



Benemérita Universidad Autónoma de Puebla

Facultad de Ciencias Físico Matemáticas

Una introducción a los métodos de Krylov mediante la
inversa de Drazin.

Tesis presentada al

Colegio de Matemáticas

como requisito parcial para la obtención del grado de

LICENCIADA EN MATEMÁTICAS APLICADAS

por

Diana Karina Jiménez Hernández

Asesorada por

Dr. Gabriel Kantún Montiel

Puebla Pue.
19 de octubre de 2022



Benemérita Universidad Autónoma de Puebla

Facultad de Ciencias Físico Matemáticas

Una introducción a los métodos de Krylov mediante la
inversa de Drazin.

Tesis presentada al

Colegio de Matemáticas

como requisito parcial para la obtención del grado de

LICENCIADA EN MATEMÁTICAS APLICADAS

por

Diana Karina Jiménez Hernández

Asesorada por

Dr. Gabriel Kantún Montiel

Puebla Pue.
19 de octubre de 2022

Título: Una introducción a los métodos de Krylov mediante la inversa de Drazin.

Estudiante: DIANA KARINA JIMÉNEZ HERNÁNDEZ

COMITÉ

Dr. Slavisa Djordjevic
Presidente

Dr. Jacobo Oliveros Oliveros
Secretario

Dr. Javier Mendoza Torres
Vocal

Vocal

Dr. Gabriel Kantún Montiel
Asesor

Agradecimientos

Deseo expresar mi gratitud a:

Al Consejo Nacional de Ciencia y Tecnología (CONACyT) por la beca parcial otorgada durante la realización del trabajo de tesis con nombre de Proyecto: **Métodos de Krylov a través de la inversa generalizada de Drazin**. Responsable Técnico: Dr. Slavisa Djordjevic.

Al Dr. Gabriel Kantún Montiel no solo por ser el asesor en este trabajo, sino por ayudarme a pulir mi formación profesional con todos los consejos que me ha brindado. Al Dr. Slavisa Djordjevic, al Dr. Jacobo Oliveros Oliveros y al Dr. Javier Mendoza Torres por sus comentarios, correcciones y por su tiempo dedicado a la finalización de este trabajo de tesis.

A mis amigos y compañeros de universidad y de trabajo, los cuales siempre han estado para apoyarme y que han hecho que los momentos difíciles sean menos pesados.

Dedicatoria

A mis padres María Patricia Hernández Tello y Armando Jiménez Pazos, por estar siempre conmigo, cuidandome y guiandome en cada uno de mis pasos, por todo el apoyo que me han dado durante todo ste tiempo y por confiar en mí, de todo corazón gracias, los amo.

A mis hermanos, Rafael Jiménez Hernández y Ana Patricia Jiménez Hernández por estar siempre al pendiente de mí, por celebrar mis logros y por su apoyo incondicional, los quiero mucho.

A mi abuelita materna María del Carmen Tello Flores, por enseñarme a ser una mujer fuerte y perseverante, gracias a sus consejos y palabras de aliento que hicieron de mi una mejor persona.

A mi abuelita paterna Julia Pazos Hernández, por siempre demostrarme su cariño incondicional, y aunque ya no este conmigo, siempre me acompaña en todos mis sueños y metas.

A mis tías, Silvia Hernández Tello, Blanca Estela Hernández Tello y primos Yael Hernández Hernández y María del Carmen Pacheco Hernández, por su amistad sincera y amor.

A mis mejores amigos Ivan Hernández Gutierrez y Jenifer Reyes Ramirez, por estar al pendiente, por creer en mí y alentarme a mejorar.

Resumen

La solución de un sistema de ecuaciones lineales de la forma $Ax = b$, con A una matriz cuadrada $n \times n$, puede resultar muy complicada de obtener utilizando algún método directo de operaciones matriciales. Sin embargo, los métodos iterativos son útiles para resolver problemas que involucran un gran número de variables. En el caso de un sistema de ecuaciones lineales, las dos clases principales de métodos iterativos son los métodos iterativos estacionarios y los más generales, métodos del subespacio de Krylov. Particularmente, uno está interesado en los métodos del subespacio de Krylov, los cuales funcionan formando una base ortogonal de la secuencia de potencias de la matriz por un residuo inicial (la secuencia de Krylov), las aproximaciones a la solución se forman minimizando el residuo en el subespacio formado. Sin embargo, cuando la matriz es singular, los métodos de Krylov pueden fallar, incluso si el sistema de ecuaciones tiene solución, esta podría no estar en un espacio de Krylov. En este tipo de casos, se puede utilizar la inversa generalizada de Drazin para obtener una solución única aproximada. Así, el objetivo principal de este trabajo de tesis es describir la solución de un sistema de ecuaciones lineales de la forma $Ax = b$, cuando el sistema tiene una matriz A cuadrada $n \times n$ singular, utilizando la inversa generalizada de Drazin en el subespacio de Krylov. Con esto en mente, la estructura de este trabajo queda definido de la siguiente manera: en el capítulo 1, se hace una breve introducción a los espacios de Krylov, se hace mención a los métodos directos e iterativos para la resolución de un sistema de ecuaciones lineales de la forma $Ax = b$ y como trabajar en un subespacio de Krylov. Finalmente se habla de manera muy general sobre la inversa de Drazin y los intereses particulares que se tienen sobre ella. El capítulo 2 esta basado en el estudio de la inversa generalizada de Drazin, donde se proporcionan algunos teoremas, proposiciones, lemas y algoritmos necesarios para el cálculo de la inversa de Drazin. Así mismo, se define el índice de una transformación lineal, importante para entender la definición algebraica y funcional de la inversa de Drazin. El capítulo 3 está dedicado al análisis y solución de sistemas de ecuaciones lineales en un espacio de Krylov utilizando la inversa generalizada de Drazin, se presenta un ejemplo particular, cuando se tiene una matriz no singular y la utilidad de la aplicación de la inversa de Drazin en el método de Krylov cuando la matriz es singular. Finalmente, en el capítulo 4 se presentan algunas conclusiones de este trabajo de tesis.

Índice general

Agradecimientos	V
Dedicatoria	VII
Resumen	IX
1. Introducción	1
1.1. Objetivos	1
1.2. Organización de la Tesis	2
1.3. El Método de Krylov	2
2. Inversas Generalizadas	5
2.1. Forma canónica de Jordan.	8
2.1.1. Ejemplos numéricos. Forma canónica de Jordan	14
2.2. El índice Drazin	17
2.3. La inversa de Drazin en en la resolución de un sistema ecuaciones lineales singular.	18
2.3.1. Ejemplos numéricos de la inversa de Drazin.	20
3. La inversa de Drazin en el método de Krylov	23
3.1. Ejemplo de un método de Krylov.	23
4. Conclusiones	31
Bibliografía	33

Capítulo 1

Introducción

Resolver un sistema de ecuaciones lineales $Ax = b$ resulta laborioso cuando la matriz A es de orden mayor a 3 y la matriz no es de tipo diagonal o dispersa. Esto se debe a que se utilizan usualmente operaciones matriciales, como ejemplo tenemos el método directo del cálculo del determinante $A - \lambda I$ [1]. Como alternativa se utilizan métodos iterativos o directos.

Los métodos iterativos son de gran utilidad para resolver sistemas de ecuaciones lineales de la forma $Ax = b$. Estos métodos también pueden resolver sistemas grandes cuando la matriz A es de coeficientes dispersa o dar información parcial de las soluciones a los métodos directos, que, de otra forma, no pueden o es muy costoso trabajar con matrices grandes, que es lo que pasa al utilizar operaciones matriciales. Además, los algoritmos directos modifican a la matriz A a lo largo del proceso y elementos que eran nulos cambian por las operaciones que se realizan a la matriz A . En cambio, en los iterativos solo se utiliza la matriz original. Los métodos iterativos nos proporcionan soluciones aproximadas después de realizar el algoritmo un cierto número de iteraciones. Estas se suelen interrumpir cuando se considera haber alcanzado los objetivos deseados. Además, cada iteración disminuye el error en una cierta razón. En general, los métodos iterativos multiplican vectores por la matriz A y trabajan con los vectores resultantes, iniciando con un vector y se calcula Ay , luego A^2y y así sucesivamente. Los algoritmos que trabajan de esta manera se llaman métodos del subespacio de Krylov [2]. Un ejemplo son el método o métodos del gradiente conjugado. Estos resuelven sistemas de ecuaciones lineales $Ax = b$ minimizando funciones cuadráticas [3]. También garantizan la solución del sistema con una matriz A hermitiana definida positiva. Esto sucede si los eigenvalores de A están alejados de cero, cercanos unos de otros y la minimización se realiza sobre espacios de Krylov.

Otro ejemplo, es el método de Lanczos, el cual de manera análoga al método del gradiente conjugado es rápido a la hora de calcular ciertos valores propios de una matriz hermitica, si los valores están bien separados del resto del espectro (el vector inicial que empieza la iteración es lo suficientemente genérico)[2].

1.1. Objetivos

El objetivo general del trabajo de tesis es describir la solución de un sistema de ecuaciones lineales, cuando el sistema tiene una matriz cuadrada A singular, utilizando la inversa generalizada Drazin en el subespacio de Krylov.

Como objetivos particulares se tienen los siguientes:

- Presentar una introducción a los espacios de Krylov.

- Estudiar la inversa Drazin.
- Analizar la solución en un espacio de Krylov utilizando la inversa generalizada de Drazin.

1.2. Organización de la Tesis

En el Capítulo 1 se describen brevemente los métodos directos e iterativos para la resolución de ecuaciones lineales de la forma $Ax = b$. También, se describe cómo se trabaja en un subespacio de Krylov y la utilización del método de Krylov en diferentes ramas. Así mismo, de manera general acerca de la inversa de Drazin y nuestro interés en ella.

El Capítulo 2 de esta tesis, acerca de la inversa generalizada de Drazin, estudiando lo concerniente de la misma, se define el índice de una transformación lineal que nos ayudará a entender la definición algebraica y funcional de la inversa de Drazin. También se proporcionan algunos teoremas, proposiciones, lemas y algoritmos que nos ayudaran a calcular la inversa de Drazin.

En el Capítulo 3 se presenta un ejemplo de los métodos de Krylov con una matriz no singular, la utilidad de la aplicación de la inversa de Drazin en el método de Krylov con una matriz singular, así como el resultado general al utilizar este método.

El Capítulo 4 se presentan las conclusiones del trabajo de tesis.

1.3. El Método de Krylov

El método de Krylov tiene la particularidad de seleccionar un valor adecuado de un vector auxiliar y si se conoce la ecuación característica, evitando el problema de realizar el cálculo del determinante. Se desea ocupar un método de Krylov cuando se trabaja con un sistema de ecuaciones $Ax = b$ de tamaño n y la matriz de coeficientes contiene muchas entradas cero. También cuando la matriz de orden n es tan grande que no puede permitirse gastar n^3 operaciones para resolver el sistema a través de eliminación gaussiana. Otro caso más sería cuando no se tiene acceso directo a la matriz.

Los métodos de Krylov se utilizan en diversos campos entre los cuales se encuentran los métodos numéricos, en numerosas aplicaciones de la ingeniería para la determinación de la ecuación característica y valores propios de una matriz A , en especial cuando los modelos matemáticos están expresados por medio de ecuaciones diferenciales. Por ejemplo, la determinación de la mayor frecuencia natural real de oscilación de sistemas masa-resorte con o sin amortiguamiento, Fig. 1.1 [4, 5].

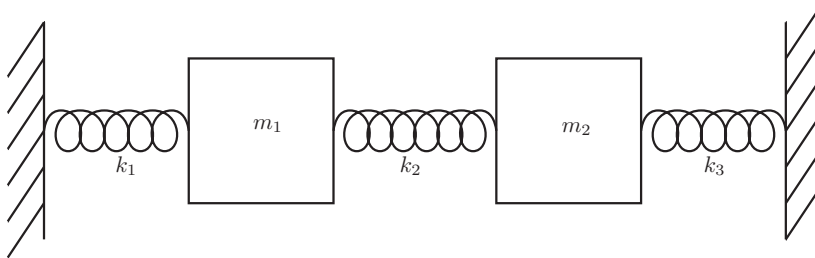


Figura 1.1: Cortés , J. (2011). Sistema de masas y resortes en donde las frecuencias natulares de la vibración son valores propios.

Se puede demostrar que toda matriz $A_{n \times n}$ tiene una sola ecuación característica, pero no se puede declarar que dada una ecuación característica existe una sola matriz A cuadrada que le de origen. Muestra de ello, es el siguiente ejemplo:

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & -1 & -3 \end{pmatrix}, \quad y \quad B = \begin{pmatrix} 0 & -2 & -\frac{5}{3} \\ 2 & 4 & \frac{5}{3} \\ -5 & -5 & -3 \end{pmatrix}. \quad (1.1)$$

Las matrices A y B tienen como ecuación característica a $(\lambda - 2)^2(\lambda + 3) = 0$.

Encontrar matrices con la misma ecuación característica se vuelve un problema difícil de resolver conforme va aumentando el orden de la matriz. Existen procedimientos del álgebra lineal para calcular distintas matrices A con la misma ecuación característica. Estos toman como base una matriz que contenga muchos ceros o unos entre sus elementos. Uno de estos procedimientos es el método de la matriz compañera, el cual toma en cuenta una matriz A de la forma:

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \cdots & \ddots & \ddots & \cdots & \cdots \\ -b_n & -b_{n-1} & -b_{n-2} & \cdots & -b_1 \end{pmatrix}, \quad (1.2)$$

donde los coeficientes de la ecuación característica son $b_1, b_2, \dots, b_{n-1}, b_n$ y tiene la forma:

$$\lambda^n + b_1 \lambda^{n-1} + b_2 \lambda^{n-2} + \cdots + b_{n-1} \lambda + b_n = 0. \quad (1.3)$$

La matriz transpuesta de A tiene la misma ec. 1.3 [5], además se puede obtener otra matriz C similar a la matriz A cuya ecuación característica ya se conoce, utilizando el producto de matrices, la igualdad de matrices y la definición de matriz inversa. Esta matriz similar se puede obtener como $C = B A B^{-1}$, donde B es una matriz no singular. La matriz C se puede calcular con un sistema de ecuaciones lineales, una vez propuesta B [6]. Pero este procedimiento se realiza a partir de conocer la matriz A .

