Comportamiento asintótico de las Cadenas de Markov

J. M. Almira

Universidad de Murcia

Cadenas de Markov irreducibles

Sea (X_n) una CM finita con espacio de estados $S = \{s_1, \cdots, s_k\}$.

Definición

Decimos que el estado s_i se comunica con el estado s_j (y lo denotamos por $s_i \rightarrow s_j$) si existe un valor $n \in \mathbb{N}$ tal que

$$P[X_{m+n} = s_j | X_m = s_i] > 0.$$

Cadenas de Markov irreducibles

Definición

Decimos que el estado s_i se comunica con el estado s_j (y lo denotamos por $s_i \to s_j$) si existe un valor $n \in \mathbb{N}$ tal que

$$P[X_{m+n} = s_j | X_m = s_i] > 0.$$

Nota: Evidentemente,

$$P[X_{m+n} = s_j | X_m = s_i] = (P^n)_{i,j}$$

(Independientemente de m).
Recuérdese que $(A)_{ij}$ representa la entrada a_{ij} de la matriz A

Cadenas de Markov irreducibles

Definición

Decimos que el estado s_i se comunica con el estado s_j (y lo denotamos por $s_i \to s_j$) si existe un valor $n \in \mathbb{N}$ tal que

$$P[X_{m+n} = s_j | X_m = s_i] > 0.$$

Definición (Cadena de Markov irreducible)

Decimos que los estados s_i, s_j se intercomunican si $s_i \to s_j$ y $s_j \to s_i$. (En tal caso usamos la notación $s_i \leftrightarrow s_j$). La cadena de Markov (X_n) es irreducible si $s_i \leftrightarrow s_j$ para todo par de estados $s_i, s_j \in S$. Si la CM no es irreducible, decimos que es reducible.

Cadenas de Markov irreducibles

Definición

Decimos que el estado s_i se comunica con el estado s_j (y lo denotamos por $s_i \to s_j$) si existe un valor $n \in \mathbb{N}$ tal que

$$P[X_{m+n} = s_j | X_m = s_i] > 0.$$

Definición (Cadena de Markov irreducible)

Decimos que los estados s_i, s_j se intercomunican si $s_i \to s_j$ y $s_j \to s_i$. (En tal caso usamos la notación $s_i \leftrightarrow s_j$). La cadena de Markov (X_n) es irreducible si $s_i \leftrightarrow s_j$ para todo par de estados $s_i, s_j \in S$. Si la CM no es irreducible, decimos que es reducible.

Teorema (Caracterización de cadenas de Markov irreducibles)

La cadena de Markov (X_n) es irreducible si y solo si para todo par de estados $s_i, s_j \in S$ existe un número natural n tal que $(P^n)_{i,j} > 0$.

Cadenas de Markov irreducibles

Definición (Grafo de transición de una CM)

Dada (X_n) *CM con espacio de estados* $S = \{s_1, \cdots, s_k\}$ *y matriz de transición* $P = (p_{ij})$*, llamamos grafo de transición asociado a la CM al* <u>*grafo dirigido*</u> $G = (V, E)$*, donde*

$$\left\{ \begin{array}{ccc} V & = & S \\ (s_i, s_j) \in E & \Leftrightarrow & p_{ij} > 0 \end{array} \right.$$

Teorema

La cadena de Markov (X_n) *es irreducible si y solo si su grafo de transición es conexo.*

Cadenas de Markov irreducibles

Ejemplo

Se puede comprobar fácilmente que la CM con matriz de transición

$$P = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/3 & 2/3 & 0 & 0 \\ 0 & 0 & 1/8 & 7/8 \\ 0 & 0 & 4/5 & 1/5 \end{bmatrix}$$

es reducible.

Aperiodicidad

Definición

Recordemos que el mcd (máximo común divisor) de un conjunto de números enteros, que denotamos por $mcd\{a_1, a_2, \cdots\}$, *es el mayor número natural* d *tal que* $d|a_i$ *para todo* i.

