Identificación de coeficientes para sistemas dinámicos lineales

Rufino Carrada Herrera

22 de marzo de 2006

Índice general

Resumen

La tesis está dedicada al estudio de la identificación simultánea de los coeficientes de un sistema dinámico estable y de las amplitudes de fuerza armónica a través de la respuesta en salida estacionaria (de este sistema dinámico estable). Este problema surge en varias ramas, particularmente en la ingeniería, como mecánica, eléctrica, etc. Estos problemas son el resultado de la modelación de sistemas mecánicos, que requieren de herramientas como álgebra lineal, cálculo de varias variables y del análisis de sistemas de ecuaciones diferenciales, y para poder simplificarlos se analizan solo las propiedades de las soluciones. Se ve sí el sistema dinámico es estable o no, o si cumple ciertas condiciones es linealizado (en el caso de ser un sistema mas complejo), entre otras cosas; ya que no hay manera de resolverlos explícitamente. El problema de identificación ayuda a determinar los parámetros que describen el estado de un sistema dinámico lineal estable o el comportamiento que éste tiene cuando sufre perturbaciones estocásticas.

Capítulo 1

Introducción

La tesis está dedicada al estudio de la identificación simultánea de los coeficientes de un sistema dinámico estable, bajo la acción de una fuerza armónica, y de las amplitudes de la fuerza armónica, a través de las amplitudes de la salida estacionaria. Estos tipos de problemas surgen, en muchas ramas, especialmente en ingeniería mecánica, eléctrica, entre otros. Una aplicación importante de la tesis, es la identificación de los coeficientes del sistema que rige el movimiento de un rotor dinámico, dicho sistema es de segundo orden y la parte derecha de este sistema de ecuaciones diferenciales es una fuerza armónica que satisface las condiciones de resolución, donde la primer componente de dicha fuerza es aleatoria y todas sus demás entradas son conocidas; entonces puede aplicarse directamente el algoritmo establecido en esta tesis para la identificación simultánea de los coeficientes del sistema. Este modelo se considera en los artículos [25] y [4], y la investigación matemática de este problema lo estudia el artículo [24].

El problema de identificación está relacionado con la teoría de control. Se demostró que, la identificación de estos coeficientes es única si se satisface la condición de controlabilidad generalizada. También consideré la influencia de perturbaciones estocásticas en este sistema dinámico lineal estable.

La tesis consta de 5 capítulos, y cada capítulo estudia temas que son necesarios para el estudio de la identificación de coeficientes en los sistemas dinámicos lineales, en el *primer capítulo* se escribe un resumen del contenido en el artículo [24], además escribo las demostraciones de éste en una forma más clara, indicando y precisando con más detalle los resultados de cuales se hacen uso, también agregué un breve recordatorio de las propiedades (y de la definición) de un sistema dinámico; en los capítulos *segundo y tercero* se dan algunas propiedades de las perturbaciones estocásticas de acuerdo con el Teorema de Límite Central, en los capítulos *cuarto y quinto* se

escriben dos nociones importantes de la teoría del control, como son la observabilidad y la controlabilidad para los sistemas dinámicos.

Capítulo 2

Identificación de Coeficientes para Sistemas Dinámicos Lineales

2.1. Definición del Problema

Un sistema dinámico da una descripción funcional de la solucion de un problema físico o de un modelo matemático que rige a un sistema físico. Por ejemplo, el movimiento de un péndulo es un un sistema dinámico en el sentido de que el movimiento del péndulo se describe por su posición y velocidad como funciones del tiempo, y de las condiciones iniciales.

Matemáticamente hablando, un sistema dinámico es una fúnción $\phi(t, x)$, definida para todo $t \in \mathbb{R}$ y $x \in E \subset \mathbb{R}^n$, que describe el movimiento o como se mueven los puntos $x \in E$ con respecto al tiempo.

Un *sistema dinámico sobre* E es una función $\phi : \mathbb{R} \times E \to E$ de clase C^1 donde E es un subconjunto abierto de $\mathbb{R}^n$ y si $\phi_t(x) = \phi(t, x)$ entonces ϕ_t satisface las siguientes propiedades; (1) $\phi_0(x) = x$ para todo $x \in E$, y (2) $\phi_t \circ \phi_s(x) = \phi_{t+s}(x)$ para cada $x \in E$ y para cada $t, s \in \mathbb{R}$. Por ejemplo, si A es una matriz $n \times n$ entonces la función $\phi(t, x) := e^{At}x$ define un sistema dinámico sobre $\mathbb{R}^n$, ya que para cada $x_0 \in \mathbb{R}^n$, $\phi(t, x_0)$ es la solución del problema de valor inicial

$$\begin{aligned} \dot{x} &= Ax, \\ x(0) &= x_0. \end{aligned}$$

En general, si $\phi(t, x)$ es un sistema dinámico sobre $E \subset \mathbb{R}^n$, entonces la función $f(x) = \frac{d}{dt}\phi(t, x) \mid_{t=0}$ define un campo vectorial (de tangentes a la curva $\phi(t, x)$) y

que para cada $x_0 \in \mathbb{R}^n$, $\phi(t, x_0)$ es la solución del problema de valores iniciales

$$\begin{aligned} \dot{x} &= f(x) \\ x(0) &= x_0. \end{aligned}$$

Poe eso, cada sistema dinámico puede verse como el conjunto solución de una ecuación diferencial ordinaria en un espacio vectorial (como por ejemplo $\mathbb{R}^n$). Recíprocamente, dada una ecuación diferencial $\dot{x} = f(x)$ con $f \in C^1(E)$ y $E \subset \mathbb{R}^n$ abierto, la solución $\phi(t, x_0)$ del problema de valor inicial

$$\begin{aligned} \dot{x} &= f(x) \\ x(0) &= x_0, \end{aligned}$$

donde $x_0 \in E$, es un sistema dinámico sobre E.

En el presente capítulo hablaré de la identificación de coeficientes de un sistema dinámico que corresponde a una ecuacion diferencial ordinaria lineal, pues como sabemos también hay ecuaciones diferenciales no lineales.

Consideremos las vibraciones de un sistema dinámico lineal estable en $\mathbb{R}^n$ bajo una fuerza armónica

$$\frac{dy}{dt} = Ay + F(\omega)e^{i\omega t} \tag{2.1}$$

donde $y(t) \in \mathbb{C}^n$, A es matriz nxn, y la amplitud $F(\omega)$ es un polinomio en la variable ω, esto es, $F(\omega) = \sum_{j=0}^{\nu} f_j \omega^j$, con $f_j \in \mathbb{C}^n$. Por ejemplo, los polinomios $F(\omega)$ ocurren en problemas de vibración que están sujetos a desbalanceos perturbaciones [1],[2].

La solución general del sistema (2.1) es [3]

$$y(t) = \tilde{y}(\omega)e^{i\omega t} + \sum_{j=1}^{q} e^{\lambda_j t} Q_j(t), \tag{2.2}$$

$$\tilde{y}(\omega) := (i\omega I - A)^{-1} F(\omega); \tag{2.3}$$

donde q es el número de valores propios diferentes de la matriz A, y el grado del vector polinomial $Q_j(t)$ es menor o igual a la multiplicidad algebraica ν_j del valor propio λ_j. Para sistemas estables $\operatorname{Re}\lambda_j \leq 0$. Sí para todos los λ_j, $\operatorname{Re}\lambda_j < 0$, cuando $t \to +\infty$ la solución del sistema (2.2) tiende a la solución estacionaria $\bar{y}(\omega) := \tilde{y}(\omega)e^{i\omega t}$ a pesar del valor inicial cuando $t \to +\infty$.

Para varios sistemas reales, el valor inicial y los coeficientes f_j son desconocidos, y solo podemos medir las amplitudes $\tilde{y}(\omega)$ de la solución estacionaria en diferentes frecuencias ω. Además, varios sistemas dinámicos reales no son lineales y pueden

representarse por el modelo lineal (2.1) después de procesos transitorios. Por ejemplo, frecuentemente estos sistemas dinámicos ocurren en los modelos de sistemas mecánicos, como un rotor que funciona con chumaceras de aceite [1],[2].

Por lo tanto, para tales sistemas, es muy natural determinar (los coeficientes de) la matriz A por medio de valores que arroja la salida estacionaria $\tilde{y}(\omega_j)$ para diferentes frecuencias $\omega_j;\ j = 1, 2, ..., l;\ l \geq n$ [4]. En general, existen tres maneras para determinar los coeficientes de un sistema dinámico lineal, las cuales se describen en seguida:

1. Si se considera que conocemos las fuerzas armónicas f_k^α entonces pueden calcularse los coeficientes de la matriz A a través de los valores que arroja la salida estacionaria $\tilde{y}(\omega_j)$,

2. Si supone que conoce las entradas de la matriz A entonces pueden calcularse las fuerzas armónicas f_k^α a través de los valores de la salida estacionaria $\tilde{y}(\omega_j)$, y

3. Si ahora se conocen únicamente los valores de la sálida estacionaria $\tilde{y}(\omega_j)$ entonces podemos calcular las entradas de la matriz A y las fuerzas armónicas f_k^α solo cuando se tengan más frecuencias ω_j y así tener más valores de la sálida estacionaria.

A este último problema vamos a llamarle el problema de identificación simúltanea.

Al extender la clase de matrices recuperables A, es recomendable considerar varios problemas (2.1) con la misma matriz A pero diferentes amplitudes $F^\alpha(\omega) = \sum_{j=0}^{\nu} f_j^\alpha \omega_j$ de la fuerza armónica. Las amplitudes $\tilde{y}_\alpha(\omega_j)$ de la salida estacionaria, se obtienen de las frecuencias $\omega_1, ..., \omega_l$ para el conjunto de coeficientes $\{f_0^\alpha, ..., f_\nu^\alpha\}$. Sean $\lambda_j;\ j = 1, ..., q$; los valores propios de la matriz A. Supongamos que existen b_j bloques de Jordan de dimensión más que cero y a_j bloques de Jordan de dimensión igual a cero, asociados al valor propio correspondiente λ_j. Note que los vectores propios correspondientes a un bloque de Jordan con dimensión cero, no tiene vectores adjuntos. Definimos a $\varkappa_A$,como

$$\varkappa_A := max\,\{(b_j + a_j)\} \tag{2.4}$$

Obviamente $\varkappa_A \leq n$. Denotemos como P_A al polinomio mínimo de la matriz A y $degP_A$ como su grado.

Definición 1 *La identificación simultánea de la matriz A y de los coeficientes $f_j^\alpha, 0 \leq j \leq \nu;\ 1 \leq \alpha \leq \varkappa$ (para $\varkappa \geq \varkappa_A$) a través de las amplitudes que resultan en la salida estacionaria $\tilde{y}_\alpha(\omega_j)$, para l diferentes frecuencias $\omega_j;\ j = 1, ..., l; l \geq \deg P_A + (\nu+1)$, se le conoce como el* ***problema inverso*** *para el sistema (2.1) con diferentes amplitudes $F^\alpha(\omega)$, $\alpha = 1, ..., \varkappa$.*

Hay muchas publicaciones sobre el problema de identificación de coeficientes [5],[6],[7],[8],[9],[10],[11], etc. Algunos de ellos se refieren a la teoría del control, otras usan procesos de transición de valores para la identificación.

La teoría del control considera sistemas de la forma

$$\begin{aligned} \frac{dx}{dt} &= Ax + Fu(t), \\ y &= Cx, \\ x(0) &= x_0 \in \mathbb{R}^n, \end{aligned} \tag{2.5}$$

aquí, A es una matriz constante $(n \times n)$; F es una matriz constante $(n \times \varkappa)$ y C es una matriz constante $(r \times n)$; con $r \leq n$. A la función vectorial $u(t) \in \mathbb{R}^{\varkappa}$, $\varkappa \leq n$, se conoce como la función de control (la entrada) y a la función $y(t) \in \mathbb{R}^r$, $r \leq n$, es llamada la salida. La función (matricial) de transferencia para el sistema (2.5) es

$$H(z) = C(zI - A)^{-1}F. \tag{2.6}$$

Evidentemente que las matrices CT^{-1}, TAT^{-1} y TF tienen la misma función matricial de transferencia para cualquiera que sea $T \in GL(n, R)$. Por lo que la teoría del control estudia a las clases de matrices C, A y F, que corresponden a una misma función matricial de transferencia [6], [7].

Así pues, para la identificación única de las matrices a través de su función de transferencia se necesita pedirles ciertas condiciones. Por ejemplo, en algunos libros como [6] y [7] afirman que las matrices C, A y F deberán satisfacer las condiciones de observabilidad y controlabilidad (teorema de Kalman).

Si $F(\omega) = F_0 + ... + F_\nu \omega^\nu$ es una matriz $(n \times \varkappa)$, cuyas columnas de las matrices F_j's son los vectores f_j^α, $\alpha = 1, ..., \varkappa$. En este caso, el problema inverso bajo consideración es equivalente al problema de identificación de las matrices A, F_j, $j = 0, ..., \nu$, a través de los valores de la función matricial de transferencia $[i\omega I - A]^{-1} F(\omega)$, en un número finitos de frecuencias ω_j (en el dominio de frecuencias). Estas matrices A y F_j, $j = 0, ..., \nu$, pueden determinarse como una solución de un sistema algebraico.

Para garantizar la unicidad de la identificación, se introduce una condición de resolución (2). Si $F(\omega) = F_0$, entonces esta condición de resolución es precisamente la condición de controlabilidad [6].

Suponga que la matriz A es fija y que las matrices F_j, $j = 0, ..., \nu$ se eligen aleatoriamente. En este caso, se demostrará en el teorema (1) que, con probabilidad igual a 1, las matrices F_j, $j = 0, ..., \nu$ satisfacen la condición de resolución y por lo tanto, bajo esta elección aleatoria de las F_j, $j = 0, ..., \nu$, el problema inverso siempre tiene solución única con probabilidad de uno.

Hay muchas definiciones del problema inverso para sistemas continuos [14],[15]. El método propuesto de identificación puede usarse para resolver algunos problemas inversos, para sistemas continuos después de su discretización. Por ejemplo este método puede aplicarse para la identificación de elasticidad y coeficientes de conducción en problemas unidimensionales y bidimensionales.

Algunos sistemas lineales dinámicos tienen matrices de particular estructura (por ejemplo las matrices de Jacobi) que pueden identificarse a partir de su vibración espectral (con matrices original y modificada) [16],[17],[18]. De cualquier modo, este método no puede aplicarse para matrices con estructuras complicadas.

2.2. Algoritmo para la Identificación

Primero se considera el problema inverso con solo una fuerza $F(\omega)$ y se describe el algoritmo para la identificación de los coeficientes de la matriz A y del conjunto de coeficientes (vectoriales) f_j, $j = 0, ..., \nu$ del polinomio $F(\omega)$, a partir de las amplitudes de la salida estacionaria $\tilde{y}(\omega_k)$, $k = 1, ..., (n + \nu + 1)$ (evidentemente, $(n + \nu + 1) \geq \deg P_A + \nu + 1$).

Como $\tilde{y}(\omega) = (i\omega I - A)^{-1} F(\omega)$, se obtiene, el siguiente sistema lineal con respecto a los elementos de la matriz A y los coeficientes vectoriales f_j

$$A\tilde{y}(\omega_k) + F(\omega_k) = i\omega_k \tilde{y}(\omega_k), \ k = 1, ..., n + \nu + 1, \tag{2.7}$$

donde los valores $\tilde{y}(\omega_k)$ son conocidos.

Se denotan por y_{km}, $m = 1, ..., n$, $k = 1, ..., (n + \nu + 1)$, a las componentes de los vectores $\tilde{y}(\omega_k)$ y sea $\|y_{km}\|$ la matriz con renglones $\tilde{y}(\omega_k)$. Entonces el sistema (2.7) puede reducirse al sistema lineal con respecto a los renglones a_{lm}, $l = 1, ..., n$, de la matriz A y las componentes f_{jm}, $m = 1, ..., n$, del vector f_j :

$$\sum_{m=1}^{n} y_{km} a_{lm} + \sum_{j=0}^{\nu} \omega_k^j f_{jl} = i\omega_k y_{kl}, \tag{2.8}$$

donde $k = 1, ..., (n + \nu + 1)$. Sea el índice l fijo, y los índices k, j y m variables; entonces el sistema (2.8) es sistema lineal con respecto a las $(n+\nu+1)$ desconocidas $a_{l1,}..., a_{ln}$, $f_{ol}, ..., f_{\nu l}$, con la matriz:

$$\begin{bmatrix} y_{11} & \cdots & y_{1n} & 1 & \omega_1 & \cdots & \omega_1^\nu \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ y_{(n+\nu+1)1} & \cdots & y_{(n+\nu+1)n} & 1 & \omega_{n+\nu+1} & \cdots & \omega_{n+\nu+1}^\nu \end{bmatrix}. \tag{2.9}$$

De esto se sigue que, se pueden determinar los coeficientes de la matriz a_{lj} y los vectores f_j variando el índice l desde 1 hasta n en el sistema (2.8). Note que para valores exactos de $\tilde{y}(\omega_j)$ el sistema (2.8)siempre tiene una solución (los coeficientes f_j y la matriz original A). El problema básico es determinar las condiciones para garantizar la unicidad de la solución.

Durante las mediciones de las amplitudes $\tilde{y}(\omega_j)$ algunos errores se generan experimentalmente. Para reducir sus influencias, uno puede resolver el sistema (2.8) (el cual es generalmente sobredeterminado) numéricamente por el método de mínimos cuadrados. Suponga que exactamente el sistema (2.8) tiene una solución única para cada diferentes $(\deg P_A + \nu + 1)$ frecuencias. En este caso, el método de mínimos cuadrados [5] en $(n + \nu + 1)$ o más frecuencias dan la misma solución única para los valores exactos de (salida) estacionaria. Por lo tanto, el método de mínimos cuadrados reduce la influencia de errores experimentales presentados en el uso de información experimental (datos). Cuando los errores experimentales son suficientemente pequeños el método de mínimos cuadrados nos da la solución, como única.

Se observa que para una matriz arbitraria A es imposible, generalmente, obtener la solución única del sistema (2.7) al restringir solo los coeficientes $f_o, ..., f_\nu$. Por lo tanto, es necesario considerar $\varkappa$ sistemas (2.7) con diferentes amplitudes de entrada $F^\alpha(\omega)$, $\alpha = 1, ..., \varkappa$; esto es,

$$A\tilde{y}_\alpha(\omega_k) + F^\alpha(\omega_k) = i\omega_k\tilde{y}_\alpha(\omega_k), \tag{2.10}$$

donde $k = 1, ..., (n + \nu + 1)$ y $1 \leq \alpha \leq \varkappa$. Este número $\varkappa$ es tal que $\varkappa \geq \varkappa_A$, donde el número $\varkappa_A$ está definido en (2.4).

Sea $\mathbb{C}^{n\varkappa(v+1)}$ el espacio de coeficientes f_j^α, $j = 0, ..., v$; $\alpha = 1, ..., \varkappa$, y sea $\Phi^\varkappa \in \mathbb{C}^{n\varkappa(v+1)}$ el vector unitario que está definido como

$$\Phi^\varkappa := \left(\frac{f_0^1}{\Psi^\varkappa}, ..., \frac{f_0^\varkappa}{\Psi^\varkappa}; ...; \frac{f_v^1}{\Psi^\varkappa}, ..., \frac{f_v^\varkappa}{\Psi^\varkappa} \right), \tag{2.11}$$

donde

$$\Psi^\varkappa = \sqrt{\sum_{\alpha=1}^{\varkappa} \sum_{j=0}^{v} (f_j^\alpha, f_j^\alpha)}, \text{ donde } f_j^\alpha \in \mathbb{C}^n \text{ para } j = 0, ..., v; \ \alpha = 1, ..., \varkappa.$$

Evidentemente, $\Phi^\varkappa \in S^{2n\varkappa(v+1)-1}$, donde $S^{2n\varkappa(v+1)-1}$ es la esfera unitaria de dimensión real $2n\varkappa(v + 1) - 1$. Definimos med como la medida uniforme sobre la esfera $S^{2n\varkappa(v+1)-1}$, esto es, si λ es la medida de Lebesgue sobre los subconjuntos borelianos de esta esfera, entonces $med\ (A) := \frac{\lambda(A)}{\lambda(S^{2n\varkappa(v+1)-1})}$ para cada $A \subset S^{2n\varkappa(v+1)-1}$ boreliano.

El determinante del sistema (2.10), es una función analítica con respecto a a_{ij}, f_j, ω_k. Más aún, este determinante, es una función homogénea de primer orden con respecto a los coeficientes f_j^α, $j = 0, ..., \nu$; $\alpha = 1, ..., \varkappa$. Por lo tanto, para determinar las condiciones tales que el determinante del sistema (2.10) no sea cero,

es suficiente considerar las componentes del vector unitario $\Phi^{\varkappa}$ en el espacio de coeficientes $\mathbb{C}^{n\varkappa(\nu+1)}$.

Al conjunto de vectores $\Phi^{\varkappa} \in S^{2n\varkappa(\nu+1)-1}$ tales que el sistema (2.10) no tiene solución única se define como $N_r(A)$. Más adelante se demuestra que, el conjunto $N_r(A)$ no depende de la elección de las frecuencias (no resonantes) $\omega_1, ..., \omega_{n+\nu+1}$ y que $med(N_r(A)) = 0$. Esto significa que, el problema inverso tiene siempre una solución única para casi todos los conjuntos de coeficientes $\left\{f_j^{\alpha}, j = 0, ..., \nu; \alpha = 1, ..., \varkappa; \varkappa \geq \varkappa_A\right\}$ que elijamos aleatoriamente y para cualquier selección de frecuencias arbitrarias $\omega_1, ..., \omega_{n+\nu+1}$.

2.3. Resultados Principales

Primero analizaré la resolución del sistema (2.8) en el caso cuando la matriz A, del sistema (2.1), tiene n diferentes valores propios. Obviamente, en este caso $\varkappa_A = 1$ y por ello, el problema inverso solo considera a los sistemas (2.7) y (2.8).

