1. Procesos estocásticos

Un proceso estocástico, también llamado aleatorio, es uno en que el los resultados son inciertos. En contraste, en un sistema determinista la salida es siempre la misma dada una entrada en particular.

Proceso estocástico: Familia de variables aleatorias definidas sobre un espacio muestral E

$$\{X_t : E \to R, t \in I\}$$

El índice t suele representar tiempo (raramente espacio) e I es el conjunto de índices del proceso. Un proceso estocástico queda determinado por su espacio de estados E que representa todos los posibles resultados de las variables aleatorias (en la mayoría de los casos es común para todas) y su conjunto de índices I.
Al fijar uno de los elementos del espacio muestral e_0 se obtiene una función determinista $X(e_0 : T \to R)$ que permite ver una trayectoria del proceso y al fijar un tiempo t_0 se obtiene una variable aleatoria particular de la familia.
Los procesos estocásticos se pueden clasificar según la contabilidad de los conjuntos de estados e índices.

	E Discreto	E Continuo
T Discreto	Cadena de Markov	Sucesión de variables aleatorias continuas
T Continuo	Proceso puntual	Proceso continuo

Nota: La simulación (en particular la Monte Carlo) cumple un rol fundamental en el análisis de procesos estocásticos dado que permite aproximar probabilidades amparándose en la ley de los grandes números.

2. Cadenas de Markov

Cadena de Markov Sea E un conjunto a lo sumo infinito numerable. Una cadena de Markov es una secuencia de variables aleatorias $X_0, X_1, ..., X_n$ que asumen los valores en E con la propiedad de que

$$P(X_{n+1} = j/X_0 = x_o, X_1 = x_1, ..., X_n = x_{n-1}, X_n = i) = P(X_{n+1} = j/X_n = i) \quad \forall x_0, ..., x_{n-1}, i, j \in E, n \geq 0 \quad (1)$$

donde $X_n = i$ significa que la cadena visita al estado i en el tiempo n. En palabras la ecuación (1) quiere decir que el estado siguiente de la cadena depende únicamente del estado actual (y no de los anteriores).
Particularmente, una cadena de Markov es homogénea en el tiempo si la ecuación (1) no depende de n, es decir

$$P(X_{n+1} = j/X_n = i) = P(X_1 = j/X_0 = i) \quad (2)$$

Como las probabilidades en la ecuación (2) solamente dependen de i y j entonces se pueden disponer en una matriz cuadrada $P_{|E|\text{x}|E|}$ de forma tal que su entrada ij sea $P(i,j) = P(X_1 = j/X_0 = i)$. A esta matriz se la suele llamar *matriz (de probabilidades) de transición en un paso* o *Matriz estocástica.*

Nota: Las entradas de la matriz deben estar entre 0 y 1 y sus filas deben sumar uno para ser consistente con la definiciones de probabilidad. En símbolos:

- $P(i,j) \geq 0 \quad \forall i, j \in E$
- $\sum_{j \in E} P(i,j) = \sum_{j \in E} P(X_{n+1} = j/N_n = i) = 1$

La matriz de transición en un paso puede ser representada por un grafo dirigido cuyos nodos son los estados de la cadena y dos estados i y j están conectados si los une un arco dirigido etiquetado con $P(i,j)$. Todos los arcos salientes de un estado deben sumar 1 y por convención con arcos etiquetados con probabilidad 0 no se grafican.

