# INSTITUTO POLITÉCNICO NACIONAL

## Escuela Superior de Física y Matemáticas

# Teorema de Perron-Frobenius Aplicado a las Matrices Estocásticas

TESIS
QUE PARA OBTENER EL TÍTULO DE
LICENCIADO EN FÍSICA Y MATEMÁTICAS

PRESENTA
Luis Ángel Filio Rivera

Director de Tesis
Manuel Robles Bernal

México, D. F. Agosto del 2006

*A mi Madre quien día con día se esfuerza por hacer de mi hermano y de mí personas de éxito, sin olvidar el profundo amor que deposita en cada una de las actividades que compartimos.*

*A mi Padre quien nos ha sabido conducir no solo durante nuestra formación académica, sino a lo largo de nuestra vida complementando siempre con el amor que lo caracteriza.*

*A mi Hermano quien con todo cario ha sabido ser pilar en los momentos de flaqueza.*

*A "El Mono" quien ha cuidado de mí durante mis enfermedades y con quien he compartido y disfrutado mi salud.*

# Agradecimientos

A mi Perica por ser un infinito momento de inspiración

A mi Profesor y Gran Amigo Manuel por haber sido soporte durante toda mi carrera y el tremendo trabajo que compartimos durante la elaboración de esta tesis (y lo que falta!).

A P. Lam por haber sido consejero durante mi carrera y su apoyo en la elaboracón de esta tesis.

A Quintín Flores, Fco. Ramírez y a E. de La Cruz por su apoyo en la elaboracón de esta tesis.

# Introducción

El trabajo de tesis tiene como objetivo principal el obtener resultados para encontrar distribuciones estacionarias de cadenas de Markov, basándose fundamentalmente en el importante *Teorema de Perron-Frobenius.* El trabajo está compuesto de tres capítulos y un apéndice; el apéndice contiene el desarrollo de Cadenas de Markov resaltando los resultados importantes en forma matricial así como la clasificación de los estados.

El capítulo 1 consiste principalmente del estudio de la *Matriz de Equilibrio* y bajo que condiciones podemos determinar su existencia dado que sus filas que son iguales representan a la distribución estacionaria asociada a la cadena. Se establece un resultado que relaciona la distribución con el tiempo esperado de retorno, que es de importancia en las aplicaciones de estos procesos y para realizar los cálculos se aplica el Teorema de la representación espectral para matrices diagonalizables.

Como se mencionó, el interés es utilizar la *Teoría de Perron-Frobenius*, que conforma el contenido del capítulo 2 en donde se desarrolla un estudio en la estructura de las matrices no negativas que da un conocimiento de sus valores y vectores propios así como de una representación canónica dada por una relación de equivalencia en su conjunto de índices. Ésto permite obtener una serie de importantes resultados que se aplicarán en el desarrollo del capítulo 3 sobre el estudio de las matrices estocásticas para la obtención de la distribución estacionaria, concluyendo con ejemplos de cadenas importantes como la *caminata aleatoria.*

# Índice general

# Capítulo 1

# Matriz de Equilibrio y Representación Espectral

En este capítulo se contempla un estudio en la estructura de la *Matriz de Equilibrio* y una serie de resultados en Álgebra Lineal para obtener el teorema de *Representación Espectral.*

## 1.1. Tiempo Esperado de Recurrencia

Bajo ciertas condiciones existe una distribución estacionaria para una cadena de Markov la cual está dada en términos del Tiempo Esperado de Recurrencia, que determina el tiempo promedio de espera del proceso para retornar. En este capítulo se presentan resultados que determinan las condiciones para la existencia y unicidad.

**Definición 1.1.1.** *Se define el* **Tiempo Esperado de Recurrencia** $\mu(i)$ *del estado* $i \in \mathcal{S}$ *como*

$$\mu(i) = \begin{cases} \infty & si\ i\ es\ Transitivo \\ \displaystyle\sum_{k=i}^{\infty} k\mathcal{P}(T_i = k \mid X_0 = i) & si\ i\ es\ Recurrente \end{cases}$$

En base a esta definición podemos clasificar a los estados recurrentes por

- El estado $i$ es llamado **Recurrente Positivo** si

$$\mu(i) < \infty$$

- El estado $i$ es llamado **Recurrente Nulo** si

$$\mu(i) = \infty$$

**Espacios de Estados Finitos**

**Teorema 1.1.1.** *Una Cadena de Markov irreducible finita con Espacio de Estados $\mathcal{S}$ tiene una única Distribución Estacionaria $\pi'$ dada a saber por:*

$$\pi'(i) = \frac{1}{\mu(i)} \tag{1.1}$$

*Demostración.* (Unicidad)

Si $\pi' = (\pi(i_1), \pi(i_2), \cdots, \pi(i_n))$ es una Distribución Estacionaria, entonces

$$\pi'\mathbf{P} = \pi'$$

esto significa que

$$\sum_{i=1}^{N} \pi(i)P(i,j) = \pi(j) \qquad \forall j \in \mathcal{S}$$

consideremos la Distribución Inicial $\mathcal{P}(X_0 = i) = \pi(i)$
demostremos que el producto

$$\pi(i)\mu(i) = 1$$

en efecto

$$
\begin{aligned}
\pi(i)\mu(i) &= \pi(i)\sum_{j=1}^{\infty} j\mathcal{P}_i(T_i = j) \\
&= \sum_{j=1}^{\infty} j\mathcal{P}_i(T_i = j)\pi(i) \\
&= \sum_{j=1}^{\infty} \mathcal{P}_i(T_i \geq= j)\pi(i) \\
&= \sum_{j=1}^{\infty} \mathcal{P}(T_i \geq j, X_0 = i) \\
&= \mathcal{P}(X_0 = i) + \sum_{j=2}^{\infty} \mathcal{P}(X_0 = i, X_1 \neq i, \cdots, X_{j-1} \neq i) \\
&= \mathcal{P}(X_0 = i) + \sum_{j=2}^{\infty} [\mathcal{P}(X_1 \neq i, \cdots, X_{j-1} \neq i) \\
& \quad -\mathcal{P}(X_0 \neq i, X_1 \neq i, \cdots, X_{j-1} \neq i)
\end{aligned}
$$

dado que la distribución inicial es Estacionaria

$$
\begin{aligned}
&= \mathcal{P}(X_0 = i) + \sum_{j=2}^{\infty} [\mathcal{P}(X_0 \neq i, \cdots, X_{j-2} \neq i)] \\
& \quad -[\mathcal{P}(X_0 \neq i, X_1 \neq i, \cdots, X_{j-1} \neq i)]
\end{aligned}
$$

esta suma es telescópica, por lo tanto

$$
\begin{aligned}
&= \mathcal{P}(X_0 = i) + \mathcal{P}(X_0 \neq i) \\
& \quad - \lim_{j\to\infty} P(X_0 \neq i, X_1 \neq i, \cdots, X_j \neq i)
\end{aligned}
$$

como el estado es recurrente se tiene

$$\begin{aligned} &= \mathcal{P}(X_0 = i) + \mathcal{P}(X_0 \neq i) \\ &= 1 \end{aligned}$$

de aquí que:

$$\pi(i) = \frac{1}{\mu(i)}$$

$\square$

Antes de demostrar la existencia de la distribución estacionaria se da la siguiente definición

**Definición 1.1.2.** *Dados los estados $i, j \in \mathcal{S}$ se define $a_i(j)$ como el número esperado de visitas al estado $i$ entre visitas al estado $j$ esto es:*

$$a_i(j) = \sum_{k=0}^{\infty} \mathcal{P}(X_k = i, T_m > k \mid X_0 = m)$$

*Demostración.* (*Existencia*)
Si la Cadena de Markov es finita e irreducible entonces existe una única distribución estacionaria la cual está dada por la ecuación (1.1)
Es de esperarse que

$$\mu(i)a_i(m) = \mu(m)$$

y por lo tanto

$$\frac{1}{\mu(i)} = \frac{a_i(m)}{\mu(m)}$$

Se tienen las siguientes dos afirmaciones.
Afirmación 1

$$\sum_{i=1}^{N} a_i(m) = \mu(m)$$

en efecto

$$\begin{aligned}
\mu(m) &= \sum_{k=0}^{\infty} \mathcal{P}(T_m > k \mid X_0 = m) \\
&= \sum_{i=1}^{N} \sum_{k=0}^{\infty} \mathcal{P}(X_k = i, T_m > k \mid X_0 = m) \\
&= \sum_{i=1}^{N} a_i(m)
\end{aligned}$$

