

1 2
2ej.



UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS

LA FORMA CANONICA DE JORDAN EN SU
APLICACION A MODELOS ECONOMICOS

T E S I S

QUE PARA OBTENER EL TITULO DE :

M A T E M A T I C O

P R E S E N T A :

JUAN ALFARO YLLESCAS

TESIS CON
FALLA DE ORIGEN



MEXICO, D. F.



FACULTAD DE CIENCIAS
SERVICIO ESCOLAR

1994



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas Tesis Digitales Restricciones de uso

DERECHOS RESERVADOS © PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis está protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

CIUDAD UNIVERSITARIA



UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MÉXICO

FACULTAD DE CIENCIAS
División de Estudios
Profesionales
Exp. Núm. 55

M. EN C. VIRGINIA ABRIN BATULE
Jefe de la División de Estudios Profesionales
Universidad Nacional Autónoma de México.
P r e s e n t e .

Por medio de la presente, nos permitimos informar a Usted, que habiendo
revisado el trabajo de tesis que realizó el pasante _____
JUAN ALFARO YLLESCAS

con número de cuenta 7798921-3 con el título: _____

"LA FORMA CANONICA DE JORDAN EN SU APLICACION A MODELOS ECONOMICOS"

Consideramos que reúne los méritos necesarios para que pueda conti-
nuar el trámite de su Examen Profesional para obtener el título de -
MATEMATICO.

GRADO NOMBRE Y APELLIDOS COMPLETOS

FIRMA

MAT. HUMBERTO SANTILLANA LOYO

Director de Tesis
M. EN C. JOSE LUIS NAVARRO URRUTIA

ACT. FRANCISCO SANCHEZ VILLARREAL

ACT. AURORA VALDES MICHEL

Suplente
MAT. LORETO CRUZ HERNANDEZ
Suplente

AGRADECIMIENTOS

Con mucha satisfacción de tener el apoyo de las apreciables personas que tienen el cargo de ser sinodales de mi trabajo de tesis profesional.

También con aprecio a todas las personas que de alguna forma directa ó indirectamente tuvieron que ver en la realización de este trabajo.

LA FORMA CANONICA DE JORDAN EN SU APLICACION
A MODELOS ECONOMICOS

	Pagina
INTRODUCCION	1
CAPITULO I	
LA FORMA CANONICA DE JORDAN	6
CAPITULO II	
SISTEMAS DE ECUACIONES DIFERENCIALES LINEALES	34
CAPITULO III	
APLICACION A MODELOS ECONOMICOS	79
CONCLUSIONES	127
BIBLIOGRAFIA	

INTRODUCCION

El hombre desde tiempos muy remotos al interrelacionarse con los fenómenos naturales, empieza a tratar de comprenderlos y en este afán va descubriendo las herramientas necesarias para poder sobrevivir y desarrollarse.

Al volverse sedentario y seguir su desarrollo en comunidades cada vez más grandes estas sociedades se vuelven más complejas; se ven en la necesidad de utilizar todo su intelecto para contrarrestar muchos de los problemas y dar una solución adecuada a estos mismos (que muchas veces no es la ideal).

En los siglos XIX y XX hubo grandes avances en la Física y las Matemáticas, se desarrollaron Teorías Económicas y Sociales, etc., pero a pesar de ello hasta hace aproximadamente unos 25 ó 20 años atrás se había dejado a un lado las Ciencias Físicas y Matemáticas, por el otro las Ciencias Sociales y Económico-Administrativas, como si fueran entes completamente ajenos.

Poco después de la Segunda Guerra Mundial, se inició la adaptación de los enfoques matemáticos que se habían aplicado a las operaciones militares, a los problemas empresariales. Este movimiento, conocido generalmente por el nombre de « investigación operativa », en su rápido crecimiento durante la década de los 50 centró principalmente su atención en problemas tales como la programación de la producción, la distribución física, la presupuestación de capital y el control de existencias. Tan sólo en años recientes sea suscitado un interés sustancial entre los especialistas de mercado por la investigación operativa, se ha llevado a cabo un esfuerzo relevante para aplicar los modelos y las técnicas matemáticas al mercado.

La descripción de la naturaleza de algunas de las aplicaciones de los modelos matemáticos a los problemas del mercado, esboza algunas de las limitaciones y posibilidades de esta nueva técnica y analiza algunos de los problemas que los especialistas en el mercado se encontrarán al utilizar estas técnicas. Se presenta en términos no matemáticos los conceptos subyacentes al empleo de los modelos matemáticos, concebido de una manera para ayudar al manejo del mercado a trabajar de un modo eficiente con los especialistas técnicos. Quizá sea todavía de mayor importancia el que se trate el empleo de los modelos matemáticos en el contexto de los procesos del mercado, tal como se presentan en la realidad. El estudio de los casos concretos no resulta una descripción abstracta de los problemas, sino que presenta todo el contexto de las situaciones de mercado en las que se emplearon los modelos.

Como tales, estos estudios tendrán también valor para aquellos especialistas en dichas técnicas, al proporcionarles una mejor apreciación del contexto en que deberán emplear sus conocimientos matemáticos.

Para que en el futuro rinda fruto la aplicación de los modelos, es necesaria la fusión de la comprensión que tiene el manejo del mercado, de la compleja realidad de los problemas del mercado y de la habilidad analítica del Investigador formado matemáticamente. Un objetivo importante de este estudio es poner en claro la necesidad de esta fusión y contribuir a su desarrollo proporcionando material relevante para las actividades de ambos grupos en esta tarea conjunta.

A partir de 1950 ha existido un interés creciente entre los ejecutivos de mercado por las posibilidades de utilizar modelos matemáticos como ayudas en la elaboración de decisiones.

Especialmente a partir de 1960, este interés ha alcanzado las proporciones de una moda, y al parecer existe una confusión considerable con respecto a la construcción de modelos, a la forma como los ejecutivos deberían utilizarlos, y al modo en que la « nueva tecnología de la información » afectará a los métodos tradicionales y a las estructuras organizativas.

Por una parte, los entusiastas prevén que muchas de las decisiones de mercado pasarán a ser « automáticas » -que se utilizarán programas para calculadoras, basados en modelos matemáticos para llegar a decisiones sobre la política a seguir con poca o ninguna necesidad de modificación por parte del juicio ejecutivo.

En el otro extremo encontramos la posición « reaccionaria », que mantiene que los problemas del mercado son demasiado complejos para que puedan representarse mediante modelos abstractos, y que la conducta de mercado resulta inherentemente demasiado caótica para que pueda seguir unos cauces regulares y predecibles.

Aunque en el uso de modelos matemáticos está el punto de buena parte de la reciente polémica, también se ha confundido con toda una serie de cuestiones conexas, algunas de las cuales constituyen de hecho verdaderos encuentros. Así, por ejemplo, se ha mantenido un muy acalorado debate en torno al empleo de computadoras en el manejo del mercado, especialmente en publicidad.

Aunque el empleo de computadoras está relacionada con el empleo de modelos matemáticos -ya que los modelos complejos requieren normalmente unos cálculos numéricos considerables-, la polémica sobre las computadoras refleja también interés por cuestiones laterales tales como el hombre versus la máquina, si las computadoras pueden « pensar », etc.

Los problemas en torno a la construcción y aplicación de los modelos también se han considerado como formando parte de la cuestión mucho más amplia sobre si el mercado es, o puede ser, una « ciencia ».

El concepto de modelo resulta básico en las ciencias físicas y biológicas, y es cierto que si el mercado debe convertirse en una ciencia, una condición necesaria es la construcción, verificación y generalización de los modelos. Sin embargo, lo opuesto no es necesariamente cierto. Los modelos pueden resultar de utilidad para los manejos aunque el mercado nunca se convierta en una ciencia.

Además, el manejo de las actividades de mercado probablemente nunca será una ciencia, al igual que la ingeniería no es una ciencia. Sin embargo, los Ingenieros se basan mucho en modelos, y es muy posible que lo mismo suceda con los expertos en ventas, aunque nunca lleguen a ser científicos.

En vista a la controversia entre los supuestos « expertos », no tiene nada de sorprendente que muchos ejecutivos de mercado estuvieran y estén inseguros por lo que concierne a la utilidad de los modelos.

Dejando de lado desacuerdos auténticos, parte de la controversia refleja envidias profesionales (pocos ejemplo, psicólogos no cuantitativos versus especialistas en investigación operativa) y/o rivalidad comercial entre empresas consultoras e investigadoras. Unos factores parecidos y una confusión de cuestiones semejante se presentó en el « gran debate » acerca de la investigación motivacional a comienzos de la década de los 50.

Con el fin de colocar las aplicaciones al mercado de los modelos matemáticos en la perspectiva adecuada, resulta necesario en primer lugar saber lo que se entiende por « modelo ». El concepto de modelo es fundamental para la investigación científica en cualquier campo. El reciente interés por la construcción de modelos en las operaciones de las empresas refleja una creencia -o al menos una esperanza- de que las ideas y los métodos de las ciencias físicas pueden aplicarse, al menos parcialmente, a los problemas con los que se enfrentan. Estos métodos y estas ideas no pueden tomarse prestadas y aplicarse de una manera fragmentaria.

ANTECEDENTES

Al estudiar y trabajar los problemas económicos me surge la idea de poder organizar, clasificar y sistematizar usando las matemáticas en una serie de problemas que se presentan con frecuencia en la economía.

La enseñanza de la Economía preocupada por revizar y actualizar sus programas de estudio ha estado haciendo una revisión que pretende actualizar y mejorar los programas de matemáticas dandoles un enfoque más práctico, para su utilización en los problemas Económicos.

La rama de la matemática que se estudia porque tiene gran aplicación es el Algebra Lineal, en una amplia gama de disciplinas, como la Estadística, Matemáticas, en aquellas en las cuales se necesita del estudio de funciones de varias Variables donde se requiere también conocimientos de Teoría de Matrices debido a la presencia de computadoras de alta velocidad y al desarrollo en la aplicación de las matemáticas en áreas tradicionalmente no técnicas.

Por lo tanto para el estudio de las Ciencias Económicas es necesario el conocimiento de las Matemáticas Aplicadas y Computación (Modelos Matemáticos aplicados a la Economía y a la Computación).

El objetivo de este trabajo es mencionar algunos de los conceptos fundamentales del Algebra Lineal, que se utilizarán para la aplicación de los Sistemas de Ecuaciones Diferenciales de Primer Orden a algunos problemas Económicos.

En el Primer Capítulo se mencionara el tema sobre Matrices llamado: La Forma Canónica de Jordan y sus diferentes facetas que se dan de acuerdo al sistema que se esta tratando cuyas características son intrínsecas al problema por resolver.

El Segundo Capítulo se mencionara parte de la Teoría de los Sistemas de Ecuaciones Diferenciales Lineales, los Teoremas de Existencia y Unicidad.

En el Tercer Capítulo se desarrollaran algunos problemas económicos resueltos mediante el Sistema de Ecuaciones Diferenciales; mediante la aplicación de la Forma Canónica de Jordan.

CAPITULO 1

FORMA CANONICA DE JORDAN

Nuestro objetivo en este capítulo es describir la forma conocida de Jordan para esto es necesario algunos conceptos que nos permitan abordarlo.

Para comenzar definiremos lo que es un vector: es un conjunto de N números (complejos) o reales que se escriben en la forma:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_N \end{bmatrix}$$

Un vector de este tipo se llama vector en columna.

Si los N números se disponen en línea horizontal como:

$$x = (x_1, x_2, \dots, x_N)$$

Se llama x un vector fila (ó renglón).

Emplearemos letras minúsculas como x, v, w , etc., para designar vectores. Cuando aparezca un conjunto particular de vectores usaremos subíndices tales como x_1, x_2 , etc.

Los elementos x_i se llaman componentes de x , mientras que N es la dimensión de ese vector. Los vectores unidimensionales se llaman escalares (ó cantidades usuales).

Si las componentes de un vector x son reales diremos que x es real.

Si las componentes de un vector x son números complejos diremos que x es un vector en los complejos.

Se puede decir que dos vectores son iguales si y sólo si sus componentes son iguales:

$$x = y \quad \text{si y sólo si} \quad x_i = y_i \quad \text{para } i = 1, 2, 3, \dots, N$$

Las operaciones que actúan sobre dos o más vectores son:

La adición: La suma de dos vectores x e y se denota por $x + y$ y se define diciendo que es el vector:

$$x + y = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_N + y_N \end{bmatrix}$$

La multiplicación por un escalar:

Es el producto de un vector x por una escalar c_1 esta definido

por la relación:

$$c_1 x = \begin{bmatrix} c_1 x_1 \\ c_1 x_2 \\ \vdots \\ c_1 x_N \end{bmatrix}$$

Producto Interno de dos vectores (ó producto escalar).

Este producto esta definido por:

$$(x_1, x_2, \dots, x_N) \cdot (y_1, y_2, \dots, y_N) = (x, y) = \sum_{i=1}^N x_i y_i$$

está es una manera de multiplicar vectores.

La importancia del producto interno estriba en el hecho de que (x, x) puede considerarse que representa el cuadrado de la longitud del vector real x . Poseemos así un método para valorar estas cantidades no numéricas que son los vectores.

Ortogonalidad. Si dos vectores reales estan ligados por la relación: $(x, y) = 0$ y también

$$(x_1, x_2, \dots, x_N) \cdot (y_1, y_2, \dots, y_N) = 0$$

Se dice que su producto interno es cero por lo tanto los vectores x e y son ortogonales. Sea (x_i) un conjunto de vectores

reales mutuamente ortogonales, esto es que, $(x_i, x_j) = 0$ para

$i \neq j$ y normalizado mediante la condición $(x_i, x_i) = 1$ se dice

que son vectores ortonormales.

Matrices. Introduzcamos ahora el concepto de matriz. Es una disposición de números reales (o complejos) escritos en la forma:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1N} \\ a_{21} & a_{22} & \dots & a_{2N} \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ a_{N1} & a_{N2} & \dots & a_{NN} \end{bmatrix}$$

Se llamará una matriz cuadrada puesto que estos son los únicos tipos de matrices que consideraremos. Las cantidades o elementos a_{ij}

se llaman los elementos de A, y el entero N es la dimensión. Los números $a_{11}, a_{12}, \dots, a_{1N}$ se dice que constituyen la fila

i-ésima de A, y los $a_{1j}, a_{2j}, \dots, a_{Nj}$, la columna j-ésima.

Multiplicación de Matrices.- Matriz por Vector.

$$y = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1N} \\ a_{21} & a_{22} & \dots & a_{2N} \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ a_{N1} & a_{N2} & \dots & a_{NN} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_N \end{bmatrix}$$

Para ser estos conceptos razonables volvemos a las transformaciones lineales de la forma

$$y_i = \sum_{j=1}^N a_{ij} x_j, \quad i = 1, 2, \dots, N$$

Matriz por Matriz.

Veamos ahora como definir el producto por una matriz por otra. Consideremos una segunda transformación lineal

$$z = By$$

que convierte las componentes de y en las componentes de z.

Para poder expresar las componentes de z en función de las componentes de x, siendo como antes $y = Ax$.

$$z_i = \sum_{k=1}^N b_{ik} y_k = \sum_{k=1}^N b_{ik} \left(\sum_{j=1}^N a_{kj} x_j \right)$$

$$z_i = \sum_{j=1}^N \left(\sum_{k=1}^N b_{ik} a_{kj} \right) x_j$$

Si introducimos ahora una nueva matriz $C = (c_{ij})$ definida por las relaciones

$$c_{ij} = \sum_{k=1}^N b_{ik} a_{kj} \quad i, j = 1, 2, \dots, N$$

Que podemos escribir como $z = Cx$.

Por lo tanto formalmente $z = By = B(Ax) = (BA)x$, con lo que llegamos a definir el producto A por B, $C = BA$.

Matriz Transpuesta definiremos ahora una importante función matricial de A la matriz transpuesta por medio de la relación

$$A^t = (a_{ji})$$

Las filas de A^t son las columnas de A y las filas de A son las columnas de A^t . Ejemplo:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad A^t = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}$$

Matriz simétrica: Las matrices que satisfacen la condición

$A = A^t$ se llaman simétricas y están caracterizadas también por la condición: $a_{ij} = a_{ji}$.

Matriz hermítica: Todas las propiedades vitales de las matrices simétricas tienen una análoga inmediata para las matrices hermíticas, por lo tanto hay ventajas en ambos procedimientos.

La matriz hermitiana A^H de A es la matriz cuadrada $q \times p$ que se obtiene tomando el complejo conjugado de sus elementos en A^t la primera columna de A^H consiste en los complejos conjugados del primer renglón de A.

Y así sucesivamente esto es:

$$\langle A^H \rangle_{ij} = \langle \bar{A} \rangle_{ji} \text{ para } 1 \leq i \leq q \text{ y } 1 \leq j \leq p$$

Ejemplo. Sean A y B como se muestra, entonces A^t , B^t , A^H , B^H como se indica:

$$A = \begin{bmatrix} -1 & 2 & 3 \\ 2 & 6 & 3 \end{bmatrix}; \quad B = \begin{bmatrix} 2 - 3i \\ 6 \end{bmatrix}$$

$$A^t = \begin{bmatrix} -1 & 2 \\ 2 & 6 \\ 4 & 3 \end{bmatrix}; \quad B^t = \begin{bmatrix} 2 - 3i & 6 \end{bmatrix}$$

$$A^H = \begin{bmatrix} -1 & 2 \\ 2 & 6 \\ 4 & 3 \end{bmatrix} \quad (= A^t); \quad B^H = \begin{bmatrix} 2 + 3i & 6 \end{bmatrix} \quad (= B^t)$$

Una matriz cuadrada A es $p \times q$ con $p = q$, se dice que es triangular superior cuando por abajo de la diagonal principal solo existen ceros por lo tanto: A de $p \times p$ satisface $\langle a \rangle_{ij} = 0$ si

$i > j$ para $1 \leq i \leq p$ y $1 \leq j \leq q$.

Matriz Inversa: Dada una matriz A existe otra matriz B conformable con A para la operación de producto de tal manera que

satisfaga la operación de producto $A \cdot B = I$. Para lo cual $B = A^{-1}$

por lo tanto $A \cdot A^{-1} = I$, a B se le llama la matriz inversa de A.

Matrices unitarias: Se plantea la posibilidad de que dada una matriz A cualquiera exista otra matriz B, conformable con A para la operación de producto que satisfaga la relación $A \cdot B = I$, donde I es una matriz identidad de orden apropiado.

Definición: Una matriz P, cuadrada $p \times p$ para la cual $P^{-1} = P^H$,

de tal modo que se dice $PP^H = P^HP = I$ es unitaria.

Una matriz ortogonal es una matriz unitaria real P, tal que

$$P^{-1} = P^T \text{ y } PP^T = P^TP = I$$

Ejemplos:

$$P_1 = 1/2 \begin{bmatrix} 1+i & -1+i \\ 1+i & 1-i \end{bmatrix} \quad \text{y} \quad P_2 = \sqrt{2}/2 \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

entonces P_1 y P_2 son unitarias, y P_2 también es ortogonal como lo

muestra el cálculo directo: $P_1^H P_1$, $P_2^H P_2$, y de $P_2^T P_2$ observamos que P_1 , no es ortogonal porque no es real.

Raíces y vectores característicos: Llamando a x al vector cuyas componentes son x_i , y $A = (a_{ij})$ podemos escribir:

$$A \cdot x = \lambda \cdot x$$

Sabemos que una condición necesaria y suficiente para que exista un vector no trivial x que satisfaga a esta relación es que λ sea una raíz de la ecuación característica de la matriz A .

La cual se representa como una $| a_{ij} - \lambda \delta_{ij} | = 0$, ó como se acostumbra a escribir $| A - \lambda I | = 0$, a esta ecuación se le llama ecuación característica de A . Como es un polinomio en λ , la ecuación posee N raíces, distintas o no, que se llaman raíces característicos o valores característicos. Si las raíces son distintas se les denomina simples, como término opuesto a múltiple.

Con mucha frecuencia también se designan con el nombre de autovalor o de valor propio. Asociado con cada uno de los distintos valores característicos λ hay un vector característico (autovector, o vector propio) determinado salvo un coeficiente de proporcionalidad.

Vectores linealmente dependientes y linealmente independientes.

Dependencia lineal: Sea x_1, x_2, \dots, x_k un conjunto de k

vectores N - dimensionales. Si existe un conjunto de escalares, c_1, c_2, \dots, c_k uno al menos los cuales es distinto de cero, con

la propiedad

$$c_1 x_1 + c_2 x_2 + \dots + c_k x_k = 0$$

donde cero representa el vector nulo, decimos que los vectores son linealmente dependientes.

Si no existe tal conjunto de escalares, decimos que los vectores son linealmente independientes.

La teoría de las matrices estará presente cada vez que se trate de resolver sistemas de ecuaciones lineales mediante sus matrices asociadas o matrices aumentadas se han desarrollado métodos como el de Cramer, Gauss, Gauss-Jordan etc., para la solución a estos.

Dentro de la teoría de matrices se tienen metodologías diferentes para trabajar con los tipos de matrices que nos son de utilidad y obtener resultados aceptables, para seguir desarrollando esta teoría o para aplicarlos a los problemas que se nos presentan.

Las matrices que son de gran ayuda son las matrices de $n \times n$.

Estas surgen en un gran número de aplicaciones que van desde una descripción de costos en Economía hasta el análisis del control de un cohete que viaja en el espacio.

Las matrices de $n \times n$ con vectores característicos linealmente independientes pueden ser llevados a una forma simple mediante una transformación de equivalencia.

Afortunadamente, las matrices dan origen a polinomios de diferentes valores característicos. Dado que en las aplicaciones que se mencionan más adelante se obtienen los valores característicos de estas matrices con los cuales se obtendrán los vectores característicos apropiados para poder establecer la matriz en su Forma Canónica de Jordan.

Se describe una relación útil e interesante que es válida entre dos matrices.

Matriz equivalente: Dos matrices A y B de $n \times n$ se llaman equivalentes (o semejantes) si existe una matriz invertible C de $n \times n$ tal que

$$B = C^{-1}AC \quad (1)$$

La función definida por (1) que asocia a la matriz A , la matriz B se llama transformación de equivalencia.

Las matrices equivalentes tienen varias propiedades importantes en común; la mayoría de las matrices son equivalentes a matrices diagonales.

Teorema 1: A y B son matrices equivalentes de $n \times n$ entonces A y B tienen la misma ecuación característica y por consiguiente tienen los mismos valores característicos.

Demostración: Puesto que A y B son equivalentes.

$$\begin{aligned} B &= C^{-1}AC \text{ y } \det. (B - \lambda I) \\ &= \det (C^{-1}AC - \lambda I) = \det \left[C^{-1}AC - C^{-1}(\lambda I)C \right] \\ &= \det \left[C^{-1}(A - \lambda I)C \right] = \det (C^{-1}) \det (A - \lambda I) \det (C) \\ &= \det (C^{-1}) \det (C) \det (A - \lambda I) = \det (C^{-1}C) \det (A - \lambda I) \\ &= \det I \det (A - \lambda I) = \det (A - \lambda I) \end{aligned}$$

Por lo tanto A y B tienen la misma ecuación característica y los mismos valores característicos.

En muchas aplicaciones es de considerable utilidad. "Diagonalizar" una matriz A, esto es, encontrar una matriz diagonal equivalente A.

Para diagonalizar una matriz nos apoyamos en los siguientes teoremas:

Teorema 2 Una matriz A cuadrada de $n \times n$ (de orden n) es diagonalizable si y sólo si tiene n vectores característicos linealmente independientes.

Demostración: Si $A = SDS^{-1}$ donde

$$D = \text{diagonal } (\lambda_1, \dots, \lambda_n) \text{ entonces}$$

$$AS^{(j)} = \lambda_j S^{(j)} \quad j = 1, \dots, n$$

Por lo tanto, las columnas de S son vectores característicos de A. Pero S es no singular, y, por lo tanto, sus columnas son linealmente independientes.

En consecuencia, A tiene n vectores característicos linealmente independientes; que son v_1, \dots, v_n ; es decir, se

verifica que $Av_j = \lambda_j v_j$, $j = 1, \dots, n$ donde v_j es diferente de 0.

Sea S la matriz $n \times n$ en la que la columna j -ésima sea igual a v_j ,
 $j = 1, \dots, n$

$$\text{Por lo tanto, } AS^{(j)} = \lambda_j S^{(j)} = \lambda_j S_j = SD^{(j)}, j = 1, \dots, n,$$

donde $D = \text{diagonal}(\lambda_1, \dots, \lambda_n)$.

Pero $AS^{(j)} = (AS)^{(j)}$ y $SD^{(j)} = (SD)^{(j)}$ por consiguiente
 $(AS)^{(j)}$, con $j = 1, \dots, n$, es decir, $AS = SD$

Ahora bien, las columnas de S son linealmente independientes por lo que el rango de S es n. Por lo tanto, S es no singular. Por consiguiente, podremos multiplicar ambos miembros de $AS = SD$ por S^{-1} para obtener $S^{-1}AS = S^{-1}SD = D$ con esto hemos demostrado que si A tiene n vectores característicos linealmente independientes (v_1, \dots, v_n) entonces A es semejante a la diagonal

$(\lambda_1, \dots, \lambda_n)$, donde λ_j es la raíz característica que corresponda a v_j con $j = 1, \dots, n$.

Para hallar las raíces del polinomio característico recordamos el conocido teorema del Algebra que afirma que las raíces racionales de un polinomio como:

$\lambda^n + (A_n - 1)\lambda^{n-1} + \dots + A_1\lambda + A_0$ con coeficientes A_0, \dots, A_{n-1} son enteros que dividen a A_0 :

$$\text{Ejemplo } A = \begin{bmatrix} 2 & -3 & 3 \\ 4 & -5 & 3 \\ 4 & -4 & 2 \end{bmatrix} \text{ la matriz característica de } A$$

$$(\lambda I_3 - A) = \begin{bmatrix} \lambda - 2 & 3 & -3 \\ -4 & \lambda + 5 & -3 \\ -4 & 4 & \lambda - 2 \end{bmatrix}$$

$$\text{Det}(\lambda I_3 - A) = \lambda^3 + \lambda^2 - 4\lambda - 4$$

Por lo tanto, las únicas raíces racionales de $\lambda^3 + \lambda^2 - 4\lambda - 4$ son divisores de -4 . Al hacer la prueba.

Con 1, -1, 2, -2, 4 y -4 encontramos que -1, 2 y -2 son las raíces si sustituimos $\lambda = -1$ en la matriz característica

$$\begin{bmatrix} -3 & 3 & -3 \\ -4 & 4 & -3 \\ -4 & 4 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

$$\begin{aligned} \text{Esto equivale a, } & -3x_1 + 3x_2 - 3x_3 = 0 \\ & -4x_1 + 4x_2 - 3x_3 = 0 \end{aligned}$$

Resolviendo este sistema de ecuaciones tenemos:

$$\begin{aligned} x_1 &= x_2 \\ x_3 &= 0 \end{aligned}$$

Por lo tanto (1, 1, 0) es un vector característico que corresponde a $\lambda = -1$.

Definición de matriz diagonalizable: una matriz A de n x n es diagonalizable si existe una matriz diagonal D tal que A es equivalente a D.

Teorema 3: una matriz A de n x n es diagonalizable si y solo si tiene n vectores característicos linealmente independientes. En este caso, la matriz diagonal D equivalente a A esta dada por:

$$\begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \cdot & \cdot & \lambda_3 & & \\ \cdot & \cdot & & & \\ \cdot & \cdot & & & \\ 0 & 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

donde $\lambda_1, \lambda_2, \dots, \lambda_n$ son los valores característicos de A si C es

una matriz cuyas columnas son vectores característicos linealmente independientes de A, entonces:

$$D = C^{-1}AC$$

Demostración: supondremos que A tiene n vectores característicos linealmente independientes, v_1, v_2, \dots, v_n (no necesariamente diferentes).

$$\text{Sea } v_1 = \begin{pmatrix} c_{11} \\ c_{21} \\ \vdots \\ c_{n1} \end{pmatrix}, v_2 = \begin{pmatrix} c_{12} \\ c_{22} \\ \vdots \\ c_{n2} \end{pmatrix}, \dots, v_n = \begin{pmatrix} c_{1n} \\ c_{2n} \\ \vdots \\ c_{nn} \end{pmatrix}$$

$$\text{y sea } C = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{pmatrix}$$

Entonces C es invertible puesto que sus columnas son linealmente independientes.

Ahora.

$$AC = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{pmatrix}$$

$$\text{y vemos que la } i\text{-ésima columna } AC \text{ es } A \begin{pmatrix} c_{1i} \\ c_{2i} \\ \vdots \\ c_{ni} \end{pmatrix} = Av_i = \lambda_i v_i.$$

De esta manera AC es la matriz cuya columna i -ésima es $\lambda_i v_i$ y

$$AC = \begin{pmatrix} \lambda_1 c_{11} & \lambda_1 c_{12} & \dots & \lambda_1 c_{1n} \\ \lambda_2 c_{21} & \lambda_2 c_{22} & \dots & \lambda_2 c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_n c_{n1} & \lambda_n c_{n2} & \dots & \lambda_n c_{nn} \end{pmatrix}$$

$$\text{pero } CD = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_1 c_{11} & \lambda_1 c_{12} & \dots & \lambda_1 c_{1n} \\ \lambda_2 c_{21} & \lambda_2 c_{22} & \dots & \lambda_2 c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_n c_{n1} & \lambda_n c_{n2} & \dots & \lambda_n c_{nn} \end{pmatrix}$$

De manera que $AC = CD$ y puesto que C es invertible podemos multiplicar por ambos lados de ($AC = CD$) por la izquierda por C^{-1} para obtener.

$$D = C^{-1}AC$$

Esto demuestra que si A tiene n vectores característicos linealmente independientes, entonces A es diagonalizable. Inversamente, suponga que A es diagonalizable. Esto es, suponga que ($D = C^{-1}AC$) es válida para alguna matriz invertible C .

Sean v_1, v_2, \dots, v_n las columnas de C .

Entonces $AC = CD$ e invirtiendo los argumentos anteriores inmediatamente vemos que $Av_i = \lambda_i v_i$ para $i = 1, 2, \dots, n$.

De esta manera v_1, v_2, \dots, v_n son vectores característicos de A y son linealmente independientes ya que C es invertible.

Corolario. Si la matriz A de $n \times n$ tiene n valores característicos distintos, entonces A es diagonalizable.

Observación. Si los coeficientes reales de un polinomio de grado n son tomados al azar, entonces, con probabilidad 1, el polinomio tendrá n raíces diferentes.

Intuitivamente, no es difícil ver que sucede esto si $n = 2$, por ejemplo, entonces la ecuación:

$$\lambda^2 + a\lambda + b = 0 \text{ tiene raíces iguales si,}$$

$a^2 = 4b$. Un caso muy improbable si a y b se escogen al azar.

Podemos escribir, por supuesto, polinomios que tengan raíces de multiplicidad algebraica mayor que 1, pero estos polinomios son excepcionales.

De modo que, sin tratar de ser matemáticamente preciso, es válido decir que la mayor parte de los polinomios tienen raíces diferentes y, por lo visto a la mayor parte de las matrices son diagonalizables.

Observación. Puesto que existe un número infinito de formas de escoger un vector característico, existe un número infinito de formas de escoger la matriz diagonalizable C . La única recomendación es escoger los vectores característicos y la matriz C tal que se simplifiquen los cálculos aritméticos. Esto significa que se deberían incluir tantos ceros y unos como sean posibles.

MATRICES SIMÉTRICAS Y DIAGONALIZABLES.

Dentro de la amplia gama de matrices diagonalizables, existe una cantidad de matrices simétricas que tienen muchas propiedades importantes; se mostrará que cualquier matriz simétrica tiene n vectores característicos reales linealmente independientes y por consiguiente por el teorema 3, es diagonalizable.

Matrices Simétricas se dice que la matriz A de $n \times n$ (cuadrada) es simétrica si $A^t = A$.

Enunciaremos algunos Teoremas para enfatizar el uso de matrices simétricas.

Teorema 4. Sea A una matriz simétrica real de $n \times n$. Entonces los valores característicos de A son reales.

Demostración: Sea α un valor característico de A con vector característico w ; esto es $Aw = \alpha w$, ahora w es un vector en C^n , y un producto interno en C^n que satisface.

$$(\beta x, y) = \beta(x, y) \text{ y } (x, \bar{\beta}y) = \bar{\beta}(x, y) \quad (1)$$

$$\text{entonces } (Aw, w) = (\beta w, w) = \beta(w, w) \quad (2)$$

Más aún, con el hecho de que $A^t = A$

$$(Aw, w) = (w, A^t w) = (w, Aw) = (w, \beta w) = \bar{\beta}(w, w) \quad (3)$$

$$\text{Igualando (2) y (3) tenemos } \beta(w, w) = \bar{\beta}(w, w) \quad (4)$$

Pero $(w, w) = |w|^2$ diferente de 0 puesto que w es un vector característico. De esta manera podemos dividir ambos lados de (4) entre (w, w) para obtener $\beta = \bar{\beta}$.

Si $\beta = a + ib$, entonces $\bar{\beta} = a - ib$ de los cual tenemos
 $a + ib = a - ib$.

Lo cual es válido únicamente si $b = 0$, esto es $\beta = a$ por lo tanto β es real.

Para matrices simétricas el resultado es mas "fuerte". Los valores característicos de una matriz simétrica correspondientes a valores característicos son ortogonales.

Teorema 5. Sea una matriz simétrica real de $n \times n$, si β_1 y β_2 son valores característicos reales de w_1 y w_2 entonces w_1 y w_2 son ortogonales.

$$\text{Demostrar } Aw_1 - w_2 = \beta_1 w_1 - w_2 = \beta_1(w_1 - w_2) \quad (1)$$

$$Aw_1 - w_2 = w_1 - Aw_2 = w_1 - (\beta_2 w_2) = \beta_2(w_1 - w_2) \quad (2)$$

Igualando 1 y 2, $\beta_1(w_1 - w_2) = \beta_2(w_1 - w_2)$ y puesto que β diferente β_2 concluimos que $w_1 - w_2 = 0$.

