Cadenas de Markov
Clase 04

Investigación Operativa UTN FRBA 2020
Curso: I4051
Docente: Martín Palazzo
Equipo: Rodrigo Maranzana, Milagros Bochor, Gabriel Boso, Juan Piro

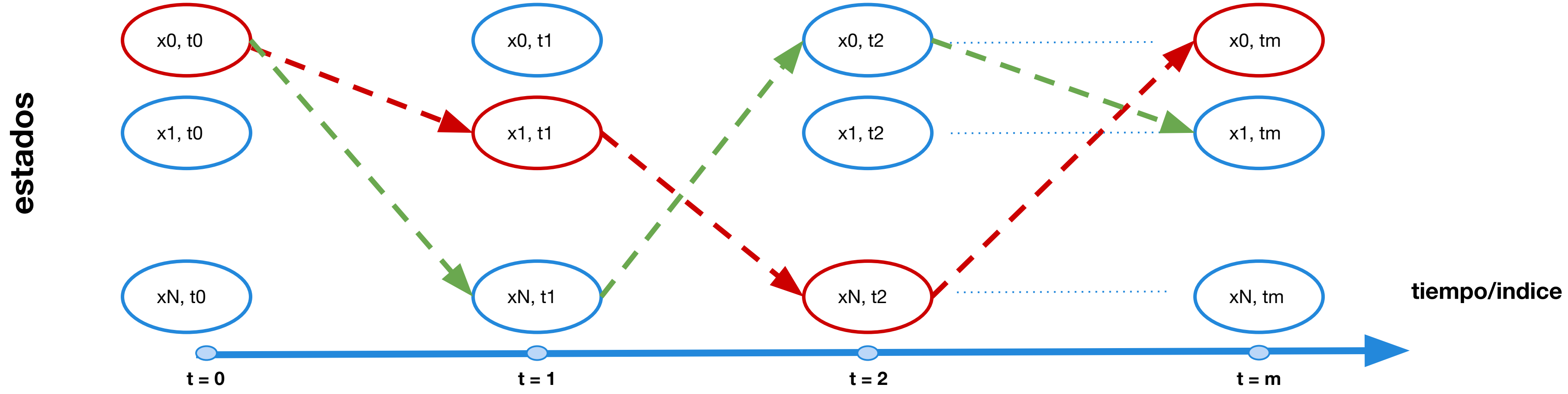
Agenda clase 04

- Repaso Cadenas de Markov
- Tipos de Cadenas de Markov
- Estado Estacionario
- Ejercicios Cadenas de Markov
- Simulación en Python de Cadenas de Markov

Procesos estocásticos

Se llama así a los **sistemas dinámicos** sometidos a un fenómeno de naturaleza **aleatoria.** El sistema adquiere distintos **estados** caracterizados por **X** que variara sujeto a cierta distribución de **probabilidad** a medida que avanza el **parámetro** (generalmente tiempo).

Realizar múltiples veces el seguimiento en un proceso estocástico de manera idéntica generará resultados distintos aunque siempre sujetos a determinadas distribuciones de probabilidad supuestas constantes.



Procesos estocásticos

Nuestro sistema estará compuesto por **variables aleatorias x = f(t)** que tomarán distintos valores a medida que evoluciona el parámetro.

Las variables aleatorias serán las **responsables** de que el **sistema tome distintos estados** bajo cierta distribución de probabilidad a medida que avanza el parámetro (tiempo).

En otras palabras, para cada estado posible del sistema, dado un instante de tiempo “t” existirá una probabilidad asociada **P(x | t)**.

Proba estados

P(x = 0 | t = 0)		**P(x = 0 | t = 1)**		**P(x = 0 | t = m)**
P(x = 1 | t = 0)		**P(x = 1 | t = 1)**		**P(x = 1 | t = m)**
P(x = 2 | t = 0)		**P(x = 2 | t = 1)**		**P(x = 2 | t = m)**
P(x = 3 | t = 0)	→	**P(x = 3 | t = 1)**	→	**P(x = 3 | t = m)**
.		.		.
.		.		.
P(x = N | t = 0)		**P(x = N | t = 1)**		**P(x = N | t = m)**

Tipos de Procesos estocásticos

- Proceso Estocástico
  - Procesos aleatorios puros
  - Procesos Markovianos
    - Procesos de Markov
    - Cadenas de Markov
      - En parámetro discreto
        - Homogénea
        - No Homogénea
      - En parámetro continuo
        - Homogénea
        - No Homogénea
  - Procesos con memoria

"Cadena de Markov" -> variable aleatoria discreta (parametro continuo o discreto)
"Proceso de Markov" -> variable aleatoria continua (parametro continuo o discreto)

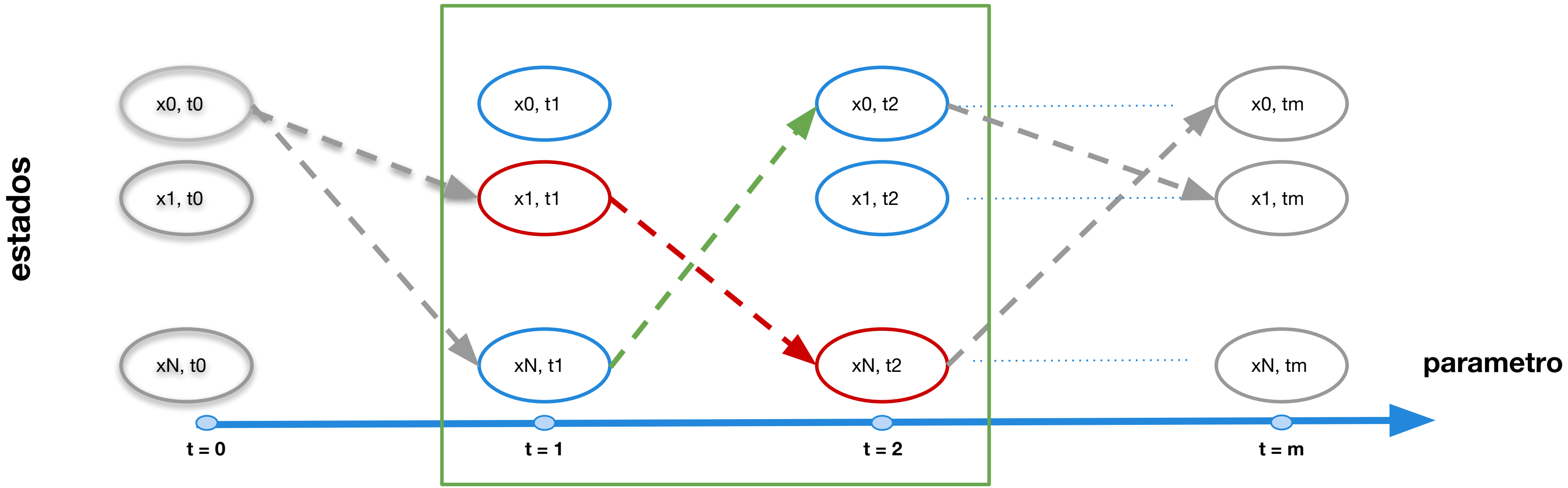
Cadenas de Markov

Las cadenas de Markov corresponden a una rama de los Procesos Estocásticos. Las mismas se caracterizan por definir sus estados actuales únicamente partiendo del estado inmediato anterior. **Es decir que las cadenas de Markov son procesos estocásticos con memoria t=-1**. Visto de otra manera, si conozco el estado X actual del sistema en el instante t, en caso de ser "markoviano" puedo estimar la probabilidad **condicional** de estado en el siguiente intervalo de tiempo t+1.

$$P\left(X_{t=i+1} \mid X_{t=i}\right)$$

Cadenas de Markov

El concepto Markoviano nace de comprender la transición de **un solo paso.** La probabilidad de cada estado depende únicamente del estado del sistema en el "paso" inmediato anterior.



