



**VNiVERSiDAD
D SALAMANCA**

CAMPUS DE EXCELENCIA INTERNACIONAL

**FACULTAD DE CIENCIAS
MÁSTER DE MODELIZACIÓN MATEMÁTICA**

**ESTUDIO DE CADENAS DE
MARKOV INFINITAS MEDIANTE
INVERSAS GENERALIZADAS DE
ENDOMORFISMOS**

AUTOR

Pablo Rodríguez García

TUTOR

Fernando Pablos Romo

Curso 2023-2024



**VNiVERSiDAD
D SALAMANCA**

CAMPUS DE EXCELENCIA INTERNACIONAL

**FACULTAD DE CIENCIAS
MÁSTER DE MODELIZACIÓN MATEMÁTICA**

**ESTUDIO DE CADENAS DE
MARKOV INFINITAS MEDIANTE
INVERSAS GENERALIZADAS DE
ENDOMORFISMOS**

AUTOR

Pablo Rodríguez García

TUTOR

Fernando Pablos Romo

Curso 2023-2024

MÁSTER UNIVERSITARIO EN MODELIZACIÓN MATEMÁTICA

D. Fernando Pablos Romo, profesor del Departamento de Matemáticas de la Universidad de Salamanca

HACE CONSTAR

Que el trabajo titulado “Estudio de cadenas de Markov infinitas mediante inversas generalizadas de endomorfismos”, que se presenta, ha sido realizado por Pablo Rodríguez García, con DNI número ***33543V y constituye la memoria del trabajo realizado para la superación de la asignatura Trabajo Fin de Máster del Máster Universitario en Modelización Matemática de la Universidad de Salamanca

Salamanca, a fecha de la firma electrónica

Firmado.- Fernando Pablos Romo

Índice

Introducción	1
1. Preliminares	2
1.1. Cadenas de Markov	2
1.2. Inversa de Drazin y aplicaciones a las cadenas de Markov . . .	6
1.3. Endomorfismos core-nilpotentes	16
2. Búsqueda de un ejemplo ilustrativo	19
3. Generalización de los resultados	22
4. Aplicación de los resultados al ejemplo	28
4.1. Verificación de las hipótesis	29
4.2. Cálculo de la inversa de Drazin	31
4.3. Cálculo de los resultados con la inversa de Drazin	32
5. Simulaciones con Mathematica	33
Referencias	36
Anexo	36
Artículo científico	36

Introducción

Es bien sabido que un endomorfismo puede no ser invertible, lo que significa que no tiene una inversa. Por lo tanto, puede ser muy útil buscar alternativas: dado un endomorfismo, encontrar otro que, aunque no sea invertible, desempeñe el papel de la inversa o, al menos, tenga características similares. Por esta razón, se crearon las pseudo-inversas: endomorfismos que tienen similitudes con las inversas. Una de las pseudo-inversas más estudiadas es la de Drazin, llamada así en honor a Michael Peter Drazin por su trabajo de 1958, [1].

La inversa de Drazin es especialmente útil en el análisis de sistemas dinámicos, donde los operadores suelen estar constituidos por reglas de transición. En particular, las cadenas de Markov modelan sistemas dinámicos estocásticos en lugar de deterministas. En este contexto, existen operadores lineales que proporcionan información sobre la distribución de probabilidad en cada instante de tiempo.

No es descabellado pensar que la inversa de Drazin puede ser útil para obtener cierta información sobre la cadena. En concreto, nos permite abordar sistemas de ecuaciones lineales asociados a operadores lineales que no son invertibles. Por ejemplo, en el cálculo de distribuciones estacionarias, comportamientos a largo plazo de la cadena, recurrencia, etc.

Otro aspecto importante de las cadenas de Markov es la clasificación de los estados en términos de recurrencia y transitoriedad. La inversa de Drazin facilita el estudio de estas propiedades en cadenas de Markov finitas al proporcionar otra perspectiva de la matriz de transición.

A lo largo de este trabajo exploraremos en detalle la definición formal de la inversa de Drazin y su generalización al caso de dimensión infinita, sus propiedades fundamentales y su aplicación en la teoría de cadenas de Markov.

Comenzaremos con una revisión a las cadenas de Markov, definiciones y resultados básicos. Seguiremos con la definición de inversa de Drazin y su aplicación a las cadenas de Markov finitas. Lo siguiente es mostrar el mayor ámbito en el que se puede definir la inversa de Drazin en la teoría de endomorfismos de espacios vectoriales de dimensión arbitraria: los endomorfismos core-nilpotentes. Se continuará con la teoría que permite generalizar los resultados a cadenas de Markov arbitrarias y finalizaremos con su aplicación a un ejemplo concreto.

El objetivo principal de este trabajo es comprobar que, a través de un conocimiento más profundo de la inversa de Drazin, podemos obtener información relevante de una cadena de Markov arbitraria. Además, este trabajo incluye la elaboración de un artículo, que tiene la intención de servir como motivación para futuras investigaciones sobre la inversa de Drazin y sus aplicaciones en diversas disciplinas que utilizan operadores lineales y matrices. El artículo, añadido en el anexo, ya se ha enviado a una revista y está bajo revisión por pares.

1. Preliminares

Antes de comenzar a desarrollar un trabajo científico, es necesario un proceso de documentación, en el que se investiga sobre los principales elementos que van a intervenir en el trabajo.

En este caso, lo principal son definiciones, propiedades y resultados de la teoría de cadenas de Markov, todo lo relacionado con la inversa de Drazin, y el marco teórico de los endomorfismos de espacios vectoriales de dimensión arbitraria donde se pueda definir la inversa de Drazin.

1.1. Cadenas de Markov

Para empezar, lo óptimo sería documentarse sobre las cadenas de Markov pues es el objetivo de estudio en este trabajo. Las principales fuentes para este tema han sido [2], [3], [4] y [5].

Un tema que da bastante controversia es si $0 \in \mathbb{N}$ o si $0 \notin \mathbb{N}$. En general, esto dependerá del contexto en el que se definan los objetos. Sin embargo, para formalizar la teoría se asumirá que $0 \in \mathbb{N}$, aunque se podría considerar lo contrario.

Definición 1. *Una cadena de Markov es una sucesión de variables aleatorias discretas $\{X_n\}_{n \in \mathbb{N}}$ verificando la propiedad de Markov, i.e.*

$$\mathcal{P}(X_{n+1} = e | X_n = e_n, \dots, X_0 = e_0) = \mathcal{P}(X_{n+1} = e | X_n = e_n),$$

para cualesquiera $e, e_0, \dots, e_n \in E$, donde $E = \bigcup_{n \in \mathbb{N}} \text{sop}(X_n)$ es el conjunto de posibles estados de la cadena.

La definición formaliza la idea de que la probabilidad de cambiar de estado solo depende del estado en el que se encuentre actualmente la cadena y no de cómo haya llegado hasta él. Para simplificar la notación se usará la siguiente nomenclatura para referirse a las probabilidades de transición:

$$p_{i,j}^{(n)} = \mathcal{P}(X_{n+1} = j | X_n = i).$$

Si E es un conjunto finito, se dice que la cadena de Markov es finita; si es infinito, entonces se dice que es infinita.

Definición 2. Se dice que una cadena de Markov $\{X_n\}_{n \in \mathbb{N}}$ es homogénea si las probabilidades de transición no dependen de la etapa, es decir,

$$p_{i,j}^{(n)} = p_{i,j}^{(0)} \quad \forall n \in \mathbb{N}.$$

En todo el trabajo solo se considerarán cadenas de Markov homogéneas, a menos que se indique lo contrario. De esta manera podemos prescindir del superíndice que representa la etapa; en adelante se escribirá $p_{i,j}$ en vez de $p_{i,j}^{(n)}$.

Proposición 1. Sea $d^{(n)} = (d_i^{(n)})_{i \in E}$ la distribución de probabilidad de la cadena en la etapa n , i.e

$$d_i^{(n)} = \mathcal{P}(X_n = i).$$

Se verifica que

$$d_j^{(n+1)} = \sum_{i \in E} d_i^{(n)} p_{i,j}.$$

Demostración. En efecto, por el Teorema de la Probabilidad Total se tiene que

$$\mathcal{P}(X_{n+1} = j) = \sum_{i \in E} \mathcal{P}(X_n = i) \mathcal{P}(X_{n+1} = j | X_n = i).$$

□

Entonces, si denotamos por $\mathcal{P} = (p_{i,j})$, el teorema anterior dice que

$$d^{(n+1)} = d^{(n)} \cdot \mathcal{P}.$$

Definición 3. Sea $\{X_n\}$ una cadena de Markov y $p_{i,j}$ sus probabilidades de transición, la matriz $\mathcal{P} = (p_{i,j})$ se denomina matriz de transición. Además, llamaremos distribución a cualquier tupla $d = (d_i)_{i \in E}$ tal que

1. $0 \leq d_i \leq 1$ para cualquier $i \in E$.

$$2. \sum_{i \in E} d_i = 1.$$

Uno puede darse cuenta de que la aplicación $\mathcal{P}(d) = \sum_{i \in E} d_i p_{i,j}$ es lineal. A lo largo de este trabajo, se utilizarán matrices como sinónimo de aplicaciones lineales; por ello, para evitar confusiones se entenderá que, si $\mathcal{P}$ representa a la matriz de transición de una cadena de Markov, entonces

$$\mathcal{P}(d) = d \cdot \mathcal{P}.$$

Definición 4. Dada una cadena de Markov, se dice que un estado $i \in E$ comunica con otro $j \in E$, y se denota por $i \rightarrow j$, si es posible pasar del estado i al j en un número finito de etapas

$$\mathcal{P}(X_n = j | X_0 = i) > 0 \quad \text{para algún } n > 0.$$

Y se dice que intercomunican si $i \rightarrow j$ y $j \rightarrow i$.

Definición 5. Sea $i \in E$ un estado de una cadena de Markov. Se dice que i es un estado trascendente si, una vez que la cadena salga de ese estado, no es seguro que vuelva a él. Y se dice que recurrente en caso contrario.

Definición 6. Se dice que un estado recurrente $i \in E$ es

- recurrente positivo si el tiempo esperado de retorno es finito. Y recurrente nulo en caso contrario.
- Periódico de periodo t si solo es posible volver en un número de pasos que sea múltiplo de t . En el caso de que $t = 1$, se dice que es un estado aperiódico.
- Se dice que ergódico si es recurrente positivo y aperiódico.

Proposición 2. Sean $i, j \in E$ dos estados de una cadena de Markov tales que $i \rightarrow j$. Si i es recurrente, j también. Además, si j es transitorio, i también.

Demostración. Si $i \rightarrow j$ e i es recurrente, entonces la cadena de Markov está en el estado i un número infinito de veces. Debido a que

$$\mathcal{P}(X_n = j | X_0 = i) > 0$$

, la cadena volverá al estado j con absoluta certeza.

En el caso de que j sea trascendente, i no puede ser recurrente porque, como acabamos de ver, j también lo sería. Por tanto, i tiene que ser trascendente. $\square$

Corolario 1. *Si dos estados de una cadena de Markov intercomunican, entonces son del mismo tipo (recurrente positivo, aperiódico o trascendente).*

Definición 7. *Una cadena de Markov se denomina ergódica si todos sus estados son ergódicos y comunican entre sí.*

Este tipo de cadenas son las más importantes y son las que se van a estudiar a continuación. El motivo de su relevancia es la existencia y unicidad de distribuciones estacionarias.

Definición 8. *Una distribución $\pi = (\pi_i)_{i \in E}$ es una distribución estacionaria de una cadena de Markov, si*

$$\mathcal{P}(\pi) = \pi,$$

donde $\mathcal{P}$ es la matriz de transición.

Teorema 1. *Dada una cadena de Markov donde todos los estados intercomunican entre sí, existe una única distribución estacionaria π si y solo si existe un estado recurrente positivo. En este caso, para cualquier distribución d se verifica que*

$$\lim_{n \rightarrow \infty} \mathcal{P}^n(d) = \pi.$$

Demostración. Supongamos que π es una distribución estacionaria. Si todos los estados fuesen trascendentes se tendría que

$$\lim_{n \rightarrow \infty} \mathcal{P}^n = 0,$$

y por tanto,

$$\pi = \lim_{n \rightarrow \infty} \pi \cdot \mathcal{P}^n = 0,$$

lo que es una contradicción. Por lo tanto, algún estado debe de ser recurrente. Como todos los estados intercomunican entre sí, todos los estados son recurrentes. En el Capítulo 6 de [2] se demuestra que, si μ_i es el tiempo medio que la cadena tarda en volver al estado $i \in E$, entonces $\pi_i \mu_i = 1$, concluyendo que π es única, que $\mu_i < \infty$ y que todos los estados son recurrentes positivos.

Recíprocamente, en el mismo capítulo, se demuestra que la distribución π definida a partir de la relación $\pi_i \mu_i = 1$ define una distribución estacionaria.

Finalmente, la demostración de que $\lim_{n \rightarrow \infty} \mathcal{P}^n(d) = \pi$ aparece con total detalle en el Capítulo 6 de [2]. La idea de la demostración es ver que la sucesión $(\mathcal{P}(X_n = i | X_0 = j))_{n \geq 0}$ es una sucesión convergente para cualesquiera $i, j \in E$. $\square$

En particular, una cadena ergódica siempre tiene una distribución estacionaria y, además, es única. Esto implica que $\dim(\text{Ker}(\text{Id} - \mathcal{P})) = 1$, hecho que tendrá relevancia más adelante.

1.2. Inversa de Drazin y aplicaciones a las cadenas de Markov

El principal documento en el que se ha basado el trabajo es el libro de S. L. Campbell y C. D. Meyer, *Generalized Inverses of Linear Transformations*, publicado en 2009 [6]. Concretamente, en los capítulos 7 y 8. En el Capítulo 7, se define la inversa de Drazin para endomorfismos de espacios vectoriales de dimensión finita y se muestran varios resultados relacionados con ella. En el capítulo 8, se aplica la teoría desarrollada en el capítulo anterior a las cadenas de Markov finitas. En lo que sigue k denotará a un cuerpo.

Teorema 2. *Consideremos E un k -espacio vectorial de dimensión finita y $A : E \rightarrow E$ un endomorfismo. Se tiene que*

$$E = \text{Ker } A^m \oplus \text{Im } A^m$$

para cierto m entero y no negativo.

Demostración. Consideremos las cadenas de subespacios vectoriales

$$0 \subseteq \text{Ker } A \subseteq \text{Ker } A^2 \subseteq \cdots \subseteq \text{Ker } A^n \subseteq \cdots \quad (1)$$

$$E \supseteq \text{Im } A \supseteq \text{Im } A^2 \supseteq \cdots \supseteq \text{Im } A^n \supseteq \cdots \quad (2)$$

Como todo espacio vectorial de dimensión finita es noetheriano y artiniiano, las dos cadenas anteriores estacionan, es decir, existe m tal que si $n \geq m$

$$\text{Ker } A^m = \text{Ker } A^n \quad \text{y} \quad \text{Im } A^m = \text{Im } A^n.$$

Sea ahora $e \in \text{Ker } A^m \cap \text{Im } A^m$, entonces $e = A^m(\bar{e})$ y $A^m(e) = A^{2m}(\bar{e}) = 0$. Concluimos que $\bar{e} \in \text{Ker } A^{2m} = \text{Ker } A^m$ y, por tanto, que $e = A^m(\bar{e}) = 0$, es decir, $\text{Ker } A^m \cap \text{Im } A^m = \{0\}$. Finalmente, por la fórmula de la dimensión se concluye que $E = \text{Ker } A^m \oplus \text{Im } A^m$. $\square$

Definición 9. *Dado $A : E \rightarrow E$ un endomorfismo de un k -espacio vectorial de dimensión finita, se denomina índice de A , denotado por $\text{Ind}(A)$, al menor número entero no negativo m que*

$$E = \text{Ker } A^m \oplus \text{Im } A^m.$$

Equivalentemente, $\text{Ind}(A)$ es el menor número entero no negativo tal que $\text{Im } A^m = \text{Im } A^{m+1}$ o que $\text{Ker } A^m = \text{Ker } A^{m+1}$.

