UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

División de Estudios de Posgrado
Facultad de Ingeniería

Procesos Markovianos Discretos Aplicados a Sistemas de Recursos Hidráulicos

TESIS DE MAESTRIA .

Que para obtener el grado de:
Maestro en Ingeniería-Investigación de Operaciones

p r e s e n t a :

JAVIER ARTURO DIAZ VARGAS

México D. F. 1980

C O N T E N I D O

1. Introducción

En el análisis y mejoramiento de sistemas en gran escala se observa que las problemáticas con que el decisor se enfrenta son cada día más complejas, pues al conceptualizar un aspecto de la realidad deben atenderse otros factores que influyen en su proceso de desarrollo. La investigación de operaciones interviene en este conflicto, proporcionando modelos y métodos de solución que facilitan la comprensión de la problemática y la toma de decisiones. Su trabajo es interdisciplinario y trata de considerar las partes que afectan al todo, así como sus interrelaciones para entender mejor el carácter dinámico de los problemas que enfrenta.

Un ejemplo típico de sistema en gran escala son los sistemas de aprovechamientos hidráulicos. Este tipo de sistemas tiene diferentes propósitos, por ejemplo, proporcionar agua para el crecimiento de cultivos, generación de energía eléctrica y también, servir como instrumento para el control de avenidas. Como puede observarse, algunos de estos propósitos son conflictivos y tienen diferentes impactos sobre los aspectos económicos y sociales de un país.

La motivación principal de este trabajo es analizar el problema de asignación de agua de un sistema de aprovechamientos hidráulicos a un distrito de riego. Asimismo, se pretende demostrar la forma en que algunas técnicas y modelos de la investigación de operaciones puede ayudarnos en el análisis de esta problemática. Especificamente, se demuestra la manera en que se manipula el caracter aleatorio de los escurrimientos a la presa y como se modela y resuelve, usando el método de Howard, el proceso de decisión periódico a que ésto da lugar.

Con el propósito de entender el método de Howard, en el siguiente capítulo damos algunos conceptos de álgebra lineal, como por ejemplo la forma normal de Jordan, forma que nos permite tratar de modo simple el concepto de convergencia en matrices. En el capítulo tres, con lo hecho en el anterior, tratamos los procesos Markovianos con un lenguaje simple que nos permite demostrar convergencia en el método de Howard de una manera sencilla. Seguidamente, en el capítulo cuatro, se explica la problemática en los distritos de riego y se presentan dos ejemplos numéricos que nos permiten entender mejor dicha problemática y mostrar la aplicación del método de Howard. En el primer ejemplo se determina una estrategia óptima anual y en el otro una semestral. Finalmente algunas conclusiones sobre lo realizado son presentadas.

2. Nociones de álgebra lineal.

En este capítulo estableceremos algunas nociones de álgebra lineal que nos serán útiles en el desarrollo de este trabajo.

2.1. Algunos conceptos básicos.

Definición. Sean V y W espacios vectoriales sobre el campo K. Una transformación lineal

$$T : V \to W$$

es una transformación que satisface que para cualesquiera u,v en V y cualesquiera escalares α, β en K

$$T(\alpha u + \beta v) = \alpha T(u) + \beta T(v)$$

Definición. El conjunto de elementos v en V tales que $T(v)=0$, donde 0 representa el elemento neutro o cero de W, recibe el nombre de núcleo de T. Al conjunto de elementos w de W para los cuales existe un elemento v de V tal que $T(v) = w$, se le conoce como la imagen de T.

A menudo abreviaremos núcleo e imagen, escribiendo simplemente Ker e Im, respectivamente.

El siguiente teorema relaciona las dimensiones del núcleo y de la imagen de una transformación lineal, con la dimensión del espacio sobre el cual está definida la transformación.

Teorema 1. Sean V, W espacios vectoriales de dimensión finita y $T:V\to W$ una transformación lineal. Entonces,

$$\dim V = \dim \operatorname{Ker} T + \dim \operatorname{Im} T$$

donde dim, es una abreviatura para dimensión.

Prueba. Si la imagen de T consta sólo del cero, entonces la afirmación es trivial. Suponga que $\{w_1,\ldots,w_s\}$ es una base de la imagen de T y sean $v_1,\ldots,v_s$ elementos de V tales que $T(v_i)=w_i$, para $i=1,..,s$. Si el núcleo de T no es $\{\Theta\}$, entonces supóngase que $\{u_1,\ldots,u_q\}$ es una base del núcleo. Si el núcleo es $\{\Theta\}$, entonces, en lo que sigue, omita toda referencia a $\{u_1,\ldots,u_q\}$. Afirmamos que $\{v_1,\ldots,v_s, u_1,\ldots,u_q\}$ es una base de V.

Para ver ésto, sea v cualquier elemento de V; entonces existen escalares $\alpha_1,\ldots,\alpha_s$ tales que

$$T(v) = \alpha_1 w_1 +,\ldots,+ \alpha_s w_s = \alpha_1 T(v)+,\ldots,+\alpha_s T(v_s)$$

ya que $\{w_1,\ldots,w_s\}$ es una base de la imagen de T. De aquí, por linealidad,

$$T(v-\alpha_1 v_1 -,\ldots,- \alpha_s v_s) = \Theta$$

Por tanto, $v-\alpha_1 v_1 -,\ldots,- \alpha_s v_s$ está en el núcleo de T y existen escalares $\beta_1,\ldots,\beta_q$ tal que

$$v - \alpha_1 v_1 -,\ldots,-\alpha_s v_s = \beta_1 u_1 +,\ldots,+ \beta_q u_q$$

Por consiguiente,

$$v = \alpha_1 v_1 +,\ldots,+ \alpha_s v_s + \beta_1 u_1 +,\ldots,+ \beta_q u_q$$

es una combinación lineal de $v_1,\ldots,v_s$, $u_1,\ldots,u_q$. Esto prueba que dichos vectores generan V. Ahora, demostraremos que son linealmente independientes y, por tanto, constituyen una base. Supóngase que existe una combinación lineal:

$$\alpha_1 v_1 +,\ldots,+ \alpha_s v_s + \beta_1 u_1 +,\ldots,+ \beta_q u_q = \Theta$$

Aplicando T a esta relación y usando el hecho de que $T(u_i)=\Theta$ para $i=1,\ldots,q$, obtenemos,

$$\alpha_1 T(v_1) +,\ldots,+ \alpha_s T(v_s) = \Theta$$

Pero $T(v_i) = w_i$, $i=1,\ldots,s$, por lo que $\alpha_1 = ,\ldots, = \alpha_s = 0$ ya que $w_1,\ldots,w_s$ son linealmente independientes. Por consiguiente,

$$\beta_1\ u_1 +,\ldots,+ \beta_q\ u_q = \Theta$$

Pero $u_1,\ldots,u_q$ constituyen una base del núcleo de T y, por tanto, $\beta_i=0$ para $i=1,\ldots,q$. Esto concluye la prueba.

Un ejemplo clásico de transformación lineal es una matriz. Por ejemplo, sea A una matriz de mxn en el campo K. Entonces, se puede asociar con A una transformación

$$T_A : K^n \to K^m$$

Haciendo

$$T_A(x) = Ax$$

para cada vector columna x de K^n. Que esta transformación es lineal se deduce simplemente de las propiedades de la multiplicación de matrices.

<u>Definición</u>. T_A es la <u>transformación lineal asociada con la matriz A</u>.

El problema contrario, o sea dada una transformación lineal $T : K^n \to K^m$ asociarle una matriz única A tal que $Ax = T(x)$ para toda x en K^n, es relativamente sencillo. Sean $E^1, \ldots, E^n$ los vectores columna unitarios en K^n, y sean $e^1, \ldots, e^m$ los vectores columna unitarios en K^m. Cualquier vector x de K^n se puede expresar como una combinación lineal

$$x = x_1 E^1 +, \ldots, + x_n E^n$$

donde x_j es la componente j de x. Por linealidad, se tiene que:

$$T(x) = x_1 T(E^1) +, \ldots, + x_n T(E^n)$$

y se puede escribir cada uno de los $T(E^j)$ en términos de $e^1, \ldots, e^m$. En otras palabras, existen números a_{ij} tales que

$$T(E^1) = a_{11} e^1 +, \ldots, + a_{m1} e^m$$

$$\cdot$$

$$T(E^n) = a_{1n} e^1 +, \ldots, + a_{mn} e^n$$

Por tanto,

$$T(x) = x_1(a_{11}e^1 +,\ldots,+ a_{m1}e^m) +,\ldots,+ x_n(a_{1n}e^1+\ldots+a_{mm}e^m)$$

$$= (a_{11}x_1 +,\ldots,+ a_{1n}x_n)e^1 +,\ldots,+ (a_{m1}x_1+,\ldots+a_{mn}x_n)e^m$$

En consecuencia, si hacemos $A = (a_{ij})$ se ve que

$$T(x) = Ax$$

De este modo $T = T_A$ y si $T = T_B$ esto implica que $A = B$ lo que demuestra la unicidad de la matriz A.

Ahora bien, un caso más general, es cuando tratamos con espacios vectoriales cualesquiera V, W sobre K, de dimensión finita. Sean $\{v_1,\ldots,v_n\}$ y $\{w_1,\ldots,w_m\}$ bases de V y W respectivamente. Para cualquier v en V, se puede expresar v en forma unívoca como una combinación lineal

$$v = x_1v_1 +,\ldots,+ x_nv_n \quad ; \quad x_i \text{ en } K$$

Análogamente para W. Si $T:V\to W$ es una transformación lineal, asociaremos a T una matriz única A tal que si x es el vector de coordenadas de un elemento v en V, relativo a la base $\{v_1,\ldots,v_n\}$, entonces Ax, es el vector de coordenadas de T(v), relativo a la base $\{w_1,\ldots,w_m\}$. Para ello, sea

$$T(v_1) = a_{11}\, w_1 +,\ldots,+ a_{m1}\, w_m$$

$$\vdots$$

$$T(v_n) = a_{1n}\, w_1 +,\ldots,+ a_{mn}w_m$$

Ciertamente se tiene que

$$T(v) = T(x_1v_1 +,\ldots,+ x_nv_n) = x_1\ T(v_1) +,\ldots,+ x_nT(v_n)$$

y por lo tanto,

$$T(v) = x_1(a_{11}w_1 +,\ldots,+a_{m1}w_n) +,\ldots,+ x_n(a_{1n}w_1+,\ldots,+a_{mn}w_n)$$

$$= (a_{11}x_1 +,\ldots,+ a_{1n}x_n)w_1 +,\ldots,+ (a_{m1}x_1 +,\ldots,+a_{mn}x_n)w_m$$

Definiendo A como

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

se tiene que Ax es el vector de coordenadas de T(v), respecto a la base $\{w_1,\ldots,w_m\}$.

<u>Definición</u>. A es <u>la matriz representando T respecto a las bases</u> <u>B</u> = $v_1,\ldots,v_n$ y <u>B'</u> = $\{w_1,\ldots,w_m\}$ y será denotada como $M_{B'}^{B}(T)$.

Un caso especial, es cuando W = V, o sea, T es una transformación de V en sí misma.

<u>Proposición 1</u>. Sea $T:V\to V$ la transformación identidad (id) y B, B' bases de V. Entonces,

$$M_{B'}^{B}(id)\ M_{B'}^{B}(id) = M_{B}^{B'}(id)\ M_{B'}^{B}(id) = I$$

donde I, representa la matriz identidad.

<u>Prueba</u>. Sea v un elemento de V y x su vector de coordenadas respecto a la base B. Entonces, el vector de coordenadas de T(v) relativo a B' es $M_{B'}^{B}(id)\,x$. De aqui, el vector de coordenadas de $T[T(v)]$ relativo a la base B es $M_{B'}^{B}(id)M_{B'}^{B}(id)x$ y ya que $T[T(v)] = v$ se sigue que $M_{B'}^{B}(id)\ M_{B'}^{B}(id)x=x$, i.e., $M_{B}^{B'}(id)\ M_{B'}^{B}(id) = I$. La otra parte de la demostración es similar.

Describiremos ahora con precisión como cambia la matriz representando una transformación lineal cuando se cambian las bases.

<u>Teorema 2</u>. Sea $T:V\to V$ y sean B, B' bases de V. Entonces, existe una matriz invertible N, tal que,

$$M_{B'}^{B}(T) = N^{-1}\ M_{B}^{B}(T)\ N$$

donde $N = M_{B}^{B'}(id)$ y N^{-1} representa la inversa de N.

<u>Prueba</u>. Sea v en V y x su vector de coordenadas respecto a la base B'. Entonces el vector de coordenadas de v respecto a la base B es $M_{B}^{B'}(id)x$ y por lo tanto el vector de coordenadas de T(v) respecto a la base B es $M_{B}^{B}(T)M_{B}^{B'}(id)x$. De aqui, el vector de coordenadas de T(v) respecto a la base B, es $M_{B'}^{B}(id)\ M_{B}^{B}(T)\ M_{B}^{B'}(id)x$, i.e.,

$$M_{B'}^{B'}(T) = M_{B'}^{B}(id)\, M_{B}^{B}(T)\, M_{B}^{B'}(id)$$

Haciendo $N = M_{B}^{B'}(id)$ y aplicando la proposición 1 anterior, el teorema queda demostrado.

En el resto del trabajo nos ocuparemos más ampliamente de las transformaciones lineales de un espacio vectorial en sí mismo.

2.2. Valores y vectores característicos

Definición. Sea T una transformación lineal de V en sí mismo. Si W es un subespacio de V tal que $T[W]$ es subconjunto (c) de W, entonces W es llamado un subespacio invariante de V bajo T. Generalmente, el problema de determinar las propiedades de T sobre V puede ser reducido al problema de determinar las propiedades de T sobre los subespacios invariantes. El caso más simple ocurre cuando el subespacio invariante W es de dimensión uno. Sea $\{\alpha_1\}$ una base para W. Entonces ya que $T(\alpha_1)$ está en W, existe un escalar λ_1 tal que $T(\alpha_1) = \lambda_1\alpha_1$. También, para cualquier α en W, $\alpha = a_1\alpha_1$ y por lo tanto $T(\alpha) = a_1T(\alpha_1) = a_1\lambda_1\alpha_1 = \lambda_1\alpha$. Vemos de este modo que λ_1 es característico del subespacio invariante W.

Definición. Un escalar λ es un valor característico de una transformación lineal $T:V\to V$ sobre el campo K, si existe un vector v en V diferente de cero tal que:

$$T(v) = \lambda v \qquad \ldots\ldots \qquad (1)$$

El vector correspondiente v recibe el nombre de vector característico de T.

La interpretación geométrica de un vector característico es que si se le aplica la transformación T, sólo cambió su medida y quizá el signo.

Note que la ecuación (1) es equivalente a la ecuación

$$[T - \lambda(id)]v = 0$$

Sabemos que si Ker $[T - \lambda(id)] \neq \{\theta\}$, entonces existe un vector v diferente de cero satisfaciendo dicha ecuación. Ahora bien, sea $\{\alpha_1,\ldots,\alpha_n\}$ cualquier base de V y A la matriz representando T con respecto a esta base.

<u>Definición</u>. El <u>polinomio</u> <u>característico</u> de la transformación lineal T es el determinante de la matriz A-xI, representado como $\det[A-xI]$.

Si $\{\alpha_1^1, \ldots, \alpha_n^1\}$ es otra base de V y A' es la matriz representando T con respecto a esta base, por el teorema 2 de la sección anterior,

$$A' = N^{-1} A N$$

relación que nos permitirá demostrar el siguiente resultado:

<u>Proposición 1</u>. $\det[A - xI] = \det[A' - xI]$

<u>Prueba</u>. Ya que $A' = N^{-1} A N$,

$$\det[A' - xI] = \det[N^{-1} A N - xI]$$

Pero $N^{-1}A N - xI = N^{-1}(A - xI)N$, de donde,

$$\det[A' - xI] = \det[N^{-1}(A - xI)N]$$

$$= (\det N^{-1}) (\det [A - xI])(\det N)$$

Es claro que $\det N^{-1} \det N = 1$ y por lo tanto concluimos que

$$\det \left[A' - xI\right] = \det \left[A - xI\right]$$

Este resultado nos muestra que el polinomio característico de la transformación T no depende de la base elegida para V. De este modo, sea

$$f(x) = \det \left[A - xI\right]$$

el polinomio característico de la transformación T y A la matriz representando T respecto a cualquier base elegida de V. Dicho polinomio puede ser escrito como

$$f(x) = k_n x^n + k_{n-1} x^{n-1} +, \ldots, + k_0$$

donde $k_n = (-1)^n$; $k_{n-1} = (-1)^{n-1} \sum_{i=1}^{n} a_{ii}$; y, el término constante k_0 es igual al determinante de A. Un teorema que tiene que ver con el polinomio característico y que nos será de mucha utilidad en la siguiente sección, es el teorema de Hamilton-Cayley.

Teorema 1. <u>(Hamilton-Cayley)</u>. Sea $T:V \to V$ una transformación lineal sobre el campo K. Sea f(x) el polinomio característico de T. Entonces $f(T) = \Theta$, donde f(T) es igual a $k_n T^n + k_{n-1} T^{n-1} +, \ldots, + k_0$ (id).

<u>Prueba</u>. Sea A la matriz que representa T respecto a la base B de V. Entonces, $f(A) = \Theta$. Sea x el vector de coordenadas de v en V respecto a B. Entonces $A^i x$ es el vector de coordenadas de $T^i(v)$ para $i=1, 2, \ldots, n$. Por lo tanto,

$$k_n A^n x +, \ldots, + k_0 x = \theta$$

es el vector de coordenadas de

$$k_n T^n(v) +, \ldots, + k_0 v$$

y de aquí

$$k_n T^n(v) +, \ldots, + k_0 v = \theta$$

Como v fue arbitraria se sigue que,

$$k_n T^n +, \ldots, + k_0 (id) = \theta$$

Esto termina la prueba.

