UNIVERSIDAD NACIONAL DE INGENIERÍA

FACULTAD DE CIENCIAS

TESIS

ESTABILIDAD DE LOS SISTEMAS LINEALES DE SALTO MARKOVIANO EN TIEMPO DISCRETO

PARA OBTENER EL TÍTULO PROFESIONAL DE:
LICENCIADO EN MATEMÁTICA

ELABORADO POR:

CAMARENA PÉREZ, VICTOR DANIEL

ASESOR:
MG. ECHEGARAY CASTILLO, WILLIAM CARLOS

LIMA-PERÚ

2019

A mi hermano Josho, con quien emprendí, a la par,
el camino del estudio aunque con recorridos distintos.

Agradecimientos:

A mis padres, Lucía Pérez y Victor Camarena, y a mis hermanos Janet Camarena, Elsa Camarena, Raquel Camarena y José Camarena por el apoyo y cariño brindado a lo largo de mis años universitarios y más.

A mi asesor William Echegaray por su consejo y auxilio durante el desarrollo del presente trabajo, y al que me presentó el problema abordado por primera vez, Jorge Mayta.

A mis amigos, por la compañía en el tránsito universitario, y a mis compañeros y conocidos.

Resumen

La meta propuesta es caracterizar la estabilidad de un sistema lineal de salto Markoviano en tiempo discreto. Específicamente, se estudia el comportamiento de las soluciones del sistema dinámico discreto en un espacio Euclideano complejo

$$x(k+1) = A_{\theta(k)}x(k)$$

cuando $\{\theta(k)\}$ es una cadena de Markov. Previamente, se estudian la teoría de estabilidad según Lyapunov en espacios Euclideanos complejos, sin considerar aleatoriedad, y algunas extensiones importantes. Luego, para un sistema lineal de salto Markoviano, se exhibe un marco estocástico natural que deriva de él y que permite el análisis conductual de las soluciones. Finalmente, una vez definida operativamente la noción de estabilidad media cuadrática se establecen diversas caracterizaciones, algunas derivadas de la teoría de Lyapunov antes estudiada, así como se revisan diversos ejemplos.

Palabras Claves: Función de Lyapunov, Cadena de Markov, Sistema Lineal de Salto Markoviano, Estabilidad Media Cuadrática.

Abstract

The proposed goal is to characterize the stability of a Discrete-Time Markov Jump Linear System. Specifically, the behavior of discrete dynamical system solutions in a complex Euclidean space is studied

$$x(k+1) = A_{\theta(k)}x(k)$$

when $\{\theta(k)\}$ is a Markov chain. Previously, stability theory according to Lyapunov is studied in complex Euclidean spaces, without considering randomness, and also some important extensions. Then, for a Markov Jump Linear System, a natural stochastic framework is exhibited that derives from it and that allows the behavioral analysis of its solutions. Finally, once the notion of mean square stability is operationally defined, several characterizations are established, some derived from the Lyapunov theory before studied, as well as several examples are reviewed.

Keywords: Lyapunov function, Markov chain, Markov Jump Linear System, Mean Square Stability.

Índice general

Índice de figuras

Introducción

Motivación

Un modelo matemático con una dinámica lineal de salto Markoviano es útil para estudiar procesos que están sujetos a cambios estocásticos en su modo de operación, la idea es determinar cuándo estas variaciones en el tiempo van afectando significativamente el comportamiento del sistema y cómo podemos cuantificar las probabilidades de los distintos escenarios posibles. Ejemplos de procesos de este tipo ocurren en sistemas económicos, sistemas en ingeniería, sistemas robóticos, sistemas de control de aeronaves, grandes estructuras flexibles para estaciones espaciales, control de receptores en centrales térmicas solares, optimización del consumo de combustible en vehículos, etc., donde los cambios abruptos pueden deberse a perturbaciones naturales abruptas, fallos o reparaciones de componentes, cambios de las interconexiones de los subsistemas, cambios violentos en el funcionamiento de una instalación a gran escala, el tránsito irregular de los vehículos, etc. Un ejemplo clásico es el modelo discreto que forma parte del modelamiento matemático del problema de control de receptores en centrales térmicas solares (Ver figura 1) descrito en [25], dicho sea de paso, este es uno de los primeros Sistemas Lineales de Salto Markoviano que aparece en la literatura y sirvió de motivación para impulsar el estudio de dichos sistemas dinámico estocásticos.

Figura 1: Proyecto Solar Crescent Dunes, Las Vegas, EE.UU.
La planta de energía solar Crescent Dunes de 110 MW es la primera instalación a escala de servicio público en el mundo que tiene capacidades avanzadas de almacenamiento de energía en torre con sales fundidas (Fuente: Solarreserve).

Antecedentes referenciales

Los primeros trabajos de Ji et al. [14], Feng et al. [11], Yang et al. [10], Fang [9], Costa and Fragoso [4], entre otros abordan el tema de la estabilidad de los sistemas lineales de salto Markoviano en tiempo discreto cuando la cadena de Markov toma valores en un espacio de estados finito.
Los libros de Czornik [7] y Costa et al. [6] muestran el estado del arte para esta familia de sistemas lineales de salto Markoviano en tiempo discreto cuando la cadena de Markov toma valores en un espacio de estados finito.

La tesis de licenciatura de Mayta [20] representa el trabajo previo más cercano pues trata el caso de Sistemas Lineales de Salto Markoviano en Tiempo Discreto sobre un espacio de estados $\mathbb{R}^d$ mientras aquí se desarrolla el caso $\mathbb{C}^d$. De hecho, nuestro trabajo se basa mayormente en el libro de Costa et al. [6].

Planteamiento de la realidad problemática

Un Sistema Lineal de Salto Markoviano en Tiempo Discreto, es un sistema dinámico descrito por la recurrencia

$$x(k+1) = A_{\theta(k)}x(k), \ \ k \geq 0, \tag{i}$$

donde $x(k)$ es un vector en $\mathbb{C}^d$, $\{\theta(k)\}_{k\geq 0}$ es una cadena de Markov con espacio de estados finito. Para el sistema (i) y cada instante $k \geq 0$: el vector $x(k)$ se denomina variable de estado del sistema, la matriz cuadrada $A_{\theta(k)}$ indica el modo de operación del sistema. Note que hay una cantidad finita de modos de operación indexada por el espacio de estados de la cadena; y que $x(k)$ es de naturaleza estocástica pues los modos de operación se intercalan aleatoriamente, de acuerdo a la dinámica de transición de la cadena de Markov. Un caso particular de este sistema dinámico estocástico en tiempo discreto es el modelo lineal clásico

$$x(k+1) = Ax(k), \ \ k \geq 0,$$

que aparece en la teoría de control al preguntarse por la estabilidad del funcionamiento del sistema. Entonces un Sistema Lineal de Salto Markoviano en Tiempo Discreto es un modelo más complejo y su estudio sigue el fin de generalizar los resultados conocidos para el modelo lineal clásico.

Objetivos

Nuestro objetivo general es establecer para un Sistema Lineal de Salto Markoviano en Tiempo Discreto un resultado análogo al que se tiene para un Sistema Lineal en Tiempo Discreto.

Los objetivos específicos son los siguientes:

1. Desarrollar un test de estabilidad de un Sistema Lineal de Salto Markoviano en Tiempo Discreto mediante el cálculo del radio espectral de una matriz asociada al sistema.

2. Determinar la ecuación de Lyapunov para un Sistema Lineal de Salto Markoviano en Tiempo Discreto.

Hipótesis

Nuestra hipótesis general es que para un Sistema Lineal de Salto Markoviano en Tiempo Discreto se puede conseguir un resultado análogo al que se tiene para un Sistema Lineal en Tiempo Discreto.

Las hipótesis específicas son las siguientes:

1. Es factible desarrollar un test de estabilidad de un Sistema Lineal de Salto Markoviano en Tiempo Discreto mediante el cálculo del radio espectral de una matriz asociada al sistema.

2. Es posible determinar la ecuación de Lyapunov para un Sistema Lineal de Salto Markoviano en Tiempo Discreto.

Estructura del marco teórico

El marco teórico está en los tres capítulos siguientes y su estructura consta de tres partes, cada una abordada en un capítulo, que se alínean de acuerdo al objetivo principal. El primer capítulo trata de la teoría de estabilidad de sistemas dinámicos en tiempo discreto desarrollada por Lyapunov con enfásis en el potente método de testeo mediante funciones de Lyapunov con el fin de conseguir una adaptación posterior al modelo de estudio, además, se establecen las notaciones matriciales y se describe un espacio vectorial de secuencia de matrices que constituye un dominio natural para ciertos operadores lineales asociados a un Sistema Lineal de Salto Markoviano en Tiempo Discreto. El segundo capítulo discute el factor aleatorio del modelo de estudio, especificamente se muestra un marco estocástico en el cual se puede describir las soluciones del Sistema Lineal de Salto Markoviano en Tiempo Discreto, también, se exhiben las técnicas probabilísticas que se usarán con frecuencia en lo restante del trabajo. El tercer capítulo desarrolla el objetivo central, primero se aprovecha la propiedad de Markov que rige los saltos entre los modos de operación del sistema de estudio para asociarle ciertos operadores que caracterizarán lo que se conoce como estabilidad media cuadrática, segundo se enuncia el resultado principal que relaciona diversas caracterizaciones para la noción de estabilidad media cuadrática asi como anotar varias observaciones y ejemplos, tercero se prueba el resultado principal donde el método de la función de Lyapunov juega un papel clave.

Capítulo 1

Preliminares

Este capítulo repasa la teoría de estabilidad de un sistema dinámico lineal en tiempo discreto siguiendo el enfoque de Lyapunov. Se inicia estableciendo las notaciones del álgebra matricial. En seguida, se muestra que, sobre un espacio Euclideano, la existencia de una función de Lyapunov es una condición suficiente para la estabilidad de un sistema dinámico en tiempo discreto, en particular para un sistema lineal se establecen múltiples equivalencias. En la última sección se estudia un espacio vectorial, isomorfo a un espacio Euclideano, que es el dominio natural de ciertos operadores asociados a un Sistema Lineal de Salto Markoviano en Tiempo Discreto.

1.1. Notaciones

Los fundamentos requeridos en esta sección se encuentran en un libro básico de análisis funcional o de teoría de operadores, por ejemplo [24] o [23].

Asuma que $\mathbb{X}$ e $\mathbb{Y}$ son espacios de Banach complejos y sea $\mathbb{B}(\mathbb{X}, \mathbb{Y})$ el espacio de Banach de los operadores lineales continuos de $\mathbb{X}$ a $\mathbb{Y}$, con la norma uniforme inducida representada por $\|\cdot\|$. Por simplicidad sea $\mathbb{B}(\mathbb{X}) = \mathbb{B}(\mathbb{X}, \mathbb{X})$. El radio espectral de un operador $T \in \mathbb{B}(\mathbb{X})$ es denotado por $\rho(T)$. Si $\mathbb{X}$ es un espacio de Hilbert el producto interno se denota por $\langle \cdot, \cdot \rangle$; y para $T \in \mathbb{B}(\mathbb{X})$, T^* denota la adjunta del operador T. La notación $T \geq 0$ (respectivamente $T > 0$) significa que el operador T es semidefinido positivo (definido positivo). En particular, $\mathbb{R}^d$ y $\mathbb{C}^d$ denotan respectivamente el espacio euclideano real y complejo d-dimensional y $\mathbb{B}(\mathbb{C}^d, \mathbb{C}^{d'})$ (respectivamente $\mathbb{B}(\mathbb{R}^d, \mathbb{R}^{d'})$) el espacio de las matrices complejas (reales) de orden $d \times d'$, con $\mathbb{B}(\mathbb{C}^d) = \mathbb{B}(\mathbb{C}^d, \mathbb{C}^d)$ (respectivamente $\mathbb{B}(\mathbb{R}^d) = \mathbb{B}(\mathbb{R}^d, \mathbb{R}^d)$). Salvo indicación contraria, $\|\cdot\|$ denotará la norma estándar en $\mathbb{C}^d$ (respectivamente $\mathbb{R}^d$), y para $M \in \mathbb{B}(\mathbb{C}^d, \mathbb{C}^{d'})$ (respectivamente $M \in \mathbb{B}(\mathbb{R}^d, \mathbb{R}^{d'})$), $\|M\|$ denota la norma inducida en $\mathbb{B}(\mathbb{C}^d, \mathbb{C}^{d'})$ (respectivamente $\mathbb{B}(\mathbb{R}^d, \mathbb{R}^{d'})$). Los superscripts $^-$, ^{t} y * indican la conjugada, transpuesta y transpuesta conjugada de una matriz, respectivamente. Luego, se denota $\mathbb{B}(\mathbb{C}^d)^+ = \{M \in \mathbb{B}(\mathbb{C}^d) \colon M^* = M \geq 0\}$. El operador identidad es denotado por I, y la matriz identidad de orden $d \times d$ por I_d (o simplemente I). Finalmente, se denotan por $\lambda_i(A)$, $i = 1, \ldots, d$ a los autovalores de la matriz $A \in \mathbb{B}(\mathbb{C}^d)$.

Observación 1. *El operador* $\operatorname{tr}(\cdot) \colon \mathbb{B}(\mathbb{C}^d) \to \mathbb{C}$ *es un funcional con las siguientes propiedades:*

1. $\operatorname{tr}(KL) = \operatorname{tr}(LK)$.

2. *Para todos* $M, N \in \mathbb{B}(\mathbb{C}^d)$ *con* $M \geq 0$, $N > 0$,

$$\left(\min_{1 \leq i \leq d} \lambda_i(N)\right) \operatorname{tr}(M) \leq \operatorname{tr}(MN) \leq \left(\max_{1 \leq i \leq d} \lambda_i(N)\right) \operatorname{tr}(M).$$

En efecto, (1) *es fácil de verificar. Para mostrar* (2)*, sean las descomposiciones espectrales (ver Corolario 4.3.5, [22]) de las matrices* M *y* N*:*

$$M = U \Sigma U^* \quad y \quad N = V \Lambda V^*$$

donde U y V son matrices unitarias en $\mathbb{C}^d$, $\Sigma = diag[\lambda_i(M)]$ y $\Lambda = diag[\lambda_i(N)]$. Se sabe que

$$\operatorname{tr}(M) = \operatorname{tr}(\Sigma) = \sum_{i=1}^{d} \lambda_i(M),$$

y de (1) *se sigue*

$$\operatorname{tr}(MN) = \operatorname{tr}(U\Sigma U^* V\Lambda V^*) = \operatorname{tr}(\Sigma U^* V\Lambda V^* U) = \operatorname{tr}\left(\Sigma U^* V\Lambda (U^* V)^*\right).$$

Entonces, si se denota por W a la matriz unitaria U^*V

$$\begin{aligned}\operatorname{tr}(MN) = \operatorname{tr}(\Sigma W\Lambda W^*) &= \sum_{i=1}^{d} \lambda_i(M) \sum_{j=1}^{d} w_{ij}\lambda_j(N) w_{ji}^* \\ &= \sum_{i=1}^{d} \lambda_i(M) \sum_{j=1}^{d} w_{ij}\lambda_j(N) \overline{w_{ij}} \\ &= \sum_{i=1}^{d} \lambda_i(M) \sum_{j=1}^{d} \lambda_j(N) |w_{ij}|^2.\end{aligned}$$

Luego, como las columnas de W son unitarias se tienen las desigualdades

$$\left(\min_{1\leq j\leq d} \lambda_j(N)\right) \leq \sum_{j=1}^{d} \lambda_j(N)|w_{ij}|^2 \leq \left(\max_{1\leq j\leq d} \lambda_j(N)\right),\ i = 1, \ldots, d,$$

y por lo tanto, multiplicando y sumando se tiene

$$\left(\min_{1\leq j\leq d} \lambda_j(N)\right) \sum_{i=1}^{d} \lambda_i(M) \leq \sum_{i=1}^{d} \lambda_i(M) \sum_{j=1}^{d} \lambda_j(N)|w_{ij}|^2 \leq \left(\max_{1\leq j\leq d} \lambda_j(N)\right) \sum_{i=1}^{d} \lambda_i(M),$$

que coincide con las desigualdades a probar.

Se define el operador φ como sigue: para $V = [v_1 \cdots v_d] \in \mathbb{B}(\mathbb{C}^d)$ con $v_j \in \mathbb{C}^d$

$$\varphi(V) = \begin{bmatrix} v_1 \\ \vdots \\ v_d \end{bmatrix} \in \mathbb{C}^{d^2}.$$

El producto Kronecker de dos matrices $K, L \in \mathbb{B}(\mathbb{C}^d)$ es la matriz por bloques

$$K \otimes L = [k_{ij}L] \in \mathbb{B}(\mathbb{C}^{d^2}).$$

Luego, se cumplen las siguientes propiedades (ver [2]):

$$\begin{aligned}
&(K \otimes L)^* = K^* \otimes L^* \\
&\varphi(HKL) = (L^t \otimes H)\varphi(K),\ H \in \mathbb{B}(\mathbb{C}^d).
\end{aligned}$$

Otra propiedad importante del operador $\mathrm{tr}(\cdot)$ asociada al operador φ es

$$\mathrm{tr}(K^*L) = \varphi(K)^*\varphi(L).$$

De esta propiedad se sigue

$$|\mathrm{tr}(K^*L)| = |\varphi(K)^*\varphi(L)| \leq \|\varphi(K)\|\|\varphi(L)\| \leq \sqrt{\mathrm{tr}(K^*K)}\sqrt{\mathrm{tr}(L^*L)}.$$

Más aún, la aplicación dada por: para K y L en $\mathbb{B}(\mathbb{C}^d)$

$$\langle K, L\rangle = \mathrm{tr}(K^*L)$$

define un producto interno sobre el espacio $\mathbb{B}(\mathbb{C}^d)$. Así, $(\mathbb{B}(\mathbb{C}^d), \langle\cdot,\cdot\rangle)$ es un espacio de Hilbert.

Por otro lado, desde que $\|K\| = \sqrt{\max_{1\leq i\leq d}\lambda_i(K^*K)}$ se tiene

$$\|K\|^2 \leq \sum_{i=1}^{d}\lambda_i(K^*K) = \mathrm{tr}(K^*K) \leq d\|K\|^2,$$

o lo que es lo mismo,

$$\|K\| \leq \sqrt{\langle K, K\rangle} \leq \sqrt{d}\|K\|.$$

Es conocido que sobre un espacio finito dimensional todas las normas son equivalentes, este resultado se usará varias veces en este trabajo. Por ejemplo, al mudar de una norma a otra se preserva la convergencia. Recuerde que dos normas $\|\cdot\|_1$, $\|\cdot\|_2$ sobre un espacio de Banach $\mathbb{X}$ son equivalentes si existen $c_1 > 0$ y $c_2 > 0$ tal que para todo $x \in \mathbb{X}$,

$$c_1\|x\|_1 \leq \|x\|_2 \leq c_2\|x\|_1.$$

De la desigualdad $\rho(K) \leq \|K\|$ se deriva la siguiente propiedad que relaciona el operador $\mathrm{tr}(\cdot)$ con la norma matricial $\|\cdot\|$:

$$\mathrm{tr}(K) = \sum_{i=1}^{d} \lambda_i(K) \leq \sum_{i=1}^{d} |\lambda_i(K)| \leq d\rho(K) \leq d\|A\|.$$

Una última relación con normas a usar viene de la desigualdad de Cauchy-Schwartz: para $x \in \mathbb{C}^d$,

$$\|xx^*\| = \max_{\|z\|=1} \|xx^*z\| = \|x\| \cdot \max_{\|z\|=1} |x^*z| \leq \|x\| \cdot \max_{\|z\|=1} \|x\|\|z\| = x^*x. \tag{1.1}$$

1.2. Estabilidad de un Sistema Lineal en Tiempo Discreto

Esta sección trata del estudio de los Sistemas Lineales de Tiempo Discreto cuya dinámica es invariante respecto del tiempo en contraste con los Sistemas Lineales de Salto Markoviano en Tiempo Discreto donde la dinámica varía de acuerdo a una cadena de Markov (que son nuestro objeto de estudio).