Probabilidad condicional de Transición de 1 paso

Dados N estados, la probabilidad de transición de un paso es aquella que indica la probabilidad de estar en en el estado “j” si en el paso anterior el sistema se encontraba en el estado “i”.

$$p_{ij}(t+1) = P(X_{t+1} = j | X_t = i)$$

Si conozco las probabilidades de transición de un paso de un sistema (entre todos los pares de estados) y supongo que las mismas se mantienen constantes en el tiempo, entonces puedo calcular las probabilidades de transición luego de M pasos.

$$p_{ij}(t+\Delta t) = P(X_{t+\Delta t} = j | X_t = i)$$

Matriz de Transición de 1 paso

Para poder modelar todas las probabilidades de transición de 1 paso en un sistema markoviano utilizamos la Matriz de Transición de 1 paso. Por ley de probabilidad, la suma de todas las probas que salgan de un estado deben sumar 1.

$$P\left(\Delta t = 1\right) = \begin{bmatrix} P_{11} & P_{12} & P_{13} & \dots & P_{1n} \\ P_{21} & P_{22} & P_{32} & \dots & P_{2n} \\ P_{31} & P_{32} & P_{33} & \dots & P_{3n} \\ \dots & \dots & \dots & P_{ij} & \dots \\ P_{n1} & \dots & \dots & \dots & P_{nn} \end{bmatrix}$$

$$0 \leq P_{ij} \leq 1$$

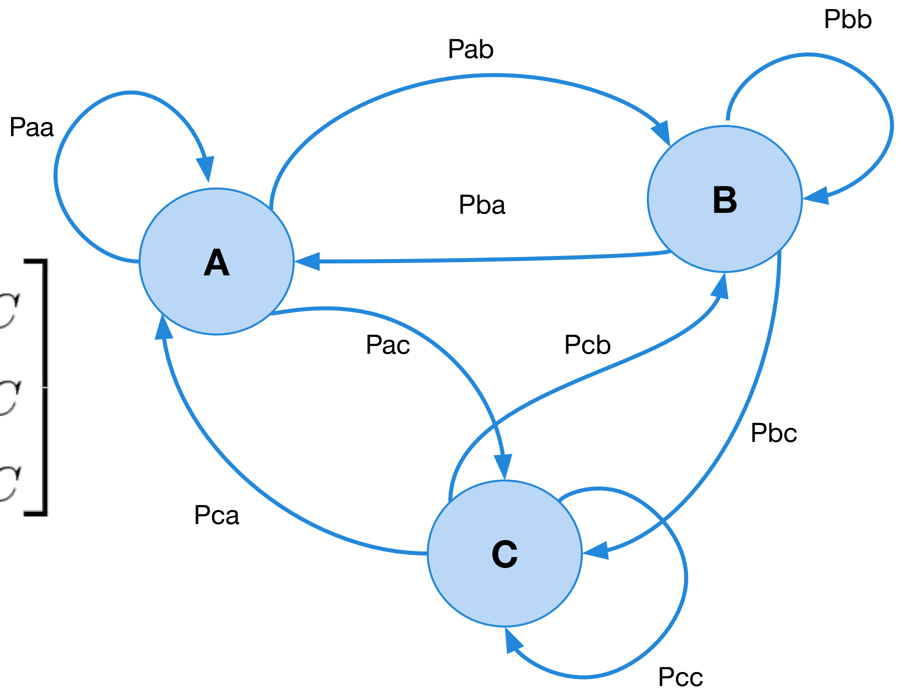
$$\sum_{j=1}^{n} P_{ij} = 1$$

Los **renglones/filas** representan los nodos/estados de origen mientras que las **columnas** los nodos/estados destino.

Grafo de Transición de 1 paso

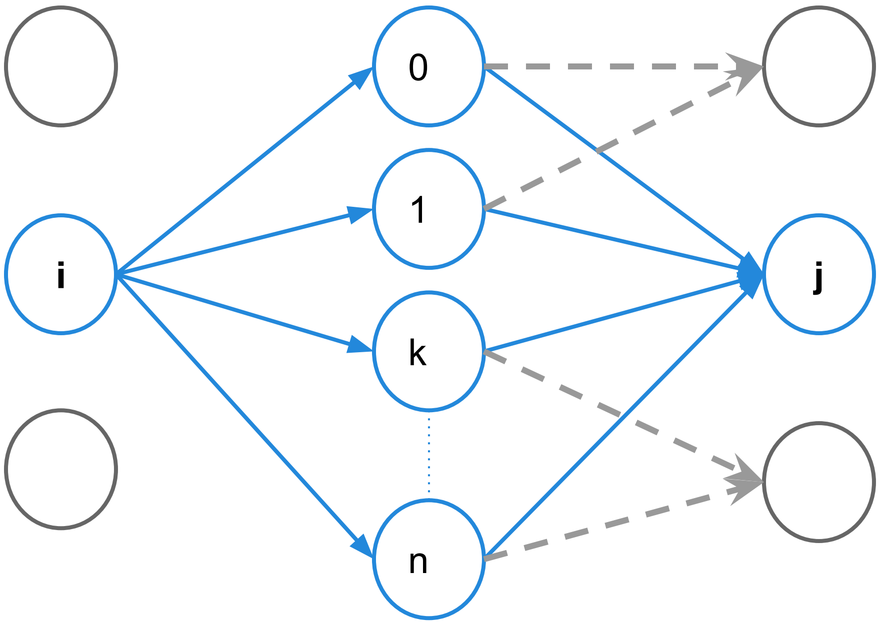
Una matriz tiene un grafo equivalente

$$P\left(\Delta t = 1\right) = \begin{bmatrix} P_{AA} & P_{AB} & P_{AC} \\ P_{BA} & P_{BB} & P_{BC} \\ P_{CA} & P_{CB} & P_{CC} \end{bmatrix}$$



Matriz de Transición de 2 pasos (Δt = 2)