Por ejemplo, se tendría que $A : E \rightarrow E$ es isomorfismo si y solo si $\text{Ind}(A) = 0$, porque $E = \text{Im } A^0 = \text{Im } \text{Id} = \text{Im } A$. Por otro lado, tomando el convenio en el que $0^0 = \text{Id}$, se ve fácilmente que $\text{Ind}(0) = 1$. En general, si A es nilpotente de orden m , $\text{Ind}(A) = m$.

El resultado siguiente es fundamental para entender el concepto de inversa de Drazin.

Teorema 3. *Sea $A : E \rightarrow E$ un endomorfismo de un k -espacio vectorial de dimensión finita con $\text{Ind}(A) = m$. Entonces $A|_{\text{Im } A^m} : \text{Im } A^m \rightarrow \text{Im } A^m$ es un isomorfismo y $A|_{\text{Ker } A^m} : \text{Ker } A^m \rightarrow \text{Ker } A^m$ es nilpotente de orden m .*

Demostración. En primer lugar debemos de observar que

$$A(\text{Im } A^m) = \text{Im } A^{m+1} = \text{Im } A^m.$$

Por tanto, la restricción $A|_{\text{Im } A^m} : \text{Im } A^m \rightarrow \text{Im } A^m$ está bien definida. Consideremos un elemento de la forma $A^m(e) \in \text{Ker } A$, entonces $A^{m+1}(e) = 0$ y $e \in \text{Ker } A^{m+1}$. Por la definición de índice concluimos que $e \in \text{Ker } A^m$, luego $A^m(e) = 0$ y, por tanto, que la restricción de A a $\text{Im } A$ es inyectiva. Al ser un endomorfismo, se concluye también que es epiyectiva.

Por otro lado, como

$$A(\text{Ker } A^m) \subseteq \text{Ker } A^{m-1} \subseteq \text{Ker } A^m,$$

la restricción $A|_{\text{Ker } A^m} : \text{Ker } A^m \rightarrow \text{Ker } A^m$ está también bien definida y es fácil ver que $(A|_{\text{Ker } A^m})^m = 0$. Por la definición de índice, $\text{Ker } A^{m-1} \subsetneq \text{Ker } A^m$, concluyendo que la restricción de A a $\text{Ker } A^m$ es nilpotente de orden m . $\square$

Es necesario recordar que, por la propiedad universal de la suma directa, para definir un endomorfismo en E basta con definir dos endomorfismos $A_1 : \text{Im } A^m \rightarrow E$ y $A_2 : \text{Ker } A^m \rightarrow E$.

Definición 10. *Sea $A : E \rightarrow E$ un endomorfismo de un k -espacio vectorial de dimensión finita con índice $\text{Ind}(A) = m$. Se define su inversa de Drazin, y se denotará por A^D , al endomorfismo definido por*

$$A^D(e) = (A|_{\text{Im } A^m})^{-1}(e) \quad \forall e \in \text{Im } A^m,$$

$$A^D(e) = 0 \quad \forall e \in \text{Ker } A^m.$$

Considerando una base de $\text{Im } A^m$ y otra de $\text{Ker } A^m$, se puede expresar el endomorfismo A en coordenadas mediante una matriz por bloques de la siguiente manera

$$A \equiv \begin{pmatrix} C & 0 \\ 0 & N \end{pmatrix}, \quad (3)$$

donde C es una matriz no singular y N es una matriz nilpotente de orden m . Es claro que la primera se refiere al isomorfismo $A|_{\text{Im } A^m}$ y la segunda a $A|_{\text{Ker } A^m}$. Por tanto, la inversa de Drazin, A^D se expresa como sigue

$$A^D \equiv \begin{pmatrix} C^{-1} & 0 \\ 0 & 0 \end{pmatrix}. \quad (4)$$

Esta formas matriciales se denominan representaciones en forma canónica de A y de A^D .

A partir de aquí, se pueden demostrar algunas propiedades elementales de la inversa de Drazin.

Proposición 3. *Sea $A : E \rightarrow E$ un endomorfismo de un k -espacio vectorial de dimensión finita con índice $\text{Ind}(A) = m$ y A^D su inversa de Drazin. Se verifican las siguientes propiedades.*

1. $AA^D = A^D A$.
2. $A^{n+1}A^D = A^n$ para $n \geq m$.
3. $A^D AA^D = A^D$.
4. $\text{Im } A^D = \text{Im } A^m$ y $\text{Ker } A^D = \text{Ker } A^m$.
5. AA^D es la proyección sobre $\text{Im } A^m$.
6. $\text{Id} - AA^D$ es la proyección sobre $\text{Ker } A^m$.
7. Si A es un isomorfismo, entonces $A^D = A^{-1}$.

Demostración. Todas las afirmaciones, salvo la número 4, son triviales a partir de la representación en forma canónica. La afirmación 4 se deduce de forma inmediata de la definición de inversa de Drazin. $\square$

Otro resultado bastante interesante es que la inversa de Drazin es caracterizada por las primeras 3 afirmaciones de la proposición anterior. Se demuestra en el siguiente teorema.

Teorema 4. Sea $A : E \longrightarrow E$ un endomorfismo de un k -espacio vectorial de dimensión finita con índice $\text{Ind}(A)$, y $X : E \longrightarrow E$ otro endomorfismo verificando

1. $A X = X A$,
2. $A^{n+1} X = A^n$ para $n \geq m$,
3. $X A X = X$.

Entonces $X = A^D$.

Demostración. Considérese una base de E para representar A en forma canónica. En dicha base, X se expresa de la siguiente manera.

$$A \equiv \begin{pmatrix} C & 0 \\ 0 & N \end{pmatrix} \quad X \equiv \begin{pmatrix} X_{1,1} & X_{1,2} \\ X_{2,1} & X_{2,2} \end{pmatrix}.$$

De 2, se obtiene que $C^{m+1} X_{1,1} = C^m$ y que $C^{m+1} X_{1,2} = 0$. Multiplicando por la izquierda a ambas igualdades por $C^{-(m+1)}$ queda

$$X_{1,1} = C^{-1} \quad \text{y} \quad X_{1,2} = 0.$$

Aplicando 1 en 2, se tiene que $X A^{m+1} = A^m$ y por un razonamiento similar anterior se tiene que $X_{2,1} = 0$. Finalmente, usando 1 y 3 se llega a la igualdad

$$(X_{2,2})^2 N = X_{2,2}. \quad (5)$$

Multiplicando a ambos lados por N^{m-1} por la derecha

$$X_{2,2} N^{m-1} = (X_{2,2})^2 N^m = 0.$$

Si ahora multiplicamos, de nuevo por la derecha, en (5) por N^{m-2}

$$X_{2,2} N^{m-2} = (X_{2,2})^2 N^{m-1} = 0.$$

Repitiendo el argumento un número finito de veces, concluimos que

$$X_{2,2} = (X_{2,2})^2 N = 0$$

□

Hay que prestar atención a que la propiedad que siempre se verifica es que $A^D A A^D = A^D$, y que en general no es cierto que $A A^D A = A$. Sin embargo, hay una situación donde sí se cumple.

Proposición 4. Dada $A : E \longrightarrow E$ un endomorfismo de un k -espacio vectorial de dimensión finita. Se verifica $A A^D A = A$ si y solo si $\text{Ind}(A) \leq 1$.

Demostración. Usando la forma canónica, tenemos que

$$A A^D A \equiv \begin{pmatrix} C & 0 \\ 0 & 0 \end{pmatrix},$$

que solo coincide con A si N no aparece en la expresión, es decir, A es invertible, o si $N = 0$. $\square$

Definición 11. Si $A : E \longrightarrow E$ un endomorfismo de un k -espacio vectorial de dimensión finita tiene $\text{Ind}(A) \leq 1$, su inversa de Drazin se denomina inversa de grupo de A y se suele denotar por $A^\#$.

El teorema que viene a continuación, deja bastante claro la razón del nombre de “inversa de grupo”.

Teorema 5. Sea $A : E \longrightarrow E$ un endomorfismo de un k -espacio vectorial de dimensión finita. A pertenece a un grupo multiplicativo $G \subseteq \text{Hom}_k(E)$ si y solo si $\text{Ind}(A) \leq 1$.

Demostración. Sea $G \subseteq \text{Hom}_k(E)$ un grupo, con la composición de aplicaciones que contenga a A . Entonces existe el inverso de A en G , A^G , verificando las hipótesis del Teorema 4 y la Proposición 4 de forma inmediata, concluyendo que $A^G = A^D$ y que $\text{Ind}(A) \leq 1$.

Recíprocamente, consideremos los endomorfismos $B \in \text{Hom}_k(E)$ tales que $B|_{\text{Im } A} : \text{Im } A \longrightarrow \text{Im } A$ es un isomorfismo y $B|_{\text{Ker } A} = 0$. El conjunto de estos endomorfismos constituye un grupo con la composición de aplicaciones al que pertenece A . $\square$

Queda claro entonces que el producto $A A^D A$ es bastante relevante. Hay que observar que este producto es la composición de A con la proyección sobre $\text{Im } A^m$ y que, por otro lado, $A - A A^D A = A(\text{Id} - A^D A)$ es la composición de A con la proyección sobre $\text{Ker } A^m$.

Definición 12. Sea $A : E \longrightarrow E$ un endomorfismo de un k -espacio vectorial de dimensión finita y A^D su inversa de Drazin. Se denomina parte core de A al endomorfismo

$$C_A = A A^D A$$

y se denomina parte nilpotente de A al endomorfismo

$$N_A = A - A A^D A = A(\text{Id} - A^D A)$$

Proposición 5. Si $A : E \longrightarrow E$ un endomorfismo de un k -espacio vectorial de dimensión finita con índice $\text{Ind}(A) = m$, entonces N_A es nilpotente de orden m e $\text{Ind}(C_A) \leq 1$.

Demostración. Si $m = 0$, entonces $N_A = 0$ y $\text{Ind}(C_A) = 0$. En caso contrario, se tiene

$$\begin{aligned} N_A^m &= A^m(\text{Id} - A^D A)^m = A^m(\text{Id} - A^D A) \\ &= A^m - A^{m+1} A^D = A^m - A^m = 0, \end{aligned} \tag{6}$$

donde se ha utilizado la propiedad 2 de la Proposición 3. Ahora, debido a que $\text{Im } C_A = \text{Im } A^m$ y $\text{Ker } C_A = \text{Ker } A^m$, se deduce que $E = \text{Im } C_A \oplus \text{Ker } C_A$ y que $\text{Ind}(C_A) = 1$. $\square$

Definición 13. Sea $A : E \longrightarrow E$ un endomorfismo de un k -espacio vectorial de dimensión finita con índice $\text{Ind}(A) = m$, C_A su core y N_A su parte nilpotente. A la descomposición $A = C_A + N_A$ se denomina descomposición core-nilpotente del endomorfismo A .

Tomando una base que proporcione la representación en forma canónica de A , se ve fácilmente que

$$C_A \equiv \begin{pmatrix} C & 0 \\ 0 & 0 \end{pmatrix}, \quad N_A \equiv \begin{pmatrix} 0 & 0 \\ 0 & N \end{pmatrix}.$$

La descomposición core-nilpotente de un endomorfismo es única en el siguiente sentido.

Teorema 6. Sea $A : E \longrightarrow E$ un endomorfismo de un k -espacio vectorial de dimensión finita con índice $\text{Ind}(A) = m$. Existe una única descomposición $A = X + Y$ donde X e Y son dos endomorfismos verificando:

1. $XY = YX = 0$.
2. $\text{Ind}(X) \leq 1$.
3. Y es nilpotente.

Demostración. Sean X e Y verificando el enunciado. Si $\text{Ind}(X) = 0$, de la primera condición se deduce que $Y = 0$. Por tanto, A es invertible y $X = A = C_A$. Supongamos, entonces que $\text{Ind}(X) = 1$.

Como Y es nilpotente, para cierto $r \geq m$

$$A^r = (X + Y)^r = X^r$$

y por tanto, debido a que $\text{Ind}(X) = 1$

$$\text{Im } A^m = \text{Im } A^r = \text{Im } X^r = \text{Im } X,$$

$$\text{Ker } A^m = \text{Ker } A^r = \text{Ker } X^r = \text{Ker } X.$$

Es claro entonces que, la primera condición implica que $Y(e) = 0$ para cualquier $e \in \text{Im } A^m$ y que $Y(e) = A(e)$ si $e \in \text{Ker } A^m$. Por este motivo, $X(e) = A(e)$ para todo $e \in \text{Im } A^m$. Esto demuestra que $Y = N_A$ y que $X = C_A$. $\square$

Proposición 6. *Con las notaciones precedentes, se verifica que*

1. $\text{Ind}(A^D) = \text{Ind}(C_A)$.
2. $(A^D)^D = C_A$.
3. $C_A^D = A^D$.

Demostración. Inmediato a partir de la representación en forma canónica. $\square$

Todos estos resultados de la inversa de Drazin son relevantes para entender su importancia y su aplicación a las cadenas de Markov.

En adelante, $\{X_n\}_{n \in \mathbb{N}}$ será una cadena de Markov finita y homogénea de m estados. El siguiente resultado es fundamental, porque es sobre el que se sustentan todos los resultados siguientes.

Teorema 7. *Sea $\mathcal{P}$ la matriz de transición de una cadena de Markov ergódica. Entonces, si $A = \text{Id} - \mathcal{P}$ se tiene que $\text{Ind}(A) = 1$.*

En realidad, el resultado es válido incluso si la cadena de Markov no es ergódica (véase teorema 8.2.1 de [6]), sin embargo, este caso no es relevante para este trabajo.

Obsérvese que resultado anterior implica que $\mathbb{R}^m = \text{Ker } A \oplus \text{Im } A$. Para el siguiente resultado es interesante recordar que $A^D A$ es la proyección sobre $\text{Im } A$ y que $\text{Id} - A^D A$ es la proyección sobre $\text{Ker } A$.

Teorema 8. *Sea $\mathcal{P}$ la matriz de transición de una cadena de Markov ergódica. Entonces*

$$\text{Id} - A^D A = \lim_{n \rightarrow \infty} \mathcal{P}^n.$$

Demostración. Sea $d \in \mathbb{R}^m$. En primer lugar, sabemos que, si $\sum_{i=1}^m d_i = 1$, $\lim_{n \rightarrow \infty} \mathcal{P}^n(d)$ existe y que coincide con π , donde π es la distribución estacionaria. Expresamos $d = v + A(\bar{d})$, con $v \in \text{Ker } A$. Entonces

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathcal{P}^n(d) &= \lim_{n \rightarrow \infty} \mathcal{P}^n(v + A(\bar{d})) = \lim_{n \rightarrow \infty} \mathcal{P}^n(v) + \lim_{n \rightarrow \infty} (\mathcal{P}^n - \mathcal{P}^{n+1})(\bar{d}) \\ &= v + s\pi - s\pi = v, \end{aligned}$$

donde $s = \sum_{i=1}^m \bar{d}_i$. Así demostramos que $\lim_{n \rightarrow \infty} \mathcal{P}^n$ es la proyección sobre $\text{Ker } A$ y por tanto, que coincide con $\text{Id} - A^D A$. $\square$

En adelante se denotará por W a $\text{Id} - A^D A$. Otro resultado bastante útil, especialmente para computar A^D de forma aproximada, es el siguiente.

Teorema 9. *Con las notaciones anteriores se tiene que*

$$A^D = \sum_{r=0}^{\infty} (\mathcal{P}^r - W)$$

Demostración. Si $v \in \text{Ker } A$, entonces

$$\sum_{r=0}^{\infty} (\mathcal{P}^r - W)(v) = \sum_{r=0}^{\infty} (v - v) = 0. \quad (7)$$

Por otro lado, si $A(d) \in \text{Im } A$, se tiene que

$$\begin{aligned} \sum_{r=0}^{\infty} (\mathcal{P}^r - W)(A(d)) &= \sum_{r=0}^{\infty} (\mathcal{P}^r(d) - \mathcal{P}^{r+1}(d)) \\ &= \lim_{r \rightarrow \infty} (d - \mathcal{P}^{r+1}(d)) = d - W(d). \end{aligned} \quad (8)$$

Obsérvese que $A(d - W(d)) = A(d)$ y que $d - W(d) \in \text{Im } A$, concluyendo, por tanto, que $A^D(d) = d - W(d)$. $\square$

El primer resultado interesante y útil que podemos obtener a partir de la inversa de Drazin está relacionado con el número de veces que la cadena pasa por un estado en concreto. Sea $N_{i,j}^{(n)}$ el número esperado de veces que la cadena de Markov pasa por el estado j durante n etapas habiendo empezado en el estado i y llamaremos a la matriz $N = (N_{i,j}^{(n)})$ matriz de recurrencia de la cadena de Markov.

