# Cadenas de Markov y Aplicaciones

Eliseo Martínez

## 1 Cadenas de Markov

### 1.1 Definición de Cadena de Markov

Sea $\{X_t; t \in I\}$ una colección de variables aleatorias de modo que $X_t$ toma valores en un conjunto $S$, que llamaremos **espacio de estado**, y al conjunto índice $I$ le llamaremos, por razones de aplicación, **espacio de tiempo**. Se dice que este conjunto de variables aleatorias constituye un proceso estocástico si para cualquier colección arbitraria $t_1; t_2; \ldots; t_n$ la distribución del vector aleatorio $(X_{t_1}; X_{t_2}; \ldots; X_{t_n})$ es conocida. Para aclarar esto último, supongamos que cada variable aaleatoria $X_t$ toma valores en el espacio de estado de los números reales $\mathbb{R}$, entonces la siguiente probabilidad debe ser conocida

$$\Pr\{X_{t_1} \le s_1; X_{t_2} \le s_2; \ldots; X_{t_n} \le s_n\}$$

en particular debe ser conocida la probabilidad para cada variable $X_t$. Un caso particular es cuando la colección de variables aleatorias son independientes, entonces se tiene que

$$\begin{aligned} & \Pr\{X_{t_1} \le s_1; X_{t_2} \le s_2; \ldots; X_{t_n} \le s_n\} \\ = \; & \Pr\{X_{t_1} \le s_1\}\Pr\{X_{t_2} \le s_2\}\cdots\Pr\{X_{t_n} \le s_n\} \end{aligned}$$

En el caso de que el espacio de tiempo $I$ sea un conjunto discreto, o con más precisión $I = \mathbb{N}$, siendo $\mathbb{N}$ el conjunto de los números naturales, el proceso estocástico se dice de tiempo **discreto**. Veamos un ejemplo de un proceso estocástico de tiempo discreto: Sea la variable aleatoria $X_n$ que denota la $n-$ésima observación de una variable que se distribuye según una Poisson de parámetro $\lambda$, y además estas variables son conjuntamente independientes. Tenemos entonces que para cada variable $X_i$

$$\Pr\{X_i = k\} = \frac{e^{-\lambda}\lambda^k}{k!}$$

y en particular, la distribución conjunta del vector estocástico $(X_1; X_2; \dots; X_n)$ está dada por

$$\begin{aligned}
\Pr\{X_1 = k_1; \dots; X_n = k_n\} &= \Pr\{X_1 = k_1\} \cdots \Pr\{X_n = k_n\} \\
&= \frac{e^{-\lambda}\lambda^{k_1}}{k_1!} \cdots \frac{e^{-\lambda}\lambda^{k_n}}{k_n!}
\end{aligned}$$

Observemos que,en este ejemplo, el espacio de estado es también el conjunto de los números naturales.

No todos los procesos estocásticos tienen la propiedad de independencia, presentaremos uno que tiene la propiedad de tener "poca memoria". Consideremos el proceso $\{X_n; n \in \mathbb{N}\}$ con espacio de estado $\mathbb{N}$, de manera tal que para cualquier elección de índices de tiempo $i_1 < \cdots < i_n$ se satisface lo siguiente

$$\begin{aligned}
\Pr\{X_{i_m} &= j \,|\, X_{i_1} = k_1;\ X_{i_2} = k_2 \cdots; X_{i_{m-1}} = i\} = \\
&= \Pr\{X_{i_m} = j \,|\, X_{i_{m-1}} = i\}
\end{aligned} \tag{1}$$

Lo que nos dice esta propiedad es que la probabilidad condicionada a un "pasado largo" sólo depende del pasado inmediatamente anterior, esto es lo que sucederá en el tiempo $i_m$ solamente dependerá de lo acontecido en el tiempo $i_{m-1}$, y además se supone que la probabilidad $\Pr\{X_{i_m} = j \,|\, X_{i_{m-1}} = i\}$ es conocida y recibe el nombre de **probabilidad de transición de un paso**. Es posible que la notación de la ecuación (1) sea complicada en virtud de los subíndices, un caso más particular, y sin mucha pérdidad de generalidad, es como sigue

$$\Pr\{X_m = j \,|\, X_1 = k_1; \dots; X_{m-1} = i\} = \Pr\{X_m = j \,|\, X_{m-1} = i\} \tag{2}$$

La propiedad (1) se llama **propiedad markoviana**, y en este caso se entrega para procesos estocásticos de tiempo y estado discreto. Y la propiedad de este tipo de procesos es lo que fundamentalmente estudiaremos.

