Cadenas de Markov
Tiempo Discreto

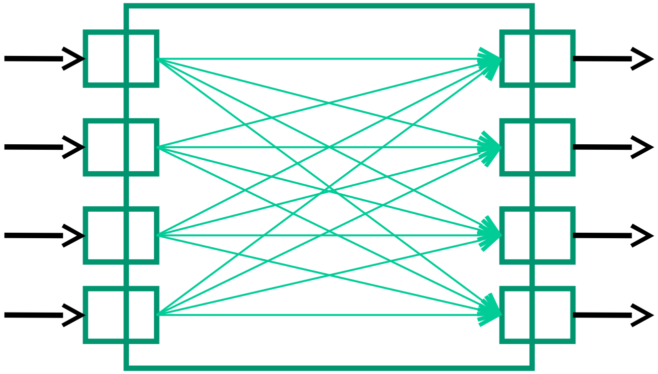
Modelado y Análisis de Redes de Telecomunicaciones

FACULTAD DE INGENIERIA

UNIVERSIDAD DE LA REPUBLICA URUGUAY

Motivación – Ejemplo 1

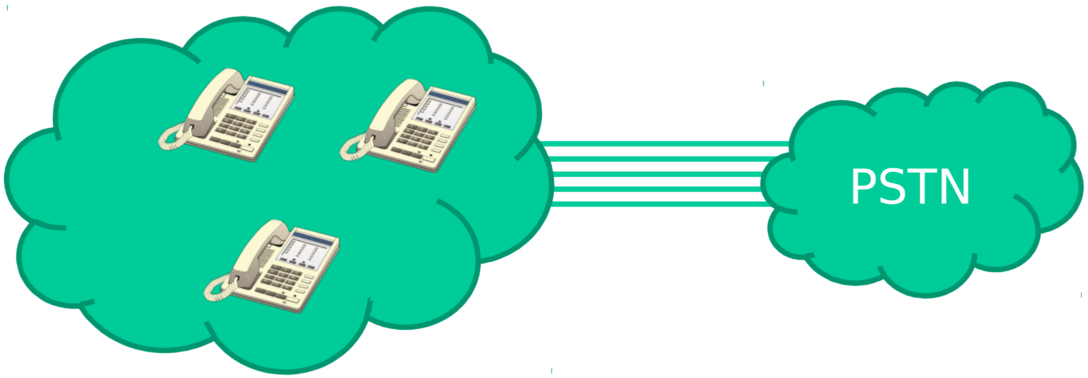
- Sea un enrutador al que arriban paquetes de otros (varios) routers



  - Cuando más de un paquete llega a la misma vez a alguno de los puertos de salida, el último en llegar se almacena en un buffer
  - Pregunta: ¿cómo dimensionar el buffer?

Motivación – Ejemplo 2

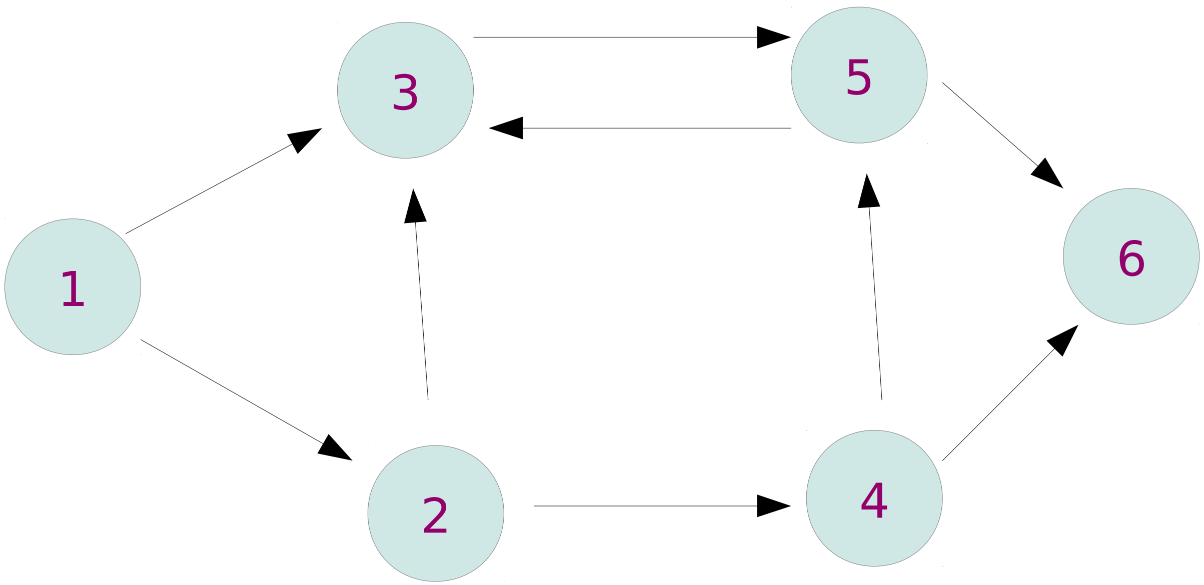
- Sean un conjunto de usuarios de teléfono



  - Cuando todas las líneas están siendo utilizadas, las nuevas llamadas son rechazadas
  - Problema: ¿cuántas líneas contratar?