Lema 1 *Suponga que $n > 1$, y que la matriz A $(n \times n)$ y su matriz complejo conjugada A^* tienen n diferentes valores propios λ_j, $j = 1, ..., n$, y $\bar{\lambda}_j$, $j = 1, ..., n$, respectivamente. Sean $\mathbf{e}_j$, $\mathbf{e}_j^*$ los vectores propios de las matrices A y A^* respectivamente, i.e. $A^*\mathbf{e}_j^* = \bar{\lambda}_j\mathbf{e}_j^*$; $A\mathbf{e}_j = \lambda_j\mathbf{e}_j$; $j = 1, ..., \nu$. Suponga que $\mathbf{e}_j$, $\mathbf{e}_i^*$ se eligen en tal forma que $(\mathbf{e}_j, \mathbf{e}_i^*) = \delta_{ji}$. Se definen en el espacio de coeficientes $\mathbb{C}^{n(\nu+1)}$ los planos π_k de dimensión $\{2n(\nu+1)-1\}$ que satisfacen la siguiente condición:*

$$\pi_k : \left(\sum_{j=0}^{\nu}(-i)^j A^j f_j, \mathbf{e}_k^*\right) = 0,\ k = 1, ..., n;\ \textit{esto es,}\ (F(-iA), \mathbf{e}_k^*) = 0\ \textit{para}\ 1 \leq k \leq n. \tag{2.12}$$

y definimos el conjunto $\tilde{N}_r(A)$ como:$\tilde{N}_r(A) := \cup_{k=1}^{n}\pi_k$.

I) Suponga que se tienen $\tilde{y}(\omega_j) = [i\omega_j I - A]^{-1}F(\omega_j)$ en $(n+\nu+1)$ frecuencias no resonantes ω_j; entonces:

1. El sistema (2.7) tiene única solución si y solamente si $\Phi^1 \notin \tilde{N}_r(A) \cap S^{2n(\nu+1)-1}$.

2. El conjunto $N_r(A) := \tilde{N}_r(A) \cap S^{2n(\nu+1)-1}$ tiene medida cero.

3. El coeficiente f_ν puede determinarse unívocamente de los sistemas (2.7) y (2.8);

4. El kernell del sistema

$$A_1\tilde{y}(\omega_j) + F^*(\omega_j) = 0,\ j = 1, ..., (n+\nu+1) \tag{2.13}$$

con respecto a A_1, $f_0^*, ..., f_\nu^*$, no depende de la elección de frecuencias no resonantes ω_j.

II) Para cualquier conjunto de valores $\tilde{y}(\omega_j)$, $j = 1, ..., (n + \nu + 1)$, *tal que si el rango de la matriz (2.9) es* $(n + \nu + 1)$, *se tiene que existe una matriz* A *única y coeficientes* $f_0^*, ..., f_\nu^*$ *únicos, tales que*

$$\tilde{y}(\omega_j) = [i\omega_j I - A]^{-1} F(\omega_j) \ j = 1, ..., (n + \nu + 1).$$

En este caso, la matriz A *y los coeficientes* $f_0^*, ..., f_\nu^*$ *satisfacen la condición* $\Phi^1 \notin N_r(A)$.

III) Si la dimensión n, *es igual a 1, entonces:*

1. La matriz $A \in \mathbb{C}$, *y la fuerza armónica* $F(\omega) \in \mathbb{C}^n = \mathbb{C}$, *y así los coeficientes* $f_0, ..., f_\nu$ *del polinomio* $F(\omega)$ *junto con la matriz* A *pueden determinarse de forma única si y solo si,* $F(-iA) \neq 0$ *en las* $(n + v + 1) = (v + 2)$ *frecuencias no resonantes* ω_j. *Es decir, la condición* $\Phi^1 \notin N_r(A)$ *se reduce en el caso unidimensional a que* $F(-iA) \neq 0$.

2. En el caso cuando el rango de la matriz (2.9) sea igual a $(\nu + 2)$ *existe un único número complejo* A *y existen únicos numeros complejos* $f_0^*, ..., f_\nu^*$ *que son los coeficientes que definen a* $F(\omega)$, *tales que:*

$(i\omega - A)F(\omega)|_{\omega=\omega_j} = y_j$ *y* $F(-iA) \neq 0$, *donde* $y_j := \tilde{y}(\omega_j)$.

Demostración. I) Es conocido el hecho [19] que, si la matriz A de tamaño $n \times n$ tiene n diferentes valores propios λ_j, $j = 1, ..., n$, entonces la matriz A^* también tiene n diferentes valores propios $\bar{\lambda}_j$, $j = 1, ..., n$. Los vectores propios $\mathbf{e}_j$, $j = 1, ..., n$ $(\mathbf{e}_j^*, j = 1, ..., n)$,de la matriz A (matriz A^*) son linealmente independientes. Más aún, ellos pueden elegirse de tal manera que $(\mathbf{e}_j, \mathbf{e}_j^*) = \delta_{ij}$.

Por lo tanto el vector $\tilde{y}(\omega) = [i\omega I - A]^{-1} F(\omega)$, se puede ver como

$$\tilde{y}(\omega) = \sum_{k=1}^{n} \frac{(F(\omega), \mathbf{e}_k^*)\mathbf{e}_k}{i\omega - \lambda_k}. \tag{2.14}$$

La matriz A y la entrada $F(\omega) = \sum_{\alpha=0}^{\nu} \omega^\alpha f_\alpha$ satisfacen los sistemas lineales (2.7) y (2.8) con los valores $\tilde{y}(\omega_k)$.

El determinante de la matriz (2.9) no es cero, si y solo si el sistema (2.8) tiene solución única. Ahora, la tarea es investigar la unicidad de las soluciones para los sistemas (2.7) y (2.8).

Suponga que el sistema (2.7) tiene dos soluciones diferentes, digamos la inicial A; $f_\alpha, \alpha = 0, .., \nu$ y alguna otra solución A_1; $f_\alpha^*, \alpha = 0, .., \nu$. Sustituyendo la expresión (2.14) con la solución A, $f_\alpha, \alpha = 0, .., \nu$ en el sistema (2.7) y sustituyendo la expresión (2.14) con la otra solución A_1; $f_\alpha^*, \alpha = 0, .., \nu$ en el sistema (2.7), nos

dan dos expresiones diferentes. Sustrayendo de la primera expresión la segunda, se obtiene

$$\sum_{k=1}^{n}(F(\omega_j),\mathbf{e}_k^*)\frac{(A-A_1)\mathbf{e}_k}{i\omega_j-\lambda_k}+(F(\omega_j)-F^*(\omega_j))=0, \tag{2.15}$$

donde $F^*(\omega_j):=\sum_{\alpha=0}^{\nu}f_\alpha^*\omega_j^\alpha$ y $j=1,...,(n+\nu+1)$.

Ahora, al tomar los siguientes polinomios

$$\Phi(i\omega):=\prod_{j=1}^{n}(i\omega-\lambda_j)\text{ y }\Phi_k(i\omega):=\prod_{j=1,j\neq k}^{n}(i\omega-\lambda_j), \tag{2.16}$$

se tiene que al multiplicar la relación (2.15) por los polinomios $\Phi(i\omega_j)$, $j=1,...,(n+\nu+1)$, el polinomio

$$\sum_{k=1}^{n}\Phi_k(i\omega)(F(\omega),\mathbf{e}_k^*)(A-A_1)\mathbf{e}_k+(F(\omega)-F(\omega^*))\Phi(i\omega)=0 \tag{2.17}$$

de grado $(n+\nu)$ tiene $(n+\nu+1)$ diferentes raíces, $\omega_1,...,\omega_{n+\nu+1}$. Por tanto, se sigue que, el polinomio del lado izquierdo en (2.17) es el polinomio idénticamente cero. Esto implica que, todos los coeficientes del polinomio en (2.17) son iguales a cero. Note que el coeficiente de este polinomio en el término $\omega^{n+\nu}$, es igual a la diferencia $(f_\nu-f_\nu^*)$, y así $f_\nu=f_\nu^*$. Esto completa la prueba del punto I.3.

Ahora se considera la ecuación (2.17) en $i\omega=\lambda_j$, $j=1,...,n$. Ya que el polinomio (2.16), es idénticamente cero, se obtiene

$$\left(\sum_{\alpha=0}^{\nu}f_\alpha(-i\lambda_j)^\alpha,\mathbf{e}_j^*\right)(A-A_1)\mathbf{e}_j=0,\ j=1,...,n. \tag{2.18}$$

Obviamente, la ecuación del plano π_k, puede escribirse en la forma

$$\begin{aligned}\left(\textstyle\sum_{j=0}^{\nu}(-i)^jA^jf_j,\mathbf{e}_k^*\right) &= \left(\textstyle\sum_{j=0}^{\nu}(-i)^jf_j,(A^*)^j\mathbf{e}_k^*\right)\\ &= \left(\textstyle\sum_{j=0}^{\nu}f_j(-i)^j\lambda_k^j,\mathbf{e}_k^*\right)=0.\end{aligned} \tag{2.19}$$

Por lo tanto, si $\Phi^1\notin N_r(A)$ entonces los primeros n factores de las n ecuaciones en (2.18) no son iguales a cero, y se obtiene

$$(A-A_1)\mathbf{e}_l=0,\ l=1,...,n. \tag{2.20}$$

Como los vectores propios $\{\mathbf{e}_1, ..., \mathbf{e}_n\}$ de la matriz A forman una base de $\mathbb{R}^n$. De (2.20) se sigue que $A = A_1$ cuando tenemos que $\Phi^1 \notin N_r(A)$. Esto termina la prueba de la parte I.1.

Se sigue de (2.18) que para $\Phi^1 \notin N_r(A)$, la ecuación (2.17) se puede reducir a

$$(F(\omega) - F^*(\omega))\Phi(i\omega) = 0, \ \omega \in \mathbb{R}^1. \tag{2.21}$$

La ecuación (2.21) implica que $F(\omega) = F^*(\omega)$, es decir, los coeficientes $f_0, ..., f_\nu$, también pueden determinarse de manera única cuando $\Phi^1 \notin N_r(A)$.

Para completar la prueba del punto I.2, es suficiente en notar que el conjunto $\tilde{N}_r(A)$ es unión de n planos, y como los planos tienen medida de Lebesgue cero en el espacio de coeficientes $\mathbb{C}^{2n(v+1)}$, pero como $N_r(A) := \tilde{N}_r(A) \cap S^{2n(\nu+1)-1}$ entonces se tiene que $med\ (N_r(A)) = 0$.

Ahora se investiga el núcleo del sistema (2.8). En seguida se probará que para $\Phi^1 \in N_r(A)$, la solución del sistema (2.8) no es única. Suponga que $\Phi^1 \in N_r(A)$, entonces

$$\begin{aligned} \left(\sum_{\beta=0}^{\nu} f_\beta(-i\lambda_k)^\beta, \mathbf{e}_j^*\right) &= 0, \text{ para } j = 1, ..., q; \\ \left(\sum_{\beta=0}^{\nu} f_\beta(-i\lambda_k)^\beta, \mathbf{e}_k^*\right) &\neq 0, \text{ para } k = (q+1), ..., n. \end{aligned} \tag{2.22}$$

De (2.22) se tiene que

$$\left(\sum_{\beta=0}^{\nu} f_\beta(\omega)^\beta, \mathbf{e}_j^*\right) = (\omega + i\lambda_j)\left(\sum_{\alpha=0}^{\nu-1} a_{j\alpha}\omega^\beta\right), \text{ para } j = 1, ..., q.$$

Aquí los números $a_{j\alpha}$ no dependen de las frecuencias $\omega_1, ..., \omega_{n+\nu+1}$. Evidentemente $\Phi_j(i\omega)(\omega + i\lambda_j) = -i\Phi(i\omega)$. Así es que, la ecuación (2.17) (la cual equivale al sistema (2.15)) es equivalente al siguiente sistema:

$$\begin{aligned} (-i)\sum_{j=1}^{q} a_{j\alpha}(A - A_1)\mathbf{e}_j + (f_\alpha - f_\alpha^*) &= 0, \\ (A - A_1)\mathbf{e}_k &= 0, \end{aligned} \tag{2.23}$$

donde $\alpha = 0, ..., \nu$; $k = (q+1), ..., n$. Este sistema determina la dimensión del núcleo (o kernel) de $(A - A_1)$, $(f_\alpha - f_\alpha^*)$ del sistema homogéneo, del problema (2.13). Evidentemente el kernel para este último sistema no depende de las frecuencias $\omega_1, ..., \omega_{n+\nu+1}$ que consideremos. Esto prueba la parte I.4.

II) Si el rango de la matriz (2.9) es igual a $(n + \nu + 1)$, entonces existe una única matriz A y coeficientes vectoriales $f_0, ..., f_\nu$ que satisfacen las ecuaciones (2.7) y (2.8). Esto implica que $\tilde{y}(\omega_j) = [i\omega_j I - A]^{-1}F(\omega_j)$. Por lo tanto, del punto I.1 de este lema,

se tiene que el vector Φ^1, correspondiente a los coeficientes $f_0, ..., f_\nu$, no pertenece a $N_r(A)$.

III) El caso unidimensional (esto es, $n = 1$) es un caso particular del punto I. ■

Ahora se considera el problema inverso para una matriz arbitraria A. En seguida recordaré algunos resultados de la teoría de matrices [19] e introduciré algunas notaciones.

Sea A una matriz que tiene q diferentes valores propios λ_j, $j = 1, ..., q$; $q \leq n$, y que existen ν_j bloques de Jordan $J_{k_j^i}$ de orden k_j^i, $i = 1, ..., \nu_j$, asociados al valor propio λ_j. Por definición, el orden del bloque de Jordan $J_{k_j^i}$, es el número de vectores adjuntos (este orden es uno menos a la dimensión de la matriz de Jordan de este bloque). Sin pérdida de generalidad puede asumirse que $\nu_1 \geq ... \geq \nu_{j-1} \geq \nu_j$, y que para un índice fijo j, los bloques se ordenan de manera que $k_j^1 \geq k_j^2 \geq ... \geq k_j^{\nu_j}$, $j = 1, ..., q$. Note que el número más grande de bloques de Jordan corresponde al valor propio λ_1 (es decir, $\nu_1 := \text{máx}_{1 \leq j \leq q}\{\nu_j\}$).

Por lo que se obtiene

$$\deg P_A = \sum_{j=1}^{q} k_j^1 + q.$$

Sea $\vartheta_j = \{\mathbf{e}_j^\beta : \beta = 1, ..., a_j;\ j \text{ fijo}\}$ el conjunto de los vectores propios de la matriz A asociados al valor propio λ_j, tal que ellos no tengan vectores adjuntos. Por construcción, el último bloque $J_{k_j^{\nu_j}}$ asociado al valor propio λ_j, le corresponden los vectores propios $\mathbf{e}_j^\beta$, $\beta = 1, ..., a_j$, y por definición $k_j^{\nu_j} = 0$ si $\vartheta_j \neq \varnothing$. Note que $k_j^i > 0$ para $i < \nu_j$ y $k_j^{\nu_j} \geq 0$. Se denota por $\mathbb{C}(\vartheta_j)$ al espacio generado por los vectores $\mathbf{e}_j^\beta$, $\beta = 1, ..., a_j$. Evidentemente, $\mathbb{C}(\vartheta_j)$ es un subespacio invariante de la matriz A. Si $\vartheta_j = \varnothing$, entonces $k_j^{\nu_j} > 0$.

Abajo, se definen los subespacios invariantes correspondientes a los bloques de Jordan de orden $k_j^i > 0$. Si $k_j^i > 0$, entonces existe un vector propio designado como $\mathbf{e}_{ij}^{k_j^i+1}$ (es decir, $A\mathbf{e}_{ij}^{k_j^i+1} = \lambda_j \mathbf{e}_{ij}^{k_j^i+1}$), y existen vectores adjuntos $\mathbf{e}_{ij}^l$, $l = 1, ..., k_j^i$ (es decir, $(A - \lambda_j I)\mathbf{e}_{ij}^l = \mathbf{e}_{ij}^{l+1}$) tal que

$$\begin{aligned} (A - \lambda_j I)^{k_j^i+1} \mathbf{e}_{ij}^1 &= 0, \\ (A - \lambda_j I)^l \mathbf{e}_{ij}^1 &= \mathbf{e}_{ij}^{l+1}, \end{aligned} \tag{2.24}$$

con $l = 1, ..., k_j^i$.

Se denota por $\mathbb{C}(i, j)$ al espacio lineal generado por los vectores $\{\mathbf{e}_{ij}^l : l = 1, ..., k_j^i + 1\}$. Obviamente, $\mathbb{C}(i, j)$ es el subespacio invariante de la matriz A, y sobre el espacio $\mathbb{C}(i, j)$ la matriz $(A - \lambda_j I)$ es la matriz bloque de Jordan $J_{k_j^i}$ de orden $k_j^i > 0$. Si

$k_j^{\nu_j} = 0$, entonces por definición $J_{k_j^{\nu_j}}$ es la matriz cero en el espacio $\mathbb{C}(\vartheta_j)$. De (2.24) se tiene que

$$\mathbf{e}_{ij}^{l+1} = J_{k_j^i}^l \, \mathbf{e}_{ij}^1, \ l = 1, ..., k_j^i, \ \text{para } k_j^i > 0. \tag{2.25}$$

Por $(\cdot, \cdot)$ se denota al nuevo producto escalar en $\mathbb{C}^n$ tal que el producto escalar de los vectores propios y los vectores adjuntos de la matriz A, sean ortonormales con respecto a este producto escalar. Si $\vartheta_j \neq \varnothing$, entonces el teorema de Jordan [19] implica que

$$\mathbb{C}^n = \oplus_{j=1}^q \oplus_{i=1}^{\nu_j - 1} \mathbb{C}(i, j) \oplus \mathbb{C}(\vartheta_j), \tag{2.26}$$

y si $\vartheta_j = \varnothing$ para todo j, entonces el teorema de Jordan implica que

$$\mathbb{C}^n = \oplus_{j=1}^q \oplus_{i=1}^{\nu_j} \mathbb{C}(i, j).$$

Ahora se denota por $I_{k_j^i}$ la matriz idéntidad sobre $\mathbb{C}(i, j)$. Se sigue del teorema de Jordan, que la matriz A relativa a la base de sus vectores propios y sus vectores adjuntos tiene la forma

$$SAS^{-1} = \begin{bmatrix} I_{k_1^1}\lambda_1 + J_{k_1^1} & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & 0 & \vdots & \vdots \\ \vdots & 0 & I_{k_1^{\nu_1}}\lambda_1 + J_{k_1^{\nu_1}} & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & 0 \\ 0 & 0 & 0 & \cdots & I_{k_q^{\nu_q}}\lambda_q + J_{k_q^{\nu_q}} \end{bmatrix} \tag{2.27}$$

Aquí S^{-1} manda los nuevos básicos a la base inicial.

Sea P_j^i la proyección ortogonal sobre el subespacio $\mathbb{C}(i, j)$. Si $\vartheta_j \neq \varnothing$, entonces $\mathbb{C}(j)$ está definido como una suma directa

$$\mathbb{C}(j) = \mathbb{C}_1(j) \oplus \mathbb{C}(\vartheta_j)$$

donde $\mathbb{C}_1(j)$ es el espacio generado por los vectores $\{e_{ij}^1 : 1 \leq i \leq \nu_j - 1$, el índice j es fijo$\}$; y si $\vartheta_j = \varnothing$, entonces por definición, $\mathbb{C}(j)$ es un espacio generado por los vectores $\{e_{ij}^1 : 1 \leq i \leq \nu_j$, el índice j es fijo$\}$. Evidentemente, la dimensión de $\mathbb{C}(j)$ es $(a_j + b_j)$.

Ahora, se considera al conjunto de amplitudes $F^\alpha(\omega) = \sum_{j=0}^{\nu} f_j^\alpha \omega^j$, donde $\alpha = 1, ..., \varkappa$; $\varkappa \geq \varkappa_A$; $f_j^\alpha \in \mathbb{C}^n$, definimos $L^\varkappa$ como el espacio lineal generado por los vectores $SF^\alpha(-iA) = S\sum_{j=0}^{\nu}(-i)^j A^j f_j^\alpha$, para $\alpha = 1, ..., \varkappa$. Se denota por π^j la proyección ortogonal de $\mathbb{C}^n$ sobre $\mathbb{C}(j)$.

Definición 2 *El conjunto de coeficientes vectoriales* $\{f_j^\alpha : \alpha = 1, ..., \varkappa;\ j = 0, ..., \nu\}$ *y la matriz A, satisface la condición de resolución, si*

$$\pi^j L^\varkappa = \mathbb{C}(j),\ \textit{para } j = 1, ..., q. \tag{2.28}$$

En [6] fue demostrado que, la condición de resolución es equivalente a la condición

$$rank[F(-iA) \vdots AF(-iA) \vdots \cdots \vdots A^{n-1}F(-iA)] = n. \tag{2.29}$$

Aquí las columnas de la matriz $F(-iA)$ son los vectores $\sum_{j=0}^{\nu}(-i)^j A^j f_j^\alpha$. Evidentemente que la condición (2.29) es invariante con respecto a las transformaciones $T^{-1}AT$, $T^{-1}F(\omega)$ para todo $T \in GL(n, \mathbb{C})$.

Enseguida se da otra definición de la condición de resolución. Sea $\{\mathbf{e}^\beta : \beta = 1, ..., (a_j + b_j)\}$ una base de $\mathbb{C}(j)$ y definamos la matriz D_j como

$$D_j = \left\|(F^\alpha(-i\lambda_j), S^*\mathbf{e}^\beta)\right\|, \tag{2.30}$$

donde $\alpha = 1, ..., \varkappa;\ \beta = 1, ..., (a_j + b_j)$. Por lo tanto, la condición de resolución es equivalente a las condiciones

$$rank\ D_j = a_j + b_j,\ j = 1, ..., q. \tag{2.31}$$

Para probar esto, basta probar que los vectores $\{S^*\mathbf{e}^\beta : \beta = 1, ..., (a_j + b_j)\}$ son los vectores propios de la matriz A^*.

Abajo se usa el vector unitario $\Phi^\varkappa$ que está definido en (2.11), y $N_r(A)$ denota al conjunto de vectores $\Phi^\varkappa$ cuyas componentes no satisfacen la condición de resolución.

Ahora estamos preparados para probar el siguiente teorema principal.