Definición. La ecuación $m(x) = 0$ grado que T satisface es llamada ecuación mínima para T y $m(x)$ es llamado el polinomio mínimo de T. Ya que T satisface su polinomio característico el grado de $m(x)$ es menor o igual que n.

Definición. Sea $f(x)$ el polinomio característico de la transformación T y A la matriz representando dicha transformación. La ecuación $f(x) = 0$ es conocida como la ecuación característica de la transformación T.

Si existe un escalar λ, en el campo K del espacio V en cuestión, que sea solución de la ecuación característica, entonces, existe un vector y en K^n diferente de cero tal que $Ay = \lambda y$. Sabemos que si el campo es el de los números complejos, existe al menos un complejo λ que satisface dicha ecuación.

$$k_n A^n x +, \ldots, + k_0 x = 0$$

es el vector de coordenadas de

$$k_n T^n(v) +, \ldots, + k_0 v$$

y de aqui

$$k_n T^n(v) +, \ldots, + k_0 v = \theta$$

Como v fue arbitraria se sigue que,

$$k_n T^n +, \ldots, + k_0(id) = \theta$$

Esto termina la prueba.

Definición. La ecuación $m(x) = 0$ de menor grado que T satisface es llamada ecuación mínima para T y $m(x)$ es llamado el polinomio mínimo de T. Ya que T satisface su polinomio característico el grado de $m(x)$ es menor o igual que n.

Definición. Sea $f(x)$ el polinomio característico de la transformación T y A la matriz representando dicha transformación. La ecuación $f(x) = 0$ es conocida como la ecuación característica de la transformación T. Si existe un escalar λ, en el campo K del espacio V en cuestión, que sea solución de la ecuación característica, entonces, existe un vector y en K^n diferente de cero tal que $Ay = \lambda y$. Sabemos que si el campo es el de los números complejos, existe al menos un complejo λ que satisface dicha ecuación.

Un resultado que relaciona la ecuación característica con los valores característicos de la transformación T es el siguiente:

<u>Proposición 2.</u> Un escalar λ es un valor característico de T sí y sólo sí es solución de la ecuación característica de T.

<u>Prueba.</u> Sea A la matriz que representa T respecto a la base $\{v_1, \ldots, v_n\}$. Si λ es un valor característico de T, entonces, por definición existe v en V diferente de cero tal que

$$T(v) = \lambda v$$

Sea y el vector de coordenadas de v respecto a la base $\{v_1, \ldots, v_n\}$. Obviamente y es diferente de cero. Entonces Ay es el vector de coordenadas de T(v) respecto a la base $\{v_1, \ldots, v_n\}$. Pero $T(v) = \lambda v$, lo que implica que,

$$Ay = \lambda y$$

i.e.,

$$[A - \lambda I] y = 0$$

para $y \neq 0$. Ahora bien, esto significa que

$$\det [A - \lambda I] = 0$$

y por lo tanto λ es solución de la ecuación característica de T.

Por otra parte, suponga que λ es solución de la ecuación característica de T. Sea A la matriz representando T a la base $\{v_1,\ldots, v_n\}$. Entonces, existe y en K^n diferente de cero tal que $[A - \lambda I]y = 0$. Sin embargo, $A - \lambda I$ es la matriz representando $T - \lambda(id)$, por lo que si definimos $v = y_1v_1 +,\ldots,+ y_nv_n$, entonces las coordenadas de $[T- \lambda(id)](v)$ son precisamente $[A - \lambda I]y=0$. O sea, $[T-\lambda(id)](v) = 0$, i.e.,

$$T(v) = \lambda v$$

De este modo, hemos encontrado un vector v distinto de cero, tal que, $T(v) = \lambda v$ por lo que λ es un valor característico de T.

Esta proposición nos permite encontrar los valores y vectores característicos de T a través de la solución de su ecuación característica.

<u>Definición</u>. El <u>espacio característico</u> de T correspondiente a λ, denotado por $S(\lambda)$, es el conjunto de vectores característicos de T correspondientes a λ junto con el elemento neutro.

Es fácil demostrar que $S(\lambda)$ es un subespacio de V.

<u>Definición</u>. La dimensión de $S(\lambda)$ es llamada la <u>multiplicidad geométrica</u> de λ.

Hemos visto que λ es solución de la ecuación característica de $f(x) = 0$. De aqui, $x-\lambda$ es un factor de $f(x)$. Si $(x-\lambda)^k$ es un factor de $f(x)$ pero $(x-\lambda)^{k+1}$ no, λ es una raíz de $f(x) = 0$ de multiplicidad k.

Definición. k es llamada la multiplicidad algebraica de λ.

Generalmente, para una transformación lineal dada, se busca una base para la cual la matriz representando la transformación tenga forma simple, tanto como sea posible. Dicha forma, sería por ejemplo una matriz diagonal. La siguiente proposición nos dará bases para buscar condiciones en que dicha matriz diagonal pueda ser encontrada.

Proposición 3. Sean $\lambda_1, \ldots, \lambda_m$ valores característicos distintos de la transformación lineal T. Entonces cualquier conjunto $e_1, e_2, \ldots, e_m$ de vectores correspondientes es linealmente independiente.

Prueba. Suponga que los vectores característicos son linealmente dependientes. Entonces existe una combinación lineal de esos vectores tal que

$$\sum_{i=1}^{m} \alpha_i e_i = 0$$

y no todas las α_i son cero. De las posibles combinaciones lineales escoja la que tiene el número mínimo de coeficientes distintos de cero. Sin pérdida de generalidad suponga que esos coeficientes corresponden a los primeros k vectores característicos y que el primer coeficiente es uno. Esto es, la combinación lineal es de la forma

$$e_1 + \sum_{i=2}^{k} \alpha_i e_i = 0 \qquad \ldots \quad (1)$$

$\alpha_i \neq 0$ para $i=2,3,\ldots,k$. Si multiplicamos esta ecuación por la transformación T obtenemos

$$Te_1 + \sum_{i=2}^{k} \alpha_i \, Te_i = \theta$$

Usando el hecho de que los e_i son valores característicos, esta última ecuación es equivalente a

$$\lambda_1 e_1 + \sum_{i=2}^{k} \alpha_i \lambda_i e_i = \theta \qquad \ldots . \quad (2)$$

Multiplicando (1) por λ_1 y restándola de (2) obtenemos

$$\sum_{i=2}^{k} \alpha_i (\lambda_i - \lambda_1) e_i = \theta$$

Sin embargo, ésto, es una combinación lineal de sólo (k-1) términos, contradiciendo la definición de k como el valor mínimo posible.

<u>Proposición 4.</u> Una transformación lineal T puede ser representada por una matriz diagonal sí y sólo sí existe una base consistiendo de vectores característicos de T.

<u>Prueba.</u> Suponga que existe un conjunto de vectores característicos $\{\xi_1,\ldots,\xi_n\}$ linealmente independientes y que $\lambda_1,\ldots,\lambda_n$ son los valores característicos correspondientes. Entonces $T(\xi_i) = \lambda_i \xi_i$ y de este modo la matriz representando T con respecto a esta base, tiene la forma:

$$\begin{bmatrix} \lambda_1 & 0 & \cdot & \cdot & 0 \\ 0 & \lambda_2 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \lambda_n \end{bmatrix}$$

esto es, T es representada por una matriz diagonal. Contrariamente, si T es representado por una matriz diagonal, los vectores en esa base son vectores característicos.

No todas las transformaciones lineales pueden ser representadas por matrices diagonales, sin embargo, cuando consideramos como campo los números complejos, existe una forma de matriz conocida como la forma normal de Jordan en la cual todas las transformaciones lineales pueden ser representadas y las matrices diagonales son un caso particular de esta forma. Esto es ya una gran ventaja y de ello nos ocuparemos en la siguiente sección.

2.3 Forma normal de Jordan.

En esta sección suponemos que el campo de escalares es el de los números complejos. Sea T una transformación lineal de un espacio vectorial V en si mismo de dimensión finita. Si f(x) es el polinomio característico de T, obtenido a través de la matriz A representando T, dicho polinomio puede ser descompuesto en factores lineales,

$$f(x) = (x - \lambda_1)^{r_1} (x - \lambda_2)^{r_2} ,\ldots, (x - \lambda_p)^{r_p}$$

ya que el campo es el de los números complejos y donde $\lambda_i \neq \lambda_j$ para $i \neq j$ y r_i es la multiplicidad algebraica del valor característico λ_i. El polinomio mínimo, que puede ser demostrado divide al polinomio característico, es entonces de la siguiente forma:

$$m(x) = (x - \lambda_1)^{s_1} (x - \lambda_2)^{s_2} ,\ldots, (x - \lambda_p)^{s_p}$$

donde $1 \leq s_i \leq r_i$.

En la proposición 4 de la sección 2.2 trabajamos con una base consistiendo de vectores característicos. Sin embargo, dado que ahora trabajamos en el caso de que dicha base no exista, debemos buscar generalizar algunas nociones dadas en la sección anterior.

Definición. Para un valor característico λ fijo, se definen los subespacios

$$M^k = \text{Ker } [T - \lambda(id)]^k$$

$$W^k = \text{Im } [T - \lambda(id)]^k$$

De esta forma, $M^0 = \{\theta\}$ y M^1 es el espacio característico de T correspondiente a λ.

Mostraremos ahora que cada M^k y W^k es invariante bajo $T-\lambda(id)$.

<u>Proposición 1</u>. (a) $[T - \lambda(id)]M^{k+1} \subset M^k \subset M^{k+1}$

(b) $[T - \lambda(id)]W^k = W^{k+1} \subset W^k$

<u>Prueba</u>. Sea α en M^{k+1}. Esto implica que

$$[T - \lambda(id)]^{k+1}\alpha = [T - \lambda(id)]^k \, [T - \lambda(id)]\alpha = \theta$$

O sea, $[T - \lambda(id)]\alpha$ está en M^k. Por otra parte, si α esta en M^k, $[T - \lambda(id)]^k\alpha = \theta$, de donde,

$$[T - \lambda(id)]^{k+1}\alpha = [T - \lambda(id)] \, [T - \lambda(id)]^k\alpha = \theta$$

esto es $M^k \subset M^{k+1}$ lo que prueba (a).

Para probar (b), sea α en W^k. Entonces, existe x en V tal que $[T - \lambda(id)]^k x = \alpha$, lo que implica que,

$$[T - \lambda(id)]\alpha = [T - \lambda(id)]^{k+1}x$$

i.e., $[T - \lambda(id)]\alpha$ está en W^{k+1}. Si β está en W^{k+1}, existe x tal que $\beta = [T - \lambda(id)]^{k+1}x$, o sea,

$$\beta = [T - \lambda(id)] \, [T - \lambda(id)]^k x$$

donde $\alpha = [T - \lambda(id)]^k x$ está en W^k, lo que implica que $\beta = [T - \lambda(id)]\alpha$ está en $[T - \lambda(id)]W^k$. Esto demuestra que $[T - \lambda(id)]W^k = W^{k+1}$. Por otra parte si α está en W^{k+1}, existe x tal que $\alpha = [T - \lambda(id)]^{k+1}x = [T - \lambda(id)]^k \, [T - \lambda(id)]x$

de donde α está en W^k. Con ésto concluimos la parte (b) y la proposición queda demostrada.

Una consecuencia directa de la proposición 1 anterior, es la siguiente:

Proposición 2. Bajo T y cualquier polinomio en $T - \lambda(id)$, M^k y W^k son subespacios invariantes. Además, si μ es cualquier otro valor característico de T, M^k y W^k son invariantes bajo $T - \mu(id)$.

Prueba. La primera parte de la proposición se sigue de que si x está en W^k, entonces $[T - \lambda(id)]x$ está en W^k. Pero λx está en W^k ya que λ es un escalar y W^k es un subespacio, lo que implica que $T(x)$ está en W^k. Por lo tanto $T[W^k] \subset W^k$. La demostración para M^k es similar. Ahora bien, $[T-\lambda(id)]^n W^k = [T - \lambda(id)]^{n-1} [T - \lambda(id)]W^k$

$$\subset [T - \lambda(id)]^{n-1} W^k$$

$$\vdots$$

$$\subset W^k$$

Por otra parte si μ es cualquier vector característico de T y x es un elemento de W^k, μx está en W^k y $T(x)$ también. Esto implica que $[T - \mu(id)]x$ está en W^k. Por lo tanto $[T - \mu(id)]W^k \subset W^k$. Con ésto, la proposición queda demostrada.

Ya que $M^0 \subset M^1 \subset M^2 \subset, \ldots, \subset V$ y V es de dimensión finita, esta sucesión debe detenerse en algún punto. Sea t el índice más pequeño tal que $M^k = M^t$ para toda $k \geq t$, y denote M^t por $M_{(\lambda)}$. Sea m_k la dimensión de M^k; por el teorema 1 de la sección 2.1 la dimensión de W^k es $n-m_k$ y $W^k = W^t$ para toda $k \geq t$. Denote W^t como $W_{(\lambda)}$.

<u>Proposición 3</u>. V es la suma directa de $M_{(\lambda)}$ y $W_{(\lambda)}$, o sea

$$V = M_{(\lambda)} \oplus W_{(\lambda)}$$

<u>Prueba</u>. De la proposición 1 anterior, parte (b), $[T-\lambda(id)]W^t = W^{t+1} = W^t$ lo que muestra que $[T-\lambda(id)]$ es una transformación uno a uno y sobre, o sea, no singular sobre $W_{(\lambda)}$. Sea α cualquier vector en V. Entonces $[T-\lambda(id)]^t\alpha = \beta$ es un elemento de $W_{(\lambda)}$. Ya que $[T-\lambda(id)]$ es no singular sobre $W_{(\lambda)}$, también lo es $[T-\lambda(id)]^t$. Esto se sigue de que

$$\begin{aligned} [T-\lambda(id)]^t W^t &= [T-\lambda(id)]^{t-1}[T-\lambda(id)]W^t \\ &= [T-\lambda(id)]^{t-1} W^t \\ &\;\;\vdots \\ &= W^t \end{aligned}$$

Entonces existe un único vector γ en W^t tal que $[T-\lambda(id)]^t\gamma = \beta$.

Sea $\alpha = \gamma + \delta$. Esto implica que δ está en $M_{(\lambda)}$. De aquí $V = M_{(\lambda)} + W_{(\lambda)}$. Ya que $n = \dim V = \dim M_{(\lambda)} + W_{(\lambda)} = \dim M_{(\lambda)} + \dim W_{(\lambda)} - \dim M_{(\lambda)} \cap W_{(\lambda)}$ donde $\cap$ representa la intersección de los subespacios, se sigue que $\dim M_{(\lambda)} \cap W_{(\lambda)} = 0$ o sea la suma es directa.

Sean $\lambda_1, \ldots, \lambda_p$ los valores característicos distintos de T. Sea M_i una notación simple para $M_{(\lambda_i)}$ y W_i para $W_{(\lambda_i)}$.

<u>Proposición 4</u>. Para $\lambda_i \neq \lambda_j$, $M_i \subset W_j$.

<u>Prueba</u>. Suponga que α está en M_i y en el núcleo de $T-\lambda_j(id)$. Entonces

$$\left[\lambda_j(id) - \lambda_i(id)\right]^{t_i}\alpha = \{\left[T - \lambda_i(id)\right] - \left[T - \lambda_j(id)\right]\}^{t_i}\alpha$$
$$= \left[T-\lambda_i(id)\right]^{t_i}\alpha + \sum_{k=1}^{t_i} (-1)^k \binom{t_i}{k} \cdot$$
$$\cdot \left[T - \lambda_i(id)\right]^{t_i-k} \left[T - \lambda_j(id)\right]^k \alpha$$

El primer término es cero porque α está en M_i, y los otros son cero porque α está en el núcleo de $T - \lambda_j(id)$. Ya que $\lambda_j - \lambda_i \neq 0$, se sigue que $\alpha = 0$. Por la proposición 2 de esta sección, $\left[T - \lambda_j(id)\right]\left[M_i\right] \subset M_i$ lo que implica que $T - \lambda_j(id)$ es no singular sobre M_i y por lo tanto $\left[T - \lambda_j(id)\right]^{t_j}$ es no singular sobre M_i. De este modo M_i está contenida en el conjunto de imágenes bajo $\left[T - \lambda_j(id)\right]^{t_j}$, i.e., $M_i \subset W_j$.

Esta proposición nos permitirá demostrar el siguiente teorema, el más importante de este capítulo, pero antes, probaremos este lema.

<u>Lema 1</u>. Sea $q(T) = \left[T - \lambda_1(id)\right] \ldots \left[T - \lambda_p(id)\right]$ y $W = W_1 \cap \ldots \cap W_p$. Entonces $q(T)$ es no singular sobre W.

Prueba. Primero demostraremos que W es invariante bajo $q(T)$, i.e., $q(T)[W] \subset W$. Para ello, sea x en W. Entonces, por la proposición 2 de esta sección $[T - \lambda_p(id)]x$ está en W_1; $[T - \lambda_p(id)]x$ está en W_2 ;...; $[T - \lambda_p(id)]x$ está en W_p. Esto implica que $[T - \lambda_p(id)]x$ está en W. De manera similar, $[T - \lambda_{p-1}(id)]\ [T - \lambda_p(id)]x$ está en W. Procediendo de esta forma, $[T - \lambda_1(id)]\ \ldots\ [T - \lambda_p(id)]x$ está en W y así W es invariante bajo $q(T)$.