En contraste, el método de Krylov aplicándolo de manera inversa, aborda el problema de obtener la matriz A cuando se conoce la ecuación característica. Sin embargo, estos métodos los utilizamos con matrices no singulares ahorrándonos tiempo al solucionar matrices grandes, pero si la matriz es singular, los métodos de Krylov pueden fallar, incluso cuando se trabaja con un sistema lineal que tenga solución, puede que no se encuentre en un espacio de Krylov. Para ello se utiliza la inversa de Drazin.

La inversa Drazin, definida por Michael P. Drazin en 1958 [7], es una inversa generalizada de una matriz A . Cuando se habla de la inversa generalizada no se refiere a la inversa que usualmente se conoce, ya que la inversa generalizada es para matrices de cualquier orden, ya sean rectangulares o cuadradas singulares, es decir, para matrices no invertibles. Esta tiene diferentes aplicaciones, por ejemplo; se emplea en la investigación de iteraciones de Cesàro-Neuman, la teoría de las cadenas finitas de Markov, criptografía, métodos iterativos en análisis numérico; solo por mencionar algunas.

En específico, nos interesa su aplicación en la solución de sistemas de ecuaciones lineales singulares, presentando la noción de inversa de Drazin solo para matrices cuadradas la cual es un tipo de inversa generalizada. Tomando en cuenta ciertas propiedades se trata de generalizar la idea de matriz inversa para matrices cuadradas singulares, buscando matrices que cumplan las condiciones que garanticen la existencia de la inversa de Drazin. Así, la inversa de Drazin no sólo nos permite dar solución a un sistema singular, sino que además nos permite trabajar en un espacio de Krylov.

Capítulo 2

Inversas Generalizadas

Sea $\mathbb{C}^{n \times n}$ el espacio que comprende las matrices de orden $n \times n$, sobre el campo complejo, esto es para que se puedan considerar todas las raíces de los polinomios característicos. Cualquier matriz A que está en $\mathbb{C}^{n \times n}$ tiene un índice que se obtiene de la inversa de Drazin de A para hacer uso de ella. Incluso si la matriz A es singular, el índice y la inversa Drazin de la matriz existen y son únicos.

Además, en este capítulo se mencionarán algunas definiciones y teoremas sobre el índice de la matriz y el inverso de Drazin. No sin antes mencionar la forma normal de Jordan, para lo cual se recordarán algunos conceptos.

Definición 1. Sea V un conjunto no vacío cuyos elementos se llaman vectores (v_1, v_2, v_3, \dots) , y \mathbb{Z} un campo cuyos elementos se denominan escalares $(\alpha, \beta, \gamma, \dots)$. Definimos en V dos leyes de composición: una interna, que denotaremos por $(+)$ y denominaremos suma o adición de vectores de V , y otra externa, a la que llamaremos producto de un vector por un escalar. En estas condiciones, se dice que V es un espacio vectorial sobre \mathbb{Z} si se verifican las siguientes condiciones:

- Para la ley interna, suma de vectores de V , el par $(V, +)$ es un grupo abeliano, por lo que cumple las propiedades siguientes:

1. Cerradura bajo adición: $\forall v_1, v_2 \in V \Rightarrow v_1 + v_2 \in V$
2. Propiedad conmutativa: $\forall v_1, v_2 \in V \Rightarrow v_1 + v_2 = v_2 + v_1$.
3. Propiedad asociatividad : $\forall v_1, v_2, v_3 \in V \Rightarrow (v_1 + v_2) + v_3 = v_1 + (v_2 + v_3)$.
4. Neutro aditivo: $\forall v_1 \in V, \exists! 0 \in V \mid v_1 + 0 = 0 + v_1 = v_1$.
5. Inverso aditivo: $\forall v_1 \in V, \exists (-v_1) \in V \mid v_1 + (-v_1) = (-v_1) + v_1 = 0$.

- Para la ley de composición externa, producto de un vector escalar, que asocia a cada escalar α , y a cada $v_1 \in V$ un vector único, llamado producto de α y v_1 , y que representa por αv_1 o $v_1 \alpha$, y satisface las siguientes propiedades:

1. Cerradura bajo multiplicación escalar: $\forall v \in V \wedge \alpha \in \mathbb{Z} \Rightarrow \alpha v \in V$
2. Propiedad distributiva: $\forall v_1, v_2 \in V \wedge \forall \alpha \in \mathbb{Z} \Rightarrow \alpha(v_1 + v_2) = \alpha v_1 + \alpha v_2$.
3. Propiedad distributiva: $\forall v_1 \in V \wedge \forall \alpha, \beta \in \mathbb{Z} \Rightarrow (\alpha + \beta)v_1 = \alpha v_1 + \beta v_1$.
4. Propiedad asociativa: $\forall v_1 \in V \wedge \forall \alpha, \beta \in \mathbb{Z} \Rightarrow (\alpha\beta)v_1 = \alpha(\beta v_1)$.
5. Identidad escalar: $\forall v_1 \in V, 1 \cdot v_1 = v_1$.

Definición 2. Sean V y E espacios vectoriales de \mathbb{C} , si E es un subconjunto no vacío de V y las operaciones que hacen de E un espacio vectorial son las mismas que las que hacen a V espacio vectorial, se dice que E es un subespacio vectorial de V .

Definición 3. Sea V espacio vectorial de \mathbb{C} . Sea $F = \{e_1, e_2, \dots, e_k\}$ un sistema finito de vectores de V (F es un conjunto de vectores dado en un cierto orden), se dice que el vector $v_1 \in V$ es combinación lineal de los vectores de F si existen k escalares $\alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{C}$ tales que:

$$v_1 = \alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_k e_k. \quad (2.1)$$

Proposición 1. Sea $F = \{e_1, e_2, \dots, e_k\} \subset V$. Todos los vectores que se pueden formar por combinación lineal de los vectores e_1, e_2, \dots, e_k pertenecen a V y, además, forman un subespacio vectorial de V nombrado subespacio generado por F , y se representa como:

$$Span(F) = \{v_1 = \alpha_1 e_1 + \dots + \alpha_k e_k \mid \alpha_i \in \mathbb{Z}, i = 1, 2, \dots, k\} \in V. \quad (2.2)$$

Asimismo, se dice que F es un sistema generador de $Span(F)$, y es el menor subespacio vectorial de V que contiene a F .

Demostración. Si V es espacio vectorial y $Span(F) \subset V$, se puede ver que $Span(F)$ cumple la condición necesaria y suficiente para ser un subespacio vectorial de V .

Dado que $0 \in Span(F)$, el vector nulo puede escribirse como combinación lineal

$$0 = 0e_1 + 0e_2 + \dots + 0e_k. \quad (2.3)$$

Por demostra que $\forall v_1, v_2 \in Span(F) \Rightarrow v_1 + v_2 \in Span(F)$.

Sean $v_1, v_2 \in Span(F)$, se tiene que:

$$\begin{aligned} v_1 &= \alpha_1 e_1 + \dots + \alpha_k e_k, \\ v_2 &= \beta_1 e_1 + \dots + \beta_k e_k, \\ \Rightarrow v_1 + v_2 &= (\alpha_1 + \beta_1)e_1 + \dots + (\alpha_k + \beta_k)e_k \in Span(F). \end{aligned} \quad (2.4)$$

Ahora, hay que demostrar que $\forall v_1 \in Span(F) \wedge \forall \alpha \in \mathbb{Z} \Rightarrow \alpha v_1 \in Span(F)$.

Sea $v_1 \in Span(F)$ y sea $\alpha \in \mathbb{Z}$, se tiene que:

$$\begin{aligned} v_1 &= \alpha_1 e_1 + \dots + \alpha_k e_k, \\ \Rightarrow \alpha v_1 &= (\alpha \alpha_1)e_1 + (\alpha \alpha_2)e_2 + \dots + (\alpha \alpha_k)e_k, \\ &\Rightarrow \alpha v_1 \in Span(F). \end{aligned} \quad (2.5)$$

Además, todo subespacio que incluya a F ha de contener a las combinaciones lineales de los vectores de F , esto es, ha de incluir a $Span(F)$. Como el $Span(F)$ es el subespacio vectorial e incluye a F , se concluye de todo ello que $Span(F)$ es el menor subespacio vectorial que incluye a F . ■

Definición 4. Sea V un espacio vectorial sobre \mathbb{C} y sea $F = \{e_1, e_2, \dots, e_k\}$ un sistema finito de vectores de V . Se dice que F es una familia o sistema libre de vectores o también que los vectores e_1, e_2, \dots, e_k son linealmente independientes si

$$\alpha_1 e_1 + \dots + \alpha_k e_k = 0, \quad (2.6)$$

se cumple sólo cuando $\alpha_1 = \alpha_2 = \dots = \alpha_k = 0$. En el caso de que existan k escalares $\alpha_1, \alpha_2, \dots, \alpha_k$, no todos nulos, tales que: $\alpha_1 e_1 + \dots + \alpha_k e_k = 0$, se dice que los vectores e_1, e_2, \dots, e_k son linealmente dependientes o que F es un sistema ligado.

Definición 5. Un espacio vectorial V sobre \mathbb{C} se dice que es de tipo finito si está generado por un número finito de vectores.

En un espacio vectorial finito, aquellos sistemas generadores y linealmente independientes se les llama bases. Todo vector del espacio se podrá expresar de forma única como combinación lineal de los vectores de la base.

Definición 6. Un conjunto finito de vectores de un espacio vectorial V sobre \mathbb{C} se dice que es base de V si es familia de V , es decir un conjunto de vectores generadores de V linealmente independiente.

Teorema 1. Representación única en una base. Sea $F = \{e_1, e_2, \dots, e_k\}$ una base de un espacio vectorial V sobre \mathbb{C} . Entonces, todo vector $v_1 \in V$ se puede expresar de forma única como combinación lineal de los vectores de F .

Demostración. Supóngase que el vector v_1 se puede expresar de las siguientes dos formas:

$$\begin{aligned} v_1 &= v'_1 e_1 + v'_2 e_2 + \dots + v'_k e_k, \\ v_1 &= v''_1 e_1 + v''_2 e_2 + \dots + v''_k e_k, \end{aligned} \tag{2.7}$$

restando ambas expresiones de la ec. (2.7), resulta que:

$$0 = (v'_1 - v''_1)e_1 + \dots + (v'_k - v''_k)e_k. \tag{2.8}$$

Luego, siendo la base una familia libre se llega a que:

$$v'_1 - v''_1 = v'_2 - v''_2 = \dots = v'_k - v''_k, \tag{2.9}$$

de donde $v'_i = v''_i \forall i = 1, 2, \dots, k$. Por lo tanto, v_1 se expresa de manera única. ■
A los escalares v'_1, \dots, v'_k se les denomina coordenadas del vector v_1 en la base F .

Teorema 2. Existencia de bases. Todo espacio vectorial V de tipo finito, distinto de $\{0\}$, tiene alguna base.

Demostración. Suponiendo que $V = \text{Span}(e_1, e_2, \dots, e_k)$. Si los vectores e_1, e_2, \dots, e_k son linealmente independientes, entonces son una base de V .

Si no son linealmente independientes, habrá al menos uno de ellos que sea combinación lineal de los restantes. Si se quitan todos estos vectores, se obtiene un conjunto linealmente independiente que sigue siendo total, pues los vectores suprimidos no ocupan un papel en la obtención de vectores de V . Es decir, se obtiene una base de V . ■

Definición 7. Se llama dimensión de un espacio vectorial V al número de vectores de una base de V . Se escribe $\dim V$.

Definición 8. Sea T una transformación lineal. Se llama núcleo de T (notación: $\ker T$) al conjunto de vectores v , tal que $T(v) = 0$.

Definición 9. Sea T una función de un espacio vectorial V a un espacio vectorial V' . Entonces T es una transformación lineal si cumple con las siguientes propiedades:

1. Para todo vector $v_1, v_2 \in V$, $T(v_1 + v_2) = T(v_1) + T(v_2)$.
2. Para todo $v_1 \in V$ y α cualquier escalar, $T(\alpha v_1) = \alpha T(v_1)$.

Definición 10. Nulidad de una transformación lineal. Sean V, V' espacios vectoriales sobre \mathbb{C} y sea $V' \in \mathcal{L}(V, V')$ (subespacio de las funciones de V en V'). La nulidad de T se define como la dimensión del núcleo de T :

$$\text{nul}(T) = \dim(\ker T). \tag{2.10}$$

Definición 11. Un número $\lambda \in \mathbb{C}$ es llamado valor propio de la matriz A si hay un vector $x \neq 0$ tal que $Ax = \lambda x$. Cualquier vector de este tipo se llama vector propio de A asociado al valor propio λ .

El conjunto $L(\lambda) = \{x \mid (A - \lambda I)x = 0\}$ forma un subespacio lineal de \mathbb{C}^n , de dimensión

$$\rho(\lambda) = n - \text{rank}(A - \lambda I). \tag{2.11}$$

El número entero $\rho(\lambda) = \dim L(\lambda)$ especifica el número máximo de vectores propios linealmente independientes asociados con el valor propio λ . Además $\phi(\mu) = \det(A - \mu I)$ es un polinomio de grado n de la forma:

$$\phi(\mu) = (-1)^n (\mu^n + \alpha_{n-1}\mu^{n-1} + \cdots + \alpha_0). \quad (2.12)$$

Este es llamado el polinomio característico de la matriz A . Sus ceros son los valores propios de A . Si $\lambda_1, \dots, \lambda_k$ son los ceros distintos de $\phi(\mu)$ entonces $\phi(\mu)$ puede ser representada en la forma:

$$\rho(\mu) = (-1)^n (\mu - \lambda_1)^{\sigma_1} (\mu - \lambda_2)^{\sigma_2} \cdots (\mu - \lambda_k)^{\sigma_k}. \quad (2.13)$$

Los números enteros σ_i , que también se denotan por $\sigma(\lambda_i)$, se denomina multiplicidad del valor propio λ_i .