Teorema 1 *I) Suponga que para alguna matriz A (que satisface (2.27)) y para los coeficientes* f_μ^α, *se tiene que los valores* $\tilde{y}_\alpha(\omega_j)$, *en* $(\deg P_A + \nu + 1)$ *o más frecuencias no resonantes* ω_j, *están dados por*

$$\tilde{y}_\alpha(\omega_j) := (i\omega_j I - A)^{-1} F^\alpha(\omega_j), \tag{2.32}$$

donde $\alpha = 1, ..., \varkappa;\ \varkappa \geq \varkappa_A;\ \mu = 0, ..., \nu$. *Si la dimensión* $n \geq 2$, *entonces:*

1. El sistema (2.10) tiene solución única A, f_μ^α *si y solo si el vector* $\Phi^\varkappa \notin N_r(A)$;

2. El conjunto $N_r(A)$ *es un cerrado y* $med(N_r(A)) = 0$;

3. El núcleo del sistema (2.10) no depende de la elección de frecuencias no resonantes ω_j.

II) Si para algún conjunto de valores $\{\tilde{y}_\alpha(\omega_j) : j = 1, ..., d;\ d \geq (\deg P_A + \nu + 1)\}$, *la matriz A y los coeficientes* f_k^α, $k = 0, ..., \nu;\ \alpha = 1, ..., \varkappa$ *satisfacen la ecuación (2.10) y esta solución es única, entonces la matriz A y los coeficientes* f_k^α *satisfacen la condición de resolución.*

Demostración. I) La matriz A satisface (2.27), por lo tanto la matriz $\frac{1}{i\omega I-A}$ puede expresarse como matriz diagonal por bloques

$$\frac{1}{i\omega I-A}=S^{-1}\begin{bmatrix}\mathbf{a}_{11} & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \mathbf{a}_{q\nu_q}\end{bmatrix}S,$$

donde

$$\mathbf{a}_{lr}=\frac{1}{(i\omega-\lambda_r)I-J_{k_r^l}},\ \text{para } r=1,...,q;\ l=1,...,\nu_r.$$

Ya que $J_{k_r^l}^s=0$ para $s>k_r^l$, se obtiene la siguiente representación por series de matrices $\mathbf{a}_{lr}$ con respecto a $J_{k_r^l}/(i\omega-\lambda_r)$:

$$\frac{1}{(i\omega-\lambda_r)I-J_{k_r^l}}=\frac{1}{(i\omega-\lambda_r)}\left(I+\frac{J_{k_r^l}}{(i\omega-\lambda_r)}+...+\frac{J_{k_r^l}^{k_r^l}}{(i\omega-\lambda_r)^{k_r^l}}\right),$$

para $r=1,...,q$; $l=1,...,\nu_r$. También, como los vectores $\tilde{y}_\beta(\omega_k)$ satisfacen (2.3), se obtiene:

$$\tilde{y}_\beta(\omega_k)=\sum_{r=1}^{q}\sum_{l=1}^{\nu_r}\left(\frac{S^{-1}P_r^lSF^\beta(\omega_k)}{i\omega_k-\lambda_r}+...+\frac{S^{-1}J_{k_r^l}^{k_r^l}(P_r^lSF^\beta(\omega_k))}{(i\omega_k-\lambda_r)^{k_r^l+1}}\right),\tag{2.33}$$

para $\beta=1,...,\varkappa$. El sistema (2.10) es equivalente a la unión de los sistemas (2.7) en $F(\omega)=F^\beta(\omega)$, $\beta=1,...,\varkappa$. Nótese que, el sistema (2.10) siempre tiene una solución: La matriz inicial A y el conjunto de vectores iniciales f_μ^β. Ahora, substituyendo la expresion (2.33) para $\tilde{y}_\beta(\omega_k)$ y suponiendo que A_1 y $f_\mu^{*\beta}$ es alguna otra solución del sistema (2.10). Substrayendo la expresión (2.10) con los valores iniciales A y coeficientes f_μ^β de (2.10) con la otra solución A_1, $f_\mu^{*\beta}$, se obtiene

$$\sum\nolimits_{r=1}^{q}\sum\nolimits_{l=1}^{\nu_r}\left(\frac{(A-A_1)S^{-1}P_r^lSF^\beta(\omega_k)}{i\omega_k-\lambda_r}+....+\frac{(A-A_1)S^{-1}J_{k_r^l}^{k_r^l-1}P_r^lSF^\beta(\omega_k)}{i\omega_k-\lambda_r}\right)+$$
$$+(F^\beta(\omega_k)-F^{*\beta}(\omega_k))=0,\tag{2.34}$$

para $\beta=1,...,\varkappa$; $\varkappa\geq\varkappa_A$, y $k=1,...,(\deg P_A+\nu+1)$. La ecuación (2.34) puede reducirse a polinomios que se hacen cero, en los puntos ω_k, $k=1,...,(\deg P_A+\nu+1)$.

Con este objetivo en mente se definen los siguientes polinomios

$$\begin{aligned} \Phi(i\omega) &: = \prod_{j=1}^{q}(i\omega-\lambda_j)^{k_r^1+1}, \\ \Phi_{r,\alpha}(i\omega) &: = (i\omega-\lambda_r)^{-\alpha+k_r^1}\prod_{j=1,j\neq r}^{q}(i\omega-\lambda_j)^{k_r^1+1}, \end{aligned} \tag{2.35}$$

donde $r=1,...,q$; $\alpha=0,...,k_r^1$. Evidentemente, los polinomios $\Phi_{r,\alpha}(i\omega)$ son linealmente independientes.

Multiplicando la ecuación (2.34) por $\Phi(i\omega_k)$ $k=1,...,(\deg P_A+\nu+1)$, se obtiene

$$\begin{aligned} &\sum_{r=1}^{q}\sum_{j=1}^{\nu_r}(\sum_{\alpha=0}^{k_j^r}(A-A_1)S^{-1}J_{k_j^r}^{\alpha}(P_r^jSF^{\beta}(\omega))\Phi_{r,\alpha}(i\omega))+ \\ &\Phi(i\omega)\sum_{\mu=0}^{\nu}(f_\mu^\beta-f_\mu^{*\beta})\omega^\mu\equiv 0, \text{ para } \beta=1,...,\varkappa;\varkappa\geq\varkappa_A. \end{aligned} \tag{2.36}$$

Del lado izquierdo de esta ecuación (2.36), se sigue que, el polinomio de orden $(\deg P_A+\nu)$ es igual a cero en $(\deg P_A+\nu+1)$ diferentes puntos ω_j. Esto implica que el polinomio del lado izquierdo de (2.36) es idénticamente igual a cero. Evidentemente, la ecuación (2.36) es equivalente a la ecuación (2.34).

Ahora se considera la ecuación (2.36) en $i\omega=\lambda_1$. Ya que k_j^i son ordenados de tal forma que $k_j^1\geq k_j^2\geq ...\geq k_j^{\nu_j}$, se tiene

$$(A-A_1)\left\{\sum_{i=1}^{\nu_1}S^{-1}J_{k_1^i}^{k_1^1}P_1^iSF^{\beta}(-i\lambda_1)\right\}=0, \text{ para } \beta=1,...,\varkappa. \tag{2.37}$$

De (2.37), se sigue que $(i\omega-\lambda_1)$ divide a (2.36). Dividiendo (2.36) por $(\lambda_1-i\omega)$ para $i\omega=\lambda_1$, se tiene

$$(A-A_1)\left\{\sum_{i=1}^{\nu_1}S^{-1}J_{k_1^i}^{k_1^1-1}P_1^iSF^{\beta}(-i\lambda_1)\right\}=0, \text{ para } \beta=1,...,\varkappa \tag{2.38}$$

De (2.38), se sigue que $(\lambda_1-i\omega)^2$ divide a (2.36), etcétera. Por inducción sobre los índices i, j se obtienen

$$(A-A_1)\left\{\sum_{i=1}^{\nu_j}S^{-1}J_{k_j^i}^{k_j^1-\alpha}P_j^iSF^{\beta}(-i\lambda_j)\right\}=0, \tag{2.39}$$

donde $\alpha=0,...,k_j^1$; $\beta=1,...,\varkappa$; $j=1,...,q$, y por definición $J_{k_j^i}^0=I_{k_j^i}$.

Ahora, de (2.39) se sigue que, la matriz $(A - A_1)$ es igual a la matriz cero sobre el espacio lineal generado por los vectores

$$\sum\nolimits_{i=1}^{\nu_j} S^{-1} J_{k_j^i}^{k_j^1-\alpha} P_j^i S F^\beta(-i\lambda_j), \tag{2.40}$$

para $\alpha = 0, ..., k_j^1$; $\beta = 1, ..., \varkappa$; $j = 1, ..., q$. Evidentemente, los vectores (2.40) pueden expresarse como

$$(A - \lambda_j)^{k_j^1-\alpha} \sum_{i=1}^{\nu_j} P_j^i S F^\beta(-i\lambda_j). \tag{2.41}$$

Se denota por $L(A, F)$ al espacio lineal generado por los vectores (2.40). Si $L(A, F) \neq \mathbb{C}^n$, entonces sea $L^\perp(A, F)$ el complemento ortogonal no trivial en $\mathbb{C}^n$. Evidentemente, en este caso, la ecuación (2.36) tiene una solución no trivial

$$f_\mu^\beta = f_\mu^{*\beta},\ (A - A_1)|_{L(A,F)} = 0 \text{ y arbitraria } (A - A_1)|_{L^\perp(A,F)}. \tag{2.42}$$

Por lo tanto, si probamos que para $\Phi^\varkappa \in N_r(A)$ el conjunto $L^\perp(A, F)$ es no vacío, entonces la ecación (2.36) tendrá un kernel no vacío. Sin pérdida de generalidad, podemos suponer que $\Phi \in N_r(A)$ y $\pi^1 L^\varkappa \neq \pi^1 \mathbb{C}^n$. Recalcando que las definiciones de proyecciones π^1, P_j^i implican que

$$\begin{aligned} \pi^1(\sum\nolimits_{i=1}^{\nu_1} P_1^i) &= (\sum\nolimits_{i=1}^{\nu_1} P_1^i)\pi^1 = \pi^1, \\ \pi^1(\sum\nolimits_{i=1}^{\nu_1} P_j^i) &= (\sum\nolimits_{i=1}^{\nu_1} P_j^i)\pi^1 = 0, \text{ para } j \neq 1. \end{aligned} \tag{2.43}$$

Por la definición de matrices $J_{k_j^i}$ se tiene que

$$(J_{k_j^i}^m P_j^i \mathbb{C}^n) \perp (\pi^1 \mathbb{C}^n), \text{ para } m > 0. \tag{2.44}$$

La condición $\pi^1 L^\varkappa \neq \pi^1 \mathbb{C}^n$ implica que, existe un vector ortogonal $\mathbf{e} \in \pi^1 \mathbb{C}^n$ ortogonal a $\pi^1 L^\varkappa$. El vector $\mathbf{e} \in \mathbb{C}^j$, por eso $\mathbf{e} = \sum_{\beta=1}^{(a_j+b_j)} \mathbf{e}^\beta c_\beta$. Como los vectores $S^*\mathbf{e}^\beta$ son los vectores propios de la matriz A^* correspondiente al valor propio $\bar{\lambda}_j$, se obtiene

$$\begin{aligned} 0 &= (\pi^1 S F^\alpha(-iA), \mathbf{e}) = \sum_{\beta=1}^{a_1+b_1} \bar{c}_\beta (S F^\alpha(-iA), S^*\mathbf{e}^\beta) \\ &= \sum_{\beta=1}^{a_1+b_1} \bar{c}_\beta (S F^\alpha(-i\lambda_j), \mathbf{e}^\beta) = (S F^\alpha(-i\lambda_j), \mathbf{e}). \end{aligned}$$

Por lo tanto, la ecuación (2.30) implica que el vector $\mathbf{e}$ es ortogonal al espacio generado de los vectores $\{\pi^1 SF^\beta(-i\lambda_1)|\beta = 1, ..., \varkappa\}$. Las relaciones (2.43) y (2.44) implican que el vector $S^*\mathbf{e}$ es ortogonal a $L(A, F)$. Esto significa que $L^\perp(A, F) \neq \varnothing$ y obtenemos que la ecuación (2.36) tiene una solución no trivial (2.42).

Ahora, suponga que la condición de resolución es cierta (esto es,$\Phi^\varkappa \notin N_r(A)$). En seguida, probaremos que el espacio generado por los vectores (2.40) es igual a $S^{-1}(\oplus_{i=1}^{\nu_j}\mathbb{C}(i, j))$ para el índice fijo j.

Suponga que $k_j^{v_j} = 0$, entonces es fácil ver que los vectores propios $\mathbf{e}_j^\beta,\ \beta = 1, ..., a_j$ (definido arriba), pueden generarse como una combinación lineal de los vectores $P_j^{v_j} SF^\alpha(-i\lambda_j)$ con $\alpha = 1, ..., \varkappa$.

Ahora, suponga que $k_j^{v_j} > 0$. Se sigue de la condición de resolución que existen coeficientes $y_{\mu,i}$ tal que para algunos coeficientes $z_{\alpha,i}$ se tiene que

$$\mathbf{e}_{i,j}^1 + \sum_{i=1}^{v_j}\sum_{\alpha=1}^{k_j^i} z_{\alpha,i} J_{k_j^i}^\alpha \mathbf{e}_{i,j}^1 = \sum_{i=1}^{v_j} P_j^i \sum_{\mu=1}^{\varkappa} y_{\mu,i} SF^\mu(-i\lambda_j). \tag{2.45}$$

Usando (2.45) y (2.41), se obtiene que el espacio generado por los vectores (2.39) contiene a los vectores de la forma

$$S^{-1} J_{k_j^i}^m \mathbf{e}_{i,j}^1 + S^{-1} \sum_{l=1}^{v_j}\sum_{\alpha=1}^{k_j^i} z_{\alpha,l} J_{k_j^i}^{\alpha+m} \mathbf{e}_{i,j}^1,\ \text{donde } m = 0, 1, ..., k_j^i;\ i = 1, ..., v_j. \tag{2.46}$$

Afirmamos que el vector de la forma

$$S^{-1}\mathbf{e}_{i,j}^1 + S^{-1} \sum_{l=1}^{v_j}\sum_{\alpha=1}^{k_j^i} z_{\alpha,l} J_{k_j^i}^{\gamma+\alpha} \mathbf{e}_{i,j}^1,\ \text{donde } \gamma = 1, ..., k_j^i;\ i = 1, ..., v_j, \tag{2.47}$$

se puede obtener como una combinación lineal de los vectores (2.46). En efecto, para $\gamma = 1$ se sigue directamente de (2.46) para $m = 0, 1$. Para $\gamma > 1$ puede probarse por inducción.

Como $J_{k_j^i}^\alpha = 0$ para $\alpha > k_j^i$, se obtiene que el espacio, generado por los vectores (2.40) contiene a los vectores $e_{i,j}^1$, obteniendo así que, la aplicación de las matrices $(A - \lambda_j I)^m$, $m = 0, 1, ..., k_j^i$, a vectores (2.40) generan a todos los espacios

$$S^{-1}(\oplus_{i=1}^{v_j}\mathbb{C}(i, j)),\ \text{con } j = 1, ..., q.$$

Por lo tanto, la ecuación (2.39) implica que la matriz $(A - A_1)$ es igual a la matriz cero en el espacio total $\mathbb{C}^n$. Por ello, $A = A_1$ si $\Phi^\varkappa \notin N_r(A)$.

Se sigue de (2.36) que, $\sum_{\mu=0}^{v}(f_\mu^\beta - f_\mu^{*\beta})\omega^\mu \equiv 0$ para $\beta = 1, ..., \varkappa$; $\varkappa \geq \varkappa_A$. Por lo tanto, $f_\mu^\beta = f_\mu^{*\beta}$ si $\Phi^\varkappa \notin N_r(A)$. Esto completa la prueba del punto I.1.

El punto I.2 no lo probare aquí, puede ver la referencia [20] para su demostración.

Ahora, considere el punto I.3. El kernel del sistema (2.10) es una base para el conjunto de soluciones de (2.36) (matrices A_1 y coeficientes f_k^α). Pero esta ecuación no depende de la elección de frecuencias no-resonantes $\omega_1, ..., \omega_l$, $l \geq (\deg P_A + v + 1)$. Por tanto el kernell no depende de la elección de las frecuencias $\omega_j's$.

II) Si el sistema (2.10) tiene solución única, entonces por el punto I.1 de este teorema se obtiene que la matriz A, y los coeficientes f_k^α satisfacen la condición de resolución. ■

2.4. Sistemas de Segundo Orden

Considere el siguiente sistema de segundo orden

$$M\frac{d^2y}{dt^2} + C\frac{dy}{dt} + Ky = F(\omega)e^{i\omega t}, \tag{2.48}$$

donde $y \in \mathbb{C}^n$; $F(\omega) = \sum_{\mu=0}^{v} f_\mu \omega^\mu$, $f_\mu \in \mathbb{C}^n$; y M, C y K son matrices $n \times n$ tales que K^{-1} y M^{-1} existan.

Es posible reducir (2.48) a un sistema de primer orden:

$$\begin{cases} \frac{dy}{dt} = M^{-1}z, \\ \frac{dz}{dt} = -CM^{-1}z - Ky + F(\omega)e^{i\omega t} \end{cases}, \; A = \begin{bmatrix} 0 & M^{-1} \\ -K & -CM^{-1} \end{bmatrix}. \tag{2.49}$$

Obviamente, si los $\varkappa$ vectores

$$\tilde{F}^\alpha(-iA) := \sum_{j=1}^{v}(-iA)^j \tilde{f}_j^\alpha, \tag{2.50}$$

satisfacen la condición de resolución del lema (1), donde $\tilde{f}_j = (0, f_j^\alpha)^\tau$; $\alpha = 1, ..., \varkappa$; $j = 1, ..., q$, aquí estos f_μ^α provienen del polinomio original $F^\alpha(\omega) = \sum_{\mu=0}^{v} f_\mu^\alpha \omega^\mu$, entonces el lema (1) es válido para el sistema (2.49) en $(2n + v + 1)$ frecuencias ω_j.

En este caso, se tiene una restricción natural sobre los coeficientes vectoriales $\tilde{f}_j^\alpha$. A saber, sus primeras n componentes son iguales a cero; y en lugar de la esfera $S^{4n\varkappa(v+1)-1}$ es necesario considerar la condición de resolución sobre la esfera $S^{2n\varkappa(v+1)-1}$.

Aquí, se considera el caso cuando la matriz A tiene $2n$ valores propios $\lambda_1, ..., \lambda_{2n}$ diferentes. Por lo tanto, se puede aplicar el lema (1). Sea $\tilde{N}_r(A)$ el conjunto definido como la unión de los $2n$ planos π_k, es decir

$$\pi_k = \left(\sum_{j=0}^{v} \tilde{f}_j (-i\lambda_k)^j, \mathbf{e}_k^* \right), \ \tilde{N}_r(A) = \cup_{k=1}^{k=2n} \pi_k. \tag{2.51}$$

Aquí $A\mathbf{e}_k^* = \lambda_k \mathbf{e}_k^*$. Ahora considere el vector Φ definido como

$$\Phi = \left(\frac{f_0}{\Psi}, ..., \frac{f_v}{\Psi} \right), \ \Psi^2 = \sum_{j=1}^{v} (f_j, f_j), \tag{2.52}$$

sobre la esfera unitaria real $S^{2n(v+1)-1}$ de dimensión $2n(v+1)-1$ en el espacio de coeficientes $f_0, ..., f_v$. Por definición, pongamos $N_r(A) = \tilde{N}_r(A) \cap S^{2n(v+1)-1}$. Evidentemente, $med(N_r(A)) = 0$ y el conjunto $N_r(A)$ es cerrado [20]. Note que, si definimos el conjunto $N_r(A)$ como $N_r(A) := \tilde{N}_r(A) \cap S^{2n(v+1)-1}$, entonces podemos aplicar el lema (1) a (2.49). Esto se prueba directo.

Capítulo 3

Funciones Características

Una solución simple para una amplia gama de problemas de la teoría de probabilidad, especialmente aquellos asociados con la suma de variables aleatorias independientes, se obtiene por medio de funciones características. Dicha teoría, la cual fue desarrollada en análisis, es conocida por el nombre de transformaciones de Fourier. En este capítulo se dan algunas propiedades básicas de funciones características.

3.1. Definición y Propiedades de las Funciones Características

La función característica de una variable aleatoria ξ es definida como la esperanza de la variable aleatoria $e^{it\xi *}$. Si $F(x)$ es la función de distribución de la variable ξ, la función característica es

$$f(t) = \int e^{itx} dF(x) \tag{3.1}$$

Acordemos ahora en adelante, para denotar a la función y la correspondiente función de distribución por las mismas letras, el caso de minúscula para el primero y el caso de mayúscula para el último.

Del hecho que $|e^{itx}| = 1 \ \forall \ t \in \mathbb{R}$, se sigue la existencia de la integral (3.1) para todas las funciones de distribución: por lo tanto, una función característica puede verse como una variable aleatoria.

Teorema 2 *Una función característica es uniformemente continua sobre toda la línea y satisface las siguientes relaciones:*

$$f(0) = 1, |f(t)| \leq 1 \qquad (-\infty < t < \infty) \tag{3.2}$$

Demostración. La relación (3.2) se sigue de la definición de una función característica. Por (3.1)

$$f(0) = \int 1 \cdot dF(x) = 1 \text{ y}$$

$$|f(t)| = \left| \int e^{itx} dF(x) \right| \leq \int \left| e^{itx} \right| dF(x) = \int dF(x) = 1.$$

Resta probar la continuidad uniforme de la función $f(t)$. Para este propósito, se considera la diferencia

$$f(t+h) - f(t) = \int e^{itx}(e^{ixh} - 1) dF(x)$$

y ahora se estima su valor absoluto. Se tiene

$$|f(t+h) - f(t)| \leq \int \left| e^{ixh} - 1 \right| dF(x)$$

Sea $\varepsilon > 0$ arbitrario; eligiendo a A suficientemente grande tal que

$$\int\limits_{|x|>A} dF(x) < \varepsilon/4$$

y seleccionamos h tan pequeño para que si,$|x| < A$

$$\left| e^{ixh} - 1 \right| < \varepsilon/2$$

Entonces

$$|f(t+h) - f(t)| \leq \int_{-A}^{A} \left| e^{ixh} - 1 \right| dF(x) + 2 \int\limits_{|x|\geq A} dF(x) \leq \varepsilon$$

Esta desigualdad prueba el teorema. ■

Teorema 3 *Si $\eta = a\xi + b$, donde a y b son constantes, entonces*

$$f_\eta(t) = f_\xi(at) e^{ibt}$$

donde $f_\eta(t)$ y $f_\xi(t)$ denotan las funciones características de las variables η y ξ, respectivamente.

Demostración. En efecto,

$$f_\eta(t) = \mathbf{M}e^{it\eta} = \mathbf{M}e^{it(a\xi+b)} = e^{ibt}\mathbf{M}e^{i(ta)\xi} = e^{ibt}f_\xi(at)$$

■

Teorema 4 *La función característica de la suma de dos variables aleatorias independientes es igual al producto de sus funciones características.*

Demostración. Sean ξ y η variables aleatorias independientes y sea $\zeta = \xi+\eta$. Entonces, claramente, $e^{it\xi}$ y $e^{it\eta}$ son también variables aleatorias con respecto a ξ y η. De esto se sigue inmediatamente que

$$\mathbf{M}e^{it\zeta} = \mathbf{M}e^{it(\xi+\eta)} = \mathbf{M}e^{it\xi}e^{it\eta} = \mathbf{M}e^{it\xi} \cdot \mathbf{M}e^{it\eta}$$

Esto prueba el teorema. ■

Corollary 5 *Si $\xi = \xi_1 + \xi_2 + ... + \xi_n$ y cada término es independiente de la suma de sus predecesores, entonces la función característica de la variable ξ es igual al producto de sus funciones características que conforman la suma.*

La aplicación de funciones características hace muchas referencias a la propiedad formulada en el teorema (4). Como sabemos, la suma de las variables aleatorias independientes conduce a una operación sumamente complicada, la convolución de la función de distribución de cada uno de los sumandos. Con referencia a las funciones características, esta operación compleja es reemplazada por una más simple, la multiplicación de las funciones características.