Teorema 10. Sea $\mathcal{P}$, A , W y N como antes. Se verifica que

$$A^D = \lim_{n \rightarrow \infty} (N^n) - (n+1)W).$$

Demostración. Es consecuencia inmediata del teorema anterior y del hecho de que

$$N^n = \sum_{r=0}^n \mathcal{P}^r.$$

□

Observación: El Teorema 8.3.2 de [6] afirma que $A^D = \lim_{n \rightarrow \infty} (N^n) - nW$. Sin embargo, esto no es cierto: si tomamos un vector $v \in \text{Ker } A$, no es difícil ver que

$$\lim_{n \rightarrow \infty} (N^n(v) - nW(v)) = v,$$

pero $A^D(v) = 0$. El autor utiliza que $\lim_{n \rightarrow \infty} N^n = \lim_{n \rightarrow \infty} N^{n-1}$, entonces concluye que

$$\lim_{n \rightarrow \infty} (N^n) - nW = \lim_{n \rightarrow \infty} (N^{n-1}) - nW). \quad (9)$$

Este argumento sería válido si los límites $\lim_{n \rightarrow \infty} N^n$ y $\lim_{n \rightarrow \infty} nW$ existiesen, pero no es el caso. Se puede entender que la motivación de este teorema es aproximar N^n mediante A^D , es decir, considerando n suficientemente grande,

$$N^n \approx A^D + (n+1)W.$$

En este caso, también se puede considerar que $n \approx n+1$ y $N^n \approx A^D + nW$. Así que, aunque en la práctica funcione, no es riguroso. □

Consideremos ahora $m_{i,j}$, que es el tiempo esperado que tarda la cadena en alcanzar el estado j por primera vez habiendo empezado en el estado i . La matriz $M = (m_{i,j})$ se denomina matriz de la media del primer paso (“*mean first passage matrix*”). Nos proponemos ahora calcular esta matriz a partir de la inversa de Drazin.

Antes de enunciar el teorema, se introducirá a continuación la notación que se usará. Si X es una matriz cuadrada, denotaremos por X_d a la matriz que se obtiene estableciendo todas las entradas de X fuera de la diagonal a 0:

$$X_d = (\delta_{i,j} X_{i,j}),$$

donde $\delta_{i,j}$ denota a la delta de Kronecker. Por otro lado J denota a la matriz enteramente de 1: $J_{i,j} = 1$ para cualquier i y j .

Observación: Usando las propiedades de la esperanza matemática, se puede llegar a la expresión siguiente [5]:

$$m_{i,j} = \sum_{r \neq j} p_{i,r} (m_{i,j} + 1) + p_{i,j}. \quad (10)$$

Sumando y restando el término $p_{i,j} m_{j,j}$ nos queda

$$m_{i,j} = 1 + \sum_{r=1}^m p_{i,r} m_{r,j} - p_{i,j} m_{j,j}, \quad (11)$$

que, en forma matricial quedaría, como sigue

$$M = J + \mathcal{P} M - \mathcal{P} M_d. \quad (12)$$

□

Teorema 11. Sea $\mathcal{P}$, A , W y M como arriba. Se verifica que

$$M = (Id - A^D + J A_d^D)(W_d)^{-1}.$$

Demostración. Se sabe que la matriz de la media del primer paso es la única solución de la ecuación siguiente [6],

$$A X = J - \mathcal{P} X_d. \quad (13)$$

Denotemos por R al miembro derecho de la igualdad del enunciado y comprobemos que es solución de esta ecuación. Obsérvese que $R_d = (W_d)^{-1}$. Entonces

$$\begin{aligned} A R &= A(Id - A^D + J A_d^D)(W_d)^{-1} = (A - A A^D)(W_d)^{-1} \\ &= (Id - \mathcal{P} - A A^D)(W_d)^{-1} = (W - \mathcal{P})(W_d)^{-1} \\ &= J - \mathcal{P}(W_d)^{-1} = J - \mathcal{P} R_d. \end{aligned} \quad (14)$$

Nótese que, como todas las filas de A suman 0, el producto matricial $A J A_d^D = 0$. □

Observación: La comunidad matemática no se pone de acuerdo con los valores que M debe de tener en la diagonal. Unos autores afirman que debe de ser 0, pues si la cadena se encuentra en el estado i no necesita ninguna etapa para acceder a él, pero unos pocos entienden que $m_{i,i}$ debe de ser el tiempo esperado que la cadena tarda en volver al estado i una vez que sale de él. Si se considera lo primero, M no verifica la ecuación (13).

1.3. Endomorfismos core-nilpotentes

El trabajo realizado consiste, en esencia, en generalizar los resultados previos para cadenas de Markov infinitas. Para ello, es necesario estudiar el marco teórico de los espacios vectoriales de dimensión arbitraria donde se pueda definir y aplicar la inversa de Drazin. Esto motiva a estudiar un tipo concreto de endomorfismos denominados core-nilpotentes. La información que sigue ha sido extraída del artículo [7].

En lo que sigue, k será un cuerpo, V un k -espacio vectorial de dimensión arbitraria y $f : V \longrightarrow V$ un endomorfismo.

Definición 14. Se dice que $f : V \longrightarrow V$ es un endomorfismo core-nilpotente (o endomorfismo CN) si existen dos subespacios suplementarios invariantes por f , U_1 y U_2 , tales que $f|_{U_1}$ es un isomorfismo y $f|_{U_2}$ es nilpotente. A la descomposición $V = U_1 \oplus U_2$ se denomina descomposición core-nilpotente (o descomposición CN).

Definición 15. Dado un endomorfismo CN $f : V \longrightarrow V$, se llama índice de f , y se denota por $\text{Ind}(f)$, al menor número entero no negativo n tal que $\text{Im } f^n = \text{Im } f^{n+1}$ y $\text{Ker } f^n = \text{Ker } f^{n+1}$, pero $\text{Im } f^n \neq \text{Im } f^{n-1}$ o $\text{Ker } f^n \neq \text{Ker } f^{n-1}$. Por convenio, si f es invertible se dice que $\text{Ind}(f) = 0$.

Teorema 12. (Caracterización de los endomorfismos CN) Sea V un k -espacio vectorial de dimensión arbitraria y $f : V \longrightarrow V$ un endomorfismo. Son equivalentes

1. Existen dos endomorfismos f_1 y f_2 tales que $f = f_1 + f_2$ con $f_1 \circ f_2 = f_2 \circ f_1 = 0$, $\text{Ind}(f_1) \leq 1$ y f_2 nilpotente.
2. Para cierto $n \in \mathbb{N}$ se tiene que $\text{Im } f^n = \text{Im } f^{n+1}$ y $\text{Ker } f^n = \text{Ker } f^{n+1}$.
3. Para cierto $n \in \mathbb{N}$, $V = \text{Ker } f^n \oplus \text{Im } f^n$.
4. f es core-nilpotente.

Demostración. $1 \Rightarrow 2$. Sea n tal que $f_2^n = 0$, entonces $f^m = f_1^m$ para cualquier $m \geq n$. Como $\text{Ind}(f_1) \leq 1$ se tiene

$$\text{Ker } f^{n+1} = \text{Ker } f_1^{n+1} = \text{Ker } f_1^n = \text{Ker } f^n,$$

y que

$$\text{Im } f^{n+1} = \text{Im } f_1^{n+1} = \text{Im } f_1^n = \text{Im } f^n.$$

$2 \Rightarrow 3$. Sea $f^n(v) \in \text{Ker } f^n \cap \text{Im } f^n$. Entonces $f^n(f^n(v)) = 0$ y $v \in \text{Ker } f^{2n} = \text{Ker } f^n$, luego $f^n(v) = 0$. Por otro lado, como $\text{Im } f^n = \text{Im } f^{2n}$, dado

$v \in V$, existe $\bar{v} \in V$ tal que $f^n(v) = f^{2n}(\bar{v})$. Por lo tanto, $v - f^n(\bar{v}) \in \text{Ker } f^n$, de donde se deduce que $V = \text{Ker } f^n + \text{Im } f^n$.

$3 \Rightarrow 4$. Como $\text{Ker } f \subseteq \text{Ker } f^n$, el hecho de que $\text{Ker } f \cap \text{Im } f^n = \{0\}$ implica que $\text{Ker } f \cap \text{Im } f^n = \{0\}$, concluyendo que $f|_{\text{Im } f^n}$ es inyectiva. Además, dado $f^n(v) \in \text{Im } f^n$, expresamos $v = f^n(v_1) + v_2$ con $v_2 \in \text{Ker } f^n$. Se tiene que $f^n(v) = f(f^{2n-1}(v_1))$ demostrando que $f|_{\text{Im } f^n}$ es epiyectiva.

$4 \Rightarrow 1$. Si $V = U_1 \oplus U_2$ verificando las condiciones de la Definición 14, para cada $v = v_1 + v_2$ con $v_1 \in U_1$ y $v_2 \in U_2$ definimos

$$f_1(v) = f(v) \quad f_2(v) = f(v_2). \quad (15)$$

Es claro que $f = f_1 + f_2$, que $f_1 \circ f_2 = f_2 \circ f_1 = 0$ y que f_2 es nilpotente. Veamos que $\text{Ind}(f_1) \leq 1$. Sea $v \in \text{Ker } f_1^2$, entonces $f_1^2(v) = f^2(v_1) = 0$ para cierto $v_1 \in U_1$. Como $f|_{U_1}$ es inyectiva, se tiene que $f(v_1) = f_1(v) = 0$. Se ha demostrado que $\text{Ker } f_1^2 = \text{Ker } f_1$. Ahora, si $f_1(v) \in \text{Im } f_1$ con $v = v_1 + v_2$, para cierto $v_1 \in U_1$ y $v_2 \in U_2$, se concluye de que $f|_{U_1}$ sea epiyectiva que $f_1(v) = f(v_1) = f_1^2(\bar{v}_1) \in \text{Im } f_1^2$ y, por lo tanto que $\text{Im } f_1^2 = \text{Im } f_1$.

Con todo lo anterior se deduce que $\text{Ind}(f_1) \leq 1$. □

Corolario 2. *La descomposición CN de un endomorfismo CN es única.*

Corolario 3. *Dado un endomorfismo CN $f : V \rightarrow V$, $\text{Ind}(f) \leq 1$ si y solo si $V = \text{Ker } f \oplus \text{Im } f$.*

Corolario 4. *Si $f : V \rightarrow V$ es un endomorfismo que admite un polinomio anulador, entonces f es un endomorfismo CN. Como consecuencia todo endomorfismo de un k -espacio vectorial de dimensión finita es CN.*

Demostración. Sea $p(x) = x^n q(x)$ el polinomio anulador de f , suponiendo que x^n y $q(x)$ son coprimos. Entonces, se sabe que $V = \text{Ker } q(f) \oplus \text{Ker } f^n$, además es fácil de ver que $f|_{\text{Ker } q(f)} : \text{Ker } q(f) \rightarrow \text{Ker } q(f)$ es un isomorfismo. □

En [1], Drazin presentó una generalización del elemento inverso en anillos y semigrupos; a partir de esta definición, se determinará la inversa de Drazin de un endomorfismo.

Definición 16. *Un endomorfismo $f : V \rightarrow V$ tiene una inversa de Drazin si existe un endomorfismo $f^D : V \rightarrow V$ que verifica*

$$1. \quad f \circ f^D = f^D \circ f.$$

2. $f^{m+1} \circ f^D = f^m$ para cierto entero no negativo m .

3. $f^D \circ f \circ f^D = f^D$.

Teorema 13. *Un endomorfismo $f : V \longrightarrow V$ tiene inversa de Drazin si y solo si es core-nilpotente.*

Demostración. Supongamos que f^D existe. Tomemos los endomorfismos

$$f_1 = f \circ f^D \circ f \quad f_2 = f - f \circ f^D \circ f$$

y veamos que verifican las hipótesis del Teorema 12. Claramente $f_1 \circ f_2 = f_2 \circ f_1 = 0$. Además, para cierto m se tiene

$$\begin{aligned} (f - f \circ f^D \circ f)^m &= \sum_{i=0}^m \binom{m}{i} (-1)^i f^{m+i} \circ (f^D)^i \\ &= f^m \sum_{i=0}^m (-1)^i \binom{m}{i} = 0 \end{aligned} \tag{16}$$

Faltaría por probar que $\text{Ind}(f_1) \leq 1$. Por el Corolario 3 basta con probar que $V = \text{Ker } f_1 \oplus \text{Im } f_1$.

En primer lugar, obsérvese que $f_1 \circ f^D \circ f_1 = f_1$ y que $f^D \circ f_1 = f_1 \circ f^D$. Por ello, si $f_1(v) \in \text{Ker } f_1$ se tiene que

$$f_1(v) = (f_1 \circ f^D \circ f_1)(v) = 0,$$

luego $\text{Ker } f_1 \cap \text{Im } f_1 = \{0\}$.

En segundo lugar, dado $v \in V$, usando que $v - f_1(f^D(v)) \in \text{Ker } f_1$ se concluye que $V = \text{Ker } f_1 + \text{Im } f_1$.

Recíprocamente, si f es core-nilpotente, entonces podemos definir el siguiente endomorfismo. Sea $V = U_1 \oplus U_2$ la descomposición CN, para cada $v = u_1 + u_2$,

$$f^D(v) = (f|_{U_1})^{-1}(u_1). \tag{17}$$

Es fácil de ver que f^D verifica la Definición 16. $\square$

Proposición 7. *Dado $f : V \longrightarrow V$ un endomorfismo, si existe una inversa de Drazin, esta es única.*

Demostración. Por el teorema anterior, f es CN y por tanto $V = U_1 \oplus U_2$ con las condiciones de la Definición 14. Por la propiedad 2 de la definición 16, se tiene que $(f|_{U_1})^{m+1} \circ f^D|_{U_1} = (f|_{U_1})^m$ y, por tanto, que $f^D|_{U_1} = (f|_{U_1})^{-1}$.

Por otro lado, al ser $f|_{U_2}$ nilpotente, $f|_{U_2} \circ f^D|_{U_2}$ también lo es, luego $\text{Id}|_{U_2} - (f|_{U_2} \circ f^D|_{U_2})$ es un isomorfismo sobre U_2 . De la Propiedad 3 de la Definición 16 se obtiene que

$$f^D|_{U_2} - f^D|_{U_2} \circ f|_{U_2} \circ f^D|_{U_2} = f^D|_{U_2}(\text{Id}|_{U_2} - f|_{U_2} \circ f^D|_{U_2}) = 0$$

Y por tanto, $f^D|_{U_2} = 0$. Con esto concluimos que f^D está unívocamente determinada. $\square$

Como se puede observar, la inversa de Drazin se puede definir en el caso de dimensión arbitraria. A partir de aquí, surgen resultados análogos a los que se mostraron en la Subsección 1.2.

Proposición 8. *Dado $f : V \longrightarrow V$ un endomorfismo CN con $\text{Ind}(f) = m$ y f^D su inversa de Drazin, se verifica las siguientes propiedades.*

1. $\text{Im } f^D = \text{Im } f^m$ y $\text{Ker } f^D = \text{Ker } f^m$.
2. $f \circ f^D$ es la proyección sobre $\text{Im } f^m$.
3. $\text{Id} - f \circ f^D$ es la proyección sobre $\text{Ker } f^m$.
4. Si f es invertible, $f^D = f^{-1}$.

A partir de la existencia de la inversa de Drazin para endomorfismos CN se puede aplicar la teoría previa a algunas cadenas de Markov infinitas (ver Sección 3).

