Scientia et Technica Año XIV, No 39, Septiembre de 2008. Universidad Tecnológica de Pereira. ISSN 0122-1701

PROPIEDADES DE LA MATRIZ P^n EN UNA CADENA DE MARKOV

Properties of the matrix P^n in a Markov chain

RESUMEN

En este artículo presentamos algunos resultados de la la matriz P^n donde P es la matriz de una cadena de Markov y daremos la forma explicita de las componentes de P^n para una matriz 2×2 y para una matriz triangular inferior 3×3.

PALABRAS CLAVES: Cadena de Markov, Martingala y Valores propios.

ABSTRACT

In this paper presents somes results of the matrix P^n where P^n is the matrix of a Markov chains and we will give the form specifies of the components the P^n for a matrix 2×2 and for a low triangular matrix 3×3.

KEYWORDS: Chains Markov, Martingales and Eigenvalues

EDGAR ALIRIO VALENCIA ANGULO
Profesor Auxiliar, Magíster en Ciencias Matemáticas
Departamento de Matemáticas
Facultad de Ciencias Básicas
Universidad Tecnológica de Pereira
evalencia@utp.edu.co

FERNANDO MESA
Profesor Titular, Magister en Instrumentación Física
Departamento de Matemáticas
Facultad de Ciencias Básicas
Universidad Tecnológica de Pereira
femesa@utp.edu.co

1. INTRODUCCIÓN

La teoría de los procesos estocástico resulta ser una herramienta importante para resolver problemas en otras ramas del conocimiento como la ingeniería genética, la estadística, la economía y en la misma matemáticas. Dentro de los procesos estocásticos ocupa un lugar importante, las llamadas cadenas de Markov, las cuales llevan este nombre en honor al matemático ruso Andrei Andreevich Markov (1856-1922) quien las definió en (1906).

En este artículo presentamos la demostración de un resultado sobre los valores propios de una matriz estocástica de una cadena de Markov y la forma como se expresa P^n en términos de los valores propios de P y en ocasiones en términos de los coeficientes de la matriz P. En el caso de una matriz 2×2 y de una matriz triangular inferior 3×3, expresamos de forma explicita los coeficientes de la matriz P^n.

Fecha de Recepción: 4 de Junio de 2008.
Fecha de Aceptación: 1 de Octubre de 2008.

Para la elaboración de este artículo seguiremos especialmente cinco referencias: Burden [1], Elaydi [2], Muñoz [3], Shiryaev [4], Williams [5].

2. CADENAS DE MARKOV CON PARÀMETRO DISCRETO

Definición 1 Un proceso estocástico es una familia de variables aleatorias $\{X_t\}_{t\in T}$, definidas todas sobre un espacio de probabilidad (Ω , $\mathcal{F}, P$) y con valores en un mismo espacio medible (S , $\mathcal{K}$). T es el conjunto de índices del proceso, S se llama espacio de estados y sus elementos se llaman estados. Usualmente S y T son subconjuntos de $\mathbf{R}$.

Definición 2 Una sucesión de variables aleatorias $\{X_n\}_{n\in\mathbf{N}}$ definidas en un conjunto de estados discreto $S \subseteq \mathbf{N}$ se llama cadena de Markov con parámetro de tiempo discreto si satisface la siguiente condición:

$$P(X_{n+1} = j / X_n = i, X_{n-1} = i_{n-1}, ..., X_0 = i_0) =$$
$$P(X_{n+1} = j / X_n = i_n)$$

para todo $n\in\mathbf{N}$ y para todo $i_0, i_1, ..., i, j\in S$ siempre y cuando las probabilidades condicionales estèn definidas.

Si la probabilidad condicional del estado i en la etapa n al estado j en la etapa $n+1$, no depende de la etapa origen de la transición, entonces las probabilidades de transición son estacionarias y la cadena de Markov se denomina homogénea.