Teorema 6 *Si una variable aleatoria ξ tiene un momento absoluto de orden n-ésimo, entonces la función característica de la variable ξ es diferenciable n veces y cuando $k \leq n$*

$$f^{(k)}(0) = i^k\mathbf{M}\xi^k. \tag{3.3}$$

Demostración. En efecto por diferenciación formal de la función característica k veces ($k \leq n$) se tiene la ecuación

$$f^{(k)}(t) = i^k \int x^k e^{itx} dF(x) \tag{3.4}$$

Pero

$$\left| \int x^k e^{itx} dF(x) \right| \leq \int \left| x^k \right| dF(x)$$

y, consecuentemente, por hipótesis del teorema, esto es acotado. Se sigue de lo último que, (3.4) existe y la diferenciación es legítima. Al tomar $t = 0$ en (3.4) se encuentra que

$$f^{(k)}(0) = i^k \int x^k dF(x).$$

La esperanza y la varianza son simplemente expresadas por medio de sus derivadas del logaritmo de la función característica. En efecto, defínase

$$\psi(t) = \log f(t)$$

Entonces

$$\psi'(t) = \frac{f'(t)}{f(t)}$$

y

$$\psi''(t) = \frac{f''(t) \cdot f(t) - [f'(t)]^2}{f^2(t)}$$

Tomando la información de la ecuación (3.3) y que $f(0) = 1$, se encuentra que

$$\psi'(0) = f'(0) = i\mathbf{M}\xi$$

y

$$\psi''(0) = f''(0) - [f'(0)]^2 = i^2\mathbf{M}\xi^2 - [i\mathbf{M}\xi]^2 = -\mathbf{D}\xi$$

De donde

$$\begin{aligned} \mathbf{M}\xi &= \frac{1}{i}\psi'(0) \\ \mathbf{D}\xi &= -\psi''(0). \end{aligned} \tag{3.5}$$

La k-ésima derivada del logaritmo de la función característica en el punto 0, multiplicado por i^k, es llamado acumulado (semi-invariante) del k-ésimo orden de la variable aleatoria. Se sigue directamente del teorema (4), se agregan variables aleatorias sus acumulados también se agregan. Hemos visto justamente que los primeros dos acumulados son la esperanza y la varianza, esto es, el primer momento y una función racional de los momentos, 1^o y 2^o orden. Esto se puede fácilmente, por medio de de operaciones, así el acumulado de cualquier orden k es enteramente una función racional de los primeros k momentos. A manera de ilustración, damos las expresiones explícitas de los acumulados 3^o y 4^o orden:

$$\begin{aligned} i\psi'''(0) &= -\{\mathbf{M}\xi^3 - 3\mathbf{M}\xi^2 \cdot \mathbf{M}\xi + 2[\mathbf{M}\xi]^3\} \\ i\psi^{iv}(0) &= \mathbf{M}\xi^4 - 4\mathbf{M}\xi^3 \cdot \mathbf{M}\xi - 3[\mathbf{M}\xi^2]^2 + 12\mathbf{M}\xi^{\mathbf{2}}[\mathbf{M}\xi]^2 - 6[\mathbf{M}\xi]^4. \end{aligned}$$

■

Consideremos unos cuantos ejemplos de funciones características.

Ejemplo 1 *Una variable aleatoria ξ tiene una distribución con ley normal de media μ y varianza σ^2. Entonces la función característica de la variable ξ es*

$$\varphi(t) = \frac{1}{\sigma\sqrt{2\pi}} \int e^{itx - \frac{(x-a)^2}{2\sigma^2}} dx.$$

Por la sustitución $z = \frac{x-a}{\sigma} - it\sigma$, $\varphi(t)$ se reduce a la forma

$$\varphi(t) = e^{iat - \frac{\sigma^2 t^2}{2}} \frac{1}{\sqrt{2\pi}} \int_{-\infty + it\sigma}^{+\infty - it\sigma} e^{-\frac{z^2}{2}} dz.$$

Esto es conocido que $\forall \alpha \in \mathbb{R}$ se cumple

$$\int_{-\infty + it\sigma}^{+\infty - it\sigma} e^{-\frac{z^2}{2}} dz = \sqrt{2\pi}.$$

Por lo tanto, $\varphi(t) = e^{iat - \frac{\sigma^2 t^2}{2}}$.

Usando el teorema (6) pueden calcularse fácilmente los momentos centrales para la distribución normal y en esta forma alternativa obtener, el resultado muy conocido

$$\mu_k = m_k = \sqrt{\frac{2}{\pi}} \sigma^k 2^{\frac{k-1}{2}} \Gamma(\frac{k+1}{2}) = \sqrt{\frac{2}{\pi}} 2^{\frac{k-1}{2}} (\frac{k-1}{2})! \sigma^k$$

Ejemplo 2 *Encontrar la función característica de una variable aleatoria ξ que tiene distribución de Poisson.*

Por hipótesis la variable toma solo valores enteros, y

$$P\{\xi = k\} = \frac{\lambda^k e^{-\lambda}}{k!} \ (k = 0, 1, 2, \ldots)$$

donde $\lambda > 0$ es una constante.

La función característica de la variable ξ es

$$f(t) = \mathbf{M} e^{it\xi} = \sum_{k=0}^{\infty} e^{ikt} P\{\xi = k\} = \sum_{k=0}^{\infty} e^{ikt} \frac{\lambda^k e^{-\lambda}}{k!}$$

$$= e^{-\lambda} \sum_{k=0}^{\infty} \frac{(e^{it}\lambda)^k}{k!} = e^{-\lambda + e^{it}\lambda}.$$

De acuerdo a (3.5), se encuentra que
$\mathbf{M}\xi = \frac{1}{i}\psi'(0) = \lambda$, $\mathbf{D}\xi = -\psi''(0) = \lambda$.

La primera de estas ecuaciones se obtiene fácilmente de la definición de esperanza y la otra análogamente.

Ejemplo 3 *Una variable aleatoria ξ tiene una distribución uniforme sobre el intervalo $(-a, a)$. La función característica es igual a:*

$$f(t) = \int_{-a}^{a} e^{itx}\frac{dx}{2a} = \frac{\sin at}{at}.$$

Ejemplo 4 *Encontrar la función característica de la variable μ, la cual es igual al número de ocurrencias de un evento A en n ensayos independientes, en cada uno la probabilidad de que A ocurra es p.*

La variable μ puede representarse en la forma de una suma

$$\mu = \mu_1 + \mu_2 + ... + \mu_n$$

de n variables independientes, cada una de las cuales toma solo uno de los dos valores,0 y 1,con probabilidad $q = 1 - p$ y p, respectivamente. La variable μ_k toma el valor 1 si el evento A ocurre en el k-ésimo ensayo y toma el valor 0 si el evento A no ocurre en el k-ésimo ensayo.

La función característica de μ_k es igual a

$$f_k(t) = \mathbf{M}e^{it\mu_k} = e^{it\cdot 0}q + e^{it\cdot 1}p = q + pe^{it}.$$

De acuerdo al teorema (4), la función característica de la variable μ es

$$f(t) = \prod_{k=1}^{n} f_k(t) = (q + pe^{it})^n.$$

Se encuentra también la función característica de la variable $\eta = \frac{\mu - np}{\sqrt{npq}}$.

Por el teorema (3), esto es

$$f_\eta(t) = e^{-it\sqrt{\frac{np}{q}}} f(\frac{t}{\sqrt{npq}}) = e^{-it\sqrt{\frac{np}{q}}}(q + pe^{i\frac{t}{\sqrt{npq}}})^n$$

$$= (qe^{-it\sqrt{\frac{p}{nq}}} + pe^{it\sqrt{\frac{q}{np}}})^n.$$

Ejemplo 5 *La función característica para una variable aleatoria satisface la ecuación*

$$f(-t) = \overline{f(t)}.$$

En efecto,

$$f(-t) = \int e^{-itx}dF(x) = \int e^{itx}dF(x) = \overline{f(t)}.$$

3.2. La Fórmula de Inversión y el Teorema de Unicidad

Hemos visto que de la función de distribución de una variable aleatoria ξ, esto es, $F(x)$, siempre es necesaria para encontrar su función característica. También es importante que la proposición recíproca se cumpla; en otras palabras, una función de distribución es determinada de manera única por su función característica.

Teorema 7 *Sean $f(t)$ y $F(x)$ la función característica y la función de distribución, respectivamente, de una variable aleatoria ξ. Si x_1 y x_2 son puntos de continuidad de la función $F(x)$, entonces*

$$F(x_2) - F(x_1) = \frac{1}{2\pi} \lim_{c\to\infty} \int_{-c}^{c} \frac{e^{itx_1} - e^{itx_2}}{it} f(t)dt \tag{3.6}$$

Demostración. De la definición de una función característica se sigue que la integral

$$J_c = \frac{1}{2\pi} \int_{-c}^{c} \frac{e^{itx_1} - e^{itx_2}}{it} f(t)dt \text{ es igual a}$$

$$J_c = \frac{1}{2\pi} \int_{-c}^{c} \int \frac{1}{it} [e^{it(z-x_1)} - e^{it(z-x_2)}] dF(z)dt$$

El orden de integración puede cambiarse en esta última integral, ya que la integral converge absolutamente con respecto a z, y los límites de integración son finitos con respecto a t. Por lo tanto,

$$J_c = \frac{1}{2\pi} \int [\int_{-c}^{c} \frac{e^{it(z-x_1)} - e^{it(z-x_2)}}{it} dt] dF(z)$$

$$\begin{aligned} &= \frac{1}{2\pi} \int [\int_{0}^{c} \frac{e^{it(z-x_1)} - e^{-it(z-x_1)} - e^{it(z-x_2)} - e^{-it(z-x_2)}}{it} dt] dF(z) \\ &= \frac{1}{\pi} \int_{-\infty}^{\infty} \int_{0}^{c} [\frac{\sin t(z-x_1)}{t} - \frac{\sin t(z-x_2)}{t}] dt dF(z) \end{aligned}$$

Del análisis es conocido el hecho, que cuando $c \to \infty$

$$\frac{1}{\pi} \int_{0}^{c} \frac{\sin \alpha t}{t} dt \to \begin{cases} 1/2 \ si \ \alpha>0 \\ -1/2 \ si \ \alpha<0 \end{cases} \tag{3.7}$$

y esta convergencia es uniforme con respecto a α en cada región $\alpha > \delta > 0$ (o $\alpha < -\delta$) y cuando $|\alpha| < \delta, \forall c$,

$$\left|\frac{1}{\pi}\int_0^c \frac{\sin \alpha t}{t} dt\right| < 1. \tag{3.8}$$

Suponga que $x_2 > x_1$ y represente la integral J_c como sigue:

$$J_c = \int_{-\infty}^{x_1-\delta} + \int_{x_1-\delta}^{x_1+\delta} + \int_{x_1+\delta}^{x_2-\delta} + \int_{x_2-\delta}^{x_2+\delta} + \int_{x_2+\delta}^{+\infty} \psi(c, z; x_1, x_2) dF(z)$$

Donde por brevedad se usa la notación

$$\psi(c, z; x_1, x_2) = \frac{1}{\pi}\int_0^c [\frac{\sin t(z - x_1)}{t} - \frac{\sin t(z - x_2)}{t}]dt$$

y $\delta > 0$ es escogido tal que $x_1 + \delta < x_2 - \delta$.

Las desigualdades $z - x_1 < -\delta$ y $z - x_2 < -\delta$ se tienen en la región $-\infty < z < x_1 - \delta$.

Por lo tanto, concluimos en base a (3.7), que cuando $c \to \infty$

$$\int_{-\infty}^{x_1-\delta} \psi(c, z; x_1, x_2) dF(z) \to 0.$$

Similarmente, cuando $x_2 + \delta < z < \infty$ y cuando $c \to \infty$,

$$\int_{x_2+\delta}^{+\infty} \psi(c, z; x_1, x_2) dF(z) \to 0.$$

Además, ya que en la región $x_1 + \delta < z < x_2 - \delta$ las desigualdades $z - x_1 > \delta$ y $z - x_2 < -\delta$ son válidas, se sigue de (3.7) que cuando $c \to +\infty$,

$$\int_{x_1+\delta}^{x_2-\delta} \psi(c, z; x_1, x_2) dF(z) \to \int_{x_1+\delta}^{x_2-\delta} dF(z) = F(x_2 - \delta) - F(x_1 + \delta).$$

Finalmente por (3.8) se pueden estimar las restantes integrales

$$\left|\int_{x_1-\delta}^{x_1+\delta} \psi(c, z; x_1, x_2) dF(z)\right| < 2\int_{x_1-\delta}^{x_1+\delta} dF(z) = 2[F(x_1 + \delta) - F(x_1 - \delta)]$$

$$\left|\int_{x2-\delta}^{x_2+\delta} \psi(c, z; x_1, x_2) dF(z)\right| < 2\int_{x_2-\delta}^{x_2+\delta} dF(z) = 2[F(x_2 + \delta) - F(x_2 - \delta)]$$

Por tanto, se encuentra que $\forall \alpha > 0$
$\lim\sup_{c\to\infty} J_c = F(x_2 - \delta) - F(x_1 + \delta) + R_1(\delta, x_1, x_2)$,
$\lim\inf_{c\to\infty} J_c = F(x_2 - \delta) - F(x_1 + \delta) + R_2(\delta, x_1, x_2)$,
donde

$$\begin{aligned} |R_i(\delta, x_1, x_2)| &< 2\{F(x_1 + \delta) - F(x_1 - \delta) + F(x_2 + \delta) - F(x_2 - \delta)\} \\ (i &= 1, 2). \end{aligned}$$

Ahora sea $\delta \to 0$. Del hecho que x_1 y x_2 son puntos de continuidad de la función $F(x)$ se sigue que

$$\lim_{\delta\to 0} F(x_1 + \delta) = \lim_{\delta\to 0} F(x_1 - \delta) = F(x_1)$$

y

$$\lim_{\delta\to 0} F(x_2 + \delta) = \lim_{\delta\to 0} F(x_2 - \delta) = F(x_2).$$

Y como J_c no depende de δ, se sigue que $\lim_{c\to\infty} J_c = F(x_2) - F(x_1)$. ■

La ecuación (3.6) es llamada la fórmula de inversión. Usaremos esto para derivar la siguiente proposición importante (el teorema de unicidad).

Teorema 8 *Una función de distribución está únicamente determinada por su función característica.*

Demostración. En efecto, se sigue directamente del teorema (7), que la fórmula se tiene en cada punto de continuidad de la función $F(x)$:

$$F(x) = \frac{1}{2\pi} \lim_{y\to-\infty} \lim_{c\to+\infty} \int_{-c}^{+c} \frac{e^{-ity} - e^{-itx}}{it} f(t)dt$$

donde el límite en y se evalúa con respecto a cualquier conjunto de puntos y que son puntos de continuidad de la función $F(x)$. ■

Como una aplicación del último teorema pueden probarse sin dificultad, las siguientes proposiciones.

Ejemplo 6 *Si las variables aleatorios independientes ξ_1 y ξ_2 son normalmente distribuidas, entonces su suma $\xi = \xi_1 + \xi_2$ tiene también distribución normal.*

En efecto, sí

$$\mathbf{M}\xi_1 = a_1, \mathbf{D}\xi_1 = \sigma_1^2, \mathbf{M}\xi_2 = a_2, \mathbf{D}\xi_2 = \sigma_2^2,$$

entonces las funciones características de las variables ξ_1 y ξ_2 son

$$f_1(t) = e^{ia_1 t - \frac{1}{2}\sigma_1^2 t^2}, f_2(t) = e^{ia_2 t - \frac{1}{2}\sigma_2^2 t^2},$$

Por el teorema (4), la función característica $f(t)$ de la suma es igual a

$$f(t) = f_1(t) \cdot f_2(t) = e^{i(a_1+a_2)t-\frac{1}{2}(\sigma_1^2+\sigma_2^2)t^2}$$

Esta es la función característica de una ley normal con esperanza $a = a_1 + a_2$ y varianza $\sigma^2 = \sigma_1^2 + \sigma_2^2$ sobre la base del teorema de unicidad, concluimos que la función de distribución de la variable ξ es normal.

La proposición recíproca debida a H. Cramer, puede establecerse como sigue en términos de funciones características: sí $f_1(t)$ *y* $f_2(t)$ son funciones características y

$$f_1(t) \cdot f_2(t) = e^{-t^2/2}$$

entonces

$$f_1(t) = e^{iat-\sigma^2\frac{t^2}{2}},\ f_2(t) = e^{-iat-\frac{(1-\sigma^2)t^2}{2}}\ (0 \leq \sigma \leq 1)$$

Ejemplo 7 *Si las variables aleatorias independientes* ξ_1 *y* ξ_2 *tienen distribución de acuerdo a la ley de Poisson, y*

$$P\{\xi_1 = k\} = \frac{\lambda_1^k e^{-\lambda_1}}{k!}, P\{\xi_2 = k\} = \frac{\lambda_2^k e^{-\lambda_2}}{k!}$$

Entonces la variable aleatoria $\xi = \xi_1 + \xi_2$ *tiene distribución de Poisson con parámetro* $\lambda = \lambda_1 + \lambda_2$.

En efecto, en el ejemplo (2) de la sección precedente se encuentra que, las funciones características de las variables aleatorias ξ_1 y ξ_2 son

$$f_1(t) = e^{\lambda_1(e^{it}-1)},\ f_2(t) = e^{\lambda_2(e^{it}-1)}$$

Por el teorema (4) de la sección precedente, la función característica de la suma $\xi = \xi_1 + \xi_2$ es

$$f(t) = f_1(t) \cdot f_2(t) = e^{(\lambda_1+\lambda_2)(e^{it}-1)},$$

esto es, su función característica es de alguna ley de Poisson. Por el teorema de unicidad, solo la distribución con $f(t)$ como función característica, es ley de Poisson, para la cual

$$P\{\xi = k\} = \frac{(\lambda_1 + \lambda_2)^k e^{-(\lambda_1+\lambda_2)}}{k!}\ (k \geq 0)$$

D. A. Raikov probó la proposición recíproca más profunda:

Si la suma de dos variables aleatorias independientes son distribuidas con una ley de Poisson, entonces cada sumando es también distribuido de acuerdo con la ley de Poisson.

Ejemplo 8 *Una función característica es real si y solo si la correspondiente función de distribución es simétrica, esto es, cuando la función de distribución satisface la ecuación*

$$F(x) = 1 - F(-x + 0)$$

para todo x.

Si la función de distribución es simétrica, entonces su función característica es real. Si ξ tiene una función de distribución simétrica, entonces ambas ξ y $-\xi$ son idénticamente distribuidos, por lo tanto la ecuación

$$f(t) = \mathbf{M}e^{it\xi} = \mathbf{M}e^{-it\xi} = f(-t) = \overline{f(t)}$$

se sigue de aquí que $f(t)$ es real.

Para probar la proposición recíproca, consideremos la variable aleatoria $\eta = -\xi$. La función de distribución de variable η es

$$G(x) = P\{\eta < x\} = P\{\xi > -x\} = 1 - F(-x + 0).$$

Las funciones características de las variables ξ y η son conectadas por la relación

$$g(t) = \mathbf{M}e^{it\eta} = \mathbf{M}e^{-it\xi} = \overline{\mathbf{M}e^{it\xi}} = \overline{f(t)}.$$

Ya que por hipótesis, $f(t)$es real, se sigue que $\overline{f(t)} = f(t)$ y por lo tanto $g(t) = f(t)$.

Del teorema de unicidad, ahora concluimos que la funciones de distribuciones de las variables ξ y η coinciden, esto es, que

$$f(x) = 1 - F(-x + 0),$$

y esto completa la prueba.

Capítulo 4

El Teorema de Límite Clásico y sus Aplicaciones

4.1. Definición del Problema

El teorema de la integral de DeMoire-Laplace servirá como el origen a un amplio rango de investigaciones de fundamental importancia ambos para la teoría de probabilidad misma y para sus múltiples aplicaciones en la ciencia natural, tecnología y la ciencia económica. Para dar una idea de la tendencia de estas investigaciones analizaremos el teorema de DeMoire-Laplace en alguna otra forma. A saber, sí, como frecuentemente tenemos el hecho, denotaremos por μ_k el número de ocurrencias de un evento A en el k-ésimo ensayo, entonces el número de ocurrencias de A en n sucesivos ensayos es igual a $\sum_{k=1}^{n} \mu_k$. Además, de probabilidad se sabe que $\mathbf{M}(\sum_{k=1}^{n} \mu_k) = np$ y $\mathbf{D}(\sum_{k=1}^{n} \mu_k) = npq$. Por lo tanto, el teorema de DeMoire-Laplace puede escribirse como sigue:

$$P\left\{a \leq \frac{\sum_{k=1}^{n}(\mu_k - \mathbf{M}\mu_k)}{\sqrt{\sum_{k=1}^{n} \mathbf{D}\mu_k}} \leq b\right\} \to \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{z^2}{2}} dz \tag{4.1}$$

cuando $n \to \infty$. En palabras: la probabilidad que la suma de desviaciones de variables aleatorias independientes —las cuales toman solo dos valores, 0 y 1, con probabilidades respectivamente igual a q y $p = 1 - q$ $(0 < p < 1)$ —de sus valores esperados (esperanzas) divididas entre la raíz cuadrada de la suma de las varianzas de los sumandos que se encuentran entre los límites de a a b tiende a la integral

$$\frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{z^2}{2}} dz$$

uniformemente en a y b. Cuando el número de sumandos crece al infinito.

Las preguntas naturales inmediatas son: ¿cómo atar estrechamente la ecuación (4.1) con la elección especial de sumandos μ_k? ¿esto no se tiene en caso de debilitar las restricciones impuestas a las funciones de distribuciones de los sumandos?. Las afirmaciones de estos problemas y también sus soluciones corresponden en principio a P. L. Chebyshev y sus alumnos A. A. Markov y A. M. Lyapunov. Sus investigaciones mostraron que debíamos imponer sobre los sumandos, solo las restricciones generales, el significado del cual depende sobre el hecho que los sumandos separados deberán ejercer un efecto significante sobre la suma. En la sección siguiente daremos una afirmación precisa de esta condición. Las razones por que estos resultados son inmensamente importantes en aplicaciones están en la gran esencia de fenómenos de escala-masa, el estudio de las regularidades de las cuales es, como ya tenemos ocasión para decirlo, es el sujeto actual de la teoría de la probabilidad.