2. Búsqueda de un ejemplo ilustrativo

Para dar valor al trabajo, sería interesante buscar un ejemplo de cadena de Markov infinita donde se pueda aplicar la teoría.

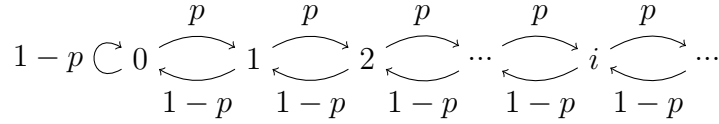
Uno de los ejemplos de cadena de Markov infinita más recurrentes en la literatura es el movimiento aleatorio. Supongamos que un objeto o una partícula que se mueve aleatoriamente a lo largo de una línea formada por los números enteros no negativos.

Sea X_n la variable aleatoria “posición de una partícula en al etapa n -ésima”. El conjunto de estados de la cadena será $\mathbb{N}$, donde consideraremos que $0 \in \mathbb{N}$. Dado $0 \leq p \leq 1$, se definen las probabilidades de transición siguientes

$$\mathcal{P}(X_{n+1} = j | X_n = i) = p_{i,j} = \begin{cases} p & \text{si } j = i + 1 \\ 1 - p & \text{si } j = i - 1 \text{ o si } j = i = 0 \\ 0 & \text{en otro caso} \end{cases} \quad (18)$$

Con otras palabras, la partícula avanzará con probabilidad p y retrocederá con probabilidad $1 - p$. En el caso de que la partícula se encuentre en la posición 0, como no puede retroceder, se mantendrá en el 0 con probabilidad $1 - p$.

Una forma de visualizar la cadena es mediante el siguiente diagrama



Denotaremos por $\mathcal{P} = (p_{i,j})$ a la matriz u operador de transición. En forma matricial se tendría lo siguiente

$$\mathcal{P} = \begin{pmatrix} 1 - p & p & 0 & 0 & \dots \\ 1 - p & 0 & p & 0 & \dots \\ 0 & 1 - p & 0 & p & \dots \\ 0 & 0 & 1 - p & 0 & \ddots \\ \vdots & \vdots & \vdots & \ddots & \ddots \end{pmatrix}. \quad (19)$$

Lo primero es averiguar si esta cadena de Markov es ergódica. Debido a que todos los estados intercomunican entre sí, se puede utilizar el Teorema 1. Entonces, si existe una distribución estacionaria, la cadena será ergódica.

Una distribución $\pi = (\pi_i)_{i \geq 0}$ será una distribución estacionaria si verifica el siguiente sistema de ecuaciones

$$\begin{aligned} \pi \cdot \mathcal{P} &= \pi \\ \sum_{i \geq 0} \pi_i &= 1 \end{aligned} \quad (20)$$

Veamos si tiene solución el sistema de ecuaciones dado por $\mathcal{P}$,

$$\left\{ \begin{array}{lcl} \pi_0 & = & (1-p)\pi_0 + (1-p)\pi_1 \\ \pi_1 & = & p\pi_0 + (1-p)\pi_2 \\ \vdots & \vdots & \vdots \\ \pi_i & = & p\pi_{i-1} + (1-p)\pi_{i+1} \\ \vdots & \vdots & \vdots \end{array} \right. \quad (21)$$

De la primera ecuación se obtiene que

$$\pi_1 = \frac{p}{1-p}\pi_0.$$

Ahora, sustituyendo la segunda y despejando, se llega a que

$$\pi_2 = \left(\frac{p}{1-p}\right)^2 \pi_0.$$

Por inducción, supondremos que

$$\pi_i = \left(\frac{p}{1-p}\right)^i \pi_0. \quad (22)$$

Considerando la ecuación correspondiente del sistema (21)

$$\pi_i = p\pi_{i-1} + (1-p)\pi_{i+1} = p\left(\frac{p}{1-p}\right)^{i-1} \pi_0 + (1-p)\pi_{i+1}.$$

Dividiendo por $(1-p)$ y restando $\frac{p^i}{(1-p)^{i-1}}\pi_0$ nos queda que

$$\pi_{i+1} = \left(\frac{p^i}{(1-p)^{i+1}} - \frac{p^i}{(1-p)^i}\right) \pi_0 = \frac{p^{i+1}}{(1-p)^{i+1}} \pi_0.$$

Concluimos que la distribución $\pi = (\pi_i)_{i \geq 0}$ definida por (22) verifica el sistema de ecuaciones infinito (21), pero también se tiene que verificar que $\sum \pi_i = 1$.

$$\sum_{i \geq 0} \pi_i = \pi_0 \sum_{i \geq 0} \left(\frac{p}{1-p}\right)^i. \quad (23)$$

La serie anterior únicamente converge si $\frac{p}{1-p} < 1$, o equivalentemente si $p < \frac{1}{2}$. En este caso, se tiene que

$$\pi_0 \sum_{i \geq 0} \left(\frac{p}{1-p}\right)^i = \pi_0 \frac{1}{1 - \frac{p}{1-p}} = \pi_0 \frac{1-p}{1-2p}. \quad (24)$$

Tomando

$$\pi_0 = \frac{1-2p}{1-p}$$

y, por tanto,

$$\pi_i = \frac{1-2p}{1-p} \left(\frac{p}{1-p} \right)^i$$

obtenemos una distribución estacionaria.

Por el Teorema 1, la cadena de Markov considerada es ergódica únicamente si $p < \frac{1}{2}$.

Este es uno de los ejemplos más representativos de una cadena de Markov infinita. En la Sección 4 la usaremos como ejercicio práctico al que se aplicará la teoría del siguiente apartado.

3. Generalización de los resultados

Sea $\{X_n\}_{n \in \mathbb{N}}$ una cadena de Markov infinita y homogénea. Sea $E = \mathbb{N}$ su conjunto de estados y $p_{i,j} = \mathcal{P}(X_{n+1} = j | X_n = i)$ sus probabilidades de transición. Dado $d^{(0)} = (d_i^{(0)})_{i \in \mathbb{N}}$ una distribución inicial, por la Proposición 1 la distribución en la siguiente etapa es

$$d^{(1)} = \left(\sum_{i \in \mathbb{N}} d_i^{(0)} p_{i,j} \right)_{j \in \mathbb{N}}.$$

Se ve fácilmente que la aplicación

$$\mathcal{P}(d) = \left(\sum_{i \in \mathbb{N}} d_i p_{i,j} \right)_{j \in \mathbb{N}}, \quad (25)$$

donde $d = (d_i)_{i \in \mathbb{N}}$ y $d_i \in \mathbb{R}$, es una aplicación lineal, en caso de existir.

Sin embargo, ¿es $\sum d_i p_{i,j}$ siempre una serie convergente? ¿Por qué? Si no es así, ¿cuándo podemos garantizar que sí lo es? Y la pregunta más importante, si $\mathcal{P}$ es un endomorfismo, ¿cuál es el espacio vectorial?

Dado un endomorfismo $f : V \longrightarrow V$, se puede hablar de matriz asociada siempre y cuando exista alguna base numerable en V . Supongamos, entonces,

que sea así y sea $\{v_i\}_{i \in I}$ una base de V . Podemos definir la matriz (finita o no) $A = (a_{i,j})$ donde $a_{i,j}$ es la coordenada j -ésima del vector $f(v_i)$,

$$f(v_i) = \sum_{j \in I} a_{i,j} v_j.$$

Aunque $\{v_i\}_{i \in I}$ tenga un número infinito de vectores, al ser una base, la suma de arriba es finita, luego $a_{i,j} = 0$ para casi todo $j \in \mathbb{N}$. Es decir, cada fila de A únicamente tiene un número finito de elementos no nulos.

Lamentablemente, es posible que no se dé esta situación en alguna cadena de Markov. Incluso en el caso de que sí sea así, no podemos considerar como espacio vectorial $V = \bigoplus_{i \in \mathbb{N}} \mathbb{R}$, porque se pueden considerar distribuciones que no pertenezcan a la suma directa. Por ejemplo, en la sección anterior se calculó una distribución estacionaria π , pero $\pi \notin \bigoplus_{i \in \mathbb{N}} \mathbb{R}$.

En resumen, aunque $\mathcal{P}$ se considere una matriz, no es una matriz asociada a un endomorfismo en el sentido habitual en la teoría de endomorfismos.

Otra alternativa es considerar $V = \prod_{i \in \mathbb{N}} \mathbb{R}$, sin embargo, se puede dar que $\sum d_i p_{i,j}$ no sea convergente y que, por tanto, el transformado de algún elemento no exista. Así que habría que considerar un subespacio de $\prod_{i \in \mathbb{N}} \mathbb{R}$, de tal manera que pueda asegurarse que las series correspondientes converjan.

Como $0 \leq p_{i,j} \leq 1$, entonces, para $d = (d_i)_{i \in \mathbb{N}}$

$$\sum_{i \in \mathbb{N}} d_i p_{i,j} \leq \sum_{i \in \mathbb{N}} d_i. \quad (26)$$

Si $\sum d_i < \infty$, entonces $\sum d_i p_{i,j} < \infty$ y, además,

$$\sum_{j \in \mathbb{N}} \sum_{i \in \mathbb{N}} d_i p_{i,j} = \sum_{i \in \mathbb{N}} \sum_{j \in \mathbb{N}} d_i p_{i,j} = \sum_{i \in \mathbb{N}} d_i \sum_{j \in \mathbb{N}} p_{i,j} = \sum_{i \in \mathbb{N}} d_i < \infty. \quad (27)$$

Por ello, el subespacio

$$V = \left\{ (d_i)_{i \in \mathbb{N}} \in \prod_{i \in \mathbb{N}} \mathbb{R} : \sum_{i \in \mathbb{N}} |d_i| < \infty \right\}$$

puede ser buen candidato.

En efecto, V es un subespacio: sean $(a_i)_{i \in \mathbb{N}}, (b_i)_{i \in \mathbb{N}} \in V$ y $\lambda \in \mathbb{R}$,

$$\sum_{i \in \mathbb{N}} |a_i + \lambda b_i| \leq \sum_{i \in \mathbb{N}} |a_i| + |\lambda| \sum_{i \in \mathbb{N}} |b_i| < \infty.$$

Obsérvese que V es, en esencia, el espacio de las sucesiones cuya serie es normalmente convergente, $\ell^1(\mathbb{R})$. En verdad, es suficiente con considerar el espacio de las sucesiones cuya serie es convergente, sin embargo es preferible tomar $\ell^1(\mathbb{R})$ porque es un espacio normado.

Proposición 9. *Sea $\{X_n\}_{n \in \mathbb{N}}$ una cadena de Markov infinita cuyo conjunto de estados es $\mathbb{N}$. Sean $p_{i,j}$ sus probabilidades de transición. El endomorfismo $\mathcal{P} : \ell^1(\mathbb{R}) \longrightarrow \ell^1(\mathbb{R})$ definido por*

$$\mathcal{P}(d) = \left(\sum_{i \in \mathbb{N}} d_i p_{i,j} \right)_{j \in \mathbb{N}}, \quad (28)$$

está bien definido.

Demostración. Es inmediato a partir de de (26) y (27) y del hecho de que una serie normalmente convergente implica convergente. $\square$

Definición 17. *Dada una cadena de Markov homogénea e infinita, se llamará al endomorfismo $\mathcal{P}$, definido a partir de la proposición anterior, operador de transición de la cadena de Markov.*

Aunque pueda no tener sentido, se utilizará la siguiente notación

$$\mathcal{P}(d) = d \cdot \mathcal{P}$$

por su analogía con el producto de matrices en el caso de dimensión finita.

Obsérvese que no es necesario que la cadena de Markov sea ergódica para que $\mathcal{P}$ exista.

Sea $A = \text{Id} - \mathcal{P}$, tal y como se demuestra en [6], $\text{Ind}(A) = 1$ en el caso de una cadena de Markov finita. Uno podría preguntarse si ocurre lo mismo para una cadena de Markov arbitraria. La respuesta es un rotundo no. Tomamos el ejemplo de la sección anterior con $p \geq \frac{1}{2}$, se ha demostrado que no existen distribuciones estacionarias, luego $\text{Ker } A = \{0\}$.

Entonces, lo siguiente que uno podría preguntarse es si para cualquier cadena de Markov ergódica, A es un endomorfismo core-nilpotente con $\text{Ind}(A) =$

1. Estudiar esto se va de los objetivos del trabajo, por lo que nos restringiremos a las cadenas de Markov donde A sea core-nilpotente con $\text{Ind}(A) = 1$.

En adelante, a menos que se diga lo contrario, se hablará únicamente de cadenas de Markov infinitas, homogéneas y ergódicas tales que $A = \text{Id} - \mathcal{P}$ sea un endomorfismo CN con $\text{Ind}(A) = 1$.

Bajo estas hipótesis, tenemos que $\ell^1(\mathbb{R}) = \text{Im } A \oplus \text{Ker } A$ y que $A \cdot A^D$ es la proyección sobre $\text{Im } A$ y $\text{Id} - A \cdot A^D$, la proyección sobre $\text{Ker } A$.

Proposición 10. *Sea una cadena de Markov cumpliendo lo de arriba y $\mathcal{P}$ su operador de transición. Para $A = \text{Id} - \mathcal{P}$, se verifica que*

$$\text{Id} - A \cdot A^D = \lim_{n \rightarrow \infty} \mathcal{P}^n$$

Demostración. Sea π la distribución estacionaria. Sabemos que si $d \in \ell(\mathbb{R})$ es una distribución, entonces

$$\lim_{n \rightarrow \infty} \mathcal{P}^n(d) = \pi.$$

Consideremos $A(d) \in \text{Im } A$, se tiene que

$$\lim_{n \rightarrow \infty} \mathcal{P}^n(A(d)) = \lim_{n \rightarrow \infty} (\mathcal{P}^n(d) - \mathcal{P}^{n+1}(d)) = s\pi - s\pi = 0,$$

donde $s = \sum_{i \in \mathbb{N}} d_i$ y $d = (d_i)_{i \in \mathbb{N}}$.

Por otro lado, si $d \in \text{Ker } A$,

$$\lim_{n \rightarrow \infty} \mathcal{P}^n(d) = \lim_{n \rightarrow \infty} d = d.$$

Concluimos que $\lim_{n \rightarrow \infty} \mathcal{P}^n$ es la proyección sobre $\text{Ker } A$ y, por tanto, que coincide con $\text{Id} - A \cdot A^D$. $\square$

Para el resto de la sección y en adelante, denotaremos por W a $\text{Id} - A \cdot A^D$ con la finalidad de simplificar la notación.

Proposición 11. *Con las notaciones anteriores, se tiene que*

$$A^D = \sum_{n=0}^{\infty} (\mathcal{P}^n - W) \tag{29}$$

Demostración. En efecto, sea $d \in \text{Ker } A$, entonces,

$$\sum_{i=0}^{\infty} (\mathcal{P}^i - W)(d) = \sum_{i=0}^{\infty} (\mathcal{P}^i(d) - W(d)) = \sum_{i=0}^{\infty} (d - d) = 0.$$

Por otro lado, si $A(d) \in \text{Im } A$, se tiene que

$$\sum_{n=0}^{\infty} (\mathcal{P}^n - W)(A(d)) = \sum_{n=0}^{\infty} (\mathcal{P}^n(d) - \mathcal{P}^{n+1}(d))$$

y

$$\lim_{r \rightarrow \infty} \sum_{n=0}^r (\mathcal{P}^n - W) = \lim_{r \rightarrow \infty} (d - \mathcal{P}^{r+1}) = d - W(d).$$

Obsérvese que $A(d - W(d)) = A(d)$ y que $d - W(d) \in \text{Im } A$, concluimos por tanto que $A^D(A(d)) = d - W(d)$. $\square$

Es posible que calcular explícitamente A^D sea muy difícil o incluso imposible. Por ello, esta serie puede ser muy útil para la computación de A^D , especialmente en algoritmos iterativos para aproximar la inversa de Drazin.

El siguiente teorema proporciona uno de los resultados más interesantes de este trabajo.

Definición 18. Sea una cadena de Markov ergódica arbitraria. Sea $N_{i,j}^{(n)}$ el número de veces esperado que la cadena pasa por el estado j en n etapas, habiendo empezado en el estado i . A la matriz $N^{(n)} = (N_{i,j}^{(n)})$ se denomina matriz de recurrencia en n etapas de la cadena de Markov.