2.1. Matriz de transición en n pasos

Un vector de probabilidad es un vector fila de número no negativos cuyas entradas suman 1. Se suelen usar letras griegas como α, λ y π para notarlos. Si X es una variable aleatoria discreta tal que $P(X = j) = \alpha_j \, \forall j = 1, 2, ...$ entonces $\alpha = (\alpha_1, \alpha_2, ...)$ es un vector de probabilidad. También es común decir que la distribución de X es α.
En particular para una cadena de Markov $\{X_0, X_1, ..\}$, la distribución de X_0 suele llamarse distribución inicial de la cadena de Markov. Si α es la distribución inicial entonces $P(X_0 = j) = \alpha_j \, \forall j$
Para cada estado i y j las probabilidades de que una cadena comience en un estado i y luego alcance j en n pasos se pueden ordenar en una matriz la cual llamaremos *matriz de transición en n pasos*. Intuitivamente, si $n = 1$ entonces dicha matriz en la matriz de transición en un paso. Uno de los principales resultados computacionales para cadenas de Markov es el siguiente (Ej: probar por inducción).

Sea $X_0, X_1, ...$ una cadena de Markov con matriz de transición (en un paso) P. La matriz P^n es la matriz de transición en n pasos de la cadena. En símbolos, para todo $n, m \geq 0$

$$P^n(i,j) = P(X_{n+m} = j / X_m = i) \quad \forall i, j \in E \tag{3}$$

2.2. Propiedad de Chapman-Kolmogorov

La siguiente propiedad se puede probar algebraicamente usando propiedades de la multiplicación entre matrices y el hecho de que estudiamos cadenas de Markov homogeneas en el tiempo.

$$P^{n+m}(i,j) = \sum_{k \in E} P^n(i,k) P^m(k,j) \quad \forall n, m \geq 0 \tag{4}$$

La interpretación probabilística de esta propiedad se sigue de que transicionar de un estado i a uno j en $m + n$ pasos es equivalente a transicionar de un estado i a un estado k en m pasos y luego transicionar de dicho estado k a un estado j en los n pasos restantes.

2.3. Distribución de X_n

Sea $X_0, X_1, ...$ una cadena de Markov con matriz de transición (en un paso) P y distribución inicial α. Para todo $n \geq 0$, la ditribución de X_n es αP^n, eso es

$$P(X_n = j) = (\alpha P^n)_j \quad \forall j \tag{5}$$

2.4. Distribución conjunta

Sea $X_0, X_1, ...$ una cadena de Markov con matriz de transición (en un paso) P y distribución inicial α. Para todo $0 \leq n_1 < n_2 < ... < n_k$ y estados $i_1, i_2, ..., i_k$

$$P(X_{n_1} = i_i, X_{n_2} = i_2, ..., X_{n_k} = i_k) = (\alpha P^{n_1})_{i1} (\alpha P^{n_2 - n_1})_{i_1 i2} ... (\alpha P^{n_k - n_{k-1}})_{ik-1 i_k} \tag{6}$$

3. Análisis de cadenas de Markov a largo plazo

En muchas casos las cadenas de Markov muestran un comportamiento límite a largo plazo. La cadena se estabiliza en una distribución límite o distribución invariante (no necesariamente lo mismo) que son independientes del estado inicial.

Sea $X_0, X_1, ...$ una cadena de Markov con matriz de transición (en un paso) P. Una distribución límite para la cadena de Markov es una distribución de probabilidad (vector fila) λ que cumple la propiedad de que para todo i,j

$$\lim_{n\to\infty} P^n(i,j) = \lambda_j \tag{7}$$

Otras formas equivalentes de enunciar lo anterior:

(i) Para cualquier distribución inicial, para todo j

$$\lim_{n\to\infty} P(X_n = j) = \lambda_j \tag{8}$$

(ii) Para cualquier distribución inicial α

$$\lim_{n\to\infty} \alpha P^n = \lambda \tag{9}$$

(iii) Sea Λ una matriz cuyas filas son iguales a λ

$$\lim_{n\to\infty} P(X_n = j) = \Lambda \tag{10}$$

En general, interpretamos a λ_j como la probabilidad de que la cadena alcance el estado j. Por la unicidad del límite, si una cadena de Markov tiene una distribución límite, entonces esta es única.
Hay diferentes formas de encontrar las distribuciones límites: un método "sucio.[es] calcular órdenes altos de la matriz P hasta que dos consecutivos coincidan y sean obviamente la distribución límite (todas filas iguales), otra forma de encontrar las distribuciones es a través de métodos numéricos (pero éstos tienen sus límites), y la última es encontrar soluciones exactas (se verán en la siguiente sección).