Teorema 6. Sea A una matriz simétrica real de $n \times n$.

Resulta entonces que A tiene n vectores característicos ortonormales.

Este teorema nos dice que si A es simétrica entonces R^n tiene una base $b = \{v_1, v_2, \dots, v_n\}$ que consiste de vectores característicos, ortonormales de A .

Sea Q la matriz de cuyas columnas son v_1, v_2, \dots, v_n .

Entonces Q es una matriz ortogonal.

Definición : Matriz ortogonalmente diagonalizable.

Una matriz A de $n \times n$ se dice que es diagonalizable ortogonalmente, si existe una matriz ortogonal Q tal que $Q^t A Q = D$.

Donde $D = \text{diag.}(\beta_1, \beta_2, \dots, \beta_n)$ y $(\beta_1, \beta_2, \dots, \beta_n)$ son los

valores característicos de A ; Q es ortogonal si $Q^t = Q^{-1}$.

Por lo tanto $Q^t A Q = D$ puede ser escrita como $Q^{-1} A Q = D$

Teorema 7. Sea A una matriz real de $n \times n$. Entonces A diagonalizable ortogonalmente si y solamente si A es simétrica.

Demostración ; sea A simétrica, entonces A es diagonalizable ortogonalmente con Q (la matriz cuya columna son los vectores característicos ortonormales). Inversamente suponga que A es diagonalizable ortogonalmente.

Entonces existe una matriz ortogonal Q tal que

$Q^t A Q = D$. Multiplicando esta ecuación por la izquierda por Q , por la derecha por Q^t y utilizando el hecho de que $Q^t Q = Q Q^t = I$, obtenemos $A = Q D Q^t$ entonces $A^t = (Q D Q^t)^t = (Q^t)^t D^t Q^t = Q D Q^t = A$.

De esta manera A es simétrica y el teorema está demostrado. Se utilizaran los hechos $(AB)^t = A^t B^t$, $(A^t)^t = A$ y $D^t = D$ para cualquier diagonal D .

Sin embargo, las matrices que no son diagonalizables (esto es que no tienen n vectores característicos linealmente independientes) aparecen en ciertas aplicaciones.

En este caso aún es posible mostrar que la matriz equivalente a otra matriz más simple, pero la nueva matriz ya no es diagonal y la matriz de transformación C es más difícil de obtener.

Dada una matriz cuadrada A , queremos escoger M tal que $M^{-1} A M$ sea lo más diagonal posible.

En el caso más sencillo, A tiene un conjunto completo de vectores propios que serán las columnas de M .

La forma de Jordan es $J = M^{-1}AM = \Lambda$, se construye completamente a partir de los bloques de 1 por 1 $J_i = \lambda_i$ y se alcanza el objetivo

al obtener una matriz totalmente diagonal. En el caso más general y más difícil, faltan algunos vectores propios y es imposible la forma diagonal.

La cuestión clave es: si A es alguna matriz de $n \times n$ (matriz cuadrada). ¿bajo qué condiciones toma la forma de Jordan (J)?

¿ Cuando existiera alguna M tal que $M^{-1}AM = J$?.

Como primera condición, cualquier matriz similar a A debe compartir los mismos valores propios. Pero esto está lejos de ser suficiente (la matriz diagonal con los mismos valores propios no sea similar a J) y nuestro objetivo está realmente relacionado con los vectores propios.

Para empezar reescribimos la relación $M^{-1}AM = J$, en la forma simple $AM = MJ$:

Admitimos cualquier matriz y el objetivo es hacer $M^{-1}AM$ tan diagonal como sea posible.

El resultado de este esfuerzo por diagonalizar es la forma de Jordan J . En el caso de que la matriz A tenga un conjunto completo de vectores propios linealmente independientes, tomamos $M = S$ y llegamos a $J = S^{-1}AS = \Delta$; La forma de Jordan coincide con la matriz diagonal Δ . Esto es imposible para una matriz defectuosa y por cada vector propio que se pierda, la forma de Jordan tendrá un 1 precisamente encima de su diagonal principal.

Los valores propios deben aparecer en la diagonal misma, ya que J es triangular.

Los valores propios distintos pueden separarse como en el caso diagonal; sólo λ repetida puede requerir o no una entrada fuera de la diagonal J .

Teorema 7. Si A una matriz tiene S vectores propios linealmente independientes entonces es similar a una matriz que está en la forma de Jordan especial.

$$J = M^{-1}AM = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_n \end{bmatrix}$$

Cada bloque de Jordan J_i es una matriz triangular con un sólo valor propio y un solo vector propio.

$$J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}$$

El mismo valor propio λ_i puede aparecer en varios bloques si corresponde a varios vectores propios independientes. Dos matrices son similares si y solo si tienen la misma forma de Jordan (J.).

FORMA DE SCHUR

Teorema 8. (forma y descomposición de Schur).

Sea A, una matriz cuadrada $p \times p$.

a) A es unitariamente semejante a una matriz triangular superior $T = P^H A P$ con P unitaria y con los eigenvalores de A (repetidos de acuerdo a sus multiplicidades algebraicas) en la diagonal principal de T.

Se llama a T una forma de Schur de A y a la descomposición de $A = P T P^H$ una descomposición de Schur de A.

b) Si A y sus eigenvalores son reales, entonces se puede considerar a P como real y por lo tanto ortogonal.

Demostración: las dos partes del Teorema son claramente ciertas para $p = 1$. Se proceda por inducción, suponiendo que el teorema es válido para $p = k$ y se buscará comprobarlo para $p = k+1$

Supongamos entonces que A es $(k+1) \times (k+1)$.

Sea λ_1 un eigenvalor de A asociado al eigenvector X_1 normalizado, de modo que $\|X_1\|_2 = 1$; Si A y λ_1 son reales;

entonces sera posible tomar a x_1 como real. Ya que $\{x_1\}$ se puede extender para formar una base para \mathbb{C}^{k+1} (ó \mathbb{R}^{k+1} si x_1 es real) y entonces usar el proceso de Gram-Schmidt para producir de $\{x_1\}$ una base ortonormal;

hay un conjunto de vectores (reales si x_1 es real) w_1, \dots, w_k

tal que $\left\{ x_1, w_1, \dots, w_k \right\}$ es ortogonal. Entonces la matriz

$$U = \begin{bmatrix} x_1 & w_1 & w_2 & \dots & w_k \end{bmatrix} = \begin{bmatrix} x_1 & w \end{bmatrix} \text{ es unitaria}$$

(ortogonal si x_1 es real)

Se calcula ahora $A^t = U^H A U$:

$$\begin{aligned} A^t &= U^H A U = \begin{bmatrix} x_1 & w \end{bmatrix}^H A \begin{bmatrix} x_1 & w \end{bmatrix} = \begin{bmatrix} x_1 & w \end{bmatrix}^H \begin{bmatrix} A x_1 & A w \end{bmatrix} \\ &= \begin{bmatrix} x_1 & w \end{bmatrix}^H \begin{bmatrix} \lambda_1 x_1 & A w \end{bmatrix} = \begin{bmatrix} \lambda_1 & x_1^H A w \\ 0 & w^H A w \end{bmatrix} = \begin{bmatrix} \lambda_1 & b^H \\ 0 & C \end{bmatrix} \end{aligned}$$

Porque $\|x_1\|^2 = 1$ y porque $w^H x_1 = 0$ debido que U es unitaria.

Como A^t es semejante a A , los eigenvalores de A y A^t son idénticos, incluyendo sus multiplicidades.

Desarrollando $\det (A^t - \lambda I)$ respecto a sus primera columna, se ve que el polinomio característico de A^t es igual a $\lambda_1 - \lambda$ veces por el de C ; de este modo los eigenvalores de A , además de λ_1 ,

son justamente los de C. Pero C de $k \times k$ (y real, si A y λ_1 lo son), por lo tanto es válida la hipótesis inductiva y será posible encontrar una U unitaria (ortogonal si A y sus eigenvalores son reales), de modo que $V^H CV = T^t$ siendo T^t triangular superior con los eigenvalores de C y por lo tanto de A es su diagonal principal.

Si ahora se define P como.

$$P = U \begin{bmatrix} I & O \\ O & V \end{bmatrix}, \text{ entonces } P^H A P = \begin{bmatrix} \lambda_1 & b^{HV} \\ O & T^t \end{bmatrix}$$

La cual tiene la forma apropiada y de este modo demuestra que la hipótesis inductiva es válida para $p = k + 1$. Por lo tanto es verdadera para toda p.

Este es un teorema de existencia, el cual dice, sin embargo, que T se calcule fácilmente; en la práctica, es necesario conocer los eigenvalor para obtener T.

Teorema 9. (FORMA DIAGONAL)

Una matriz A, $p \times p$, es normal (por ejemplo simétrica real, hermitiana o unitaria) si y solo si A es unitariamente semejante a una matriz (forma de Schur) diagonal $D = P^H A P$, en donde P es unitaria y D es diagonal con los eigenvalores de A (repetidos de acuerdo a sus multiplicidades algebraicas) en la diagonal principal de D; Si A y sus eigenvalores son reales, entonces P se puede considerar como real y por lo tanto ortogonal.

Demostración (forma diagonal \Rightarrow normal). Si existe P unitaria tal que $P^H A P = D$ diagonal, entonces ciertamente A es normal porque las matrices diagonales conmutan:

$$\begin{aligned} A^H A &= (PDP^H)^H (PDP^H) = PD^H P^H PDP^H = PD^H DP^H = PDD^H P^H \\ &= PDP^H PD^H P^H = (PDP^H)(PDP^H)^H = AA^H \end{aligned}$$

Demostración (normal \Rightarrow forma diagonal). Suponga que es normal, y sea T una forma de Schur, donde $P^H A P = T$, T también es normal:

$$T^H T = (P^H A P)^H (P^H A P) = P^H A^H A P = P^H A A^H P = (P^H A P)(P^H A P)^H = T T^H$$

Como se necesitaba. Si t_{ij} es $\langle T \rangle_{ij}$, se tiene que $t_{ij} = 0$ para $i > j$. Como $TT^H = T^HT$, sus elementos (i,i) son iguales se tiene entonces $\langle TT^H \rangle_{ij} = |t_{ij}|^2 + |t_{i,i+1}|^2 + \dots + |t_{ip}|^2$

Igualando las dos expresiones y restando el término común $|t_{ii}|^2$ de cada lado, nos da $|t_{i,i+1}| + \dots + |t_{ip}|^2 = |t_{ii}|^2 + \dots + |t_{i-1,i}|^2$ para toda i . Si se hace $i = 1$ en esta

igualdad, se obtendrá 0 en el lado derecho ya que no tiene término; entonces en el lado izquierdo $t_{12} = t_{13} = \dots = t_{1p} = 0$. Si

entonces se hace $i = 2$, se obtendrá $|t_{21}|^2$ del lado derecho, que es igual a cero; entonces $t_{23} = t_{24} = \dots = t_{2p} = 0$. Continuando de este modo se demuestra, que en efecto, T es diagonal, y así se podrá considerar $D = T$.

FORMA DE JORDAN

Una forma especial a la que se puede reducir toda matriz mediante una transformación de semejanza se desarrolla en tres etapas: la forma de Schur, la forma diagonal en bloques triangulares superiores y finalmente la forma de Jordan (en casos sencillos es posible obtener esta forma directamente) en lugar de hacerlo atravez de las tres etapas.

1.- Formas diagonales en bloques triangulares superiores supongase que A es $p \times p$ se sabe que A es ortogonalmente semejante a su forma de Schur T : una matriz triangular superior con los eigenvalores de A en su diagonal principal. Por el modo en que se contruyo T para la forma de Schur, es claro que se puede construir T de tal modo que cada eigenvalor distinto λ_i se repita $(m_i \text{ veces})$ en posiciones consecutivas a lo largo de la diagonal principal de T .

Para alcanzar la segunda etapa en el desarrollo de la forma de Jordan se demostrara que T a su vez, es semejante a otra matriz triangular superior con la misma estructura y con la propiedad adicional de ser diagonal en bloques: se llamara la forma diagonal en bloques triangulares superiores.

Lema (forma diagonal en bloques triangulares superiores)
 suponga que T es triangular superior y que además, tiene la forma

$$T = \begin{pmatrix} T_{11} & T_{12} & \dots & T_{1s} \\ 0 & T_{22} & \dots & T_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & T_{ss} \end{pmatrix}$$

en donde cada T_{ii} es $m_i \times m_i$ y triangular superior con todos los elementos en la diagonal principal de T_{ii} iguales a λ_i y las λ_i distintas para $1 \leq i \leq s$.

Entonces, T es semejante a una matriz diagonal en bloques triangulares superiores

$$V = \begin{pmatrix} V_1 & 0 & \dots & 0 \\ 0 & V_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & V_s \end{pmatrix}$$

En donde cada v_i es $m_i \times m_i$ y triangular superior; todos los elementos de la diagonal principal de v_i son iguales a λ_i , y desde luego las λ_i son distintas.

Demostración. La idea básica es usar transformaciones de semejanza basadas en las matrices elementales $E_{ij}(c)$ recordando que $E_{ij}(c)$ es no singular y que su inversa es igual a $E_{ij}(-c)$. Suponga que $i < j$ y considere la transformación de semejanza.

$$E_{ij}(-c) T E_{ij}(c) = T'$$

Esta reemplaza el elemento (i, j) de T por

$$\langle T \rangle_{ij} + c (\langle T \rangle_{ii} - \langle T \rangle_{jj})$$

y por lo demás, sólo modifica a los elementos en el i -ésimo renglón (a la derecha del elemento (i, j)) y en la j -ésima columna (arriba de ese elemento).

Si se escogen a i y j para que correspondan a renglones en distintos bloques T_{mm} y T_{nn} entonces $\langle T \rangle_{ii} - \langle T \rangle_{jj} = \lambda_m - \lambda_n \neq 0$.

Por lo tanto, es posible escoger a $c = -\langle T \rangle_{ij} / (\lambda_m - \lambda_n)$ y el elemento (i, j) de la matriz semejante T' sera igual a cero.

Se puede hacer una sucesión de tales transformaciones de semejanza que reemplace los bloques

$$T_{s-1,s}, T_{s-2,s}, T_{s-3,s}, \dots, T_{12}, T_{13}, \dots, T_{1s}$$

en ese orden por bloques cero (siguiendo de abajo hacia arriba y de izquierda a derecha dentro de cada bloque). Esto da v .

Recordando que la diagonal principal de cada v_i tiene todos los elementos iguales a λ_i : $\langle v_i \rangle_{kk} = \lambda_i$ para toda k .

Es posible demostrar que cada una de estas matrices v_i es semejante a una matriz diagonal en bloque triangulares superiores J_i , cada uno de cuyos bloques es un bloque de Jordan $J(\lambda_i)$.

Definición un bloque de Jordan en una matriz cuadrada triangular superior $J(\lambda)$ tal que:

a) Todos sus elementos en la diagonal principal son iguales a λ : $\langle J(\lambda) \rangle_{ii} = \lambda$.

b) Todos sus elementos en la primera sobre la diagonal son iguales a 1: $\langle J(\lambda) \rangle_{i,i+1} = 1$

c) Todos los demás elementos son iguales a cero.
De este modo

$$J = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ & & \dots & & \\ 0 & 0 & 0 & \dots & \lambda \end{pmatrix}$$

Ya que la combinación de semejanzas sucesivas también es una semejanza, las transformaciones de A a T , de T a v y después a bloques de Jordan se pueden expresar como una sola transformación de semejanza.

Teorema 10 (forma de Jordan). Toda matriz A de $p \times p$, que es semejante a una matriz J en la forma de Jordan.

$$J = Q^{-1}AQ, \text{ y } A = QJQ^{-1} \text{ con}$$

$$Q^{-1}AQ = \begin{pmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & \dots & 0 \\ & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_\mu \end{pmatrix} = J$$

En donde cada J_r es un bloque de Jordan $n_r \times n_r$, $\mu = \mu_1 + \dots + \mu_r$ es igual a la suma de las multiplicidades geométricas de los diferentes eigenvalores de A. El mismo eigenvalor diferente puede estar en distintos bloques de Jordan J_r , pero el número total de bloques con ese eigenvalor es igual a su multiplicidad geométrica μ_i , mientras que el número total de elementos en la diagonal principal con ese eigenvalor es igual a su multiplicidad algebraica m_i . Los números n_r y el número total de bloques quedan determinados unívocamente por A.

Este es un teorema de existencia: no se da ningún método para encontrar J en la práctica; sin embargo, la misma existencia de una J en su forma especial permite su cálculo.

Ya que $Q^{-1}AQ = J$, se tiene que $AQ = QJ$. Escribiendo Q en términos de sus columnas como

$$Q = [q_1 \quad q_2 \quad \dots \quad q_p]$$

Se ve que $AQ = QJ$ es equivalente a $Aq_i = \lambda q_i + v_i q_{i-1}$, en

donde λ es el eigenvalor en el bloque de Jordan que afecta a q_i , y v_i es igual a cero o a 1. De manera más precisa, ya que las J_r son $n_r \times n_r$, las columnas de Q afectadas por el bloque de J_r en el producto $QJ = AQ$.

Son exactamente las n_r numeradas.

$n_1 + n_2 + \dots + n_{r-1}$ hasta $n_1 + n_2 + \dots + n_r$.

Por conveniencia, se designarán esas columnas de Q como v_{r1}, \dots, v_{rn} .

Resulta entonces de $AQ = QJ$ que

* $Av_{r1} = \lambda v_{r1}$ y $Av_{rj} = \lambda v_{rj} + v_{rj-1}$ para $j = 2, \dots, n_r$ en donde v_{rj} es la columna de Q numerada de $n_1 + \dots + n_{r-1} + j$ y donde

$Q^{-1}AQ = J$ y $AQ = QJ$ en donde el r -ésimo bloque de Jordan de J es $n_r \times n_r$.

La observación crucial es que (*) significa que v_{r1} es un eigenvector de A; por lo tanto se encuentran los eigenvectores v_{r1} y después se les usa para ayudar a encontrar las v_{rj} restantes.

Se tiene dos bases de un espacio vectorial V de dimensión finita sobre F y una transformación lineal T en V , entonces las matrices de T con respecto a las bases A y B están relacionadas de una manera muy precisa a saber,

$B = C^{-1}AC$, en donde C es la matriz de transición.

Esto condujo a establecer que dos matrices A y B son equivalentes o semejantes si para cierta matriz invertible C ,

$$B = C^{-1}AC.$$

La equivalencia de A y B se denota escribiendo $A \approx B$. Se aprecia que esta relación de equivalencia es muy parecida a la igualdad es decir:

- 1.- $A \approx A$
- 2.- $A \approx B$ implica $B \approx A$
- 3.- $A \approx B$ y $B \approx C$ implica $A \approx C$

Esta tres propiedades nos conducen luego a definir la clase de equivalencia de A mediante

$$cl(A) = \{ B \in M_n(F) \mid B \approx A \}$$

y se descubre que dos equivalencias dadas o son iguales o bien carecen de elementos en común. Surge ahora la pregunta: dada una clase $cl(A)$ ¿existen en ella matrices particulares que indiquen si cierta matriz B pertenece a $cl(A)$? Tales matrices particulares reciben el nombre de formas canónicas. Hay varios tipos de formas canónicas; se aludirá a una de ellas conocida como la forma canónica de Jordan.

Puesto que se necesitan las raíces características de todas las matrices, se trabaja en $M_n(C)$ en vez de $M_n(R)$. Sin embargo, para simplificar la notación se utilizarán aquí literales latinas en lugar de griegas para representar números complejos.

Se señaló que si T pertenece a $M_n(C)$, entonces para cierta matriz invertible C ,

$$C^{-1}TC = \begin{pmatrix} a_1 & \dots & 0 \\ \cdot & \cdot & \cdot \\ 0 & & a_n \end{pmatrix}$$

o sea, $C^{-1}TC$ es triangular superior. Es posible precisar aún más la naturaleza de $C^{-1}TC$. En definitiva se pondrá de manifiesto que T es equivalente a una matriz canónica de Jordan, en el sentido de la.

Definición. Una matriz canónica de Jordan es una matriz de $n \times n$

$$\begin{pmatrix} a_1 & b_1 & \dots & 0 \\ & \cdot & & \\ & & \cdot & \\ & & & b_{n-1} \\ 0 & & & \cdot & a_n \end{pmatrix}$$

tal que para cada s , $1 \leq s \leq n-1$, se tiene que $b_s = 0$ o bien $b_s = 1$ y $a_s = a_{s+1}$.

De ahí que una matriz canónica de Jordan es una matriz de $n \times n$ cuyos elementos distintos de cero están contenidos en la diagonal y en la supradiagonal de modo tal que los componentes diagonales son iguales en bloques para los cuales los componentes supradiagonales son equivalentes a 1.

Por ejemplo, el primer bloque de componentes diagonales iguales es $a_1 = a_2 = \dots = a_k$, en donde $b_r = 1$ para $1 \leq r \leq k-1$ y $b_k = 0$ o bien $k = n$ por lo cual es el bloque de $k \times k$

$$\begin{pmatrix} a & 1 & \dots & 0 \\ & \cdot & & \\ & & \cdot & \\ & & & 1 \\ 0 & & & \cdot & a \end{pmatrix}$$

en donde $a = a_r$ para $1 \leq r \leq k$.

Tomese los bloques A_1, \dots, A_k de componentes diagonales iguales de una matriz canónica de Jordan A , y sean I_1, \dots, I_k las matrices identidad del tamaño correspondientes. Entonces

$$A = \begin{pmatrix} A_1 & \dots & 0 \\ & \cdot & \\ & & \cdot & \\ & & & I_k \\ 0 & & & \cdot & A_k \end{pmatrix}$$

y

$$A_i = \begin{pmatrix} a_i & 1 & \dots & 0 \\ & \ddots & \ddots & \vdots \\ & & \ddots & 1 \\ 0 & & & a_i \end{pmatrix} = a_i I_i + N_i \text{ en donde}$$

$$a_i I_i = \begin{pmatrix} a_i & \dots & 0 \\ & \ddots & \vdots \\ & & \ddots & \vdots \\ 0 & & & a_i \end{pmatrix} \quad \text{y} \quad N_i = \begin{pmatrix} 0 & 1 & \dots & 0 \\ & \ddots & \ddots & \vdots \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix}$$

Se puede intercambiar las posiciones de dos bloques diferentes A_r y A_s y la matriz resultante permanecerá en la misma clase de equivalencia.

Para obtener una matriz de Jordan (J) de una matriz general T a otras muy particulares, se efectúan una serie de reducciones.

El primer paso consiste en encontrar una matriz invertible $C \in M_n(\mathbb{C})$ tal que

$$C^{-1}TC = \begin{pmatrix} A_1 & 0 \\ 0 & A_k \end{pmatrix}$$

En donde cada bloque A_i es una matriz triangular superior que posee únicamente una raíz característica a_i ($1 \leq i \leq k$). Esto significa que la matriz $N_i = A_i - a_i I_i$ es nilpotente, es decir

$N_i^m = 0$ para algún m . Entonces solo es menester hallar

matrices invertibles D_i (de tipo apropiado) tales que la matriz

$D_i^{-1}A_iD_i$ presentan la forma deseada para todo i ya que

$$\begin{pmatrix} D_1^{-1} & 0 \\ & \ddots \\ 0 & D_k^{-1} \end{pmatrix} \begin{pmatrix} A_1 & 0 \\ & \ddots \\ 0 & A_k \end{pmatrix} \begin{pmatrix} D_1 & 0 \\ & \ddots \\ 0 & D_k \end{pmatrix} = \begin{pmatrix} D_1^{-1}A_1D_1 & 0 \\ & \ddots \\ 0 & D_k^{-1}A_kD_k \end{pmatrix}$$

tiene dicha forma. Si $D_i^{-1}N_iD_i$ posee la forma deseada entonces

$D_i^{-1}A_iD_i$ también la posee pues las dos están estrechamente relacionadas:

$$D_i^{-1}A_iD_i = D_i^{-1}(N_i + a_{ii})D_i = D_i^{-1}N_iD_i + a_{ii}$$

Cuando se aborda este aspecto comienza la parte central del razonamiento.

Según se menciona todo esto no es fácil pero tiene su recompensa en primer lugar se vea como la forma canónica de Jordan permite resolver un sistema homogéneo de ecuaciones diferenciales lineales de un modo elegante y eficiente.

Se sustituyen las raíces características en la ecuación $(A - \lambda I) = 0$ de esta manera se encuentran los vectores característicos correspondientes con los cuales se obtiene la matriz Q . A la cual se le obtiene su inversa para después realizar

el producto $Q^{-1}AQ = J$ así obtenemos la matriz de Jordan.

Ejemplos para obtener una matriz de Jordan

1.- Reduzca a la forma de Jordan la matriz

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 6 & -1 & -4 \end{pmatrix}$$

Obtenemos el determinante de $(A - \lambda I) = 0$.

$$\begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 6 & -1 & -4 \end{vmatrix} - \begin{vmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{vmatrix} = \begin{vmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ 6 & -1 & -4 - \lambda \end{vmatrix} \Rightarrow \begin{vmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ 6 & -1 & -4 - \lambda \end{vmatrix} = 0$$

del determinante $(A - \lambda I) = 0$ obtenemos el polinomio característico

$-\lambda^3 - 4\lambda^2 - \lambda + 6 = 0$ el cual factorizamos de la siguiente manera $(\lambda - 1)(-\lambda - 2)(\lambda + 3) = 0$ ya que por el teorema fundamental del álgebra para las raíces de los polinomios; tenemos que las posibles soluciones para este polinomio es 1, 2, 3.

Por lo tanto podemos establecer que: $\lambda - 1 = 0$ implica que $\lambda = 1$; $-\lambda - 2 = 0$ eso implica que $\lambda = -2$; $\lambda + 3 = 0$.

Sustituyendo cada uno de los valores de λ en la matriz: $(A - \lambda I)X = 0$

Sustituimos a $\lambda_1 = 1$

$$(A - \lambda_1 I)x = 0 \text{ eso implica } \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 6 & -1 & -5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

de este producto se obtiene el vector característico $x_1 = [1 \ 1 \ 1]^t$

Sustituimos a $\lambda_2 = -2$

$$(A - \lambda_2 I)x = 0 \text{ eso implica } \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 6 & -1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

de este producto se obtiene el vector característico $x_2 = [1 \ -2 \ 4]^t$

Se sustituye $\lambda_3 = -3$

$$(A - \lambda_3 I)x = 0 \text{ eso implica } \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 1 \\ 6 & -1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

de este producto se obtiene el vector característico $x_3 = [1 \ -3 \ 9]^t$
con los vectores x_1 , x_2 y x_3 obtenemos la matriz siguiente.

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & -2 & -3 \\ 1 & 4 & 9 \end{pmatrix} = Q \text{ obtenemos la inversa de esta matriz.}$$

$$Q^{-1} = \begin{pmatrix} 1/2 & 42/100 & 8/100 \\ 1 & -67/100 & -33/100 \\ -1/2 & 1/4 & 1/4 \end{pmatrix}$$

De esta manera obtenemos $Q^{-1}AQ = J$.

$$\begin{pmatrix} 1/2 & 42/100 & 8/100 \\ 1 & -67/100 & -33/100 \\ -1/2 & 1/4 & 1/4 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 6 & -1 & -4 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & -2 & -3 \\ 1 & 4 & 9 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -3 \end{pmatrix}$$

CAPITULO 2

ECUACIONES LINEALES DE ORDEN SUPERIOR

Una ecuación diferencial lineal n-ésimo orden es una ecuación de la forma:

$$P_0(x) \frac{d^n y}{dx^n} + P_1(x) \frac{d^{n-1} y}{dx^{n-1}} + P_{n-1}(x) \frac{dy}{dx} + P_n(x)y = G(x) \quad (1)$$

Supondremos, a menos que se indique otra cosa, que las funciones P_0, \dots, P_n , y G son funciones continuas de valores reales sobre algún intervalo $\alpha < x < \beta$ y que P_0 no es igual a cero en parte alguna del intervalo.

Si dividimos entonces la ecuación (1) entre $P_0(x)$ obtendremos:

$$L[Y] = \frac{d^n y}{dx^n} + P_1(x) \frac{d^{n-1} y}{dx^{n-1}} + \dots + P_{n-1}(x) \frac{dy}{dx} + P_n(x)Y = g(x) \quad (2)$$

donde se introduce el operador diferencial Lineal L .

Ya que en el desarrollo de la teoría de las ecuaciones diferenciales lineales es útil introducir la notación de operadores diferenciales. Sean P_0, \dots, P_n funciones continuas sobre un intervalo abierto $\alpha < x < \beta$.

Entonces para cualquier función ϕ que tiene n-ésimo derivada sobre $\alpha < x < \beta$, definimos el operador diferencial L por la ecuación:

$$L[\phi] = \phi^n + P_1 \phi^{n-1} + \dots + P_n \phi \quad (3)$$

Note que $L[\phi]$ también es una función sobre $\alpha < x < \beta$; de hecho el propio L , puede considerarse a la vez, funciones de una sola variable real.

El operador L a menudo se escribe como:

$$L = D^n + P_1 D^{n-1} + P_2 D^{n-2} + \dots + P_n D \quad (4)$$

donde D es el operador derivada.

El valor de la función $L[\phi]$ en el punto x es:

$$L[\phi](x) = \phi^n(x) + P_1(x)\phi^{n-1}(x) + \dots + P_n(x)\phi(x). \quad (5)$$

Trataremos primero la ecuación lineal homogénea de n-ésimo orden.

$L[\phi](x) = \phi^n(x) + P_1(x) \phi^{n-1}(x) + \dots + P_n(x) \phi(x) = 0$ (6)
 donde considerando que las funciones $P_1(x), \dots, P_n(x)$ son continuas sobre el intervalo $\alpha < x < \beta$, recordemos que se acostumbra usar el símbolo Y para denotar a $\phi(x)$ por lo tanto se escribirá a menudo:

$$L[Y] = \frac{d^n}{dx^n} + P_1(x) \frac{d^{n-1}y}{dx^{n-1}} + \dots + P_{n-1}(x) \frac{dy}{dx} + P_n(x) Y = 0 \quad (7)$$

Enfatizamos que $L[\phi]$ que está dado en la ecuación (3) es una función, mientras que $L[Y]$ como está dado en la ecuación (7), es el valor de la función $L[\phi]$ en el punto (x) .

Notaremos primero que en virtud de que la ecuación (2) contiene a la e-nésima derivada de Y con respecto a x , se puede decir que se requerirán (n) integraciones para resolver la ecuación (2). Cada una de estas integraciones introduce una constante arbitraria.

Por lo tanto, podemos esperar que para obtener una solución única es necesario especificar (n) condiciones iniciales digamos:

$$Y(x_0) = Y_0, Y'(x_0) = Y'_0, \dots, Y^{(n-1)}(x_0) = Y_0^{(n-1)} \quad (8)$$

donde x_0 puede ser cualquier punto en el intervalo $\alpha < x < \beta$ y

$Y_0, Y'_0, \dots, Y_0^{(n-1)}$ es cualquier conjunto de constantes reales especificadas. Que tal solución existe y que es única esto lo asegura el siguiente teorema:

Teorema 1 de Existencia y Unicidad.

Si las funciones P_1, P_2, \dots, P_n y g son continuas sobre el intervalo abierto $\alpha < x < \beta$, entonces existe una y solo una función $Y = \phi(x)$ que satisface:

$$f(x, y, y', \dots, y^n) = g(x)$$

$$L[x] \frac{d^n y}{dx^n} + P_1(x) \frac{d^{n-1} y}{dx^{n-1}} + P_{n-1}(x) \frac{dy}{dx} + P_n(x) Y = g(x)$$

sobre el intervalo $\alpha < x < \beta$ y las condiciones iniciales especificadas.

$$Y(x_0) = Y_0, Y'(x_0) = Y'_0, \dots, Y^{(n-1)}(x_0) = Y_0^{(n-1)}$$

Teorema 2 de Existencia. Sea f una función continua de variable real, definida en el rectángulo

$$R: |x - x_0| \leq a, |y - y_0| \leq b \quad (a, b, > 0),$$

y sea $|f(x, y)| \leq M$ para toda (x, y) en R .

Además supongamos que f satisface una condición de Lipschitz en R , cuya constante es K . Entonces las aproximaciones sucesivas:

$$Q_0(x) = Y_0, \quad Q_{k+1}(x) = Y_0 + \int_{x_0}^x f(t, Q_k(t))$$

converge en el intervalo $(K = 0, 1, 2, \dots)$,

$$I \quad |x - x_0| < \alpha = \min(a, b/M)$$

a una solución Q del problema con valor inicial:

$$Y' = f(x, y), \quad y(x_0) = Y_0$$

Prueba del Teorema:

a) **Convergencia de $(\phi_k(x))$.** La clave para demostrar este teorema es la observación de que ϕ_k puede escribirse de la siguiente manera:

$\phi_k = \phi_0 + (\phi_1 - \phi_0) + (\phi_2 - \phi_1) + \dots + (\phi_k - \phi_{k-1})$, y en consecuencia, $\phi_k(x)$ es una suma parcial de la serie:

$$\phi_0(x) + \sum_{p=1}^{\infty} [\phi_p(x) - \phi_{p-1}(x)] \quad (1)$$

Por lo tanto demostrar que la secuencia $(\phi_k(x))$ converge es igual a demostrar que (1) es convergente. Para demostrar esto

último necesitamos estimar los términos $\phi_p(x) - \phi_{p-1}(x)$ de (1)

además todas las funciones ϕ_p existen como funciones continuas en I

y $(x, \phi_p(x))$ esta R para x en I , de la misma manera:

$$|\phi_1(x) - \phi_0(x)| \leq M |x - x_0| \quad (2)$$

para x en I .