Teorema 14. Dada una cadena de Markov homogénea y ergódica, donde $\mathcal{P}$ es su operador de transición, $A = \text{Id} - \mathcal{P}$ verifica que $\text{Ind}(A) = 1$ y $N^{(n)}$ denota a su matriz de recurrencia en n etapas, se tiene que

$$A^D = \lim_{n \rightarrow \infty} (N^{(n)} - (n+1)W).$$

Demostración. Se sabe que

$$N^{(n)} = \sum_{r=0}^n \mathcal{P}^r. \quad (30)$$

Entonces, aplicando la Proposición 11,

$$\lim_{n \rightarrow \infty} (N^{(n)} - (n+1)W) = \lim_{n \rightarrow \infty} \left(\sum_{r=0}^n (\mathcal{P}^r - W) \right) = A^D.$$

$\square$

A partir de este teorema podemos concluir que, si n es suficientemente alto, podemos aproximar la matriz de recurrencia a través de la inversa de Drazin

$$N^{(n)} \approx A^D + (n+1)W. \quad (31)$$

Si la cadena de Markov fuese finita, $N^{(n)}$ sería muy fácil de calcular mediante la expresión (30). Sin embargo, en el caso de que $\mathcal{P}$ sea una matriz infinita, no es tan inmediato calcular $\mathcal{P}^n$. Por ello, si se sabe calcular A^D , conocer $N^{(n)}$ es bastante sencillo a partir de este teorema, siempre y cuando n sea muy grande.

Se ha hablado del número de veces que pasa la cadena por un determinado estado, ahora se hablará del tiempo que tarda hasta alcanzarlo por primera vez. Sea $m_{i,j}$ el tiempo esperado que la cadena tarda en alcanzar el estado j por primera vez, habiendo empezado en el estado i . Es posible también calcular estos coeficientes mediante A^D .

Teorema 15. *En las hipótesis del teorema anterior, sea $m_{i,j}$ el tiempo esperado que la cadena tarda en alcanzar el estado j por primera vez, habiendo empezado en el estado i . Entonces se tiene que*

$$m_{i,j} = \frac{1}{\pi_j} (\delta_{i,j} + A^D(\mu^j - \mu^i)_j), \quad (32)$$

donde $\pi = (\pi_j)_{j \in \mathbb{N}}$ es la distribución estacionaria de la cadena, $\delta_{i,j}$ denota a la delta de Kronecker, $\mu^i = (\delta_{i,j})_{j \in \mathbb{N}}$ y $A^D(\mu^i)_j$ es la coordenada j -ésima de $A^D(\mu^i)$.

Demostración. Sabemos que se verifica la ecuación

$$m_{i,j} = 1 + \sum_{r \in \mathbb{N}} p_{i,r} m_{r,j} - p_{i,j} m_{j,j}. \quad (33)$$

Tenemos que comprobar que la expresión del enunciado verifica (33).

$$\begin{aligned} & 1 + \frac{1}{\pi_j} \sum_{r \in \mathbb{N}} p_{i,r} (\delta_{r,j} + A^D(\mu^j - \mu^r)_j) - p_{i,j} \frac{1}{\pi_j} \\ &= 1 + \frac{1}{\pi_j} p_{i,j} + \frac{1}{\pi_j} A^D(\mu^j)_j \sum_{r \in \mathbb{N}} p_{i,r} - \frac{1}{\pi_j} \sum_{r \in \mathbb{N}} p_{i,r} A^D(\mu^r)_j - p_{i,j} \frac{1}{\pi_j} \\ &= 1 + \frac{1}{\pi_j} A^D(\mu^j)_j - \frac{1}{\pi_j} \sum_{r \in \mathbb{N}} p_{i,r} A^D(\mu^r)_j. \end{aligned} \quad (34)$$

Obsérvese que el sumatorio $\sum_{r \in \mathbb{N}} p_{i,r} A^D(\mu^r)_j$ es el producto de la fila i -ésima de la matriz $\mathcal{P}$ con la columna j -ésima de A^D , es decir, el elemento (i, j)

de la matriz $\mathcal{P} \cdot A^D$, o equivalentemente, $A^D(\mathcal{P}(\mu^i))_j$. Además, recordando que $A = \text{Id} - \mathcal{P}$ y que $W = \text{Id} - A \cdot A^D$,

$$A^D(\mathcal{P}(\mu^i)) = A^D(\mu^i) - A^D(A(\mu^i)) = A^D(\mu^i) + W(\mu^i) - \mu^i. \quad (35)$$

Sustituyendo en (34) nos queda

$$\begin{aligned} & 1 + \frac{1}{\pi_j} A^D(\mu^j)_j - \frac{1}{\pi_j} A^D(\mu^i)_j - 1 + \frac{1}{\pi_j} \delta_{i,j} \\ &= \frac{1}{\pi_j} (\delta_{i,j} + A^D(\mu^j - \mu^i)_j), \end{aligned} \quad (36)$$

como queríamos probar. $\square$

Obsérvese que si $i = j$, entonces $m_{i,i} = \frac{1}{\pi_i}$, que se sabe que coincide con el tiempo esperado que tarda la cadena en volver al estado i .

Es fascinante que algo tan complejo de calcular como $m_{i,j}$, pues hay que resolver un sistema de ecuaciones infinito, se reduzca a una cuenta tan simple usando la inversa de Drazin.

4. Aplicación de los resultados al ejemplo

Ahora que se tiene demostrado los resultados teóricos, en esta sección se van a aplicar a un caso concreto. En la Sección 2 se consideró una cadena de Markov infinita ergódica y homogénea que usaremos de ejemplo.

Lo primero es comprobar que estamos en las hipótesis de las Sección 3. Después, se hará el cálculo de la inversa de Drazin correspondiente. Finalmente, se calculará la matriz de recurrencia y el tiempo medio hasta alcanzar un cierto estado.

Durante toda la sección, consideraremos que $\{X_n\}_{n \geq 0}$ es la cadena de Markov de la Sección 2, utilizando $0 < p < \frac{1}{2}$, las probabilidades de transición son

$$p_{i,j} = \begin{cases} p & \text{si } j = i + 1 \\ 1 - p & \text{si } j = i - 1 \text{ o si } j = i = 0 \\ 0 & \text{en otro caso} \end{cases} \quad (37)$$

y $\pi = (\pi_i)_{i \geq 0}$ donde

$$\pi_i = \frac{1 - 2p}{1 - p} \left(\frac{p}{1 - p} \right)^i, \quad (38)$$

es la distribución estacionaria. Denotaremos por $\mathcal{P} : \ell^1(\mathbb{R}) \longrightarrow \ell^1(\mathbb{R})$ al operador de transición y sea $A = \text{Id} - \mathcal{P}$, A^D su inversa de Drazin y $W = \text{Id} - A \cdot A^D$. En forma matricial tenemos

$$\mathcal{P} = \begin{pmatrix} 1-p & p & 0 & 0 & \cdots \\ 1-p & 0 & p & 0 & \cdots \\ 0 & 1-p & 0 & p & \cdots \\ 0 & 0 & 1-p & 0 & \ddots \\ \vdots & \vdots & \vdots & \ddots & \ddots \end{pmatrix}, \quad A = \begin{pmatrix} p & -p & 0 & 0 & \cdots \\ p-1 & 1 & -p & 0 & \cdots \\ 0 & p-1 & 1 & -p & \cdots \\ 0 & 0 & p-1 & 1 & \ddots \\ \vdots & \vdots & \vdots & \ddots & \ddots \end{pmatrix}.$$

4.1. Verificación de las hipótesis

Para poder aplicar la teoría, hay que comprobar que la cadena de Markov cumple con las hipótesis que se comentaron en la Sección 3. Sabemos que es una cadena de Markov homogénea y ergódica. Veamos que A es un endomorfismo core-nilpotente y que $\text{Ind}(A) = 1$.

Teorema 16. *En esta situación se verifica que*

$$\ell^1(\mathbb{R}) = \text{Im } A \oplus \text{Ker } A.$$

Demostración. En la Sección 2 se ha demostrado que $\text{Ker } A = \langle \pi \rangle$. Supongamos por reducción al absurdo que $\text{Im } A \cap \text{Ker } A \neq \{0\}$. Entonces, el sistema de ecuaciones siguiente tiene solución

$$p d_0 + (p-1) d_1 = 1 \tag{39}$$

$$-p d_{i-1} + d_i + (p-1) d_{i+1} = \left(\frac{p}{1-p} \right)^i \quad i > 0 \tag{40}$$

para $(d_i)_{i \geq 0} \in \ell^1(\mathbb{R})$. Sumando las ecuaciones (39) y (40) con $i = 1$ obtenemos

$$p d_1 + (p-1) d_2 = 1 + \frac{p}{1-p}. \tag{41}$$

Si ahora, a la ecuación (41) le sumamos la ecuación (40) con $i = 2$ queda

$$p d_2 + (p-1) d_3 = 1 + \frac{p}{1-p} + \left(\frac{p}{1-p} \right)^2. \tag{42}$$

En general, procediendo por inducción, se puede demostrar que para cualquier $i \geq 0$ se verifica

$$p d_i + (p-1) d_{i+1} = \sum_{k=0}^i \left(\frac{p}{1-p} \right)^k. \tag{43}$$

Dado que $\sum d_i$ converge, la sucesión $(d_i)_{i \geq 0}$ tiende a 0, concluyendo que el miembro izquierdo de (43) también tiende a 0. Sin embargo, el miembro derecho converge a $\frac{1-p}{1-2p}$, lo que es una contradicción que aparece al suponer que $\text{Im } A \cap \text{Ker } A \neq \{0\}$.

Sea $(\lambda_i)_{i \geq 0} \in \ell^1(\mathbb{R})$. Veamos si el siguiente sistema de ecuaciones tiene solución.

$$p d_0 + (p-1) d_1 + \Lambda = \lambda_0 \quad (44)$$

$$-p d_{i-1} + d_i + (p-1) d_{i+1} + \Lambda \left(\frac{p}{1-p} \right)^i = \lambda_i \quad i > 0 \quad (45)$$

para $\Lambda \in \mathbb{R}$ y $(d_i)_{i \geq 0} \in \ell^1(\mathbb{R})$. Con un argumento similar al anterior se llega a la expresión

$$p d_i + (p-1) d_{i+1} + \Lambda \sum_{k=0}^i \left(\frac{p}{1-p} \right)^k = \sum_{k=0}^i \lambda_k \quad \forall i \geq 0. \quad (46)$$

y, despejando, se tiene que

$$\Lambda = \frac{1}{\sum_{k=0}^i \left(\frac{p}{1-p} \right)^k} \left(\sum_{k=0}^i \lambda_k - p d_i - (p-1) d_{i+1} \right). \quad (47)$$

Por esto, la sucesión definida por el miembro derecho de (47) ha de ser constante y, además, debe de coincidir con su límite, luego

$$\Lambda = \frac{1-2p}{1-p} \sum_{k \geq 0} \lambda_k. \quad (48)$$

Finalmente, a partir de (46), conociendo d_0 , la sucesión $(d_i)_{i \geq 0}$ definida por la siguiente fórmula de recurrencia

$$d_{i+1} = \frac{1}{p-1} \left(\sum_{k=0}^i \lambda_k - \Lambda \sum_{k=0}^i \left(\frac{p}{1-p} \right)^k - p d_i \right) \quad \forall i > 0, \quad (49)$$

donde Λ viene dado por (48), verifica las ecuaciones (44) y (45). Para acabar la demostración, veamos que $(d_i)_{i \geq 0} \in \ell^1(\mathbb{R})$.

Como se tiene que

$$\lim_{i \rightarrow \infty} \frac{|d_{i+1}|}{|d_i|} = \lim_{i \rightarrow \infty} \frac{\left| \sum_{k=0}^i \lambda_k - \Lambda \sum_{k=0}^i \left(\frac{p}{1-p} \right)^k - p d_i \right|}{|d_i| |p-1|} = \frac{p}{1-p} < 1, \quad (50)$$

por el criterio del cociente, la serie $\sum |d_i|$ converge. Se concluye, por tanto, que $\text{Im } A + \text{Ker } A = \ell^1(\mathbb{R})$. $\square$

Por este teorema y por el Teorema 12, se demuestra que A es un endomorfismo CN y que $\text{Ind}(A) = 1$. Entonces, se tienen las condiciones para poder aplicar la teoría de la inversa de Drazin.

4.2. Cálculo de la inversa de Drazin

Naturalmente, para poder utilizar y aplicar los resultados de la Sección 3, es necesario calcular antes A^D .

Sea $(\lambda_i)_{i \geq 0} \in \ell^1(\mathbb{R})$. Se ha demostrado en el Teorema 16 que existe una única descomposición

$$(\lambda_i)_{i \geq 0} = A(d) + a\pi,$$

donde $d = (d_i)_{i \geq 0} \in \ell^1(\mathbb{R})$ y $a \in \mathbb{R}$. Además, como se sabe que $A \cdot A^D = \text{Id} - W$

$$A^D((\lambda_i)_{i \geq 0}) = A^D(A(d) + a\pi) = A^D(A(d)) = d - W(d). \quad (51)$$

En la demostración del Teorema 16, se determina d a partir de la fórmula de recurrencia

$$d_{i+1} = \frac{1}{p-1} \left(\sum_{r=0}^i \lambda_r - \bar{\lambda} \frac{1-2p}{1-p} \sum_{r=0}^i \left(\frac{p}{1-p} \right)^r - p d_i \right) \quad \forall i > 0, \quad (52)$$

donde $\bar{\lambda} = \sum \lambda_r$. Realizando los cálculos correspondientes se obtiene la siguiente expresión explícita

$$\begin{aligned} d_i = & \frac{1}{p-1} \left[\sum_{r=0}^{i-1} \lambda_r \sum_{s=0}^{i-1-r} \left(\frac{p}{1-p} \right)^s \right. \\ & \left. - \bar{\lambda} \frac{1-2p}{1-p} \sum_{r=0}^{i-1} (r+1) \left(\frac{p}{1-p} \right)^r \right] + \left(\frac{p}{1-p} \right)^i d_0. \end{aligned} \quad (53)$$

Para simplificar la notación, definimos

$$\bar{d}_i = \frac{1}{p-1} \left[\sum_{r=0}^{i-1} \lambda_r \sum_{s=0}^{i-1-r} \left(\frac{p}{1-p} \right)^s - \bar{\lambda} \frac{1-2p}{1-p} \sum_{r=0}^{i-1} (r+1) \left(\frac{p}{1-p} \right)^r \right]. \quad (54)$$

Por otro lado, es fácil de ver que $W(d) = \sum_{r \geq 0} d_r \pi$ debido a la Proposición 10.

Si la coordenada j -ésima de $W(d)$ coincide con

$$W(d)_j = \left(\sum_{r \geq 0} \bar{d}_r + d_0 \frac{1-p}{1-2p} \right) \frac{1-2p}{1-p} \left(\frac{p}{1-p} \right)^j,$$

entonces se concluye que la coordenada j -ésima de $A^D((\lambda_r)_{r \geq 0})$ es

$$A^D((\lambda_r)_{r \geq 0})_j = \bar{d}_j - \frac{1-2p}{1-p} \sum_{r \leq 0} \bar{d}_r \left(\frac{p}{1-p} \right)^j. \quad (55)$$

Obsérvese que la expresión final no depende del d_0 elegido, como ha de ser. Recuérdesse que $A^D((\lambda_r)_{r \geq 0})$ no puede depender ni ser diferente en función de la sucesión d escogida para calcularlo.