Motivación – Ejemplo 3

- Sean un conjunto de páginas en una web cualquiera (la internet o un conjunto de ella)



¿Cuál es la importancia de cada página?

- ¿cantidad de páginas que apuntan a ella?
- ¿es lo mismo que el link venga de cualquier página?

Motivación

- En ambos ejemplos, hacer un análisis determinístico no es una opción
  - Al ser un sistema asíncrono, los paquetes pueden llegar en cualquier momento y tienen un tamaño cualquiera (al menos dentro de cierto rango)
  - Las llamadas telefónicas pueden llegar en cualquier momento y su duración es totalmente arbitraria

- Sin embargo…
  - Los paquetes de 1500B son más frecuentes (o con **probabilidad** mayor) que los de 10B
  - Las llamadas largas son menos frecuentes (o con **probabilidad** menor) que las llamadas cortas
- Por lo tanto, un enfoque probabilístico del problema nos dará mejores resultados

Proceso Estocástico

- Hay veces que lo importante no es un valor único, sino una sucesión temporal (aleatoria)
  - Cuántos paquetes hay en el buffer en cada instante *t*
  - Cuántas llamadas están siendo cursadas en cada instante *t*
- Un proceso estocástico le asigna a cada elemento del espacio muestral S una función en *t* (sucesión de variables aleatorias)
  - A cada elemento ω de S se le asocia una función $X_t(\omega)$
    - Para un valor de ω, $X_t(\omega)$ es una función temporal
    - Para un valor de t, $X_t(\omega)$ es una variable aleatoria
    - Para un valor de t y ω, $X_t(\omega)$ es un valor fijo
  - Generalmente la función $X_t(\omega)$ para un ω dado, es llamado realización del proceso o trayectoria

Proceso Estocástico

- Espacio de estados
  - El conjunto de posibles valores que puede tomar $X_t(\omega)$ (continuo o discreto)

- Espacio de parámetros
  - El conjunto de posibles valores que puede tomar t (continuo o discreto)

- Algunos parámetros de interés:
  - Distribución temporal: la distribución de la v.a. X_t para un t dado
  - Distribución estacionaria: la distribución de la v.a. X_t cuando $t \rightarrow \infty$ (si existe)

Proceso Estocástico

- Algunos parámetros de interés (cont...)
  - Estadísticas de orden n (o distribuciones finito-dimensionales)

$$F_{X_{t_1},\ldots,X_{t_n}}(x_1,\ldots,x_n) = P(X_{t_1} \leq x_1,\ldots,X_{t_n} \leq x_n)$$

para todos los posibles conjuntos $\{t_1,\ldots,t_n\}$

  - Ejemplos:
    - Estadística de orden 1:
      - la distribución estacionaria
      - la esperanza para un *t* dado
    - Estadística de orden 2: la covarianza

$$R_{t,s} = E\{(X_t - \bar{X}_t)(X_s - \bar{X}_s)\}$$

Procesos Estacionarios y Ergódicos

- **Proceso estacionario**
  - El proceso es estacionario cuando las estadísticas de todos los órdenes permanecen incambiadas por una traslación temporal

$$F_{X_{t_1+\tau},\ldots,X_{t_n+\tau}}(x_1,\ldots,x_n) = F_{X_{t_1},\ldots,X_{t_n}}(x_1,\ldots,x_n) \quad \forall n,\ \forall t_1,\ldots,t_n$$

  - Más sencillo de verificar es un proceso estacionario en el sentido amplio (la media y la covarianza son invariantes)

$$\bar{X}_t = \text{constante}, \quad R_{t+\tau,s+\tau} = R_{t,s}\ \forall\tau$$

- **Proceso ergódico**
  - La estadística de $X_t(\omega)$ para un ω dado es la misma que la de X_t para un t dado
  - Consecuencia práctica: con una única realización (infinita) del proceso puedo obtener toda la estadística de $X_t(\omega)$

Procesos de Markov

- Un proceso estocástico es de Markov si:

$$P(X_{t_{n+1}} \leq x_{n+1} | X_{t_n} = x_n, \ldots, X_{t_1} = x_1) = P(X_{t_{n+1}} \leq x_{n+1} | X_{t_n} = x_n)$$

$$\forall n,\ \forall t_1 < \ldots < t_n$$

  - La condición se puede interpretar de varias maneras:
    - El futuro de un proceso de Markov depende únicamente del estado actual (y no de cómo se llegó a él)
    - El estado actual de un proceso de Markov es toda la información que se necesita para caracterizar el futuro
    - Dado el estado actual, el futuro y el pasado son indeptes.