Considere las siguentes ecuaciones en diferencias

$$x(k+1) = f(x(k)) \tag{1.2}$$

y

$$x(k+1) = Ax(k) \tag{1.3}$$

con $k \in \mathbb{N}$, $x(k) \in \mathbb{C}^d$, $f\colon \mathbb{C}^d \to \mathbb{C}^d$, $A \in \mathbb{B}(\mathbb{C}^d)$. Una sucesión $x(0), x(1), \ldots$ generada de acuerdo a (1.2) o (1.3) es llamada la trayectoria del sistema. La segunda ecuación es un caso particular de la primera ecuación y es de mayor interés para nuestro estudio.

Primero recuerde que un punto $x_e \in \mathbb{C}^d$ se llama punto de equilibrio del Sistema (1.2), si $f(x_e) = x_e$ (Esta definición es análoga a punto singular para una ecuación diferencial). En particular, $x_e = 0$ es un punto de equilibrio del Sistema (1.3). Las siguientes definiciones se aplican a los Sistemas (1.2) y (1.3).

Definición 1 (Estabilidad según Lyapunov)**.** *Un punto de equilibrio $x_e \in \mathbb{C}^d$ se dice estable en el sentido de Lyapunov si para cada $\epsilon > 0$ existe $\delta > 0$ tal que $\|x(k) - x_e\| \leq \epsilon$ para todo $k \geq 0$ siempre que $\|x(0) - x_e\| \leq \delta$.*

Definición 2 (Estabilidad asintótica)**.** *Un punto de equilibrio $x_e \in \mathbb{C}^d$ se dice estable asintóticamente si es estable en el sentido de Lyapunov y existe $\delta > 0$ tal que para $\|x(0) - x_e\| \leq \delta$ se tiene $x(k) \to x_e$ cuando $k \to \infty$. Además, el punto x_e es estable global asintóticamente si él es estable asintóticamente y $x(k) \to x_e$ cuando $k \to \infty$ para todo $x(0) \in \mathbb{C}^d$.*

La definición dice que un punto de equilibrio es estable si dada cualquier bola con centro en el punto de equilibrio, se puede encontrar alguna bola (presumiblemente más pequeña) de modo que todas las trayectorias del sistema que inicien en un punto de la segunda bola se mantengan dentro de la primera bola. Además, si todas las trayectorias convergen al punto de equilibrio entonces este es estable asintóticamente. Entonces, el problema a resolver es hallar condiciones suficientes para tener ya sea estabilidad según Lyapunov o estabilidad asintótica.

Definición 3 (Función de Lyapunov)**.** *Sea x_e es un punto de equilibrio para el Sistema* (1.2)*. Una función positiva $\phi\colon D \to \mathbb{R}$, donde D es un abierto conexo tal que $x_e \in D \subset \mathbb{C}^d$, se dice función de Lyapunov para el Sistema* (1.2) *en x_e si*

1. *ϕ es continua.*

2. *$\phi(x_e) < \phi(x)$ para todo $x \in D$ tal que $x \neq x_e$.*

3. *$\Delta\phi(x) := \phi(f(x)) - \phi(x) \leq 0$ para todo $x \in D$.*

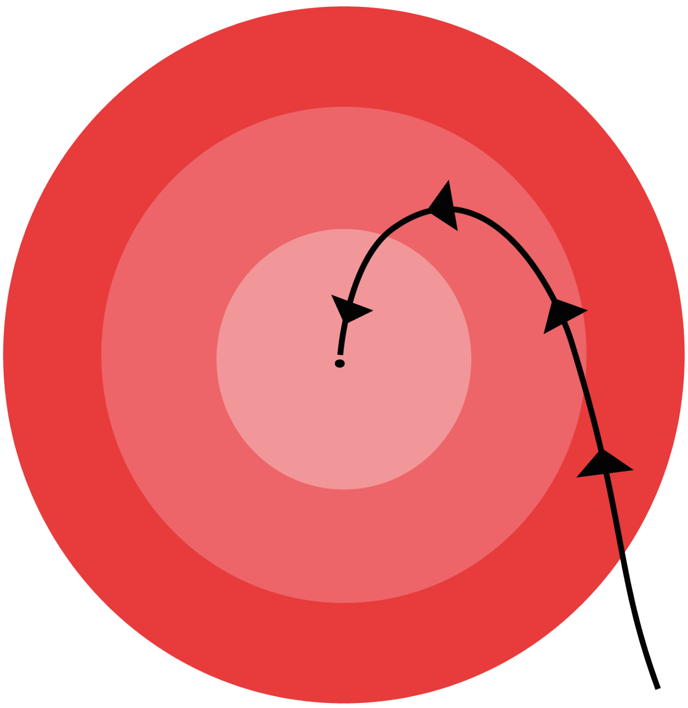

Figura 1.1: Trayectoria subordinada a una función de Lyapunov.
La intensidad del color de las superficies de nivel de la función de Lyapunov es proporcional a los valores asociados, la trayectoria atraviesa el espacio en la dirección de menores valores.

Las propiedades de una función de Lyapunov se interpretan como que varía continuamente en una vecindad del punto de equilibrio (1), que tiene al punto de equilibrio como único mínimo (2) y que la función no crece a lo largo de las trayectorias del sistema (3). De esta última, si la función de Lyapunov solo puede disminuir con el tiempo a medida que el sistema evoluciona entonces debe tender a su valor mínimo, y necesariamente las trayectorias tenderán a quedarse cerca al punto de equilibrio (Ver figura 1.1). Es decir, el punto de equilibrio es estable, esto se precisa en el siguiente teorema.

Teorema 1 (Teorema de Lyapunov, Sección 9.6 de [19]). *Suponga que* $f\colon \mathbb{C}^d \to \mathbb{C}^d$ *es continua. Si existe una función de Lyapunov* ϕ *para el Sistema* (1.2) *en el punto de equilibrio* x_e*, entonces* x_e *es estable en el sentido de Lyapunov. Más aún, si* $\Delta\phi(x) < 0$ *para todo* $x \neq x_e$*, entonces él es estable asintóticamente. Además, si* ϕ *está definido sobre* $\mathbb{C}^d$ *y* $\phi(x)$ *tiende al infinito cuando toda componente de* x *se vuelve arbitrariamente grande en magnitud entonces el punto de equilibrio* x_e *es estable global asintóticamente.*

Demostración. Primero, por simplicidad y sin pérdida de generalidad se puede considerar $D = \{x \in \mathbb{C}^d\colon \|x - x_e\| \leq R_0\}$ y $0 < \epsilon < R_0$. Desde que f es continua en x_e entonces se puede elegir $\delta_1 < \epsilon$ tal que

$$\text{si } \|x - x_e\| < \delta_1 \text{ se verifica } \|f(x) - x_e\| < \epsilon.$$

Sea $m = \min\{\phi(x)\colon \delta_1 \leq \|x - x_e\| \leq R_0\}$, por la continuidad de ϕ en x_e existe $\delta < \delta_1$ tal que

$$\text{si } \|x - x_e\| < \delta \text{ entonces } \phi(x_e) \leq \phi(x) < m.$$

Así, si $\|x(0) - x_e\| \leq \delta$ de la no positividad de $\Delta\phi$ se tiene que $(\phi(x(k)))$ es una sucesión no negativa y no creciente, más áun

$$m > \varphi(x(0)) \geq \varphi(x(1)) \geq \varphi(x(2)) \geq \cdots \geq \varphi(x_e).$$

Se tiene que $\|x(1) - x_e\| < \epsilon$, si se asume que $\|x(1) - x_e\| \geq \delta_1$ entonces $\phi(x(1)) \geq m$, luego $\|x(1) - x_e\| < \delta_1$. Por inducción, si se tiene que $\|x(k) - x_e\| < \delta_1$

entonces $\|x(k+1) - x_e\| < \epsilon$ de modo que por un razonamiento similar al anterior se concluye que $\|x(k+1) - x_e\| < \delta_1$. Por lo tanto toda la sucesión $(x(k))$ está a una distancia menor a $\delta_1 < \epsilon$ del punto de equilibrio x_e, es decir, x_e es estable en el sentido de Lyapunov.

Segundo, se tiene que $\Delta\phi$ es continua no negativa y solo se anula en x_e. Para $\|x(0) - x_e\| < \delta$ de lo anterior

$$\overline{m} = \lim_{k\to\infty} \phi(x(k))$$

está bien definido. Desde que

$$\Delta\phi(x(k)) = \phi(f(x(k))) - \phi(x(k)) = \phi(x(k+1)) - \phi(x(k)) \to 0,$$

y $(x(k))$ admite una subsucesión convergente a $\overline{x}$, ocurre que $\Delta\phi(x(k)) \to \Delta\phi(\overline{x})$, así $\overline{x} = x_e$ y

$$\phi(x(k)) \to \phi(\overline{x}) = \overline{m}.$$

Ahora, suponga que que $x(k) \nrightarrow x_e$ de modo que exista una subsucesión que converge a $\hat{x} \neq x_e$, luego

$$\phi(x(k)) \to \phi(\hat{x}) = \overline{m}$$

lo cual es una contradicción pues de lo anterior $\overline{m} = \phi(x_e)$. Por lo tanto, $x(k) \to x_e$ y x_e es estable asintóticamente.

Tercero, tenemos que ϕ está definido sobre $\mathbb{C}^d$. Sea $x(k)$, $k \geq 0$, con $x(0) \in \mathbb{C}^d$ distinto de x_e, una solución de (1.2). Suponga que la solución $(x(k))$ es acotada considere $\overline{m} = \lim_{k\to\infty} \phi(x(k))$ y razone de forma análoga a como se hizo en el segundo paso para concluir que $x(k) \to x_e$ cuando $k \to \infty$. Ahora, suponga que una solución $(x(k))$ no es acotada, entonces existe una subsucesión $(x(k_j))$ tal que $x(k_j) \to \infty$ cuando $j \to \infty$. Por hipótesis se tiene que $\phi(x(k_j)) \to \infty$, de donde

$$\phi(x(k_i)) > \varphi(x(0))$$

para algún k_i suficientemente grande. Lo último es una contradicción pues $\phi(x(k))$ es no creciente, por lo tanto la solución $(x(k))$ es acotada. $\square$

El Teorema de Lyapunov se aplica, en particular al Sistema (1.3). Considere la posible función de Lyapunov para el Sistema (1.3) dada por

$$\phi(x) = x^* V x \tag{1.4}$$

con $V > 0$. Entonces

$$\begin{aligned} \Delta\phi(x) &= \phi(Ax) - \phi(x) \\ &= (Ax)^* V (Ax) - x^* V x \\ &= x^* A^* V A x - x^* V x \\ &= x^* (A^* V A - V) x. \end{aligned}$$

Con esto se puede presentar el siguiente teorema que establece la relación entre la estabilidad del Sistema (1.3) y la llamada Ecuación de Lyapunov. Este es el teorema central de los sistemas lineales de tiempo discreto, (1.3), que muestra también otras afirmaciones equivalentes para la estabilidad.

Teorema 2 (Teorema de Estabilidad)**.** *La siguiente afirmaciones son equivalentes.*

1. $x_e = 0$ *es el único punto de equilibrio estable global asintóticamente del Sistema* (1.3).

2. $$\rho(A) < 1.$$

3. *Para cada* $U > 0$ *existe una única* $V > 0$ *tal que*

$$V - A^* V A = U. \tag{1.5}$$

4. Para algún $V > 0$, se tiene

$$V - A^* V A > 0.$$

Demostración. Se sigue la secuencia

$$(1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (4) \Rightarrow (1).$$

Primero, para probar que $\rho(A) < 1$ basta ver que el módulo de todo autovalor de A es menor que uno. Sea λ un autovalor de A, considere $x(0)$ un autovector asociado, $Ax(0) = \lambda x(0)$, de modo que

$$x(k) = A^k x(0) = \lambda^k x(0) \to 0$$

cuando $k \to \infty$, y por ello se requiere que $|\lambda| < 1$.

Segundo, la Ecuación de Lyapunov admite solo una solución. Al aplicar el operador φ a la ecuación (1.5) se consigue

$$\begin{aligned} \varphi(U) &= \varphi(V) - \varphi(A^* V A) \\ &= \varphi(V) - (A^t \otimes A^*)\varphi(V) \\ &= (I_{d^2} - A^t \otimes A^*)\varphi(V) \end{aligned} \tag{1.6}$$

Un autovalor de $A^t \otimes A^*$ es de la forma $\lambda_i \overline{\lambda_j}$ con λ_i, λ_j autovalores de A entonces de la hipótesis $\rho(A^t \otimes A) < 1$, más aún $I_{d^2} - A^t \otimes A^*$ es no singular y la solución de (1.6) está dada por

$$\begin{aligned} \varphi(V) &= \sum_{k=0}^{\infty} \left(A^t \otimes A^*\right)^k \varphi\left(U\right) \\ &= \sum_{k=0}^{\infty} \varphi\left((A^*)^k U A^k\right) \\ &= \varphi\left(\sum_{k=0}^{\infty} (A^*)^k U A^k\right). \end{aligned}$$

Como el operador φ transforma soluciones de (1.5) en soluciones de (1.6) biúnivocamente, se verifica que

$$V = \sum_{k=0}^{\infty} (A^*)^k U A^k$$

converge absolutamente y es la solución de (1.5).

Ahora, asuma que U es definida positiva. Luego, como la operación conjugación es continua y aditiva se sigue

$$\left(\sum_{k=0}^{\infty} (A^*)^k U A^k \right)^* = \sum_{k=0}^{\infty} \left((A^*)^k U A^k \right)^* = \sum_{k=0}^{\infty} (A^*)^k U^* A^k,$$

así, V es hermitiana. Para $x \in \mathbb{C}^n$ se tiene

$$\begin{aligned} x^* V x &= \sum_{k=0}^{\infty} x^* (A^*)^k U A^k x \\ &= \sum_{k=0}^{\infty} (A^k x)^* U (A^k x) \geq 0, \end{aligned}$$

además, si $x^* V x = 0$ entonces $x^* U x = 0$ y así $x = 0$. Por lo tanto, V es definida positiva.

Tercero, (4) es concecuencia directa de la afirmación (3).

Cuarto, suponga que $V > 0$ cumple $V - A^* V A > 0$. Claramente $\phi(x) = x^* V x$ define una función de Lyapunov para el Sistema (1.3), del Teorema de Lyapunov se sigue que $x_e = 0$ es estable global asintóticamente. Si x_e es un punto de equilibrio estable global asintóticamente se tiene que $x_e = A x_e$, así

$$0 = \phi(x_e) - \phi(A x_e) = -\Delta\phi(x_e) = {x_e}^* (V - A^* V A) x_e,$$

y por hipótesis $x_e = 0$.

$\square$

Desde que (1.3) tiene un único punto de equilibrio cuando este es estable, es común decir en este caso que el Sistema (1.3) es estable.

Este teorema será extendido para el caso en el que la matriz A evolucione en el tiempo, según una dinámica de Markov, en el Teorema 4 del capítulo 3.

1.3. El espacio $\mathbb{H}^d$

Sea el conjunto de índices $\mathbb{I} = \{1, 2, \ldots, I\}$. Es de interés estudiar las secuencias de I-matrices, donde S indica el número de estados de la cadena de Markov asociada al Sistema (3.1). Entonces, es conveniente introducir el espacio $\mathbb{H}^{d,d'}$, como el espacio vectorial de las secuencias de I-matrices $V = (V_1, \ldots, V_I)$ con $V_i \in \mathbb{B}(\mathbb{C}^d, \mathbb{C}^{d'})$, $i \in \mathbb{I}$. Por simplicidad, sea $\mathbb{H}^d = \mathbb{H}^{d,d}$. Para $V = (V_1, \ldots, V_I) \in \mathbb{H}^{d,d'}$ se definen las siguientes normas equivalentes en el espacio finito dimensional $\mathbb{H}^{d,d'}$:

$$\begin{aligned}
\|V\|_1 &= \sum_{i\in\mathbb{I}} \|V_i\| \\
\|V\|_2 &= \left(\sum_{i\in\mathbb{I}} \operatorname{tr}(V_i^* V_i)\right)^{1/2} \\
\|V\|_\infty &= \text{máx}\{\|V_i\| \colon i \in \mathbb{I}\}.
\end{aligned}$$

Es fácil verificar que $\mathbb{H}^{d,d'}$ es un espacio de Banach con cualquiera de estas normas y, de hecho, $\left(\mathbb{H}^{d,d'}, \|\cdot\|_2\right)$ es un espacio de Hilbert, con el producto interno dado, para $U = (U_1, \ldots, U_I)$ y $V = (V_1, \ldots, V_I)$ en $\mathbb{H}^{d,d'}$, por

$$\langle U, V\rangle = \sum_{i\in\mathbb{I}} \operatorname{tr}({U_i}^* V_i).$$

Es también conveniente definir las siguientes normas equivalentes inducidas $\|\cdot\|_1$ y $\|\cdot\|_2$ en el espacio finito dimensional $\mathbb{B}(\mathbb{H}^d)$. Para $\mathcal{Z}\in\mathbb{B}(\mathbb{H}^d)$,

$$\|\mathcal{Z}\|_1 = \sup_{\|V\|_1=1}\|\mathcal{Z}(V)\|_1, \quad \|\mathcal{Z}\|_2 = \sup_{\|V\|_2=1}\|\mathcal{Z}(V)\|_2$$

Para $V=(V_1,\dots,V_I)\in\mathbb{H}^{d,d'}$ se denota $V^*=(V_1^*,\dots,V_I^*)\in\mathbb{H}^{d',d}$ y se dice que $V\in\mathbb{H}^d$ es hermitiano si $V^*=V$. Sean

$$\mathbb{H}^{d*}=\{V\in\mathbb{H}^d\colon V^*=V\}$$

y

$$\mathbb{H}^{d+}=\{V=(V_1,\dots,V_I)\in\mathbb{H}^{d*}\colon V_i\geq 0, i\in\mathbb{I}\}$$

y se escribe, para $U=(U_1,\dots,U_I)$ y $V=(V_1,\dots,V_I)$ en $\mathbb{H}^d$, que $U\geq V$ si $U-V=(U_1-V_1,\dots,U_I-V_I)\in\mathbb{H}^{d+}$, y que $U>V$ si $U_i-V_i>0$ para todo $i\in\mathbb{I}$. Se dice que un operador $\mathcal{Z}\in\mathbb{B}(\mathbb{H}^d)$, es hermitiano si $\mathcal{Z}(V)\in\mathbb{H}^{d*}$ para cualquier $V\in\mathbb{H}^{d*}$, y que es positivo si $\mathcal{Z}(V)\in\mathbb{H}^{d+}$ para cualquier $V\in\mathbb{H}^{d+}$.