Afirmación 2

$$a_j(m) = \sum_{i=1}^{N} a_i(m) p_{ij}$$

supongamos primero que $j \neq m$ entonces

$$\begin{aligned}
a_j(m) &= \sum_{k=0}^{\infty} \mathcal{P}(X_k = j, T_m > k \mid X_0 = m) \\
&= \sum_{k=1}^{\infty} \mathcal{P}(X_k = j, T_m > k \mid X_0 = m)
\end{aligned}$$

dado que $j \neq m$

$$\begin{aligned}
&= \sum_{k=1}^{\infty} \mathcal{P}(X_k = j, T_m > k-1 \mid X_0 = m) \\
&= \sum_{k=1}^{\infty} \sum_{i=1}^{N} \mathcal{P}(X_k = j, X_{k-1} = i, T_m > k-1 \mid X_0 = m)
\end{aligned}$$

por la propiedad Markoviana

$$
\begin{aligned}
&= \sum_{k=1}^{\infty}\sum_{i=1}^{N} p(i,j)\mathcal{P}(X_{k-1}=i, T_m > k-1 \mid X_0 = m) \\
&= \sum_{i=1}^{N} p(i,j)\sum_{k=0}^{\infty} \mathcal{P}(X_k = i, T_m > k \mid X_0 = m) \\
&= \sum_{i=1}^{N} p(i,j)a_i(m)
\end{aligned}
$$

si $i = j$ entonces

$$a_m(m) = 1$$

por recurrencia

$$
\begin{aligned}
&= \sum_{k=1}^{\infty} \mathcal{P}(T_m = k \mid X_0 = m) \\
&= \sum_{k=1}^{\infty}\sum_{i=1}^{N} \mathcal{P}(X_{k-1}=i, T_m = k \mid X_0 = m) \\
&= \sum_{k=1}^{\infty}\sum_{i=1}^{N} p(i,k)\mathcal{P}(T_m > k-1 \mid X_0 = m) \\
&= \sum_{k=1}^{N} p(i,k)a_i(m)
\end{aligned}
$$

en base a estas dos afirmaciones se tiene la existencia de la distribución estacionaria la cual esta dada por la ecuación (1.1) $\square$

## 1.2. Matriz de Equilibrio

Dado que los coeficientes de la matriz $\mathbf{P}^m$ representa las probabilidades de transición en *m-pasos* y las filas distribuciones condicionales. Nuestro interés es analizar el comportamiento de éstas cuando $m \to \infty$
Si existe el límite

$$\lim_{m\to\infty} \mathbf{P}^m = \mathbf{H}$$

$\mathbf{H} = (h_{ij})$ es llamada la **Matriz Límite** , dado que

$$h_{ij} = \lim_{m\to\infty} p_{ij}^{(m)}$$

se tiene

- $0 \leq h_{ij} \leq 1$
- $\sum_j h_{ij} = 1$

**Definición 1.2.1.** *Si las filas son idénticas, esto es:*

$$h_{ij} = h_{kj} \qquad \forall i \ \ \forall k$$

*y dado que*

$$p_{ij}^{(m)} = \mathcal{P}(X_m = j | X_0 = i)$$

*se concluye que la distribución Límite es independiente de la distribución inicial y en este caso se le llama* **La Matriz de Equilibrio.**
*Se define*

$$\pi_i(j) \doteq h_{ij}$$

Supóngase que la distribución límite $\mathbf{H}$ satisface:

$$\mathbf{HP} = \mathbf{H}$$

entonces se tiene la igualdad

$$\pi'_{\mathbf{i}}\mathbf{P} = \pi'_{\mathbf{i}}$$

esto implica que la distribución $\pi'_i$ dada por la fila i es una distribución estacionaria.

**Valores Propios de una Matriz Estocástica**

La matriz $\mathbf{P}$ tiene como vector propio al vector columna $(1, 1 \cdots, 1)^t$ con valor propio 1 ya que

$$\mathbf{P}\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

Los valores propios de una matriz estocástica tienen la siguiente característica.

**Proposición 1.2.1.** *Si $\lambda$ es un valor propio de la matriz estocástica $\mathbf{P}$ entonces*

$$|\, \lambda \,| \leq 1$$

*Demostración.* Sea $(x_1, x_2, \cdots, x_n)^t$ el vector propio asociado al valor propio $\lambda$ entonces

$$\mathbf{P}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

para cada $i$ se tiene

$$\sum_j p(i,j)x_j = \lambda x_i$$

por lo tanto

$$|\, \lambda \,||\, x_i \,| \leq \sum_j p(i,j) \,|\, x_j \,|$$

consideremos

$$| x_k |= max\{| x_1 |, | x_2 |, \cdots, | x_n |\}$$

entonces

$$| \lambda || x_k |\leq \sum_j p(k,j) | x_j | \leq| x_k | \sum_j p(k,j) =| x_k |$$

por lo tanto

$$| \lambda |\leq 1$$

□

En base a este resultado se tiene

**Proposición 1.2.2.** *Si la matriz* $\mathbf{P}$ *es diagonalizable entonces la distribución límite existe si el único valor propio con magnitud* $1$ *es* $\lambda = 1$

por lo tanto

$$\mathbf{H} = \lim_{m\to\infty} \mathbf{C}^{-1}\mathbf{D}^m\mathbf{C} = \mathbf{C}^{-1}\mathbf{E}\mathbf{C}$$

donde $\mathbf{E} = diag\{e_{11}, e_{22}, \cdots, e_{nn}\}$ con

$$e_{ii} = \begin{cases} 0 & si\ d_{ii} < 1 \\ 1 & si\ d_{ii} = 1 \end{cases}$$

Dado que $\mathbf{P} = \mathbf{C}^{-1}\mathbf{D}\mathbf{C}$ entonces

$$\mathbf{P}\mathbf{C}^{-1} = \mathbf{C}^{-1}\mathbf{D}$$

de aquí que las columnas de $\mathbf{C}^{-1}$ son los vectores propios por la derecha de

**P** y como

$$\mathbf{CP} = \mathbf{DC}$$

las filas de **C** son los vectores propios por la izquierda de **P**

Supóngase que **P** tiene como único valor propio de magnitud 1 a $\lambda = 1$ y que los vectores propios por la izquierda asociado a $\lambda = 1$ está generado por el vector

$$\nu = (\nu_1, \nu_2, \cdots, \nu_n) \quad 0 \leq \nu_i \leq 1 \forall i$$

y además

$$\sum_i \nu_i = 1$$

entonces **P** tiene una distribución de equilibrio dada por $\nu$ en efecto

$$\begin{pmatrix} 1 & \cdots \\ 1 & \cdots \\ \vdots & \ddots \\ 1 & \cdots \end{pmatrix} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} \nu_1 & \cdots & \nu_n \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \end{pmatrix} = \begin{pmatrix} \nu_1 & \nu_2 & \cdots & \nu_n \\ \nu_1 & \nu_2 & \cdots & \nu_n \\ \vdots & \vdots & \ddots & \vdots \\ \nu_1 & \nu_2 & \cdots & \nu_n \end{pmatrix}$$

## 1.3. Representación Espectral

Si la matriz $\mathbf{P}$ es Diagonalizable entonces

$$\mathbf{P} = \mathbf{C}^{-1}\mathbf{D}\mathbf{C}$$

donde $\mathbf{D}$ es una matriz diagonal formada por los valores propios, esto es:

$$\mathbf{D} \doteq diag\{\lambda_{11}, \lambda_{22}, \cdots, \lambda_{nn}\}$$

entonces

$$\mathbf{P}^m = \mathbf{C}^{-1}\mathbf{D}^m\mathbf{C}$$

donde

$$\mathbf{D}^m \doteq diag\{\lambda_{11}^m, \lambda_{22}^m, \cdots, \lambda_{nn}^m\}$$