Por otra parte, sea $x \varepsilon W$ tal que $q(T)x = \theta$, i.e., $[T - \lambda_1(id)]$ $\ldots$ $[T - \lambda_p(id)]x = \theta$. Pero $[T - \lambda_2(id)]\ .\ .\ .\ [T - \lambda_p(id)]x$ está en W y de aquí en W_1. Como $[T - \lambda_1(id)]$ es no singular sobre W_1, esto implica que $[T - \lambda_2(id)]\ .\ .\ .\ [T - \lambda_p(id)]x=\theta$. Sin embargo, $[T - \lambda_2(id)]\ .\ .\ .\ [T - \lambda_p(id)]x$ está en W_2 y de nuevo implicamos que $[T-\lambda_3(id)]..[T-\lambda_p(id)]x=\theta$. Procediendo de manera similar, concluimos que $x=\theta$. De aquí, $q(T)$ es no singular sobre W.

Teorema 1. $V = M_1 \oplus M_2 \oplus ,\ldots, \oplus M_p$.

Prueba. Ya que $V = M_2 \oplus W_2$ y $M_2 \subset W_1$, tenemos que

$$\begin{aligned} V &= M_1 \oplus W_1 \\ &= M_1 \oplus (W_1 \cap V) \\ &= M_1 \oplus [W_1 \cap (M_2 \oplus W_2)] \\ &= M_1 \oplus [(W_1 \cap M_2) \oplus (W_1 \cap W_2)] \\ &= M_1 \oplus [M_2 \oplus (W_1 \cap W_2)] \end{aligned}$$

continuando de este modo, obtenemos

$$V = M_1 \oplus M_2 \oplus , \ldots , \oplus M_p \oplus (W_1 \cap \ldots \cap W_p)$$

Si probamos que $W = W_1 \cap \ldots \cap W_p = \{\theta\}$, el teorema queda demostrado. Por el lema anterior, $q(T)$ es no singular sobre W y en base a ésto fácilmente se sigue que $[q(T)]^k$ es también no singular sobre W para k arbitraria. Pero $q(x)$ contiene cada factor del polinomio característico $f(x)$, tal que, para k adecuada, $f(x)$ divide a $[q(x)]^k$. Esto es, $[q(x)]^k = g(x)f(x)$ para algún polinomio g. Por el teorema de Hamilton-Cayley (teorema 1 de la sección 2.2), $[q(T)]^k = g(T)f(T)=\theta$. Como $[q(T)]^k$ es no singular sobre W, esto implica que $W = \{\theta\}$ con lo cual el teorema queda demostrado.

Una consecuencia inmediata de este teorema es el siguiente corolario.

<u>Corolario 1</u>. $t_i = s_i$ para $i=1,\ldots,p$.

<u>Prueba</u>. Ya que $V = M_1 \oplus M_2 \oplus \ldots \oplus M_p$ y $[T - \lambda_i(id)]^{t_i} = \theta$ sobre M_i, se sigue que $[T - \lambda_1(id)]^{t_1} \ldots [T - \lambda_p(id)]^{t_p} = \theta$ sobre todo V. De este modo $(x-\lambda_1)^{t_1} \ldots (x-\lambda_p)^{t_p}$ es divisible por el polinomio mínimo y $s_i \leq t_i$.

Por otra parte, si para algún i, tenemos que $s_i<t_i$, entonces existe α en M_i tal que $[T - \lambda_i(id)]^{s_i}\alpha\neq\theta$. Para toda $\lambda_j\neq\lambda_i$, $[T - \lambda_j(id)]$ es no singular sobre M_i y también $[T-\lambda_j(id)]^k$ para cualquier k. De aquí, existe γ en M_i tal que $[T-\lambda_{i+1}(id)]^{s_{i+1}} \ldots [T-\lambda_p(id)]^{s_p}\gamma=\alpha$, i.e., $m(T)\gamma\neq\theta$,

lo que es una contradicción y por lo tanto $t_i = s_i$.

Retornamos ahora a la situación donde, para el valor característico λ, M^k es el núcleo de $[T - \lambda(id)]^k$ y W^k su imagen. En vista de lo demostrado en el corolario 1, $M^k = M^s$ para toda $k \geq s$. Por inducción podemos construir una base $\{\alpha_1,\ldots,\alpha_{m_s}\}$ de $M_{(\lambda)}$ tal que $\{\alpha_1,\ldots,\alpha_{m_k}\}$ es una base de M^k.

Procedemos ahora a modificar esta base. El conjunto $\{\alpha_{m_{s-1}+1},\ldots,\alpha_{m_s}\}$ consiste de aquellos elementos básicos de M^s que no están en M^{s-1}. Estos elementos no serán cambiados, pero por consistencia en la notación, cambiaremos sus nombres; sea $\alpha_{m_{s-1}+\nu} = \beta_{m_{s-1}+\nu}$. Ahora sea $[T - \lambda(id)]\beta_{m_{s-1}+\nu} = \beta_{m_{s-2}+\nu}$ y considere el conjunto $\{\alpha_1,\ldots,\alpha_{m_{s-2}}, \beta_{m_{s-2}+1},\ldots,\beta_{m_{s-2}+m_s-m_{s-1}}\}$.

<u>Proposición 5</u>. Este conjunto es linealmente independiente.

<u>Prueba</u>. Suponga que es linealmente dependiente. Entonces existe una combinación lineal no trivial en la cual al menos alguna de las β_i tiene un coeficiente distinto de cero ya que el conjunto $\{\alpha_1,\ldots,\alpha_{m_{s-2}}\}$ es linealmente independiente. Sea pues dicha combinación

$$\sum_{i=1}^{m_{s-2}} a_i\,\alpha_i + \sum_{\nu=1}^{m_s-m_{s-1}} b_{m_{s-2}+\nu}\,\beta_{m_{s-2}+\nu} = 0$$

con al menos una $b_{m_{s-2}+\nu} \neq 0$.

De esta ecuación, vemos que una combinación lineal no trivial de las β_i es una combinación lineal de las α_i y por lo tanto está en M^{s-2}, i.e.,

$$[T - \lambda(id)]^{s-2} \sum_{\nu=1}^{m_s-m_{s-1}} b_{m_{s-2}+\nu}\, \beta_{m_{s-2}+\nu} = \theta$$

Pero, entonces

$$[T - \lambda(id)]^{s-1} \sum_{\nu=1}^{m_s-m_{s-1}} b_{m_{s-2}+\nu}\, \alpha_{m_{s-1}+\nu} =$$

$$= [T - \lambda(id)]^{s-2} [T - \lambda(id)] \cdot$$

$$\cdot \sum_{\nu=1}^{m_s-m_{s-1}} b_{m_{s-2}+\nu}\, \alpha_{m_{s-1}+\nu}$$

$$= [T - \lambda(id)]^{s-2} \sum_{\nu=1}^{m_s-m_{s-1}} b_{m_{s-2}+\nu}\, \beta_{m_{s-2}+\nu}$$

$$= \theta$$

lo que implica que una combinación lineal no trivial de $\{\alpha_{m_{s-1}}, \ldots, \alpha_{m_s}\}$ está en M^{s-1}. Esto contradice la independencia lineal de $\{\alpha_1, \ldots, \alpha_{m_s}\}$. Por lo tanto, el conjunto $\{\alpha_1, \ldots, \alpha_{m_{s-2}}, \beta_{m_{s-2}+1}, \ldots, \beta_{m_{s-2}+m_s-m_{s-1}}\}$ es linealmente independiente.

Este subconjunto de M^{s-1} puede ser expandido a una base de M^{s-1} y usamos β para denotar esos elementos adicionales, si esos elementos son necesarios. De esta forma tenemos una nueva base $\{\alpha_1, \ldots, \alpha_{m_{s-2}}, \beta_{m_{s-2}+1}, \ldots, \beta_{m_{s-1}}\}$ de M^{s-1}.

Si ahora $[T - \lambda(id)]\ \beta_{m_{s-2}+\nu} = \beta_{m_{s-3}+\nu}$ y procedemos de manera similar obtenemos una nueva base $\{\alpha_1,\ldots,\alpha_{m_{s-3}+1},\beta_{m_{s-3}+1},\ldots,\beta_{m_{s-2}}\}$ de M^{s-2}.

Siguiendo este procedimiento obtenemos una base nueva $\{\beta_1,\ldots,\beta_{m_s}\}$ de $M_{(\lambda)}$ tal que $\{\beta_1,\ldots,\beta_{m_k}\}$ es una base de M^k y $[T - \lambda(id)]\beta_{m_k+\nu} = \beta_{m_{k-1}+\nu}$ para $k\geq 1$. Esta relación puede ser escrita en la forma

$$T\ \beta_{m_k+\nu} = \lambda\ \beta_{m_k+\nu} + \beta_{m_{k-1}+\nu} \qquad \text{para } k\geq 1$$

$$T\ \beta_{\nu} = \lambda\ \beta_{\nu} \qquad \text{para } \nu\leq m_1$$

Vemos que, en cierto sentido, $\beta_{m_k+\nu}$ es "casi" un vector característico.

Reordenando los vectores básicos como $\{\beta_1, \beta_{m_1+1},\ldots,\beta_{m_{s-1}+1}\}$, luego los vectores $\{\beta_2, \beta_{m_1+2},\ldots,\}$, etc., un nuevo orden en la base es obtenido. Con la base de $M_{(\lambda)}$ en este orden y suponiendo que $M_{(\lambda)}$ es toda V, la matriz representando T toma la forma

$$\begin{array}{r}
\\ \\ \\ \text{s renglones} \\ \\ \\ \\ \\ \\ \leq \text{s renglones} \\ \\ \\ \\
\end{array}
\left[\begin{array}{ccccccc|ccccc|c}
\lambda & 1 & 0 & \cdots & 0 & 0 & & & & & & & \\
0 & \lambda & 1 & \cdots & 0 & 0 & & & & & & & \\
0 & 0 & \lambda & \cdots & 0 & 0 & & & & & & & \\
\cdot & \cdot & \cdot & \cdots & \cdot & \cdot & & & \text{ceros} & & & & \text{ceros} \\
\cdot & \cdot & \cdot & \cdots & \cdot & \cdot & & & & & & & \\
0 & 0 & 0 & \cdots & \lambda & 1 & & & & & & & \\
0 & 0 & 0 & \cdots & 0 & \lambda & & & & & & & \\
\hline
& & & & & & & \lambda & 1 & \cdot & \cdot & 0 & \\
& & & & & & & 0 & \lambda & \cdot & \cdot & 0 & \\
& & \text{ceros} & & & & & \cdot & \cdot & \cdot & \cdot & \cdot & \text{ceros} \\
& & & & & & & \cdot & \cdot & \cdot & \cdot & \cdot & \\
& & & & & & & 0 & 0 & \cdot & \cdot & \lambda & \\
\hline
& & \text{ceros} & & & & & & \text{ceros} & & & & \text{etc.}
\end{array}\right]$$

<u>Teorema 2</u>. Sea $f(x) = (x - \lambda_1)^{r_1}, \ldots, (x - \lambda_p)^{r_p}$ el polinomio característico de T y $m(x) = (x - \lambda_1)^{s_1}, \ldots, (x - \lambda_p)^{s_p}$ su polinomio mínimo. Sea A la matriz representando T asociada a cualquier base de V. Entonces $A = N^{-1}J\,N$ donde J es una matriz con submatrices de la forma

$$B_i = \begin{bmatrix}
\lambda_i & 1 & 0 & \cdot & \cdot & 0 & 0 \\
0 & \lambda_i & 1 & \cdot & \cdot & 0 & 0 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & 0 & \cdot & \cdot & \lambda_i & 1 \\
0 & 0 & 0 & \cdot & \cdot & 0 & \lambda_i
\end{bmatrix}$$

a lo largo de la diagonal principal. Todos los otros elementos de J son cero. Para cada λ_i existe al menos una B_i de orden s_i. Todos los otros B_i correspondientes a λ_i son de orden menor que o igual a s_i. El número de B_i correpondientes a este λ_i es igual a la multiplicidad geométrica de λ_i. Además, la suma de las órdenes de todas las B_i correspondientes a λ_i es r_i. J es llamada la <u>forma normal de Jordan</u> correspondiente a T.

<u>Prueba</u>. Del teorema 1 de esta sección tenemos que $V = M_1 \oplus M_2 \oplus \ldots \oplus M_p$. En la discusión precedente a este teorema, mostramos que M_i tiene una base de tipo especial. Ya que V es la suma directa de las M_i, la unión de esas bases es una base para V. Esto muestra que una matriz J del tipo descrito en este teorema representa T y por el teorema 2 de la sección 2.1, para cualquier matriz A representando T con respecto a otra base de V, $A = N^{-1}J\ N$.

Ya que, por la proposición 1 de la sección 2.2, el polinomio característico de T no depende de la base elegida, el número de veces que $(x - \lambda_i)$ aparece como factor del polinomio característico obtenido usando J debe ser igual a r_i. Como ésto depende sólo del número de veces que λ_i aparece en la diagonal principal de J, la suma de las órdenes de las B_i correspondientes a λ_i debe ser r_i.

La forma normal de Jordan nos permitirá tratar en forma simple el concepto de convergencia en matrices, tema que desarrollamos en la siguiente sección.

2.4 Norma y convergencia en matrices

Una de las propiedades fundamentales de un espacio normado es el manejo del concepto de convergencia y la definición de límite. En el caso del espacio de las matrices es posible establecer el concepto de norma usando dos enfoques. El primer enfoque esta basado en la postulación de una función $||\cdot||$ en el espacio de matrices que satisfaga los requisitos de la norma. Un ejemplo típico es:

$$||B|| = \max \{ \sum_{J=1}^{n} |b_{ij}| \ ; \ i=1,\ldots,m\}$$

donde B pertenece al espacio de matrices reales (ó complejas) de orden mxn. Como puede esperarse dicha función satisface los postulados de norma, esto es,

a. $||B|| \geq 0$ y $||B||=0$ sí y sólo sí $B = 0$

b. $||\lambda B|| = |\lambda| \ ||B||$ $\qquad \lambda \varepsilon R \ (\lambda \varepsilon C)$

c. $||A+B|| \leq ||A|| + ||B||$

donde a es la condición de no-negatividad; b, la condición de homogeneidad; y c, la desigualdad del triangulo. Asimismo, para esta norma, en particular, se satisface la desigualdad de Schwartz:

d. $||AB|| \leq ||A|| \ \ ||B||$

donde A es una matriz real (o compleja) de orden pxm y B es otra matriz real (o compleja) de orden mxn.

El otro enfoque usado para establecer el concepto de norma esta basado en la consideración de que una matriz A de orden mxn es una transformación lineal de espacio R^m a R^n. Supondremos que m=n para facilitar la discusión, sin embargo, las propiedades y resultados que se establecen son ciertos para $m \neq n$. En este caso, la manera de establecer el concepto de norma de A queda dado por la expresión

$$|||A||| = \sup \{ ||Ax|| / ||x|| \ ; \ 0 \neq x \varepsilon R^n \}$$

donde $||\cdot||$ es una norma en el espacio R^n. Pero, las propiedades del espacio R^n y la definición de norma de A permiten establecer las siguientes expresiones equivalentes:

$$|||A||| = \max \{ ||Ax|| / ||x|| \quad ; \quad 0 \neq x \varepsilon R^n \}$$

$$= \max \{ ||Ax|| \quad ; \quad ||x|| \leq 1 \}$$

$$= \max \{ ||Ax|| \quad ; \quad ||x|| = 1 \}$$

Conviene señalar que el cálculo de la norma de A depende del tipo de norma usada en R^n. Asimismo, es necesario establecer formas concretas del cálculo de la norma de A que sean fáciles de manipular, pues hasta ahora lo único que puede garantizarse es que $|||\cdot|||$ satisface los postulados de norma.

En relación a la representación de la norma $|||\cdot|||$ en forma explícita se tiene lo siguiente

<u>Lema 1</u>. Sea M el espacio de matrices de orden nxn.

$$|||A||| = \max \{ \sum_{i=1}^{n} |a_{ij}| \ ; \ j=1,..,n \}$$

si la norma en R^n es $||x||_1 = \sum_{i=1}^{n} |x_i|$. Asimismo

$$|||A||| = \max \{ \sum_{j=1}^{n} |a_{ij}| \ ; \ i=1,...,n\}$$

si la norma en R^n es $||x||_\infty = \max \{|x_i| \ ; \ i=1,...,n\}$.

<u>Prueba</u>. Sea $x \varepsilon R^n$ tal que $||x||_1 = \sum_{i=1}^{n} |x_i| = 1$. Entonces

$$||Ax||_1 = \sum_{i=1}^{n} | \sum_{j=1}^{n} a_{ij}x_j |$$

$$\leq \sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}| \, |x_j|$$

$$\leq \sum_{j=1}^{n} |x_j| \ . \ (\sum_{i=1}^{n} |a_{ij}|)$$

$$\leq \max \{ \sum_{i=1}^{n} |a_{ij}| \ . \ j=1,..,n\}$$

De donde $||A|| \leq \max \{ \sum_{i=1}^{n} |a_{ij}| \ ; \ j=1,...,n$. Recíprocamente, si hacemos $x = e_j$ $(j=1,...,n)$ donde e_j es vector de ceros, excepto en la posición j en que tiene coeficiente igual a uno. note que Ae_j es el j-ésimo vector columna de A. Asimismo, $||e_j|| = 1$ y $||Ae_j||_1 = \sum_{i=1}^{n} |a_{ij}|$. Esto demuestra que

$$|||A||| = \max \{ \sum_{i=1}^{n} |a_{ij}| \ ; \ j=1,...,n\}$$

La segunda parte es similar y se pide al lector.