Uno de los más importantes proyectos usados para explotación de los resultados de teoría de la probabilidad, en la ciencia natural y tecnología, consiste en lo siguiente. Se supone que un proceso ocurre bajo la influencia de un número grande de factores aleatorios operando independientemente, cada uno de los cuales solo causan una modificación insignificante en el curso del fenómeno o proceso. El investigador quien es interesado en el proceso, como un todo y no en la operación de separar los factores, observa solo la operación general de estos factores. Ilustraremos con dos ejemplos típicos.

Ejemplo 9 *Sea a una cantidad medible. En cuyo resultado es inevitable la influencia por gran número de factores que generan errores en la medición. Estos errores se deben a las condiciones en las que están los instrumentos de medida, los cuales podrían cambiar en forma abrupta las mediciones bajo los efectos de varios factores atmosféricos y mecánicos. Hay errores que se cometen por el humano, causados por la visión y el oído, y también estos pueden alterarse insignificantemente por lo físico o el estado físico del observador y cosas así. Cada uno de estos factores, debería generar un error insignificante. Pero la medición es afectada en una sola vez por todos estos errores, el resultado es un error total general. En otras palabras, los errores observados en mediciones actualmente serán una variable aleatoria —la suma de un enorme número de cantidades pequeñas que son variables aleatorias. Y aunque estas cantidades son desconocidas, como también sus distribuciones, sus efectos sobre los resultados de las mediciones son notables y por esta razón deberá ser el sujeto de estudio.*

Ejemplo 10 *En muchas grandes hornadas o lotes industriales de artículos idénticos se producen por los procesos de masa-producción. Consideramos alguna característica*

numérica del producto que estemos interesados. En la medida en que el artículo se ajusta a cierta técnica estándar, hay un cierto valor estándar de la característica que elegimos. En la actualidad, de cualquier modo, siempre existe una cierta desviación de este valor estándar. En un proceso de producción propiamente organizado tales desviaciones pueden ser generadas debido a factores aleatorios, cada uno de los cuales produce un efecto innotable. La acción total, de cualquier modo genera una desviación notable de la normal.

Cualquier número de tales ejemplos serán mencionados, por lo tanto, así nos deja el problema de estudiar las peculiaridades de las regularidades de la suma de un gran número de variables aleatorias independientes, cada una de las cuales ejerce un efecto insignificante sobre la suma. Después sobre la marcha daremos el significado de los requerimientos más precisos. En lugar de estudiar la suma de un gran número finito de sumandos, consideramos la suma de gran número de elementos y supongamos que las soluciones de los problemas son dadas por límites de funciones de distribuciones por una sucesión de funciones de distribución de la suma. Esta clase del paso de una condición finita del problema a una condición de límite es una costumbre ambos de matemáticas modernas y en varias divisiones de la ciencia natural.

Hemos arribado el siguiente problema: dada una sucesión de variables aleatorias independientes $\xi_1, \xi_2, ..., \xi_n, ...$ las cuales se suponen que tiene esperanzas y varianzas finitas. De ahora en adelante adherimos las siguientes notaciones:

$$a_k = \mathbf{M}\xi_k,\ b_k^2 = \mathbf{D}\xi_k, B_n^2 = \sum_{k=1}^{n} b_k^2 = \mathbf{D}\sum_{k=1}^{n}\xi_k.$$

La pregunta es: ¿qué condiciones se deberán imponerse a las variables ξ_k tal que las funciones de distribuciones o distribuciones de la suma

$$\frac{1}{B_n}\sum_{k=1}^{n}(\xi_k - a_k) \tag{4.2}$$

converja a una ley de distribución normal. En la siguiente sección veremos que para esta proposición es suficiente que la condición de Lindeberg se satisfaga: para cualquier $\tau > 0$

$$\lim_{n\to\infty}\frac{1}{B_n^2}\sum_{k=1}^{n}\int\limits_{|x-a_k|>\tau B_n}(x-a_k)^2 dF_k(x) = 0$$

donde $F_k(x)$ denota a la función de distribución de la variable ξ_k.

Clasificaremos el significado de esta condición.

Denotemos por A_k el evento el que consiste en el hecho

$$|\xi_k - a_k| > \tau B_n \ (k = 1, 2, ..., n)$$

y estimamos la probabilidad

$$P\{\max_{1\leq k\leq n} |\xi_k - a_k| > \tau B_n\},$$

ya que

$$P\{\max_{1\leq k\leq n} |\xi_k - a_k| > \tau B_n\} = P\{A_1 \cup A_2 \cup \cdots \cup A_n\}$$

y

$$P\{A_1 \cup A_2 \cup \cdots \cup A_n\} \leq \sum_{k=1}^{n} P\{A_k\}.$$

Por notación que

$$P\{A_k\} = \int\limits_{|x-a_k|>\tau B_n} dF_k(x) \leq \frac{1}{(\tau B_n)^2} \int\limits_{|x-a_k|>\tau B_n} (x - a_k)^2 dF_k(x),$$

se puede encontrar la desigualdad

$$P\{\max_{1\leq k\leq n} |\xi_k - a_k| \geq \tau B_n\} \leq \frac{1}{(\tau B_n)^2} \sum_{k=1}^{n} \int\limits_{|x-a_k|>\tau B_n} (x - a_k)^2 dF_k(x)$$

En virtud de la condición de Lindeberg, la última suma tiende a cero cuando $n \to \infty$

para cualquier constante $\tau > 0$. Por lo tanto, la condición de Lindeberg es una clase peculiar de demanda o condición para la convergencia uniforme de los pequeños términos$\frac{1}{B_n}(\xi_k - a_k)$en la suma (4.2).

Notemos una vez más que el significado de las condiciones que garantizan la convergencia de las distribuciones de la suma (4.2) a la ley normal están totalmente eludidas en las investigaciones de A. A. Markov y A. M. Lyapunov.

4.2. Teorema de Lyapunov

Comenzaremos por la demostración de la suficiencia de la condición de Lindeberg.

Teorema 9 *Si una sucesión de variables aleatorias mutuamente independientes* $\xi_1, ..., \xi_n, ...$ *tales que para cualquier constante* $\tau > 0$ *satisface la condición de Lindeberg*

$$\lim_{n\to\infty} \frac{1}{B_n^2} \sum_{k=1}^{n} \int\limits_{|x-a_k|>\tau B_n} (x-a_k)^2 dF_k(x) = 0 \tag{4.3}$$

entonces cuando $n \to \infty$

$$P\{\frac{1}{B_n}\sum_{k=1}^{n}(\xi_k - a_k) < x\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{z^2}{2}} dz \tag{4.4}$$

uniformemente en x.

Demostración. Por brevedad se introducen las siguientes notaciones:

$$\begin{aligned} \xi_{nk} &= \frac{\xi_k - a_k}{B_n}, \\ F_{nk} &= P\{\xi_{nk} < x\}. \end{aligned}$$

Es obvio que $\mathbf{M}\xi_{nk} = 0$, $\mathbf{D}\xi_{nk} = \frac{1}{B_n^2}\mathbf{D}\xi_k$, y consecuentemente,

$$\sum_{k=1}^{n} \mathbf{D}\xi_k = 1. \tag{4.5}$$

Es fácil ver que en estas notaciones, la condición de Lindeberg está dada por

$$\lim_{n\to\infty} \sum_{k=1}^{n} \int\limits_{|x|>\tau} x^2 dF_{nk}(x) = 0. \tag{4.6}$$

La función característica de la suma

$$\frac{1}{B_n}\sum_{k=1}^{n}(\xi_k - a_k) = \sum_{k=1}^{n} \xi_{nk}$$

es igual a $\varphi_n(t) = \Pi_{k=1}^{n} f_{nk}(t)$.

Se tiene que probar que

$$\lim_{n\to\infty} \varphi_n(t) = e^{-\frac{t^2}{2}}.$$

Para este propósito primero se establece que los factores $f_{nk}(t)$ tienden a 1 uniformemente respecto a k $(1 \le k \le n)$ cuando $n \to \infty$. En efecto, tomando de acuerdo a la ecuación $\mathbf{M}\xi_{nk} = 0$, se encuentra que

$$f_{nk}(t) - 1 = \int (e^{itx} - 1 - itx) dF_{nk}(x)$$

Ya que para cualquier real α de la fórmula de Taylor, se sigue[1]

$$\left|e^{i\alpha} - 1 - i\alpha\right| \leq \frac{\alpha^2}{2}. \tag{4.8}$$

De esto se sigue que

$$|f_{nk}(t) - 1| \leq \frac{t^2}{2} \int x^2 dF_{nk}(x)$$

Sea $\varepsilon > 0$ arbitrario; entonces, claramente

$$\int x^2 dF_{nk}(x) = \int\limits_{|x|\leq\varepsilon} x^2 dF_{nk}(x) + \int\limits_{|x|>\varepsilon} x^2 dF_{nk}(x) \leq \varepsilon^2 + \int\limits_{|x|>\varepsilon} x^2 dF_{nk}(x)$$

Para n suficientemente grande y el último sumando puede, por (4.6), hacerse menos que ε^2. Por lo tanto, para n suficientemente grande y t en cualquier intervalo finito $|t| \leq T$, se llega a que

$$|f_{nk}(t) - 1| \leq \varepsilon^2 T^2$$

uniformemente en $k(1 \leq k \leq n)$. De esto se concluye que

$$\lim_{n\to\infty} f_{nk}(t) = 1 \tag{4.9}$$

uniformemente en $k(1 \leq k \leq n)$. Entonces para n suficientemente grande, para t dentro de un intervalo finito $|t| \leq T$,se tiene la desigualdad:

$$|f_{nk}(t) - 1| < \frac{1}{2}. \tag{4.10}$$

[1]Esta desigualdad y una serie de pasos similares puede deducirse como lo que sigue,

$$\left|e^{i\alpha} - 1\right| = \left|\int_0^\alpha e^{ix} dx\right| \leq \alpha \ (\alpha > 0)$$

tenemos la desigualdad

$$\left|e^{i\alpha} - 1 - i\alpha\right| = \left|\int_0^\alpha (e^{ix} - 1) dx\right| \leq \frac{\alpha^2}{2}$$

de esta última desigualdad se sigue que

$$\begin{aligned} \left|e^{i\alpha} - 1 - i\alpha + \frac{\alpha^2}{2}\right| &= \left|\int_0^\alpha (e^{ix} - 1 - ix) dx\right| \leq \int_0^\alpha \left|e^{ix} - 1 - ix\right| dx \\ &\leq \int_0^\alpha \frac{x^2}{2} dx = \frac{\alpha^3}{6} \end{aligned} \tag{4.7}$$

Podemos, por lo tanto, en el intervalo $|t| \leq T$ escribir la expansión (log representa el valor principal de logaritmo natural):

$$\begin{aligned} \log \varphi_n(t) &= \sum_{k=1}^{n} \log f_{nk}(t) = \sum_{k=1}^{n} \log[1 + (f_{nk}(t) - 1)] \\ &= \sum_{k=1}^{n} (f_{nk}(t) - 1) + R_n, \end{aligned} \tag{4.11}$$

donde

$$R_n = \sum_{k=1}^{n} \sum_{s=2}^{\infty} \frac{(-1)^s}{s!} (f_{nk}(t) - 1)^s.$$

En virtud de (4.10) se obtiene

$$\begin{aligned} |R_n| &\leq \sum_{k=1}^{n} \sum_{s=2}^{\infty} \frac{1}{2} |f_{nk}(t) - 1|^s = \frac{1}{2} \sum_{k=1}^{n} \frac{|f_{nk}(t) - 1|^2}{1 - |f_{nk}(t) - 1|} \\ &\leq \sum_{k=1}^{n} |f_{nk}(t) - 1|^2 . \end{aligned}$$

Ya que

$$\sum_{k=1}^{n} |f_{nk}(t) - 1| = \sum_{k=1}^{n} \left| \int (e^{itx} - 1 - itx) dF_{nk}(x) \right| \leq \frac{t^2}{2} \sum_{k=1}^{n} \int x^2 dF_{nk}(x) = \frac{t^2}{2}$$

de aquí se sigue que

$$|R_n| \leq \frac{t^2}{2} \max_{1 \leq k \leq n} |f_{nk}(t) - 1| .$$

De (4.9), se sigue que, en un intervalo finito arbitrario $|t| \leq T$, cuando $n \to \infty$

$$R_n \to 0 \tag{4.12}$$

uniformemente en t. Pero

$$\sum_{k=1}^{n} (f_{nk}(t) - 1) = -\frac{t^2}{2} + \rho_n \tag{4.13}$$

donde

$$\rho_n = \sum_{k=1}^{n} \int (e^{itx} - 1 - itx) dF_{nk}(x) + \frac{t^2}{2}.$$

Sea $\varepsilon > 0$ arbitrario; entonces por (4.5)

$$\rho_n = \sum_{k=1}^{n} \int_{|x|\leq\varepsilon} (e^{itx} - 1 - itx - \frac{(itx)^2}{2}) dF_{nk}(x) + \sum_{k=1}^{n} \int_{|x|>\varepsilon} (\frac{t^2x^2}{2} + e^{itx} - 1 - itx) dF_{nk}(x).$$

Las desigualdades (4.8) y (4.7) permiten obtener la siguiente estimación

$$\begin{aligned} |\rho_n| &\leq \frac{|t|^3}{6} \sum_{k=1}^{n} \int_{|x|\leq\varepsilon} |x|^3 \, dF_{nk}(x) + t^2 \sum_{k=1}^{n} \int_{|x|>\varepsilon} x^2 dF_{nk}(x) \\ &\leq \frac{|t|^3}{6} \varepsilon \sum_{k=1}^{n} \int_{|x|\leq\varepsilon} x^2 dF_{nk}(x) + t^2 \sum_{k=1}^{n} \int_{|x|>\varepsilon} x^2 dF_{nk}(x) \\ &= \frac{|t|^3}{6} \varepsilon + t^2 (1 - \frac{|t|}{6} \varepsilon) \sum_{k=1}^{n} \int_{|x|>\varepsilon} x^2 dF_{nk}(x). \end{aligned}$$

De acuerdo a la ecuación (4.6), el segundo sumando puede hacerse mas pequeño que $\eta > 0$ para cualquier $\varepsilon > 0$, así elegimos n suficientemente grande, y como $\varepsilon > 0$ es arbitrario, podemos elegirlo tan pequeño, sin importar como son $\eta > 0$ y T, tal que la siguiente desigualdad se satisface para todo t dentro del intervalo $|t| \leq T$:

$$|\rho_n| < 2\eta \ (n \geq n_0(\varepsilon, \eta, T)).$$

Esta desigualdad muestra que

$$\lim_{n\to\infty} \rho_n = 0 \tag{4.14}$$

uniformemente en cualquier intervalo finito de valores t. Reuniendo las relaciones (4.11), (4.12), (4.13) y (4.14),finalmente encontramos $\lim_{n\to\infty} \log \varphi_n(t) = -\frac{t^2}{2}$ uniformemente en cualquier intervalo finito de t. Queda finalmente demostrado el teorema. ■

Corollary 10 *Si las variables aleatorias mutuamente independientes $\xi_1, ..., \xi_n, ...$son idénticamente distribuidas y tienen varianza diferente de cero, entonces cuando n tiende a infinito*

$$P\{\frac{1}{B_n} \sum_{k=1}^{n} (\xi_k - a_k) < x\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{z^2}{2}} dz$$

uniformemente en x.

Demostración. Es suficiente probar que la condición de Lindeberg se satisface bajo la suposiciones dadas. Para este propósito se nota que en nuestro caso $B_n = b\sqrt{n}$ donde b^2 es la varianza de uno de los sumandos. Si $\mathbf{M}\xi_{\mathbf{k}} = a$ puede escribirse la siguiente ecuación obvia:

$$\begin{aligned}\sum_{k=1}^{n} \frac{1}{B_n^2} \int\limits_{|x-a|>\tau B_n} (x-a)^2 dF_k(x) &= \frac{1}{nb^2} \cdot n \int\limits_{|x-a|>\tau B_n} (x-a)^2 dF_1(x) \\ &= \frac{1}{b^2} \int\limits_{|x-a|>\tau B_n} (x-a)^2 dF_1(x).\end{aligned}$$

De la suposición que la varianza es finita y positiva concluimos que la integral del lado derecho de esta ecuación tiende a cero cuando $n \to \infty$. ■

Teorema 11 *(Teorema de Lyapunov) Si para una sucesión de variables aleatorias mutuamente independientes $\xi_1, ..., \xi_n, ...$ es posible escoger un $\delta > 0$ tal que cuando $n \to \infty$*

$$\frac{1}{B_n^{2+\delta}} \sum_{k=1}^{n} \mathbf{M} \left|\xi_k - a_k\right|^{2+\delta} \to 0, \tag{4.15}$$

entonces cuando $n \to \infty$

$$P\{\frac{1}{B_n} \sum_{k=1}^{n} (\xi_k - a_k) < x\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{z^2}{2}} dz$$

uniformemente en x.

Demostración. Nuevamente, será suficiente que la condición de Lyapunov [condición (4.15)] implica la condición de Lindeberg. Pero, esto es claro del siguiente cambio de desigualdades

$$\begin{aligned}\frac{1}{B_n^2} \sum_{k=1}^{n} \int\limits_{|x-a_k|>\tau B_n} (x-a_k)^2 dF_k(x) &\leq \frac{1}{B_n^2 (\tau B_n)^{\delta}} \sum_{k=1}^{n} \int\limits_{|x-a_k|>\tau B_n} \left|x-a_k\right|^{2+\delta} dF_k(x) \\ &\leq (\frac{1}{\tau^{\delta}}) \frac{\sum_{k=1}^{n} \int \left|x-a_k\right|^{2+\delta} dF_k(x)}{B_n^{2+\delta}} \to 0.\end{aligned}$$

■

En el capítulo (1) fue considerado el sistema lineal (2.1) bajo la acción de fuerzas aleatorias. Es por eso la necesidad de estudiar el siguiente sistema. Sin pérdida de generalidad, se puede considerar la ecuación de n-ésimo orden

$$x^{(n)} + a_1 x^{(n-1)} + ... + a_n x = f(t) \tag{4.16}$$

con coeficientes $a_k = a_k(t)$.

Sea $w(t,s)$ es la solución del problema

$$\frac{d^n}{dt^n}w(t,s) + a_1\frac{d^{n-1}}{dt^{n-1}}w(t,s) + ... + a_n w(t,s) = 0$$

$$\frac{d^k w}{dt^k}(t,s)|_{t=s} = 0,\ k = 0,1,...,n-2;\ \frac{d^{n-1}}{dt^{n-1}}w(t,s)|_{t=s} = 1. \tag{4.17}$$

Los cálculos directos [22] muestran que la solución del problema no-homogéneo (4.16) con valores iniciales cero, puede presentarse de la siguiente manera

$$x(t) = \int_0^t w(t,s)f(s)ds. \tag{4.18}$$

Ahora se considera el sistema (4.16) bajo la acción de fuerzas aleatorias

$$f = \sum_{k=1}^{\infty} \Delta\eta(t_k)\delta(t - t_k), \tag{4.19}$$

donde $\Delta\eta(t_k)$ para $k = 1, 2, ...$ son valores aleatorios independientes que tienen las propiedades siguientes

$$\begin{aligned} \mathbf{M}\Delta\eta(t_k) &= a_k\Delta t_k,\ \mathbf{D}\Delta\eta(t_k) = b_k\Delta t_k \\ \Delta t_k &: = t_k - t_{k-1}. \end{aligned}$$

Entonces, en el límite, el estado de muestra del sistema lineal (4.16) será la función aleatoria $\xi(t)$ que tiene esperanza

$$\mathbf{M}\xi_{\Delta t}(t) = \lim_{\Delta t\to 0}\sum_{0<t_k<t} w(t,t_k)a(t_k)\Delta t_k = \int_0^t w(t,s)a(s)ds \tag{4.20}$$

y varianza

$$\mathbf{D}\xi_{\Delta t}(t) = \lim_{\Delta t\to 0}\sum_{0<t_k<t} w^2(t,t_k)b(t_k)\Delta t_k = \int_0^t w^2(t,s)b(s)ds. \tag{4.21}$$

$$\mathbf{M}[\Delta\eta(t_k)]^2 = b_k\Delta t_k + [a(t_k)\Delta t_k]^2 \to 0 \text{ si } \Delta t \to 0. \tag{4.22}$$

En este caso

$$\sum_{0<t_k<t} \mathbf{D}\Delta\eta(t_k) = \sum_{0<t_k<t} b(t_k)\Delta t_k \to \int_0^t b(s)ds.$$

Suponiendo que

$$\sum_{0<t_k<t} \mathbf{M}[\Delta\eta(t_k)]^3 \to 0, \tag{4.23}$$

según el *teorema de límite central,* se obtiene que

$$P\{x' \le \xi_{\Delta t}(t) \le x''\} = \frac{1}{\sqrt{2\pi B}} \int_{x'}^{x''} e^{-\frac{(x-A)^2}{2B}} dx \tag{4.24}$$

donde

$$A = \int_0^t w(t,s)a(s)ds \text{ y } B = \int_0^t w^2(t,s)b(s)ds.$$

4.3. Definición Exacta del Integrando Estocástico

El proceso estocástico $\eta(t)$ se llama un proceso con *incremento sin correlación* si los valores aleatorios $\eta_1 = \eta(t_1) - \eta(s_1)$ y $\eta_2 = \eta(t_2) - \eta(s_2)$ tal que $s_1 \le t_1 \le s_2 \le t_2$, satisfacen la relación

$$\mathbf{M}(\eta_1 - \mathbf{M}\eta_1)(\eta_2 - \mathbf{M}\eta_2) = 0.$$

Ahora considere el proceso estocástico $\eta(t)$, para $c_1 \le t \le c_2$, con incrementos sin correlación y tal que

$$\mathbf{M}[\eta(t)-\eta(s)] = \int_s^t a(u)du;\ \mathbf{D}[\eta(t)-\eta(s)] = \int_s^t b(u)du.$$

Sea primero $a \equiv 0$. Entonces, para la función simple $\varphi(t)$ en $[c_1, c_2]$ que toma el valor de y_{k-1} en $[t_{k-1}, t_k]$ la integral estocástica es

$$\int_{c_1}^{c_2} \varphi(t)d\eta(t) = \sum_{k=1}^{n} y_k\Delta\eta(t_k). \tag{4.25}$$

Es evidente que, para la función simple φ, el integrando estocástico (4.25) no depende de la elección de los intervalos $[t_j, t_{j+1}]$, donde $\varphi(t)$ es constante.