De igual manera podemos estimar la probabilidad de transición de “2” utilizando como unidad básica la probabilidad de transición de 1 paso. Para el caso de 2 pasos la formulación es:



$$P_{ij}^{(2)} = P(X_{t+2} = j | X_t = i)$$

$$P_{ij}^{(2)} = \sum_{k=0}^{k=n} (P_{ik} \times P_{kj})$$

Ecuación de Chapman-Kolmogorov (Δt = 2)

La manera de expresar las probabilidades de transición de una cadena de Markov en “n” pasos es a través de la ecuación de Chapman-Kolmogorov. Estas indican que la probabilidad de pasar del estado “i” al “j” en n pasos (en este caso n=2) es primero sumando todas las probabilidades de pasar del “i” a cualquiera de los estados intermedios “k” en 1 paso y luego sumar todas las probabilidades de pasar del estado “k” al destino “j” en el siguiente paso.

$$P_{ij}^{(2)} = \sum_{k=0}^{k=n} (P_{ik} \times P_{kj})$$

Ecuación de Chapman-Kolmogorov (matricial, Δt = 2)

El producto matricial de las matrices de transición de 1 paso da como resultado la matriz de transición de 2 pasos.

$$P^{(2)} = \begin{bmatrix} P_{11}^{(2)} & P_{12}^{(2)} & \dots & P_{1n}^{(2)} \\ P_{21}^{(2)} & P_{22}^{(2)} & \dots & P_{2n}^{(2)} \\ P_{31}^{(2)} & \dots & \dots & \dots \\ \dots & \dots & P_{ij}^{(2)} & \dots \\ P_{n1}^{(2)} & \dots & \dots & P_{nn}^{(2)} \end{bmatrix} = \begin{bmatrix} P_{11}^{(1)} & P_{12}^{(1)} & \dots & P_{1n}^{(1)} \\ P_{21}^{(1)} & P_{22}^{(1)} & \dots & P_{2n}^{(1)} \\ P_{31}^{(1)} & \dots & \dots & \dots \\ \dots & \dots & P_{ij}^{(1)} & \dots \\ P_{n1}^{(1)} & \dots & \dots & P_{nn}^{(1)} \end{bmatrix} \times \begin{bmatrix} P_{11}^{(1)} & P_{12}^{(1)} & \dots & P_{1n}^{(1)} \\ P_{21}^{(1)} & P_{22}^{(1)} & \dots & P_{2n}^{(1)} \\ P_{31}^{(1)} & \dots & \dots & \dots \\ \dots & \dots & P_{ij}^{(1)} & \dots \\ P_{n1}^{(1)} & \dots & \dots & P_{nn}^{(1)} \end{bmatrix}$$

$$P^{(2)} = P^{(1)} \times P^{(1)}$$

Ecuación de Chapman-Kolmogorov (m pasos)

Expresión genérica de las ecuaciones de Chapman-kolmogorov (izquierda) y la misma expresión a nivel matricial (derecha).

$$p_{ij}^{(m)} = \sum_{k=0}^{n} p_{ik} . p_{kj}^{(m-1)}$$

$$\begin{aligned}
P^{(2)} &= P^{(1)} . P^{(1)} \\
P^{(3)} &= P^{(1)} . P^{(2)} \\
P^{(4)} &= P^{(1)} . P^{(3)} \\
P^{(5)} &= P^{(1)} . P^{(4)} \\
\cdots & \quad \cdots \\
P^{(m)} &= P^{(1)} . P^{(m-1)}
\end{aligned}$$

Probabilidad incondicional de estado

La probabilidad incondicional de estado es aquella que indica la chance que hay de que el sistema se encuentre en el estado “i” dado el instante “Ti”.

$$p_i\left(t\right) = p_{x=i}\left(t\right)$$

El conjunto de probabilidades de estado individuales de cada uno de los estados del sistema definen el **vector de probabilidades de estado p(t)**

$$p\left(t\right) = \left[p_1(t), p_2(t), \ldots, p_i(t), \ldots, p_n(t)\right]$$

¿Qué condiciones debe cumplir este vector?

Probabilidad incondicional de estado

Cada “paso” o instante de tiempo tendrá un vector de probabilidad de estado asociado. Este vector puede ser calculado partiendo de las ecuaciones de Chapman-Kolmogorov.

$$p(t_0) = [p_1(t_0), p_2(t_0), \ldots, p_i(t_0), \ldots, p_n(t_0)]$$
$$p(t_1) = [p_1(t_1), p_2(t_1), \ldots, p_i(t_1), \ldots, p_n(t_1)]$$
$$p(t_2) = [p_1(t_2), p_2(t_2), \ldots, p_i(t_2), \ldots, p_n(t_2)]$$
$$\ldots$$
$$p(t_m) = [p_1(t_m), p_2(t_m), \ldots, p_i(t_m), \ldots, p_n(t_m)]$$

Probabilidad incondicional de estado

Para calcular por ejemplo el vector de probabilidad de estado en el “t=1” teniendo el vector de probabilidad de estado en “t=0” (a.k.a. vector de estado inicial) y la matriz de transición de 1 paso

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

$$[p_0(1), p_1(1), p_2(1), ..., p_n(1)] = [p_0(0), p_1(0), p_2(0), ..., p_n(0)] \cdot \begin{bmatrix} p_{00} & p_{01} & p_{0n} \\ p_{10} & p_{11} & p_{1n} \\ p_{n0} & p_{n10} & p_{nn} \end{bmatrix}$$

Probabilidad incondicional de estado

De forma genérica, las probabilidades de estado para cualquier tiempo "t" pueden definirse tanto de manera genérica (izquierda) como matricial (derecha)

$$p_j^{(m)} = \begin{cases} \sum_{i=1}^{n} p_i^{(0)} . p_{ij}^{(m)} \\ \sum_{i=1}^{n} p_i^{(m-1)} . p_{ij}^{(1)} \end{cases} \qquad p^{(m)} = \begin{cases} p^{(0)} . P^{(m)} \\ p^{(m-1)} . P^{(1)} \end{cases}$$

Las expresiones del renglón superior equivalen a las del inferior, llegando al mismo resultado con transiciones de "m" pasos o con transiciones de "1" paso siendo el vector de estado en el instante "m" o "0".

Cadenas de Markov en régimen permanente

Si deseamos entender como se comportaría una cadena de Markov en el "**estado estacionario**" o "**régimen permanente**" entonces debemos observar el comportamiento de la cadena luego de "m" pasos de tiempo con "m" tendiendo a infinito.

$$\lim_{m \to \infty} P^{(m)} = \begin{bmatrix} p_{11}^{(m)} = \boxed{p(1)^{(m)}} & p_{12}^{(m)} = \boxed{p(2)^{(m)}} & \cdots & p_{1n}^{(m)} = \boxed{p(n)^{(m)}} \\ p_{21}^{(m)} = p(1)^{(m)} & p_{22}^{(m)} = p(2)^{(m)} & \cdots & p_{2n}^{(m)} = p(n)^{(m)} \\ \cdots & \cdots & \cdots & \cdots \\ p_{n1}^{(m)} = p(1)^{(m)} & p_{n2}^{(m)} = p(2)^{(m)} & \cdots & p_{nn}^{(m)} = p(n)^{(m)} \end{bmatrix}$$

Lo que sucederá es que las probabilidades de transición de "m" pasos con "m" tendiendo a infinito serán igual no importa desde que estado se inicie la transición.