**Definición 1.3.1.** *Consideremos la matriz* $\mathbf{P}$ *diagonalizable entonces existe una matriz no singular* $\mathbf{C}$ *y una matriz diagonal* $\mathbf{D}$ *que satisface:*

$$\mathbf{P} = \mathbf{C}^{-1}\mathbf{D}\mathbf{C}$$

Denotemos por $\alpha(j)$ la j-ésima columna de $\mathbf{C}^{-1}$ y por $\beta(i)$ la i-ésima fila de $\mathbf{C}$,dado que

$$\mathbf{C}\mathbf{P} = \mathbf{D}\mathbf{C}$$

entonces

$$\beta(i)\mathbf{P} = \lambda_i\beta(i)$$

donde $\mathbf{D} = \{\lambda_1, \cdots, \lambda_n\}$ por lo tanto la matriz $\mathbf{C}$ está formada por los vectores propios fila de $\mathbf{P}$y como

$$\mathbf{P}\mathbf{C}^{-1} = \mathbf{C}^{-1}\mathbf{D}$$

entonces

$$\mathbf{P}\alpha(j) = \lambda_j\alpha(j)$$

esto nos dice que $\mathbf{C}^{-1}$ está formada por los vectores propios columna de $\mathbf{P}$

**Observación 1.3.1.** *Dado que* $\mathbf{P} = \mathbf{C}^{-1}\mathbf{D}\mathbf{C}$ *entonces*

$$\mathbf{P}^k = \mathbf{C}^{-1}\mathbf{D}^k\mathbf{C}$$

tenemos los siguientes dos lemas

**Lema 1.3.1.**

$$tr(AB) = tr(BA)$$

**Lema 1.3.2.**

$$tr(\mathbf{P}^k) = \sum_i \lambda_i^k$$

Si $v$ es un vector propio columna de $\mathbf{P}$con $\lambda_1$ como valor propio y $w$ es un eigenvector fila con valor propio $\lambda_2$ con $\lambda_1 \neq \lambda_2$ entonces los vectores propios son ortogonales, en efecto

$\mathbf{P}v = \lambda_1 v \;\Rightarrow\; w\mathbf{P}v = \lambda_2 wv = \lambda_1 wv \;\Rightarrow\; (\lambda_1 - \lambda_2)wv = 0 \;\Rightarrow\; wv = 0 \;\Rightarrow$ $w$ *&* $v$ *son ortogonales*

La representación espectral que estableceremos está considerada para matrices diagonalizables.

$$\mathbf{P} = \mathbf{C}^{-1}\mathbf{D}\mathbf{C}$$

dado que

$$\mathbf{C}\mathbf{C}^{-1} = I \;\Rightarrow\; \beta(i)\alpha(j) = \delta_{ij}$$

**Definición 1.3.2.** *definamos las matrices* $\mathbf{B}_k$ $k = 1, 2, \cdots, n$ *como*

$$\mathbf{B}_k \doteq \alpha(k)\beta(k)$$

ahora bien:

$$\begin{aligned} \mathbf{B}_k\mathbf{B}_l &= \alpha(k)\beta(k)\alpha(l)\beta(l) \\ &= \alpha(k)\delta_{kl}\beta(l) \end{aligned}$$

por lo tanto

$$\mathbf{B}_k\mathbf{B}_l = \begin{cases} 0 & si\ k \neq l \\ \mathbf{B}_k & si\ k = l \end{cases}$$

notemos que

$$p_{ij} = \sum_{l} \sum_{k} c_{ik}^{(-1)} d_{kl} c_{lj} = \sum_{k} \lambda_k c_{ik}^{(-1)} c_{kj}$$

se tiene que $c_{ik}^{(-1)} c_{kj}$ es el coeficiente en la fila $i$ columna $j$ de $\alpha(k)\beta(k)$ y por consiguiente se tiene la representación de $\mathbf{P}$ como:

$$\mathbf{P} = \sum_{k=1}^{n} \lambda_k \mathbf{B}_k$$

que es llamada la *representación espectral* de la matriz $\mathbf{P}$

**Teorema 1.3.1.** *Sea* $\mathbf{P}$ *una matriz diagonalizable con valores propios asociados* $\lambda_1, \cdots, \lambda_n$ *y tal que*

$$\mathbf{P} = \mathbf{C}^{-1}\mathbf{D}\mathbf{C}$$

*donde* $\mathbf{D} = diag\{\lambda_1, \cdots, \lambda_n\}$
*Sea* $\mathbf{B}_k$ *la matriz que se obtiene al múltiplicar la k-ésima columna de* $\mathbf{C}^{-1}$ *con la k-ésima fila de la matriz* $\mathbf{C}$ *entonces se tiene que:*

$$\mathbf{P} = \lambda_1 \mathbf{B}_1 + \cdots + \lambda_n \mathbf{B}_n$$

*y además*

$$\mathbf{P}^k = \lambda_1^k \mathbf{B}_1 + \cdots + \lambda_n^k \mathbf{B}_n$$

Este resultado nos permite saber de la existencia y obtener la matriz límite $\mathbf{H}$

## Ejemplo

Consideremos una C.M. $(X_n)_{n\geq 0}$ cuya matriz de transición es

$$\mathbf{P} = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$$

Los valores propios de la matriz son $\lambda_1 = \lambda_2 = 1,\ \lambda_3 = \lambda_4 = 0$ por lo tanto la matriz $\mathbf{P}$ se representa por

$$\mathbf{P} = B_1 \ + \ B_2$$

donde

$$B_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \qquad B_2 = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

de aquí que la matriz límite es

$$H = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

# Capítulo 2

# Matrices positivas

> En el presente capítulo se desarrolla una serie de resultados sobre matrices cuadradas positivas, sobre todo las importantes *Matrices Primitivas*, sobre las cuales presentamos el teorema de *Perron-Frobenius*, se define a su vez las llamadas *Matrices Irreducibles* sobre las cuales se tiene una versión del teorema de *Perron-frobenius.* Para realizar el estudio se dará una relación de equivalencia entre los elementos de un conjunto $I$ de enteros llamados los índices de la matriz y mediante esta relación daremos la forma canónica de la matriz en términos de submatrices irreducibles y otras submatrices no irreducibles.

Consideremos una matriz $\mathbf{T} = (t_{ij})$ de orden $n \times n$ se dice no negativa si

$$t_{ij} \geq 0 \quad \forall i, j \in \{1, \cdots, n\}$$

Dadas la matrices no negativas $\mathbf{A} = (a_{ij})$ $y$ $\mathbf{B} = (b_{ij})$ decimos que

$$\mathbf{A} \leq \mathbf{B} \Longleftrightarrow \ a_{ij} \leq b_{ij} \quad \forall i, j \in \{1, \cdots, n\}$$

por $\mathbf{x}' = (x_1, \cdots, x_n)$ representamos un vector fila y por $\mathbf{x}$ el vector columna. $\mathbf{T}^k = (t_{ij}^{(k)})$ es la k-ésima potencia de $\mathbf{T}$

**Definición 2.0.3.** *La matriz A se dice que es una* Matriz Primitiva *si*

$$\mathbf{A}^k > 0 \quad \textit{para algún entero } k \geq 1$$

El hecho de que una matriz sea no negativa no implica que sea primitiva.

La matriz no negativa $\mathbf{A} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ es ejemplo de que no es primitiva ya que

$$\mathbf{A}^k = \mathbf{A} \quad k \geq 1$$

Para cada matriz no negativa $\mathbf{T}$ le asociamos una matíz $\tilde{\mathbf{T}}$ que consta de unos en los lugares donde los $t_{ij} > 0$ esta matriz es llamada la *matriz incindente* asociada a $\mathbf{T}$. $\mathbf{T}$ es primitiva si y solo si $\tilde{\mathbf{T}}$ lo es.