En el caso concreto de tener $\mu^i = (\delta_{i,j})_{j \geq 0}$, se tiene que

$$A^D(\mu^i)_j = \overline{d^{(i)}}_j - \frac{1-2p}{1-p} \sum_{r \leq 0} \overline{d^{(i)}}_r \left(\frac{p}{1-p} \right)^j \quad (56)$$

donde

$$\overline{d^{(i)}}_j = \begin{cases} \frac{1-2p}{(1-p)^2} \sum_{r=0}^{j-1} (r+1) \left(\frac{p}{1-p} \right)^r & \text{if } j \leq i \\ \frac{1}{p-1} \sum_{r=0}^{j-i-1} \left(\frac{p}{1-p} \right)^r + \frac{1-2p}{(1-p)^2} \sum_{r=0}^{j-1} (r+1) \left(\frac{p}{1-p} \right)^r & \text{if } j > i \end{cases}$$

Si usamos el convenio de que $\sum_{r=0}^j d_i = 0$ si $j < 0$, entonces la fórmula anterior se simplifica a

$$\overline{d^{(i)}}_j = \frac{1}{p-1} \sum_{r=0}^{j-i-1} \left(\frac{p}{1-p} \right)^r + \frac{1-2p}{(1-p)^2} \sum_{r=0}^{j-1} (r+1) \left(\frac{p}{1-p} \right)^r. \quad (57)$$

4.3. Cálculo de los resultados con la inversa de Drazin

El objetivo de este apartado es calcular los valores de $N_{i,j}^{(n)}$ y $m_{i,j}$, definidos en la sección anterior, a partir de los resultados de los Teoremas 14 y 15.

En primer lugar, si n es suficientemente alto, se puede aproximar $N^{(n)}$ mediante

$$N^{(n)} \approx A^D + (n+1)W.$$

Si, por ejemplo, queremos conocer el número de veces que la cadena pasa por el estado 0 en n etapas, habiendo empezado desde otro estado i arbitrario, debemos calcular la coordenada 0-ésima de la sucesión $N^{(n)}(\mu^i)$,

$$N_{i,0}^{(n)} \approx A^D(\mu^i)_0 + (n+1)W(\mu^i)_0. \quad (58)$$

Como $W(\mu^i) = \pi$ y $\overline{d^{(i)}}_0 = 0$ para cualquier $i \geq 0$, se tiene que

$$N_{i,0}^n \approx -\frac{1-2p}{1-p} \sum_{r \leq 0} \overline{d^{(i)}}_r + (n+1) \frac{1-2p}{1-p}. \quad (59)$$

Una vez calculado el número esperado de veces que pasa el sistema por el 0, calculemos el tiempo esperado que tarda en llegar al él.

$$m_{i,0} = \frac{1}{\pi_0} (\delta_{i,0} + A^D(\mu^0)_0 - A^D(\mu^i)_0). \quad (60)$$

Si $i > 0$, tenemos entonces que

$$\begin{aligned} m_{i,0} &= \frac{1-p}{1-2p} \left(-\frac{1-2p}{1-p} \sum_{r \leq 0} \overline{d^{(0)}}_r + \frac{1-2p}{1-p} \sum_{r \leq 0} \overline{d^{(i)}}_r \right) \\ &= \sum_{r \leq 0} \overline{d^{(i)}}_r - \sum_{r \leq 0} \overline{d^{(0)}}_r, \end{aligned} \quad (61)$$

y si $i = 0$, tenemos

$$m_{0,0} = \frac{1}{\pi_0} = \frac{1-p}{1-2p}.$$

En general, si $i \neq j$ se tiene

$$m_{i,j} = \frac{1-p}{1-2p} \left(\frac{1-p}{p} \right)^j \left(\overline{d^{(j)}}_j - \overline{d^{(i)}}_j \right) + \sum_{r \leq 0} \left(\overline{d^{(i)}}_r - \overline{d^{(j)}}_r \right), \quad (62)$$

y, en caso contrario,

$$m_{i,i} = \frac{1-p}{1-2p} \left(\frac{1-p}{p} \right)^i. \quad (63)$$

5. Simulaciones con Mathematica

En último lugar, usaremos el software *Wolfram Mathematica* para realizar algunas cómputos y verificar experimentalmente que los cálculos teóricos funcionan.

La primera celda (Figure 1) consta de la definición de una función para calcular los coeficientes $\overline{d^{(i)}}_j$. Como argumentos de entrada recibe, aparte de i y j , la probabilidad p .

Después, usando las expresiones (59) y (61), se calcula el número esperado de veces que la cadena pasa por el estado 0 en 1000 etapas y el tiempo

$$\text{In}[1]:= d[p_ , i_ , j_] := \frac{1}{p-1} \sum_{k=0}^{j-i-1} \left(\frac{p}{1-p} \right)^k + \frac{1-2p}{(1-p)^2} \sum_{k=0}^{j-1} (k+1) \left(\frac{p}{1-p} \right)^k$$

Figura 1: Función para computar $\overline{d^{(i)}}_j$.

esperado hasta llegar a él por primera vez, cuando la cadena se sitúa inicialmente en el estado 41, es decir, los valores de $N_{41,0}^{1000}$ y $m_{41,0}$ respectivamente (Figura 2). Los resultados que se obtienen, si $p = 0,1$, son $N_{41,0}^{1000} = 844,361$ y $m_{41,0} = 51,25$.

```
In[12]:= p = 0.10;
i = 41;
n = 1000;
jmax = 2000;
(*Expected number of times the M.c. enters in 0th state.*)
- (1 - 2 p) / (1 - p) * Sum[d[p, i, j], {j, 0, jmax}] + (n + 1) * (1 - 2 p) / (1 - p)
(*Expected time to reach 0th state.*)
Sum[d[p, i, j] - d[p, 0, j], {j, 0, jmax}]

Out[16]= 844.361

Out[17]= 51.25
```

Figura 2: Cálculo teórico de $N_{41,0}^{1000}$ y $m_{41,0}$. Se ha utilizado $p = 0,1$.

Para reducir los tiempos de cómputo, las series $\sum_{r \geq 0} \overline{d^{(i)}}_r$ han sido aproximadas por una suma parcial.

El código que se muestra en la Figura 3 pretende aproximar $N_{41,0}^{1000}$ experimentalmente. Con él se simula 1000 etapas de la cadena de Markov empezando en el estado 41. El programa cuenta cuantas veces la cadena se sitúa en el 0. Después repite el experimento 2000 veces y calcula la media de los resultados. Se puede observar que el resultado obtenido es prácticamente idéntico al obtenido a priori, pues coinciden al aproximar por un número entero.

De la misma manera y mediante un proceso análogo, se aproxima $m_{41,0}$ de manera experimental. El código puede verse en la Figura 4. En este caso, lo que se cuenta es el número de etapas que tarda en alcanzar el estado 0.

```
In[18]= (*Experimental computation of expected number of times
the M.c. enters in 0th state.*)
meanRecurrenceTimes = {};
Do[
  in = i;
  count = 0;
  Do[
    in = If[in == 0, in + RandomChoice[{1 - p, p} -> {0, 1}],
      in + RandomChoice[{1 - p, p} -> {-1, 1}]];
    If[in == 0, count++];
  , n];
  AppendTo[meanRecurrenceTimes, count];
, 2000]
Mean[meanRecurrenceTimes] // N
Out[20]= 844.426
```

Figura 3: Cálculo experimental para $N_{41,0}^{1000}$.

```
(*Experimetal computation of expected time to reach 0
th state*)
meanPassageTimes = {};
Do[
  in = i;
  time = 0;
  While[in != 0,
    in = in + RandomChoice[{1 - p, p} -> {-1, 1}];
    time++;
  ];
  AppendTo[meanPassageTimes, time];
, 2000]
Mean[meanPassageTimes] // N
Out[23]= 51.292
```

Figura 4: Cálculo experimental para $m_{41,0}$.

Con esto se ve que la teoría funciona y, que operando en el ámbito de los endomorfismos CN, todos los resultados conocidos de la aplicación de la inversa de Drazin a las cadenas de Markov finitas se pueden usar en el caso infinito. Los experimentos realizados han confirmado que las propiedades útiles de la inversa de Drazin se mantienen intactas cuando se extienden a cadenas de Markov infinitas. Este hallazgo es particularmente relevante porque permite la generalización de técnicas y métodos previamente aplicados

en cadenas de Markov finitas, ampliando así el alcance y la aplicabilidad de la teoría de Markov.

En resumen, la investigación ha demostrado que los conceptos y herramientas teóricas desarrollados son sólidos y aplicables. La validación experimental ha sido crucial para confirmar que los resultados teóricos no solo son correctos, sino también útiles y prácticos para aplicaciones reales. Esto establece una base sólida para futuras investigaciones y aplicaciones en la teoría de cadenas de Markov infinitas.

Referencias

- [1] M. P. Drazin, «Pseudo-inverses in associative rings and semigroups,» *The American Mathematical Monthly*, vol. 64, n.º 7, págs. 506-514, 1958.
- [2] G. R. Grimmett y D. Stirzaker, *Probability and Random Processes. Fourth Edition*. Oxford University Press, 2020.
- [3] J. G. Kemeny, J. L. Snell y A. W. Knapp, *Denumerable Markov Chains*. Springer-Verlag, 1976.
- [4] D. A. Levin e Y. Peres, *Markov Chains and Mixing Times, second edition*. American Mathematical Society, 2009.
- [5] C. Grinstead y J. Snell, *Introduction to Probability*. American Mathematical Society, 1997.
- [6] S. L. Campbell y C. D. Meyer, *Generalized Inverses of Linear Transformations*. Society for Industrial y Applied Mathematics, 2009.
- [7] F. Pablos Romo, «Core-Nilpotent Endomorphisms of Infinite-dimensional Vector Spaces,» *Mediterr. J. Math.*, vol. 18, 84, 2021.

Anexo

Artículo científico

Se adjunta a continuación el artículo científico que se ha elaborado a partir de este trabajo, el cual ya ha sido enviado a una revista de carácter internacional para su publicación.

AN APPROACH TO THE STUDY OF INFINITE MARKOV CHAINS FROM THE DRAZIN INVERSE OF A CORE-NILPOTENT OPERATOR

PABLO RODRÍGUEZ GARCÍA

ABSTRACT. The Drazin inverse is the most important psuedo-inveserse of lineal operators of finite-dimensional vector spaces. One of its applications is in finite Markov chains theory, where they are used to calculate certain properties such as the average number of times the chain passes through a particular state or the expected time it takes for the chain to enter it. Since it has been possible to generalize Drazin's inverse to linear operators in arbitrary dimension (Core-Nilpotent Operators), the aim of this paper is to use it to obtain the same results for infinite Markov chains. Futhermore, the results will be applied to a concrete example.

CONTENTS

1. Introduction	1
2. Preliminaries	2
2.1. Drazin inverses and Core-Nilpotent operators	2
2.2. Markov Chains	3
2.3. Aplication of the Drazin inverse to finite Markov chains	5
3. Aplication of Drazin Inverses to Infinite Markov Chains	6
4. Ilustrative Example	8
References	13

1. INTRODUCTION

It is well known that a linear operator may not be invertible, meaning its inverse operator does not exist. Therefore, it can be very useful to look for alternatives: given a linear operator, find another operator that, even if not invertible, plays the role of the inverse operator or, at least, has similar characteristics. For this reason, generalized inverses (or pseudo-inverses) were created: linear operators that have similarities with inverse operators. One of the most studied generalized inverses is the Drazin inverse, named in honor of Michael Peter Drazin for his work in 1958, [3].

The Drazin inverse is particularly useful in the analysis of dynamic systems, where operators are often constituted by transition rules. In particular, Markov chains model dynamic systems that are stochastic rather than deterministic. In this context, there are linear operators that provide information about the probability distribution at each stage.

It is not unreasonable to think that the Drazin inverse can be useful in obtaining certain information about the chain. Specifically, it allows us to approach systems

Key words and phrases. Core-nilpotent decomposition, linear map, infinite linear system, Drazin inverse, group inverse, infinite Markov chains, infinite random walks.

This work is partially supported by a collaboration-fellowship of the Spanish Goverment.

of linear equations associated with linear operators that are not invertible. For example, in the calculation of stationary distributions, long-term behaviors of the chain, recurrence, etc.

Another important aspect of Markov chains is the classification of states in terms of recurrence and transience. The Drazin inverse facilitates the study of these properties in finite Markov chains by providing another perspective on the transition matrix.

Throughout this work, we will explore in detail the formal definition of the Drazin inverse and its generalization to infinite-dimensional case, its fundamental properties, and its application in the theory of Markov chains.

We will start with a review of linear operators in spaces with arbitrary dimension and the definition of the Drazin inverse, followed by an introduction to Markov chains and the application of the pseudo-inverse to chains with a finite number of states. We will continue with the generalization of some results to infinite Markov chains and, finally, we will end with the application of these results to an illustrative example.

The main objective of this work is to see that, through a deeper understanding of the Drazin inverse, we can obtain relevant information from an arbitrary Markov chain. We hope that this exploration will serve as a motivation for future research about Drazin inverses and its applications in various disciplines that use linear operators and matrices.

2. PRELIMINARIES

2.1. Drazin inverses and Core-Nilpotent operators. Let k be a field and let V be a k -vector space. It is possible for a lineal operator to not have an inverse. For this reason, one can try to find another operator with similar properties to inverses. In this section, we will define a generalized inverse, called Drazin inverse, and study when its characterisation.

Definition 2.1. *A lineal operator $f : V \longrightarrow V$ is core-nilpotent (CN operator) when there exists two f -invariant subspaces, U_1 and U_2 , such that $f|_{U_1}$ is an automorphism, $f|_{U_2}$ is nilpotent and*

$$V = U_1 \oplus U_2.$$

The decomposition $V = U_1 \oplus U_2$ is called CN decomposition of f .

Definition 2.2. *A CN operator $f : V \longrightarrow V$ has index n , and it will be denoted as $\text{Ind}(f) = n$, when $\text{Ker}f^n = \text{Ker}f^{n+1}$ and $\text{Im}f^n = \text{Im}f^{n+1}$, but $\text{Ker}f^n \neq \text{Ker}f^{n-1}$ or $\text{Im}f^n \neq \text{Im}f^{n-1}$.*

Theorem 2.1. (Characterization of CN operators). *If $f : V \longrightarrow V$ is a lineal operator, then the following conditions are equivalent:*

- (1) f is CN operator.
- (2) $\text{Ker}f^n = \text{Ker}f^{n+1}$ and $\text{Im}f^n = \text{Im}f^{n+1}$ for a certain $n \in \mathbb{N}$.
- (3) $V = \text{Ker}f^n \oplus \text{Im}f^n$ for a certain $n \in \mathbb{N}$.
- (4) There exists two lineal operators $f_1, f_2 : V \longrightarrow V$ such that $f = f_1 + f_2$, $\text{Ind}(f_1) \leq 1$, f_2 is nilpotent and $f_1 \circ f_2 = f_2 \circ f_1 = 0$.

Proof. See Theorem 3.6 in [6]. □

As an immediate consequence of the previous theorem and with the previous notations

Corollary 2.1. *Given $f : V \longrightarrow V$ a lineal operator*

$$\text{Ind}(f) \leq 1 \text{ if and only if } V = \text{Ker}f \oplus \text{Im}f.$$

Proposition 2.1. *If V is finite-dimensional, then every lineal operator is CN.*

Proof. One can find a explanation in [6]. $\square$

Now, we will introduce the definition of the Drazin inverse.

Definition 2.3. *A lineal operator $f : V \longrightarrow V$ has a Drazin inverse when there exists $f^D : V \longrightarrow V$ such that*

- $f^{m+1} \circ f^D = f^m$ for $m \geq n$ where n is a certain non-negative integer,
- $f^D \circ f \circ f^D = f^D$,
- $f^D \circ f = f \circ f^D$.

Notice that if f is an isomorphism, then his Drazin inverse exists and coincides with his inverse.

Proposition 2.2. *A lineal operator $f : V \longrightarrow V$ is a CN operator if and only if there exists an unique Drazin inverse f^D where $n = \text{Ind}(f)$.*

Proof. We know that $V = U_1 \oplus U_2$ where $f|_{U_1}$ is an isomorphism. We define

$$(1) \quad f^D(u_1 + u_2) = (f|_{U_1})^{-1}(u_1).$$

It is easy to see that it verifies the previous definition. Conversely, one can take $f_1 = f \circ f^D \circ f$ and $f_2 = f - f \circ f^D \circ f$. It is easy to prove that they verify condition 4 of the Theorem 2.1. A detailed proof appear in [6]. $\square$

Proposition 2.3. *Let $f : V \longrightarrow V$ a CN operator and f^D his Drazin inverse. The following properties are satisfied:*

- $(f^D)^D = f$ if and only if $\text{Ind}(f) \leq 1$.
- $f \circ f^D = \text{Id}|_{U_1}$ and $\text{Id} - f \circ f^D = \text{Id}|_{U_2}$ where $V = U_1 \oplus U_2$ is the CN decomposition.
- $(f^m)^D = (f^D)^m$ for all $m \in \mathbb{N}$.