Definición (Periodo de un estado de la CM)

Sea (X_n) *CM con matriz de transición* P *y sea* $s_i \in S$ *uno de sus estados. Llamamos periodo del estado* s_i *al número*

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

Si $d(s_i) = 1$ *decimos que el estado* s_i *es aperiódico.*

Definición (Cadena de Markov aperiódica)

La CM (X_n) *se dice aperiódica si todos sus estados son aperiódicos. En otro caso (si contiene estados periódicos) decimos que la CM es periódica.*

Aperiodicidad

Dadas las matrices de transición P_1, P_2, P_3, decidir -de manera justificada- si las Cadenas de Markov asociadas son periódicas o aperiódicas:

$$P_1 = \begin{bmatrix} 1/2 & 1/2 \\ 1/3 & 2/3 \end{bmatrix}; P_2 = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/3 & 2/3 & 0 & 0 \\ 0 & 0 & 1/8 & 7/8 \\ 0 & 0 & 4/5 & 1/5 \end{bmatrix}$$

$$P_3 = \begin{bmatrix} 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \end{bmatrix}$$

Aperiodicidad

Teorema

Sea (X_n) *una CM con espacio de estados* $S = \{s_1, \cdots, s_k\}$ *y matriz de transición* P*. Las siguientes afirmaciones son equivalentes:*

- (X_n) *es irreducible y aperiódica.*
- *Existe* $N < \infty$ *tal que*

$$(P^n)_{i,j} > 0 \text{ para todo } i, j \in \{1, \cdots, k\} \text{ y todo } n \geq N.$$

Distribuciones estacionarias

Definición

Sea (X_n) *una CM finita con espacio de estados*

$$S = \{s_1, \cdots, s_k\}$$

y matriz de transición P*.*
Sea $\mu = (\mu_1, \cdots, \mu_k)$ *una distribución de probabilidad:*

$$0 \leq \mu_i \leq 1 \; \textit{para todo } i \,, \; \sum_{i=1}^{k} \mu_i = 1$$

Decimos que μ *es una distribución estacionaria (o de equilibrio) para la cadena de Markov* (X_n) *si*

$$\mu P = \mu$$

Distribuciones estacionarias

Teorema (Fundamental de las Cadenas de Markov)

(a) *Toda cadena de Markov admite al menos una distribución estacionaria.*

(b) *Si la cadena de Markov es irreducible y aperiódica, entonces la distribución estacionaria μ es única y, además, satisface:*

$$lim_{n\to\infty} d_{TV}(\mu^{(n)}, \mu) = 0,$$

para toda distribución de probabilidad inicial $\mu^{(0)}$.

Nota: Dadas $\mu = (\mu_1, \cdots, \mu_k)$, $\eta = (\eta_1, \cdots, \eta_k)$ dos distribuciones de probabilidad, se define su distancia (en variación total) como:

$$d_{TV}(\mu, \eta) = \frac{1}{2}\sum_{i=1}^{k} |\mu_i - \eta_i|$$

Distribuciones estacionarias

Obsérvese que d_{TV} es una buena forma de medir distancias entre probabilidades porque, si μ, η son distribuciones de probabilidad , entonces

- $\mu = \eta$ si y solo si $d_{TV}(\mu, \eta) = 0$,
- $0 \leq d_{TV}(\mu, \eta) \leq 1$,
- $d_{TV}(\mu, \eta) = 1$ si y solo si ambas distribuciones tienen soportes disjuntos.
- $$d_{TV}(\mu, \eta) = \max_{A \subseteq S} |P_\mu(A) - P_\eta(A)|$$

Existencia de distribuciones estacionarias (o de equilibrio)

Vamos a dar una demostración basada en un Teorema de Punto Fijo

- Obsérvese que si definimos $T(x) = xP$, entonces las distribuciones estacionarias de la CM son puntos fijos de T.
- Concretamente, necesitamos que $T(\mu) = \mu$ con las condiciones adicionales: $\mu \geq 0$ (es decir, todas sus componentes son ≥ 0) y $\sum_{i=1}^{k} \mu_i = 1$. Ahora bien, esto significa que

$$\mu \in K = \{x \in \mathbb{R}^k : \sum_{i=1}^{k} x_i = 1 \ , 0 \leq x_i \text{ para todo i}\}$$

Existencia de distribuciones estacionarias (o de equilibrio)

Obsérvese que K es un subconjunto (no vacío) de $\mathbb{R}^k$ convexo y compacto.