Definición 3 Se define la matriz de transición de la cadena de Markov como la matriz de probabilidades de transición en una etapa $P=(P_{i,j})$, $i,j \in S$ donde $P_{i,j}=P\left(X_{n+1}=j / X_n=i\right)$. La matriz de transición de una cadena de Markov $[(X]_n)_{n\in\mathbb{N}}$ es estocástica si

$P_{i,j} \geq 0$ y $\sum_{j\in S} P_{i,j}$ =1 para todo $i \in S$.

Un resultado que demostramos a continuación en este articulo y que se propone en Shiryaev **[4]** es el siguiente:

Proposición 4 Sea $P=(P_{i,j})$, $1 \leq i,j \leq r$ una matriz estocástica y λ un valor propio de la matriz P . Muestre que $\lambda_0 = 1$ es un valor propio y que todos los otros valores propios tienen modulo no mayor a uno 1. Si todos los valores propios son distintos, entonces $P_{i,j}{}^{(n)}$ admite la representación

$P_{i,j}{}^{(n)} = \pi_j + \cdots + a_{i,j}(r)\lambda_r^n$

donde π_j, $a_{i,j}(2), \ldots, a_{i,j}(r)$ pueden ser expresados en términos de los elementos de P.

Demostración Sea $v=\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}$. Veamos que $Pv=\lambda v$. En efecto

$$Pv=\begin{pmatrix}P_{11} & \cdots & P_{1r}\\ \vdots & \ddots & \vdots\\ P_{r1} & \cdots & P_{rr}\end{pmatrix}\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}=\begin{pmatrix}P_{11}+\cdots+P_{1r}\\ \vdots\\ P_{r1}+\cdots+P_{rr}\end{pmatrix}=\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}.$$

Por lo tanto que $\lambda_0 = 1$ es un valor propio de P . Sea λ un valor propio de P y $v \neq 0$, definamos $|v_i| = max(|v_1|, \ldots, |v_n|)$ como λ es un valor propio, $Pv=\lambda v$ y la componente $i-$ésima de $Pv=\lambda v$ es

$$\sum_{j=1}^{r} P_{ij}v_j = \lambda v_i$$

, por lo tanto

$$\left|\sum_{j=1}^{r} P_{ij}v_j\right| = |P_{i1}v_1+\cdots+P_{ir}v_r| = |\lambda||v_i|.$$

De aquí que

$|\lambda||v_i| \leq P_{i1}|v_1|+\cdots+P_{ir}|v_r|$ luego

$|\lambda| \leq P_{i1}\dfrac{|v_1|}{|v_i|}+\ldots+P_{ir}\dfrac{|v_r|}{|v_i|}$.

Ahora, como $|v_i| = max(|v_1|, \ldots, |v_n|)$, $\dfrac{v_j}{v_i} < 1$ para $1 \leq j \leq r$, $j \neq i$ por lo tanto $|\lambda| \leq P_{i1}+\ldots+P_{ir}=1$. Como λ es un valor propio arbitrario, entonces el resultado se tiene para cualquier λ.

Supongamos ahora que los valores propios $\lambda_1, \ldots, \lambda_r$ son distintos, por consiguiente la matriz P es diagonalizable, por lo cual existe una matriz invertible A tal que $APA^{-1}=D$ donde D es una matriz diagonal cuya diagonal son los valores propios de P . Por lo tanto $P=A^{-1}DA$ luego $P^n=A^{-1}D^nA$.

Sean $A=\begin{pmatrix}A_1\\ \vdots\\ A_r\end{pmatrix}$, $A^{-1}=(B_1 \;\cdots\; B_r)$ donde A_i y B_j representan las filas y las columnas de A y A^{-1} respectivamente

$$P=(B_1 \;\cdots\; B_r)\begin{pmatrix}1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \lambda_r\end{pmatrix}\begin{pmatrix}A_1\\ \vdots\\ A_r\end{pmatrix}$$

$$P = 1B_1A_1+\cdots+\lambda_rB_rA_r = I_1A_1+\cdots+\lambda_rB_rA_r$$

$P^n = I_1A_1+\cdots+\lambda_r{}^nB_rA_r$, luego

$$P_{i,j}{}^{(n)} = (I_1A_1)_{ij}+\cdots+\lambda_r{}^n(B_rA_r)_{ij}$$

donde $I_1A_1=\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}(a_{11} \;\cdots\; a_{12})=\begin{pmatrix}a_{11} & \cdots & a_{1r}\\ \vdots & \ddots & \vdots\\ a_{11} & \cdots & a_{1r}\end{pmatrix}$.