- ¿Se pueden analizar mediante procesos de Markov los ejemplos que dimos al comienzo?

Cadena de Markov -Tiempo Discreto

- Una cadena de Markov en Tiempo Discreto (CMTD) es un proceso de Markov con:
  - Espacio de estados *E* discreto (finito o no)
    - Un sub-conjunto de los naturales
  - Espacio de parámetros discreto también (*t* = 0,1,2...)
- Es decir $X = \{X_n : n = 0, 1, 2, \ldots\}$

$$\text{o}$$

$$X = X_n;\, n \in \mathcal{N} \text{ tal que } X_n(w) \in \{0, 1, 2, \ldots\}$$

- La condición resulta:

$$P(X_{n+1} = j | X_n = i_n, \ldots, X_0 = i_0) = P(X_{n+1} = j | X_n = i_n)$$

CMTD - Definición

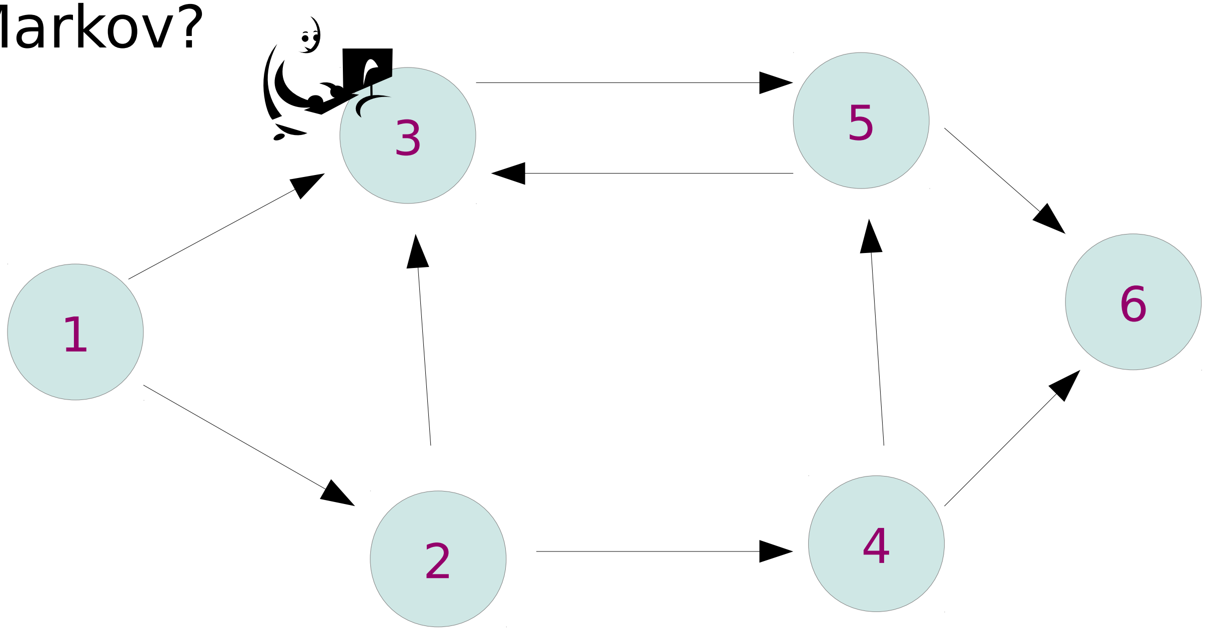
- El proceso está caracterizado por las probabilidades de transición del estado *i* al *j* en el instante *n*

$$P(X_{n+1} = j | X_n = i) = p_{ij}(n)$$

- Estas probabilidades describen la evolución del sistema
- De aquí en adelante asumiremos homogeneidad: $p_{ij}(n) = p_{ij}$
- Interesa estudiar $P(X_n = i)$
  - Probabilidad de que el sistema se encuentre en el estado i en tiempo n