Considere el espacio de matrices de orden mxn cuyos elementos están definidos en el conjunto de números complejos. Suponga que en dicho conjunto se define la norma

$$||B|| = \max \{ \sum_{j=1}^{n} |b_{ij}| \; ; \; i = 1,...,m \}$$

Se dice que una sucesión de matrices $\{A_m\}$, converge a una matriz A si lim $||A_m - A|| = 0$ y se denota por $A = \lim A_m$. Asimismo, se dice que una serie de matrices $\sum_{i=0}^{\infty} B_i$ converge a una matriz S si la sucesión de matrices $\{S_k\}$ donde $S_k = \sum_{i=0}^{k} B_i$ converge a S y se denota por $S = \sum_{i=0}^{\infty} B_i$.

Proposición 1: Una sucesión de matrices $\{A_m\}$ converge a la matriz A sí y sólo si cada elemento (i,j) de la sucesión $\{A_m\}$ converge al elemento (i,j) de A, esto es, $\lim_{m} (A_m)_{ij} = A_{ij}$.

Prueba. Es inmediata

La equivalencia de convergencia puntual y convergencia en el sentido de la norma permite generalizar la mayoría de las propiedades de límites para escalares y números complejos.

Proposición 2: Suponga que $\{A_m\}$ y $\{B_m\}$ son sucesiones de matrices que convergen a las matrices A y B, respectivamente. Entonces $\{A_m B_m\}$ converge a AB y $\{\alpha A_m + \beta B_m\}$ converge a la matriz $\alpha A + \beta B$ para cualquier α, β complejo.

Prueba. Fije un elemento (i,j) de $(A_m B_m)$. Entonces

$$(A_m B_m)_{ij} = \sum_{k=1}^{n} (A_m)_{ik} (B_m)_{kj}$$

Sin embargo, por tratarse de sucesiones de números complejos, se tiene inmediatamente que

$$\lim (A_m B_m)_{ij} = \sum_{k=1}^{n} \lim (A_m)_{ik} \cdot \lim (B_m)_{kj}$$

$$= \sum_{k=1}^{n} A_{ik} B_{kj} = (AB)_{ij}$$

y $\lim_m A_m B_m = AB$. La prueba restante es análoga.

Sea A una matriz compleja de orden nxn. El conjunto de valores característicos de A se denota por $\sigma(A)$ y el máximo valor absoluto de estos elementos se denomina el radio espectral de A, denotándose por $|\sigma(A)|$. Por otra parte una matriz J de orden sxs se dice Jordan si

$$J = \begin{bmatrix} \lambda & 1 & & \\ 0 & \lambda & \ddots & \\ \vdots & & \ddots & 1 \\ & & & \lambda \end{bmatrix}$$

o bien $J = \lambda I + E$ donde es un número complejo; I es la matriz identidad de orden sxs; y E es una matriz con elementos igual a uno arriba de la diagonal principal y cero en otras posiciones.

Lema 2. Sea J una matriz Jordan de orden sxs. Entonces $0 = \lim J^n$ sí y sólo si $|\lambda|<1$.

Prueba. Es sencillo verificar que $E^s = 0$. Asimismo

$$J^n = (\lambda I+E)^n = \sum_{k=0}^{n} \binom{n}{k} \lambda^{n-k} E^k$$

De donde, si $n \geq s$ la fórmula anterior se reduce a

$$J^n = \sum_{k=0}^{s-1} \binom{n}{k} \lambda^{n-k} E^k$$

Sin embargo, si fijamos k $(0 \leq k \leq s-1)$ y comparamos los coeficientes de E^k en las matrices J^n y J^{n+1} se tiene que dicha relación es igual a

$$q = \frac{\binom{n+1}{k} \lambda^{n+1-k}}{\binom{n}{k} \lambda^{n-k}} = \frac{n+1}{n+1-k} \lambda$$

que implica $|q|<1$ sí y sólo si $|\lambda|<1$ y n es suficientemente grande. De esta relación podemos concluir que cada uno de los coeficientes de E^k converge a cero cuando n tiende infinito sí y sólo si $|\lambda|<1$. Equivalentemente $0 = \lim J^n$ si y sólo si $|\lambda|<1$ y la prueba termina.

Teorema 1. Suponga que B es una matriz cuadrada compleja. Entonces, los postulados siguientes son equivalentes:

a. $|\sigma(B)| < 1$

b. $\lim B^N = 0$

c. $||B^N|| < 1$ para algún $N \geq 1$

d. $\sum_{N=0}^{\infty} B^N$ es absolutamente convergente

Prueba. El teorema de Jordan implica que $B = QJQ^{-1}$ donde Q es una matriz no-singular y J es de la forma

$$J = \begin{bmatrix} J_1 & & & & \\ & \cdot & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & J_n \end{bmatrix}$$

Asimismo, cada $J_i (i=1,...,n)$ es una matriz de Jordan, i.e., $J_i = \lambda_i I + E$, donde cada una de las matrices es de orden $n_i x n_i$ y λ_i pertenece a $\sigma(B)$. Asimismo, se observa que $B^N = QJ^NQ^{-1}$ para toda N y se cumple que

$$\begin{aligned} ||B^N|| &= ||QJ^NQ^{-1}|| \\ &\leq ||Q|| \quad ||J^N|| \quad ||Q^{-1}|| \\ &= ||Q|| \quad ||Q^{-1}|| \quad ||J^N|| \end{aligned}$$

donde $||.||$ es la norma en el espacio de matrices. De aquí se concluye que $0 = \lim B^N$ sí y sólo si $0 = \lim J^N$ (pues también $J^N = Q^{-1}B^NQ$). Sin embargo, $0 = \lim J^N$ sí y sólo si $0 = \lim J_i^N$ para toda $i = 1,\ldots,n$. Por otra parte, el Lema 2 demuestra que $0 = \lim J_i^N$ sí y sólo si $|\lambda_i| < 1$. De donde se concluye que $0 = \lim B^N$ sí y sólo si $|\sigma(B)| < 1$.

La prueba de b implica c es inmediata. Para demostrar que c implica d observe que: $\sum_{k=0}^{\infty} B^k$ es absolutamente convergente sí y sólo si

$$\sum_{k=0}^{\infty} ||B^k|| < \infty$$

Sin embargo, se observa que

$$\begin{aligned}\sum_{k=0}^{\infty} ||B^k|| &= \sum_{j=0}^{\infty} \sum_{i=0}^{n-1} ||B^{jN+i}|| \\ &\leq \sum_{j=0}^{\infty} \sum_{i=0}^{n-1} ||B^N||^j \; ||B^i|| \\ &= \sum_{i=0}^{n-1} ||B^i|| \sum_{j=0}^{\infty} ||B^N||^j < \infty\end{aligned}$$

si $||B^N|| < 1$ para algún $N>1$. Finalmente, se observa que d implica b pues la convergencia absoluta de la serie $\sum_{N=0}^{\infty} B^N$ equivale a $\sum_{N=0}^{\infty} ||B^N|| < \infty$ y esto implica que $||B^N||$ converge a cero cuando N tiende a infinito. Equivalente a $0 = \lim B^N$ y la prueba termina.

Corolario 1. Si cualquiera de las equivalencias del teorema se cumple, entonces la inversa de I-B existe y es igual a

$$(I - B)^{-1} = \sum_{k=0}^{\infty} B^k$$

3. Procesos de decisión markovianos con descuento.

Los procesos de decisión markovianos son modelos matemáticos simples de interés no sólo para investigadores de operaciones y matemáticos, sino también para economistas e ingenieros. Estos procesos poseen resultados elegantes y en este capítulo nos ocuparemos de uno de los tipos de procesos de decisión markovianos más usuales, específicamente los discretos con descuento. El propósito de describir y analizar las principales propiedades de estos procesos es establecer el marco de referencia para la determinación de políticas de asignación de agua en sistemas de recursos hidráulicos.

Este capítulo se desarrolla como sigue: primero se establecen los conceptos y definiciones básicas de un proceso markoviano. A continuación, se plantea el problema general usando un lenguaje más formal, lo que permite enfatizar el concepto de estrategias estacionarias. En la sección 3 se demuestra la existencia de estrategias estacionarias que son óptimas y se establece un resultado que permite proponer un algoritmo para su determinación. Dicho algoritmo, debido a Howard, se presenta en la sección 4. Finalmente, se presenta una prueba alternativa del algoritmo de Howard usando los resultados del capítulo 2.

3.1 Conceptos y definiciones

Para entender lo que es un proceso markoviano es necesario definir antes lo que es una cadena de Markov.

<u>Definición</u>. Una <u>cadena</u> de <u>Markov</u> es una sucesión de variables aleatorias X_0, $X_1, \ldots,$ tal que la distribución condicional de X_{n+1}, dadas $X_0, X_1, \ldots, X_n$ depende solamente del valor de X_n, pero no de los valores X_0, $X_1, \ldots, X_{n-1}$, i.e., para toda i,h, ...,j y n,

$$P_r \{X_{n+1} = j \mid X_0 = h, \ldots, X_n = i\} = P_r\{X_{n+1} = j \mid X_n = i\}$$

En otras palabras, para decir como se desarrolla el proceso después del tiempo n, sólo necesitamos saber en dónde se encuentra el tiempo n. Dicha propiedad es conocida como la propiedad markoviana. Usualmente,

$$P_r \{X_{n+1} = j \mid X_n = i\}$$

se representa como $p_{ij}(n,n+1)$ y para facilitar el lenguaje diremos que el sistema está en el <u>estado</u> i al tiempo n si $X_n = i$. De este modo, $p_{ij}(n,n+1)$ representa la probabilidad de pasar del estado i al estado j en una <u>transición</u>. Trabajaremos exclusivamente con cadenas de Markov en que el número de estados es finito y diremos que el conjunto de estados es discreto digamos N.

Las probabilidades de transición asociada con una cadena de Markov, son consideradas más convenientemente, como los elementos de una matriz.

Una matriz P de orden N x N se denomina <u>matriz de transición</u> ó <u>estocástica</u> si sus elementos son no negativos y la suma de los elementos de cada hilera es igual a uno, específicamente, si

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1N} \\ p_{21} & p_{22} & \cdots & p_{2N} \\ - & - & - & - \\ - & - & - & - \\ p_{N1} & p_{N2} & & p_{NN} \end{bmatrix}$$

donde se cumple que

a. $p_{ij} \geq 0 \qquad i,j = 1,\ldots,N$

b. $\sum_{j=1}^{N} p_{ij} = 0 \qquad i = 1,\ldots,N$.

Es claro, que todos los elementos de una matriz P asociados con una cadena de Markov [i.e., una matriz cuyos elementos son $p_{ij}(n,n+1)$], independientemente del tiempo n en que nos encontremos, cumplen con esas propiedades. Por otra parte, es conveniente puntualizar que si P y Q son matrices de transición, PQ también es una matriz de transición. De aquí, por inducción, P^m es una matriz de transición, donde m es un entero positivo cualquiera.

Estamos ahora en posibilidad de entender lo que es un proceso markoviano discreto.

<ins>Definición</ins>. Un <ins>proceso de decisión markoviano discreto</ins> con N estados es un proceso estocástico que puede ser descrito a través de una cadena de Markov, en donde se tienen asociados ciertos beneficios al pasar de un estado a otro. Se dice que

el proceso es de decisión, porque los beneficios y las probabilidades de transición de un estado a otro dependen de la decisión, que sobre distintas alternativas se elija.

Bajo esta situación, nos preguntamos: ¿Cuál debería ser la estrategia, o sea la sucesión de decisiones en cada etapa, que debemos seguir de modo tal que hagamos máximo el beneficio esperado total si empezamos en algún estado? Si el horizonte de planeación, o sea el número de etapas,es finito, el problema puede ser formulado directamente a través de programación dinámica, y una estrategia óptima puede ser obtenida por medio de una relación de recurrencia. El problema interesante surge cuando el número de etapas es infinito, en cuyo caso, no podemos aplicar programación dinámica porque las estrategias consideradas son infinitas. Además, el beneficio esperado total podría diverger (en el sentido matemático).

Para atacar dicha problemática, introduciremos un factor de descuento $\beta (0 \leq \beta < 1)$, por eso se dice que el proceso markoviano es con descuento, que representa qué tan deseable es una unidad monetaria en el presente. Con ésto, el beneficio esperado total siempre converge, y nuestra tarea será desarrollar un algoritmo que busque la estrategia que lo hace máximo. Cabe recordar que el número de estrategias es infinito y por ello definiremos de entre esas estrategias un tipo especial, que vienen a jugar el mismo papel que juegan los puntos extremos en la programación lineal; el problema se reduce a elegir entre esas estrategias, cuyo número es finito, la óptima.

3.2 Planteamiento del problema

Considere un sistema cuyo conjunto de estados es finito. Sea S dicho conjunto formado por los estados etiquetados por los enteros $i = 1,2,\ldots,N$, i.e., $S = \{1,2,\ldots,N\}$. Para cada i en S, tenemos un conjunto finito K_i de acciones (o alternativas) etiquetadas por los enteros $k = 1,2,\ldots,K_i$. El conjunto de políticas es denotado por el producto cartesiano de cada conjunto de políticas, esto es, $K = K_1 \times K_2 \times \ldots \times K_N$. Seguidamente, considere un problema de decisión en el que periódicamente observamos uno de los estados y realizamos una acción. Cuando el sistema está en el estado i en S en el tiempo n y realizamos una acción k en K_i, suceden dos cosas:

(1) El sistema obedece la ley de probabilidad $p_{ij}^k(n,n+1)$ (j en S) en el período siguiente, donde $p_{ij}^k(n,n+1)$ es la probabilidad de transición de que el sistema esté en el estado j en el siguiente período dado que el sistema se encuentra en el estado i y una acción k es hecha.

(2) Obtenemos el beneficio inmediato R_{ij}^k de pasar del estado i al estado j realizando la acción k.

Como consecuencia de estos dos sucesos, obtenemos un beneficio inmediato esperado, que denotaremos como

$$r_i^k = \sum_{j \text{ en } S} p_{ij}^k(n,n+1)\, R_{ij}^k$$

Aquí, suponemos que el beneficio r_i^k es acotado para toda i en S y k en K_i.

Como dijimos antes, el proceso es con descuento, siendo $\beta (0 \leq \beta < 1)$ el factor de descuento, considerado como el recíproco de la unidad más la tasa de interés (t.i.), i.e.,

$$\beta = \frac{1}{1+t.i.}$$

Con objeto de plantear nuestra problemática, en un lenguaje más adecuado, daremos las siguientes definiciones.

Definición. $F = \{f \mid f : S \longrightarrow K,\ f \text{ función}\}$.

Ya que S y K son conjuntos finitos, F es un conjunto finito.

Definición. Sean f ó g funciones en F. Una estrategia Π es definida por una sucesión $\{f_n,\ n=1,2,\ldots\}$ y puede ser escrita como

$$\Pi = \{f_1,\ f_2,\ldots,f_n,\ldots\}$$

donde f_n es el vector de decisión para cada estado en el período n, i.e., $f_n(i)$, el elemento i-ésimo de f_n es una acción en el estado i en S en el período n.

Una estrategia puede ser defasada un período ó varios períodos y en ese caso escribimos (g,Π) ó (g^n,Π) respectivamente, donde $(g,\Pi) = (g,f_1,f_2,\ldots)$ y g en F y, similarmente,

$$(g^n,\Pi) = (\underbrace{g,g,\ldots,g}_{n},f_1,f_2,\ldots)\ .$$

También, una estrategia puede ser independiente del período y sólo depender del estado en que el sistema se encuentra.

Definición. Sea f en F. Una estrategia $(f,f,\ldots,f,\ldots)$ es llamada una estrategia estacionaria y es denotada por f^∞.

Por otra parte, escribiremos la matriz de probabilidades de transición en n pasos como

$$P_n(\Pi) = P(f_1)\ P(f_2)\ \ldots\ P(f_n) \qquad \text{para } n = 1,2,\ldots$$

donde $P(f_n)$ es una matriz de transición de orden N x N cuyo elemento i-jésimo es p_{ij}^k , $k = f_n(i)$ en K_i. Para $n = 0$, definimos $P_0(\Pi) = I$ y para cualquier f en F, escribimos el N x 1 vector de beneficios inmediatos esperados $r(f)$ cuyo i-ésimo elemento es r_i^k , $k=f(i)$ en K_i. Bajo la notación definida arriba, tenemos que el vector N x 1 de beneficio total esperado es:

$$V_\beta(\Pi) = \sum_{n=0}^{\infty} \beta^n\ P_n(\Pi)\ r(f_{n+1})$$

Este vector es acotado y para ver ésto, sea $r^U = \max_{i,k} r_i^k$, $r^L = \min_{i,k} r_i^k$. Entonces

$$\frac{r^L}{(1-\beta)}\ 1 \ \leq\ V_\beta(\Pi)\ \leq \frac{r^U}{(1-\beta)}\ 1$$

donde $\underset{\sim}{1}$ es el vector N x 1 cuyos elementos son la unidad.

Además, si $\Pi = f^{\infty}$, entonces

$$V_\beta(\Pi) = \sum_{n=0}^{\infty} \beta^n P^n(f) r(f)$$

converge. Para ver ésto, sea $||\cdot||$ la norma en el espacio de matrices definida como

$$||A|| = \max \{ \sum_{j=1}^{n} |a_{ij}| \quad ; \quad i=1,\ldots,n\}$$

y considere la matriz $\beta P(f)$, entonces

$$|| \beta P(f) || = \beta$$

Pero $0 \leq \beta < 1$ lo que implica, por el teorema 1 de la sección 2.4, que

$$\sum_{n=0}^{\infty} \beta^n P^n(f)$$

es absolutamente convergente. De este modo, $V_\beta(\Pi)$ es convergente para cualquier estrategia estacionaria y, por el corolario 1 de la sección 2.4,

$$V_\beta(\Pi) = [I - \beta P(f)]^{-1} r(f)$$

Como consecuencia de lo que acabamos de discutir, si P es una matriz de transición cualquiera y β es un escalar tal que $|\beta|<1$ entonces la inversa de $I - \beta P$ existe. De esta forma podemos enunciar este hecho como:

<u>Proposición 1</u>. Si P es una matriz de transición y β es un escalar tal que $|\beta|<1$, entonces $[I - \beta P]^{-1}$ existe y es igual a $\sum_{n=0}^{\infty} \beta^n P^n$.

