos utilizados regularmente, los cuales se basan en el algoritmo "simplex", es que no permiten la atención de soluciones en los que se exprese explícitamente un conjunto de valores de las variables de decisión en función de los parámetros.[2]

Los problemas de decisión "secuencial" envuelven la toma de por lo menos dos decisiones en diferentes tiempos y en los cuales la(s) última(s) decisión(es) depende(n) no sólo de la decisión anterior sino que también de algunos parámetros estocásticos cuyos valores serán conocidos antes de tomar decisiones posteriores. De esto se desprende que la característica esencial de los problemas de decisiones en secuencia es que por lo menos una de las decisiones no se tomará sino hasta que uno o más de los parámetros estocásticos del sistema hayan sido observados. Es decir, que la solución óptima no será sencillamente un número sino que una función de una o

[1] Hillier F., Lieberman, G., *Introduction to Operations Research*, Holden Day, San Francisco, 1968, p. 402.

[2] Hadley, G., *Non Linear and Dynamic Programming*, Addison-Wesley, Massachusetts, 1964, p. 158.

más variables relacionadas con las variables aleatorias que se observaron antes de tomar la decisión.[3]

Precesos Markov

Un proceso markov es un modelo matemático que se presta para el estudio de sistemas complejos. Los dos conceptos fundamentales de los procesos Markov son el de "estado" y el de "transición".

Un estado es un vector de parámetros que describe la condición del sistema. Así pues, se dice que un sistema ocupa cierto "estado" cuando éste es completamente descrito por los valores de las variables que definen ese "estado". Lógicamente, el sistema hará una **transición** cuando pasa a ser descrito por los valores de las variables de otro estado. R. Howard, explicando el concepto de procesos Markov, hace una analogía que difícilmente se puede superar en cuanto a simpleza. Dice que el proceso se puede representar por un sapo que descansa sobre la hoja de un lirio en un lago, y con el pasar del tiempo el sapo salta de una hoja a otra, según se le antoje de momento. El estado del sistema estaría entonces representado por el número de la hoja sobre la cual se encuentra el sapo; y cada salto sería una transición. Si el número de hojas de lirio es finito entonces se tiene un proceso con un número finito de estados, y el proceso tendría un número infinito de estados si hubiera un número infinito de lirios. El concepto de estado puede estar caracterizado por propiedades cualitativas o cuantitativas.

Para estudiar procesos que no son continuos con respecto al tiempo, se debe indicar la naturaleza probabilística de las transiciones. Se puede entonces asumir que las decisiones se toman sólo en puntos discretos de tiempo igualmente distanciados unos de otros, donde el tiempo entre dos decisiones sucesivas se denomina como un período. Ahora bien, podría especificarse un conjunto de probabilidades condicionales, p_{ij}, las cuales representen la probabilidad de que el sistema se encuentre en el estado j en la próxima transición, dado que ahora ocupa el estado i. Ya que el sistema deberá ocupar algún estado al final de la siguiente transición, y asumiendo que hay N estados posibles.

$$\sum_{J=1}^{N} p_{ij} = 1$$

Y como p_{ij} son probabilidades,

$$0 \leq p_{ij} \leq 1$$

[3] Hadley, *op. cit.*, p. 174.

Las probabilidades de transición pueden entonces representarse convenientemente en forma de matriz:

$$P = [p_{ij}] = \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ " & " & & " \\ " & " & & " \\ " & " & & " \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{bmatrix}$$

Esta matriz proporciona una descripción completa del proceso y se puede utilizar para el análisis de muchos aspectos del proceso. Por ejemplo, se puede calcular la probabilidad de que el sistema se encuentre en el estado i después de n transiciones si su estado actual es conocido, sencillamente computando la enésima potencia de la matriz de transición:

$$P^{(n)} = P.P.\dots.P = P^n$$

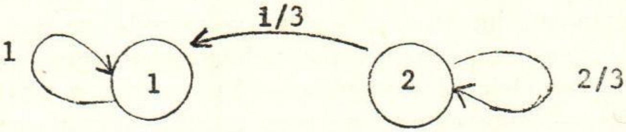
donde $P^{(n)}$ es la probabilidad antes mencionada. El proceso se comporta de diferentes maneras de acuerdo con las características que tenga la matriz de transiciones. Primeramente, hay estados que pueden llamarse **transitorios,** ya que es seguro que no estarán ocupados por el sistema después de muchas transiciones. Luego están los que podrían denominarse estados **absorbentes** o capturadores, ya que una vez que el sistema ocupa dicho estado, no puede salir de él; es decir, la probabilidad de transición es cero o, equivalentemente, su probabilidad de permanencia es uno ($P_{ii} = 1$). Un estado se llama **recurrente** si una vez que el sistema ha ocupado dicho estado se sabe que volverá a ocuparlo más adelante. Un estado recurrente puede exhibir la propiedad de **periodicidad** si el sistema ocupara ahora un estado y luego lo vuelve a ocupar después de p, 2p, 3p, 4p... transiciones, donde p es un número entero que describe la periodicidad del sistema. Es importante hacer notar que un vector de estados o "cadena" puede exhibir las características mencionadas. Por ejemplo, puede haber una "cadena recurrente". En este caso, el sistema cambiaría de un estado a otro entre los estados de una cadena, pero sin salirse de la misma. Para aclarar estos conceptos se presentarán algunos ejemplos simples.