Por eso, para cualesquiera dos funciones simples φ y ψ, se obtiene:

$$\int_{c_1}^{c_2} [\lambda_1\varphi(t) + \lambda_2\psi(t)]d\eta(t) = \lambda_1 \int_{c_1}^{c_2} \varphi(t)d\eta(t) + \lambda_2 \int_{c_1}^{c_2} \psi(t)d\eta(t).$$

Evidente que

$$\mathbf{M}\int_{c_1}^{c_2} \varphi(t)d\eta(t) = 0.$$

Ahora se toman en consideración los procesos aleatorios complejos. Se toma

$$\mathbf{D}\eta = \mathbf{M}|\eta - \mathbf{M}\eta|^2.$$

Las funciones aleatorias η_1 y η_2 están no-correlacionadas si

$$\mathbf{M}(\eta_1 - \mathbf{M}\eta_1)(\overline{\eta_2 - \mathbf{M}\eta_2)} = 0.$$

En adelante, se considerará el proceso aleatorio como complejo y con incrementos sin correlación tal que $\mathbf{M}[\eta(t)-\eta(s)] = \int_s^t a(u)du$ y $\mathbf{D}[\eta(t)-\eta(s)] = \int_s^t b(u)du$.

Sean φ y ψ funciones complejas definidas en $[c_1,c_2]$. Entonces

$$\mathbf{M}[\int_{c_1}^{c_2} \varphi(t)d\eta(t)\overline{\int_{c_1}^{c_2} \psi(t)d\eta(t)}] = \int_{c_1}^{c_2} \varphi(t)\overline{\psi(t)}b(t)dt. \tag{4.26}$$

En particular, se obtiene

$$\mathbf{M}\left|\int_{c_1}^{c_2} \varphi(t)d\eta(t)\right|^2 = \int_{c_1}^{c_2} |\varphi(t)|^2 b(t)dt.$$

Ahora, es introducida la norma

$$\left\|\int_{c_1}^{c_2} \varphi(t)d\eta(t)\right\|^2 = \mathbf{M}\left|\int_{c_1}^{c_2} \varphi(t)d\eta(t)\right|^2 = \int_{c_1}^{c_2} |\varphi(t)|^2 b(t)dt. \tag{4.27}$$

Entonces, la norma de $(\xi - \eta)$, donde

$$\xi = \int_{c_1}^{c_2} \varphi(t)d\eta(t),\ \eta = \int_{c_1}^{c_2} \psi(t)d\eta(t),$$

resulta

$$\|\xi - \eta\|^2 = \int_{c_1}^{c_2} |\varphi(t) - \psi(t)|^2 b(t)dt.$$

En este momento, se introduce el espacio de funciones en $L_2[b(t)dt]$ con la norma

$$\|\varphi\|_{L_2} = \left(\int_{c_1}^{c_2} |\varphi(t)|^2 \, b(t)dt \right)^{\frac{1}{2}}. \tag{4.28}$$

Suponga que para la función $\varphi \in L_2[b(t)dt]$ existe la sucesión de funciones simples $\varphi_n(t) \to \varphi(t)$ en *norma,* es decir $\|\varphi_n - \varphi\|_{L_2} \to 0$ $(n \to \infty)$ uniformemente. (Si $b(t)$ es acotada uniformemente y medible, esta propiedad se cumple para cualquier función $f \in L_2[b(t)dt]$.) La sucesión de integrales correpondientes satisface que, si

$$\xi_n = \int_{c_1}^{c_2} \varphi_n(t)d\eta(t),$$

se tiene

$$\|\xi_n - \xi_m\|_{L_2} = \int_{c_1}^{c_2} |\varphi_n(t) - \varphi_m(t)|^2 \, b(t)dt \to 0 \text{ si } n, m \to \infty.$$

En virtud de la completes del espacio de funciones aleatorias, con norma (4.28), existe una función aleatoria ξ, para la cual

$$\|\xi - \xi_m\|_{L_2} \to 0 (n \to \infty).$$

Este límite ξ se le llama *la integral estocástica* y se designa como

$$\xi = \int_{c_1}^{c_2} \varphi(t)d\eta(t).$$

La función aleatoria ξ está definida para cada evento elemental y con probabilidad 1 y no depende de la elección de la sucesión de funciones $\varphi_n(t) \to \varphi(t)$.

Es un buen momento ahora para determinar la integral estocástica cuando $\mathbf{M}\eta(t) \neq 0$.

Sea

$$\mathbf{M}\eta(t) = \int_{c_1}^{t} a(u)du$$

y sea $\varphi(t) \in L_2[b(t)dt]$ una función, tal que

$$\int_{c_1}^{c_2} |\varphi(t)a(t)| \, dt < \infty \text{ y } \int_{c_1}^{c_2} |\varphi(t)|^2 \, b(t)dt < \infty,$$

se definen

$$\int_{c_1}^{c_2} \varphi(t)d\eta(t) = \int_{c_1}^{c_2} \varphi(t)d\eta^o(t) + \int_{c_1}^{c_2} \varphi(t)a(t)dt,$$

donde $\eta^o(t) = \eta(t) - \mathbf{M}\eta(t)$.

4.4. Convergencia a Proceso Estacionario

Sean los coeficientes del sistema (4.16) no dependientes del tiempo y sea el sistema (4.16) estable. Es decir, para los valores característicos λ del polinomio característico $\lambda^n + a_1\lambda^{n-1} + ... + a_n = 0$ son todos de parte real negativa. ($\mathrm{Re}\,\lambda < 0$.)

Entonces, en [22] fue demostrado que, el proceso aleatorio

$$\xi^o(t) = \int_{t_0}^{t} w(t-s)d\eta(s)$$

para el caso $a(t) = a$ y $b(t) = b$, converge cuando $t \to \infty$ en probabilidad al proceso

$$\xi(t) = \int_{-\infty}^{\infty} w(t-s)d\eta(s),$$

es decir, $\|\xi - \xi^o\| \underset{t\to\infty}{\longrightarrow} 0$.

Capítulo 5

Controlabilidad

5.1. Estado Controlable y el Criterio de Kalman

En varios problemas, se llega a un sistema dinámico que necesitamos conocer a su estado y su salida, en otras palabras, queremos saber si estos sistemas son controlables. Estos sistemas son de la forma:

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + Du, \end{aligned} \tag{5.1}$$

donde el vector de *estado* x, es un n-vector; el vector *control* u, es un r-vector; y la *salida* y, es un m-vector; es decir, $x,\ y,\ u$ en $\mathbb{R}^n$, $\mathbb{R}^m$, $\mathbb{R}^r$, respectivamente o en $\mathbb{C}^n$, $\mathbb{C}^m$, $\mathbb{C}^r$ y las matrices A, B, C y D son matrices con entradas reales o complejas. Nos cuestionamos:

1.¿Puede el estado inicial $x(t_0)$ transferirse a cualquier estado deseado en un intervalo de tiempo finito?

2.¿Puede el estado inicial $x(t_0)$ determinarse por observaciones de la salida $y(t)$ sobre un intervalo de tiempo finito?

Para responder a estas preguntas, Kalman introdujó los conceptos de controlabilidad y observabilidad, así como condiciones que involucran las matrices A, B, C; las cuales sirven para dar respuesta o no. Estos conceptos, son de gran importancia en la teoría de control moderna; en varios casos ellos preven condiciones necesarias y suficientes para la existencia de una solución en algunos problemas de control.

Un sistema representado por (5.1), es llamado de estado completamente controlable, si para cualquier tiempo t_o es siempre posible contruir un vector control $u(t)$ sin restricciones el cual tranfiere al estado inicial $x(t_0)$ a un estado final arbitrario $x(t_f)$en un intervalo de tiempo finito, $t_0 \leq t \leq t_f$.

Si el sistema (5.1) es de tiempo-invariante, puede desarrollarse un simple criterio para el cual determinarse si el sistema es o no es completamente de estado controlable. Esta condición es establecida en el siguiente teorema.

Teorema 12 *(Criterio de Kalman) Un sistema de tiempo continuo descrito por la ecuación (5.1), donde A y B son matrices constantes. El estado es completamente de estado controlable si, y solo si, la matriz $n \times nr$ compuesta*

$$P = [B|AB|\cdots|A^{n-1}B] \tag{5.2}$$

es de rango n.

Demostración. La solución de (5.1), suponiendo $t_0 = 0$, es

$$x(t) = e^{At}[x(0) + \int_0^t e^{-A\tau}Bu(\tau)d\tau] \tag{5.3}$$

Se demuestra primero que, se puede tomar el estado final como el origen del *espacio estado,* sin pérdida de generalidad. Para el caso que el tiempo final $t_f = T$, entonces ya que e^{At} es no singular, se puede escribir (5.3), en la forma de

$$e^{-At}x(t) - x(0) = \int_0^t e^{-A\tau}Bu(\tau)d\tau$$

Ya que para vectores $x(0)$ y $x(T)$ arbitrarios, el vector $e^{-At}x(t) - x(0)$ es también arbitrario, entonces puede tomarse $x(T) = 0$. El problema es entonces reducido a determinar condiciones sobre A y B, tal que se pueda construir un vector de control $u(t)$, tal que

$$-x(0) = \int_0^T e^{-A\tau}Bu(\tau)d\tau \tag{5.4}$$

para cualquier arbitrario $x(0)$.

Considérese el caso donde el sistema es de estado completamente controlable. Como sabemos de álgebra lineal, podemos escribir

$$e^{-A\tau} = \sum_{i=0}^{p-1} \alpha_i(\tau)A^i, \tag{5.5}$$

donde $p \leq n$, es el grado del polinomio mínimo de A, y α_i son escalares. En términos de vectores $\mathbf{e}_i$, puede escribirse

$$u(t) = \sum_{i=1}^{r} u_i(t)\mathbf{e}_i, \tag{5.6}$$

donde $u_i(t)$ es la i-ésima componente de $u(t)$. Substituyendo (5.5) y (5.6) en (5.4), se tiene

$$\begin{aligned} -x(0) &= \int_0^T (\sum_{i=0}^{p-1} \alpha_i(\tau) A^i) B (\sum_{j=1}^{r} u_j(\tau) \mathbf{e}_j) d\tau, \\ &= \sum_{i=0}^{p-1} \sum_{j=1}^{r} [\int_0^T \alpha_i(\tau) u_j(\tau) d\tau] A^i B \mathbf{e}_j. \end{aligned}$$

Note que $B\mathbf{e}_j = B_j$, es la j-ésima columna de B, y al definir la cantidad

$$\gamma_{ij} = \int_0^T \alpha_i(\tau) u_j(\tau) d\tau,$$

puede escribirse

$$-x(0) = \sum_{i=0}^{p-1} \sum_{j=1}^{r} \gamma_{ij} A^i B_j. \tag{5.7}$$

Ya que el sistema es completamente de estado controlable, y cualquier estado inicial puede transferirse al origen.esto significa que el conjunto de pr vectores $A^i B_j$,$i = 0, 1, 2, ..., p-1; j = 1, 2, ..., r$,deberá generar el espacio n-dimensional. Consecuentemente este contiene un subconjunto de n vectores linealmente independientes y forma una base para el espacio de estado. El hecho que un subconjunto de pr vectores forman una base para el espacio estado n-dimensional, esto significa que la matriz $n \times pr$

$$Q = [B|AB|...|A^{p-1}B]$$

tiene rango no menos que n.Ya que el rango máximo de cualquier matriz con n renglones es n. Por lo tanto la matriz Q tiene rango n, y por eso la matriz P está dada por

$$P = [B|AB|...|A^{n-1}B]$$

Ahora suponga que la matriz P tiene rango n. Esto significa que P tiene n columnas linealmente independientes. Estas columnas también serán columnas de Q. Si $p < n$, entonces, se sigue que

$$A^p = \sum_{i=0}^{p-1} \beta_i A^i,$$

ya que el polinomio mínimo de A tiene grado p. Por lo tanto las columnas de $A^k B$, $k = p, p+1, ..., n-1$ son combinación lineal de $B, AB, ..., A^{p-1}B$ y consecuentemente Q deberá contener n columnas linealmente independientes.

Regresando a (5.7), los vectores columna A^iB_j, $i = 1, 2, ..., p-1$, $j = 1, 2, ..., r$ tiene un subconjunto el cual forma una base del espacio estado n-dimensional y por tanto cualquier n-vector puede representarse como (5.7). En otras palabras cualquier $x(0)$ puede escribirse en la forma (5.7). Ahora solo se necesitan encontrar las funciones $u_j(\tau)$ tal que la matriz γ_{ij} cumple (5.7). Por lo que, el sistema es de estado completamente controlable.

Esto completa la prueba. ■

Ejemplo 11 *Dado el sistema de ecuaciones diferenciales*

$$\dot{x} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} x + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u.$$

La matriz $P = [B|AB]$ *es*

$$P = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

cuyo rango es 2. Como $n = 2$ *el sistema es completamente de estado controlable.En otras palabras, por el teorema anterior, el estado del sistema puede tranferirse de cualquier estado inicial al origen en un intervalo de tiempo finito* T*. Pero, no nos da información a cerca del vector control que hace esta tranferencia.*

Ejemplo 12 *Consideremos ahora el siguiente sistema*

$$\dot{x} = \begin{bmatrix} -1 & -1 \\ 0 & -2 \end{bmatrix} x + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u \tag{5.8}$$

La matriz P *que se encuentra es*

$$P = \begin{bmatrix} 1 & -2 \\ 1 & -2 \end{bmatrix}.$$

Ya que el rango de P *es 1, se sigue de aquí que el sistema no es completamente de estado controlable. Este criterio nos da información explícita acerca de que estados son controlables. Una examinación de (5.7) revela que los vectores* B *y* AB *generan un espacio unidimensional generado por el vector* $B = [1\ 1]^T$*.*

Ejemplo 13 *Consideremos el sistema obtenido de la ecuación diferencial con coeficientes constantes*

$$\dddot{x} + a_2\ddot{x} + a_1\dot{x} + a_0x = u(t),$$

en donde las variables de estado son seleccionadas para ser variables fase $x_i(t) = D^{i-1}x(t)$, $i = 1, 2, 3$.

Las matrices A y B en la ecuación de estado son

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_0 & -a_1 & -a_2 \end{bmatrix}, \; B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

La matriz $P = [B|AB|A^2B]$ es dada por

$$P = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & -a_2 \\ 1 & -a_2 & -a_1 + a_2^2 \end{bmatrix}$$

Su determinante es diferente de cero, y así el sistema es de estado completamente controlable.

Podemos generalizar la ecuación diferencial en este último ejemplo y considerar la ecuación diferencial de orden n.

$$\sum_{k=0}^{n} a_k D^k x(t) = u(t), \; a_n = 1.$$

Nuevamente usando las variables fase

$$x_i(t) = D^{i-1}x(t), \; i = 1, 2, ..., n$$

Las matrices A y B son

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-2} & -a_{n-1} \end{bmatrix}, \; B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

Encontramos que la matriz P tiene la forma

$$\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 1 & -a_{n-1} \\ \cdots & \cdots & \cdots\cdots\cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 1 & \cdots & p_{n-2,n-1} & p_{n-2,n} \\ 0 & 1 & -a_{n-1} & \cdots & p_{n-1,n-1} & p_{n-1,n} \\ -1 & -a_{n-1} & -a_{n-2} + a_{n-1}^2 & \cdots & p_{n,n-1} & p_{n,n} \end{bmatrix}$$

donde los $p'_{ij}s$ son determinados de los $a'_k s$. Esto puede verse totalmente fácil, que el determinante de P es $(-1)^{n(n+3)/2}$ y consecuentemente el sistema está en estado completemente controlable.

5.2. El criterio de Gilbert

Gilbert define controlabilidad en alguna forma diferente a la definición previa, dada. Como se mostrará en la sección (5.3), las definiciones son equivalentes, cuando los valores propios de la matriz A son distintos. Al considerar esta definición de Gilbert, que los valores propios λ_i, $i = 1, 2, ..., n$, son distintos, entonces se puede encontrar una matriz no singular o de orden n-ésimo, la cual diagonaliza a A. Definiendo las coordenadas normales de las componentes del vector estado n-dimensional z (visto como un vector columna)

$$x = \rho z, \tag{5.9}$$

la nueva ecuación de estado, está dada por

$$\dot{z} = \Lambda z + \beta u \tag{5.10}$$

donde

$$\begin{aligned} \Lambda &= \rho^{-1} A \rho = diag(\lambda_1, \lambda_{2,}..., \lambda_n), \\ \beta &= \rho^{-1} B. \end{aligned}$$

Según Gilbert, el sistema representado por (5.10) es controlable si β *no tiene* renglones que son cero. Las coordenadas z_i correspondientes a renglones no cero de β, son llamadas controlables; las coordenadas correspondientes a renglones cero de β son llamadas incontrolabes. Cada coordenada z_i satisface la ecuación diferencial

$$\dot{z}_i = \lambda_i z_i + \sum_{j=1}^{r} \beta_{ij} u_j(t), \ i = 1, 2, ..., n. \tag{5.11}$$

Si el i-ésimo renglón de β es no cero, entonces, como puede verse de (5.11) la i-ésima coordenada puede ser influenciada por la entrada y consecuentemente controlada. Si el i-ésimo renglón de β es cero, entonces la i-ésima coordenada normal no es afectada por la entrada pero es determinada únicamente por su valor inicial.

Ejemplo 14 *Considerese como un ejemplo, el siguiente sistema dado en (5.8),*

$$\dot{x} = \begin{bmatrix} -1 & -1 \\ 0 & -2 \end{bmatrix} x + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u.$$

Como la matriz A es triangular sus valores propios son los de la diagonal,$\lambda = -1$, $\lambda = -2$. *Los correspondientes vectores propios son*

$$\rho_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ \rho_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

y por lo tanto la matriz de transformación la cual diagonaliza el sistema es

$$\rho = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

Ya que

$$\rho^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}^t = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix},$$

el sistema diagonalizado es

$$\dot{z} = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix} z + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u.$$

Como podemos ver el sistema diagonalizado, la coordenada normal z_1 no puede ser controlada. Ya que la coordenada z_1 satisface la ecuación diferencial:

$$\dot{z}_1 = -z_1,$$

tenemos

$$z_1(t) = z_1(0)e^{-t}.$$

Solo el estado que puede transferirse al origen son estados para el cual $z_1(0) = 0$. Ya que la coordenada z_2 es controlable, podemos encontrar una función control $u(t)$ para la cual cambia z_2 de cualquier valor inicial a cer en cualquier tiempo finito.

Es posible mostrar que, en esta función control se puede considerar como constante sobre el intervalo específico. La ecuación diferencial la cual determina el comportamiento de $z_2(t)$ es

$$\dot{z}_2(t) = 2z_2(t) + u(t).$$

Resolviendo esta ecuación, tenemos

$$\dot{z}_2(T) = e^{-2T}[z_2(0) + \int_0^T e^{2t}u(t)dt].$$

Ahora suponiendo que $z_2(T) = 0$ y $u(t) = K$, una contante, para $0 \leq t \leq T$.

Entonces lo siguiente se tiene

$$\begin{aligned} -z_2(0) &= K\int_0^T e^{2t}dt \\ &= \frac{1}{2}K(e^{2T} - 1). \end{aligned}$$

Por lo tanto se tiene

$$u(t) = \frac{2z_2(0)}{1 - e^{2T}},\ 0 \leq t \leq T.$$

Consideremos ahora qué le sucede al vector de estado original $\mathbf{x}$. *Ya que de (5.9) tenemos*

$$z = \rho^{-1}x$$

o

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

las coordenadas son dadas por las ecuaciones

$$\begin{aligned} z_1 &= x_1 - x_2, \\ z_2 &= x_2 \end{aligned} \tag{5.12}$$

Ya que $z_1(0) = 0$ *es la primera coordenada del estado controlable en el sistema de coordenadas normal, se sigue de (5.12), que las coordenadas del estado controlable en el sistema original de coordenadas satisfacen*

$$x_1(0) - x_2(0) = 0.$$

Este conjunto de estados iniciales, es un espacio de una dimensión generado por el vector $[11]^T$, *el cual es el mismo resultado usando el criterio de Kalman.*

Muy pronto, se podrá decir que un sistema es de estado controlable usando el criterio de Kalman o el de Gilbert. De cualquier modo, no puede darse información acerca de la función de control, pero que, es muy distinto, preguntarnos sí la función control existe, si no se consideran ningunas restricciones sobre las magnitudes de sus componentes. Ahora mostrare que sí, el sistema es controlable en el sentido de Gilbert, y sí $u(t)$ es escalar, entonces puede encontrarse una función de control (escalonada), la cual tranferirá el estado del sistema, de cualquier estado inicial al origen, en algún tiempo finito T. Para este caso, las ecuaciones diferenciales en (5.11) tienen la forma

$$\dot{z}_i = \lambda_i z_i + \beta_i u,\ i = 1, 2, ..., n,$$

donde β_i es la i-ésima componente de β. La solución entonces es

$$z_i(t) = e^{\lambda_i t}[z_i(0) + \beta_i \int_0^t e^{-\lambda_i \tau} u(\tau) d\tau],\ i = 1, 2, ..., n.$$

Si $z_i(T) = 0$, $i = 1, 2, ..., n$, entonces como $e^{\lambda_i T} \neq 0$ deberá determinarse $u(t)$ de la expresión:

$$-z_i(0) = \beta_i \int_0^t e^{-\lambda_i \tau} u(\tau) d\tau \tag{5.13}$$

Ahora mostrare que $u(\tau)$ puede ser función escalonada, definida por

$$u(\tau) = K_j, \ (j-1)\frac{T}{n} \leq \tau \leq j\frac{T}{n}, \ j = 1, 2, ..., n. \tag{5.14}$$

Sustituyendo esta expresión de $u(t)$ sobre (5.13), tenemos:

$$-z_i(0) = \beta_i \sum_{j=1}^{n} K_j \int_{(j-1)\frac{T}{n}}^{j\frac{T}{n}} e^{-\lambda_i \tau} d\tau$$

Haciendo el cambio

$$\tau = t + (j-1)\frac{T}{n},$$

la ecuación de arriba

$$\begin{aligned} -z_i(0) &= \beta_i \sum_{j=1}^{n} K_j \int_0^{\frac{T}{n}} e^{-\lambda_i t} e^{-\lambda_i (j-1)\frac{T}{n}} dt \\ &= \beta_i \sum_{j=1}^{n} K_j e^{-\lambda_i (j-1)\frac{T}{n}} \int_0^{\frac{T}{n}} e^{-\lambda_i t} dt \\ &= \beta_i \int_0^{\frac{T}{n}} e^{-\lambda_i t} dt \sum_{j=1}^{n} K_j e^{-\lambda_i (j-1)\frac{T}{n}}. \end{aligned} \tag{5.15}$$

Ya que $\beta_i \neq 0$, $i = 1, 2, ..., n$ puede escribirse el sistema de n ecuaciones dada en (5.15),para $i = 1, 2, ..., n$ en la forma

$$\mathbf{MK} = \boldsymbol{\gamma} \tag{5.16}$$

Donde $\mathbf{M} = [m_{ij}]$, $i, j = 1, 2, ..., n$, con $m_{ij} = e^{-\lambda_i (j-1)\frac{T}{n}}$, $\mathbf{K} = [K_1 \ K_2 ... K_n]^t$, y $\boldsymbol{\gamma} = [\gamma_1 \ \gamma_2 \ ... \ \gamma_n]^{\mathbf{t}}$ con

$$\gamma_i = -\beta_i^{-1} z_i(0) [\int_o^{\frac{T}{n}} e^{-\lambda_i t} dt]^{-1}.$$

Este sistema tiene siempre solución única para cualquier estado inicial $z(0)$ arbitrario cuando el determinante de la matriz de coeficientes M no es cero. Éste es siempre el

caso, ya que M es la transpuesta de la matriz de Vandermond

$$M = \begin{bmatrix} 1 & e^{-\lambda_1 \frac{T}{n}} & e^{-\lambda_1 \frac{2T}{n}} & \dots & e^{-\lambda_1 (n-1)\frac{T}{n}} \\ 1 & e^{-\lambda_2 \frac{T}{n}} & e^{-\lambda_2 \frac{2T}{n}} & \dots & e^{-\lambda_2 (n-1)\frac{T}{n}} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & e^{-\lambda_n \frac{T}{n}} & e^{-\lambda_n \frac{2T}{n}} & \dots & e^{-\lambda_n (n-1)\frac{T}{n}} \end{bmatrix}$$

$$\det(M) = \prod_{\substack{i,j=1 \\ i>j}}^{n} [e^{-\lambda_i \frac{T}{n}} - e^{-\lambda_j \frac{T}{n}}]$$

el cual no es cero cuando los valores propios son distintos.Esta es una manera de resolver (5.16) para las $K_i's$, y sustituyendo estos valores sobre (5.14) se obtiene la función de control, la cual, transferirá el estado del sistema $\mathbf{z}(0)$ a la posición de equilibrio (al origen).