Ahora definimos una función $L(f)$ que mapea cualquier vector $N\times 1$ w en $L(f)w = r(f) + \beta P(f)w$. De este modo,

$V_\beta(f,\Pi) = L(f)V_\beta(\Pi)$. También definimos, para cualesquiera vectores columna w_1, w_2, $w_1 \geq w_2$ si cada elemento de w_1 no es menor que el correspondiente elemento de w_2, y escribimos $w_1 > w_2$ si $w_1 \geq w_2$ y $w_1 \neq w_2$.

Definición. Una estrategia Π^* es llamada β-óptima si $V_\beta(\Pi^*) \geq V_\beta(\Pi)$ para toda Π, donde $\beta (0 \leq \beta < 1)$ es fija.

Esta definición significa que una estrategia óptima es alcanzada simultáneamente para cada estado, hecho que no es trivial como mostraremos más adelante.

Podemos formular ahora muestra problemática como: "encontrar la estrategia Π^* que maximiza el beneficio total esperado $V_\beta(\Pi^*)$ para una β fija".

3.3 Existencia de estrategias estacionarias.

Nos prepararemos para demostrar la existencia de una estrategia estacionaria que es β-óptima.

Lema 1. $L(f)$ es monótona; esto es, $w_1 \geq w_2$ implica que $L(f)w_1 \geq L(f)w_2$.

Prueba: $L(f)w_1 - L(f)w_2 = \beta P(f)(w_1 - w_2) \geq 0$ si $w_1 \geq w_2$.

Con este lema, probaremos los siguientes teoremas.

Teorema 1. $V_\beta(\Pi^*) \geq V_\beta(g,\Pi^*)$ para toda g en F implica que Π^* es β-óptima.

<u>Prueba</u>: Nuestra hipótesis es que $L(g)\ V_\beta(\Pi^*) \leq V_\beta(\Pi^*)$ para toda g en F. Para cualquier estrategia

$\Pi = (f_1, f_2, \ldots, f_n, \ldots)$, haciendo $g = f_n$, tenemos $L(f_n)\ V_\beta(\Pi^*) \leq V_\beta(\Pi^*)$. Aplicando el operador monótono $L(f_1) \ldots L(f_{n-1})$, tenemos

$$L(f_{n-1})\ L(f_n)\ V_\beta(\Pi^*) \leq L(f_{n-1})\ V_\beta(\Pi^*) \leq V_\beta(\Pi^*)$$

$$L(f_{n-2})\ L(f_{n-1})\ L(f_n)\ V_\beta(\Pi^*) \leq L(f_{n-2})\ V_\beta(\Pi^*) \leq V_\beta(\Pi^*)$$

$$\vdots$$

$$L(f_1)\ .\ .\ .\ L(f_n)\ V_\beta(\Pi^*) \leq V_\beta(\Pi^*)$$

para toda n. Pero

$$L(f_1) \ldots L(f_n) V_\beta(\Pi^*) = \sum_{i=0}^{n-1} \beta^i\ P_i(\Pi)\ r(f_{i+1}) + \beta^n P_n(\Pi) V_\beta(\Pi^*)$$

y tomando el límite cuando n tiende a infinito,

$$\lim_{n \to \infty} L(f_1) \ldots L(f_n) V_\beta(\Pi^*) = V_\beta(\Pi)$$

de donde

$$V_\beta(\Pi) \leq V_\beta(\Pi^*)$$

para toda Π.

Este teorema nos dice que si existe una estrategia Π^* tal que defasándola un período e introduciendo una función de decisión cualquiera, se cumple que la estrategia Π^* tiene un valor esperado total mayor o igual que la estrategia defasada, entonces Π^* es óptima.

<u>Teorema 2.</u> $V_\beta(f,\Pi) > V_\beta(\Pi)$ implica $V_\beta(f^\infty) > V_\beta(\Pi)$.

<u>Prueba:</u> La hipótesis es que $L(f)V_\beta(\Pi) > V_\beta(\Pi)$. Aplicando el operador monótono $L(f)$ repetidamente, obtenemos

$$L(f)\ L(f)\ V_\beta(\Pi) \geq L(f)\ V_\beta(\Pi) > V_\beta(\Pi)$$

$$L(f)\ L(f)\ L(f)V_\beta(\Pi) \geq L(f)V_\beta(\Pi) > V_\beta(\Pi)$$

$$\vdots$$

$$L^n(f)\ V_\beta(\Pi) > V_\beta(\Pi)$$

para $n > 0$. Pero

$$L^n(f)V_\beta(\Pi) = \sum_{i=0}^{n-1} \beta^i P^i(f) r(f_{i+1}) + \beta^n P^n(f) V_\beta(\Pi)$$

y tomando límite cuando n tiende a infinito

$$\lim_{n \to \infty} L^n(f)V_\beta(\pi) = V_\beta(f^\infty)$$

de donde

$$V_\beta(f^\infty) \geq V_\beta(\Pi)$$

Ahora bien, suponga

$$V_\beta(f^\infty) = V_\beta(\Pi)$$

Por hipótesis, $L(f)\ V_\beta(\Pi) > V_\beta(\Pi)$ de donde

$$L(f)\ V_\beta(f^\infty) > V_\beta(f^\infty)$$

Pero $L(f)\ V_\beta(f^\infty) = V_\beta(f^\infty)$ y por lo tanto

$$V_\beta(f^\infty) > V_\beta(f^\infty)$$

que es una contradicción, lo que implica que

$$V_\beta(f^\infty) > V_\beta(\Pi)$$

La consecuencia de estos teoremas, es el teorema principal de esta sección.

Teorema 3. Tome cualquier f en F. Para cada i en S, denote G(i,f) el conjunto de todas las k en K_i para las cuales

$$r_i^k + \beta \sum_{j \text{ en } S} p_{ij}^k v_j > v_i$$

donde v_i es el i-ésimo elemento de $V_\beta(f\infty)$. Si G(i,f) es vacío para toda i en S, entonces f∞ es β-óptima. Para cualquier g en F tal que a) g(i) en G(i,f) para alguna i y b) g(i) = f(i) cuando g(i) no está en G(i,f) tenemos que $V_\beta(g\infty) > V_\beta(f\infty)$.

Prueba: El i-ésimo elemento de $V_\beta(g,f\infty)$ es $r_i^k + \beta \sum_{j \in S} p_{ij}^k v_j$ donde k = g(i). Esto puede exceder v_i si y sólo si g(i) está en G(i,f) y será igual a v_i si g(i) = f(i). De este modo si G(i,f) es vacío para toda i en S, entonces $V_\beta(f\infty) \geq V_\beta(g,f\infty)$ para toda g en F, lo que implica que f∞ es β-óptima (por el teorema 1). Mientras que, para cualquier g en F satisfaciendo (a) y (b), tenemos que $V_\beta(g,f\infty) > V_\beta(f\infty)$, de donde $V_\beta(g\infty) > V_\beta(f\infty)$ (por el teorema 2).

Una consecuencia directa de este teorema es el siguiente corolario, objeto de esta sección.

Corolario 1. Existe una estrategia β-óptima que es estacionaria.

Prueba. De acuerdo al teorema anterior, si tomamos cualquier estrategia estacionaria f∞ suceden dos cosas:

I) f^∞ es β-óptima [caso $G(i,f)$ vacío para toda i en S].

II) Existe una estrategia estacionaria g^∞ que mejora el valor esperado total [caso $G(i,f)$ no vacío para alguna i en S].

Ya que el número de estrategias estacionarias es finito, existe una que no puede ser mejorada en un número finito de iteraciones; de aquí debe ser β-óptima.

3.4 El método de mejoramiento de políticas

Este método, basado en los teoremas de la sección anterior, fue desarrollado primero por Howard y por ello es llamado a veces el algoritmo de iteración de políticas de Howard. El algoritmo tiene dos partes, como sigue:

I. Determinación de valores. Tome cualquier f en F. Resuelva.

$$v_i = r_i^k + \beta \sum_{j \in S} p_{ij}^k \, v_j$$

para v_i (i en S), donde $k = f(i)$ corresponde a la estrategia seleccionada f^∞.

II. Mejoramiento de políticas. Usando los valores v_i (i en S), encuentre el elemento de $G(i,f)$ para cada i en S tal que

$$r_i^k + \beta \sum_{j \in S} p_{ij}^k \, v_j > v_i$$

y k en K_i. Si $G(i,f)$ es vacío para toda i en S, f^∞ es β-óptima y $V_\beta(f^\infty)$ es el beneficio total esperado. Si al menos alguna $g(i)$ está en $G(i,f)$ para alguna i, definiremos una nueva estra tegia g^∞ tal que $g(i)$ está en $G(i,f)$ para alguna i y $g(i) = f(i)$ para $G(i,f)$ vacío; entonces retornamos al paso I.

Como una estrategia inicial, f^∞ puede ser tomada a ser, por ejemplo, $\max_{k \text{ en } k_i} r_i^k$ para cada i en S.

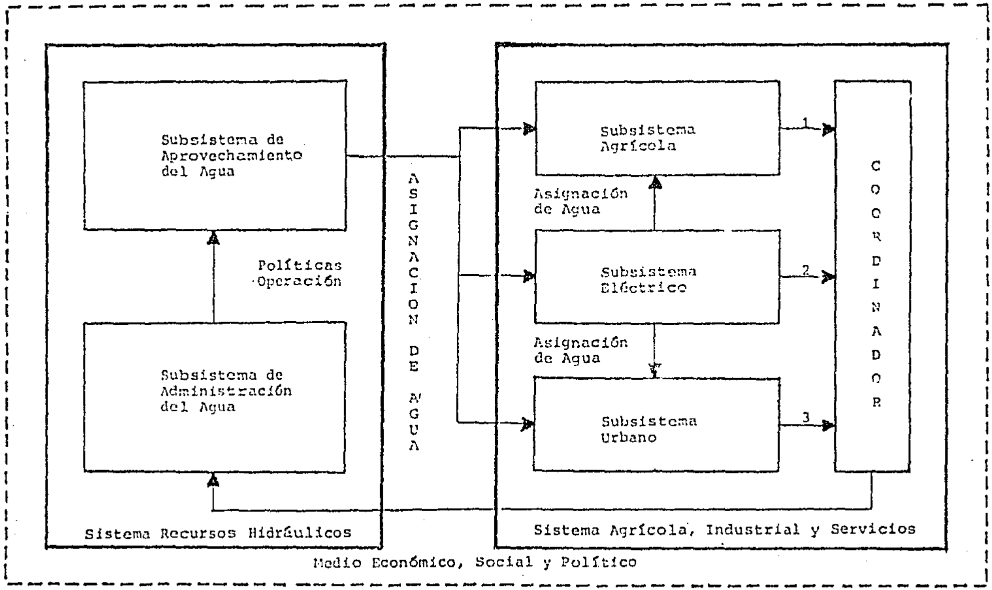
Este algoritmo es simple y elegante y la convergencia es bastante rápida.

4. Los sistemas de recursos hidráulicos

Una de las problemáticas más importantes en el sector agrícola es la determinación de las políticas de operación para el uso adecuado de los recursos hidráulicos. En los denominados Distritos de Riego las políticas de asignación de agua han sido tomadas, en general, siguiendo dos grandes tendencias. La primera consiste en tener grandes almacenamientos y utilizar, únicamente, pequeños volúmenes de agua con el propósito de asegurar el suministro de la misma en años futuros. La otra tendencia, basada en el valor actual del agua, consiste en favorecer su uso inmediato. Dichas tendencias pueden llegar a ser contraproducentes pues puede ocurrir que en años de sequía no sea posible suministrar agua o que en años abundantes se tengan grandes volumenes desperdiciados, que podrían ser aprovechados de otro modo. Estos desequilibrios conducen a pérdidas económicas y problemas sociales.

Un sistema de recursos hidráulicos que permite situar la problemática descrita en un contexto más general se describe en la figura 1. El diagrama tiene como propósito identificar las partes afectadas y sus interrelaciones, así como mostrar algunos objetivos globales de este sistema.

En este capítulo se desarrollan algunos modelos que permiten evaluar y jerarquizar las distintas políticas del uso del agua. El método de Howard es usado para determinar las políticas de asignación de agua, sobre una base anual y otra semestral, del



1. Productos Agrícolas 2. Energía Generada 3. Calidad del Agua

Fig. 1. El Sistema General de Recursos Hidráulicos

sistema de recursos hidráulicos al subsistema de riego, bajo el criterio de maximización de beneficios obtenidos de la producción de cultivos. La naturaleza periódica de los ciclos agrícolas permite modelar las problemáticas que nos ocupan como procesos de decisión Markovianos con descuento.

Respecto a la base (anual o semestral) en que se determinan las políticas de asignación de agua, su elección depende de la naturaleza de los escurrimientos en la región. En muchos casos, una política anual es adecuada; sin embargo, el conflicto surge cuando una base anual resulta poco práctica, pues en muchas regiones es posible observar un comportamiento de los escurrimientos que no puede ser reflejado sobre esta base. Por ejemplo, de lo que escurre anualmente, es factible que una parte considerable se concentre en unos cuantos meses y tomar decisiones anuales no contemplaría este hecho. Por otra parte, hacer decisiones en períodos muy cortos, es difícil, pues los beneficios obtenidos de la producción agrícola dependen de la duración del ciclo productivo, que generalemente, no es menor de tres meses para muchos productos. Un sistema de recursos hidráulicos que describe el proceso anterior, cuando se tiene un período húmedo y un período seco, se tiene en la figura 2.

En este trabajo, se supone, que han sido tomados en cuenta los desequilibrios que surgen al determinar la periodicidad en que se van a establecer las políticas de operación del sistema en cuestión y dicha periodicidad ha sido fijada. Por otra parte,

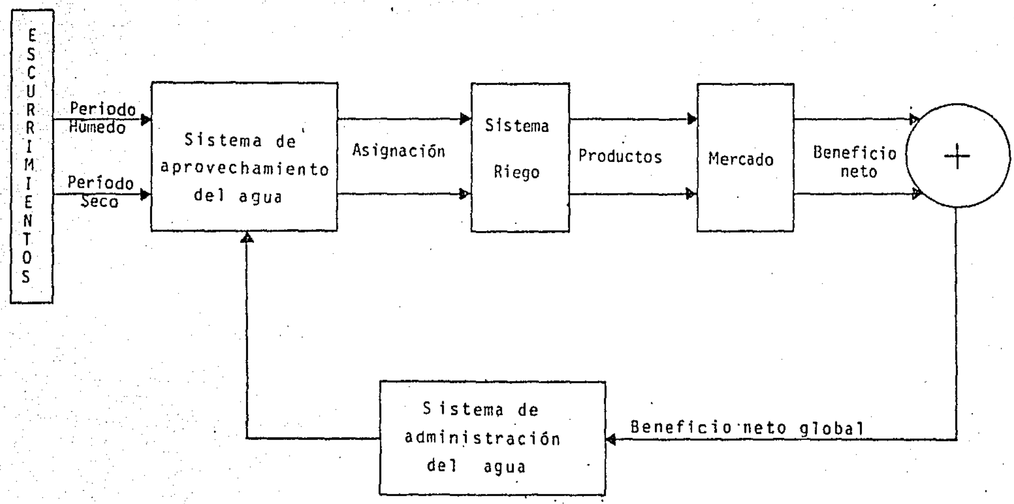


Fig. 2. Sistema de Recursos Hidráulicos (Estacional)

dado que el método de Howard, supone que el número de transiciones que realice el sistema es grande, esta suposición puede ser garantizada para los casos que se modelan si se satisfacen las condiciones:

a. La tecnología agrícola permanece constante.
b. Se mantiene la relación entre el costo de los insumos y el precio de los productos agrícolas.

Una justificación de esta hipótesis es decir que el efecto del incremento en tecnología se anula con el incremento de los costos de los insumos y el precio de los productos.

A continuación se presentan las aplicaciones de los procesos Markovianos a sistemas de recursos hidráulicos en los cuales se determinan estrategias óptimas de asignación de agua para riego sobre una base de decisión anual y otra semestral.

