INTRODUCCION ELEMENTAL A LAS CADENAS DE MARKOV

Por Mariela Gómez

Una de las principales tareas de las matemáticas consiste en formular de manera abstracta lo esencial de una o más situaciones concretas y hacer demostraciones de hechos fundamentales independientemente de las propiedades accidentales del problema concreto.

Un concepto básico en el proceso de unificación de la ciencia es el de probabilidad que ha permitido relacionar campos científicos aparentemente desvinculados.

La teoría tuvo su origen en 1654 en los estudios realizados por los matemáticos franceses Blas Pascal y Pierre Fermat relativos a la solución de problemas a que dan lugar juegos de azar. En el siglo XVIII los cultivadores más destacados de la teoría fueron Bernoulli (1654 - 1705) y De Moivre (1667 1754). Solo hasta 1812 en un trabajo de Laplace se mostró la posibilidad de que la teoría podía ser aplicada a problemas prácticos y científicos.

En el desarrollo ulterior de la teoría han tenido especial importancia Tchebycheff, Von Mises, Kolmogoroff y Markov (1856 1922)

La múltiple gama de aplicaciones de las probabilidades ha propiciado un gran avance en la teoría, el que a su vez, ha ampliado el marco de su influencia.

Los procesos gobernados por leyes probabilísticas se denominan estocásticos o aleatorios.

La teoría de los procesos estocásticos hace referencia al estudio de fami lias de variables aleatorias X_t donde t varía en un conjunto T interpretado generalmente como tiempo Los valores de X_t pueden ser unidimensionales o n-dimensionales

Los principales elementos de un proceso aleatorio son :

i) El espacio estado o sea el conjunto S en el que toma valores X_t .

S puede ser real valuado o vector valuado.

ii) El conjunto T de subíndices que puede ser discreto o contínuo.

iii) La relación de dependencia entre las variables aleatorias.

Un tipo clásico de proceso estocástico es el que hace referencia con los denominados procesos de Markov, que se caracteriza por constituir un proceso sin memoria del pasado, lo cual significa que la probabilidad de un evento futuro del proceso no depende de la forma como se llegó al estado presente, o sea que la probabilidad del evento no se altera por conocimiento adicional de la conducta del proceso en el pasado. Sólo es necesario conocer exactamente el estado presente .

Para un proceso X_t Markoviano se tiene :

$$P\{ a < X_t \leq b \mid X_{t_1} = X_1 , \dots , X_{t_n} = X_n] = P [a < X_t \leq b \mid X_{t_n} = X_n \}$$

con $t_1 < t_2 < \dots < t_n < t$.

Una cadena de Markov de tiempo discreto es un proceso estcástico de Markov cuyo espacio estado es un conjunto numerable para el que :

$T = \{ 0 , 1 , 2 , \dots \}$

X_n es el resultado de la enésima prueba

X_n se dice estar en el estado i si $X_n = i$; la probabilidad de que

X_{n+1} esté en el estado J dado que X_n esté en estado i es :

$P^{n,\,n+1}_{iJ} = P\{\,X_{n+1} = J \mid X_n = i\,\}$ Tal probabilidad es función, no solo de los estados inicial y final, sino también en general del tiempo de transición. Si la probabilidad es independiente del tiempo, el proceso se clasifica como de transición estacionaria o de tiempo homogéneo.

$P^{n,n+1}_{iJ} = P_{ij}$ probabilidad de que una prueba pase del estado i al estado J en un paso.

La matriz $(P_{ij}) = P$ que organiza las probabilidades del proceso es :

$$P = \begin{bmatrix} P_{00} & P_{01} & P_{02} & \cdots \\ P_{10} & P_{11} & P_{12} & \cdots \\ P_{20} & P_{21} & P_{22} & \cdots \\ \cdot & \cdot & \cdot & \cdots \\ \cdot & \cdot & \cdot & \cdots \\ \cdot & \cdot & \cdot & \cdots \end{bmatrix} \qquad \begin{matrix} P_{iJ} \geq 0;\ i, J = 0,\ 1,\ 2,\ \ldots \\ \\ \sum_{j=1}^{\infty} P_{iJ} = 1;\ \ i = 0,\ 1,\ 2,\ \ldots \end{matrix}$$

P es la matriz de probabilidad estocástica de transición o matriz de Markov. Si el proceso es finito, la matriz tendría r filas, siendo r el número de estados del proceso.

La $(i+1)$ fila de P es la distribución de las probabilidades de X_{n+1} bajo la condición que $X_n = i$.