Un proceso de dos estados posibles podría ser representado así:

$$P = \begin{bmatrix} 1 & 0 \\ 1/3 & 2/3 \end{bmatrix}$$

Así, si el sistema se encuentra en el estado 2, tiene una probabilidad de encontrarse en la próxima transición en el estado 1 de un tercio ($p_{21} =$

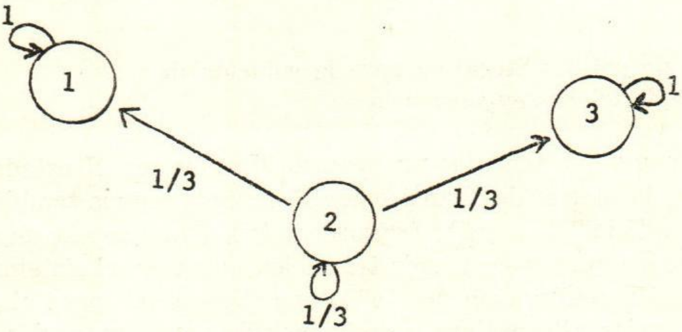
1/3). Sin embargo, si el sistema llegara a ocupar el estado 1 no podría volver a abandonarlo, ya que la probabilidad de quedarse en el mismo es de uno ($p_{11} = 1$). Por lo tanto, el estado 2 es transitorio mientras que el estado 1 es absorbente. Gráficamente,



Un ejemplo de un proceso de tres estados con 2 cadenas recurrentes es el siguiente:

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 1/3 & 1/3 & 1/3 \\ 0 & 0 & 1 \end{bmatrix}$$

que gráficamente podría representarse



Como se observa, los estados 1 y 3 constituyen estados recurrentes y a su vez son absorbentes. Su comportamiento no cambiaría si cada uno estuviera constituido por una serie de estados (cadena). El estado 2 es transitorio ya que puede enviar al sistema al estado 1 o al 3, o sea a cualquiera de las "cadenas recurrentes". Por último es importante mencionar una condición del proceso que llamaremos ergodicidad.[4] Se dice que un

[4] Este término es una adaptación liberal del inglés "ergodicity". Para mayor detalle con respecto a esta condición, ver Howard, *op. cit.*, o Parzen, E., *Modern Probability Theory and Its Applications*, Wiley, New York, 1960, pp. 139-140.

proceso Markov es completamente ergódico cuando la probabilidad de que el sistema se encuentre en cierto estado después de un número grande de transiciones es independiente del estado inicial. Es decir, el valor límite al que tienden las probabilidades de encontrarse en diferentes estados se hace independiente del estado donde se encontraba el sistema inicialmente. Este concepto está ligado al de cadena recurrente, ya que todo proceso Markov debe tener por lo menos una cadena recurrente, y si tiene sólo una entonces es completamente ergódico. El propósito de revisar —aunque muy brevemente— todos estos conceptos, es el de lograr una mejor comprensión de cómo encaja esto en la toma de decisiones al nivel nacional e internacional. Se puede notar, por ejemplo, que si se consideran las condiciones que caracterizan a un país dado en cierto momento como su "estado" y a su vez se consideran varios estados posibles para ese país, entonces se podrían aplicar los conceptos expuestos anteriormente. Así, por ejemplo, podría haber estados absorbentes, como podría serlo el caso de una dictadura fuerte o un estado de coloniaje del cual no haya probabilidad de salir, al menos por un tiempo; el concepto de cadenas recurrentes podría interpretarse como ciclos inflacionarios o recesiones económicas; un caso de periodicidad en el plano político sería la elección presidencial alterna entre partidos conservadores y liberal como fue el caso colombiano hasta hace poco.

El método de política iterativa para la solución de procesos de decisiones en secuencia

Supongamos que tenemos un proceso Markov con N estados posibles, descrito por la matriz de **Transiciones** P. Introduzcamos también una matriz de "beneficio" $R = (r_{ij})$, (reward matrix), que asocia un pago —positivo, negativo o cero— con cada transición que haga el sistema. Es decir, cada transición conlleva un "costo" o una "ganancia" para el sistema. Si existiesen varias **alternativas** o decisiones de entre las cuales se pudiese escoger una cuando el sistema se encuentre en un estado determinado, entonces las decisiones que se tomen afectarán los beneficios del sistema. O sea que, si denominamos al estado en el que se encuentra el sistema con i, al estado al que pasará luego con j, al beneficio que derivará el sistema en su transición de i a j con r_{ij}, y a las alternativas asociadas con cada estado del sistema con k (el número de alternativas posibles en cada estado puede ser diferente) entonces podríamos pensar en determinar aquel conjunto de decisiones que proporcionen el mayor beneficio para el sistema. Aquí empezamos a topar con dificultades, ya que si la situación es una en la que las funciones de densidad de las variables estocásticas y de los costos pueden cambiar, entonces sólo se podría utilizar un "horizonte de planeamiento" limitado. Así, por ejemplo, si este fuera el caso, sólo sería de interés el determinar la decisión inicial que sea óptima, ya que a la hora de tomar la segunda decisión la situación puede haber cambiado considerablemente y por lo tanto necesitaría una nueva solución con los datos nuevos.[5] Aún si se asumiera que dichas densidades no cambian pero que el proceso tiene una duración corta estaríamos bastante limitados, ya que en la realidad no es difícil encontrar gran cantidad de procesos que, para propósitos de análisis, podrían considerarse de duración indefinida. Además, como dice Howard, es importante el reconocer que a menudo un proceso no tiene que tener muchas etapas o períodos para que un análisis de "larga duración" sea significativo. Y tratándose este ensayo de aplicar métodos decisionarios en secuencia a decisiones gubernamentales, pareciera que es este último tipo de análisis el que sería de más utilidad, ya que raramente los planes gubernamentales son de duración corta y de terminación inminente. Nos referimos aquí al proceso de desarrollo de un país o sistema. Por ejemplo, dado que un determinado país se encuentra en una situación particular y cuenta con varias alternativas a tomar que podrían conducir al país a otro estado (con beneficios o costos implícitos en la transición) y así sucesivamente por varias etapas, no es de suponer que el proceso de transiciones —que en el mejor de los casos conducen al sistema al logro de sus objetivos— encuentre terminación sin antes tomar un número de decisiones que para su análisis podría considerarse infinito. Además, es importante hacer notar que puede haber un número infinito de decisiones para un proceso ya sea porque el número de períodos es infinito o porque la duración del tiempo entre decisiones sucesivas es insignificante[6] y entonces, a pesar de que el número de etapas sea finito, el número de decisiones a tomar es infinito, o por lo menos muy grande. Lo que se necesita pues, es un método de tomar decisiones para cualquier estado en que se encuentre el sistema, de manera que sus beneficios se maximicen. Es evidente que no podemos utilizar el "beneficio total esperado" como criterio en el caso de procesos infinitos, ya que dicho beneficio tendería al infinito entre mayor sea el número de transiciones.[7] Un