El primer resultado que veremos a continuación es que todo proceso estocástico que satisface la propiedad de Markov está completamente definido, es decir la probabilidad conjunta de cualquier vector aleatorio está absolutamente definida si ocurre una pequeña condición: debemos conocer la distribución inicial de $X_0$. Es decir $\Pr\{X_0 = k\} = p_k \; \forall k \in N$ es conocido.

**Teorema (1.1)** Supongamos que tenemos un proceso estocástico $\{X_n; n \in N\}$ con espacio de estado $N$ de manera que satisface las siguientes propiedades

1. la propiedad de Markov
2. la distribución inicial de $X_0$ es conocida

Entonces la distribución de $(X_{j_1}, X_{j_2}, \ldots, X_{j_n})$ está determinada.

**Demostración.** A efectos de no complicar en demasía la notación y sin pérdida de generalidad, vamos a calcular la distribución del vector aleatorio $(X_0, X_1, \ldots, X_n)$. Para esto debemos recordar que

$$\Pr\{A \cap B\} = \Pr\{A \mid B\}\Pr\{B\},$$

con esto en mente podemos deducir lo siguiente,

$$\begin{aligned} \Pr\{X_0 &= i_0, \ldots, X_n = i_n\} = \\ \Pr\{X_n &= i_n \mid X_0 = i_0, \ldots, X_{n-1} = i_{n-1}\}\Pr\{X_0 = i_0, \ldots, X_{n-1} = i_{n-1}\} \end{aligned}$$

y en virtud de la propiedad markoviana se tiene que

$$\begin{aligned} \Pr\{X_0 &= i_0, \ldots, X_n = i_n\} = \\ \Pr\{X_n &= i_n \mid X_{n-1} = i_{n-1}\}\Pr\{X_0 = i_0, \ldots, X_{n-1} = i_{n-1}\} \end{aligned} \tag{3}$$

de esta última ecuación, para el segundo factor del miembro derecho podemos aplicar nuevamente la propiedad anterior, esto es

$$\begin{aligned} \Pr\{X_0 &= i_0, \ldots, X_{n-1} = i_{n-1}\} = \\ \Pr\{X_{n-1} &= i_{n-1} \mid X_0 = i_0, \ldots, X_{n-2} = i_{n-2}\}\Pr\{X_0 = i_0, \ldots, X_{n-2} = i_{n-2}\} \\ &= \Pr\{X_{n-1} = i_{n-1} \mid X_{n-2} = i_{n-2}\}\Pr\{X_0 = i_0, \ldots, X_{n-2} = i_{n-2}\} \end{aligned} \tag{4}$$

Procediendo de manera similar a (3) y (4) podemos concluir que

$$\begin{aligned} \Pr\{X_0 &= i_0, \ldots, X_n = i_n\} = \\ \Pr\{X_n &= i_n / X_{n-1} = i_{n-1}\} \cdots \Pr\{X_1 = i_1 / X_0 = i_0\} \Pr\{X_0 = i_0\} \end{aligned} \tag{5}$$

y puesto que todas las cantidades del lado derecho de (5) son conocidas se determina, entonces, la distribución del vector aleatorio $(X_0, X_1, \ldots, X_n)$ $\blacksquare$

La definición dada en (1) necesita algunos comentarios. Todo proceso estocastico de tiempo discreto y estado discreto que satisface esta propiedad suele llamarsele Cadena de Markov, y fundamentalmente nos dice que lo que acontecerá en el futuro depende del presente y no de lo que aconteció en el pasado. Existen bastantes fenómenos en la naturaleza[1] y las ciencias sociales que se modelan según esta propiedad. Sin mucho temor podemos decir, entonces, que la naturaleza, en varios casos, es markoviana.