Proposición 1) El proceso está totalmente definido si

$$P^{n,\,n+1}_{iJ} = P\{\,X_{n+1} = J \mid X_n = i\,\} \quad \text{y} \quad P_i = P\{\,X_0 = i\,\}$$

Demostración.

$P\{X_o = i_0, \ldots, X_n = i_n\} = \{X_n = i_n \mid X_o = i_o, \ldots, X_{n-1} = i_{n-1}\}$

$P\{X_0 = i_0, \ldots, X_{n-1} = i_{n-1}\} = P\{X_n = i_n \mid X_{n-1} = i_{n-1}\} P\{X_o = i_o, \ldots, X_{n-1}$

$= i_{n-1}\} = P_{i_{n-1}, i_n} \cdot P\{X_0 = i_0, \ldots, X_{n-1} = i_{n-1}\} = \ldots$

$= P_{i_{n-1}, i_n} \cdot P_{i_{n-2}, i_{n-1}} \ldots \cdot P_{i_0, i_1} \cdot P_i$

Para un proceso que hace referencia a los eventos x_{ji}, x_{jk} con $J_1 < J_2 < J_k$ se obtienen análogos resultados, usando términos de la forma $P\{X_0 = i, \ldots, X_n = i_n\}$ y efectuando las sumaciones correspondientes.

Una cadena de Markov está completamente definida por la matriz de probabilidad de transición de un paso y la distribución de las probabilidades del estado del proceso para $t = 0$. Al análisis de la cadena de Markov concierne principalmente el cálculo de las probabilidades de las posibles realizaciones del proceso.

Central en estos cálculos son las matrices de probabilidad de transición de n pasos.

$P^{(n)} = (P_{ij}^{n})$; P_{ij} es la probabilidad de que el proceso va del estado i al J en n pasos.

$$P_{ij}^{n} = P_r [X_{n+m} = J \mid X_m = i]$$

Supuesto el proceso homogéneo.

Proposición 2) Si la matriz de probabilidad de transición de un paso de una cadena de Markov es $P = (P_{ij})$.

$\Longrightarrow P^{n}_{ij} = \sum_{k=0}^{\infty} P^{r}_{ik} \; P^{s}_{kj}$ para cualquier par de enteros nonegativos fijos.

$$r + s = n, y, P^{0}_{ij} = \begin{cases} 1 & \text{si} \quad i = J \\ 0 & \text{si} \quad i \neq J \end{cases}$$

o sea que los números P^{n}_{ij} pueden ser tomados como los elementos de la matriz P^{n}, n-ésima potencia de P.

Demostración.

(Para $n = 2$). El evento de ir del estado i al estado J en dos pasos puede ser realizado en los caminos mutuamente excluyentes : Ir de i a K, y, luego de K a J por el estado intermedio K; $K = (0, 1, 2 ..)$

De la hipótesis Markoviana, la probabilidad de tales transiciones son P_{ik}, P_{kj} respectivamente. La ley de la probabilidad total permite establecer.

$$P^{2}_{ij} = \sum_{k=0}^{\infty} P^{r}_{ik} \cdot P^{s}_{kj} = \sum_{k=0}^{\infty} P_{ik} \; P_{kj}$$

Si la probabilidad del proceso inicialmente en el estado J es P_j, y, la ley de distribución de X_0 es $P\{X_0 = J\} = P_j$ entonces la probabilidad de que el proceso esté en estado K en el tiempo n es :

$$P^{(n)}_{k} = \sum_{j=0}^{\infty} P_j \; P^{n}_{Jk} = P[X_n = K]$$

Ejemplos :

1) A un paciente se han prescrito los medicamentos a y b con la condición de tomar uno cada día y de que a no debe ser tomado 2 días seguidos. El proceso determinado por el ejemplo es de Markov.

El espacio estado es $\{ a, b \}$.

La matriz de Markov asociada al proceso es :

$$\begin{array}{c c} & \begin{array}{cc} a & \quad b \end{array} \\ \begin{array}{c} a \\ \\ b \end{array} & \begin{bmatrix} 0 & 1 \\ \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \end{array} = P$$

La primera fila indica que a no es tomado 2 días seguidos, la segunda fila que después de tomar b puede ser tomada cualquiera : a ó b.