[5] Hadley, *op. cit.*, p. 175.

[6] Nemhauser, *Introduction to Dynamic Programming*, Wiley, New York, 1967, pp. 210-212.

[7] Es de notar que cuando se considera relevante el costo del dinero o cuando existe incertidumbre respecto a la duración del proceso, se puede utilizar el beneficio total esperado como criterio. La corrección que se hace para evitar que los beneficios crezcan sin límite es el introducir un factor de descuento que permite computar el "valor presente" del beneficio total esperado. Ver Howard, *op. cit.*, capítulo 7.

criterio más relevante sería el "beneficio promedio" del proceso por unidad de tiempo.

Antes de seguir adelante, es importante el que estemos conciente de los supuestos que conlleva el análisis de procesos infinitos. Se asume que las funciones de densidad de las variables aleatorias y de los costos no cambian de períodos a período. También se asume que las decisiones se toman únicamente en puntos discretos de tiempo, igualmente espaciados. De esta manera, consideraremos que el proceso durará indefinidamente y tomaremos en cuenta el efecto de las decisiones que se tomarán en el futuro.

Aun cuando aceptemos todo lo anterior, queda otro aspecto por considerar. Este se refiere a la condición de ergodicidad del proceso. Esto es relevante, ya que el procedimiento de solución es un poco diferente para casos en los que el proceso es completamente ergódico y los que no lo son. Debido a que ambos casos podrían ser útiles en la toma de decisiones al nivel nacional, e internacional, los trataremos separadamente.

El caso de una sola cadena recurrente (completamente ergódico)

Como se dijo anteriormente, la condición de completa ergodicidad se da cuando el valor de la probabilidad de que el sistema se encuentre en un estado después de muchas transiciones, es independiente del estado inicial. Vamos a esbozar a continuación el método de Política-iterativa. Consideremos un proceso Markov completamente ergódico de N períodos, descrito por una matriz de transiciones $P = (p_{ij})$, y una matriz de beneficios $R = (r_{ij})$. El proceso se observa en puntos de tiempo discretos $t = 0, 1, 2, 3, \ldots$. En cada uno de estos tiempos hay que tomar una decisión k de todas las posibles alternativas disponibles en el momento. Lo que se quiere es determinar aquel "set" de decisiones (que especifican exactamente cuál alternativa escoger bajo todas las circunstancias posibles) que maximice los beneficios promedio por transición. Este "set" de decisiones, que sería la política óptima, se puede representar como un vector,

$$d = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_i \\ \vdots \\ d_n \end{bmatrix}$$

que especifica que si el sistema se encuentra en el estado 1 se debe tomar la decisión d_1, si se encuentra en el estado i se tomará la decisión d_i, etc.

Para dar una idea de lo tedioso que sería el cómputo necesario si hubiera que evaluar todas las alternativas posibles, supongamos que un proceso consta de 30 estados y 60 alternativas. El número de políticas posibles sería de 30^{60}, o sea, un número en el orden de ochenta y nueve cifras. El método de política-iterativa desarrollado por R. Howard, sin embargo, permite hallar la política óptima en un número pequeño de iteraciones. Este método consta de 2 partes: la de evaluación y la de mejoramiento de política.

La evaluación

En esta operación sencillamente se evalúa el beneficio para el sistema de una política dada. Sin pretender presentar un desarrollo matemático riguroso, vamos a definir algunas variables de importancia.

$v_i\ (n)$ = beneficio total esperado que el sistema obtendrá al cabo de n transiciones si empieza en el estado i bajo una política determinada.

g_i = beneficio esperado por transición si se empieza en el estado i (también es el beneficio del período i).

q_i = beneficio esperado de una sola transición desde el estado i (beneficio inmediato esperado) bajo la alternativa k.

π_i = valor al que tiende la probabilidad del sistema de encontrarse en el estado i cuando n tiende a infinito.