Una cadena de Markov, como lo vimos en el teorema anterior, está determinada si conocemos las probabilidades de transición a un paso $\Pr\{X_{n+1} = j / X_n = i\}$. En rigor esta probabilidad dependerá de los estados $i$ y $j$, y de los tiempos $n$ y $n+1$, eso es

$$\Pr\{X_{n+1} = j / X_n = i\} = p_{i,j}^{n,n+1} \tag{6}$$

Ahora si esta probabilidad de transición no depende del tiempo se dice que la cadena de Markov es **estacionaria**, y si este es el caso, la expresión (6) queda como

$$\Pr\{X_{n+1} = j / X_n = i\} = p_{i,j} \tag{7}$$

obteniendo el agradable resultado siguiente

$$\Pr\{X_{m+1} = j / X_m = i\} = \Pr\{X_{n+1} = j / X_n = i\} = p_{i,j} \; ; \quad m \neq n$$

lo que resalta la particularidad de que el valor del "salto" desde el estado $i$ al estado $j$ no depende del tiempo en que ocurre el salto. Deciamos que esta propiedad de estacionaridad es muy agradable puesto que el valor de $p_{i,j}$ se puede tratar como un elemento matricial, donde el estado $i$, denota la $i-$ésima columna, y el estado de llegada $j$, denota la fila $j-$ésima[2], siendo $p_{ij}$

[1] Lamentablemente exentos de complejidad, como hubiesemos deseado.

[2] Notese que utilizamos la notación inversa al sentido clásico de la notación matricial, esto es la entrada $a_{ij}$ se ubica en la $i-$ésima fila y $j-$ésima columna.

el valor de la probabilidad de que el proceso pase desde el estado $i$ al estado $j$ en una unidad de tiempo.

En lo que sigue vamos a considerar solamente cadenas de Markov, y estas pueden ser de estado finito o infinito (numerable). Si el espacio de estado de una cadena de Markov es $S = \{0, 1, \ldots, m\}$ entonces las probabilidades de transición se ubican en la matriz[3]

$$\begin{pmatrix} p_{00} & p_{10} & \cdots & p_{m0} \\ p_{01} & p_{11} & \cdots & p_{m1} \\ \vdots & \vdots & \vdots & \vdots \\ p_{0i} & p_{1i} & \cdots & p_{mi} \\ \vdots & \vdots & \vdots & \vdots \\ p_{0m} & p_{1m} & \cdots & p_{mm} \end{pmatrix} \tag{8}$$

que es una matriz cuadrada de orden $m + 1 \times m + 1$. Ahora si el espacio de estado es $S = \{0, 1, \ldots\}$ entonces la matriz infinita asociada a la cadena de Markov es

$$\begin{pmatrix} p_{00} & p_{10} & \cdots & p_{m0} & \cdots \\ p_{01} & p_{11} & \cdots & p_{m1} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \cdots \\ p_{0i} & p_{1i} & \cdots & p_{mi} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \cdots \\ p_{0m} & p_{1m} & \cdots & p_{mm} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Observemos que las columnas de estas matrices deben sumar uno, en efecto es claro que

$$\sum_{j=0}^{M} \Pr\{X_{n+1} = j / X_n = i\} = 1$$

con $M = \infty$ o $M = m$ según sea el caso, puesto que la función $\Pr\{\cdot / X_n = i\}$ es una función de probabilidad.

[3] Que obviamente llamaremos *Matriz de Markov*.