Teorema 3. Si λ es valor propio de la matriz $A \in \mathbb{C}^{n \times n}$, entonces:

$$1 \leq \rho(\lambda) \leq \sigma(\lambda) \leq n. \quad (2.14)$$

La demostración a este teorema se puede encontrar en la referencia [8].

2.1. Forma canónica de Jordan.

Definición 12. *Diagonalización de una matriz.* Suponga que la matriz $A_{n \times n}$ tiene n vectores propios linealmente independientes. Si estos vectores propios son las columnas de una matriz S , entonces $S^{-1}AS$ es una matriz diagonal Λ . Los valores propios de A están sobre la diagonal de Λ :

$$\text{Diagonalización: } \Lambda = S^{-1}AS = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}. \quad (2.15)$$

Teorema 4. Si A es una matriz de tamaño $n \times n$, entonces las proposiciones siguientes son equivalentes:

1. A es diagonalizable,
2. A tiene n vectores propios linealmente independientes.

Demostración. $1 \Rightarrow 2$. Supongamos que A es diagonalizable, así que existe una matriz invertible

$$S = \begin{pmatrix} s_{11} & \cdots & s_{1n} \\ \vdots & \cdots & \vdots \\ s_{n1} & \cdots & s_{nn} \end{pmatrix}, \quad (2.16)$$

tal que $\Lambda = S^{-1}AS$ es diagonal, ec. (2.15). Por lo tanto, $AS = S\Lambda$; es decir:

$$AS = \begin{pmatrix} s_{11} & \cdots & s_{1n} \\ \vdots & \cdots & \vdots \\ s_{n1} & \cdots & s_{nn} \end{pmatrix} \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} = \begin{pmatrix} \lambda_1 s_{11} & \cdots & \lambda_n s_{1n} \\ \vdots & \cdots & \vdots \\ \lambda_1 s_{n1} & \cdots & \lambda_n s_{nn} \end{pmatrix} \quad (2.17)$$

Si ahora se denotan los vectores columna de S por s_1, s_2, \dots, s_n , entonces por la ec. (2.17), las columnas sucesivas de AS son: $\lambda_1 s_1, \lambda_2 s_2, \dots, \lambda_n s_n$. Sin embargo, las columnas sucesivas de AS son As_1, As_2, \dots, As_n . Así entonces, se tiene que:

$$As_1 = \lambda_1 s_1, As_2 = \lambda_2 s_2, \dots, As_n = \lambda_n s_n. \quad (2.18)$$

Dado que S es invertible, no todos sus vectores columna son cero. Luego, por la ec. (2.18), $\lambda_1, \dots, \lambda_n$ son valores propios de A , y s_1, \dots, s_n son vectores propios correspondientes. Ya

que S es invertible. Por el Teorema 13 de la ref.[9], se deduce que: s_1, \dots, s_n son linealmente independientes. Por lo tanto, A tiene n vectores propios linealmente independientes.

$2 \Rightarrow 1$. Supongamos que A tiene n vectores propios linealmente independientes, s_1, \dots, s_n , con los valores propios correspondientes $\lambda_1, \dots, \lambda_n$, y la matriz S definida como en la ec. (2.16). La matriz cuyos vectores columna son s_1, \dots, s_n , las columnas del producto AS son As_1, As_2, \dots, As_n , pero

$$As_1 = \lambda_1 s_1, As_2 = \lambda_2 s_2, \dots, As_n = \lambda_n s_n, \quad (2.19)$$

de modo que:

$$AS = \begin{pmatrix} \lambda_1 s_{11} & \cdots & \lambda_n s_{1n} \\ \vdots & \dots & \vdots \\ \lambda_1 s_{n1} & \cdots & \lambda_n s_{nn} \end{pmatrix} = \begin{pmatrix} s_{11} & \cdots & s_{1n} \\ \vdots & \dots & \vdots \\ s_{n1} & \cdots & s_{nn} \end{pmatrix} \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} = S\Lambda. \quad (2.20)$$

En donde Λ es la matriz diagonal que tiene los valores propios $\lambda_1, \dots, \lambda_n$ en la diagonal principal. Supuesto que los vectores de S son linealmente independientes, S es invertible, así entonces, por la ec. (2.20), se puede reescribir como: $S^{-1}AS = \Lambda$; es decir, A es diagonalizable. ■

Proposición 2. Sea $A_{n \times n}$ y $\{\lambda_1, \dots, \lambda_k\} \subseteq \Lambda(A)$, donde $\Lambda(A)$ es el conjunto de valores propios (distintos) de A . Si los vectores propios u_1, \dots, u_k de A corresponden a valores propios distintos $\lambda_1, \dots, \lambda_k$, entonces los vectores u_1, \dots, u_k son linealmente independientes.

Demostración. Demostración por contradicción. Sea $\Lambda(A)$ el conjunto de valores propios (distintos) de A , entonces $\lambda_i \neq \lambda_j$ para $i \neq j$. Luego, con los vectores u_1, \dots, u_k se construye la matriz $U = [u_1 \cdots u_k]$. Si u_1, \dots, u_k no son linealmente dependientes, entonces $rank U = c \leq k$ y hay c columnas de U , diganse u_{i_1}, \dots, u_{i_c} que son linealmente independientes y vectores propios asociados a los valores propios $\lambda_{i_1}, \dots, \lambda_{i_c}$. Sea u_j una columna de X tal que $j \notin \{i_1, \dots, i_c\}$, entonces

$$u_j = \alpha_1 u_{i_1} + \cdots + \alpha_r u_{i_c}. \quad (2.21)$$

Como $u_j \in \ker(\lambda_j I_n - A)u_j$ y $Au_{i_d} = \lambda_{i_d} u_{i_d}$, con $d = 1, \dots, k$, resulta que:

$$\begin{aligned} 0 &= (\lambda_j I_n - A)u_j = (\lambda_j I_n - A)(\alpha_1 u_{i_1} + \cdots + \alpha_c u_{i_c}), \\ &= \alpha_1 (\lambda_j - \lambda_{i_1}) u_{i_1} + \cdots + \alpha_c (\lambda_j - \lambda_{i_c}) u_{i_c}, \end{aligned} \quad (2.22)$$

y como u_{i_1}, \dots, u_{i_c} son linealmente independientes

$$\alpha_d (\lambda_j - \lambda_{i_d}) = 0, \quad d = 1, \dots, c. \quad (2.23)$$

Pero $u_j \neq 0$, de modo que al menos un $\alpha_d \neq 0$. Para ese índice s se concluye que necesariamente $\lambda_j = \lambda_{i_s}$, contradiciendo que todos los valores propios son distintos. ■

Corolario 1. Si una matriz $A_{n \times n}$ tiene n valores propios distintos, entonces A es diagonalizable.

Demostración. Si u_1, \dots, u_k son vectores propios correspondientes a los valores propios distintos $\lambda_1, \dots, \lambda_k$ entonces, por el Teorema 3, u_1, \dots, u_k son linealmente independientes. Por lo tanto, A es diagonalizable. ■

Definición 13. Para $\lambda_0 \in \Lambda(A)$, se define la multiplicidad geométrica de λ_0 como la dimensión del subespacio propio de A asociado a λ_0 , y se denota por $mg_A(\lambda_0)$:

$$mg_A(\lambda_0) = \dim \ker(\lambda_0 I_n - A). \quad (2.24)$$

Definición 14. La multiplicidad algebraica de λ_0 como valor propio de A , se representa por $ma_A(\lambda_0)$, es la multiplicidad como raíz del polinomio característico tal como se explica después de la Definición 11.

Proposición 3. Para $A \in \mathbb{C}^{n \times n}$ y $\lambda_0 \in \Lambda(A)$, se tiene que: $ma(\lambda_0) \geq mg(\lambda_0)$.

Demostración. Sea $mg(\lambda_0) = l$ y $\{t_1, \dots, t_n\}$ una base del subespacio propio $O_{\lambda_0} = \ker(\lambda_0 I_n - A)$. Al aplicar esta base $\{t_1, \dots, t_l, t_{l+1}, \dots, t_n\}$ para obtener una base de \mathbb{C}^n . Sea $T = [t_1 t_2 \dots t_n]$ la matriz cuyas columnas son los vectores de la base escogida. Y sea $B = T^{-1} A T$, se tiene que:

$$\begin{aligned} AT &= [A t_1 \dots A t_l \ A t_{l+1} \dots A t_n] = [\lambda_0 t_1 \dots \lambda_0 t_l \ \lambda_0 t_{l+1} \dots \lambda_0 t_n] \\ &= [t_1 \dots t_l \ t_{l+1} \dots t_n] \begin{pmatrix} \lambda_0 I_l & C \\ 0 & \Lambda \end{pmatrix} = T B. \end{aligned} \quad (2.25)$$

Así,

$$\lambda I_n - B = \begin{pmatrix} (\lambda - \lambda_0) I_l & -C \\ 0 & \lambda I_{n-l} - \Lambda \end{pmatrix}, \quad (2.26)$$

con

$$\det(\lambda I_n - B) = (\lambda - \lambda_0)^l \det(\lambda I_{n-l} - \Lambda). \quad (2.27)$$

Como A y B tienen para cada valor propio las mismas multiplicidades algebraicas y geométricas y λ_0 podría ser valor propio de Λ , se concluye que:

$$l = mg(\lambda_0) \leq ma(\lambda_0). \blacksquare \quad (2.28)$$

Definición 15. Sea $A \in \mathbb{C}^{n \times n}$ y $\lambda_0 \in \Lambda(A)$

1. Si $mg(\lambda_0) < ma(\lambda_0)$, se dice que λ_0 es un valor propio defectuoso de A , en caso contrario se dice que es no defectuoso. La matriz A se dice que es defectuosa si tiene algún valor propio defectuoso y, en caso contrario, se dice que es no defectuosa.
2. Si $ma(\lambda_0) = 1$, se dice que λ_0 es un valor propio simple de A . La matriz A se dice que es simple si sus valores propios son simples.
3. Los valores propios no defectuosos, que no son simples, se llaman semisimples. La matriz A se dice que es semisimple si sus valores propios son simples o semisimples.

Teorema 5. A es una matriz diagonalizable si y solo si para cada valor propio λ se verifica que:

$$ma(\lambda) = mg(\lambda). \quad (2.29)$$

Demostración. Sean $\lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{C}$ los valores propios distintos de A y sea T un operador lineal asociado a la matriz A . Suponiendo que T es diagonalizable, entonces la unión de las bases de los subespacios propios $s_{\lambda_1}, s_{\lambda_2}, \dots, s_{\lambda_k}$ da lugar a una base S . Así que:

$$l = \sum_{i=1}^h \dim(s_{\lambda_i}) = \sum_{i=1}^l mg(\lambda_i). \quad (2.30)$$

Por otro lado,

$$\sum_{i=1}^h mg(\lambda_i) \leq \sum_{i=1}^h \leq \sum_{\lambda} ma(\lambda) = l, \quad (2.31)$$

donde λ es la raíz del polinomio característico de A , entonces:

$$l = \sum_{i=1}^h mg(\lambda_i) = \sum_{i=1}^h ma(\lambda_i), \quad (2.32)$$

de donde

$$\sum_{i=1}^h (ma(\lambda_i) - mg(\lambda_i)) = 0. \quad (2.33)$$

Como $ma(\lambda_i) - mg(\lambda_i) \geq 0$, $\forall i = 1, \dots, h$, la igualdad anterior, ec. (2.33), implica que:

$$ma(\lambda_i) - mg(\lambda_i) = 0 \quad \forall i = 1, \dots, h,$$

es decir,

$$ma(\lambda_i) = mg(\lambda_i).$$

Por otro lado, la igualdad

$$\sum_{i=1}^h ma(\lambda_i) = \sum_{\lambda} ma(\lambda) = l, \quad (2.34)$$

implica que toda raíz del polinomio característico de A se encuentra en \mathbb{C} . (Explicando esto último, supongase que en esta ocasión S es un espacio vectorial real y que el polinomio característico de A tiene una raíz compleja, llamada λ_0 . Entonces se tiene que $ma(\lambda_0)$ no aparece en la suma $\sum_{i=1}^h ma(\lambda_i)$, por lo cual $\sum_{i=1}^h ma(\lambda_i) < \sum_{\lambda} ma(\lambda)$, y esto no es cierto.).

Ahora suponiendo que toda raíz del polinomio de A se encuentra en \mathbb{C} y que todo valor propio $\lambda \in \mathbb{C}$ verifica $mg(\lambda) = ma(\lambda)$. Esto implica que:

$$\sum_{i=1}^h mg(\lambda_i) = \sum_{i=1}^h ma(\lambda_i) = \sum_{\lambda} ma(\lambda) = l. \quad (2.35)$$

Como $\sum_{i=1}^h mg(\lambda_i) = l$, se tiene que la suma de las dimensiones de los espacios propios s_{λ_i} coincide con la dimensión de S , por lo que las bases de los s_{λ_i} forman una base de A y por lo tanto S , es diagonalizable.

Observación 1. *Valores propios de A^k son $\lambda_1^k, \dots, \lambda_n^k$, y cada vector propio de A sigue siendo un vector propio de A^k . Una vez que S diagonaliza a A , también diagonaliza a A^k :*

$$\Lambda^k = (S^{-1} A S)(S^{-1} A S) \cdots (S^{-1} A S) = S^{-1} A^k S, \quad (2.36)$$

con $S^{-1} S = 1$ y solo la primera S^{-1} y la última S sobreviven.