Bajo la condición de completa ergodicidad, el beneficio del sistema está dado por

[8] La metodología que se presenta a continuación es básicamente la que presenta Howard, con algunas simplificaciones. Por tanto, quienes están interesados en el desarrollo matemático riguroso del método de política iterativa, son referidos a Howard, R., *op. cit.*

* Nemhauser, *op. cit.*, justifica esta relación informalmente. En condiciones de equilibrio (steady state) el beneficio esperado de la transición de i a j es $r_{ij}\ P_{ij}\ \pi_i$. Sumando sobre i y j se obtiene

$$g = \sum_{i=1}^{N} \sum_{j=1}^{N} r_{ij}\ P_{ij}\ \pi_i$$

si sustituimos

$$q_i = \sum_{j=1}^{N} r_{ij}\ P_{ij}$$

obtenemos

$$g = \sum_{t=1}^{N} \pi_i\ q_i$$

$$g = \sum_{i=1}^{N} \pi_i q_i \quad (*)$$

que es la cantidad que se requiere maximizar. De igual manera, v_i (n) debe satisfacer la relación.

$$v_i (n) = \sum_{j=1}^{N} P_{ij} [r_{ij} + v_j (n-1)] \quad i = 1, 2, \ldots, N \quad n = 1, 2, 3, \ldots.$$

Es decir, que al efectuarse una transición de i a j, el sistema obtiene un beneficio r_{ij}, pero además hay que considerar que después de la primera transición todavía le quedan (n-1) transiciones, las cuales comenzaría encontrándose en el estado j. Entonces hay que añadirle al beneficio inmediato esperado el beneficio que espera de los (n-1) períodos que le quedarían, todo "pesado" por la probabilidad de que la transición de i a j ocurra, p_{ij}. De la ecuación anterior se obtiene

$$vi (n) = \sum_{j=1}^{N} p_{ij} r_{ij} + \sum_{j=1}^{N} p_{ij} v_j (n-1) \quad i = 1, 2, \ldots, N \quad n = 1, 2, 3, \ldots$$

y si se define

$$q_i = \sum_{j=1}^{N} p_{ij} r_{ij} \quad i = 1, 2, \ldots, N$$

se obtiene

$$v_i (n) = q_i + \sum_{j=1}^{N} p_{ij} v_j (n-1) \quad i = 1, 2, \ldots, N \quad n = 1, 2, 3, \ldots \qquad (1)$$

Se puede demostrar que cuando n tiende a infinito

$$v_i (n) = ng + v_i \quad i = 1, 2, \ldots, N \qquad (2)$$

Sustituyendo la ecuación (2) en la (1) se obtiene

$$g + v_i = q_i + \sum_{j=1}^{N} p_{ij} v_j \quad i = 1, 2, \ldots, N$$

Esta es la relación que deseábamos obtener, ya que en ella se relacionan g y v_i con la estructura de probabilidades y la de beneficios, del proceso. Este sistema de ecuaciones tiene N ecuaciones, pero hay N + 1 incógnitas que son las N v_j y una g. Entonces no queda más que igualar una de las v's, por ejemplo v_n, a cero y resolver el sistema de N ecuaciones con N incógnitas para obtener lo que se ha llamado "valores relativos" de la política en cuestión. La diferencia entre los valores relativos de dos estados (ej. $v_1 - v_2$) sería la ventaja relativa de empezar en 1 y no en 2, o sea, lo que racionalmente uno daría por empezar el proceso desde el estado 1 y no del 2.

Hasta aquí hemos logrado evaluar una política determinada por medio del beneficio esperado. Sin embargo, no tenemos ningún método para escoger otra política que sea superior a la primera. Esta es precisamente la contribución de la fase de Mejoramiento de Política, que a continuación explicamos.

Mejoramiento de política

Para determinar la mejor alternativa k en el estado i, basta evaluar la siguiente cantidad

$$q_i^k + \sum_{j=1}^{N} p_{ij}^k v_j \quad (*)$$

con respecto a las alternativas disponibles en ese estado y utilizando los valores relativos (v_j) encontrados anteriormente. La alternativa que maximice esta cantidad se convierte entonces en d_i y la decisión en el estado i. Repitiendo esto para cada estado se determina una nueva política, que asimismo se puede evaluar, luego mejorar, etc.

Howard ha combinado las fases de Evaluación y Mejoramiento de Política en lo que él llama el "Ciclo-Iterativo", que se presenta a continuación, en la fig. 1:

* Para hallar la mejor alternativa en el estado i para el período n + 1, dado que tenemos una política óptima hasta el período n, maximizamos

$$q_i^k + \sum_{j=1}^{N} p_{ij}^k v_j \ (n)$$

Para n muy grande y sustituyendo la ecuación (2) en la anterior, se tiene

$$q_i^k + \sum_{j=1}^{N} p_{ij}^k (ng + V_j)$$

Pero $\sum_{j=1}^{N} p_{ij}^k = 1$, lo que haría la contribución de ng independiente de k,

por lo que sencillamente se puede maximizar

$$q_i^k + \sum_{j=1}^{N} p_{ij}^k vj$$

Evaluación

Use p_{ij}, q_i en una política dada para evaluar

$$g + v_i = q_i + \sum_{j=1}^{N} p_{ij} v_j \quad i = 1, 2, \ldots, N$$

y encontrará los valores relativos v_i y g, igualando v_n a cero.

Mejoramiento de Política

Para cada estado i encuentre la alternativa k′ que maximice

$$q_i^{k} + \sum_{j=1}^{N} p_{ij}^{k} v_j$$

utilizando los v_j, encontrados. Entonces k′ se convierte en d_i, $q_i^{k'}$ en q_i y $p_{ij}^{k'}$ en p_{ij}

Fig. 1

Es decir, en la primera fase se encuentran los beneficios en función de una política y en la segunda se encuentra una política en función de los beneficios. Es interesante ver que el ciclo se puede iniciar en cualquiera de las dos fases. De no existir ningún criterio a priori de qué valores o qué política utilizar para comenzar, es generalmente satisfactorio usar los $v_j = 0$ y empezar en la fase de mejoramiento de política.