3.1. Distribución límite

La distribución límite da la probabilidad a largo plazo de visitas cada estado. También puede ser interpretada como la proporción a largo plazo del tiempo que la cadena visita cada estado.

Estas afirmaciones se pueden mostrar analíticamente valiéndose del lema de Cesaro: si una secuencia de números converge a un límite entonces la secuencia de promedios parciales también convergen a ese límite. En símbolos: si $x_n \to x$ mientras $n \to \infty$ entonces $\frac{x_1+x_2+...+x_n}{n} \to x$ si $n \to \infty$

3.2. Distribución invariante / estacionaria / de equilibrio

Sea $X_0, X_1, ...$ una cadena de Markov con matriz de transición (en un paso) P. Una distribución invariante es una distribución de probabilidad π que satisface

$$\pi = \pi P \tag{11}$$

o lo que es lo mismo:

$$\pi_j = \sum_i \pi_i P(i,j) \quad \forall j \tag{12}$$

3.2.1. Relación entre distribución estacionaria y distribución invariante

Lema: Si π es la distribución limite de una cadena de Markov con matriz de transición P entonces π es una distribución invariante.

Nota: el recíproco no es cierto. Contraejemplo: $P_{11} = P_{22} = 0$ y $P_{12} = P_{21} = 1$ (la cadena no converge, oscila) o la matriz I_{2x2} (el estado a largo plazo de la cadena depende del estado inicial y todos los vectores de probabilidad son distribuciones invariantes).

3.2.2. Matrices regulares

Matriz positiva Una matriz de transición P es positiva si todas sus entradas son mayores a 0

Matriz regular Una matriz de transición P es regular si alguna potencia de P es positiva. En símbolos,

$$P \quad \text{Regular} \leftrightarrow \exists\, n \geq 1 \ / \ P^n > 0 \tag{13}$$

Teorema (Distribución límite para matrices regulares): Una cadena de Markov cuya matriz de transición P es regular tiene una distribución límite que es única, positiva e invariante. En símbolos: existe un vector de probabilidad $\pi > 0$ tal que

$$\lim_{n\to\infty} P^n(i,j) = \pi_j \quad \forall i,j \ \text{ donde } \sum_i \pi_i P(i,j) = \pi_j \tag{14}$$

O de forma equivalente. Existe una matriz estocástica positiva Π tal que

$$\lim_{n\to\infty} P^n = \Pi \tag{15}$$

donde Π tiene filas iguales π y π es el único vector de probabilidad que satisface

$$\pi P = \pi$$

Nota: Si para alguna potencia n, todos los ceros en la matriz P^n aparecen en las mismas posiciones que los ceros en la matriz P^{n+1} entonces la matriz **no** es regular (porque los ceros seguirán apareciendo en las mismas posiciones en potencias mayores a n).

3.3. Técnicas para encontrar matrices invariantes

3.3.1. Tecnica 1

Sea π una distribución invariante para una cadena de Markov con matriz de transición P entonces verifica la ecuación (12), lo cual determina un sistema de ecuaciones lineales. Si P es una matriz de tamaño k$_x$k, el sistema tiene k ecuaciones con k incógnitas. Más aún, como las filas de P suman 1 entonces el sistema contiene una ecuación redundante.

$$\begin{cases} P_{11}\pi_1 + P_{12}\pi_1 + ... + P_{12}\pi_1 = \pi_1 \\ P_{21}\pi_2 + P_{22}\pi_2 + ... + P_{2k}\pi_2 = \pi_2 \\ P_{31}\pi_3 + P_{32}\pi_3 + ... + P_{3k}\pi_3 = \pi_3 \\ ... \\ P_{k1}\pi_k + P_{k2}\pi_k + ... + P_{kk}\pi_k = \pi_k \\ \pi_1 + \pi_2 + ... + \pi_k = 1 \end{cases}$$

El método consiste en:

1. Eliminar una ecuación redundante del sistema
2. Resolver el sistema de ecuaciones resultante para $x = (1, x_2, ...)$ en donde la primera (o cualquier) componente de x se reemplaza por un 1.