Cadenas de Markov en régimen permanente

La probabilidad de estado en “t=m” cuando t tiende a infinito será igual que la probabilidad de estado en “t = m-1”. De esta manera multiplicando el vector de probabilidad de estado por la matriz de transición de un paso dará por resultado el vector de probabilidad de estado en régimen permanente.

$$\lim_{m \to \infty} p^{(m)} = \lim_{m \to \infty} p^{(m-1)} . P^{(1)}$$

$$\boxed{\begin{gathered} \rightarrow p^{(m)} = p^{(m)} . P^{(1)} \\ \sum_{i=1}^{n} p_i = 1 \end{gathered}}$$

Tipos de Cadenas de Markov

- Estados
    - Estados Accesibles
    - Estados Comunicantes
- Clases
    - Clase Recurrente
    - Clase Transitoria
    - Cadena Ergodica

Tipos de Cadenas de Markov

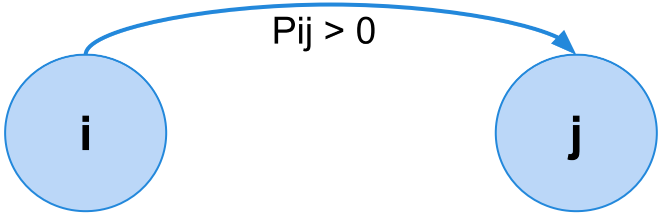
- Estado “j” es **Accesible** desde “i” en algún paso n si pij(n) > 0.

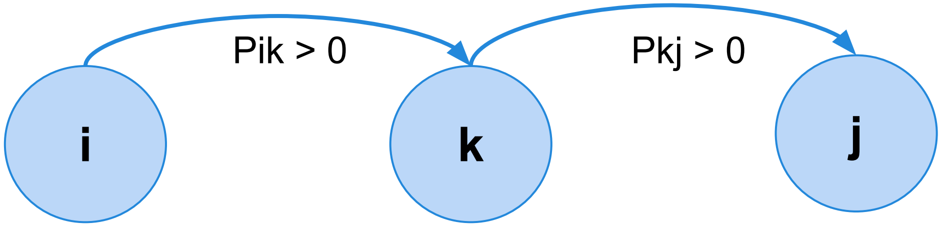
- Dos estados “i” y “j” son **Comunicantes** si “j” es accesible desde “i” y viceversa.

- Estado **absorbente**, caso particular de clase recurrente (una vez que la cadena lo ha alcanzado no puede abandonarlo).

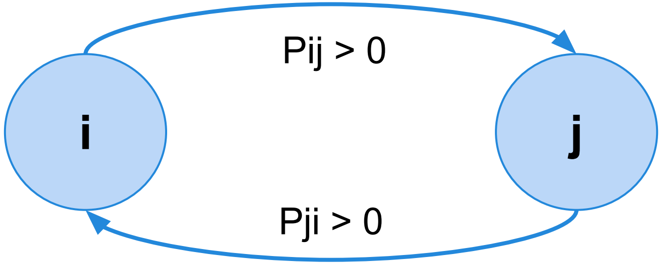
Tipos de Cadenas de Markov

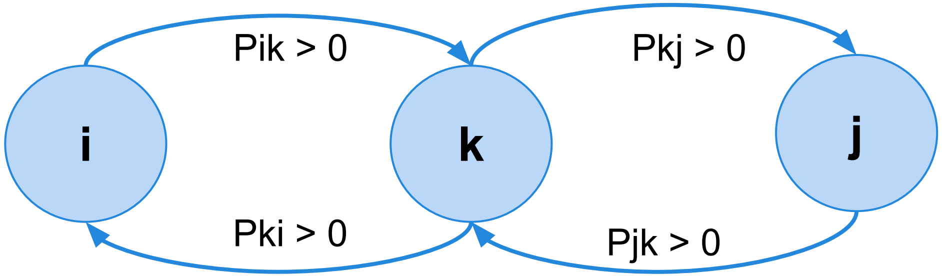
Estado “j” <u>accesible</u> desde “i” en “n” pasos (n es finito)





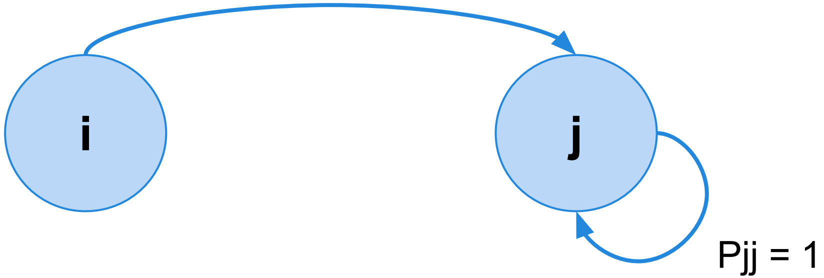
Estado “i” es <u>comunicante</u> con “j” en “n” pasos (n es finito)