## 2.1. Teorema de Perron-Frobenius para Matrices Primitivas

**Teorema 2.1.1.** *Consideremos* $\mathbf{T} = (t_{ij})$ *matriz positiva primitiva de orden* $n$*.* $\mathbf{T}$ *tiene un valor propio* $r$ *que satisface:*

*a)* *$r$ es real positivo y es raíz simple de su polinomio característico*

*b)* *$r$ tiene asociados dos vectores propios estrictamente positivos por la derecha y por la izquierda esto es.*

$$\exists\ \mathbf{w}' > 0 \ y\ \mathbf{v} > 0 \ tal\ que\ \mathbf{w}'\mathbf{T} = r\mathbf{w}' \ y\ \mathbf{T}\mathbf{v} = r\mathbf{v}$$

*c)* *si $\lambda$ es un valor propio $\lambda \neq r$ entonces $\mid \lambda \mid < r$*

*d)* *Si $\mathbf{O} \leq \mathbf{A} \leq \mathbf{T}$ y $\alpha$ es un valor propio de $\mathbf{A}$ entonces $\mid \alpha \mid \leq r$. Si $\mid \alpha \mid = r$ entonces $\mathbf{A} = \mathbf{T}$*

*Demostración.* consideremos un vector fila $\mathbf{x}' \neq \mathbf{0}$, $\mathbf{x}' \geq \mathbf{0}$ tenemos

$$\mathbf{x}'\mathbf{T} = \left( \sum_i x_i t_{ij} \right)_{j=1\cdots n}$$

se define

$$r(\mathbf{x}) = \begin{cases} \text{mín}_j \dfrac{\sum\limits_i x_i t_{ij}}{x_j} & x_j \neq 0 \\ \\ \infty & x_j = 0 \end{cases}$$

para cada $j = 1 \cdots n$ se tiene:

$$r(\mathbf{x})x_j \leq \sum_i x_i t_{ij}$$

en forma matricial

$$\mathbf{x}' r(\mathbf{x}) \leq \mathbf{x}'\mathbf{T}$$

si $\mathbf{1}$ es la matriz columna de 1's tenemos

$$\mathbf{x}'\mathbf{1} r(\mathbf{x}) \leq \mathbf{x}'\mathbf{T}\mathbf{1}$$

Si $K = \text{máx}_i \sum\limits_j t_{ij}$ entonces se tiene $\mathbf{T1} \leq K\mathbf{1}$ se sigue que

$$r(\mathbf{x}) \leq \frac{\mathbf{x}'\mathbf{1}K}{\mathbf{x}'\mathbf{1}} = K$$

así $r(\mathbf{x})$ está uniformemente acotada para toda $\mathbf{x}$
ya que la matriz $\mathbf{T}$ es primitiva no debe de tener una columna de ceros y por lo tanto $r(\mathbf{1}) > 0$ por lo que se sigue que:

$$r = \sup_{\substack{\mathbf{x} \geq \mathbf{0} \\ \mathbf{x} \neq \mathbf{0}}} \min_j \frac{\sum_i x_i t_{ij}}{x_j}$$

satisface

$$0 < r(\mathbf{1}) \leq r \leq K < \infty$$

más aún, la igualdad no se altera si normalizamos el vector $\mathbf{x}$ esto es

$$r = \sup_{\substack{\mathbf{x} \geq \mathbf{0} \\ \mathbf{x}'\mathbf{x} = \mathbf{1}}} \min_j \frac{\sum_i x_i t_{ij}}{x_j}$$

la región $\{\mathbf{x} \mid \mathbf{x}'\mathbf{x} = \mathbf{1}\}$ es un compacto y la función $\mathbf{r}(\mathbf{x})$ es semi-continua superior en este conjunto, por lo tanto el supremo se alcanza en un $\mathbf{x}$ digamos que sea $\hat{\mathbf{x}}$

$$r = \min_j \frac{\displaystyle\sum_i \hat{x}_i t_{ij}}{\hat{x}_j}$$

se tiene

$$\sum_i \hat{x}_i t_{ij} \geq r\hat{x}_j \quad o \quad \hat{\mathbf{x}}'\mathbf{T} \geq r\hat{\mathbf{x}}'$$

para cada $j = 1, \cdots, n$;con igualdad para algún elementode $\hat{\mathbf{x}}$
consideremos el vector

$$\mathbf{z}' = \hat{\mathbf{x}}'\mathbf{T} - r\hat{\mathbf{x}}' \geq \mathbf{0}'$$

Mostremos que $\mathbf{z} = \mathbf{0}$ tenemos que $\mathbf{T}^k > \mathbf{0} \quad k > k_0$ por ser una matriz primitiva ahora bien si $\mathbf{z} > \mathbf{0}$ entonces

$$\mathbf{z}'\mathbf{T}^k = \hat{\mathbf{x}}'\mathbf{T}^{k+1} - r\hat{\mathbf{x}}'\mathbf{T}^k > \mathbf{0}'$$

esto implica que

$$\frac{(\hat{\mathbf{x}}'\mathbf{T}^k\mathbf{T})_j}{(\hat{\mathbf{x}}'\mathbf{T}^k)_j} > r \quad \forall j = 1, \cdots, n$$

esto es una contradicción a la definición de $r$, por lo tanto **z=0** y por lo tanto

$$\boxed{\hat{\mathbf{x}}'\mathbf{T} = r\hat{\mathbf{x}}'}$$

lo cual prueba que $r$ es un valor propio positivo, llamado el valor propio de *Perron* y $\mathbf{x}'$ es su vector propio positivo
Consideremos un valor propio $\lambda$ de la matriz $\mathbf{T}$ entonces existe un vector $\mathbf{x} \neq \mathbf{0}$ tal que

$$\mathbf{x}'\mathbf{T} = \lambda\mathbf{x}' \quad y \quad \mathbf{x}'\mathbf{T^k} = \lambda^k\mathbf{x}'$$

en términos de sus coordenadas

$$\sum_i x_i t_{ij} = \lambda x_j \qquad \sum_i x_i t_{ij}^{(k)} = \lambda^k x_j \tag{2.1}$$

ahora bien

$$\mid \lambda x_j \mid = \mid \sum_i x_i t_{ij} \mid \leq \sum_i \mid x_i \mid t_{ij}$$

de aqui que

$$\mid \lambda \mid \leq \frac{\sum_i x_i t_{ij}}{x_j} \quad x_j \neq 0 \quad \infty \quad si\ x_j = 0$$

y por lo tanto de la definición de $r$ se tiene

$$\mid \lambda \mid \leq r$$

Si $\mid \lambda \mid = r$ entonces

$$r \mid x_j \mid = \mid \lambda \mid\mid x_j \mid \leq \sum_i \mid x_i \mid t_{ij}$$

en una situación similar se tiene la siguientes identidades

$$\sum_i \mid x_i \mid t_{ij} = r \mid x_j \mid, \ > 0; \quad j = 1, 2, \cdots, n$$

y por lo tanto

$$\sum_i \mid x_i \mid t_{ij}^{(k)} = r^k \mid x_j \mid, \; > 0; \quad j = 1, 2, \cdots, n$$

es decir

$$\mid \sum_i x_i t_{ij} \mid = \mid \lambda^k x_j \mid \tag{2.2}$$

donde $k$ puede ser seleccionado tan grande tal que $T^k > 0$ ; escribiendo $x_j = |x_j| \exp i\theta_j$ por la ecuación (1.2) implica que $\theta_j = \theta$ es independiente de $j$ cancelando la exponencial y de (1.1) se tiene $\lambda = r$

La obtención del vector propio por la derecha es analogo a la demostración anterior.