Este método reduce el problema a un sistema de ecuaciones de (k-1)$_x$(k-1). Si la cadena original tiene una distribución invariante única entonces el sistema de ecuaciones reducido tiene una solución única pero esta solución no es necesariamente un vector de probabilidades (no necesariamente suma 1). El resultado final $\pi = (\pi_1, \pi_2, ..., \pi_k)$ se puede obtener calculando

$$\pi = \frac{1}{1 + x_2 + ... + x_k}(1, x_2, ..., x_k)$$

3.3.2. Tecnica 2

La técnica de adivinar y luego probar si el vector de probabilidades encontrado por intuición satisface las restricciones del problema no debe ser subestimada. Si se cree que un vector π es la distribución invariante de una cadena de Markov con matriz de transición P entonces debe satisfacer la ecuación (11).

Complementario: ver la relación entre las distribuciones invariantes y los autovalores de la matriz de transiciones de una cadena de Markov.

3.4. Clasificación de estados

Alcanzabilidad Un estado j es alcanzable por un estado i si $P^n(i,j) \geq 0$, $p.a.\, n \geq 0$. El palabras, existe probabilidad positiva de llegar a j desde i en un número finito de pasos.

Comunicación Se dice que i y j se comunican o están comunicados si i es alcanzable desde j y j es alcanzable desde i.

Ejercicio: Probar que la çomunicación.^{es} una relación de equivalencia (reflexividad, simetría y transitividad).

Como la comunicación es una relación de equivalencia entonces el espacio de estados puede ser dividido en subconjuntos disjuntos en los cuales hay comunicación entre cualquier par de nodos pertenecientes a un subconjunto (una clase) pero no hay comunicación con nodos externos a dicho subconjunto.

Irreducibilidad Una cadena de Markov se dice irreducible si tiene exactamente una clase de comunicación (un par de nodos cualquiera está comunicado).

3.4.1. Recurrencia y transitoriedad

Tiempo (mínimo) hasta la primer visita Dada una cadena de Markov $X_0, X_1, ...$ se define T_j como

$$T_j = min\{n > 0 : X_n = j\} \tag{16}$$

T_j representa la mínima cantidad de transiciones que se requieren para llegar a j. Si $X_n \neq j \quad \forall n > 0$ entonces $T_j = \infty$

Probabilidad de revisitar Se define f_j como la probabilidad de que una cadena comience en un estado j y eventualmente lo visite de nuevo (por primera vez). En símbolos, $f_j = P(T_j < \infty / X_0 = j)$

Estados recurrentes Un estado j se dice recurrente si una cadena de Markov comienza en j y eventualmente visita a j de nuevo. En símbolos: $f_j = 1$

Estados transitorios Un estado j se dice transitorio si existe probabilidad positiva de que una cadena de Markov comience en j y nunca revisite j. En símbolos: $f_j < 1$

Con el objetivo de encontrar la relación entre las distribuciones a las que tienden las cadenas de Markov y la frecuencia con la que se visitan los estados de la misma definimos algunos coneptos.

Llegar de un nodo i a un nodo j en un número finito de pasos Definimos I_n como

$$I_n = \begin{cases} 1 & X_n = j \\ 0 & \text{si no} \end{cases} \tag{17}$$

para algún $n \geq 0$. Y resulta que $\sum_{n=0}^{\infty} I_n$ es el número de visitas que recibe j durante la evolución de la cadena.