Definición

Sea $K \subseteq \mathbb{R}^k$ convexo. Decimos que $f : K \to K$ es una aplicación afín si

$$f(tx + (1-t)y) = tf(x) + (1-t)f(y)$$

siempre que $x, y \in K$ y $0 \leq t \leq 1$.

Teorema (del punto fijo de Markov-Kakutani)

Sea $K \subseteq \mathbb{R}^k$ convexo, compacto y no vacío. Supongamos que $f : K \to K$ es una aplicación continua y afín. Entonces f tiene al menos un punto fijo en K.

Existencia de distribuciones estacionarias (o de equilibrio)

Es evidente que $T(x) = xP$ es lineal (y, por tanto, afín, cuando se restringe a cualquier convexo) y continua. Pero debemos probar que $T(K) \subseteq K$. Sea $x = (x_1, \cdots, x_k) \in K$. Entonces:

- Como $x, P \geq 0$, concluimos que también $xP \geq 0$.
- La componente i-ésima de $T(x) = xP$ viene dada por:

$$(xP)_i = \sum_{j=1}^{k} x_j p_{ji}$$

Por tanto,

$$\sum_{i=1}^{k} (xP)_i = \sum_{i=1}^{k} \left(\sum_{j=1}^{k} x_j p_{ji} \right) = \sum_{j=1}^{k} x_j \left(\sum_{i=1}^{k} p_{ji} \right) = \sum_{j=1}^{k} x_j = 1$$

Existencia de distribuciones estacionarias (o de equilibrio)

Demostración del Teorema del punto fijo de Markov-Kakutani:

- Sea $x \in K$. Consideremos los promedios $x_N = \frac{1}{N}\sum_{i=0}^{N-1} f^i(x)$.
- Como K convexo y $x, f(x), f(f(x)) = f^2(x), \cdots, f^N(x) \in K$, está claro que $x_N \in K$ para todo N.
- Como K compacto, existe una subsucesión convergente, $\{x_{n_k}\} \to x^* \in K$.

Existencia de distribuciones estacionarias (o de equilibrio)

- Veamos que $f(x^*) = x^*$

$$
\begin{aligned}
\|x_{n_k} - f(x_{n_k})\| &= \|\frac{1}{n_k}\sum_{i=0}^{n_k-1} f^i(x) - f\left(\frac{1}{n_k}\sum_{i=0}^{n_k-1} f^i(x)\right)\| \\
&= \frac{1}{n_k}\|\sum_{i=0}^{n_k-1} f^i(x) - \sum_{i=0}^{n_k-1} f^{i+1}(x)\| \text{ (por ser } f \text{ afín)} \\
&= \frac{1}{n_k}\|x - f^{n_k}(x)\| \\
&\leq \frac{1}{n_k} 2 \sup_{z\in K}\|z\| \to 0 \text{ para } k \to \infty.
\end{aligned}
$$

Se sigue que $\|x^* - f(x^*)\| = \lim_{k\to\infty}\|x_{n_k} - f(x_{n_k})\| = 0$
Esto finaliza la prueba.

Comportamiento asintótico de cadenas de Markov irreducibles y aperiódicas

Ya hemos probado:

Teorema (Fundamental de las Cadenas de Markov-Parte (a))

Toda cadena de Markov admite al menos una distribución estacionaria μ.