## 1.2 Ejemplos de Cadenas de Markov

### Paseo aleatorio unidimensional

Una "partícula" puede estar en cualquier estado de $\{0; 1; 2 \ldots\}$. Ahora, toda vez que se encuentre, en el tiempo $n$, en el estado $i$ para el próximo tiempo $n+1$ solo puede quedarse en el mismo estado o "saltar" a los estados adyacentes $i+1$, $i-1$. es decir

$$
\begin{aligned}
\Pr\{X_{n+1} &= i+1 / X_n = i\} = p_i \ ; \ i \geq 1 \\
\Pr\{X_{n+1} &= i-1 / X_n = i\} = q_i \ ; \ i \geq 1 \\
\Pr\{X_{n+1} &= i / X_n = i\} = r_i \ ; \ i \geq 1 \\
\Pr\{X_{n+1} &= j / X_n = i\} = 0 \ ; |j-i|\ i > 1
\end{aligned}
$$

para el caso en que $i = 0$ se tiene

$$
\begin{aligned}
\Pr\{X_{n+1} &= 0 / X_n = 0\} = r_0 \\
\Pr\{X_{n+1} &= 1 / X_n = 1\} = p_0
\end{aligned}
$$

donde $p_i + q_i + r_i = 1$ para $i \geq 1$ y $r_0 + p_0 = 1$.

La designación de paseo aleatorio viene del hecho de que se asemeja a una trayectoria de una persona que aleatoriamente avanza hacia atrás o hacia adelante. La matriz de Markov o de transición de este proceso es

$$
\begin{pmatrix}
r_0 & q_1 & 0 & 0 & \cdots \\
p_0 & r_1 & q_2 & 0 & \cdots \\
0 & p_1 & r_2 & q_3 & \cdots \\
0 & 0 & p_2 & r_3 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix}
$$

Este modelo sirve para describir, entre otros, el siguiente juego

### La ruina del jugador

Supongamos que un jugador A con $k$ unidades de dinero juega con otro infinitamente rico, y tiene una probabilidad $p_k$ de ganar una unidad y $q_k = 1 - p_k$, con $k \geq 1$, de perder una unidad en cada respuesta o decisión (la decisión o respuesta en cada etapa del juego puede depender de su fortuna acumulada hasta entonces), y además $r_0 = 1$ (es decir, toda vez que el jugador

A se queda sin dinero, su contendor no le da crédito). Definamos el proceso $\{X_n\}$, donde $X_n$ representa la fortuna de A despues de n jugadas. Notemos que una vez que el estado 0 es alcanzado el proceso jamás sale de este estado. Este proceso es claramente un paseo aleatorio.

El paseo aleatorio correspondiente a $p_k = p$ y $q_k = 1 - p = q$ para $k \geq 1$ y $r_0 = 1$, con $p > q$, describe la situación de decisiones idénticas con una ventaja definida para el jugador A en cada jugada. Más adelante demostraremos que con probabilidad $(q/p)^{X_0}$ el jugador A se arruinará, mientras que con probabilidad $1 - (q/p)^{X_0}$ su fortuna aumentará en una larga serie de jugadas sin límite, donde $X_0$ representa la fortuna inicial. Si $p < q$ entonces la ventaja está a favor del casino, y con probabilidad 1 el jugador A se arruinará si persiste en jugar tanto como le sea posible. Lo mismo ocurrirá si $p = q = 1/2$.

Ahora si el jugador A juega contra un adversario no infinitamente rico, obtenemos el siguiente ejemplo

**Los jugadores con poco dinero.**

Supongamos dos jugadores que juegan un juego donde se pierde o se gana, y en cada jugada se apuesta un dólar por jugador. El jugador A tiene $a$ dólares y el jugador B tiene $b$ dólares. Supongamos que no hay ventajas para ningún jugador, esto es cada uno tiene las mismas opciones de ganar (o perder). Desde el punto de vista de la ganancia del jugador A, definamos el proceso $X_n$ como la ganancia obtenida por A en la $n-$ésima jugada.