Teorema 6. *Dos matrices diagonalizables A y B , son conmutables $AB = BA$ si y solo si comparten la misma matriz vector propio S .*

Demostración. Demostración por la izquierda. Si la misma matriz S diagonaliza tanto $A = S \Lambda_1 S^{-1}$ como a $B = S \Lambda_2 S^{-1}$, multiplicando A por B y B por A en cualquier orden:

$$\begin{aligned} AB &= S \Lambda_1 S^{-1} S \Lambda_2 S^{-1} = S \Lambda_1 \Lambda_2 S^{-1}, \\ BA &= S \Lambda_2 S^{-1} S \Lambda_1 S^{-1} = S \Lambda_2 \Lambda_1 S^{-1}, \end{aligned} \quad (2.37)$$

como $\Lambda_1 \Lambda_2 = \Lambda_2 \Lambda_1$ (las matrices diagonales conmutan), se tiene que:

$$AB = BA. \quad (2.38)$$

Demostración por la derecha. Sea $AB = BA$, partiendo de $Ax = \lambda x$, se tiene que:

$$ABx = BAx = B\lambda x = \lambda Bx. \quad (2.39)$$

Así, x y Bx son vectores propios de A , que comparten el mismo valor propio λ (o bien $Bx = 0$). Se asume por conveniencia que los valores propios de A son distintos (los espacios propios son todos unidimensionales), entonces Bx debe ser múltiplo de x . En otras palabras, x es un vector propio de B y de A . ■

Observación 2. Suponga que $B = M^{-1} A M$, entonces A y B tienen los mismos valores propios. Todo vector propio x de A corresponde a un vector propio $M^{-1} x$ de B .

Definición 16. La matriz N es normal si conmuta con su traspuesta conjugada, esto es:

$$N N^H = N^H N, \quad (2.40)$$

con N^H la matriz traspuesta conjugada.

Si la matriz A no es diagonalizable se puede encontrar una matriz M , no singular, tal que $M^{-1} A M$ es casi diagonal.

Definición 17. Dada A que tiene s vectores propios independientes, es semejante a una matriz con s bloques:

$$\text{Forma de Jordan: } J = M^{-1} A M = \begin{pmatrix} J_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & J_s \end{pmatrix}. \quad (2.41)$$

Cada bloque de Jordan J_i se llama bloque elemental de Jordan, de valor propio λ_i , es decir, es una matriz triangular que sólo tiene un valor propio λ_i y solo un vector propio:

$$\text{Bloque de Jordan: } J_i = \begin{pmatrix} \lambda_i & 1 & \cdots & 0 \\ 0 & \lambda_i & \cdots & \cdots \\ \cdots & \cdots & \cdots & 1 \\ 0 & \cdots & 0 & \lambda_i \end{pmatrix} \quad (2.42)$$

El mismo λ_i aparece en varios bloques, si tiene varios vectores propios independientes. Dos matrices son semejantes si y sólo si comparten la misma forma de Jordan J .

Teorema 7. Los r vectores w_i , los p vectores y_i y los $n - r - p$ vectores z_i forman cadenas de Jordan para la matriz A , y estos vectores son linealmente independientes. Van en las columnas de M , y $J = M^{-1} A M$ está en forma de Jordan.

Demostración. Por inducción matemática, empezando con el hecho de que toda matriz de 1×1 ya está en forma de Jordan. Puede suponerse que la construcción se logra para todas las matrices de orden menor que n y luego explicar los pasos para una matriz de orden n . La demostración se hará en tres pasos:

- **Paso 1:** Si se supone que A es singular, entonces su espacio columna tiene dimensión $r < n$. Observando solo dentro de este espacio más pequeño, la hipótesis de inducción garantiza que una forma de Jordan es posible: en el espacio columna debe haber r vectores independientes w_i tales que:

$$A w_i = \lambda_i w_i, \quad A w_i = \lambda_i w_i + w_{i-1}. \quad (2.43)$$

- **Paso 2.** Suponga que el espacio nulo y el espacio columna de A tienen una intersección de dimensión p . Es claro que, todo vector en el espacio nulo es un vector propio correspondiente a $\lambda = 0$. Por consiguiente, por Paso 1, se deben hacer p cadenas que comiencen en este valor propio, y se tiene interés en los vectores w_i que vienen al final de estas cadenas. Cada uno de estos p vectores está en el espacio columna, por lo que cada uno es una combinación de las columnas de A :

$$w_i = A y_i,$$

para alguna y_i .

- **Paso 3.** La dimensión del espacio nulo siempre es $n - r$. En consecuencia, de forma independiente con respecto a su intersección p -dimensional con el espacio columna, debe contener $n - r - p$ vectores básicos adicionales z_i fuera de esa intersección.

Ahora, juntando estos pasos para obtener el teorema de Jotdan.

Los r vectores w_i , los p vectores y_i y los $n - r - p$ vectores z_i forman cadenas de Jordan para la matriz A , y estos vectores son linealmente independientes. Si estos vectores conforman las columnas de la matriz M , entonces M es no singular y $J = M^{-1} A M$ está en forma de Jordan.

Si se desea reenumerar estos vectores como u_1, u_2, \dots, u_n , entonces cada y_i debería insertarse inmediatamente después del w_i del que provienen; esto completa una cadena en la que $\lambda_i = 0$. Los z_i vienen hasta el final, cada una sola en su propia cadena; de nuevo el valor propio es cero, ya que las z_i están en el espacio nulo. Los bloques con valores propios diferentes de cero ya se habían terminado en el Paso 1, los bloques con valores propios iguales a cero aumentan por una fila y una columna por Paso 2 y finalmente el Paso 3 contribuye con cualesquiera de los bloques de tamaño 1×1 , $J_i = [0]$.

Bajo esta construcción, hay un único punto técnico que verificar; la independencia de toda la colección de vectores w_i, y_i y z_i . Esto se logra suponiendo que alguna combinación de ellos es cero, es decir:

$$\sum_{i=1}^r \alpha_i w_i + \sum_{i=1}^p \beta_i y_i + \sum_{i=1}^{n-r-p} \gamma_i z_i = 0, \quad (2.44)$$

Al multiplicar por A , y usando las relaciones dadas en la ec. (2.43) para w_i , se tiene que:

$$\sum_{i=1}^r \alpha_i \begin{bmatrix} w_i \\ 0 \\ \lambda_i w_i + w_{i-1} \end{bmatrix} + \sum_{i=1}^p \beta_i A y_i = 0, \quad (2.45)$$

dado que los $A y_i$ son los w_i especiales al final de la cadena correspondiente a λ_0 , no aparecen en la primera suma, además la ec. (2.1) es un tipo de combinación de los w_i , los cuales son independientes por la hipótesis de inducción (dan la forma de Jordan en el espacio columna). Por tanto, se concluye que cada β_i debe ser cero, así, de la ec. (2.44), se tiene que:

$$\sum_{i=1}^r \alpha_i w_i = - \sum_{i=1}^{n-r-p} \gamma_i z_i. \quad (2.46)$$

Como el lado izquierdo está en el espacio columna y los vectores z_i son independientes de ese espacio, entonces cada γ_i debe ser igual a cero. Por lo tanto,

$$\sum_{i=1}^r \alpha_i w_i = 0, \quad (2.47)$$

y es claro que de la independencia de los vectores w_i , se tiene $\alpha_i = 0$.

Por otra parte, si la matriz A inicial no es singular, entonces se pueden aplicar los tres pasos a la transformación $\hat{A} = A - c I_n$ (la constante c se elige de manera que \hat{A} sea singular y que pueda ser cualquiera de los valores propios de A). El algoritmo pasa \hat{A} a su forma de Jordan, es decir, $P^{-1} \hat{A} P = \hat{J}$, al producir las cadenas u_i de las w_i, y_i y z_i . Entonces la forma de Jordan de A utiliza las mismas cadenas y la misma matriz P , tal que

$$P^{-1} A P = P^{-1} (\hat{A} + c I_n) P = \hat{J} + P^{-1} c P = \hat{J} + c I_n = J, \quad (2.48)$$

lo que completa la demostración de que cada matriz A es similar a alguna matriz de Jordan J , salvo el reordenamiento de los bloques, es similar a solo una J , por lo que se concluye que A tiene una forma de Jordan única. ■

2.1.1. Ejemplos numéricos. Forma canónica de Jordan

A continuación, se realizarán unos ejemplos que servirán para entender mejor los resultados.

Ejemplo 1. Sea A una matriz cuadrada de la forma:

$$A = \begin{pmatrix} 0 & 3 & 1 \\ 2 & -1 & -1 \\ -2 & -1 & -1 \end{pmatrix}. \quad (2.49)$$

La ecuación característica de esta matriz es:

$$|A - \lambda I| = \lambda^3 - 2\lambda^2 + 4\lambda + 8 = (\lambda - 2)(\lambda + 2)^2 = 0, \quad (2.50)$$

con valores propios $\lambda_1 = 2$ y $\lambda_2 = -2$ (doble).

El subespacio propio asociado a $\lambda_1 = 2$ es $\ker(A - 2I)$, que es el subespacio definido por las ecuaciones:

$$\begin{aligned} \begin{pmatrix} -2 & 3 & 1 \\ 2 & -3 & -1 \\ -2 & -1 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} &\Leftrightarrow \begin{pmatrix} 0 & 0 & 0 \\ 2 & -3 & -1 \\ 0 & -4 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \\ &\Leftrightarrow \begin{cases} 2x_1 - 3x_2 - x_3 = 0 \\ -4x_2 - 4x_3 = 0 \end{cases}, \\ &\Leftrightarrow \begin{cases} x_1 = x_2 \\ x_3 = -x_2 \end{cases}, \end{aligned} \quad (2.51)$$

donde

$$\ker(A - 2I) = \{x \in \mathbb{R} \mid (A - 2I)x = 0\} = \left\{ \begin{pmatrix} x_1 \\ x_1 \\ -x_1 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \mid x_1 \in \mathbb{R} \right\}, \quad (2.52)$$

entonces, $u_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$ es un vector propio asociado a $\lambda_1 = 2$.

Ahora, para el valor propio $\lambda_2 = -2$, el $\ker(A + 2I)$, que es el subespacio vectorial definido por las ecuaciones:

$$\begin{aligned} \begin{pmatrix} 2 & 3 & 1 \\ 2 & 1 & -1 \\ -2 & -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} &\Leftrightarrow \begin{pmatrix} 2 & 3 & 1 \\ 2 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \\ &\Leftrightarrow \begin{pmatrix} 2 & 3 & 1 \\ 0 & -2 & -2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \end{aligned} \quad (2.53)$$

Luego,

$$\ker(A + 2I) = \left\{ x_1 \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \mid x_1 \in \mathbb{R} \right\},$$

la multiplicidad algebraica de λ_1 es 1 y la de λ_2 es 2, lo que implica que A no es diagonalizable, por lo que a la matriz A se le calcula su forma reducida de Jordan.

Si uno calcula el subespacio vectorial $\ker(A + 2I)^2$, sus ecuaciones son:

$$\begin{pmatrix} 2 & 3 & 1 \\ 2 & 1 & -1 \\ -2 & -1 & 1 \end{pmatrix}^2 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 8 & 8 & 0 \\ 8 & 8 & 0 \\ -8 & -8 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad (2.54)$$

$$\Leftrightarrow x_1 = -x_2.$$

Por lo tanto,

$$\ker(A + 2I)^2 = \left\{ x_1 \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \mid x_1, x_3 \in \mathbb{R} \right\}, \quad (2.55)$$

es el subespacio vectorial de dimensión 2.

Sea $u_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ un vector del $\ker(A + 2I)^2$ que no pertenece al $\ker(A + 2I)$, tomando

$$u_2 = (A + 2I)u_3 = \begin{pmatrix} 2 & 3 & 1 \\ 2 & 1 & -1 \\ -2 & -1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \quad (2.56)$$

se tiene que $\{u_1, u_2, u_3\}$ es una base de \mathbb{R}^3 formada por valores propios generalizados, donde la matriz de Jordan de A es:

$$J = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & -2 \end{pmatrix}, \quad (2.57)$$

y la matriz de paso es:

$$M = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ -1 & 1 & 1 \end{pmatrix}. \quad (2.58)$$

Lo cual cumple que:

$$J = M^{-1} A M. \quad (2.59)$$

Ejemplo 2. Calcular la forma de Jordan de la siguiente matriz:

$$A = \begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & 2 & -1 \\ -1 & -1 & -1 & 0 \\ 2 & 1 & 0 & 3 \end{pmatrix}. \quad (2.60)$$

Sea $\lambda^4 - 4\lambda^3 + 8\lambda^2 - 8\lambda + 4$ el polinomio característico de A , al factorizar, la ecuación característica es:

$$(\lambda(\lambda - 2) + 2)^2 = 0 \Leftrightarrow \lambda(\lambda - 2) + 2 = 0 \Leftrightarrow \lambda^2 - 2\lambda + 2 = 0. \quad (2.61)$$

Así, los valores propios son:

$$\begin{aligned} \lambda_1 &= 1 + i \text{ (doble)} \\ \lambda_2 &= 1 - i \text{ (doble)}. \end{aligned} \quad (2.62)$$

El subespacio propio asociado a $\lambda_1 = 1 + i$ es $\ker(A - \lambda_1 I)$, definido por las ecuaciones:

$$\begin{aligned} \begin{pmatrix} -i & 0 & 1 & -1 \\ 0 & -i & 2 & -1 \\ -1 & -1 & -2-i & 0 \\ 2 & 1 & 0 & 2-i \end{pmatrix}^2 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} &\Leftrightarrow \begin{pmatrix} -2i & 0 & 0 & -1-i \\ 0 & -i & 0 & -i \\ 0 & 0 & 2 & i-1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \\ &\Leftrightarrow \begin{cases} -2ix_1 + (-1-i)x_4 = 0 \\ -ix_2 - ix_4 = 0 \\ 2x_3 + (i-1)x_4 = 0 \end{cases} \Leftrightarrow \begin{cases} x_1 = \frac{(1+i)}{2i}x_4 = \frac{i-1}{2}x_4 \\ x_2 = -x_4 \\ x_3 = \frac{1-i}{2}x_4 \end{cases}, \end{aligned} \tag{2.63}$$

entonces $v_1 = \begin{pmatrix} i-1 \\ -2 \\ 1-i \\ 2 \end{pmatrix}$ es un vector propio asociado a $\lambda_1 = 1 + i$.