Comportamiento asintótico de cadenas de Markov irreducibles y aperiódicas

Vamos a probar:

Teorema (Fundamental de las Cadenas de Markov-Parte (b))

Si la cadena de Markov es irreducible y aperiódica, entonces solo admite una distribución estacionaria μ y, además, se satisface que:

$$lim_{n\to\infty} d_{TV}(\mu^{(n)}, \mu) = 0,$$

para toda distribución de probabilidad inicial $\mu^{(0)}$.

Comportamiento asintótico de cadenas de Markov irreducibles y aperiódicas

Recordemos: Si (X_n) es un cadena de Markov con $X_0 \sim \mu^{(0)}$ y matriz de transición P, se pueden definir funciones ϕ, ψ de modo que:

$$\begin{cases} X_0 = \phi_{\mu^{(0)}}(U_0) \\ X_k = \psi(X_{k-1}, U_k),\ k = 1, 2, \cdots . \end{cases}$$

donde (U_n) es una sucesión de v.a. independientes, $U_n \sim \text{Unif}[0, 1]$, $n \in \mathbb{N}$.

Comportamiento asintótico de cadenas de Markov irreducibles y aperiódicas

Vamos a estudiar $\mu^{(n)}$ cuando (X_n) es irreducible y aperiódica:

- Denotemos por η una distribución estacionaria de la Cadena de Markov (cuya existencia acabamos de probar).
- Podemos asumir que (X_n) se ha obtenido mediante las fórmulas:

$$\begin{cases} X_0 = \phi_{\mu^{(0)}}(U_0) \\ X_k = \psi(X_{k-1}, U_k), \; k = 1, 2, \cdots . \end{cases}$$

Comportamiento asintótico de cadenas de Markov irreducibles y aperiódicas

- Consideremos ahora una nueva Cadena de Markov (X'_n) que construimos tomando como distribución de probabilidad inicial la distribución estacionaria η.
- Usamos una nueva sucesión (U'_n) de v.a. índependientes, idéntic. distrib. con distribución $U([0,1])$:

$$\begin{cases} X'_0 = \phi_\eta(U'_0) \\ X'_k = \psi(X'_{k-1}, U'_k) \end{cases}$$

Comportamiento asintótico de cadenas de Markov irreducibles y aperiódicas

$$\begin{cases} X_0 = \phi_{\mu^{(0)}}(U_0) \\ X_k = \psi(X_{k-1}, U_k) \end{cases} \quad \begin{cases} X'_0 = \phi_\eta(U'_0) \\ X'_k = \psi(X'_{k-1}, U'_k) \end{cases}$$

- Como η es estacionaria, $X'_n \sim \eta$ para todo n.
- Las cadenas de Markov (X_n) y (X'_n) son independientes por serlo (U_n) y (U'_n) entre ellas.

Comportamiento asintótico de cadenas de Markov irreducibles y aperiódicas

Consideremos la v.a.

$$T = \min\{n : X_n = X'_n\}.$$

($T = +\infty$ si las cadenas nunca se encuentran).

Lema

Se tiene que:

$$P[T < \infty] = 1$$

O, lo que es lo mismo,

$$lim_{n\to\infty} P(T > n) = 0$$

Comportamiento asintótico de cadenas de Markov irreducibles y aperiódicas

Demostración del Lema: Como (X_n) irreducible y aperiódica, sabemos que existe $M > 0$ tal que

$$(P^M)_{i,j} > 0 \text{ para todo } i,j \in \{1, \cdots, k\}.$$

Sea $0 < \alpha = \min_{i,j}(P^M)_{i,j}$.