La probabilidad de que exactamente después de 3 días se pase de a a b puede ser visualizada en la matriz .

$$P^3 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \\ \frac{1}{4} & \frac{3}{4} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & \frac{3}{4} \\ \\ \frac{3}{8} & \frac{5}{8} \end{bmatrix}$$

Tal probabilidad es $\frac{3}{4}$

Supuesto que $P^{(0)} = [\frac{1}{2}, \frac{1}{2}]$.

$$P^{(1)} = P^{0} P = [\frac{1}{2}, \frac{1}{2}] \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} = [\frac{1}{4}, \frac{3}{4}]$$

$$P^{(2)} = P^{(1)} P = [\frac{1}{4}, \frac{3}{4}] \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} = [\frac{3}{8}, \frac{5}{8}]$$

$$P^{(3)} = P^2 P = [\frac{3}{8}, \frac{3}{8}] \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} = [\frac{5}{16}, \frac{11}{16}]$$

Después de 3 días la probabilidad que tome a es : $\frac{5}{16}$ y que tome b es : $\frac{11}{16}$.

Ejemplo 2 .

Una partícula está en uno de los puntos de : $[0, 1, 2, 3, 4]$; cambia de lugar de modo que se desplaza en una unidad a la derecha con probabilidad p, o ; una a la izquierda con probabilidad q, a menos que esté en 0 o en 4. Si está en 0 solo puede moverse a la derecha, si está en cuatro sólo a la izquierda. El proceso es de Markov.

Sea X_t la posición después de t pasos. $[X_0, X_1, \ldots, X_4]$ es el espacio estado. La matriz de Markov asociada al proceso es :

$$\begin{array}{c c} & \begin{array}{ccccc} X_0 & X_1 & X_2 & X_3 & X_4 \end{array} \\ \begin{array}{c} X_0 \\ X_1 \\ X_2 \\ X_3 \\ X_4 \end{array} & \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 \\ 0 & q & 0 & p & 0 \\ 0 & 0 & q & 0 & p \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \end{array}$$

El proceso se dice que constituye un camino aleatorio con barrera de reflexión.

Si se supone que el proceso se inicia con la partícula en el punto 2

$$P^{(0)} = (0 \quad 0 \quad 1 \quad 0 \quad 0)$$

$$P^{(1)} = P^{(0)}\, p = (\, 0 \quad q \quad 0 \quad 0 \quad p \quad 0\,)$$

$$P^{(2)} = P^{(1)}\, P = (\, q^2 \,,\, 0 \,,\, pq + qp \,,\, 0 \,,\, p^2\,)$$

Después de 2 pasos se tendrá :

La probabilidad de que la partícula esté en 0 es q^2 ; en 1 es 0; en 2 es $pq+qp$; en 3 es : 0; en 4 es p^2.

Ejemplo 3.

Un jugador tiene x pesos , apuesta un peso cada vez, ganando con probabilidad p y perdiendo con probabilidad q. El juego termina cuando ha perdido los x pesos o cuando ha ganado $y - x$.

Para $x = 2$, $y = 6$ la matriz de transición del proceso de Markov es :

$$\begin{array}{c|ccccccc} & X_0 & X_1 & X_2 & X_3 & X_4 & X_5 & X_6 \\ \hline X_0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ X_1 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \\ X_2 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ X_3 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ X_4 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ X_5 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ X_6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}$$

El espacio estado es $S = [X_0, X_1, \dots, X_6]$, X_i corresponde a tener i pesos.

Este es un ejemplo de camino aleatorio con barrera de absorción : X_0 y X_6 son estados absorbentes. El sistema permanece en el estado X_0 ó X_6 cuando entran en ellos. La distribución de probabilidad inicial es : $(0, 0, 1, 0, 0, 0, 0)$.

Un camino aleatorio unidimensional es una cadena de Markov, cuyo espacio estado es un subconjunto finito o infinito de enteros en los que se supone se realiza el movimiento de una partícula de modo que si la partícula está en el estado i puede en un paso o transición quedarse en i o colocarse en las posiciones : $(i - 1)$, ó $(i + 1)$.