Tipos de Cadenas de Markov

Estado “j” es <u>absorbente</u>



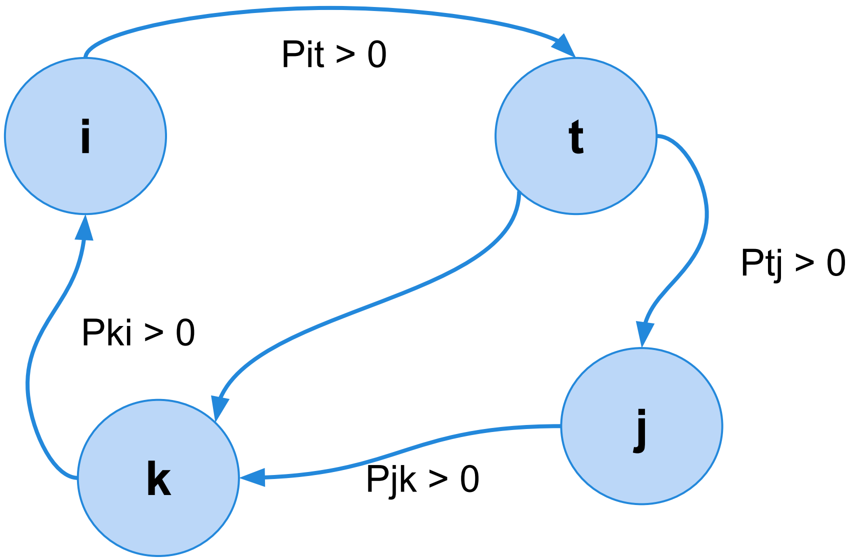
Tipos de Cadenas de Markov

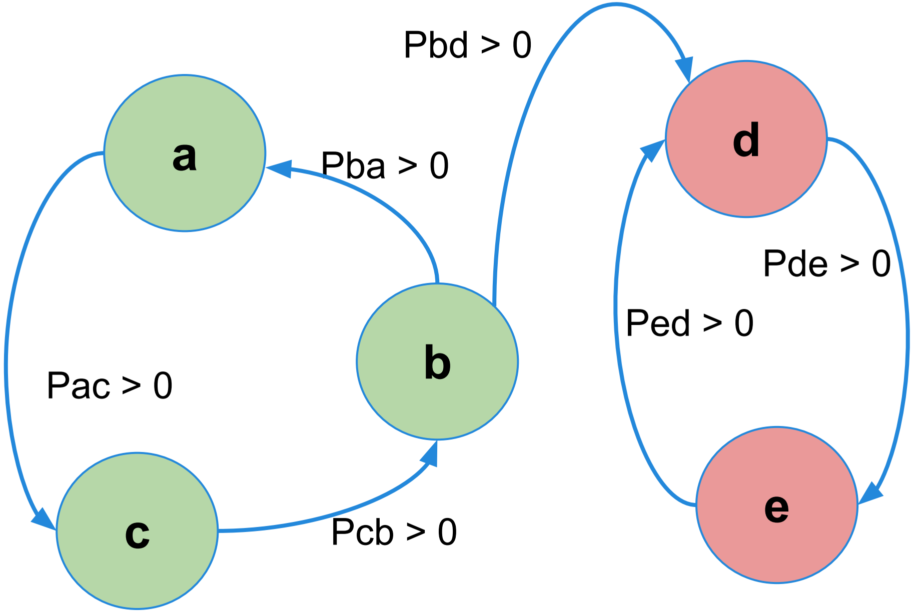
- **Clase**: conjunto de estados en una cadena de markov
- Una clase es **comunicante** si sus estados se comunican todos entre si.
- Clase **recurrente** es aquella cuando la probabilidad de que la cadena se encuentre en un estado de dicha clase después de ∞ transicíones es positiva.
- Una clase es **Transitoria** si la probabilidad de que la cadena se encuentre en un estado de dicha clase después de ∞ transiciones es nula;es decir que una vez que se ha alcanzado dicha clase, hay una probabilidad de nunca regresar.

Tipos de Cadenas de Markov

La clase "i-t-j-k" es comunicante y recurrente

La clase "a-b-c" es solo comunicante. La clase "d-e" es comunicante y recurrente.





Cadenas de Markov Ergodicas

- Son aquellas cadenas irreducibles (no pueden sub-dividirse en más clases) de clase recurrente donde todos sus estados son comunicantes.
- En el estado estacionario las probabilidades de estado se estabilizan a ciertos valores

Estado Estacionario

Existen 3 maneras de calcular el vector de probabilidad en el estado estacionario:

1. **Metodo directo**
2. **Chapman-Kolmogorov**
3. **Flujo Probabilistico**

Estado Estacionario: método directo

A medida que incremento el paso de la matriz de transición, comenzará a suceder que los valores de las columnas convergerán al mismo valor. Es decir que las probabilidades de transición de “n” pasos con “n” tendiendo a infinito serán igual no importa desde que estado se inicie la transición.

$$\lim_{n \mapsto \infty} P^n = \begin{bmatrix} p_0 & p_1 & p_2 & p_n \\ p_0 & p_1 & p_2 & p_n \\ p_0 & p_1 & p_2 & p_n \\ p_0 & p_1 & p_2 & p_n \end{bmatrix}$$

Estado Estacionario: chapman-kolmogorov

La probabilidad de estado en “n” cuando tiende a infinito será igual que la probabilidad de estado en “n-1”. De esta manera multiplicando el vector de probabilidad de estado por la matriz de transición de un paso dará por resultado el vector de probabilidad de estado en regimen permanente.

$$\lim_{n \mapsto \infty} p^{(n)} = \lim_{n \mapsto \infty} p^{(n-1)} . P_{ij}^{(1)}$$

$$\rightarrow p^{(n)} . P^{(1)} = p^{(n)}$$

$$\sum p_j = 1$$

Estado Estacionario: Flujo Probabilístico

Para cada estado “j” se establece que la probabilidad de arribar a dicho estado desde otro será igual que la probabilidad de estar en dicho estado y transicionar a otro. La suma de flujos probabilísticos que confluyen al nodo “j” será igual a la suma de flujos que salen del mismo.

Probabilidad de estar en el nodo “i” y transicionar al “j”

$$\sum_{\forall i \neq j} p_i \cdot p_{ij} = p_j \cdot \sum_{\forall k \neq j} p_{jk}$$

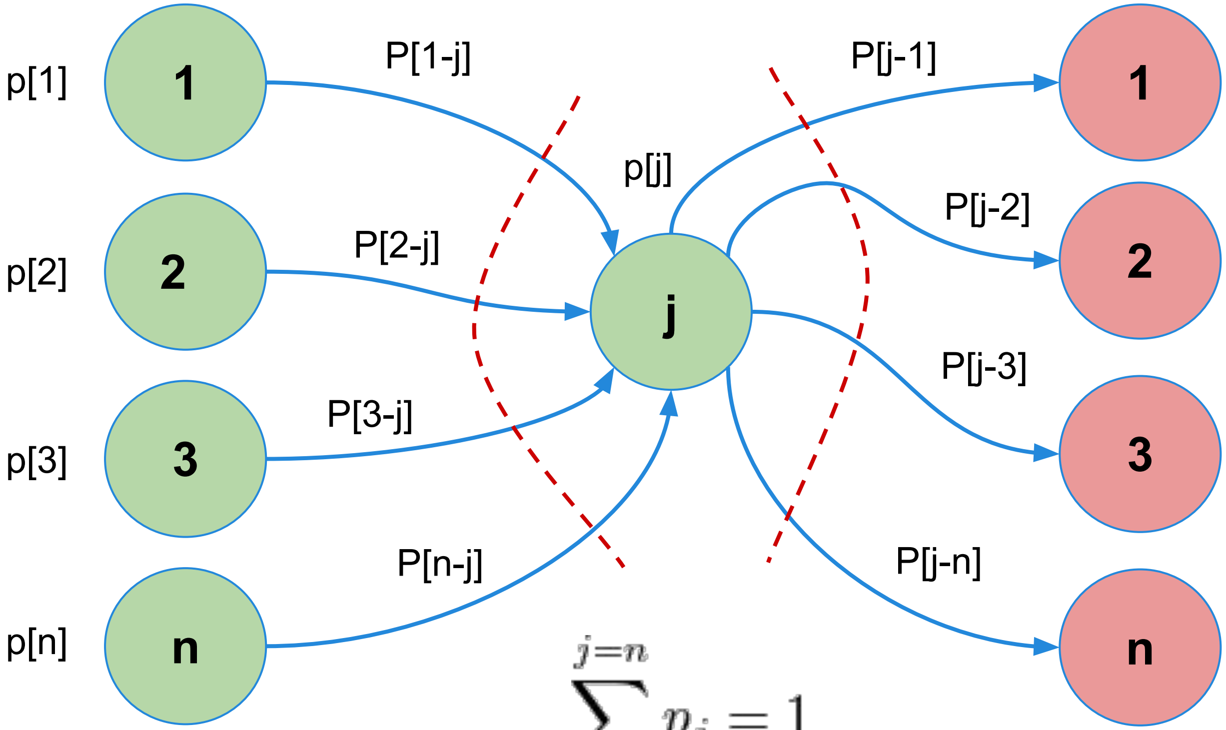
Probabilidad de estar en el nodo “j” y transicionar al “k”