Comportamiento asintótico de cadenas de Markov irreducibles y aperiódicas

$$
\begin{aligned}
P(T \leq M) &\geq P(X_M = X'_M) \geq P(X_M = s_1, X'_M = s_1) \\
&= P(X_M = s_1)P(X'_M = s_1) \text{ por la independencia de } X_M \text{ y } X'_M \\
&= \left(\sum_{i=1}^{k} P(X_M = s_1, X_0 = s_i)\right)\left(\sum_{i=1}^{k} P(X'_M = s_1, X'_0 = s_i)\right) \\
&= \left(\sum_{i=1}^{k} P(X_M = s_1 | X_0 = s_i) P(X_0 = s_i)\right) \times \\
&\quad \times \left(\sum_{i=1}^{k} P(X'_M = s_1 | X'_0 = s_i) P(X'_0 = s_i)\right) \\
&\geq \left(\alpha \sum_{i=1}^{k} P(X_0 = s_i)\right)\left(\alpha \sum_{i=1}^{k} P(X'_0 = s_i)\right) = \alpha^2
\end{aligned}
$$

Comportamiento asintótico de cadenas de Markov irreducibles y aperiódicas

Acabamos de probar que: $P(T \leq M) \geq \alpha^2$. Por tanto:

$$P(T > M) \leq 1 - \alpha^2$$

También hemos probado que $P(X_M = X'_M) \geq \alpha^2$. Pero entonces es claro que

$$P(X_{2M} = X'_{2M} | T > M) \geq \alpha^2$$

Y, por tanto,

$$P(X_{2M} \neq X'_{2M} | T > M) \leq 1 - \alpha^2$$

Comportamiento asintótico de cadenas de Markov irreducibles y aperiódicas

Y, por tanto,

$$P(X_{2M} \neq X'_{2M} | T > M) \leq 1 - \alpha^2$$

Se sigue que

$$\begin{aligned} P(T > 2M) &= P(T > 2M | T > M) P(T > M) \\ &\leq P(T > 2M | T > M)(1 - \alpha^2) \\ &\leq P(X_{2M} \neq X'_{2M} | T > M)(1 - \alpha^2) \\ &\leq (1 - \alpha^2)^2 \end{aligned}$$

Repitiendo el argumento, concluimos que

$$P(T > rM) \leq (1 - \alpha^2)^r \to 0 \text{ cuando } r \to \infty$$

$$P(T < \infty) = 1 - P(T = \infty) = 1 - 0 = 1.$$

Comportamiento asintótico de cadenas de Markov irreducibles y aperiódicas

Consideremos ahora la Cadena de Markov: (X''_n) definida por:

$$X''_0 = X_0$$

$$X''_{n+1} = \begin{cases} \psi(X''_n, U_{n+1}) & \text{si} \quad X''_n \neq X'_n \\ \psi(X''_n, U'_{n+1}) & \text{si} \quad X''_n = X'_n \end{cases}$$

- (X''_n) es una Cadena de Markov con matriz de transición P y distribución inicial de probabilidad $\mu^{(0)}$. Por tanto,

$$X''_n \sim \mu^{(n)}.$$

Comportamiento asintótico de cadenas de Markov irreducibles y aperiódicas

Sea $i \in \{1, 2, \cdots, k\}$. Entonces

$$\begin{aligned}
\mu_i^{(n)} - \eta_i &= P(X_n'' = s_i) - P(X_n' = s_i) \\
&\leq P(X_n'' = s_i, X_n' \neq s_i) \\
&\leq P(X_n'' \neq X_n') \\
&= P(T > n) \to 0 \text{ cuando } n \to \infty.
\end{aligned}$$

Analogamente (cambiando los roles de X_n'' y X_n'),

$$\eta_i - \mu_i^{(n)} \leq P(T > n) \to 0 \text{ cuando } n \to \infty.$$

Se sigue que

$$\lim_{n\to\infty} d_{TV}(\mu^{(n)}, \eta) = 0$$

Ley de los grandes números

Una CM se dice ergódica si es irreducible, aperiódica y los tiempos de espera para volver a cada uno de sus estados son todos finitos. En particular, si la CM es finita, será ergódica sii es irreducibe y aperiódica.