Escribiendo las relaciones que definen ϕ_2 y ϕ_1 y restandolas entre sí, obtenemos.

$$\phi_2(x) - \phi_1(x) = \int_{x_0}^x [f(t, \phi_2(t)) - f(t, \phi_1(t))] dt.$$

por consiguiente:

$$|\phi_2(x) - \phi_1(x)| \leq \left| \int_{x_0}^x \{f(t, \phi_1(t)) - f(t, \phi_0(t))\} dt \right|$$

y dado que f satisface la condición de Lipschitz.

" La condición de Lipschitz sea f una función definida para (x, y) en un conjunto S , decimos que f satisface una condición de Lipschitz en S , si existe una constante $K > 0$ tal que

$$|f(x, y_1) - f(x, y_2)| \leq K |y_1 - y_2|$$

para toda $(x, y_1), (x, y_2)$ en S .

La constante K se llama constante de Lipschitz. Si f es continua y satisface esta condición en el rectángulo R .

Entonces las aproximaciones sucesivas converge a una solución del problema con valor inicial en el intervalo cerrado $|x - x_0| \leq \alpha$.

$$|f(x, y_1) - f(x, y_2)| \leq k |Y_1 - Y_2|$$

tenemos que:

$$|\phi_2(x) - \phi_1(x)| \leq k \left| \int_{x_0}^x \phi_1(t) - \phi_0(t) dt \right|$$

Usando 2, obtenemos:

$$|\phi_2(x) - \phi_1(x)| \leq kM \left| \int_{x_0}^x (t - x_0) dt \right| = kM \frac{(x - x_0)^2}{2}$$

Así, si $x \geq x_0$

$$|\phi_2(x) - \phi_1(x)| \leq kM \int_{x_0}^x (t - x_0) dt = kM \frac{(x - x_0)^2}{2}$$

El mismo resultado es válido para el caso en que $x \leq x_0$. Necesitamos demostrar por inducción que:

$$|\phi_p(x) - \phi_{p-1}(x)| \leq \frac{Mk^{p-1} |x - x_0|^p}{p!} \quad (3)$$

Ya hemos visto que esta igualdad es válida para $p = 1$ y $p = 2$, supongamos que $x \leq x_0$. Supongamos que 3 es válida para $p = m$. Usando la definición de ϕ_{m+1} y ϕ_m , obtenemos.

$$\phi_{m+1}(x) - \phi_m(x) = \int_{x_0}^x [f(t, \phi_m(t)) - f(t, \phi_{m-1}(t))] dt$$

y así,

$$|\phi_{m+1}(x) - \phi_m(x)| \leq \int_{x_0}^x |f(t, \phi_m(t)) - f(t, \phi_{m-1}(t))| dt$$

Usando la condición de Lipschitz, obtendremos.

$$|\phi_{m+1}(x) - \phi_m(x)| \leq k \int_{x_0}^x |\phi_m(t) - \phi_{m-1}(t)| dt$$

Dado que hemos supuesto (3) es válida para $p = m$, ello da lugar a lo siguiente:

$$|\phi_{m+1}(x) - \phi_m(x)| \leq \frac{Mk^m}{m!} \int_{x_0}^x |t - x_0|^m dt = \frac{Mk^m}{(m+1)!} |x - x_0|^{m+1}$$

Esto es precisamente (3) con $p = m + 1$, y, en consecuencia, (3) es válida para $p = 1, 2, \dots$, por el principio de inducción matemática.

De (3) se sigue que la serie infinita.

$$\phi_0(x) + \sum_{p=1}^{\infty} \{\phi_p(x) - \phi_{p-1}(x)\} \quad (1)$$

es absolutamente convergente en I , esto es, que la serie.

$$|\phi_0(x) + \sum_{p=1}^{\infty} |\phi_p(x) - \phi_{p-1}(x)| \quad (4)$$

es convergente en I . Más todavía, puede observarse que

$$|\phi_p(x) - \phi_{p-1}(x)| \leq \frac{M}{K} \frac{K^p |x - x_0|^p}{p!}$$

lo cual demuestra que el p -ésimo término de la serie dada en (4) es menor o igual que M/K veces el p -ésimo término del desarrollo en serie de potencia de $e^{k|x-x_0|}$. Dado que ésta última serie es convergente, entonces la serie (4) es convergente para $x \in I$. Esto implica que la serie (1) es convergente en I .

Por consiguiente, la k -ésima suma parcial de (1), que es

precisamente $\phi_k(x)$, tiende a un límite $\phi(x)$ cuando $k \rightarrow \infty$, para cada x en I .

b) Propiedades del límite ϕ . Esta función límite es una solución en I para nuestro problema. Primero, vamos a demostrar que ϕ es continua en I . Esto puede hacerse de la siguiente manera.

Si x_1, x_2 , están en I ,

$$|\phi_{k+1}(x_1) - \phi_{k+1}(x_2)| = \left| \int_{x_2}^{x_1} f(t, \phi_k(t)) dt \right| \leq M |x_1 - x_2|$$

lo cual implica, haciendo que $k \rightarrow \infty$.

$$|\phi(x_1) - \phi(x_2)| \leq M |x_1 - x_2| \quad (5)$$

Esto demuestra que cuando $x_2 \rightarrow x_1$,

$\phi(x_2) \rightarrow \phi(x_1)$ esto es, que ϕ es continua en I .

También sustituyendo en (5) $x_1 = x$, $x_2 = x_0$, obtenemos.

$$|\phi(x) - y_0| \leq M |x - x_0|, \quad (x \text{ en } I),$$

lo cual implica que los puntos $(x, \phi(x))$; están en \mathbb{R} para toda x en I .

c) Estimación de $|\phi(x) - \phi_k(x)|$. Vamos a estimar ahora

$$|\phi(x) - \phi_k(x)|.$$

Tenemos

$$\phi(x) = \phi_0(x) + \sum_{p=1}^{\infty} [\phi_p(x) - \phi_{p-1}(x)]$$

y

$$\phi_k(x) = \phi_0(x) + \sum_{p=1}^k [\phi_p(x) - \phi_{p-1}(x)]$$

Por consiguiente, usando la desigualdad (3), vemos que:

$$|\phi(x) - \phi_k(x)| = \left| \sum_{p=k+1}^{\infty} [\phi_p(x) - \phi_{p-1}(x)] \right|$$

$$\begin{aligned}
&\leq \sum_{p=k+1}^{\infty} |[\phi_p(x) - \phi_{p-1}(x)]| \\
&\leq \frac{M}{k} \sum_{p=k+1}^{\infty} \frac{(k \alpha)^p}{p!} \\
&\leq \frac{M}{k} \frac{(k \alpha)^{k-1}}{(k+1)!} \sum_{p=0}^{\infty} \frac{(k \alpha)^p}{p!} \\
&= \frac{M}{k} \frac{(k \alpha)^{k-1}}{(k+1)!} e^{k\alpha} \tag{6}
\end{aligned}$$

Haciendo $c_k = (k \alpha)^{k+1}/(k+1)!$, vemos que $c_k \rightarrow 0$ cuando $k \rightarrow \infty$,

ya c_k es el término general del desarrollo en serie de $e^{k\alpha}$. Según

esto (6) puede escribirse en términos de c_k , de la siguiente manera:

$$|\phi(x) - \phi_k(x)| \leq \frac{M}{K} e^{k\alpha} c_k, \quad (c_k \rightarrow 0, k \rightarrow \infty) \tag{7}$$

d) El límite ϕ es una solución. Para terminar la demostración, necesitamos comprobar que:

$$\phi(x) = Y_0 + \int_{x_0}^x f(t, \phi(t)) dt \tag{8}$$

para toda x en I . El miembro derecho de (8) tiene sentido, siempre y cuando ϕ sea continua en I , y f sea continua en R ; así que la función F dada por $F(t) = f(t, \phi(t))$ es continua en I . Ahora:

$$\phi_{k+1}(x) = Y_0 + \int_{x_0}^x f(t, \phi_k(t)) dt,$$

y

$$\phi_{k+1}(x) \rightarrow \phi(x), \text{ cuando } k \rightarrow \infty.$$

Por consiguiente, para demostrar la ecuación (8), necesitamos comprobar que para cada x en I se cumple lo siguiente:

$$\int_{x_0}^x f(t, \phi_k(t)) dt \rightarrow \int_{x_0}^x f(t, \phi(t)) dt, \quad (k \rightarrow \infty) \quad (9)$$

Tenemos

$$\begin{aligned} & \left| \int_{x_0}^x f(t, \phi(t)) dt - \int_{x_0}^x f(t, \phi_k(t)) dt \right| \\ & \leq \left| \int_{x_0}^x |f(t, \phi(t)) - f(t, \phi_k(t))| dt \right| \leq k \int_{x_0}^x |\phi(t) - \phi_k(t)| dt \end{aligned} \quad (10)$$

Vallendonos del hecho de que f satisface una condición de Lipschitz. La estimación dada por (7). Ahora puede sustituirse en (10) para obtener:

$$\left| \int_{x_0}^x f(t, \phi(t)) dt - \int_{x_0}^x f(t, \phi_k(t)) dt \right| \leq M e^{k\alpha} c_k |x - x_0|$$

Expresión que tiende a cero cuando $k \rightarrow \infty$, para cada x en I . Así queda demostrada la condición (9) y en consecuencia, el hecho de que ϕ satisface la ecuación (8). Con esto queda completa la demostración del teorema de existencia.

De acuerdo a las hipótesis dadas en el teorema de existencia (1), podemos demostrar que la solución obtenida aquí es la única que satisface en I el problema con valor inicial. El método para demostrarlo puede adaptarse para obtener información adicional de importancia en relación con las aproximaciones a soluciones.

Supongamos que tenemos dos problemas con valores iniciales:

$$y' = f(x, y), \quad y(x_0) = y_1, \quad (11)$$

y

$$y' = g(x, y), \quad y(x_0) = y_2, \quad (12)$$

donde f, g , son ambas funciones continuas de variable real, definidas en:

$$R: |x - x_0| \leq a, \quad |y - y_0| \leq b, \quad (a, b > 0),$$

$(x_0, y_1), (x_0, y_2)$ son puntos que están en R . Vamos a demostrar

que si g es próximo a f , y y_2 es próximo a y_1 , entonces toda solución ψ de (12) que esté definida en un intervalo I que contenga a x_0 , se aproxima en I a una solución ϕ de (11).

Supongamos que existen ciertas constantes no negativas c, δ , tales que:

$$|f(x,y) - g(x,y)| \leq c, \quad [(x,y) \in R] \quad (13)$$

$$y \quad |y_1 - y_2| \leq \delta \quad (14)$$

Entonces tenemos el siguiente resultado.

Teorema 3 Sean f, g , funciones continuas en R y supongamos que f satisface aquí una condición de Lipschitz, cuya constante de Lipschitz es K . Sean ϕ, ψ soluciones de (11) y (12), respectivamente, en un intervalo I que tenga a x_0 , tales que sus gráficas estén contenidas en R .

Si valen las desigualdades (13) y (14), entonces:

$$|\phi(x) - \psi(x)| \leq \delta e^{K|x-x_0|} + \frac{c}{K} (e^{K|x-x_0|} - 1) \quad (15)$$

para todo x en I .

En la última desigualdad (15) si hacemos $g = f$ y $y_0 = y_1 = y_2$ vemos que en este caso podemos escoger $c = 0, \delta = 0$, y entonces

Tenemos

Corolario. 1. (Teorema de Unicidad). Sea f continua y tal que satisfaga una condición de Lipschitz.

Si ϕ y ψ son dos soluciones de

$$y' = f(x,y), \quad y(x_0) = y_0,$$

Definidas en un intervalo I que contenga a x_0 , entonces

$$\phi(x) = \psi(x) \quad \text{para todo } x \text{ en } I.$$

Observece que para garantizar la unicidad es necesario imponer a f una restricción adicional a la continuidad. El problema de valores iniciales.

Por intuición puede verse que si tenemos una secuencia de funciones $g_k \rightarrow f$ definidas en R , y una secuencia $y_k \rightarrow y_0$, podemos esperar que las soluciones ψ_k de

$$y' = g_k(x,y), \quad y(x_0) = y_k \quad (16)$$

tiendan a la solución ϕ de:

$$y' = f(x, y) \quad y(x_0) = y_0 \quad (17)$$

esta es una consecuencia directa de (15) supongamos que las g_k son continuas en R y que existen ciertas constantes c_k tales que

$$|f(x, g) - g_k(x, y)| \leq c_k \quad (\text{toda } (x, y) \text{ en } R) \quad 18$$

y otras constantes δ_k tales que

$$|y_k - y_0| \leq \delta_k.$$

donde c_k y δ_k tienden a 0 cuando $k \rightarrow \infty$ aplicando (15) obtenemos.

Corolario 2 Sea f continua y tal que satisfaga una condición de Lipschitz en R :

Supongamos que las g_k ($k = 1, 2, \dots$) son continuas en R y además satisfacen (18) para ciertas constantes,

$c_k \rightarrow 0$, ($k \rightarrow \infty$) y, hagamos que $y_k \rightarrow y_0$ ($k \rightarrow \infty$).

Si ψ_k es una solución de (16) definida en un intervalo I que contenga a x_0 y ϕ es la solución de (17) en I , entonces $\psi_k(x) \rightarrow \phi(x)$ en I .

Demostración del Teorema 2

De acuerdo con (11) y (12)

Vemos que

$$\phi(x) = y_1 + \int_{x_0}^x f(t, \phi(t)) dt$$

$$\psi(x) = y_2 + \int_{x_0}^x g(t, \psi(t)) dt$$

y en consecuencia,

$$\begin{aligned} \phi(x) - \psi(x) &= y_1 - y_2 + \int_{x_0}^x [f(t, \phi(t)) - g(t, \psi(t))] dt \\ &= y_1 - y_2 + \int_{x_0}^x [f(t, \phi(t)) - f(t, \psi(t))] dt \end{aligned}$$

$$+ \int_{x_0}^x [f(t, \psi(t)) - g(t, \psi(t))] dt.$$

Aplicando (13) y (14) y el hecho de que f satisface una condición de Lipschitz, cuya constantes K , obtenemos, para $x \geq x_0$.

$$|\Phi(x) - \Psi(x)| \leq \delta + K \int_{x_0}^x |\Phi(t) - \Psi(t)| dt + \epsilon (x - x_0) \quad (19)$$

$$\text{Si } E(x) = \int_{x_0}^x |\Phi(t) - \Psi(t)| dt$$

Vemos que (18) puede escribirse de la siguiente manera:

$$E'(x) - K E(x) \leq \delta + \epsilon (x - x_0) \quad (20)$$

Esta es una desigualdad diferencial de primer orden que podemos "resolver" las ecuaciones diferenciales lineales de primer orden.

Multiplcando (20) por $e^{-K(x-x_0)}$

obtenemos, después de sustituir x por t , lo siguiente:

$$[e^{-K(t-x_0)} E]'(t) \leq \delta e^{-K(t-x_0)} + \epsilon (t - x_0) e^{-K(t-x_0)}$$

Integrando desde x_0 hasta x obtenemos *

$$e^{-K(x-x_0)} E(x) \leq \frac{\delta}{K} [1 - e^{-K(x-x_0)}] \\ + \frac{\epsilon}{K^2} [-K(x - x_0) - 1] e^{-K(x-x_0)} + \frac{\epsilon}{K^2},$$

Multiplcando ambos miembros de ésta desigualdad por $e^{K(x-x_0)}$
Hallamos que:

$$E(x) \leq \frac{\delta}{K} [e^{K(x-x_0)} - 1] - \frac{\epsilon}{K^2} [K(x - x_0) + 1] + \frac{\epsilon}{K^2} e^{K(x-x_0)}$$

y sustituyendo esto en (19) obtenemos finalmente.

$$|\Phi(x) - \Psi(x)| \leq \delta e^{K(x-x_0)} + \frac{\epsilon}{K} [e^{K(x-x_0)} - 1]$$

Que es precisamente la desigualdad 15 para $x \geq x_0$. Una demostración similar es válida para el caso en que $x \leq x_0$.

* observece que si $c \neq 0$; $\int t e^{ct} dt = \frac{1}{c^2} (ct - 1)e^{ct}$

También se emplea el hecho de que $E(x_0) = 0$

Notemos que si las funciones P_1, P_2, \dots, P_n son constantes, realmente controlaremos la solución de la ecuación.

$$L[y] = \frac{d^n y}{dx^n} + P_1(x) \frac{d^{n-1} y}{dx^{n-1}} + P_{n-1}(x) \frac{dy}{dx} + P_n(x) y = g(x)$$

Que satisfaga las condiciones iniciales.

$$Y(x_0) = Y_0, Y'(x_0) = Y'_0, \dots, Y^{(n-1)}(x_0) = Y_0^{(n-1)}$$

Discutiremos primero el problema de resolver la ecuación homogénea o complementaria.

$$Y^{(n)} + P_1(x) Y^{(n-1)} + \dots + P_{n-1}(x) Y' + P_n(x) Y = 0$$

Teorema 1

* Teoría general de las ecuaciones lineales de n-ésimo orden.

Si las funciones Y_1, Y_2, \dots, Y_n son soluciones de la ecuación diferencial lineal homogénea de n-ésimo orden.

$$L[Y] = Y^{(n)} + P_1(x) Y^{(n-1)} + \dots + P_n(x) Y = 0 \quad (21)$$

Se deduce por cálculo directo, que la combinación lineal.

$$Y = C_1 Y_1(x) + C_2 Y_2(x) + \dots + C_n Y_n(x) \quad (22)$$

donde C_1, \dots, C_n son constantes arbitrarias, también es una solución de (21).

Es lógico preguntarse si toda solución de la ecuación (21) puede expresarse como una combinación lineal de Y_1, Y_2, \dots, Y_n . Esto será cierto si, independientemente de las condiciones iniciales.

$$Y(x_0) = Y_0, Y'(x_0) = Y'_0, \dots, Y^{(n-1)}(x_0) = Y_0^{(n-1)} \quad (23)$$

Que se especifiquen, es posible elegir las constantes C_1, \dots, C_n de tal manera que la combinación lineal (22) satisfaga las condiciones iniciales.

Específicamente, para cualquier elección del punto x_0 en :
 $\alpha < x < \beta$ y para cualquier elección de,

$Y_0, Y_0', Y_0'', \dots, Y_0^{(n-1)}$ debe ser
 posible determinar C_1, \dots, C_n de tal manera que las ecuaciones:

$$\begin{aligned} C_1 Y_1(x_0) + \dots + C_n Y_n(x_0) &= Y_0 \\ C_1 Y_1'(x_0) + \dots + C_n Y_n'(x_0) &= Y_0' \\ &\vdots \\ C_1 Y_1^{(n-1)}(x_0) + \dots + C_n Y_n^{(n-1)}(x_0) &= Y_0^{(n-1)} \end{aligned} \quad (24)$$

se satisfagan. Las ecuaciones (24) pueden resolverse siempre para las constantes C_1, C_2, \dots, C_n que el determinante de los coeficientes no se anule por otra parte, si éste determinante se anula, siempre se puede escoger valores de $Y_0, Y_0', \dots, Y_0^{(n-1)}$ tales que las ecuaciones (24) no tengan una solución.

Por lo tanto, una condición necesaria y suficiente para la existencia de una solución de las ecuaciones (24) para valores arbitrarios de $Y_0, Y_0', Y_0'', \dots, Y_0^{(n-1)}$ es que el wronskiano

$$W(Y_1, Y_2, \dots, Y_n) = \begin{vmatrix} Y_1 & Y_2 & \dots & Y_n \\ Y_1' & Y_2' & \dots & Y_n' \\ Y_1^{(n-1)} & Y_2^{(n-1)} & \dots & Y_n^{(n-1)} \end{vmatrix} \quad (25)$$

no se anule en $x = x_0$

Supuesto que x_0 puede ser cualquier punto en el intervalo $\alpha < x < \beta$ es necesario y suficiente que $w(Y_1, Y_2, \dots, Y_n)$ no se anule en todo punto del intervalo.

Tomaremos como ejemplo las ecuaciones lineales homogéneas de 2do. orden.

$$L[\phi](x) = \phi''(x) + P(x) + \phi'(x) + g(x) \phi(x) = 0 \quad (26)$$

donde consideramos las funciones p y q son continuas en un intervalo $\alpha < x < \beta$.

Recordemos que se acostumbra usar el símbolo $\{Y\}$ para denotar a $\phi(x)$ y a menudo escribiremos en lugar de la ecuación (26).

$$L\{Y\} = y'' + p(x)y' + q(x)y = 0 \quad (27)$$

Enfatizaremos que $L[\phi]$ es función mientras que $L[Y]$ como está dado en (26), es el valor de la función $L[\phi]$ en el punto x .

Tomando en cuenta μ_1, μ_2 son funciones derivables, entonces:

$$\{C_1\mu_1(x) + C_2\mu_2(x)\}' = C_1\mu_1'(x) + C_2\mu_2'(x)$$

donde C_1 y C_2 son constantes arbitrarias de lo cual se puede deducir el importante teorema.

Teorema 4 Si $y = Y_1(x)$ y $Y = Y_2(x)$ son soluciones de la ecuación diferencial (2)

$$L[Y] = Y' + p(x)Y' + q(x)Y = 0$$

Entonces la combinación lineal $Y = C_1Y_1(x) + C_2Y_2(x)$ donde C_1 y C_2 son constantes arbitrarias, también es una solución de la ecuación (27).

Demostración.

Debemos demostrar que si:

$$\begin{aligned} L\{Y_1\} &= Y_1'' + pY_1' + qY_1 = 0 & y \\ L\{Y_2\} &= Y_2'' + pY_2' + qY_2 = 0 \end{aligned}$$

entonces

$$\begin{aligned} L\{C_1Y_1 + C_2Y_2\} &= 0 \text{ pero} \\ L\{C_1Y_1 + C_2Y_2\} &= (C_1Y_1 + C_2Y_2)'' + p(C_1Y_1 + C_2Y_2)' + q(C_1Y_1 + C_2Y_2) \\ &= C_1(Y_1'' + pY_1' + qY_1) + C_2(Y_2'' + pY_2' + qY_2) \\ &= C_1L\{Y_1\} + C_2L\{Y_2\} = 0 \end{aligned}$$

lo cual prueba el teorema.

Si hacemos $C_2 = 0$ en la demostración anterior se obtiene el resultado de que la función Y_1 es una solución de la ecuación (2), entonces cualquier múltiplo constante de Y_1 también es una solución de dicha ecuación.

La prueba del teorema 2, demuestra que para cualesquiera dos funciones μ_1, μ_2 que tengan segunda derivadas continuas, y para cualesquiera constantes arbitrarias C_1, C_2 ,

$$L\{C_1\mu_1 + C_2\mu_2\} = C_1L\{\mu_1\} + C_2L\{\mu_2\}$$

un operador con esta propiedad se le conoce como operador lineal, y en este caso es un operador lineal de 2do. orden. El hecho de que una combinación lineal de soluciones de una ecuación lineal homogénea sea también solución de la ecuación, tiene una importancia fundamental y a menudo, se menciona como el principio de Superposición.

Este principio se ilustra en los siguientes ejemplos sencillos:

Ejemplo 1. Verificar, por medio del cálculo directo que

$\phi(x) = C_1 \cos x + C_2 \sin x$ es una solución de la ecuación diferencial $y'' + y = 0$

sustituyendo y por $\phi(x)$ tenemos que

$$\begin{aligned}\phi''(x) + \phi(x) &= (C_1 \cos x + C_2 \sin x)'' + (C_1 \cos x + C_2 \sin x) \\ &= C_1 [(\cos x)'' + \cos x] + C_2 [(\sin x)'' + \sin x] = 0 \\ &= C_1 [-\cos x + \cos x] + C_2 [-\sin x + \sin x] = 0\end{aligned}$$

Ejemplo 2.- Verificar que $\phi(x) = x + 1$ es una solución de la ecuación diferencial $y'' + 3y' + y = x + 4$ pero que $\psi(x) = 2\phi(x)$ no es solución.

Ya que $\phi'(x) = 1$, $\phi''(x) = 0$ Tenemos que

$$\phi''(x) + 3\phi'(x) + \phi(x) = 0 + 3(1) + (x + 1) = x + 4$$

Sin embargo

$$\psi''(x) + 3\psi'(x) + \psi(x) = 0 + 3(2) + 2(x + 1)$$

no es una contradicción puesto que la ecuación diferencial no es homogénea.

Ejemplo 3.- Demostrar que si las funciones Y_1 y Y_2 son solución de la ecuación $L[Y] = Y'' + Y^2 = 0$ no necesariamente se deduce que la combinación lineal $C_1Y_1 + C_2Y_2$ es una solución.

Sustituyendo tenemos:

$$\begin{aligned}L[C_1Y_1 + C_2Y_2] &= (C_1Y_1 + C_2Y_2)'' + (C_1Y_1 + C_2Y_2)^2 \\ &= C_1Y_1'' + C_2Y_2'' + C_1^2 Y_1^2 + 2 C_1C_2Y_1Y_2 \\ &\quad + C_2^2 Y_2^2 \neq C_1 [Y_1] + C_2 L[Y_2]\end{aligned}$$

coeficientes, no sea cero es decir:

$$\begin{pmatrix} Y_1(x_0) & Y_2(x_0) \\ Y_1'(x_0) & Y_2'(x_0) \end{pmatrix} = Y_1(x_0)Y_2'(x_0) - Y_1'(x_0)Y_2(x_0) \neq 0$$

entonces, con C_1 y C_2 determinados por las ecuaciones (3) las funciones $C_1Y_1 + C_2Y_2$ y ϕ satisfacen la mismas condiciones iniciales en $X = X_0$ así como la misma ecuación diferencial (2) por el teorema siguiente.

Teorema 5 .- Si las funciones p , q y g son continuas sobre el intervalo abierto $\alpha < x < \beta$, entonces existe una función y solo una $y = \phi(x)$ que satisface la ecuación diferencial.

$$\frac{d^2y}{dx^2} + p(x) \frac{dy}{dx} + q(x)y = g(x)$$

$$y'' + p(x)y' + q(x)y = g(x)$$

sobre el intervalo completo $\alpha < x < \beta$ y las condiciones iniciales $Y(x_0) = Y_0$, $Y'(x_0) = Y_0'$ en un punto particular X_0 en el intervalo.

Siempre podemos elegir un x_0 tal que se cumpla (4) si sabemos que se cumple para todo x_0 en $\alpha < x < \beta$. En este caso o siempre pueden resolverse las ecuaciones (3) independientemente de los valores x_0 , $\phi'(x_0)$.

De esta forma probamos el siguiente teorema.

Teorema 6.- Si las funciones p y q son continuas sobre el intervalo abierto $\alpha < X < \beta$ y si Y_1 y Y_2 son soluciones de la ecuación diferencial (2).

$$L\{Y\} = Y'' + p(X)Y' + q(X)Y = 0$$

que satisfacen la condición.

$$Y_1(X)Y_2'(X) - Y_1'(X)Y_2(X) \neq 0$$

para todo punto en $\alpha < X < \beta$ entonces cualquier solución de la ecuación (2) sobre el intervalo $\alpha < X < \beta$ puede expresarse como una combinación lineal de Y_1 y Y_2 .

A la combinación lineal $C_1Y_1 + C_2Y_2$ con C_1 y C_2 arbitrarias se le conoce como solución general de la ecuación (2).

Los siguientes teoremas nos sirven para ver como determina si las funciones Y_1 y Y_2 forman un conjunto fundamental esto es como se determina si se anula $w(Y_1, Y_2)$ en el intervalo $\alpha < X < \beta$.

Teorema 7.- Si las funciones p y q son continuas sobre el intervalo abierto $\alpha < X < \beta$ y si las funciones Y_1 y Y_2 son soluciones de la ecuación diferencial (2).

$$L \{Y\} = Y'' + p(x) Y' + q(x)Y = 0$$

sobre $\alpha < x < \beta$ entonces $w\{Y_1, Y_2\}$ se anula idénticamente, o bien, no se anula en $\alpha < x < \beta$.

Teorema 8.- Si las funciones p y q son continuas sobre el intervalo abierto $\alpha < x < \beta$ y si las funciones Y_1 y Y_2 son soluciones de la Ecuación Diferencial (2).

$$L \{Y\} = Y'' + p(x) Y' + q(x)Y = 0$$

sobre el intervalo $\alpha < x < \beta$ donde $w\{Y_1, Y_2\}$ no es cero, entonces, cualquier solución $y = \phi(x)$ de la ecuación (2) puede expresarse en la forma:

$$\phi(x) = C_1 Y_1(x) + C_2 Y_2(x)$$

Finalmente el siguiente teorema nos demuestra que realmente existe un conjunto fundamental de soluciones de la ecuación (2).

$$L \{Y\} = Y'' + p(x)Y' + q(x)Y = 0$$

Teorema 9.- Si las funciones p y q son continuas sobre el intervalo abierto $\alpha < x < \beta$ entonces existe un conjunto fundamental de soluciones de la ecuación diferencial (2).

$$L \{Y\} = Y'' + p(x) Y' + q(x)Y = 0$$

sobre el intervalo $\alpha < x < \beta$.

Del teorema 6.- se deduce que existe Y_1 y Y_2 para los problemas con valores iniciales.

$$Y'' + p(x)Y' + q(x)Y = 0 \quad Y(C) = 1, \quad Y'(C) = 0$$

$$Y'' + p(x)Y' + q(x)Y = 0 \quad Y(C) = 0 \quad Y'(C) = 1$$

sobre el intervalo $\alpha < x < \beta$, fácilmente se ve que $w\{Y_1, Y_2\}(C) = 1 \neq 0$ por lo tanto el teorema 9 se deduce que Y_1 y Y_2 forman un conjunto fundamental de soluciones de la ecuación 2.

$$L \{Y\} = Y'' + p(x)Y' + q(x)Y = 0$$

Teorema 10.- Si las funciones P_1, P_2, \dots, P_n son continuas sobre el intervalo abierto $\alpha < x < \beta$, si las funciones Y_1, Y_2, \dots, Y_n son soluciones de la ecuación.

$$L \{Y\} = Y'' + P_1(x) Y^{(n-1)} + \dots + P_n(x) Y = 0 \text{ y si}$$

$w(Y_1, \dots, Y_n)(x) \neq 0$ o al menos en un punto $\alpha < x < \beta$.

entonces cualquier solución de la ecuación (1) puede expresarse como una combinación lineal de las soluciones Y_1, Y_2, \dots, Y_n .

Tal conjunto se conoce como conjunto fundamental de soluciones de (1). La existencia de un conjunto fundamental puede probarse. Ya que de todas las soluciones de la ecuación (1) son de la forma:

$$Y = C_1 Y_1(x) + C_2 Y_2(x) + \dots + C_n Y_n(x)$$

se le llama solución general para referirse a una combinación lineal arbitraria y se dice que las funciones Y_1, \dots, Y_n son linealmente independientes sobre $\alpha < x < \beta$ si no existe un conjunto de constantes C_1, C_2, \dots, C_n (excepto $C_1 = C_2 = \dots = C_n = 0$) tal que $C_1 Y_1(x) + C_2 Y_2(x) + \dots + C_n Y_n(x) = 0$ para todo x en $\alpha < x < \beta$. Si Y_1, \dots, Y_n son soluciones de la ecuación (1).

También puede demostrarse que una condición necesaria y suficiente para que sean linealmente independientes es que $w(Y_1, \dots, Y_n)$ no se anule sobre $\alpha < x < \beta$ por lo tanto las funciones que forman un conjunto fundamental de soluciones de la ecuación.

$$L\{Y\} = Y^n + P_1(x)Y^{(n-1)} + \dots + P_n(x)Y = 0$$

son linealmente independientes y un conjunto linealmente independiente de n soluciones de la ecuación.

$$L\{Y\} = Y^n + P_1(x)Y^{n-1} + \dots + P_n(x)Y = 0$$

forman un conjunto fundamental de soluciones de dicha ecuación.

Problema no homogéneo.

Consideramos la ecuación.

$$L\{Y\} = Y^n + P_1(x)Y^{(n-1)} + \dots + P_n(x)Y = g(x) \quad (1)$$

Inmediatamente se deduce a partir de la linealidad del operador L , que $L\{Y_{P1} - Y_{P2}\} = g - g = 0$.

Por lo tanto, la diferencia de cualquier par de soluciones de la ecuación no homogénea (1) es una solución de la ecuación homogénea.

$$L\{Y\} = Y^n + p_1(x)Y^{(n-1)} + \dots + P_n(x)Y = 0 \quad (2)$$

puesto que cualquier solución de la ecuación homogénea puede

expresarse como una combinación lineal de un conjunto fundamental de soluciones Y_1, Y_2, \dots, Y_n , se deduce que cualquier solución de la ecuación (1) puede escribirse como:

$$Y = Y_0(x) + Y_p(x) \\ = C_1 Y_1(x) + C_2 Y_2(x) + \dots + C_n Y_n(x) + Y_p(x) \quad (3)$$

donde Y_p es cualquier solución de la ecuación no homogénea (1).

La combinación lineal (3) generalmente se conoce como solución general de la ecuación no homogénea (1).

El problema principal es determinar un conjunto fundamental de soluciones Y_1, Y_2, \dots, Y_n .

Si los coeficientes en la ecuación diferencial son constantes, este es un problema muy simple.

Si los coeficientes no son constantes se utilizan métodos numéricos, o métodos de series.

Existe una importante relación entre los sistemas de ecuaciones y las ecuaciones aisladas de orden arbitrario, en efecto, una ecuación de n -ésimo orden.

$$Y^{(n)} = F(t, Y, Y', \dots, Y^{(n-1)}) \quad (1)$$

siempre puede reducirse a un sistema de n ecuaciones y las ecuaciones de primer orden que tiene una forma un tanto especial.