Por otro lado, la cantidad de visitas a un estado esperada se puede calcular como

$$E(\sum_{n=0}^{\infty} I_n) = \sum_{n=0}^{\infty} E(I_n) = \sum_{n=0}^{\infty} P(X_n = j/X_0 = i) = \sum_{n=0}^{\infty} P^n(i,j)$$

Observaciones:

1. La sumatoria infinita puede diverger a $+\infty$
2. Si la cadena comienza en el estado i entonces la cantidad esperada de visitas al estado j, que se puede notar como N_j es la entrada ij de la matriz $\sum\limits_{n=0}^{\infty} P^n$

Supongamos que el estado j es recurrente. La cadena que comienza en el estado j va a revisitar el estado j con probabilidad 1. Una vez que lo hace, la cadena comienza de nuevo y se comporta exactamente igual que al comienzo. Se suele decir que la cadena se regenera y esto se corresponde con *La propiedad fuerte de Markov*. En símbolos:

$$\sum_{n=0}^{\infty} P^n(j,j) = \infty$$

Ahora supongamos que j es transitorio. Empezando en j, la probabilidad de revisitar j es f_j y la probabilidad de nunca más visitar j es su complemento, osea $1 - f_j$. Si una cadena visita j, el hecho de que revisite j es independiente de los estados anteriores. Entonces la secuencia de visitas sucesivas a j se puede modelar como una secuencia de variables aleatorias idénticamente distribuidas de tiros de moneda en los cuales sale 'cara' si j va a ser revisitado y 'cruz' en caso contrario. Más aún, la cantidad de veces que j es visitado es la cantidad de tiros de moneda hasta que salga 'cruz', lo cual tiene una distribución geométrica de parámetro $1 - f_j$. Por lo tanto, el número esperado de visitas al estado j es $\frac{1}{1-f_j}$ y

$$\sum_{n=0}^{\infty} P^n(j,j) = \frac{1}{1-f_j} < \infty$$

En palabras, un estado transitorio va a ser visitado una cantidad finita de veces.

Estados recurrentes Un estado j se dice recurrente sii

$$\sum_{n=0}^{\infty} P^n(j,j) = \infty$$

Estados transitorios Un estado j se dice transitorio sii

$$\sum_{n=0}^{\infty} P^n(j,j) = \frac{1}{1-f_j} < \infty$$

Observaciones:

1. Si un estado j es recurrente y alcanzable desde un estado i entonces hay probabilidad positiva de transicionar de i a j y luego desde j la cantidad esperada de revisitas es ∞ por lo tanto el número esperado de visitas a j desde i es ∞.
2. Con un razonamiento análogo, si j es transitorio y alcanzable desde i entonces la cantidad esperada de visitas al estado j comenzando desde i es finita. Más aún $\lim\limits_{n\to\infty} P^n(i,j) = 0$. En palabras, la probabilidad a largo plazo de que la cadena de markov eventualmente visite un estado transitorio es 0.

Teorema (La recurrencia y la transitoriedad son propiedades de clase) Los estados de una clase de comunicación son o bien todos recurrentes o bien todos transitorios.

Corolario En una cadena de Markov irreducible finita, todos los estados son recurrentes.

Nota: El corolario no vale para cadenas de Markov infinitas. Contraejemplo: caminata aleatoria en una recta numérica. Este resultado solo vale para 2 dimensiones, no vale para dimensiones mayores iguales a 3. 'Una persona ebria siempre encuentra el camino de regreso pero un pájaro ebrio puede perderse para siempre'.

3.4.2. Descomposición canónica

Conjunto cerrado Un conjunto de estados C se dice cerrado si ningún estado fuera de C es accesible desde algún estado de C. En símbolos, $P(i,j) = 0 \quad \forall i \in C, j \notin C$

Lema: Una clase de comunicación es cerrada si consiste en estados recurrentes únicamente. Una clase de comunicación finita es cerrada si y sólo si está formada solamente por estados recurrentes.