Por otra parte, las ecuaciones para el subespacio vectorial $\ker(A - \lambda_1 I)^2$ son:

$$\begin{aligned} \begin{pmatrix} -i & 0 & 1 & -1 \\ 0 & -i & 2 & -1 \\ -1 & -1 & -2-i & 0 \\ 2 & 1 & 0 & 2-i \end{pmatrix}^2 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} &\Leftrightarrow \begin{pmatrix} 1 & 0 & 0 & \frac{1}{2} - i\frac{1}{2} \\ 0 & 1 & 1+i & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \\ &\Leftrightarrow \begin{cases} x_1 = \frac{i-1}{2}x_4 \\ x_2 = (-1-i)x_3 \end{cases}, \end{aligned} \tag{2.64}$$

cuya base es: $\left\{ \begin{pmatrix} 0 \\ -1-i \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} i-1 \\ 0 \\ 0 \\ 2 \end{pmatrix} \right\}$.

Sea $u_1 = \begin{pmatrix} 0 \\ -1-i \\ 1 \\ 0 \end{pmatrix}$ un vector de $\ker(A - \lambda_1 I)^2$ que no pertenece a $\ker(A - \lambda_1 I)$, tomando

$$u_2 = (A - \lambda_1 I)u_1 = \begin{pmatrix} 2 & 3 & 1 \\ 2 & 1 & -1 \\ -2 & -1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ -1-i \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ i+1 \\ -1 \\ -1-i \end{pmatrix}, \tag{2.65}$$

la base para este bloque de Jorda es: $\left\{ \begin{pmatrix} 0 \\ -1-i \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ i+1 \\ -1 \\ -1-i \end{pmatrix} \right\}$. Así, el bloque de Jordan

$\begin{pmatrix} 1+i & 1 \\ 0 & 1+i \end{pmatrix}$ tiene asociada esta base y para $\lambda_2 = 1 - i$, que es el conjugado de λ_1 , el bloque

de Jordan es: $\begin{pmatrix} 1-i & 1 \\ 0 & 1-i \end{pmatrix}$ con base: $\left\{ \begin{pmatrix} 0 \\ -1+i \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1-i \\ -1 \\ -1+i \end{pmatrix} \right\}$.

Finalmente la matriz de Jordan es:

$$J = \begin{pmatrix} 1+i & 1 & 0 & 0 \\ 0 & 1+i & 0 & 0 \\ 0 & 0 & 1-i & 1 \\ 0 & 0 & 0 & 1-i \end{pmatrix}, \tag{2.66}$$

con matriz de paso:

$$M = \begin{pmatrix} 1 & 0 & 1 & 0 \\ i+1 & -1-i & 1+i & -1+i \\ -1 & 1 & -1 & 1 \\ -1-i & 0 & -1+i & 0 \end{pmatrix}, \quad (2.67)$$

cumpléndose que:

$$J = M^{-1} A M. \quad (2.68)$$

2.2. El índice Drazin

Definición 18. Sea $A \in \mathbb{C}^{n \times n}$, se dice que el número entero no negativo k será el índice de la matriz A , si k es el número entero no negativo más pequeño, tal que:

$$\text{rank}(A^k + 1) = \text{rank}(A^k), \quad (2.69)$$

Lo que es equivalente a la dimensión del bloque de Jordán más grande correspondiente al valor propio cero de A . El índice de la matriz A es denotado por $\text{ind}(A)$. Para cualquier matriz $A \in \mathbb{C}^{n \times n}$ la única forma normal de Jordán se puede construir. Por lo tanto, para alguna matriz $A \in \mathbb{C}^{n \times n}$, el índice de A existe y es único.

Definición 19. Sea $A \in \mathbb{C}^{n \times n}$, con $\text{ind}(A) = k$. La matriz X de orden n es la inversa de Drazin de A , denotada por A^D , si X satisface las siguientes condiciones:

$$\begin{aligned} A X &= X A, \\ X A X &= X, \\ A^k X A &= A^k, \end{aligned} \quad (2.70)$$

Cuando $\text{ind}(A) = 1$, A^D se llama grupo inverso de A , y se denota por A_g .

En el siguiente teorema se muestra que para cualquier matriz $A \in \mathbb{C}^{n \times n}$, el inverso de Drazin A^D de A existe y es único [10, 11].

Teorema 8. Sea $A \in \mathbb{C}^{n \times n}$, con $\text{ind}(A) = k$, $\text{rank}(A^k) = r$. Se puede asumir que la forma normal de Jordán de A se define de la siguiente manera:

$$A = P \begin{pmatrix} D & 0 \\ 0 & N \end{pmatrix} P^{-1}, \quad (2.71)$$

entonces,

$$X = A^D = P \begin{pmatrix} D^{-1} & 0 \\ 0 & N \end{pmatrix} P^{-1} \quad (2.72)$$

es la inversa de Drazin, donde P es una matriz no singular, D es una matriz no singular de orden r y N es una matriz nilpotente. Cuando el $\text{ind}(A) = 1$, es obvio que $N = 0$ [10, 11].

Demostración. Sea

$$A^D := P \begin{pmatrix} D^{-1} & 0 \\ 0 & N \end{pmatrix} P^{-1} \quad (2.73)$$

Mediante un cálculo directo se verifica que:

$$\begin{aligned} A A^D &= P \begin{pmatrix} D & 0 \\ 0 & N \end{pmatrix} P^{-1} P \begin{pmatrix} D^{-1} & 0 \\ 0 & N \end{pmatrix} P^{-1} = P \begin{pmatrix} D & 0 \\ 0 & N \end{pmatrix} \begin{pmatrix} D^{-1} & 0 \\ 0 & N \end{pmatrix} P^{-1}, \\ &= P \begin{pmatrix} D D^{-1} & 0 \\ 0 & N N \end{pmatrix} P^{-1} = P \begin{pmatrix} D^{-1} D & 0 \\ 0 & N N \end{pmatrix} P^{-1} \\ &= P \begin{pmatrix} D^{-1} & 0 \\ 0 & N \end{pmatrix} P^{-1} P \begin{pmatrix} D & 0 \\ 0 & N \end{pmatrix} P^{-1} = A^D A, \end{aligned} \quad (2.74)$$

Las otras dos propiedades, ec. (2.70), se verifican de manera similar. Por lo tanto, A^D es la inversa de Drazin. ■

Definición 20. Una matriz $N \in \mathbb{C}^{n \times n}$ es nilpotente si existe $k \in \mathbb{N}$ tal que $N^k = 0$.

Corolario 2. Sea $\text{ind}(A) = 1$ y $D = I$. Está claro que $D^1 = I$, por lo tanto $A = A_g$.

Teorema 9. Si $A \in \mathbb{C}^{n \times n}$ es una matriz no singular, entonces $\text{ind}(A) = 0$.

Demostración. Sea $A \in \mathbb{C}^{n \times n}$ una matriz no singular, recordando que $\text{rank}(A) = \text{rank}(I_n) = n$, entonces el $\text{ind}(A) = 0$. Además, los valores propios de A no son cero. Por lo tanto, si A es una matriz no singular, entonces $\text{ind}(A) = 0$ y $A^D = A^{-1}$, lo cual satisface las condiciones de la Definición 19.

Teorema 10. Si $Ax = b$ es un sistema lineal singular consistente o inconsistente, donde $\text{ind}(A) = k$, entonces el sistema lineal de ecuaciones

$$A^k Ax = A^k b, \tag{2.75}$$

es consistente

Demostración. El sistema lineal $Ax = b$ tiene solución si y solo si $\text{rank}(A) = \text{rank}[A|b]$ [12]. De la Definición 18, ec. (2.69), se tiene que:

$$\text{rank}(A^{k+1}) = \text{rank}(A^{k+1} | A^k b).$$

Por lo tanto, $A^k Ax = A^k b$ es consistente.

De acuerdo con las propiedades de la inversa de Drazin, ec. (2.70), y que $A^k Ax = A^k b$, para obtener la inversa de Drazin, el método de proyección resuelve el sistema lineal consistente o inconsistente $Ax = b$, donde $\text{ind}(A) = k$ a través de la resolución del sistema lineal singular consistente $A^k Ax = A^k b$ [13]. ■

Teorema 11. $A^D b$ es una solución de $Ax = b$, $k = \text{ind}(A)$ si y solo si $b \in R(A^k)$, y $A^D b$ es una solución única de $Ax = b$, siempre que $x \in R(A^k)$.

2.3. La inversa de Drazin en en la resolución de un sistema ecuaciones lineales singular.

En esta sección se explicarán algunas de las aplicaciones de la inversa de Drazin en la resolución de un sistema de ecuaciones lineal singular. Los sistemas de ecuaciones lineales singulares surgen en varias aplicaciones científicas distintas. Una de ellas son las ecuaciones diferenciales parciales discretizadas con los métodos de diferencias finitas o del elemento finito. Los sistemas de ecuaciones lineales singulares grandes se pueden solucionar con métodos de descomposición escasa o con métodos iterativos. Para sistemas lineales singulares consistentes, estos dos enfoques pueden combinarse también en un método que emplea un preconditionamiento de descomposición aproximado para un método iterativo. Pero no se puede utilizar el método de preconditionamiento iterativo para sistemas lineales singulares inconsistentes [13].

Los sistemas de ecuaciones lineales singulares con el índice unitario, surgen en varias aplicaciones, como el modelado de la cadena de Markov y el experimento numérico en la ecuación de Navier-Stokes perturbada. Así, para el sistema lineal singular con índice uno, se debe resolver el sistema

$$A Ax = Ab, \tag{2.76}$$

la cual es consistente. Así se puede elegir todo tipo de métodos preconditionados, por ejemplo: el método del gradiente conjugado preconditionado (PCG), el preconditionamiento de métodos de Residuo Mínimo Generalizado (PGMRES), entre otros, para resolver el sistema $A Ax = Ab$.

Ahora se darán algunos resultados sobre el índice de matriz y el inverso Drazin.

Teorema 12. Si $A \in \mathbb{C}^{n \times n}$ es una matriz con índice unitario, entonces $\text{rank}(A) = \text{rank}(A_g)$.

Demostración. De $\text{rank}(A_g) = \text{rank}(A_g A A_g) \leq \text{rank}(A A_g) \leq \text{rank}(A)$, por transitividad se tiene que:

$$\text{rank}(A_g) \leq \text{rank}(A). \quad (2.77)$$

Por otra parte, el $\text{rank}(A) = \text{rank}(A A_g A) \leq \text{rank}(A_g A) \leq \text{rank}(A_g)$, entonces

$$\text{rank}(A) \leq \text{rank}(A_g). \quad (2.78)$$

Por lo tanto,

$$\text{rank}(A) = \text{rank}(A_g). \blacksquare \quad (2.79)$$

Corolario 3. Sea $A \in \mathbb{C}^{n \times n}$ donde $\text{ind}(A) > 1$. Se tiene que $\text{rank}(A^D) < \text{rank}(A)$.

Teorema 13. Sea $A \in \mathbb{C}^{n \times n}$, entonces A_g existe si y solo si $\text{rank}(A) = \text{rank}(A^2)$ [2].

Definición 21. Sea $P \in \mathbb{C}^{n \times n}$, se dice que es idempotente si $P^2 = P$.

Corolario 4. El índice de cualquier matriz idempotente es igual a uno. El recíproco no siempre se cumple.

Demostración. Sea $P \in \mathbb{C}^{n \times n}$ es idempotente, entonces $P^2 = P$ y $\text{rank}(P^2) = \text{rank}(P)$, por lo que: $\text{ind}(P) = 1$.

Ahora, se puede ver un ejemplo cuando el recíproco no se cumple. Sea

$$A = \begin{pmatrix} 3 & 0 \\ 0 & 0 \end{pmatrix}, \quad (2.80)$$

es fácil ver que:

$$A^2 = \begin{pmatrix} 9 & 0 \\ 0 & 0 \end{pmatrix}, \quad (2.81)$$

así, se tiene que $\text{rank}(A) = \text{rank}(A^2)$, pero $A^2 \neq A$. \blacksquare

Teorema 14. Si $A \in \mathbb{C}^{n \times n}$ es una matriz simétrica con índice uno, entonces $A_g = (A_g)^T$ (el superíndice T se usa para denotar a la matriz transpuesta).

Demostración. Dado que A_g es un grupo inverso de A , se tiene que:

$$A A_g = A_g A \Rightarrow A_g A A_g = A_g, \quad A A_g A = A, \quad (2.82)$$

como $A = A^T$, entonces se tiene que:

$$(A_g)^T A = A (A_g)^T \Rightarrow (A_g)^T A (A_g)^T = (A_g)^T, \quad A (A_g)^T A = A. \quad (2.83)$$

De las anteriores igualdades, por Definición 19, $(A_g)^T$ es grupo inverso de A . El grupo inverso de la matriz A es único, por lo tanto $A_g = (A_g)^T$. \blacksquare

Corolario 5. Para cualquier matriz $A \in \mathbb{C}^{n \times n}$ con índice uno, $(A_g)_g = A$.

Teorema 15. Si $A \in \mathbb{C}^{n \times n}$ donde el $\text{ind}(A) = k$, entonces $\text{ind}(A) = \text{ind}(A^T)$.