A modo de ilustración, considérese el sistema diagonalizado

$$\dot{x} = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix} x + \begin{bmatrix} 1 \\ 2 \end{bmatrix} u \tag{5.17}$$

y elija una función de control constante por trozos:

$$\begin{aligned} u(t) &= K_1, 0 \leq t \leq 1, \\ &= K_2, 1 \leq t \leq 2, \end{aligned}$$

la cual tranfiere el estado inicial $x(0) = [a\ b]^T$ al origen en 2 segundos. Para este caso, las ecuaciones en (5.15) son

$$\begin{aligned} -a &= (1) \int_0^1 e^t dt [K_1 + K_2 e] \\ -b &= (2) \int_0^1 e^{2t} dt [K_1 + K_2 e^2] \end{aligned}$$

o en forma simplificada

$$\begin{aligned} K_1 + K_2 e &= -a(e-1)^{-1} \\ K_1 + K_2 e^2 &= -b(e^2-1)^{-1}. \end{aligned}$$

Resolviendo estas ecuaciones para K_1 y K_2, encontramos la función control requerida y es:

$$\begin{aligned} u(t) &= [-ae + b(e+1)^{-1}]/(e-1)^2,\ 0 \leq t \leq 1, \\ &= [a - b(e+1)^{-1}]/e(e-1)^2,\ 1 \leq t \leq 2. \end{aligned}$$

5.3. Equivalencia de criterios de Kalman y Gilbert

Cuando los valores propios de A son diferentes, el criterio de Kalman y Gilbert son equivalentes. Este resultado se establecerá en la forma de un teorema.

Teorema 13 *Sean $B_i, i = 1, 2, ..., r$, las columnas de B. Un sistema como (5.1) es controlable en el sentido de Gilbert si y solo si la matriz A tiene diferentes valores propios y vectores $\mathbf{e}_{ki} = A^k B_i, i = 1, 2, ..., r,\ k = 0, 1, ..., n-1$, que generan el espacio n-dimensional.*

Demostración. Suponga que el sistema es controlable en el sentido de Gilbert. Ya que $\Lambda = \rho^{-1} A \rho$, se sigue que

$$\begin{aligned} \mathbf{e}_{ki} &= A^k B_i = (\rho \Lambda \rho^{-1})^k B_i \\ &= \rho \Lambda^k \rho^{-1} B_i \\ &= \rho \Lambda^k \beta_i \end{aligned}$$

donde β_i es la i-ésima columna de $\beta = \rho^{-1} B$. Ya que β no tienen renglones cero, es posible formar un vector

$$\beta^+ = \sum_{i=1}^{r} K_i \beta_i$$

el cual no tiene componentes cero.

Primero se probará que, el conjunto de vectores $\mathbf{e}_k^+ = \rho \Lambda^k \beta^+$, $k = 0, 1, ..., n-1$, son linealmente independientes y por lo tanto generan el espacio n-dimensional. Esto es posible, si podemos verificar que $\det[\mathbf{e}_0^+ ... \mathbf{e}_{n-1}^+] \neq 0$. Ya que puede escribirse

$$\begin{aligned} [\mathbf{e}_0^+ ... \mathbf{e}_{n-1}^+] &= [\rho\beta^+ | \rho\Lambda\beta^+ | \cdots | \rho\Lambda^{n-1}\beta^+] \\ &= \rho \begin{bmatrix} \beta_1^+ & \lambda_1\beta_1^+ & \cdots & \lambda_1^{n-1}\beta_1^+ \\ \cdots & \cdots & \cdots & \cdots \\ \beta_n^+ & \lambda_n\beta_n^+ & \cdots & \lambda_n^{n-1}\beta_n^+ \end{bmatrix} \\ &= \rho[diag(\beta_1^+, \beta_2^+, ..., \beta_n^+)]\mathbf{V} \end{aligned}$$

donde $\mathbf{V}$ es la transpuesta de la matriz de Vandermonde

$$\mathbf{V} = \begin{bmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{n-1} \\ \cdots & \cdots & \cdots & \cdots \\ 1 & \lambda_n & \cdots & \lambda_n^{n-1} \end{bmatrix}.$$

Usando la propiedad de que el determinante de un producto de matrices $n \times n$ es igual al producto de los determinantes, se obtiene

$$\det[\mathbf{e}_0^+ ... \mathbf{e}_{n-1}^+] = (\det \boldsymbol{\rho})(\prod_{i=1}^{n} \beta_i^+) \det \mathbf{V} \neq 0$$

ya que P y V son no singulares y $\beta_i^+ \neq 0$ para $i = 1, 2, ..., n$. En seguida, se prueba que el espacio generado por los vectores $\mathbf{e}_k^+$, $k = 0, 1, ..., n-1$, es un subespacio del espacio que generan los vectores $\mathbf{e}_{ki}$, $k = 0, 1, ..., n-1$, $i = 1, 2, ..., r$. Estos espacios son generados, respectivamente, de los elementos de la forma

$$\sum_{k=0}^{n-1} \alpha_k \mathbf{e}_k^+ = \sum_{k=0}^{n-1} \sum_{i=1}^{r} \alpha_k K_i \rho \Lambda^k \beta_i \tag{5.18}$$

y

$$\sum_{k=0}^{n-1} \sum_{i=1}^{r} \alpha_{ki} \mathbf{e}_{ki} = \sum_{k=0}^{n-1} \sum_{i=1}^{r} \alpha_{ki} \rho \Lambda^k \beta_i, \tag{5.19}$$

donde los K_i, $i = 1, 2, ..., r$ son constantes. Claramente, cada vector que pertenece al espacio determinado por (5.18) para arbitrarios α_k está en el espacio determinado por (5.19) para arbitrarios α_{ki}. Ya que los vectores $\mathbf{e}_k^+$, $k = 0, 1, ..., n-1$, son linealmente independientes y generan un subespacio del espacio generado por los vectores $\mathbf{e}_{ki}$, $k = 0, 1, ..., n-1$, $i = 1, 2, ..., r$ se sigue que, el espacio generado por $\mathbf{e}_{ki}$ es el espacio de coordenadas n-dimensional. (Consecuentemente la matriz

$$P = [B|AB| \cdots |A^{n-1}B]$$

deberá tener rango n).

Suponga que, los $\mathbf{e}_{ki}$ generan el espacio n-dimensional. Entonces, para cualquier $R \neq 0$,

$$R = (\rho^T)^{-1} \begin{bmatrix} 1 & 0 & ... & 0 \end{bmatrix}^T,$$

dice que el producto interno (R, e_{ki}) no puede ser cero para toda k y toda i. Para ver porque esto es cierto, suponga que $a_1, a_2, ..., a_n$ es un subconjunto linealmente independiente de e_{ki}. Tal conjunto existe, pues e_{ki} genera el espacio n-dimensional. Ahora, sí

$$(R, a_i) = 0 \; i = 1, 2, .., n,$$

entonces

$$R^T \begin{bmatrix} a_1 & a_2 & ... & a_n \end{bmatrix} = 0$$

o en diferente forma

$$\sum_{i=1}^{n} R_i a_{ji} = 0,\ j = 1, 2, ..., n,$$

donde R_i es la i-ésima componente de R y a_{ji} es la i-ésima componente de a_j. Como los a_i son linealmente independientes,

$$\det \begin{bmatrix} a_1 & a_2 & ... & a_n \end{bmatrix} \neq 0$$

y por eso $R_i = 0$, $i = 1, 2, ..., n$. Como $R_i \neq 0$, se sigue que $(R, e_{ki}) \neq 0$ para algún e_{ki}, como establecimos primero. Ahora

$$\begin{aligned} (R, e_{ki}) &= R^T e_{ki} = \begin{bmatrix} 1 & 0 & ... & 0 \end{bmatrix} \rho^{-1} e_{ki} \\ &= \begin{bmatrix} 1 & 0 & ... & 0 \end{bmatrix} \rho^{-1} \rho \Lambda^k \beta_i = \begin{bmatrix} 1 & 0 & ... & 0 \end{bmatrix} \Lambda^k \beta_i \\ &= \lambda_1^k \beta_{i1} \end{aligned}$$

donde β_{i1} es la primera componente de β_i. Si todos los $\beta_{i1} = 0$, se sigue que $(R, e_{ki}) = 0$ y por eso, tenemos demostrado que $R = 0$, lo cual no es posible por como tomamos R. De cualquier modo, como $R \neq 0$, se sigue que, al menos uno de los β_{i1} no es cero y el sistema es controlable en el sentido de Gilbert. Tenemos esencialmente demostrado, que si la matriz

$$P = [B|AB| \cdots |A^{n-1}B]$$

tiene rango n (una cosecuencia directa de la definición de controlabilidad de Kalman), y si los valores característicos de la matriz A son distintos, entonces el sistema es de estado completamente controlable respecto a la definición de Gilbert. ■

Se sigue que, para un sistema, donde la matriz A, tiene valores propios distintos, las definiciones de controlable, dada por Kalman y por Gilbert, son equivalentes. No obstante, el procedimiento de Gilbert, da información directa de cuales de las variables de estado, en el sistema diagonalizado, son incontrolables. Esencialmente, la misma información puede obtenerse de la definición de Kalman y determinar, cuales son las columnas de P que son linealmente independientes y consecuentemente, generen un subespacio de variables controlables.

Capítulo 6

Observabilidad

Un concepto relacionado con controlabilidad es el de observabilidad. Para cada sistema, las variables de estado no son generalmente accesibles, para medirlas directamente. Las cantidades que son posibles medirlas, son la entrada y la salida del sistema. Una pregunta natural es que, ¿cuando pueden determinarse las variables de estado por la observación de las variables de entrada y salida del sistema sobre un intervalo de tiempo?

Esto nos llevará al concepto de observabilidad de un sistema.

Formalmente, se define la observabilidad como sigue. Un sistema con entrada u, salida y, y el estado x, es llamado completamente observable si y solo si, el estado inicial $x(t_0)$ puede obtenerse por mediciones de $y(t)$ sobre el intervalo de tiempo finito $t_0 \leq t \leq t_0 + T$. Notamos, que si el sistema es observable, puede obtenerse $x(t_0)$ para varios valores de t_0 y así, obtener una medición de $x(t)$.

Restringiremos nuestra atención en esta sección a sistemas lineales, tiempo invariantes y continuos que se describen por

$$\dot{x} = Ax + Bu \tag{6.1}$$

$$y = Cx + Du \tag{6.2}$$

Como el sistema es lineal y de tiempo invariante, podemos tomar $t_0 = 0$, en cuyo caso, tenemos

$$y(t) = C\phi(t)x(0) + C\int_0^t \phi(t-\tau)Bu(\tau)d\tau + Du. \tag{6.3}$$

Podemos reescribir el resultado en la forma

$$z(t) = C\phi(t)x(0) \tag{6.4}$$

donde

$$z(t) = y(t) - C\int_0^t \phi(t-\tau)Bu(\tau)d\tau - Du. \tag{6.5}$$

Ya suponemos que y y u son conocidos, entonces z es conocido, y la pregunta de observabilidad del sistema, es entonces, sí podemos encontrar $x(0)$ de (6.4) para un z conocido sobre el intervalo $0 \leq t \leq T$. Esta pregunta, es claramente independiente de u, para el caso $u = 0$ y $u \neq 0$, simplemente corresponde a diferentes valores de z, como vemos de (6.5). Por lo que, sin pérdida de generalidad, podemos suponer, para propósitos de investigación, que $u = 0$, en cuyo caso el sistema de ecuaciones, está dado por

$$\dot{x}(t) = Ax(t) \tag{6.6}$$

$$y(t) = Cx(t) = C\phi(t)x(0). \tag{6.7}$$

Por lo tanto, la observabilidad de un sistema, depende solo de A y C, y el caso de que C sea no singular, el sistema es obviamente observable, ya que $x(t)$ se obtiene directamente de (6.7).

Consideremos primero el caso donde la salida y es un escalar $y(t)$.

Suponiendo que el grado del polinomio mínimo de A es p, podemos escribir

$$\phi(t) = \sum_{k=0}^{p-1} \alpha_k(t)A^k$$

y consecuentemente, se sigue de (6.7), que

$$y(t) = \sum_{k=0}^{p-1} \alpha_k(t)CA^k x(0). \tag{6.8}$$

Como, por pruebas resolvemos para $x(0)$, un vector de n componentes, necesitamos ordinariamente n ecuaciones involucrando las cantidades $x_i(0)$, $i = 1, 2, ..., n$. Una técnica bastante usada, para obtener ecuaciones adicionales de (6.8), la cual involucra a p coeficientes $\alpha_k(t)$, $k = 0, 1, ..., p-1$, es formando productos internos de $y(t)$ y $\alpha_k(t)$. El producto interno que utilizamos, está definido por

$$(\alpha_i(t), \alpha_j(t)) = \int_0^T \bar{\alpha}_i(t)\alpha_j(t)dt \tag{6.9}$$

donde $\bar{\alpha}_i(t)$ es complejo conjugado de $\alpha_i(t)$ y T es la longitud del tiempo que $y(t)$ es observado. Por lo tanto, tomando el producto interno de $\alpha_i(t)$ y $y(t)$, obtenemos

$$\begin{aligned} (\alpha_j(t), y(t)) &= \sum_{k=0}^{p-1} (\alpha_j(t), \alpha_k(t))CA^k x(0), \\ j &= 0, 1, ..., p-1. \end{aligned} \tag{6.10}$$

Como $y(t)$ es escalar, la cantidad $CA^k x(0)$ es también un escalar. Si ponemos,

$$CA^k x(0) = \beta_k, \ \ k = 0, 1, ..., p-1. \tag{6.11}$$

$$(\alpha_j(t), y(t)) = \mu_j, \ \ j = 0, 1, ..., p-1 \tag{6.12}$$

$$(\alpha_j(t), \alpha_k(t)) = \gamma_{jk}, j, k = 0, 1, ..., p-1 \tag{6.13}$$

entonces podemos escribir (6.10) en la forma matricial

$$\mu = \gamma\beta \tag{6.14}$$

donde β, μ, y γ son matrices con elementos dados por (6.11), (6.12) y (6.13), respectivamente. La matriz γ está definida como una matriz de Gram y puede ser demostrado que es no singular, ya que los $\alpha_k(t)$ son linealmente independientes. Por lo tanto la solución de (6.14), está dada por

$$\beta = \gamma^{-1}\mu. \tag{6.15}$$

Substituyendo estos valores de β_k sobre (6.11), de un sistema de p ecuaciones en las n desconocidas $x_1(0), x_2(0), ..., x_n(0)$. En la forma matricial

$$\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{p-1} \end{bmatrix} x(0) = \beta, \tag{6.16}$$

donde los coeficientes sobre la matriz en la izquierda es una matriz $p \times n$. Esta ecuación puede resolverse de forma única para $x(0)$ si y solo si $p = n$ y el rango de los coeficientes de la matriz es n. Esto equivale pedir que, la matriz $n \times n$

$$P = [C^T | A^T C^T | ... | (A^T)^{n-1} C^T] \tag{6.17}$$

tiene rango n. Más precisamente, este resultado se establece como un teorema.

Teorema 14 *El sistema representado por*

$$\dot{x} = Ax, \ \ y = Cx,$$

donde x es n-vector y y es un escalar, es observable si y solo si P dada por (6.17) es no singular.

6.1. Un Ejemplo

Anteriormente precedimos al caso general en que y es un m-vector, daremos un ejemplo en donde el vector estado será determinado para el caso en que y sea un escalar. Consideremos el sistema

$$\begin{aligned} \dot{x}(t) &= \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} x(t) + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u(t) \\ y(t) &= \begin{bmatrix} 1 & 1 \end{bmatrix} x(t) \end{aligned} \tag{6.18}$$

con $u(t) = 1$. La matriz P dada por (6.17) para el sistema, es

$$P = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

la cual tiene rango 2. Consecuentemente el sistema es completamente observable. En la práctica no requerimos tener una expresión analítica para $y(t)$, pero para ilustrar el principio complicado, supondremos que

$$y(t) = \frac{1}{2}e^{2t} + 2e^{t} - 3/2.$$

La matriz de transición puede escribirse, en dos maneras

$$\begin{aligned} \phi(t) &= diag\{e^{t}, e^{2t}\}, \\ \phi(t) &= (2e^{t} - e^{2t})I + (e^{2t} - e^{t})A \end{aligned}$$

que puede probarse con fácilidad. Como el sistema es forzado, podemos calcular $z(t)$ dada por

$$\begin{aligned} z(t) &= y(t) - C\int_0^t \phi(t-\tau)Bu(\tau)d\tau \\ &= \frac{1}{2}e^{2t} + 2e^{t} - 3/2 - \begin{bmatrix} 1 & 1 \end{bmatrix} \int_0^t \begin{bmatrix} e^{t-\tau} & 0 \\ 0 & e^{2(t-\tau)} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} d\tau \\ &= e^{t}. \end{aligned}$$

De la expresión para $\phi(t)$ obtenemos

$$\begin{aligned} \alpha_0(t) &= 2e^{t} - e^{2t} \\ \alpha_1(t) &= e^{2t} - e^{t} \end{aligned}$$

y por tanto

$$\begin{aligned}
\gamma_{00} &= (\alpha_0(t), \alpha_0(t)) = \int_0^T (2e^t - e^{2t})^2 dt \\
\gamma_{01} &= (\alpha_0(t), \alpha_1(t)) = \int_0^T (2e^t - e^{2t})(e^{2t} - e^t) dt \\
\gamma_{11} &= (\alpha_1(t), \alpha_1(t)) = \int_0^T (e^{2t} - e^t)^2 dt
\end{aligned}$$

$$\begin{aligned}
\mu_0 &= (\alpha_0(t), z) = \int_0^T (2e^t - e^{2t}) e^t dt \\
\mu_1 &= (\alpha_1(t), z) = \int_0^T (e^{2t} - e^t) e^t dt
\end{aligned}$$

usando (6.15), la matriz β está dada por

$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \frac{1}{\gamma_{00}\gamma_{11} - \gamma_{12}^2} \begin{bmatrix} \gamma_{11} & -\gamma_{10} \\ -\gamma_{10} & \gamma_{00} \end{bmatrix} \begin{bmatrix} \mu_0 \\ \mu_1 \end{bmatrix}.$$

Como A y C son matrices reales, (6.16), puede escribirse como

$$P^T x(0) = \beta$$

y para nuestro caso, se sigue que

$$x(0) = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}^{-1} \beta = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \beta.$$

Combinando las relaciones anteriores, encontramos que

$$\begin{aligned}
x_1(0) &= 2\beta_0 - \beta_1 \\
&= \frac{2(\gamma_{11}\mu_0 - \gamma_{10}\mu_1) - (-\gamma_{10}\mu_0 + \gamma_{00}\mu_1)}{\gamma_{00}\gamma_{11} - \gamma_{10}^2}
\end{aligned}$$

y

$$\begin{aligned}
x_2(0) &= -\beta_0 + \beta_1 \\
&= \frac{-(\gamma_{11}\mu_0 - \gamma_{10}\mu_1) + (-\gamma_{10}\mu_0 + \gamma_{00}\mu_1)}{\gamma_{00}\gamma_{11} - \gamma_{10}^2}
\end{aligned}$$

Llevando a cabo los cálculos necesarios, obtenemos

$$\begin{aligned}
\gamma_{00} &= \frac{1}{4}e^{4T} - \frac{4}{3}e^{3T} - 2e^{2T} - \frac{11}{12} \\
\gamma_{01} &= -\frac{1}{4}e^{4T} + e^{3T} - e^{2T} + \frac{1}{4} \\
\gamma_{11} &= \frac{1}{4}e^{4T} - \frac{2}{3}e^{3T} + \frac{1}{2}e^{2T} - \frac{1}{12}
\end{aligned}$$

$$\begin{aligned}
\mu_0 &= -\frac{1}{3}e^{3T} + e^{2T} - \frac{2}{3} \\
\mu_1 &= \frac{1}{3}e^{3T} - \frac{1}{2}e^{2T} + \frac{1}{6}.
\end{aligned}$$

Las cantidades necesitadas para determinar $x(0)$ son encontradas:

$$\begin{aligned}
\gamma_{11}\mu_0 - \gamma_{10}\mu_1 &= -\gamma_{10}\mu_0 + \gamma_{00}\mu_1 = \gamma_{00}\gamma_{11} - \gamma_{10}^2 \\
&= \frac{1}{72}e^{6T} - \frac{1}{8}e^{4T} + \frac{2}{9}e^{3T} - \frac{1}{8}e^{2T} + \frac{1}{72}.
\end{aligned}$$

Usando estos valores, tenemos

$$x(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Es fácil verificar que este sistema con su estado inicial encontrado, su salida es la antes dada.