Una observación es que la matriz P^n es una matriz que converge a una matriz estocástica constante, es decir, $\lim_{n\to\infty} P^n = I_1A_1$.

Ejemplo 1. Si P es una matriz 2×2 calculemos los coeficientes de P^n.

Sea $P=\begin{pmatrix}a & b\\ c & d\end{pmatrix}$ donde $a+b=1$ y $c+d=1$ y $0 \leq a,b,c,d \leq 1$.

$|P-\lambda I| = \begin{vmatrix}a-\lambda & b\\ c & d-\lambda\end{vmatrix} = (a-\lambda)(d-\lambda)-cd = 0$,

Resolviendo esta ecuación, usando las hipótesis $a+b=1$ y $c+d=1$ y $0 \leq a,b,c,d \leq 1$, llegamos a que $\lambda = 1$ y $\lambda = a+d-1 \leq 1$.

Observemos que el segundo valor propio de P esta expresado en términos de a y d .

Para calcular el vector propio correspondiente al valor propio $\lambda = 1$ resolvemos el sistema matricial
$\begin{pmatrix} a-1 & b \\ c & d-1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

y usamos el hecho de que $c = 1 - d$ y $b = 1 - a$,

para obtener el vector propio $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.
Para calcular el vector propio correspondiente al valor propio valor propio de $\lambda = a + d - 1$, resolvemos el sistema matricial

$$\begin{pmatrix} 1-d & 1-a \\ 1-d & 1-a \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

El vector propio es $\begin{pmatrix} \frac{-b}{c} \\ 1 \end{pmatrix}$. Por lo tanto P se puede ver como
$$P = \begin{pmatrix} 1 & \frac{-b}{c} \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & a+d-1 \end{pmatrix}\begin{pmatrix} 1 & \frac{b}{c} \\ -1 & 1 \end{pmatrix}\frac{1}{1+\frac{b}{c}}$$
donde la primera matriz esta formada por los vectores propio encontrados, la segunda matriz es la matriz diagonal, cuya diagonal son los valores propios de P y la tercera matriz es la inversa de la primera matriz, es decir,
$P = \Lambda D \Lambda^{-1}$.
Esto implica que
$P^n = \Lambda D^n \Lambda^{-1}$,
esto es,
$$P^n = \begin{pmatrix} 1 & \frac{-b}{c} \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & (a+d-1)^n \end{pmatrix}\begin{pmatrix} 1 & \frac{b}{c} \\ -1 & 1 \end{pmatrix}\frac{1}{1+\frac{b}{c}}.$$
Resolviendo este producto de matrices llegamos a que

$$P^n = \begin{pmatrix} \frac{(a+d-1)^n b + c}{c+b} & \frac{b - b(a+d-1)^n}{c+b} \\ \frac{c - c(a+d-1)^n}{c+b} & \frac{b + c(a+d-1)^n}{c+b} \end{pmatrix}$$ y su convergencia viene dada por la matriz
$$\lim_{n\to\infty} P^n = \begin{pmatrix} \frac{c}{c+b} & \frac{b}{c+b} \\ \frac{c}{c+b} & \frac{b}{c+b} \end{pmatrix}.$$

Ejemplo 2. Si P es una matriz estocástica triangular inferior 3×3 calculemos los coeficientes de P^n.