Para mostrar esto, introduciremos las variables x_1, x_2, \dots, x_n definidas por:

$$x_1 = Y, x_2 = Y', x_3 = Y'', \dots, x_n = Y^{(n-1)} \quad (2)$$

independientemente se deduce que

$$x_1' = x_2, x_2' = x_3, \dots, x_{n-1}' = x_n \quad (3)$$

y de la ecuación (1)

$$x_n' = F(t, x_1, x_2, \dots, x_n) \quad (4)$$

Las ecuaciones (3) y (4) son un caso especial del sistema más general.

$$\begin{aligned}
 x_1' &= F_1(t, x_1, \dots, x_n), \\
 x_2' &= F_2(t, x_1, \dots, x_n), \\
 &\vdots \\
 x_n' &= F_n(t, x_1, \dots, x_n),
 \end{aligned}
 \tag{5}$$

se dice que el sistema de ecuaciones (5) tiene una solución sobre el intervalo $\alpha < x < \beta$ si existe un conjunto de (n) funciones $x_1 = \phi_1(t), \dots, x_n = \phi_n(t)$ que sean diferenciales en todos los puntos en $\alpha < x < \beta$ y que satisfagan el sistema (5) en todos los puntos en este intervalo.

Además del sistema dado de ecuaciones diferenciales, también pueden darse condiciones iniciales de la forma:

$$x_1(t_0) = x_1^0, x_2(t_0) = x_2^0, \dots, x_n(t_0) = x_n^0 \tag{6}$$

donde t_0 es un valor específico de t en $\alpha < x < \beta$, y

x_1^0, \dots, x_n^0 son números preescritos.

Las ecuaciones diferenciales (5) y las condiciones iniciales (6), juntas, forman un problema con valores iniciales.

Para garantizar que el problema con valores iniciales (5) y (6), tienen una solución y es única, es necesario imponer ciertas condiciones sobre las funciones F_1, F_2, \dots, F_n .

Teorema 11. - Suponga que cada una de las funciones F_1, \dots, F_n y cada una de las derivadas parciales $\delta F_1 / \delta x_1, \dots, \delta F_1 / \delta x_n, \dots, \delta F_n / \delta x_1, \dots, \delta F_n / \delta x_n$ son continuas en una región R del espacio t, x_1, x_2, \dots, x_n que contiene al punto $(t_0, x_1^0, x_2^0, \dots, x_n^0)$.

Entonces, existe un intervalo $|t - t_0| < \eta$ en el que existe una solución única $x_1 = \phi_1(t), \dots, x_n = \phi_n(t)$ del sistema de ecuaciones diferenciales (5), la cual también satisface las condiciones iniciales (6).

Si cada una de las funciones F_1, \dots, F_n en las ecuaciones (5) es una función lineal de las variables dependientes x_1, \dots, x_n entonces se dice que el sistema de ecuaciones es lineal.

Por lo tanto, el sistema más general de (n) ecuaciones lineales de primer orden tiene la forma:

$$\begin{aligned}
 x'_1 &= P_{11}(t)x_1 + \dots + P_{1n}(t)x_n + g_1(t) \\
 &\cdot \phantom{= P_{11}(t)x_1} \phantom{+ P_{1n}(t)x_n} \\
 &\cdot \phantom{= P_{11}(t)x_1} \phantom{+ P_{1n}(t)x_n} \\
 x'_n &= P_{n1}(t)x_1 + \dots + P_{nn}(t)x_n + g_n(t)
 \end{aligned}
 \tag{7}$$

Si cada una de las funciones g_1, \dots, g_n es idénticamente cero, entonces se dice que el sistema (7) es homogéneo; de lo contrario es no homogéneo.

Para el sistema lineal (7), el teorema de existencia y unicidad es más sencillo y tiene una conclusión más fuerte.

Teorema 12.- Si las funciones $P_{11}, \dots, P_{nn}, g_1, \dots, g_n$ son continuas sobre un intervalo abierto $\alpha < x < \beta$, que contiene al punto $t = t_0$, entonces existe una solución única $x_1 = \Phi_1(t), \dots, x_n = \Phi_n(t)$ del sistema de ecuaciones diferenciales (7), que también satisface las condiciones iniciales (6).

Esta solución es válida en todo intervalo $\alpha < x < \beta$.

CAPITULO 2.1

TEORIA BASICA DE LOS SISTEMAS DE ECUACIONES LINEALES DE PRIMER ORDEN

La teoría general de un sistema de n ecuaciones lineales de primer orden.

$$\begin{aligned} X'_1 &= P_{11}(t)X_1 + \dots + P_{1n}(t)X_n + g_1(t) \\ &\vdots \\ X'_n &= P_{n1}(t)X_1 + \dots + P_{nn}(t)X_n + g_n(t) \end{aligned} \tag{1}$$

esta estrechamente ligada con la de una sola ecuación lineal de n -ésimo orden.

Para analizar el sistema (1) más efectivamente lo escribiremos en notación matricial.

Considerando $X_1 = \phi_1(t)$, ..., $X_n = \phi_n(t)$ como componentes de un vector $X = \phi(t)$; de forma semejante, $g_1(t)$, ..., $g_n(t)$ son componentes de un vector $g(t)$ y $P_{11}(t)$, ..., $P_{nn}(t)$ son elementos de una matriz $P(t)$ de $n \times n$.

Entonces la ecuación (1) toma la forma.

$$X' = P(t)X + g(t) \tag{2}$$

se dice que un vector $X = \phi(t)$ es solución de la ecuación (2) si sus componentes satisfacen el sistema de ecuaciones (1).

En toda esta sección supondremos que P y g son continuas sobre algún intervalo $\alpha < t < \beta$ o sea que P_{11} , ..., P_{nn} ; g_1 , ..., g_n cada una de estas funciones escalares es continua allí.

De acuerdo al teorema 12, esto es suficiente para garantizar la existencia de soluciones de la ecuación (2), sobre el intervalo $\alpha < t < \beta$.

Es conveniente considerar primero la ecuación homogénea,

$$X' = P(t)X \tag{3}$$

que se obtiene de (2) haciendo $g(t) = 0$

Usaremos la notación.

$$X^{(1)}(t) = \begin{pmatrix} X_{11}(t) \\ X_{21}(t) \\ \vdots \\ X_{n1}(t) \end{pmatrix}, \dots, X^{(k)}(t) = \begin{pmatrix} X_{1k}(t) \\ X_{2k}(t) \\ \vdots \\ X_{nk}(t) \end{pmatrix}; \quad (4)$$

para indicar soluciones específicas del sistema (3) observamos que $X_{ij}(t) = X_i^{(j)}(t)$ se refiere a la i -ésima componente de la j -ésima solución $X^{(j)}(t)$.

Los principales hechos acerca de la estructura de los sistemas (3) se establecen en los siguientes teoremas.

Teorema 13.- Si las funciones vectoriales $X^{(1)}$ y $X^{(2)}$ son soluciones del sistema (3), entonces la combinación lineal $C_1 X^{(1)} + C_2 X^{(2)}$, también es una solución, para cualquier constantes C_1 y C_2 .

Esto es el principio de superposición: se prueba simplemente derivando $C_1 X^{(1)} + C_2 X^{(2)}$ y tomando en cuenta que $X^{(1)}$ y $X^{(2)}$ satisfacen la ecuación (3) por aplicaciones repetidas del teorema (13), se llega a la conclusión de que si $X^{(1)}, \dots, X^{(k)}$ son soluciones de la ecuación (3) entonces.

$$X = C_1 X^{(1)}(t) + \dots + C_k X^{(k)}(t) \quad (5)$$

también es una solución, para cualquier constantes C_1, \dots, C_k .

Como un ejemplo, puede verificarse que:

$$X^{(1)}(t) \begin{pmatrix} e^{3t} \\ 2e^{3t} \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} e^{3t}, \quad X^{(2)}(t) = \begin{pmatrix} e^{-t} \\ -2e^{-t} \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \end{pmatrix} e^{-t} \quad (6)$$

Satisface la ecuación:

$$X' = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix} X \quad (7)$$

De acuerdo con el teorema 13.

$$X = C_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} e^{3t} + C_2 \begin{pmatrix} 1 \\ -2 \end{pmatrix} e^{-t} = C_1 X^{(1)}(t) + C_2 X^{(2)}(t) \quad (8)$$

también satisface la ecuación 7.

Como se indicó antes, aplicando repetidas veces el teorema 13 se deduce que toda combinación lineal infinita de soluciones de la ecuación (3) es nuevamente una solución.

Por analogía con los casos anteriores, resulta razonable esperar que, basta formar combinaciones lineales de n soluciones adecuadamente seleccionadas. Por lo tanto, sean $X^{(1)}, \dots, X^{(n)}$, n soluciones del sistema (3) de n -ésimo orden y considere la matriz $X(t)$ cuyas columnas son los vectores $X^{(1)}(t), \dots, X^{(n)}(t)$:

$$X(t) = \begin{pmatrix} X_{11}(t) & \dots & X_{1n}(t) \\ \vdots & & \vdots \\ X_{n1}(t) & \dots & X_{nn}(t) \end{pmatrix} \quad (9)$$

Recuerde que las columnas de $X(t)$ son linealmente independientes para un valor dado de t si de $t \neq 0$ para ese valor de t . Este determinante se llama Wronskiano de las n soluciones:

$X^{(1)}, \dots, X^{(n)}$, también se denota por

$$W [X^{(1)}, \dots, X^{(n)}]; \text{ esto es,}$$

$$W [X^{(1)}, \dots, X^{(n)}] = \det. X \quad (10)$$

Entonces las soluciones $X^{(1)}, \dots, X^{(n)}$ son linealmente independientes en un punto si y solo si,

$$W [X^{(1)}, \dots, X^{(n)}] \text{ no es cero allí.}$$

Teorema 14.- Si las funciones vectoriales $X^{(1)}, \dots, X^{(n)}$ son soluciones linealmente independientes del sistema (3), para cada punto en el intervalo $\alpha < t < \beta$, entonces cada solución $X = \Phi(t)$ del sistema (3) puede expresarse como una combinación lineal de $X^{(1)}, \dots, X^{(n)}$,

$$\Phi(t) = C_1 X^{(1)}(t) + \dots + C_n X^{(n)}(t), \quad (11)$$

exactamente de esta manera.

Antes de probar el teorema 14, observe que, de acuerdo con el teorema 13, todas las expresiones de la forma (11) son soluciones del sistema (3), mientras que por el teorema 14, todas las soluciones de la ecuación (3) puede escribirse en la forma (11).

Si las constantes C_1, \dots, C_n se imaginan como arbitrarias, entonces la ecuación (11) incluye todas las soluciones del sistema (3) y se acostumbra llamarla solución general.

Cualquier conjunto de soluciones $X^{(1)}, \dots, X^{(n)}$ de la ecuación (3) linealmente independiente en cada punto del intervalo $\alpha < t < \beta$, se dice que es un conjunto fundamental de soluciones para ese intervalo.

Para probar el teorema 14 demostraremos, dada cualquier solución Φ de la ecuación (3), que:

$$\Phi(t) + C_1 X^{(1)}(t) + \dots + C_n X^{(n)}(t)$$

para valores adecuados de C_1, \dots, C_n .

Sea $t = t_0$ algún punto en el intervalo $\alpha < t < \beta$ y $\xi = \Phi = \Phi(t_0)$.

Ahora deseamos determinar si existe solución de la forma:

$X = C_1 X^{(1)}(t) + \dots + C_n X^{(n)}(t)$ que también satisfaga la misma condición inicial $X(t_0) = \xi$. Estos es, deseamos saber si existen valores de C_1, \dots, C_n tales que

$$C_1 X^{(1)}(t_0) + \dots + C_n X^{(n)}(t_0) = \xi, \quad (12)$$

o en forma escalar,

$$\begin{aligned} C_1 X_{11}(t_0) + \dots + C_n X_{1n}(t_0) &= \xi_1, \\ &\vdots \\ C_1 X_{n1}(t_0) + \dots + C_n X_{nn}(t_0) &= \xi_n. \end{aligned} \quad (13)$$

La condición necesaria y suficiente para que las ecuaciones (13) posean una solución única C_1, \dots, C_n es precisamente que no se anule el determinante de los coeficientes, el cual es el wronskiano.

$W [X^{(1)}, \dots, X^{(n)}]$ evaluando en $t = t_0$. La hipótesis de que

$X^{(1)}, \dots, X^{(n)}$ son linealmente independientes en todo intervalo

$\alpha < t < \beta$ garantiza $W [X^{(1)}, \dots, X^{(n)}]$ no es cero en $t = t_0$, y, por lo tanto, existe una solución (única) de la ecuación (3) de la

forma $X = C_1 X^{(1)}(t) + \dots + C_n X^{(n)}(t)$ que también satisface la condición inicial (12). Por la parte de unicidad del teorema 12, esta solución es idéntica a $\Phi(t)$, y, por lo tanto,

$$\Phi(t) = C_1 X^{(1)}(t) + \dots + C_n X^{(n)}(t).$$

Como se quería probar.

Teorema 15.- Si $X^{(1)}, \dots, X^{(n)}$ son soluciones de la ecuación (3), sobre el intervalo $\alpha < t < \beta$, entonces, en este

intervalo, $W [X^{(1)}, \dots, X^{(n)}]$ es idénticamente cero o nunca se anula. El significado de este teorema se basa en el hecho de que

nos evita la necesidad de examinar $W [X^{(1)}, \dots, X^{(n)}]$ en todos los puntos en el intervalo de interés y nos permite determinar si:

$[X^{(1)}, \dots, X^{(n)}]$ forman un conjunto fundamental de soluciones, simplemente evaluando su Wronskiano en cualquier punto conveniente en el intervalo.

El teorema 15 se prueba estableciendo el Wronskiano de

$X^{(1)}, \dots, X^{(n)}$ satisface la ecuación.

$$\frac{dW}{dt} = (P_{11} + P_{22} + \dots + P_{nn})W \quad (14)$$

Por lo tanto, W es una función exponencial y se llega de inmediato a la conclusión del teorema. Alternativamente, puede establecerse el teorema (15) demostrando que si n soluciones

$X^{(1)}, \dots, X^{(n)}$ de la ecuación (3) son linealmente dependientes en un punto $t = t_0$, entonces deben ser linealmente dependientes en cada punto en $\alpha < t < \beta$.

Consecuentemente, $X^{(1)}, \dots, X^{(n)}$ son linealmente independientes en un punto, deben de ser linealmente independientes en cada punto en el intervalo.

El teorema 16.- Establece que el sistema (3) siempre tiene al menos un conjunto fundamental de soluciones.

Teorema 14. Sea.

$$e^{(1)} \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e^{(2)} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, e^{(n)} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{pmatrix};$$

además sean $X^{(1)}, \dots, X^{(n)}$ las soluciones del sistema (3) que satisfacen las condiciones iniciales.

$$X^{(i)}(t_0) = e^{(i)}, \dots, X^{(n)}(t_0) = e^{(n)}, \dots \quad (15)$$

respectivamente, donde t_0 es cualquier punto en $\alpha < t < \beta$. Entonces

$X^{(1)}, \dots, X^{(n)}$ forman un conjunto fundamental de soluciones del sistema (3). Para probar este teorema, observe que la existencia y

unicidad de las soluciones $X^{(1)}, \dots, X^{(n)}$ que se mencionan en el teorema (16) esta asegurada por el teorema (12). No es difícil ver que el Wronskiano de estas soluciones es igual a uno, cuando

$t = t_0$; por lo tanto, $X^{(1)}, \dots, X^{(n)}$ son un conjunto fundamental de soluciones.

EL PRINCIPIO DE SUPERPOSICION

[es la propiedad característica de los sistemas lineales].

Consideremos el sistema lineal.

$$(1) X'(t) = A(t) X(t) + f(t)$$

donde suponemos que la función matricial A y la función vectorial f son continuas sobre $-\infty, \infty$.

El principio de superposición es, como veremos una consecuencia directa de la linealidad de:

$$A(t)X'(t) = A(t)[C_1X^1(t) + C_2X^2(t)] = C_1A(t)X^1(t) + C_2A(t)X^2(t)$$

Teorema 17 (El principio de superposición).

Si Y^1 es una solución de $X'(t) = A(t)X(t) + f^1(t)$ y Y^2 es una solución de $X'(t) = A(t)X(t) + f^2(t)$ entonces $Y = C_1Y^1 + C_2Y^2$ es una solución de $X'(t) = A(t)X(t) + C_1f^1(t) + C_2f^2(t)$

Prueba

$$\begin{aligned} Y' &= C_1Y^{1'} + C_2Y^{2'} \stackrel{2}{=} C_1AY^1 + C_1f^1 + C_2AY^2 + C_2f^2 \\ &= A(C_1Y^1 + C_2Y^2) + C_1f^1 + C_2f^2 \\ &= AY + C_1f^1 + C_2f^2 \end{aligned}$$

En la hipótesis de que A y f sean continuas sobre $\langle -\infty, \infty \rangle$ o sobre cualquier intervalo J se sabe que (1) tiene solución única X sobre $\langle -\infty, \infty \rangle$ (o sobre J) que satisface $X(t_0) = X_0$ ($t_0 \in J$)

Hemos mostrado esto para coeficientes constantes y aceptamos el resultado más general sin prueba.

Supongamos que el sistema comienza en su posición de reposo $X(0) = 0$. Entonces la ecuación (1) tiene una solución única (y) que satisface esta condición inicial. Considerando a $A(t)$ como una función matricial fija asociamos entonces en cada función f continua sobre $\langle -\infty, \infty \rangle$ la solución (Y).

Denotamos esta correspondencia por $L = \{f, y\}$ y $Y = L(f)$. L es una transformación definida sobre el espacio de funciones f continuas sobre $\langle -\infty, \infty \rangle$.

La función f se llama a veces entrada y la solución (Y) salida.

El principio de superposición afirma que L es una transformación lineal:

$$\begin{aligned} L(C_1f^1 + C_2f^2) &= C_1L(f^1) + C_2L(f^2) \\ f &\rightarrow L \rightarrow Lf \\ f^1 + f^2 &\rightarrow L \rightarrow L(f^1) + L(f^2) \\ Cf &\rightarrow L \rightarrow CL(f) \end{aligned}$$

Es fácil ver esto. Las funciones $Y^1 = L(f^1)$ y $Y^2 = L(f^2)$ son soluciones de $X' = Ax + f^1$ y $X' = Ax + f^2$ que satisfacen $X(0) = 0$ por el principio de superposición $C_1Y^1 + C_2Y^2$ es una solución de $X = Ax + C_1f^1 + C_2f^2$ y satisface $X(0) = 0$.

Por lo tanto,

$$L(C_1f^1 + C_2f^2) = C_1Y^1 + C_2Y^2 + C_1L(f^1) + C_2L(f^2).$$

En términos de entradas y salidas el principio de superposición afirma que: si las entradas son añadidas, las salidas se añaden

$[L(f^1 + f^2) = L(f^1) + L(f^2)]$; si la entrada es multiplicada por una constante, la salida es multiplicada por la misma constante.

$$[L(cf) = cL(f)]$$

SISTEMAS LINEALES HOMOGÉNEOS CON COEFICIENTES CONSTANTES

Empezaremos a mostrar como elaborar la solución general de un sistema de ecuaciones lineales homogéneas con coeficientes constantes -esto es un sistema de la forma:

$$X' = AX \quad (1)$$

donde A es una matriz constante de $n \times n$.

Por analogía con el tratamiento de las ecuaciones lineales de segundo orden, buscaremos soluciones de la ecuación (1) de la forma:

$$X = \xi e^{rt} = A\xi e^{rt} \quad (2)$$

donde deben determinarse r y el vector constantes ξ , sustituyendo x en el sistema (1) por su expresión dada en la ecuación (2) da

$$r\xi e^{rt} = A\xi e^{rt}$$

Después de cancelar el factor escalar diferente de cero e^{rt} obtenemos $A\xi = r\xi$, o bien

$$(A - rI)\xi = 0 \quad (3)$$

donde I es la matriz identidad de $n \times n$.

Por tanto, para resolver el sistema de ecuaciones diferenciales (1), debemos resolver el sistema de ecuaciones algebraicas (3).

Este último problema es precisamente el que determina los valores característicos y vectores característicos de la matriz A.

Por lo tanto, el vector X dado por la ecuación (2) es una solución de la (1), siempre que r sea un eigenvalor y ξ un eigen vector asociado de la matriz de los coeficientes A.

En el sistema general (1) procedamos para encontrar soluciones de la ecuación diferencial, debemos encontrar los eigenvalores y eigenvectores de A, apartir del sistema algebraico asociado (3). Los eigenvalores r_1, \dots, r_n (que no necesitan ser diferentes) son raíces de la ecuación polinomial.

$$\det(A - rI) = 0 \quad (4)$$

La naturaleza de los eigenvalores y los eigenvectores correspondientes determinan la naturaleza de la solución del sistema (1).

SISTEMAS HERMITIANOS.

La situación es más sencilla cuando A es una matriz hermitiana.

"Una importante clase especial de matrices llamadas matrices

autoadjuntas o hermitianas, son aquellas para las cuales $A^* = A$;

esto es $\bar{A}_{ji} = A_{ij}$.

Las matrices hermitianas incluyen como subclase a las matrices simétricas reales; esto es, matrices que tienen elementos reales para las cuales $A^T = A$. Los eigenvalores y eigenvectores de las matrices hermitianas siempre tienen las útiles propiedades siguientes:

1.- Todos los eigenvalores son reales.

2.- Siempre existe un conjunto completo de n eigenvectores linealmente independientes, sin importar las multiplicidades de los eigenvalores.

3.- Si $X^{(1)}$ y $X^{(2)}$ son eigenvectores que corresponden a

eigenvalores diferentes, entonces $(X^{(1)}, X^{(2)}) = 0$. Por lo tanto, si todos los eigenvalores son simples, entonces los eigenvectores asociados forman un conjunto ortogonal de vectores.

4.- Correspondiendo a un eigenvalor de multiplicidad m , es posible elegir m eigenvectores que sean mutuamente ortogonales, como consecuencia, siempre puede elegirse el conjunto completo de n eigenvectores que sean ortogonales, así como linealmente independientes."

Los eigenvalores r_1, \dots, r_n son todos reales en este caso.

Además, incluso si alguno de los eigenvalores están repetidos, siempre existe un conjunto completo de n eigenvectores linealmente

independientes $\xi^{(1)}, \dots, \xi^{(n)}$. En consecuencia, las soluciones correspondientes del sistema diferencial (1) son

$$X^{(1)}(t) = \xi^{(1)} e^{r_1 t}, \dots, X^{(n)}(t) = \xi^{(n)} e^{r_n t} \quad (5)$$

Para poder demostrar que estas soluciones forman un conjunto fundamental, evaluamos su Wronskiano:

$$\begin{aligned}
 W[X^{(1)}, \dots, X^{(n)}] &= \begin{vmatrix} \xi_1^{(1)} e^{r_1 t} & \dots & \xi_1^{(n)} e^{r_n t} \\ \vdots & & \vdots \\ \xi_n^{(1)} e^{r_1 t} & \dots & \xi_n^{(n)} e^{r_n t} \end{vmatrix} = \\
 &= e^{(r_1 + \dots + r_n)t} \begin{vmatrix} \xi_1^{(1)} & \dots & \xi_1^{(n)} \\ \vdots & & \vdots \\ \xi_n^{(1)} & \dots & \xi_n^{(n)} \end{vmatrix} \quad (5')
 \end{aligned}$$

Primero observamos que la función exponencial nunca es cero.

A continuación, como los eigenvectores $\xi^{(1)}, \dots, \xi^{(n)}$ son linealmente independientes, el determinante en el último término de la ecuación (5) es diferente de cero. Como consecuencia, el

Wronskiano de $\{X^{(1)}, \dots, X^{(n)}\}$ forman un conjunto fundamental de soluciones.

Por tanto, cuando A es una matriz hermitiana, la solución general de la ecuación (1) es

$$X = C_1 \xi^{(1)} e^{r_1 t} + \dots + C_n \xi^{(n)} e^{r_n t} \quad (6)$$

Una subclase importante de las matrices hermitianas es la clase de las matrices reales simétricas. Si A es real y simétrica entonces los eigenvectores $\xi^{(1)}, \dots, \xi^{(n)}$ así como los eigenvalores r_1, \dots, r_n son todos reales.

De aquí que todas las soluciones dadas por la ecuación (4) son de valores reales. Sin embargo, si la matriz hermitiana A no es real, entonces, en general, los eigenvectores tienen partes imaginarias diferentes de cero y las soluciones de (4) son de valores complejos.

Sistemas no hermitianos. Si la matriz de coeficientes A en el sistema (1).

$$X' = AX$$

no es hermitiana, entonces la situación referente a la solución es más complicada.

Supongamos primero que A es real.

Entonces se tienen tres posibilidades para los eigenvalores de A :

- 1.- Todos los eigenvalores son reales y distintos.
- 2.- Algunos eigenvalores ocurren en parejas de complejos conjugados.
- 3.- Algunos eigenvalores están repetidos.

El primer caso no conduce a dificultad alguna.

Si tiene un solo eigenvector real linealmente independiente por cada eigenvalor y como consecuencia, existen n soluciones linealmente independientes de la forma (4).

Por consiguiente, la solución general aún está dada por la ecuación (6), donde ahora se sobre entiende que r_1, \dots, r_n son todos diferentes.

Si alguno de los eigenvalores ocurren en parejas de complejos conjugados, entonces aún se tienen n soluciones linealmente independientes, de la forma (4), siempre que todos los eigenvalores sean diferentes. Por supuesto, las soluciones que provienen de eigenvalores complejos son de valores complejos. Se presentan dificultades más serias si un eigenvalor está repetido.

En este evento el número de eigenvectores correspondientes, linealmente independientes, puede ser menor que la multiplicidad del eigenvalor. si es así, el número de soluciones linealmente

Independientes de la forma ξe^{rt} será menor que n . Entonces, para construir un conjunto fundamental de soluciones, es necesario buscar soluciones adicionales de otra forma.

La situación es un tanto análoga a la de una lineal de n -ésimo orden con coeficientes constantes; una raíz repetida de la ecuación característica dio lugar a soluciones de la forma.

$$e^{rx}, xe^{rx}, x^2 e^{rx}, \dots$$

Finalmente, si A es compleja, pero no hermitiana, entonces los eigenvalores complejos no necesariamente ocurren en parejas de conjugadas y, normalmente, los eigenvectores son de valores complejos, aún cuando el eigenvalor asociado sea real.

Las soluciones de la ecuación diferencial (1) aún son de la forma (4), siempre que los eigenvalores son distintos pero en general, todas las soluciones son de valores complejos.

EIGENVALORES COMPLEJOS

Consideremos nuevamente un sistema de n ecuaciones lineales homogéneas, con coeficientes constantes,

$$X' = AX \quad (1)$$

donde ahora se supone que la matriz de los coeficientes A es de valores reales.

Si buscamos soluciones de la forma $X = \xi e^{rt}$, entonces, como r es un eigenvalor y ξ un eigenvector correspondiente de la matriz de coeficientes de A , ya que los eigenvalores r_1, \dots, r_n de A son las raíces de la ecuación.

$$\det (A - rI) = 0 \quad (2)$$

y que los eigenvectores correspondientes satisfacen

$$(A - rI)\xi = 0$$

Si A es real, entonces los coeficientes en la ecuación polinomial (2) para r son reales y cualesquiera eigenvalores complejos deben ocurrir en los pares conjugados. Por ejemplo, si $r_1 = \lambda + i\mu$, donde λ y μ son reales, es un eigenvalor de A , entonces lo es $r_2 = \lambda - i\mu$.

Además, los eigenvectores correspondientes $\xi^{(1)}$ y $\xi^{(2)}$ también son conjugados complejos.

Para verlo, suponga que r_1 y $\xi^{(1)}$ satisfacen

$$(A - r_1 I) \xi^{(1)} = 0 \quad (4)$$

Tomando el conjugado complejo de esta ecuación y observando que A e I son de valores reales, obtenemos:

$$(A - \bar{r}_1 I) \xi^{(1)} = 0 \quad (5)$$

donde \bar{r}_1 y $\xi^{(1)}$ son los conjugados complejos de r_1 y $\xi^{(1)}$ respectivamente.

En otras palabras, $r_2 = \bar{r}_1$ también es un eigenvalor y $\xi^{(2)} = \bar{\xi}^{(1)}$ es el eigenvector correspondiente. Entonces, las soluciones correspondientes.

$$X^{(1)}(t) = \xi^{(1)} e^{r_1 t}; X^{(2)}(t) = \bar{\xi}^{(1)} e^{\bar{r}_1 t} \quad (6)$$

de la ecuación diferencial (1) son conjugadas complejas entre sí.

Por lo tanto, podemos encontrar dos soluciones de valores reales de la ecuación (1), correspondientes a los eigenvectores r_1

y r_2 tomando las partes reales o imaginarias de $X^{(1)}(t)$ o $X^{(2)}(t)$ dadas por las (6).

Escribamos $\xi^{(1)} = a + ib$, donde a y b son reales; entonces tenemos.

$$\begin{aligned} X^{(1)}(t) &= (a + ib) e^{(\lambda + i\omega t)} \\ &= (a + ib) e^{\lambda t} (\cos \omega t + i \sin \omega t) \end{aligned} \quad (7)$$

Después de separar a $X^{(1)}(t)$ en sus partes real e imaginaria, obtenemos.

$$\begin{aligned} X^{(1)}(t) &= e^{\lambda t} (a \cos \omega t - b \sin \omega t) + \\ &+ i e^{\lambda t} (a \sin \omega t + b \cos \omega t) \end{aligned} \quad (8)$$

Si escribimos $X^{(1)}(t) = u(t) + iv(t)$, entonces los vectores.

$$u(t) = e^{\lambda t} (a \cos \omega t - b \sin \omega t) \quad (9a)$$

$$v(t) = e^{\lambda t} (a \sin \omega t + b \cos \omega t) \quad (9b)$$

son soluciones de valores reales de la ecuación (1) es posible demostrar que u y v son soluciones linealmente independientes.

Por ejemplo, suponga que $r_1 = \lambda + i\mu$, $r_2 = \lambda - i\mu$ y que r_3, \dots, r_n son todas reales y distintas.

Sean los eigenvalores correspondientes

$$\xi^{(1)} = a + ib, \xi^{(2)} = a - ib, \xi^{(3)}, \dots, \xi^{(n)}$$

Entonces la solución general de la ecuación (1) es

$$X = C_1 U(t) + C_2 V(t) + C_3 \xi^{(3)} e^{r_3 t} + \dots + C_n \xi^{(n)} e^{r_n t} \quad (10)$$

donde $U(t)$ y $V(t)$ están dadas por las ecuaciones (9). Conviene hacer énfasis en que este análisis sólo se aplica si la matriz de los coeficientes A , en la ecuación (1), es real, porque sólo entonces los eigenvalores y eigenvectores complejos ocurren en pares conjugados.

EIGENVALORES REPETIDOS

Concluirémos la consideración del sistema lineal homogéneo con coeficiente constante.

$$X' = AX \quad (1)$$

con un estudio del caso en el que la matriz A tiene un eigenvalor repetido.

La discusión en esta sección es válida para A real o compleja.

Si $r = p$ es una raíz de multiplicidad R de la ecuación:

$$\det(A - rI) = 0 \quad (2)$$

se dice que entonces p es un eigenvalor de multiplicidad R de la matriz A .

En este evento, existen dos posibilidades: se tienen R eigenvectores linealmente independientes que corresponden al eigenvalor p , o bien, hay menos de R tales eigenvectores.

En el primer caso, sean $\xi^{(1)}, \dots, \xi^{(R)}$ R eigenvectores linealmente independientes, asociados con el eigenvalor p de multiplicidad R , entonces

$$X^{(1)}(t) = \xi^{(1)} e^{pt}, \dots, X^{(R)}(t) = \xi^{(R)} e^{pt}$$

son R soluciones linealmente independientes de la ecuación (1).

Por tanto, en este caso, no existe diferencia en que se repita el eigenvalor $r = p$ aún se tiene un conjunto fundamental de soluciones de la ecuación (1), de la forma ξe^{rt} .

Siempre ocurre este caso si la matriz de los coeficientes A es hermitiano.

Sin embargo si la matriz de los coeficientes no es hermitiana, entonces puede haber menos de R eigenvalores independientes, que correspondan aun eigenvalor de multiplicidad R y, si es así, habrá

menos de R soluciones de la ecuación (1), de la forma ξe^{rt} asociadas con este eigenvalor. Por lo tanto, para construir la solución general de la (1) es necesario encontrar otras soluciones de alguna forma diferente. Por analogía con los resultados previos para las ecuaciones lineales de orden n , resulta natural buscar soluciones adicionales que comprendan productos de polinomios y funciones exponenciales.

MATRICES FUNDAMENTALES

La teoría de los sistemas de ecuaciones diferenciales lineales puede aclararse aún más, introduciendo la idea de matriz fundamental. Este concepto nos será particularmente útil en la acción que sigue, donde extenderemos el método de variación de parámetros hacia los sistemas de ecuaciones lineales de primer

orden. Suponga que $X^{(1)}, \dots, X^{(n)}$ forman un conjunto fundamental de soluciones para la ecuación.

$$X' = P(t)X \quad (1)$$

sobre algún intervalo $\alpha < t < \beta$, entonces se dice que la matriz

$$\Psi(t) = \begin{pmatrix} X^{(1)}(t) & \dots & X^{(n)}(t) \\ \vdots & & \vdots \\ X_n^{(1)} & \dots & X_n^{(n)}(t) \end{pmatrix} \quad (2)$$

cuyas columnas son los vectores $X^{(1)}, \dots, X^{(n)}$, es una matriz fundamental para el sistema (1) observe que cualquier matriz fundamental es no singular, ya que las columnas son vectores linealmente independientes.