Para este ejemplo el vector estado también puede determinarse totalmente fácil de (6.4). En este caso tenemos

$$\begin{aligned}
e^t &= \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} e^t & 0 \\ 0 & e^{2t} \end{bmatrix} \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} \\
&= e^t x_1(0) + e^{2t} x_2(0).
\end{aligned}$$

Comparando los coeficientes de ambos lados vemos que $x_1(0) = 1$ y $x_2(0) = 0$.

6.2. El Caso General

Es buen momento para probar un teorema, dando una condición de observabilidad cuando la salida $y(t)$ es un m-vector. Este teorema es una extensión del teorema previo que toma la sálida escalar y, como uno debe esperar, la prueba es muy similar.

Teorema 15 *El sistema representado por*

$$\begin{aligned} \dot{x} &= Ax + Bu, \\ y &= Cx + Du, \end{aligned} \tag{6.19}$$

donde x es un n-vector y y es un m-vector, es completamente observable si y solo si, la matriz

$$P = [C^T | A^T C^T | \cdots | (A^T)^{n-1} C^T]$$

es de rango n.

Demostración. Podemos tomar $u = 0$ sin pérdida de generalidad, como hemos visto antes, para dicho caso y, usando la expansión de e^{At}, está dada por

$$y(t) = \sum_{k=0}^{p-1} \alpha_k(t) C A^k x(0).$$

Si denotamos al i-ésimo renglón de C por C_i, la i-ésima coordenada $y_i(t)$ de $y(t)$ es

$$y(t) = \sum_{k=0}^{p-1} \alpha_k(t) C_i A^k x(0), \ i = 1, 2, ..., m.$$

Como antes, podemos tomar productos internos de $\alpha_j(t)$ con $y(t)$ y así obtener pm ecuaciones en las pm desconocidas $C_i A^k x(0)$. Definiendo las cantidades

$$C_i A^k x(0) = \beta_{ki},$$

$$(\alpha_j(t), y_i(t)) = \mu_{ji}$$

$$(\alpha_j(t), \alpha_k(t)) = \gamma_{jk}$$

las ecuaciones mencionadas arriba, pueden escribirse

$$\begin{aligned} \sum_{k=0}^{p-1} \gamma_{jk} \beta_{ki} &= \mu_{ji}, j = 0, 1, ..., p-1; \\ i &= 1, 2, ..., m. \end{aligned}$$

Si denotamos por β_i y μ_i las i-ésimas columnas de las matrices $\beta = [\beta_{ki}]$ y $\mu = [\mu_{ji}]$, respectivamente, la última ecuación puede escribirse en la forma matricial

$$\gamma \beta_i = \mu_i, i = 1, 2, ..., m.$$

Como un caso previo de la matriz de Gram, γ es no singular y podemos escribir

$$\beta_i = \gamma^{-1}\mu_i.$$

Teniendo determinadas las cantidades β_{ki}, estaremos dispuestos ahora considerar nuevamente las ecuaciones

$$C_i A^k x(0) = \beta_{ki};\ i = 1, 2, ..., m;\ j = 0, 1, ..., p-1,$$

o la ecuación matricial equivalente

$$\begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{p-1} \end{bmatrix} x(0) = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ \vdots \\ b_{p-1} \end{bmatrix}$$

donde b_j es la tranpuesta del (j+1)-ésimo renglón de β. Este es un sistema de mp ecuaciones con n desconocidas y el vector $x(0)$ puede determinarse únicamente si y solo si, el rango de la matriz de coeficientes $mp \times n$ es n. En este caso $p = n$, la matriz de coeficientes es P^T donde P está dada en (6.17). Cuando $p < n$ se sigue que las últimas $m(n-p)$ columnas de

$$P = [C^T | A^T C^T | \cdots | (A^T)^{p-1} C^T | \cdots | (A^T)^{n-1} C^T]$$

son combinaciones lineales de las columnas $(A^T)^k C^T$, $k = 0, 1, ..., p-1$. Por lo tanto, el rango de la matriz P es n, la cual es la condición de controlabilidad. Esto completa la prueba del teorema. ■

Ejemplo 15 *Considere el sistema*

$$\begin{aligned} \dot{x} &= \begin{bmatrix} 0 & 1 \\ -3 & -4 \end{bmatrix} x + \begin{bmatrix} 1 \\ -3 \end{bmatrix} u \\ y &= \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} x. \end{aligned}$$

La matriz P está dada como

$$P = \begin{bmatrix} 2 & 1 & -6 & -3 \\ 2 & 1 & -6 & -3 \end{bmatrix}.$$

El rango de P es 1 y consecuentemente el sistema, no es completamente observable.

6.3. Definición de Gilbert

El concepto de observabilidad, considerado en la sección previa se debe a Kalman, también Gilbert tiene una definición de observabilidad (la cual es diferente a la que introduce Kalman). De acuerdo a Gilbert, el sistema diagonalizado con distintos valores propios, dado por

$$\begin{aligned} \dot{x} &= \Lambda x + \beta u \\ y &= \gamma x + Du \end{aligned} \tag{6.20}$$

donde $\Lambda = diag(\lambda_1, \lambda_2, ..., \lambda_n)$, es observable si γ tiene todas sus columnas son diferentes de cero. Las x_i correspondientes a una columna cero de γ se llaman inobservables, y las no cero, observables. Claramente una coordenada inobservable, no aparece en la salida y por lo tanto su valor no puede determinarse, por observaciones de la salida.

Ejemplo 16 *Como un simple ejemplo, consideremos el sistema*

$$\begin{aligned} \dot{x} &= \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix} x + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u \\ y &= \begin{bmatrix} 0 & 1 \end{bmatrix} x. \end{aligned}$$

De acuerdo a la definición de Gilbert, la coordenada x_1 es inobservable y x_2 es observable ya que la primera columna de la matriz γ es cero y la segunda columna no es cero. La matriz P para este sistema, usando la definición de Kalman, es

$$P = \begin{bmatrix} 0 & 0 \\ 1 & -2 \end{bmatrix}.$$

Esta, tiene rango 1, por eso el sistema es inobservable.

Ejemplo 17 *Como otro ejemplo, el sistema (6.18), es observable. Ya que $A = \Lambda$ y $\gamma = [1\ 1]$ no tiene columnas cero.*

Hemos visto que, para sistemas son distintos valores propios, las dos definiciones de controlabilidad debidas a Kalman y a Gilbert, son equivalentes. Las definicones de Gilbert sobre controlabilidad y sobre observabilidad, son totalmente similares, con los renglones de β jugando el mismo papel que las columnas de γ en observabilidad, y el criterio en controlabilidad de Kalman envuelve los rangos de las matrices dadas en (5.2) y (6.17), son muy similares.

Por tanto, podríamos esperar que las dos definiciones de observabilidad son equivalentes, como fue el caso para controlabilidad. Es un hecho el caso, y podemos establecer un teorema que trata de observabilidad, el cual es análogo al teorema (13).

Teorema 16 *Sean c_i, $i = 1, 2, ..., m$ las columnas de C^T. El sistema (6.19) es observable en el sentido de Gilbert, si y solo si, A tiene valores propios distintos y los vectores*

$$e_{ki} = (A^T)^k c_i;\ i = 1, 2, ..., m;\ k = 0, 1, ..., n-1,$$

generan el espacio n-dimensional.

Demostración. La demostración de este teorema, es muy similar a la dada para el caso de controlabilidad en el teorema (13). Consecuentemente, solo indicaremos la dirección tomada y dejamos los detalles, por ser muy parecidos, en este caso

$$\begin{aligned} e_{ki} &= (A^T)^k c_i = [(\rho\Lambda\rho^{-1})^T]^k(\gamma_i^T\rho^{-1})^T \\ &= (\rho^{-1})^T(\Lambda^T)^k\rho^T(\rho^{-1})^T\gamma_i \\ &= (\rho^{-1})^T(\Lambda^T)^k\gamma_i \end{aligned}$$

donde $\gamma^T = [\gamma_1\ \gamma_2 ... \gamma_m]$. Como γ no tiene columnas cero, entonces γ^T no tiene renglones cero y podemos formar el vector $\gamma^+ = k_1\gamma_1 + k_2\gamma_2 + ... + k_m\gamma_m$, ninguna de las componentes, es cero. Los pasos tratados, son los mismos, como el teorema (13). Como los e_{ki} son las columnas de P en (6.17) y generan el espacio n-dimesional, la matriz P deberá tener rango n. Por ello, en el caos de que A tiene distintos valores propios, los criterios de Gilbert y Kalman son equivalentes. ■

6.4. Principio de Dualidad

Como previamente notamos, la estructura de las matrices las cuales se necesitan para determinar si un sistema es observable y controlable, son muy similares. Esta similaridad, sugiere que hay una analogía entre controlabilidad y observabilidad. Está analogía es claramente exhíbida por el principio de dualidad debida a Kalman, el cual se establece como un teorema.

Teorema 17 *Considere el sistema S_1 definido por*

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx, \end{aligned}$$

donde x es un n-vector, u es un r-vector,y y es un m-vector, y el sistema dual S_2 definido por

$$\begin{aligned} \dot{\xi} &= A^T\xi + C^T v \\ \eta &= B^T\xi, \end{aligned}$$

donde ξ es un n-vector,v es un m-vector, y η es un r-vector. El principio de dualidad establece el sistema S_1 es completamente de estado controlado (observable) si y solo si el sistema S_2 es completamente observable (de estado controlable).

Para probar este teorema, escribimos abajo las condiciones necesarias y suficientes de los teoremas (12) y (15) que debemos satisfacer. El sistema S_1 es completamente de estado controlable, si y solo si, la matriz $n \times nr$

$$[B|AB|...|A^{n-1}B]$$

tenga rango n. Esto es precisamente, la condición necesaria y suficiente, que el sistema S_2 sea completamente observable. La condición necesaria y suficiente, para que el sistema S_1 sea completamente observable, es que la matriz $n \times mn$

$$[C^T|A^TC^T|...|(A^T)^{n-1}C^T]$$

tenga rango n. Esto es, también la condición necesaria y suficiente, para que el sistema S_2 sea de estado completamente controlable y el teorema es probado.

En seguida, se menciona un sistema de ecuaciones que ocurren muy frecuentemente, en algunos problemas de sistemas dinámicos, a la hora de discretizar dicho sistema.

6.5. La Ecuación TA-BT=C

En esta sección se encuentra un resultado de la teoría de matrices,que se usará adelante en un ejemplo.

La ecuación matricial

$$TA - BT = C \tag{6.21}$$

juega un papel muy importante en la teoría de estabilidad y en observabilidad, así como otras líneas de investigación, como análisis de sistemas. En esta ecuación A y B son matrices cuadradas de orden m y n, respectivamente, y consecuentemente T es una matriz $n \times m$. Demostrare que sí A y B no tienen valores propios comunes, entonces (6.21), tiene una solución única T.

1. Probare que $TA - BT = 0$ tiene solo la solución trivial.

2. Entonces probaremos que $TA - BT = C$, tiene solución única.

Consideremos la parte 1. Por álgebra lineal, sabemos que las matrices A y B son similares, a matrices en su forma canónica de Jordan. Supongamos que la forma canónica de Jordan, de la matriz A es dada por

$$J_A = P^{-1}AP = diag(J_{m1}, J_{m2}, ..., J_{mp}) \tag{6.22}$$

donde el orden de J_{m_k} es m_k y $m_1 + m_2 + ... + m_p = m$, y B es similar a

$$J_B = Q^{-1}BQ = diag(J_{n1}, J_{n2}, ..., J_{nq}) \tag{6.23}$$

donde el orden de J_{n_k} es n_k y $n_1 + n_2 + ... + n_q = n$. El valor propio de A asociado a J_{m_i} es λ_i y el valor propio de B asociado a J_{n_j} es μ_j. Podemos resolver para A y B en (6.22) y (6.23), y entonces sustituir estos valores en la ecuación

$$TA - BT = 0. \tag{6.24}$$

La ecuación por tanto, que se obtiene es

$$TPJ_AP^{-1} - QJ_BQ^{-1}T = 0.$$

Multiplicando esta ecuación por Q^{-1} sobre el lado izquierdo y por P sobre el lado derecho, tenemos

$$Q^{-1}TPJ_A - J_BQ^{-1}TP = 0.$$

Sea ahora

$$X = Q^{-1}TP, \tag{6.25}$$

entonces, puede escribirse

$$XJ_A - J_BX = 0. \tag{6.26}$$

Notemos que (6.26), es meramente un caso especial de (6.24), donde en (6.26), las matrices conocidas, están en su forma canónica de Jordan. Si podemos resolver (6.26) para X, entonces podemos obtener T, de la ecuación

$$T = QXP^{-1}. \tag{6.27}$$

Como J_A y J_B son esencialemente matrices particionadas, será conveniente particionar la matriz X, tal que los productos XJ_A y J_BX están definidos. Consecuentemente, será una matriz particionada $q \times p$; esto es, la submatriz X_{ij} será matriz $n_i \times n_j$. Cuando llevamos la multiplicación como en (6.26), usando las matrices particionadas, obtenemos las ecuaciones

$$X_{ij}J_{m_j} = J_{n_i}X_{ij};\ i = 1, 2, ..., q,\ j = 1, 2, ..., p \tag{6.28}$$

La matriz J_{m_j}, $m_j \times m_j$, tiene la forma

$$\begin{bmatrix} \lambda_j & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_j & 1 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & \lambda_j & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & \lambda_j \end{bmatrix}$$

mientras la matriz $n_i \times n_i$, J_{n_i}, tiene la forma

$$\begin{bmatrix} \mu_i & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \mu_i & 1 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & \mu_i & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & \mu_i \end{bmatrix}.$$

Introduciendo la matriz $s \times s$

$$\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix},$$

podemos entonces escribir

$$J_{m_j} = \lambda_j I + H_{m_j},\ j = 1, 2, ..., p,$$

y

$$J_{n_i} = \mu_i I + H_{n_i},\ i = 1, 2, ..., q.$$

Substituyendo esta expesión sobre (6.28), tenemos

$$(\lambda_j - \mu_i)X_{ij} = H_{n_i}X_{ij} - X_{ij}H_{m_j.} \tag{6.29}$$

Ahora multipliquemos esta expresión, por $(\lambda_j - \mu_i)$,obtenemos

$$(\lambda_j - \mu_i)^2 X_{ij} = H_{n_i}(\lambda_j - \mu_i)X_{ij} - (\lambda_j - \mu_i)X_{ij}H_{m_j.} \tag{6.30}$$

Ahora, substituyendo (6.29) en (6.30), obtenemos

$$(\lambda_j - \mu_i)^2 X_{ij} = (H_{n_i})^2 X_{ij} - 2H_{n_i}X_{ij}H_{m_j} + X_{ij}(H_{m_j})^2.$$

Continuando este proceso, puede mostrarse, por inducción que

$$(\lambda_j - \mu_i)^r X_{ij} = \sum_{k=0}^{r} (-1)^{r-k} \binom{r}{r-k} (H_{n_i})^k X_{ij} (H_{m_j})^{r-k}.$$

AL tomar $r \geq n_i + m_j - 1$, entonces, como $r = k + (r - k)$, se sigue que, o bien $k \geq n_i$ o $r - k \geq m_j$. Como H_s es una matriz nilpotente de índice s, o bien $(H_{n_i})^k = 0$ o $(H_{m_j})^{r-k} = 0$. Consecuentemente, tenemos

$$(\lambda_j - \mu_i)^r X_{ij} = 0$$

y como $\lambda_j \neq \mu_i$ para $i = 1, 2, ..., q;\ j = 1, 2, ..., p$,debemos tener

$$X_{ij} = 0, i = 1, 2, ..., q;\ j = 1, 2, ..., p.$$

Por lo tanto, solo la solución trivial, es la única que tiene (6.26). Sí $X = 0$, se sigue que, de (6.27), $T = 0$ es solo la solución de (6.24).

Ahora que tenemos establecida la parte 1, estaremos listos para tomar la parte 2. Probaremos que sí, A y B, no tienen valores propios comunes, el sistema (6.21), tiene una única solución T. Tomando el elemento c_{ij} de cada lado de (6.21), tenemos

$$\sum_{k=1}^{m} t_{ik}a_{kj} - \sum_{k=1}^{n} b_{ik}t_{kj} = c_{ij}, \tag{6.31}$$

$$i = 1, 2, ..., n;\ j = 1, 2, ..., m.$$

Este es el sistema lineal de mn ecuaciones desconocidas $t_{rs},\ r = 1, 2, ..., n;\ s = 1, 2, ..., m$. Este sistema no es homogéneo, pero tiene solución única, si y solo si, el determinante de los coeficientes no es cero. El correspondiente sistema homogéneo es (6.24), el cual tiene la misma matriz como (6.21). Ya que la solución del sistema homogéneo es la solución trivial, se sigue que el determinante de la matriz de coeficientes no es cero. por lo tanto, la matriz de coeficientes no es singular y el sistema (6.31) o (6.21) tiene solución única T para cada matriz C.

Ejemplo 18 *Considere*

$$A = [2],\ B = \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix},\ C = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

El valor propio de A es 2, y valores propios de B, son 1 y -1. Por tanto, hay una solución única de (6.21), para este caso. Resolviendo el sistema

$$\begin{bmatrix} t_{11} \\ t_{21} \end{bmatrix} [2] - \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} t_{11} \\ t_{21} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

para t_{11} *y* t_{21}. *Ecuación correspondiente de elementos que se obtinen*

$$\begin{aligned} t_{11} - t_{21} &= 1 \\ 3t_{21} &= 0. \end{aligned}$$

Por tanto $t_{21} = 0$ *y* $t_{11} = 1$ *y la solución única es*

$$T = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Capítulo 7

Conclusiones

Los resultados de la tesis muestran que podemos aplicar el método de identificación propuesto a sistemas dinámicos lineales estables. Esto, primero se probó para los sistemas cuyas matrices de tamaño $n \times n$ tenían n valores propios diferentes, luego abordamos el caso donde las matrices eran más generales, de valores propios no necesariamente todos diferentes entre sí. Fue demostrado que la probabilidad de tener identificación no única es igual a cero, esto es, que la identificación de los coeficientes del sistema dinámico propuesto (2.1) es única, con probabilidad de uno. Estas ideas surgieron gracias a la similitud que existe con los problemas de la Teoría del Control, en nuestro caso, el problema en su forma discreta requiere de condiciones similares impuestas a las de la teoría del Control.

Bibliografía

[1] Rao J. S. *Rotor dynamics* 1996 (New Delhi: New Age International Publisher).

[2] Vance J. M. *Rotordynamics of Turbomachinery* 1998 (Jon Wiley and Sons, Inc.).

[3] Earl A. Coddington, Norman Levinson 1984 *Theory of ordinary differential equations* (Malabar, Fla. :Klieger Publishing Company).

[4] A. Antonio-García, J. Gómez-Manzilla, V. V. Kucherenko 2002 *Identification of rotordynamics parameters for the Jefcott rotor-bearing model* (Honolulu, Hawaii, USA: The 9th International Simposium on Transport Phenomena and Dynamics of Rotating Machinery (ISROMAC-9)).

[5] Alberto Tarantola 1987 *Inverse Problem Theory* (Methods for Data Fitting and Model Parameter Estimation) (Amsterdam-Oxford-New York_Tokyo: Elsevier).

[6] Frank M. Callier and Charles A. Desoer 1994 *Linear System Theory* (New York, Berlin, Heidelbery London: Springer-Verlag (Springer Texts in Electrical Engineering))

[7] R. E. Kalman, P. L. Falb and M. A. Arbib 1969 *Topics Mathematical System Theory* (McGraw- Hill Book Company).

[8] J. Schoukens and R. Pintelon 1981 *Indentification of Linear System* (Pergamon Press).

[9] Pieter Eykhoff 1981 *Trends and Progess in System Identification* (Pergamon Press, IFAC series Vol. I).

[10] Chen C. T. 1987 *Technique for Identification of Linear Invariant Multivariable Systems* (Control and Dynamics Sistems, V. 26, Part 2, p. 1-34).

[11] Y. Ito, T. Hagiwara, H. Maeda and M. Araki 2001 *Bisection algorithm for computing the frequency response gain of sampled-data system* (IEEE Trans. Automat. Control, Vol. 46, N3, p. 369-381).

[12] Datte K. B. and Gangopadhlyay 1995 *Identification of MIMO systems from input output data* (Control Theory Adn. Tech., Vol. 10, N4, Part. 5, p. 2109-2123).

[13] Win ADRE *Analysis and Diagnosis for Rotating for Equipment, user's manual* (Bently 2000, Nevada).

[14] Victor Isakov 1997 I*nverse Problems for partial differential equations* (New York. Springer).

[15] Banks H. T. and K. Kunish 1989 *Estimation Technique for Distributed Parameter Systems* (Boston, Basel, Berlin: Birkhauser).

[16] L. Z. Lu and V. W. Sun 1999 *On necessary conditions for reconstruction of a special structured Jacobi matrix from eigenvalues* (Linear Algebra and Aplications, V. 15, N4, p. 977-988).

[17] De Boor C. and G. H. Golub 1978 *The numerical stable reconstruction of a Jacobi matrix from spectral data* (Linear Algebra and Aplications, V. 21, N2, p. 45-60).

[18] Vladimir A. Marchenko 1988 *Methods of Investigation of the Inverse Spectral Problem for Finite Dimensional Operators* (in russian) (Kharkov University, p. 46).

[19] John T. Moore 1968 E*lements of linear algebra and matrix theory* (New York: McGraw-Hill)

[20] Shlomo Sternberg 1964 *Lectures on differential geometry* (Englewood Cliffs, NJ: Prentice-Hall).

[21] A. Krivkov and V. V. Kucherenko *Identification of linear dynamics system* (International Conference ISAAC Italy, 2005).

[22] Yu. V. Prohorov and Yu. A. Rozanov *Probability Theory* (Spinger Verlag, 1969)

[23] H. Kwakernaak and R. Sivan *Linear Optimal Control Systems* (Wiley-Interscience 1972).

[24] Andriy Kryvko and Valeri V. Kucherenko Coeficient identification for the linear dynamic system (en revisión en SIAM Optimization and Control)

[25] Alberto Antonio García, Julio GómezManzilla, Valeri V. Kucherenko 2004 *Coeficient Identification of linear dinamical system with harmonic input aplication to rotordinamic vibration* (Cientifica vol. 8 N. 2 pp 79-86).