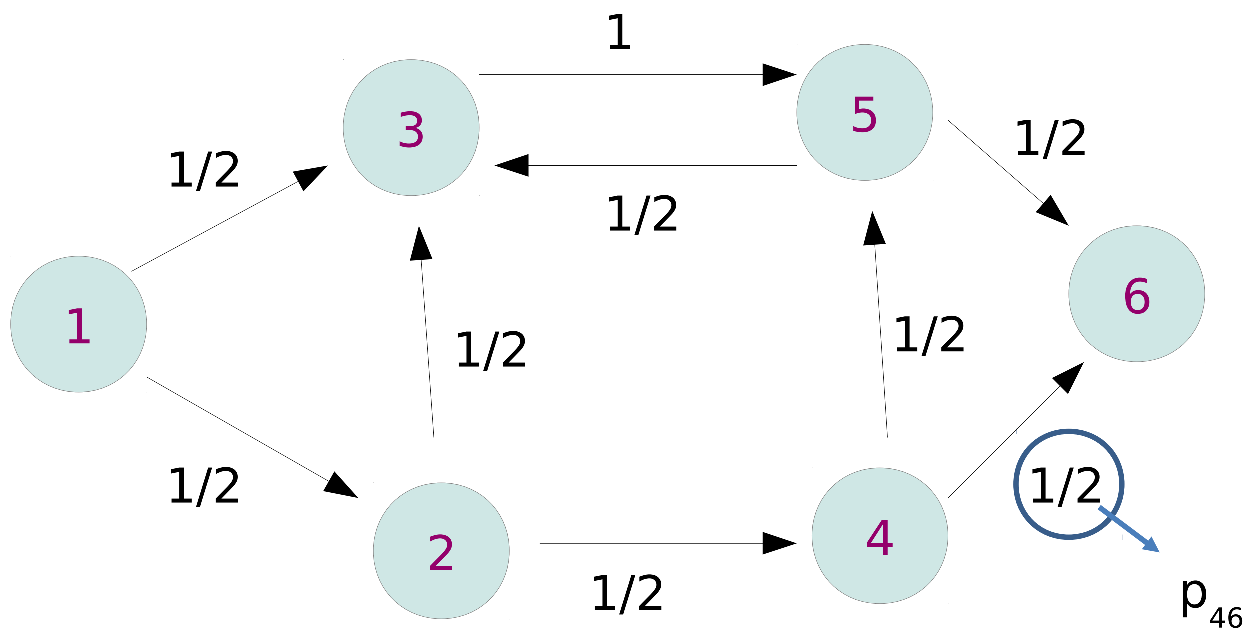
CMTD - Ejemplo

- Una persona navegando la web:
  - El usuario sólo hace click en los links que le aparecen en cada página (por ahora...)
  - En cada página, elige entre los links al azar sin recordar su elección anterior
  - ¿Xn={página elegida por la persona} es una cadena de Markov?



CMTD – Grafo Asociado

- La manera clásica de representar la cadena es mediante el denominado *grafo asociado*



CMTD – Matriz de Transiciones

- Otra manera de representar la cadena es mediante una matriz, llamada matriz de transiciones: $P = (p_{ij})$
- Es una matriz estocástica :

$$P_{ij} \geq 0$$

$$\sum_{j \in E} P_{ij} = 1 \quad \forall i \in E$$

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

Columna *j*: probabilidades de llegar a *j* por los otros estados

Fila *i*: probabilidades de pasar a los otros estados desde *i* (debe sumar 1)

CMTD – Distribución

- Se llama ley o distribución de X en el instante n al vector π_n (quizá infinito) $\pi_n=(\pi_n(j))_{j\in E}$

$$\pi_n(j) = P(X_n = j), \quad j = 0, 1, \ldots$$

- π_n es un vector de probabilidad: $\sum\pi_n(j)=1$ y $\pi_n(j)\geq 0$
- En particular se llama ley o distribución inicial a π_0
- Conociendo la ley inicial y las probabilidades de transición queda completamente determinada la probabilidad de una trayectoria
  - En el ejemplo del navegador, si arranca en 6 queda ahí por siempre

CMTD – Distribución

- ¿Cuánto vale π_{n+1}?

$$\pi_{n+1}(j) = P(X_{n+1} = j) = \sum_i P(X_{n+1} = j | X_n = i) P(X_n = i)$$

$$\Rightarrow \pi_{n+1}(j) = \sum_i \pi_n(i) p_{ij}$$

$$\therefore \pi_{n+1} = \pi_n \mathbf{P}$$

- Por recurrencia vale: $\boxed{\pi_n = \pi_0 \mathbf{P}^n}$
- Probabilidad de pasar de i a j en n pasos