Es claro que la ganancia de este jugador antes de una determinada jugada, dependerá de la ganancia que haya obtenido en la jugada inmediatamente anterior (esto es, la propiedad markoviana). Además existe cierta regla de parada, toda vez que el jugador A pierda todo su dinero, o le gane todo el dinero a su contrincante, el juego se detiene. Bajo estas condiciones el espacio de estado de este proceso es $S = \{0, 1, 2, \ldots, a + b\}$, donde las probabilidades de transición son:

$$
Pr = \begin{cases} \frac{1}{2} \text{ si } j = i + 1 \text{ y } 0 < i < a + b \\ \frac{1}{2} \text{ si } j = i - 1 \text{ y } 0 < i < a + b \\ 1 \quad \text{si } i = j = 0 \\ 1 \quad \text{si } i = j = a + b \\ 0 \quad \text{en cualquier otro caso} \end{cases}
$$

Esta definición no debe extrañarnos, nos dice que si el jugador A lleva una ganancia (o dólares) que no son ni cero ni $a + b$ entonces en la próxima jugada tiene opción de ganar o perder con probabilidad $1/2$, ahora si queda con cero dólares entonces se acaba el juego y con probabilidad 1 se queda en ese estado, y si felizmente llega al estado $a + b$ es porque el jugador B se quedo sin banca y el juego se detiene quedando A en ese estado. Como en cada jugada se apuesta un dolar no es posible obtener ganancias mayores que la unidad por cada jugada.

La matriz de Markov para este proceso es

$$
\begin{pmatrix} 1 & \frac{1}{2} & 0 & \cdots & 0 \\ 0 & 0 & \frac{1}{2} & \cdots & 0 \\ 0 & \frac{1}{2} & 0 & \cdots & 0 \\ 0 & 0 & \frac{1}{2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}
$$

## Un modelo de inventario

Supóngase un stock que se maneja a dos niveles, $0 < s < S$, de manera que si lo almacenado es menor o igual a $s$ entonces inmediatamente se repone hasta el nivel $S$, en caso contrario ninguna reposición se hace. la inspección se realiza en una cierta unidad de tiempo discreta (por ejemplo al final de cada semana). Ahora bien, la demanda para cada período de tiempo se distribuye según una Poisson de parámetro $\lambda$ en forma independienre, esto es si $\xi_n$ es la demanda en la $n-$ésima semana, entonces

$$
Pr\{\xi_n = i\} = e^{-\lambda} \frac{\lambda^i}{i!}; \qquad \forall n
$$

Se supone además que si el stock, en un determinado período, no alcanza para una demanda en el mismo período entonces ésta es satisfecha parcialmente y no se difiere el resto (se entrega lo que haya sin compromiso para el próximo período). Para modelar esta situación definamos como $X_n$ el nivel del stock al final delperíodo n. Notemos que este proceso es claramente markoviano, puesto que lo que ocurra en la semana $n-$ésima dependerá de lo que ha sucedido en la semana $n-1$. Para fijar ideas calculemos la probabilidad de transición $p_{00}$, esto es

$$p_{00} = Pr\{X_n = 0 / X_{n-1} = 0\} = Pr\{\xi_n \geq S\} = \sum_{i=S}^{\infty} e^{-\lambda} \frac{\lambda^i}{i!}$$

Calculemos ahora $p_{03}$, esto es

$$p_{03} = Pr\{X_n = 3 / X_{n-1} = 0\} = Pr\{\xi_n = S - 3\} = e^{-\lambda} \frac{\lambda^{S-3}}{(S-3)!}$$

en este último caso suponiendo que $3 < S$. Notemos que $p_{ij} = 0$ si $i < j$ siendo $s < i < S$. Con estos antecedentes se puede construir facilmente la matriz de Markov asociado a este modelo de inventario. Una manera elegante de definir este proceso de Markov es como sigue

$$X_{n+1} = \begin{cases} (X_n - \xi_{n+1})^+ & \text{si } s < X_n < S \\ (S - \xi_{n+1})^+ & \text{si } X_n \leq s \end{cases}$$

donde $f^+ = \max\{f, 0\}$.

### El paseo del borracho o aproximación al Movimiento Browniano

Los paseos aleatorios también sirven para la aproximación discreta a procesos físicos que describen el movimiento de difusión de partículas. Si una partícula está sujeta a colisiones o impulsos aleatorios, entonces su trayectoria fluctúa aleatoriamente. Si la futura posición de una partícula depende solo de su posición actual, entonces el proceso puede ser markoviano. Una versión clásica discreta de un movimiento Browniano es proporcionada por el paseo aleatorio simétrico bidimensional.