4.1 Aplicación I: políticas de asignación anual.

Considere un sistema de abastecimientos hidráulicos consistente de un vaso y un distrito de riego (fig. 1). Suponga que los escurrimientos que llegan al vaso en cada periodo (i.e., período = un año) son estocásticamente independientes y pueden representarse por una función de densidad discreta como sigue: $Q = 1/4$ con probabilidad igual a 1/2 y $Q = 3/4$ con probabilidad igual a 1/2. Asimismo considere que la capacidad del vaso es igual a uno y suponga que los beneficios obtenidos de asignar un volumen $\underset{\sim}{x}$ de agua al distrito de riego son:

$\underset{\sim}{x}$	0	1/4	1/2	3/4	1
$B(\underset{\sim}{x})$	0	2	7/2	9/2	5

Por otra parte, si x denota el volumen de agua prometido a el distrito y se entrega un volumen $\underset{\sim}{x}$, la penalización debida al deficit de agua $x-\underset{\sim}{x}\geq 0$ es dada por:

$x-\underset{\sim}{x}$	0	1/4	1/2	3/4	1
$T(x-\underset{\sim}{x})$	0	-4	-7	-9	-10

De manera semejante la penalización debida a el derrame de un volumen $w\geq 0$ de agua del vaso es dada por:

w	0	1/4	1/2	3/4	1
$D(w)$	0	-3/2	-4	-4	-4

Finalmente suponga que el factor de descuento para actualizar

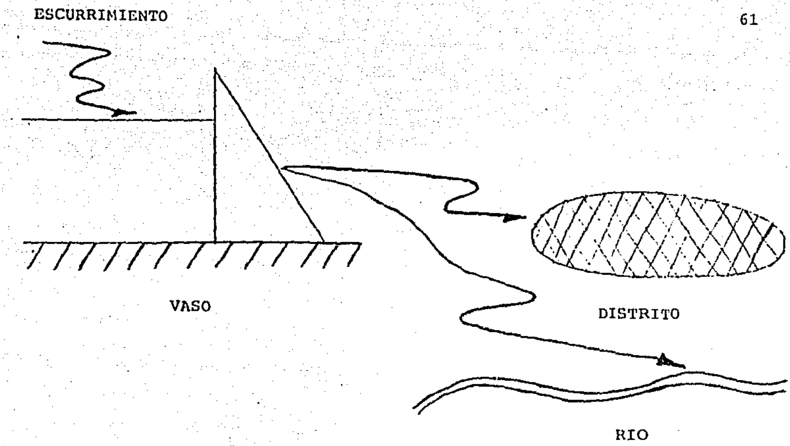


Fig. 1. El sistema de aprovechamientos

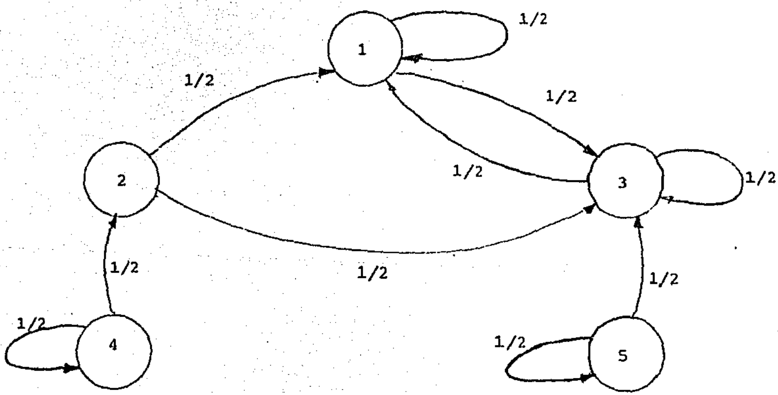


Fig. 2. El proceso de Markov (política óptima)

un peso de un período al anterior es 0.83.

¿Cual debería ser la asignación de agua al distrito, para un número grande de períodos?.

<u>Solución</u>:

Con el propósito de analizar la problemática descrita se presenta, en la siguiente hoja, un modelo que resume las restricciones del sistema de aprovechamientos, así como el criterio de jerarquización para comparar alternativas de operación.
En el modelo, (1) representa la ecuación de balance de entradas y salidas de agua a el vaso; (2) especifica que el almacenamiento en el vaso debe eatar entre $\underline{S}=0$ y $\overline{S}=1$; (3) especifica que el volumen de agua entregada es no-negativo y menor igual a el volumen de agua prometida; y, (4) especifica que el volumen de agua derramado es no-negativo. Conviene mencionar que todos los eventos en las restricciones se efectuan al principio de cada período. Finalmente, la función objetivo considera los beneficios de asignación de agua, la penalización por deficits en la entrega y la debida a derrames, todos ellos convenientemente actualizados.

El estado del sistema, puede ser caracterizado por los diferentes niveles de almacenamiento del vaso. Con objeto de hacer discreta esta variable, se escogen N=5 valores que serán etiquetados por los enteros 1, 2, 3, 4 y 5 correspondientes a los niveles del vaso 0, 1/4, 1/2, 3/4 y 1 respectivamente. Así el conjunto de estados es finito. Ahora bien, en cada estado

MODELO DE ASIGNACION DE AGUA

$$\underset{\{x_i\}}{\text{Maximice}} \ \underset{\{Q_i\}}{E} \left[\sum_{i=1}^{\infty} \beta^{i-1} \{ B(\underset{\sim}{x}_i) + T(x-\underset{\sim}{x}_i) + D(w_i) \} \right]$$

sujeto a:

(1) $S_{i+1} = S_i + Q_i - \underset{\sim}{x}_i - w_i$

(2) $\underline{S} \leq S_i \leq \overline{S}$

(3) $0 \leq \underset{\sim}{x}_i \leq x_i$

(4) $0 \leq w_i$

para toda i.

podemos realizar una acción k, k=1, 2, 3, 4 y 5 donde el entero especifica que el agua prometida para riego es $x_i = 0$, 1/4, 1/2, 3/4 y 1, respectivamente. Como consecuencia podemos calcular el beneficio inmediato esperado como:

$$r_i^k = \sum_{j=1}^{5} P_{ij}^k \; R_{ij}^k$$

donde P_{ij}^k representa la probabilidad de transición del nivel de almacenamiento i al almacenamiento j prometiendo k volumen de agua para riego. Dicha probabilidad puede ser calculada a través de la ecuación de balance de agua, representada por la restricción (1) del modelo planteado anteriormente. Similarmente R_{ij}^k representa el beneficio obtenido al pasar del nivel de almacenamiento i al j prometiendo k vólumen de agua para riego. Su cálculo es por medio de las funciones de beneficios y penalizaciones dadas anteriormente. Los valores de las matrices de transición $[P_{ij}^k]$ y de beneficios inmediatos $[R_{ij}^k]$ se dan en la tabla 1.

Con esto se logra reformular la problemática analizada como un proceso de Markov con descuento (fig. 2). Queremos, así, buscar la estrategia estacionaria que sea β-óptima, i.e., que maximice

$$V_\beta(f^\infty) = \sum_{n=0}^{\infty} \beta^n \; P^n(f) \; r(f)$$

donde r(f) es el vector de beneficios inmediatos esperados cuyo i-ésimo elemento es r_i^k, k=f(i).

La solución del modelo Markoviano, usando la técnica de Howard se describe en la siguiente hoja.

Aplicación del método de Howard

Si se considera como política inicial:

Estado	1	2	3	4	5
Decisión	2	3	4	5	5

Tenemos que la matriz de transición y el vector de beneficios esperados es

$$P = \begin{bmatrix} 1/2 & 0 & 1/2 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \end{bmatrix} \quad r(f)= \begin{bmatrix} 2 \\ 7/2 \\ 9/2 \\ 5 \\ 5 \end{bmatrix}$$

De donde se plantea el sistema de ecuaciones

$$v_1 = 2 + 0.83\ (0.5\ v_1 + 0.5\ v_3)$$

$$v_2 = 3.5 + 0.83\ (0.5\ v_1 + 0.5\ v_3)$$

$$v_3 = 4.5 + 0.83\ (0.5\ v_1 + 0.5\ v_3)$$

$$v_4 = 5 + 0.83\ (0.5\ v_1 + 0.5\ v_3)$$

$$v_5 = 5 + 0.83\ (0.5\ v_2 + 0.5\ v_4)$$

cuya solución es:

v_1	v_2	v_3	v_4	v_5
17.87	19.37	20.37	20.87	21.70

Usando la rutina de mejoramiento de políticas (tabla 2) se tiene que es necesario cambiar, en los estados 4 y 5, las políticas de decisión (extracción) originales 5 y 5, respectivamente, por las políticas 4 y 4, respectivamente.

En esta segunda iteración la matriz de transición y vector de beneficios inmediatos es

$$p = \begin{bmatrix} 1/2 & 0 & 1/2 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 \end{bmatrix} \qquad r(f) = \begin{bmatrix} 2 \\ 7/2 \\ 9/2 \\ 9/2 \\ 9/2 \end{bmatrix}$$

De donde el nuevo sistema que se plantea es:

$$v_1 = 2 + 0.83\ (0.5\ v_1 + 0.5\ v_3)$$

$$v_2 = 3.5 + 0.83\ (0.5\ v_1 + 0.5\ v_3)$$

$$v_3 = 4.5 + 0.83\ (0.5\ v_1 + 0.5\ v_3)$$

$$v_4 = 4.5 + 0.83\ (0.5\ v_2 + 0.5\ v_4)$$

$$v_5 = 4.5 + 0.83\ (0.5\ v_3 + 0.5\ v_5)$$

cuya solución es:

v_1	v_2	v_3	v_4	v_5
17.83	19.37	20.37	21.43	22.14

Usando nuevamente la rutina de mejoramiento de políticas (tabla 1) se tiene que es necesario cambiar, en el estado 3, la política de decisión (extracción) 4 por 3.

En esta tercera iteración la matriz de transición y vector de beneficios inmediatos es:

$$P = \begin{bmatrix} 1/2 & 0 & 1/2 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 \end{bmatrix} \quad r(f) = \begin{bmatrix} 2 \\ 7/2 \\ 7/2 \\ 9/2 \\ 9/2 \end{bmatrix}$$

y el sistema de ecuaciones es

$$v_1 = 2 + 0.83\ (0.5\ v_1 + 0.5\ v_3)$$

$$v_2 = 3.5 + 0.83\ (0.5\ v_1 + 0.5\ v_3)$$

$$v_3 = 3.5 + 0.83\ (0.5\ v_2 + 0.5\ v_4)$$

$$v_4 = 4.5 + 0.83\ (0.5\ v_2 + 0.5\ v_4)$$

$$v_5 = 4.5 + 0.83\ (0.5\ v_3 + 0.5\ v_5)$$

cuya solución es

v_1	v_2	v_3	v_4	v_5
17.96	19.46	20.50	21.50	22.23

como puede observarse de la tabla 2, esta es la política óptima.

Tabla 1. Matrices de transición y beneficios inmediatos asociados al problema

Estado (Almacenamiento)	Política (Extracción)	Probabilidad $\left[P_{ij}^{k}\right]$					Beneficio inmediato $\left[R_{ij}^{k}\right]$					Beneficio inmediato esperado $\sum_{j=1}^{5} P_{ij}^{k} R_{ij}^{k}$
1(0)	1(0)	0	1/2	0	1/2	0	0	0	0	0	0	0
	2(1/4)	1/2	0	1/2	0	0	2	0	2	0	0	→ 2
	3(2/4)	1/2	1/2	0	0	0	-2	7/2	0	0	0	3/4
	4(3/4)	1	0	0	0	0	-1/4	0	0	0	0	-1/4
	5(1)	1	0	0	0	0	-13/4	0	0	0	0	-13/4
2(1/4)	1(0)	0	0	1/2	0	1/2	0	0	0	0	0	0
	2(1/4)	0	1/2	0	1/2	0	0	2	0	2	0	2
	3(1/2)	1/2	0	1/2	0	0	7/2	0	7/2	0	0	→ 7/2
	4(3/4)	1/2	1/2	0	0	0	-1/2	9/2	0	0	0	2
	5(1)	1	0	0	0	0	3/4	0	0	0	0	3/4
3(1/2)	1(0)	0	0	0	1/2	1/2	0	0	0	0	-3/2	-3/4
	2(1/4)	0	0	1/2	0	1/2	0	0	2	0	2	2
	3(1/2)	0	1/2	0	1/2	0	0	7/2	0	7/2	0	7/2
	4(3/4)	1/2	0	1/2	0	0	9/2	0	9/2	0	0	→ 9/2
	5(1)	1/2	1/2	0	0	0	1/2	5	0	0	0	11/4
4(3/4)	1(0)	0	0	0	0	1	0	0	0	0	-2	-2
	2(1/4)	0	0	0	1/2	1/2	0	0	0	2	1/2	5/4
	3(1/2)	0	0	1/2	0	1/2	0	0	7/2	0	7/2	7/2
	4(3/4)	0	1/2	0	1/2	0	0	9/2	0	9/2	0	9/2
	5(1)	1/2	0	1/2	0	0	5	0	5	0	0	→ 5
5(1)	1(0)	0	0	0	0	1	0	0	0	0	-11/4	-11/4
	2(1/4)	0	0	0	0	1	0	0	0	0	0	0
	3(1/2)	0	0	0	1/2	1/2	0	0	0	7/2	2	11/4
	4(3/4)	0	0	1/2	0	1/2	0	0	9/2	0	9/2	9/2
	5(1)	0	1/2	0	1/2	0	0	5	0	5	0	→ 5

Estado (Almacenamiento)	Decisión (Extracción)	$r_i^k + \beta \sum_{j=1}^{5} p_{ij}^k v_j$		
		Iteración 1	Iteración 2	Iteración 3
1	1	16.7	16.93	17.0
	2	17.87 *	17.87 *	17.96 *
	3	16.21	16.21	16.28
	4	14.58	14.58	17.6
	5	11.58	11.58	15.2
2	1	17.46	17.46	17.73
	2	18.93	18.93	19.0
	3	19.37 *	19.37 *	19.46 *
	4	17.45	17.45	17.53
	5	15.58	15.58	19.20
3	1	17.25	17.33	17.4
	2	19.00	19.64	19.73
	3	20.20	20.43	20.5 *
	4	20.37 *	20.37 *	20.46
	5	18.20	18.20	18.28
4	1	16.01	16.38	16.45
	2	18.92	19.33	19.75
	3	20.96	21.14	21.23
	4	21.20	21.43 *	21.5 *
	5	20.87 *	20.87	20.96
5	1	15.26	15.63	15.7
	2	18.01	18.38	18.45
	3	20.42	20.83	21.25
	4	21.96	22.14 *	22.23 *
	5	21.70 *	21.93	22

* Indica la política que se proponía como óptima

Tabla 2. Resumen de resultados de la rutina de mejoramiento de políticas

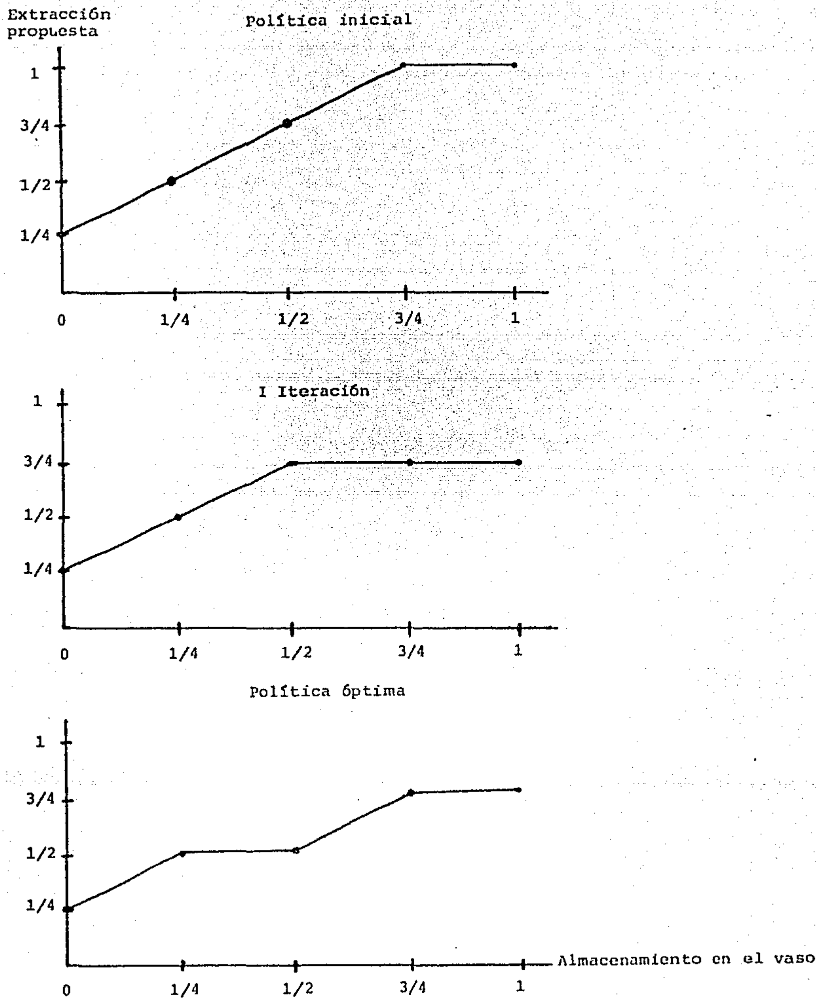


Fig. 3. Políticas de extracción de agua

4.2 Aplicación II: políticas de asignación estacionales

Considere un sistema de aprovechamientos hidráulicos, consistente de un vaso y un distrito de riego (fig. 1). En dicho sistema, los escurrimientos que llegan al vaso en un año quedan caracterizados por medio de dos períodos semestrales denominados: período húmedo y período seco. En el primero los escurrimientos siguen una función de densidad discreta dada por:

$$Q = \begin{cases} 1/4 & \text{con probabilidad } 1/2 \\ 3/4 & \text{con probabilidad } 1/2 \end{cases}$$

Asimismo, suponga que los escurrimientos que llegan al vaso en el período seco pueden representarse por una función de densidad discreta, como sigue:

$$G = \begin{cases} 0 & \text{con probabilidad } 1/2 \\ 1/2 & \text{con probabilidad } 1/2 \end{cases}$$

Se supone además que los escurrimientos entre los distintos períodos y año con año, son estocásticamente independientes.