Teorema (Ley de los grandes números para CM)

Sea (X_n) una CM ergódica y sea μ su (única) distribución estacionaria. Sea X una v.a. tal que $X \sim \mu$. Sea $f : \mathbb{R} \to \mathbb{R}$ una función acotada. Entonces, con probabilidad 1, se tiene que

$$lim_{n\to\infty}\frac{1}{n}(f(X_1)+\cdots f(X_n)) = E(f(X)).$$

Nota: Como solo consideramos aquí CM finitas, arriba podemos asumir que X toma valores en $\{1,\cdots,k\}$ y por tanto

$$E(f(X)) = \sum_{j=1}^{k} f(j)\mu_j$$

donde $\mu = (\mu_1,\cdots,\mu_k)$.

Distribuciones reversibles, Cadenas de Markov reversibles

Definición

Sea (X_n) *una CM finita con espacio de estados*

$$S = \{s_1, \cdots, s_k\}$$

y matriz de transición P*.*
Sea $\mu = (\mu_1, \cdots, \mu_k)$ *una distribución de probabilidad:*

$$0 \leq \mu_i \leq 1 \text{ para todo } i \, , \, \sum_{i=1}^{k} \mu_i = 1$$

Decimos que μ *es una distribución reversible para la cadena de Markov* (X_n) *si*

$$\mu_i P_{i,j} = \mu_j P_{ji} \text{ para todo } i, j$$

Distribuciones reversibles, Cadenas de Markov reversibles

Definición

La CM se dice reversible si admite alguna distribución de probabilidad reversible.

Teorema (Cadenas de Markov reversibles)

Supongamos que μ es una distribución reversible para la CM (X_n). Entonces μ es también una distribución de equilibrio para la CM.

Si (X_n) es una CM irreducible y aperiódica con distribución reversible μ, entonces μ es la única distribución estacionaria de (X_n) y la distribución de probabilidad de X_n converge a μ en la distancia d_{TV}:

$$\lim_{n\to\infty} d_{TV}(\mu_{X_n}, \mu) = 0.$$

Distribuciones reversibles, Cadenas de Markov reversibles

Demostración del Teorema: Basta comprobar que $\mu P = \mu$. De hecho,

$$(\mu P)_i = \sum_{j=1}^{k} \mu_j P_{ji} = \sum_{j=1}^{k} \mu_i P_{ij} = \mu_i \sum_{j=1}^{k} P_{ij} = \mu_i$$

$\square$

Simulación de v.a. con fdp dada, mediante el uso de CM (Método MCMC)

- Queremos simular una v.a. X con espacio de estados $S = \{s_i\}_{i=1}^k$ y fdp η.

Definición (Algoritmo de Metrópolis del Método Monte Carlo)

- *Construimos un grafo conexo $G = (S, E)$.*
  Dado $s_i \in S$, $d_i = \deg(s_i)$ es el número de nodos adyacentes a s_i.
  Es importante lograr imponer que d_i sea pequeño para lograr una simulación práctica de la cadena de Markov.
- *La CM de Metrópolis asociada a G (y a η) es la que tiene matriz de transición:*

$$P_{i,j} = \begin{cases} \frac{1}{d_i} min\{\frac{\eta_j d_i}{\eta_i d_j}, 1\} & si\ s_i \sim s_j \\ 0 & si\ s_i, s_j\ no\ son\ vecinos \\ 1 - \sum_{t\ ,\ s_t \sim s_i} \frac{1}{d_i}\ min\left\{\frac{\eta_t d_i}{\eta_i d_t}, 1\right\} & si\ i = j. \end{cases}$$

Simulación de v.a. con fdp dada, mediante el uso de CM (Método MCMC)

Teorema

La CM de Metrópolis (X_n) *asociada a* G *(y a* η*) es una CM aperiódica, irreducible y reversible cuya distribución estacionaria es* η*. Por tanto, la simulación de la misma sirve para obtener muestras de la fdp* η*, pues la distribución de* X_n *converge en* d_{TV} *a* η*.*

Simulación de v.a. con fdp dada, mediante el uso de CM (Método MCMC)