El espacio de estados E de una cadena de Markov finita puede ser particionado en estados recurrentes y transitorios como sigue

$$E = T \cup R_1 \cup R_2 \cup ... \cup R_m \tag{18}$$

donde T es el conjunto de todos los estados transitorios y R_i es uno de los m conjuntos de estados recurrentes cerrados.

Dada una descomposición canónica, el espacio de estados se puede reordenar de forma que la matriz de transición quede de la forma

$$P = \begin{array}{c c} & \begin{array}{cccc} T & R_1 & \dots & R_m \end{array} \\ \begin{array}{c} T \\ R_1 \\ \vdots \\ R_m \end{array} & \begin{pmatrix} * & * & \dots & * \\ 0 & P_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & P_m \end{pmatrix} \end{array}$$

Donde cada submatriz $P_1, ..., P_m$ es una matriz cuadrada estocástica que se correcponde con una clase de counicación cerrada. La descomposición canónica permite visualizar fácilmente el comportamiento a largo plazo de una cadena de Markov. Para cada $n \geq 1$ se tiene

$$P^n = \begin{array}{c c} & \begin{array}{cccc} T & R_1 & \dots & R_m \end{array} \\ \begin{array}{c} T \\ R_1 \\ \vdots \\ R_m \end{array} & \begin{pmatrix} * & * & \dots & * \\ 0 & P_1^n & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & P_m^n \end{pmatrix} \end{array}$$

$$\lim_{n\to\infty} P^n = \begin{array}{c c} & \begin{array}{cccc} T & R_1 & \dots & R_m \end{array} \\ \begin{array}{c} T \\ R_1 \\ \vdots \\ R_m \end{array} & \begin{pmatrix} * & * & \dots & * \\ 0 & \lim\limits_{n\to\infty} P_1^n & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lim\limits_{n\to\infty} P_m^n \end{pmatrix} \end{array}$$

Nota: Las clases de comunicación cerrada recurrentes $R_1, ..., R_m$ se comportan como pequeñas cadenas de Markov irreducibles donde todo par de estados está comunicado.

3.5. Cadenas de Markov Irreducibles

Teorema (Teorema del límite para cadenas de Markov irreducibles finitas) Sea $X_0, X_1, ...$ una cadena de Markov con matriz de transición (en un paso) P. Por cada estado j sea $\mu_j = E(T_j.X_0 = j)$ el tiempo esperado de retorno a j. Entonces, μ_j es finito y existe una distribución de probabilidad positiva, única π tal que

$$\pi_j = \frac{1}{\mu_j} \quad \forall j \tag{19}$$

Más aún, para todos los estados i

$$\pi_j = \lim_{n\to\infty} \frac{1}{n} \sum_{m=0}^{n-1} P^n(i,j) \tag{20}$$

Observaciones:

1. Si se visita al estado j cada μ_j pasos entonces la proporción de visitas a j es $\frac{1}{\mu_j}$.
2. π no necesariamente es una distribución límite. (Ver sección de cadenas de Markov ergódicas)
3. El teorema se puede extender a cadenas de Markov infinitas irreducibles.

Los estados recurrentes se suelen subclasificar según $E(T_j/X_0 = j)$

Un estado j recurrente puede ser

Positivo recurrentes Si el tiempo promedio/esperado entre dos visitas consecutivas a él es finito. En símbolos $E(T_j/X_0 = j) < \infty$

Nulo recurrentes Si el tiempo promedio/esperado que pasa entre dos visitas consecutivas a él es infinito. En símbolos $E(T_j/X_0 = j) = \infty$

El tiempo esperado de retorno a un estado se puede calcular

1. como el recíproco de la probabilidad invariante π_j
2. condicionando el primer paso de la cadena y haciendo uso de la propiedad de variables aleatorias $E(Y) = E(E(Y/X))$ (análisis del primer paso)

3.6. Periodicidad

Las cadenas de Markov tienen distribuciones invariantes positivas únicas. Aunque no tienen distribuciones límite, tienen un comportamiento de tipo límite en el sentido de que para todos los estados i,j los promedios parciales $\frac{1}{n}\sum_{m=0}^{n-1} P^n(i,j)$ convergen.