La solución de un problema con valores iniciales puede abreviarse en términos de una matriz fundamental. La solución general de la ecuación es

$$X = C_1 X^{(1)}(t) + \dots + C_n X^{(n)}(t) \quad (3)$$

$$\text{o, en términos de } \Psi(t); X = \Psi(t)C \quad (4)$$

donde C es un vector constante con componentes arbitrarias C_1, \dots, C_n . Para un problema con valores iniciales que consiste de la ecuación diferencial (1) y la condición inicial.

$$X(t_0) = X^0 \quad (5)$$

donde t_0 es un punto dado en el intervalo $\alpha < t < \beta$ y X^0 es un vector inicial, dado; sólo es necesario seleccionar el vector C en la ecuación (4) de modo que satisfaga la condición inicial (5).

De ahí que C debe satisfacer.

$$\Psi(t_0)C = X^0 \quad (6)$$

Por lo tanto, ya que $\Psi(t)$ es no singular,

$$C = \Psi^{-1}(t_0)X^0 \quad (7)$$

y

$$X = \Psi(t)\Psi^{-1}(t_0)X^0 \quad (8)$$

es la solución del problema con valores iniciales, dado.

Sin embargo haremos énfasis en que para resolver un problema con valores iniciales dado, por el común se resolvería la ecuación (4) por reducción de renglones y, a continuación se sustituiría C en la ecuación (4) en lugar de calcular $\Psi^{-1}(t_0)$ y usar la ecuación (8).

A veces es conveniente usar la matriz fundamental especial,

denotada por $\Phi(t)$, cuyas columnas son los vectores $X^{(1)}, \dots, X^{(n)}$ que se como el conjunto fundamental de soluciones del sistema homogéneo.

Además de la ecuación diferencial (1), estos vectores, satisfacen las condiciones iniciales $X^{(j)}(t_0) = e^{(j)}$ (9)

donde $e^{(j)}$ es el vector unitario, con un uno en la j -ésima posición y ceros en las demás. Por tanto $\Phi(t)$ tiene la propiedad de que

$$\Phi(t_0) = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & & 1 \end{pmatrix} = I \quad (10)$$

siempre reservamos el símbolo Φ para denotar la matriz fundamental que satisface la condición inicial (10) y usaremos $\Psi(t)$ cuando se pliese en una matriz fundamental arbitraria. En términos de $\Phi(t)$ la solución del problema con valores iniciales (1) y (5) es aún más

sencilla aparentemente, ya que $\Phi^{-1}(t_0) = I$ se deduce de la ecuación

$$(8) \text{ que } X = \phi(t)X^0 \quad (11)$$

Aún cuando la matriz fundamental $\Phi(t)$ a menudo es más complicada que $\Psi(t)$, es especialmente útil si el mismo sistema de ecuaciones diferenciales debe resolverse repetidas veces, sujeto a muchas condiciones iniciales diferentes.

Esto corresponde a un sistema físico dado que puede arrancarse desde muchos estados iniciales diferentes. Si se ha determinado la matriz fundamental $\Phi(t)$ entonces puede encontrarse la solución para cada conjunto de condiciones iniciales, simplemente por una multiplicación de matrices, como lo indica la ecuación (11).

En consecuencia, la matriz $\Phi(t)$ representa una transformación de las condiciones iniciales X^0 hacia la solución $X(t)$, en un instante arbitrario t . Comparando las ecuaciones.

$$X = \Psi(t)\Psi^{-1}(t_0)X^0 \text{ y } X = \Phi(t)X^0$$

es obvio que $\Phi(t) = \Psi(t)\Psi^{-1}(t_0)$.

Ahora evaluamos nuevamente al sistema.

$$X' = AX$$

donde A es una matriz constante dada, no necesariamente real. Ya que describieron dos métodos para resolver tales sistemas: por el de

eliminación, y por la suposición de $X = \xi e^{rt}$. Aquí se proporcionará aún otro punto de vista. La razón básica por la que un sistema de ecuaciones presenta cierta dificultad es que las ecuaciones comunmente están ligadas, en otras palabras, algunas de ellas, o todas contienen a más de una, tal vez a todas, de las variables dependientes. Esto ocurre siempre que la matriz de los coeficientes A no sea una matriz diagonal; de donde, las ecuaciones en el sistema deben resolverse simultáneamente, en lugar de consecutivamente. Esta observación sugiere que, una manera de resolver un sistema de ecuaciones podría ser transformándolo a un sistema equivalente no ligado, en el cual cada ecuación sólo contenga una variable dependiente. Esto corresponde a transformar la matriz de los coeficientes A en una matriz diagonal.

De acuerdo con los resultados obtenidos anteriormente es posible hacer esto siempre que A tenga un conjunto completo de n eigenvectores linealmente independientes.

Sean $\xi^{(1)}, \dots, \xi^{(n)}$ eigenvectores de A que corresponden a los eigenvalores r_1, \dots, r_n , y formemos la matriz de la transformación T , cuyas columnas son:

$\xi^{(1)}, \dots, \xi^{(n)}$. Entonces

$$T = \begin{pmatrix} \xi^{(1)} & \dots & \xi^{(n)} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \xi^{(1)} & \dots & \xi^{(n)} \end{pmatrix} \quad (12)$$

Definiendo una nueva variable dependiente y por la relación.

$$X = Ty,$$

Tenemos de la ecuación ($X' = AX$), que

$$T y' = ATy$$

o bien

$$y' = (T^{-1}AT)y$$

de la ecuación $C_1 X^{(1)} + \dots + C_k X^{(k)} = 0$

conjunto de R vectores linealmente dependientes o linealmente independientes.

$$T^{-1}AT = D = \begin{pmatrix} r_1 & 0 & \dots & 0 \\ 0 & r_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & r_n \end{pmatrix}$$

es la matriz diagonal cuyos elementos diagonales son los eigenvalores de A . Esto es, y satisface el sistema de ecuaciones no ligado.

$$Y' = DY$$

para el que una matriz fundamental es la matriz diagonal.

$$Q(t) = \begin{pmatrix} e^{r_1 t} & 0 & \dots & 0 \\ 0 & e^{r_2 t} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{r_n t} \end{pmatrix}$$

Entonces se encuentra una matriz fundamental Ψ para el sistema

$X' = AX$ a partir de Q , por la transformación $X = TY$.

$$\Psi = TQ$$

esto es

$$\Psi = \begin{pmatrix} \xi_1^{(1)} e^{r_1 t} & \dots & \xi_1^{(n)} e^{r_n t} \\ \vdots & & \vdots \\ \xi_1^{(1)} e^{r_1 t} & \dots & \xi_1^{(n)} e^{r_n t} \end{pmatrix}$$

Esta ecuación de $\Psi(t)$ es el mismo resultado que se obtuvo anteriormente.

Este procedimiento de diagonalización no proporciona ventaja de cálculo alguno sobre el método, puesto que en cualquiera de los dos casos es necesario calcular los eigenvalores y eigenvectores de la matriz de los coeficientes en el sistema de ecuaciones diferenciales. Sin embargo, es notable que el problema de resolver un sistema de ecuaciones diferenciales y el diagonalizar una matriz sean matemáticamente el mismo.

Sistemas Lineales Homogeneos con Coeficientes Constantes
un sistema de la forma

$$X' = AX$$

cuyas soluciones son la de la forma

$$X = \xi e^{rt}$$

donde deben determinar r y el vector constante ξ para esto se debe resolver el sistema de ecuaciones algebraicas dada por

$$(A - rI)\xi = 0$$

siempre que sea un eigenvalor y ξ un eigenvector asociado de la matriz de coeficientes constantes A .

$$\det(A - rI) = 0$$

Los eigenvalores r_1, \dots, r_n son raíces de la ecuación polinomial.

Los cuales determinan la naturaleza de la solución del sistema.

$$X' = AX$$

SISTEMAS LINEALES NO HOMOGENEOS

Principiemos con sistemas de la forma

$$X' = Ax + g(t) \quad (1)$$

donde A es una matriz constante diagonalizable de $n \times n$ y $g(t)$ es continua para $\alpha < t < \beta$.

Procederemos por el método indicado al final. Sea t la matriz cuyas columnas sean los eigenvectores $\xi^{(1)}, \dots, \xi^{(n)}$ de A y definamos una nueva variable dependiente Y por

$$X = TY \quad (2)$$

Entonces, sustituyendo X en la ecuación (1) obtenemos

$$TY' = ATY + g(t)$$

Multiplicado por T^{-1} se deduce que

$$\begin{aligned} Y' &= (T^{-1}AT)Y + T^{-1}g(t) \\ &= DY + h(t) \end{aligned} \quad (3)$$

donde $h(t) = T^{-1}g(t)$ y D es la matriz diagonal cuyos elementos en la diagonal son los eigenvalores r_1, \dots, r_n de A, arreglados en el mismo orden en el que los eigenvectores correspondientes

$\xi^{(1)}, \dots, \xi^{(n)}$ aparecen como columnas de T.

La ecuación (3) es un sistema de n ecuaciones no ligadas para $Y_1(t), \dots, Y_n(t)$; como consecuencia, las ecuaciones pueden resolverse por separado.

En la forma escalar, la ecuación (3) queda

$$Y'_j(t) = r_j Y_j(t) + h_j(t) \quad j = 1, \dots, n \quad (4)$$

donde $h_j(t)$ es una cierta combinación lineal de $g_1(t), \dots, g_n(t)$,

la ecuación (4) es una ecuación lineal de primer orden y puede resolverse por los métodos de integración.

En efecto, tenemos:

$$Y_j(t) = e^{r_j t} \int_{t_0}^t e^{-r_j s} h_j(s) ds + C_j \cdot e^{r_j t} \quad j = 1, \dots, n, \quad (5)$$

donde las C_j son constantes arbitrarias por último, la solución X de la ecuación (1) se obtiene de la ecuación (2).

Consideremos ahora los problemas más generales, en los que la matriz de los coeficientes es no constante o no diagonalizable. Sea

$$X' = P(t)X + g(t) \quad (6)$$

donde $P(t)$ y $g(t)$ son continuas sobre $\alpha < t < \beta$, suponga que sea encontrado una matriz fundamental $\psi(t)$ para el sistema homogéneo correspondientes

$$X' = P(t)X \quad (7)$$

Usaremos el método de variación de parámetros para construir una solución particular Y , de aquí la solución general del sistema no homogéneo (6).

Como la solución del sistema homogéneo (7) es $\Psi(t)c$, resulta natural proceder con el método de variación de parámetros y buscar una solución del sistema no homogéneo (6) reemplazando el vector constante c por una función vectorial $u(t)$. Por lo tanto supongamos que

$$X = \Psi(t)u(t) \quad (8)$$

donde $u(t)$ es un vector que debe encontrarse. Después de derivar X , como se da en la ecuación (8), y requiriendo que se satisfaga la ecuación (6) obtenemos

$$\Psi'(t)u(t) + \Psi(t)u'(t) = P(t)\Psi(t)u(t) + g(t) \quad (9)$$

supuesto que, $\Psi(t)$ es una matriz fundamental.

$\Psi' = P(t)\Psi(t)$, y por lo tanto la ecuación (9) se reduce a

$$\Psi(t)u'(t) = g(t) \quad (10)$$

Recordemos que $\Psi(t)$ es no singular sobre cualquier intervalo en donde P sea continua.

De donde $\Psi^{-1}(t)$ existe y, por lo tanto

$$u'(t) = \Psi^{-1}(t)g(t) \quad (11)$$

Así que para $u(t)$ podemos seleccionar cualquier vector de la clase de vectores que satisfagan la ecuación (11); estos vectores están determinados solo hasta una constante arbitraria aditiva (vector); por lo tanto, denotamos

$$u(t) = \int^t \Psi^{-1}(s)g(s)ds + c \quad (12)$$

donde el vector constante c es arbitrario.

Finalmente sustituyendo $u(t)$ en la ecuación (8), da la solución X del sistema (6).

$$X = \Psi(t)c + \Psi(t)\int^t \Psi^{-1}(s)g(s)ds \quad (13)$$

Ya que c es arbitraria, cualquier condición inicial en un punto $t = t_0$ puede satisfacerse por una elección apropiada de c . Así toda solución del sistema (6) está contenida en la expresión dada por la ecuación (13) por lo tanto, es la solución general de la

ecuación (6). Note que el primer término del segundo miembro de la ecuación (13) es la solución general del sistema homogéneo correspondiente (7) y el segundo término es una solución particular del propio sistema no homogéneo (6) consideremos ahora el problema con valores iniciales que consiste de la ecuación diferencial (6) y la condición inicial:

$$X(t_0) = X^0 \quad (14)$$

Podemos escribir de modo más conveniente la solución de este problema, si seleccionamos para la solución particular en la ecuación (13) la específica que es cero cuando $t = t_0$. Este puede hacerse usando $t = t_0$ como el límite inferior de la integración en la ecuación (13), de manera que la solución general de la ecuación diferencial toma la forma:

$$X = \Psi(t)c + \Psi(t) \int_{t_0}^t \Psi^{-1}(s)g(s)ds \quad (15)$$

la condición inicial (14) también puede satisfacerse siempre que

$$c = \Psi^{-1}(t_0)X^0 \quad (16)$$

Por lo tanto

$$X = \Psi(t)\Psi^{-1}(t_0)X^0 + \Psi(t) \int_{t_0}^t \Psi^{-1}(s)g(s)ds \quad (17)$$

es la solución del problema con valor inicial dado. Una vez más, aún cuando es útil usar Ψ^{-1} para escribir las soluciones (13) y (17), por lo general es mejor en los casos particulares, resolver las ecuaciones necesarias por reducción de los renglones en lugar de calcular Ψ^{-1} y sustituir en las ecuaciones (13) y (17).

La solución (14) toma una forma ligeramente más sencilla, si usamos la matriz fundamental $\Phi(t)$ que satisface $\Phi(t_0) = I$.

En este caso, tenemos

$$X = \Phi(t)X^0 + \Phi(t) \int_{t_0}^t \Phi^{-1}(s)g(s)ds \quad (18)$$

La ecuación (18) puede simplificarse aún más si la matriz de los coeficientes $P(t)$ es una matriz constante.

En algunos casos, una manera más conveniente de encontrar una solución particular del sistema no homogéneo (6) es el método de los coeficientes indeterminados. Para usar este método, se supone la forma de la solución, con alguno o todos los coeficientes no especificados, y entonces se intenta determinar estos coeficientes, de tal manera, que satisfagan la ecuación diferencial.

Como algo práctico, este método solo es aplicable si la matriz de los coeficientes P es una matriz constante y si las componentes de g son funciones polinomiales exponenciales o senoidales, o sumas de productos de estas.

Solo en estos casos, es cuando se puede pronosticar la forma correcta de la solución de una forma sencilla y sistemático. El procedimiento para seleccionar la forma de la solución es sustancialmente el mismo que el problema de las ecuaciones no homogéneas para las ecuaciones lineales de segundo orden. La diferencia principal queda ilustrado por el caso de un término no

homogéneo de la forma $Ve^{\lambda t}$, de donde λ es una raíz simple de la ecuación característica. En esta situación, en lugar de suponer una

solución de la forma $ate^{\lambda t}$, es necesario usar $ate^{\lambda t} + be^{\lambda t}$, donde a y b se determinan sustituyendo en la ecuación diferencial.

CAPITULO 3.0

Para que el concepto de modelo sea relevante, es preciso que se considere en el contexto de la forma de pensar científica sobre los acontecimientos del « mundo real », los métodos de representar, medir y analizar estos acontecimientos, y la validez de los resultados derivados del análisis.

¿Qué es un modelo?

En el sentido más general, un « modelo » es cualquier cosa empleada para representar otra cosa -un mapa utilizado para representar una sección de la corteza terrestre es un modelo, y también lo es un diagrama utilizado para representar los cambios concurrentes en las ventas de combustible y la temperatura. En este sentido tan general, todas las percepciones humanas de la realidad y los modos de pensar sobre ellas son « modelos ».

Considérese por ejemplo la noción de contar y el sistema de números basados en este concepto. Los orígenes de contar se pierden en la antigüedad, pero al parecer se necesitaron muchos miles de años para que apareciera esta idea y llegara a desarrollarse plenamente en su forma actual.

El sistema numérico es un modelo mediante el cual se pueden representar los objetos físicos (hombres, ganado, etc.). Se basa en el reconocimiento de que exista algo en común entre un grupo de 3 hombres y un grupo de 3 vacas. Hablamos así de la propiedad común referente al número de objetos de cada grupo. Por supuesto que existen muchas diferencias entre 3 hombres y 3 vacas; al asignar el número 3 a ambos estamos abstrayendo una propiedad común entre las muchas que pueden tomarse en consideración.

Uno de los propósitos de realizar esta abstracción es para poder hacer uso de las operaciones de la aritmética. Así sabemos que $3 + 2 = 5$, tanto si los objetos representados son hombres, vacas o cualquier otra cosa.

Supóngase que dos grupos de individuos quieren ir juntos de excursión ¿cabrán en un automóvil?. Una de las formas de resolver el problema consiste en intentar si dichos individuos pueden entrar todos en el vehículo. Pero puede resultar mucho más fácil representar a cada grupo por un número por medio de la operación de contar; pongamos por ejemplo que hay 3 personas en uno de los grupos y 2 en el otro. Para determinar el tamaño de los dos grupos combinados efectuamos el familiar proceso de la adición, $3 + 2 = 5$.

Al igual que existe una correspondencia entre cada grupo y el número empleado para representarlo, también existe una correspondencia entre la operación abstracta de la adición y la operación física de combinar los dos grupos. Basándonos en esta correspondencia, sabemos que los resultados de efectuar las operaciones en el modelo corresponderán a los resultados de combinar los dos grupos. Basándonos en esta correspondencia, sabemos que los resultados de efectuar las operaciones en el modelo corresponderán a los resultados de la combinación física, es decir, el tamaño del grupo combinado será igual a 5.

Si ahora disponemos de una información adicional -la capacidad del automóvil en términos del número de pasajeros-, entonces el análisis del modelo nos permite resolver el problema y adoptar cualquier acción que se considere adecuada.

El ejemplo que hemos utilizado aquí resulta tan familiar que se acerque a la trivialidad, pero de todas formas nos sirven para sacar a la superficie la mayor parte de las cuestiones claves que se plantean en la construcción y el empleo de modelos por complejos que puedan ser.

Una cuestión fundamental la constituye, por supuesto, la naturaleza de la « correspondencia » entre el modelo y los acontecimientos y objetos reales que se supone que el modelo representa.

Nadie duda de la validez del sistema numérico como modelo para representar una amplia variedad de cosas; es evidente que serían pocos los que pondrían en duda la validez del resultado obtenido en el ejemplo anterior.

Los modelos se construyen en vistas a toda una serie de propósitos; estos comprenden generalmente la comprensión, la predicción y el control de los acontecimientos del mundo real.

Finalmente puede observarse que el valor del modelo deriva de su eficiencia por llegar a un resultado. Los argumentos reales para utilizar modelos matemáticos son:

1.- La determinación de resultados es más rápida, menos cara, o más precisa con la ayuda de un modelo;

2.- Esta operación no es posible efectuarla directamente en los acontecimientos u objetos físicos.

Clases de modelos:

Puesto que el concepto de modelos resulta tan amplio, es de esperar que existan muchas clases de modelos. Resulta útil clasificar los modelos de varias maneras.

Modelos implícitos y explícitos.

Modelos matemáticos.

Modelos descriptivos, predictivos y de control.

VARIABLES Y FUNCIONES COMO INSTRUMENTOS DE LAS CIENCIAS APLICADAS:

SISTEMAS Y MODELOS

El conocimiento humano comenzó a fraguarse como conocimiento científico en el momento en que empezaron a utilizarse estructuras lógicas y matemáticas -modelos- para explicar los fenómenos que acontecen en la realidad. Esto no tuvo lugar sino hasta bien entrada la Edad Moderna, siendo el pionero de esta práctica el matemático, físico y filósofo Galileo.

El punto de vista actual sobre los fenómenos consiste en considerarlos como acontecimientos que suceden en el seno de algún sistema. De esta manera los sistemas se han convertido en los objetos de estudio preferente de las ciencias aplicadas.

Abstraído de sus características peculiares, un sistema no es más que un agregado de componentes de uno o varios tipos, particularmente relacionados, capaces de interactuar entre sí (a través de campos, estímulos, etc.) y de provocar reacciones de respuesta de acuerdo con la naturaleza y el estado de los elementos reaccionantes.

Enfoque generalista versus enfoque analítico.

Existen dos maneras de afrontar el estudio de los sistemas. Según la primera (enfoque generalista o macro), los sistemas deben contemplarse en su globalidad, prescindiendo del análisis de sus componentes. Según la segunda (enfoque analítico o micro), los sistemas deben estudiarse descendiendo al detalle de sus componentes.

De este modo, mientras los generalistas dirigen su atención hacia características globales de los sistemas (los físicos, se fijan en el color, la temperatura, etc; los economistas, en el consumo público, el acervo de capital, etc.), los analistas se

Interesan antes que nada por las características de sus componentes (los físicos se interesan por el spin, la carga, el impulso, etc; los economistas, por la renta individual, el sistema de precios de los bienes, las preferencias de los individuos, etc.).

Pero el objetivo que persiguen unos y otros es idéntico: elegir las características patentemente importantes para la comprensión del fenómeno que se trata de explicar y estructurarlas lógicamente o matemáticamente. Los generalistas elaboran así lo que se conoce como *macro-modelos*, y los analistas, *micro-modelos*.

Entre todos los modelos -tanto si son macro como si son micro -destacan los que se articulan exclusivamente mediante características que pueden cuantificarse utilizando criterios intersubjetivos (masa, temperatura, acervo de capital, inversión pública, son características de este tipo: no lo son el grado de democracia o el nivel de cultura de una sociedad). De este modo, la actividad científica traduce las características cuantificables de los sistemas en sencillas variables matemáticas (la masa, por ejemplo, se traduce en una variable sobre $[0, +\infty] \subset \mathbb{R}$ y el sistema de precios de una economía de n bienes en una variable sobre $[0, +\infty] \times \dots \times [0, +\infty] \subset \mathbb{R}^n$.

Macrovariables y microvariables.

Estas variables, según estén asociadas a características globales del sistema o a características de sus componentes, reciben el nombre de *macrovariables* o de *microvariables* del sistema.

Así, pues las macro y microvariables de un sistema no son más que la traducción en términos matemáticos de las características cuantificables del sistema considerado como un todo, o de sus componentes.

Macroestados: factibles y no factibles.

Si x_1, x_2, \dots, x_n son n macrovariables que describan en algún modelo la situación (global) de un sistema, se llama (macro) estado del sistema a una n -epia de valores $(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$ uno por uno por cada una de las macrovariables. Existen estados que no corresponden a ninguna situación real posible del sistema; recibe el nombre de *estados no factibles*. Por el contrario, los estados

que se corresponden por situaciones reales posibles reciben el nombre de *estados factibles*. Los estados factibles de un sistema no

son más que puntos de R^n que pertenecen a un subconjunto determinado de este espacio.

Microestados.

Análogamente se definen las (*micro*) estados de un sistema.

Supuesto los componentes del sistema ordenados y asociados a los números naturales 1, 2, ..., N, sean $x_{11}, x_{12}, \dots, x_{1n_1}$ las n_1 microvariables de cada uno de los componentes. De uno por cada una de las microvariables de cada uno de los componentes. De forma obvia se definen también en este caso los *estados factibles* y *no factibles*, y se puede repetir la observación anterior de que los estados factibles de un sistema son puntos de un subconjunto de un espacio euclídeo.

Modelos cuantitativos.

Se llaman *modelos cuantitativos* (deterministas) a aquellos que contienen macro o microvariables - o una u otra cosa, se entiende - *engarzadas por relaciones de tipo matemático*. La forma más frecuente como estas variables se relacionan es por medio de funciones, y sólo en casos excepcionales lo hacen mediante correspondencias que no son funciones.

Leyes.

No todas estas relaciones matemáticas gozan de la misma generalizada.

Las hay que son prácticamente exclusivas de un modelo particular, por expresar circunstancias peculiares del correspondiente sistema, y las hay que se presentan con cierta generalidad en los modelos que caen dentro de una misma teoría.

Estas últimas reciben el nombre genérico de *leyes*.

Tampoco las leyes son igualmente generales. Algunas pueden utilizarse prácticamente sin restricciones, mientras que otras pueden aplicarse sólo después de haberse comprobado que el sistema cumple determinadas condiciones: son leyes más restringidas.

Características de las leyes aplicables a los modelos cuantitativos.

Las leyes limitan siempre « salidas » posibles para los fenómenos: [las formas como los fenómenos pueden suceder] y se expresan a través de *relaciones funcionales* entre las variables de estado que se presentan en macro y micromodelos correspondientes a un hipotético sistema.

Estados de equilibrio, modelos estáticos.

Entre los posibles estados de un sistema los científicos manifiestan especial interés por los estados de equilibrio. Un sistema se dice que está en un estado de equilibrio si, manteniendo constantes en el tiempo sus intercambios materiales, energéticos e informacionales con su entorno, no cambia de estado. Los científicos construyen modelos estáticos para caracterizar los estados de equilibrio de los sistemas. En los *modelos estáticos* las variables de estado son independientes del tiempo y las leyes que en ellas se pueden utilizar reciben el calificativo de estáticas.

Las leyes estáticas de una teoría son, pues, relaciones matemáticas muy generales, aptas para estructurar las variables de los modelos construidos para dar cuenta de los equilibrios de una clase de sistemas (de los que la citada teoría se ocupa).

Manera general de construir modelos cuantitativos estáticos.

La forma típica de construir un modelo cuantitativo (determinista) estático consiste en : a) elegir las N variables de estado del sistema: x_i variable en U_i , ..., x_n variable en U_n , y

b) localizar $k \geq 1$ funciones F_1, \dots, F_k , todas ellas de,

$U_1 \times \dots \times U_n$ en \mathbb{R} , de forma que se verifique:

$$(\bar{x}_1, \dots, \bar{x}_n) \text{ es un estado de equilibrio } \left\{ \begin{array}{l} F_1(\bar{x}_1, \dots, \bar{x}_n) = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ F_k(\bar{x}_1, \dots, \bar{x}_n) = 0 \end{array} \right.$$

Por consiguiente, los estados de equilibrio del sistema constituyen un subconjunto de soluciones de las relaciones (funcionales) matemáticas:

$$\begin{array}{l} F_1(x_1, \dots, x_n) = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ F_k(x_1, \dots, x_n) = 0 \end{array}$$

Modelos dinámicos.

Un sistema que no está en equilibrio se caracteriza por cambiar de estado en el transcurso del tiempo. Para describir su evolución y descubrir el estado de equilibrio al que eventualmente vendrá a dar el sistema con el paso del tiempo, los científicos construyen *modelos dinámicos*.

En estos modelos las variables de estado dependen, en general, de forma desconocida del tiempo.

Decir que se conoce la evolución en el tiempo de un sistema equivale a decir que se sabe como evolucionan en el tiempo sus variables de estado x_1, \dots, x_N . Como el tiempo es una variable en

R , la evolución en el tiempo de las variables de estado x_1, \dots, x_N tiene que ser concebida por medio de funciones de R en U_1, \dots, U_N .

Como el tiempo es una variable en R , la evolución en el tiempo de las variables de estado x_1, \dots, x_N tiene que ser concebida por medio de funciones de R en U_1, \dots, U_N .

Por costumbre generalizada, estas funciones se representan mediante la misma letra con que se representan las variables, así, no se escribe $x_1 = f_1(t), \dots, x_N = f_N(t)$, sino

$$x_1 = x_1(t), \dots, x_N = x_N(t).$$

Forma general de los modelos cuantitativos.

Los aspectos dinámicos del sistema resultan perfectamente capturados por las funciones $t \rightarrow x_1(t), \dots, t \rightarrow x_N(t)$.

Pero lo más frecuente es que no se conozcan tales funciones. Los modelos cuantitativos (deterministas) dinámicos establecen relaciones matemáticas a las que tales funciones obedecen.

Generalmente asumen la siguiente forma:

k funciones F_1, \dots, F_k establecen las relaciones existentes entre las velocidades de variación de las variables de estado, en un instante temporal dado, con los valores que en ese mismo instante tomen dichas variables de estado. Formalmente:

$$\begin{array}{l} \text{velocidad de variación } x_1 \text{ en } t = F_1(x_1(t), \dots, x_N(t)) \\ \text{.....} \\ \text{velocidad de variación } x_N \text{ en } t = F_k(x_1(t), \dots, x_N(t)) \end{array}$$

Leyes dinámicas.

La evolución del sistema, esto es, el curso de las variables de estado a lo largo del tiempo, conocido el estado inicial del sistema, se obtienen « resolviendo » las ecuaciones anteriores.

Para construir sistemas dinámicos los científicos utilizan *leyes dinámicas*: relaciones matemáticas entre las variables de estado en las que la variable tiempo desempeña un papel esencial.

Los estados de equilibrio se pueden determinar a partir de modelos dinámicos.

Es interesante advertir que los modelos dinámicos sirven también para determinar los estados de equilibrio de los sistemas.

En efecto; como el equilibrio las tasas de variación de las variables de estado se anulan, los estados de equilibrio se obtienen anulando los primeros miembros de las ecuaciones (*) y

viendo para qué valores constantes (\bar{x} , \bar{y} , ... , \bar{w}), se satisfacen las ecuaciones así corregidas.

MODELOS DE PROCESOS O CADENAS DE MARKOV

Los modelos más sencillos de cambios de marca se basan en una analogía entre el comportamiento de compra del consumidor y determinados procesos físicos que pueden describirse mediante una relación conocida por el nombre de procesos o cadenas de Markov. El rasgo clave de un proceso de Markov reside en que, el « estado » del « sistema en cualquier momento del tiempo, depende únicamente de su estado en el momento inmediatamente anterior y no de otros sucesos pasados ». En términos estrictos, un proceso de Markov es « un proceso estocástico tal, que si conocemos el resultado del último experimento podemos prescindir de toda la demás información del pasado con el objeto de predecir el futuro ».

En el contexto del comportamiento de compra del consumidor, el consumidor es un « sistema » y la marca que compra es el « experimento » y la marca específica es el « resultado » del experimento. El rasgo que distingue un proceso de Markov de otros procesos estocásticos (probabilísticos), reside en el supuesto de que la « memoria » del sistema únicamente tiene un período de duración.

En un modelo del proceso de Markov se centra la atención en los cambios o transiciones de la selección de marcas por parte del consumidor.

En la mayoría de las aplicaciones que se han realizado hasta el presente, se ha supuesto que los consumidores compra a intervalos regulares, por ejemplo mensualmente. Esto constituye un proceso de Markov discreto.

También resulta posible operar con procesos continuos en los que no se efectúa supuesto alguno de compra a intervalos regulares.

El « estado » del consumidor en cada mes es, por consiguiente, la marca comprada en dicho mes.

El modelo representa cambios de marca que vienen determinados por probabilidades (conocidas o supuestas) de « transiciones » de una marca a otra. Por ejemplo, si sólo existen dos marcas de un producto (A y B), entonces los cambios de marca podrían predecirse sobre la base de probabilidades transitivas P_{AA}, P_{AB}, P_{BA} y P_{BB} .

La cantidad P_{AA} representa la probabilidad de que un consumidor, una vez elegida la marca A en un período, compre de nuevo la misma marca en el período siguiente; P_{AB} es la probabilidad de que un consumidor cambie de marca A a la marca B, etc. Si se supone un número de consumidores constante, entonces todos los consumidores de la marca A deben comprar algo $P_{AA} + P_{AB} = 1$. Por la misma razón,

$$P_{BA} + P_{BB} = 1.$$

El modelo de los procesos de Markov se ha aplicado a muchos tipos distintos de procesos físicos. Más recientemente ha sido utilizado por los científicos sociales para representar procesos sociales como el cambio de ocupaciones, cambios en los tamaños de las empresas de una industria y casos parecidos.

Al menos parte del atractivo que tiene el modelo de los procesos de Markov para los científicos sociales e investigadores de mercado, se debe a su simplicidad científica. Si los supuestos del modelo resultan ser adecuados y útiles para una determinado problema, entonces pueden llevarse a cabo una multitud de análisis utilizando sus « bien sabidas » propiedades (para los matemáticos).

Estas propiedades devienen convenientes para tratar con procesos de Markov, pero debe señalarse una vez más que la conveniencia de cálculo no está necesariamente ligada al realismo.

Los supuestos del más sencillo de los modelos discretos del proceso de Markov con probabilidades transitivas (de cambio) constantes deja mucho que desear en cuanto a descripción del comportamiento del consumidor.

Entre las críticas que se han hecho de este modelo, son de señalar:

1.- Los consumidores raramente a intervalos regulares; la frecuencia de las compras pueden variar, y en la misma transacción pueden muy bien comprarse dos o más marcas del mismo producto.

2.- Además de los cambios de una marca a otra, una descripción realista del comportamiento del consumidor debe tener en cuenta los consumidores que entran y salen del mercado.

Por ejemplo, en un estudio de un producto alcohólico destilado, se averiguó que el número de nuevos clientes que cada año entraban en el mercado superaba al número de aquellos que cambiaban de marca.

3.- Las probabilidades de transición de una marca a otra no se mantienen constantes; el objetivo y efecto de las actividades promocionales es el de cambiar estas probabilidades.

4.- Las elecciones de las marcas pueden muy bien estar influidas por toda una serie de compras anteriores, y no solamente por la del período anterior. Se han sugerido que los consumidores « aprenden » sus preferencias.

Aunque críticas ponen en duda la validez general del modelo simple de las cadenas de Markov, en determinadas situaciones, sin embargo, puede dar lugar a unas predicciones razonablemente precisas del comportamiento del mercado.