Demostración. Sea $A \in \mathbb{C}^{n \times n}$, dado que $\text{ind}(A) = k$, con k número entero no negativo, por la Definición 18 tenemos que $\text{rank}(A^k + 1) = \text{rank}((A^T)^k)$. De $(A^n)^T = (A^T)^n$, se tiene que:

$$(A^{k+1})^T = (A^T)^{k+1} \Rightarrow (A^k)^T = (A^T)^k, \quad (2.84)$$

así,

$$\text{rank}((A^T)^{k+1}) = \text{rank}((A^T)^k). \quad (2.85)$$

Por lo tanto,

$$\text{ind}(A) = \text{ind}(A^T). \blacksquare \quad (2.86)$$

Teorema 16. Si λ es un eigenvalor de $A \in \mathbb{C}^{n \times n}$, entonces $\frac{1}{\lambda}$ es un eigenvalor de A^D .

Demostración. Partiendo de que $Ax = \lambda x$, con x diferente de cero, se tiene que:

$$\begin{aligned} AA^D x &= \lambda A^D x, \\ A^D A A^D x &= \lambda A^D A^D x. \end{aligned} \tag{2.87}$$

De la Definición 19, se puede obtener que $A^D x = \lambda A^D A^D x$. Ahora, si se pone $A^D x = y$, se tiene que $\frac{1}{\lambda} y = A^D y$. ■

Teorema 17. Si $A \in \mathbb{C}^{n \times n}$ es una matriz nilpotente tal que $A^k = 0$, entonces $ind(A) = k$.

Demostración. De $A^k = 0$, se tiene que:

$$rank(A^k) = rank(0). \tag{2.88}$$

De $A^{k+1} = 0$, con $k \in \mathbb{N}$, se puede obtener que $rank(A^{k+1}) = rank(A^k)$. Por lo tanto,

$$ind(A) = k. \blacksquare \tag{2.89}$$

Teorema 18. Si $A \in \mathbb{C}^{n \times n}$ entonces $0 \leq ind(A) \leq n$.

Demostración. Por el Teorema 9, el $ind(A) = 0$ para una matriz A no singular y por el Teorema 17, el $ind(A) = n$ para una matriz A nilpotente tal que $A^n = 0$.

Sean $\lambda_1, \dots, \lambda_k$ los ceros distintos del polinomio característico de la matriz $A \in \mathbb{C}^{n \times n}$ y $\lambda_1 = 0$ con multiplicidad $\sigma(\lambda_1) = k$. De la Definición 11, se tiene que:

$$\rho(\mu) = (-1)^n (\mu - \lambda_1)^{\sigma_1} (\mu - \lambda_2)^{\sigma_2} \dots (\mu - \lambda_k)^{\sigma_k}. \tag{2.90}$$

Cualquier matriz $A \in \mathbb{C}^{n \times n}$ tiene exactamente n valores propios. Por el Teorema 3, se tiene que:

$$\rho(\lambda_1) \leq \sigma(\lambda_1) < n, \tag{2.91}$$

y por Definición 18, se puede obtener que $0 < ind(A) < n$. Por lo tanto,

$$0 \leq ind(A) \leq n. \blacksquare \tag{2.92}$$

2.3.1. Ejemplos numéricos de la inversa de Drazin.

Ahora se darán unos ejemplos para ilustrar la inversa de Drazin en un espacio de Krylov.

Ejemplo 3. Determinar el índice y la inversa de Drazin de la siguiente matriz

$$A = \begin{pmatrix} 2 & -3 & -5 \\ -1 & 4 & 5 \\ 1 & -3 & -4 \end{pmatrix}. \tag{2.93}$$

Teniendo en cuenta de que:

$$A^2 = \begin{pmatrix} 2 & -3 & -5 \\ -1 & 4 & 5 \\ 1 & -3 & -4 \end{pmatrix}, \tag{2.94}$$

luego $rank(A) = 2 = rank(A^2)$, por lo que $ind(A) = 1$. Además, la matriz A tiene valores propios $\lambda_1 = 0$, $\lambda_2 = 1$ con multiplicidad

$$\begin{aligned} \sigma(\lambda_1) &= 1, & \rho(\lambda_1) &= 1, \\ \sigma(\lambda_2) &= 1, & \rho(\lambda_2) &= 1. \end{aligned} \tag{2.95}$$

Así, la forma normal de Jordán para la matriz A tiene la siguiente forma:

$$J = P A P^{-1} = \begin{pmatrix} [1] & 0 & 0 \\ 0 & [1] & 5 \\ 0 & 0 & [0] \end{pmatrix}, \quad P = \begin{pmatrix} -1 & 3 & 4 \\ -2 & 3 & 5 \\ -1 & 3 & 5 \end{pmatrix}. \quad (2.96)$$

P es una matriz no singular. La dimensión del bloque de Jordán más grande correspondiente al cero valor propio de J y P es igual a uno. Como $A = A^2$, entonces A es una matriz idempotente, por el Corolario 4 se tiene que: $ind(A) = 1$. Por otro lado, por el Teorema 8, se tiene que:

$$A = P^{-1} \begin{pmatrix} D & 0 \\ 0 & N \end{pmatrix} P = \begin{pmatrix} 2 & -3 & -5 \\ -1 & 4 & 5 \\ 1 & -3 & -4 \end{pmatrix}, \quad (2.97)$$

en la que $D = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, es una matriz no singular de orden 2, y $N = 0$, es una matriz nilpotente. Por lo tanto, el grupo inverso de A es:

$$A_g = P^{-1} \begin{pmatrix} D & 0 \\ 0 & N \end{pmatrix} P = \begin{pmatrix} 2 & -3 & -5 \\ -1 & 4 & 5 \\ 1 & -3 & -4 \end{pmatrix}. \quad (2.98)$$

Como $D = I$, por el Corolario 2, se tiene que A_g es grupo inverso de A . Así, el sistema de ecuaciones

$$\begin{cases} 2x_1 - x_2 + x_3 = 1 \\ -3x_1 + 4x_2 - 3x_3 = 1 \\ -5x_1 + 5x_2 - 4x_3 = 2 \end{cases}, \quad (2.99)$$

es inconsistente, $ind(A^T) = 1$.

Está claro que el siguiente sistema es consistente

$$\begin{cases} 2x_1 - x_2 + x_3 = 3 \\ -3x_1 + 4x_2 - 3x_3 = -5 \\ -5x_1 + 5x_2 - 4x_3 = 8 \end{cases}, \quad (2.100)$$

Ejemplo 4. Determinar el índice y la inversa de Drazin de la siguiente matriz:

$$B = \begin{pmatrix} 1 & 1 & 3 \\ 5 & 2 & 6 \\ -2 & -1 & -3 \end{pmatrix}. \quad (2.101)$$

Se tiene que $rank(B^3) = rank(B^4)$, por lo que $ind(B) = 3$. Además, la matriz B tiene los valores propios $\lambda = 0$ con multiplicidad

$$\sigma(\lambda) = 3, \quad \rho(\lambda) = 1. \quad (2.102)$$

La forma normal de Jordán de la matriz B tiene la siguiente forma:

$$J = P B P^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad P = \begin{pmatrix} -1 & 0 & -1 \\ 1 & 0 & 0 \\ 1 & 1 & 3 \end{pmatrix}, \quad (2.103)$$

con P es una matriz no singular. La dimensión del bloque de Jordán más grande correspondiente al valor propio cero de J y P es igual a 3. Además $B^3 = 0$, entonces $ind(B) = 3$. Por el Teorema 8, se tiene que:

$$B = P^{-1} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} P = \begin{pmatrix} 1 & 1 & 3 \\ 5 & 2 & 6 \\ -2 & -1 & -3 \end{pmatrix}. \quad (2.104)$$

Por lo tanto, la inversa de Drazin de B es:

$$B^D = P^{-1} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, P = 0. \quad (2.105)$$

Además, $B \in \mathbb{C}^{3 \times 3}$ es una matriz nilpotente de índice 3, por el Corolario 3 se tiene que $B^D = 0$. Además, las matrices B, B^D satisfacen las condiciones de la Definición 19, B^D es la inversa de Drazin de B .

Ejemplo 5. Considere la siguiente matriz simétrica:

$$C = \begin{pmatrix} -1 & -1 & -1 \\ -1 & \frac{1}{3} & -1 \\ -1 & -1 & -1 \end{pmatrix}. \quad (2.106)$$

La matriz C tiene los valores propios $\lambda_1 = 0, \lambda_2 = 1$ y $\lambda_3 = -\frac{8}{3}$. El índice de la matriz C es igual a uno, porque $\text{rank}(C) = \text{rank}(C^2)$. Así, la forma normal de Jordán de la matriz C tiene la siguiente forma:

$$J = P C P^{-1} = \begin{pmatrix} [1] & 0 & 0 \\ 0 & [-\frac{8}{3}] & 0 \\ 0 & 0 & [0] \end{pmatrix}, \quad P = \begin{pmatrix} \frac{1}{11} & -\frac{3}{11} & \frac{1}{9} \\ -\frac{22}{11} & -\frac{11}{11} & -\frac{11}{11} \\ \frac{1}{2} & 0 & -\frac{1}{2} \end{pmatrix}, \quad (2.107)$$

donde P es una matriz no singular. La dimensión del bloque de Jordán más grande correspondiente al valor propio cero de J , P es igual a uno. Por el Teorema 8 se tiene que:

$$C_g = P^{-1} J P = \frac{1}{16} \begin{pmatrix} -1 & -6 & -1 \\ -6 & 12 & -6 \\ -1 & -6 & -1 \end{pmatrix}. \quad (2.108)$$

C, C_g satisfacen las condiciones de la Definición 19, por lo que C_g es grupo inverso de C . Está claro que $\text{rank}(C) = \text{rank}(C_g)$ y $(C_g)_g = C$. Ahora, considere el siguiente sistema de ecuaciones lineales singular con índice uno

$$\begin{cases} -x_1 + x_2 - x_3 = 3 \\ -x_1 + \frac{1}{3}x_2 - x_3 = 1 \\ -x_1 - x_2 - x_3 = 3 \end{cases}, \quad (2.109)$$

tomando $b = \begin{pmatrix} 3 \\ 1 \\ 3 \end{pmatrix}$, la solución del sistema lineal singular con índice uno es:

$$x = C_g b = \begin{pmatrix} -\frac{1}{16} & -\frac{6}{16} & -\frac{1}{16} \\ -\frac{1}{16} & \frac{12}{16} & -\frac{1}{16} \\ -\frac{1}{16} & -\frac{6}{16} & -\frac{1}{16} \end{pmatrix} \begin{pmatrix} 3 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} -\frac{3}{4} \\ \frac{3}{4} \\ -\frac{3}{4} \end{pmatrix}. \quad (2.110)$$

Capítulo 3

La inversa de Drazin en el método de Krylov

Se ha hablado de la inversa de Drazin de manera analítica y se han presentado algunos ejemplos para determinar el índice y la inversa de Drazin. Como se ha mencionado antes, trabajar en un espacio de Krylov nos da la oportunidad de converger rápidamente a una solución cuando se trata de un sistema $Ax = b$ de grado n . Se ha visto que los valores propios juegan un papel central cuando se trata de asegurar la existencia y la singularidad de las soluciones de Krylov. En este caso, se describe una clase de lado derecho para los que la solución se encuentra en un espacio de Krylov. En este caso, sólo hay una solución que se encuentra en un espacio de Krylov, y se puede obtener a partir de la inversa de Drazin.

Sea un espacio Krylov generado por un vector c ,

$$\mathcal{K}_k(A, c) \equiv \text{span} \{c, Ac, \dots, A^{k-1}c\}. \quad (3.1)$$

Uno se limita a los espacios $\mathcal{K}_k(A, c)$ de Krylov que son generados por el lado derecho b de un sistema lineal $Ax = c$.

3.1. Ejemplo de un método de Krylov.

El método de residuos mínimos generalizados (GMRES) fue publicado por Saad y Schultz en 1986 [14]. En la iteración $k \geq 1$, GMRES escoge la "mejor" solución x_k del espacio $\mathcal{K}_k(A, b)$ de Krylov. "Mejor" significa que el residuo es lo más pequeño posible sobre $\mathcal{K}_k(A, b)$, es decir, x_k resuelve el problema de los mínimos cuadrados

$$\text{mín} \|b - Az\|, \quad z \in \mathcal{K}_k(A, b), \quad (3.2)$$

en la norma euclídeana $\|\cdot\|$.

GMRES resuelve este problema de mínimos cuadrados construyendo una base ortonormal $\{v_1, v_2, \dots, v_k\}$ para $\mathcal{K}_k(A, b)$ usando el método de Arnoldi, que es una versión del procedimiento de Gram-Schmidt adaptado a los espacios de Krylov. Comenzando con el lado derecho normalizado $v_1 = \frac{b}{\|b\|}$ como base para $\mathcal{K}_1(A, b)$, el método de Arnoldi construye recursivamente una base ortonormal para $\mathcal{K}_{j+1}(A, b)$ a partir de una base ortonormal para $\mathcal{K}_j(A, b)$. Ortogonalizando el vector Av_j de $\mathcal{K}_{j+1}(A, b)$ contra el espacio anterior $\mathcal{K}_j(A, b)$, es decir,

$$\hat{v}_{j+1} = Av_j - (h_{1j}v_1 + \dots + h_{jj}v_j), \quad (3.3)$$

La inversa de Drazin en el método de Krylov
3.1 Ejemplo de un método de Krylov.

donde $h_{ij} = v_i^* A v_j$ y $*$ se usa para denotar la transpuesta conjugada. El nuevo vector base es:

$$v_{j+1} = \frac{\hat{v}_{j+1}}{\|\hat{v}_{j+1}\|}. \quad (3.4)$$

Si se recogen los vectores de base ortonormal para $\mathcal{K}_j(A, b)$ en una matriz, $V_j = (v_1 \cdots v_j)$, se obtiene que la descomposición asociada al método de Arnoldi es:

$$A V_j = V_{j+1} H_j, \quad (3.5)$$

donde H_j es una matriz superior de Hessenberg de tamaño $(j+1) \times j$ (una matriz triangular superior con una subdiagonal adicional debajo de la diagonal).