Consideremos $\mathbf{Y} \neq \mathbf{0}$ un vector propio por la derecha de $\mathbf{B}$ con valor propio $\beta$ si $\mathbf{Y}_+ \doteq (|y_i|)$ entonces

$$|\beta|\mathbf{Y}_+ \leq \mathbf{B}\mathbf{Y}_+ \leq \mathbf{T}\mathbf{Y}_+$$

se tiene

$$|\beta|\hat{\mathbf{X}}'\mathbf{Y}_+ \leq \hat{\mathbf{X}}'\mathbf{T}\mathbf{Y}_+ = r\hat{\mathbf{X}}'\mathbf{Y}_+$$

y dado que $\hat{\mathbf{X}}'\mathbf{Y}_+ > 0$ entonces

$$|\beta| \leq r$$

Si se cumple $|\beta| = r$ entonces

$$r\mathbf{Y}_+ \leq \mathbf{T}\mathbf{Y}_+$$

de donde se sigue que $\mathbf{T}\mathbf{Y}_+ = r\mathbf{Y}_+ > \mathbf{0}$ y así

$$r\mathbf{Y}_+ = \mathbf{B}\mathbf{Y}_+ = \mathbf{T}\mathbf{Y}_+$$

por lo tanto se sigue que $\mathbf{B} = \mathbf{T}$

Tenemos las siguientes identidades

$$(x\mathbf{I}-\mathbf{T})adj(x\mathbf{I}-\mathbf{T}) = det(x\mathbf{I}-\mathbf{T})\mathbf{I} \tag{2.3}$$

$$adj(x\mathbf{I}-\mathbf{T})(x\mathbf{I}-\mathbf{T}) = det(x\mathbf{I}-\mathbf{T})\mathbf{I}$$

si ponemos $x = r$ se tiene que cualquier fila de $adj(x\mathbf{I}-\mathbf{T})$ cumple alguna de las siguientes afirmaciones

a) es un vector propio por la derecha correspondiente a $r$

b) es una fila de ceros

lo mismo para las columnas
también se cumple alguna de las siguentes afirmaciones sobre $adj(x\mathbf{I}-\mathbf{T})$

a) no tiene elementos ceros

b) es la matriz nula

Se probara que que se tiene un elemento diferente de cero y por lo tanto todos lo serán.
sea $\mathbf{T}_{n-1}$ la matriz de orden $n-1$ que se obtiene al omitir la fila y columna $n$ de $\mathbf{T}$ lo mismo $\mathbf{I}_{n-1}$.
Se tiene

$$0 \leq \begin{pmatrix} \mathbf{T}_{n-1} & \mathbf{0} \\ \mathbf{0}' & 0 \end{pmatrix} \leq \mathbf{T} \quad \neq \mathbf{T}$$

y dado que $\mathbf{T}$ es primitiva no tiene columnas cero por lo tanto se tiene

$$det(r\mathbf{I}_{n-1}-\mathbf{T}_{n-1}) > 0$$

por lo tanto se deduce que $adj(x\mathbf{I}-\mathbf{T})$ tiene todos sus elementos positivos.
Se define $\varphi(x) \doteq (x\mathbf{I}-\mathbf{T})$ diferenciando (1.3)

$$adj(x\mathbf{I}-\mathbf{T})+(x\mathbf{I}-\mathbf{T})\frac{d}{dx}adj(x\mathbf{I}-\mathbf{T})=\varphi'(x)\mathbf{I}$$

sustituyendo $x = r$ y multipicando por $\hat{\mathbf{X}}'$ se obtiene

$$(\mathbf{0}>)\quad \hat{\mathbf{X}}'adj(r\mathbf{I}-\mathbf{T})=\varphi'(r)\hat{\mathbf{X}}'$$

por lo tanto $\varphi'(r) > 0$ y así $r$ es simple

□

**Proposición 2.1.1.** *se tiene la siguiente afirmación*

$$\boxed{\text{mín}_i \sum_j t_{ij} \leq r \leq \max_i \sum_j t_{ij}}$$

*Demostración.* Tenemos de la definición de $r(\mathbf{X})$

$$0 < r(\mathbf{1}) = \min_{\mathbf{j}} \sum_{\mathbf{i}} \mathbf{t_{ij}} \leq \mathbf{r} \leq \mathbf{K} = \max_{\mathbf{i}} \sum_{\mathbf{j}} \mathbf{t_{ij}} < \infty$$

como la tanspuesta de $\mathbf{T}$ es primitiva y tiene el mismo r entonces

$$\min_j \sum_i t_{ji} \leq r \leq \max_i \sum_j t_{ji}$$

combinando estas dos desigualdades se tiene el resultado.
Observación: si una de las desigualdades es una igualdad entonces la otra también lo sera.

□

**Corolario 2.1.1.** *Sean* $\mathbf{v}'$ *y* $\mathbf{w}$ *los vectores propios por la derecha e izquierda correspondientes al valor de perron* $r$ *los cuales estan normalizados* $(\mathbf{v}'\mathbf{w} = 1)$ *entonces*

$$adj(r\mathbf{I}-\mathbf{T})=\varphi'(r)\mathbf{w}\mathbf{v}'$$

Consideremos los valores propios de $\mathbf{T}$

$$r > |\lambda_2| \geq |\lambda_3| \geq \cdots$$

si $m_2$ es la multiplicidad de $\lambda_2$ se tiene la siguiente proposición

**Proposición 2.1.2.** *Para una matriz primitiva* $\mathbf{T}$ *se tienen las siguientes afirmaciones:*

- *Si* $\lambda_2 \neq 0$ *entonces para* $k \gg 0$ *se tiene:*

$$\mathbf{T}^k = r^k \mathbf{w}\mathbf{v}' + O(k^{m_2-1}|\lambda_2|^k)$$

- *Si* $\lambda_2 = 0$ *entonces para* $k \geq n-1$

$$\mathbf{T}^k = r^k \mathbf{w}\mathbf{v}'$$

## 2.2. Estructura de las Matrices Positivas

*en esta sección se presenta un forma de estudiar la estructura de una matriz positiva en terminos de una relación de equivalencia entre entre los indices de sus terminos y a su vez definimos los bloques de matrices irreducibles que la conforman.*

*Consideremos la matriz no negativa* $\mathbf{T} = (t_{ij})$ *de orden* $n$ *donde* $I = \{1, 2, \cdots, n\}$ *es llamado el conjunto de indices.*

**Definición 2.2.1.** *Decimos que el indice* $i$ *se comunica con el indice* $j$ *si exite una sucesión finita de elenentos de* $\mathbf{T}$

$$t_{ii_1}, t_{i_1i_2}, t_{i_2i_3}, \cdots, t_{i_{n-1}j}$$

*tal que el producto*

$$t_{ii_1}t_{i_1i_2}t_{i_2i_3}\cdots t_{i_{n-1}j} > 0$$

*pedimos que la longitud* $n$ *de esta sucesión se mínima*

*también se representa por una* $n - ada$

$$(i, i_1, i_2, \cdots, i_{n-1}, j)$$

*o por el díagrama*

$$i \curvearrowright i_1 \curvearrowright i_2 \curvearrowright \cdots \curvearrowright i_{n-1} \curvearrowright j$$

*cuando* $i = j$ *decimos que tenemos un ciclo*

*esto permite dar una clasificación de indices.*

# Clasificación de indices

**Definición 2.2.2.** *Decimos que el indice* $i$ *se dirige al indice* $j$ *si*

$$t_{ij}^{(k)} > 0 \qquad k \geq 1$$

*lo simbolizamos por* $i \to j$

*Decimos que están comunicados si* $i \to j$ *y* $j \to i$ *y lo simbolizamos por*

$$i \longleftrightarrow j$$

*El conjunto de indices se pueden clasificar y agrupar de la siguiente forma.*

- *Si* $i \to j$ *pero* $j \nrightarrow i$ *para algún indice* $j$ *entonces a* $i$ *le llamamos un estado* no escencial *en caso contrario le llamamos* escencial

- *Si $i$ es escencial y $i \to j$ entonces $i \leftrightarrow j$*
- *La relación de comunicación $\leftrightarrow$ da una partición del conjunto de indices en* clases escenciales *y* clases no escenciales