Se define el operador $\hat{\varphi}$ como sigue: para $V=(V_1,\dots,V_I)\in\mathbb{H}^{d,d'}$

$$\hat{\varphi}(V)=\begin{bmatrix}\varphi(V_1)\\ \vdots\\ \varphi(V_I)\end{bmatrix}\in\mathbb{C}^{d'dI}.$$

Observación 2. *Es fácil de verificar que la aplicación $\hat{\varphi}\colon\mathbb{H}^{d,d'}\to\mathbb{C}^{d'dI}$ es un isomorfismo lineal, así todo operador $\mathcal{Z}\in\mathbb{B}(\mathbb{H}^{d,d'})$ pueden ser representado en $\mathbb{B}(\mathbb{C}^{d'dI})$ como la aplicación $\hat{\varphi}[\mathcal{Z}]$ definida por*

$$\hat{\varphi}[\mathcal{Z}](V)=\mathcal{Z}\left(\hat{\varphi}^{-1}(V)\right),\ V\in\mathbb{H}^{d,d'}.$$

Claramente se tiene $\rho(\hat{\varphi}[\mathcal{Z}])=\rho(\mathcal{Z})$ desde que

$$\lambda\in\mathbb{C}\text{ es autovalor de }\hat{\varphi}[\mathcal{Z}]\ \Leftrightarrow\ \lambda\in\mathbb{C}\text{ es autovalor de }\mathcal{Z}.$$

Observación 3. *Para todo* $W \in \mathbb{B}(\mathbb{C}^d)$ *existen* $W^{ij} \in \mathbb{B}(\mathbb{C}^d)$*,* $i, j = 1, 2$*, tales que* $W^{ij} \geq 0$ *y* $\|W^{ij}\| \leq \|W\|$ *para* $i, j = 1, 2$*, y*

$$W = \left(W^{11} - W^{12}\right) + \sqrt{-1}\left(W^{21} - W^{22}\right).$$

De hecho, se puede escribir

$$W = W^1 + \sqrt{-1}W^2$$

donde

$$W^1 = \frac{1}{2}(W^* + W) \qquad W^2 = \frac{1}{2}(W^* - W).$$

Desde que W^1 *y* W^2 *son autoadjuntos, y todo operador autoadjunto en* $\mathbb{C}^d$ *puede descomponerse en su partes positiva y negativa (ver la observación del Teorema 4.2.11 en [23]), se tiene que existen* $W^{ij} \in \mathbb{B}(\mathbb{C}^d)^+$*,* $i, j = 1, 2$*, tales que*

$$W^1 = W^{11} - W^{12} \qquad W^2 = W^{21} - W^{22}.$$

Entonces para todo $V = (V_1, \ldots, V_I) \in \mathbb{H}^d$*, se puede hallar* $V^{ij} \in \mathbb{H}^{d+}$*,* $i, j = 1, 2$ *tales que* $\|V^{ij}\|_1 \leq \|V\|_1$ *y*

$$V = \left(V^{11} - V^{12}\right) + \sqrt{-1}\left(V^{21} - V^{22}\right).$$

La siguiente proposición se deduce de la decomposición de matrices cuadradas en matrices semi-definida positiva dada en la Observación 3.

Proposición 1 ([16], Lema 1)**.** *Sea* $\mathcal{Z} \in \mathbb{B}(\mathbb{H}^d)$*. Las siguientes afirmaciones son equivalentes*

1. *Para todo* $V \in \mathbb{H}^{d+}$ *se cumple*

$$\|\mathcal{Z}^k(V)\|_1 \to 0 \quad \textit{cuando} \quad k \to \infty.$$

2.

$$\rho(\mathcal{Z}) < 1.$$

3. *Para algunos* $0 < \eta < 1 \leq \beta$ *se tiene*

$$\|\mathcal{Z}^k\|_1 \leq \beta\eta^k, \quad k = 0, 1, 2 \ldots.$$

4. *Para todo* $V \in \mathbb{H}^{d+}$

$$\sum_{k=0}^{\infty} \|\mathcal{Z}^k(V)\|_1 < \infty.$$

Demostración. De la Observación 3 para cada $V = (V_1, \ldots, V_S) \in \mathbb{H}^d$, se puede hallar $V^{ij} \in \mathbb{H}^{d+}$, $i, j = 1, 2$ tales que $\|V^{ij}\|_1 \leq \|V\|_1$ y

$$V = \left(V^{11} - V^{12}\right) + \sqrt{-1}\left(V^{21} - V^{22}\right).$$

Desde que $\mathcal{Z}$ es un operador lineal se sigue

$$\begin{aligned}
\|\mathcal{Z}^k(V)\|_1 &= \|\mathcal{Z}^k\left(V^{11}\right) - \mathcal{Z}^k\left(V^{12}\right) + \sqrt{-1}\left(\mathcal{Z}^k\left(V^{21}\right) - \mathcal{Z}^k\left(V^{22}\right)\right)\|_1 \\
&\leq \sum_{i,j=1}^{2} \|\mathcal{Z}^k(V^{ij})\|_1.
\end{aligned}$$

Por lo tanto, tener convergencia para todo $V \in \mathbb{H}^{d+}$ equivale a tener convergencia para todo $V \in \mathbb{H}^d$.

Como en el Teorema 2 se sigue la secuencia

$$(1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (4) \Rightarrow (1).$$

Del razonamiento anterior es fácil verificar todas las implicaciones salvo $(2) \Rightarrow (3)$. Asuma que $\rho(\mathcal{Z}) < 1$, de la fórmula del radio espectral para operadores acotados (Teorema 3.1.10 de [24]), $\lim_{k\to\infty} \|\mathcal{Z}^k\|_1^{1/k} = \rho(\mathcal{Z})$, se consigue que

$$\|\mathcal{Z}^k\|_1^{1/k} < \frac{1 + \rho(\mathcal{Z})}{2} < 1, \quad k = k_0 + 1.k_0 + 2, \ldots$$

con k_0 suficientemente grande. Por lo tanto si se toma

$$\alpha = \frac{1 + \rho(\mathcal{Z})}{2}$$

y

$$\eta \geq \text{máx}\left\{1, \frac{\|\mathcal{Z}^k\|_1}{\alpha}, \ldots, \frac{\|\mathcal{Z}^{k_0}\|_1^{1/k_0}}{\alpha^{k_0}}\right\}$$

se verifica (3).

$\square$

Proposición 2 ([17], Lema 1)**.** *Sea $\mathcal{Z} \in \mathbb{B}(\mathbb{H}^d)$. Si $\rho(\mathcal{Z}) < 1$ entonces para todo $U \in \mathbb{H}^d$ existe una única $V \in \mathbb{H}^d$ tal que*

$$V = \mathcal{Z}(V) + U. \tag{1.7}$$

Más aún

$$V = \hat{\varphi}^{-1}\left((I_{d^2S} - \hat{\varphi}[\mathcal{Z}])^{-1}\hat{\varphi}(U)\right) = \sum_{k=0}^{\infty} \mathcal{Z}^k(U).$$

Además, si $\mathcal{Z}$ es un operador hermitiano entonces

$$U = U^* \Leftrightarrow V = V^*$$

y si $\mathcal{Z}$ es un operador positivo

$$U \geq 0 \Rightarrow V \geq 0$$

$$U > 0 \Rightarrow V > 0.$$

Demostración. Aplicando el operador $\hat{\varphi}$ a la Ecuación (1.7) se tiene

$$\hat{\varphi}\,(V) = \hat{\varphi}[\mathcal{Z}]\hat{\varphi}\,(V) + \hat{\varphi}\,(U)\,. \tag{1.8}$$

De la Observación 2 se tiene que $\rho\left(\hat{\varphi}[\mathcal{Z}]\right) < 1$, así, la ecuación previa admite una única solución. Si V es solución de (1.7) entonces $\hat{\varphi}(V)$ es solución (1.8), por lo tanto (1.8) admite una única solución dada por

$$V = \hat{\varphi}^{-1}\left((I_{d^2S} - \hat{\varphi}[\mathcal{Z}])^{-1}\hat{\varphi}(U)\right).$$

Además, de la proposición anterior se tiene que $\sum_{k=0}^{\infty} \mathcal{Z}^k(U)$ es convergente y además verifica la Ecuación (1.8), luego

$$V = \sum_{k=0}^{\infty} \mathcal{Z}^k(U). \tag{1.9}$$

Ahora, suponga que $\mathcal{Z}$ es hermitiano entonces si $V = V^*$ tomando tranpuesta conjugada a la Ecuación (1.7) se concluye $U = U^*$. Por otro lado, por inducción se verifica que $\mathcal{Z}^k$ también es hermitiano para cada $k \geq 0$ de modo que si $U = U^*$ de la fórmula de sumatoria para V, Ecuación (1.9), se tiene que $V = V^*$.

Por último, de manera similar las dos últimas implicaciones se derivan usando la fórmula de sumatoria para V. $\square$

Capítulo 2

El modelo estocástico

Aqui se trata la construcción de un marco estocástico que deriva de un Sistema Lineal de Salto Markoviano en Tiempo Discreto, en el cual se puede estudiar las soluciones de este modelo estocástico. Para ello previamente se adicionan las notaciones de la teoría de probabilidad haciendo enfásis en el concepto de esperanza condicional y se revisa la teoría de cadenas de Markov tratando especialmente el problema de la construcción como proceso estocástico sobre el espacio de realizaciones.

2.1. Teoría de la probabilidad

Los fundamentos de la teoría de la probabilidad se pueden hallar en libros clásicos como [3], [8] o [21].

Suponga que $(\Omega, \mathcal{F}, \mathbb{P})$ es un espacio de probabilidad. Este es el modelo matemático para el espacio de las realizaciones o resultados de un experimento aleatorio. Aquí Ω es un conjunto, $\mathcal{F}$ es un σ-álgebra sobre Ω y $\mathbb{P}$ es una medida en $\mathcal{F}$ con $\mathbb{P}(\Omega) = 1$. Los elementos de $\mathcal{F}$ se llaman eventos, y se entiende que $\mathcal{F}$ es la información provista por el experimento aleatorio. Un elemento arbitrario de Ω es denotado por ω. Cada elemento ω es interpretado como una realización del experimento aleatorio. Se dice que una Propiedad ocurre *casi seguramente*, escrito c.s., cuando la probabilidad del evento definido por todas las realizaciones ω que satisfacen la Propiedad tiene medida 1. Suponga que $(\mathbb{S}, \mathcal{S})$ es un espacio medible. Este es el modelo del espacio de estados de un experimento aleatorio. Típicamente $\mathbb{S}$ es un espacio métrico completo y separable con σ-álgebra de Borel $\mathcal{S} = \mathcal{B}_{\mathbb{S}}$, particularmente el espacio de estados puede ser $\mathbb{R}^d$ (o $\mathbb{C}^d$) y $\mathbb{B}(\mathbb{R}^d)$ (o $\mathbb{B}(\mathbb{C}^d)$). Un elemento de $\mathbb{S}$ es un posible estado del fenómeno estocástico.

Las funciones medibles $\mathbb{S}$-valuadas se denominan *elementos aleatorios* o *variables de estado* y generalmente se indican con X, Y, Z u otras letras mayúsculas. Particularmente, si el espacio de estados es $\mathbb{R}$ ($\mathbb{C}$) un elemento aleatorio es conocido como variable aleatoria real (compleja) o simplemente variable aleatoria; si es $\mathbb{R}^d$ ($\mathbb{C}^d$), un elemento aleatorio es conocido como vector aleatorio real (complejo); y si es $\mathbb{B}(\mathbb{R}^d)$, un elemento aleatorio es conocido como matriz aleatoria real (compleja). La integral de una variable aleatoria real respecto a $\mathbb{P}$ es llamada *esperanza* o *valor esperado*, y se escribe $\mathbb{E}[X]$ para $\int X \, d\mathbb{P}$. La esperanza de una variable aleatoria real X es un promedio generalizado del conjunto de valores que puede tomar X, interpretación igualmente válida para las definiciones de esperanza que siguen. La integral de una variable aleatoria compleja se define como la suma compleja de la integral de la parte real más $\sqrt{-1}$ veces la integral de la parte imaginaria. La integral de una vector aleatorio (o una matriz aleatoria) se define componente a componente. Siempre que esté definido la integral de un elemento aleatorio X se mantiene el nombre de esperanza así como la notacion $\mathbb{E}[X]$ para la integral $\int X \, d\mathbb{P}$. La variable aleatoria 1_A es la función que vale 1 sobre A y 0 en otro caso. Esta es llamada la *indicadora* de A. Los eventos tales como $\{\omega \colon X(\omega) \in A\}$, con $A \in \mathcal{S}$, son usualmente abreviados por

$\{X \in A\}$. Para seguir el mismo espíritu de la interpretación de σ-álgebra, la información asociada al elemento aleatorio X es la colección de todos los eventos $\{X \in A\}$, esto resulta un σ-álgebra conocido como *σ-álgebra generado* por X y denotado por $\sigma(X)$. Además, los eventos de la forma $\{\omega\colon X(\omega) > a\}$ (con $a \in \mathbb{R}$) se abrevian por $\{X > a\}$.

Dado un elemento aleatorio X, se puede definir una probabilidad sobre el espacio de estados $\mathbb{S}$ por

$$\mathbb{P}_X(A) = \mathbb{P}(X \in A), \quad A \in \mathcal{S}.$$

La probabilidad $\mathbb{P}_X$ es llamada la *ley* de X o la *distribución* de X, para indicar esto se denota $X \sim \mathbb{P}_X$. Cuando X es una variable aleatoria real se define $F_X\colon \mathbb{R} \to [0,1]$ por

$$F_X(x) = \mathbb{P}_X(]-\infty, x]) = \mathbb{P}(X \leq x), \quad x \in \mathbb{R}.$$

La función F_X es llamada *función de distribución* de X. Está función caracteriza a la ley de una variable aleatoria en el sentido siguiente:

$$\mathbb{P}_X = \mathbb{P}_Y \text{ si y solo si } F_X = F_Y.$$

Por esta razón, también se usa la notación $X \sim F_X$ a cambio de $X \sim \mathbb{P}_X$. Algunos ejemplos de distribuciones de variables aleatorias son las siguientes.

1. *Uniforme discreta*: Se denota $X \sim \mathrm{Uni}(x_1, \ldots, x_n)$ para indicar que la variable aleatoria X toma valores sobre $\{x_1, \ldots, x_n\}$ y

$$\mathrm{Uni}(x_1, \ldots, x_n)(\{x_k\}) = \mathbb{P}(X = x_k) = \frac{1}{n}, \ \ k = 1, \ldots, n.$$

2. *Binomial* de parámetros n y p: Se denota $X \sim \mathrm{Bin}(n,p)$, $n \geq 1, 0 \leq p \leq 1$, para indicar que X tomas valores sobre $\{0, 1, \ldots, n\}$ y

$$\mathrm{Bin}(n,p)(\{k\}) = \mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \ \ k = 0, 1, \ldots, n.$$

Si F es absolutamente continua, se dice que $f = F'$ es la *densidad* de F. Si tal F es la función de distribución de una variable aleatoria X, entonces

$$\mathbb{P}_X(A) = \mathbb{P}(X \in A) = \int_A f(x)dx.$$

Aquí, sigue siendo válido escribir $X \sim f_X$ en lugar de $X \sim F_X$. Algunos ejemplos de distribuciones con densidad son las siguientes.

1. *Uniforme* sobre un abierto (o un cerrado) E de $\mathbb{R}$: Se denotada $X \sim \mathrm{Uni}(E)$ para indicar que X toma valores sobre E y admite función de densidad igual a

$$\frac{1}{\lambda(E)} 1_E(x), \quad x \in \mathbb{R},$$

donde λ es la medida de Lebesgue en $\mathbb{R}$.

2. *Gaussiana* de parámetros μ y σ^2: Se denota $X \sim N(m, \sigma^2)$, $m \in \mathbb{R}, \sigma^2 \geq 0$, para indicar que X toma valores sobre $\mathbb{R}$ y admite una función de densidad igual a

$$\frac{1}{\sqrt{2\pi}\sigma} e^{-(x-m)^2/(2\sigma^2)}, \quad x \in \mathbb{R}.$$

Cuando $m = 0$ y $\sigma^2 = 1$, tal X se llama *Gaussiana estándar*.

Es posible utilizar la ley de una variable aleatoria para computar esperanzas: Si g es medible acotada o no negativa, entonces

$$\mathbb{E}[g(X)] = \int g(x)\mathbb{P}_X(dx).$$

Si F_X tiene una densidad f entonces

$$\mathbb{E}[X] = \int x f(x)dx.$$

La media de una variable aleatoria es su valor esperado. El p-ésimo momento de X es $\mathbb{E}[X^p]$ donde p es un entero positivo. La varianza de una variable aleatoria X es el segundo momento de la variable centrada $X - \mathbb{E}[X]$. Para el cálculo de los momentos de una variable aleatoria es útil la siguiente proposición: Si $X \geq 0$ y $p > 0$, entonces

$$\mathbb{E}[X^p] = \int px^{p-1}\mathbb{P}(X > x)dx.$$

La prueba de esta afirmación muestra que la igualdad también es válida si se reemplaza $\mathbb{P}(X > x)$ por $\mathbb{P}(X \geq x)$. Una primera desigualdad elemental es la llamada desigualdad de Markov: Si $X \geq 0$ y $a > 0$,

$$\mathbb{P}(X \geq a) \leq \frac{1}{a}\mathbb{E}X.$$

Una segunda desigualdad elemental es la desigualdad de Jensen: Si g es una función convexa sobre $\mathbb{R}$ y $\mathbb{P}(X = a) = 1 - \mathbb{P}(X = b) = p \geq 0$ entonces la desigualdad de convexidad

$$g(pa + (1-p)b) \leq pg(a) + (1-p)g(b)$$

se escribe

$$g(\mathbb{E}X) \leq \mathbb{E}g(X). \tag{2.1}$$

Esta desigualdad sigue siendo válida cuando X es un vector aleatorio en $\mathbb{R}^d$ integrable y g es una función convexa sobre $\mathbb{R}^d$ tal que $g(X)$ es integrable (Lema 3.5 en [15]).

Se dice que dos eventos A y B son *independientes* si cumplen el principio multiplicativo $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. Los eventos $A_1, \ldots, A_n$ son *independientes* si

$$\mathbb{P}(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_m}) = \mathbb{P}(A_{i_1})\mathbb{P}(A_{i_2}) \cdots \mathbb{P}(A_{i_m})$$

para cada subconjunto $\{i_1, \ldots, i_m\}$ de $\{1, \ldots, n\}$ con $1 \leq i_1 < \cdots < i_m \leq n$. Si A y B son independientes entonces A^c y B son independientes, esto motiva la siguiente definición. Se dice que dos σ-álgebras $\mathcal{G}$ y $\mathcal{H}$ son *independientes* si A y B son independientes para todos $A \in \mathcal{G}$ y $B \in \mathcal{H}$. Dos elementos aleatorios X e Y son *independientes* si lo son sus σ-álgebras generadas $\sigma(X), \sigma(Y)$. La definición de independencia de n σ-álgebras o de n elementos aleatorios es hecha de manera similar. Si f y g son funciones medibles y X e Y son independientes, entonces $f(X)$ y $g(Y)$ son independientes. Esto se sigue desde que la información proporcionada por $f(Z)$ es una subinformación de la proporcionada por Z.

La función $F_{X,Y}(x, y) = \mathbb{P}(X \leq x, Y \leq y)$ denota la llamada *función distribución conjunta* de las variables aleatorias X e Y (La coma dentro de los conjuntos significa "intersección"; esta es una convención estándar en probabilidad). Dada esta definición, se tiene la siguiente caracterización:

$$F_{X,Y}(x, y) = F_X(x)F_Y(y),\ x, y \in \mathbb{R} \quad \Leftrightarrow \quad X \text{ e } Y \text{son independentes.}$$

Es conocido que si X e Y son independientes y X e Y son integrables o no negativas, entonces

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y].$$

De manera similar se puede definir la función de distribución conjunta de un vector aleatorio real de dimensión $d \geq 2$. Incluso, se puede definir una función de distribución sobre $\mathbb{C}$ y, por lo tanto, una función de distribución conjunta de un vector complejo. Por esta razón, se tienen las siguientes generalizaciones de una distribución Gaussiana a espacios Euclideanos.

1. Distribución Gaussiana multivariante en $\mathbb{R}^d$: Se denota $X \sim N(m, \Sigma)$, $m \in \mathbb{R}^d$, $\Sigma \in \mathbb{B}(\mathbb{R}^d), \Sigma \geq 0$, para indicar que X toma valores en $\mathbb{R}^d$ y admite una función de densidad igual a

$$\frac{1}{\sqrt{2\pi \det(\Sigma)}} \exp\left\{-\frac{1}{2}(x - m)^t \Sigma^{-1}(x - m)\right\}, \quad x \in \mathbb{R}^d.$$

Aquí, $m = \mathbb{E}X$ es llamado *vector de medias* y

$$\Sigma = \mathbb{E}[(X - m)(X - m)^t]$$

es llamada *matriz de covarianzas*. Cuando $m = 0$ y $\Sigma = Id$, tal X se llama *vector Gaussiano estándar*.