Proof. One can easily see all of them from (1). $\square$

2.2. Markov Chains. Let $\{X_i\}_{i \in \mathbb{N}}$ be a Markov chain and let E be the states set (finite o countable) of the chain. We denote by $p_{i,j}^{(n)}$ the probability to change to j -state from the i -estate in n -stage; and by $d^{(n)} = (d_i^{(n)})_{i \in E}$ the probability distribution of the chain. That is,

$$(2) \quad p_{i,j}^{(n)} = \mathcal{P}(X_{n+1} = j | X_n = i), \quad d_i^{(n)} = \mathcal{P}(X_n = i).$$

Because of the law of total probability, it is satisfied:

$$(3) \quad d_j^{(n+1)} = \sum_{i \in E} d_i^{(n)} p_{i,j}^{(n)},$$

or equivalently

$$(4) \quad d^{(n+1)} = d^{(n)} \cdot \mathcal{P}^{(n)},$$

where $\mathcal{P}^{(n)} = (p_{i,j}^{(n)})_{i,j \in E}$ is the transition matrix.

Definition 2.4. *A Markov chain is homogeneous when the probability of change state does not depend of the stage, that is*

$$p_{i,j}^{(n)} = p_{i,j}^{(0)} \quad \forall n \in \mathbb{N}.$$

Unless otherwise indicated, from here on we will always consider homogeneous Markov chains.

Definition 2.5. If E is a finite set, one says that a Markov chain is finite. On the contrary, one says that it is infinite.

In the next section, we will focus in a specific type of Markov chain. So, it is important to know the states and Markov chain classification.

But first, a definition will be introduced to aid in the states classification.

Definition 2.6. Let $e_1, e_2 \in E$ be two states of a Markov chain, it is said that e_1 communicates with e_2 , and it will be denoted by $e_1 \rightarrow e_2$, if it is possible to go from e_1 to e_2 in a finite number of steps. That is

$$\mathcal{P}(X_n = e_2 | X_0 = e_1) > 0 \quad \text{for some } n > 0,$$

and it is said that e_1 and e_2 intercommunicate, denoted by $e_1 \leftrightarrow e_2$, if $e_1 \rightarrow e_2$ and $e_2 \rightarrow e_1$.

The first and most important distinction between states is based on the system's ability to return to a specific state.

Definition 2.7. Let $e \in E$. We say that e is transient if it is not certain that, once the state exits, the chain will return to it. Otherwise, it is called recurrent.

Definition 2.8. Let $e \in E$ be a recurrent state:

- it is positive recurrent if the mean time it takes to return to it is finite. Otherwise, it is null recurrent.
- it is said to be periodic with period t if it is only possible to return to that state after a number of steps that is a multiple of t . In the case where $t = 1$, the state is said to be aperiodic.
- it is ergodic if it is positive recurrent and aperiodic.

Proposition 2.4. Given a Markov chain, and $e_1, e_2 \in E$ such that $e_1 \rightarrow e_2$. If e_1 is recurrent, e_2 is also recurrent and if e_2 is transient, e_1 too.

Proof. If $e_1 \rightarrow e_2$ and e_1 is recurrent, then the Markov chain is in state e_1 an infinite number of times and, due to $\mathcal{P}(X_n = e_2 | X_0 = e_1) > 0$, the chain returns to e_2 with absolute certainty.

Now, suppose that e_2 is transient. If e_1 was recurrent, we just prove that e_2 is recurrent too, but this leads a contradiction. $\square$

Corollary 2.2. If two states intercommunicate, they are of the same type (positive recurrent, null recurrent, aperiodic or transient).

During this work, we will only considerate ergodic Markov chains. Let us define this concept.

Definition 2.9. A Markov chain is said to be ergodic if all states are ergodic and they intercommunicate with each other.

Theorem 2.2. Let us consider a ergodic Markov chain. Then there is a unique stationary distribution, that is, there exists a unique $\pi = (\pi_i)_{i \in E}$ where $0 \leq \pi_i \leq 1$ and $\sum \pi_i = 1$, such that

$$(5) \quad \pi = \pi \cdot \mathcal{P},$$

and, in that case,

$$(6) \quad \pi = \lim_{n \rightarrow \infty} d^{(0)} \cdot \mathcal{P}^n$$

for all initial probability distribution $d^{(0)}$.

Proof. One can find the demonstration in Section 1.5 of [5]. $\square$

There is a similar theorem, but more useful in some later proofs.

Theorem 2.3. *Given a Markov chain whose states all intercommunicate with each other, there exists a stationary distribution π if and only if there is a positive recurrent state. In that case, all states are positive recurrent and π is unique.*

Proof. See Section 6 of [4]. $\square$

2.3. Application of the Drazin inverse to finite Markov chains. Let $\mathcal{P}$ a transition matrix of a Markov chain with m states. Then, for $A = \text{Id} - \mathcal{P}$, $\text{Ind}(A) = 1$, as demonstrated by Campbell in Section 8 of [2].

Theorem 2.4. *Let us consider an ergodic Markov chain and A like before. One has*

$$(7) \quad \text{Id} - A A^D = \lim_{n \rightarrow \infty} \mathcal{P}^n,$$

where A^D is the Drazin inverse of A .

Proof. See Theorem 8.2.2 in [1]. $\square$

Henceforth, we will write $W = \text{Id} - A A^D$. It is easier to calculate W than A^D , because each row of W is the stationary distribution. For this reason, the next result is useful.

Theorem 2.5. *With the notation of above, one has that*

$$(8) \quad A^D = \sum_{n=0}^{\infty} (\mathcal{P}^n - W).$$

Proof. See Theorem 8.3.1 in [1]. $\square$

Another characterization for A^D using W is given in the next theorem.

Theorem 2.6. *For $A = \text{Id} - \mathcal{P}$, A^D is the unique solution of the equations $W X = 0$ and $A X = I - W$.*

Proof. See Theorem 8.5.5 in [1]. $\square$

Theorem 2.7. *Let $N^{(n)}$ be a matrix whose (i, j) th entry is the expected number of times the chain is in j th state in n stages when the chain was initially in i th state. Then*

$$(9) \quad A^D = \lim_{n \rightarrow \infty} (N^{(n)} - (n+1)W).$$

Proof. See Theorem 8.3.2 in [1]. $\square$

Theorem 2.8. *Let M be the matrix whose (i, j) th entry is the number of steps before entering j th state for first time from i th state. Then,*

$$(10) \quad M = (\text{Id} - A^D + J A_d^D) D,$$

where J is a matrix of all 1's, A_d is the diagonal matrix of A , obtained setting all off diagonal entries of A equal to zero and $D = (W_d)^{-1}$

Proof. See Theorem 8.4.1 in [1]. $\square$

3. APPLICATION OF DRAZIN INVERSES TO INFINITE MARKOV CHAINS

Throughout this section, we will consider a homogeneous Markov chain $\{X_n\}_{n \geq 0}$. We will assume that the state space is $E = \mathbb{N}$ (whether $0 \in \mathbb{N}$ or $0 \notin \mathbb{N}$ will depend on the context in which the Markov chain is defined).

Let $V = \left\{ (d_i)_{i \in \mathbb{N}} \in \prod_{i \in \mathbb{N}} \mathbb{R} : \sum_{i \in \mathbb{N}} |d_i| < \infty \right\}$. There exists an operator $\mathcal{P} : V \rightarrow V$ called the transition operator, defined as

$$(11) \quad \mathcal{P}((d_i)_{i \in \mathbb{N}}) = \left(\sum_{i \in \mathbb{N}} d_i p_{i,j} \right)_{j \in \mathbb{N}}.$$

Lemma 3.1. *The transition operator defined as (11) is well defined.*

Proof. Since $p_{i,j} \leq 1$, one has

$$\sum_{i \in \mathbb{N}} d_i p_{i,j} \leq \sum_{i \in \mathbb{N}} d_i < \infty.$$

On the other hand,

$$\sum_{j \in \mathbb{N}} \left| \sum_{i \in \mathbb{N}} d_i p_{i,j} \right| \leq \sum_{j \in \mathbb{N}} \sum_{i \in \mathbb{N}} |d_i p_{i,j}| = \sum_{i \in \mathbb{N}} \sum_{j \in \mathbb{N}} |d_i| p_{i,j} = \sum_{i \in \mathbb{N}} |d_i| < \infty$$

□

We will call $A = \text{Id} - \mathcal{P}$. We will restrict ourselves to the case where the chain is ergodic and where A is a CN operator of index 1. Then $V = \text{Ker } A \oplus \text{Im } A$ such that A over $\text{Im } A$ is an isomorphism.

Furthermore, since the chain is ergodic, there exists a unique stationary distribution $\pi = (\pi_i)_{i \in \mathbb{N}}$.

In this section, we will attempt to generalize the existing results for finite Markov chains to arbitrary Markov chains using the theory of CN operators.

The first result deals with the relationship between the stationary distribution and A^D .

Theorem 3.1. *Let $\{X_n\}_{n \geq 0}$ be a homogeneous ergodic Markov chain and let $\mathcal{P}$ his transition operator. For $A = \text{Id} - \mathcal{P}$*

$$(12) \quad \text{Id} - A A^D = \lim_{n \rightarrow \infty} \mathcal{P}^n.$$

Proof. First, we know that if $\sum_{i \in \mathbb{N}} d_i = 1$, $\lim_{n \rightarrow \infty} \mathcal{P}^n(d)$ exists and coincides with π .

By proposition 2.3, $\text{Id} - A A^D$ vanishes over $\text{Im } A$ and it is the identity over $\text{Ker } A$. Let us see that the limit in the right side of the equation (12) verifies the same. Let $v \in \text{Ker } A$, then $\mathcal{P}(v) = v$, so

$$(13) \quad \lim_{n \rightarrow \infty} \mathcal{P}^n(v) = \lim_{n \rightarrow \infty} v = v.$$

Now, let $A(d) \in \text{Im } A$, then

$$(14) \quad \lim_{n \rightarrow \infty} \mathcal{P}^n(\text{Id} - \mathcal{P})(d) = \lim_{n \rightarrow \infty} (\mathcal{P}^n - \mathcal{P}^{n+1})(d) = s\pi - s\pi = 0,$$

where $s = \sum_{i \in \mathbb{N}} d_i$.

□

This result is significant because it connects the long-term behavior of the Markov chain, as described by the limit of $\mathcal{P}^n$, with the Drazin inverse.

For the rest of the section, we will denote by W the projection onto $\text{Ker } A$, $W = I - A A^D$.

Theorem 3.2. *Given an ergodic arbitrary Markov chain, it holds that*

$$(15) \quad A^D = \sum_{k=0}^{\infty} (\mathcal{P}^k - W).$$

Proof. Let $v \in \text{Ker} A$, then

$$(16) \quad \sum_{k=0}^{\infty} (\mathcal{P}^k - W)(v) = \sum_{k=0}^{\infty} (v - v) = 0.$$

On the other hand, if $A(d) \in \text{Im} A$,

$$(17) \quad \begin{aligned} \sum_{k=0}^{\infty} (\mathcal{P}^k - W)(A(d)) &= \sum_{k=0}^{\infty} (\mathcal{P}^k(d) - \mathcal{P}^{k+1}(d)) \\ &= \lim_{k \rightarrow \infty} (d - \mathcal{P}^{k+1}(d)) = d - W(d) \end{aligned}$$

Note that $A(d - W(d)) = A(d)$ and that $d - W(d) \in \text{Im} A$, we conclude therefore that $A^D(d) = d - W(d)$. $\square$

This series can be very useful in practical computations, especially in iterative algorithms for approximating the Drazin inverse, because calculating Drazin inverse explicitly may be expensive in terms of time computing or even impossible.

The following theorem gives one of the most interesting uses of the Drazin inverse.

Theorem 3.3. *Let $\{X_n\}_{n \geq 0}$ be an ergodic Markov chain and let $N_{i,j}^{(n)}$ be the expected number of times the chain enters state j in n steps, having started from state i . If $N^{(n)} = (N_{i,j}^{(n)})$, then it holds that*

$$(18) \quad A^D = \lim_{n \rightarrow \infty} (N^{(n)} - (n+1)W).$$

Proof. It is known that $N^{(n)} = \sum_{k=0}^n \mathcal{P}^k$. Applying the previous theorem, we get

$$(19) \quad \lim_{n \rightarrow \infty} (N^{(n)} - (n+1)W) = \lim_{n \rightarrow \infty} \sum_{k=0}^n (\mathcal{P}^k - W) = A^D.$$

$\square$

This result implies that if n is large enough, we can approximate $N^{(n)}$ by $A^D + (n+1)W$.

$$(20) \quad N^{(n)} \approx A^D + (n+1)W.$$

We have talked about the number of times the chain passes through a certain state, now the question is, how long will it take the chain to reach a certain state? The following theorem allows us to calculate it using the Drazin inverse.

Theorem 3.4. *For an arbitrary ergodic Markov chain, let $m_{i,j}$ be the expected number of steps it takes for the system to reach state j when it initially starts in state i . Then, it holds that*

$$(21) \quad m_{i,j} = \frac{1}{\pi_j} (\delta_{i,j} + A^D(\mu^j - \mu^i)_j),$$

where $\delta_{i,j}$ is the Kronecker delta, $\mu^i = (\delta_{i,k})_{k \in \mathbb{N}}$, $A^D(d)_j$ is the j th coordinate of $A^D(d)$, and $\pi = (\pi_j)_{j \in \mathbb{N}}$ is the stationary distribution of the Markov chain.

Proof. The expected number of steps to go from state i to state j satisfies the following system of equations

$$(22) \quad m_{i,j} = 1 + \sum_{k \in \mathbb{N}} p_{i,k} m_{k,j} - p_{i,j} m_{j,j}.$$

Let's see that the expression in the statement satisfies it. Taking into account that $\sum_{k \in \mathbb{N}} p_{i,k} A^D(\mu^k)_j = A^D(\mathcal{P}(\mu^i))_j$ and that $A A^D(\mu^i) = (\text{Id} - W)(\mu^i) = \mu^i - \pi$, we have that

$$\begin{aligned} & 1 + \sum_{k \in \mathbb{N}} p_{i,k} \frac{1}{\pi_j} (\delta_{k,j} + A^D(\mu^j - \mu^k)_j) - p_{i,j} \frac{1}{\pi_j} \\ (23) \quad &= 1 + \frac{1}{\pi_j} A^D(\mu^j)_j - \frac{1}{\pi_j} A^D(\mathcal{P}(\mu^i))_j = 1 + \frac{1}{\pi_j} A^D(\mu^j)_j - \frac{1}{\pi_j} A^D(\text{Id} - A)(\mu^i)_j \\ &= 1 + \frac{1}{\pi_j} A^D(\mu^j)_j - \frac{1}{\pi_j} A^D(\mu^i)_j + \frac{1}{\pi_j} (\delta_{i,j} - \pi_j) = \frac{1}{\pi_j} (\delta_{i,j} + A^D(\mu^j - \mu^i)_j) \end{aligned}$$

□

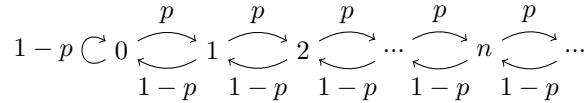
This theorem provides a powerful method for computing mean passage times using the Drazin inverse in arbitrary Markov chains. It proves that the Drazin inverse not only helps in understanding the stationary behavior of Markov chains but also in calculating important dynamical quantities like the expected passage times between states.