*Consideremos el siguiente ejemplo*

Ejemplo *La matriz de incidencia de* $\mathbf{T}$ *es*

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

*se observa que:*

$$(1 \to 2) \wedge (2 \to 1) \Rightarrow 1 \leftrightarrow 2$$

## 2.3. Forma Canónica

**Definición 2.3.1.** *La matriz se dice* Irreducible *si todos sus indices están comunicados entre ellos, es decir solo hay una clase de equivalencia.*

*En base a la descomposición en clases de equivalencia del conjunto de indices podemos restructurar la matriz* $\mathbf{T}$ *de tal manera que cada clase* *de indices escenciales genera una submatriz irreducible, esto permite dar una nueva estructura de la matriz, a saber*

$$\mathbf{T} = \left(\begin{array}{ccccc|c} \mathbf{T_1} & \mathbf{0} & \cdots & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{T_2} & \cdots & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \mathbf{0} & \ddots & \mathbf{0} & \cdots & \vdots \\ \mathbf{0} & & & \ddots & \mathbf{T_r} & \mathbf{0} \\ \hline & & \mathbf{R} & & & \mathbf{Q} \end{array}\right)$$

*donde* $\mathbf{T_i}$ *es la submatriz asociada a la clase* $C_i$ *de indices escenciales y la matriz* $\mathbf{Q}$ *es*

$$\mathbf{Q} = \begin{pmatrix} \mathbf{Q_1} & & & & \\ & \mathbf{Q_2} & \cdots & \mathbf{0} & \\ & \mathbf{S} & \ddots & & \\ & & & & \mathbf{Q_r} \end{pmatrix}$$

*donde* $\mathbf{Q_j}$ *es la submatriz asociada a la clase* $j$ *de indices no escenciales.*

*La forma canonica de la matriz* $\tilde{\mathbf{T}}$ *en el ejemplo*

*El conjunto de indices es* $I = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ *las clases escenciales son:*

$$C_1 = \{5\},\ C_2 = \{4, 9\},\ C_3 = \{3, 7\}$$

*Las clases no escenciales son:*

$$C_4 = \{1, 2\},\ C_5 = \{6\},\ C_6 = \{8\}$$

*por lo tanto la forma canónica se puede escribir como:*

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \end{pmatrix}$$

**Definición 2.3.2.** *Una matriz formada por una sola clase es llamada una* matriz irreducible

*En nuestro ejemplo tenemos tres submatrices irreducibles que corresponden a las clases escenciales.*

## Periodo de un Índice

**Definición 2.3.3.** *Dado el índice $i$ se define su periodo como el máximo común divisor de los $k \geq 1$ tal que $t_{ii}^{(k)} > 0$ y lo denotamos por $d(i)$*

$$d(i) = mcd\{\ k \ \mid \ t_{ii}^{(k)} > 0\ \}$$

*tenemos el siguiente resultado*

**Proposición 2.3.1.** *Si $i \leftrightarrow j$ entonces $d(i) = d(j)$*

*El periodo de una clase se define como el periodo de un elemento de la clase.*

*Considerando nuestro ejemplo se tiene:*

Clases Escenciales:

$$\{5\} \quad \textit{tiene periodo 1, dado que} \quad t_{55} > 0$$

$$\{4,9\} \quad \textit{tiene periodo 1, dado que} \quad t_{44} > 0$$

$$\{3,7\} \quad \textit{tiene periodo 2, dado que} \quad t_{33}^{(k)} > 0$$

*para k par, y es cero para k impar.*

Clases No Escenciales:

$$\{1,2\} \quad \textit{tiene periodo 1, dado que} \quad t_{22} > 0$$

$$\{6\} \quad \textit{tiene periodo 1, dado que} \quad t_{66} > 0$$

$$\{8\} \quad \textit{tiene periodo 1, dado que} \quad t_{88} > 0$$

*Si* $d(i) = 1$ *se dice que es aperiodico (aciclico)*

*El siguiente teorema nos caracteriza las matrices primitivas en términos de su periodo.*

**Teorema 2.3.1.** *La matriz* **T** *es primitiva si y solo si es irreducible aperiodica.*

*En el siguiente y último capítulo aplicamos estos resultados a las* Matrices Estocasticas.

# Capítulo 3

# Matrices Estocasticas

*Las matrices estocasticas que caracterizán a los procesos estocasticos conocidos como cadenas de markov homogéneas seran de estudio usando los resultados obtenidos en el capítulo 2 para lo cual se hara una analogía entre el conjunto de índices y el espacio de estados de la cadena con el fin de obtener resultados de la distribución estacionaria.*

## 3.1. Matriz Estocastica

*Una Cadena de Markov homogenea con espacio de estados*

$$\mathcal{S} = \{s_1, s_2, \cdots, s_n\}$$

*tiene asociada una matriz cuadrada* $\mathbf{P} = (p_{ij})$ *no negativa de orden el número de estados donde:*

$$p_{ij} \doteq \mathcal{P}(X_{n+1} = s_j \mid X_n = s_i)$$

*y que satisface*

$$\sum_j p_{ij} = 1 \quad \forall i \tag{3.1}$$

*Estas matrices no negativas se les conocen como* Matrices Estocasticas o Markovianas

*Para su estudio haremos la siguiente analogía; Dando un orden a los estados de la cadena establecemos:*

$$s_i \leftrightarrow i$$

*La relación entre indices es equivalente a la relación entre los estados establecida en el apéndice del trabajo. Esto determina la identificación siguiente:*

$$\begin{array}{ccc} Estado & & \acute{I}ndice \\ s_i & \leftrightarrow & j \\ recurrente & \leftrightarrow & escencial \\ transitivo & \leftrightarrow & no\ escencial \end{array}$$

*En base a esta identificación se tiene*

$$\begin{array}{rcl} Clase\ de\ Recurrencia & \leftrightarrow & Clase\ Escencial \\ Clase\ Transitiva & \leftrightarrow & Clase\ no\ Escencial \end{array}$$

## 3.2. Cadena Markov Irreducible

*La Cadena de Markov se dice Irreducible si su matriz de transición* **P** *lo es. De la ecuación (3.1) se tiene*

$$\mathbf{P1} = \mathbf{1}$$

*de la proposición (2.1.1) se tiene*

$$1 = \min_i \sum_j P_{ij} \leq r \leq \max_i \sum_j P_{ij} = 1$$

*esto implica que el valor propio de Perron-Frobenius $r$ de una matriz estocastica es*

$$r = 1 \quad \textit{con vector propio por la derecha } \mathbf{1}$$

*Por el teorema de PERRON FROBENIUS se tiene un único vector propio* $\mathbf{v}'$ *tal que*

$$\mathbf{v}'\mathbf{P} = \mathbf{v}'$$

*Lo anterior se resúme en el siguiente teorema*

**Teorema 3.2.1.** *Una cadena de markov irreducible tiene una única distribución estacionaria dada por la solución*

$$\mathbf{v}'\mathbf{P} = \mathbf{v}' \quad \mathbf{v}'\mathbf{1} = 1$$

*De la proposición (2.1.2) se tiene el siguiente teorema*

**Teorema 3.2.2.** *Teorema Ergódico para cadenas de Markov primitivas*

$$\lim_{k \to \infty} \mathbf{P}^k = \mathbf{1}\mathbf{v}'$$

*esto es, existe la matriz de equilibrio* **H** *cuyas filas son dadas por la distribución* **v'**

*Presentamos los siguiente ejemplos donde mostramos diferentes técnicas de compúto para obtener la distribución estacionaria.*

*Ejemplo 1 (Caminata Aleatoria con barreras reflejantes)*

*Consideremos la cadena de markov con matriz de transición* **P**

$$\mathbf{P} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

*Esta matriz es irreducible y aperiodica por lo tanto tiene una única distribución estacionaria que es vector propio por la derecha. Si diagonalizamos se tiene:*

$$\mathbf{P} = \begin{pmatrix} \frac{1}{\sqrt{6}} & * & * & * & * & * \\ \frac{1}{\sqrt{6}} & * & * & * & * & * \\ \frac{1}{\sqrt{6}} & * & * & * & * & * \\ \frac{1}{\sqrt{6}} & * & * & * & * & * \\ \frac{1}{\sqrt{6}} & * & * & * & * & * \\ \frac{1}{\sqrt{6}} & * & * & * & * & * \end{pmatrix} Diag\{1, \cdots\} \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ * & * & * & * & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \end{pmatrix}$$