Suponga usted que una partícula se ncuentra en una malla cuadriculada y en cada intervalo de tiempo regular toma una nueva posición vecina ubicada en uno de los cuatro puntos cardinales. Para fijar ideas supongamos que en el tiempo n la partícula se encuentra en el vértice $(i; j)$, entonces en el tiempo $n+1$ solamente podrá estar en una y solamente una de las posiciones siguientes $(i+1; j)$, $(i-1; j)$, $(i; j-1)$ y $(i; j+1)$. Ahora bien, cada nueva posición es igualmente preferible, esto es la probabilidad de tomar una nueva posición es $1/4$. De otra forma si $X_n$ denota la posición de la partícula en el tiempo $n$, entonces las probabilidades de transición estan dadas por

$$
\begin{aligned}
\Pr\{X_{n+1} &= (k; l) / X_n = (i; j)\} = \\
&= \begin{cases} \frac{1}{4} & \text{si } (k; l) \in \{(i-1; j); (i+1; j); (i; j-1); (i; j+1)\} \\ 0 & \text{en cualquier otro caso} \end{cases}
\end{aligned}
$$

El modelo de Ehrenfest.

Un modelo clásico matemático de difusión de partículas a través de una membrana es el modelo de Ehrenfest, que consiste en un paseo aleatorio sobre estados finitos y donde los estados fronteras son refractantes. Este paseo aleatorio se restringe a los estados $S = \{-a; -a+1; \ldots; -1; 0; 1; 2; \ldots; a\}$ y las probabilidades de transición se definen como

$$
p_{ij} = \begin{cases} \frac{a-i}{2a} & \text{si } j = i+1 \\ \frac{a+i}{2a} & \text{si } j = i-1 \\ 0 & \text{en cualquier otro caso} \end{cases} \tag{9}
$$

La interpretación física de este modelo es la siguiente: Tenemos dos contenedores con un total de $2a$ bolitas. Supongamos que el primer contenedor tiene $k$ bolitas y el segundo las restantes $2a-k$. Se elige aleatoriamente una de entre las $2a$ bolitas. Según del contenedor que pertenezca esta bolita se deposita en el contenedor contrario. Cada selección genera una transición del proceso. Claramente las bolitas fluctuarán entre ambos contenedores, siendo más probable el traslado del contenedor de mayor concentración. En efecto, supongamos que $X_n$ es el número de bolitas del primer contenedor, si tenemos que $X_n = k$ entonces puede ocurrir uno y solo uno de los siguintes sucesos: $X_{n+1} = k+1$ o $X_{n+1} = k-1$. Y es claro que

$$
\Pr\{X_{n+1} = k+1 / X_n = k\} = \frac{2a-k}{2a} \qquad k = 0; 1; \ldots; 2a
$$

que es la probabilidad de sacar una bolita del segundo contenedor, luego

$$\Pr\{X_{n+1} = k - 1 / X_n = k\} = 1 - \frac{2a - k}{2a} = \frac{k}{2a} \qquad k = 0, 1, \ldots, 2a$$

que es la probabilidad de sacar una bolita del rpimer contenedor. Como $2a - k = a + a - k$ hagamos $a - k = i$, entonces $i \in \{-a, -a-1, \ldots, -1, 0, 1, \ldots, a-1, a\}$ y se concluye la relación (9).

### Un modelo de colas markoviano discreto.

Llegan cllientes para un servicio y toman lugar en una linea de espera. Durante cada período de tiempo se atiende a un solo cliente, siempre, por supuesto, que al menos un cliente esté en la linea de espera. Si ningún cliente espera servicio entonces durante este período no opera el servicio (podemos pensar, por ejemplo, en una parada de taxis en la cual estos llegan a intervalos de tiempos fijos para recoger pasajeros, si nadie está presente el taxi parte inmediatamente). Durante un período de servicio, pueden llegar nuevos clientes. Vamos a suponer que el número de llegadas en el n-ésimo período es una variable aleatoria $\xi_n$, cuya función de distribución es independiente del período y está dada por