En el contexto del problema de los mínimos cuadrados esto significa que: Si $z \in \mathcal{K}_k(A, b)$, entonces $z = V_k y$ para alguna y , entonces:

$$A z = A V_k y = V_{k+1} H_k y \quad y \quad b = \beta v_1 = \beta V_{k+1} e_1, \quad (3.6)$$

donde $\beta = \|b\|$ y e_1 es la primera columna de la matriz de identidad. El problema de los mínimos cuadrados en la iteración k de GMRES se reduce a:

$$\min \|b - A z\| = \min_y \|\beta e_1 - H_k y\|, \quad z \in \mathcal{K}_j(A, b). \quad (3.7)$$

Así pues, el GMRES procede de la siguiente manera:

1. **Iteración 0:** Iniciar $x_0 = 0$, $v_1 = \frac{b}{\beta}$ y $V_1 = v_1$,
2. **Iteración $k \geq 1$:**
 - a) Ortogonaliza: $\hat{v}_{k+1} = A v_k$, donde $h_k = V_k^* A v_k$.
 - b) Normalizar: $v_{k+1} = \frac{\hat{v}_{k+1}}{\|\hat{v}_{k+1}\|}$.
 - c) Actualizar: $V_{k+1} = (V_k \ v_{k+1})$, $H_k = \begin{pmatrix} H_{k+1} & h_k \\ 0 & \|\hat{v}_{k+1}\| \end{pmatrix}$, donde la primera columna de H_k se omite cuando $k = 1$.
 - d) Resuelve el problema de los mínimos cuadrados $\min_y \|\beta e_1 - H_k y\|$, y llama a la solución y_k .
 - e) La solución aproximada es: $x_k = V_k y_k$.

El GMRES se detiene cuando produce un vector cero. Sea s el primer índice para el cual $\hat{v}_{s+1} = 0$. Si $s = 0$ entonces claramente $b = 0$ y $x_0 = 0$. En este caso, GMRES ha encontrado la solución a $A x = b$.

Si $s > 0$ entonces la última fila de H_s es cero. Se debe dejar que \hat{H}_s sea H_s sin su última fila. El método de Arnoldi implica que $A V_s = V_s \hat{H}_s$. Esto significa que las columnas de V_s abarcan un subespacio invariante de A y los valores propios de \hat{H}_s son valores propios de A . Como A no tiene valores propios cero, tampoco \hat{H}_s . Por lo tanto, \hat{H}_s es no singular, y el problema de los mínimos cuadrados se reduce a un sistema lineal no singular $\hat{H}_s y_s = \beta e_1$. De $A V_s = V_s \hat{H}_s$ se sigue que:

$$A V_s y_s = V_s \hat{H}_s y_s = \beta V_s e_1 = b, \quad (3.8)$$

y $x_s = V_s y_s$ es la solución a $A x = b$. De nuevo, GMRES ha encontrado la solución.

Nóte que s no puede sobrepasar n porque el espacio de dimensión n puede acomodar muchos vectores linealmente independientes. Por lo tanto, GMRES trabaja correctamente (uno se restringe a la aritmética de punto flotante).

La inversa de Drazin en el método de Krylov

3.1 Ejemplo de un método de Krylov.

En la práctica, un método Krylov como el GMRES no se ejecuta hasta su finalización, sino que se termina prematuramente cuando se considera que el número de iteraciones es lo suficientemente bueno para lo que nos interesa. Esto puede significar que la norma residual $\|Ax_k - b\|$ es suficientemente pequeña o que algún otro criterio de convergencia se cumple.

Al igual que el GMRES, existen otros métodos de Krylov, por ejemplo: el gradiente conjugado, el residual conjugado, la biortogonalización de Lanczos, el residual cuasimínimo (QMR), el gradiente biconjugado y los métodos de dirección A -conjugada, los cuales tienden a proporcionar soluciones aceptables en un número de iteraciones mucho menor que el orden de A . El número de iteraciones necesarias depende de los valores propios de A , y la naturaleza de esta dependencia es crucial para comprender los métodos de Krylov.

La resolución de $Ax = b$ tiene una representación natural como miembro de un espacio de Krylov $\mathcal{K}_k(A, b)$, entonces se puede entender por qué se construyen aproximaciones a x desde este espacio. Si la dimensión de $\mathcal{K}_k(A, b)$ es pequeña, entonces el método de Krylov tiene la oportunidad de encontrar x en pocas iteraciones. Por eso se selecciona como indicador de convergencia a la dimensión del espacio de Krylov más pequeño $\mathcal{K}_k(A, b)$ que contiene x . Si esta dimensión es pequeña, se tiene una razón plausible para esperar una rápida convergencia (en la práctica, la convergencia puede juzgarse no sólo por el número de iteraciones sino también por alguna estimación de la reducción del error).

Cuando se trata con matrices no singulares, se debe utilizar el polinomio mínimo de la matriz de coeficientes A para expresar A^{-1} en términos de potencias de A . Esto arroja la solución $x = A^{-1}b$ automáticamente como miembro de un espacio Krylov. La dimensión de este espacio es el grado del polinomio mínimo de A .

A continuación, se consideran los sistemas lineales cuya matriz de coeficientes A es singular. Para estar seguros de una solución que se encuentra en un espacio Krylov $\mathcal{K}_k(A, b)$, se debe confinar el lado derecho b a la *parte no singular* de A y mantenerlo alejado de la *parte nilpotente*. Como resultado, la dimensión del espacio Krylov se reduce: Es el grado del polinomio mínimo de A menos el índice del valor propio cero. También resulta que sólo hay una única solución que se encuentra en el espacio de Krylov $\mathcal{K}_n(A, b)$.

El polinomio mínimo $q(t)$ de A es el único polinomio monico de grado mínimo tal que $q(A) = 0$. Se construye a partir de los valores propios de A de la siguiente manera. Si los distintos valores propios de A son λ_j , y si λ_j tiene el índice m_j (del tamaño del mayor bloque de Jordan asociado con λ_j), entonces la suma de todos los índices es:

$$m = \sum_{j=1}^d m_j, \quad \text{y} \quad q(t) = \prod_{j=1}^d (t - \lambda_j)^{m_j}. \quad (3.9)$$

Ejemplo 6. Considere la matriz

$$A = \begin{pmatrix} 3 & 1 & \cdots & \cdots \\ \vdots & 3 & \ddots & \vdots \\ \vdots & \ddots & 4 & \vdots \\ \cdots & \cdots & \cdots & 4 \end{pmatrix}, \quad (3.10)$$

la cual tiene un valor propio 3 de índice 2 y un valor propio 4 de índice 1, así que $m = 3$ y $q(t) = (t - 3)^2(t - 4)$. Cuando A es diagonalizable, m es el número de valores propios distintos de A . Cuando A es un bloque de Jordan de orden n , entonces $m = n$.

Está claro por ec. (3.9) que al escribir:

$$q(t) = \sum_{j=0}^m \alpha_j t^j, \quad (3.11)$$

donde el término constante es $\alpha_0 = \prod_{j=1}^d (-\lambda_j)^{m_j}$. Por lo tanto, $\alpha_0 \neq 0$ si y solo si A es no singular.

Lo importante aquí es usar el polinomio mínimo para representar la inversa de una matriz no singular A en término de las potencias de A , como:

$$0 = q(A) = \alpha_0 I + \alpha_1 A + \cdots + \alpha_m A^m, \quad (3.12)$$

donde I es la matriz identidad y $\alpha_0 \neq 0$, deduciendo que:

$$A^{-1} = -\frac{1}{\alpha_0} \sum_{j=0}^{m-1} \alpha_{j+1} A^j. \quad (3.13)$$

En consecuencia, cuanto menor sea el grado del polinomio mínimo, más corta será la expresión de A^{-1} . Esta expresión de A^{-1} indica inmediatamente que $x = A^{-1}b$ es un miembro de un espacio Krylov.

Teorema 19. *Si el polinomio mínimo de la matriz no singular A tiene el grado m , entonces la solución de $Ax = b$ se encuentra en el espacio $\mathcal{K}_m(A; b)$.*

Por lo tanto, en ausencia de cualquier información sobre b , se tiene que asumir que la dimensión del espacio de Krylov más pequeño que contiene x es m , el grado del polinomio mínimo de A . Si el polinomio mínimo tiene un grado bajo, entonces el espacio Krylov que contiene la solución es pequeño y un método Krylov tiene la oportunidad de converger rápidamente.

Ejemplo 7. *El Teorema 19 sugiere que un espacio de Krylov debería tener una dimensión máxima cuando la matriz es un bloque no singular de Jordan, porque en este caso el polinomio mínimo tiene un grado máximo. Es necesario averiguar que hace GMRES con $Ax = b$ cuando*

$$A = \begin{pmatrix} 2 & 1 & \cdots & \cdots \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ \cdots & \cdots & \cdots & 2 \end{pmatrix} \quad y \quad b = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}. \quad (3.14)$$

Supóngase que A tiene orden n , y se denotan las columnas de la matriz identidad de orden n por e_1, \dots, e_n . Entonces $b = e_n$.

1. **Iteración 0:**

$$v_1 = b = e_n. \quad (3.15)$$

2. **Iteración 1:**

$$h_{11} = v_1^* A v_1 = e_n^* A e_n = 2 \quad (3.16)$$

y,

$$v_2 = \hat{v}_2 = (A - h_{11} I) v_1 = (A - 2I) e_n = e_{n-1}. \quad (3.17)$$

3. **Iteración 2:**

$$h_{12} = v_1^* A v_2 = e_n^* A e_{n-1} = 0, \quad h_{22} = v_2^* A v_2 = e_{n-1}^* A e_{n-1} = 2, \quad (3.18)$$

y

$$v_3 = \hat{v}_3 = (A - h_{22} I) v_2 = (A - 2I) e_{n-1} = e_{n-2}. \quad (3.19)$$

4. Iteración n:

$$h_{1n} = \dots = h_{n-1n} = 0, \quad h_{nn} = v_n^* A v_n = e_1^* A e_1 = 2, \quad (3.20)$$

y

$$v_{n+1} = \hat{v}_{n+1} = (A - h_{nn} I) v_n = (A - 2I) e_1 = 0. \quad (3.21)$$

Saad y Schultz [14], demostraron que el número máximo de iteraciones en el GMRES no excede el grado del polinomio mínimo de A .

En este punto, uno puede preguntarse ¿Por que $\mathcal{K}_k(A, b)$ suele ser un buen espacio para construir una solución aproximada? y ¿Por qué los valores propios son importantes para los métodos de Krylov para las matrices no singulares? $\mathcal{K}_m(A, b)$ es un buen espacio para construir soluciones aproximadas de un sistema lineal de ecuaciones no singular $Ax = b$, ya que está íntimamente ligado al inverso de la matriz. Los valores propios son importantes para los métodos de Krylov porque la dimensión del espacio de solución es determinado por el grado del polinomio mínimo de la matriz.

No obstante, para complementar la respuesta a las preguntas, es necesario mirar las matrices singulares. Aunque los sistemas singulares no son tan abundantes en la práctica como los sistemas no singulares, sí se dan en la ref. [4].

Suponiendo que un sistema lineal tiene una matriz de coeficientes singulares. Incluso si existe una solución, puede que no esté en el espacio de Krylov $\mathcal{K}_n(A, b)$.

Ejemplo 8. Sea $Nx = c$ un sistema lineal consistente, donde N es una matriz nilpotente y $c \neq 0$, existe un i tal que $N^i = 0$ pero N^{i-1} distinto de cero. Suponiendo que una solución a $Nx = c$ es una combinación lineal de vectores de Krylov, es decir, $x = \epsilon_0 c + \epsilon_1 Nc + \dots + \epsilon_{i-1} N^{i-1}c$, entonces:

$$c = Nx = \epsilon_0 Nc + \dots + \epsilon_{i-2} N^{i-1}c, \quad y \quad (I - \epsilon_0 N - \dots - \epsilon_{i-2} N^{i-1})c = 0. \quad (3.22)$$

La matriz entre paréntesis es no singular. Sus valores propios son todos iguales a uno, porque la suma de los términos que contienen N es nilpotente. En consecuencia, $c = 0$. En otras palabras, una solución a un sistema nilpotente con un lado derecho de distinto de cero no puede estar en el espacio de Krylov $\mathcal{K}_n(A, b)$. Esta observación es importante porque sugiere que si se quiere que la solución de un sistema general cuadrado $Ax = b$ se encuentre en un espacio Krylov, se debe restringir b manteniéndolo de alguna manera alejado de la *parte nilpotente* de A .

Para esto se debe descomponer el espacio en $C^n = R(A^i) \oplus N(A^i)$, donde i es el índice del eigenvalor cero de $A \in \mathbf{C}^{n \times n}$, y donde $R(\cdot)$ y $N(\cdot)$ denotan rango y espacio nulo. Asumiendo que A es una matriz de Jordan con todos los eigenvalores cero en la parte inferior. Entonces la descomposición de espacio induce la descomposición de la matriz

$$A = \begin{pmatrix} C & 0 \\ 0 & N \end{pmatrix}, \quad (3.23)$$

donde C es no singular y N es nilpotente de índice i .

