Lecturas de Clase:

Introducción a Cadenas de Markov

Fredy Vides
Departamento de Matemática Aplicada, UNAH

1 Definición de Cadenas de Markov

- Una **cadena de Markov** es un proceso estocástico en tiempo discreto con un espacio de estados discreto, donde la probabilidad de transitar entre estados depende únicamente del estado actual.
- La **propiedad de Markov** establece que:

$$P(X_{t+1} = x_{t+1}|X_t = x_t, X_{t-1} = x_{t-1}, \dots) = P(X_{t+1} = x_{t+1}|X_t = x_t).$$

- **Homogeneidad:** Si la probabilidad de transición no depende del tiempo:

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

2 Matrices de Probabilidades de Transición

- Una cadena de Markov homogénea está caracterizada por su matriz de transición P, donde:

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix},$$

con $p_{ij} \geq 0$ y $\sum_j p_{ij} = 1$ para cada i.

- La probabilidad de transición en n pasos se calcula como P^n.
- La matriz P es estocástica (por filas): cada fila suma 1.

Notación.

Cuando no se especifica si una matriz A es estocástica por renglones o columnas, se asumirá que es estocástica por renglones o columnas, dependiendo del contexto. Cuando una matriz es estocástica por renglones y por columnas, se escribirá que es doblemente estocástica.

3 Propiedades Importantes

- **Clases de Comunicación:** Un estado j es accesible desde i si existe un $n \geq 1$ tal que $P_{ij}^n > 0$.

  Para dos estados i y j en el espacio de estados S, diremos que el estado j es accesible desde el estado i y escribiremos

  $$i \to j \quad \text{si} \quad \exists n \in \mathbb{Z}^+ \quad \text{tal que} \quad p_{ij}(n) > 0.$$

  Si $i \to j$ y $j \to i$, entonces diremos que el estado i se comunica con el estado j y escribiremos

  $$i \leftrightarrow j.$$

  La propiedad "$\leftrightarrow$" induce una *relación de equivalencia*. Esta relación produce una partición del espacio de estados. A estas clases de equivalencia las llamaremos *clases de comunicación*.

  Dado un estado $i \in S$, denotaremos a su clase de comunicación como $\mathcal{C}(i)$, por lo que

  $$i \leftrightarrow j \iff \mathcal{C}(i) = \mathcal{C}(j).$$

  Si $\mathcal{C}(i) = S$, entonces se dice que la cadena es *irreducible*.

- **Irreducibilidad:** Todos los estados se comunican entre sí.

- **Periodicidad:** El *período* de un estado $i \in S$ se define como:

  $$d(i) = \text{mcd}\{n \geq 1 : p_{ii}(n) > 0\}$$

  donde mcd denota el *máximo común divisor*.

  - Si $d(i) = 1$, diremos que i es un estado *aperiódico*.
  - Si $d(i) = k \geq 2$, diremos que i tiene período k.

Una cadena de Márkov se dice *aperiódica* si todos sus estados son aperiódicos, es decir, si

$$d(i) = 1 \quad \forall\, i \in S.$$

- **Tiempo de Primera Visita:** Si $\mathcal{C} \subset S$, definimos el *tiempo de primera visita* a $\mathcal{C}$ como la *variable aleatoria*

$$\tau_{\mathcal{C}} = \begin{cases} \min\{n > 0 \mid X_n \in \mathcal{C}\} & \text{si } \{n > 0 \mid X_n \in \mathcal{C}\} \neq \emptyset, \\ 1 & \text{si } \{n > 0 \mid X_n \in \mathcal{C}\} = \emptyset. \end{cases}$$

  Esto es, $\tau_{\mathcal{C}}$ denota la primera vez que la cadena entra al conjunto de estados $\mathcal{C}$.

- **Probabilidad de Primera Visita:**

  Se define

$$f_{ij}(n) = \mathbb{P}[X_n = j, X_{n-1} \neq j, \ldots, X_1 \neq j \mid X_0 = i]$$

  como la probabilidad de que una cadena que inicia en el estado i llegue al estado j por primera vez en n pasos, donde $f_{ij}(0) = 0$.

  En particular, cuando $i = j$, $f_{ii}(n)$ denota la probabilidad de regresar por primera vez al estado i en n pasos.

  Y se definen

$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}(n)$$

  como la probabilidad de una eventual visita a partir del estado i al estado j, y

$$f_{ii} = \sum_{n=1}^{\infty} f_{ii}(n)$$

  como la probabilidad de partir del estado i y regresar a él mismo en un tiempo finito.

- **Recurrencia:**

  En una cadena de Markov con espacio de estados S, diremos que:

  - i es un estado *recurrente* si $f_{ii} = 1$.
  - i es *transitorio* si $f_{ii} < 1$.