*por la teoría de PERRON FROBENIUS se tiene que la matriz límite* **H** *es:*

$$\mathbf{H} = \begin{pmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{pmatrix}$$

*y la distribución estacionaria (vector de Perron) es*

$$\pi = \left(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\right)$$

*Ejemplo 2 Consideremos la matriz estocástica*

$$\mathbf{P} = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$$

*Los valores propios de la matriz son* $\lambda_1 = \lambda_2 = 1, \ \lambda_3 = \lambda_4 = 0$ *por el teorema de Perron-Frobenius ésta no es primitiva. La matriz incidente es:*

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

*Tiene dos clases de equivalencia escenciales a saber:*

$$C_1 = \{1, 3\} \quad y \quad C_2 = \{2, 4\}$$

*por por lo tanto la nueva matriz al reordenar los índices se tienen* $I = \{1, 3\} \cup \{2, 4\}$ *y dos bloques de matrices primitiva a saber:*

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

*Aplicando el teorema de Perron-Frobenius a cada bloque se tiene la disribución estacionarias*

$$\mathbf{v_1} = (\frac{1}{2}, \frac{1}{2}) \quad \textit{para } C_1$$

$$\mathbf{v_2} = (\frac{1}{2}, \frac{1}{2}) \quad \textit{para } C_2$$

*que generan las distribuciones*

$$\pi_{\mathbf{1}} = (\frac{1}{2}, 0, \frac{1}{2}, 0)$$

$$\pi_{\mathbf{2}} = (0, \frac{1}{2}, 0, \frac{1}{2}, 0)$$

*Para concluir se tiene la siguiente afirmación: Dado que el conjunto de Distribuciones Estacionarias es un conjunto convexo se tiene que la cadena tiene una única o una infinidad de distribuciones.*

# Apéndice A

# Cadenas de Markov

**Definición A.0.1.** *Una* **Cadena de Markov** *es un* Proceso de Markov discreto *cuyo* Espacio de Estados $\mathcal{S}$ *es también discreto y el cual lo podemos representar por una sucesión de variables aleatorias* $(X_n)_{n\geq 0}$

$(X_n)_{n\geq 0}$ *una cadena de Markov, entonces tenemos que:*

$$\mathcal{P}(X_{n+1} = j \mid X_0, X_1, \cdots, X_{n-1}, X_n = i) = \mathcal{P}(X_{n+1} = j \mid X_n = i) \quad \forall n \ , \ \forall i, j \in \mathcal{S}$$

$\mathcal{P}(X_{n+1} = j \mid X_n = i)$ *es llamada la probabilidad de transición en un paso.*

*Si las probabilidades de transición en un paso son independientes de* $n$ *la Cadena de Markov se dice* Estacionaria u Homogénea *y denotamos por*

$$p_{ij} \doteq \mathcal{P}(X_{n+1} = j \mid X_n = i)$$

*que satisface:*

$$\sum_{j\in\mathcal{S}} p_{ij} = 1 \qquad \forall i \in \mathcal{S}$$

*Se determina la matriz no negativa*

$$\mathbf{P} = (p_{ij})$$

*llamada matriz de transición asociada a la C.M.*
*Estas Matrices son tambien llamadas* Estocásticas o Markovianas

**Definición A.0.2.** *Denotamos por $\pi'_n$ la Disribución de la v.a. $X_n$, la cual se representara por un vector fila*

$$\pi'_n = (\pi_n(i))_{i \in \mathcal{S}}$$

*donde*

$$\pi_n(i) = \mathcal{P}(X_n = i)$$

$\pi'_0$ *es llamada la* **Distribución inicial** *del Proceso*

## La Distribución Conjunta

*La Distribución Conjunta de las variables $X_1, X_2, \cdots X_n$ esta determinada en términos de las probabilidades de transición y la Distribución inicial a saber:*

$$\mathcal{P}(X_0 = i_0 \cdots X_n = i_n) =$$

$$\mathcal{P}(X_n = i_n \mid X_0 = i_0, \cdots, X_{n-1} = i_{n-1})\mathcal{P}(X_0 = i_0 \cdots X_{n-1} = i_{n-1})$$

*que es una relación recursiva que al resolverla nos da la siguiente identidad*

$$\mathcal{P}(X_0 = i_0, \cdots X_n = i_n) = \pi_0(i_0) p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n} \tag{A.1}$$

*La Distribución Condicional del sistema en varias situaciones futuras al conocer el pasado y el presente esta dada por:*

$$\mathcal{P}(X_{n+1} = i_{n+1}, \cdots, X_{n+m} = i_{n+m} \mid X_0, \cdots, X_n = i_n) = p_{i_n i_{n+1}} p_{i_{n+1} i_{n+2}} \cdots p_{i_{n+m-1} i_{n+m}}$$

## La Distribución de la v.a. $X_n$

*Observemos que*

$$\pi_n(j) = \sum_{i \in \mathcal{S}} p_{ij} \pi_{n-1}(i)$$

*en forma matricial*

$$\pi'_n = \pi'_{n-1} \mathbf{P}$$

*que es una relación recursiva que al resolver da la siguiente identidad*

$$\pi_n(j) = \sum_{i_0 \in \mathcal{S}} \sum_{i_1 \in \mathcal{S}} \cdots \sum_{i_{n-1} \in \mathcal{S}} \pi_0(i) p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{n-1} j}$$

*en forma maticial es:*

$$\pi'_n = \pi'_0 \mathbf{P}^n$$

## Probabilidades de transición en n-pasos

*Se define la probabilidad de transición del sistema después de n evoluciones como:*

$$p_{ij}^{(n)} \doteq \mathcal{P}(X_n = j \mid X_0 = i)$$

*se obtiene la siguiente identidad*

$$p_{ij}^{(n)} = \sum_{i_1 \in \mathcal{S}} \sum_{i_2 \in \mathcal{S}} \cdots \sum_{i_{n-1} \in \mathcal{S}} p_{ii_1} p_{i_1 i_2} \cdots p_{i_{n-1} j}$$

*esta identidad nos dice que*

$$\mathbf{P}^n = \left(p_{ij}^{(n)}\right)$$

## La Ecuación de Chapmann-Kolmogorov

*La siguiente identidad es conocida como la ecuación de Chapmann-Kolmogorov que es de gran importancia en los procesos markovianos*

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

*en forma matricial*

$$\mathbf{P}^{m+n} = \mathbf{P}^m \mathbf{P}^n$$

**Definición A.0.3.** **Distribución Estacionaria**

*Una Distribución $\pi$ sobre el Espacio de Estados $\mathcal{S}$ de una C.M. con matriz asociada $\mathbf{P}$ es llamada una* **Distribución Estacionaria** *si satisface:*

$$\pi' \mathbf{P} = \pi'$$

**Observación A.0.1.** *Si la Distribución inicial $\pi_{\mathbf{0}}$ de una C.M. es Estacionaria entonces se tiene que*

$$\pi'_{\mathbf{n}} = \pi'_{\mathbf{0}} \qquad \forall\ \mathbf{n}$$

## A.1. Clasificación de Estados

**Tiempo de Éxito** $(X_n)_{n\geq 0}$ *una* Cadena de Markov *con Espacio de Estados* $\mathcal{S}$

**Definición A.1.1.** *Se define el* Primer Tiempo de Éxito *del estado* $i \in \mathcal{S}$ *como la variable*

$$T_i \doteq inf\{n > 0 \mid X_n = i\}$$

*para cada* $n > 1,\ i, j \in \mathcal{S}$ *se definen las siguientes probabilidades*

$$f_{ij}^{(n)} \doteq \mathcal{P}_i(X_1 \neq j, \cdots, X_{n-1} \neq j, X_n = j) = \mathcal{P}_i(T_j = n)$$

$$f_{ij} \doteq \sum_{n=1}^{\infty} f_{ij}^{(n)} = \mathcal{P}_i(T_j < \infty)$$

$f_{ij}^{(n)}$ *representa la probabilidad del primer éxito del estado* $j$ *en el tiempo* $n$ *cuando la C.M. inicia en el estado* $i$