Período Para una cadena de Marcov con matriz de transición P, el período de un estado i, $d(i)$ es el máximo común divisor del conjunto de todos los posibles tiempos de retorno al nodo i. En símbolos,

$$d(i) = mcd(n > 0 : P^n(i,i) > 0) \tag{21}$$

Si $d(i) = 1$ entonces se dice que el estado i es *aperiodico*. Si el conjunto de los posibles tiempos de retorno al nodo es vacío entonces $d(i) = +\infty$

En otras palabras, el período de un estado i indica que solamente se puede revisitar el nodo i en múltiplos de $d(i)$ pasos. Además $d(i)$ es el mayor número que cumple dicha propiedad.

Lema (El período es una propiedad de clase): Los estados de una misma clase de comunicación tienen el mismo período.

Nota: En particular, todos los estados de una cadena de Markov irreducible tienen el mismo período.

Cadenas de Markov periódicas / aperiódicas Una cadena de Marcov se dice periódica si es irreducible y todos sus estados tienen período mayor a 1. Una cadena de Marcov se dice aperiódica si es irreducible y todos sus estados tienen período igual a 1.

Observación: cualquier estado i que satisfaga $P(i,i) > 0$ es aperiódico. Más aún $P(i,i) > 0 \;\; p.a\, i$ es condición suficiente para que una cadena de Markov irreducible sea aperiódica. En otras palabras, si al menos una de las entradas en las diagonales de la matriz es distinta de cero entonces la cadena de Marcov irreducible es aperiódica.

3.7. Cadenas de Markov ergódicas

Ergodicidad Una cadena de Markov es ergódica si es irreducible, aperiódica y todos los estados tienen un tiempo de retorno finito. La última condición es siempre verdadera en el caso de cadenas finitas. Entonces en particular, una cadena de Markov finita es ergódica si es irreducible y aperiódica.

Teorema fundamental del límite para cadenas de Markov ergódicas Sea $X_0, X_1, ...$ una cadenas de Markov ergódica. Existe una distribución invariante positiva única π que es la distribución límite de la cadena. En símbolos,

$$\pi_j = \lim_{n\to\infty} P^n(i,j) \quad \forall i,j \tag{22}$$

3.8. Cadenas de Markov absorventes

Estado absorvente: Un estado i es absorvente si $P(i,i) = 1$

Cadena de Markov absorvente: Una cadena de Markov es absorvente si si contiene al menos 1 estado absorvente.

Descomposición canónica II Se considera una cadena de Markov absorvente de k estados de los cuales t son estados transitorios y $(k-t)$ estados absorventes. Los estados pueden ser reordenados (como en la descomposición canónica de la sección Recurrencia y Transitoriedad) de la siguiente forma

$$P = \left(\begin{array}{c|c} Q & R \\ \hline 0 & I \end{array}\right)$$

donde Q es una matriz de txt, R es una matriz de tx(k-t), 0 es una matriz de (k-t)xt de ceros e I es la matriz identidad de (k-t)x(k-t). Este orden particular es útil para calcular potencias de la matriz. Para $n \geq 1$ se tiene

$$P^n = \left(\begin{array}{c|c} Q^n & (1 + Q + ... + Q^{n-1})R \\ \hline 0 & I \end{array}\right)$$