Si realizamos el balance de flujos para cada nodo más la ecuación de cierre, podremos obtener el estado estacionario.

$$\sum p_j = 1$$

Estado Estacionario: Flujo Probabilístico

$$\sum_{\forall i \neq j} p_i \cdot p_{ij}$$



$$p_j \cdot \sum_{\forall k \neq j} p_{jk}$$

$$\sum_{j=1}^{j=n} p_j = 1$$

Caso de estudio: marketing digital

Toy example de Cadenas de Markov

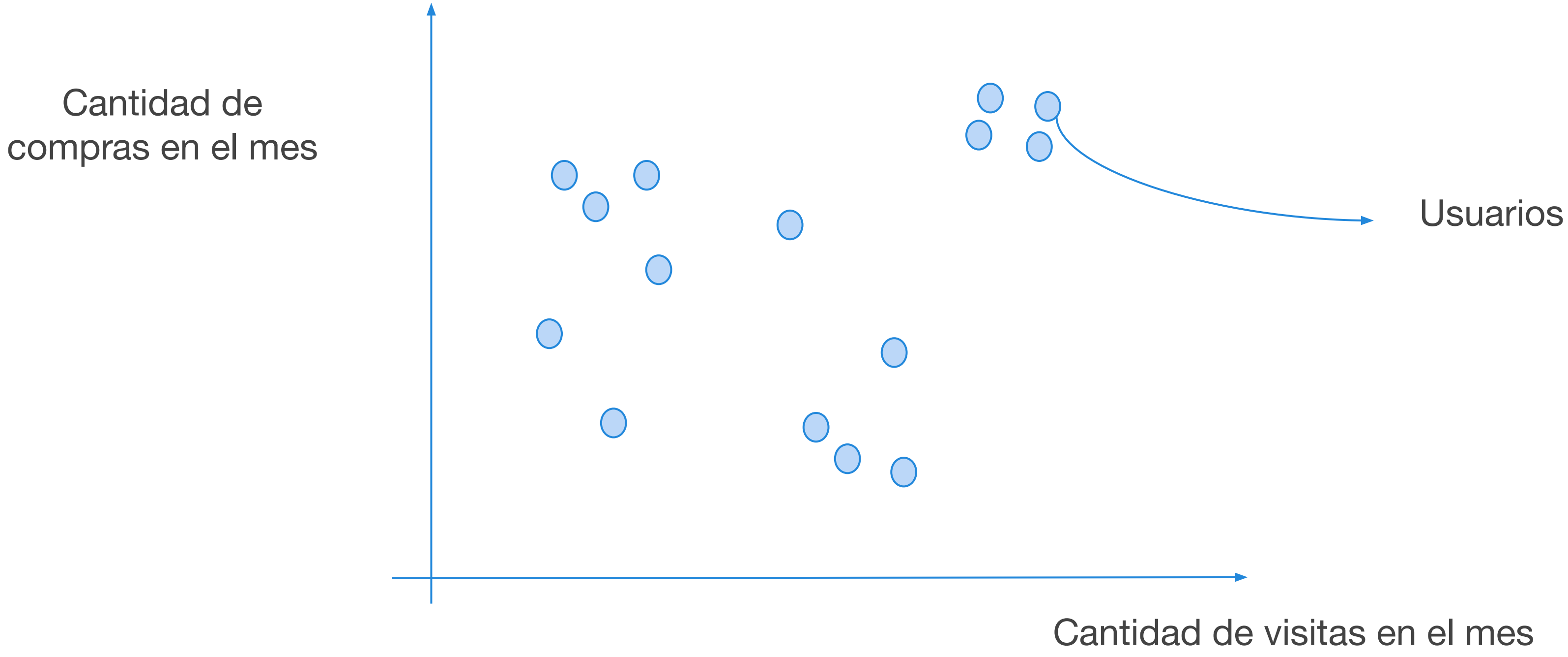
Marketing online (aka growth hacking)

User/marketing funnel

Métricas importantes en marketing online

- DAUs (Daily Active Users)
- MAUs (Montly active users)
- Daily Installs (#)
- Retention (days/months)
- Life time value ($)
- Impressions (#)
- Clicks (#)
- Cost per click ($)
- Conversion rate (%)
- Cantidad de visitas