Por otra parte considere que la capacidad del vaso es igual a uno y suponga que los beneficios de asignar un volumen de agua $\tilde{x}(y)$ al distrito de riego en el período húmedo (seco) son:

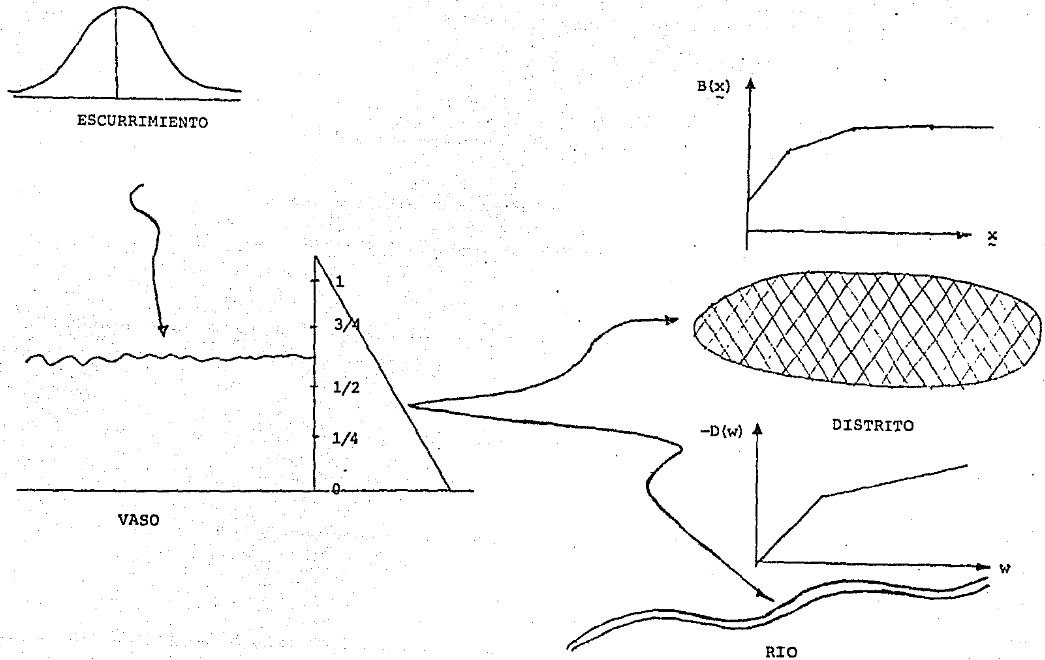


Fig. 1. El sistema de aprovechamientos hidráulicos.

$\underset{\sim}{x}/\underset{\sim}{y}$	0	1/4	1/2	3/4	1
$B_1(\underset{\sim}{x})$	0	2	7/2	9/2	5
$B_2(\underset{\sim}{y})$	0	3/2	3	4	9/2

Asimismo, si x(y) denota el volumen de agua prometida al distrito en el período húmedo (seco) y se entrega un volumen $\underset{\sim}{x}(\underset{\sim}{y})$, la penalización debida al deficit de agua $x-\underset{\sim}{x}\geq 0$ $(y-\underset{\sim}{y}\geq 0)$, es dada por

$x-\underset{\sim}{x}/y-\underset{\sim}{y}$	0	1/4	1/2	3/4	1
$T_1(x-\underset{\sim}{x})$	0	-4	-7	-9	-10
$T_2(y-\underset{\sim}{y})$	0	-3	-6	-8	-9

De manera semejante, la penalización debida al derrame de un vólumen $w\geq 0$ $(z\geq 0)$ de agua del vaso en el período húmedo (seco) es dada por:

w/z	0	1/4	1/2	3/4	1
$D_1(w)$	0	-3/2	-4	-4	-4
$D_2(z)$	0	-1	-3/2	-4	-4

Finalmente suponga que el factor de descuento, para actualizar un peso de un período al anterior es 0.911.

¿Cual debería ser la asignación de agua al distrito, para cada uno de los niveles del vaso (S=0, 1/4, 1/2, 3/4, 1) en cada período, para un número grande de años?

Solución:

El modelo que representa la problemática descrita, se tiene en la siguiente hoja, donde (1) y (2) representan la ecuación de balance de entradas y salidas de agua al vaso, de un período al otro; (3) especifica que el almacenamiento en el vaso, en ambos períodos, debe estar entre $\underline{S}=0$ y $\overline{S}=1$; (4) denota que el volumen de agua entregada, en ambos períodos es no negativo y menor o igual al volumen de agua prometido; y por último, (5) especifica que el volumen de agua derramado es no negativo. Conviene mencionar que todos los eventos en las restricciones se efectúan al principio de cada período. En la función objetivo se consideran los beneficios por asignación de agua, la penalización por déficits en la entrega y la debida a derrames, en cada período, convenientemente actualizados.

Considerando que se desea establecer la extracción de agua del vaso para cada nivel propuesto de almacenamiento en cada período es conveniente establecer:

a. Período húmedo

Los niveles S = 0, 1/4, 1/2, 3/4 y 1 se denotan por los índices r = 1, 2, 3, 4 y 5 respectivamente. Asimismo, las alternativas de extracción correspondientes, x=0, 1/4, 1/2, 3/4 y 1 se denotan por los índices k = 1, 2, 3, 4 y 5, respectivamente.

MODELO DE ASIGNACION DEL AGUA
(ESTACIONAL)

$$\max_{\{x_i,\ y_i\}} \ \underset{\{Q_i,\ G_i\}}{E} \left\{ \sum_{i=1}^{\infty} \beta^{2(i-1)} \left[B_1(\underset{\sim}{x}_i) + T_1(x_i - \underset{\sim}{x}_i) + D_1(w_i) + \beta\{B_2(\underset{\sim}{y}_i) + T_2(y_i - \underset{\sim}{y}_i) + D_2(z_i)\} \right] \right\}$$

sujeto a:

(1) $H_i = S_i + Q_i - \underset{\sim}{x}_i - w_i$

(2) $S_{i+1} = H_i + G_i - \underset{\sim}{y}_i - z_i$

(3) $\underline{S} \leq S_i \leq \overline{S} \ ; \ \underline{S} \leq H_i \leq \overline{S}$

(4) $0 \leq \underset{\sim}{x}_i \leq x_i \ ; \ 0 \leq \underset{\sim}{y}_i \leq y_i$.

(5) $0 \leq w_i \ ; \ 0 \leq z_i$.

para toda i.

b. Período seco

Los niveles $H = 0, 1/4, 1/2, 3/4$ y 1 se denotan por los índices 6, 7, 8, 9 y 10, respectivamente. Las alternativas de extracción $y = 0, 1/4, 1/2, 3/4$ y 1, se denotan por 6, 7, 8, 9 y 10, respectivamente.

Como consecuencia podemos calcular el beneficio inmediato esperado como:

$$r^k(i) = \sum_{j=1}^{10} p_{ij}^k R_{ij}^k \qquad i=1,\dots,10$$

donde p_{ij}^k es la probabilidad de pasar del estado (almacenamiento) i al estado j realizando la acción (extracción) k. Note que si $1 \le i \le 5$, entonces $p_{ij}^k = 0$ para $j = 1, 2, \dots, 5$. Similarmente, $p_{ij}^k = 0$ para $j = 6, \dots, 10$ cuando $6 \le i \le 10$. Ahora bien, una vez que se ha fijado una estrategia f, el vector de beneficios esperados r(f), será ahora igual a $r^k(i)$, $k=f(i)$ para $1 \le i \le 10$. Por otra parte, la matriz de transición P(f) tiene la forma:

$$P(f) = \left[\begin{array}{c|c} 0 & A(f) \\ \hline B(f) & 0 \end{array}\right]$$

donde cada una de las submatrices es de orden 5x5, y, A(B) es la matriz de transición del primer (segundo) período del año.

Con estas observaciones, y con objeto de actualizar debidamente los beneficios esperados, debemos buscar la estrategia estacionaria β óptima, i.e., que maximice

$$V_{\beta}(f^{\infty}) = \sum_{n=0}^{\infty} \beta^{n} P^{n}(f) \; r(f)$$

Implantación del método de Howard

Con el objeto de adecuar el método de Howard al caso de que se tengan dos períodos de decisión en el año, se propone la siguiente metodología:

a. Solución del sistema de ecuaciones

Una vez que se ha especificado una estrategia, debemos encontrar las v(i), tal que,

$$v(i) = r(i) + \beta \sum_{j=6}^{10} p_{ij} \, v(j) \qquad i=1,\ldots,5$$

$$v(i) = r(i) + \beta \sum_{j=1}^{5} p_{ij} \, v(j) \qquad i=6,\ldots,10$$

Dichas fórmulas pueden ser vistas como un sistema de 10 ecuaciones con 10 incógnitas con la estructura que se muestra en la figura 2. Sin embargo, si definimos:

$$v_1 = \begin{bmatrix} v(1) \\ \cdot \\ \cdot \\ \cdot \\ v(5) \end{bmatrix} \qquad\qquad r_1 = \begin{bmatrix} r(1) \\ \cdot \\ \cdot \\ \cdot \\ r(5) \end{bmatrix}$$

$$
\begin{bmatrix} v(1) \\ v(2) \\ \cdot \\ \cdot \\ \cdot \\ v(5) \\ v(6) \\ v(7) \\ \cdot \\ \cdot \\ \cdot \\ v(10) \end{bmatrix}
=
\begin{bmatrix} r(1) \\ r(2) \\ \cdot \\ \cdot \\ \cdot \\ r(5) \\ r(6) \\ r(7) \\ \cdot \\ \cdot \\ \cdot \\ r(10) \end{bmatrix}
+ \beta
\begin{bmatrix}
0 \cdots 0 & p_{1,6} \cdots p_{1,10} \\
0 \cdots 0 & p_{2,6} \cdots p_{2,10} \\
\cdots & \cdots \\
\cdots & \cdots \\
\cdots & \cdots \\
0 \cdots 0 & p_{5,6} \cdots p_{5,10} \\
p_{6,1} \cdots p_{6,5} & 0 \cdots 0 \\
p_{7,1} \cdots p_{7,5} & 0 \cdots 0 \\
\cdots & \cdots \\
\cdots & \cdots \\
\cdots & \cdots \\
p_{10,1} \cdots p_{10,5} & 0 \cdots 0
\end{bmatrix}
\begin{bmatrix} v(1) \\ v(2) \\ \cdot \\ \cdot \\ \cdot \\ v(5) \\ v(6) \\ v(7) \\ \cdot \\ \cdot \\ \cdot \\ v(10) \end{bmatrix}
$$

Fig. 2. El sistema de ecuaciones.

$$v_2 = \begin{bmatrix} v(6) \\ v(7) \\ \cdot \\ \cdot \\ \cdot \\ v(10) \end{bmatrix} \qquad r_2 = \begin{bmatrix} r(6) \\ r(7) \\ \cdot \\ \cdot \\ \cdot \\ r(10 \end{bmatrix}$$

el sistema de ecuaciones puede ser representado matricialmente como:

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} + \beta \left[\begin{array}{c|c} \theta & A \\ \hline B & \theta \end{array}\right] \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

donde A es la matriz de transición del período uno al dos y B es la del período dos al uno. Resolviendo este sistema de ecuaciones se tiene que:

$$v_1 = r_1 + \beta A v_2 \qquad \ldots \quad (1)$$

$$v_2 = r_2 + \beta B v_1 \qquad \ldots \quad (2)$$

Sustituyendo la ecuación (2) en la ecuación (1), tenemos que,

$$v_1 = r_1 + \beta A r_2 + \beta^2 AB v_1$$

de donde

$$[I - \beta^2 AB] v_1 = r_1 + \beta A r_2$$

Haciendo $\quad C = [I - \beta^2 AB]^{-1},$

$$v_1 = Cr_1 + \beta CA r_2$$

y sustituyendo este valor en la ecuación (2) podemos encontrar v_2.

Hemos demostrado de esta manera, que basta con calcular la matriz C (la existencia de C está garantizada ya que $0<\beta^2<1$ y AB es una matriz de transición (ver proposición 1 de la sección 3.2), para que con operaciones matriciales simples encontremos la solución del sistema de ecuaciones.

b) <u>Mejoramiento de políticas</u>

una vez resuelto el sistema de ecuaciones para una estrategia específica debemos, de acuerdo a Howard, maximizar:

$$r^k(i) + \beta \sum_{j=1}^{10} p_{ij}^k \, v(j) \qquad \text{para } i=1,\ldots,5$$

$$r^k(i) + \beta \sum_{j=1}^{10} p_{ij}^k \, v(j) \qquad \text{para } i=6,\ldots,10$$

respecto a todas las alternativas en el i-ésimo nivel de almacenamiento. En el caso de dos períodos, dada la estructura de la matriz de transición P, el problema se reduce a maximizar

$$r^k(i) + \beta \sum_{j=6}^{10} p_{ij}^k \, v(j) \qquad \text{si } 0 \leq i \leq 5 \ (k=1,\ldots,5)$$

$$r^k(i) + \beta \sum_{j=1}^{5} p_{ij}^k \, v(j) \qquad \text{si } 6 \leq i \leq 10 \ (k=6,\ldots,10)$$

Los valores de la matriz de transición $\left[p_{ij}^k\right]$ y la de beneficios inmediatos se dan en la tabla 1. Con ésto, se logra reformular la problemática analizada como un proceso de Markov con descuento (fig. 3). La solución del modelo Markoviano, usando la técnica de Howard se describe en la siguiente hoja.

Aplicación del método de Howard

Si se considera como política inicial:

Estado	1	2	3	4	5	6	7	8	9	10
Decisión	2	3	4	5	5	6	7	8	9	10

tenemos que la matriz de transición y el vector de beneficios esperados es:

$$P = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad r(f) = \begin{bmatrix} 2 \\ 7/2 \\ 9/2 \\ 5 \\ 5 \\ 0 \\ 6/4 \\ 3 \\ 4 \\ 9/2 \end{bmatrix}$$

De donde se plantea el sistema de ecuaciones

$$v(1) = 2 + 0.911(0.5v(6) + 0.5v(8))$$

$$v(2) = 7/2 + 0.911(0.5v(6) + 0.5v(8))$$

$$v(3) = 9/2 + 0.911(0.5v(6) + 0.5v(8))$$

$$v(4) = 5 + 0.911(0.5v(6) + 0.5v(8))$$

$$v(5) = 5 + 0.911(0.5v(7) + 0.5v(9))$$

$$v(6) = 0 + 0.911(0.5v(1) + 0.5v(3))$$

$v(7) = 6/4 + 0.911(0.5v(1) + 0.5v(3))$

$v(8) = 3 + 0.911(0.5v(1) + 0.5v(3))$

$v(9) = 4 + 0.911(0.5v(1) + 0.5v(3))$

$v(10) = 9/2 + 0.911(0.5v(1) + 0.5v(3))$

cuya solución es:

$v(1) = 25.90$ $v(2) = 27.40$ $v(3) = 28.40$ $v(4) = 28.90$

$v(5) = 30.04$ $v(6) = 24.74$ $v(7) = 26.24$ $v(8) = 27.74$

$v(9) = 28.74$ $v(10) = 29.24$

Usando la rutina de mejoramiento de políticas (tabla 2) se tiene que es necesario cambiar, en los estados 3, 4, 5 y 10, las políticas de decisión (extracción) originales 4, 5, 5 y 10, respectivamente, por las políticas 3, 4, 4 y 9 respectivamente. En esta segunda iteración la matriz de transición y vector de beneficios inmediatos es:

$$P = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad r(f) = \begin{bmatrix} 2 \\ 7/2 \\ 7/2 \\ 9/2 \\ 9/2 \\ 0 \\ 6/4 \\ 3 \\ 4 \\ 4 \end{bmatrix}$$

y el sistema de ecuaciones es

$$v(1) = 2 + 0.911(0.5v(6) + 0.5v(8))$$
$$v(2) = 7/2 + 0.911(0.5v(6) + 0.5v(8))$$
$$v(3) = 7/2 + 0.911(0.5v(7) + 0.5v(9))$$
$$v(4) = 9/2 + 0.911(0.5v(7) + 0.5v(9))$$
$$v(5) = 9/2 + 0.911(0.5v(8) + 0.5v(10))$$
$$v(6) = 0 + 0.911(0.5v(1) + 0.5v(3))$$
$$v(7) = 6/4 + 0.911(0.5v(1) + 0.5v(3))$$
$$v(8) = 3 + 0.911(0.5v(1) + 0.5v(3))$$
$$v(9) = 4 + 0.911(0.5v(1) + 0.5v(3))$$
$$v(10) = 4 + 0.911(0.5v(2) + 0.5v(4))$$

cuya solución es:

$v(1) = 26.24$ $v(2) = 27.74$ $v(3) = 28.88$ $v(4) = 29.88$

$v(5) = 31.09$ $v(6) = 25.11$ $v(7) = 26.61$ $v(8) = 28.11$

$v(9) = 29.11$ $v(10) = 30.25$

Usando nuevamente la rutina de mejoramiento de políticas (tabla 2) se tiene que es necesario cambiar en los estados 4, 9 y 10, las políticas de extracción 4, 9 y 9 por 3, 8 y 8 respectivamente.

En esta tercera iteración, la matriz de transición y vector de beneficios inmediatos es:

$$P = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad r(f) = \begin{bmatrix} 2 \\ 7/2 \\ 7/2 \\ 7/2 \\ 9/2 \\ 0 \\ 6/4 \\ 3 \\ 3 \\ 3 \end{bmatrix}$$

y el sistema de ecuaciones es:

$$v(1) = 2 + 0.911(0.5v(6) + 0.5v(8))$$
$$v(2) = 7/2 + 0.911(0.5v(6) + 0.5v(8))$$
$$v(3) = 7/2 + 0.911(0.5v(7) + 0.5v(9))$$
$$v(4) = 7/2 + 0.911(0.5v(8) + 0.5v(10))$$
$$v(5) = 9/2 + 0.911(0.5v(8) + 0.5v(10))$$
$$v(6) = 0 + 0.911(0.5v(1) + 0.5v(3))$$
$$v(7) = 6/4 + 0.911(0.5v(1) + 0.5v(3))$$
$$v(8) = 3 + 0.911(0.5v(1) + 0.5v(3))$$
$$v(9) = 3 + 0.911(0.5v(2) + 0.5v(4))$$
$$v(10) = 3 + 0.911(0.5v(3) + 0.5v(5))$$

cuya solución es:

$v(1) = 26.49$ $v(2) = 27.99$ $v(3) = 29.22$ $v(4) = 30.37$
$v(5) = 31.37$ $v(6) = 25.38$ $v(7) = 26.88$ $v(8) = 28.38$
$v(9) = 29.57$ $v(10) = 30.61$

que como puede observarse en la tabla 2, ésta es la política óptima.