Si dos políticas sucesivas son iguales entonces se habrá encontrado la política óptima y el ciclo iterativo terminaría. Howard ha probado que el beneficio de la nueva política determinada por el "mejoramiento de política" es superior al de la anterior. Dicho de otra manera, si empezando con la política A se obtiene la política B que sea diferente de A, entonces $g^B > g^A$. No sólo esto, sino que es imposible que exista una política superior sin que ésta sea encontrada después de cierto número de iteraciones por este método. Resumiendo, el método de Política — Iterativa se reduce a

a) resolver sistemas de ecuaciones simultáneas y compararlos.
b) encontrar políticas sucesivas, cada una de las cuales será superior a la anterior.
c) terminar el ciclo iterativo cuando dos políticas sucesivas sean iguales.

Ejemplo:

Imaginémenos un caso hipotético que a todas luces es una simplificación exagerada de la realidad, pero que para efectos de ilustración sirve su propósito. Supongamos que un país dado se puede encontrar en tres condiciones o estados, a través del tiempo. Estos estados son 1) Coloniaje, 2) Dependencia concentrada y 3) Interdependencia. Encontrándose en cada uno de estos estados, el país tiene varias alternativas a escoger:

si se encuentra en estado de coloniaje puede:

1. Efectuar una revolución.
2. Diversificar sus exportaciones (exportar a más países).
3. No hacer nada.

Si se encuentra en un estado de dependencia concentrada podría:

1. Llevar a cabo un programa de sustitución de importaciones
2. Diversificar sus exportaciones

Por último, si se encuentra en un estado de interdependencia podría:

1. Llevar a cabo programa de sustitución de importaciones
2. Aumentar sus exportaciones
3. Integrarse en un mercado común.

Para cada estado existe una probabilidad de que el sistema (país) pueda pasar a otro de los estados. Asimismo, existe un "beneficio" asociado con cada transición del país bajo cada una de las alternativas. Aquí tropezamos otra vez con dificultades. ¿Cómo determinar dichas probabilidades de transición y los beneficios de las transiciones para el sistema? Obviamente esto no es nada sencillo e inevitablemente contendrá elementos subjetivos. Las probabilidades deberán estimarse en base a toda la información histórica económica, social y política de país o países en cuestión, que sea relevante. La "matriz de beneficios" puede constituirse en base a diferentes criterios. Por ejemplo, se podría medir el impacto que una transición bajo una determinada política tendría sobre índices sociales o económicos, tales como el producto nacional bruto, el bienestar, la taza anual de crecimiento, el poder en la toma de decisiones, el comercio exterior, la mortalidad, la esperanza de vida, etc. Algunos de estos índices son fácilmente obtenibles. En Puerto Rico existe un modelo de crecimiento económico que permite computar el impacto que una política de sustitución de importaciones o de un cierto porcentaje de ahorros tendría en el PBN, en el empleo, en el comercio exterior, etc. Por supuesto, el grado de dificultad que se encuentre en este proceso depende en gran medida de los flujos de información del sistema, así como de la capacidad predictiva de quienes hagan las estimaciones. En nuestro ejemplo, asumimos que el impacto o "beneficio" de las diferentes transiciones bajo cada política se ha determinado en base a un índice que combina aspectos sociales, económicos y políticos con unidades arbitrarias. Los datos para el problema se presentan a continuación, en el cuadro No. 1.

Estado i	Alternativa k	probabilidad p_{ij}^k			beneficio r_{ij}^k			beneficio inmediato esperado
		$j=1$	2	3	$j=1$	2	3	$q_i^k = \sum_{j=1}^{N} p_{ij}^k r_{ij}^k$
1	1	0	1/2	1/2	—10	3	12	7.5 ←
	2	3/4	1/8	1/8	2	4	5	2.62
	3	1	0	0	0	1	3	0
2	1	1/4	1/4	1/2	— 4	3	5	2.5
	2	0	1/4	3/4	— 3	0	6	4.5 ←
3	1	1/16	1/16	7/8	—13	— 6	8	5.8
	2	1/8	1/8	3/4	—10	— 5	8	4.12
	3	1/8	1/8	3/4	—12	— 6	13	7.5 ←

Cuadro Nº 1

Lo que perseguimos es encontrar aquella política que nos indique qué hacer en todas las posibles circunstancias, de manera que se maximice el beneficio del país después de muchas transiciones. No teniendo criterios a-priori, igualamos v_1, v_2 y v_3 a cero y empezamos con aquella política que maximice los beneficios inmediatos esperados, que de acuerdo a los datos sería

$$d = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Es decir, escoger la alternativa 1 si el país se encuentra en el estado 1, la alternativa 2 si se encuentra en el estado 2 y la alternativa 3 si se encuentra en el tercer estado. La matriz de probabilidades y los beneficios inmediatos esperados correspondientes a esta política son

$$p = \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/4 & 1/4 & 1/2 \\ 1/8 & 1/8 & 3/4 \end{bmatrix} \qquad q = \begin{bmatrix} 7.5 \\ 2.5 \\ 7.5 \end{bmatrix}$$