Un camino aleatorio es simétrico en los enteros si es una cadena de Markov cuyo espacio estado son los enteros y cuya matriz de transición tiene sus elementos definidos así :

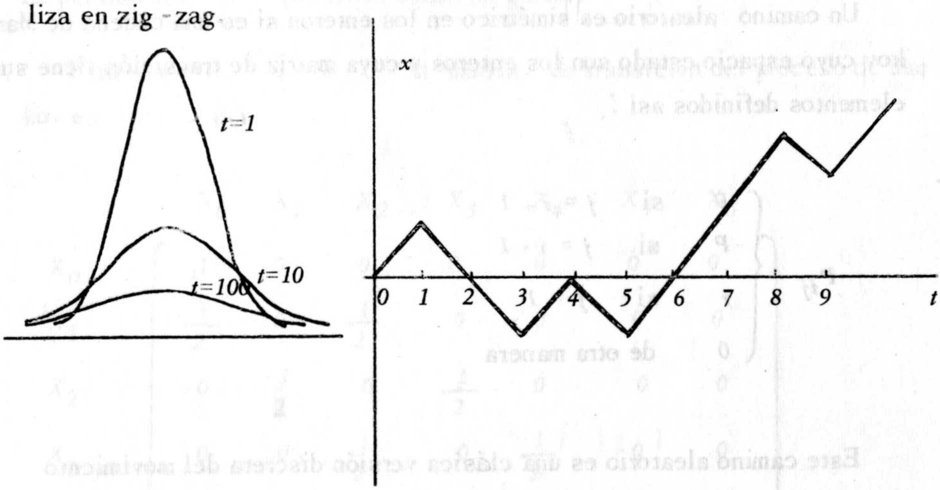
$$P_{ij} = \begin{cases} P & \text{si } j = i + 1 \\ P & \text{si. } j = i - 1 \\ r & \text{si } J = 1 \\ 0 & \text{de otra manera} \end{cases}$$

Este camino aleatorio es una clásica versión discreta del movimiento Browniano.

Una de las aplicaciones de mayor importancia de las cadenas de Markov ha sido la aparición del concepto del potencial probabilístico que representa la fusión de dos grandes ramas de la Física Teórica : la teoría probabilística de los procesos aleatorios como el movimiento Browniano movimiento de una partícula azotada por un bombardeo de moléculas y la teoría del potencial o de las funciones armónicas que aparecen en el análisis de estados de equili brio de un medio homogéneo como la distribución del calor en un sólido en equilibrio térmico.

Un primer paso en la interpretación probabilística del movimiento Browniano fue dado por Einstein en 1905, quien interpretó el fenómeno por medio de la mecánica estadística. Se estableció que la representación de la posición sucesiva observada en la partícula es una campana de Gauss.

En la gráfica se señalan, en el eje horizontal las distancias recorridas por la partícula en una dirección dada, supuesto que para el tiempo $t = 0$ la partícula está en $x = 0$. Para $t = 1$, $t = 10$, $t = 100$, las campanas de Gauss del movimiento Browniano tienen las representaciones que se observan a la izquierda. A la derecha una gráfica del modelo del movimiento Browniano debido a Wiener. Se deduce de él la continuidad del camino que se realiza en zig-zag.



De la teoría del flujo de calor se sabe que en un sólido homogéneo la temperatura en un punto P tiende a bajar si el promedio de temperaturas es inferior a la de P y tiende a subir en caso contrario. Si el cuerpo está en equilibrio térmico o sea que la temperatura en cualquier punto no cambia con el tiempo, la temperatura en P es el promedio de las temperaturas sobre la superficie de una pequeña esfera que rodea a P, lo que significa que la temperatura es una función armónica de las coordenadas de P.

La principal conexión entre las dos teorías hace relación con el problema denominado de Dirichlet. Supuesto que se mide la temperatura en todos los puntos de la superficie de un cuerpo en equilibrio térmico. Unos puntos están calientes, otros fríos, si el cuerpo es mantenido en el mismo estado es de esperar que el equilibrio térmico sea alcanzado en el interior. La temperatura en el interior varía de punto a punto, pero en cada punto no varía con el tiempo.

Con estas hipótesis es posible calcular la temperatura interior ? Con el problema se busca una función armónica definida en el interior del cuerpo, con ciertos valores conocidos en la superficie. La solución al problema de Dirichlet fue dada por el japonés Kakutani quien relacionó la situación con el movimiento de una partícula en movimiento Browniano y eventos esperados en un movimiento Browniano pueden ser visualizados en la configuración de un conductor de calor.

El punto de vista probabilístico ha clarificado y unificado los principios fundamentales de la teoría del potencial. Inversamente conceptos originados en la teoría del potencial aplicados a la teoría de las probabilidades, han mostrado la estructura analítica de los procesos de Markov.

BIBLIOGRAFIA

Brownian Motion and Potencial Theory Rev. Scientific American 1968	R. Hersche - R. Griego
A first Course in Stochastic Processes	Karlin
Calcul des Probabilités	A. Kaufmann
Random Walks	S. Kemeny
Probability	L. Lipschutz
Modern Probability Theory and its applications	E. Parzen