Lema (álgebra lineal) Sea A una matriz cuadrada que verifica $A^n \to 0$ si $n \to \infty$ entonces

$$\sum_{n=0}^{\infty} A^n = (I - A)^{-1} \tag{23}$$

Usando (23) y la descomposición canónica, es posible calcular la probabilidad a largo plazo de una cadena de Marcov como

$$\lim_{n\to\infty} P^n = \left(\begin{array}{c|c} 0 & (I-Q)^{-1}R \\ \hline 0 & I \end{array}\right)$$

La submatriz $(I-Q)^{-1}R$ está indexada por filas transitorias y columnas absorventes. Se puede interpretar como que la entrada ij es la probabilidad de que una cadena comience en el estado i y sea absorvido en el estado j. En particular, si la cadena tiene un único estado absorvente entonces la submatriz $(I-Q)^{-1}R$ es un vector columna de $(k-1)$ unos.

Nota: Se le suele llamar matriz fundamental a $(I-Q)^{-1}$

3.8.1. Numero esperado de visitas a estados transitorios y tiempo esperado de absorción

Teorema (Número esperado de visitas a estados transitorios) Se considera una cadena de Markov con t estados transitorios. Sea F una matriz de t filas y t columnas indexada por estados transitorios donde $F(i, j)$ es el número esperado de visitas al estado j dado que la cadena comenzó en i. Entonces

$$F = (I - Q)^{-1} \tag{24}$$

Tiempo esperado de absorción Si una cadena de Markov comienza en el estado i, se define a_i como el tiempo esperado de absorción. La cantidad de transiciones desde un estado transitorio i a un estado absorvente es la suma de las transiciones desde i a cada uno de los estados absorventes hasta la eventual absorción. La cantidad esperada de transiciones desde un estado i a un estado transitorio j es $F(i, j)$. Luego

$$a_i = \sum_{k \in T} F(i, k) \tag{25}$$

O equivalentemente $a = F1$ donde 1 es un vector columna de unos. En otras palabras, el tiempo de absorción esperado es la suma de las filas de la matriz fundamental.

Resumen En una cadena de MArkov absorvente todos los estados son transitorios y/o absorventes. Sea $F = (I - Q)^{-1}$

1. (Probabilidad de absorción) La probabilidad de que un estado j absorva la cadena a partir de un estado transitorio i es $(FR)_{i,j}$
2. (Tiempo de absorción) La cantidad esperada de transiciones desde un estado transitorio i hasta que la cadena es absorvida por un estado absorvente es $(F1)_i$

3.8.2. Tiempo esperado de visita para cadenas de Markov irreducibles

Se quiere encontrar la cantidad de transiciones esperadas hasta que un estado i se visite por **primera vez** en una cadena de Markov irreducible.
Sea P la matriz de transición de una cadena de Markov. Se considera una nueva cadena en donde el estado i es absorvente. En esta nueva cadena

1. La matriz de transición $\tilde{P}$ va a estar dada por la matriz P con ceros en la i-ésima fila e i-ésima columna excepto por la posición ii que contiene un 1.
2. La matriz resultante Q (con los estados transitorios) se va a obtener de eliminar de P la i-ésima fila y la i-ésima columna de P (o de $\tilde{P}$).

Entonces la cantidad de transiciones esperadas hasta que un estado i se visite por primera vez en la cadena original es igual a la cantidad de transiciones esperadas hasta que se absorva la nueva cadena en i.

3.9. Miscelánea

3.9.1. La recurrencia positiva y la recurrencia nula son propiedades de clase

Todos los estados en una clase de comunicación recurrente son o bien positivos recurrentes o bien nulos recurrentes (pero no ambos).

3.9.2. Proposición

Sea P la matriz de transición de una cadena de Markov finita. La cadena es ergódica si y solo si P es regular.

Lema Si i es un estado aperiódico entonces existe un entero positivo N tal que $P^n(i, i) \geq 0 \quad \forall n \geq N$

3.10. Temas complementarios

1. Reversibilidad en el tiempo
2. Regeneración y la propiedad (fuerte) de las cadenas de Markov

Índice