  O utilizando las probabilidades de transición en n pasos:

- i es recurrente si
$$\sum_{n=1}^{\infty} p_{ii}(n) = \infty.$$
- i es transitorio si
$$\sum_{n=1}^{\infty} p_{ii}(n) < \infty.$$

La recurrencia es una propiedad de clase, pues:

- Si i es recurrente e $i \leftrightarrow j$, entonces j es recurrente.
- Si i es transitorio e $i \leftrightarrow j$, entonces j es transitorio.

- **Tiempo Medio de Recurrencia:**

Se define como el *tiempo medio de recurrencia* de un estado recurrente j a partir del estado i como la *esperanza* de

$$\tau_{ij} = \min\{n \geq 1 : X_n = j \mid X_0 = i\}$$

y se denota por μ_{ij}:

$$\mu_{ij} = \mathbb{E}[\tau_{ij}] = \sum_{n=1}^{\infty} n f_{ij}(n),$$

donde $f_{ij}(n)$ representa la probabilidad de primera visita en n pasos.

Esta *esperanza* representa el número de pasos promedio que a la cadena le toma regresar al estado recurrente j.

En particular, cuando $i = j$, escribimos μ_i en lugar de μ_{ij}.

Se dice que un estado recurrente i es:

- *recurrente nulo* si $\mu_i = \infty$.
- *recurrente positivo* si $\mu_i < \infty$.

La recurrencia positiva es una propiedad de clase, pues:

- Si i es recurrente positivo e $i \leftrightarrow j$, entonces j es recurrente positivo.
- Si i es recurrente nulo e $i \leftrightarrow j$, entonces j es recurrente nulo.

- **Distribución Estacionaria:** Es un vector π tal que:

$$\pi P = \pi \quad \text{y} \quad \sum_{i} \pi_i = 1.$$

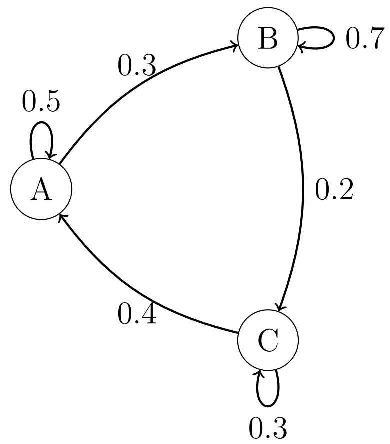
Representa una distribución invariante en el tiempo.

4 Ejemplo: Grafo de Transición

Consideremos un proceso con tres estados $\{A, B, C\}$ y la siguiente matriz de transición:

$$P = \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.1 & 0.7 & 0.2 \\ 0.4 & 0.3 & 0.3 \end{bmatrix}.$$

El grafo de transición asociado es:



5 Matrices Estocásticas

Las matrices estocásticas (por renglones o por columnas), fundamentales en el modelado de procesos estocásticos como las cadenas de Markov, tienen aplicaciones en comunicación de redes, sistemas de transporte y algoritmos como el PageRank de Google. Una matriz estocástica P es una matriz cuadrada donde $p_{ij} \geq 0$ y la suma de las entradas de cada columna (o fila) es igual a 1, es decir, $\sum_i p_{ij} = 1$ (o $\sum_j p_{ij} = 1$). En esta sección se exploran algunas de sus propiedades y aplicaciones.

5.1 Propiedades de las Matrices Estocásticas

Definición

Una matriz estocástica (por columnas) P satisface:

$$p_{ij} \geq 0 \quad \text{y} \quad \sum_{i=1}^{n} p_{ij} = 1 \quad \forall j.$$

Propiedades

1. Preservación de Vectores de Probabilidad: Si $\mathbf{v}$ es un vector de probabilidad ($\mathbf{v} \geq 0$ y $\sum_i v_i = 1$), entonces:

$$P\mathbf{v} \quad \text{también es un vector de probabilidad.}$$

2. Producto de Matrices Estocásticas: Si P y Q son matrices estocásticas, su producto PQ también es estocástico.

3. Invarianza con respecto a permutaciones: El resultado permutar renglones y/o columnas de una matriz estocástica A, da como resultado una matriz estocástica.

4. Dominancia del Eigenvalor Principal: El mayor valor propio de P es $\lambda_{\max} = 1$, y siempre existe un vector propio asociado que es no negativo.

5. Convergencia Iterativa: Para P regular, las potencias P^k convergen:

$$P^k \to \mathbf{v}\mathbf{w}^T \quad \text{cuando } k \to \infty,$$

   donde $\mathbf{v}$ es el vector propio dominante y $\mathbf{w}^T$ normaliza las columnas.

6. Combinaciones Convexas: Una combinación convexa de matrices estocásticas $\alpha P + (1-\alpha)Q$ con $0 \leq \alpha \leq 1$ también es estocástica.

7. Teorema de Perron-Frobenius para una matriz estocástica irreducible:
   - El valor propio dominante $\lambda = 1$ es único.
   - El vector propio correspondiente es positivo ($v_i > 0$ para todo i).