2. Distribución Gaussiana multivariante compleja en $\mathbb{C}^d$: Se denota $X \sim CN(m, \Sigma, K)$, $m \in \mathbb{C}^d, \Sigma, K \in \mathbb{B}(\mathbb{C}^d), \Sigma \geq 0$, para indicar que $X = (X_1, \ldots . x_d)$ toma valores en $\mathbb{C}^d$ y el vector
$$(\mathrm{Re}X_1, \ldots, \mathrm{Re}X_d, \mathrm{Im}X_1, \ldots, \mathrm{Im}X_d)$$
tiene distribución Gaussiana en $\mathbb{R}^{2d}$ con vector de medias y matriz de convarianzas únicamente determinadas por las matrices m, Σ, K. Aquí, $m = \mathbb{E}X$ es llamado *vector de medias* y
$$\Sigma = \mathbb{E}[(X - m)(X - m)^*] \tag{2.2}$$
es llamada *matriz de covarianzas* y
$$K = \mathbb{E}[(X - m)(X - m)^t]$$
es llamada *matriz de pseudo-covarianzas*. Para más detalle revisar la sección 3.7 de [12].

Si $\{A_k\}_{k\geq 1}$ es una sucesión de conjuntos, se define $\{A_k \ i.v.\}$, leído "A_k infinitas veces", por
$$\{A_k \ i.v.\} = \bigcap_{m\geq 1} \bigcup_{k\geq m} A_k.$$
Este conjunto consiste de los elementos ω que están en infinitos A_k's. Una simple pero importante proposición es el lema Borel–Cantelli. Este consta de dos partes, considere que $A_1, A_2, \ldots$ es una sucesión de eventos. Primero,
$$\sum_{k=1}^{\infty} \mathbb{P}(A_k) < +\infty \quad \Rightarrow \quad \mathbb{P}(A_k \ i.v.) = 0.$$
Segundo, si además la sucesión de eventos es independiente
$$\sum_{k=1}^{\infty} \mathbb{P}(A_k) = +\infty \quad \Rightarrow \quad \mathbb{P}(A_k \ i.v.) = 1.$$

También, se define $\{A_k\ eventualm.\}$, leído "A_k eventualmente", por

$$\{A_k\ eventualm.\} = \bigcup_{m\geq 1}\bigcap_{k\geq m} A_k.$$

Este conjunto consiste de los elementos ω que están en los A_k's a partir de cierto índice k_0. Se tiene la relación natural

$$\{A_k\ eventualm.\} \subset \{A_k\ i.v.\}.$$

Además, usando las leyes de Morgan se tiene que

$$\{{A_k}^c\ eventualm.\} = \{A_k\ i.v.\},$$

y por lo tanto, cuando $A_1, A_2, \ldots$ es una sucesión de eventos,

$$\mathbb{P}(A_k\ i.v.) = 0 \quad \Rightarrow \quad \mathbb{P}({A_k}^c\ eventualm.) = 1$$

(Aquí, A^c denota el complemento del evento A).

Ahora, existen varios tipos de respuestas a la pregunta ¿la sucesión de elementos aleatorios X_k converge? Primero, asuma que el espacio de estados $\mathbb{S}$ es un espacio métrico separable con métrica d. Una manera natural de responder es decir que *X_k converge a X casi seguramente*, es decir, si el evento $\{\mathrm{d}(X_k, X) \to 0\}$ tiene probabilidad 1. Otra forma de respuesta es decir que *X_k converge a X en probabilidad* si para cada $\epsilon > 0$,

$$\mathbb{P}(\mathrm{d}(X_k, X) > \epsilon) \to 0$$

cuando $k \to \infty$. Segundo, cuando $\mathbb{S}$ es un espacio de Banach, para $p \geq 1$, se dice que X_k *converge a X en L_p* si

$$\mathbb{E}\left[\|X_k - X\|^p\right] \to 0$$

cuando $k \to \infty$. Y tercero, deje que $\mathbb{S}$ sea solo un espacio métrico con métrica d, se dice que X_k *converge a X en ley* (o *en distribución*) si

$$\mathbb{E}[h(X_k)] \to \mathbb{E}[h(X)]$$

para toda función $h\colon \mathbb{S} \to \mathbb{R}$ continua y acotada. El siguiente enunciado muestra las relaciones entre estos tipos de convergencia.

1. Si $X_k \to X$ casi seguramente, entonces $X_k \to X$ en probabilidad.
2. Si $X_k \to X$ en L_p, entonces $X_k \to X$ en probabilidad.
3. Si $X_k \to X$ en probabilidad, entonces existe una subsucesión n_j tal que $X_{k_j} \to X$ casi seguramente.
4. Si $X_k \to X$ en probabilidad, entonces $X_k \to X$ en distribución.

El tipo de convergencia que nos interesa en el presente trabajo es la convergencia en L^2 también llamada convergencia en "media cuadrática". Note que la convergencia en media cuadrática implica la convergencia casi segura, algo similar también ocurre cuando se habla de estabilidad de un Sistema Lineal de Salto Markoviano en Tiempo Discreto.

Esperanza condicional

Considere que $\mathcal{G} \subset \mathcal{F}$ es un σ-álgebra y X una variable aleatoria integrable, la *esperanza condicional de X dado $\mathcal{G}$*, denotada $\mathbb{E}[X|\mathcal{G}]$ y leído "la esperanza de X dado $\mathcal{G}$", es una variable aleatoria W, $\mathcal{G}$-medible, tal que

$$\mathbb{E}[W1_A] = \mathbb{E}[X1_A], \quad A \in \mathcal{G}. \tag{2.3}$$

La *probabilidad condicional de $A \in \mathcal{F}$ dado $\mathcal{G}$* es definida por $\mathbb{P}(A|\mathcal{G}) = \mathbb{E}[1_A|\mathcal{G}]$.

Si Y_1, Y_2 son dos variables aleatorias $\mathcal{G}$-medibles con

$$\mathbb{E}[Y_1 1_A] = \mathbb{E}[Y_2 1_A], \quad A \in \mathcal{G},$$

entonces $Y_1 = Y_2$, c.s., y así la esperanza condicional es única bajo equivalencia casi segura.

En el caso que X sea $\mathcal{G}$-medible, $\mathbb{E}[X|\mathcal{G}] = X$. Si X es independiente de G, $\mathbb{E}[X|\mathcal{G}] = EX$. Ambos casos se siguen inmediatamente de la definición. Para otro ejemplo, si $\{A_i\}_i$ es una colección finita (o numerable) de eventos disjuntos dos a dos cuya unión es Ω, $\mathbb{P}(A_i) > 0$ para todo i, y $\mathcal{G}$ es el σ-álgebra generada por los A_i's, entonces

$$\mathbb{P}(A|\mathcal{G}) = \sum_i \frac{\mathbb{P}(A \cap A_i)}{\mathbb{P}(A_i)} 1_{A_i}.$$

Esto se sigue desde que el lado derecho de la igualdad anterior es $\mathcal{G}$-medible y tomando esperanza se verifica la Ecuación (2.3). Este candidato proviene de la definición de probabilidad condicional elemental: Si $\mathbb{P}(B) > 0$, entonces

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Una propiedad importante y fácil de verificar a partir de (2.3) es la ley de la esperanza iterada:

$$\mathbb{E}\left[\mathbb{E}[X|\mathcal{G}]\right] = \mathbb{E}[X]. \tag{2.4}$$

De hecho, la esperanza condicional es una generalización de la esperanza. Fijado un σ-álgebra $\mathcal{G}$, la existencia de la esperanza condicional de X dado $\mathcal{G}$ está garantizada por la integrabilidad de X. Se usa la notación $\mathbb{E}[X|\mathcal{G}]$ para referirnos a alguna W que verifique las propiedades que definen la esperanza condicional. El operador $\mathbb{E}[\cdot|\mathcal{G}]$ definido sobre el espacio de las variables aleatorias integrables es lineal y monótono:

$$\mathbb{E}[aX + Y|\mathcal{G}] = a\mathbb{E}[X|\mathcal{G}] + \mathbb{E}[Y|\mathcal{G}] \ c.s.,$$
$$X \leq Y \Rightarrow \mathbb{E}[X|\mathcal{G}] \leq \mathbb{E}[Y|\mathcal{G}] \ c.s.$$

Se puede verificar que los teoremas límites tales como el teorema de convergencia monótona y convergencia dominada tienen su versión con esperanza condicional, así como las desigualdades de Markov y de Jensen. Así, por ejemplo, se tiene la desigualdad de Jensen para esperanza condicional: Si g es convexa

$$g\left(\mathbb{E}[X|\mathcal{G}]\right) \leq \mathbb{E}[g(X)|\mathcal{G}] \ c.s.$$

siempre que X y $g(X)$ sean integrables. Un hecho clave es el siguiente: Si X y XY son integrables e Y es $\mathcal{G}$-medible, entonces

$$\mathbb{E}[XY|\mathcal{G}] = Y\mathbb{E}[X|\mathcal{G}] \; c.s.$$

Una generalización de (2.4) es la ley del condicionamiento iterado: Si $\mathcal{H} \subset \mathcal{G} \subset F$ son σ-álgebras, entonces

$$\mathbb{E}\left[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}\right] = \mathbb{E}[X|\mathcal{H}] = \mathbb{E}\left[\mathbb{E}[X|\mathcal{H}]|\mathcal{G}\right]. \tag{2.5}$$

Como caso particular se tiene que

$$\mathbb{E}\left[\mathbb{E}[X|\mathcal{G},\mathcal{H}]|\mathcal{H}\right] = \mathbb{E}[X|\mathcal{H}]$$

donde $\mathbb{E}[X|\mathcal{G},\mathcal{H}] := \mathbb{E}[X|\sigma(\mathcal{G},\mathcal{H})]$.

Cuando $\mathcal{G} = \sigma(Y)$, usualmente se escribe $\mathbb{E}[X|Y]$ y $\mathbb{P}(A|Y)$ para $\mathbb{E}[X|\mathcal{G}]$ y $\mathbb{P}(A|\mathcal{G})$, respectivamente (Es decir, se acostumbra escribir el elemento aleatorio a cambio de su σ-álgebra generada). Una notación que es comúnmente usada es $\mathbb{E}[X|Y = y]$. Para dar significado a esto se apela al lema de Doob-Dynkin: Si W es $\sigma(Y)$-medible entonces $W = h(Y)$ para alguna función h Borel medible. Sea $W = E[X|Y]$ y elija una función h Borel medible tal que $W = h(Y)$, entonces $E[X|Y = y]$ es definido como $h(y)$.

Si $X \in L_2$ y $M = \{W \in L_2 \colon W \text{ es } \mathcal{G}\text{-medible}\}$, se puede mostrar que $E[X|\mathcal{G}]$ es igual a la proyección de X sobre el subespacio M. Esto da nuevas luces en la interpretación de la esperanza condicional, ahora se puede entender $\mathbb{E}[X|\mathcal{G}]$ como la mejor aproximación de X en media, dada la información proporcionada por $\mathcal{G}$. Así muchas propiedades ya mencionadas se vuelven intuitivas. También, de la misma manera como se hizo con la integral, se puede definir la esperanza condicional de variables aleatorias complejas y vectores o matrices aleatorias.

Un concepto importante es la independencia condicional: Sea $\mathcal{G} \subset \mathcal{F}$ un σ-álgebra, se dice que dos eventos $A, B \in \mathcal{F}$ son independientes dado $\mathcal{G}$ si

$$\mathbb{P}(A \cap B|\mathcal{G}) = \mathbb{P}(A|\mathcal{G})\mathbb{P}(B|\mathcal{G}),$$

y se dice que dos elementos aleatorios X e Y son independientes dado $\mathcal{G}$ si

$$\mathbb{P}(X \in A, Y \in B|\mathcal{G}) = \mathbb{P}(X \in A|\mathcal{G})\mathbb{P}(Y \in B|\mathcal{G}) \tag{2.6}$$

para todo par de medibles A y B. Luego, X e Y son independientes dado $\mathcal{G}$ si y solo si

$$\mathbb{E}[g(X)h(Y)|\mathcal{G}] = \mathbb{E}[g(X)|\mathcal{G}]\mathbb{E}[h(Y)|\mathcal{G}]$$

para todo par de funciones g, h medibles y acotadas. Otra forma útil de la independencia condicional es la siguiente: X e Y son independientes dado $\mathcal{G}$ si y solo si

$$\mathbb{E}[h(Y)|\mathcal{G}, X] = \mathbb{E}[h(Y)|\mathcal{G}]$$

para toda función h medible y acotada, o lo que es lo mismo,

$$\mathbb{P}(Y \in B|\mathcal{G}, X) = \mathbb{P}(Y \in B|\mathcal{G})$$

para todo medible B. Cabe mencionar que esta noción de independencia condicional generaliza la noción previa de independencia.

2.2. Cadenas de Markov

Suponga que T es un conjunto de índices. Dado un experimento aleatorio, T denotará el espacio temporal asociado a este experimento. Usualemente $T = \mathbb{N}$ (tiempo discreto), $T = [0, 1]$ o $T = [0, +\infty[$ (tiempo continuo). Un proceso estocástico es entendido como una colección de observaciones de la realización de un sistema que evoluciona a lo largo del tiempo, esto motiva la definición dada a continuación.

Definición 4. *Una colección de elementos aleatorios $(Z_t)_{t\in T}$ sobre un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ con valores en un espacio medible $(\mathbb{S}, \mathcal{S})$ es llamado* proceso estocástico *$\mathbb{S}$-valuado con conjunto de índices T.*

Recuerde que la noción de información es capturada por el concepto de σ-álgebra, más aún, una cadena de información en aumento es lo que se conoce como filtración. Una familia de σ-álgebras $(\mathcal{F}_t)_{t\in T} \subset \mathcal{F}$ es llamada *filtración* si $\mathcal{F}_t \supset \mathcal{F}_u$ para todo $t, u \in T$ con $t \geq u$. Se dice que un proceso estocástico $Z = (Z_t)_{t\in T}$ es *adaptado* a una filtración $(\mathcal{F}_t)_{t\in T}$ si $\sigma(Z_t) \subset \mathcal{F}_t$ para todo $t \in T$. El proceso de observación de un sistema que evoluciona con el tiempo induce una cadena de información dada por la recolección de las observaciones hasta el tiempo t. Así, todo proceso estocástico $(Z_t)_{t\in T}$ siempre admite una filtración que lo adapte, llamada *filtración canónica*, dada por: $\mathcal{F}_t = \sigma(Z_u \colon u \leq t)$, para cada $t \in T$.

Definición 5 (Propiedad de Markov)**.** *Un proceso estocástico $(Z_t)_{t\in T}$ tiene la* Propiedad de Markov *si para todos $t > u$ en T y para todo $B \in \mathcal{S}$ se cumple que*

$$\mathbb{P}(Z_t \in B|\mathcal{F}_u) = \mathbb{P}(Z_t \in B|Z_u) \tag{2.7}$$

En adelante, se cuando se trate con procesos de Markov en tiempo discreto, en este caso $T = \mathbb{N}$, el proceso estocástico se acostumbra a denotar $(Z_k)_{k\geq 0}$ y la propiedad de Markov se escribe: para todo $k \geq 0$ y $B \in \mathcal{S}$ se verifica

$$\mathbb{P}(Z_{k+1} \in B|Z_k, \dots, Z_0) = \mathbb{P}(Z_{k+1} \in B|Z_k).$$

Según la noción de independencia condicional, la propiedad de Markov nos dice que el futuro Z_{k+1} es independiente del pasado $Z_{k-1}, \dots, Z_0$ dado el presente Z_k, es decir,

$$\mathbb{E}[g(Z_{k+1})h(Z_{k-1}, \dots, Z_0)|Z_k] = \mathbb{E}[g(Z_{k+1})|Z_k]\mathbb{E}[h(Z_{k-1}, \dots, Z_0)|Z_k]$$

para todo par de funciones medibles y acotadas g, h (tales que la igualdad anterior tenga sentido).

Ahora, note que la dinámica de transición de un estado al estado posterior inmediato está descrita por

$$\mathbb{P}(Z_{k+1} \in B | \mathcal{F}_k),$$

de modo que depende del instante k, el estado Z_k y los posibles estados posteriores B (esto por la Propiedad de Markov). Cuando esta dinámica no muda respecto a la variable temporal k, se dice que el proceso tiene la Propiedad de homogeneidad temporal. La precisión de estas nociones se hacen en seguida.

Definición 6. *Una función* $P\colon \mathbb{S} \times \mathcal{S} \to \mathbb{R}$ *se llama* probabilidad de transición *si*

a. Para cada $z \in \mathbb{S}$, $A \mapsto P(z, A)$ es una medida de probabilidad sobre $(\mathbb{S}, \mathcal{S})$.

b. Para cada $A \in \mathcal{S}$, $z \mapsto P(z, A)$ es una función medible.

Se dice que un proceso estocástico $(Z_k)_{k \geq 0}$ es una proceso de Markov en tiempo discreto *(adaptado a la filtración $(\mathcal{F}_k)_{k \geq 0}$) con probabilidad de transición P y* distribución inicial *μ si*

1. *μ es una probabilidad sobre $(\mathbb{S}, \mathcal{S})$ y $Z_0 \sim \mu$.*

2. *Para cada $k \geq 0$ y cada $B \in \mathcal{S}$,*

$$\mathbb{P}(Z_{k+1} \in B | \mathcal{F}_k) = P(Z_k, B).$$

Ejemplo 1. *Sea (X_k) una sucesión de variables aleatorias Bernoulli idénticamente distribuidas con $\mathbb{P}(X_1 = 1) = 1 - \mathbb{P}(X_1 = -1) = p \geq 0$. Se define un camino aleatorio en $\mathbb{S} = \mathbb{Z}$ como la sucesión de variables aleatorias reales dada por*

$$Z_k = x + \sum_{i=1}^{k} X_i, k \geq 0.$$

Entonces, la familia (Z_k) es un proceso de Markov. Fijado ω, la aplicación $k \mapsto Z_k(\omega)$ describe el camino de una partícula que parte desde el punto x haciendo saltos a sus vecinos más cercanos con una preferencia p de ir hacia la derecha. Note que para el camino aleatorio la distribución inicial es una medida delta de Dirac concentrada en x y la probabilidad de transición está determinada por el arreglo

$$Q(i,j) = \begin{cases} p & j = i+1 \\ 1-p & j = i-1 \\ 0 & \textit{c.o.c} \end{cases},$$

de la siguiente manera

$$P(i,A) = \sum_{j \in A} Q(i,j),\ i \in \mathbb{Z},\ A \subset \mathbb{Z}.$$

En general, para toda realización $\omega \in \Omega$, la aplicación $t \mapsto Z_t(\omega)$ se llama *trayectoria* del proceso, el espacio de las trayectorias del proceso es el espacio funcional $\mathbb{S}^T$ (ver figura 2.1). Ahora, se puede pensar al proceso estocástico como un elemento aleatorio que para cada realización muestre la trayectoria asociada, es decir, como la función $Z\colon \Omega \to \mathbb{S}^T$, $\omega \mapsto Z_{(\cdot)}(w)$. Entonces, surge la cuestión de saber si esta función es medible (respecto del σ-álgebra producto $\mathcal{S}^T$), o en menor exigencia, si induce de manera única una probabilidad en $\mathbb{S}^T$. Una respuesta afirmativa requiere que la colección de elementos aleatorios que conforman el proceso estocástico satisfaga cierta condición de consistencia. Los procesos de Markov en tiempo discreto si gozan de cierta consistencia que deriva de la Propiedad de Markov.