Al proyectar el porcentaje de mercado a conseguir por marcas nuevas, sobre la base de resultados de pruebas del mercado,

No obstante, un uso similar de un modelo simple de los procesos de Markov para proyectar los resultados de pruebas de mercado para un producto de lavandería, resultó un fracaso. El modelo se utilizó para predecir el porcentaje de mercado que obtendría una marca si se comercializara durante un período de tiempo considerable al mismo precio, con la misma publicidad, etc., como se hace en la pruebas o tests de mercado. También se averiguó que las evaluaciones de los consumidores con respecto a la marca nueva cambiaban después de haberla utilizado durante varios meses.

En consecuencia, las pruebas de mercado, y la marca no consiguió mantenerse al nivel que se esperaba.

Refinamientos del modelo de los procesos de Markov.

Con el fin de superar las debilidades del modelo simple de los procesos de Markov, pueden hacerse varias modificaciones.

Quizás el mejor modo de tratar una frecuencia de compra variable consiste en eliminar el supuesto de los intervalos de compra regulares. Las probabilidades transitivas pueden tratarse sobre una base de compra a otra, y la frecuencia de compra puede representarse como una variable separada. Este enfoque nos conduce a un modelo de los « procesos semi-Markov ». Herniter y Howard han trabajado con este tipo de modelo y han concluido que su mayor complejidad se ve más que compensada por su mayor realismo.

En el modelo simple de los procesos de Markov, el problema de las entradas y salidas del mercado puede tenerse en consideración permitiendo que los consumidores cambien en ambos sentidos a un estado artificial de « no compra ».

El mismo propósito se consigue más fácilmente en el modelo de los procesos semi-Markov, permitiendo que la frecuencia de compra tome el valor cero en un intervalo de tiempo determinado.

Los cambios en las probabilidades transitivas pueden representarse como resultados de las actividades proporcionales de las empresas competidoras. Si se sabe o se supone que las probabilidades cambian con frecuencia, entonces es posible que las pautas de compra nunca se acerquen al « equilibrio » hacia el que se mueven si las tasas de cambio de marcas se mantiene constantes.

La mayor parte de los teoremas matemáticos de los procesos de Markov tratan de la predicción de las condiciones de equilibrio, de la tasa de aproximación al equilibrio, etc., y son de muy poco uso si cambian las probabilidades transitivas. Un mercado en cambio constante puede analizarse mejor mediante la simulación de su comportamiento durante un período de tiempo. Aunque la simulación es un enfoque relativamente caro y que lleva mucho tiempo, parece que es necesario si se intentan hacer predicciones significativas del comportamiento.

UN MODELO DE COMPETENCIA ENTRE TRES LECHERIAS

Se desea modelar y analizar el movimiento de los clientes entre los proveedores, suponiendo por simplicidad, que la misma fracción de los clientes cambiara de una lechería cualquiera a otra durante cada período de tiempo (un mes).

Se supone que al comenzar el modelo las lecherías 1, 2 y 3 controlan las fracciones X_0 , Y_0 , Z_0 del mercado. También supone que son las únicas repartidoras cuyas fracciones deben ascender a 1 ; $X_0 + Y_0 + Z_0 = 1$ suponga que despues de un mes la lechería 1 ha logrado mantener la fracción A_{11} de sus propios clientes y además ha atraído las fracciones A_{12} de los clientes de la lechería 2 y A_{13} de la lechería 3. Suponga que el número de clientes permanece constante por lo tanto se establece ecuaciones similares para cada una de las lecherías.

Siendo X_1 la fracción del mercado que tiene la lechería 1 despues de un mes.

$$X_1 N = A_{11}(X_0 N) + A_{12}(Y_0 N) + A_{13}(Z_0 N)$$

Entonces

$$X_1 = A_{11}X_0 + A_{12}Y_0 + A_{13}Z_0$$

$$Y_1 = A_{21}X_0 + A_{22}Y_0 + A_{23}Z_0$$

$$Z_1 = A_{31}X_0 + A_{32}Y_0 + A_{33}Z_0$$

Se considera un ejemplo numérico para ilustrar las ideas anteriores. Considerando a X_0 como las partes iniciales del mercado y la matriz de transición A es

$$X_0 = \begin{bmatrix} 0.2 \\ 0.3 \\ 0.5 \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} 0.8 & 0.2 & 0.1 \\ 0.1 & 0.7 & 0.3 \\ 0.1 & 0.1 & 0.6 \end{bmatrix}$$

obtenemos que $X_1 = Ax_0$, $X_2 = Ax_1$, ... $X_{17} = Ax_{16}$ donde el significado de hecho es que todas las X_r para $r \geq 16$ son iguales por lo tanto las partes del mercado despues de 16 meses el mercado esta en equilibrio cuando se tiene que

$$X_{16} = \begin{bmatrix} 0.450 \\ 0.350 \\ 0.200 \end{bmatrix}$$

Después se quiere averiguar el comportamiento límite y equilibrio. Se examina si las matrices columnas X_r pudieran tender a alguna matriz columna X_∞ a medida que r crece.

Ya que $X_{r+1} = AX_r$, si X_k tiende a X_∞ se sigue que $X_\infty = AX_\infty$; de

modo similar, ya que se tiene $1_3^T X_r = [1] \forall r$, se sigue que, también

$1_3^T X_\infty = [1]$. En resumen se tiene lo siguiente.

Si la X_r tienden a un límite $X_\infty = [X_\infty, Y_\infty, Z_\infty]^T$,

entonces

$$AX_\infty = X_\infty \text{ y } X_\infty + Y_\infty + Z_\infty = 1$$

Como A es de 3×3 , la ecuación $AX_\infty = X_\infty$ representa tan solo un sistema de 3 ecuaciones combina con el sistema anterior para dar una incognita, X_∞ , Y_∞ y Z_∞ . Esto es si las X_r tienden a un límite.

Los candidatos posibles se encuentran al resolverse este sistema.

$$0.8X_\infty + 0.2Y_\infty + 0.1Z_\infty = X_\infty$$

$$0.1X_\infty + 0.7Y_\infty + 0.3Z_\infty = Y_\infty$$

$$0.1X_\infty + 0.1Y_\infty + 0.6Z_\infty = Z_\infty$$

$$X_\infty + Y_\infty + Z_\infty = 1$$

concluyendo que al resolver este sistema se encuentra que

$$X_\infty = 0.45, Y_\infty = 0.35, Z_\infty = 0.20.$$

Lo que demuestra que es el único límite posible.

Que de ningún modo depende de las contribuciones iniciales del mercado.

Esto quiere decir que si el suministro de leche permanece constante y las ventas también son constantes se establece un equilibrio con las proporciones de

$$X_r = \begin{bmatrix} 0.45 \\ 0.35 \\ 0.20 \end{bmatrix} \text{ y esto puede suceder para una economía estable.}$$

Es decir se inicia con una proporción del mercado de

$$X_0 = \begin{bmatrix} 0.2 \\ 0.3 \\ 0.5 \end{bmatrix} \text{ y se llega a un equilibrio cuando } X_r = \begin{bmatrix} 0.45 \\ 0.35 \\ 0.20 \end{bmatrix}$$

para $r = 17$ meses y permanece en equilibrio para $X_\infty, Y_\infty, Z_\infty$;
 $X_\infty = 0.45, Y_\infty = 0.35, Z_\infty = 0.20$

O sea que no importa con que fracción del mercado inician al paso de los meses se establece un equilibrio con las proporciones del mercado.

$$X_\infty = \begin{bmatrix} 0.45 \\ 0.35 \\ 0.20 \end{bmatrix}$$

La forma de Jordan es extremadamente importante.

Es éste ejemplo modeló como evolucionaban de mes a mes las partes del mercado controlado por tres lecherías; el resultado fue $x_{r+1} = Ax_r$ en donde los tres componentes de x_r representan las

fracciones del mercado en manos de cada lechería en el mes r y en donde la matriz A de transición describe el efecto de las fuerzas del mercado.

La matriz de transición A del modelo de competencia entre lecherías tenía eigenvalores $\lambda_1 = 0.5, \lambda_2 = 0.6$ y $\lambda_3 = 1.0$, que se demostró así como un conjunto linealmente independiente de tres eigenvectores asociados v_1, v_2 y v_3 .

Entonces como $p(A) = 1$, se sigue del teorema clave (potencias de matrices no defectuosas). Sea $A, p \times p$ una matriz no defectuosa.

Entonces, a medida que i tiende a más infinito:

a) A^i converge a cero si y sólo si el radio espectral $\rho(A) < 1$.

b) $A^1 x_0$ converge a cero para toda x_0 si y sólo si $p(A) < 1$.

c) A^1 está acotado si y sólo si $p(A) \leq 1$.

d) $A^1 x_0$ está acotada para toda x_0 si y sólo si $p(A) \leq 1$.

e) $\|A^1\|$ tiende a infinito en alguna norma matricial (y por lo tanto en todas las normas matriciales) si y sólo si $p(A) > 1$.

f) $\|A^1 x_0\|$ tiende a infinito para alguna x_0 en alguna norma vectorial (por lo tanto en todas las normas vectoriales) si y sólo si $p(A) > 1$.

Retomando lo anterior: Que A^1 y $A^1 x_0$ están acotadas para toda x_0 .

De hecho, como $v_1, v_2,$ y v_3 forman una base para \mathbb{R}^3 , x_0 se puede escribir como

$$x_0 = a_1 v_1 + a_2 v_2 + a_3 v_3$$

Entonces,

$$A^1 x_0 = a_1 (0.5)^1 v_1 + a_2 (0.6)^1 v_2 + a_3 (1.0)^1 v_3$$

lo que demuestra que $A^1 x_0$ converge a un múltiplo de $a_3 v_3$ para toda x_0 . En ese modelo, la suma de los elementos de cada estado

($x_1 = A^1 x_0$) es igual a 1, de modo que esto deber ser cierto en el límite $a_3 v_3$ también el límite entonces debe ser igual a

$$[0.45 \quad 0.35 \quad 0.20]^T$$

para toda x_0 .

MATRICES NO NEGATIVAS Y DE MARKOV

En el ejemplo se analizó el modelo de competencia entre lecherías y mostró que en ese caso específico $A^l x_0$ converge a un múltiplo del eigenvector v_3 asociado con $\lambda_3 = 1$, sin importar la x_0 inicial qué se tome.

«Definición. Se dice que una matriz A , $p \times q$, es no negativa si y sólo si todos sus elementos son no negativos: $A_{ij} \geq 0$.

Una matriz de Markov es una matriz $p \times p$ no negativa tal que la suma de los elementos en cada una de sus columnas es igual a 1.

Si se define la matriz $p \times 1$, $1 = [1 \ 1 \ \dots \ 1]^T$, entonces $1^T A = 1^T$, porque los elementos de cada columna de una matriz de Markov A suman 1; de modo de $\lambda = 1$ es un eigenvector de A . Ya que $\|A\|_1 = 1$, $|\lambda| \leq 1$ para cada eigenvector λ de A debido al teorema clave (7.10) e); estos dos hechos juntos implican que $\rho(A) = 1$.

Además, como $\|A^l\|_1 = \|A\|_1^l = 1$, se sabe que A^l esta acotada entonces implica que cada eigenvector de magnitud 1 tiene multiplicidades algebraicas y geométricas iguales -esto es, que A tiene un "conjunto completo" de eigenvectores asociados con tales eigenvalores. Sin embargo, esto no es suficiente para demostrar que $A^l x_0$ converge para toda x_0 , por ejemplo; la competencia de las lecherías nos ha llevado a una matriz de Markov.

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

tiene eigenvalores $+1$ y -1 , y eigenvectores asociados $v_1 = [1 \ 1]^T$

y $[1 \ -1]^T$; si se escribe $x_0 = a_1 v_1 + a_2 v_2$, entonces $A^l x_0$ es igual ya sea a x_0 (para l par) o $a x_0 - 2a_2 v_2$ (para l impar), por lo tanto,

$A^l x_0$ converge (a x_0 de hecho) si y sólo si $a_2 = 0$ -esto es, cuando

x_0 es 0 o algún eigenvector asociado con $\lambda = 1$.

Si fuera posible de algún modo tener la seguridad de que todos los eigenvalores de A distintos de $\lambda = 1$ tuvieran una magnitud estrictamente menor que 1, no se daría la situación que se acaba de describir; esto requiere, sin embargo, de algunas condiciones adicionales en la matriz A de Markov como lo demuestra el ejemplo de la matriz 2 x 2 que se acaba de presentar.

La teoría desarrollada para tratar este tema se aplica más en general y no sólo a las matrices de Markov.

POBLACION DE COMPETENCIA

En este modelo se considera la situación un poco complejo de dos poblaciones que compiten una contra otra. Los números en estas poblaciones se denotan mediante Z_i y G_i que pueden visualizarse como el conteo de zorros y gallinas.

Suponga que las gallinas, sin zorros que las molesten tienen una tasa de natalidad que exceda a la tasa de mortalidad; para ser más específico, suponga que en esta situación se tiene $G_{i+1} = 1.2G_i$.

Si gallinas para alimentarse, sería de esperar que los zorros comenzaran a extinguirse digamos que $Z_{i+1} = 0.6Z_i$.

Se quiere poder un modelo a lo que sucede con los zorros tienen éxito devorando cierto número de gallinas en cada período de tiempo, suponiendo que esto permitiera un incremento en la población de zorros proporcional al número de gallinas devoradas.

Para ser más específico, suponga que $Z_{i+1} = 0.6Z_i + 0.5G_i$.

La población de gallinas obviamente comenzara a decrecer debido a los zorros, de modo que se toma.

$G_{i+1} = 1.2G_i - kZ_i$, donde K representa la tasa de gallinas devoradas por zorros; k permanece como variable para estudiar el efecto de diferentes tasas de mortalidad.

Suponiendo que existe un número inicial de 100 gallinas y de 100 zorros, se obtiene como modelo.

$$\begin{aligned} Z_{i+1} &= 0.6Z_i + 0.5G_i \\ (1) \quad G_{i+1} &= -kZ_i + 1.2G_i \text{ para } i = 1, 2, 3 \end{aligned}$$

con $Z_1 = 100$ y $G_1 = 1000$.

Se utilizan matrices para analizar el comportamiento de estas poblaciones conforme pasa el tiempo. Sean

$$A = \begin{bmatrix} 0.6 & 0.5 \\ -k & 1.2 \end{bmatrix} \quad x_1 = \begin{bmatrix} Z_1 \\ G_1 \end{bmatrix} \text{ así que } x_1 = \begin{bmatrix} 100 \\ 1000 \end{bmatrix}$$

el modelo 1 se convierte en

$$(2) \quad x_{i+1} = Ax_i \quad \text{para } i = 1, 2, \dots, \text{ con } x_1 = [100 \ 1000]^T$$

se sigue de (2) que $x_{i+1} = Ax_i = A(Ax_{i-1}) = A^2x_{i-1}$ y así sucesivamente de modo que de hecho,

$$(3) \quad x_{i+1} = A^i x_1 \quad \text{para } i = 0, 1, 2, \dots$$

De acuerdo con (3) el estudiar el comportamiento de x_i conforme i crece es equivalente a estudiar el comportamiento de las potencias de A^i de A de forma parecida como se realizó en el modelo de competencia de lecherías, sin embargo para este caso no se tienen las condiciones adicionales para ese modelo.

De hecho en éste modelo, A^i puede comportarse de manera diferente para diferentes valores de k .

Ejemplo 1:

Si se toma el modelo para $k = 0.1$ de manera que

$$A = \begin{bmatrix} 0.6 & 0.5 \\ -0.1 & 1.2 \end{bmatrix}$$

En la tabla siguiente se muestran los números aproximados de zorros y gallinas en el punto i -ésimo de tiempo para varios valores de i .

El modelo verifica de manera experimental lo que se esperaría: para una tasa de mortalidad k baja, la población de gallinas crece sin límite y esto permite que la población de zorros también crezca sin límite.

La tabla indica que finalmente las dos poblaciones se igualan.

i	1	2	3	4	5	6	8	12	16	20	30	100
Z _i	100	560	931	1244	1523	1783	2292	3470	5107	7483	19409	15328199
G _i	1000	1190	1372	1553	1739	1934	2367	3488	5111	7483	19409	15328199

Ejemplo 2: Considere una tasa de mortalidad de gallinas significativamente mayor, $k = 0.18$; se verá que esto causa que la población de gallinas tienda a quedar exterminada, lo cual a su vez lleva a la muerte de los zorros - ambas poblaciones se extinguen. Con $k = 0.18$, se tiene:

$$A = \begin{bmatrix} 0.6 & 0.5 \\ -0.18 & 1.2 \end{bmatrix}$$

La tabla, abajo, muestra los números aproximados de zorros y gallinas en el punto i -ésimo de tiempo para varios valores de i . El modelo verifica de manera experimental lo que se esperaría: ambas poblaciones tienden a desaparecer.

i	1	2	3	4	5	8	12	16	20	30	40	60	80	100
Z _i	100	560	927	1214	1434	1808	1854	1654	1371	713	312	43	3	0
G _i	1000	1182	1317	1413	1477	1530	1400	1177	940	459	193	25	2	0

Estos dos ejemplos muestran que el modelo puede dar dos tipos de comportamiento radicalmente diferentes, para diferentes valores de k .

Se quiere saber:

Como se puede determinar el comportamiento de x_1 a medida que i crece para cualquier valor específico de k . Ya que:

$$x_{i+1} = A^i x_1 \text{ para } i = 0, 1, 2, \dots \text{ dice que el comportamiento de}$$

x_1 se determina por el comportamiento de las potencias

A^i de igual se quiere saber como se puede determinar el comportamiento de las potencias A^i a medida que i crece para cualquier valor específico de k .

La forma de Jordan es extremadamente importante en las matemáticas aplicadas porque da la clave de la respuesta a un conjunto de preguntas concernientes a una amplia gama de aplicaciones.

En el ejemplo de las lecherías se modeló cómo evolucionaban de mes a mes las partes del mercado controlado por tres lecherías; el resultado fue $(2,4) - x_{r+1} = Ax_r$ - en donde las tres componentes de x_r representan las fracciones del mercado en manos de cada lechería en el mes r y en donde la matriz A de transición describe el efecto de las fuerzas del mercado.

De manera similar se modeló la evolución de poblaciones de zorros y gallinas en competencia con $x_{i+1} = Ax_i$ - en donde las dos componentes de x_i representan los números de zorros y gallinas en el tiempo i , y la matriz A describe su competencia.

En general, los modelos matemáticos conducen a menudo a una matriz $p \times 1, x_1$, que describe el estado de algún sistema complicado en el tiempo i a una matriz A $p \times p$, que representan los procesos internos del sistema, de tal modo que

$$x_{i+1} = Ax_i$$

modela la evolución del sistema en el tiempo. Denotando el estado inicial por x_0 , es posible escribir lo anterior en forma

equivalente

$$x_i = A^i x_0 \text{ para } i \geq 0$$

Las preguntas importantes para esos ejemplos específicos, y generalmente de importancia modular, son: ¿qué pasa conforme transcurre el tiempo?, ¿tiende el sistema al equilibrio?

Los estados x_i ¿llegan a ser arbitrariamente grandes?

¿Tienden a cero?, ¿oscilan?

Gracias a $x_i = A^i x_0$ para $i \geq 0$, la forma de Jordan de A puede manejar tales asuntos.

Escribiendo $A = QJQ^{-1}$, en donde J es una forma de Jordan A , es posible aplicar el teorema (semejanza y potencias).

Suponiendo que B es semejante a A con $B = P^{-1}AP$, entonces:

a) Para cada entero k positivo, B^k es semejante a A^k siendo

$$B^k = P^{-1}A^kP.$$

b) $\det B = \det A$.

c) B es no singular si y sólo si A es no singular.

d) Si A y B son no singulares, entonces B^k es semejante a A^k siendo $B^k = P^{-1}A^kP$ también para enteros k negativos de modo que en especial $B^{-1} = P^{-1}A^{-1}P$.

e) Si f es un polinomio: $f(x) = a_0x^m + \dots + a_m$ y si F(X) para una matriz cuadrada X es

$$a_0X^m + \dots + a_mI,$$

entonces $f(B)$ es semejante a $f(A) = P^{-1}f(A)P$.

Demostración

a) $B^k = (P^{-1}AP)(P^{-1}AP) \dots (P^{-1}AP)$ - con k términos - y la ley asociativa permite la eliminación de los paréntesis de modo que P^{-1} repetidamente da I y el producto se colapsa a $P^{-1}A^kP$.

b) $\det B = \det (P^{-1}AP) = \det (P^{-1}) \det (A) \det (P) = \det A$.

c) Esto es consecuencia de b) y del hecho de que una matriz es no singular si y sólo si su determinante es diferente de cero.

$$d) B^{-1} = (P^{-1}AP)^{-1} = P^{-1}A^{-1}(P^{-1})^{-1} = P^{-1}A^{-1}P.$$

El resultado para una k negativa arbitraria sigue de aplicar

a) a B^{-1} y a A^{-1} con $|k|$.

$$e) f(B) = a_0P^{-1}A^mP + a_1P^{-1}A^{m-1}P + \dots + a_mP^{-1}P =$$

$$= P^{-1}(a_0 A^m + a_1 A^{m-1} + \dots + a_m I) P = P^{-1}f(A)P.$$

como se afirmó.

Observe que la relación $B^k = P^{-1}A^kP$ puede ser un extremo útil si es necesario calcular o analizar B^k y si A^k se analiza con mayor facilidad, y entonces

$$A^k = QJQ^{-1} \text{ y por lo tanto}$$

$$x_1 = QJ^k(Q^{-1}x_0)$$

Esto permite estudiar el comportamiento de J^k -un problema mucho más simple que para A^k directamente. Esto se vuelve todavía más sencillo debido al teorema clave en el que J misma es una matriz diagonal en bloques con bloques de Jordan J_i en la diagonal:

entonces, sólo es necesario estudiar las potencias J_i^k , ya que

$$J^k = \begin{bmatrix} J_1^k & 0 & \dots & 0 \\ 0 & J_2^k & \dots & 0 \\ 0 & 0 & \dots & J_\mu^k \end{bmatrix}$$

Ejemplo. La matriz A del modelo de competencia entre poblaciones con $k = 0.18$ en el ejemplo de presa depredador, tenía un eigenvalor doble $\lambda_1 = \lambda_2 = 0.9$ como se demostro en el ejemplo sobre

las poblaciones en competencia, (aunque entonces no estaba disponible esta terminología). La matriz defectuosa porque $(A - 0.9) x = 0$ necesita que

$$\begin{bmatrix} -0.30 & 0.5 \\ -0.18 & 0.3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

para la cual toda solución es sólo un múltiplo de $[5 \ 3]^T$. Ya que

$p(A) = 0.9 < 1$, sabemos que A^k tiende a cero como lo hace $A^k x_0$ para

x_0 .

Esto prueba lo que se indicó experimentalmente en el ejemplo: presa depredador las poblaciones en este caso mueren, sin importar la distribución de zorros y gallinas.

UNA APLICACION IMPORTANTE: ECUACIONES DIFERENCIALES MATRICIALES

Sea $x = f(t)$ una función que representa una cantidad física tal como el volumen de una sustancia, la población de determinada especie, la masa de una sustancia radiactiva o el número de pesos invertidos en bonos. Entonces el crecimiento de $f(t)$ está dado por su derivada $f'(t) = dx/dt$. Si $f(t)$ crece a una razón constante, entonces $dx/dt = k$ y $x = kt + C$; esto es $x = f(t)$ es una recta.

Con frecuencia es más interesante y más apropiado el considerar la tasa de crecimiento relativo definida por

$$\begin{aligned} \text{Tasas de crecimiento relativo} &= \frac{\text{tamaño real del crecimiento}}{\text{tamaño de } f(t)} \\ &= \frac{f'(t)}{f(t)} = \frac{x'(t)}{x(t)} \end{aligned} \quad (1)$$

Si la tasa de crecimiento relativo es constante, entonces

$$\frac{x'(t)}{x(t)} = a \quad (2)$$

o bien

$$x'(t) = ax(t) \quad (3)$$

La Ecuación (3) se llama ecuación diferencial ya que es una ecuación que tiene una derivada. No es difícil demostrar que las únicas soluciones de (3) son de la forma

$$x(t) = ce^{at} \quad (4)$$

donde c es una constante arbitraria. Sin embargo, si $x(t)$ representa una cantidad física, entonces normalmente se especifica un valor inicial $x_0 = x(0)$ de la cantidad.

Entonces, sustituyendo $t = 0$ en (4), tenemos $x_0 = x(0) = ce^{a \cdot 0}$.

$= c$, o bien

$$x(t) = x_0 e^{at} \quad (5)$$

La función $x(t)$ dada por (5) es la única solución de (3) que satisface la condición $x(0) = x_0$.

La Ecuación (3) surge en muchas aplicaciones interesantes. Algunas de ellas aparecen indudablemente en los textos de Cálculo, en el capítulo relativo a la función exponencial. En esta sección consideraremos una generalización de la Ecuación (3).

En el modelo discutido anteriormente, buscamos una función desconocida. A menudo se tiene que aparecen varias funciones vinculadas por varias ecuaciones diferenciales. Algunos ejemplos se dan más adelante. Considere el siguiente sistema de n ecuaciones diferenciales en n funciones incógnitas

$$\begin{aligned} x_1'(t) &= a_{11}x_1(t) + a_{12}x_2(t) + \dots + a_{1n}x_n(t) \\ x_2'(t) &= a_{21}x_1(t) + a_{22}x_2(t) \dots + a_{2n}x_n(t) \\ &\vdots \\ &\vdots \\ &\vdots \\ x_n'(t) &= a_{n1}x_1(t) + a_{n2}x_2(t) \dots + a_{nn}x_n(t) \end{aligned} \quad (6)$$

donde los a_{ij} son números reales. A (6) se le llama un sistema de

ecuaciones diferenciales lineales de primer orden de $n \times n$. El término "primer orden" significa que en el sistema únicamente aparecen las primeras derivadas.

Ahora, sea;

$$x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix}$$

Aquí $x(t)$ se llama función vectorial.

Definimos

$$x'(t) = \begin{pmatrix} x'_1(t) \\ x'_2(t) \\ \vdots \\ x'_n(t) \end{pmatrix}$$

Entonces, si definimos la matriz de $n \times n$.

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

el sistema (6) puede escribirse como

$$x'(t) = Ax(t) \quad (7)$$

Note que la ecuación (7) es casi idéntica a la Ecuación (3).

La única diferencia es que ahora tenemos una función vectorial y una matriz, mientras que antes teníamos una función "escalar" y un número (matriz de 1×1 .)

Para resolver la Ecuación (7) debemos conjeturar que una solución tendría la forma e^{At} . Pero ¿qué significa e^{At} ?

Responderemos a esta pregunta en un momento. Primero, recordemos el desarrollo en serie de la función e^t .

$$e^t = 1 + t \frac{t^2}{2!} + \frac{t^3}{3!} + \frac{t^4}{4!} + \dots \quad (8)$$

Esta serie converge para todo número real t . Entonces, para todo número real a .

Definición 1

La matriz e^A . Sea A una matriz de $n \times n$ con componentes reales (o complejas). Entonces e^A es una matriz de $n \times n$ definida por

$$e^A = I + A \frac{A^2}{2!} + \frac{A^3}{3!} + \frac{A^4}{4!} + \dots \quad (10)$$

Observación. No es difícil demostrar que la serie de matrices en la Ecuación (10) converge para cada matriz A , pero hacerlo nos desviaría de nuestro objetivo.

Sin embargo, podemos dar una indicación de por qué sucede esto. Primero definimos $|A|_1$ como la suma de los valores absolutos

de los componentes en el renglón i -ésimo de A . Entonces definimos la norma de A , denotada $|A|$, por

$$|A| = \max_{1 \leq i \leq n} |A|_i \quad (11)$$

Se puede mostrar que

$$|AB| \leq |A| |B| \quad (12)$$

y
$$|A + B| \leq |A| + |B| \quad (13)$$

Entonces, usando (12) y (13) en (10), obtenemos

$$|e^A| \leq 1 + |A| \frac{|A|^2}{2!} + \frac{|A|^3}{3!} + \frac{|A|^4}{4!} + \dots = e^{|A|}$$

Puesto que $|A|$ es un número real, $e^{|A|}$ es finito. Esto demuestra que la serie en (10) converge para toda matriz A .

Ahora veremos la utilidad de la serie en la Ecuación (10).

Teorema 1

Para cualquier vector constante c , $x(t) = e^{At}c$ es una solución de (7). Además, la solución de (7) dada por $x(t) = e^{At}x_0$ satisface $x(0) = x_0$.

Demostración Calculamos, usando (10):

$$x(t) = e^{At}c = \left[I + At + A^2 \frac{t^2}{2!} + A^3 \frac{t^3}{3!} + \dots \right] c \quad (14)$$

Pero como A es una matriz constante y c es un vector constante tenemos

$$\frac{d}{dt} A^k \frac{t^k}{k!} = \frac{d}{dt} \frac{t^k}{k!} A^k = \frac{k t^{k-1}}{k!} A^k = \frac{A^k t^{k-1}}{(k-1)!} = A \left[A^{k-1} \frac{t^{k-1}}{(k-1)!} \right] \quad (15)$$

Entonces, combinando (14) y (15), obtenemos (debido a que c es un vector constante).

$$\begin{aligned} x'(t) &= \frac{d}{dt} e^{At} c = A \left[I + At + A^2 \frac{t^2}{2!} + A^3 \frac{t^3}{3!} + \dots \right] c \\ &= A e^{At} c = Ax(t) \end{aligned}$$

Finalmente, en virtud de que $e^{A(0)} = e^0 = I$

$$x(0) = e^{A(0)} x_0 = I x_0 = x_0$$

Definición 2

Matriz solución principal. La matriz e^{At} se llama matriz solución principal del sistema $x' = Ax$.

Queda pendiente un problema mayor (y obvio). ¿Cómo calculamos e^{At} de una manera práctica? Comencemos con dos ejemplos.

Ejemplo 1

Sea $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$. Entonces

$$A^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2^2 & 0 \\ 0 & 0 & 3^2 \end{pmatrix}, A^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2^3 & 0 \\ 0 & 0 & 3^3 \end{pmatrix}, \dots, A^m = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2^m & 0 \\ 0 & 0 & 3^m \end{pmatrix}$$

$$y e^{At} = I + At + \frac{A^2 t^2}{2!} + \frac{A^3 t^3}{3!} + \dots = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2t & 0 \\ 0 & 0 & 3t \end{pmatrix}$$

$$+ \begin{pmatrix} \frac{1}{2!} & 0 & 0 \\ 0 & \frac{2^2 t^2}{2!} & 0 \\ 0 & 0 & \frac{3^2 t^2}{2!} \end{pmatrix} + \begin{pmatrix} \frac{1}{3!} & 0 & 0 \\ 0 & \frac{2^3 t^3}{3!} & 0 \\ 0 & 0 & \frac{3^3 t^3}{3!} \end{pmatrix} + \dots$$

$$= \begin{pmatrix} 1 + t \frac{1^2}{2!} + \frac{1^3}{3!} + \dots & 0 & 0 \\ 0 & 1 + (2t) + \frac{(2t)^2}{2!} + \frac{(2t)^3}{3!} + \dots & 0 \\ 0 & 0 & 1 + (3t) + \frac{3t^2}{2!} + \frac{(3t)^3}{3!} + \dots \end{pmatrix}$$

$$= \begin{pmatrix} e^t & 0 & 0 \\ 0 & e^{2t} & 0 \\ 0 & 0 & e^{3t} \end{pmatrix}$$

Ejemplo 2.

Sea $A = \begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix}$. Entonces, como es fácil de verificar,

$$A^2 = \begin{pmatrix} a^2 & 2a \\ 0 & a^2 \end{pmatrix}, A^3 = \begin{pmatrix} a^3 & 3a^2 \\ 0 & a^3 \end{pmatrix}, \dots, A^m = \begin{pmatrix} a^m & ma^{m-1} \\ 0 & a^m \end{pmatrix}, \dots$$

por lo que
$$e^{At} = \begin{pmatrix} \sum_{m=0}^{\infty} \frac{(at)^m}{m!} & \sum_{m=1}^{\infty} \frac{ma^{m-1} t^m}{m!} \\ 0 & \sum_{m=0}^{\infty} \frac{(at)^m}{m!} \end{pmatrix}$$

Ahora
$$\sum_{m=1}^{\infty} \frac{ma^{m-1} t^m}{m!} = \sum_{m=1}^{\infty} \frac{a^{m-1} t^m}{(m-1)!} = t + at^2 + \frac{a^2 t^3}{2!} + \frac{a^3 t^4}{3!} + \dots$$

$$= t(1 + at + \frac{a^2 t^2}{2!} + \frac{a^3 t^3}{3!} + \dots) = te^{at}$$

$$\text{En consecuencia } e^{At} = \begin{pmatrix} e^{\lambda_1 t} & te^{\lambda_1 t} \\ 0 & e^{\lambda_2 t} \end{pmatrix}$$

Como lo muestra el Ejemplo 1, es fácil calcular e^{At} es una matriz diagonal. El Ejemplo 1 muestra que si $D = \text{diag}(\lambda_1, \dots, \lambda_n)$,

entonces $e^{Dt} = \text{diag}(e^{\lambda_1 t}, e^{\lambda_2 t}, \dots, e^{\lambda_n t})$.

En el ejemplo 2, calculamos e^{At} para una matriz A en la forma canónica de Jordan. Resulta claro, pues que esto es todo lo que necesitamos hacer, tal como lo sugiere el siguiente teorema.