Ahora, suponiendo que $Ax = b$ tiene una solución de Krylov

$$x = \sum_{j=0}^p \alpha_j A^j b = \sum_{j=0}^p \alpha_j \begin{pmatrix} C^j & 0 \\ 0 & N \end{pmatrix} b. \quad (3.24)$$

Partiendo de los vectores de acuerdo con la matriz,

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \quad (3.25)$$

La inversa de Drazin en el método de Krylov
3.1 Ejemplo de un método de Krylov.

se tiene que:

$$x_1 = \sum_{j=0}^p \alpha_j C^j b_1, \quad y x_2 = \sum_{j=0}^p \alpha_j N^j b_2, \quad (3.26)$$

pero $Ax = b$ implica que $Nx_2 = b_2$, así que:

$$N \left(\sum_{j=0}^p \alpha_j N^j b_2 \right) = b_2, \quad y \quad \left(I - \sum_{j=0}^p \alpha_j N^{j+1} \right) b_2 = 0. \quad (3.27)$$

La matriz entre paréntesis es no singular y $b_2 = 0$. Por lo tanto,

$$b = \begin{pmatrix} b_1 \\ 0 \end{pmatrix} \in R(A^i), \quad y \quad x = \begin{pmatrix} C^{-1} b_1 \\ 0 \end{pmatrix}, \quad (3.28)$$

es la solución de $Ax = b$.

Como se ha confinado el lado derecho de la *parte no singular* de A , se puede aplicar la idea del ejemplo después del Teorema 19 a la matriz C . El polinomio mínimo para C tiene el grado $m - i$, y hay un polinomio $p(x)$ del grado $m - i - 1$ tal que $C^{-1} = p(C)$. Sustituyendo este polinomio en la expresión para x , se tiene que:

$$x = \begin{pmatrix} C^{-1} b_1 \\ 0 \end{pmatrix} = \begin{pmatrix} p(C) & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} b_1 \\ 0 \end{pmatrix} = \begin{pmatrix} p(C) & 0 \\ 0 & p(N) \end{pmatrix} b = p(A) b \in \mathcal{K}_{m-i}(A, b). \quad (3.29)$$

Por lo tanto, $b \in R(A^i)$ garantiza la existencia de una solución de Krylov. El siguiente teorema resume nuestro hallazgo hasta ahora.

Teorema 20. (*Existencia de una solución de Krylov*) *Un sistema lineal $Ax = b$ tiene una solución de Krylov si y solo si $b \in R(A^i)$, donde i es el índice del valor propio cero de A .*

En otras palabras, un sistema lineal tiene una solución de Krylov si y solo si el lado derecho se mantiene alejado de la *parte nilpotente* y es confinado a la *parte no singular*.

En el caso especial en que A no es un singular, $i = 0$ y la condición en b es vacía. Cuando A tiene un valor propio cero no defectuoso, $i = 1$ y la condición en b se reduce a la condición conocida de consistencia $b \in R(A)$. Esto ocurre, por ejemplo, cuando A es diagonalizable. En este caso un sistema consistente $Ax = b$ tiene una solución:

$$x \in \begin{cases} \mathcal{K}_{d-1}(A, b) & \text{sí } A \text{ es singular} \\ \mathcal{K}_d(A, b) & \text{sí } A \text{ es no singular} \end{cases}, \quad (3.30)$$

donde d es el número de eigenvalores de A .

Comparado con el caso no singular, el mayor espacio de Krylov para el caso singular se ha reducido. Su dimensión es menor por i que el grado del polinomio mínimo. El índice del valor propio cero afecta la dimensión del espacio de búsqueda, así como la dimensión del espacio de los lados derechos que admiten una solución en $\mathcal{K}_n(A, b)$. En particular, a medida que aumenta la deficiencia del valor propio cero, el espacio de búsqueda se reduce, y también lo hace el espacio de los lados derechos deseables. Esto responde también por qué son importantes los eigenvalores para los métodos de Krylov. Los valores propios son importantes para los métodos de Krylov porque el índice del valor propio cero afecta la existencia decuna solución en $\mathcal{K}_n(A, b)$.

Un espacio de Krylov es un buen espacio para construir una solución aproximada para $Ax = b$, en el caso no singular se argumenta que A^{-1} puede ser expresado como un polinomio en A y por lo tanto está íntimamente ligado a $\mathcal{K}_n(A, b)$. Pero cuando la matriz es singular no se tiene un inverso,

así que se busca un pseudoinverso adecuado. En primera instancia uno podría pensar en el inverso de Moore-Penrose de A , pero esto no funciona, dado que el inverso de Moore-Penrose generalmente no puede expresarse como un polinomio en A , ref.[15] sección 7.5, es por esta razón que se probó con la inversa de Drazin. Así, para matrices singulares, un espacio de Krylov $\mathcal{K}_k(A, b)$ es un buen espacio a partir del cual construir una solución aproximada a un sistema singular $Ax = b$ porque cuando es lo suficientemente grande, contiene una solución pseudoinversa única (siempre que b se encuentre en $R(A^i)$).

Combinando todos los resultados se obtiene una declaración completa sobre las soluciones de Krylov en $\mathcal{K}_n(A, b)$. Sea m el grado del polinomio mínimo para $A \in \mathbf{C}^{n \times n}$, y sea i el índice del valor propio cero de A .

- El sistema lineal $Ax = b$ tiene una solución de Krylov en $\mathcal{K}_n(A, b)$ si y sólo si $b \in R(A^i)$.
- Cuando existe una solución de Krylov, es única y es la solución inversa de Drazin,

$$x = A^D b \in \mathcal{K}_{m-i}(A, b). \quad (3.31)$$

- Cada sistema consistente $Ax = b$, con matriz de coeficientes diagonalizable A tiene una solución de Krylov,

$$x = A^D b \in \begin{cases} \mathcal{K}_{d-1}(A, b) & \text{si } A \text{ es singular} \\ \mathcal{K}_d(A, b) & \text{si } A \text{ es no singular} \end{cases}, \quad (3.32)$$

con d el número de los distintos valores propios de A .

La discusión anterior no explica completamente la popularidad de los métodos de Krylov. En la práctica, no basta con saber que la dimensión de un espacio de Krylov está limitada por n , porque n puede ser muy grande y la dimensión del espacio de búsqueda puede ser igual a n . Por ejemplo, las matrices almacenadas en aritmética de precisión finita tienden a ser no singulares con eigenvalores distintos, lo que da lugar a un espacio de búsqueda de dimensión máxima. En el caso de los grandes sistemas lineales, no es práctico ejecutar ni siquiera cerca de n iteraciones. En consecuencia, los algoritmos de Krylov se utilizan como métodos iterativos. Esto significa que se terminan prematuramente, mucho antes de que se hayan completado todas las n iteraciones. La otra mitad de la historia gira en torno a la cuestión de cómo asegurar que un pequeño número de iteraciones entregue una solución aproximada que sea razonablemente precisa.

La afirmación dada en ec. (3.30) proporciona la pista. Supóngase que fuera posible encontrar una matriz no singular M , que hace a la MA diagonalizable con sólo unos pocos valores propios distintivos. Entonces, se espera encontrar una solución a $MAx = Mb$ en un espacio Krylov de pequeña dimensión. Premultiplicar (o postmultiplicar) el sistema lineal para reducir el número de iteraciones en un método de Krylov se llama preconditionamiento. Por supuesto, hay una delicada compensación entre la reducción del espacio de búsqueda contra el costo de obtener el preconditionador M . Considere, por ejemplo, el caso extremo $M = A^{-1}$. El espacio de búsqueda es mínimo (tiene dimensión uno), pero la construcción del preconditionador es tan costosa como la solución del sistema original.

Aunque encontrar una matriz MA diagonalizable con pocos valores propios distintivos puede no ser barato, uno puede ser capaz de explotar la estructura del problema físico subyacente para construir preconditionadores que entreguen una MA diagonalizable cuyos valores propios caigan en unos pocos grupos. Por ejemplo: si los diámetros de los grupos son lo suficientemente pequeños, entonces MA se comporta numéricamente como una matriz con t valores propios distintos. Como resultado, se espera que las iteraciones del método de Krylov produzcan una aproximación razonablemente precisa. Mientras que la intuición es simple, los argumentos rigurosos no siempre son fáciles de establecer. Diferentes algoritmos requieren diferentes técnicas, y esto ha sido el foco

La inversa de Drazin en el método de Krylov
3.1 Ejemplo de un método de Krylov.

de mucho trabajo. Las ideas para GMRES pueden ser consultadas en la ref.[4]. Construir buenos preconditionadores y luego probar que realmente funcionan es la otra mitad de la historia de Krylov, y esto sigue siendo un área activa de investigación en el análisis numérico.

Capítulo 4

Conclusiones

Finalmente, se presentan algunas conclusiones sobre los resultados obtenidos a partir del Método de Krylov mediante la inversa de Drazin.

Matrices Hermitianas. Se da sólo una respuesta superficial a ¿Por qué los métodos de Krylov suelen funcionar tan bien para las matrices hermitianas? En primer lugar, un sistema lineal consistente $Ax = b$ con la matriz de coeficientes hermitiana A siempre tiene una solución de Krylov. Segundo, la matriz de vectores propios de una matriz hermitiana puede ser elegida unitaria, por lo tanto, está bien condicionada. Si la matriz hermitiana A también es definida positiva, el número de iteraciones necesarias para producir una solución satisfactoria tiende a ser pequeño.

Otra razón es la eficiencia. Tomando el GMRES, por ejemplo, cuando A es Hermitiana, $V_j^* A V_j$ es también Hermitiano y H_j es tridimensional. Por consiguiente, el recuento de operaciones de una iteración del GMRES es independiente del número de iteraciones. Por lo tanto, el costo de las iteraciones t del GMRES es proporcional al costo de sólo t productos de la matriz-vector. Como el GMRES, muchos otros métodos de Krylov son igualmente baratos cuando se aplican a una matriz hermitiana.

Una conjetura inicial no cero. Muchos métodos de Krylov expresan los iterados como $x_k = x_0 + p_k$, donde x_0 (no necesariamente cero) es una suposición inicial y p_k es un vector de dirección.

Uno puede mantener el contexto de la discusión precedente incorporando la conjetura inicial en el lado derecho, $r_0 = b - Ax_0$. En lugar de resolver $Ax = b$, se debe resolver $A_p = r_0$ y recuperar la solución de $x = x_0 + p$. Así, r_0 reemplaza a b , p reemplaza a x , y p_k reemplaza a x_k .

En general el sistema lineal $Ax = b$ tiene una solución de Krylov en $A^D b$, si y solo si b pertenece a $R(A^k)$, con $k = \text{ind}(A)$.

Cuando existe una solución de Krylov, es única y es la solución inversa de Drazin $x = A^D b$, siempre que x pertenezca a $R(A^k)$, con $k = \text{ind}(A)$. De esta forma, queda caracterizada la inversa de Drazin de cualquier matriz singular $n \times n$.

Se ha utilizado y ejemplificado la inversa de Drazin para una matriz cuadrada.

A cada matriz $M \in \mathbf{C}^{n \times n}$, se le puede asignar un entero no negativo k , al que llamamos índice Drazin y este es útil para definir la inversa Drazin.

Se presentó un algoritmo con pasos detallados para calcular la inversa de Drazin y se da un ejemplo particular.

Bibliografía

- [1] Solar G.E., Speziale De G.L. Apuntes de álgebra lineal, México, Limusa Noriega Editores, Facultad de Ingeniería, UNAM, pp. 291-866,(2001).
- [2] Nikuie,M & Mirnia,M.K, Mahmoudi. Some results about the index of matrix and Drazin inverse, Mathematical Sciences Quarterly Journal. 4, (2010).
- [3] Owe, A. Iterative Solution Methods, New York, USA: Cambridge University Press, (1994).
- [4] Ipsen, Ilse & Meyer, Carl. The Idea Behind Krylov Methods, The Mathematical Monthly. 105. 10.2307/2589281,(1997).
- [5] Cortes R.J.J & Arganis M.L. Aplicación inversa del método de Krilov para obtener una matriz de orden tres. México: Ingeniería, investigación y tecnología, 13(2), 169-174, (2012).
- [6] Chen C.T. Introduction to Linear Systems Theory, New York, Holt, Rinehart And Wiston Inc., (1970).
- [7] Drazin, M. P. "Pseudo-inverses in associative rings and semigroups". The American Mathematical Monthly. 65 (7): 506–514 (1958).
- [8] Stoer J., Bulirsch R. The Jordan Normal Form of a Matrix. pp.335, Introduction to Numerical Analysis,New York, USA: Springer,(1993).
- [9] ANTON, H., Introducción al Álgebra Lineal. Pp. 191. Limusa, México.
- [10] Chen J., Xu Z. Comment on condition number of Drazin inverse and their condition number of singular systems. Applied Mathematics and computacion, 199,512-526,(2008).
- [11] Zhang L. A characterization of the Drazin inverse. Linear Algebra and its Aplications, 355, 183-188,(2001).
- [12] Wei Y. Index splitting for the Drazin inverse and the singular linear system. Applied Mathematics and Computation. 95,115-124, (1998).
- [13] Wei Y., Zhou J. A two-step algorithm for solving singular linear systems with index one. Applied Mathematics and Computation, 175, 472-485,(2006).
- [14] Y. Saad and M.H. Schultz, "GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems", SIAM J. Sci. Stat. Comput., 7:856–869, (1986).
- [15] S.L. Campbell and C.D. Meyer. Generalized Inverses of Linear Transformations. Dover, New York, (1979).
- [16] S.L. Campbell, C.D. Jr. Meyer, N.J. Rose. Application of the Drazin inverse of linear systems od differential equations with singular constant coefficients, SIAM J. Appl. Math 31, 411-425, (1983).

- [17] Wang G. A Cramer rule for finding the solution of class of singular equations. *Linear Algebra and its Applications*, 116,27-34, (1989).
- [18] V.N. Faddeeva. *Computational Methods of Linear Algebra*. Dover, New York, (1959).
- [19] Ion Zaballa. El álgebra de los valores propios. pp.181, *Análisis matricial aplicado y ampliación de métodos numéricos*, Departamento de Matemática Aplicada y Estadística e Investigación Operativa, Euskal Herriko Unibertsitatea.