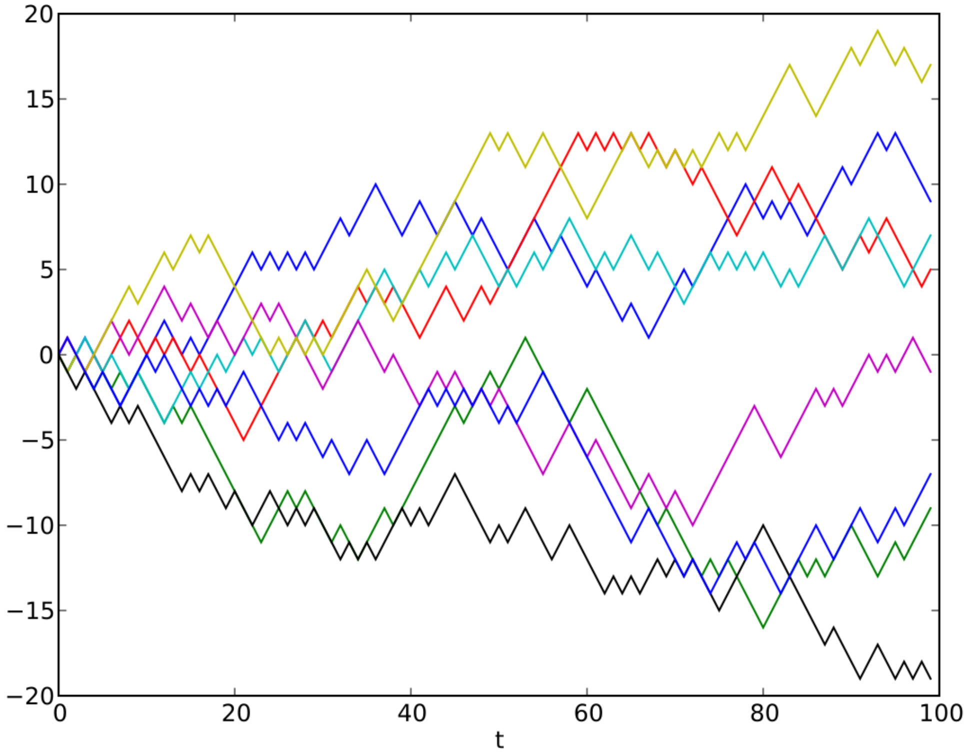

Figura 2.1: Camino aleatorio en $\mathbb{Z}$.

Evolución temporal de las trayectorias del camino aleatorio (Fuente: Wikipedia).

Sea $(Z_k)_{k\geq 0}$ un proceso de Markov en tiempo discreto con probabilidad de transición P y distribución inicial μ. La familia de probabilidades definidas por

$$\mathbb{P}(Z_0 \in B_0, \ldots, Z_k \in B_k) = \int_{B_0} \mu(dz_0) \int_{B_1} P(z_0, dz_1) \cdots \int_{B_k} P(z_{k-1}, dz_k), \tag{2.8}$$

donde $B_0, \ldots, B_k \in \mathcal{S}$ y $k \geq 0$, tiene la siguiente Propiedad: para todo $k \geq 0$,

$$\mathbb{P}((Z_0, \ldots, Z_k) \in B, Z_{k+1} \in \mathbb{S}) = \mathbb{P}((Z_0, \ldots, Z_k) \in B), \quad B \in \mathcal{S}^{k+1}. \tag{2.9}$$

El espacio producto $\left(\Pi_{l=0}^k \mathbb{S}, \otimes_{l=0}^k \mathcal{S}\right)$ se abrevia por $(\mathbb{S}^{k+1}, \mathcal{S}^{k+1})$. Recuerde que una familia de probabilidades finito dimensionales $(\nu_k)_{k\geq 0}$, donde ν_k es una probabilidad sobre

$(\mathbb{S}^{k+1}, \mathcal{S}^{k+1})$, se llama *consistente* si para todo $k \geq 0$

$$\nu_{k+1}(B \times \mathbb{S}) = \nu_k(B), \quad B \in \mathcal{S}^{k+1}.$$

Teorema 3 (Teorema de Extensión de Kolmogorov)**.** *Sea $(\nu_k)_{k\geq 0}$ una familia consistente de probabilidades finito dimensionales. Si $\mathbb{S}$ es un espacio métrico completo y separable entonces existe una única medida de probabilidad sobre $(\mathbb{S}^{\mathbb{N}}, \mathcal{S}^{\mathbb{N}})$ tal que las proyecciones $\pi_{0,\dots,k}\colon \mathbb{S}^{\mathbb{N}} \to \mathbb{S}^{k+1}$ tienen como ley a ν_k, $k \geq 0$.*

Demostración. Considere el álgebra de los cilindros finito dimensionales

$$\mathcal{A} = \{\pi_{0,\dots,k}^{-1}(B)\colon B \in \mathcal{S}^{k+1}, k \geq 0\}.$$

Desde que el σ-álgebra producto $\mathcal{S}^{\mathbb{N}}$ es generado por el álgebra de los cilindros finito dimensionales, el Teorema de Extensión de Caratheodory (Teorema A.1.3 de [8]) nos dice que es suficiente definir una aplicación sobre $\mathcal{A}$ que sea σ-aditiva. En nuestro caso, se puede probar aditividad y continuidad en el vacío que implican la σ-aditividad.

Note que para $B \in \mathcal{S}^{k+1}$ se puede imaginar la siguiente cadena

$$\nu_k(B) = \nu_{k+1}(B \times \mathbb{S}) = \nu_{k+2}\left(B \times \mathbb{S}^2\right) = \cdots = \nu\left(B \times \mathbb{S}^{\{k+2,k+3,\dots\}}\right),$$

donde ν sería la probabilidad que queremos definir. Esto sugiere que efectivamente tal ν verificaría $\pi_{0,\dots,k} \sim \nu_k$, y que es natural definir

$$\nu(\pi_{0,\dots,k}^{-1}(B)) = \nu_k(B), \quad \text{si } B \in \mathcal{S}^{k+1},$$

como la aplicación candidata. En efecto, esta aplicación está bien definida: Si $\pi_{0,\dots,k}^{-1}(B) = \pi_{0,\dots,n}^{-1}(A)$ con $k \leq n$ entonces $A = B \times \mathbb{S}^{n-k}$ y así

$$\nu\left(\pi_{0,\dots,n}^{-1}(A)\right) = \nu_n\left(B \times \mathbb{S}^{n-k}\right) = \cdots = \nu_k(B) = \nu\left(\pi_{0,\dots,k}^{-1}(B)\right).$$

Más aún, ν es aditiva: Sean $\pi^{-1}_{0,\dots,k}(B), \pi^{-1}_{0,\dots,n}(A)$ disjuntos con $k \leq n$, entonces

$$\begin{aligned}
\nu\left(\pi^{-1}_{0,\dots,n}(A) \cup \pi^{-1}_{0,\dots,k}(B)\right) &= \nu\left(\pi^{-1}_{0,\dots,n}(A \cup B \times \mathbb{S}^{n-k})\right) \\
&= \nu_n\left(A \cup B \times \mathbb{S}^{n-k}\right) \\
&= \nu_n(A) + \nu_n\left(B \times \mathbb{S}^{n-k}\right) \\
&= \nu\left(\pi^{-1}_{0,\dots,n}(A)\right) + \nu\left(\pi^{-1}_{0,\dots,k}(B)\right).
\end{aligned}$$

Resta ver que ν es continua en el vacío. Sea $\tilde{B}_n = \pi^{-1}_{0,\dots,k_n}(B_n)$, $B_n \in \mathcal{S}^{k_n+1}$, una sucesión tal que $\tilde{B}_n \downarrow \emptyset$ se quiere probar que

$$\nu(\tilde{B}_n) \downarrow 0.$$

Por contradicción, sin pérdida de generalidad se puede suponer que k_n es creciente y que existe $\delta > 0$ de modo que

$$\delta \leq \nu(\tilde{B}_n) = \nu_{k_n}(B_n), \quad n \geq 1.$$

Por la Propiedad de rigidez de una medida finita sobre un espacio métrico completo y separable (Proposición 2.7.18 de [21]) se tiene que

$$\nu(\tilde{B}_n \setminus \tilde{K}_n) = \nu_{k_n}(B_n \setminus K_n) \leq \frac{1}{2^{n+1}}\delta$$

para algún compacto $K_n \subset B_n$, $n \geq 1$. Sea $\tilde{C}_n = \bigcap_{i=1}^n \tilde{K}_i$ entonces

$$\begin{aligned}
\nu_{k_n}(C_n) = \nu(\tilde{C}_n) = \nu(\tilde{B}_n) - \nu\left(\tilde{B}_n \setminus \tilde{C}_n\right) &\geq \delta - \nu\left(\bigcup_{i=1}^n (\tilde{B}_n \setminus \tilde{K}_i)\right) \\
&\geq \delta - \nu\left(\bigcup_{i=1}^n (\tilde{B}_i \setminus \tilde{K}_i)\right) \\
&\geq \delta - \sum_{i=1}^n \nu\left(\tilde{B}_i \setminus \tilde{K}_i\right) \\
&\geq \delta - \sum_{i=1}^n \frac{1}{2^{i+1}}\delta \geq \frac{1}{2}\delta.
\end{aligned}$$

Dado que los $\tilde{C}_n$'s son conjuntos no vacíos asociados a los compactos K_n's y el espacio de estados es completo, se puede esperar que

$$\bigcap_{n=1}^{\infty} \tilde{C}_n = \bigcap_{n=1}^{\infty} \tilde{K}_n \neq \emptyset,$$

lo cual claramente es una contradicción con la hipótesis $\tilde{B}_n \downarrow \emptyset$. En efecto, para cada $n \geq 1$ elija

$$(z_0^n, z_1^n, \dots) \in \tilde{C}_n = \bigcap_{i=1}^{n} \pi^{-1}_{0,\dots,k_i}(K_i).$$

Luego, como la secuencia k_n es creciente, la sucesión $(z_0^n, \dots z_{k_1}^n)$ está el compacto K_1 y admite una subsucesión convergente a un punto $(z_0, \dots z_{k_1}) \in K_1$, por simplicidad,

$$(z_0^n, \dots z_{k_1}^n) \to (z_0, \dots z_{k_1}) \quad \text{cuando } n \to \infty.$$

De igual manera, prácticamente se puede considerar que la sucesión $(z_0^n, \dots z_{k_2}^n)$ en el compacto K_2 es tal que

$$(z_0^n, \dots z_{k_2}^n) \to (z_0, \dots z_{k_1}, z_{k_1+1} \dots, z_{k_2}) \in K_2 \quad \text{cuando } n \to \infty.$$

Este procedimiento se puede seguir replicando, de este modo queda definido recursivamente el punto $z = (z_0, z_1, \dots) \in \mathbb{S}^{\mathbb{N}}$ con la Propiedad

$$\pi_{0,\dots,k_n}(z) = (z_0, \dots, z_{k_n}) \in K_n, \quad n \geq 1.$$

Es decir, se tiene un punto

$$z \in \bigcap_{n=1}^{\infty} \tilde{K}_n.$$

□

El Teorema de Extensión de Kolmogorov resuelve el problema de garantizar la existencia de un proceso estocástico. Este teorema muestra la construcción de un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P}) = \left(\mathbb{S}^{\mathbb{N}}, \mathcal{S}^{\mathbb{N}}, \nu\right)$ llamado *espacio canónico* en el cual se puede definir un proceso estocástico dado por las proyecciones canónicas $\pi_k \colon \mathbb{S}^{\mathbb{N}} \to \mathbb{S}$, $k \geq 0$, llamado *proceso canónico*. En particular, cuando se trabaja con un proceso de Markov en tiempo discreto no es díficil verificar de las ecuaciones (2.8) y (2.9) que la familia de probabilidades finito dimensionales

$$\nu_k = \mathbb{P}_{(Z_0,\ldots,Z_k)}, \quad k \geq 0$$

es consistente, por lo tanto, existe el espacio canónico

$$(\Omega_\mu, \mathcal{F}_\mu, P_\mu) = \left(\mathbb{S}^{\mathbb{N}}, \mathcal{S}^{\mathbb{N}}, \nu\right)$$

y el proceso de Markov canónico (π_k). De hecho, lo que vale para un proceso de Markov canónico también vale para un proceso de Markov sobre un espacio de probabilidad arbitrario, por tal razón se puede asumir que se esta trabajando con el proceso de Markov canónico. Cuando la distribución inicial es determinística $\mu = \delta_{\{x\}}(\cdot)$, con $x \in \mathbb{S}$, se denota E_x para la esperanza respecto a $P_x := P_\mu$. Por último, cuando un proceso de Markov en tiempo discreto toma valores en un espacio de estados numerable se acostumbra a denominar cadena de Markov en tiempo discreto.

2.3. Marco estocástico de un Sistema Lineal de Salto Markoviano en Tiempo Discreto

Sea el modelo estocástico dado por la ecuación

$$x(k+1) = A_{\theta(k)} x(k) \tag{2.10}$$

donde $\theta(k)$ es una cadena de Markov tomando valores en un conjunto numerable $\mathbb{I}$ (En el capítulo siguiente $\mathbb{I} = \{1, \ldots, I\}$). Luego la colección de vectores acoplados $(x(k), \theta(k))$ toma valores en $\mathbb{C}^d \times \mathbb{I}$, para cada $k \geq 0$, este proceso se llamará *proceso acoplado*. Como ocurrió en la sección anterior con un proceso de Markov en tiempo discreto, surge la pregunta ¿se puede garantizar la existencia de un proceso estocástico de esta naturaleza?. Si, de hecho, el proceso acoplado es un proceso de Markov en tiempo discreto con espacio de estados $\mathbb{C}^d \times \mathbb{I}$.

Sea el espacio de las trayectorias del proceso acoplado

$$\Omega = \left(\mathbb{C}^d \times \mathbb{I}\right)^{\mathbb{N}}$$

con σ-álgebra producto

$$\mathcal{F} = \left(\mathcal{B}_{\mathbb{C}^d} \times 2^{\mathbb{I}}\right)^{\mathbb{N}}.$$

Se sabe que el espacio de estados $\mathbb{C}^d \times \mathbb{I}$ provisto con la topología produto es un espacio métrico completo separable (Combinar el Corolario 14.1.8 y la Proposición 13.5.1 de [13]). Luego, para mostrar que el proceso acoplado define una familia consistente de probabilidades finito dimensionales es suficiente ver que efecctivamente es un proceso de Markov en tiempo discreto, esto se basa en la interpretación de la propiedad de Markov como una independencia entre futuro y pasado dado el presente, como se muestra en la siguiente proposición.

Proposición 3. *Sea $\{\theta(k)\}_{k\geq 0}$ una cadena de Markov sobre $\mathbb{I}$, y sea $\{x(k)\}_{k\geq 0}$ un proceso estocástico definido por la Ecuación (2.10) con estado inicial x_0 independiente de la cadena de Markov, ambos procesos definidos sobre un espacio de probabilidad $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mathcal{P})$. Entonces el proceso acoplado $\{(x(k), \theta(k))\}_{k\geq 0}$ es un proceso de Markov en tiempo discreto sobre $\mathbb{C}^d \times \mathbb{I}$.*

Demostración. Para $k \geq 0$, sea $\mathcal{G}_k = \sigma(\theta(l) \colon l \leq k)$ la filtración canónica de la cadena de Markov $\theta(k)$ entonces la filtración

$$\mathcal{F}_k = \sigma(\mathcal{G}_k, x(0))$$

adapta al proceso acoplado $(x(k), \theta(k))$.

Sea $k \geq 0$, desde que $x(k+1) = A_{\theta(k)}x(k)$ es $\sigma(x(k), \theta(k))$-medible se sigue

$$\begin{aligned} E\left[1_{B^1 \times B^2}(x(k+1), \theta(k+1)) \mid \mathcal{F}_k\right] &= E\left[1_{B^1}(x(k+1))1_{B^2}(\theta(k+1)) \mid \mathcal{F}_k\right] \\ &= 1_{B^1}(x(k+1))E\left[1_{B^2}(\theta(k+1)) \mid \mathcal{G}_k, x(0)\right] \end{aligned}$$

de la independencia entre $x(0)$ y $\theta(k+1)$, dado el σ-álgebra $\mathcal{G}_k$,

$$\begin{aligned} &= 1_{B^1}(x(k+1))E\left[1_{B^2}(\theta(k+1)) \mid \mathcal{G}_k\right] \\ &= 1_{B^1}(x(k+1))E\left[1_{B^2}(\theta(k+1)) \mid \theta(k)\right] \end{aligned}$$

de la independencia entre $\theta(k+1)$ y el vector $x(k)$, dado $\theta(k)$,

$$\begin{aligned} &= 1_{B^1}(x(k+1))E\left[1_{B^2}(\theta(k+1)) \mid \theta(k), x(k)\right] \\ &= E\left[1_{B^1 \times B^2}(x(k+1), \theta(k+1)) \mid \theta(k), x(k)\right] \ c.s., \end{aligned}$$

para todo $B^1 \times B^2 \in \mathcal{B}_{\mathbb{C}^d} \times 2^{\mathbb{I}}$.

Sea la familia de medibles en $\mathcal{B}_{\mathbb{C}^d} \times 2^{\mathbb{I}}$ que satisfacen la Propiedad de Markov para el proceso acoplado

$$\mathcal{M} = \left\{B \in \mathcal{B}_{\mathbb{C}^d} \times 2^{\mathbb{I}} \colon E[1_B(x(k+1), \theta(k+1)) \mid \mathcal{F}_k] = E\left[1_B(x(k+1), \theta(k+1)) \mid \theta(k), x(k)\right] \ c.s.\right\},$$

entonces previamente se ha mostrado que

$$\mathcal{C} \subset \mathcal{M}$$

donde

$$\mathcal{C} = \left\{B_1 \times B_2 \in \mathcal{B}_{\mathbb{C}^d} \times 2^{\mathbb{I}} \colon B_1 \in \mathcal{B}_{\mathbb{C}^d}, B_2 \in 2^{\mathbb{I}}\right\}$$

es la familia de rectángulos medibles de $\mathcal{F}$. Desde que la familia de rectángulos medibles es un π-sistema que genera el σ-álgebra producto, por el Lema de Dynkin (Teorema A.1.4 del Apéndice de [8]), basta demostrar que $\mathcal{M}$ es un λ-sistema para concluir lo requerido.