4. ILLUSTRATIVE EXAMPLE

The most typical example of an infinite Markov chain is the random walk with reflection. Consider a particle that moves randomly on the natural numbers, forward, with probability p , or backward with probability $1 - p$. In the special case of 0, since it cannot move further backward, we consider that it can stay in place with probability $1 - p$. Then

$$(24) \quad \mathcal{P}(X_{n+1} = j | X_n = i) = p_{i,j} = \begin{cases} p & \text{if } j = i + 1 \\ 1 - p & \text{if } j = i - 1 \text{ or } j = i = 0 \\ 0 & \text{in other case} \end{cases}$$

We can represent this chain with the following diagram.



In matrix form, one has the following

$$\mathcal{P} = \begin{pmatrix} 1-p & p & 0 & \cdots \\ 1-p & 0 & p & \cdots \\ 0 & 1-p & 0 & \ddots \\ \vdots & \vdots & \ddots & \ddots \end{pmatrix}$$

It is easily seen that any state can be accessed from any other, but it is not so immediate that all states are ergodic.

Theorem 4.1. *If $p < \frac{1}{2}$, the previous Markov chain is ergodic and, therefore, it has a unique stationary distribution.*

Proof. One can find a similar proof of this result in [7]. Let us see if there exists a stationary distribution, $\pi = (\pi_i)_{i \in E}$. This distribution should be a solution of the following system,

$$(25) \quad \pi \cdot \mathcal{P} = \pi$$

therefore, it has to verify

$$(26) \quad \begin{cases} \pi_0 = \pi_0(1-p) + \pi_1(1-p) \\ \pi_i = \pi_{i-1}p + \pi_{i+1}(1-p) \end{cases} \quad \forall i > 0$$

From first equation, it follows that

$$(27) \quad \pi_1 = \frac{p}{1-p} \pi_0,$$

and from the second with $i = 1$ we have that

$$(28) \quad \pi_2 = \frac{1}{1-p}(\pi_1 - \pi_0 p) = \left(\frac{p}{1-p}\right)^2 \pi_0.$$

Applying induction, we can conclude that

$$(29) \quad \begin{aligned} \pi_{i+1} &= \frac{1}{1-p}(\pi_i - \pi_{i-1}(1-p)) = \frac{1}{1-p} \left(\frac{p^i}{(1-p)^i} \pi_0 - \frac{p^i}{(1-p)^{i-1}} \pi_0 \right) \\ &= \left(\frac{p}{1-p} \right)^{i+1} \pi_0, \end{aligned}$$

and that $\left(\frac{p^i}{(1-p)^i} \pi_0 \right)_{i \in E}$ is a solution of the system (26). For it to indeed be a solution, it must satisfy

$$(30) \quad \sum_{i=0}^{\infty} \left(\frac{p}{1-p} \right)^i \pi_0 = 1.$$

Since $p < \frac{1}{2}$, then $\frac{p}{1-p} < 1$ and

$$(31) \quad \sum_{i=0}^{\infty} \left(\frac{p}{1-p} \right)^i \pi_0 = \frac{\pi_0}{1 - \frac{p}{1-p}}.$$

Taking $\pi_0 = 1 - \frac{p}{1-p} = \frac{1-2p}{1-p}$, we determinate a stationary distribution. By theorem 2.3 all states are positive recurrent and, therefore, the chain is ergodic due to the fact that the states are aperiodic. $\square$

Theorem 4.2. *If $\mathcal{P}$ is the transition operator of the Markov chain of the random walk with reflection, defined in (24), the operator $A = Id - \mathcal{P}$ has index 1.*

Proof. Let us see $V = \{(d_i)_{i \geq 0} : \sum d_i < \infty\}$ decompose into $V = \text{Ker} A \oplus \text{Im} A$. Fistly, one has that

$$(32) \quad A((d_i)_{i \geq 0}) = (p d_0 + (p-1)d_1, \dots, -p d_{i-1} + d_i + (p-1)d_{i+1}, \dots)$$

Furthermore, in the previous proof we just demonstrate that $\text{Ker} A = \left\langle \left(\left(\frac{p}{1-p} \right)^i \right)_{i \geq 0} \right\rangle$.

Suppose by contradiction that $\text{Ker} A \cap \text{Im} A \neq 0$, so there exists $(d_i)_{i \geq 0} \in V$ such that

$$(33) \quad \begin{cases} p d_0 + (p-1)d_1 = 1 \\ -p d_{i-1} + d_i + (p-1)d_{i+1} = \frac{p^i}{(1-p)^i} \end{cases} \quad \forall i > 0$$

Adding first equation with second one when $i = 1$, we get that

$$(34) \quad p d_1 + (p-1)d_2 = 1 + \frac{p}{1-p}.$$

Applying induction, we obtain that

$$(35) \quad p d_i + (p-1)d_{i+1} = \sum_{k=0}^i \left(\frac{p}{1-p} \right)^k.$$

Since $\sum_{i \geq 0} d_i < \infty$, $\lim_{i \rightarrow \infty} d_i = 0$. Then, the left member of the expression above goes to 0, but the right-hand side goes to $\frac{1}{1-\frac{p}{1-p}}$, which leads to a contradiction.

Now, let $(\mu_i)_{i \geq 0} \in V$. To see if it is in $\text{Ker} A + \text{Im} A$, let us see if the following system has a solution.

$$(36) \quad \begin{cases} p d_0 + (p-1)d_1 + a = \mu_0 \\ -p d_{i-1} + d_i + (p-1)d_{i+1} + a \frac{p^i}{(1-p)^i} = \mu_i \quad \forall i > 0 \end{cases}$$

Reasoning analogously to before, we conclude that for any $i > 0$ it holds that

$$(37) \quad p d_i + (p-1)d_{i+1} + a \sum_{k=0}^i \left(\frac{p}{1-p} \right)^k = \sum_{k=0}^i \mu_k.$$

And as $i \rightarrow \infty$, we obtain that

$$(38) \quad a = \frac{1-2p}{1-p} \sum_{k=0}^{\infty} \mu_k.$$

Finally, solving from (37), one determines $(d_i)_{i \geq 0}$ through the recurrence formula

$$(39) \quad d_{i+1} = \frac{1}{p-1} \left(\sum_{k=0}^i \mu_k - a \sum_{k=0}^i \left(\frac{p}{1-p} \right)^k - p d_i \right).$$

We need to check if $(d_i)_{i \geq 0} \in V$. For this, we use the ratio criterion.

$$(40) \quad \lim_{i \rightarrow \infty} \frac{|d_{i+1}|}{|d_i|} = \lim_{i \rightarrow \infty} \frac{\left| \sum_{k=0}^i \mu_k - a \sum_{k=0}^i \left(\frac{p}{1-p} \right)^k - p d_i \right|}{|d_i|(p-1)} = \frac{p}{1-p} < 1.$$

The series $\sum |d_i|$ converges.

With all the above, it has been proved that $V = \text{Ker} A \oplus \text{Im} A$ and by Theorem 2.1, A has index 1. $\square$

Now that we have shown that we are under the hypotheses of the theory from the previous section, we can apply it. For this, it is necessary to calculate the Drazin inverse of A . Let $\mu = (\mu_i)_{i \geq 0}$. We know that there exist $d = (d_i)_{i \geq 0} \in V$ and $v \in \text{Ker} A$ such that $\mu = A(d) + v$. In the same way, we can write $d = A(\bar{d}) + \bar{v}$, so $d - \bar{v} \in \text{Im} A$ and $A(d) = A(d - \bar{v})$. Consequently, $A^D(\mu) = A^{-1}(A(d)) = d - \bar{v}$. It is easy to check that $\bar{v} = \left(\frac{1-2p}{1-p} \sum d_j \right) \left(\left(\frac{p}{1-p} \right)^i \right)_{i \geq 0}$ and

$$(41) \quad A^D(\mu) = \left(d_i - \left(\frac{1-2p}{1-p} \sum_{j=0}^{\infty} d_j \right) \left(\frac{p}{1-p} \right)^i \right)_{i \geq 0}$$

where $(d_i)_{i \geq 0}$ is recurrently calculated by

$$(42) \quad d_{i+1} = \frac{1}{p-1} \left(\sum_{k=0}^i \mu_k - a \sum_{k=0}^i \left(\frac{p}{1-p} \right)^k - p d_i \right).$$

By Theorem 3.3, the expected number of times the chain is in j th state in n stages when the chain was initially in i th state, $N_{i,j}^{(n)}$, can be approximated by (i, j) th entry

of $A^D + (n+1)W$. For example, if we want to know how many times the chain is in 0 state, we need to calculate the 0th coordinate of $A^D(\mu^i) + (n+1)W(\mu^i)$ where $\mu_j^i = \delta_{i,j}$, and $\delta_{i,j}$ is the Kronecker delta.

With patience, one can find

$$(43) \quad d_j = \begin{cases} \frac{1-2p}{(1-p)^2} \sum_{k=0}^{j-1} (k+1) \left(\frac{p}{1-p}\right)^k + \left(\frac{p}{1-p}\right)^j d_0 & \text{if } j \leq i \\ \frac{1}{p-1} \sum_{k=0}^{j-i-1} \left(\frac{p}{1-p}\right)^k + \frac{1-2p}{(1-p)^2} \sum_{k=0}^{j-1} (k+1) \left(\frac{p}{1-p}\right)^k + \left(\frac{p}{1-p}\right)^j d_0 & \text{if } j > i \end{cases}$$

Since $W = \lim_{n \rightarrow \infty} \mathcal{P}^n$, it is clear that $W(\mu^i) = \pi$, where π is the stationary distribution.

Consequently, the 0th coordinate of $A^D(\mu^i)$ matches with

$$(44) \quad d_0 - \frac{1-2p}{1-p} \sum_{j=0}^{\infty} \left(\frac{1}{p-1} \sum_{k=0}^{j-i-1} \left(\frac{p}{1-p}\right)^k + \frac{1-2p}{(1-p)^2} \sum_{k=0}^{j-1} (k+1) \left(\frac{p}{1-p}\right)^k + \left(\frac{p}{1-p}\right)^j d_0 \right) \\ = d_0 - \frac{1-2p}{1-p} \sum_{j=0}^{\infty} \bar{d}_j^{(i)} - \frac{1-2p}{1-p} \sum_{j=0}^{\infty} \left(\frac{p}{1-p}\right)^j d_0 = -\frac{1-2p}{1-p} \sum_{j=0}^{\infty} \bar{d}_j^{(i)},$$

where $\bar{d}_j^{(i)} = \frac{1}{p-1} \sum_{k=0}^{j-i-1} \left(\frac{p}{1-p}\right)^k + \frac{1-2p}{(1-p)^2} \sum_{k=0}^{j-1} (k+1) \left(\frac{p}{1-p}\right)^k$ and we used the convention $\sum_{k=0}^j a_k = 0$ if $j < 0$. Notice that it does not depend on the $(d_i)_{i \geq 0}$ we used to calculate it.

Finally, one has that

$$(45) \quad N_{i,0}^{(n)} = -\frac{1-2p}{1-p} \sum_{j=0}^{\infty} \bar{d}_j^{(0)} + (n+1) \frac{1-2p}{1-p}.$$

Now, we can consider the expected number of steps the chain needs for reaching the state j from the state i . By Theorem 3.4, this number matches with $\frac{1}{\pi_j} (\delta_{i,j} + A^D(\mu^j - \mu^i)_j)$. Since

$$(46) \quad A^D(\mu^i)_j = \bar{d}_j^{(i)} - \frac{1-2p}{1-p} \sum_{k \geq 0} \bar{d}_k^{(i)},$$

one can get

$$(47) \quad m_{i,j} = \frac{1-p}{1-2p} \left(\frac{1-p}{p}\right)^j \left(\delta_{i,j} + \bar{d}_j^{(j)} - \bar{d}_j^{(i)} - \frac{1-2p}{1-p} \sum_{k \geq 0} (\bar{d}_k^{(j)} - \bar{d}_k^{(i)}) \right).$$

To finish the section, we will use the software *Wolfram Mathematica* to do some computation and verify experimentally that the theoretical calculations work.

First, a function to compute $\bar{d}_j^{(i)}$ is created (Figure 1). Then, using $p = 0.1$, we compute the expected number of times the chain is in 0th state in 1000 stages when the chain was initially in 41th state, $N_{41,0}^{(1000)}$, and the expected time to reach 0th state for first time, $m_{41,0}$ (Figure 2). The results are 844.361 and 51.25 respectively.

In Figure 3, one can see the code for approximating $N_{41,0}^{(1000)}$. For this propose, 2000 simulations of the random walk have been done, and the number of times it passes thought 0th state has been measured. Finally, the average of this data has been calculated.

$$\text{In}[1]:= d[p_ , i_ , j_] := \frac{1}{p-1} \sum_{k=0}^{j-i-1} \left(\frac{p}{1-p} \right)^k + \frac{1-2p}{(1-p)^2} \sum_{k=0}^{j-1} (k+1) \left(\frac{p}{1-p} \right)^k$$

FIGURE 1. Function to compute $\overline{d}_j^i$.

```

In[12]:= p = 0.10;
i = 41;
n = 1000;
jmax = 2000;
(*Expected number of times the M.c. enters in 0th state.*)
- \frac{1-2p}{1-p} \sum_{j=0}^{jmax} d[p, i, j] + (n+1) \frac{1-2p}{1-p}
(*Expected time to reach 0th state.*)
\sum_{j=0}^{jmax} (d[p, i, j] - d[p, 0, j])

Out[16]= 844.361

Out[17]= 51.25

```

FIGURE 2. Computation of $N_{41,0}^{(1000)}$ and $m_{41,0}$. To reduce computing time, the series $\sum \overline{d}_j^i$ is approximated by a partial sum.

```

In[18]:= (*Experimental computation of expected number of times
the M.c. enters in 0th state.*)
meanRecurrenceTimes = {};
Do[
  in = i;
  count = 0;
  Do[
    in = If[in == 0, in + RandomChoice[{1-p, p} -> {0, 1}],
    in + RandomChoice[{1-p, p} -> {-1, 1}]];
    If[in == 0, count++];
  , n];
  AppendTo[meanRecurrenceTimes, count];
, 2000]
Mean[meanRecurrenceTimes] // N

Out[20]= 844.426

```

FIGURE 3. Experimental calculation for $N_{41,0}^{(1000)}$.

Similarly, the analogous process has been carried out to approximate $m_{41,0}$. The code can be seen in Figure 4.

The results obtained experimentally coincide with the theoretical ones.

```

(*Experimental computation of expected time to reach 0
th state*)
meanPassageTimes = {};
Do[
  in = i;
  time = 0;
  While[in ≠ 0,
    in = in + RandomChoice[{1 - p, p} → {-1, 1}];
    time++;
  ];
  AppendTo[meanPassageTimes, time];
, 2000]
Mean[meanPassageTimes] // N
Out[23]= 51.292

```

FIGURE 4. Experimental calculation for $m_{41,0}$.

REFERENCES

- [1] Campbell, S. L., “The Drazin Inverse of an Infinite Matrix”, *SIAM Journal on Applied Mathematics* (1976).
- [2] Campbell, S.L. and Meyer, C.D., “Generalized Inverses of Linear Transformations”, *Society for Industrial and Applied Mathematics* (2009).
- [3] Drazin, M.P., “Pseudo-inverses in associative rings and semigroups”, *The American Mathematical Monthly* **64**(7) (1958) 506-514.
- [4] Grimmett, G. R. and Stirzaker, D.R., “Probability and Random Processes. Fourth Edition”, *Oxford University Press* (2020).
- [5] Levin, D. A. and Peres, Y., “Markov Chains and Mixing Times, second edition”, *American Mathematical Society* (2009).
- [6] Pablos Romo, F., “Core-Nilpotent Endomorphisms of Infinite-dimensional Vector Spaces”, *Mediterr. J. Math.* **18**(84) (2021).
- [7] “LECTURE 26: Infinite Markov chains. Continuous time Markov chains.”, MIT OpenCourseWare, https://dspace.mit.edu/bitstream/handle/1721.1/121170/6-436j-fall-2008/contents/lecture-notes/MIT6_436JF08_lec25.pdf

DEPARTAMENTO DE MATEMÁTICAS (STUDENT), UNIVERSIDAD DE SALAMANCA, PLAZA DE LA MERCED 1-4, 37008 SALAMANCA, ESPAÑA
 Email address: pablo291101@usal.es