Sea $P = \begin{pmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$, $a_{11} = 1$, $a_{21} + a_{22} = 1$ y

$a_{31} + a_{32} + a_{33} = 1$. Es claro que los valores propios de la matriz P son $\lambda = a_{11} = 1$, $\lambda = a_{22} = 1 - a_{21}$ y , $\lambda = a_{33} = 1 - a_{31} - a_{32}$.

Para $\lambda = a_{11} = 1$ obtenemos el vector propio $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ resolviendo el sistema matricial

$$\begin{pmatrix} 0 & 0 & 0 \\ a_{21} & -a_{21} & 0 \\ a_{31} & a_{32} & -a_{31} - a_{32} \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Para $\lambda = a_{22}$ obtenemos el vector propio

$\begin{pmatrix} 0 \\ \frac{a_{22} - a_{33}}{a_{32}} \\ 1 \end{pmatrix}$ resolviendo el sistema

$$\begin{pmatrix} a_{11} - a_{22} & 0 & 0 \\ a_{12} & 0 & 0 \\ a_{31} & a_{32} & a_{33} - a_{22} \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Finalmente para $\lambda = a_{33}$ obtenemos el vector propio $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ resolviendo el sistema

$$\begin{pmatrix} a_{11} - a_{33} & 0 & 0 \\ a_{12} & a_{22} - a_{33} & 0 \\ a_{31} & a_{32} & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$
La matriz ortogonal es por lo tanto

$$\Lambda = \begin{pmatrix} 1 & 0 & 0 \\ 1 & \frac{a_{22} - a_{33}}{a_{32}} & 0 \\ 1 & 1 & 1 \end{pmatrix}$$
y su inversa es

$$\Lambda^{-1} = \begin{pmatrix} 1 & -\frac{a_{32}}{a_{22} - a_{33}} & \frac{-a_{22} + a_{32} + a_{33}}{a_{22} - a_{33}} \\ 0 & \frac{a_{32}}{a_{22} - a_{33}} & -\frac{a_{32}}{a_{22} - a_{33}} \\ 0 & 0 & 1 \end{pmatrix}.$$
Ahora, como P es una matriz diagonalizable, entonces $P = \Lambda D \Lambda^{-1}$
donde D es una matriz diagonal cuyo elementos son los valores propios de P , es decir,

$$\mathrm{D} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix}.$$

Por con siguiente

$$P^n = \Lambda \mathrm{D}^n \Lambda^{-1}$$

donde

$$\mathrm{D}^n = \begin{pmatrix} 1 & 0 & 0 \\ 0 & a_{22}{}^n & 0 \\ 0 & 0 & a_{33}{}^n \end{pmatrix}.$$

Resolviendo este producto de matrices obtenemos

$$P^n = \begin{pmatrix} 1 & 0 & 0 \\ 1 - \dfrac{a_{32}a_{22}{}^n}{a_{31}} & \dfrac{a_{32}a_{22}{}^n}{a_{31}} & 0 \\ 1 - \dfrac{a_{32}a_{22}{}^n - a_{33}{}^n(a_{33} + a_{32} - a_{22})}{(a_{22} - a_{33})} & \dfrac{a_{32}(a_{22}{}^n - a_{33}{}^n)}{(a_{22} - a_{33})} & a_{33}{}^n \end{pmatrix}.$$

La convergencia de esta matriz es una la matriz constante

$$\lim_{n\to\infty} P^n = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Ejemplo 3 Sea $[(X]_k)_{1\le k\le n}$ una cadena de Markov homogénea con espacio de estado S y $P = (P_{x,y})$ una matriz de transición. Sea

$$T(\varphi(i)) = E\left(\varphi(X_1)/_{X_0} = i\right) = \sum_j \varphi(j) P_{x,y}$$

donde φ satisface $T(\varphi(x)) = \varphi(x)$ para todo $x \in S$.
Muestre que $\{\lambda_k, \mathcal{D}_k{}^X\}_{1\le k\le n}$ es una martingala, donde $\lambda_k = \varphi(X_k)$ y $\mathcal{D}_k{}^X = \mathcal{A}(X_1, \dots X_n)$ es el algebra generada por la sucesión de variables aleatoria $[(X]_k)_{1\le k\le n}$.