En efecto, primero note que $\mathbb{C}^d \times \mathbb{I} \in \mathcal{M}$, segundo si $B \subset A$ en $\mathcal{M}$ entonces por linealidad de esperanza condicional se tiene

$$\begin{aligned}
E[1_{A\setminus B}(x(k+1),\theta(k+1)) \mid \mathcal{F}_k] &= E[(1_A - 1_B)(x(k+1),\theta(k+1)) \mid \mathcal{F}_k] \\
&= E[1_A(x(k+1),\theta(k+1)) \mid \mathcal{F}_k] \\
&\quad - E[1_B(x(k+1),\theta(k+1)) \mid \mathcal{F}_k] \\
&= E[1_A(x(k+1),\theta(k+1)) \mid \theta(k), x(k)] \\
&\quad - E[1_B(x(k+1),\theta(k+1)) \mid \theta(k), x(k)] \\
&= E[(1_A - 1_B)(x(k+1),\theta(k+1)) \mid \theta(k), x(k)] \\
&= E\left[1_{A\setminus B}(x(k+1),\theta(k+1)) \mid \theta(k), x(k)\right] \ c.s.,
\end{aligned}$$

y tercero si B_n en $\mathcal{M}$ con $B_n \uparrow B$ entonces por la versión condicional del teorema de convergencia monótono

$$\begin{aligned}
E[1_B(x(k+1),\theta(k+1)) \mid \mathcal{F}_k] &= \lim_{n\to\infty} E[1_{B_n}(x(k+1),\theta(k+1)) \mid \mathcal{F}_k] \\
&= \lim_{n\to\infty} E[1_{B_n}(x(k+1),\theta(k+1)) \mid \theta(k), x(k)] \\
&= E[1_B(x(k+1),\theta(k+1)) \mid \theta(k), x(k)] \ c.s.
\end{aligned}$$

□

Como el proceso acoplado es Markoviano, el Teorema de Extensión de Kolmogorov garantiza la existencia del espacio canónico $(\Omega, \mathcal{F}, \mathbb{P})$ y el proceso acoplado canónico definido por las proyecciones $\pi_k \colon \Omega \to \mathbb{C}^d \times \mathbb{I}$. La filtración asociada al proceso acoplado canónico es la filtración definida por

$$\mathcal{F}_k = \sigma(\pi_0, \ldots, \pi_k), \quad k \geq 0.$$

Entonces queda determinado el espacio de probabilidad filtrado canónico $(\Omega, \mathcal{F}, (\mathcal{F}_k), \mathbb{P})$ sobre el cual se estudiará al Sistema (2.10). De hecho, si $\pi_k = (r(k), \delta(k))$

con $r(k)\colon :\Omega\to\mathbb{C}^d$, $\delta(k)\colon :\Omega\to\mathbb{I}$ entonces $\{\delta(k)\}$ es una cadena de Markov y $\{r(k)\}$ es un proceso que cumple la ecuación

$$r(k+1)=A_{\delta(k)}r(k),\ k\geq 0.$$

Sobre el espacio de Hilbert $C^d=L_2(\Omega,\mathcal{F},\mathbb{P},\mathbb{C}^d)$ de las variables aleatorias $\mathcal{F}$-medibles $\mathbb{C}^d$-valuadas cuadrado integrables se define el producto interno

$$\langle X,Y\rangle=\mathbb{E}[X^*Y]=\sum_{i=1}^{d}\mathbb{E}\left[\overline{X_i}Y_i\right],$$

donde $\mathbb{E}[\cdot]$ es la esperanza de variables aleatorias complejas, y la norma asociada denotada por $\|\cdot\|_2$. Sea $l_2(C^d)=\Pi_{k\geq 0}C^d$ el espacio producto de las sucesiones en C^d, $r=\{r(k)\colon k\geq 0\}$, con $r(k)\in C^d$ para cada $k\geq 0$, y tales que

$$\|r\|_2^2=\sum_{k\geq 0}\|r(k)\|_2^2=\mathbb{E}\left[\sum_{k\geq 0}r(k)^*r(k)\right]<\infty.$$

Para cada $r,s\in l_2(C^d)$, el producto interno $\langle r,s\rangle$ es dado por

$$\langle r,s\rangle:=\sum_{k\geq 0}\langle r(k),s(k)\rangle\leq\|r\|_2\|s\|_2.$$

Se define el subespacio $\mathcal{C}^d\subset l_2(C^d)$ de la siguiente manera: $r\in\mathcal{C}^d$ si $r\in l_2(C^d)$ y $r(k)\in L_2(\Omega,\mathcal{F}_k,\mathbb{P},\mathbb{C}^d)$ para cada $k\geq 0$. Se tiene que $\mathcal{C}^d$ es un subespacio vectorial cerrado de $l_2(C^d)$ y por lo tanto un espacio de Hilbert, este es el espacio que contiene a las soluciones del Sistema (2.10). También se define $\mathcal{C}_0^d$ como el formado por los elementos $r_0\in L_2(\Omega,\mathcal{F}_0,\mathbb{P},\mathbb{C}^d)$, este es el espacio de los vectores de estado inicial del Sistema (2.10). Finalmente se define Θ_0 como el conjunto de todas las variables $\mathcal{F}_0$-medibles $\mathbb{I}$-valuadas, este es el espacio de los modos de operación inicial del Sistema (2.10).

Capítulo 3

Estudio de los Sistemas Lineales de Salto Markoviano en Tiempo Discreto

Ahora se aborda el estudio del comportamiento de las trayectorias definidas por un Sistema lineal de Salto Markoviano en Tiempo Discreto, llamado simplemente Sistema Lineal de Salto Markoviano. Un enfoque moderno es el tratamiento hecho en la referencia [6], en el cual está basado nuestro trabajo. Primero, se define la notación a usar y la noción de estabilidad del Sistema Lineal de Salto Markoviano como un comportamiento asintótico de las soluciones del sistema. Segundo, se revisan algunos ejemplos para luego concluir con la prueba del teorema central que versa sobre las distintas caracterizaciones de la estabilidad del Sistema Lineal de Salto Markoviano.

3.1. Definición y operadores asociados

En adelante, se fija el marco estocástico $(\Omega, \mathcal{F}, \{\mathcal{F}_k\}, \mathbb{P})$, por ejemplo el marco estocástico descrito en la tercera sección del capítulo anterior. Aquí $\{\theta(k)\colon k \geq 0\}$ es una cadena de Markov adaptada a la filtracción $(\mathcal{F}_k)$ tomando valores sobre $\mathbb{I} = \{1, 2, \ldots, I\}$, en este caso, su probabilidad de transición está dada por una matriz (de transición) $P = [p_{ij}]$. Las condiciones iniciales del sistema son el *vector de estado inicial* $x_0 \in \mathcal{C}_0^d$ y la *variable salto* inicial $\theta_0 \in \Theta_0$ se asumen independientes. Los *modos de operación* del sistema están dados por la secuencia de I-matrices $A = (A_1, \ldots, A_I) \in \mathbb{H}^d$. Así, se define el Sistema lineal de salto Markoviano en Tiempo Discreto mediante la dinámica siguiente

$$\begin{cases} x(k+1) = A_{\theta(k)} x(k), \quad k \geq 0 \\ x(0) = x_0, \quad \theta(0) = \theta_0 \end{cases}. \tag{3.1}$$

Se sabe que, en general, el proceso $\{x(k)\colon k \geq 0\}$ no es un cadena de Markov, sin embargo, el proceso acoplado $\{(x(k), \theta(k))\colon k \geq 0\}$ si lo es. Esto es justamente la Proposición 3, haciendo un razonamiento similar se llega a concluir el importante lema que viene a continuación.

Lema 1 ([6], Sección 2.3). *Dados $k \geq 0$ y $j \in \mathbb{I}$,*

$$\mathbb{P}(\theta(k+1) = j | x(k), \theta(k)) = \mathbb{P}(\theta(k+1) = j | \theta(k)) = p_{\theta(k)j},$$

donde

$$p_{\theta(k)j} = \sum_{i \in \mathbb{I}} p_{ij} 1_{\{\theta(k)=i\}}.$$

Demostración. Como $\sigma(x(k), \theta(k)) \subset \sigma(x(0), \theta(0), \ldots, \theta(k))$ se sigue

$$\begin{aligned} \mathbb{P}(\theta(k+1) = j | x(k), \theta(k)) &= \mathbb{E}[\mathbb{E}[1_{\{\theta(k+1)=j\}} | x(k), \theta(k)] | x(0), \theta(0), \ldots, \theta(k)] \\ &= \mathbb{E}[\mathbb{E}[1_{\{\theta(k+1)=j\}} | x(0), \theta(0), \ldots, \theta(k)] | x(k), \theta(k)], \end{aligned}$$

de la independencia de $\theta(k+1)$ y $x(0)$,

$$= \mathbb{E}[\mathbb{E}[1_{\{\theta(k+1)=j\}}|\theta(0),\ldots,\theta(k)]|x(k),\theta(k)]$$

de la Propiedad de Markov,

$$\begin{aligned} &= \mathbb{E}[\mathbb{E}[1_{\{\theta(k+1)=j\}}|\theta(k)]|x(k),\theta(k)] \\ &= \mathbb{E}[1_{\{\theta(k+1)=j\}}|\theta(k)]. \end{aligned}$$

$\square$

Ahora antes de pasar a estudiar la estabilidad del sistema que nos hemos propuesto necesitamos estudiar primero los valores esperados de la variable de estado en cada período temporal.

Primero, note que

$$\mathbb{E}[x(k)] = \sum_{i\in\mathbb{I}} \mathbb{E}\left[x(k)1_{\{\theta(k)=i\}}\right]$$

y

$$\mathbb{E}[x(k)x(k)^*] = \sum_{i\in\mathbb{I}} \mathbb{E}\left[x(k)x(k)^* 1_{\{\theta(k)=i\}}\right].$$

Luego, para cada $k \geq 0$ e $i \in \mathbb{I}$ se consideran las notaciones:

$$q(k) = \begin{bmatrix} q_1(k) \\ \vdots \\ q_I(k) \end{bmatrix} \in \mathbb{C}^{dS}, \tag{3.2}$$

$$q_i(k) = \mathbb{E}\left[x(k)1_{\{\theta(k)=i\}}\right] \in \mathbb{C}^d, \tag{3.3}$$

$$Q(k) = (Q_1(k),\ldots,Q_I(k)) \in \mathbb{H}^{d+}, \tag{3.4}$$

$$Q_i(k) = \mathbb{E}\left[x(k)x(k)^* 1_{\{\theta(k)=i\}}\right] \in \mathbb{B}(\mathbb{C}^d)^+, \tag{3.5}$$

y

$$\mu(k) = \mathbb{E}[x(k)] = \sum_{i\in\mathbb{I}} q_i(k) \in \mathbb{C}^d, \tag{3.6}$$

$$\mathbb{M}(k) = \mathbb{E}\left[x(k)x(k)^*\right] = \sum_{i\in\mathbb{I}} Q_i(k) \in \mathbb{B}(\mathbb{C}^d)^+. \tag{3.7}$$

A continuación, se deducen ecuaciones recursivas para (3.3) y (3.5).

Proposición 4 ([6], Proposición 3.1)**.** *Para todo* $k \geq 0$ *y* $j \in \mathbb{I}$ *se verifica*

1. $q_j(k+1) = \sum_{i\in\mathbb{I}} p_{ij} A_i q_i(k)$.
2. $Q_j(k+1) = \sum_{i\in\mathbb{I}} p_{ij} A_i Q_i(k) A_i^*$.
3. $\mathbb{E}\left[\|x(k)\|^2\right] = \mathbb{E}\left[\|A_{\theta(k-1)} \cdots A_{\theta(0)} x(0)\|^2\right] \leq d\|Q(k)\|_1$.
4. $\|Q(k)\|_1 \leq \mathbb{E}\left[\|x(k)\|^2\right]$.

Demostración. Dados $k \geq 0$ y $j \in \mathbb{I}$. Para (1), de la definición de $q_j(k+1)$ se sigue

$$\begin{aligned}
q_j(k+1) &= \mathbb{E}\left[A_{\theta(k)} x(k) 1_{\{\theta(k+1)=j\}}\right] \\
&= \mathbb{E}\left[\sum_{i\in\mathbb{I}} A_i x(k) 1_{\{\theta(k)=i\}} 1_{\{\theta(k+1)=j\}}\right] \\
&= \sum_{i\in\mathbb{I}} A_i \mathbb{E}\left[x(k) 1_{\{\theta(k)=i\}} 1_{\{\theta(k+1)=j\}}\right] \\
&= \sum_{i\in\mathbb{I}} A_i \mathbb{E}\left[\mathbb{E}\left[x(k) 1_{\{\theta(k)=i\}} 1_{\{\theta(k+1)=j\}} | x(k), \theta(k)\right]\right] \\
&= \sum_{i\in\mathbb{I}} A_i \mathbb{E}\left[x(k) 1_{\{\theta(k)=i\}} \mathbb{P}\left(\theta(k+1) = j | x(k), \theta(k)\right)\right] \\
&= \sum_{i\in\mathbb{I}} A_i \mathbb{E}\left[x(k) 1_{\{\theta(k)=i\}} p_{ij}\right] \\
&= \sum_{i\in\mathbb{I}} p_{ij} A_i q_i(k).
\end{aligned}$$

Analógamente para (2), de la definición de $Q_j(k+1)$ se tiene

$$\begin{aligned}
Q_j(k+1) &= \mathbb{E}\left[A_{\theta(k)}x(k)x(k)^*A_{\theta(k)}{}^*1_{\{\theta(k+1)=j\}}\right] \\
&= \sum_{i\in\mathbb{I}} A_i\mathbb{E}\left[x(k)x(k)^*1_{\{\theta(k)=i\}}1_{\{\theta(k+1)=j\}}\right]A_i^* \\
&= \sum_{i\in\mathbb{I}} A_i\mathbb{E}\left[\mathbb{E}\left[x(k)x(k)^*1_{\{\theta(k)=i\}}1_{\{\theta(k+1)=j\}}|x(k),\theta(k)\right]\right]A_i^* \\
&= \sum_{i\in\mathbb{I}} A_i\mathbb{E}\left[x(k)x(k)^*1_{\{\theta(k)=i\}}p_{ij}\right]A_i^* \\
&= \sum_{i\in\mathbb{I}} p_{ij}A_iQ_i(k)A_i^*.
\end{aligned}$$

Dado $k \geq 0$. Para (3),

$$\begin{aligned}
\mathbb{E}[\|x(k)\|^2] = \mathbb{E}\left[x(k)^*x(k)\right] &= \sum_{i\in\mathbb{I}} \mathbb{E}\left[x(k)^*x(k)1_{\{\theta(k)=i\}}\right] \\
&= \sum_{i\in\mathbb{I}} \mathbb{E}\left[\text{tr}(x(k)x(k)^*1_{\{\theta(k)=i\}})\right] \\
&= \sum_{i\in\mathbb{I}} \text{tr}\left(\mathbb{E}\left[x(k)x(k)^*1_{\{\theta(k)=i\}}\right]\right) \\
&= \text{tr}\left(\sum_{i\in\mathbb{I}} \mathbb{E}\left[x(k)x(k)^*1_{\{\theta(k)=i\}}\right]\right) \\
&= \text{tr}\left(\sum_{i\in\mathbb{I}} Q_i(k)\right) \\
&\leq d\left\|\sum_{i\in\mathbb{I}} Q_i(k)\right\| \\
&\leq d\sum_{i\in\mathbb{I}} \|Q_i(k)\| = d\|Q(k)\|_1.
\end{aligned}$$

Para (4), de la desigualdad del Teorema I.11, capitulo 3, en [18] y de la desigualdad (1.1), se tienen las dos desigualdades siguientes

$$\begin{aligned}
\|Q(k)\|_1 &= \sum_{j\in\mathbb{I}} \left\| \mathbb{E}\left[x(k)x(k)^* 1_{\{\theta(k)=i\}} \right] \right\| \\
&\leq \sum_{j\in\mathbb{I}} \mathbb{E}\left[\|x(k)x(k)^*\| \, 1_{\{\theta(k)=i\}} \right] \\
&\leq \sum_{j\in\mathbb{I}} \mathbb{E}\left[x(k)^* x(k) 1_{\{\theta(k)=i\}} \right] = \mathbb{E}\left[\|x(k)\|^2 \right].
\end{aligned}$$

□

La Proposición 4 sugiere definir los siguientes operadores lineales, que tienen como dominio el espacio de las sucesiones de I-matrices $\mathbb{H}^d$ (el cual fue estudiado en la última sección del capítulo 1),

$$\begin{aligned}
\mathcal{T}(\cdot) &= (\mathcal{T}_1(\cdot), \ldots, \mathcal{T}_S(\cdot)) \in \mathbb{B}(\mathbb{H}^d) \\
\mathcal{L}(\cdot) &= (\mathcal{L}_1(\cdot), \ldots, \mathcal{L}_S(\cdot)) \in \mathbb{B}(\mathbb{H}^d)
\end{aligned}$$

como sigue: para $V = (V_1, \ldots, V_L) \in \mathbb{H}^d$ e $i, j \in \mathbb{I}$

$$\mathcal{T}_j(V) = \sum_{i\in\mathbb{I}} p_{ij} A_i V_i A_i^* \in \mathbb{B}(\mathbb{C}^d) \tag{3.8}$$

$$\mathcal{L}_i(V) = \sum_{j\in\mathbb{I}} p_{ij} A_i^* V_j A_i \in \mathbb{B}(\mathbb{C}^d). \tag{3.9}$$

Con estas definiciones, de la Proposición 4 se tiene que la conexión entre el operador $\mathcal{T}$ y la matriz de segundos momentos definida en (3.4) está dada por

$$Q(k+1) = \mathcal{T}(Q(k)) \tag{3.10}$$

y así

$$Q(k) = \mathcal{T}^k(Q(0)).$$

De las definición de $\mathcal{T}$, para $V \in \mathbb{H}^d$ se tiene $\mathcal{T}_j(V)^* = \mathcal{T}_j(V^*)$ y por lo tanto $\mathcal{T}(V)^* = \mathcal{T}(V^*)$, luego, el operador $\mathcal{T}$ como aplicación de $\mathbb{H}^{d*}$ a $\mathbb{H}^{d*}$ es hermitiano, similarmente para $\mathcal{L}$. Además, desde que una aplicación $V \mapsto AVA^*$ preserva la propiedad de ser semidefinido positivo y $p_{ij} \geq 0$ se tiene que $\mathcal{T}$ como aplicación de $\mathbb{H}^{d+}$ a $\mathbb{H}^{d+}$ es positivo, similarmente para $\mathcal{L}$. Más aún, la siguiente proposición muestra la relación entre ambos operadores.

Proposición 5 ([6], Proposición 3.2)**.** *El operador adjunto de $\mathcal{T}$ es $\mathcal{L}$, es decir,*

$$\mathcal{T}^* = \mathcal{L}.$$

Demostración. Sean $U, V \in \mathbb{H}^d$ entonces se sigue

$$\begin{aligned}
\langle \mathcal{T}(V), U \rangle &= \sum_{j \in \mathbb{I}} \operatorname{tr}\left(\mathcal{T}_j(V)^* U_j\right) \\
&= \sum_{j \in \mathbb{I}} \operatorname{tr}\left(\mathcal{T}_j(V^*) U_j\right) \\
&= \sum_{j \in \mathbb{I}} \sum_{i \in \mathbb{I}} p_{ij} \operatorname{tr}\left(A_i V_i^* A_i{}^* U_j\right) \\
&= \sum_{i,j \in \mathbb{I}} p_{ij} \operatorname{tr}\left(V_i^* A_i{}^* U_j A_i\right) \\
&= \sum_{i \in \mathbb{I}} \operatorname{tr}\left(V_i^* \left(\sum_{i \in \mathbb{I}} p_{ij} A_i{}^* U_j A_i\right)\right) \\
&= \sum_{i \in \mathbb{I}} \operatorname{tr}\left(V_i^* \mathcal{L}_i(U)\right) \\
&= \langle V, \mathcal{L}(U) \rangle.
\end{aligned}$$

□

Para cada $M = (M_1, \ldots, M_I) \in \mathbb{H}^d$, $\operatorname{diag}[M] = \operatorname{diag}[M_i] \in \mathbb{B}(\mathbb{C}^{dI})$ se denota la matriz diagonal por bloques de M dada por

$$\operatorname{diag}[M] = \begin{bmatrix} M_1 & 0 & \cdots & 0 \\ 0 & M_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & M_I \end{bmatrix}.$$

Luego, se define

$$\mathcal{B} = \left(P^t \otimes I_d\right) \operatorname{diag}\left[A\right] \in \mathbb{B}(\mathbb{C}^{dI}) \tag{3.11}$$

$$\mathcal{A} = \left(P^t \otimes I_{d^2}\right) \operatorname{diag}\left[\overline{A_i} \otimes A_i\right] \in \mathbb{B}(\mathbb{C}^{d^2 I}) \tag{3.12}$$

Para cada $j \in \mathbb{I}$,

$$\begin{aligned} \mathcal{B}_{j\cdot}q(k) &= \begin{bmatrix} p_{j1}^t A_1 & \cdots & p_{jS}^t A_I \end{bmatrix} \begin{bmatrix} q_1(k) \\ \vdots \\ q_I(k) \end{bmatrix} \\ &= \sum_{i \in \mathbb{I}} p_{ij} A_i q_i(k). \end{aligned}$$

Así de la Proposición 4 se tiene que la matriz $\mathcal{B}$ y el vector de primeros momentos definido en (3.2) se relacionan como sigue

$$q(k+1) = \mathcal{B}q(k) \tag{3.13}$$

y así

$$q(k) = \mathcal{B}^k q(0).$$

Proposición 6 ([6], Proposición 3.4)**.** *Para cada $Q \in \mathbb{H}^d$ se tiene*

$$\hat{\varphi}(\mathcal{T}(Q)) = \mathcal{A}\hat{\varphi}(Q).$$

Demostración. Sea $j \in \mathbb{I}$,

$$\begin{aligned}
\mathcal{A}_{j\cdot}\hat{\varphi}(Q) &= \left[\begin{array}{ccc} p_{j1}^t\overline{A_1} \times A_1 & \cdots & p_{jI}^t\overline{A_I} \times A_I \end{array}\right] \left[\begin{array}{c} \varphi(Q_1) \\ \vdots \\ \varphi(Q_I) \end{array}\right] \\
&= \sum_{i\in\mathbb{I}} p_{ji}\overline{A_i} \times A_i\varphi(Q_i) \\
&= \varphi\left(\sum_{i\in\mathbb{I}} p_{ji}A_iQ_iA_i^*\right) \\
&= \varphi\left(\mathcal{T}_j(Q)\right).
\end{aligned}$$

□

Por lo tanto de la Proposición 6 y la Proposición 5 se consigue

$$\rho(\mathcal{A}) = \rho(\mathcal{A}^*) = \rho(\mathcal{T}) = \rho(\mathcal{L}).$$

La siguiente proposición garantiza que la estabilidad del operador segundo momento garantiza la estabilidad del operador primer momento.