Modelado del problema



Simulación Cadena de Markov en Python

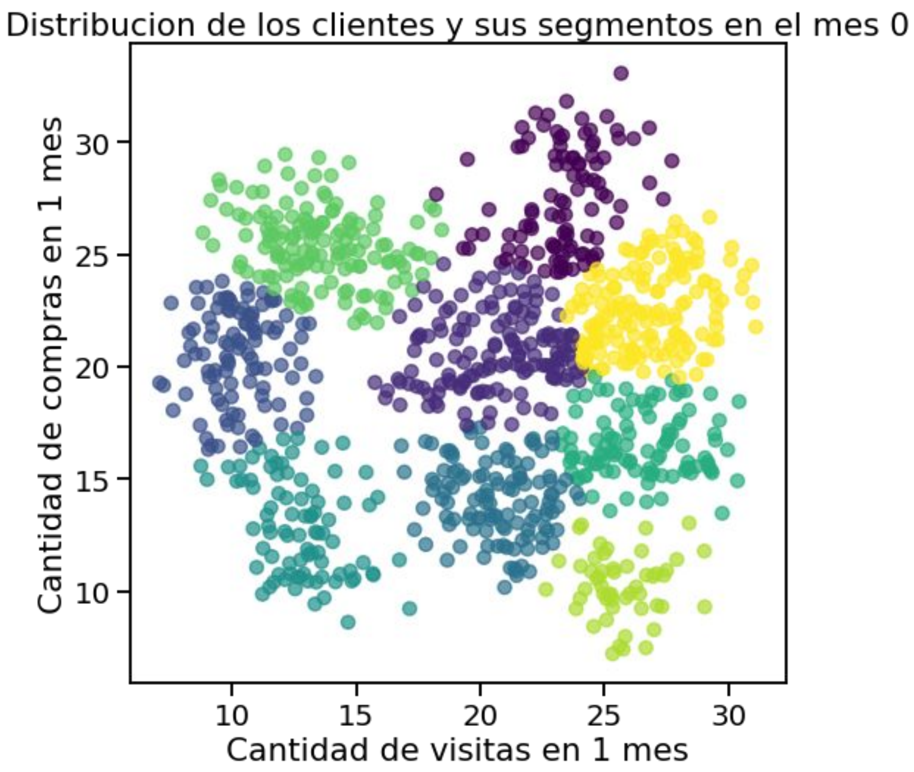
Pre-procesamiento

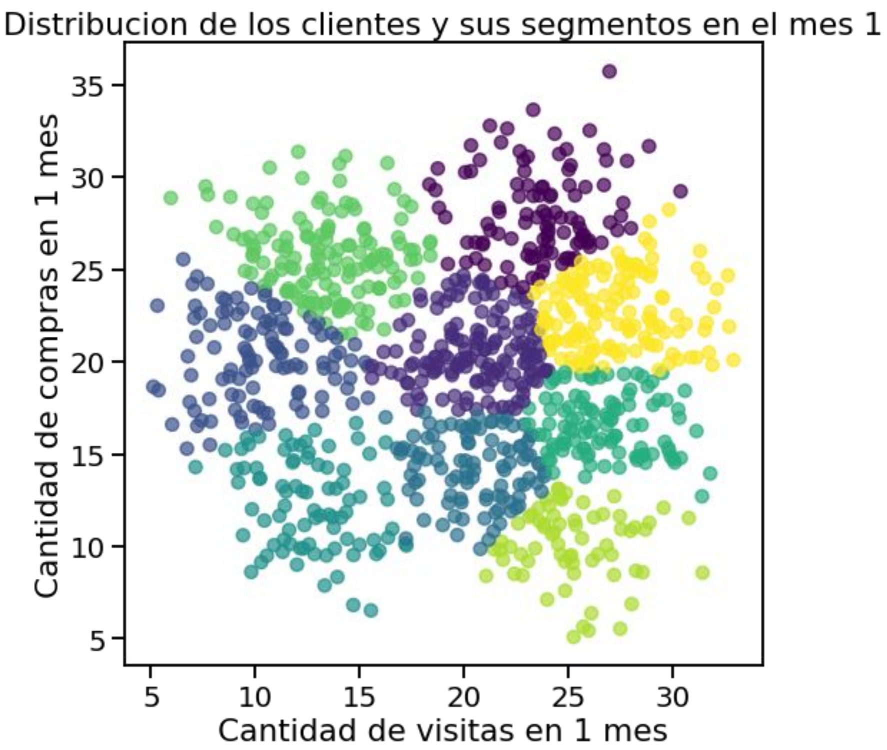
1. *Partimos de una base de datos de 1000 clientes. A cada cliente se le mide la cantidad de veces que visita el sitio web vs la cantidad de compras que realiza.*
2. *Localizamos los clientes en un espacio de 2 dimensiones (una dimension por indicador) para ver cómo se distribuyen durante el mes 0 y el mes 1.*
3. *Con esos mediciones decidimos segmentar los clientes en "k" categorías con un algoritmo de machine learning llamado "k-means". Las segmentaciones formarán grupos de clientes similares considerando los 2 indicadores mencionados. Realizamos la segmentación en ambos meses.*
4. *Definimos que cada segmento/grupo obtenido es un **estado** que los clientes pueden tomar.*

Modelado con Cadenas de Markov

5. Calculamos la transición de clientes entre el mes 0 y el 1 entre estados. Es decir que construimos la matriz de transición basándonos en la cantidad de clientes que migraron de un estado al otro entre el mes 0 y 1.
6. Con la distribución de clientes en los diferentes estados del mes 0 armamos el vector de probabilidad de estado inicial.
7. Simulamos la cadena de markov m meses hacia adelante.

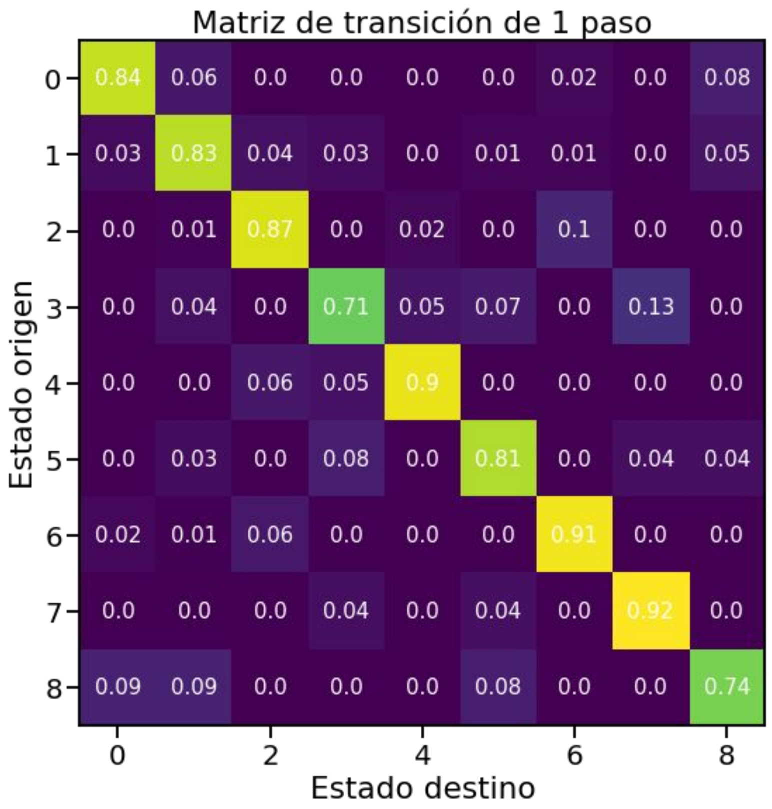
Segmentamos los usuarios en mes 0 y 1





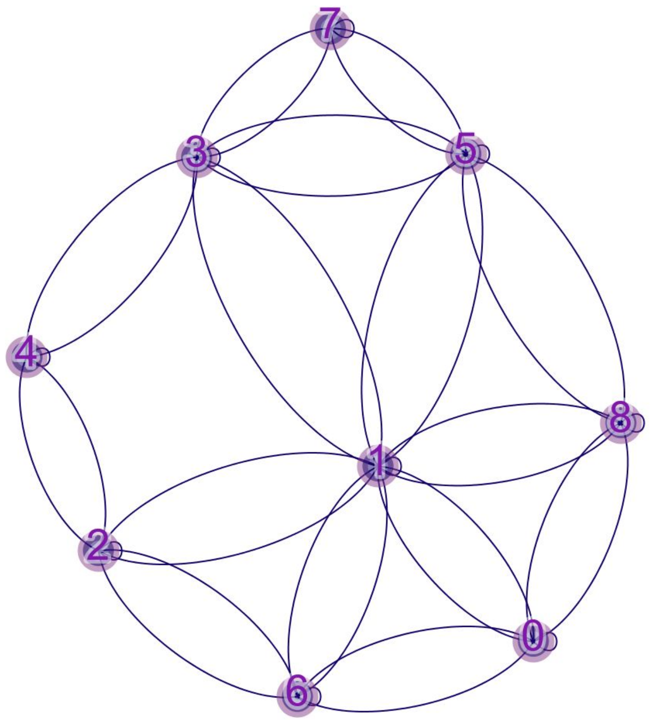
Calculamos la matriz de transición de 1 paso

Sabiendo en qué estado estaba cada cliente en el mes 0 y en cual estaba en el mes 1 podemos calcular la fracción de clientes que migran de un estado “i” a otro “j” en 1 mes.