Teorema 2 Sea J la forma canónica de Jordan una matriz A sea

$$J = C^{-1}AC. \text{ Entonces } A = CJC^{-1} \text{ y}$$

$$e^{At} = Ce^{Jt}C^{-1} \quad (16)$$

Demostración

Primero notemos que

$$\begin{aligned} A^n &= (CJC^{-1})^n = \overbrace{(CJC^{-1})(CJC^{-1}) \dots (CJC^{-1})}^{n \text{ veces}} \\ &= CJ(C^{-1}C)J(C^{-1}C)J(C^{-1}C) \dots (C^{-1}C)JC^{-1} \\ &= C J^n C^{-1} \end{aligned}$$

De aquí se deduce que

$$(At)^n = C (Jt)^n C^{-1} \quad (17)$$

Por lo tanto

$$e^{At} = I + (At) + \frac{(At)^2}{2!} + \dots + C^{-1}C + C(Jt)C^{-1} + C \frac{(Jt)^2}{2!} C^{-1} + \dots$$

$$= C \left[I + (Jt) + \frac{(Jt)^2}{2!} \dots \right] C^{-1} = Ce^{Jt}C^{-1}$$

El teorema 2 nos dice que para evaluar e^{At} lo único que realmente necesitamos calcular es e^{Jt} . Cuando J es una diagonal (como sucede en la mayoría de los casos, sabemos entonces cómo determinar e^{Jt} . si A es una matriz de 2×2 que no es diagonalizable, entonces

$$J = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \text{ y } e^{Jt} = \begin{pmatrix} e^{\lambda t} & te^{\lambda t} \\ 0 & e^{\lambda t} \end{pmatrix}$$

como lo calculamos en el ejemplo 2. De hecho no es difícil evaluar

e^{Jt} donde J es cualquier matriz de Jordan. Primero es necesario

calcular e^{Jt} para una matriz bloque de Jordan B . Un método para lograrlo se da en los problemas.

Apliquemos ahora nuestros cálculos a un simple modelo biológico de crecimiento poblacional. suponga que en un ecosistema existen dos especies S_1 y S_2 relacionadas entre sí. Denotamos las

poblaciones de las especies en el momento t por $x_1(t)$ y $x_2(t)$.

Un sistema que gobierna el crecimiento relativo de las dos especies es

$$\begin{aligned} x_1'(t) &= ax_1(t) + bx_2(t) \\ x_2'(t) &= cx_1(t) + dx_2(t) \end{aligned} \quad (18)$$

Podemos interpretar las constantes a , b , c y d del siguiente modo. Si las especies están en competencia, entonces es razonable tener $b < 0$ y $c < 0$. Esto es verdadero porque al incrementarse la población de una de las especies disminuirá el crecimiento de la otra. Un segundo modelo es una relación entre depredador y presa.

Si S_1 es la presa y S_2 es el depredador (S_2 se como a S_1),

entonces es razonable tener $b < 0$ y $c > 0$ puesto que un incremento en la especie depredadora causará una disminución en la especie de la presa, mientras que un incremento en la especie de la presa causará un incremento en la especie depredadora (puesto que tendrán más alimento). Finalmente, en una relación simbiótica (cada especie vive de la otra) tendríamos $b > 0$ y $c > 0$. Por supuesto, las constantes a, b, c , y d dependen de una gran variedad de factores incluyendo alimento disponible, temporada del año, clima, límites debido a la sobrepoblación, competencia con otras especies y muchas otras. Analizaremos cuatro modelos diferentes usando el material en esta sección. Supongamos que t está medido en años.

Ejemplo 3

Un modelo de especies en competencia, considere el sistema.

$$x_1'(t) = 3x_1(t) - x_2(t)$$

$$x_2'(t) = -2x_1(t) + 2x_2(t)$$

Este sistema describe la influencia de las poblaciones de dos especies en competencia, sobre sus respectivas tasas de crecimiento. Suponga que las poblaciones iniciales son $x_1(0) = 90$ y

$x_2(0) = 150$. Determine las poblaciones de ambas especies para $t > 0$.

Solución: Tenemos $\begin{pmatrix} 3 & -1 \\ -2 & 2 \end{pmatrix}$. Los valores característicos de A son

$\lambda_1 = 1$ y $\lambda_2 = 4$ con sus correspondientes vectores característicos

$v_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ y $v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Entonces

$$C = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} \quad C^{-1} = -\frac{1}{3} \begin{pmatrix} -1 & -1 \\ -2 & 1 \end{pmatrix} \quad J = D = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} \quad e^{Jt} = \begin{pmatrix} e^t & 0 \\ 0 & e^{4t} \end{pmatrix}$$

$$e^{\lambda t} = Ce^{Jt}C^{-1} = -\frac{1}{3} \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} e^t & 0 \\ 0 & e^{4t} \end{pmatrix} \begin{pmatrix} -1 & -1 \\ -2 & 1 \end{pmatrix}$$

$$= -\frac{1}{3} \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} -e^t & -e^{4t} \\ -2e^{4t} & e^{4t} \end{pmatrix}$$

$$= -\frac{1}{3} \begin{pmatrix} -e^t & -2e^{4t} & -e^t & +e^{4t} \\ -2e^t & +2e^{4t} & -2e^t & -e^{4t} \end{pmatrix}$$

Por ejemplo, después de 6 meses ($t = 1/2$ año), $x_1(t) = 80e^{1/2} +$

$$10e^2 = 206 \text{ individuos, mientras que } x_2(t) = 160e^{1/2} - 10e^2 = 190$$

individuos. Más significativo aún es $160e^t - 10e^{4t} = 0$ cuando

$16e^t = e^{4t}$ o $16 = e^{3t}$ o $3t = \ln 16$ y $t = (\ln 16)/3 = 2.77/3 \approx 11$ meses. De esta manera la segunda especie será eliminada después de escasamente 11 meses, aun cuando comenzó con una población más grande. En los problemas 10 y 11 se le pide mostrar que ninguna población será eliminada si $x_2(0) = 2x_1(0)$ y que la primera

población será eliminada si $x_2(0) > 2x_1(0)$. Por lo tanto, como

Darwin sabía muy bien, la supervivencia en este modelo muy simple depende de los tamaños relativos de las especies competitivas cuando comienza la competencia.

Ejemplo 4

Un modelo depredador-presa. Consideremos el siguiente sistema en el que la especie 1 es la presa y la especie 2 es el depredador.

$$x_1'(t) = 2x_1(t) - x_2(t)$$

$$x_2'(t) = x_1(t) + 4x_2(t)$$

Encuentre las poblaciones de dos especies para $t > 0$ si las poblaciones iniciales son $x_1(0) = 500$ y $x_2(0) = 100$.

Solución

Aquí $A = \begin{pmatrix} 2 & -1 \\ 1 & 4 \end{pmatrix}$ y el único valor característico es $\lambda = 3$

con el único vector característico $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Una solución a la

ecuación $(A - 3I)v_2 = v_1$ da como resultado a $v_2 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$.

Entonces

$$C = \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix} \quad C^{-1} = \begin{pmatrix} 2 & 1 \\ -1 & -1 \end{pmatrix} \quad J = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$$

$$e^{Jt} = \begin{pmatrix} e^{3t} & te^{3t} \\ 0 & e^{3t} \end{pmatrix} = e^{3t} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \quad (\text{del Ejemplo 2})$$

$$= e^{3t} \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix} \begin{pmatrix} 2-t & 1-t \\ -1 & -1 \end{pmatrix} = e^{3t} \begin{pmatrix} 1-t & -t \\ t & 1+t \end{pmatrix}$$

De esta manera la solución al sistema es

$$x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} e^{At} x_0 = e^{3t} \begin{pmatrix} 1-t & -t \\ t & 1+t \end{pmatrix} \begin{pmatrix} 500 \\ 100 \end{pmatrix} = e^{3t} \begin{pmatrix} 500-600t \\ 100+600t \end{pmatrix}$$

Es obvio que la especie presa será eliminada después de $5/6$ año = 10 meses, aun cuando comenzó con una población cinco veces más grande que la especie depredadora. De hecho es fácil mostrar que independiente del tamaño inicial de la especie presa, ésta será eliminada en menos de un año.

Ejemplo 5

Otro modelo depredador-presa. Consideremos el modelo depredado-presa regido por el sistema

$$x_1'(t) = x_1(t) - x_2(t)$$

$$x_2'(t) = x_1(t) + x_2(t)$$

Si las poblaciones iniciales son $x_1(0) = x_2(0) = 100$,

determine las poblaciones de las dos especies para $t > 0$.

Solución

Aquí $A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ con ecuación característica $\lambda^2 - 2\lambda + 2 = 0$ raíces complejas $\lambda_1 = 1 + i$ y $\lambda_2 = 1 - i$ y vectores característicos

$$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{y} \quad v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Nótese que $\lambda_2 = \lambda_1$ y $v_2 = v_1$. Esto no debe sorprender puesto

que los valores propios de matrices reales aparecen en pares complejos conjugados y sus vectores propios correspondientes son complejos conjugados.

Entonces

$$C = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad C^{-1} = -\frac{1}{2i} \begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

$$J = D = \begin{pmatrix} 1+i & 0 \\ 0 & 1-i \end{pmatrix} \quad \text{y} \quad e^{Jt} = \begin{pmatrix} e^{(1+i)t} & 0 \\ 0 & e^{(1-i)t} \end{pmatrix}$$

Ahora por la fórmula de Euler, $e^k = \cos t + i \sin t$.

De esta manera $e^{(1+i)t} = e^t e^{it} = e^t (\cos t + i \sin t)$.

Análogamente, $e^{(1-i)t} = e^t e^{-it} = e^t (\cos t - i \sin t)$.

En consecuencia

$$e^{Jt} = e^t \begin{pmatrix} \cos t + i \sin t & 0 \\ 0 & \cos t - i \sin t \end{pmatrix}$$

y

$$e^{At} = C e^{Jt} C^{-1} = \frac{e^t}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \cos t + i \sin t & 0 \\ 0 & \cos t - i \sin t \end{pmatrix}$$

$$\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \frac{e^t}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \cos t + i \sin t & -1 \cos t + \sin t \\ \cos t - i \sin t & 1 \cos t + \sin t \end{pmatrix}$$

Finalmente

$$x(t) = e^{At} x(0) = e^t \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix} \begin{pmatrix} 1000 \\ 1000 \end{pmatrix} =$$

$$\begin{pmatrix} 1000e^t(\cos t + \sin t) \\ 1000e^t(\cos t - \sin t) \end{pmatrix}$$

La especie presa comienza a extinguirse cuando $1000e^t(\cos t - \sin t) = 0$ o cuando $\sin t = \cos t$. Se tiene que la primera solución positiva de esta última ecuación es $t = \pi/4 = 0.7854$ año = 9.4 meses.

Ejemplo 6

Un modelo de cooperación entre especies (Simbiosis).

Consideremos el modelo simbiótico gobernado por el sistema.

$$\begin{aligned}x_1'(t) &= 1/2 x_1(t) + x_2(t) \\x_2'(t) &= 1/4 x_1(t) - 1/2 x_2(t)\end{aligned}$$

Note que en este modelo la población de cada especie incrementa proporcionalmente a la población de la otra y decrece proporcionalmente con respecto a su propia población. Suponga que $x_1(0) = 200$ y $x_2(0) = 500$. Determine la población de cada especie para $t > 0$.

Solución

$$\text{Aquí } A = \begin{pmatrix} -1/2 & 1 \\ 1/4 & -1/2 \end{pmatrix}, \text{ con valores característicos}$$

$\lambda_2 = 0$ y $\lambda_2 = -1$ y sus correspondientes vectores característicos

$$v_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad \text{y} \quad v_2 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}. \quad \text{Entonces} \quad C = \begin{pmatrix} 2 & 2 \\ 1 & -1 \end{pmatrix},$$

$$C^{-1} = -1/4 \begin{pmatrix} -1 & -2 \\ -1 & 2 \end{pmatrix}, \quad J = D = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}, \quad \text{y } e^{Jt} = \begin{pmatrix} e^{0t} & 0 \\ 0 & e^{-t} \end{pmatrix} =$$

$$\begin{pmatrix} 1 & 0 \\ 0 & e^{-t} \end{pmatrix}. \text{ De esta manera}$$

$$e^{At} = -1/4 \begin{pmatrix} 2 & 2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & e^{-t} \end{pmatrix} \begin{pmatrix} -1 & -2 \\ -1 & 2 \end{pmatrix} =$$

$$-1/4 \begin{pmatrix} 2 & 2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & -2e^{-t} \\ -e^{-t} & 2e^{-t} \end{pmatrix} = -1/4 \begin{pmatrix} -2 & -2e^{-t} & -4 & +4e^{-t} \\ -1 & +e^{-t} & -2 & -2e^{-t} \end{pmatrix}$$

y así

$$x(t) = e^{At} x(0) = -1/4 \begin{pmatrix} -2 & -2e^{-t} & -4 & +4e^{-t} \\ -1 & +e^{-t} & -2 & -2e^{-t} \end{pmatrix} \begin{pmatrix} 200 \\ 500 \end{pmatrix}$$

$$= 1/4 \begin{pmatrix} -2400 & +1600e^{-t} \\ -1200 & -800e^{-t} \end{pmatrix} \begin{pmatrix} 600^{-t} & -400e^{-t} \\ 300 & +200e^{-t} \end{pmatrix}$$

Note que $e^{-t} \rightarrow 0$ cuando $t \rightarrow \infty$. Esto significa que conforme pase el tiempo, las dos especies simbióticas se aproximan a las poblaciones de equilibrio de 600 y 300, respectivamente. Ninguna población es eliminada.

ECUACIONES EN DIFERENCIAS Y MULTIPLICADORES DINAMICOS

Esta sección concierne al trabajo de sistemas de ecuaciones lineales en diferencias y varios tipos de multiplicadores Keynesianos dentro de los sistemas.

Keynes, -John Maynard (1883-1946).

Nació en Cambridge, Inglaterra.- Estudió en la Universidad de su ciudad natal.

Presentó con sus ideas una nueva dimensión de la economía, que ha quedado pantetizada en su rica producción como investigador y autor. Su obra más importante es *Teoría General de la Ocupación el Interés y el Dinero*, publicada en 1936. Las teorías que encierra la obra, según sus críticos, no estaban lo suficientemente elaboradas al detalle para constituir un nuevo cuerpo general de análisis económico completo. La formulación de las teorías de Keynes en un principio resultó para él mismo una novedad. Su principal contenido discrepaba en algunos aspectos tratados por la escuela clásica; fue por ello que dedico parte de sus obras a aclarar conceptos que no habían quedado debidamente precisados por los clásicos; por ejemplo: Ingreso, Consumo, Ahorro, e Inversión.

El papel armonizador de Keynes lo convirtió en representante distintivo de la Teoría Económica Moderna y modificador de la metodología científica de esta ciencia. Desarrollo una teoría macroeconómica con un equilibrio dinámico, es decir, con diferentes niveles de empleo. Algunos conceptos importantes utilizados por Keynes son: Demanda Efectiva, Multiplicador de la Inversión, Propensión Marginal al Consumo, Eficacia Marginal al Capital, etc.

Effecto Multiplicador: término acuñado originalmente por Keynes; significa las consecuencias que ocurren en la inversión, en el empleo y en general en toda la economía, como consecuencia de un incremento en la inversión inicial. El efecto multiplicador se representa matemáticamente así:

$$K = \frac{1}{1 - \frac{\Delta C}{\Delta Y}}$$

donde:

K = Multiplicador

ΔC = Incremento del Consumo

ΔY = Incremento del Ingreso

ΔC

— = Propensión Marginal al Consumo

ΔY

El multiplicador cambia directamente con las variaciones de la propensión marginal al consumo, aunque el multiplicador es siempre mayor que 1 y la propensión marginal al consumo es menor que 1.

Eficacia marginal del capital.

Término utilizado originalmente por Keynes; significa el tipo previsto de beneficios que se espera obtener por el incremento de una unidad más de capital. La eficacia marginal del capital es muy importante por que condiciona, junto con el tipo de interés, la inversión, aunque es muy inestable y se encuentra influida por el mercado de valores, la confianza comercial, las ventas, etc.

La eficacia marginal del capital esta condicionada por la previsión de beneficios y por el costo de reposición o precio de oferta de los bienes de capital; se caracteriza por su inestabilidad a corto plazo y por una tendencia así la disminución en el largo plazo.

Una ecuación en diferencias de orden r (≥ 1) con coeficientes constantes es expresado para un tiempo t discreto como:

$$P(G) X(t) = b(t), \quad (1)$$

donde $P(G) = \sum_{i=0}^r A_i G^i$, G es el operador diferencial tal que

$$G X(t) = X(t+1) \text{ con } G^0 = 1, \quad (2)$$

A_i ($i = 0, 1, \dots, r$) son constantes, $a_r \neq 0$, b es una función

conocida de t . La ecuación (1) es no homogénea si y solamente si $b(t)$ no es idénticamente cero, y su solución general es obtenida análogamente a la ecuación en diferencias lineal que la ecuación

$$X(t) = Xp(t) + \sum_{i=1}^k \sum_{s=1}^{w_i-1} C_{i,s} t^{s-1} \lambda_i^t, \quad (3)$$

donde $Xp(t)$ es una solución particular de la ecuación (1), λ_i ($i = 1, \dots, K$) establecido para toda raíz distinta de la ecuación característica.

$$P(\lambda) = A_0 + A_1 \lambda + \dots + A_{r-1} \lambda^{r-1} + A_r \lambda^r = 0 \quad (4)$$

con multiplicidad w_i satisfaciendo $\sum_{i=1}^k w_i = r$ y $C_{|s|}$ para constantes arbitrarias.

También una aplicación de la solución (3), analizaremos un problema económico titulado "La interacción entre el análisis multiplicador y el principio de aceleración". El cual es debido a Samuelson (1939) y Hicks (1950).

Una serie de notación es introducida como sigue: Y (ingreso nacional); C (consumo); I (inversión); K (provisión nacional de capital); v (coeficiente de capital); s (propensión marginal al ahorro); \bar{C} (nivel fijo de consumo); g (propensión de crecimiento de inversiones autónomas); y t un arbitrario período discreto de tiempo.

Haciendo a un lado el gobierno y la actividad del comercio extranjero. Establecemos el ingreso nacional definitivamente igual a la suma del consumo y la inversión.

$$Y_t = C_t + I_t + \bar{C} \quad (5)$$

El consumo y la inversión son supuestos a depender como se establecieron las variables arriba dentro del período previo en la siguiente forma

$$C_t = (1 - s) Y_{t-1} + \bar{C} \quad (6)$$

Con la suposición $1 - s > 0$

$$I_t = v Y_{t-1} - K_{t-1} + \alpha (1 + g)^t, \quad (7)$$

donde α es una constante.

Otra definición idénticamente es

$$K_t = K_{t-1} + I_t \quad (8)$$

substituyendo de (7) a (8) y entonces considerando un período largo, tenemos

$$K_{t-1} = v Y_{t-2} + \alpha(1+g)^{t-1},$$

el cual es sustituido en (7) da Y sustituyendo de (6) y (9) en (5) obtenemos una ecuación diferencial lineal de orden dos.

$$Y_t - (1-s+v)Y_{t-1} + vY_{t-2} = \bar{C} + \alpha g(1+g)^{t-1} \quad (10)$$

Estas ecuaciones no homogéneas serían resueltas como indicado por la solución (3).

Primero damos con la ecuación homogénea asociada de (10).

$$Y_t - (1-s+v)Y_{t-1} - vY_{t-2} = 0 \quad (11)$$

La ecuación característica de (11)

$$\lambda^2 - (1-s+v)\lambda + v = 0$$

cuyas raíces son:

$$\lambda_1 = 1/2(1-s+v + [(1-s+v)^2 - 4v]^{1/2})$$

$$\lambda_2 = 1/2(1-s+v - [(1-s+v)^2 - 4v]^{1/2})$$

Ya que $\lambda_1\lambda_2 > 0$ y $\lambda_1 + \lambda_2 > 0$, los siguientes tres casos son distinguidos.

Caso 1: λ_1 y λ_2 son distintos reales y positivos.

Caso 2: $\lambda_1 = \lambda_2$ es real y positivo.

Caso 3: λ_1 es un número complejo y λ_2 es su conjugado.

Estos casos serían verificados y examinados uno a uno.

El caso 1 compuesto de dos subcasos, caso 1a y caso 1b.

En el caso 1a, $v > (1+s^{1/2})^2$ y aquí $\lambda_1 > 1$, $\lambda_2 > 1$

$$\text{Así } Y_t = C_1\lambda_1^t + C_2\lambda_2^t \quad (12)$$

Incrementaría monotamente y eventualmente tiende a cero dependiendo de las condiciones iniciales.

En el caso 1b, $v < (1 - s^{1/2})^2$ y aquí ambos λ_1 y λ_2 son menos que uno, implicando que Y_t en (12) converge monotamente a cero.

Caso 2 también contiene dos subcasos. Caso 2a y caso 2b.

En el caso 2a, $v = (1 + s^{1/2})^2$ y $\lambda_1 = \lambda_2 > 1$.

Así el comportamiento de Y_t es de la misma clase como el caso 1a.

En el caso 2b donde $v = (1 - s^{1/2})^2$, así ocurre una situación similar al caso 1b.

En el caso 3, $(1 - s^{1/2})^2 < v < (1 + s^{1/2})^2$

Dado $\lambda_n = \alpha + i\beta$ y $\lambda_2 = \alpha - i\beta$,

donde $\alpha = 1/2(1 - s + v)$ y $\beta = 1/2[4v - (1 - sv)^2]^{1/2}$

ya que $\alpha \pm i\beta = p(\cos w \pm i \sin w)$,

donde $p = (\alpha^2 + \beta^2)^{1/2} = v^{1/2}$ y w es el ángulo entre λ y el eje real x , la solución de (11) llega a ser.

$$Y_t = (v^{1/2})^t (C_1 \cos t w + C_2 \sin t w) \quad (13)$$

donde C_1 y C_2 son constantes reales. Así si $v = 1$, Y_t en (13) oscila con una amplitud fija.

Si $v > 1$, Y_t oscila con una amplitud expandida; y si $v < 1$, Y_t oscila con una amplitud restringida.

En suma tenemos cinco diferentes casos:

a) Y_t converge monotamente a cero en caso $v \leq (1 - s^{1/2})^2$;

b) Y_t oscila con amplitud restringida en caso $(1 - s^{1/2})^2 < v < 1$;

c) Y_t oscila con amplitud fija en el caso $1 < v < (1 + s^{1/2})^2$;

d) Y_t expande monotamente o tiende eventualmente a cero dependiendo

sobre las condiciones iniciales en el caso $(1 - s^{1/2})^2 \leq v$.

También sera analizado el comportamiento de Y_t en la ecuación homogenea (11). Proximamnte una particular solución de (10) es un esbozo.

Partimos la ecuación en dos:

$$Y_t - (1 - s + v) Y_{t-1} + vY_{t-2} = \bar{C} \quad (10a)$$

$$Y_t - (1 - s + v) Y_{t-2} = \alpha g(1 + g)^{t-1} \quad (10b)$$

Y_t satisface (10a) es igual una constante:

$$\bar{Y} = \bar{C}/s. \text{ Estableciendo } Y_t = Z(1 + g)^{t-1}$$

en (10b) obtenemos $Z = \alpha g / (s + gs + g - vg + g^2)$. Aqui tenemos una solución particular de la ecuación (10) como sigue:

$$\bar{Y}_t = \frac{\bar{C}}{s} + \frac{\alpha g(1 + g)^{t+1}}{s + (1 + s - v)g + g^2} \quad (14)$$

La cual puede ser considerada como una trayectoria dinamica de equilibrio. Finalmente, la solución general de la ecuación (10) es obtenida por la suma de (14) a (12) o por (13). El primer término de (14) da el recorrido largo del nivel de equilibrio por lo cual el ingreso nacional aproximado en el caso Y_t en (12) es convergente mientras el segundo término de (14) es decir sera el super multiplicador el cual levanta, el largo recorrido del nivel de equilibrio en una proporción constante.

Ahora retornamos a un sistema de ecuaciones lineal en diferencias de orden r (≥ 1) con coeficientes constantes el cual, puede ser expresado generalmente como:

$$P(G)x(t) = b(t) \quad (15)$$

donde G es el operador en diferencias definida por (2):

$$P(G) = \begin{bmatrix} P_{11}(G) & \dots & P_{1n}(G) \\ \vdots & & \vdots \\ P_{n1}(G) & \dots & P_{nn}(G) \end{bmatrix}; \quad x(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}; \quad b(t) = \begin{bmatrix} b_1(t) \\ \vdots \\ b_n(t) \end{bmatrix}$$

$P_{hj}(G) = \sum_{l=0}^r A_l^{(hj)} G^l$ ($h, j = 1, \dots, n$) $A_l^{(hj)}$ es una constante al

menos uno de los $A_l^{(hj)}$ puede ser no cero, b_j es una función

conocida de t . La solución general de la ecuación (15) toma la forma.

$$x(t) = x_p(t) + \sum_{i=1}^k \sum_{s_i=0}^{w_i-1} C_{s_i} V_{s_i} t^{s_i} \lambda_i^{t_i} \quad (16)$$

donde $x_p(t)$ es una solución particular de (15) C_{s_i} ($i = 1, \dots, k$)

establecido para una constante arbitraria, λ_i es una raíz de la ecuación característica.

$$\{p(\lambda)\} = \{A_0 + \lambda A_1 + \dots + \lambda^{r-1} A_{r-1} + \lambda^r A_r\} = 0 \quad (17)$$

Con multiplicidad w_i satisfaciendo $\sum_{i=1}^k w_i = r \times n$

$$A_l = \begin{bmatrix} P_l^{(11)} & \dots & P_l^{(1n)} \\ \vdots & & \vdots \\ P_l^{(n1)} & \dots & P_l^{(nn)} \end{bmatrix}; \quad (l = 0, 1, \dots, r)$$

y v_{s_i} ($s_i = 0, 1, \dots, w_{i-1}$) establecido por la independencia

lineal de los eigenvectores de A definida abajo asociada con λ_i ;

es decir cuyos vectores v satisfacen

$$\{\lambda_i I - A\} v = 0$$

en donde A esta definida como.

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & & & \ddots & \\ \vdots & & & & \vdots \\ 0 & & & & 1 \\ A_0 & A_1 & \dots & A_{r-1} & \end{bmatrix} \quad (18)$$

con $A_i = -A_r^{-1} A_i$ ($i = 0, 1, \dots, r-1$) sobre la suposición que A_r es no singular. Claramente, el sistema (15) es igual a

$$a_0 x(t) + a_1 x(t+1) + a_2 x(t+2) + \dots + a_r x(t+r) = bt \quad (15')$$

como un ejemplo del sistema (15), considere el siguiente sistema Keynesiano (19) donde: C: consumo; V: Inversión; F: otra demanda efectiva; Y: Ingreso nacional y R: la razón del interés (o proporción del); y subscripto t designado a un período de tiempo.

$$\begin{aligned} C_t &= 0.4Y_t + 0.3C_{t-1} \\ V_t &= 0.1Y_t + 0.5(Y_{t-1} - Y_{t-2}) - 2Rt \\ Y_t &= C_t + V_t + F_t \end{aligned} \quad (19)$$

Este sistema de ecuaciones puede ser escrito como:

$$\begin{bmatrix} 1 & 0 & -.4 \\ 0 & 1 & -.1 \\ -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} C_t \\ V_t \\ Y_t \end{bmatrix} = \begin{bmatrix} .3 & 0 & 0 \\ 0 & 0 & .3 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} C_{t-1} \\ V_{t-1} \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -.5 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} C_{t-2} \\ V_{t-2} \\ Y_{t-2} \end{bmatrix} \\ + \begin{bmatrix} 0 & 0 \\ -2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} R_t \\ F_t \end{bmatrix}$$

El cual es de la forma de (15') con un tiempo lento de dos periodos. Premultiplicando este sistema por la inversa de la matriz de coeficientes de lado izquierdo obtenemos.

$$x(t) = A_1 x(t-1) + A_0 x(t-2) + N_z(t) \quad (20)$$

donde

$$x(t) = \begin{bmatrix} C_t \\ V_t \\ Y_t \end{bmatrix}, A_1 = \begin{bmatrix} .54 & 0 & .4 \\ .06 & 0 & .6 \\ .6 & 0 & 1 \end{bmatrix}, A_0 = \begin{bmatrix} 0 & 0 & -.4 \\ 0 & 0 & -.6 \\ 0 & 0 & -1 \end{bmatrix},$$

$$N = \begin{bmatrix} -1 & .6 & .8 \\ -2 & .4 & .2 \\ -4 & & 2 \end{bmatrix}, Z(t) = \begin{bmatrix} R_t \\ F_t \end{bmatrix}$$

$$\text{Definiendo } x(t) = \begin{bmatrix} x(t-1) \\ x(t) \end{bmatrix}, A = \begin{bmatrix} 0 & I \\ A_0 & A_1 \end{bmatrix},$$

$$w = \begin{bmatrix} 0 \\ N \end{bmatrix}, \text{ y } Z(t) = \begin{bmatrix} 0 \\ Z(t) \end{bmatrix}$$

Transformamos (20) en su forma canónica

$$x(t) = Ax(t-1) + wz(t) \quad (21)$$

Una sustitución iterativa representada sobre (21) produce.

$$x(t) = A^t x(0) + \sum_{s=0}^{t-1} A^s wz(t-s) \quad (22)$$

Así obtenemos los efectos del multiplicador dinámico sobre x de z como sigue

$$\frac{\delta x(t)}{\delta z(t-s)} = A^s w \quad \text{para } s = 0, 1, 2, \dots, t-1$$

o equivalentemente los efectos del multiplicador de x de z son:

$$\frac{\delta x(0)}{\delta z(0)} = \begin{bmatrix} \delta C_0 / \delta R_0 & \delta C_0 / \delta F_0 \\ \delta V_0 / \delta R_0 & \delta V_0 / \delta F_0 \\ \delta Y_0 / \delta R_0 & \delta Y_0 / \delta F_0 \end{bmatrix} = N \text{ El multiplicador de impacto}$$

$$\frac{\delta x(1)}{\delta z(0)} = A_1 N \text{ (el multiplicador de un período lento)}$$

$$\frac{\delta x(2)}{\delta z(0)} = (A_0 + A_1^2) N \text{ (el multiplicador retardado de dos períodos)}$$

$$\frac{\delta x(3)}{\delta z(0)} = A_0 \frac{\delta x(1)}{\delta z(0)} + A_1 \frac{\delta x(2)}{\delta z(0)} \text{ (el multiplicador retardado de tres tiempos)}$$

⋮

$$\frac{\delta x(t)}{\delta z(0)} = A_0 \frac{\delta x(t-2)}{\delta z(0)} + A_1 \frac{\delta x(t-1)}{\delta z(0)} \text{ (el multiplicador de t-ésimo período de retardo)}$$

Si la variable autónoma del vector $z(t)$ permanece constante sobre el tiempo t , el efecto del multiplicador de acumulación de $z(0)$ sobre $x(t)$ sería:

$$[I + A + A^2 + \dots + A^{t-1}] = [I - A]^{-1} [I - A^t] w \quad (23)$$

Sean las series en (23) convergentes o no lo cual se vera en seguida.

Dado λ sea un eigenvalor de una matriz A de $n \times n$. Entonces λ^q es un eigenvalor de A^q donde q es establecida para un entero positivo arbitrario. Denotado por b_{ij} el (i, j) ésimo elemento de A^q . Ya que

$$|\lambda|^q = |\lambda^q| \leq \max \sum_{i=1}^n |b_{ij}|$$

Si la matriz A^q tiende a cero cuando q se aproxima al infinito, cada eigenvalor de A sería menor que la unidad en modulo.

El comentario también sería verificado como se establece abajo.
Para una matriz A cuadrada donde M es una matriz no singular tal

que $J = M^{-1}AM$, toma la forma canónica de Jordan es decir J es particionada en la diagonal en bloques.

$$J_1 = \begin{bmatrix} \lambda_1 & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & \dots & \dots & \lambda_1 \end{bmatrix} = \lambda_1 I + E_1 \quad (24)$$

|-----|
w₁

y cero abajo de la diagonal de los bloques, donde λ_i denota un eigenvalor de A con multiplicidad $w_i (i = 1, \dots, k)$.

$$y E_1 = \begin{bmatrix} 0 & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & \dots & \dots & 0 \end{bmatrix}$$

|-----|
w₁

Notamos que algunos de los unos en J_1 y E_1 serian ceros, pero esto no influencia la secuencia del analisis en algo sustancioso conforme a la transformación de A en J, el sistema (22) sería:

$$X^*(t) = J^t X^*(0) + \sum_{s=0}^{t-1} J^s Z^*(t-s) \quad (25)$$

donde $X^* = M^{-1}X$ y $Z^* = M^{-1}WZ$. Note que J^t tiene la diagonal en bloques J_1^t y cero abajo de los bloques de la diagonal. En vista de (24).

$$J_1^t = \lambda_1^t I + \lambda_1^{t-1} E_1 + 1/2t (t-1) \lambda_1^{t-2} E_1^2 + 1/3!t (t-1)(t-2) \lambda_1^{t-3} E_1^3 +$$

$$\dots + 1/(w_1 - 1)! f_1(t) E_1^{w_1-1},$$

$$\text{donde } f_1(t) = \sum_{s_1=0}^{w_1-2} (t-s_1) \lambda_1^{t-w_1+s_1},$$

por ejemplo si $w_1 = 3$

$$J_1^t = \begin{bmatrix} \lambda_1^t & t\lambda_1^{t-1} & (1/2)t(t-1)\lambda_1^{t-2} \\ 0 & \lambda_1^t & t\lambda_1^{t-1} \\ 0 & 0 & \lambda_1^t \end{bmatrix}$$

Ya que:

$$\frac{f_1(t+1)}{f_1(t)} = \frac{1 + 1/t}{1 - (w_1-2)/t} \lambda_1, \dots \frac{|f_1(t+1)|}{|f_1(t)|} \text{ tiende a } |\lambda_1| \text{ cuando}$$

t se aproxima a infinito.