$$\Pr\{\xi_n = k\} = a_k \ ; \ a_k \geq 0 \quad k = 0, 1, \ldots$$

y además $\sum_{k=0}^{\infty} a_k = 1$. También vamos a suponer que las variables aleatorias $\xi_n$ son independientes. El estado del sistema en el inicio de cada período n se define como el número de clientes en las lineas de espera para tomar el servicio, $S = \{0, 1, \ldots\}$, y lo denotaremos por $X_n$. Ahora bie, si el estado tiene el valor de $X_n$, entonces en el próximo período de tiempo se tendrá que

$$X_{n+1} = (X_n - 1)^+ + \xi_{n+1}.$$

La matriz de Markov para este proceso es como sigue

$$\begin{pmatrix} a_0 & a_0 & 0 & 0 & 0 & \cdots \\ a_1 & a_1 & a_0 & 0 & 0 & \cdots \\ a_2 & a_2 & a_1 & a_0 & 0 & \cdots \\ a_3 & a_3 & a_2 & a_1 & a_0 & \cdots \\ a_4 & a_4 & a_3 & a_2 & a_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

y ésta se obtiene al evaluar $p_{ij} = \Pr\{X_{n+1} = j / X_n = i\}$.

## 1.3 Dinámica en las Cadenas de Markov

Existe un resultado clásico en Teoría de Probabilidades que es necesario recordarlo. Si $\{B_j\}$ con $j = 1, 2, \ldots$ es una partición del espacio de probabilidad, y si $A$ es un suceso cualquiera, entonces $\Pr\{A\} = \Pr\{A \cap \{\cup_i \{B_i\}\}\} = \Pr\{\cup_j (A \cap B_j)\} = \sum_j \Pr\{A \cap B_j\} = \sum_j \Pr\{A|B_j\}\Pr\{B_j\}$. Para nuestro interés hagamos $A = \{X_{n+1}\}$ y $B_j = \{X_n = j\}$, con el convencimiento d que los $\{X_n = j\}$ constituyen una partición. Con esto obtenemos una igualdad fundamental

$$\Pr\{X_{n+1}\} = \sum_j \Pr\{X_{n+1} = i | X_n = j\} \Pr\{X_n = j\} \tag{10}$$

y puesto que $\Pr\{X_{n+1} = i | X_n = j\}$ son las probabilidades de transición $p_{ji}$, lo que tenemos es un producto vectorial del tipo

$$\Pr\{X_{n+1} = i\} = (p_{0i}\ p_{1i}\ p_{2i} \cdots p_{ji} \cdots) \begin{pmatrix} \Pr\{X_n = 0\} \\ \Pr\{X_n = 1\} \\ \Pr\{X_n = 2\} \\ \vdots \\ \Pr\{X_n = j\} \\ \vdots \end{pmatrix}$$

de manera que llegamos a la siguiente igualdad matricial

$$\begin{pmatrix} \Pr\{X_{n+1} = 0\} \\ \Pr\{X_{n+1} = 1\} \\ \Pr\{X_{n+1} = 2\} \\ \vdots \\ \Pr\{X_{n+1} = j\} \\ \vdots \end{pmatrix} = \begin{pmatrix} p_{00} & p_{10} & \cdots & p_{j0} & \cdots \\ p_{01} & p_{11} & \cdots & p_{j1} & \cdots \\ p_{02} & p_{12} & \cdots & p_{j2} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \cdots \\ p_{0j} & p_{1j} & \cdots & p_{jj} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \cdots \end{pmatrix} \begin{pmatrix} \Pr\{X_n = 0\} \\ \Pr\{X_n = 1\} \\ \Pr\{X_n = 2\} \\ \vdots \\ \Pr\{X_n = j\} \\ \vdots \end{pmatrix} \tag{11}$$