8. Normalización: Si P tiene entradas no negativas pero las columnas no suman 1, se puede normalizar dividiendo cada columna por su suma:

$$P'_{ij} = \frac{P_{ij}}{\sum_k P_{kj}}.$$

9. Preservación de la Uniformidad: Una matriz estocástica (por renglones) preserva el vector uniforme:

$$P\mathbf{u} = \mathbf{u}, \quad \text{donde } \mathbf{u} = \frac{1}{n}[1, 1, \ldots, 1]^T.$$

5.2 Aplicaciones de las Matrices Estocásticas

5.2.1 Comunicación en Redes

Las matrices estocásticas modelan sistemas de comunicación donde los nodos representan dispositivos y las aristas, las probabilidades de comunicación. La matriz de transición P describe las probabilidades de mover un paquete de un nodo a otro.

Evolución de Paquetes

Si $\mathbf{v}(t)$ representa la distribución de paquetes en t, su evolución está dada por:

$$\mathbf{v}(t+1) = P\mathbf{v}(t).$$

En redes fuertemente conexas, $P^k\mathbf{v}(0)$ converge a la distribución estacionaria $\mathbf{v}_\infty$.

Redundancia y Tolerancia a Fallos

En redes robustas, se reconfiguran las probabilidades en caso de fallos para garantizar conectividad:

$$\hat{p}_{ij} = \begin{cases} p_{ij} + \delta_j & \text{si } j \text{ redirige a } i, \\ p_{ij} & \text{en caso contrario,} \end{cases}$$

donde δ_j asegura que $\sum_i \hat{p}_{ij} = 1$.

5.3 Sistemas de Transporte

En transporte, las matrices estocásticas modelan probabilidades de tránsito entre estaciones. Por ejemplo, en una red de transporte con n estaciones:

$$\mathbf{v}(t+1) = P\mathbf{v}(t),$$

donde $\mathbf{v}(t)$ es la distribución de pasajeros en t.

5.4 Algoritmo PageRank de Google

El grafo web se representa como un grafo dirigido, donde los nodos son páginas y las aristas son enlaces. La matriz de transición P está definida como:

$$p_{ij} = \begin{cases} \frac{1}{\text{grado de salida}(j)} & \text{si } (j,i) \in E, \\ 0 & \text{en caso contrario.} \end{cases}$$

Modelo del Navegador Aleatorio

Para garantizar la conectividad fuerte, se introduce un factor de teletransportación α:

$$P_{\text{Google}} = \alpha P + (1-\alpha)\frac{\mathbf{1}\mathbf{1}^T}{n},$$

donde $\mathbf{1}$ es el vector de unos y n es el número de páginas.

Distribución Estacionaria

El vector PageRank $\mathbf{r}$ es la distribución estacionaria de P_{Google}, definida como:

$$\mathbf{r} = P_{\text{Google}}\mathbf{r}.$$

Este vector clasifica las páginas según su importancia.

6 Ejercicios

1. Demostrar las propiedades de la sección §5.1. Cuando su argumento requiera conceptos o proposiciones adicionales de dominio público del álgebra lineal intermedia o avanzada, no requiere demostrar dichos conceptos o proposiciones adicionales.
2. Determinar si una matriz de transición dada describe una cadena de Markov irreducible.
3. Calcular P^2 y P^3 para una matriz de transición simple y analizar las probabilidades de alcanzar ciertos estados.
4. Identificar la distribución estacionaria para una matriz de transición con tres estados.
5. Simular una cadena de Markov con datos sintéticos, graficar la evolución de las probabilidades de los estados y validar la convergencia hacia la distribución estacionaria.

References

[1] Goswami, A. y Rao, B. V. *A Course in Applied Stochastic Processes.* Hindustan Book Agency, 2006.

[2] Ito, Kiyosi. *Essentials for Stochastic Processes.* Association of Mathematical Society, 2006.

[3] Shapiro, A., Dentcheva, D. y Ruszczynski, A. *Lectures on Stochastic Programming: Modeling and Theory*. Society of Industrial and Applied Mathematics, 2009.

[4] Stroock, D. W. *An Introduction to Markov Processes*. Graduate Texts in Mathematics Series, Springer Verlag, 2005.

[5] Van Kampen, N. G. *Stochastic Processes in Physics and Chemistry*. Third Edition, North Holland, 2007.

[6] Saad, Y. *Iterative Methods for Sparse Linear Systems*. SIAM, 2003.

[7] Vides, F. *Métodos Numéricos y Modelación Computacional*. Disponible en: `https://fredyvides.github.io/MNMC.pdf`, 2019.

[8] Markovsky, I., Van Huffel, S., Willems, J. C., De Moor, B. *Exact and Approximate Modeling of Linear Systems: A Behavioral Approach*. SIAM, 2005.

[9] Boyd, S., Vandenberghe, L. *Introduction to Applied Linear Algebra: Vectors, Matrices, and Least Squares*. Cambridge University Press, 2018.

[10] Quarteroni, A., Saleri, F., Gervasio, P. *Scientific Computing with MATLAB and Octave*. Fourth Edition, Springer, 2014.