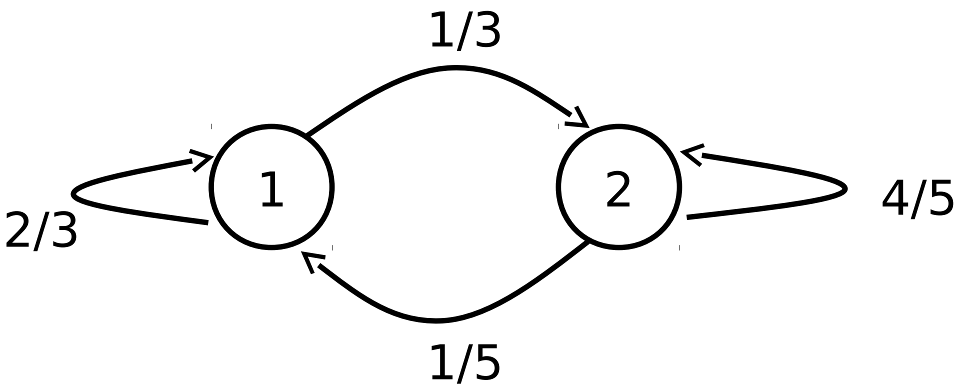
$$p_{ij}^{(n)} = P\left(X_{m+n} = j | X_m = i\right) \Rightarrow P^n = \left(p_{ij}^{(n)}\right)_{ij \in E}$$

- Ecuaciones de Chapman-Kolmogorov

$$p_{ij}^{(n+m)} = \sum_{k \in E} p_{ik}^{(m)} p_{kj}^{(n)}$$

CMTD -Distribución

- Tomemos la siguiente cadena sencilla:



$$\mathbf{P} = \begin{pmatrix} 0.666 & 0.333 \\ 0.2 & 0.8 \end{pmatrix}$$

$$\mathbf{P}^2 = \begin{pmatrix} 0.511 & 0.489 \\ 0.293 & 0.707 \end{pmatrix}$$

$$\mathbf{P}^3 = \begin{pmatrix} 0.439 & 0.561 \\ 0.337 & 0.663 \end{pmatrix}$$

$$\ldots$$

$$\mathbf{P}^{10} = \begin{pmatrix} 0.375 & 0.625 \\ 0.375 & 0.625 \end{pmatrix}$$

Después de un cierto tiempo, la distribución de la cadena converge a un único valor sin importar la condición inicial

- ¿Esto es siempre verdad? Si no, ¿cuándo?
  - Veremos que tiene relación con la ergodicidad

CMTD – Distribución Límite

- Se quiere estudiar el sistema (las leyes de probabilidad del sistema) al cabo de un tiempo largo
- En particular, saber si hay un "régimen asintótico", es decir si

$$\lim_{n\to\infty} \pi_n = \mu$$

- μ es ley límite de *X* si existe π_0 tal que

$$\mu = \lim_{n\to\infty} \pi_n = \lim_{n\to\infty} \pi_0 P^n$$

$$\mu(j) = \lim_{n\to\infty} \pi_n(j) = \lim_{n\to} \sum_{i\in E} \pi_0(i) p_{ij}^{(n)}$$

- En principio las leyes límites (si existen) dependen de π_0 y del comportamiento asintótico de P^n
- Para E finito:
  - Si P^n converge término a término a M, $\mu = \pi_0 M$ es ley límite
  - Si $\mu = \lim \pi_n$ entonces μ es vector de probabilidad: puede no ser cierto en el caso infinito

CMTD – Distribución Límite

- Ejemplo anterior (cadena sencilla): única distribución límite
- Ejemplo del navegador: si hubiera más de una página sin links (como 6) habría más de una distribución límite:
  - Dependerá de π_0 y/o del azar
- Otro ejemplo:

$$\mathbf{P} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

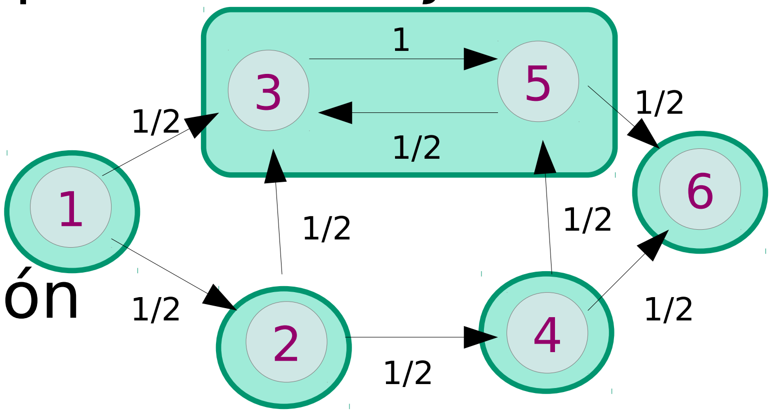
  - no existe lim P^n :

$$\mathbf{P}^{2n} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \mathbf{P}^{2n+1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

  - pero sí existe ley límite para $\pi_0 = (1/2\ 1/2)$
- Entonces ¿bajo qué condiciones se puede asegurar que exista y que sea única? Y en ese caso ¿cómo se calcula?