Solución. Recordemos el concepto de martingala, como aparece en Williams [5], $\{\lambda_k, \mathcal{D}_k{}^X\}_{1\le k\le n}$ es una martingala si λ_k es integrable, λ_k es $\mathcal{D}_k{}^X - medible$ y $E(\lambda_{k+1}/\mathcal{D}_k{}^X) = \lambda_k$.

Como la cadena de Markok es homogénea

$$T(\varphi(x)) = E\left(\varphi(X_1)/_{X_0} = x\right) = E\left(\varphi(X_k)/_{X_{k-1}} = x\right).$$

Ahora, como $T(\varphi(x)) = \varphi(x)$ por tanto

$$\lambda_k = \varphi(X_k) = T(\varphi(X_k)) = \sum_{i=1}^{j} T(\varphi(x_i)) I_{\{X_k = x_i\}}$$

$$\lambda_k = \sum_{i=1}^{j} E(\varphi(x_{k+1})/X_k = x_i) I_{\{X_k = x_i\}}.$$

Calculemos

$$E(\lambda_{k+1}/\mathcal{D}_k{}^X) = E\left(\varphi((x_{k+1}))/_{X_1,\dots,X_k}\right).$$

$$E(\lambda_{k+1}/\mathcal{D}_k{}^X) = \sum_{i=1}^{j} E\left(\varphi((x_i)) I_{\{X_{k+1}=x_i\}}/_{X_1,\dots,X_k}\right)$$

$$E(\lambda_{k+1}/\mathcal{D}_k{}^X) = E\left(E(\varphi(x_{k+1})/X_k = x_i) I_{\{X_{k+1}=x_i\}}/_{X_1,\dots,X_k}\right)$$

$$= E(\varphi(x_{k+1}))/X_k = x_i)\, E\left(I_{\{X_{k+1}=x_i\}}/_{X_1,\dots,X_k}\right).$$

Ahora, como

$$E\left(I_{\{X_{k+1}=x_i\}}/_{X_1,\dots,X_k}\right) = P(X_{k+1} = x_i/X_1, \dots X_k) = P(X_{k+1} = x_i$$

por ser cadena de Markov, luego

$$E\left(\lambda_{k+1}/_{\mathcal{D}_k{}^X}\right) = \sum_{i=1}^{j} T(\varphi(x_i)) P\left(X_{k+1} = x_i/_{X_k}\right) =$$

$x_{\square\downarrow}$

cambiando s por i se tiene que

$$E(\lambda_{k+1}/\mathcal{D}_k{}^X) = \sum_{i=1}^{j} I_{\{X_k = x_i\}}\ T(\varphi(x_i)) = \lambda_k.$$

1. CONCLUSIONES

Presentamos una forma diferente la demostración de la Proposición 4 sobre como son los valores propios de una matriz estocástica P . Como se expresa la matriz estocástica P^n en términos de los valores propios de P , y la convergencia de P^n. Hacemos la demostración de un resultado que relaciona procesos de Markov y martingalas. Finalmente calculamos explícitamente los coeficientes de la matriz de P^n y su limite para los casos en que P es una matriz 2×2 y cuando P es una matriz triangular inferior 3×3

2. BIBLIOGRAFÍA

[1] R. L. Burden, J. D. Faires. *Análisis Numérico.* Grupo Editorial Iberoamérica, 1985.

[2] S. N. Elaydi. *An Introduction to difference Equations.* Springer, Second Edition, 1999.

[3] M. Muñoz, L. Blanco. *Introducción a la teoría avanzada de la probabilidad.* Universidad Nacional de Colombia, Primera edición 2002.

[4] A. N. Shiryaev. *Probability*. Second Edition. Academic Press, 1975.

[5] D. Williams. *Probability Theory with Martingales.* Cambridge University Press, 1997.