Si denotamos por $P_{n+1}$ al vector de probabilidad del lado izquierdo de (11) y por $M$ a la matriz de Markov de la misma ecuación (11), obtenemos la fórmula compacta

$$P_{n+1} = M \cdot P_n; \tag{12}$$

luego si sabemos la distribución inicial de este proceso, es decir el valor del vector $P_0$ entonces el vector de probabilidad $P_1$, que es la distribución de encontrar el sistema en el tiempo 1, está dado por

$$P_1 = M \cdot P_0 \; ;$$

y dado que estamos considerando cadenas de Markov estacionarias, se tiene el cálculo del vector $P_2$ mediante

$$P_1 = M \cdot P_0$$

y dado que estamos considerando cadenas de Markov estacionarias, se tiene el cálculo del vector $P_2$ mediante

$$P_2 = M \cdot P_1$$

esto es

$$P_2 = M \cdot P_1 = M \cdot M \cdot P_0 = M^2 \cdot P_0$$

No resulta complicado deducir que

$$P_n = M^n \cdot P_0 \tag{13}$$

y esta ecuación dinámica nos entrega el vector de probabilidad en cualquier tiempo, con la condición de saber la distribución inicial del proceso.

**Ejemplo 1** Supongamos una cadena de Markov con dos estados, el 0 y el 1, de manera que toda vez que el proceso se encuentre en el estado 0, en el próximo tiempo, salte al estado 1 con probabilidad $0 < \alpha < 1$ y obviamente se puede quedar en el mismo estado con probabilidad $1 - \alpha$. Ahora si el proceso se encuentra en el estado 1 entonces con probabilidad 1 pasa al estado 0[4]. De otra forma $p_{00} = 1 - \alpha$, $p_{01} = \alpha$, $p_{10} = 1$ y $p_{11} = 0$, entomces la matriz de Markov de este proceso luce de la siguiente forma:

$$\begin{pmatrix} 1 - \alpha & 1 \\ \alpha & 0 \end{pmatrix}$$

Utilizando un lenguaje formal diremos que el proceso $X_n$ tiene como espacio de estado a $S = \{0; 1\}$ y para cualquier tiempo n se tine que $\Pr\{X_{n+1} =$

[4]Cuando un estado tiene esta propiedad, esto es que con probabilidad 1 salta a otro estado, se dice que es un estado *refractario*.

$0/X_n = 0\} = 1 - \alpha$, $Pr\{X_{n+1} = 1/X_n = 0\} = \alpha$, $Pr\{X_{n+1} = 0/X_n = 1\} = 1$ y $Pr\{X_{n+1} = 1/X_n = 1\} = 0$.

En un lenguaje más informal pero intuitivo diremos que la "partícula" se encuentra en el tiempo n en algún estado, y puede iniciar un salto, en el próximo tiempo conforme a lo establecido anteriormente. Supongamos finalmente que la partícula, en el tiempo 0 se encuentra en el estado 0, esto significa que el vector de distribución inicial es

$$P_0 = \begin{pmatrix} Pr\{X_0 = 0\} \\ Pr\{X_0 = 1\} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

de manera que la distribución en el tiempo $n = 1$, $P_1$, está dada por

$$P_1 = \begin{pmatrix} Pr\{X_1 = 0\} \\ Pr\{X_1 = 1\} \end{pmatrix} = \begin{pmatrix} 1 - \alpha & 1 \\ \alpha & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 - \alpha \\ \alpha \end{pmatrix}$$

El cálculo para $P_2$ lo podemos hacer según la igualdad (12) o (13), en cualquier caso

$$\begin{aligned} P_2 &= \begin{pmatrix} 1 - \alpha & 1 \\ \alpha & 0 \end{pmatrix}^2 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 - \alpha & 1 \\ \alpha & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 - \alpha \\ \alpha \end{pmatrix} = \\ &= \begin{pmatrix} 1 - (\alpha - \alpha^2) \\ \alpha - \alpha^2 \end{pmatrix} \end{aligned}$$

Un problema interesante es saber si el vector de probabilidad $P_n$ para un n suficientemente grande se estabilizará, de otra forma si existe el $\lim_{n \to \infty} P_n$. En nuestro ejemplo la respuesta es positiva, sin embargo no siempre ocurre. Las condiciones de existencia para una distribución "asintótica" se estudiará más adelante.