Procedemos entonces a la fase de evaluación. Para esto es necesario resolver el sistema de ecuaciones

$$g + v_i = q_i + \sum_{j=1}^{N} p_{ij}\, v_j \qquad i = 1, 2, \ldots, N$$

que nos da

$$\begin{aligned} g + v_1 &= 7.5 + o v_1 + 1/2\, v_2 + 1/2\, v_3 \\ g + v_2 &= 4.5 + 1/4\, v_1 + 1/4\, v_2 + 1/2\, v_3 \\ g + v_3 &= 7.5 + 1/8\, v_1 + 1/8\, v_2 + 3/4\, v_3 \end{aligned}$$

Si asumimos que $v_3 = 0$ y resolvemos las ecuaciones obtenemos

$$\begin{aligned} v_1 &= -\ 1.2 \\ v_2 &= -\ 3.6 \\ v_3 &= 0 \\ g &= 6.9 \end{aligned}$$

Esto quiere decir que el país obtendría un beneficio de 6.9 unidades por transición si sigue la política mencionada. Ahora entramos a la fase de **mejoramiento de política.** El resultado aparece en el cuadro 2.

Estado i	Alternativa k	Cantidad para comparar $q_i^k + \sum_{j=1}^{N} p_{ij}^k v_j$
1	1	5.70
	2	1.27
	3	—1.20
2	1	1.30
	2	3.60
3	1	5.50
	2	3.52
	3	6.90

TABLA 2

De estas comparaciones, donde se utilizaron los v_j obtenidos en la fase de evaluación, se observa que a pesar de que los beneficios esperados de las diferentes alternativas variaron, los cambios no fueron suficientes para el utilizamiento de una política diferente. Es decir, la política que maximizó el beneficio sigue siendo la de escoger la 1ra. alternativa si el país se encuentra en el primer estado, la segunda si se encuentra en el segundo, y la 3ra. alternativa de encontrarse en el tercer estado. O sea, que nuevamente se ha obtenido

$$d = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

y por lo tanto se ha hallado la política óptima. De no haber sido este el caso, y de haberse encontrado otra política diferente a la primera, que tendría que ser superior a la misma, entonces hubiéramos procedido a entrar en otro ciclo iterativo. Así pues, hubiéramos buscado los v_j para la nueva política, utilizando una nueva matriz de probabilidades y un nuevo vector "q" de beneficios inmediatos. Luego, entraríamos de nuevo a la fase de mejoramiento de política, y así sucesivamente, hasta obtener la misma política en iteraciones consecutivas.

El caso de cadenas recurrentes múltiples

Hasta el momento se han considerado casos en los que se suponía que el proceso era completamente ergódico; es decir, que sólo tenía una cadena recurrente. El método desarrollado es generalmente satisfactorio, ya que muchas situaciones reales se pueden adaptar al mismo. Sin embargo, no es difícil encontrar situaciones que no se ajustan a la condición de completa ergodicidad. Así, por ejemplo, podría darse la situación de que un sistema pueda encontrarse en varios estados, donde más de uno sea un estado absorbente o donde haya varias cadenas recurrentes. De entre los estados posibles en que un sistema se puede encontrar, o a los que podría llegar si sigue determinadas políticas, podrían figurar los de dictadura fuerte y de colonia, que se podrían considerar como dos cadenas recurrentes en el sentido de que una vez que el sistema cae en ellos se le hace casi imposible salir. Pero no entremos ahora en mucho detalle de ejemplos para poder concentrarnos en el método de solución requerido para estas situaciones.

Una diferencia básica entre los dos casos es que cuando el proceso es completamente ergódico se tiene un beneficio único (g), y por lo tanto el problema se reducía a buscar aquella política que tuviera el mayor beneficio esperado. Mientras que en el caso de cadenas múltiples, el beneficio depende no sólo del proceso sino que también del estado en que se empieza, de tal manera que si se empieza en el estado i se obtiene un beneficio g_i. Entonces el nuevo objetivo debe ser el de encontrar aquella política que maximice el beneficio de **todos** los estados del sistema.

La evaluación

A pesar de que cada estado posee un beneficio g_i se puede probar que el beneficio de todos los elementos de una cadena recurrente es idéntico.[9] Para hallar g_i podríamos utilizar la relación (1), que ya se discutió

$$v_i(n+1) = q_i + \sum_{j=1}^{N} p_{ij} v_j(n) \qquad i = 1, 2, \ldots, N$$

donde n se ha sustituido por (n+1). Reemplazando v_i (n) por su valor cuando n es muy grande, que es

$$v_i = ng_i + v_i \qquad i = 1, 2, \ldots, N$$

y reemplazando en esta última relación n por (n+1), da

$$(n+1) g_i + v_i = q_i + \sum_{j=1}^{N} p_{ij} (ng_{ij} + v_j)$$

[9] Howard, R., *op. cit.*, p. 24.

o sea

$$ng_i + g_i + v_i = q_i + n \sum_{j=1}^{N} p_{ij}\ g_{ij} + \sum_{j=1}^{N} p_{ij}\ v_j \qquad (3)$$

Para que esta ecuación sea satisfecha para n grande

$$g_i = \sum_{j=1}^{N} p_{ij}\ g_j \qquad i = 1, 2, \ldots, N \qquad (4)$$

por lo que la ecuación (3) se reduce a

$$g_i + v_i = q_i + \sum_{j=1}^{N} p_{ij}\ v_j \qquad i = 1, 2, \ldots, N \qquad (5)$$