En efecto: la conexidad de G garantiza que la cadena es irreducible y aperiodica. Veamos que es reversible y que η es una distribución reversible para la CM. En efecto: Si $i = j$ no hay nada que estudiar. Si $i \neq j$ pero los vértices s_i, s_j no son adyacentes, tampoco es necesario estudiar nada. Veamos qué sucede si $i \neq j$ y $s_i \sim s_j$. Entonces

$$\eta_i P_{i,j} = \frac{\eta_i}{d_i} \min\{\frac{\eta_j d_i}{\eta_i d_j}, 1\} = \begin{cases} \frac{\eta_j}{d_j} & \text{si } \frac{\eta_j}{d_j} \leq \frac{\eta_i}{d_i} \\ \frac{\eta_i}{d_i} & \text{si } \frac{\eta_j}{d_j} > \frac{\eta_i}{d_i} \end{cases} = min\{\frac{\eta_i}{d_i}, \frac{\eta_j}{d_j}\}$$

es una función simétrica en i, j, por lo que

$$\eta_i P_{i,j} = \eta_j P_{j,i},$$

que es lo que buscábamos.

Simulación de v.a. con fdp dada, mediante el uso de CM (Método MCMC)

Teorema

Consideremos una CM con espacio de estados S tal que, si $X_n = s_i$, X_{n+1} se obtiene del siguiente modo:

- *Se elige un nuevo estado s_j al azar, con distribución de probabilidad uniforme entre los vecinos de s_i (i.e. con probabilidad $\frac{1}{d_i}$ para todo s_j vecino de s_i).*
- *Se toma*

$$X_{n+1} = \begin{cases} s_j & \textit{con probabilidad } min\{\frac{\eta_j d_i}{\eta_i d_j}, 1\} \\ s_i & \textit{con probabilidad } 1 - min\{\frac{\eta_j d_i}{\eta_i d_j}, 1\} \end{cases}$$

Entonces (X_n) tiene matriz asociada la de la CM de Metrópolis. Además, η es la única distribución estacionaria para (X_n) y la distribución de X_n converge en d_{TV} a η.

Simulación de v.a. con fdp dada, mediante el uso de CM (Método MCMC)

En efecto. Supongamos que $i \neq j$. Entonces:

- Si s_j no es vecino de s_i, $P_{i,j} = P(X_{n+1} = s_j | X_n = s_i) = 0$ porque en el primer paso solo tomamos vecinos de s_i.
- Ahora bien, si $s_i \sim s_j$, entonces en el primer paso elegimos s_j con probabilidad $\frac{1}{d_i}$ y, a continuación, nos quedamos con $X_{n+1} = s_j$ con probabilidad $\min\{\frac{\eta_j d_i}{\eta_i d_j}, 1\}$. Por tanto, en este caso,

$$P_{i,j} = \frac{1}{d_i} \min\{\frac{\eta_j d_i}{\eta_i d_j}, 1\}.$$

Simulación de v.a. con fdp dada, mediante el uso de CM (Método MCMC)

- Finalmente, una vez elegido $s_j \sim s_i$ (con probabilidad $1/d_i$), nos quedamos en s_i con probabilidad $1 - \min\{\frac{\eta_j d_i}{\eta_i d_j}, 1\}$, lo que arroja un valor $\frac{1}{d_i} - \frac{1}{d_i}\min\{\frac{\eta_j d_i}{\eta_i d_j}, 1\}$, pero esto puede pasar con cualquier $s_j \sim s_i$, lo que nos conduce a que

$$
\begin{aligned}
P_{i,i} &= \sum_{j\,,\,s_j \sim s_i} \left(\frac{1}{d_i} - \frac{1}{d_i}\min\{\frac{\eta_j d_i}{\eta_i d_j}, 1\} \right) \\
&= 1 - \sum_{j\,,\,s_j \sim s_i} \frac{1}{d_i}\min\{\frac{\eta_j d_i}{\eta_i d_j}, 1\}
\end{aligned}
$$