Proposición 7 ([6], Proposición 3.6)**.** *Si* $\rho(\mathcal{A}) < 1$ *entonces* $\rho(\mathcal{B}) < 1$.

Demostración. Es suficiente mostrar la siguiente afirmación: para todo vector e de $\mathbb{C}^{dI}$ se tiene

$$\mathcal{B}^k e \to 0 \quad \text{cuando} \quad k \to \infty.$$

Fije las bases $\{e_i \colon i = 1, \ldots, dI\}$ y $\{v_i \colon i = 1, \ldots, d\}$ de $\mathbb{C}^{dI}$ y $\mathbb{C}^d$, respectivamente. Para $1 \leq i \leq d$ y $j \in \mathbb{I}$, asuma las condiciones iniciales $x_0 = v_i$ y $\theta_0 \sim \delta_{\{j\}}(\cdot)$ en el Sistema homogéneo (3.1). Entonces $q(0) = e_{i+(j-1)d}$ y de la Ecuación (3.13) se sigue

$$\begin{aligned}
\left\|\mathcal{B}^k e_{i+(j-i)d}\right\|^2 = \left\|\mathcal{B}^k q(0)\right\|^2 &= \|q(k)\|^2 \\
&= \sum_{l\in\mathbb{I}} \|q_l(k)\|^2 \\
&= \sum_{l\in\mathbb{I}} \|\mathbb{E}\left[x(k)1_{\{\theta(k)=l\}}\right]\|^2 \\
&\leq \sum_{l\in\mathbb{I}} \mathbb{E}\left[\|x(k)1_{\{\theta(k)=l\}}\|^2\right] \\
&= \sum_{l\in\mathbb{I}} \mathrm{tr}\left(Q_l(k)\right) \\
&\leq \sum_{l\in\mathbb{I}} d\|Q_l(k)\| = d\|Q(k)\|_1.
\end{aligned}$$

Ahora, de la Proposición 6 se tiene

$$\hat{\varphi}\left(Q(k)\right) = \mathcal{A}^k \hat{\varphi}\left(Q(0)\right)$$

con $Q_j(0) = v_i {v_i}^*$ y $Q_l(0) = 0$ para $l \neq j$. Luego, como $\hat{\varphi}$ es un isomorfismo y $\rho(\mathcal{A}) < 1$ entonces $\|Q(k)\|_1 \to 0$ cuando $k \to \infty$ y así $\|\mathcal{B}^k e_{i+(j-i)d}\| \to 0$ cuando $k \to \infty$. Desde que i y j son arbitrarios se cumple la afirmación para todo vector en la base de $\mathbb{C}^{dI}$. $\square$

Observación 4 ([6], Observación 3.7)**.** *Sin embargo, la estabilidad del operador primer momento, $\mathcal{B}$, no garantiza la estabilidad del operador segundo momento $\mathcal{A}$. Por ejemplo, considere $d = 1$, $I = 2$,*

$$A_1 = 0{,}7, \quad A_2 = 1{,}25$$

y

$$P = \begin{bmatrix} 0{,}5 & 0{,}5 \\ 0{,}5 & 0{,}5 \end{bmatrix}.$$

En este caso

$$\mathcal{B} = \begin{bmatrix} 0{,}35 & 0{,}625 \\ 0{,}35 & 0{,}625 \end{bmatrix}, \qquad \mathcal{A} = \begin{bmatrix} 0{,}245 & 0{,}78125 \\ 0{,}245 & 0{,}78125 \end{bmatrix}.$$

Luego $\rho(\mathcal{B}) = 0{,}975 < 1$ *pero* $\rho(\mathcal{A}) = 1{,}02625 \geq 1$.

3.2. Estabilidad

Definición 7 (Estabilidad Media Cuadrática)**.** *Se dice que el sistema* (3.1) *es* estable en media cuadrática (EMC) *si para cualesquiera condiciones iniciales* $x_0 \in \mathcal{C}_0^d$*,* $\theta_0 \in \Theta_0$ *existen* $\mu \in \mathbb{C}^d$ *y* $\mathbb{M} \in \mathbb{B}(\mathbb{C}^d)^+$ *tales que*

1. $\|\mu(k) - \mu\| \to 0$ *cuando* $k \to \infty$.

2. $\|\mathbb{M}(k) - \mathbb{M}\| \to 0$ *cuando* $k \to \infty$.

A continuación se presenta el resultado principal, el cual establece condiciones necesarias y suficientes para la EMC del sistema (3.1).

Teorema 4 ([6], Teorema 3.9)**.** *Las siguientes proposiciones son equivalentes:*

1. *El sistema* (3.1) *es EMC.*

2. $$\rho(\mathcal{A}) < 1.$$

3. *Para cada* $U \in \mathbb{H}^{d+}$*,* $U > 0$*, existe un único* $V \in \mathbb{H}^{d+}$*,* $V > 0$*, que resuelve*
$$V - \mathcal{T}(V) = U. \tag{3.14}$$

4. *Para algún* $V \in \mathbb{H}^{d+}$*,* $V > 0$*, se satisface*
$$V - \mathcal{T}(V) > 0. \tag{3.15}$$

5. *Existen* $0 < \eta < 1 \leq \beta$*, de modo que para todo* $x_0 \in \mathcal{C}_0^d$ *y todo* $\theta_0 \in \Theta_0$*,*

$$\mathbb{E}\left[\|x(k)\|^2\right] \leq \beta\eta^k \mathbb{E}\left[\|x_0\|^2\right], \quad k \geq 0. \tag{3.16}$$

6. *Para todo* $x_0 \in \mathcal{C}_0^d$ *y todo* $\theta_0 \in \Theta_0$ *se tiene*

$$\sum_{k=0}^{\infty} \mathbb{E}\left[\|x(k)\|^2\right] < \infty. \tag{3.17}$$

Este resultado sigue siendo válido si se reemplaza, en (3.14) *y* (3.15)*,* $\mathcal{T}$ *por* $\mathcal{L}$*.*

Observación 5. *Del Teorema 4 es fácil ver que el sistema* (3.1) *es EMC equivale a ver que* $\mathbb{M}(k) \to 0$ *cuando* $k \to \infty$*. Esto será evidente luego de la revisión de las pruebas de la Proposición 10 y la Proposición 12. Ademá, de los items 3 y 4 de la Proposición 4 se sigue que el sistema* (3.1) *es EMC si y solo si la solución* $\{x(k) \colon k \geq 0\}$ *converge al origen en media cuadrática.*

Definición 8 (Estabilidad Estocástica)**.** *Se dice que el sistema* (3.1) *es* estable estocásticamente (EE) *si para cualesquiera condiciones iniciales* $x_0 \in \mathcal{C}_0^d$*,* $\theta_0 \in \Theta_0$

$$\sum_{k=0}^{\infty} \mathbb{E}\left[\|x(k)\|^2\right] < \infty \tag{3.18}$$

Corolario 1. *Si el sistema* (3.1) *es EMC entonces para todo* $x_0 \in \mathcal{C}_0^d$ *y todo* $\theta_0 \in \Theta_0$ *se tiene que existe* $\gamma > 0$ *tal que*

$$\limsup_{k} \frac{1}{k} \ln \|x(k)\| \leq -\gamma \quad \mathbb{P}\text{-c.s.}.$$

Demostración. Si el sistema (3.1) es EE existen constantes $0 < \eta < 1 \leq \beta$ tales que para todo $k \geq 0$ se tiene

$$\mathbb{E}\left[\|x(k)\|^2\right] \leq \beta\eta^k \mathbb{E}\left[\|x(0)\|^2\right].$$

Sea $0 < \gamma < -\frac{1}{2}\ln \eta$ entonces, por la desigualdad de Markov, se tiene

$$\mathbb{P}\left(\|x(k)\| > e^{-\gamma k}\right) \leq \frac{1}{e^{-2\gamma k}}\mathbb{E}\left[\|x(k)\|^2\right], \; k \geq 0,$$

y por lo tanto, sumando sobre $k \in \mathbb{N}$, se consigue

$$\begin{aligned}
\sum_{k=0}^{\infty}\mathbb{P}\left(\|x(k)\| > e^{-\gamma k}\right) &\leq \frac{1}{e^{-2\gamma k}}\sum_{k=0}^{\infty}\mathbb{E}\left[\|x(k)\|^2\right] \\
&\leq \sum_{k=0}^{\infty}\frac{\beta\eta^k}{e^{-2\gamma k}}\mathbb{E}\left[\|x(0)\|^2\right] \\
&= \beta\mathbb{E}\left[\|x(0)\|^2\right]\sum_{k=0}^{\infty}\left(e^{\ln\eta+2\gamma}\right)^k < +\infty. \qquad (3.19)
\end{aligned}$$

De esta desigualdad, el lema de Borel-Cantelli nos asegura que

$$\mathbb{P}\left(\|x(k)\| > e^{-\gamma k} \; i.v.\right) = 0.$$

Es decir,

$$\mathbb{P}\left(\frac{1}{k}\ln\|x(k)\| \leq -\gamma \; eventualm.\right) = 1,$$

y así

$$\mathbb{P}\left(\limsup_k \frac{1}{k}\ln\|x(k)\| \leq -\gamma\right) = 1.$$

□

Definición 9 (Estabilidad Exponencial Casi Segura)**.** *Se dice que el sistema* (3.1) *es* estable exponencial casi seguramente (EECS) *si para cualesquiera condiciones iniciales* $x_0 \in \mathcal{C}_0^d$, $\theta_0 \in \Theta_0$ *si existe* $\gamma > 0$ *tal que*

$$\limsup_k \frac{1}{k}\ln\|x(k)\| \leq -\gamma \quad \mathbb{P}\text{-c.s.}. \qquad (3.20)$$

Observación 6. *A continuación se esquematizan las relaciones entre los tipos de estabilidad definidos:*

$$\textit{EE} \Leftrightarrow \textit{EMC} \Rightarrow \textit{EECS}.$$

Por otro lado, en [5] muestran que cuando la cadena de Markov $\{\theta(k)\colon k \geq 0\}$ *toma valores en un espacio infinito numerable entonces se pierde la equivalencia entre EE y EMC.*

3.3. Ejemplos

A continuación se listan algunos ejemplos que muestran la diversidad de comportamientos que pueden tener los sistemas lineales de salto Markoviano. En los ejemplos, se usa el test del radio espectral para determinar si hay estabilidad por ser el más conocido y fácil de computar.

Ejemplo 2 (EMC no implica estabilidad de los modos de operación, Ejemplo 3.18 de [6])**.** *Sea* $\{\theta(k)\}_{k\geq 0}$ *una cadena de Markov que toma solo dos estados y tiene matriz de transición*

$$P = \begin{pmatrix} 0{,}1 & 0{,}9 \\ 0{,}9 & 0{,}1 \end{pmatrix}.$$

Sean los dos posibles modos de operación

$$A_1 = \begin{pmatrix} 2 & -1 \\ 0 & 0 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 0 & 1 \\ 0 & 2 \end{pmatrix}.$$

Observe que cada modo de operación es inestable. Sin embargo, el sistema es EMC pues *el operador segundo momento*

$$\mathcal{A} = \left[\begin{array}{cccc|cccc} 0{,}4 & -0{,}2 & -0{,}2 & 0{,}1 & 0 & 0 & 0 & 0{,}9 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1{,}8 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1{,}8 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3{,}6 \\ \hline 3{,}6 & -1{,}8 & -1{,}8 & 0{,}9 & 0 & 0 & 0 & 0{,}1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0{,}2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0{,}2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0{,}4 \end{array}\right]$$

es estable ($\rho(\mathcal{A}) = 0{,}4 < 1$). Es decir, la estabilidad del Sistema Lineal de Salto Markoviano no implica la estabilidad de los modos de operación del sistema.

Ejemplo 3 (Estabilidad de los modos de operación no implica EMC, Ejemplo 3.17 de [6])**.** *Sea $\{\theta(k)\}_{k\geq 0}$ una cadena de Markov que toma solo dos estados y tiene matriz de transición*

$$P = \begin{pmatrix} 0{,}5 & 0{,}5 \\ 0{,}5 & 0{,}5 \end{pmatrix}.$$

Sean los dos posibles modos de operación

$$A_1 = \begin{pmatrix} 0 & 2 \\ 0 & 0{,}5 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 0{,}5 & 0 \\ 2 & 0 \end{pmatrix}.$$

Observe que cada modo de operación es inestable. Sin embargo, el sistema es EMC pues *el operador segundo momento*

$$\mathcal{A} = \left[\begin{array}{cccc|cccc} 0 & 0 & 0 & 2 & 0{,}125 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0{,}5 & 0{,}5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0{,}5 & 0{,}5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0{,}125 & 2 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 2 & 0{,}125 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0{,}5 & 0{,}5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0{,}5 & 0{,}5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0{,}125 & 2 & 0 & 0 & 0 \end{array}\right]$$

no es estable ($\rho(\mathcal{A}) = 2{,}125 \geq 1$). Es decir, la estabilidad de los modos de operación del sistema no implica la estabilidad del Sistema Lineal de Salto Markoviano.

Ejemplo 4 (EECS no implica EMC)**.** *Sea $\{\theta(k)\}_{k\geq 0}$ una cadena de Markov que toma solo dos estadosy matriz de transición*

$$P = \begin{pmatrix} 0{,}6 & 0{,}4 \\ 0{,}1 & 0{,}9 \end{pmatrix}.$$

Sean los dos posibles modos de operación

$$A_1 = \begin{pmatrix} 0{,}2 & 1 \\ 0 & 0{,}2 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 0{,}9 & 0{,}4 \\ 0{,}5 & 0{,}2 \end{pmatrix}.$$

Como se muestra en el ejemplo de la sección V de [1], el Sistema Lineal de Salto Marko-

viano es EECS. Sin embargo, el sistema no es EMC pues el operador segundo momento

$$\mathcal{A} = \left[\begin{array}{cccc|cccc} 0{,}024 & 0{,}12 & 0{,}12 & 0{,}6 & 0{,}004 & 0{,}02 & 0{,}02 & 0{,}10 \\ 0 & 0{,}024 & 0 & 0{,}12 & 0 & 0{,}004 & 0 & 0{,}02 \\ 0 & 0 & 0{,}024 & 0{,}12 & 0 & 0 & 0{,}004 & 0{,}02 \\ 0 & 0 & 0 & 0{,}024 & 0 & 0 & 0 & 0{,}004 \\ \hline 0{,}324 & 0{,}144 & 0{,}144 & 0{,}064 & 0{,}729 & 0{,}324 & 0{,}324 & 0{,}144 \\ 0{,}18 & 0{,}072 & 0{,}08 & 0{,}032 & 0{,}405 & 0{,}162 & 0{,}18 & 0{,}072 \\ 0{,}18 & 0{,}08 & 0{,}072 & 0{,}032 & 0{,}405 & 0{,}18 & 0{,}162 & 0{,}072 \\ 0{,}10 & 0{,}04 & 0{,}04 & 0{,}016 & 0{,}225 & 0{,}09 & 0{,}09 & 0{,}036 \end{array}\right]$$

no es estable ($\rho(\mathcal{A}) = 1{,}14375 \geq 1$). Es decir, la estabilidad exponencial casi segura del sistema no implica la estabilidad media cuadrática.

Ejemplo 5 (EECS, un contraejemplo, Ejemplo 4.6 de [10]). *En $d = 1$, considere $A_1, A_2, \ldots A_I$ números positivos, y $\{\theta(k) \colon k \geq 0\}$ una cadena de variables aleatorias independientes e idénticamente distribuidas de tal manera que*

$$\mathbb{E}\left[\ln \|A_{\theta(0)}\|\right] = 0.$$

Luego, el conjunto de variables aleatorias $\{\ln \|A_{\theta(k)}\| \,:\, k \geq 0\}$ son independientes, idénticamente distribuidas, centradas y tienen varianza finita de modo que por la Ley del Logaritmo Iterado (Teorema 3.25 en [26])

$$\mathbb{P}\left(\limsup_k \frac{\ln \|A_{\theta(0)}\| + \cdots + \ln \|A_{\theta(k-1)}\|}{\sqrt{2\sigma^2 k \ln \ln k}} = 1\right) = 1.$$

Luego,

$$\limsup_k \ln \|A_{\theta(k-1)} \cdots A_{\theta(0)}\| = +\infty \quad \mathbb{P} - c.s.$$

y así

$$\limsup_k \|x(k)\| = \limsup_k \|A_{\theta(k-1)} \cdots A_{\theta(0)} x(0)\| = +\infty \quad \mathbb{P} - c.s.$$

Análogamente, se llega a mostrar que

$$\liminf_k \|x(k)\| = \liminf_k \|A_{\theta(k-1)} \cdots A_{\theta(0)} x(0)\| = -\infty \quad \mathbb{P}-c.s.$$

Por lo tanto, no hay EECS pues las trayectorias del Sistema Lineal de Salto Markoviano oscilan en todo el espacio $\mathbb{C}$.

Ejemplo 6 (EMC no implica EE, Observación 6 de [5])**.** *Cuando la cadena de Markov* $\{\theta(k)\colon k \geq 0\}$ *toma valores en un espacio de estados numerable, los conceptos de EMC y EE no son equivalentes. De hecho, considere el Sistema Lineal de Salto Markoviano* (3.1) *con* $d = 1$,

$$P_{i,i+1} = 1 \quad y \quad A_i = \sqrt{-\frac{i}{i+1}}, \;\; i = 1, 2, 3, \ldots$$

En este caso

$$\mathbb{E}\left[\|x(k)\|^2 \mid x(0), \theta(0)\right] = \frac{\theta_0}{\theta_0 + k} {x_0}^2 \to 0, \; \textit{cuando } k \to \infty,$$

y $x_0 \in \mathcal{C}_0^d$ *domina a la sucesión de esperanzas condicionales. Entonces, por el Teorema de Convergencia Dominada se tiene que*

$$\mathbb{E}\left[\|x(k)\|^2\right] = \mathbb{E}\left[\mathbb{E}\left[\|x(k)\|^2 \mid x(0), \theta(0)\right]\right] \to 0, \; \textit{cuando } k \to \infty,$$

para cualesquiera condiciones iniciales $\theta_0 \in \Theta_0, x_0 \in \mathcal{C}_0^d$*, es decir, el Sistema Lineal de Salto Markoviano es EMC. Sin embargo, cuando* $x_0 = 1$ *y* $\theta_0 = 1$ *se tiene que*

$$\sum_{k=0}^{+\infty} \mathbb{E}\left[\|x(k)\|^2\right] = \sum_{k=0}^{+\infty} \frac{1}{1+k} = +\infty,$$

es decir, el sistema no es EE.

3.4. Demostración del Teorema de Estabilidad

Teorema 5 ([6], Teorema 3.19)**.** *Si existen* $U, V \in \mathbb{H}^{d+}$, $U, V > 0$, *cummpliendo*

$$V = \mathcal{T}(V) + U$$

entonces

$$\rho(\mathcal{A}) < 1.$$

Este resultado sigue siendo válido si se reemplaza $\mathcal{T}$ *por* $\mathcal{L}$*.*

Demostración. Como $\rho(\mathcal{A}) = \rho(\mathcal{L})$ se estudia el sistema lineal

$$\begin{cases} Y(k+1) = \mathcal{L}(Y(k)) \\ Y(0) = Y_0 \in \mathbb{H}^{d+} \end{cases} . \tag{3.21}$$

Ahora, si el origen es estable global asintóticamente entonces de la Proposición 1 se tiene lo requerido.