Así, si $|\lambda_1| < 1$, donde existe un entero t_0 tal que $t \geq t_0$,

$$\frac{|f_1(t+1)|}{|f_1(t)|} < \delta, \text{ o equivalentemente } |f_1(t+1)| < \delta |f_1(t)|,$$

donde $\delta \in (|\lambda_1|, 1)$.

En general, para un entero positivo τ obtenemos

$$|f_1(t + \tau)| < \delta^\tau |f_1(t)| \text{ para } t \geq t_0 \text{ ya que } 0 < \delta < 1,$$

$|f_1(t + \tau)|$ tiende a cero cuando τ se incrementa. Por lo tanto, cuando λ_1 es, más pequeña que la unidad en modulo, $f_1(t)$ tiende a cero y aquí J_1^t la matriz converge a cero cuando t se aproxima a infinito.

Si $|\lambda_i| < 1$ para cada i , J^t y aquí A^t la matriz converge

a cero ya que $A^t = MJ^tM^{-1}$. Por el análisis anterior, obtenemos un teorema sobre estabilidad de matrices.

Teorema dado A sea una matriz cuadrada real.

Entonces el $\lim_{t \rightarrow \infty} A^t = 0$ si y solamente si cada eigenvalor de A

sea una matriz cuadrada real, las series:

$I + A + A^2 + A^3 + \dots$ convergen a $(I - A)^{-1}$ si y solamente si cada eigenvalor de A es menor que el modulo de la unidad.

Recordando, si el modulo de cada eigenvalor de A es menor que la unidad con un $Z(t)$ fijo en (21) para toda t , dado Z_0 , el nivel de $x(t)$ converge en equilibrio.

$$[I - A]^{-1} W Z_0 \quad (26)$$

El cual sería llamado multiplicador de equilibrio largo. En un similar sentido el sistema (15) de ecuación lineal en diferencias tiende a estar estable si cada raíz característica de (17) es menor que el modulo de la unidad.

MODELO DE COMPETENCIA ENTRE TRES LECHERIAS

$$A = \begin{bmatrix} 0.8 & 0.2 & 0.1 \\ 0.1 & 0.7 & 0.3 \\ 0.1 & 0.1 & 0.6 \end{bmatrix}$$

$$A - \lambda I = \begin{bmatrix} 0.8 - \lambda & 0.2 & 0.1 \\ 0.1 & 0.7 - \lambda & 0.3 \\ 0.1 & 0.1 & 0.6 - \lambda \end{bmatrix}$$

Para que $A - \lambda I$ sea singular es necesario que $\det(A - \lambda I) = 0$.

Por el cálculo directo se obtiene:

$$\begin{aligned} \det(A - \lambda I) &= \lambda^3 + 2.1\lambda^2 - 1.4\lambda + 0.3 \\ &= -(\lambda - 0.5)(\lambda - 0.6)(\lambda - 1.0) \end{aligned}$$

y entonces los eigenvalores λ deben ser

$$\lambda = 0.5, \lambda = 0.6, \lambda = 1.0$$

Para encontrar los eigenvalores asociados considere, por ejemplo la ecuación $(A - \lambda I)x = 0$ por $\lambda = 0.5$ en término de los elementos x_1 , x_2 , y x_3 de x esto es

$$0.3x_1 + 0.2x_2 + 0.1x_3 = 0$$

$$0.1x_1 + 0.2x_2 + 0.3x_3 = 0$$

$$0.1x_1 + 0.1x_2 + 0.1x_3 = 0$$

La técnica acostumbrada por los sistemas lineales de ecuaciones reducen lo anterior a

$$x_1 + 0x_2 - x_3 = 0$$

$$0x_1 + x_2 + 2x_3 = 0$$

$$0x_1 + 0x_2 + 0x_3 = 0$$

de aquí es posible hacer que x_3 sea un α arbitrario,

$$x_1 = \alpha, \text{ y } x_2 = -2\alpha \text{ y entonces } x = \alpha \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \text{ es un eigenvector}$$

asociado con el eigenvalor $\lambda = 0.5$ para todo $\alpha \neq 0$. Por métodos semejantes se encontrara que existen múltiplos arbitrarios de los vectores :

$$\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 9 \\ 7 \\ -1 \end{bmatrix}; \text{ son eigenvectores asociados con los eigenvalores}$$

$\lambda = 0.6$ y $\lambda = 1.0$ respectivamente, los tres subespacios unidimensionales de \mathbb{R}^3 , cada uno generado por uno de esos tres eigenvectores, son, de este modo subespacios invariantes de A.

$$Q = \begin{bmatrix} 1 & 1 & 9 \\ -2 & -1 & 7 \\ 1 & 0 & -1 \end{bmatrix}; \quad Q^{-1} = \begin{bmatrix} 0.07 & 0.07 & 1.07 \\ 0.33 & -0.67 & -1.67 \\ 0.07 & 0.07 & 0.07 \end{bmatrix}$$

$$Q^{-1}AQ = J.$$

$$\begin{bmatrix} 0.07 & 0.07 & 1.07 \\ 0.33 & -0.67 & -1.67 \\ 0.07 & 0.07 & 0.07 \end{bmatrix} \begin{bmatrix} 0.8 & 0.2 & 0.1 \\ 0.1 & 0.7 & 0.3 \\ 0.1 & 0.1 & 0.6 \end{bmatrix} \begin{bmatrix} 1 & 1 & 9 \\ -2 & -1 & 7 \\ 1 & 0 & -1 \end{bmatrix} = J$$

$$J = \begin{bmatrix} 0.5 & 0 & 2 \\ 0 & 0.60 & -2.5 \\ 0 & 0 & 1.0 \end{bmatrix}$$

Forma una base para \mathbb{R}^3 , de tal manera que:

$$x_0 = a_1 v_1 + a_2 v_2 + a_3 v_3$$

$$\text{Entonces } A^k x_0 = a_1 (0.5)^k v_1 + a_2 (0.6)^k v_2 + a_3 (1.0)^k v_3$$

lo que demuestra que $A^k x_0$ converge a un múltiplo de $a_3 v_3$ para todo x_0 .

Ejemplo 7 los modelos de crecimiento de población surgieron esencialmente de la hipótesis de que el crecimiento es proporcional a la población presente:

$$p_{t+1} - p_t = (b - d)p_t$$

El modelo continuo análogo para pequeños intervalos de tiempo o para poblaciones con rápido crecimiento usaría la derivada con respecto al tiempo \dot{p} :

$$p'(t) = (\beta - \delta)p(t)$$

En donde β y δ reflejan las tasas de nacimientos y muertes para las poblaciones en competencia como las del modelo gallinas - zorros.

$$F'(t) = -0.4 F(t) + 0.5 C(t), F(t_0) = 100$$

$$C'(t) = -k F(t) + 0.2 C(t), C(t_0) = 1000$$

$$O, \text{ en notación matricial } x' = Ax, x(t_0) = x_0$$

En donde

$$x = \begin{bmatrix} F \\ C \end{bmatrix}, A = \begin{bmatrix} -0.4 & 0.5 \\ -k & 0.2 \end{bmatrix}, x_0 = \begin{bmatrix} 100 \\ 1000 \end{bmatrix}$$

Si $k = 0.16$ en este modelo de competencia entre poblaciones obtenemos la matriz A :

$$A = \begin{bmatrix} -0.4 & 0.5 \\ -0.16 & 0.2 \end{bmatrix}$$

que tiene eigenvalores $\lambda_1 = 0$, $\lambda_2 = -0.2$, y se reduce a la forma

diagonal de Jordan por :

$$P = \begin{bmatrix} 5 & 5 \\ 4 & 2 \end{bmatrix}, P^{-1} = \begin{bmatrix} -0.2 & 0.5 \\ 0.4 & -0.5 \end{bmatrix}$$

$$P^{-1}AP = J.$$

Los elementos diagonales de $J(t)$ son entonces $\exp(0t) = 1$ y $\exp(-0.2t)$ de lo cual obtenemos:

$$x(t) = [F(t) \ C(t)]^t$$

en donde

$$F(t) = 2400 - 2300 \exp(-0.2t) \text{ y}$$

$$C(t) = 1920 - 920 \exp(-0.2t).$$

Conforme pasa el tiempo las poblaciones tienden a valores estables de 2400 y 1920.

Si ahora $k = 0.18$ la matriz A toma la forma :

$$A = \begin{bmatrix} -0.4 & 0.5 \\ -0.18 & 0.2 \end{bmatrix}$$

tiene un eigenvalor doble $\lambda_1 = -0.1$ y todo eigenvector es múltiplo

de $\begin{bmatrix} 5 & 3 \\ 1 & 1 \end{bmatrix}$. Esto significa que A tiene un bloque de Jordan 2×2 en su forma de Jordan que se encuentra fácilmente:

$$J = P^{-1}AP = \begin{bmatrix} 0.02 & 0.3 \\ -0.06 & 0.1 \end{bmatrix} \begin{bmatrix} -0.40 & 0.5 \\ -0.18 & 0.2 \end{bmatrix} \begin{bmatrix} 5 & -15 \\ 3 & 1 \end{bmatrix} =$$

$$= \begin{bmatrix} -0.1 & 1 \\ 0 & -0.1 \end{bmatrix}; \quad x(t) = [F(t) C(t)]^T, \text{ en donde}$$

$$F(t) = (100 + 470t) \exp(-0.1t),$$

$$C(t) = (1000 + 282t) \exp(-0.1t).$$

Ambas poblaciones mueren cuando t tiende a infinito.

Ejemplo: una matriz A 2×2 y su forma de Jordan

$$J = \begin{bmatrix} -0.1 & 1 \\ 0 & -0.1 \end{bmatrix} \text{ para } A = \begin{bmatrix} -0.40 & 0.5 \\ -0.18 & 0.2 \end{bmatrix}$$

Usando P y P^{-1} resulta

$$\exp(tA) = \exp(-0.1t) \begin{bmatrix} 1 - 0.3t & 0.5t \\ -0.18t & 1 + 0.3t \end{bmatrix};$$

haciendo $t = 1$, se obtiene

$$\exp(A) = \exp(1A) = \exp(-0.1) \begin{bmatrix} 0.7 & 0.5 \\ -0.18 & 1.3 \end{bmatrix}.$$

Una vez encontradas las soluciones de las ecuaciones diferenciales estudiadas en los ejemplos anteriores se podría ver inmediatamente cómo se comportaron cuando t tendía a infinito.

Se pueden escribir esas soluciones como $x(t) = \exp(t\lambda)x_0$ esos, resultados deberían ser consecuencia de las propiedades de A y de $\exp(tA)$. La fórmula explícita para $\exp(tA)$ da una prueba inmediata de esos resultados una vez que se ha entendido el comportamiento de $\exp(\lambda t)$ para una λ posiblemente compleja.

Si $\lambda = \alpha + \beta i$ con α y β reales, entonces la serie infinita para

$\exp(\lambda t)$ se divide en la parte real y la parte imaginaria y resulta $\exp((\alpha + \beta i)t) = \exp(\alpha t) (\cos \beta t + i \sin \beta t)$. Si α , la parte real de λ es negativa, entonces $\exp(\lambda t)$ tiende a cero a medida que t tiende a infinito; si la parte real de λ es positiva, $|\exp(\lambda t)|$ tiende a infinito, y la parte real es cero, $\exp(\lambda t)$ está acotada.

Ejemplo 8. suponga que $v(t)$ es el ingreso total en el tiempo t de una empresa dada (o quizás el PIB de una economía), y suponga también que, si $i(t)$ de este ingreso se reinvierte, la tasa de crecimiento de v será proporcional a i de modo que:

$$v'(t) = g i(t)$$

para determinada razón de crecimiento g . Se propone reinvertir una parte fija r de los ingresos, de modo que $i(t) = r v(t)$, y se supone que el resto del ingreso, $(1 - r)v(t)$, deberá ser distribuido como ganancias entre los accionistas después de que se han deducido impuestos y otros gastos, de modo de que la tasa de crecimiento de la ganancia o utilidad p sea proporcional a $(1 - r)v$, esto es,

$$p' = s(1 - r)v.$$

Entonces, en resumen, se obtiene el modelo

$$v' = g \cdot r \cdot v,$$

$$p' = s(1 - r)v.$$

o en notación matricial,

$$x = \begin{bmatrix} v \\ p \end{bmatrix}, \quad A = \begin{bmatrix} gr & 0 \\ s(1 - r) & 0 \end{bmatrix}$$

$$A = \begin{bmatrix} (0.08)(0.2) & 0 \\ 0.7(1 - 0.2) & 0.2 \end{bmatrix}$$

con $v(0) = 100$ y $p(0) = 5$

Aquí, g , r y s son constantes.

CONCLUSIONES

En las aplicaciones del Algebra Lineal y la Teoría de Matrices juegan un papel muy importante puesto que estos dos entes se conjugan de una manera tan extraordinaria que una de tantas utilidades que se les puede dar en este caso es: la Matriz en su Forma Canónica de Jordan.

Las matrices que no son diagonalizables o sea que no tienen n vectores linealmente independientes, suelen aparecer en ciertas aplicaciones. Para lo cual todavía es posible mostrar que la matriz es equivalente a otra más simple, pero la matriz ya no es diagonal y la matriz que se desea obtener es un poco más difícil.

La matriz J se llama Forma Canónica de Jordan de A .

1.- Si A es diagonalizable, entonces:

$$J = D \text{ diagonal } (\lambda_1, \lambda_2, \dots, \lambda_n)$$

cuando $\lambda_1, \lambda_2, \dots, \lambda_n$ son los valores característicos de A (no es

necesariamente distintos). Cada componente de la diagonal es una matriz bloque de Jordan de 1×1 .

2.- Dada una matriz que no es diagonalizable se puede llevar mediante una transformación de semejanza a una Forma Canónica de Jordan para poder analizarla.

Dada una matriz cuadrática A , queremos escoger M tal que $M^{-1}AM$ sea lo más diagonalizable.

En el caso más sencillo, A tiene un conjunto completo de vectores propios que serán las columnas de M (se conoce como S).

La Forma de Jordan es $J = M^{-1}AM = \Lambda$, se construye completamente a partir de los bloques de 1 por 1 por 1 $J_i = \lambda_i$ se alcanza

el objetivo de obtener una matriz totalmente diagonal. En el caso más general y más difícil, faltan algunos vectores propios y es imposible la forma diagonal. Ese caso es ahora nuestro objetivo.

Teorema que se pretende demostrar:

Si una matriz S vectores propios linealmente independientes, entonces es similar a una matriz que está en la forma de Jordan especial.

$$J = M^{-1}AM = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_s \end{bmatrix}$$

con bloques $J_1 = \begin{bmatrix} \lambda_1 & 1 & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_1 \end{bmatrix}$

Un ejemplo de dicha matriz es

$$J = \begin{bmatrix} 8 & 1 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 8 & 1 \\ 0 & 8 \end{bmatrix} & & & \\ & \begin{bmatrix} 8 & 1 \\ 0 & 8 \end{bmatrix} & & \\ & & [0] & \\ & & & [0] \end{bmatrix} = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & J_3 & \\ & & & \end{bmatrix}$$

El vector propio doble $\lambda = 8$ tiene un solo vector propio, en dirección de la primera coordenada $e_1 = (1, 0, 0, 0, 0)^T$; esto hace que $\lambda = 8$ aparezca solamente en un bloque J_1 .

El valor propio Triple $\lambda = 0$ tiene dos vectores propios e_3 y e_5 , los cuales corresponden a los bloques de Jordan J_2 y J_3 .

La cuestión clave es: Si A es alguna otra matriz de S por S , ¿ bajo que condiciones su forma de Jordan sera la misma matriz J ?

¿ Cuándo existirá alguna M tal que $M^{-1}AM = J$?

Como primera condición, cualquier matriz similar A debe compartir los mismos valores propios $8, 8, 0, 0, 0$.

Pero esto esta lejos de ser suficiente (la matriz diagonal con estos valores propios no es similar a J), y nuestra pregunta está realmente relacionada con los vectores propios.

Para contestala reescribimos la relación $M^{-1}AM = J$ en la forma simple $AM = MJ$

$$A \begin{bmatrix} x_1 & x_2 & \dots & x_5 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & \dots & x_5 \end{bmatrix} \begin{bmatrix} 8 & 1 & & & \\ 0 & 8 & & & \\ & & 0 & 1 & \\ & & & 0 & 0 \\ & & & & 0 \end{bmatrix}$$

Realizando las multiplicaciones una columna a la vez:

$$Ax_1 = 8x_1 \text{ y } Ax_2 = 8x_2 + x_1 \tag{1}$$

$$Ax_3 = 0x_3 \text{ y } Ax_4 = 0x_4 + x_3 ; Ax_5 = 0x_5 \tag{2}$$

Podemos ahora reconocer las condiciones en A.

Debe tener tres vectores propios autenticos, tal como los tiene J. Aquél con $\lambda = 8$ irá en la primera columna de M, exactamente como hubiera ido en la primera columna de S: $Ax_1 = 8x_1$.

Los otros dos, que llamaremos x_3 y x_5 , van en la tercera y quinta columna de M: $Ax_3 = Ax_5 = 0$.

Finalmente, debe otros dos vectores especiales, los [vectores propios generalizados] x_2 y x_4 .

Consideramos x_2 como elemento de una hilera de vectores encabezados por x_1 y descritos por la ecuación (1). De hecho x_2 es el único vector adicional en la hilera, y el bloque correspondiente J_1 es de orden dos. La ecuación (2) describe dos hileras diferentes, una en la que x_4 sigue a x_3 y la otra en la que x_5 está solo; los bloques J_2 y J_3 son de 2 por 2 y de 1 por 1.

La búsqueda de la Forma de Jordan de A se convierte en la búsqueda de estas hileras de vectores, cada una encabezada por algún vector propio: para cada i, o

$$Ax_i = \lambda_i x_i \quad \text{ó} \quad Ax_i = \lambda_i x_i + x_{i-1} \quad (3)$$

Los vectores x_i van en las columnas de M, y cada hilera produce un solo bloque en J. Esencialmente, tenemos que mostrar de qué manera pueden construirse estas hileras para cada matriz A.

Entonces, si las hileras corresponden a las ecuaciones particulares (1) y (2), nuestra J será la Forma de Jordan de A.

Existen tres pasos.

Después de dar una descripción general los aplicaremos a un ejemplo específico.

Paso 1. Si suponemos que A es singular, entonces su espacio columna tiene dimensión $r < n$. En lo que respecta solamente a este espacio pequeño, la hipótesis de inducción garantiza que una Forma de Jordan es posible; debe haber r vectores independientes w_i en el

espacio columna tales que:

$$A w_i = \lambda_i w_i \quad \text{ó} \quad A w_i = \lambda_i w_i + w_{i-1} \quad (4)$$

Paso 2. Supongamos que el espacio nulo y el espacio columna de A tienen una intersección de dimensión P.

Desde luego, cada vector del espacio nulo es un vector propio correspondiente a $\lambda = 0$. Por lo tanto, debe haber p hileras en el paso 1 que comienzan a partir de este valor propio, y nos interesan los vectores w_i que tienen al final de estas hileras.

Cada uno de estos vectores está en el espacio columna, así que cada uno es una combinación de las columnas de A: $w_i = Ay_i$ para

alguna y_i .

Paso 3. El espacio nulo siempre tiene dimensión $n - r$. Por lo tanto, independiente de su intersección p - dimensional con el espacio columna, debe contener $n - r - p$ vectores básicos adicionales z_i fuera de esa intersección.

Juntemos estos pasos para obtener el Teorema de Jordan.

Los r vectores w_i , los p vectores y_i y los $n - r - p$ vectores

z_i forman las hileras de Jordan para la matriz A: y estos vectores

son linealmente independientes. Van en las columnas de:

M y $J = M^{-1}AM$ esta en la Forma de Jordan.

El resultado de un esfuerzo supremo por diagonalizar una matriz A de $n \times n$ es la Forma Canónica de Jordan.

Tabla de Transformación de Similitud.

Para matrices A de $n \times n$.

1.- A es diagonalizable: las columnas de S son los vectores propios de A y $S^{-1}AS = \Lambda$ es Diagonal.

2.- A es arbitraria: las columnas de M son los vectores propios y los vectores propios generalizados de A, y la Forma de Jordan $M^{-1}AM = J$ es diagonal por bloque.

3.- A es arbitraria y U es unitaria; podemos escoger U tal que $U^{-1}AU = T$ sea triangular superior.

4.- A es normal $AA^H = A^HA$: podemos elegir U tal que $U^{-1}AU = \Lambda$
Casos especiales, todos con vectores propios ortonormales.

- a : Si A es Hermitiana, entonces Λ es real.
 a' : Si A es real y simétrica, entonces Λ es real y $U = Q$ es ortonormal.
 b : Si A es Anti-Hermitiana, entonces Λ es imaginaria.
 c : Si A es unitaria, entonces todas los $|\lambda_i| = 1$

Sea la matriz A de 2×2 con un valor característico λ de multiplicidad algebraica 2 y multiplicidad geométrica correspondiente a λ . Entonces existe un vector V_2 que satisface la

la ecuación:

$$(A - \lambda I) v_2 = v_1 \quad (1)$$

Demostración:

Sea $x \in C^2$ un vector constante que no sea múltiplo de v_1 por lo cual x no es un vector característico de A. Primero mostraremos que:

$$w = (A - \lambda I)x \quad (2)$$

es un vector característico de A. Esto es, demostraremos que

$w = cv_1$ para alguna constante c. Ya que $w \in C^2$ y v_1, x son linealmente independientes, existen constantes c_1, c_2 tales que

$$w = c_1 v_1 + c_2 x \quad (3)$$

Para mostrar que w es un vector característico de A, debemos mostrar que $c_2 = 0$. De (2) y (3) encontramos que:

$$(A - \lambda I)x = c_1 v_1 + c_2 x \quad (4)$$

Sea $B = A - (\lambda + c_2)I$ entonces (4).

$$Bx = [A - (\lambda + c_2)I]x = c_1 v_1 \quad (5)$$

si suponemos que c_2 diferente de cero entonces $\lambda + c_2 \neq \lambda$ y $\lambda + c_2$

no es un valor característico de A (ya que λ es el único valor característico de A). Se tiene de esta forma que

$\det B = \det [A - (\lambda + c_2)I] \neq 0$ lo cual significa que B es

invertible.

De aquí que (5) puede escribirse como

$$x = B^{-1}c_1v_1 = c_1B^{-1}v_1 \quad (6)$$

Entonces multiplicando ambos lados de (6) por λ ,

$$\lambda x = \lambda c_1 B^{-1}v_1 = c_1 B^{-1}\lambda v_1 = c_1 B^{-1}Av_1 \quad (7)$$

Pero $B = A - (\lambda + c_2)I$ de manera que

$$A = B + (\lambda + c_2)I \quad (8)$$

Sustituyendo (8) en (7)

$$\begin{aligned} \lambda x &= c_1 B^{-1}[B + (\lambda + c_2)I]v_1 = c_1[I + (\lambda + c_2)B^{-1}]v_1 \\ &= c_1v_1 + (\lambda + c_2)c_1B^{-1}v_1 \end{aligned} \quad (9)$$

Pero usando de nuevo (5), $c_1B^{-1}v_1 = x$, por lo que (9) se convierte en:

$$\begin{aligned} \lambda x &= c_1v_1 + (\lambda + c_2)x = c_1v_1 + c_2x + \lambda x \\ \text{o bien } 0 &= c_1v_1 + c_2x \end{aligned} \quad (10)$$

Pero v_1 y x son linealmente independientes, así $c_1 = c_2 = 0$

Esto contradice la hipótesis de que $c_2 \neq 0$. De esta manera

$c_2 = 0$ y, por (3), w es un múltiplo de v_1 por lo que $w = c_1v_1$ es un vector característico de A .

Más aún, $w \neq 0$ ya que si $w = 0$ entonces (2) nos dice que x es un vector característico de A por consiguiente $c_1 \neq 0$.

$$\text{Sea } v_2 = \frac{1}{c_1}x \quad (11)$$

$$\text{Entonces } (A - \lambda I)v_2 = \left(\frac{1}{c_1}\right)(A - \lambda I)x = \left(\frac{1}{c_1}\right)w = v_1$$

Definición de: vector característico generalizado.

Sea A una matriz de 2×2 con el único valor característico λ con multiplicidad geométrica 1. Sea v_1 característico de A .

Entonces el vector v_2 definido por $(A - \lambda I)v_2 = v_1$ se llama vector característico generalizado de A correspondiente al valor característico de λ .

Ejemplo sea $A = \begin{bmatrix} 3 & -2 \\ 8 & 5 \end{bmatrix}$ la ecuación característico de A es

$\lambda^2 + 2\lambda + 1 = (\lambda + 1)^2 = 0$ en lo que $\lambda = -1$ es un valor característico de multiplicidad algebraica 2. Entonces:

$$(A - \lambda I)v = (A + I)v = \begin{bmatrix} 4 & -2 \\ 8 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

De aquí se obtiene el vector característico $v_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

No existe otro vector característico linealmente independiente. Para encontrar un vector característico v_2 calculamos

$(A + I)v_2 = v_1$ o bien $\begin{bmatrix} 4 & -2 \\ 8 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ lo cual nos da el sistema

$$\begin{aligned} 4x_1 - 2x_2 &= 1 \\ 8x_1 - 4x_2 &= 2 \end{aligned}$$

La segunda ecuación es el doble de la primera, por lo que x_2

puede escogerse arbitrariamente y $x_1 = \frac{(1 + 2x_2)}{4}$

Por consiguiente una posible elección de v_2 es

$$v_2 = \begin{bmatrix} 1/4 \\ 0 \end{bmatrix}.$$

Teorema: Sea A, λ , v_1 y v_2 como en el teorema anterior y sea C la matriz cuya columna son v_1 y v_2 , entonces:

$C^{-1}AC = J$ donde $J = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$ es la Forma Canónica de Jordan de A.

Demostración:

Puesto que v_1 y v_2 son linealmente independientes, vemos que C es invertible.

Ahora nótese que $AC = A(v_1, v_2) = (Av_1, Av_2) = (\lambda v_1, Av_2)$ pero por la ecuación (1), $Av_2 = v_1 + \lambda v_2$ de modo que $AC = (\lambda v_1, v_1 + \lambda v_2)$, asimismo

$$CJ = (v_1, v_2) \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} = (\lambda v_1, v_1 + \lambda v_2). \text{ De esta manera}$$

$AC = CJ$, lo cual significa que $C^{-1}AC = J$.

Del ejemplo anterior

$$v_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \text{ y } v_2 = \begin{bmatrix} 1/4 \\ 0 \end{bmatrix}, \text{ entonces } C = \begin{bmatrix} 1 & 1/4 \\ 2 & 0 \end{bmatrix},$$

$$C^{-1} = -2 \begin{bmatrix} 0 & -1/4 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1/2 \\ 4 & -2 \end{bmatrix}, \text{ y}$$

$$\begin{aligned} C^{-1}AC &= \begin{bmatrix} 0 & 1/2 \\ 4 & -2 \end{bmatrix} \begin{bmatrix} 3 & -2 \\ 8 & 5 \end{bmatrix} \begin{bmatrix} 1 & 1/4 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1/2 \\ 4 & -2 \end{bmatrix} \begin{bmatrix} -1 & 3/4 \\ -2 & 2 \end{bmatrix} \\ &= \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix} = J. \end{aligned}$$

El método antes descrito puede generalizarse para obtener la Forma Canónica de Jordan de cada matriz. Aunque no se demuestra aquí este resultado siempre es posible determinar el número de unos arriba de la diagonal en la Forma Canónica de Jordan de una matriz de $n \times n$. Sea λ_i un valor característico de A con multiplicidad

algebraica r_i y multiplicidad geométrica s_i , $s_1, \lambda_1, \lambda_2, \dots, \lambda_k$

son los valores característicos de A .

El número de unos arriba de la diagonal de la Forma Canónica de Jordan de A ,

$$= (r_1 - s_1) + (r_2 - s_2) + \dots + (r_k - s_k)$$

$$= \sum_{i=1}^k r_i - \sum_{i=1}^k s_i = n - \sum_{i=1}^k s_i$$

Si conocemos la ecuación característica de una matriz A , entonces podemos determinar las posibles Formas Canónicas de Jordan de A .

Aplicación: Evolución de Sistemas Continuos y Exponenciales de Matrices.

El valor analítico de la Forma de Jordan se desarrolla para ver la evolución de sistemas generales cuyo estado en un punto específico del tiempo se da como una transformación lineal del estado en un tiempo específico anterior.

Si el intervalo de tiempo en cuestión es extremadamente pequeño o si el cambio en el estado es extremadamente pequeño en relación con el estado, a menudo es útil suponer que el estado x se define para todo tiempo t por la función $x(t)$ y que la información

de los cambios de estado se da en términos de la derivada x' en

donde x' es

$$x'(t) = \begin{bmatrix} \frac{dx_1}{dt}(t) \\ \frac{dx_2}{dt}(t) \\ \cdot \\ \cdot \\ \frac{dx_p}{dt}(t) \end{bmatrix}$$

Si x es $p \times 1$. En otros casos, particularmente en las ciencias físicas y en la Ingeniería, se considera al tiempo como una variable muestreada continuamente de modo que las descripciones naturales de la evolución del estado x de un sistema usan la derivada x' y ecuaciones diferenciales ordinarias.

Se dieron algunos ejemplos para ilustrar la forma en que algunos sistemas de ecuaciones diferenciales ordinarias que surgen en las aplicaciones se pueden expresar en una forma estándar

$$x' = Ax + f.$$

Como las matrices que no son diagonalizables suelen darse en las aplicaciones en la realidad.

La finalidad de hallar la Forma Canónica de Jordan es que en estas matrices se pueden expresar como

$$e^{At} = \left[I + At + \frac{2t^2}{2!} + A^2 \frac{3t^3}{3!} + \dots \right].$$

Como en los sistemas de ecuaciones diferenciales lineales se encuentra una matriz fundamental a la cual se le obtiene su Wronskiano y debe tener n vectores linealmente independientes la Forma Canónica de Jordan que se obtiene de un sistema de Ecuaciones Diferenciales cuyo matriz no es fundamental o sea que no tiene un conjunto completo de n vectores linealmente independientes se realiza la transformación para obtener sus vectores linealmente independientes generalizados se puede tomar esta matriz de Jordan como la matriz fundamental que sería

$$J = D + N$$

donde D es la diagonal y N es una matriz nulpotente.

Como las matrices que no son diagonalizables suelen darse en las aplicaciones reales.

La finalidad de hallar la Forma Canónica de Jordan es que estas matrices se pueden expresar como:

$$e^{At} = I + A^{n-1} \frac{e^{2t}}{2!} + \dots \text{ para que puedan dar información de}$$

los problemas donde estas aparezcan.

BIBLIOGRAFIA

- 1.- ECONOMIA MATEMATICA, TAKAMA, 1985
- 2.- ESSENTIAL MATEMATICS FOR ECONMICS, 1980
- 3.- ECONOMIC DYNAMICS METHODS AND MODELS
- 4.- MODELOS ECONOMETRICOS, INTRILIGATOR,
- 5.- INTRODUCCION A LAS ECUACIONES DIFERENCIALES ORDINARIAS.
EARL A. CODDINGTON. EDITORIAL CECSA.
- 6.- MATHEMATICS FOR STABILITY AND OF ECONOMICS SYSTEMS.
YASUO MURATA, ECONOMIA MATEMATICA, 1960.
- 7.- LINEAR ALGEBRA AND MATRIZ THEORY, SECOND, EDITION, 1963.
EVAR, D. NERIN G., EDIT. JHON WILEY.
- 8.- DIFFERENTIAL EQUATIONS, DINAMICAL SYSTEMS, DINAMICAL
SYSTEMS AND LINEAR ALGEBRA, HIRSCH, MORRIS.
- 9.- ECUACIONES DIFERENCIALES Y PROBLEMAS CON VALORES A LA
FRONTERA, WILLIAM E. BOYCE, RICHARD C. DIPRINA, 3ERA.
EDICION, 1983, LIMUSA.
- 10.- ALGEBRA LINEAL Y SUS APLICACIONES, GILBERT STRANG ADDISON
- WESLEY IBEROAMERICANA, 1986.
- 11.- ALGEBRA LINEAL APLICADA: BEN NOBLE, JAMES W. DANIEL,
TERCERA EDICION, 1989, EDIT. PRENTICE - HALL
HISpanoAMERICANA, CAP. 9, PAG. 339-394. EIGENSISTEMAS DE
MATRICES ARBITRARIAS GENERALES, CON APLICACIONES.
- 12.- ALGEBRA LINEAL, STANLEY I. GROSSMAN, SEGUNDA EDICION,
1988, EDIT. GRUPO EDITORIAL IBEROAMERICANA, CAP. 6.7 UNA
APLICACION IMPORTANTE: ECUACIONES DIFERENCIALES MATRICIALES
PAG. 362-373.
- 13.- MATEMATICAS PARA ECONOMISTAS, JULIO GRAFE, 2da. EDICION
1991, EDIT, MC GRAW HILL.
- 14.- ECUACIONES DIFERENCIALES Y CALCULO VARIACIONAL, L.
ELSGOLTZ, EDICION 1975, EDICIONES DE CULTURA POPULAR.
PAG.269 PARTE VARIABLES Y FUNCIONES HASTA PAG. 272 LEYES
DINAMICAS.
- 15.- FIRST COURSE IN NUMERICAL METHODS, W. JENNIGS, EDIT. MC.
MILLAR, NUEVA YORK, 1964, PAG. 143-147.

- 16.- ANALISIS MATEMATICO I, II
AUTORES: NORMAN B.HASSER, JOSEPH P. LASALLE y JOSEPH A.
SULLIVAN.
EDITORIAL TRILLAS 1979.
- 17.- INTRODUCCION AL ANALISIS MATRICIAL.
AUTOR: BELLMAN