Ahora tenemos dos sistemas de N ecuaciones cada uno. (ecuaciones (4) y (5)), para encontrar los valores de los N $g_{i's}$ y de los N $v_{i's}$. Sin embargo, ya que sólo habrá tantos $g_{i's}$ independientes como haya cadenas recurrentes, utilizaremos la ecuación (4) para obtener los g_i correspondientes a los estados transitorios una vez que hallemos los correspondientes a las cadenas recurrentes. La razón por la que no podemos usar las ecuaciones (4) para directamente hallar los valores de los $g_{'s}$ es porque la matriz (I-P) es singular. Entonces recurrimos a las ecuaciones (5) para obtener los L $g_{i's}$ independientes y los N valores v_i. Pero otra vez surge el problema de tener más incógnitas que ecuaciones; en este caso hay L incógnitas de más. Aplicando el mismo procedimiento que antes, arbitrariamente igualamos L $v_{i's}$ a cero correspondientes a un estado en cada cadena recurrente, que podría ser el de mayor i. Ahora sí es posible resolver para los L g_i independientes y para los (N-L) v_i que quedan.

El mejoramiento de política

Una vez que hemos obtenido los beneficios y los valores relativos para cada estado del proceso, el siguiente paso es el de encontrar otra política que mejore los beneficios obtenidos. Dicho de otra manera, si hemos adaptado una política determinada durante los primeros n períodos, queremos encontrar una mejor para cualquier estado i en el período n + 1. Esto se puede lograr escogiendo aquella alternativa k que maximice

$$q_i^{k} + \sum_{j=1}^{N} p_{ij}^{k}\ v_j\ (n).$$

Otra vez, valiéndonos de que para n grande $v_i(n) = ng + v_i$ obtenemos

$$q_i^k + \sum_{j=1}^{N} p_{ij}^k (ng_j + v_j)$$

o sea

$$q_i^k + n \sum_{j=1}^{N} p_{ij}^k g_j + \sum_{j=1}^{N} p_{ij}^k v_j$$

Obviamente, cuando n es muy grande, la expresión anterior es maximizada por aquella alternativa que maximice el factor

$$\sum_{j=1}^{N} p_{ij}^k g_i$$

ya que este está multiplicado por n. En caso de empate, entonces se escogerá aquella alternativa que maximice la otra parte de la expresión, o sea

$$q_i^k + \sum_{j=1}^{N} p_{ij}^k v_j$$

donde los v_j son los obtenidos en la fase anterior, de evaluación. Resumiendo en forma esquemática, el ciclo iterativo para el caso de cadenas múltiples aparece en la fig. 2.

Evaluación

Para un set de p_{ij} y q_i y una política dada resuelva

$$qi + v_i = q_i + \sum_{j=1}^{N} p_{ij}\, g_i \qquad i = 1, 2, \ldots, N$$

$$g_i = \sum_{j=1}^{N} p_{ij}\, g_i \qquad 1 = 1.2, \ldots, N$$

para todo v_i y g_i igualando un v_i a cero en cada cadena recurrente (L en total)

$$p_{ij} = p_{ij}^{k}$$
$$q_i = q_i^{k}$$

Mejoramiento de Política

Determinar la alternativa k, para cada estado i que maximice

$$\sum_{j=1}^{N} p_{ij}^{k}\, g_j$$

En caso de empate, utilice la alternativa que maximice

$$q_i^{k} + \sum_{j=1}^{N} p_{ij}^{k}\, v_j$$

Fig. Núm. 2

El ciclo se termina cuando dos políticas sucesivas son iguales.

Ejemplo

Otro ejemplo hipotético, que más que representar una situación posible trata de ilustrar el procedimiento, sería el de un país con tres estados posibles y tres alternativas en cada estado. Los posibles estados son:

1. Una dictadura
2. Subdesarrollo crónico
3. Colonia económica

Las alternativas disponibles en cada uno de los casos son:

1. No hacer nada
2. Llevar a cabo política tendiente a reducir el grado de dependencia
3. Efectuar revolución

Supongamos que los datos para este problema son los presentados en el cuadro Núm. 2.[10]

Estado i	Alternativa k	Probabilidades p_{i1}^k	p_{i2}^k	p_{i3}^k	Beneficio inmediato esperado g_i^k
1	1	1	0	0	1
	2	0	1	0	2
	3	0	0	1	3
2	1	1	0	0	6
	2	0	1	0	4
	3	0	0	1	5
3	1	1	0	0	8
	2	0	1	0	9
	3	0	0	1	7

Cuadro Núm. 3

La primera política sería aquella que maximice los beneficios inmediatos esperados, o sea

$$d = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} \qquad p = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \qquad q = \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix}$$

Inmediatamente empezamos la tarea de evaluación. La ecuación (4) nos da

$$g_1 = g_3 \qquad g_2 = g_1 \qquad g_3 = g_2$$

lo cual indica que en realidad sólo hay una cadena recurrente a la cual pertenecen los tres estados y por lo tanto existe sólo una ganancia $g = g_1 = g_2 = g_3$. Podríamos proseguir, pero si lo hacemos no se exhibiría el efecto de un proceso con cadenas múltiples. Sin embargo, si escogemos como política inicial otra, como por ejemplo,

$$d = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \qquad p = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \qquad q = \begin{bmatrix} 3 \\ 4 \\ 8 \end{bmatrix}$$

entonces sí tendríamos dos cadenas, que lo demuestra el resultado de la ecuación (4):

$$g_1 = g_3 \qquad g_2 = g_2 \qquad g_3 = g_1$$

[10] Este ejemplo es una adaptación de uno dado por Howard, *op. cit.*, y utiliza los mismos datos numéricos.