Defina la función $\phi \colon \mathbb{H}^{d+} \to \mathbb{R}$ como sigue: para $Y \in \mathbb{H}^{d+}$

$$\begin{aligned} \phi(Y) := \langle V, Y \rangle &= \sum_{j \in \mathbb{I}} \operatorname{tr}(Y_j V_j) \\ &= \sum_{j \in \mathbb{I}} \operatorname{tr}(Y_j V_j^{1/2} V_j^{1/2}) \\ &= \sum_{j \in \mathbb{I}} \operatorname{tr}(V_j^{1/2} Y_j V_j^{1/2}) \geq 0. \end{aligned}$$

Del Teorema 1 (de Lyapunov) basta que ϕ sea una función de Lyapunov del Sistema (3.21) en el origen, para ello es necesario verificar las siguientes propiedades:

a) ϕ es continua.

b) $\phi(Y) > \phi(0) = 0$ para todo $Y \neq 0$ en $\mathbb{H}^{d+}$.

c) $\Delta\phi(Y) = \phi(\mathcal{L}(Y)) - \phi(Y) < 0$ para todo $Y \neq 0$ en $\mathbb{H}^{d+}$.

d) $\phi(Y) \to \infty$ cuando $\|Y\|_2 \to \infty$ en $\mathbb{H}^{d+}$.

En efecto. Para cada $W \in \mathbb{H}^{d+}$, denote

$$\lambda_m(W) = \min_{j\in\mathbb{I}} \min_{1\leq i\leq d} \lambda_i(W_j) \geq 0$$

$$\lambda_M(W) = \max_{j\in\mathbb{I}} \max_{1\leq i\leq d} \lambda_i(W_j) \geq 0.$$

Más aún, como $V, U > 0$ se tiene que $\lambda_M(V) \geq \lambda_m(V) > 0$, $\lambda_M(U) \geq \lambda_m(U) > 0$. Luego, para $Y = (Y_1, \ldots, Y_I) \in \mathbb{H}^{d+}$, de la Observación 1 se satisface

$$\lambda_m(V) \sum_{i=1}^{d} \lambda_i(Y_j) \leq \mathrm{tr}(V_j Y_j) \leq \lambda_M(V) \sum_{i=1}^{d} \lambda_i(Y_j)$$

y así

$$\lambda_m(V) \sum_{j\in\mathbb{I}} \sum_{i=1}^{d} \lambda_i(Y_j) \leq \phi(Y) \leq \lambda_M(V) \sum_{j\in\mathbb{I}} \sum_{i=1}^{d} \lambda_i(Y_j). \tag{3.22}$$

Además,

$$\|Y\|_2^2 = \sum_{j\in\mathbb{I}} \mathrm{tr}(Y_j^2) = \sum_{j\in\mathbb{I}} \sum_{i=1}^{d} \lambda_i(Y_j)^2$$

y

$$\sum_{j\in\mathbb{I}} \sum_{i=1}^{d} \lambda_i(Y_j)^2 \leq \left(\sum_{j\in\mathbb{I}} \sum_{i=1}^{d} \lambda_i(Y_j) \right)^2 \leq Id \sum_{j\in\mathbb{I}} \sum_{i=1}^{d} \lambda_i(Y_j)^2,$$

por lo que

$$\|Y\|_2 = 0 \quad \Leftrightarrow \quad \lambda_i(Y_j) = 0,\ i = 1, \ldots, d,\ j \in \mathbb{I}$$

y

$$\|Y\|_2 \to \infty \quad \Leftrightarrow \quad \sum_{j\in\mathbb{I}} \sum_{i=1}^{d} \lambda_i(Y_j) \to \infty.$$

Entonces, (b) y (d) son consecuencia de la Ecuación (3.22). La continuidad de ϕ en (a) es consecuencia de su definición como producto interno. Por último, para mostrar (c), de la Proposición 5 se sigue

$$\begin{aligned}\Delta\phi(Y) = \phi(\mathcal{L}(Y)) - \phi(Y) &= \langle V, \mathcal{L}(Y)\rangle - \langle V, Y\rangle \\ &= \langle \mathcal{T}(V), Y\rangle - \langle V, Y\rangle \\ &= \langle V - U, Y\rangle - \langle V, Y\rangle \\ &= -\langle U, Y\rangle \\ &\leq -\lambda_m(U) \sum_{j\in\mathbb{I}} \sum_{i=1}^{d} \lambda_i(Y_j) < 0\end{aligned}$$

cuando $Y \neq 0$ en $\mathbb{H}^{d+}$. $\square$

De manera similar, desde que $\rho(\mathcal{T}) = \rho(\mathcal{A})$, de la Proposición 2 se obtiene el siguiente resultado.

Proposición 8 ([6], Proposición 3.20)**.** *Si* $\rho(\mathcal{A}) < 1$ *entonces existe una única* $V \in \mathbb{H}^d$*, tal que*

$$V = \mathcal{T}(V) + U$$

para cada $U \in \mathbb{H}^d$*. Más aún,*

- $V = \hat{\varphi}^{-1}\left((I_{d^2S} - \hat{\varphi}[\mathcal{T}])^{-1}\hat{\varphi}(U)\right) = \sum_{k=0}^{\infty} \mathcal{T}^k(U)$.
- $U = U^* \Leftrightarrow V = V^*$.
- $U \geq 0 \Rightarrow V \geq 0$.
- $U > 0 \Rightarrow V > 0$.

Este resultado sigue siendo válido si se reemplaza $\mathcal{T}$ *por* $\mathcal{L}$*.*

Proposición 9 ([6], Proposición 3.25)**.** *Si* $\rho(\mathcal{A}) < 1$ *entonces existen* $0 < \eta < 1 \leq \beta$*, de* *modo que para todo $x_0 \in \mathcal{C}_0^d$ y todo $\theta_0 \in \Theta_0$,*

$$\mathbb{E}\left[\|x(k)\|^2\right] \leq \beta\eta^k \mathbb{E}\left[\|x_0\|^2\right], \quad k \geq 0.$$

Demostración. Suponga $\rho(\mathcal{T}) = \rho(\mathcal{A}) < 1$. Entonces, de la Proposición 1, existen $0 < \eta < 1 \leq \overline{\beta}$ de modo que

$$\|\mathcal{T}^k\|_1 \leq \overline{\beta}\,\eta^k, \quad k = 0, 1, \ldots$$

Desde que $Q(k) = \mathcal{T}^k(Q(0))$, de las desigualdades de la Proposición 4 se sigue

$$\begin{aligned}
\mathbb{E}[\|x(k)\|^2] &\leq d\|Q(k)\|_1 \\
&\leq d\|\mathcal{T}^k\|_1 \|Q(0)\|_1 \\
&\leq d\overline{\beta}\,\eta^k \mathbb{E}\left[\|x_0\|^2\right].
\end{aligned}$$

Claramente se tiene lo requerido con $\beta = d\overline{\beta}$. □

La siguiente proposición muestra que EE implica EMC.

Proposición 10 ([6], Proposición 3.24)**.** *Si el Sistema (3.1) es EE entonces también es EMC.*

Demostración. De la Ecuación (3.17) y de la tercera desigualdad en la Proposición 4 se deduce que

$$\|Q(k)\|_1 \leq \mathbb{E}[\|x(k)\|^2] \to 0$$

cuando $k \to \infty$ para cualesquiera condiciones iniciales $x_0 \in \mathcal{C}_0^d$ y $\theta(0) \in \Theta_0$. Luego, desde que $\rho(\mathcal{A}) = \rho(\mathcal{T}) < 1$, la Proposición 7 implica que $\rho(\mathcal{B}) < 1$. Así,

$$Q(k) \to 0, \quad q(k) = \mathcal{B}^k q(0) \to 0,$$

y por lo tanto

$$\mathbb{M}(k) \to 0, \quad \mu(k) \to 0$$

cuando $k \to \infty$ para cualesquiera condiciones iniciales $x_0 \in \mathcal{C}_0^d$ y $\theta(0) \in \Theta_0$. $\square$

Ahora, se establece la conexión entre el radio espectral de $\mathcal{A}$ y la EMC. Primero, de la Ecuación (3.10) se tiene

$$Q(k) = \mathcal{T}^k(Q(0))$$

y así

$$\mathbb{M}(k) = \sum_{j \in \mathbb{I}} Q_j(k) = \sum_{j \in \mathbb{I}} \mathcal{T}_j^k(Q(0)).$$

Como $\rho(\mathcal{T}) = \rho(\mathcal{A})$ queda establecido el siguiente resultado.

Proposición 11 ([6], Proposición 3.22)**.** *Si* $\rho(\mathcal{A}) < 1$ *entonces el Sistema* (3.1) *es EMC.*

Segundo, a continuación se enuncia la afirmación reversa.

Proposición 12 ([6], Proposición 3.23)**.** *Si el Sistema* (3.1) *es EMC entonces* $\rho(\mathcal{A}) < 1$.

Demostración. Desde que $\mathcal{A} = \mathcal{T}$, es suficiente mostrar que $\rho(\mathcal{T}) < 1$. Por hipótesis, existe $\mathbb{M} \geq 0$ tal que $\mathbb{M}(k) \to \mathbb{M}$ cuando $k \to \infty$ para todo $x_0 \in \mathcal{C}_0^d$ y $\theta_0 \in \Theta_0$. Para $x_0 = 0$ se tiene $Q(0) = 0$ y de ello

$$\mathbb{M}(k) = \sum_{j \in \mathbb{I}} \mathcal{T}_j^k(Q(0)) \to 0,$$

es decir $\mathbb{M} = 0$. Entonces, para todo $Q(0) \in \mathbb{H}^{d+}$

$$\mathcal{T}^k(Q(0)) \to 0.$$

Además, las matrices componente de cualquier $Q = (Q_1, \ldots, Q_I) \in \mathbb{H}^{d+}$ corresponden a la matriz convarianza de algún vector aleatorio complejo con distribución Gaussiana compleja multivariada, es decir,

$$Q_i = \mathbb{E}[x_0 {x_0}^*] = Q_i(0)$$

con condiciones iniciales $x_0 \sim CN(0, Q_i, 0)$ y $\theta_0 \sim \delta_{\{i\}}(\cdot)$. Luego, desde que T es lineal se tiene que

$$\mathcal{T}^k(Q) \to 0$$

para todo $Q \in \mathbb{H}^{d+}$. Así, por la Proposición 1, $\rho(\mathcal{T}) < 1$. □

Finalmente se realiza la prueba del Teorema 4.

Demostración. Se sigue la secuencia

$$(2) \Rightarrow (3) \Rightarrow (4) \Rightarrow (2) \Rightarrow (5) \Rightarrow (6) \Rightarrow (1) \Rightarrow (2)$$

dada en la secuencia de resultados: Proposición 8, directo, Teorema 5, Proposición 9, Directo, Proposición 10, Proposición 12.

□

Análisis de resultados y conclusiones

Análisis de resultados y constrastación de hipótesis

En cada item siguiente se analiza puntualmente las partes del presente trabajo.

- Respecto del buen planteamiento del problema es necesario construir un marco matemático para el estudio del sistema dinámico estocástico. Específicamente, en la última sección del capítulo 2, para un Sistema Lineal de Salto Markoviano en Tiempo Discreto se pueden estudiar las trayectorias solución como parte de una cadena de Markov asociada al sistema (Proposición 3) con lo cual queda resuelta esta cuestión.

- Una metodología general para el estudio de un sistema dinámico estocástico es la teoría de estabilidad según Lyapunov. Clásicamente, esta metodología es útil para un sistema dinámico en tiempo discreto (Teorema 1) y, en particular, para un Sistema Lineal en Tiempo Discreto (Teorema 2). Lo interesante es que se puede aplicar, con alguna adecuación, a un Sistema Lineal de Salto Markoviano en Tiempo Discreto (Teorema 5).

- La Estabilidad Media Cuadrática (EMC, Definición 7) de un Sistema Lineal de Salto Markoviano en Tiempo Discreto admite un paisaje completo (Teorema 4) análogo a la caracterización de la estabilidad de un Sistema Lineal en Tiempo Discreto

(Teorema 2). Es decir, la hipótesis general se cumple y EMC puede considerarse como la generalización natural de la noción de estabilidad para un Sistema Lineal en Tiempo Discreto. Específicamente, la segunda afirmación del Teorema 4 es un test de radio espectral para un Sistema Lineal de Tiempo Discreto asociado al Sistema Lineal de Salto Markoviano y la tercera afirmación del Teorema 4 es un conjunto de ecuaciones de Lyapunov en forma compacta para el Sistema Lineal de Salto Markoviano. Por lo tanto, las hipótesis específicas también son satisfechas.

- La noción de estabilidad para un Sistema Lineal de Salto Markoviano en Tiempo Discreto tiene diversos sabores: EMC, Estabilidad Estocástica (EE, Definición 8) y Estabilidad Exponencial Casi Segura (EECS, Definición 9). Cuando la cantidad de modos de operación del Sistema Lineal de Salto Markoviano en Tiempo Discreto es finita (nuestro caso de estudio) se tiene la equivalencia entre EMC y EE, las cuales implican EECS (Observación 6). Sin embargo, la infinitud de modos de operación posibilita que se pierda la equivalencia entre EMC y EE (Ejemplo 6) haciendo notar que la cantidad de modos de operación influye sobre la estabilidad del sistema.

- La breve lista de ejemplos de Sistemas Lineales de Salto Markoviano en Tiempo Discreto muestran la diversidad de comportamientos de las trayectorias solución, contradiciendo la ingenua intuición de que la estabilidad de los modos de operación y la estabilidad del Sistema Lineal de Salto Markoviano están directamente relacionadas. Esta riqueza, o complejidad, proviene de la acción de la aleatoriedad en el sistema, y también se pone en evidencia en los enunciados del Teorema 4.

Conclusiones

Del presente trabajo se obtienen las siguientes conclusiones.

- La estabilidad media cuadrática (EMC) de un Sistema Lineal de Salto Markoviano en Tiempo Discreto admite un paisaje completo como se esperaba. De hecho, en el Teorema 2 se exhiben los enunciados análogos al test de radio espectral y la ecuación de Lyapunov del modelo clásico listados en el Teorema 4. Cumpliéndose así los obejtivos generales y específicos.

- Cuando la cantidad de modos de operación del Sistema Lineal de Salto Markoviano en Tiempo Discreto es finita (nuestro caso de estudio) se tiene la equivalencia entre los tipos de Estabilidad Media Caudrática (EMC) y la Estabilidad Estocástica (EE), las cuales implican la Estabilidad Exponencial Casi Segura (EECS).

- La estabilidad de los modos de operación y la estabilidad del Sistema Lineal de Salto Markoviano no están directamente relacionadas. Además, la noción de estabilidad depende de la cantidad modos de operación del sistema y la dinámica aleatoria de cambio entre estos.

Recomendaciones

Cabe mencionar que hay muchos puntos aún por estudiar, se recomienda para trabajos futuros los listados a continuación.

- Una caracterización completa para la Estabilidad Exponencial Casi Segura de un Sistema Lineal de Salto Markoviano en Tiempo Discreto.

- Explotar la propiedades que tiene una cadena de Markov para el caso del proceso estocástico de la Proposición 3.

- Explorar casos más generales de Sistemas Lineales de Salto Markoviano en Tiempo Discreto, por ejemplo, cuando hay una cantidad numerable de modos de operación para el sistema.

- Simular por una solución de un Sistema Lineal de Salto Markoviano en Tiempo Discreto. El método a usar sería el algoritmo de Monte Carlo con Cadenas de Markov.

- Desarrollar la teoría de estabilidad de un Sistema Lineal de Salto Markoviano en Tiempo Continuo.

Bibliografía

[1] P. Bolzern, P. Colaneri, and G. D. Nicolao. On almost sure stability of discrete-time markov jump linear systems. In *2004 43rd IEEE Conference on Decision and Control (CDC) (IEEE Cat. No.04CH37601)*, volume 3, pages 3204–3208, December 2004.

[2] J. Brewer. Kronecker products and matrix calculus in system theory. *IEEE Transactions on Circuits and Systems*, 25(9):772–781, September 1978.

[3] K. L. Chung. *A Course in Probability Theory*. Academic Press, 3 edition, 2001.

[4] O. Costa and M. Fragoso. Stability results for discrete-time linear systems with markovian jumping parameters. *Journal of Mathematical Analysis and Applications*, 179(1):154 – 178, 1993.

[5] O. L. V. Costa and M. D. Fragoso. Discrete-time lq-optimal control problems for infinite markov jump parameter systems. *IEEE Transactions on Automatic Control*, 40(12):2076–2088, December 1995.

[6] O. L. V. Costa, M. D. Fragoso, and R. P. Marques. *Discrete-Time Markov Jump Linear Systems*. Probability and Its Applications. Springer Science & Business Media, 2005.

[7] A. Czornik. On control problems for jump linear systems.

[8] R. Durrett. *Probability: Theory and Examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 4 edition, 2010.

[9] Y. Fang. *Stability analysis of linear control systems with uncertain parameters*. PhD thesis, Case Western Reserve University, 1994.

[10] Y. Fang, K. Loparo, and X. Feng. Stability of discrete time jump linear systems. *Journal of Mathematical Systems, Estimation and Control*, 5(3):275–321, 1995.

[11] X. Feng, K. A. Loparo, Y. Ji, and H. J. Chizeck. Stochastic stability properties of jump linear systems. *IEEE Transactions on Automatic Control*, 37(1):38–53, Jan 1992.

[12] R. G. Gallager. *Stochastic Processes: Theory for Applications*. Cambridge University Press, 2013.

[13] D. J. H. Garling. *A Course in Mathematical Analysis*, volume 2. Cambridge University Press, 2014.

[14] Y. Ji, H. Chizeck, X. Feng, and K. Loparo. Stability and control of discretetime jumplinear systems. *Control, theory and advanced technology*, 7:247–270, 06 1991.

[15] O. Kallenberg. *Foundations of Modern Probability*. Probability and Its Applications. Springer-Verlag New York, 2006.

[16] C. Kubrusly. Mean square stability for discrete bounded linear systems in Hilbert space. *SIAM Journal on Control and Optimization*, 23(1):19–29, 1985.

[17] C. Kubrusly and O. Costa. Mean square stability conditions for discrete stochastic bilinear systems. *IEEE Transactions on Automatic Control*, 30(11):1082–1087, November 1985.

[18] D. Li and H. Queffélec. *Introduction à L'étude des Espaces de Banach: Analyse et Probabilités*, volume 12. Société Mathématique de France, 2004.

[19] D. G. Luenberger. *Introduction to Dynamic Systems: Theory, Models, and Applications*. John Wiley & Sons, 1979.

[20] J. E. Mayta Guillermo. Sistemas lineales asociados a una cadena de Markov en tiempo discreto. Tesis de licenciatura, Universidad Nacional de Ingeniería, 2016. Accedido: 2018-04-06.

[21] K. R. Parthasarathy. *Introduction to Probability and Measure*, volume 33 of *Texts and Readings in Mathematics*. Hindustan Book Agency, 2005.

[22] P. Petersen. *Linear Algebra*. Undergraduate Texts in Mathematics. Springer-Verlag New York, 2012.

[23] B. Simon. *Operator Theory: A Comprehensive Course in Analysis*, volume 4. American Mathematical Society, 2015.

[24] V. S. Sunder. *Functional Analysis: Spectral Theory*, volume 13 of *Texts and Readings in Mathematics*. Hindustan Book Agency, 1997.

[25] D. Sworder and R. Rogers. An lq-solution to a control problem associated with a solar thermal central receiver. *IEEE Transactions on Automatic Control*, 28(10):971–978, October 1983.

[26] S. R. Varadhan. *Probability Theory*, volume 7 of *Courant Lecture Notes in Mathematics*. AMS and the Courant Institute of Mathematical Sciences at New York University, 2001.