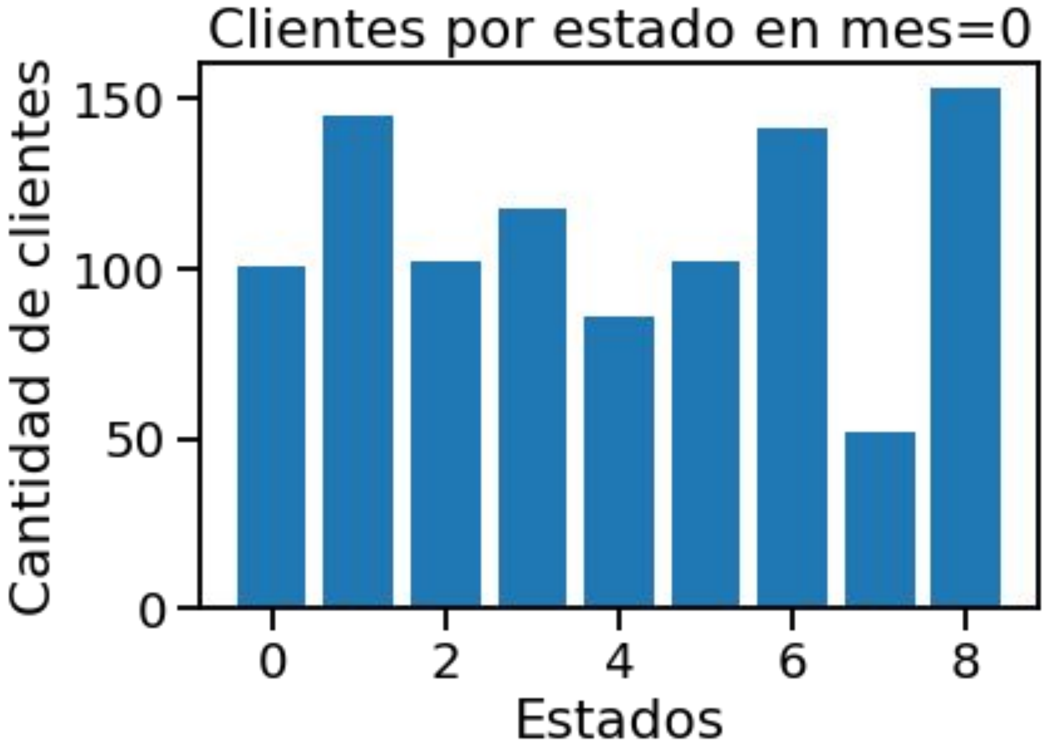
Calculamos la matriz de transición de 1 paso

Partiendo de la matriz de transición podemos construir un grafo donde el grado de la matriz es la cantidad de nodos (estados). Cada elemento de la matriz indica el peso del arco, la fila el origen y la columna el destino.



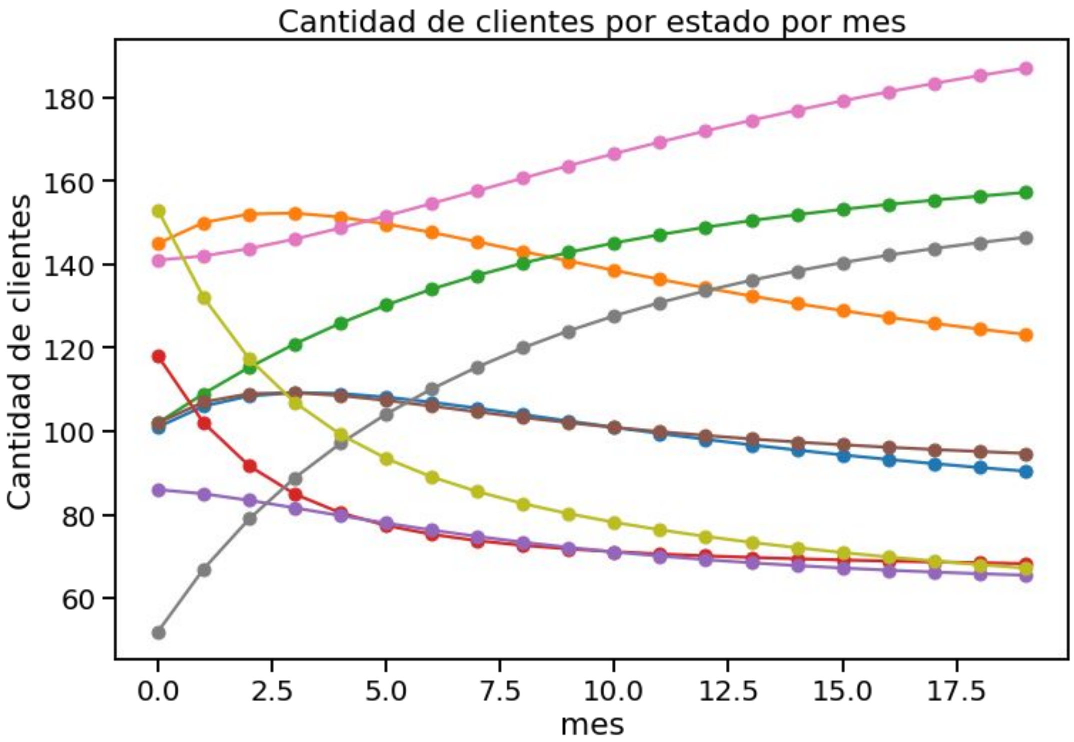
Determinamos el vector estado inicial

Con los datos iniciales del problema podríamos determinar la cantidad de clientes que hay por estado en el mes 0 y considerarlo como el vector de estado inicial. Si bien tenemos las cantidades absolutas y trabajaremos con ellas en este problema, podríamos pasar estas a valores de probabilidad determinando la fracción del total de clientes por estado.



Simulamos la cantidad de clientes por estado utilizando la cadena de markov definida

Simulamos la cantidad de clientes por estado considerando la matriz de transición obtenida y el vector de estado inicial en el mes 0. El resultado indicará como evolucionaría la cantidad de clientes por estado en cada mes considerando el vector de estado inicial y la matriz de transiciòn de 1 paso y dado una cantidad “m” de meses (o pasos). En el ejemplo realizamos la simulación con 20 meses (20 pasos). Observamos que el sistema converge.



Simulamos la cantidad de clientes por estado utilizando la cadena de markov definida

Recordar que en estas visualizaciones mostramos la cantidad de clientes en cada estado luego de “m” pasos dada una matriz de transiciòn y un estado inicial. Tranquilamente con estos datos podríamos transformar las cantidades absolutas de cada estado en probabilidades de estado.