Clasificación de Estados

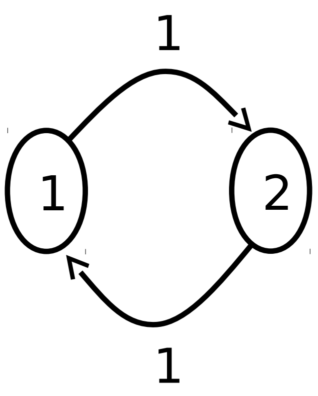
- Comunicación entre dos estados ($i \leftrightarrow j$)
  - Los estados i y j se comunican si existen m y n positivos tales que $p_{ij}^{(m)} > 0$ y $p_{ji}^{(n)} > 0$
  - Es decir, existe un recorrido que va de *i* a *j* en el grafo asociado y visceversa
- Ejemplo del navegador:
  - $3 \leftrightarrow 5$



- La comunicación es una relación de equivalencia
  - Reflexiva, simétrica y cumple la transitiva
  - Se puede hacer una partición en conjuntos con la relación (cada conjunto se denomina clase de equivalencia)
- Si existe una única clase (todos se comunican entre sí), la cadena se dice **irreducible**

Periodicidad

- Sea $d(i)=\text{mcd}\{n\colon p_{ii}^{(n)}>0\}$. El estado i es **periódico** de período $d(i)$ sii $d(i)>1$; en caso contrario se dice que es **aperiódico**
- Ejemplos:
  - Se vuelve al estado después de 2, 4, 6, ... pasos, entonces es periódico de período 2



  - En el navegador existen periodicidades
- Si $p_{ii}>0$, entonces i es aperiódico
- La periodicidad es propia de la clase

Recurrencia

- Sea f_{ii}^{n} (o $f_n(i)$) la probabilidad de, saliendo del estado *i*, volver a i por primera vez en *n* pasos
- La probabilidad f_{ii} de volver al estado i en tiempo finito es:

$$f(i) = \sum_{n \geq 1} f_n(i)$$

- El estado *j* es **recurrente** sii f_{ii} =1
- El estado *i* es **transitorio** sii f_{ii}<1
  - hay probabilidad no nula de nunca volver al estado *i*
- Los estados transitorios son aperiódicos

Recurrencia

- Si el espacio de estados es finito, siempre existe al menos un estado recurrente
- Si el espacio de estados es infinito, puede no haber estados recurrentes (e.g. $p_{i,i+1}=1$)

- En términos de las transiciones, un estado es recurrente sii $\sum p_{ii}^{(n)}=\infty$
- Si j es transitorio $\sum p_{ij}^{(n)}<\infty$ para todo i y entonces

$$\lim_{n\to\infty} p_{ij}^{(n)} = 0$$

Recurrencia Positiva o Nula

- Sea $M(i)$ el tiempo medio de retorno al estado:

$$M(i) = \sum_{n \geq 1} n f_n(i)$$

  - Si $M(i)=\infty$ el estado se denomina **recurrente nulo**
  - Si $M(i)<\infty$ el estado se denomina **recurrente positivo**
- La recurrencia es propia de la clase (usando la caracterización anterior)
  - Por ejemplo, en una cadena irreducible, si un estado es recurrente positivo, todos los estados lo serán
- Este valor está relacionado con la distribución límite

CMTD Irreducible y Aperiódica

Se da únicamente una de tres posibilidades:

1. Es transitoria:

$$\lim_{n\to\infty} p_{ij}^{(n)} = \lim_{n\to\infty} \pi_n(j) = 0$$

$$\text{Más aún } \sum_{n=1}^{\infty} p_{ij}^{(n)} < \infty \quad \forall i,j \in E$$

2. Es recurrente nula:

$$\lim_{n\to\infty} p_{ij}^{(n)} = \lim_{n\to\infty} \pi_n(j) = 0$$

$$\text{Más aún } \sum_{n=1}^{\infty} p_{ij}^{(n)} = \infty \quad \forall i,j \in E$$

3. Es recurrente positiva:

$$\lim_{n\to\infty} p_{ij}^{(n)} = \lim_{n\to\infty} \pi_n(j) = \frac{1}{M(j)}$$

- Obs: Si E es finito, siempre estoy en el caso 3.