Aquí la primera cadena está compuesta por los estados 1 y 3 y la segunda por el estado 2. Si llamamos ${}^{i}g$ al beneficio de la cadena i, entonces

$$g_1 = g_3 = {}^{1}g, \text{ y } g_2 = {}^{2}g$$

Igualando $v_2 = v_3 = 0$ y utilizando la ecuación (5),

$${}^{1}g + v_1 = 3 \qquad {}^{2}g = 4 \qquad {}^{1}g = 8 + v_1$$

que da:

$$g_1 = 11/2 \qquad g_2 = 4 \qquad g_3 = 11/2$$
$$v_1 = -5/2 \qquad v_2 = 0 \qquad v_3 = 0$$

Una vez obtenidos los g_i y los v_i, empezamos la fase de mejoramiento de política, que aparece en el cuadro 4.

Estado i	Alternativa k	Prueba de beneficios $\sum_{j=1}^{N} p_{ij}^{k} g_j$	Prueba de valores $g_i^{k} + \sum_{j=1}^{N} p_{ij}^{k} v_j$	
1	1	11/2	— 3/2	
	2	4	2	
	3	11/2	3	←
2	1	11/2	7/2	
	2	4	4	
	3	11/2	5	←
3	1	11/2	11/2	
	2	4	9	
	3	11/2	7	←

Cuadro Núm. 4

Aquí vemos cómo en el caso de que más de una alternativa dé el mismo valor para la cantidad de comparación de beneficios, se decide en base a la comparación de valores, que aparecen en la última columna de la tabla.

La nueva política es

$$d = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix} \qquad p = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \qquad q = \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix}$$

Para evaluar esta nueva política usamos las ecuaciones 4:

$$g_1 = g_3 \qquad g_2 = g_3 \qquad g_3 = g_3$$

lo que reduce de nuevo el problema a una sola cadena. Entonces podemos hacer $g_1 = g_2 = g_3 = g$, $v_3 = 0$ y por medio de las ecuaciones (5) obtener

$$g + v_1 = 3 \qquad g + v_2 = 5 \qquad g = 7$$

cuya solución da

$$\begin{array}{ll} g_1 = 7 & v_1 = -4 \\ g_2 = 7 & v_2 = -2 \\ g_3 = 7 & v_3 = 0 \end{array}$$

Utilizando estos valores y beneficios entramos de nuevo al mejoramiento de política, cuyos resultados aparecen en el cuadro número 5.

Estado i	Alternativa k	Prueba de beneficios $\sum_j p_{ij}^k g_j$	Prueba de valores $q_i^k + \sum_j p_{ij}^k v_j$
1	1	7	— 3
	2	7	0
	3	7	3 ←
2	1	7	2
	2	7	2
	3	7	5 ←
3	1	7	4
	2	7	7
	3	7	7 ←

Cuadro Núm. 5

En el estado 3 hay un empate entre las alternativas 2 y 3, pero siempre se deja la anterior en estos casos, lo que da la política,

$$d = \begin{bmatrix} 3 \\ 3 \\ 2 \end{bmatrix}$$

que era la misma anterior y por lo tanto se ha obtenido la política óptima. Esta política dice que si el país se encuentra bajo una dictadura debe haber una revolución, si se encuentra en un estado de subdesarrollo crónico debe hacer lo mismo y en el caso de encontrarse bajo el coloniaje debe llevar a cabo una política tendiente a la reducción de la dependencia. Los beneficios esperados de esta política, después de muchos períodos es de 7 unidades, que son arbitrarias. Claro, que estos resultados son meramente ilustrativos. Lo que se ha querido presentar aquí son ejemplos sencillos para ilustrar el procedimiento de análisis en forma simple. Los problemas reales tendrían más variables, las alternativas serían más complejas y variadas y la asignación de probabilidades y de beneficios se dificultaría bastante. A pesar de que los ejemplos dados tienen un enfoque un tanto internacional, el método se puede utilizar en la toma de decisiones de carácter nacional.[11]

Una de las implicaciones de más importancia de este modelo —dentro del contexto de la planificación en sistemas abiertos— es la necesidad de un sistema de información, o más propiamente de "inteligencia", que permita estimados razonables de las probabilidades de transición del sistema, considerando las posibles reacciones exógenas que se producirían en respuesta a las decisiones que se tomen para el sistema en cuestión.

Indudablemente que las premisas implícitas en el modelo (especialmente la de que las probabilidades de transición inmediata no varían) deberán estudiarse más detenidamente con respecto a su validez en el contexto de la toma de decisiones aquí tratada, así como deberá ahondarse en otros aspectos de la aplicación del modelo. Sin embargo, como se estableció al principio, el propósito de este ensayo no ha sido el de presentar un estudio exhaustivo del tema sino que el de estimular la exploración de nuevos enfoques analíticos para la toma de decisiones en sistemas nacionales.

[11] El concepto de procesos Markov ha sido ya aplicado en contextos similares. Paul Smith lo utilizó para estudiar el crecimiento y la asignación de recursos al nivel regional (Smith, P., "Markov Matrices and Regional Development", *Journal of Regional Science*, Vol. 3, No. 1, Summer, 1961, pp. 27-36), en tanto que A. Rogers lo usó para analizar los movimientos migratorios Interregionales y su control (Rogers, A., *Matrix Analysis of Interregional Population Growth and Distribution*, University of California Press, Berkeley, 1968).