$f_{ij}$ *es la probabilidad de tener éxito en el estado* $j$ *cuando la C.M. inicia el el estado* $i$

*Las probabilidades* $p^{(n)}(i, j)$ *se pueden dar en términos de los tiempos de éxito*

**Proposición A.1.1.** *Dados los estados* $i, j \in \mathcal{S} \quad n \geq 1$ *se tiene la siguiente identidad*

$$p^{(n)}(i, j) = \sum_{k=1}^{n} f_{ij}^{(k)} p^{(n-k)}(j, j)$$

**Definición A.1.2.** *Dado el estado* $j \in \mathcal{S}$ *se clasifica como:*

- **Estado Transitivo** *si:*

$$f_{jj} < 1$$

- **Estado Recurrente** *si:*

$$f_{jj} = 1$$

## A.2. Distribución Estacionaria

*Bajo condiciones generales existe una única distribución estacionaria para una cadena de Markov la cual está dada en términos del Tiempo Esperado de Recurrencia $\mu(i)$, que determina el promedio de tiempo esperado del proceso que inicia en el estado i, para retornar. En éste cápitulo se presentan resultados para determinar condiciones para la existencia y unicidad de la distribución estacionaria.*

**Tiempo Esperado de Recurrencia**

**Definición A.2.1.** *Se define el* **Tiempo Esperado de Recurrencia** *$\mu(i)$ del estado $i \in \mathcal{S}$ como*

$$\mu(i) = \begin{cases} \infty & si\ i\ es\ Transitivo \\ \displaystyle\sum_{k=i}^{\infty} k\mathcal{P}(T_i = k \mid X_0 = i) & si\ i\ es\ Recurrente \end{cases}$$

*En base a esta definición podemos clasificar a los estados recurrentes por*

- *El estado i es llamado* **Recurrente Positivo** *si*

$$\mu(i) < \infty$$

- *El estado $i$ es llamado* **Recurrente Nulo** *si*

$$\mu(i) = \infty$$

*En base a estas clasificación de los estados podemos dar condiciones para la existencia y unicidad de la distribución estacionaria,analizamos primeramente cuando el espacio Ses finito.*

*Consideremos una Cadena de Markov irreducible finita con Espacio de Estados $\mathcal{S} = \{x_1, x_2, \cdots, x_n\}$ entonces los estados son* recurrentes.*Sea $\mu(i)$ el tiempo Esperado de Recurrencia entonces tenemos una única Distribución Estacionaria $\pi$ dada a saber por:*

$$\pi(i) = \frac{1}{\mu(i)}$$

**Tiempos de Espera**

> *En ésta sección se dan elementos que nos permitan establecer condiciones para la existencia y unicidad de la distribución estacionaria en términos de sus estados y de la propiedad de periodicidad*

**Definición A.2.2.** *La transitividad o recurrencia de un estado se puede caracterizar en términos de la* **Matriz de Potencial G** *definida por*

$$\mathbf{G} = \sum_{k=1}^{\infty} \mathbf{P}^k$$

*cuyas entradas son los valores esperados de visitas*

$$\mathcal{V}(i,j) = \sum_{k=1}^{\infty} p^k(i,j)$$

**Ergodicidad**

*Un proceso es llamado* ergódico *si es posible encontrar alguna propiedad estadística del proceso se puede ver con tan solo una trayectoria muestral.Entre los teoremas ergódicos básicos más importantes esta la* **Ley Fuerte de los Grandes Números** *la cual afirma:*

*Si* $X_1, X_2, \cdots, X_n$ *son variables aleatorias con idéntica distribución, entonces existe* $\mu$ *tal que es casi seguro*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} X_k = \mu$$

*si y solo si* $\mathcal{E}(| X_1 |) < \infty$ *y* $\mu = \mathcal{E}(X_1)$

*Decimos que la Cadena de Markov* $(X_n)_{n\geq 0}$ *satisface el* **Teorema Ergódico Medio** *si tiene una única distribución estacionaria* $\pi$ *tal con probabilidad uno se tiene*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} I_j(X_k) = \pi(\mathbf{j})$$

*es decir, la frecuencia de visitas de la cadena se aproxima a la distribución estacionaria.*

**Tiempos del Exito y Tiempos de Espera**

**Definición A.2.3.** *Se define el* Tiempo del r Exito *al estado j como la variable aleatoria*

$$T_j(r) = min\{n \mid N_n(j) = r\}$$

*se tiene* $T_j(1) = T_j$

**Definición A.2.4.** *Se define la variable aleatoria* $W_j(r)$ *como el tiempo de espera entre el r-1 éxito y el r éxito cuando* $r \geq 2$ *y* $W_1$ *como el tiempo de espera para el primer éxito esto es:*

$$W_j(1) = T_j \qquad W_j(r) \doteq T_j(r) - T_j(r-1)$$

*las variables* $W_j(1), W_j(2), \cdots$ *son independientes y con idéntica distribución y con* $\mathcal{E}_j(T_j) = \mu(j)$ *entoces por la Ley Fuerte de los Grandes Números se tiene*

$$\lim_{n\to\infty} \frac{W_j(1) + \cdots + W_j(n)}{n} = \mu(j)$$

*además*

$$W_j(1) + \cdots + W_j(n) = T_j(n)$$

*por lo tanto*

$$\lim_{n\to\infty} \frac{T_j(n)}{n} = \mu(j)$$

*Dado* $N_j(n)$ *es el número de visitas en el tiempo* $n$ *entonces se tiene*

$$T_j(N_j(n)) \leq n < T_j(N_j(n) + 1)$$

*asi que*

$$\frac{T_j(N_j(n))}{N_j(n)} < \frac{n}{N_j(n)} < \frac{T_j(N_j(n) + 1)}{N_j(n)}$$

*por recurrencia cuando* $n \to \infty$ *sabemos que es casi seguro que* $N_j(n) \to \infty$ *y por lo tanto*

$$\lim_{n\to\infty} \frac{N_j(n)}{n} = \frac{1}{\mu(j)} \tag{A.2}$$

*Esto dice que una cadena irreducible recurrente positiva satisface el Teorema Medio Ergódico. Tomando el valor esperado se tiene*

$$\lim_{n\to\infty} \frac{\mathcal{V}_n(i,j)}{n} = \frac{1}{\mu(j)} \tag{A.3}$$

# Conclusiones

*Las Cadenas de Markov son procesos de mucha importancia por sus aplicaciones en diferentes áreas de la Ciencias (Exactas,Sociales,Económicas,..) y la Ingeniería ya que permiten tener información previa sobre las probabilidades de que ocurran o se repitan ciertos eventos de fenómeno, eso da elementos para realizar toma de decisiones. El trabajo de tesis da elementos y formas de como obtener la distribuci—'on estacionaria y si ésta es única con base en el importante* Teorema de Perron-Frobenius *dado que en base al conocimento de la distibución obtenemos información sobre los tiempos promedios de recurrencia de los eventos.*

# Bibliografía

[1] *Bellman,R.,* Introduction to Matrix Analysis, *McGraw-Hill, New York, Third Edition, 1960.*

[2] *Seneta,E.,* Non Negative Matrices and Markov Chains, *Second Edition, Editorial Springer Verlagé, S. A., 1973.*

[3] *CHung K.L. ,* Markov Chains with stationary transition probabilities, *Second Edition, Editorial Springer é, S. A., 1967*

.

[4] *Kemeny J.G.,Snell L.J.,* Finite Markov Chains, *Editorial Van Nostrand ,Pricenton,N.J.*