Ergodicidad

- Una cadena es ergódica si existe un único vector de probabilidad π_∞ tal que cualquiera sea la ley inicial π_0 se cumple que:

$$\pi_\infty = \lim_{n\to\infty} \pi_n = \lim_{n\to\infty} \pi_0 \mathbf{P}^n$$

- Una cadena irreducible es ergódica sii todos sus estados son:
  - Recurrentes positivos
  - Aperiódicos
- Por ser irreducible basta verificar las condiciones en un solo estado

Ergodicidad

- En una cadena ergódica, la porción de tiempo que la cadena está en el estado *i* es $\pi_\infty(i)$
- Además, tiene la siguiente relación con el tiempo medio de retorno *M*(*i*)

$$\pi_\infty(i) = 1/M(i)$$

Intuición: Cada vez que salgo del estado *i* vuelvo en *M*(*i*) instantes => estoy 1 de cada *M*(*i*) instantes en *i*

Ergodicidad – Caso *E* Finito

- Si el espacio de estados *E* es finito entonces las siguientes son condiciones necesarias y suficientes
  - *X* es ergódica con distribución límite π_∞ sii

$$\lim_{n\to\infty} \mathbf{P}^n = \begin{bmatrix} \pi_\infty \\ \vdots \\ \pi_\infty \end{bmatrix}$$

  - *X* es ergódica sii $\lambda=1$ es valor propio dominante de P con multiplicidad 1 (y además el vector propio asociado es π_∞)
    - Se puede probar tomando en cuenta que por la ecuación anterior, π_∞ debe cumplir

$$\boxed{\pi_\infty = \lim_{n\to\infty} \mathbf{P}^{n+1} = \left(\lim_{n\to\infty} \mathbf{P}^n\right)\mathbf{P} = \pi_\infty \mathbf{P}}$$

  - Una manera de calcular π_∞ es resolviendo la ecuación anterior junto a la normalización $\sum \pi_\infty(i)=1$

Ergodicidad – Caso *E* infinito

- Si el espacio de estados *E* es infinito, pero la cadena es irreducible y aperiódica
  - Si existe una distribución estacionaria de la cadena; i.e.

$$\pi_\infty(j) = \sum_{i \in E} p_{ij} \pi_\infty(i) \quad (\pi_\infty = \pi_\infty \mathbf{P})$$

    entonces la cadena es recurrente positiva
  - Por lo tanto, la distribución es única y la cadena es ergódica
  - Consecuencia práctica: cómo probar ergodicidad

    1 – Verificar irreducibilidad y aperiodicidad

    2 – Tratar de encontrar una distribución estacionaria

Ecuaciones de Balance

- Si π_∞ es la distribución límite entonces es invariante:

$$\pi_\infty(j) = \sum_{i \in E} p_{ij} \pi_\infty(i) \quad (\pi_\infty = \pi_\infty \mathbf{P})$$

- Además:

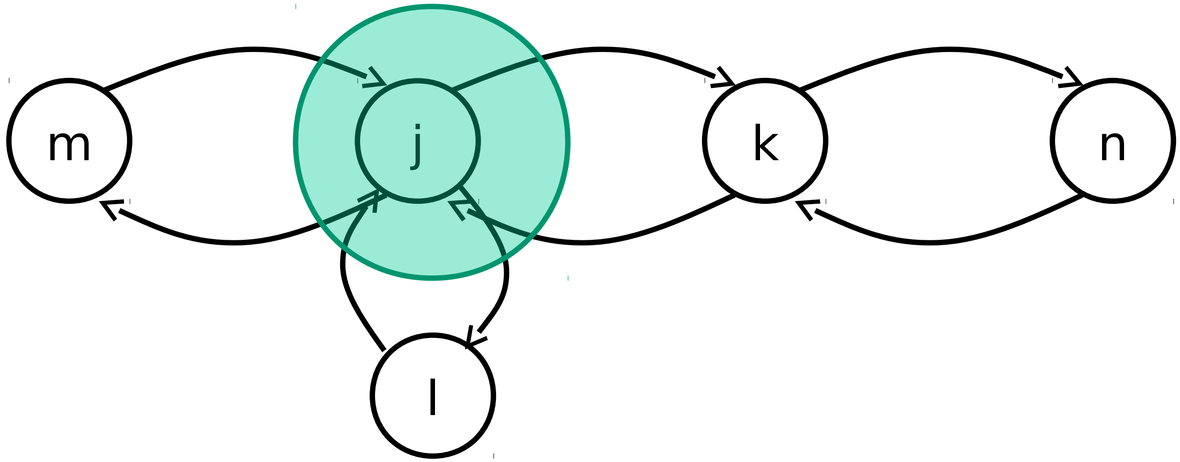
$$\sum_{i \in E} p_{ji} = 1$$

- Multiplicando las dos ecuaciones y eliminando el término repetido obtenemos:

$$\pi_\infty(j) \sum_{i \neq j} p_{ji} = \sum_{i \neq j} p_{ij} \pi_\infty(i)$$

Ecuaciones de Balance

$$\pi_\infty(j) \sum_{i \neq j} p_{ji} = \sum_{i \neq j} p_{ij} \pi_\infty(i)$$



- Todo lo que sale (pesado por la probabilidad de estar ahí) es igual a lo que entra (también pesado por la probabilidad de estar ahí)

Ecuaciones de Balance

- Las ecuaciones de balance se pueden extender a conjuntos *S* de estados
  - Se toma la ecuación

$$\pi_\infty(j)\sum_{i\in E} p_{ji} = \sum_{i\in E} p_{ij}\pi_\infty(i)$$

  - Se suma en los estados pertenecientes a *S*

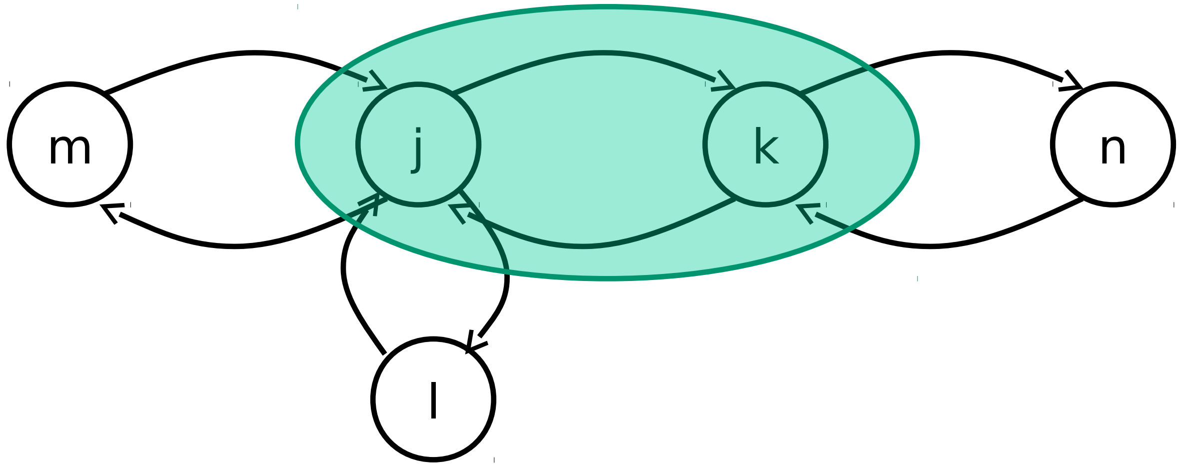
$$\sum_{j\in S}\pi_\infty(j)\sum_{i\in E} p_{ji} = \sum_{j\in S}\sum_{i\in E} p_{ij}\pi_\infty(i)$$

  - Y cancelando los términos repetidos resulta:

$$\sum_{j\in S}\pi_\infty(j)\sum_{i\notin S} p_{ji} = \sum_{i\notin S}\pi_\infty(i)\sum_{j\in S} p_{ij}$$

Ecuaciones de Balance

$$\sum_{j \in S} \pi_\infty(j) \sum_{i \notin S} p_{ji} = \sum_{i \notin S} \pi_\infty(i) \sum_{j \in S} p_{ij}$$



- Es una ecuación sumamente útil para encontrar la distribución estacionaria
  - Escribimos todas las $\pi_\infty(i)$ en función de un $\pi_\infty(j)$ en particular, y luego normalizamos

Bibliografía

- Esto no fue más que un vistazo a las cadenas de Markov en tiempo discreto. Por más info consultar la web del curso y/o el siguiente material:

  - R. Gallager. Notas de su curso en línea “Discrete Stochastic Processes” o su libro “Stochastic Processes: Theory for applications”
  - “Teoría de Probabilidad” de Valentín Petrov y Ernesto Mordecki, editorial DIRAC – Facultad de Ciencias, 2008.
  - K. L. Chung. Markov chains with stationary transition probabilities. Springer, 1960-67.
  - P. Brémaud. Markov chains. Springer, 1999.
  - S. Resnick. Adventures in stochastic processes. Birkhäuser, 1994.
  - W. Feller. Introducción a la teoría de probabilidades y sus aplicaciones. Vol 1. Limusa, 1986.