Tabla 1. Matrices de transición y beneficios inmediatos asociados al problema

Estado (Almacenamiento)	Política (Extracción)	Probabilidad $\left[P_{ij}^{k}\right]$					Beneficio inmediato $\left[R_{ij}^{k}\right]$					Beneficio inmediato esperado $\sum_{j=6}^{10} P_{ij}^{k} R_{ij}^{k}$
1(0)	1(0)	0	1/2	0	1/2	0	0	0	0	0	0	0
	2(1/4)	1/2	0	1/2	0	0	2	0	2	0	0	→ 2
	3(2/4)	1/2	1/2	0	0	0	-2	7/2	0	0	0	3/4
	4(3/4)	1	0	0	0	0	-1/4	0	0	0	0	-1/4
	5(1)	1	0	0	0	0	-13/4	0	0	0	0	-13/4
2(1/4)	1(0)	0	0	1/2	0	1/2	0	0	0	0	0	0
	2(1/4)	0	1/2	0	1/2	0	0	2	0	2	0	2
	3(1/2)	1/2	0	1/2	0	0	7/2	0	7/2	0	0	→ 7/2
	4(3/4)	1/2	1/2	0	0	0	-1/2	9/2	0	0	0	2
	5(1)	1	0	0	0	0	3/4	0	0	0	0	3/4
3(1/2)	1(0)	0	0	0	1/2	1/2	0	0	0	0	-3/2	-3/4
	2(1/4)	0	0	1/2	0	1/2	0	0	2	0	2	2
	3(1/2)	0	1/2	0	1/2	0	0	7/2	0	7/2	0	7/2
	4(3/4)	1/2	0	1/2	0	0	9/2	0	9/2	0	0	→ 9/2
	5(1)	1/2	1/2	0	0	0	1/2	5	0	0	0	11/4
4(3/4)	1(0)	0	0	0	0	1	0	0	0	0	-2	-2
	2(1/4)	0	0	0	1/2	1/2	0	0	0	2	1/2	5/4
	3(1/2)	0	0	1/2	0	1/2	0	0	7/2	0	7/2	7/2
	4(3/4)	0	1/2	0	1/2	0	0	9/2	0	9/2	0	9/2
	5(1)	1/2	0	1/2	0	0	5	0	5	0	0	→ 5
5(1)	1(0)	0	0	0	0	1	0	0	0	0	-11/4	-11/4
	2(1/4)	0	0	0	0	1	0	0	0	0	0	0
	3(1/2)	0	0	0	1/2	1/2	0	0	0	7/2	2	11/4
	4(3/4)	0	0	1/2	0	1/2	0	0	9/2	0	9/2	9/2
	5(1)	0	1/2	0	1/2	0	0	5	0	5	0	→ 5

PERIODO HUMEDO

PERÍODO SECO

Estado (Almacenamiento)	Política de extracción	Probabilidad $\left[p^k_{ij}\right]$					Beneficio inmediato $\left[R^k_{ij}\right]$					Beneficio inmediato esperado $\sum_{j=1}^{5} p^k_{ij} R^k_{ij}$
6 (0)	6 (0)	1/2	0	1/2	0	0	0	0	0	0	0	→ 0
	7 (1/4)	1/2	1/2	0	0	0	-3	3/2	0	0	0	-3/4
	8 (1/2)	1	0	0	0	0	-3/2	0	0	0	0	-3/2
	9 (3/4)	1	0	0	0	0	-4	0	0	0	0	-4
	10 (1)	1	0	0	0	0	-6	0	0	0	0	-6
7 (1/4)	6 (0)	0	1/2	0	1/2	0	0	0	0	0	0	0
	7 (1/4)	1/2	0	1/2	0	0	3/2	0	3/2	0	0	→ 6/4
	8 (1/2)	1/2	1/2	0	0	0	-3/2	3	0	0	0	3/4
	9 (3/4)	1	0	0	0	0	-1/4	0	0	0	0	-1/4
	10 (1)	1	0	0	0	0	-1/4	0	0	0	0	-11/4
8 (1/2)	6 (0)	0	0	1/2	0	1/2	0	0	0	0	0	0
	7 (1/4)	0	1/2	0	1/2	0	0	3/2	0	3/2	0	6/4
	8 (1/2)	1/2	0	1/2	0	0	3	0	3	0	0	→ 3
	9 (3/4)	1/2	1/2	0	0	0	0	4	0	0	0	2
	10 (1)	1	0	0	0	0	3/4	0	0	0	0	2/4
9 (3/4)	6 (0)	0	0	0	1/2	1/2	0	0	0	0	-1	-1/2
	7 (1/4)	0	0	1/2	0	1/2	0	0	3/2	0	3/2	6/4
	8 (1/2)	0	1/2	0	1/2	0	0	3	0	3	0	3
	9 (3/4)	1/2	0	1/2	0	0	4	0	4	0	0	→ 4
	10 (1)	1/2	1/2	0	0	0	5/2	9/2	0	0	0	14/4
10 (1)	6 (0)	0	0	0	0	1	0	0	0	0	-3/2	-3/2
	7 (1/4)	0	0	0	1/2	1/2	0	0	0	3/2	1/2	1
	8 (1/2)	0	0	1/2	0	1/2	0	0	3	0	3	3
	9 (3/4)	0	1/2	0	1/2	0	0	4	0	4	0	4
	10 (1)	1/2	0	1/2	0	0	9/2	0	9/2	0	0	→ 9/2

Estado (Almacenamiento)	Decisión (Extracción)	Período Húmedo $\max\limits_{\{1,2,3,4,5\}} r^k(i) + \beta \sum\limits_{j=6}^{10} p_{ij}^k\, v(j)$ $(1 \leq i \leq 5)$	Período Seco $\max\limits_{\{6,7,8,9,10\}} r^k(i) + \beta \sum\limits_{j=1}^{5} p_{ij}^k\, v(j)$ $(6 \leq i \leq 10)$	
		Iteración 1	Iteración 2	Iteración 3
1(0)	1 (0)	25.04	25.38	25.71
	2(1/4)	25.90*	26.23*	26.49*
	3(1/2)	23.97	24.31	24.56
	4(3/4)	22.29	22.62	22.88
	5 (1)	19.29	19.62	19.88
2(1/4)	1 (0)	25.95	26.58	26.87
	2(1/4)	27.04	27.38	27.71
	3(1/2)	27.40*	27.74*	27.99*
	4(3/4)	25.22	25.55	25.81
	5 (1)	23.29	23.62	23.88
3(1/2)	1 (0)	25.66	26.28	26.66
	2(1/4)	27.95	28.58	28.87
	3(1/2)	28.54	28.87*	29.22*
	4(3/4)	28.40*	28.74	28.99
	5 (1)	25.97	26.31	21.06
4(3/4)	1 (0)	24.64	25.55	25.88
	2(1/4)	27.66	28.28	28.66
	3(1/2)	29.45	30.08	30.37*
	4(3/4)	29.54	29.88*	30.21
	5 (1)	28.90*	29.24	29.49
5(1)	1 (0)	23.89	24.80	25.13
	2 (1/4)	26.64	27.55	27.88
	3 (1/2)	29.16	29.78	30.16
	4 (3/4)	30.45	31.08*	31.37*
	5 (1)	30.04*	30.38	30.71

6(0)	1 (0)	24.74*	25.10*	25.38*
	2(1/4)	23.53	23.83	24.07
	3(1/2)	22.10	22.40	22.64
	4(3/4)	19.60	19.90	20.14
	5 (1)	17.60	17.90	18.14
7(1/4)	1 (0)	25.65	26.24	26.59
	2(1/4)	26.24*	26.60*	26.88*
	3(1/2)	25.03	25.33	25.57
	4(3/4)	23.35	23.65	23.89
	5 (1)	20.85	21.15	21.39
8(1/2)	1 (0)	26.62	27.31	27.60
	2(1/4)	27.15	27.74	28.09
	3(1/2)	27.74*	28.10*	28.38*
	4(3/4)	26.28	26.58	26.82
	5 (1)	24.10	24.40	24.64
9(3/4)	6 (0)	26.35	27.27	27.63
	7(1/4)	28.12	28.81	29.10
	8(1/2)	28.65	29.24	29.59*
	9(3/4)	28.74*	29.10*	29.38
	10(1)	27.78	28.08	28.32
10(1)	6 (0)	25.87	26.82	27.08
	7(1/4)	27.85	28.77	29.13
	8(1/2)	29.62	30.31	30.60*
	9(3/4)	29.65	30.24*	30.59
	10(1)	29.24*	29.60	29.88

* Indica la política que se proponía como óptima

TABLA 2. RESUMEN DE RESULTADOS DE LA RUTINA DE MEJORAMIENTO DE POLÍTICAS

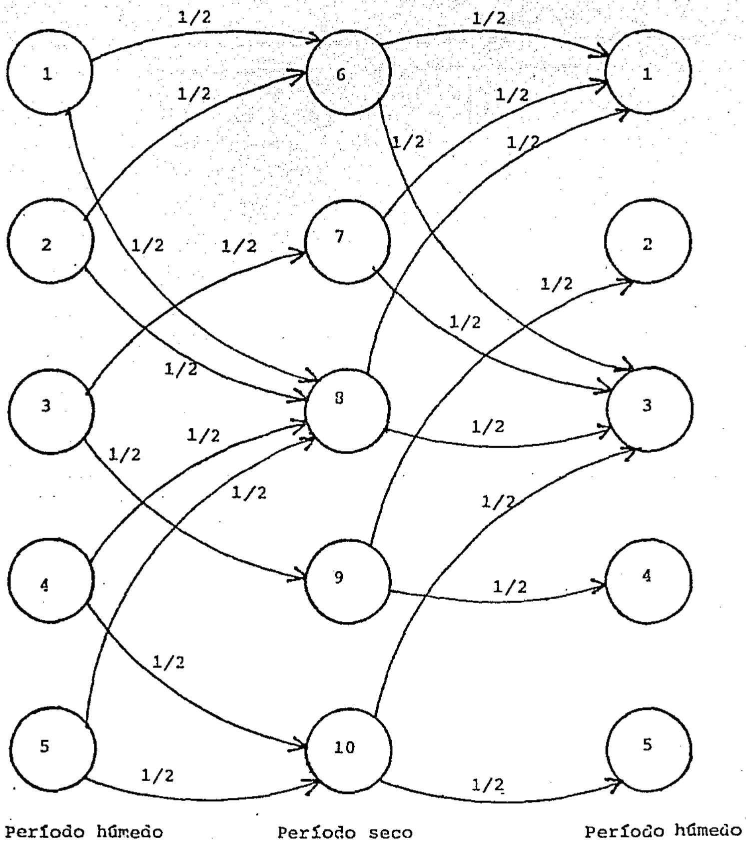


Fig. 3. El proceso de Markov (Política óptima).

PERIODO HUMEDO
(semestral)

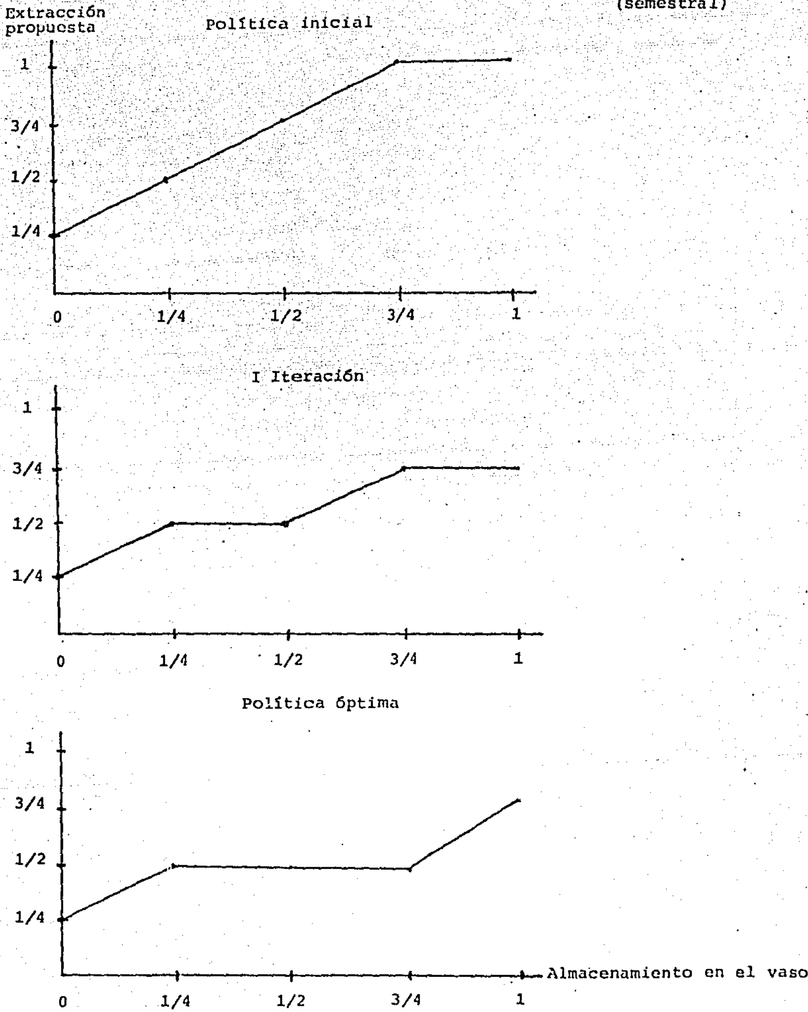


Fig. 4. Políticas de extracción de agua

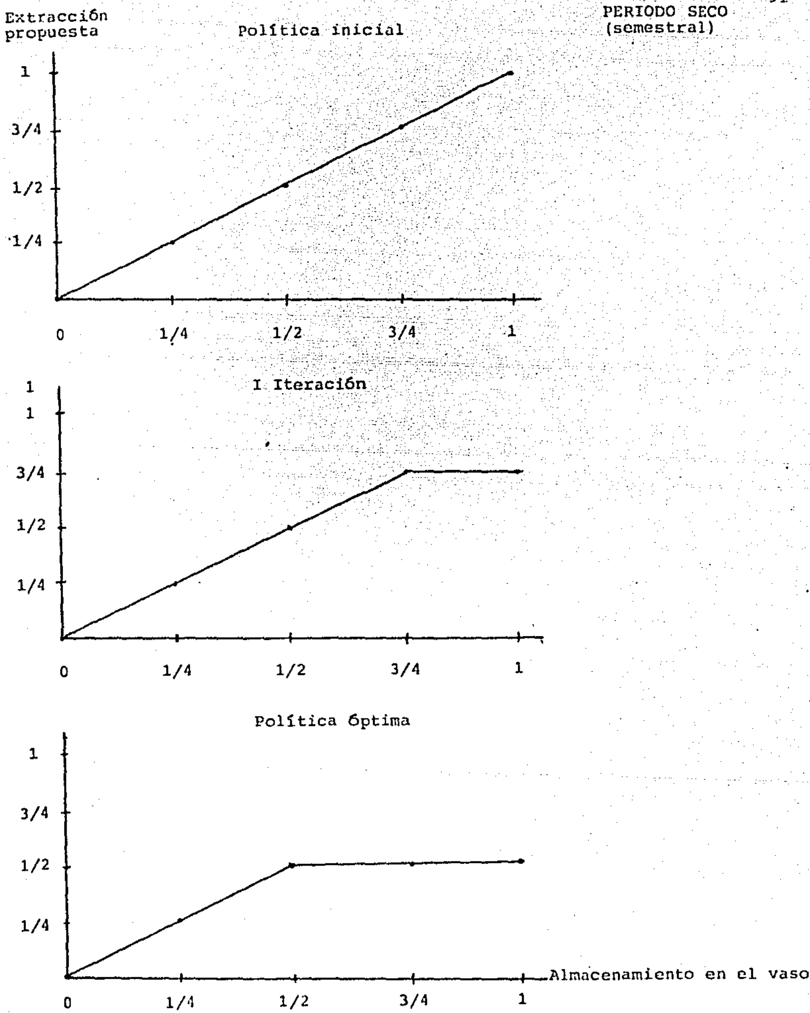


Fig. 4. Políticas de extracción de agua

5. Conclusiones

Se aplicó el método de Howard y para ellos se demostró su convergencia con solo la introducción de una notación adecuada. También dicho método fue aplicado para resolver ejemplos numéricos sencillos de sistemas de recursos hidráulicos.

La extensión de dichos ejemplos a casos específicos puede obtenerse sin dificultad y, en particular, para la aplicación en que se determinó una estrategia estacional, si se toman en cuenta más períodos de decisión en en el año, e.g., tres cuatro o más basta una simple adecuación al ejemplo tratado.

Tal vez la dificultad mayor, cuando de casos concretos se trata, el la implantación en computadora de método de Howard y principalmente en lo que respecta a la solución del sistema de ecuaciones. Por esto, en el caso estacional se propone una metodología para encontrar la solución a través de simples operaciones matriciales.

Finalmente, un análisis de sensibilidad respecto al factor de descuento sería también recomendable .

BIBLIOGRAFIA

Lang, S., "Algebra Lineal", Fondo Educativo Interamericano, (1976).

Luenberger, D. G., "Introduction to Dynamic Systems, Theory, Models, and Applications", Wiley (1979).

Mine, H. , Osaki, S., "Markovian Decision Process", American Elsevier, (1970).

Nering, E. D., "Linear Algebra and Matrix Theory", Wiley, (1963).